On the Virtues of Parameterized Uniform Crossover

William M. Spears
Naval Research Laboratory
Washington, D.C. 20375 USA
spears@aic.nrl.navy.mil

Kenneth A. De Jong
George Mason University
Fairfax, VA 22030 USA
kdejong@aic.gmu.edu

Abstract

Traditionally, genetic algorithms have relied upon 1 and 2-point crossover operators. Many recent empirical studies, however, have shown the benefits of higher numbers of crossover points. Some of the most intriguing recent work has focused on uniform crossover, which involves on the average $L/2$ crossover points for strings of length L. Theoretical results suggest that, from the view of hyperplane sampling disruption, uniform crossover has few redeeming features. However, a growing body of experimental evidence suggests otherwise. In this paper, we attempt to reconcile these opposing views of uniform crossover and present a framework for understanding its virtues.

1 Introduction

One of the unique aspects of the work involving genetic algorithms (GAs) is the important role that recombination plays. In most GAs, recombination is implemented by means of a crossover operator which operates on pairs of individuals (parents) to produce new offspring by exchanging segments from the parents' genetic material. Traditionally, the number of crossover points (which determines how many segments are exchanged) has been fixed at a very low constant value of 1 or 2. Support for this decision came from early work of both a theoretical and empirical nature [Holland, 1975; DeJong, 1975]. However, there continue to be indications that there are situations in which having a higher number of crossover points is beneficial [Syswerda, 1989; Eschelman, 1989]. Perhaps the most surprising result (from a traditional perspective) is the effectiveness on some problems of uniform crossover, an operator which produces on the average $L/2$ crossings on strings of length L [Syswerda, 1989].

Recent work by [Spears and De Jong, 1990] has extended the theoretical analysis of n-point and uniform crossover with respect to disruption of sampling distributions. However, they pointed out that disruption analysis alone is not sufficient in general to predict and/or select optimal forms of crossover. In particular, they have shown that the population size must also be taken into account [DeJong and Spears, 1990]. This paper extends that work by looking at the properties of a parameterized uniform crossover operator and by considering two other aspects of crossover operators, namely, their recombination potential and their exploratory power. In this context, a surprisingly positive view of uniform crossover emerges.

2 Disruption Analysis

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 used is not too severe [Holland, 1975]. Since mutation is typically run at a very low rate, it is generally ignored as a significant source of disruption. However, crossover is usually applied at a very high rate. So, considerable attention has been given to estimating P_d, the probability that a particular application of crossover will be disruptive.

Holland's initial analysis of the sampling disruption of 1-point crossover [Holland, 1975] has been extended to n-point and uniform crossover [DeJong, 1975; Spears and DeJong, 1990]. These results are in the form of estimates of the likelihood that the sampling of a kth order hyperplane (H_k) will be disrupted by a particular form of crossover. It turns out to be easier mathematically to estimate the complement of disruption: the likelihood of a sample surviving crossover (which we denote as P_s). As one might expect, the results are a function of both the order k of the hyperplane and its defining length (see the Appendix and [Spears and DeJong, 1990] for more precise details).

We provide in Figure 1 a graphical summary of a typical instance of these results for the case of 3rd order hyperplanes. The non-horizontal curves represent the survival

of 3rd order hyperplanes under *n*-point crossover (n = 1...6). The horizontal line represents the probability of survival under uniform crossover. Figure 1 highlights two important points. First, if we interpret the area *above* a particular curve as a measure of the cumulative disruption potential of its associated crossover operator, then these curves suggest that 2-point crossover is the least disruptive of the *n*-point crossover family, and less disruptive than uniform crossover. Finally, unlike *n*-point crossover, uniform crossover disrupts all hyperplanes of order k with equal probability, regardless of how long or short their defining lengths are.

3 A Positive View of Crossover Disruption

A recurring theme in Holland's work is the importance of a proper balance between exploration and exploitation when adaptively searching an unknown space for high performance solutions [Holland, 1975]. The disruption analysis of the previous section implicitly assumes that disruption of the sampling distributions is a bad thing and to be avoided (e.g., a high disruption may stress exploration at the expense of exploitation). However, this is not always the case. There are important situations in which minimizing disruption hinders the adaptive search process by overemphasizing exploitation at the expense of needed exploration. One of the clearest examples of this is when the population size is too small to provide the necessary sampling accuracy for complex search spaces [DeJong and Spears, 1990].

