

Running Races with Fraser's Recombination

David B. Fogel

Natural Selection, Inc.
3333 N. Torrey Pines Ct., Suite 200
La Jolla, CA 92037
dfogel@natural-selection.com

Alex S. Fraser

Professor, Emeritus
University of Cincinnati
Dept. of Biological Sciences
2624 Clifton Ave.
Cincinnati, OH 45221

Abstract - Several recombination operators have been proposed in evolutionary computation. Standard procedures include one-point, two-point, and uniform crossover. Little attention, however, has been given to a recombination operator that preceded each of these, which was offered in 1957 by Alex Fraser. Fraser's recombination assigns a variable probability for crossing over between two solutions at each locus. This operator subsumes the three standard forms of crossover. Experiments are conducted on a set of 10 test functions where individuals in a population are tagged to recombine in one of the four above-mentioned ways. Attention is focused on the relative performance of the various methods of recombination. Uniform crossover is seen to offer the greatest utility on the test functions, and Fraser's recombination repeatedly outperforms one-point and two-point crossover. Several issues of the experimental design are discussed.

1 Introduction

Fraser (1957a) offered one of the earliest computer simulations of genetic systems. Simulations were conducted using diploid "organisms" represented by binary strings. (Let the length of the string be described by n .) Each bit in a string corresponded to an "allele" (dominant or recessive) and the "phenotype" (expressed behavior) of each string was determined by its "genetic" composition. Fraser suggested a recombination operator that was a variant of what would now be described as n -point crossover. When two strings were "mated," each position along each string was assigned a probability for crossing over to the other string. Interactions between genes could be addressed by forming *linkage groups* between alleles based on the probabilities of crossing over at each locus. Figure 1 illustrates Fraser's recombination. Fraser's procedure included a population of P parents that gave rise to P' offspring via recombination. Selection then eliminated all but P of the offspring, and all of the parents. (Today, this is referred to as (μ, λ) selection, following

the efforts of Rechenberg and Schwefel, see Bäck et al., 1997.)

Fraser and colleagues studied this general system for over a decade and published a series of papers (Fraser, 1957b; Barker, 1958; Fraser, 1960; and many others), culminating in the book *Computer Models in Genetics* (Fraser and Burnell, 1970). Looking back to these publications, Fraser's algorithm was essentially equivalent to what was later termed a genetic algorithm, following Holland (1975), even including the inversion operator which was introduced in Fraser (1968). Despite the obvious similarities, comparatively little effort has been made with crossover operators that involve probabilities of alternating between solutions, let alone methods that allow such probabilities to evolve as a function of the evolutionary search. One-point and two-point crossover, two common forms (Bäck et al., 1997), choose crossing points uniformly at random. Uniform crossover (Reed et al., 1967; Syswerda, 1989) applies a fixed probability, typically 0.5, to crossing between parents at each locus. Spears and De Jong (1991) suggested that the probability of crossing under uniform crossover could be set to values other than 0.5 (i.e., a fixed probability $p \neq 0.5$). They provided analysis comparing the disruptiveness, in terms of schema processing, of such a parameterized uniform crossover to one- and two-point crossover. Fraser's method of recombination subsumes all of the above methods: By setting the probabilities of crossing over at each locus, the technique can implement one-point, two-point, uniform, or parameterized uniform crossover. The current work describes a series of experiments performed on standard test functions that compares the efficiency of one-point, two-point, and uniform crossover to Fraser's recombination. A procedure is offered to self-adapt the probabilities of crossing over at each locus for individuals that undergo Fraser's recombination. The results indicate an advantage for uniform crossover, but also indicate that Fraser's recombination outperformed one-point and two-point crossover consistently.

$$\text{Individual} \begin{cases} [x_1 & x_2 & \dots & x_n] \longleftarrow \text{Genetic Composition} \\ [p_1 & p_2 & \dots & p_n] \longleftarrow \text{Crossover Vector} \end{cases}$$

Figure 1. In Fraser’s recombination, each individual in the population is represented by a vector \mathbf{x} of genes with an associated crossover vector \mathbf{p} that indicates the probability of switching over to the other parent. Each locus is considered in turn from 1 to n . A sequence of small-valued entries (i.e., close to zero) in the vector \mathbf{p} indicates a linked group of genes because it is unlikely that recombination will disrupt that group.

