University of Montana

ScholarWorks at University of Montana

Computer Science Faculty Publications

Computer Science

2002

Group Properties of Crossover and Mutation

Jonathan E. Rowe

Michael D. Vose

Alden H. Wright
University of Montana - Missoula, alden.wright@umontana.edu

Recommended Citation

Rowe, Jonathan E.; Vose, Michael D.; and Wright, Alden H., "Group Properties of Crossover and Mutation" (2002). *Computer Science Faculty Publications*. 11.
https://scholarworks.umt.edu/cs_pubs/11

Group Properties of Crossover and Mutation

Jonathan E. Rowe   J.E.Rowe@cs.bham.ac.uk
School of Computer Science, University of Birmingham, Birmingham B15 2TT, UK

Michael D. Vose   vose@cs.utk.edu
Computer Science Department, University of Tennessee, Knoxville, TN 37996, USA

Alden H. Wright   wright@cs.umt.edu
Department of Computer Science, University of Montana, Missoula, MT 59812, USA

Abstract

It is supposed that the finite search space Ω has certain symmetries that can be described in terms of a group of permutations acting upon it. If crossover and mutation respect these symmetries, then these operators can be described in terms of a mixing matrix and a group of permutation matrices. Conditions under which certain subsets of Ω are invariant under crossover are investigated, leading to a generalization of the term *schema*. Finally, it is sometimes possible for the group acting on Ω to induce a group structure on Ω itself.

1 Introduction

This is the first of a series of papers that sets out to investigate the relationship between the choice of representation used for a search problem and the genetic operators (crossover and mutation) that act upon such representations. The main insight is that there are often certain structural symmetries within combinatorial search spaces, and it is possible to construct genetic operators that have interesting properties with respect to those symmetries. The mathematics of group theory is used throughout to describe these properties. Recall that a group is a set L together with a binary operation $\circ$ defined on elements of L such that:

1. $a \circ b \in L$ for all $a, b \in L$.

2. $a \circ (b \circ c) = (a \circ b) \circ c$ for all $a, b, c \in L$.

3. There exists an element $0 \in L$, called the identity, such that $a \circ 0 = 0 \circ a = a$ for all $a \in L$.

4. For each element $a \in L$, there exists an inverse element $a^{-1} \in L$ such that $a \circ a^{-1} = a^{-1} \circ a = 0$.

If the group operator is commutative, so that $a \circ b = b \circ a$ for all $a, b \in L$, then the group is said to be *Abelian*.

The paper is organized as follows. Section 2 formally defines the simple genetic algorithm, which will be the focus of this work. Section 3 looks at some general implications of having a group acting on the search space. That is, we have a set of permutations of the search space that forms a group under function composition. This leads to conditions under which the genetic operators (crossover and mutation) commute with the natural group action on the space of populations. This in turn leads to a general definition of the mixing scheme in terms of a *mixing matrix*. Section 4 incorporates Radcliffe's concepts of *pure* and *respectful* crossover, by studying invariant sets. We then present a generalization of the term *schema* and look for conditions under which schemata are respected by crossover.

It is possible that the group action on the search space Ω can induce a group structure on Ω itself. That is, we can turn the search space into a group in its own right. This possibility is studied in Section 5. One way in which such symmetries might arise is through considering the neighborhood structure of a search landscape and the automorphisms that act upon it. This will be the theme of a future paper.

The sequel to this paper (Rowe et al., 2001) will concentrate on the case where Ω is itself a group and will introduce the class of *structural* operators for which important invariance results can be proved. The special case when Ω is Abelian will be examined (as in the case of fixed-length binary strings, for example) and results concerning the effects of the Fourier transform presented.

To summarize, the main goals of these papers are:

- To study the interplay between search space structure and the genetic operators.

- To show how such relationships affect the genetic algorithm at the population level.

- To give conditions under which crossover is invariant on certain subsets of the search space.

- To provide methods by which genetic operators with certain desirable properties can be constructed.

Notation

Angle brackets $\langle \cdots \rangle$ denote a tuple that is to be regarded as a column vector. Indexing of vectors and matrices begins with 0. Transpose is indicated with superscript T. The column vector of all 1s is denoted by $\mathbf{1}$. The $n \times n$ identity matrix is I, and the jth column of the identity matrix is the vector e_j. For vector x, $\mathrm{diag}(x)$ denotes the square diagonal matrix with i, ith entry x_i. The number of nonzero components of x is $\#x$.

Composition of functions f and g is $f \circ g(x) = f(g(x))$. Square brackets $[\cdots]$ are, besides their standard use as specifying a closed interval of real numbers, used to denote an indicator function: if *expr* is an expression that may be true or false, then

$$[expr] = \begin{cases} 1 & \text{if } expr \text{ is true} \\ 0 & \text{otherwise} \end{cases}$$

2   The Simple Genetic Algorithm

This section introduces the *Simple Genetic Algorithm* as a specific instance of random heuristic search (Vose, 1999).

Let Ω be a finite set of cardinality n, referred to as the *search space*. To simplify notation, Ω is identified with the set of integers $\{0, 1, \ldots, n-1\}$. A *fitness function* f satisfies

$$f : \Omega \to \Re^+$$

where $\Re^+$ is the set of positive reals. A *population P* is a collection (multiset) of elements from Ω of size r, the *population size*.

The particular type of genetic algorithm (GA) considered in this paper produces the next generation from the current population as follows:

1. Obtain two parents by a selection method that can be modeled by random heuristic search (such as proportional, ranking, or tournament selection).

2. Mutate the parents.

3. Crossover the two mutated parents to form an offspring.

4. Put the offspring into the next generation.

5. If the next generation is not full, go to step 1.

In particular, there are few constraints placed on mutation and crossover.[1] For more details on modeling selection by random heuristic search, see Vose (1999).

The behavior of the GA is modeled using the following state space, which is said to *correspond* to the search space,

$$\Lambda = \left\{ p \in \Re^n : p_k \geq 0, \ \sum p_k = 1 \right\}$$

A population P corresponds to the point $p \in \Lambda$ defined by

$$p_k = \text{the proportion of } P \text{ occupied by } k \in \Omega$$

The progression from one generation to the next is stochastic due to the randomness inherent in selection (and also, perhaps, inherent in mutation and crossover). This is represented by the operator

$$\mathcal{G} : \Lambda \to \Lambda$$

where, given $p \in \Lambda$, the vector $\mathcal{G}(p)$ has kth component equal to the probability that $k \in \Omega$ will appear in the next generation. This operator is usually divided into two components,

$$\mathcal{G} = \mathcal{M} \circ \mathcal{F}$$

where the *selection scheme $\mathcal{F}$* represents the effect of selection, and the *mixing scheme $\mathcal{M}$* represents the combined effects of crossover and mutation. For proportional selection, $\mathcal{F}$ is defined by

$$\mathcal{F}(p) = \frac{\text{diag}(f)\, p}{f^T p}$$

where f is the fitness vector $f_k = f(k)$, and $\text{diag}(f)$ is the diagonal matrix

$$\text{diag}(f)_{k,k} = f_k$$

---

[1] Mutation is a stochastic unary operator, crossover is a stochastic binary operator.

Other methods of selection such as ranking or tournament selection could also be used. See Vose (1999) for more details.

This paper and its sequel are concerned mostly with properties of the mixing scheme $\mathcal{M}$. In particular, properties of $\mathcal{M}$ with respect to any group structure that may exist on Ω will be studied; if the search space has some structural symmetries (which in practice, it often does), it is of interest to consider the effects that crossover and mutation have on those symmetries and, moreover, how genetic operators can be designed that respect such properties.

By placing the analysis of genetic operators into such a general setting, theorems will be proved at an abstract level. General results will be obtained for: representing crossover and mutation via a mixing matrix, studying sets for which crossover is invariant, and defining genetic operators in terms of landscape structure. In the sequel, further results will be given for: implementing genetic operators via probability distributions over binary masks, properties of the Fourier transform, and implicit parallelism (which may be thought of as a schema theorem). Whereas similar results have already been obtained in particular cases (for example, when the search space comprises fixed-length binary strings), this paper establishes results in a more general context, which will therefore be more widely applicable.

3 Group Action and Quadratic Operators

3.1 Symmetries in the Search Space

Suppose $(L, \circ)$ is a finite group that acts on Ω. That is, there is a mapping $L \times \Omega \to \Omega$, such that, denoting the image of (k, w) by $k(w)$,

$$(a \circ b)(w) = a(b(w))$$

for all $a, b \in L$, $w \in \Omega$, and

$$(a^{-1} \circ a)(w) = (a \circ a^{-1})(w) = w$$

where a inverse is denoted by a^{-1}. In particular, each element of L corresponds to a permutation (bijection) of the set Ω, and the identity element of L corresponds to the identity function. In so far as notation is concerned, the group operation behaves like function composition (which, for convenience, is why the group operation of L is denoted by $\circ$) and, like composition, it is not necessarily Abelian ($a \circ b$ need not equal $b \circ a$).

The group action of L on Ω is denoted by $L(\Omega)$, and it is assumed to be *transitive*, i.e., for every $u, v \in \Omega$, there exists $a \in L$ such that $a(u) = v$. By abuse of notation, L will be called transitive when $L(\Omega)$ is. The *isotropy group* Ω_L *of* $L(\Omega)$ is the normal subgroup (of L)

$$\Omega_L = \{k \in L \ : \ \forall w \,.\, k(w) = w\}$$

The quotient group L/Ω_L acts naturally on Ω by

$$(a \circ \Omega_L)(w) = a(w)$$

A transitive group action $L(\Omega)$ is called *reduced* if Ω_L contains only the identity. It is assumed throughout that $L(\Omega)$ is reduced—if not, simply replace L by L/Ω_L.