To illustrate this we have selected a 30 bit problem with 6 peaks from [DeJong and Spears, 1990]. The measure of performance is simply the best individual found by the genetic algorithm. This is plotted every 100 evaluations. Since we are maximizing, higher curves represent better performance. Figures 2 and 3 illustrate the effect of population size on GA performance. Notice how uniform crossover dominates 2-point crossover on the 6-Peak problem with a small population, but just the opposite is true with a large population.

One conclusion of these results might be that we should maintain a portfolio of crossover operators and study the effects of various combinations. We have been examining another approach: achieving a better balance of exploration and exploitation using only uniform crossover. We are intrigued by this possibility for two reasons: its simplicity (only one crossover form) and its potential for increased robustness because the disruptive effect of uniform crossover is not influenced by hyperplane defining length.

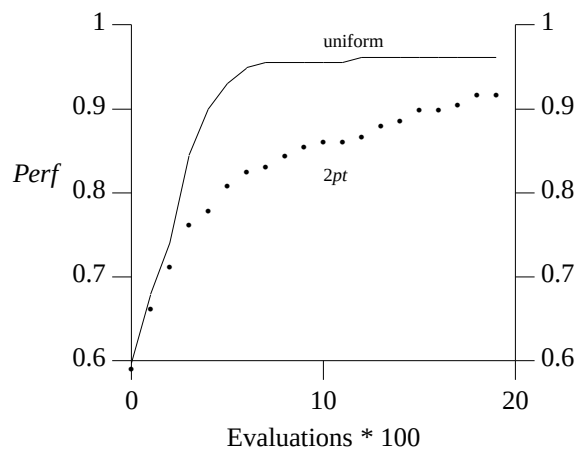

Figure 2: 6-Peak (30 bits) - Population 20

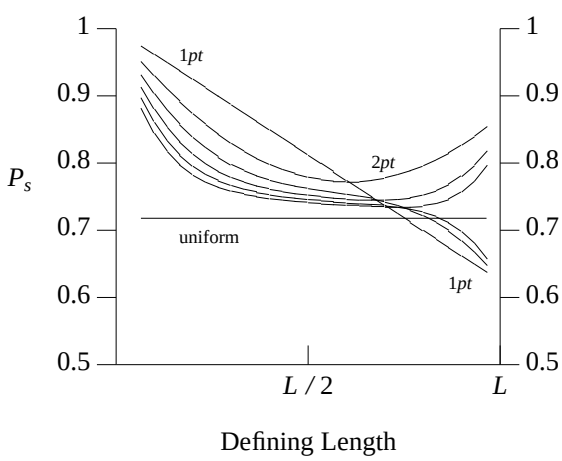

Figure 1. Survival of 3rd Order Hyperplanes

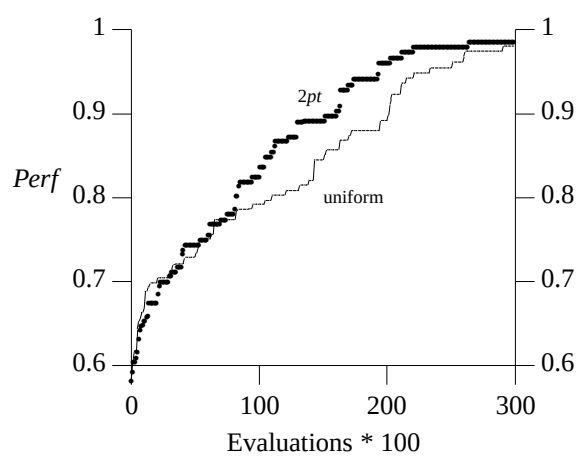

Figure 3: 6-Peak (30 bits) - Population 1000

4 A Closer Look at Uniform Crossover

It is clear that the level of disruption provided by uniform crossover is too high in many cases (e.g., when large populations are used). This standard form of uniform crossover swaps two parents' alleles with a probability of 0.5. Suppose, however, that we parameterize uniform crossover, where P_0 denotes the probability of swapping. We can now consider the effect of decreasing P_0.[1] Figure 4 illustrates this for 3rd order hyperplanes. Notice how the disruption of uniform crossover can be controlled by lowering P_0, *without* affecting the property that the disruption has no defining length bias. In particular, note that by simply lowering P_0 to .1, uniform crossover is less disruptive (overall) than 2-point crossover and has no defining length bias! This suggests a much more positive view of the potential of uniform crossover, namely, an unbiased recombination operator whose disruption potential can be easily controlled by a single parameter P_0.