2 Method and Results

Experiments were conducted using the first 10 functions offered in Yao et al. (1999). These are shown in Table I. The objective was to find the minimum of each function. In each case, a population of 2000 parent solutions was sampled uniformly at random from the bounding area associated with each function. Each solution was randomly assigned a tag that designated the type of recombination that it could perform: one-point crossover, two-point crossover, uniform crossover (with a parameter $p = 0.5$), or Fraser’s recombination. Each solution that was tagged with Fraser’s recombination also carried an additional vector of probabilities for switching between itself and another solution at each locus when mating. The values in this vector were sampled uniformly at random over $[0, 1]$. Evolution was conducted over 50 generations where each solution mated with another solution of its same type, chosen at random from the population. If no other solution of the same type could be found in 100 attempts, the solution served as its own mate. Each mating generated a single solution. For solutions that were tagged with Fraser’s recombination, each element of the associated probability vector for the new offspring was created by copying the first parent’s probability vector. All individuals that were tagged with Fraser’s recombination were then subjected to having the elements in their probability vector modified by:

$$p'_i = p_i \exp(N(0,1))$$

where p_i is the i th component of the vector of probabilities, and $N(0,1)$ is a standard Gaussian random variable, this following a common form of mutation applied in self-adaptive evolutionary algorithms. If the value of p'_i was lower than 0.0 or greater than 1.0, it was reset to the bound it exceeded.

After each parent in turn participated in generating an offspring, all 4000 individuals competed in a tournament for survival. Each individual was paired against 10 randomly chosen opponents from the population: if the (error) score of the individual was less than or equal to its

opponent, it received a “win.” Those solutions with the greatest number of wins were selected to be parents of the next generation.

After the 50th generation, the population was assessed to determine the proportion of each type of tag on the surviving individuals. In the case where no single tag type had completely dominated the population, the trial was recorded as “inconclusive.” A series of 100 independent trials was conducted on each function. In essence, each trial was a “race” between the various forms of crossover to determine which could locate solutions that came to dominate the population.

The results are shown in Table II. Uniform crossover dominated the population in the vast majority of cases across all functions. All results are statistically significant using a chi-square test ($P < 0.01$) in favor of uniform crossover, with the exception of function 3, for which the results are statistically significant in favor of one-point crossover. None of the trials with function 7 converged to a single tag type in 50 generations. Each trial was examined and it was found that of the 100 trials, 17 were dominated by one-point crossover, 20 were dominated by two-point crossover, 39 were dominated by uniform crossover, and 24 were dominated by Fraser’s recombination. Interestingly, with the exceptions of functions 3 and 5, Fraser’s recombination dominated the population more often than either one-point or two-point crossover. Further independent experimentation will be required to determine the statistical significance of that relationship.

Figure 2 shows the results of a trial on function 7 where the number of individuals of each tag type are plotted as a function of the generation number. The most common occurrence was for uniform crossover to rapidly increase and swamp out the other forms of recombination. In this case, however, uniform crossover was surpassed by Fraser’s recombination at generation 14. This apparently came at the expense of one-point crossover, which suffered a steep reduction in its representation in the population. Near the end of the 50 generations, uniform crossover was again gaining representation and it is unclear what the final converged population would have com-

Table I. The 10 functions used for the experiments conducted here, taken from Yao et al. (1999). The ranges for each function are $[-100, 100]^n$, $[-10, 10]^n$, $[-100, 100]^n$, $[-100, 100]^n$, $[-30, 30]^n$, $[-100, 100]^n$, $[-1.28, 1.28]^n$, $[-500, 500]^n$, $[-5.12, 5.12]^n$, and $[-32, 32]^n$, respectively.