3.2 Examples

Throughout this paper, we will illustrate our results with a number of examples. Some of these examples will be familiar to readers, but others will be new. We hope that this variety will illustrate the generality and power of our work.

Binary Strings An important special case of a group action on a search space is when Ω itself has a group structure. The action of Ω on itself is defined by

$$u(v) = u \circ v$$

for all $u, v \in \Omega$. The group action is naturally transitive. This situation arises, for example, when the search space comprises the set of fixed-length binary strings of length ℓ. The group operator is bitwise addition modulo 2. For example, if $\ell = 5$ and $u = 01101$, $v = 11110$, then

$$u(v) = 01101(11110) = 01101 \oplus 11110 = 10011$$

The identity element is 00000.

Traveling Salesman Another example of where Ω is a group acting on itself is when it is the set of all permutations of a set of points, as is the case for the traveling salesman problem. The group operation is simply the composition of permutations. For example, if there are five cities $\{A, B, C, D, E\}$, we can denote a tour by a vector $\tau = \langle \tau(1), \tau(2), \tau(3), \tau(4), \tau(5) \rangle$, where $\tau(k)$ is the kth city visited in the tour. Given two tours τ_1 and τ_2, we combine them as if they were functions:

$$(\tau_1 \circ \tau_2)(k) = \tau_1(\tau_2(k))$$

Thus $\tau_1 = \langle B, C, E, D, A \rangle$ and $\tau_2 = \langle E, D, C, B, A \rangle$ combine to give the tour $\tau_1 \circ \tau_2 = \langle A, D, E, C, B \rangle$. The group of all possible permutations of m objects is known as the *symmetric group* (denoted S_m), and it contains $m!$ elements.

Assigning Jobs to Processors Suppose we have four processors connected in a ring: A is connected to B that is connected to C that is connected to D that is connected to A (see Figure 1). The connections have different bandwidths. We have four jobs numbered 1, 2, 3 and 4 that we need to assign to the processors (one job per processor), subject to the following constraints:

>   1 and 2 must be directly connected
>   2 and 3 must be directly connected
>   3 and 4 must be directly connected
>   4 and 1 must be directly connected

The aim is to find an assignment of jobs to processors that minimizes overall communication time. Because of the constraints, not all assignments are valid. The set of all valid assignments forms our search space Ω.

Now consider the set of permutations of the processors. This set forms a group. Any permutation can be interpreted as a way of specifying a method for re-assigning jobs. For example, the permutation $\langle D, A, B, C \rangle$ specifies that the job on processor A should be moved to D; the job on B should be moved to A; the job on C to B; and that on D to C. This corresponds to rotating the jobs through
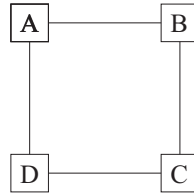
Figure 1: A network of processors.

90°. Applying this permutation to the assignment $\langle 1 \rightarrow A, 2 \rightarrow B, 3 \rightarrow C, 4 \rightarrow D \rangle$ produces another assignment $\langle 1 \rightarrow D, 2 \rightarrow A, 3 \rightarrow B, 4 \rightarrow C \rangle$.

Not all permutations will give rise to valid assignments, however. We will therefore restrict our attention to the set of permutations that always transforms valid assignments into valid assignments. Such a permutation is called an *automorphism* and the set of automorphisms forms a subgroup of the group of permutations, called the *automorphism group*. This will be our group action $L(\Omega)$.

There are three possible rotations of the network. Other automorphisms come from swapping certain jobs, by reflections of the network. If we include the identity (don't move anything!), this gives a total of eight possible re-arrangements that preserve the constraints in our network.

A	B	C	D	group element action
A	B	C	D	the identity
D	A	B	C	90° rotation
C	D	A	B	180° rotation
B	C	D	A	270° rotation
C	B	A	D	reflection through B-D diagonal
A	D	C	B	reflection through A-C diagonal
B	A	D	C	reflection through vertical
D	C	B	A	reflection through horizontal

This set forms a group (known as the *dihedral* group D_4) that acts on our search space of possible assignments. Given one valid assignment of jobs, we can apply any of our automorphisms to get another. It can be checked that the group action is transitive.

3.3 Symmetries in Population Space

We now look at how symmetries in the search space Ω give rise to corresponding symmetries in the space of possible populations. We will then be able to describe the relationship between the group action on Ω and the genetic operators.

Since the population state space Λ is comprised of vectors whose components are in one-to-one correspondence with elements of the search space, Ω may be identified with those components, and, under this identification, each $a \in L$ corresponds to a permutation matrix σ_a, where $\sigma_a : \Lambda \longrightarrow \Lambda$ by

$$\sigma_a \langle v_0, \ldots, v_{n-1} \rangle = \langle v_{a^{-1}(0)}, \ldots, v_{a^{-1}(n-1)} \rangle$$

Alternatively—and equivalently—define for each $a \in L$ the permutation matrix σ_a by

$$(\sigma_a)_{i,j} = [i = a(j)]$$

It is easily checked that $\sigma_{a^{-1}} = \sigma_a^{-1} = \sigma_a^T$ and $\sigma_a \sigma_b = \sigma_{a \circ b}$. It follows that these permutations form a group (under matrix multiplication), which because $L(\Omega)$ is reduced, is isomorphic to L. Note that

$$\sigma_a e_i = e_{a(i)}$$

and for any matrix M (of suitable dimension),

$$(\sigma_a^T M \sigma_a)_{i,j} = M_{a(i), a(j)}$$

The effects of crossover and mutation on a population can be described in terms of operators that map points in the population state space Λ to other such points. To characterize their relationship with the group action, we first define two classes of operator: quadratic and linear. It will turn out that crossover and mutation, defined quite generally, belong to these classes.

An operator $A : \Re^n \longrightarrow \Re^n$ is *quadratic* if there exist matrices $M_0, \ldots, M_{n-1}$ (depending on A) such that

$$A(x) = \langle x^T M_0 x, \ldots, x^T M_{n-1} x \rangle$$

In the notation above, a quadratic operator A does not uniquely determine the associated matrices $M_0, \ldots, M_{n-1}$. However, they are uniquely determined if they are symmetric. Because replacing M_k by $(M_k + M_k^T)/2$ does not alter the quadratic operator, both symmetry and uniqueness of the associated matrices will, without loss of generality, be assumed.

An operator $B : \Re^n \longrightarrow \Re^n$ is *linear* if there exists a matrix B such that

$$B(x) = Bx$$

The following theorem also appears in Rowe (2001).

THEOREM 1: *If A is quadratic and B is linear, then $A \circ B$ and $B \circ A$ are quadratic.*

PROOF: Suppose A has associated matrices M_k. Then $B \circ A$ has associated matrices M_k' defined by

$$
\begin{aligned}
(B \circ A(x))_k &= \sum_j B_{k,j} \, x^T M_j x \\
&= x^T (\sum_j B_{k,j} M_j) x \\
&= x^T M_k' x
\end{aligned}
$$

Similarly $A \circ B$ has associated matrices M_k'' defined by

$$
\begin{aligned}
(A \circ B(x))_k &= (Bx)^T M_k (Bx) \\
&= x^T (B^T M_k B) x \\
&= x^T M_k'' x
\end{aligned}
$$

$\square$

Note that the identity function $\mathcal{I} : \Lambda \longrightarrow \Lambda$ is an example of a quadratic operator with associated matrices $(M_k)_{i,j} = ([i = k] + [j = k])/2$. By the preceding theorem, every linear operator $B : \Lambda \longrightarrow \Re^n$ is quadratic since

$$B = B \circ \mathcal{I}$$

This means that the class of linear operators on Λ is a subset of the class of quadratic operators.

The important relationship between quadratic operators and the group action on Ω can now be given. A quadratic operator $\mathcal{C}$ is said to *commute with* $L(\Omega)$ if, for all a,

$$\mathcal{C} \circ \sigma_a = \sigma_a \circ \mathcal{C}$$

In the following theorem, we give necessary and sufficient conditions for a quadratic operator to commute with the group action. In it, we identify a particular element of Ω as a reference point. This choice is arbitrary, and for convenience we call the chosen element 0.

THEOREM 2: *A quadratic operator $\mathcal{C}$ commutes with $L(\Omega)$ if and only if its associated matrices $M_0, \ldots, M_{n-1}$ satisfy $M_{a(0)} = \sigma_a M_0 \sigma_a^T$ for all $a \in L$.*

PROOF: Suppose that $\mathcal{C}$ commutes with $L(\Omega)$. That is, $\mathcal{C}(\sigma_a x) = \sigma_a(\mathcal{C}(x))$ for all $a \in L(\Omega)$ and for all $x \in \Lambda$. Now

$$(\mathcal{C}(\sigma_a x))_k = (\sigma_a x)^T M_k(\sigma_a x) = x^T \sigma_a^T M_k \sigma_a x$$

and

$$(\sigma_a(\mathcal{C}(x)))_k = x^T M_{a^{-1}(k)} x$$

So choosing k such that $a(0) = k$ we get

$$\sigma_a^T M_{a(0)} \sigma_a = M_0$$

and so

$$M_k = \sigma_a M_0 \sigma_a^T$$

for all $a \in L(\Omega)$.

