Crossover or Mutation?

William M. Spears
Navy Center for Applied Research in Artificial Intelligence
Code 5510
Naval Research Laboratory
Washington, D.C. 20375-5320
E-mail: SPEARS@AIC.NRL.NAVY.MIL

Abstract

Genetic algorithms rely on two genetic operators - crossover and mutation. Although there exists a large body of conventional wisdom concerning the roles of crossover and mutation, these roles have not been captured in a theoretical fashion. For example, it has never been theoretically shown that mutation is in some sense "less powerful" than crossover or vice versa. This paper provides some answers to these questions by theoretically demonstrating that there are some important characteristics of each operator that are not captured by the other.

1 INTRODUCTION

One of the major issues in genetic algorithms (GAs) is the relative importance of two genetic operators: mutation and crossover. In the 1960's, L. Fogel *et al.* (1966) illustrated how mutation and selection can be used to evolve finite state automatons for a variety of tasks. Simultaneously, in Europe, Rechenberg (1973) investigated "evolution strategies" that again concentrate on mutation as the key genetic operator. Sophisticated versions of evolution strategies, with adaptive mutation rates, proved quite useful for function optimization tasks (Baeck *et al.*, 1991; Schwefel, 1977). Recent studies confirm this view, illustrating the power of mutation (Schaffer *et al.*, 1989). D. Fogel has continued the earlier work of L. Fogel and makes an even stronger claim - that crossover has no general advantage over mutation (Fogel & Atmar, 1990).

On the other hand, proponents of the Holland (1975) style of genetic algorithm believe that crossover is the more powerful of the two operators. Considerable effort has been spent in analyzing crossover and its effects on performance (e.g., De Jong, 1975; Spears & De Jong, 1991; Vose & Liepins, 1991). In most of these analyses mutation is considered to be a background operator and of secondary importance. To support these views, experimental results have been presented, illustrating the power of crossover (e.g., De Jong, 1975). Most recently, Schaffer & Eshelman (1991) empirically compare mutation and crossover, and conclude that mutation alone is not always sufficient.

Unfortunately, empirical comparisons can often be disputed or may be misleading. For example, Schaffer & Eshelman speculate that implementation and representation may explain Fogel's results. Similarly, it can be speculated that Schaffer & Eshelman did not implement mutation reasonably (e.g., with an adaptive rate). To date, there has been no theoretical justification to support either camp's beliefs. It has never been theoretically shown that crossover is in any sense more powerful than mutation, or that mutation is more powerful than crossover. Similarly, no theoretical basis exists for supposing that both operators are necessary and perform different roles within the GA.

In this paper we show that, in a general sense, both camps are correct, although we dispute the stronger claim of Fogel and Atmar. We define two potential roles of any genetic operator, disruption and construction, and consider how well mutation and crossover perform these roles. Our results show that in terms of disruption, mutation is more powerful than crossover, although it lacks crossover's ability to preserve alleles common to individuals. However, in terms of construction, crossover is more powerful than mutation.

2 DISRUPTION THEORY

Holland provided the initial formal analysis of the behavior of GAs by showing how they allocate trials in a near optimal way to competing low order hyperplanes if the disruptive effects of the genetic operators is not too severe (Holland, 1975). Considerable attention has been given to estimating the *disruption rate* of crossover, i.e., the probability that a particular application of crossover will be disruptive. As has been pointed out, however, another important consideration is not just how *often* a sample will be disrupted, but *how* it will be disrupted (Eshelman *et al.*, 1989). In this section we will first consider a theory of disruption rates for crossover, and show how we can compare this with a disruption rate theory for mutation. We then briefly review both mutation and crossover with respect to how they disrupt hyperplane samples.

2.1 DISRUPTION RATES

Holland's initial analysis of the sampling disruption of 1-point crossover (Holland, 1975) has been extended to n-point crossover and a parameterized (P_0) uniform crossover (De Jong, 1975; Syswerda, 1989; Spears & De Jong, 1991), where n is the number of crossover points and P_0 represents the probability of swapping alleles between two parents. These results estimate the likelihood that the sampling of a kth-order hyperplane (H_k) will be disrupted by a particular form of crossover.

For example, given a 3rd-order hyperplane (H_3), one can compute the probability that an application of n-point or uniform crossover will disrupt the sampling of that hyperplane.

It turns out to be easier mathematically to estimate the complement of disruption, the likelihood that a hyperplane sample will survive crossover, which we denote as P_s. We can also refer to the survival and disruption of individuals within a hyperplane H_k. If an individual survives with respect to H_k, it remains within H_k. If an individual is disrupted with respect to H_k, it is no longer within H_k. Finally, it should be noted that if each application of crossover is independent, we can interpret P_s as the probabilistic survival of an individual within a hyperplane H_k. For example, if N individuals are within some H_k, we expect roughly $N{\cdot}P_s$ individuals to remain (survive) in H_k after crossover.

Figure 1 illustrates P_s for 3rd-order hyperplanes. For n-point crossover the probability that a sample will survive depends on the order of the hyperplane, its defining length, and the number of crossover points n. For uniform crossover the probability of survival depends on the order of the hyperplane and the probability of swapping alleles, P_0.[1] The reader should note that uniform crossover is labelled as "P_0 uniform" in Figure 1. For example, ".1 uniform" indicates that $P_0 = 0.1$.