$$f_1(\mathbf{x}) = \sum_{i=1}^n x_i^2$$

$$f_2(\mathbf{x}) = \sum_{i=1}^n |x_i| + \prod_{i=1}^n |x_i|$$

$$f_3(\mathbf{x}) = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2$$

$$f_4(\mathbf{x}) = \max_i \left\{ |x_i|, i = 1, \dots, n \right\}$$

$$f_5(\mathbf{x}) = \sum_{i=1}^{n-1} \left[100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2 \right]$$

$$f_6(\mathbf{x}) = \sum_{i=1}^n \left(|x_i + 0.5| \right)^2$$

$$f_7(\mathbf{x}) = \sum_{i=1}^n ix_i^4 + \text{random}[0,1)$$

$$f_8(\mathbf{x}) = \sum_{i=1}^n -x_i \sin \left(\sqrt{|x_i|} \right)$$

$$f_9(\mathbf{x}) = \sum_{i=1}^n \left[x_i^2 - 10 \cos(2\pi x_i) + 10 \right]$$

$$f_{10}(\mathbf{x}) = -20 \exp \left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \right) - \exp \left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i) \right) + 20 + e$$

prised. Figure 3 shows a trial on function 7 where one-point crossover increased its representation in the population after an initial slow beginning.

3 Discussion

For the problems studied, the relative performance offered by the respective variation operators suggests that maintaining building blocks of linked elements within solutions may not be particularly important. (Several of the test problems are known to have independent components.) Only in one case did one-point crossover come to dominate the population with the greatest frequency; two-point crossover never dominated the population with the greatest frequency. In contrast, uniform crossover and Fraser's recombination, both of which are more prone to disrupt linkages between elements than are one- or two-point crossover, offered a more effective search by locating solutions with improved performance that overtook the population. Further investigation will be taken to examine if the evolved probability vectors associated with Fraser's recombination ever converged to values that would suggest the discovery of tightly linked elements (e.g., if the probability vector for a solution that was tagged with Fraser's recombination was $[0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0]$, that would indicate two sets of tightly linked elements).

The use of a lognormal perturbation on the probabilities in the vector of elements controlling Fraser's recombination follows traditional methods of self-adaptation; however, no scaling factor was applied to this perturbation. One possibility for examining the efficacy of Fraser's recombination is to use the scaling factor $\exp(\tau N(0,1))$,

where $\tau = (2n)^{-0.5}$ (see Bäck et al., 1997) or perhaps other values. The sensitivity of the procedure remains to be determined.

When taken across all possible problems, the no free lunch theorems (Wolpert and Macready, 1997; English 1996; and others) mandate that none of the crossover operators tested here will be advantageous in general. Fraser's recombination, however, offers a versatility not found in the standard versions of crossover. It may be that this versatility yields a robustness across a range of problems that would not be found with the standard one-point, two-point, or uniform crossover operators. Discovering such problems remains for future investigation.

The choice of executing each of the trials for 50 generations was arbitrary, but served as a useful termination point in 9 of the 10 cases. For function 7, however, evaluating 100,000 solutions was insufficient. This raises the interesting possibility that the utility of different operators may cycle or present nonstationary patterns as a function of time and races such as those conducted here may, in fact, never converge to a single type of variation operator. Furthermore, it is interesting to consider the possibility for mutating the tag type, thereby re-entering different forms of crossover even after they have gone "extinct." These possibilities are being examined currently. It would also be interesting to examine the quality of evolved solutions under different forms of recombination, including Fraser's recombination, or in the presence of mutation or other variation operators.

The authors hope that the preliminary experiments offered here serve to increase interest in applying Fraser's recombination to other interesting problems in diverse areas.