Suppose that $M_{a(0)} = \sigma_a M_0 \sigma_a^T$ for all $a \in L$. Let $b(0) = i$. Then

$$\begin{aligned}
\mathcal{C}(\sigma_a x)_i &= (\sigma_a x)^T M_i (\sigma_a x) \\
&= x^T \sigma_a^T (\sigma_b M_0 \sigma_b^T) \sigma_a x \\
&= x^T (\sigma_a^{-1} \sigma_b) M_0 (\sigma_a^{-1} \sigma_b)^T x \\
&= x^T M_{(a^{-1} \circ b)(0)} x \\
&= x^T M_{a^{-1}(i)} x \\
&= (\sigma_a \mathcal{C}(x))_i
\end{aligned}$$

$\square$

Given any element $k \in \Omega$, an important subgroup of $L(\Omega)$ is the set of all elements of the group that leave k unchanged. That is, we define

$$\mathbf{Fix}(k) = \{a \in L(\Omega) : a(k) = k\}$$

As a corollary to the theorem we have:

COROLLARY 3: *If a quadratic operator $\mathcal{C}$ with associated matrices $M_0, \ldots, M_{n-1}$ commutes with $L(\Omega)$, then M_0 commutes with $\mathbf{Fix}(0)$, for all $k \in L$.*

PROOF: Let $a \in \mathbf{Fix}(0)$. Then

$$M_0 = M_{a(0)} = \sigma_a M_0 \sigma_a^T$$

and so $M_0 \sigma_a = \sigma_a M_0$. □

Notice that if Ω is itself a group, then $\mathbf{Fix}(0)$ only contains the identity, and so naturally any matrix commutes with it. However, if this is not the case, then the above theorem says, among other things, that not all matrices are suitable for constructing quadratic operators that commute with L. For example, suppose that $\Omega = \{0, 1, 2\}$ and $L(\Omega)$ is the group of all possible permutations on Ω. Then

$$\mathbf{Fix}(0) = \{(), (12)\}$$

(where we are writing permutations in cycle form). For a matrix M to commute with this subgroup requires

$$M_{i,j} = M_{(12)(i),(12)(j)}$$

for all i, j. Since, moreover, we require our matrices to be symmetric, suitable matrices must have the form

$$\begin{pmatrix} A & B & B \\ B & D & E \\ B & E & D \end{pmatrix}$$

The theorem further points out that if $\mathcal{C}$ commutes with $L(\Omega)$, then all its matrices are permutations of one another, the permutations being given by the group action.

The following theorem shows that all reduced Abelian groups have $\mathbf{Fix}(0)$ containing only the identity, and so again all matrices commute with this subgroup. A much stronger implication will be discussed in Section 5.

THEOREM 4: *If $L(\Omega)$ is a reduced Abelian group action on Ω, then $\mathbf{Fix}(k) = \{0\}$ for all $k \in \Omega$.*

PROOF: Let $a \in \mathbf{Fix}(k)$, and let $g \in L$. If L is Abelian, then $g \circ a = a \circ g$. For any $i \in \Omega$, we can pick g so that $g(k) = i$, since $L(\Omega)$ is transitive. Then

$$a(i) = a \circ g(k) = g \circ a(k) = g(k) = i$$

and therefore a fixes all elements of Ω. But L is reduced and so a must be the identity. □

3.4 Crossover and Mutation

In this section, we show that any crossover and mutation operators, defined quite generally, are quadratic and linear operators, respectively.

Given any crossover operator $\mathcal{X}$, let $r(i, j, k)$ be the probability that parents i and j have child k,

$$r(i, j, k) = \text{Prob}\{\mathcal{X}(i, j) = k\} \tag{1}$$

The *crossover scheme* $\mathcal{C} : \Lambda \to \Lambda$ is defined by

$$\mathcal{C}(x)_k = \text{Prob}\{\mathcal{X}(i, j) = k \mid i \text{ and } j \text{ are chosen uniformly from } x\}$$

where the choice of i and j is independent and with replacement. This definition is symmetric with respect to i and j, suggesting that the natural quantities associated with $\mathcal{C}$ are

$$s(i, j, k) = (r(i, j, k) + r(j, i, k))/2 \tag{2}$$

THEOREM 5: $\mathcal{C}$ *is a quadratic operator with associated matrices*

$$(M_k)_{i,j} = s(i, j, k)$$

Moreover, $\mathcal{C}$ commutes with $L(\Omega)$ if and only if

$$s(a(i), a(j), a(k)) = s(i, j, k)$$

for all $a \in L$.

PROOF: By definition,

$$\mathcal{C}(x)_k = \sum_{i,j} x_i x_j r(i, j, k) = x^T M_k \, x$$

Assume $s(a(i), a(j), a(k)) = s(i, j, k)$ for all $i, j, k \in \Omega$, and $a \in L$. Let $k = a(0)$, then

$$(M_k)_{i,j} = s(i, j, k) = s(a^{-1}(i), a^{-1}(j), 0) = (M_0)_{a^{-1}(i), a^{-1}(j)}$$

and therefore

$$M_{a(0)} = M_k = \sigma_{a^{-1}}^T M_0 \, \sigma_{a^{-1}} = \sigma_a M_0 \, \sigma_a^T$$

so that Theorem 2 applies. Conversely, if

$$M_{a(0)} = \sigma_a M_0 \, \sigma_a^T$$

for all $a \in L$, then the preceding equalities show

$$s(i, j, k) = s(a(i), a(j), a(k))$$

$\square$

Given any mutation operator μ, let $U_{i,j}$ be the probability that j mutates to i,

$$U_{i,j} = \text{Prob}\{\mu(j) = i\}$$

The *mutation scheme $\mathcal{U} : \Lambda \to \Lambda$* is defined by

$$\mathcal{U}(x)_k = \text{Prob}\{\mu(i) = k \mid i \text{ is chosen uniformly from } x\}$$

THEOREM 6: $\mathcal{U}$ *is a linear operator. Moreover, it commutes with $L(\Omega)$ if and only if*

$$U_{a(i), a(j)} = U_{i,j}$$

PROOF: By definition, $\mathcal{U}(x)_i = \sum_j x_j \, U_{i,j}$, hence $\mathcal{U}$ is linear. Note that

$$\sigma_a U = U \sigma_a$$
$$\Longleftrightarrow \quad U = \sigma_a^T U \sigma_a$$
$$\Longleftrightarrow \quad U_{i,j} = U_{a(i),\, a(j)}$$

$\square$

The *mixing scheme* $\mathcal{M}$ is the composition

$$\mathcal{M} = \mathcal{C} \circ \mathcal{U}$$

By Theorems 1, 5, and 6, the mixing scheme is quadratic; let its associated matrices be $M_0, \ldots, M_{n-1}$. The *mixing matrix* is M_0. By Theorem 2, the mixing matrix determines the mixing scheme if crossover and mutation commute with $L(\Omega)$, since in that case

$$\mathcal{M} \circ \sigma_k = \mathcal{C} \circ \mathcal{U} \circ \sigma_k = \mathcal{C} \circ \sigma_k \circ \mathcal{U} = \sigma_k \circ \mathcal{C} \circ \mathcal{U} = \sigma_k \circ \mathcal{M}$$

Each matrix in the mixing scheme is defined by the mixing matrix and the group action by

$$M_{a(0)} = \sigma_a M_0 \sigma_a^T$$

for all $a \in L(\Omega)$.

As has already been observed, the identity function $\mathcal{I} : \Lambda \longrightarrow \Lambda$ is quadratic. It is also linear. Because the concept of crossover allows $\mathcal{C} = \mathcal{I}$, and the concept of mutation allows $\mathcal{U} = \mathcal{I}$, crossover and mutation are special cases of mixing and therefore have corresponding mixing matrices.

The matrix that has $r(i, j, 0)$ as its i, jth entry will be referred to as the *crossover matrix*. The mixing matrix has $s(i, j, 0)$ as its i, jth entry and is a symmetrized version of the crossover matrix. Two different crossover operators may have different crossover matrices but the same mixing matrix. This means that although they may be defined to act differently on individuals, the net effect on the population is the same.

3.5 Examples

Binary Strings An early example of mixing commuting with a group action on Ω was given by Vose (1990) for the case of 1-point crossover (with crossover rate), mutation corresponding to a mutation rate, and fixed-length binary strings. In this situation, Ω is itself a group under bitwise addition modulo 2. More generally, all the standard crossovers (one-point, n-point, uniform) defined on fixed-length strings of cardinality c commute with $L(\Omega)$ (where the group operator is component-wise addition modulo 2) as do mutation operators defined by way of masks (Vose, 1999). More specific examples of these types of crossover and mutation are given later in the paper.

Traveling Salesman A paper by Vose and Whitley (1999) gives general conditions under which crossover defined on the search space Ω of permutations (such as for the traveling salesman problem) commute with $L(\Omega)$, where $L = \Omega$ considered as a group under composition. As a specific example of the crossovers they considered (which commute with $L(\Omega)$), consider order crossover 1, as defined by Davis (1985) and described in Whitley and Yoo (1995) and Vose and Whitley (1999). Given parents $P1$ and $P2$, pick a contiguous "crossover section" from $P1$. (The crossover section does not consist of the whole string and does not wrap around.)