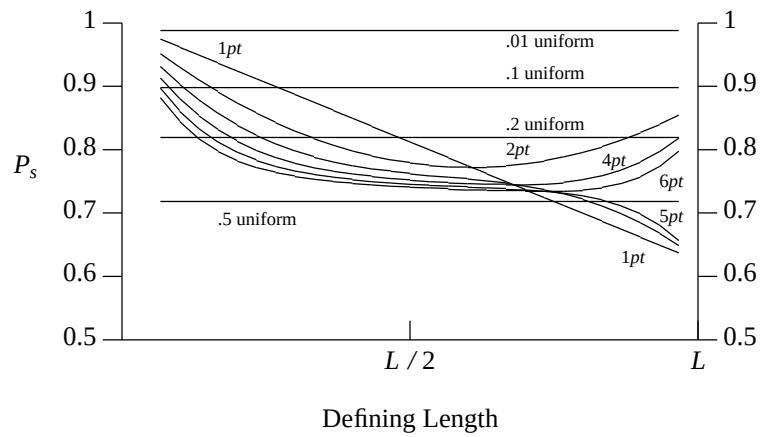

Figure 1: Crossover Survival, 3rd-Order, $P_{eq} = 0.5$

These results are time dependent in the sense that they are affected by the degree to which the population has converged. To see this, we need to define what we mean by convergence. In our theory we denote $P_{eq}(d)$ to be the probability that two hyperplanes will have the same allele at a particular defining position d (De Jong & Spears, 1992). As a useful simplification we also denote P_{eq} to be the average of the $P_{eq}(d)$'s over all d. Since we assume a bit level representation, $P_{eq} = 0.5$ represents the condition when the population is first randomly initialized and each allele has an equal probability of being a 1 or 0. When P_{eq} is close to 1, the population is nearly converged and lacks diversity. Figure 2 illustrates how crossover is affected by the convergence of the population. The horizontal axis represents the convergence of the population ($0.5 \le P_{eq} \le 0.9$). For the sake of simplicity we illustrate only uniform crossover, where P_0 ranges from 0.1 to 0.5 in increments of 0.1. These values are useful because the levels of disruption provided

[1] See Spears & De Jong (1991) for more precise details.

by n-point crossover are roughly bounded by the disruption levels of uniform crossover when $0.1 \le P_0 \le 0.5$ (see Figure 1). Note that disruption decreases as the diversity of the population decreases.

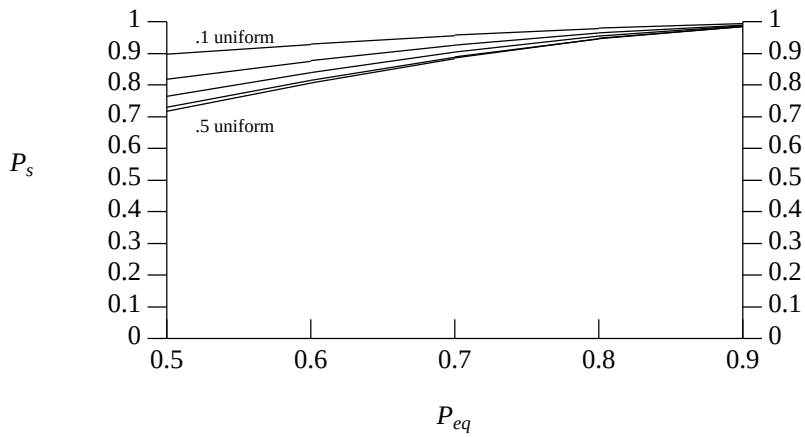

Figure 2: Crossover Survival as a Function of Convergence, 3rd-Order

In the previous paragraphs we provided a review of disruption rate theory for n-point and P_0 uniform crossover. Can we now provide a similar theory for mutation? At first blush this would appear to be difficult, since crossover is a function of two individuals, while the mutation of one individual is not affected by the mutation of another. More precisely, how can P_s represent the probability of survival of one individual within a hyperplane H_k, given that two individuals are involved in the crossover operation?

The answer lies in the (often hidden) assumption that crossover is used to create two offspring, as opposed to one. Consider the situation where one parent individual is within H_k, while the other parent individual is not. Then, after crossover, at most one offspring will also be within H_k, and P_s represents the probability of that event. Equivalently, if there are N individuals within H_k, there will be roughly $N \cdot P_s$ individuals within H_k after crossover. Suppose, however, that both parents are in H_k. Then, after crossover, both offspring are guaranteed to be in H_k. In this case $P_s = 1$ and $N \cdot P_s = N$ individuals will survive crossover. If only one offspring were created, this analysis would not be correct. In summary, the assumption that both offspring are created is necessary to ensure that P_s correctly represents the independent survival of one individual within a hyperplane.

Since, for crossover, P_s really represents the probability of one individual within a hyperplane H_k surviving, we can compare this with a similar analysis for mutation. For mutation we want P_s to represent the probability that an individual in H_k will survive mutation. In this case, independence is trivial, since the mutation of one individual is not affected by the mutation of another. Again, if there are N individuals within some H_k, roughly $N \cdot P_s$ individuals will survive mutation. In this paper mutation is defined to be the operator that probabilistically selects a bit and flips that bit (recall that we are assuming a bit level representation). Again, suppose we have a 3rd-order hyperplane. Then the probability that an individual within that hyperplane will survive mutation is

given by:

$$P_s(H_3) = (1 - P_m)^3$$

where P_m is the probability of mutating an allele. In general, we have:

$$P_s(H_k) = (1 - P_m)^k$$

for kth-order hyperplanes.

