

ADAPTIVE SEARCH USING A REPRODUCTIVE META-PLAN

R. E. MERCER and J. R. SAMPSON

Department of Computing Science, University of Alberta, Edmonton, Alberta (Canada)

(Received May 30, 1977)

A reproductive plan is a type of adaptive procedure devised by Holland¹ which embodies many principles found in the adaptation of natural systems through evolution. This study develops a reproductive meta-plan, an adaptive procedure that controls the modification of parameter values in a reproductive plan. After a brief review of Holland's formalization of reproductive plans, a non-reproductive meta-plan devised by Cavicchio is discussed. In response to some of the limitations of Cavicchio's² approach, a reproductive meta-plan is developed. Experiments with computer implementations of both meta-plans show improvement in a reproductive plan's utility gain when the reproductive meta-plan is used.

1 INTRODUCTION

Adaptation appears in both natural situations, such as evolution, and man-made instances, such as artificial intelligence, control theory, and economic planning. Different contexts involve different kinds of devices, different methods of device modification, different environments, and different performance criteria. As a result of this diversity the terms "adaptive" and "adaptation" have been used with many, sometimes inconsistent, meanings. These inconsistencies have arisen not from fundamental disagreement about the nature of adaptation, but rather from the diverse realizations of adaptation as implementations that *behave* adaptively.

Theoretical unification thus requires an acceptable general paradigm, or *functional* definition, of adaptation. Recognizing this requirement, John H. Holland and his colleagues at the University of Michigan have begun to formalize adaptation. The theory is intended to be adequate for describing natural adaptation and for developing new algorithms for artificial adaptation. The framework permits direct comparison of widely differing adaptive systems.

The function of an *adaptive plan* is to make successive modifications to a device, producing a trajectory through the space of all possible devices. Due to incomplete information about the environment with which the device is interacting, it is uncertain whether a modification will be beneficial or deleterious. The essence of adaptation is the use of information received from the environment (feedback) to modify the device. The feedback function can have many local optima. An *efficient* plan converges quickly to devices whose

feedback corresponds to these optima. An *effective* plan's search prevents entrapment on these early local optima by continuously locating better local optima in search of the global optimum.

Holland¹ has developed the *reproductive plan* as a candidate for an adaptive plan which is both efficient and effective in a wide range of environments. Much of the development has been influenced by observing and generalizing natural adaptive methods found in systems of reproducing organisms. In the operation of a reproductive plan, an initial set of devices is encoded into strings of attributes called structures. In each reproductive cycle, the current structures are tested in the environment. Structures are reproduced (copied) a number of times according to their relative performance. Modification procedures (operators) are then employed to produce new structures. In the type of reproductive plan of interest to Holland, the operators are analogous to the genetic mechanisms of crossover, inversion, and (point) mutation. This study is concerned with using a reproductive meta-plan to study meta-adaptation of the genetic operator application rates.

The next section of this paper introduces Holland's theoretical framework and mentions some previous studies inspired by it. In some of these studies different stages of adaptation have required different rates (probabilities) of operator application. Section 3 considers one approach to introducing such variation, Cavicchio's² parameter modification scheme. This non-reproductive meta-plan leaves a number of meta-adaptation problems unsolved. The fourth section of this paper responds to these problems by proposing a reproductive meta-plan. Section 5 discusses experimental

results obtained from testing an implementation of the proposed reproductive meta-plan and an implementation of Cavicchio's parameter modification scheme. The final section assesses what has been accomplished in this work and concludes with a discussion of avenues for further research.

2 REPRODUCTIVE PLANS

The following discussion is necessarily terse and incomplete. Readers unfamiliar with Holland's work are encouraged to consult his book¹ for further details and examples. Explication of the theory begins with specific formal realizations of the general terms in the following (loose) definition of adaptation: some *thing* is, by some *means*, *improving* with respect to some *extrinsic criterion*.

The *thing* Holland calls a *structure*, denoted A . In ecology structures would represent organisms, in genetics, chromosomes, and in artificial intelligence, programs. The set of all possible structures is denoted \mathcal{A} . Since adaptation is a modification procedure, those tasks of interest have structures with identifiable and replaceable substructural units, or more simply *subunits*. For example in Samuel's Checker Player³ the structures would correspond to all sets of weights (coefficients in the polynomial evaluation function), while subunits would be the individual weights. Chromosome subunits are genetic loci. Associated with each structure is an information vector I , an element of \mathcal{I} . The structures can be expanded to include memory, in the form of an explicit finite record of past information vectors.

The *extrinsic criterion* for adaptation is the *environment*, denoted E and exemplified by the ecosystem of an organism or the data for a program. E communicates the outcome of testing a structure through an instance of the information vector. In interesting adaptation problems it is uncertain what E will confront the plan (or which set of information vectors will describe the outcomes of the tests). In Samuel's Checker Player for example, the sequence of the opponent's moves is unknown at the beginning of a game.

The means of adaptation, the procedure which modifies the structures, is the *adaptive plan*, \mathcal{T} . A trajectory through the set of structures is produced by successive generation of new structures. To be adaptive the trajectory must be influenced by the environmental input. Hence if $\mathcal{A}(t)$, the set of structures tried at time t , is the state of the plan at time t , then \mathcal{T} is a state transition function, $\mathcal{A}(t+1) =$

$\mathcal{T}(\mathcal{I}(t), \mathcal{A}(t))$. The plan can also be realized as a stochastic process, so that $\mathcal{A}(t+1)$ is derived according to the distribution $P(t+1) = \mathcal{T}(\mathcal{I}(t), \mathcal{A}(t))$, where $P(t+1)$ is a probability distribution over the set of structures.

New structures are generated by the application of *operators* to existing structures. Operators are chosen (by some specified mechanism) from the set of all available operators. Examples of operators are learning rules in artificial intelligence (e.g. Samuel's weight adjustment function) and mutation and cross-over in genetics.

The *improvement* of the structure(s) corresponds to an increase in *fitness (utility, payoff)*, a function which maps structures in a particular environment into the reals. For many tasks, this function can be multi-dimensional and, in any or all dimensions, non-linear, discontinuous, and multimodal, creating intriguing problems that cannot be solved by analytic techniques. The utility measure is often the only component of I , which is the case throughout this paper.

In addition to providing a general framework for investigating adaptive processes, Holland has also proposed a class of adaptive plans, called reproductive plans, which he has shown to be robust. Informally, a robust plan is efficient in all environments. Robustness is associated with optimal allocation of trials to structures (for all feedback functions).