To test this hypothesis, we have run a number of experiments in which P_0 varied. As expected, we can increase and decrease performance on a given problem with a fixed population size simply by varying P_0. Figure 5 illustrates this on the 6-Peak problem. Note that in this particular case, a value of $P_0 = 0.2$ produced the best results. Referring back to Figures 3 and 4, we can now see why. For the 6-Peak problem, a population size of 1000 has sufficient sampling capacity to require only the disruption level provided by 2-point crossover. Uniform crossover with $P_0 = 0.2$ provides approximately the same level of disruption but without the length bias.

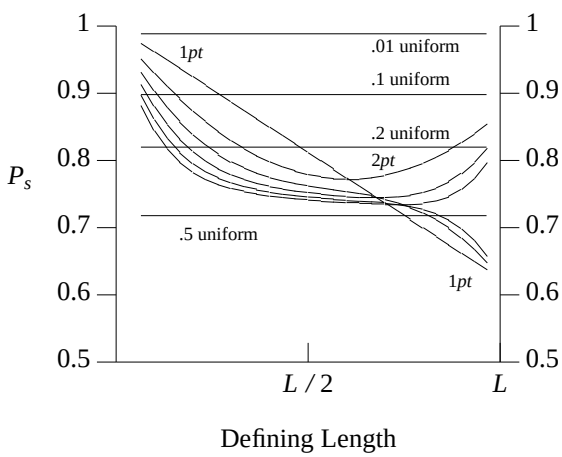

Figure 4. Survival of 3rd Order Hyperplanes

[1] Note that we do not need to consider the possibility of increasing P_0, due to the symmetry of uniform crossover.

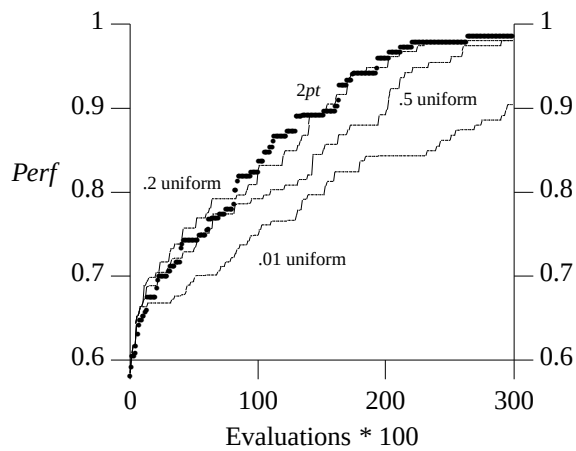

Figure 5: 6-Peak (30 bits) - Population 1000

Is this lack of length bias really important? Intuitively, it should help overcome representation problems in which important hyperplanes happen to have defining lengths which are adversely affected by the particular *n*-point crossover operator in use. Syswerda illustrated this clearly with his "sparse 1-max" problem in which 270 fake bits were appended to a 30-bit problem [Syswerda, 1989]. One can show similar results with almost any problem. Figure 6 illustrates this on our 6-Peak problem appended with 270 fake bits and the same evaluation function. Notice that, in comparison to the original 30-bit problem shown in Figure 5, the performance of 2-point crossover is worse, while the performance of uniform crossover ($P_0 = .2$) remains essentially unchanged.