Given this analysis, we can now compare the disruptive effects of mutation with those of crossover. Figure 3 illustrates this with mutation rates of 0.01, 0.1, and 0.5. The curves for uniform crossover are the same as those in Figure 2.

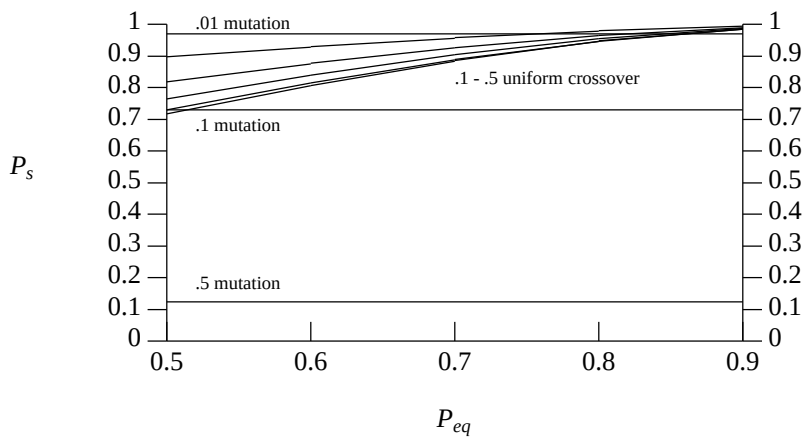

Figure 3: Survival Comparison, 3rd-Order

Figure 3 highlights several points. First, as expected, the disruptive effects of crossover are time dependent, while those of mutation are time independent. Second, the highest level of disruption for uniform crossover occurs when $P_0 = 0.5$. More interestingly, though, mutation can provide the same level of disruption as uniform crossover, if we allow the mutation rate to vary as a function of P_{eq}, k, and P_0.

In summary, then, we have introduced a disruption rate theory for mutation, and have compared this theory with our disruption rate theory for crossover. This comparison indicates that every level of disruption provided by crossover can be achieved with mutation alone. In fact, the comparison further indicates that crossover can not achieve the high levels of disruption that can be provided by mutation.

2.2 DISRUPTION DISTRIBUTIONS

Disruption rate theory estimates the likelihood that a genetic operator will disrupt a hyperplane sample. Again, we can also interpret this as the likelihood that individuals within a hyperplane will leave that hyperplane. It does not, however, indicate where those individuals will go. In other words, disruption rate theory does not indicate the distribution of disruptions, simply the likelihood that disruptions will occur.

Previous researchers (e.g., Eshelman *et al.*, 1989) have discussed the *exploratory power* of crossover operators, namely, the manner in which crossover disrupts individuals within hyperplanes. For example, suppose we consider the crossover of individuals from the two hyperplanes (the "#" denotes the "don't care" symbol):

1: 1####1
2: 0####0.

Uniform crossover will produce individuals from "######", while 1-point crossover will produce individuals from "1####0" and "0####1". In general, uniform crossover is more "explorative" than 1-point crossover. Eshelman *et al.* (1989) and Booker (1992) provide analyses of other biases in crossover operators.

What is the explorative power of mutation? Recall that our model of mutation assumes that a bit is flipped if it is chosen for mutation. We do not disrupt an individual within any hyperplane if the mutation rate is 0.0. If the mutation rate is 1.0, we always disrupt the individual, and produce the complement of the individual. For a mutation rate of 0.5, an individual will be disrupted with high probability, possibly creating any other individual. In summary, we can control the amount of exploration that mutation performs by adjusting the mutation rate. Mutation can provide any *amount* of exploration that crossover can provide.

Let us now compare the *type* of exploration that crossover and mutation provide. Suppose we consider individuals from the two hyperplanes:

1: 10####
2: 11####.

Crossover will only produce individuals from the hyperplane "1#####". The first "1" is guaranteed because it is common to the first defining position of both hyperplanes. Mutation, however, will *not* necessarily honor that guarantee, since it is a one individual operator and does not determine the commonality of alleles. Crossover, then, *preserves* alleles that are common to the individuals within the two hyperplanes (Radcliffe (1991) refers to this as "respect"). Preservation limits the type of exploration that crossover can perform. This limitation becomes more acute as the population loses diversity, since the number of common alleles will increase.

In summary, disruption analysis is the traditional analysis for describing the behavior of GAs in general, and crossover in particular. We have shown that mutation can provide any level of disruption that crossover can provide. We have also considered the form of disruption for both operators, by considering their exploratory power. Again, crossover has no advantage over mutation in terms of the amount of exploration that can be performed. They do differ, however, in the type of exploration. Crossover guarantees preservation of common alleles, while mutation does not. Given this evidence, then, we might suppose that there is some theoretical support for disputing Fogel's claim that crossover has no general advantage over mutation.[2] In the next section we will consider another potential difference between crossover and mutation.

---

[2] We are not implying that mutation has no advantage over crossover, however.