The search process poses two conflicting methods of finding better structures, (1) exploitation of known subunits in new contexts, and (2) search for alternative subunits that may lead to optimal performance. The adaptive plan must therefore continue to test and incorporate new and old subunits associated with better performance. To identify useful subunits, structures must be compared, which leads to the problem of apportioning, to the subunits, credit for changes in the performance by structures. Since subunit contributions to performance may combine non-linearly, a method for associating combinations of subunits is necessary. Holland introduces the notion of a *schema* for this purpose.

Schemata can best be studied if the subunits comprise a set of l "detectors," each mapping an effectively describable feature of a structure into a set of values (attributes) for that feature. The representation of a structure thus becomes a string of values. Henceforth structures and their associated representations will be considered equivalent.

A schema is a subset of structures having some common attributes. Certain attributes are fixed while the remainder may vary independently. The *length* of a schema is the distance between the two extreme

fixed attributes. The set of schemata decomposes \mathcal{A} into subsets (if k positions are fixed the schema could also be considered as a $1 - k$ dimensional hyperplane in \mathcal{A}). All schemata with the same attributes fixed partition \mathcal{A} . From the complementary viewpoint, each structure represents 2^{**1} schemata (for binary representations). Associated with each schema is a utility equal to the average utility of all structures that are instances of that schema.

A reproductive plan allocates tests to structures according to their relative fitness in the environment. Given a finite sample of structures $\mathcal{B}(t) = \{A_1(t), A_2(t), A_3(t), \dots\}$ at time t , the following is a basic algorithm for reproductive plans.

- 1) Select one structure from $\mathcal{B}(t)$ probabilistically, after assigning each structure a probability proportional to its observed performance at time t .
- 2) Copy the structure and apply operators to produce a new structure.
- 3) Either (3.1) select another element of $\mathcal{B}(t)$ (all elements equally likely) and replace it with the new structure, or (3.2) store the new element returning to (1) until the number of stored elements equals the number of elements in $\mathcal{B}(t)$ and replace all of them with the new structures.
- 4) Observe and record the performance of the new structure(s).
- 5) Increment t and return to (1).

Although the operators in a general reproductive plan can be any which satisfy the criterion of the framework, genetic operators confer the advantages of *intrinsic parallelism*, in that a great number of schemata are affected, and of compact storage and use of information resulting from previous observations of schemata. Three operators are necessary and sufficient to produce any structure: crossover, inversion, and mutation. To describe the operators, the following terminology is borrowed from genetics. A *gene* is a functional unit, the counterpart of a detector. An *allele* is a particular instance of a gene, analogous to a value for a detector. A *chromosome* is a string of genetic loci. A *locus* is a physical position on a chromosome. Loci which are closer together have a greater degree of *linkage*.

Crossover recombines alleles by exchanging segments between pairs of chromosomes. After two structures are chosen from the current population (in reproductive plans the choice is proportional to utility), an inter-locus boundary is randomly selected, and two new structures are formed by exchanging the segments to the right of the boundary. Crossover

affects schemata in two important ways: new instances (different contexts) of schemata already in the population are generated; and new schemata are generated. Each crossover event affects a large number of schemata (intrinsic parallelism), while each schema is affected independently.

Inversion alters the linkage between loci by changing the length of schemata. A new structure is formed by inverting the segment between two randomly selected locus boundaries in a chosen structure. This new structure (and the affected schemata) differs from the original one only in the order of the loci that the genes occupy. Inversion alone can produce all permutations of a structure and, like crossover, is intrinsically parallel.

Mutation modifies structures by randomly inserting new alleles for the genes at one or more loci. The probability of mutation at a locus is usually quite small. Mutation guarantees a finite expected time between the loss and reappearance of an allele in the population. Mutation is also a constant source of loss for any schema and can thus be viewed as a disturbance to prevent entrapment on local optima.

During the past decade's evolution of the theory of reproductive plans, several of Holland's students undertook experimental studies employing computer implementations of such plans. The study by Cavicchio² is of special relevance to the present work, and will be considered in the next section. Highlights of two other studies are mentioned here.

Bagley⁴ compared the performance of reproductive plans and correlation algorithms. A correlation algorithm is a prototype of many weight adjusting techniques such as that used in the Samuel Checker Player.³ Bagley demonstrated that the reproductive plan outperformed the correlation algorithm in most cases. He also appears to have been one of the first to suggest the use of a meta-adaptive plan to control values of parameters used by an adaptive plan.

De Jong's⁵ results encompass the most complete analysis of reproductive plans to date. To test the robustness of the plans, the environments consisted of continuous, discontinuous, convex, non-convex, unimodal, multimodal, low-dimensional, and high-dimensional functions. In his basic reproductive plan, De Jong varied four parameters: population size, crossover rate, mutation rate, and generation gap (the proportion of parents replaced by offspring). De Jong evolved several variations of his basic plan in an attempt to improve performance and reduce the allele loss rate, a problem arising from small populations and thus not an issue in Holland's general

theory. The *elitist* plan ensured that the best individual to date was kept in the population. The *expected value* plan reduced sampling error by forcing the actual number of offspring to be within ± 1 of the expected value. A combined *elitist expected value* plan performed better than any of the previous plans. De Jong also compared his reproductive plans with two function optimization methods. These methods performed much better than the plans for the functions for which they were developed, and much worse for the others.

3 CAVICCHIO'S META-PLAN

Cavicchio's² research was concerned with developing efficient and effective reproductive plans. His work is of special relevance to this study because he developed a meta-adaptive plan that adjusted the genetic operator rates during adaptation. The system that Cavicchio made adaptive was the pattern recognition scheme of Bledsoe and Browning.⁶ The environment was composed of two 16-letter alphabets which were represented by a 1 in every grid point that the letter touched, and 0 elsewhere. The system "learned" one of the alphabets (each letter of the alphabet was accepted as an example of that letter and was used for comparison purposes) and then tried to match correctly the letters of the other alphabet with this "learned" alphabet. The haploid individuals were strings of detectors (n -tuples ($n = 2, 3, \dots, 6$) of grid points). Each individual in the initial population had the same number of grid points randomly distributed in a variable number of genes (detectors). The pattern recognizer used the detectors to classify unknown letters as examples of known letters according to the number of matches made. To achieve a match between an unknown letter and a known letter the two letters had to coincide exactly on those grid points represented by an n -tuple.