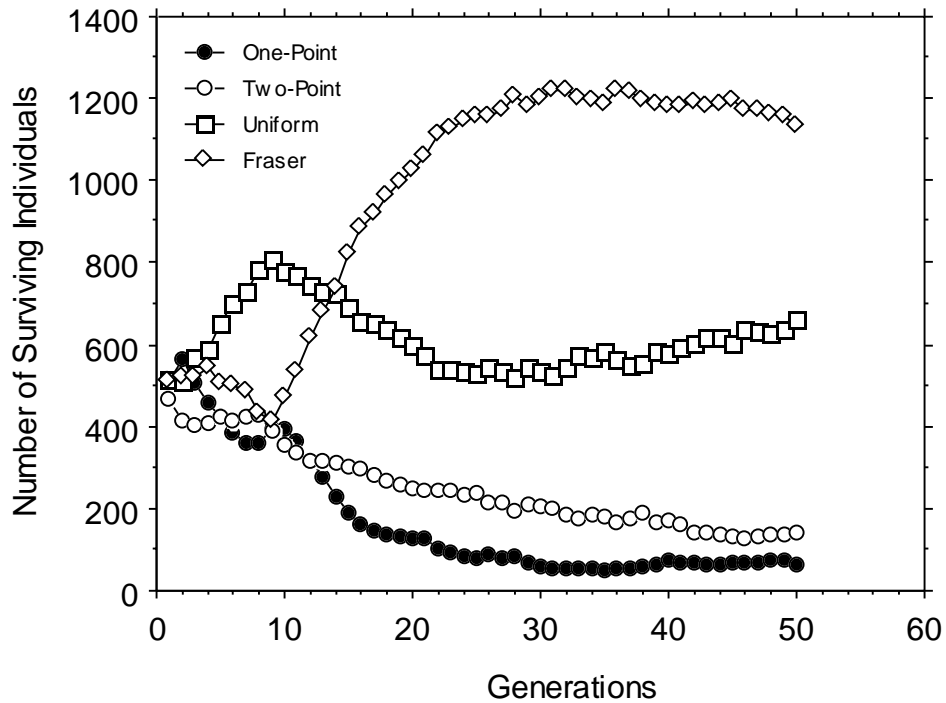


Figure 2. The results of a trial on function 7 where the number of individuals of each tag type is plotted as a function of the generation number. Uniform crossover was surpassed by Fraser’s recombination at generation 14. As Fraser’s recombination increased, one-point crossover decreased, indicating its ineffectiveness at that point in time. After generation 30, uniform crossover was once again increasing its representation.

Table II. The tabulated results from 100 trials on each of the 10 functions. The entries in the table indicate the number of times that the population was completely taken over by a particular type of crossover operator. If the population did not converge to a single type within the maximum of 50 generations, the results were tabulated as “inconclusive.”

Function	1-Point	2-Point	Uniform	Fraser’s	Inconclusive
1	2	11	60	25	2
2	0	13	62	23	1
3	64	11	0	2	23
4	3	18	48	21	10
5	1	17	61	13	8
6	0	5	66	20	9
7	0	0	0	0	100
8	0	9	64	26	1
9	1	11	61	27	0
10	2	11	61	24	2