The crossover section is copied to the offspring in the same absolute position that it occupied in $P1$. The remaining elements are copied from $P2$ based on their relative order in $P2$. We refer to the *filler block* as the ordered set of elements from $P2$ with elements also found in the crossover section deleted. The order of the filler block is determined starting at the beginning of $P2$. Elements are added to the offspring from the filler block starting at the end of the crossover section. The following example illustrates.

```
Parent P1:  A B C D E F G H I    crossover-section: _ _ C D E F _ _ _
Parent P2:  d i f g b e c h a    filler-block: i g b h a
Offspring:  h a C D E F i g b
```

If $S = \{\text{A}, \text{B}, \text{C}\}$, and the permutations of S are ordered as

$$\langle \text{A}, \text{B}, \text{C}\rangle, \langle \text{B}, \text{A}, \text{C}\rangle, \langle \text{B}, \text{C}, \text{A}\rangle, \langle \text{A}, \text{C}, \text{B}\rangle, \langle \text{C}, \text{A}, \text{B}\rangle, \langle \text{C}, \text{B}, \text{A}\rangle$$

in function notation or as

$$(), (\text{AB}), (\text{ABC}), (\text{BC}), (\text{ACB}), (\text{AC})$$

in cycle notation, then the crossover matrix for order crossover 1 is:

$$
\begin{bmatrix}
4/5 & 3/5 & 4/5 & 3/5 & 4/5 & 3/5 \\
1/5 & 0 & 0 & 1/5 & 1/5 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
1/5 & 1/5 & 1/5 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/5 & 0 & 1/5 & 1/5
\end{bmatrix}
$$

The i, j entry of this matrix corresponds to $r(i, j, 0)$ of Equation (1). The matrix can be made symmetric by using Equation (2). The resulting symmetric matrix is

$$
\begin{bmatrix}
4/5 & 2/5 & 2/5 & 2/5 & 2/5 & 3/10 \\
2/5 & 0 & 0 & 1/5 & 1/10 & 0 \\
2/5 & 0 & 0 & 1/10 & 0 & 1/10 \\
2/5 & 1/5 & 1/10 & 0 & 0 & 0 \\
2/5 & 1/10 & 0 & 0 & 0 & 1/10 \\
3/10 & 0 & 1/10 & 0 & 1/10 & 1/5
\end{bmatrix}
$$

This matrix is the mixing matrix for the crossover scheme defined by order crossover 1, and for no mutation.

Assigning Jobs to Processors Consider again the problem of placing four jobs onto four processors A, B, C, and D arranged in a ring (subject to the same constraints as before). A possible mutation that preserves the constraints is to swap the jobs on A and C and simultaneously swap B and D. This corresponds to performing a $180°$ rotation. One could assign a certain probability μ of doing this. The resulting

mutation then commutes with the group action. The reason for this is that this automorphism commutes with all the other automorphisms. That is, if we denote the $180°$ rotation as ρ then $\rho \circ a = a \circ \rho$ for all automorphisms $a \in L(\Omega)$. Then we set

$$U_{i,j} = \mu[\rho(j) = i]$$

and so for any automorphism a

$$U_{a(i),a(j)} = \mu[\rho(a(j)) = a(i)] = \mu[a(\rho(j)) = a(i)] = \mu[\rho(j) = i] = U_{i,j}$$

More generally, for any group L, we define its *center* $Z(L)$ to be the set containing those elements that commute with every element in the group:

$$Z(L) = \{a \in L : a \circ b = b \circ a, \text{ for all } b \in L\}$$

The center of a group forms a normal subgroup. We can place a probability distribution μ over the center and define mutation to be:

$$U_{i,j} = \sum_{a \in Z(L)} \mu(a)[a(j) = i]$$

and again we get a mutation operator that commutes with $L(\Omega)$. Notice that this always includes the identity, so we can have a given probability that no mutation will take place.

Defining a crossover operator that commutes with the group action (and preserves the constraints) is a much trickier job. In fact, one of the main goals of this paper and its sequel is to develop a systematic approach to this problem in general. The next subsection will characterize all the possible mixing matrices that are allowable. Conversely, if we have already defined a crossover operator, then we can characterize the natural group actions that commute with it.

3.6 Symmetries of Genetic Operators

We now characterize all possible mixing schemes (crossovers and mutations) that will commute with the group action $L(\Omega)$ in terms of their mixing matrices. We start, however, with the converse problem: Given a particular crossover and mutation operator, what are the groups with which it commutes?

Given a particular definition of crossover and mutation, the reduced group actions $L(\Omega)$ that commute with the corresponding mixing scheme can be characterized. First, note that the mixing scheme is quadratic. Second, note that by what has already been observed, any given reduced group action $L(\Omega)$ induces a group action $L(\Lambda)$ of permutation matrices acting on Λ. Moreover, by identifying Ω with the vertices of Λ through the correspondence

$$i \in \Omega \longleftrightarrow e_i \in \Lambda$$

a group action $L(\Omega)$ is recovered from a group action $L(\Lambda)$. Third, in view of the equivalence of $L(\Omega)$ and $L(\Lambda)$ as described above, L is therefore without loss of generality a group of permutation matrices that can be thought of as acting on both Λ and Ω. For $j \in \Omega$ and permutation matrix σ, therefore, $\sigma(j)$ is that element $k \in \Omega$ satisfying $\sigma(e_j) = e_k$. In particular, $\sigma_a(j) = a(j)$.

Given a quadratic operator $\mathcal{C}$ with associated matrices M_k, let $H(\mathcal{C})$ be the set of all permutation matrices σ acting on Λ (and therefore acting on Ω as described above)—where σ need not be an element of L—such that for all k

$$M_{\sigma(k)} = \sigma M_k \sigma^T$$

THEOREM 7: *The set $H(\mathcal{C})$ is a group, and $L(\Omega)$ is a reduced group action that commutes with $\mathcal{C}$ if and only if L is a transitive subgroup of $H(\mathcal{C})$.*

PROOF: Suppose $M_{\sigma(k)} = \sigma M_k \sigma^T$ and $M_{\sigma'(k)} = \sigma' M_k \sigma'^T$ for all k.

$$
\begin{aligned}
M_{\sigma \circ \sigma'(k)} &= \sigma M_{\sigma'(k)} \sigma^T \\
&= \sigma \sigma' M_k \sigma'^T \sigma^T \\
&= \sigma \sigma' M_k (\sigma \sigma')^T \\
&= (\sigma \circ \sigma') M_k (\sigma \circ \sigma')^T
\end{aligned}
$$

Since H is a collection of permutations of a finite set and is closed under composition, it follows that $(H, \circ)$ is a group. By definition, a reduced group action is transitive. Given σ and k, let $\sigma' \in L$ be such that $\sigma'(0) = \sigma(k)$. If $L(\Omega)$ commutes with $\mathcal{C}$, then by Theorem 2

$$
\begin{aligned}
M_{\sigma(k)} &= M_{\sigma'(0)} \\
&= \sigma' M_0 \sigma'^T \\
&= \sigma'(\sigma'^T \sigma \sigma^T \sigma') M_0 (\sigma'^T \sigma \sigma^T \sigma') \sigma'^T \\
&= \sigma(\sigma^T \circ \sigma') M_0 (\sigma^T \circ \sigma')^T \sigma^T \\
&= \sigma M_{\sigma^{-1} \circ \sigma'(0)} \sigma^T \\
&= \sigma M_{\sigma^{-1} \circ \sigma(k)} \sigma^T \\
&= \sigma M_k \sigma^T
\end{aligned}
$$

Hence L is a subgroup of H.

Conversely, if L is a subgroup of H then Theorem 2 now applies. $\qquad\square$

Conversely, we now suppose that a reduced group action $L(\Omega)$ is given. We seek to characterize those mixing schemes that commute with it. The term *mixing matrix* was previously defined with reference to a mixing scheme. It should not be confused with the following definition of *mixing matrix with respect to*:

A matrix M is a *mixing matrix with respect to* $L(\Omega)$ if the following conditions hold

1. M commutes with **Fix**(0).

2. $M_{i,j} = M_{j,i}$ for all i, j.

3. $M_{i,j} \geq 0$ for all i, j.

4. $\frac{|L|}{n} = \sum_{\sigma \in L} M_{\sigma(i), \sigma(j)}$.

If M is a mixing matrix with respect to $L(\Omega)$, then its *corresponding quadratic operator* is defined to have associated matrices $\{M_{\sigma(0)} = \sigma M \sigma^T : \sigma \in L\}$. Note that these associated matrices are well defined because

$$
\begin{aligned}
\sigma(0) = \sigma'(0) &\implies \sigma^T \sigma'(0) = 0 \\
&\implies \sigma^T \sigma' M = M \sigma^T \sigma' \\
&\implies \sigma' M \sigma'^T = \sigma M \sigma^T
\end{aligned}
$$

Note also that the mixing matrix of the quadratic operator corresponding to M is M. The technical definition of *mixing matrix* requires that the quadratic operator be a

mixing scheme; that, however, is the case, as the following theorem shows.

THEOREM 8: *If M is a mixing matrix with respect to $L(\Omega)$, then its corresponding quadratic operator is a mixing scheme that commutes with $L(\Omega)$.*

PROOF: Let $\mathcal{C}$ be the corresponding quadratic operator. The proof of Theorem 7 shows that L is a subgroup of $H(\mathcal{C})$. Since L is transitive, Theorem 7 applies, and so $\mathcal{C}$ commutes with $L(\Omega)$. It remains to show that $\mathcal{C}$ is a mixing scheme. This will be done by constructing a crossover operator $\mathcal{X}$ with mixing matrix M.