Cavicchio's adaptive plan searched for a structure that could have correctly matched the unknown letters with the known ones. Each individual's utility was measured by its ability to classify the unknown letters correctly (actually a function of its ability to discriminate between the two likeliest candidates). Each generation was divided into the following phases: (1) preselection of offspring (from the previous generation) and selection of parents based on utility, (2) readjustment of genetic operator values using the meta-plan, (3) reproduction of the parents and formation of offspring using crossover, double crossover (interchange of two segments), inversion, and three

types of mutation, and (4) assignment of a utility to the offspring according to their fitness in the environment. (Mutation 1 replaced the complete n -tuple with n randomly chosen grid points. Mutation 2 changed one grid point in the n -tuple. Mutation 3 changed two genes into one or one gene into two depending on the number of grid points found in the gene to be mutated.)

Experimentally, Cavicchio developed some techniques to increase the effectiveness of adaptive plans. A major cause of ineffective search was decrease in population variance; many individuals began to contain a large proportion of common alleles. This problem was especially prevalent in the small (12 to 40 individual) populations Cavicchio used. A *selection scheme* limited the number of offspring a good parent could contribute, to prevent its subunits from dominating the population.

Cavicchio also developed a *preselection scheme* that prohibited an offspring from entering the current population unless it was superior to its parent (if formed by a mutation operator) or one of its parents (if formed by a crossover operator), even though it may have been better than some other member of the current population. If the offspring became a new population member, the inferior parent was removed from the current population.

To increase both efficiency and effectiveness Cavicchio developed a *parameter modification scheme* which altered the rates of application of the genetic operators. This scheme was designed to increase efficiency by allowing more large modifications (through crossover) early in the search, and to increase effectiveness by enhancing the contribution of small modifications (through mutation) later. Cavicchio wanted the crossover rate to decrease as the plan converged on an optimal structure, allowing a larger proportion of the offspring to be modified solely by mutation, thus facilitating a local search.

Subsequent investigators of reproductive plans have confirmed Cavicchio's intuitions that fixed rates of genetic operator application may be detrimental, although some results are at variance with his proposed modification scheme. For example, Bethke *et al.*⁷ found that high mutation rates at the outset produced rapid convergence and kept the population variance high. However as the population neared the optimal structure high mutation rates slowed convergence drastically, but still kept the variance high.

Thus even though there undoubtedly exist sets of genetic operator rates that are optimal for the entire adaptive period, an attempt to find these rates

would be counterproductive in at least two ways. First, every environment could require a different set of optimal rates. Second, the plan that uses the optimal set of rates is probably not as efficient or as effective as the same plan with modifiable rates, since optimal rates could fluctuate. The possibility exists that the modifiable plan could be better overall even if it did not use currently optimal rates at each modification step.

Any scheme that generates a sequence of genetic operator rates during adaptation is a *meta-plan*. Such a plan searches the meta-space (\mathcal{A}^1) to find the meta-structure, a string of rates, which interacts favorably with the meta-environment, the adaptive plan, (E^1). (Here and subsequently adaptive plan symbols are superscripted with 0, and meta-plan symbols with 1, unless no confusion could arise.)

Cavicchio's parameter modification scheme independently adjusted each current rate on the basis of its involvement in creating new population members. The function was: $P(t+n) = P(t) + [O(t+n) - P(t)]A$, where $P(t)$ is the rate at time t ; $O(t+n)$ is the observed frequency of the particular operator's involvement in creating new population members between times t and $t+n$; n is determined by requiring a certain number of new population members to be formed between modifications; and A is a term ($0 \leq A \leq 1$) which assures that $m \leq P(t+n) \leq 1$, where m is the minimum rate allowed for the particular genetic operator. Other modifications were used by the meta-plan whenever the adaptive plan began to stagnate. Two such adjustments were: (1) halving the number of new population members required for modification to occur, and (2) halving the distance between an operator's present rate and its minimum rate.

Cavicchio achieved significant gains by incorporating the above parameter modification scheme in the original reproductive plan. But even though the scheme was based on the reasonable assumption that $O(t+n)$ gave a better estimate of the optimal genetic operator rate than $P(t)$, there were a number of problems with the meta-plan. First, Cavicchio's meta-plan used a serial search which lacks the scope of a parallel search. Second, the search lacked generality because it was restricted to narrowly predefined modification methods. Third, searching the space of meta-structures depended on the number of new population members produced by the adaptive plan; very low or inappropriate rates thus changed slowly. Fourth, interactions between operators were not considered in calculating new rates. All operators were given equal credit (discredit) for the success (failure) of producing a new population

member. In fairness to Cavicchio, it should be noted that since his research was not primarily concerned with the development of a meta-plan, he attempted to solve only a few of the problems that his meta-plan engendered.

4 A REPRODUCTIVE META-PLAN

The genetic operator rate modification scheme proposed here is a *reproductive* meta-plan. It is applied to Cavicchio's reproductive plan for several reasons. First, the operation of his meta-plan can be used as a benchmark for the performance of the proposed meta-plan. Second, difficulties with Cavicchio's meta-plan can be reassessed in the context of the new meta-plan. Third, Cavicchio used an environment requiring a lengthy period of adaptation, which facilitates the study of a meta-plan. The ease with which this environment can be modified slightly also facilitates study of the effects of changing environments on both types of meta-plan.

In the operation of the meta-plan proposed here, an M/N population specifies the maximum size of the population (M), the number of individuals kept as potential parents (N), and the number of offspring produced at each generation ($M - N$). The N best individuals are the potential parents. The selection scheme chooses parents from these N individuals (based on utility) and grants a reproduction quota to each parent (based on relative utility) such that the sum of the individual quotas is $M - N$. Each parent can reproduce a maximum of two offspring per generation and a maximum of three offspring over all generations. The latter condition is relaxed whenever the sum of the reproduction quotas is less than $M - N$ and all parents have been allotted their maximum quotas.

In the reproductive meta-plan experiments reported in the next section, the preselection scheme and the reproductive phase of the original reproductive plan are not changed from those described in Section 3. Cavicchio's selection scheme is modified so that the sum of the reproduction quotas will be $(M - N)/n$, where n is the number of meta-structures. Each meta-structure will use the same parents to produce a subset of the offspring. Although this may not be the best selection method, the main concern is to see how different meta-structures operate. The selection method used will remove some biases that could occur if different parents were used. A utility is assigned to each meta-structure according to how well its subset of offspring performs.

A meta-structure is a string of values representing rates (probabilities) of application of crossover, double crossover, and the three types of mutation. The utility, I^1 , of a meta-structure is the number of possible new population members (in the meta-environment) found in the meta-structure's subset of offspring. The term "possible" is used since this number is obtained irrespective of the other subsets of offspring. When the new population members are chosen by the adaptive plan all offspring compete; therefore some possible new population members may not survive. Once the meta-structures have been ranked according to utility, some subset of best structures is selected to reproduce in the next generation.