3 CONSTRUCTION THEORY

In the traditional theory, crossover is analyzed as a disruptive operator. However, more recently, Syswerda (1989) hypothesized that a more positive theory of crossover is constructive in nature.[3] For example, instead of calculating the probability that an existing hyperplane sample will be disrupted, we now calculate the probability that an individual within a hyperplane will be constructed from existing individuals within lower order hyperplanes. Syswerda's theory was extended by Spears & De Jong (1991) to include *n*-point and P_0 uniform crossover. This theory indicated that highly disruptive crossover operators are also highly constructive. Unfortunately, however, there was no theoretical evidence to indicate that mutation is not as constructive as crossover. In this section we will show that an analysis of the constructive abilities of mutation will provide us with that evidence.

Suppose, then, that we wish to create a theory of construction for mutation. More specifically, imagine that we wish to construct an individual within the 5th-order hyperplane "11111###" from an individual within another 5th-order hyperplane "11100###". This can be accomplished by mutating the 0's, while not mutating the 1's. In general, suppose we wish to construct an individual within a *k*th-order hyperplane H_k from an individual within another *k*th-order hyperplane H_s, when the two hyperplanes match on *m* alleles and do not match on *n* alleles (i.e., $k = n + m$). Then the probability of construction (denoted as P_{con}) is given by:

$$P_{con}(H_k, H_s) = (1 - P_m)^m \, (P_m)^n$$

In order to compare this with the constructive effects of crossover we again need to be careful about our assumptions. In this case we wish to compute the probability that crossover will construct an individual within a *k*th-order hyperplane from an individual in another hyperplane with *m* correct and *n* incorrect alleles, given an arbitrary mate. As an illustration, let us again imagine that we wish to construct an individual in the 5th-order hyperplane "11111###" from an individual in "11100###", using crossover. The individual in the hyperplane "11100###" will be crossed with an arbitrary individual from one of the four following hyperplanes:

1: ###00###
2: ###01###
3: ###10###
4: ###11###.

Of these four situations, only the last can result in the construction of an individual in the hyperplane "11111###". Each of these situations is not equally likely, unless $P_{eq} = 0.5$. For example, given the hyperplane "11100###" and the fact that $P_{eq} = 0.8$, we can compute that an individual from "11100###" will be crossed with an individual from hyperplane "###11##" with probability 0.04. In general, the probability that two hyperplanes differ in *n* defining positions is:

$$\prod_{d \in N} (1 - P_{eq}(d))$$

[3] In prior work we refer to this as "recombination" theory. Since we wish to extend this theory to mutation, the term "construction" seems more appropriate.

where N is the set of n defining positions.

We have now calculated the probability that a potentially successful recombination can occur. However, since this does not guarantee success, we also need to determine the probability that crossover will yield an individual within the hyperplane "11111###", given individuals from:

1: 11100###
2: ###11###.

This can be done in a straightforward fashion by using the earlier recombination theory of Spears & De Jong (1991), that deals with the construction of individuals within a hyperplane from individuals within two non-overlapping lower order hyperplanes (see the Appendix for details). In this theory, for example, it is possible to compute the probability that an individual within "11111###" will be constructed from the crossover of individuals from:

1: 111#####
2: ###11###.

This is a more general case of the above situation, in which an individual from "11100###" is crossed with an individual from "###11###" (i.e., because the "#" is more general than a "0"). Specifically, the two situations are identical if we state that $P_{eq}(d) = 0$ for the last two defining positions.[4]

In general, if we are interested in constructing an individual within a kth-order hyperplane from an individual within a hyperplane that has n incorrect alleles, we first need to compute the probability that some other individual contains the n correct alleles, to ensure that a potentially successful recombination can occur. Given these two individuals, we then compute the probability that construction will occur, by using a specific case of the earlier recombination theory, where $P_{eq}(d) = 0$ for those n defining positions. In summary, we can use the earlier recombination theory to create a construction theory for crossover that can be compared with the construction theory for mutation.

Using these theories, we can compare mutation and crossover from the viewpoint of construction. Figure 4 presents the comparison for 3rd-order hyperplanes, while Figure 5 presents the comparison for 8th-order hyperplanes. Mutation rates of 0.01, 0.1, and 0.5 are again compared with uniform crossover. Again, for the sake of simplicity, we illustrate uniform crossover where P_0 ranges from 0.1 to 0.5, because this roughly bounds the levels of construction provided by n-point crossover. It is important to note that a mutation rate of 0.5 yields the highest probability of construction. Due to symmetry, mutation rates above 0.5 yield lower probabilities.

Figures 4 and 5 illustrate several interesting points. First, when the population is diverse, mutation can not match the levels of construction that crossover can achieve. In fact, for 3rd-order hyperplanes, crossover has higher constructive levels until the population is 70% converged. Second, this advantage increases as the order of the hyperplane increases. For example, with 8th-order hyperplanes, crossover has higher

[4] Again, the theory allows one to define distinct $P_{eq}(d)$'s for each defining position.