Define $\mathcal{X}$ by $\mathcal{X}(i, j) = k$ with probability $(M_k)_{i,j}$, where $\mathcal{C}$ has associated matrices M_k. It follows that $\mathcal{C}$ is the crossover scheme for $\mathcal{X}$ if $(M_k)_{i,j}$ are indeed probabilities. Given k and j, let σ' be such that $\sigma'(k) = j$. Note that

$$\sum_\sigma [\sigma(0) = k] = \sum_\sigma [\sigma'^{-1}\sigma(0) = k]$$
$$= \sum_\sigma [\sigma(0) = \sigma'(k)]$$
$$= \sum_\sigma [\sigma(0) = j]$$

Hence each element of Ω is repeated the same number of times by $\sigma(0)$ as σ ranges over L, and that number is $|L| \, / \, |\Omega|$. Therefore,

$$\sum_k (M_k)_{i,j} = \frac{n}{|L|} \sum_\sigma (M_{\sigma(0)})_{i,j}$$
$$= \frac{n}{|L|} \sum_\sigma (\sigma M \sigma^T)_{i,j}$$
$$= \frac{n}{|L|} \sum_\sigma M_{\sigma^{-1}(i), \sigma^{-1}(j)}$$
$$= \frac{n}{|L|} \sum_\sigma M_{\sigma(i), \sigma(j)}$$
$$= 1$$

Moreover, since $M_{i,j} \geq 0$ for all i, j, then these quantities are, in fact, probabilities. $\quad\square$

The next theorem shows that *every* mixing scheme that commutes with $L(\Omega)$ arises as the quadratic operator corresponding to a mixing matrix with respect to $L(\Omega)$.

THEOREM 9: *If $\mathcal{M}$ is a mixing scheme that commutes with $L(\Omega)$, then its mixing matrix is a mixing matrix with respect to $L(\Omega)$.*

PROOF: Let $\mathcal{M}$ have mixing matrix M. By Theorem 2,

$$\sigma^{-1}\sigma'(0) = 0 \implies \sigma(0) = \sigma'(0)$$
$$\implies M_{\sigma(0)} = M_{\sigma'(0)}$$
$$\implies \sigma' M \sigma'^T = \sigma M \sigma^T$$
$$\implies \sigma^{-1}\sigma' M = M \sigma^{-1}\sigma'$$

J. Rowe, M. Vose, and A. Wright

Since $\mathcal{M}$ is a mixing scheme, then M is symmetric and contains only non-negative entries. Moreover,

$$
\begin{aligned}
1 &= \sum_k (M_k)_{i,j} \\
&= \frac{n}{|L|} \sum_\sigma (M_{\sigma(0)})_{i,j} \\
&= \frac{n}{|L|} \sum_\sigma (\sigma M \sigma^T)_{i,j} \\
&= \frac{n}{|L|} \sum_\sigma M_{\sigma^{-1}(i),\sigma^{-1}(j)} \\
&= \frac{n}{|L|} \sum_\sigma M_{\sigma(i),\sigma(j)}
\end{aligned}
$$

$\square$

3.7 Example

Assigning Jobs to Processors The theorems in the previous subsection have shown that every mixing scheme that commutes with the group action arises as a quadratic operator corresponding to a mixing matrix (with respect to $L(\Omega)$). Continuing our example of four jobs on four processors, we will use this result to construct every possible crossover operator (that commutes with the group). The search space has eight elements in it, as does the group action. We therefore seek to construct an 8 by 8 matrix M that is a mixing matrix with respect to the automorphism group. First, note that in this example, **Fix**$(0) = \{0\}$, and so every matrix commutes with it. The third condition tells us that we can only use non-negative entries for our matrix. The final condition is

$$
\sum_{\sigma \in L} M_{\sigma(i),\sigma(j)} = 1
$$

since $|L| = n = 8$. This suggests that we start by picking a pair i,j and a probability distribution and then fill the entries of the matrix corresponding to each pair $(\sigma(i), \sigma(j))$ with the corresponding probability from that distribution. The correct way to do this will be explored in detail in Section 5. We then finish by symmetrizing the resulting matrix by replacing $M_{i,j}$ with $(M_{i,j} + M_{j,i})/2$.

4 Invariant Subsets and Schemata

This section deals with subsets of Ω. To simplify exposition, the notation

$$
x \subset A
$$

where $x \in \Lambda$ and $A \subset \Omega$, will be used to indicate

$$
x_i > 0 \Longrightarrow i \in A
$$

This means that we will say that a population is a subset of A if it only contains individuals that are members of A.

4.1 Pure Crossover

In Radcliffe's (1992) work on crossover, he introduces the concepts of *purity* and *respect*, which are restated to fit the framework of this paper.

DEFINITION 1: *A crossover operator $\mathcal{C}$ is said to be* pure *if for all i*

$$\mathcal{C}(e_i) = e_i$$

It is easy to see that a crossover operator is pure if and only if the result of crossing any $i \in \Omega$ with itself is always i.

4.2 Examples

Binary Strings In the case of fixed-length binary strings, a crossover operator is pure if crossing a string with itself always produces the original string. It is easy to see that the usual crossovers (one-point, two-point, uniform) are all pure.

Traveling Salesman In the traveling salesman problem, the order crossover 1 described at the end of the previous section is not pure. For example, the result of crossing $\langle \texttt{A}, \texttt{B}, \texttt{C} \rangle$ with itself using crossover section B is $\langle \texttt{C}, \texttt{B}, \texttt{A} \rangle$. A variant described in Whitley and Yoo (1995) and Vose and Whitley (1999) is pure. This variant of order crossover 1 selects the elements from parent P2 for the filler block starting at the position where the crossover section ends in parent P1, whereas the previously defined order crossover 1 started at the beginning of the string. The following example illustrates.

```
Parent P1:  A B C D E F G H I    crossover-section: _ _ C D E F _ _ _
Parent P2:  d i f g b e c h a       filler-block: h a i g b
Offspring:  g b C D E F h a i
```

If $S = \{\texttt{A}, \texttt{B}, \texttt{C}\}$, and if the permutations are listed in the same order as before, the crossover matrix for this variant of order crossover 1 is:

$$
\begin{bmatrix}
1 & 3/5 & 2/5 & 3/5 & 1 & 3/5 \\
1/5 & 0 & 0 & 1/5 & 1/5 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
1/5 & 0 & 0 & 0 & 1/5 & 1/5 \\
0 & 0 & 0 & 0 & 0 & 0 \\
1/5 & 1/5 & 0 & 0 & 1/5 & 0
\end{bmatrix}
$$

Note that the upper left entry is a 1, and all of the remaining diagonal entries are 0. This matrix can be made into a symmetric mixing matrix as was done with the previous example of a crossover matrix.

4.3 Respectful Crossover

DEFINITION 2: *A crossover operator $\mathcal{C}$ is said to* respect $A \subset \Omega$ *if*

$$x \subset A \Longrightarrow \mathcal{C}(x) \subset A$$

In other words, A is invariant under crossover.

The idea here is that respectful crossovers preserve genetic information in the population. If all members of a population have a certain property, then it is guaranteed that all the offspring will also have that property.

4.4 Examples

Binary Strings For fixed-length binary strings, the usual crossovers all respect schemata. For example, consider the two parents 00111 and 10110. They both share the schema *011*. Using any of the usual crossovers, the offspring that these parents can produce will also have this schema. This example shows that the notion of a schema is connected with that of respect. This concept will be generalized to any search space.

Traveling Salesman For the traveling salesman problem, let Ω be the group of permutations on the finite set S. For each subset b of S, let Ω_b be the subgroup

$$\Omega_b = \{k \in \Omega : k(b) = b\}$$

Let $\mathcal{N} = \{\Omega_b : b \subset S\}$. (Using notation introduced in the sequel paper, the set b would be specified by a binary mask that serves as an indicator function for b. This binary mask will also be denoted by b.) The previously described variant of order crossover 1 does not respect these subgroups. This is shown by the previous example, where the crossover illustrated does not respect the subgroup of permutations that are fixed on the set $\{A\} \subset S$. We can define yet another variant of order crossover 1 that does respect these subgroups. Under this variant, for any positions where the parents are equal, the offspring is equal to the parents. When selecting the elements from parent P2 for the filler block, these positions are skipped.

```
Parent P1:   A B C D E F G H I    crossover-section: _ _ C D E F _ _ _
Parent P2:   d i f g b e c h a          filler-block: a i g b
Offspring:   g b C D E F a h i
```

The crossover matrix for this variant of order crossover 1 is:

$$\begin{bmatrix} 1 & 4/5 & 2/5 & 4/5 & 1 & 4/5 \\ 1/5 & 0 & 0 & 1/5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1/5 & 0 & 0 & 0 & 0 & 1/5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1/5 & 1/5 & 0 & 0 & 0 & 0 \end{bmatrix}$$

4.5 Conditions for Respectfulness and Purity

We start our analysis of respect and purity by showing how properties of a population are affected by the group action. That is, if a population x has property A, then we show that the population $\sigma(x)$ has the property $\sigma(A)$.

THEOREM 10: *Let $x \in \Lambda$, $\sigma \in L$, and $A \subset \Omega$. Then*

$$x \subset A \iff \sigma(x) \subset \sigma(A)$$

PROOF: Let $\sigma = \sigma_a$. Note that

$$
\begin{aligned}
x \subset A &\iff x_i > 0 \Rightarrow i \in A \\
&\iff x_i > 0 \Rightarrow a(i) \in a(A) \\
&\iff (\sigma_a x)_{a(i)} > 0 \Rightarrow a(i) \in a(A) \\
&\iff \sigma_a x \subset a(A) \\
&\iff \sigma(x) \subset \sigma(A)
\end{aligned}
$$

$\square$

The following theorem shows that once a subset A is respected, then respect for all sets of the form $\sigma(A)$ is automatic.

THEOREM 11: *Let crossover commute with $L(\Omega)$. If A is a subset of Ω, which crossover respects, then crossover also respects $\sigma(A)$ for all $\sigma \in L$.*

PROOF: Suppose $x \subset \sigma(A)$. By Theorem 10, $\sigma^{-1}x \subset A$. Since $\mathcal{C}$ respects A and commutes with $L(\Omega)$,

$$
\sigma^{-1}x \subset A \implies \mathcal{C}(\sigma^{-1}x) \subset A \implies \sigma^{-1}\mathcal{C}(x) \subset A \implies \mathcal{C}(x) \subset \sigma(A)
$$

$\square$

The relationship between respectfulness and purity can now be given.