How do we explain the drop in performance of 2-point crossover? In this case, the 30 important bits are all

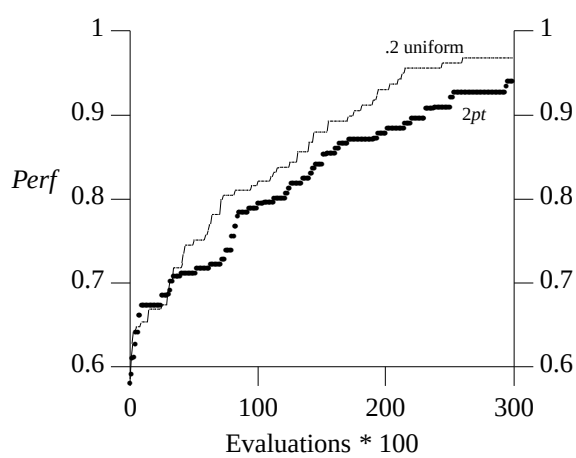

Figure 6: 6-Peak (300 bits) - Population 1000

within a distance of $L/10$ of each other (where L is the length of the string). If we examine Figure 4, we note that 2-point crossover is less disruptive within that range (0 to $L/10$) of defining lengths. In other words, the addition of 270 additional bits effectively decreases the disruption of the important hyperplanes under 2-point crossover. This effect is most obvious towards the end of the runs (see Figure 6), where disruption is increasingly useful (due to the increasing homogeneity of the population). Uniform crossover is not influenced by the added 270 bits, since it is insensitive to defining length.

In summary, we see two important virtues of uniform crossover. The first is the ease with which the disruptive effect of uniform crossover can be controlled by varying P_0. This is useful in achieving the proper balance between exploration and exploitation. The second virtue is that the disruptive potential of uniform crossover does not depend on the defining length of hyperplanes. This allows uniform crossover to perform equally well, regardless of the distribution of important alleles.

5 Recombination Potential

Another possible virtue of uniform crossover that has been discussed in the literature is its recombination potential. In comparing uniform, 1 and 2-point crossover, Syswerda felt that uniform crossover gained significant advantage from its ability to combine small building blocks into larger ones [Syswerda, 1989]. He defined recombination potential as the ability of crossover to create higher order hyperplanes when the parents contain the necessary lower order hyperplanes. He provided an analysis showing uniform crossover (P_0 = .5) to have a higher recombination potential than 1 and 2-point crossover.

Syswerda pointed out that recombination can be considered to be a specialized form of survival, in which two lower order hyperplanes survive onto the same string, resulting in a higher order hyperplane. This observation allowed Syswerda to construct a recombination analysis from his survival analysis. However, since his survival analysis was limited to 1 and 2-point crossover, and to uniform crossover with a P_0 of .5, his recombination analysis was similarly limited. This motivated us to create a new recombination analysis in a similar vein, since our survival analysis includes all of n-point crossover and a parameterized uniform crossover.

In [Spears and DeJong, 1990], we developed a survival analysis for n-point crossover and a parameterized (P_0) uniform crossover. Details of this analysis, and our recombination analysis, are presented in the Appendix. Figure 7 illustrates the relationships of the crossover operators in terms of their recombination potential (we denote P_r as the probability of recombination). Note specifically that there is evidence to support the claim that uniform crossover (P_0 = .5) has a higher

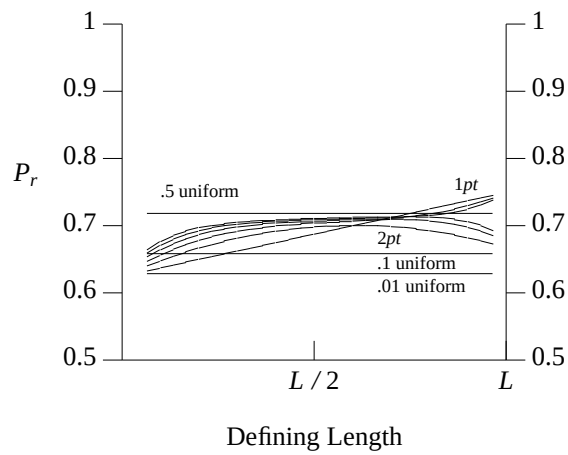

Figure 7: 3rd Order Hyperplane Recombination

recombination potential than the other crossover operators. However, it is even more interesting to note that these relationships are qualitatively identical to those shown in Figure 4. In other words, if one operator is better than another for survival, it is worse for recombination (and vice versa). This observation appears to hold for all k, and suggests very strongly that the recombination analysis tells us nothing new about crossover.