Despite many plausible alternatives, the above utility function was chosen because the number of possible new population members created by each meta-structure is a reasonable estimate of the ability of each meta-structure to help generate an efficient and effective search. This utility implicitly defines a criterion to estimate the goodness of the individuals produced by the adaptive plan. It seems reasonable to rank the meta-structure higher (in the sense of generating a better search) if it is used to create many "average" new population members rather than a few exceptional ones. (A modification of the function was actually introduced after some experimentation, as described in Section 5.)

The meta-plan \mathcal{T}^1 generates the next set of meta-structures, attempting to augment the efficiency and effectiveness of the reproductive plan. New meta-structures should provide a wide range of rates for each genetic operator (allowing quick response to changes in the environment, E^0) and converge on an optimal set of genetic operator rates.

The meta-plan uses one type of genetic operator that has not been previously discussed, called frame-shift. Frame-shift is an insertion or deletion of a basic building block in an allele. If an operator rate is viewed as a quotient of binary numbers then a deletion of the rightmost numerator bit and an insertion of a new leftmost numerator bit would be two frame-shifts. If the denominator is all 1's then a 0 inserted in the leftmost bit after a rightmost deletion would be equivalent to division by two (with rounding up).

The meta-plan forms new genetic operator rates in the meta-structures using multiple crossing over with frame-shifts (operators 2, 3, 4, 5 below) and without (operator 6), and two types of mutation (operators 1, 7). One of the meta-operators is chosen at random to generate each new genetic

operator value in each new meta-structure, using $v(1)$, the operator rate in the best meta-structure, and $v(2)$, the rate in the second best meta-structure. The meta-operators are:

- (1) $v = (v(1) + v(2))/2$,
the average of the rates,
- (2) $v = (v(1) + 1.0)/2$,
half way between $v(1)$ and 1.0,
- (3) $v = (v(1) + 0.0)/2$,
half way between $v(1)$ and 0.0,
- (4) $v = (v(2) + 1.0)/2$,
half way between $v(2)$ and 1.0,
- (5) $v = (v(2) + 0.0)/2$,
half way between $v(2)$ and 0.0,
- (6) $v = v(1)$,
the best rate,
- (7) $v =$ a random rate.

The criteria used to compare Cavicchio's parameter modification scheme and the proposed meta-plan involve two attributes of the meta-plans: (1) the ability of the meta-plans to search the meta-space in certain situations, and (2) the rate of increase in utility of \mathcal{B}^0 .

5 EXPERIMENTAL RESULTS

This section presents results of experiments with computer implementations of both meta-plans, and proposes some conclusions based on these results. In what follows, "meta-plan with (without) preselection" is to be interpreted as "the meta-plan applied to the *reproductive plan* that uses (does not use) the preselection scheme."

Two experiments were run using Cavicchio's meta-plan, one with and one without his preselection scheme. These studies provided some benchmarks for the reproductive meta-plan, and also some indication of the decrease in utility that can be encountered when preselection is used with Cavicchio's meta-plan. Three experiments tested the reproductive meta-plan with preselection. One objective was to determine how changing E^0 affected both meta-plans. Environmental changes occurred when 250 or 260 offspring had been produced (unknown letters were shifted one column right) and when 500 offspring had been produced (unknown letters were shifted one column left). Finally, a special experiment was run to test the performance of Cavicchio's meta-plan with preselection at a "crisis point."

Figures 1 and 2 show the changes in genetic operator rates for a period of 75 generations for the present

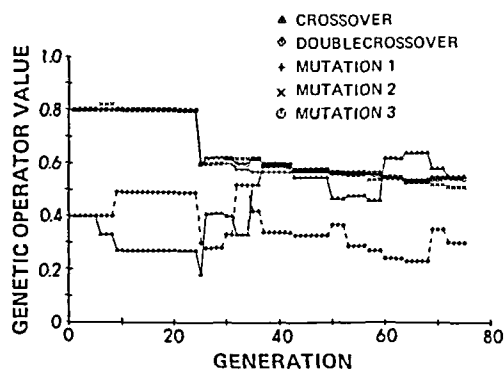


FIGURE 1 Genetic operator values of Cavicchio's meta-plan, with preselection, for 75 generations.

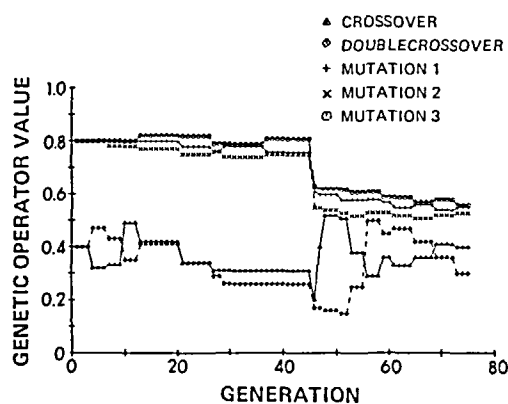


FIGURE 2 Genetic operator values of Cavicchio's meta-plan, without preselection, for 75 generations.

implementation of Cavicchio's meta-plan with and without preselection respectively. Figure 3 shows the utility curves for both experiments. In this and all subsequent graphs of utility, the average utility of the 10 best structures is plotted.

In replication of Cavicchio's findings the crossover rates in Figures 1 and 2 show no trends (except that they tend to shift in opposite directions when modified), while the mutation operators converge to rates between 0.5 and 0.6. (Cavicchio's results show convergence to slightly lower rates; but he ran his experiments for over 200 generations and did not modify the environment.) Figure 3 shows that the reproductive plan which does not use preselection outperforms the plan that does, for the first 25 generations and, if the amount of increase in utility is considered, the last 25 generations as well.

Two observations can be made about these results.

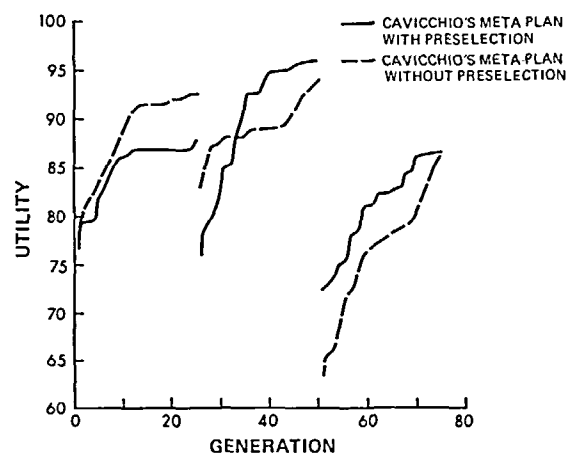


FIGURE 3 Average utility of 10 best structures of Cavicchio's meta-plan with and without preselection.