THEOREM 12: *A crossover operator that commutes with $L(\Omega)$ is pure if and only if it respects the set $\{0\} \subset \Omega$.*

PROOF: Assume $\mathcal{C}$ is pure. Then

$$
x \subset \{0\} \implies x = e_0 \implies \mathcal{C}(x) = e_0 \subset \{0\}
$$

Conversely, assume the set $\{0\}$ is respected, i.e., assume $\mathcal{C}(e_0) = e_0$. Let $a \in L$ be such that $a(0) = i$. Then

$$
\mathcal{C}(e_i) = \mathcal{C}(\sigma_a e_0) = \sigma_a \mathcal{C}(e_0) = \sigma_a e_0 = e_i
$$

$\square$

4.6 Examples

Binary Strings For binary strings, this result shows that crossover is pure if and only if it respects the string containing all zeros. That is, as long as we already know that crossover commutes with the group, we only need to check what happens when crossing $00\ldots0$ with itself to find out if it is pure. For example, we know that one-point crossover commutes with the group action. Now we check that crossing $00\ldots0$ with itself again produces $00\ldots0$. It clearly does, and therefore we conclude that one-point crossover is pure.

Traveling Salesman The previous two variants of order crossover 1 were pure. It can be checked that crossing the identity permutation $\langle A, B, C, D, E, F, G, H, I \rangle$ with itself again produces the identity. This is the only case we need to check to guarantee that the crossover is pure.

4.7 Generalized Schemata

We have seen in the example of binary strings that schemata are important because all the usual forms of crossover respect them. This leads us to generalize the concept to all search spaces that have a group action.

DEFINITION 3: *Let $\mathcal{N}$ be a set of subgroups of L. For each $A \in \mathcal{N}$ and each $b \in L$, the set*

$$b \circ A(0) = \{i \in \Omega \mid i = (b \circ a)(0) \text{ for some } a \in A\}$$

is called a $\mathcal{N}$-schema. When the context is not ambiguous, $\mathcal{N}$-schemata are referred to simply as schemata. *In particular, if i, j are contained in a schema, then i and j are said to* share *that schema.*

Note that if A is a normal subgroup of L, then the schemata $b \circ A(0)$ is $A \circ b(0) = A(b(0))$. Hence in that case, $\mathcal{N}$-schema are simply orbits of A.

4.8 Examples

Binary Strings An early example of $\mathcal{N}$-schemata are the M-schemata introduced by Battle and Vose (1993) for the case of length ℓ binary strings under component-wise addition modulo 2. He took $\mathcal{N}$ to be the set of subgroups generated by rows of an invertible $\ell \times \ell$ matrix M. If M is the identity matrix, "schemata" in the traditional sense result as follows.

Suppose our string length is five. Subgroups (under bitwise addition modulo 2) correspond to sets of strings that contain a zero in certain fixed positions. Thus the set

$$\{00000, 00001, 00010, 00011, 00100, 00101, 00110, 00111\}$$

is a subgroup. Because Ω is itself the group, then $A(0) = A$ for any set $A \subseteq \Omega$. If we take A to be the subgroup just described, and we add the string $b = 11001$ to each element, we get the set

$$\{11001, 11000, 11011, 11010, 11101, 11100, 11111, 11110\}$$

which is a *coset* of the subgroup. You can see that it corresponds to the schema 11*** in the traditional sense. All traditional schemata arise in this way as cosets of subgroups. The subgroup itself corresponds to 00***.

Traveling Salesman For another example of $\mathcal{N}$-schemata, let Ω be the group of permutations of S, assume $0 \in \Omega$ is the identity permutation, and let $\mathcal{N}$ be the collection $\{\Omega_b : b \subset S\}$ defined previously. Two permutations are in the same left coset of Ω_b if they agree on b, and the $\mathcal{N}$-schemata are the left cosets (of the Ω_b). So, for example, one such subgroup is the set of tours for a traveling salesman in which the fourth, fifth, and sixth cities visited are cities D, E, and F. We can compose this set with an appropriate permutation to get the set of tours in which the fourth, fifth, and sixth cities visited are A, B, and C (or indeed any other three cities). Our schemata here are thus tours in which certain cities always appear at certain points on the tour.

Assigning Jobs to Processors In the example of placing jobs on processors, one subgroup is the following set:

$$\{\langle A, B, C, D\rangle, \langle A, D, C, B\rangle\}$$

Choosing the assignment $1 \to A, 2 \to B, 3 \to C, 4 \to D$ to be the element 0, this subgroup gives us the schema

$$\{\langle 1 \to A, 2 \to B, 3 \to C, 4 \to D\rangle, \langle 1 \to A, 4 \to B, 3 \to C, 2 \to D\rangle\}$$

which is the schema containing all permissible assignments with job 1 on processor A and job 3 on processor C. If we act on this set with the group element $\langle D, A, B, C\rangle$ that rotates the jobs through $90°$ we get the schema

$$\{\langle 4 \to A, 1 \to B, 2 \to C, 3 \to D\rangle, \langle 2 \to A, 1 \to B, 4 \to C, 3 \to D\rangle\}$$

which is now the schema containing all assignments with job 1 on B and job 3 on D.

4.9   When Does Crossover Respect Schemata?

Given a collection of subgroups $\mathcal{N}$ of L, the conditions under which $\mathcal{N}$-schemata are respected are of interest. Toward that end, the following concept will be useful:

DEFINITION 4: *Given a collection of subgroups $\mathcal{N}$ of L, elements $i, j \in \Omega$ are said to be* compatible *(with respect to $\mathcal{N}$) if whenever they share a $\mathcal{N}$-schema, it is also shared by the element $0 \in \Omega$.*

In Vose (1999), the concept of a *separative* matrix is defined with respect to the space of binary strings. This concept can now be generalized:

DEFINITION 5: *Given a collection of subgroups $\mathcal{N}$ of L, a matrix A is* separative *with respect to $\mathcal{N}$ if $A_{i,j} \neq 0$ implies that i and j are compatible.*

We can now provide necessary and sufficient conditions for a crossover operative to respect all schemata.

THEOREM 13: *Suppose crossover commutes with $L(\Omega)$ and has mixing matrix M. Let $\mathcal{N}$ be a collection of subgroups of L. Then crossover respects all $\mathcal{N}$-schemata if and only if M is separative with respect to $\mathcal{N}$.*

PROOF: Assume that $\mathcal{C}$ commutes with $L(\Omega)$ and respects all $\mathcal{N}$-schemata. Assume $M_{i,j} > 0$. It needs to be shown that i and j are compatible. Suppose they share the schema $b \circ A(0)$, that is,

$$\begin{aligned} i &= b \circ a(0) \\ j &= b \circ a'(0) \end{aligned}$$

for some $a, a' \in A$. Using Theorem 5,

$$\begin{aligned} 0 &< M_{i,j} \\ &= s(b \circ a(0), b \circ a'(0), 0) \\ &= s(a(0), a'(0), b^{-1}(0)) \end{aligned}$$

Thus $b^{-1}(0) \in A(0)$ since $\mathcal{C}$ respects $A(0)$, and therefore $0 \in b \circ A(0)$, which means that i and j are compatible.

Conversely, suppose that $M_{i,j} > 0 \implies i, j$ compatible. Let $A \in \mathcal{N}$ and suppose $s(i, j, k) > 0$, where $i, j \in A(0)$. Let $b \in L$ be such that $b(0) = k$. Then

$$
\begin{aligned}
0 \quad &< \quad s(i, j, b(0)) \\
&= \quad s(b^{-1}(i), b^{-1}(j), 0) \\
&= \quad M_{b^{-1}(i), b^{-1}(j)}
\end{aligned}
$$

and therefore, $b^{-1}(i)$ and $b^{-1}(j)$ are compatible. Since they share the schemata $b^{-1}A(0)$ (because i and j share $A(0)$), it follows that $0 \in b^{-1}A(0)$, and therefore $k = b(0) \in A(0)$. What has been shown is that crossover respects $A(0)$. By Theorem 11, it respects all $\mathcal{N}$-schemata. □

4.10 Examples

Binary Strings As an example of Theorem 13, consider the group $\Omega = \{0, 1\}^\ell$ under componentwise addition modulo 2. For each subset b of string positions, let Ω_b be the subgroup of Ω consisting of strings that have zeros in the positions of b. Let $\mathcal{N} = \{\Omega_b : b$ is a set of string positions$\}$. Then, as shown in the sequel paper, any crossover based on masks will respect $\mathcal{N}$. Two strings are compatible if whenever they agree in a string position, they are both zero in that position.

Traveling Salesman For another example, let Ω the group of permutations of a set S, and let $\mathcal{N}$ be as defined in the previous permutation example. Two permutations are compatible if whenever they agree on an element of S, they are the identity on that element of S. For example, if M is the mixing matrix of the traveling salesman example in Section 4.3, then $M_{3,2}$ is forced to be zero because the corresponding permutations $\langle B, C, A \rangle$ and $\langle B, A, C \rangle$ are incompatible.

4.11 Purity and Compatibility

The following observations can be made concerning the examples given above. The previous two examples have the following in common: an element i is compatible with itself precisely when i is the identity element. The first two example crossover matrices have an analogous pattern: their i, ith entry is nonzero precisely when i is the identity element. These observations are explained by the following results.