6 Exploration Power

It has also been pointed out that disruption does not necessarily mean useful exploration. Crossover disruption simply implies that a hyperplane sample has been modified by crossover in some way so as to no longer be a member of that hyperplane, without any indication as to the possible forms that change might take. The potential number of ways in which a crossover operator can effect a change has been called its exploratory power. It has been pointed out that uniform crossover has the additional property that it has more exploratory power than n-point crossover [Eschelman, 1989].

To see that this is true, consider the extreme case in which one parent is a string of all 0s and the other all 1s. Clearly uniform crossover can produce offspring anywhere in the space while 1 and 2-point crossover are restricted to rather small subsets. In general, uniform crossover is much more likely to distribute its disruptive trials in an unbiased manner over larger portions of the space.

The difficulty comes in analyzing whether this exploratory power is a virtue. If we think of exploitation as the biased component of the adaptive search process, it makes sense to balance this with unbiased exploration. Clearly, this exploratory power can help in the early generations, particularly with smaller population sizes, to make sure the whole space is well sampled. At the same

time, some of this exploratory power can be achieved over several generations via repeated applications of 1 and 2-point crossover. Unfortunately, our current analysis tools do not allow us to make comparisons of properties which span generations and are strongly affected by selection. Hopefully we will develop such tools and resolve questions of this type in the near future.

7 Conclusions and Further Work

The extensions to the analysis of *n*-point and uniform crossover presented in this paper open up an interesting and positive view of the usefulness of uniform crossover. There appear to be three potentially important virtues of uniform crossover. First, the disruption of hyperplane sampling under uniform crossover does not depend on the defining length of the hyperplanes. This reduces the possibility of representation effects, since there is no defining length bias. Second, the disruption potential is easily controlled via a single parameter P_0. This suggests the need for only one crossover form (uniform crossover), which is adapted to different situations by adjusting P_0. Finally, when a disruption does occur, uniform crossover results in a minimally biased exploration of the space being searched.

The first two virtues have been confirmed both theoretically and experimentally. At the same time, it should be emphasized that the empirical studies presented are limited to a carefully controlled experimental setting. The authors are currently working on expanding these experiments and on developing an exploration theory for recombination operators. Our goal is to understand these interactions well enough so that GAs can be designed to be self-selecting with respect to such decisions as optimal population size and level of disruption.

Acknowledgements

We would like to thank Diana Gordon for pointing out flaws in our preliminary recombination analysis.

References

De Jong, Kenneth 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. (1990). An Analysis of of the Interacting Roles of Population Size and Crossover in Genetic Algorithms, *Proceedings of the First Int'l Conf. on Parallel Problem Solving from Nature*, Dortmund, Germany, October 1990.

Eschelman, L., Caruana, R. & Schaffer, D. (1989). Biases in the Crossover Landscape, *Proc. 3rd Int'l Conference on Genetic Algorithms*, Morgan Kaufman Publishing.

Holland, John H. (1975). *Adaptation in Natural and Artificial Systems,* The University of Michigan Press.

Spears, W. & De Jong, K. A. (1990). An Analysis of Multi-point Crossover, *Proceedings of the Foundations of Genetic Algorithms Workshop*, Indiana, July 1990.

Syswerda, Gilbert. (1989). Uniform Crossover in Genetic Algorithms, *Proc. 3rd Int'l Conference on Genetic Algorithms*, Morgan Kaufman Publishing.

Appendix

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 *k*th 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 *k*th order hyperplane. The effect of the recursion and summation is to consider every possible placement of *n* crossover points within the *k*th order hyperplane. The correction factor C_s computes the probability that the hyperplane will survive, based on that placement of crossover points. Suppose that crossover results in *x* of the *k* defining positions being exchanged. Then the hyperplane will survive if: 1) the parents match on all *x* positions being exchanged, or 2) if they match on all $k - x$ positions not being exchanged, or 3) they match on all *k* defining positions. Hence, the general form of the correction is:

$$C_s = P_{eq}{}^{x} + P_{eq}{}^{k-x} - P_{eq}{}^{k}$$

where P_{eq} is the probability of two parents sharing an allele at each locus, and the $P_{eq}{}^{k}$ reflects an overlap within the 3 possibilities (and hence must be subtracted).