First, the meta-plan seems to be "unaware" of any changes in the environment. Since the only method for generating new detectors in reproductive plans is mutation, it would seem reasonable to expect environmental changes to result in increased mutation rates, followed by an increase in both crossover rates. This procedure would inject new alleles into the population and then employ a high rate of recombination to search for the structure that performs best. The steady decline in mutation rates thus seems particularly inappropriate.

Second, the meta-plan is unable to cope adequately with periods of stagnation in the reproductive plan's search. Because a minimum number (N) of the new population members must be generated before the genetic operator rates can be modified, these periods of stagnation can cause a loss of efficiency. The method used by the meta-plan to escape this predicament is to halve the required number of new population members. However this action is only taken after the increase in utility over the past 15 generations drops below 3.0. In this experiment that is much too late. At the same time the genetic operator rates are also reduced. (Cavicchio decided that stagnation of the reproductive plan's search meant the operator rates were too high.) In both instances (Figure 1 generation 25, and Figure 2 generation 46) this appears to be the wrong strategy, since in both cases the meta-plan increases the (single) crossover rate in the following generation.

Some indication of the problems with Cavicchio's meta-plan (mentioned at the end of Section 3) has become apparent in these two experiments. The lack of generality is exemplified when the meta-plan

reduces the two crossover rates during periods of stagnation, although they should have been increased as the subsequent modification indicates. Also, requiring a minimum number of new population members to be generated before modification of operator rates causes a slow and limited search. Even though the meta-plan halves the number required, this process can happen too late. The second experiment loses 9 generations because the meta-plan is unable to generate new genetic operator rates during this period of time.

In all experiments with the reproductive meta-plan the meta-population consisted of 5 meta-structures, of which the best 2 remained as meta-parents in the next meta-generation. The 5 crossover-double crossover combinations were initialized to (0.5, 0.5), (0.4, 0.4), (0.3, 0.3), (0.2, 0.2), and (0.1, 0.1), inversion rates were set at 0.5, and each mutation rate was set to give an average of 2 mutations per structure per generation (which appears as a numerical value of 0.8 in the results below because of Cavicchio's involved scheme for applying this parameter). Table I shows the best two (parent) meta-structures in each meta-generation (and the entire starting population) for an experiment involving a 60/10(10) population. (In an $M/N(p)$ population each of the n meta-structures is used by the reproductive plan to generate p offspring. Thus $pn = M - N$.) This experiment was run for 5 generations and revealed the variations in genetic operator rates that can occur.

The 60/10(10) population was used as a test case to see if different combinations of genetic operator application rates would in fact produce a significantly different number of possible new population members. As the results tabulated in Table II indicate, certain combinations produce a significantly larger number of (possible) new population members.

Populations of this size are not economical for simulation because there are many more (possible) new population members generated than openings for new population members in the next generation. Those that do not become parents in the following generation are lost forever, because the 10 best individuals are replaced almost every generation.

TABLE II

Number of (possible) new population members/meta-structure and a summary of the average utility of the 10 best structures in a 60/10(10) population

Meta-generation	Average utility (10 best population members)	Number of (possible) new population members/meta-structure				
		1	2	3	4	5
1	76.8	4	5	3	5	5
2	82.8	1	3	1	1	2
3	84.5	3	3	3	3	3
4	86.6	2	1	2	3	5
5	87.9	1	1	0	1	2

TABLE I

Initial meta-population and meta-parents for the remaining meta-generations of an experiment using a 60/10(10) population

Meta-generation	Meta-structure	Cross-over	Double cross-over	Inversion	Mutation		
					1	2	3
1	1	0.50	0.50	0.50	0.80	0.80	0.80
	2	0.40	0.40	0.50	0.80	0.80	0.80
	3	0.30	0.30	0.50	0.80	0.80	0.80
	4	0.20	0.20	0.50	0.80	0.80	0.80
	5	0.10	0.10	0.50	0.80	0.80	0.80
2	1	0.40	0.40	0.50	0.80	0.80	0.80
	2	0.20	0.20	0.50	0.80	0.80	0.80
3	1	0.20	0.20	0.50	0.80	0.80	0.80
	2	0.60	0.40	0.75	0.90	0.80	0.90
4	1	0.20	0.20	0.50	0.80	0.80	0.80
	2	0.60	0.40	0.75	0.90	0.80	0.90
5	1	0.60	0.88	0.88	0.45	0.40	0.85
	2	0.40	0.10	0.46	0.90	0.90	0.90

Further, since the results will be compared on an offspring basis with the data produced by the experiments with Cavicchio's plan, this meta-plan would be at a great disadvantage. Since the generations are in a 5 to 1 ratio, the reproductive plan using Cavicchio's meta-plan produces 40 out of 50 offspring using successively better sets of parents. Most of the utility gain is attributable to the increase in parental utility, not to good genetic operator rates. Table II reveals the relatively poor performance of this population size when compared with Cavicchio's meta-plan without preselection at generation 25.

For reasons of economy and fairer comparison it was decided to use a smaller population (30/10(4)) in the remaining experiments. Figure 4 shows some surprising results concerning the best meta-structure.

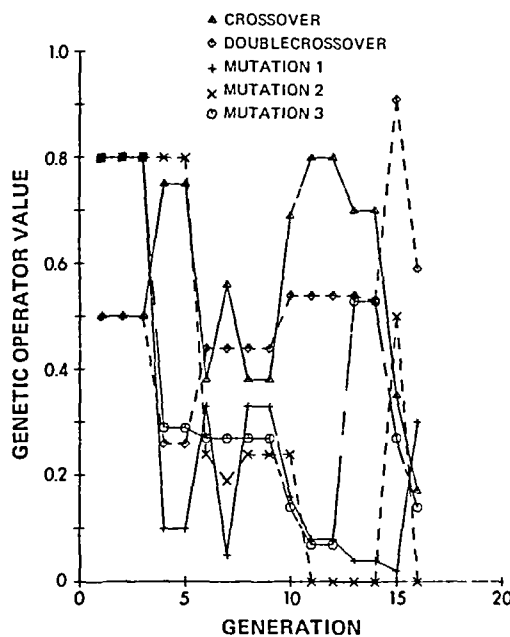


FIGURE 4 Genetic operator rates of original reproductive meta-plan for 16 generations.

(This and all subsequent presentations of genetic operator rates will be concerned with the best meta-structure only.) As the population of structures approaches a matching score of 100 (see the utility rates for this experiment in Figure 5) the crossover rates stay very high and the mutation rates drop in some cases to 0. This seemingly inappropriate action could have a good explanation. The population at this point has been partitioned into four distinct groups.