THEOREM 14: *Suppose crossover $\mathcal{C}$ commutes with $L(\Omega)$ and has mixing matrix M. The following are equivalent,*

- $M_{0,0} = 1$

- $M_{i,i} = [i = 0]$

- *Crossover is pure*

PROOF: A simple computation from the definitions shows that

$$
\mathcal{C}(e_0)_k = M_{a^{-1}(0), a^{-1}(0)}
$$

where $a(0) = k$. Hence if $M_{i,i} = [i = 0]$, then choosing a to be the identity yields

$$
\mathcal{C}(e_0)_0 = M_{0,0} = [0 = 0] = 1
$$

Therefore,

$$
x \subset \{0\} \implies x = e_0 \implies \mathcal{C}(x) = e_0 \subset \{0\}
$$

so by Theorem 12, crossover is pure. Conversely, if crossover is pure, then

$$x = e_0 \implies x \subset \{0\} \implies \mathcal{C}(x) \subset \{0\} \implies \mathcal{C}(x) = e_0$$

Thus $M_{0,0} = \mathcal{C}(e_0)_0 = (e_0)_0 = 1$. As shown in the proof of Theorem 8, each element of Ω is repeated $|L| \, / \, |\Omega|$ times by $\sigma(0)$ as σ ranges over L. Therefore,

$$1 = \frac{|\Omega|}{|L|} \sum_{\sigma} s(0,0,\sigma(0)) = \frac{|\Omega|}{|L|} \sum_{\sigma^{-1}} s(\sigma(0),\sigma(0),0) = \sum_k M_{k,k}$$

If $M_{0,0} = 1$, the other diagonal entries must be zero. □

DEFINITION 6: *A collection of $\mathcal{N}$-schemata is called separating if and only if*

$$\bigcap_{A \in \mathcal{N}} A(0) \;=\; \{0\}$$

and for every $A \in \mathcal{N}$ and $b, b' \in L$

$$b \circ A(0) \,\cap\, b' \circ A(0) \neq \emptyset \implies b \circ A(0) \;=\; b' \circ A(0).$$

When $\mathcal{N}$ is a set of normal subgroups, the second condition is automatic, as the following argument shows. Suppose

$$z \;\in\; b \circ A(0) \,\cap\, b' \circ A(0)$$

Assuming A is normal, $b \circ A = A \circ b$, and $b' \circ A = A \circ b'$, so that

$$z = a \circ b(0) = a' \circ b'(0)$$

for some $a, a' \in A$. Thus

$$b(0) = a^{-1} \circ a' \circ b'(0)$$

and

$$b \circ A(0) = A(b(0)) = A(a^{-1} \circ a' \circ b'(0)) = A(b'(0)) = b' \circ A(0)$$

DEFINITION 7: *$\mathcal{N}$-schemata U and V are called complementary iff they can be written as*

$$\begin{aligned} U &= b \circ A(0) \\ V &= b \circ A'(0) \end{aligned}$$

where $A, A' \in \mathcal{N}$, and the $\{A, A'\}$-schemata are separating.

THEOREM 15: *If there exists a pair of complementary $\mathcal{N}$-schemata, then*

$$i \text{ is compatible with } i \iff i = 0$$

Moreover, if crossover respects a pair of complementary schemata and commutes with $L(\Omega)$, then crossover is pure.

PROOF: Let $U = b \circ A(0)$ and $V = b \circ A'(0)$ be complementary schemata, where $A, A' \in \mathcal{N}$ and the $\{A, A'\}$-schemata are separating. Given $i \neq 0$, therefore, $i \notin A(0) \cap A'(0)$.

Without loss of generality, suppose $i \notin A(0)$. Since $L(\Omega)$ is transitive, choose b such that $i \in b \circ A(0)$. If i were compatible with i, then $0 \in b \circ A(0)$. But since the $\{A, A'\}$-schemata are separating and $0 \in A(0)$, it would follow that $i \in b \circ A(0) = A(0)$, which is a contradiction.

Note that if crossover $\mathcal{C}$ respects U, V and commutes with $L(\Omega)$, then crossover respects all $\{A, A'\}$-schemata: Given $x, c \in \Omega$, choose σ such that $\sigma(c) = b$. For any $C \in \{A, A'\}$,

$$
\begin{aligned}
x \subset c \circ C(0) &\implies \sigma(x) \subset \sigma(c) \circ C(0) \\
&\implies \sigma(x) \subset b \circ C(0) \\
&\implies \mathcal{C}(\sigma(x)) \subset b \circ C(0) \\
&\implies \sigma \circ \mathcal{C}(x) \subset \sigma(c) \circ C(0) \\
&\implies \mathcal{C}(x) \subset c \circ C(0)
\end{aligned}
$$

In particular, crossover respects $A(0) \cap A'(0) = \{0\}$. □

5 Inducing a Group Structure on the Search Space

5.1 The Search Space as a Group

A special case of a group action $L(\Omega)$ is when $\Omega = L$ and the action is given by $k(w) = k \circ w$. Given a group L acting on Ω, one might wonder whether this special case could be recovered. Perhaps a group operation could be induced on Ω such that $L(\Omega)$ and $\Omega(\Omega)$ were equivalent.

If L is a group of permutations acting on Ω, then Ω is said to have a *group structure compatible with* $L(\Omega)$ if there exists a binary operation $\circ$ with identity element defined on Ω, and a function $\phi : L \longrightarrow \Omega$ such that for all $k \in L, w \in \Omega$,

$$k(w) = \phi(k) \circ w$$

In that case, for notational convenience, denote the identity elements of L and Ω by 0.

THEOREM 16: *If $L(\Omega)$ is reduced and Ω has a group structure compatible with it, then $(\Omega, \circ)$ is a group isomorphic to L by ϕ.*

PROOF: Suppose ϕ were not surjective. Then there exists $w \in \Omega$ such that for all $k \in L$,

$$w \neq \phi(k) = \phi(k) \circ 0 = k(0)$$

contradicting that L is transitive. Note that, for all $w \in \Omega$,

$$\phi(k) = \phi(k') \implies k(w) = \phi(k) \circ w = \phi(k') \circ w = k'(w)$$

Thus $k = k'$ so ϕ is also injective. Moreover,

$$
\begin{aligned}
\phi(k \circ k') &= \phi(k \circ k') \circ 0 \\
&= (k \circ k')(0) \\
&= k(k'(0)) \\
&= \phi(k) \circ k'(0) \\
&= \phi(k) \circ (\phi(k') \circ 0) \\
&= \phi(k) \circ \phi(k')
\end{aligned}
$$

Therefore ϕ is a group isomorphism. □

Note that the only possibility for ϕ is $\phi(k) = \phi(k) \circ 0 = k(0)$. Moreover, when $L(\Omega)$ is reduced, the group operation on Ω is completely determined by $k(w) = \phi(k) \circ w$. Therefore, Theorem 16 speaks to a situation wherein both the isomorphism ϕ and the group operation on Ω are completely determined by what $L(\Omega)$ and the identity element of Ω are.

5.2   Example

Assigning Jobs to Processors Consider again the example of placing jobs on processors. We write out the complete search space of valid assignments:

A	B	C	D
1	2	3	4
4	1	2	3
3	4	1	2
2	3	4	1
3	2	1	4
1	4	3	2
2	1	4	3
4	3	2	1

Now we write out the group that acts on this set:

A	B	C	D
A	B	C	D
D	A	B	C
C	D	A	B
B	C	D	A
C	B	A	D
A	D	C	B
B	A	D	C
D	C	B	A

By examining the two tables it is clear that we can define a bijection $\phi : L \to \Omega$ by

$$A \mapsto 1, B \mapsto 2, C \mapsto 3, D \mapsto 4$$

Notice that this gives us $\phi(k) = k(0)$, where we choose 0 to be the assignment $1 \to A, 2 \to B, 3 \to C, 4 \to D$. We can then define a group operator on Ω by setting $i \circ j = \phi^{-1}(i)(j)$. We can easily see that, in accordance with Theorem 16, this group is isomorphic to L.

It would be nice if we could have done this without having to draw up complete tables for the two sets. Theorem 17, which follows, tells us exactly when this can be done.

5.3   Left and Right Group Actions

Define the set R of permutations of Ω—where a permutation of Ω need not be an element of L—by the condition that, for all $k \in L$,

$$g \in R \Longleftrightarrow g \circ k = k \circ g$$

It follows that R is a group acting on Ω by function application, and $\Omega_R = \{0\}$.

5.4 Examples

Binary Strings The group of fixed-length binary strings under addition is Abelian. Clearly, in this case, $L \subseteq R$. In fact, the following theorem will show that $L = R$.

Traveling Salesman Consider the traveling salesman problem on three cities, A, B and C. There are six possible routes that we can label 0 to 5 as follows:

$$
\begin{array}{c|ccc}
0 & A & B & C \\
1 & B & A & C \\
2 & B & C & A \\
3 & A & C & B \\
4 & C & A & B \\
5 & C & B & A
\end{array}
$$

We now consider the group action to simply specify functions from the set $\{0, 1, 2, 3, 4, 5\}$ to itself. For example,

$$\langle C, B, A \rangle (1) = \langle C, B, A \rangle \circ \langle B, A, C \rangle = \langle B, C, A \rangle = 2$$

The definition of $R(\Omega)$ invites us to consider the set of all permutations of the set $\{0, 1, 2, 3, 4, 5\}$. There are 6! of these. The ones that are in R are those that commute with L. That is, $g \in R$ if, for all $k \in L$,

$$g \circ k(x) = k \circ g(x)$$

The following theorem will show that these permutations can be found as follows: for each $i = 0, 1, \ldots 5$, define

$$g_i(x) = x \circ i$$

in which on the right hand side we interpret x and i as again being permutations.

Assigning Jobs to Processors The jobs on processors problem will illustrate the full power of the next theorem, and so it will be treated below.