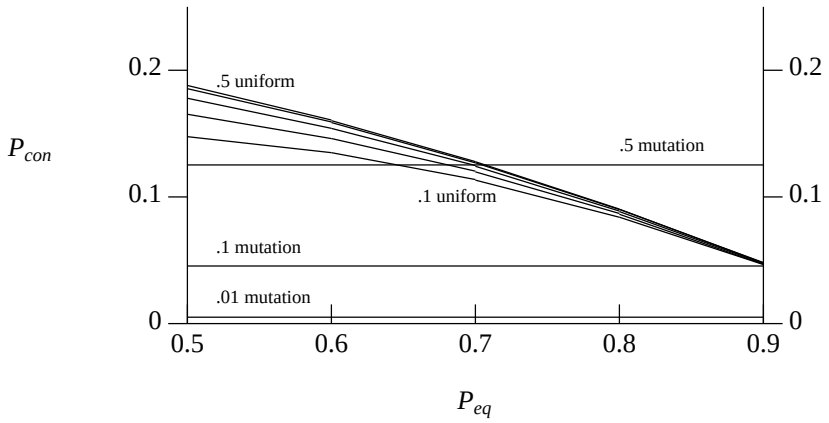

Figure 4: Construction Comparison, 3rd-Order

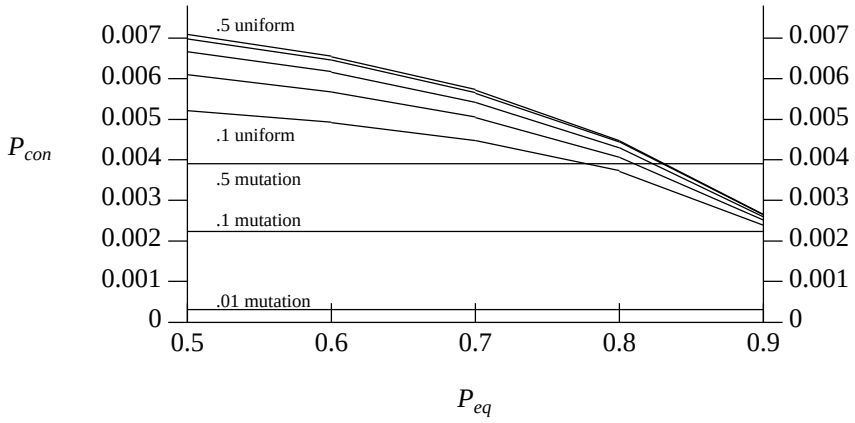

Figure 5: Construction Comparison, 8th-Order

constructive levels until the population is 80% converged.

Finally, note that it is impossible for mutation to simultaneously achieve high levels of construction and survival. This would appear to be important, since one without the other may not be extremely useful (i.e., it is nice if some of the constructions survive!). High construction levels with mutation are accomplished at the expense of survival (see 0.5 mutation), while good survival is at the expense of construction (see 0.01 mutation). In fact, crossover can simultaneously achieve higher levels of construction and survival than any particular amount of mutation.

4 SUMMARY AND DISCUSSION

These results provide a theoretical justification for Holland's belief that the role of crossover is to construct high order building blocks (hyperplanes) from low order building blocks. Mutation can not perform the role as well as crossover. Clearly, the role of crossover is construction and, in this case, crossover provides an advantage over mutation. In terms of disruption, mutation can provide higher levels of disruption and exploration, but at the expense of preserving alleles common to particular defining positions.

The disruption and construction theories presented here all concentrate on hyperplane (building block) analysis. Since the concept of a building block is often central to genetic algorithm research, it is important to connect our work with other relevant findings. First, this work does not assume a condition referred to as the *Static Building Block Hypothesis*. Next, we tie our work to the exploration and exploitation tradeoff, and indicate that our findings are consistent with experimental results. We conclude with the observation that our current distinction between crossover and mutation may not be necessary.

4.1 BUILDING BLOCKS

Since a role of crossover is construction, we would expect crossover to be useful on problems that have appropriate building blocks. What exactly is an appropriate building block? One possible answer lies in the following hypothesis (Grefenstette, 1992a):

> The *Static Building Block Hypothesis* (SBBH): Given any short, low order hyperplane partition, a GA is expected to converge to the hyperplane with the best static average fitness.

This hypothesis is often used as a base for theoretical and experimental work in genetic algorithms and implies that appropriate building blocks should have the highest average fitness. Unfortunately, as Grefenstette (1992a) indicates, the hypothesis is flawed in that a genetic algorithm is unlikely to determine the actual average fitness of a hyperplane, because the sampling of hyperplanes is biased. Although construction theory is concerned with the building of higher order hyperplanes from lower order hyperplanes, we do *not* make use of the SBBH. Rather, construction theory is consistent with what we will call the DBBH:

> The *Dynamic Building Block Hypothesis* (DBBH): Given any short [5], low order hyperplane partition, a GA is expected to converge to the hyperplane with the best dynamic (observed) average fitness.

In other words, a GA estimates the static average fitness from a dynamic biased sampling. As can be expected, the observed average fitness of a hyperplane can be quite different from its actual average fitness, implying that the GA may not converge to the hyperplane with the best static average fitness. Crossover, then, constructs higher order hyperplanes from lower order hyperplanes that have higher observed average fitness. These higher order hyperplanes may or may not bias search appropriately. Crossover works well with problems that have building blocks conducive to the creation of higher order building blocks that bias search correctly (see Vose & Liepins (1991) for a

[5] Actually, length is irrelevant for uniform crossover.

theoretical treatment of the relationship between crossover and building blocks). Although the appropriateness of building blocks is dynamic, and not well understood, some progress has been made in understanding the underlying issues. In the next section we outline some of the recent work. This work helps us to understand the roles of crossover and mutation in genetic algorithm search.

4.2 EXPLOITATION VS EXPLORATION

The issue concerning the relative importance of mutation and crossover can be viewed at a higher level. Mutation serves to create random diversity in the population, while crossover serves as an accelerator that promotes emergent behavior from components. The meta-issue, then, is the relative importance of diversity and construction. For the GA community, this is also related to the balance between exploration and exploitation. This meta-issue is the key to the difference in philosophy between Holland and Fogel. Specifically, Fogel *et al*. question the importance of recombination. They do not believe that natural selection selects individual traits (or, presumably, combinations of traits). Recombination is considered to be a third order factor, since it does not appear to occur frequently in nature (Atmar, 1992).