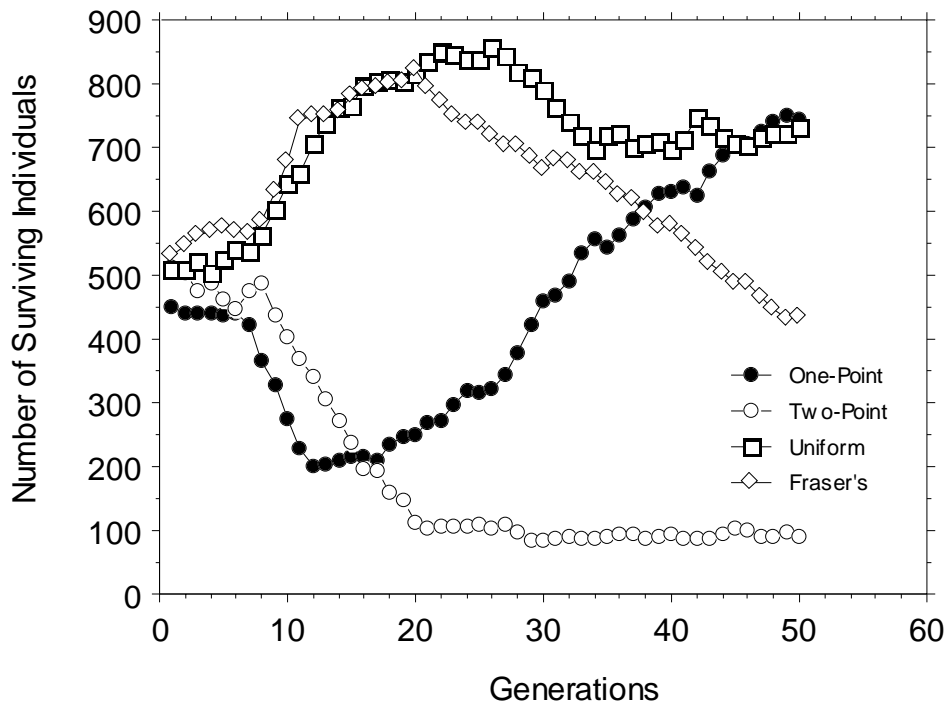


Figure 3. A trial on function 7 where one-point crossover increased its representation in the population after an initial slow beginning. This was an unusual circumstance. In the vast majority of cases, one-point and two-point crossover did not compare well with uniform crossover or Fraser's recombination.

Acknowledgments

The authors thank Dick Lewontin, Gary Fogel, Larry Fogel, and the reviewers for their helpful comments and criticisms of this paper. Thanks also go to Dan Ashlock for organizing the special session.

References

1. T. Bäck, U. Hammel, and H.-P. Schwefel, "Evolutionary computation: Comments on the history and current state," *IEEE Transactions on Evolutionary Computation*, Vol. 1:1, pp. 3-17, 1997.
2. J.S.F. Barker, "Simulation of genetic systems by automatic digital computers. III. Selection between alleles at an autosomal locus," *Australian J. Biological Sciences*, Vol. 11, pp. 603-612, 1958.
3. A.S. Fraser, "Simulation of genetic systems by automatic digital computers. I. Introduction," *Australian J. Biological Sciences*, Vol. 10, pp. 484-491, 1957.
4. A.S. Fraser, "Simulation of genetic systems by automatic digital computers. II. Effects of linkage or rates of advance under selection," *Australian J. Biological Sciences*, Vol. 10, pp. 492-499, 1957.
5. A.S. Fraser, "Monte Carlo analysis of genetic models," *Nature*, Vol. 181, pp. 208-209, 1958.
6. A.S. Fraser, "The evolution of purposive behavior," *Purposive Systems*, H. von Foerster, J.D. White, L.J. Peterson, and J.K. Russell (eds.), Spartan Books, Washington DC, pp. 15-23, 1968.
7. A.S. Fraser and D. Burnell, *Computer Models in Genetics*, McGraw-Hill, NY, 1970.
8. J.H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, MI, 1975.
9. J. Reed, R. Toombs, and N.A. Barricelli, "Simulation of biological evolution and machine learning: I. Selection of self-reproducing numeric patterns by data processing machines, effects of hereditary control, mutation type and crossing," *J. Theoret. Biol.*, Vol. 17, pp. 319-342, 1967.

10. W.M. Spears and K.A. De Jong, "On the virtues on parameterized uniform crossover," *Proc. 4th Intern. Conf. on Genetic Algorithms*, R.K. Belew and L.B. Booker (eds.), Morgan Kaufmann, San Mateo, CA, pp. 230-236, 1991.

11. G. Syswerda, "Uniform crossover in genetic algorithms," *Proc. 3rd Intern. Conf. Genetic Algorithms*, J.D. Schaffer (ed.), Morgan Kaufmann, San Mateo, CA, pp. 2-9, 1989.

12. X. Yao, Y. Liu, and G. Lin, "Evolutionary programming made faster," *IEEE Transactions on Evolutionary Computation*, Vol. 3:2, pp. 82-102, 1999.