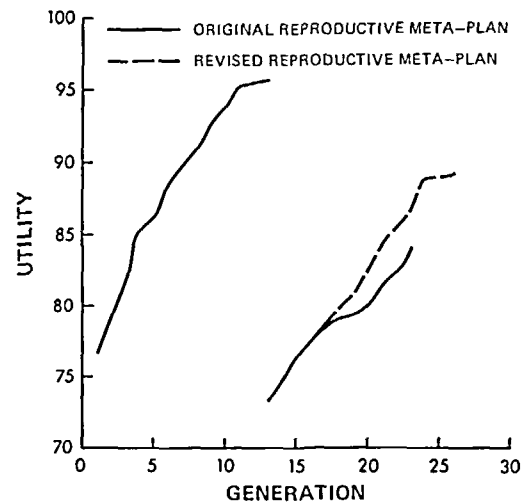


FIGURE 5 Average utility of 10 best structures for original and revised reproductive meta-plans for 27 generations.

Each partition includes a number of individuals with minor variations in genetic content. The effects of crossover and double crossover may have become similar to those of mutation. The pieces exchanged between two structures are so similar that the effect is like mutation, except that it has a better than random chance of improving individuals. Apparently, the "problem" Cavicchio associated with late high crossover rates may in some cases be an advantage.

Figure 5 reveals a comparatively worse response to the environmental change (at generation 14) by the reproductive meta-plan than by Cavicchio's meta-plan with preselection. One explanation is that at this point the meta-plans are working with different meta-environments. With Cavicchio's meta-plan θ^0 retained much more of its initial variance than with the reproductive meta-plan. (In the latter case, the reproductive plan was working with only 82 different alleles, out of a possible 430.)

An oversight in the design of the reproductive meta-plan is suggested by these results. When more than two meta-structures generated the same number of (possible) new population members, these ties were broken by taking the first two from the list of meta-structures that were tied. The current meta-parents were thus retained if none of the meta-offspring were better. However, as the instance presented in Table III shows, a meta-structure involved in a tie can be discarded even though it is used to generate better offspring.

The reproductive meta-plan was therefore revised

TABLE III

Meta-generation 17 showing the selection method for the original and revised reproductive meta-plans

Rank of A^1 at time t	Number of possible new population members produced using A^1	Utility of possible new population member(s)	Rank of A^1 at time $t + 1$ (original meta-plan)	Rank of A^1 at time $t + 1$ (revised meta-plan)
1	2	77.7 77.3	1	2
2	1	77.3	—	—
3	2	77.3 77.3	2	—
4	0	—	—	—
5	2	78.1 77.8	—	1

so that the meta-structure chosen in case of a tie was the one which produced the structure with the highest utility. Table III shows the first instance (meta-generation 17) of a different choice of meta-structures. In the original reproductive meta-plan the meta-structure that is labelled "1" was never replaced for the remainder of the experiment even though it was involved in a tie in half of the remaining meta-generations. In the revised meta-plan however this meta-generation was the beginning of a trend towards higher mutation rates.

The resulting gains in utility obtained by this revised reproductive meta-plan are evident in the results presented in Figure 5. A significant improvement is apparent after the first few generations following the environmental change. The increase in utility is not as large as for Cavicchio's meta-plan with preselection, possibly because the reproductive plan using the reproductive meta-plan starts with a tremendous disadvantage (the low population variance). But the revised reproductive meta-plan makes a somewhat greater gain in utility than does Cavicchio's meta-plan without preselection. The mutation rates, shown in Figure 6, increase as expected (after a slight time lag), possibly showing the meta-plan's "awareness" of the problem at hand. The meta-plan causes the reproductive plan to generate new genetic material in response to the "crisis."

The other four meta-structures show some patterns in their genetic operator rates. Since meta-structures 3, 4 and 5 are new for every meta-generation the individual genetic operator rates tended to be random *except* when the particular rates in the parent meta-structures (1 and 2) are similar. In this case the rate is biased slightly towards the rate in the parents, since one of the meta-operators averages the two

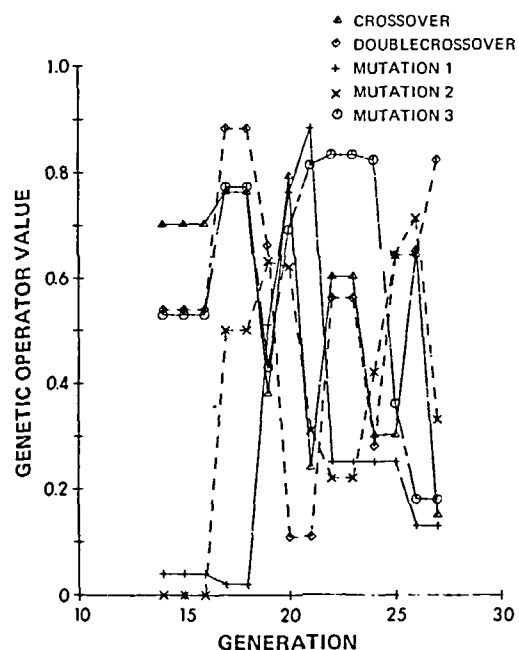


FIGURE 6 Genetic operator rates of revised reproductive meta-plan for 14 generations.

parental rates. And the rate is even more biased when the rates in both meta-parents are close to 0 or 1, since four meta-operators force the rates closer to 0 or 1. The patterns of rates for meta-structure 2 generally follow those for meta-structure 1, usually with some time lag. This similarity may be attributed to two factors: (1) meta-structure 1 occasionally was replaced but was good enough to be retained as the second parent; and (2) an offspring of meta-structure 1 sometimes became meta-structure 2. Although this discussion emphasizes the "drift" of the meta-population in the same direction as the best meta-structure, the equally important ability of the meta-plan to explore different areas of the space of meta-structures coexists with the ability to follow the best meta-structure's trends, because some of the meta-operators force the search in different directions.

Throughout Cavicchio's discussion of the merits of modifiable genetic operator rates he frequently observed that "reasonable" operator rate settings were appropriate at most times. In the following discussion the exceptions will be called "crisis points." The crisis point of main concern to Cavicchio occurred when the reproductive plan neared its potential capability. He believed that unduly high crossover rates were destroying potentially good mutations. His

meta-plan successfully lengthened the reproductive plan's effectiveness even though crossover rates were not actually decreased. One crisis point not considered by Cavicchio was a change in the environment, E^0 . In the present experiments run using Cavicchio's meta-plan all instances of environmental changes occurred when the genetic operator rates were still reasonably high. The worst case that could face the meta-plan at an environmental change is if the genetic operator rates were low.