As an example, consider Figure 8. The two parents are denoted by P1 and P2. In this figure, we represent the survival of a 4th order hyperplane. The hyperplane defining positions are depicted with circles. Since 1 of the defining positions will be exchanged (under the 2-point crossover shown), the probability of survival is:

$$C_s = P_{eq}{}^1 + P_{eq}{}^3 - P_{eq}{}^4$$

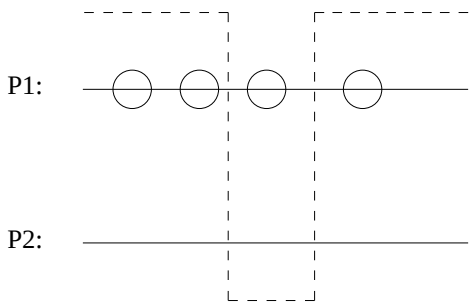

Figure 8: 4th Order Hyperplane Survival

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

$$P_{k,s}(H_k) =$$

$$\sum_{i=0}^{k} \begin{bmatrix} k \\ i \end{bmatrix} (P_0)^i \ (1 - P_0)^{k-i} \ (P_{eq}{}^i + P_{eq}{}^{k-i} - P_{eq}{}^k)$$

where P_0 is the probability of swapping two parents' alleles at each locus. A graphical representation of these equations has been shown previously in Figure 4.

Recombination 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 hyperplanes survive to form a higher order hyperplane. The difference is that the two lower order hyperplanes (each of which exists in one 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 x of the k defining positions surviving in the same individual (i.e., x is a subset of the $m + n$ defining positions), then recombination will occur if: 1) the parents match on all of the x positions, or 2) if they match on all $k - x$ positions, or 3) they match on all k defining positions. Hence, the general form of the recombination correction C_r is:

$$C_r = P_{eq}{}^x + P_{eq}{}^{k-x} - P_{eq}{}^k$$

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.

As an example, consider Figure 9. In this figure, we represent the recombination of 2 2nd order hyperplanes. One hyperplane is depicted with circles, and the other with rectangles. Since 3 of the defining positions will survive onto the same individual (under the 2-point crossover shown), the probability of survival is:

$$C_r = P_{eq}{}^3 + P_{eq}{}^1 - P_{eq}{}^4$$

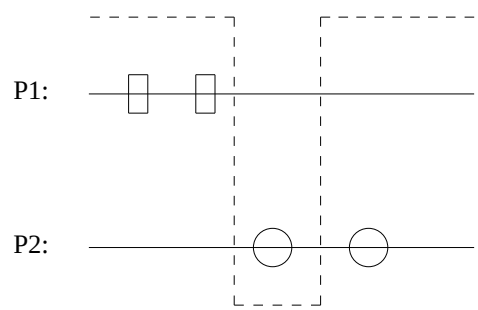

Figure 9: 2nd Order Hyperplane Recombination

Recombination 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:

$$P_{k,s,orig}(H_k) = \sum_{i=0}^{k} \begin{bmatrix} k \\ i \end{bmatrix} (P_0)^i \ (1 - P_0)^{k-i} \ (P_{eq}{}^{k-i})$$

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

$$P_{k,s,other}(H_k) = \sum_{i=0}^{k} \begin{bmatrix} k \\ i \end{bmatrix} (P_0)^i \ (1 - P_0)^{k-i} \ (P_{eq}{}^i)$$

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

$$P_{k,s,both}(H_k) = \sum_{i=0}^{k} \begin{bmatrix} k \\ i \end{bmatrix} (P_0)^i \ (1 - P_0)^{k-i} \ (P_{eq}{}^k) = P_{eq}{}^k$$

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. Finally, it is interesting to note that the last term is equivalent to $P_{eq}{}^m P_{eq}{}^n = P_{eq}{}^k$.