Of course, this does not necessarily imply that recombination is not useful for problems we wish to solve. Atmar is correct to remind us that "uncritical advocacy of a particular phenomenon" promotes "a blindness in perspective that is very difficult to dispel" (Atmar, 1992). Neither mutation nor crossover should be uncritically advocated or dismissed. Each operator plays a different role in the search process. *A priori*, it is difficult to specify the relative importance of each operator, for each problem. The appropriate balance of exploration and exploitation required for good performance depends on the amount of diversity in the population, the style of genetic algorithm used, and the purpose for which it is used.

For example, although GAs are often used as optimizers, our current understanding is that they attempt to maximize accumulated payoff (Holland, 1975). In this sense, they are greedy and should not necessarily be expected to find optimal solutions. Crossover can enhance this effect. Fogel and Atmar (1990) report that although the mean behavior of a GA with crossover outperformed the mean behavior of a GA without crossover (albeit insignificantly), they regard the winner to be the algorithm that found superior individual solutions. The GA without crossover had a much higher variance and found superior solutions in 6 of 10 trials. This effect has also been noted by Spears and Anand (1991), although they found that the results were dependent on population size. This indicates, then, that the GA practitioner should be clear about his or her goals. If optimality is sought, crossover may be deleterious. If the maximization of accumulated payoff is sought, mutation may be insufficient.

Similarly, greater disruption is more important for steady state genetic algorithms, since they suffer a higher allele loss than do their generational counterparts (De Jong & Sarma, 1992). It is also more important in non-stationary environments, where the optimal solution changes over time (Grefenstette, 1992b). We can conclude, then, that mutation will play an important role in these situations. Figures 4 and 5 support these ideas by suggesting that mutation becomes more important relative to crossover as the population loses diversity. Experiments with adaptive operator probabilities (Davis, 1989) support this analysis.

Crossover, however, will play an important role when construction and survival are required for good performance. This occurs when the population is diverse and problems consist of appropriate building blocks. Recent work suggests that fitness correlation (Lipsitch, 1991; Manderick *et al.*, 1991) and epistasis (Schaffer & Eshelman, 1991; Davidor, 1990) provide useful measures for determining the usefulness of crossover. For example, crossover appears to work well with functions that are highly correlated or have mild epistasis.

4.3 WHAT IS MUTATION?

Although our discussion of mutation and crossover stresses the differences between the two operators, it is also important to note that mutation can be greatly modified, minimizing those differences. The reason that crossover can exhibit high simultaneous levels of preservation, survival, and construction is that crossover shares information between fit individuals. Mutation, on the other hand, is often implemented with a parameter that is constant during genetic algorithm search. No information is shared when mutation is implemented in this fashion.

It is possible to implement mutation with a parameter that is adapted during genetic algorithm search. Population statistics, such as population convergence, are often used to adapt the mutation rate (Davis, 1989). The European community (e.g., Baeck *et al.*, 1991) go further, and explicitly adjust the mutation of each parameter, for every chromosome. One can easily imagine, then, a situation in which these mutation rates are based on finer grained population statistics, such as column convergence (De Jong, 1975). At this point, information can be communicated in a fashion similar to that of crossover. For example, we could measure the allele loss for each defining position and only mutate at defining positions with small allele loss, thus preserving common alleles. This would give mutation a disruption distribution more similar to that of crossover (see Section 2.2). At what point do we no longer call this mutation?

This leads us to the realization that standard mutation and crossover are simply two forms of a more general exploration operator, that can perturb alleles based on any available information (e.g., see Syswerda, 1992). It is not clear that the current distinction between crossover and mutation is necessary, or even desirable, although it may be convenient. The creation of more general operators, however, may lead to more robust biases. For example, it may be possible to implement one general operator that can specialize to mutation, crossover, or any variation in between. In our future work we intend to investigate this alternative.

Acknowledgements

I would like to thank Lawrence Fogel for suggesting this work in a dream. In this dream, Dr. Fogel approached me and said "You know, in terms of disruption theory, mutation can do everything crossover can. Crossover isn't necessary." I rushed to work and immediately began preparing this paper.

I would also like to thank Ken De Jong and John Grefenstette for many constructive discussions on methods for comparing crossover with mutation. Finally, I also wish to thank Diana Gordon, Alan Schultz, and four anonymous reviewers for helping to clarify earlier versions of this paper.

References

Atmar, W. (1992) The philosophical errors that plague both evolutionary theory and simulated evolutionary programming. In D. Fogel & J. Atmar (eds.), *Proceedings of the First Annual Conference on Evolutionary Programming*, 27-34. San Diego, CA: Evolutionary Programming Society.

Baeck, T., Hoffmeister, F., & Schwefel, H.-P. (1991) A survey of evolution strategies. *Proceedings of the Fourth International Conference on Genetic Algorithms*, 2-9. La Jolla, CA: Morgan Kaufmann.

Booker, L. B. (1992) Recombination distributions for genetic algorithms. *Proceedings of the Foundations of Genetic Algorithms Workshop*. Vail, CO: Morgan Kaufmann.