An experiment was therefore devised to test the effects of this crisis point on Cavicchio's meta-plan. The crossover and inversion rates were set at the 0.1 minimum, the mutation rates at the 0.2 minimum. The population of structures that the reproductive plan used was the population that existed at generation 50 in the second experiment above. Figure 7 shows that Cavicchio's meta-plan is incapable of coping with this crisis point. There is no increase in

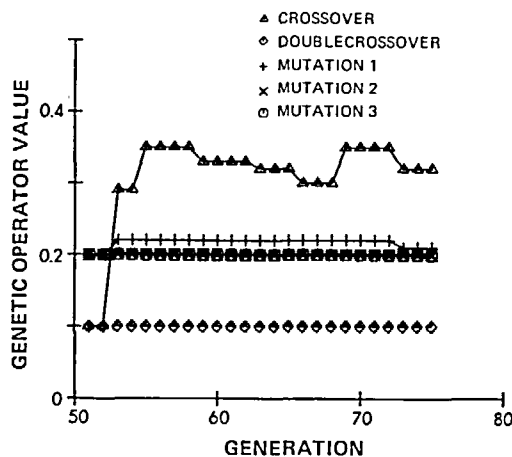


FIGURE 7 Genetic operator rates of Cavicchio's meta-plan with preselection, at a crisis point.

mutation rates, a step which is quite important after an environmental change. Figure 8 shows the inability of this meta-plan to increase the utility of the population as quickly as the same meta-plan that starts with more "reasonable" genetic operator rates.

During the experiment with the reproductive meta-plan, the revised reproductive meta-plan proved itself capable of adjusting to a similar "crisis" point (see Figures 5 and 6). In this case the initial rates for mutations 1, 2 and 3 were approximately 0, 0 and 0.5 respectively. Not only were these rates low but the population had also lost nearly all of its variance. Yet the meta-plan was able to raise the mutation rates enough to allow the reproductive plan to

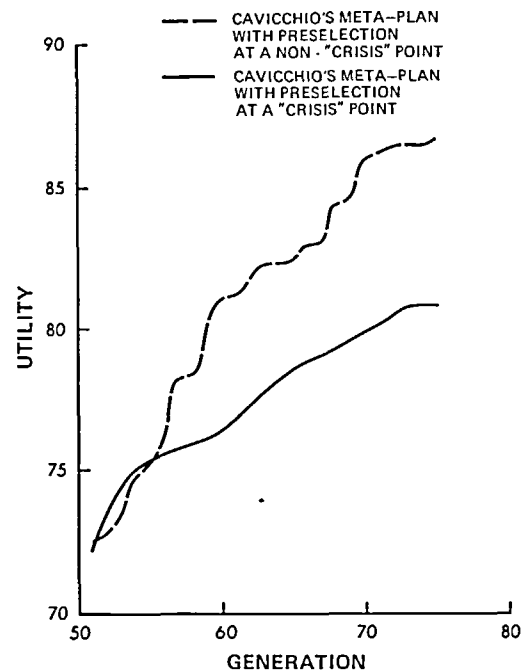


FIGURE 8 Comparison of average utility of Cavicchio's meta-plan, with preselection, at a crisis point and a non-crisis point.

increase the population variance and to adapt significantly, given the initial problem of low variance.

Figure 9 summarizes some of the results obtained from the experiments described above. The x-axis has

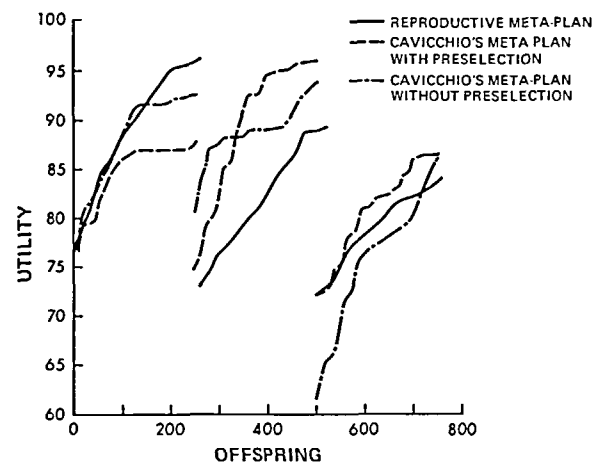


FIGURE 9 Summary of average utility of the 10 best structures for Cavicchio's meta-plan with and without preselection and the original and revised reproductive meta-plans.

been changed from generations to number of offspring since the reproductive plan produces 20 offspring/generation with the reproductive meta-plan but only 10 with Cavicchio's meta-plan. The utility curves for Cavicchio's meta-plan are the same as those in Figure 3. The curve for the reproductive meta-plan is the same as Figure 5 during phase 1 (between 0 and 260 offspring). Phase 2 (between 260 and 520) shows the performance of the revised reproductive meta-plan (also from Figure 5). The final phase (the last 260 offspring) represents the results of an experiment which used the revised reproductive meta-plan, starting with the same meta-environment and genetic operator rates as Cavicchio's meta-plan with preselection used at generation 51.

As previously discussed the reproductive meta-plan outperforms the non-reproductive meta-plan during the first phase. In the second phase it appears that the reproductive meta-plan outperforms the non-reproductive meta-plan without preselection despite the initial difficulty of a population with little variance. However in the third phase the reproductive plan does very poorly. The reproductive plan seems to suffer more from the initial low variance and thus does not achieve the sorts of gains obtained by Cavicchio's meta-plan in this phase. This problem might be solved by changing the selection scheme used, as discussed in the next section.

It appears that there is an overall improvement in searching the space of meta-structures when using a reproductive meta-plan. Initial constraints placed on the non-reproductive meta-plan that cause poorer results show the importance of generality in reproductive plans. For example the reliance of the non-reproductive meta-plan on the number of new population members was one cause of a limited search by the meta-plan. And a lack of generality is apparent when the non-reproductive meta-plan decreases operator values whenever there is a stagnation of the adaptive plan.

The first differences plotted in Figure 10(b) and Figure 10(c) better reveal this lack of generality. In most cases the peaks in the first difference curves for Cavicchio's meta-plan coincide with a change in the genetic operator rates. This is followed by an immediate drop in utility gain. Note that a large increase in utility is more often the cause rather than the effect of a modification of the operator rates.