Davidor, Y. (1990) Epistasis variance: a viewpoint on GA-hardness. *Proceedings of the Foundations of Genetic Algorithms Workshop*, 23-35. Indiana University: Morgan Kaufmann.

Davis, L. (1989) Adapting operator probabilities in genetic algorithms. *Proceedings of the Third International Conference on Genetic Algorithms*, 60-69. La Jolla, CA: Morgan Kaufmann.

De Jong, K. A. (1975) *An analysis of the behavior of a class of genetic adaptive systems.* Doctoral Thesis, Department of Computer and Communication Sciences. University of Michigan, Ann Arbor.

De Jong, K. A. & Spears, W. M. (1992) A formal analysis of the role of multi-point crossover in genetic algorithms. *Annals of Mathematics and Artificial Intelligence Journal* 5, 1-26. J. C. Baltzer A. G. Scientific Publishing Company.

De Jong, K. A. & Sarma, J. (1992) Generation gaps revisited. *Proceedings of the Foundations of Genetic Algorithms Workshop*. Vail, CO: Morgan Kaufmann.

Eshelman, L. J., Caruana, R. A., & Schaffer, J. D. (1989) Biases in the crossover landscape. *Proceedings of the Third International Conference on Genetic Algorithms*, 10-19. La Jolla, CA: Morgan Kaufmann.

Fogel, L. J., Owens, A. J., & Walsh, M. J. (1966) *Artificial Intelligence Through Simulated Evolution.* New York: Wiley Publishing.

Fogel, D. B. & Atmar, J. W. (1990) Comparing genetic operators with gaussian mutations in simulated evolutionary processes using linear systems. *Biological Cybernetics* 63, 111-114.

Grefenstette, J. G. (1992a) Deception considered harmful. *Proceedings of the Foundations of Genetic Algorithms Workshop*. Vail, CO: Morgan Kaufmann.

Grefenstette, J. G. (1992b) Genetic algorithms for changing environments. Submitted to the *Parallel Problem Solving from Nature (PPSN) Workshop*. Brussels, Belgium.

Holland, J. H. (1975) *Adaptation in Natural and Artificial Systems*. Ann Arbor, Michigan: The University of Michigan Press.

Lipsitch, M. (1991) Adaptation on rugged landscaped generated by iterated local interactions of neighboring genes. *Proceedings of the Fourth International Conference on Genetic Algorithms*, 128-135. La Jolla, CA: Morgan Kaufmann.

Manderick, B., de Weger, M., & Spiessens, P. (1991) The genetic algorithm and the structure of the fitness landscape. *Proceedings of the Fourth International Conference on Genetic Algorithms*, 143-149. La Jolla, CA: Morgan Kaufmann.

Radcliffe, N. J. (1991) Forma analysis and random respectful recombination. *Proceedings of the Fourth International Conference on Genetic Algorithms*, 222-229. La Jolla, CA: Morgan Kaufmann.

Rechenberg, I. (1973) *Evolutionsstrategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution*. Frommann-Holzboog, Stuttgart.

Schaffer, J. D., Caruana, R. A., Eshelman, L. J., & Das, R. (1989) A study of control parameters affecting online performance of genetic algorithms for function optimization. *Proceedings of the Third International Conference on Genetic Algorithms*, 51-60. La Jolla, CA: Morgan Kaufmann.

Schaffer, J. D., Eshelman, L. J. (1991) On crossover as an evolutionarily viable strategy. *Proceedings of the Fourth International Conference on Genetic Algorithms*, 61-68. La Jolla, CA: Morgan Kaufmann.

Schwefel, H.-P. (1977) *Numerical Optimization of Computer Models*. New York: John Wiley & Sons.

Spears, W. M. & Anand, V. (1991) A study of crossover operators in genetic programming. *Proceedings of the Sixth International Symposium on Methodologies for Intelligent Systems*, 409-418. Charlotte, NC: Springer-Verlag.

Spears, W. M. & De Jong, K. A. (1991) On the virtues of uniform crossover. *Proceedings of the Fourth International Conference on Genetic Algorithms*, 230-236. La Jolla, CA: Morgan Kaufmann.

Syswerda, G. (1989) Uniform crossover in genetic algorithms. *Proceedings of the Third International Conference on Genetic Algorithms*, 1-9. Fairfax, VA: Morgan Kaufmann.

Syswerda, G. (1992) Simulated crossover in genetic algorithms. *Proceedings of the Foundations of Genetic Algorithms Workshop*. Vail, CO: Morgan Kaufmann.

Vose, M. D., & Liepins, G. E. (1991) Schema disruption. *Proceedings of the Fourth International Conference on Genetic Algorithms*, 237-242. La Jolla, CA: Morgan Kaufmann.

Appendix[6]

Summary of the Survival Analysis

For n-point crossover, P_s is expressed in the order dependent form ($P_{k,s}$):

$$P_{2,s}(n, L, L_1) = \sum_{i=0}^{n} \binom{n}{i} \left(\frac{L_1}{L} \right)^i \left(\frac{L - L_1}{L} \right)^{n-i} C_s$$

and

$$P_{k,s}(n, L, L_1, \ldots, L_{k-1}) = \sum_{i=0}^{n} \binom{n}{i} \left(\frac{L_1}{L} \right)^i \left(\frac{L - L_1}{L} \right)^{n-i} P_{k-1,s}(i, L_1, \ldots, L_{k-1})$$

Note that the survival of a kth-order hyperplane under n-point crossover is recursively defined in terms of the survival of lower order hyperplanes. L refers to the length of the individuals. The $L_1 \cdots L_{k-1}$ refer to the defining lengths between the defining positions of the kth-order hyperplane. The effect of the recursion and summation is to consider every possible placement of n crossover points within the kth-order hyperplane. The correction factor C_s computes the probability that the hyperplane will survive, based on that placement of crossover points.