Even though there are peaks in the first differences of the reproductive meta-plan (Figure 10(a)), these peaks usually are not followed by a dramatic decrease which is the more common occurrence in Figures 10(b)

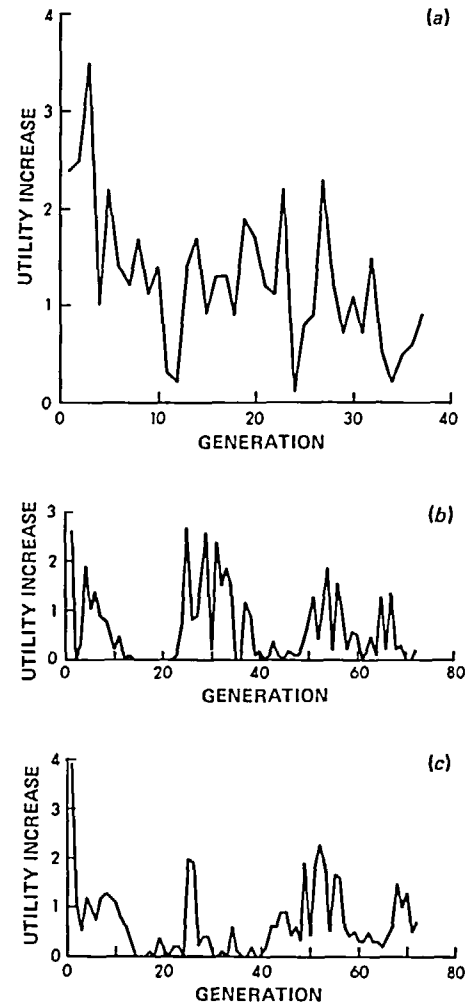


FIGURE 10 First differences of the average utilities summarized in Figure 4.9. (a) Reproductive meta-plan. (b) Cavicchio's meta-plan with preselection. (c) Cavicchio's meta-plan without preselection.

and (c). (The peaks immediately following the environmental change are normal because the greatest utility increase usually occurs when the utility is low.) Figure 10(a) does reveal a steady increase in utility caused by the use of nearly optimal genetic operator rates every generation with the reproductive meta-plan, instead of the erratic increases evident in Figures 10(b) and (c) caused by genetic operator rate modification every three generations (approximately) with Cavicchio's meta-plan.

Another improvement associated with the reproductive meta-plan is that it seems to be "aware" of environmental changes. The crisis point experiment shows that

the non-reproductive meta-plan is almost totally unable to increase the mutation rates. The reproductive meta-plan on the other hand is able to modify the operator rates so that genetic variance increases significantly in the second phase.

6 CONCLUDING REMARKS

With emphasis on the superiority of the reproductive meta-plan at crisis points, which are the situations most detrimental to the first level reproductive plan, and on the lack of significant overall superiority of either meta-plan in non-"crisis" situations, it would be fair to say that the reproductive meta-plan developed here is an improvement over Cavicchio's meta-plan. Despite this improvement there remain some significant problems with the results obtained and with the reproductive meta-plan itself.

As a result of the extremely high cost of simulation, all of the experiments discussed in Section 5 were run only once. Each experiment should ideally have been replicated a number of times with different initial populations and different parameter settings, so that more reliable performance estimates could have been obtained. A dedicated computer would be convenient for such lengthy studies. A microprogrammable machine would be ideal, since it could be configured to appear like a population of structures and to have high level instructions like "REPRODUCE," and "MUTATE."

A second suggestion is that, under the reproductive meta-plan, the selection method used by the reproductive plan should allow the $M - N$ offspring to be allocated to all N parents. The present selection method allocates p offspring to some subset of the N parents, a subset which is used n times, once for each meta-structure. As mentioned previously, this selection method is probably what caused the considerable loss in population variance. One of the main purposes of Cavicchio's selection scheme was to limit the number of offspring of any parent. But in the case of the reproductive meta-plan a parent in the reproductive plan can contribute to $2n$ offspring. Also, if each parent has a reproduction quota of one, then in Cavicchio's reproductive plan each parent contributes to one offspring. Under the reproductive meta-plan each of the p parents contributes to n offspring.

The reason for using the same set of parents for each meta-structure was to keep any biases attributable to parental "goodness" from entering into the choice of the next meta-generation's meta-parents. If meta-

structures could be associated with different parents, the development of a different feedback function to counteract the biases attributable to parental "goodness" would be required. One approach would be to change the ratio of generations to meta-generations from 1-to-1 to 2-to-1 or more, thus increasing the likelihood of the meta-structures being compared on a less biased choice of parents. However, once the meta-generation is lengthened the meta-plan becomes less sensitive to meta-environmental changes.

Other facets of the reproductive meta-plan that do not arise from the above problems could also be studied. For example, biasing the number of offspring produced using a meta-structure according to whether or not it is a meta-parent could result in some improvement. Cavicchio's meta-plan uses the "best" set of genetic operator values to produce all of the offspring. On the other hand the reproductive meta-plan uses the superior meta-structures to produce fewer than half of the offspring and the other meta-structures (which could be inferior) to produce the rest.

Finally, it is tempting to consider the possibility of meta-meta-adaptation. For example, the seven meta-operators, used equiprobably in this study, could have their usage frequencies under the control of a reproductive plan. Although there is no theoretical limit to the number of levels of adaptation, the extra computational costs would quickly begin to outweigh significantly any possible performance improvements, perhaps even at the third level.

ACKNOWLEDGEMENTS

The authors are grateful to Drs. K. Morgan, L. Schubert and K. Wilson of the University of Alberta for their many helpful suggestions. This research was supported in part by the National Research Council of Canada, through Grant A-7643 to J. R. Sampson.

REFERENCES

1. J. H. Holland, *Adaptation in Natural and Artificial Systems* (Ann Arbor: The University of Michigan Press, 1975).
2. D. J. Cavicchio, "Adaptive search using simulated evolution," *Technical Report 03296-4-T* (Computer and Communication Sciences Department, University of Michigan, 1970).
3. A. L. Samuel, "Some studies in machine learning using the game of checkers," In E. A. Feigenbaum and J. Feldman (eds.), *Computers and Thought* (New York: McGraw-Hill, 1963).

4. J. D. Bagley, *The Behavior of Adaptive Systems Which Employ Genetic and Correlation Algorithms* (Ph.D. Dissertation, University of Michigan, 1967).
5. K. A. De Jong, "Analysis of the behavior of a class of genetic adaptive systems," *Technical Report 185* (Logic of Computers Group, University of Michigan, 1975).
6. W. W. Bledsoe, and I. Browning, "Pattern recognition and reading by machine," In *Proceedings of the Eastern Joint Computer Conference* (1959).
7. A. D. Bethke, B. P. Ziegler and D. M. Strauss, "Convergence properties of simple genetic algorithms," *Technical Report 159* (Logic of Computers Group, University of Michigan, 1974).