For the ease of presentation we denote K to be the set of k defining positions, while X denotes a subset of K. Suppose that crossover results in a subset X of defining positions being exchanged. Then the hyperplane will survive if: 1) the parents match on the subset X, or 2) if they match on the subset $K - X$, or 3) they match on the set K. Hence, the general form of the correction is:

$$C_s = \prod_{d \in X} P_{eq}(d) + \prod_{d \in K-X} P_{eq}(d) - \prod_{d \in K} P_{eq}(d)$$

where $P_{eq}(d)$ is the probability that two parents have the same alleles on a particular defining position d.

For parameterized uniform crossover, P_s is also expressed in an order dependent form ($P_{k,s}$):

$$P_{k,s}(H_k) = \sum_{I \subseteq K} (P_0)^{|I|} \; (1 - P_0)^{|K-I|} \left[\prod_{d \in I} P_{eq}(d) + \prod_{d \in K-I} P_{eq}(d) - \prod_{d \in K} P_{eq}(d) \right]$$

where I is a subset of K, and P_0 is the probability of swapping two parents' alleles at each defining position. A graphical representation of these equations has been shown previously in Figure 1.

Recombination (Construction) Analysis for N-Point Crossover

In our definition of survival, it is possible for a hyperplane to survive in either child. Recombination can be considered a restricted form of survival, in which two lower order

[6] This Appendix is a compendium of crossover theory from De Jong & Spears (1992).

hyperplanes survive to form a higher order hyperplane. The difference is that the two lower order hyperplanes (each of which exists in a different parent) must survive in the same individual, in order for recombination to occur.

In the remaining discussion we will consider the creation of a kth-order hyperplane from two hyperplanes of order m and n. We will restrict the situation such that the two lower order hyperplanes are non-overlapping, and $k = m + n$. Each lower order hyperplane is in a different parent. We denote the probability that the kth-order hyperplane will be recombined from the two hyperplanes as $P_{k,r}$.

An analysis of recombination under n-point crossover is simple if one considers the correction factor C_s defined earlier for the survival analysis. Recall that recombination will occur if both lower order hyperplanes survive in the same individual. If an n-point crossover results in a subset X of the k defining positions surviving in the same individual, then recombination will occur if: 1) the parents match on the subset X, or 2) if they match on the subset $K - X$, or 3) they match on the set K. Hence, the general form of the recombination correction C_r is:

$$C_r = \prod_{d \in X} P_{eq}(d) + \prod_{d \in K-X} P_{eq}(d) - \prod_{d \in K} P_{eq}(d)$$

Note the similarity in description with the survival correction factor C_s (the only difference is in how X is defined). In other words, given a kth-order hyperplane, and two hyperplanes of order n and m, $P_{k,r}$ is simply $P_{k,s}$ with the correction factor redefined as above.

Recombination (Construction) Analysis for Uniform Crossover

The analysis of recombination under uniform crossover also involves the analysis of the original survival equation. Note that, due to the independence of the operator (each allele is swapped with probability P_0), the survival equation can be divided into three parts. The first part expresses the probability that a hyperplane will survive in the original string:

$$P_{k,s,orig}(H_k) = \sum_{I \subseteq K} (P_0)^{|I|} \; (1 - P_0)^{|K-I|} \prod_{d \in K - I} P_{eq}(d)$$

The second part expresses the probability that a hyperplane will survive in the other string:

$$P_{k,s,other}(H_k) = \sum_{I \subseteq K} (P_0)^{|I|} \; (1 - P_0)^{|K-I|} \prod_{d \in I} P_{eq}(d)$$

The final part expresses the probability that a hyperplane will exist in both strings:

$$P_{k,s,both}(H_k) = \sum_{I \subseteq K} (P_0)^{|I|} \; (1 - P_0)^{|K-I|} \prod_{d \in K} P_{eq}(d) = \prod_{d \in K} P_{eq}(d)$$

Then:

$$P_{k,s}(H_k) = P_{k,s,orig}(H_k) + P_{k,s,other}(H_k) - P_{k,s,both}(H_k)$$

Note, however, that this formulation allows us to express recombination under uniform crossover. Again, assuming the recombination of two non-overlapping hyperplanes of order n and m into a hyperplane of order k:

$$
\begin{aligned}
P_{k,r}(H_k) = {} & P_{m,s,orig}(H_m)\, P_{n,s,other}(H_n) + \\
& P_{m,s,other}(H_m)\, P_{n,s,orig}(H_n) - \\
& P_{m,s,both}(H_m)\, P_{n,s,both}(H_n)
\end{aligned}
$$

This equation reflects the decomposition of recombination into two independent survival events. The first term is the probability that H_m will survive on the original string, while H_n switches (i.e., both hyperplanes survive on one parent). The second term is the probability that both hyperplanes survive on the other parent. The third term reflects the joint probability that both hyperplanes survive on both strings, and must be subtracted. It should also be noted that the third term is equivalent to $\prod_{d \in K} P_{eq}(d)$.