5.5 Conditions for Group Compatibility

The next theorem tells us exactly when the search space Ω can be given a group structure compatible to L. There are three equivalent conditions. One of these relates to the group action $R(\Omega)$ just defined and asks that R be reduced. In fact, R just has to be transitive, since $\Omega_R = \{0\}$ by construction. That R might not be transitive is shown in the following example. Let $\Omega = \{0, 1, 2\}$ and suppose our group action $L(\Omega)$ is the set of all permutations of Ω. Then it can be checked that R contains just the identity and so is not transitive. The other conditions relate directly to L. One states that **Fix**$(x) = \{0\}$ for all $x \in \Omega$. That is, for any x, the only element that has $k(x) = x$ is the case when k is the identity. The last condition states that there is at least one element of Ω on which different elements of L produce different results. These conditions are easy to check for a particular group. The corollary to the theorem states that these conditions are met if the group L is Abelian.

THEOREM 17: *If $L(\Omega)$ is reduced, then the following are equivalent:*

1. *Ω has a group structure compatible with $L(\Omega)$.*

2. *There exists an element $0 \in \Omega$ such that, for all $k, k' \in L$,*

$$k(0) = k'(0) \implies k = k'$$

3. **Fix**$(w) = \{0\}$, *for all $w \in \Omega$*

4. *$R(\Omega)$ is reduced.*

Moreover, in that case (i.e., everything above holds), then L is anti-isomorphic to R, and the group operation induced by R (on Ω) may be chosen to be the reverse of the group operation induced by L.

PROOF: $(1 \Leftrightarrow 2)$: If Ω has a group structure compatible with $L(\Omega)$, then ϕ is injective. Hence,

$$k(0) = k'(0) \implies \phi(k) = \phi(k') \implies k = k'$$

Conversely, let $0 \in \Omega$ be such that for all $k, k' \in L$,

$$k(0) = k'(0) \implies k = k'$$

Define $\circ$ on Ω so that 0 is an identity element, and define $\phi : L \to \Omega$ by

$$\phi(k) = k(0)$$

Since L acts transitively, ϕ is surjective. Complete the definition of $\circ$ on Ω by

$$\phi(k) \circ w = k(w)$$

The operation $\circ$ is well-defined since

$$\phi(k) = \phi(k') \implies k(0) = k'(0) \implies k = k'$$

$(2 \Leftrightarrow 3)$: By hypothesis,

$$k \circ h(0) = k' \circ h(0) \implies k \circ h = k' \circ h \implies k = k'$$

By transitivity, $h(0)$ is an arbitrary element of Ω. Thus, for all w,

$$k(w) = k'(w) \implies k = k'$$

which is equivalent to

$$k^{-1}k'(w) = w \implies k^{-1}k' = 0$$

which is equivalent to

$$k^{-1}k' \neq 0 \implies k^{-1}k'(w) \neq w$$

$(4 \Rightarrow 3)$: Since R is a permutation group on Ω, it is reduced if and only if it is transitive. Suppose R is transitive, and that $k(w) = w$. Then for all $g \in R$,

$$k(g(w)) = g(k(w)) = g(w)$$

Since R is transitive, $g(w)$ is an arbitrary element of Ω. Hence $k = 0$.

$(1 \Rightarrow 4)$: By hypotheses and Theorem 16, $(\Omega, \circ)$ is a group. Given $x \in \Omega$, define the permutation g_x by

$$g_x(z) = z \circ x$$

Now $g_x \in R$ since for all $k \in L$,

$$
\begin{aligned}
g_x(k(w)) &= g_x(\phi(k) \circ w) \\
&= \phi(k) \circ w \circ x \\
&= k(w \circ x) \\
&= k(g_x(w))
\end{aligned}
$$

Given any $u, v \in \Omega$, let $x = u^{-1} \circ v$ and then $g_x(u) = u \circ u^{-1} \circ v = v$. Hence $R(\Omega)$ is transitive and therefore reduced.

If $L(\Omega)$ is reduced and conditions 1 through 4 hold, then condition 2 holds with R in place of L since, for all $k \in R$,

$$
g(0) = g'(0) \implies k \circ g(0) = k \circ g'(0) \implies g(k(0)) = g'(k(0)) \implies g = g'
$$

Hence there is a group operation $\circ'$ induced on Ω and an isomorphism $\phi' : R \longrightarrow \Omega$ such that, for all $g \in R$, $w \in \Omega$,

$$
g(w) = \phi'(g) \circ' w
$$

The argument of $(1 \Rightarrow 4)$ shows that for every $w \in \Omega$, the permutation g_w defined by

$$
g_w(z) = z \circ w
$$

is an element of R. Thus $\phi'(g_w) = g_w(0) = 0 \circ w = w$, provided that the identity elements of $(\Omega, \circ)$ and $(\Omega, \circ')$ are the same. That can be arranged, however, since the arguments of $(1 \Leftrightarrow 2)$ and $(2 \Leftrightarrow 3)$ show that the choice of identity is arbitrary. It follows that

$$
\begin{aligned}
w \circ' z &= \phi'(\phi'^{-1}(w)) \circ' z \\
&= (\phi'^{-1}(w))(z) \\
&= (\phi'^{-1}(\phi'(g_w)))(z) \\
&= (g_w)(z) \\
&= z \circ w
\end{aligned}
$$

Finally, the map $k \mapsto g_k$ is an anti-isomorphism,

$$
g_{k \circ k'}(z) = z \circ k \circ k' = (g_k(z)) \circ k' = g_{k'}(g_k(z)) = g_{k'} \circ g_k(z)
$$

□

Notice that under the conditions of Theorem 17, we can explicitly write down the group $R(\Omega)$. Each element of the group corresponds to an element $x \in \Omega$ with group action $z \mapsto z \circ x$. The action of R on Ω therefore corresponds to a permutation acting from the right (R), whereas the original group acts from the left (L).

Note also that $L = R$ whenever L is Abelian, because in that case multiplication on the left (L) and multiplication on the right (R) coincide. Moreover, we have the following:

COROLLARY 18: *If $L(\Omega)$ is a reduced Abelian group action on Ω, then Ω has a group structure compatible with L.*

PROOF: Since $L = R$ in this case, R must be reduced and the theorem applies. □

5.6 Example

Assigning Jobs to Processors We can now apply our theorem to the jobs on processors example. The simplest condition to check is number 2. If we look at the arrangement of jobs $\langle 1 \rightarrow A, 2 \rightarrow B, 3 \rightarrow C, 4 \rightarrow D \rangle$ and consider the effect of each group member on it, it is clear that each automorphism maps it to a different assignment. Condition 2 is satisfied, and we therefore know that we can give Ω a group structure compatible with L. This group is given by the bijection $\phi : L \rightarrow \Omega$ defined by $\phi(k) = k(0)$ and then carrying over group operator in a natural way. Moreover, the group R is anti-isomorphic to L and can be used as an alternative (but equivalent) group action.

5.7 Characterizing Crossover and Mutation

For the remainder of this section, we assume that the conditions in Theorem 17 hold. In this case, consider the action of L on the set of pairs of elements Ω^2, defined in the natural way. Define an equivalence relation between pairs as follows:

$$(i, j) \sim (x, y) \iff x = a(i) \text{ and } y = a(j) \text{ for some } a \in L(\Omega)$$

From Theorem 17 it follows that there are n equivalence classes, each containing n pairs. Theorems 8 and 9 show that all mixing matrices have the property

$$\sum_{a \in L} M_{a(i), a(j)} = 1$$

for each pair $(i, j) \in \Omega^2$. We can therefore determine all possible mixing matrices with respect to L.

THEOREM 19: *Assuming the conditions of Theorem 17 hold, we can determine all mixing matrices with respect to L by choosing n probability distributions $w_k \in \Lambda$ (where $k \in \Omega$) and assigning*

$$M_{i,j} = (w_{i^{-1} \circ j})_i$$

and then symmetrizing (replace $M_{i,j}$ with $(M_{i,j} + M_{j,i})/2$).

PROOF: The preceding discussion shows that the values $M_{a(i),a(j)}$ must form a probability distribution as a ranges over L, for fixed i, j within each equivalence class. To construct a mixing matrix, we can therefore take any n probability distributions $w_k \in \Lambda$ (where $k \in \Omega$) and set

$$M_{i,j} = (w_k)_i$$

where $(i, j) \sim (0, k)$. That is, $k = a(j)$, where $a(i) = 0$. Because of the isomorphism between the group action and $(\Omega, \circ)$, we have

$$k = a(j) = a(0) \circ j$$

and

$$0 = a(i) = a(0) \circ i$$

which means that $k = i^{-1} \circ j$. $\qquad\qquad\square$

We can also determine all the possible mutation operators (that commute with L) as follows:

THEOREM 20: *Assuming the conditions of Theorem 17 hold, a mutation matrix U commutes with $L(\Omega)$ if and only if there exists $\alpha \in \Lambda$ such that*

$$U = \sum_{k \in \Omega} \alpha_k \pi_k$$

where the π_k are permutation matrices representing group R.

PROOF: That the condition is sufficient is clear from the definition of $R(\Omega)$. To show that it is necessary, assume that U commutes with $L(\Omega)$. That is, for all $a \in L$,

$$U_{i,j} = U_{a(i),a(j)}$$

Define the matrix π_k to have i, jth component

$$[(i,j) \sim (0,k)]$$

for each $k \in \Omega$, and let

$$\alpha_k = U_{0,k}$$

Then

$$U = \sum_{k \in \Omega} \alpha_k \pi_k$$

We now need to show that the set $\{\pi_k : k \in \Omega\}$ is a representation of R. Let $a \in L$, and let σ_a be the permutation matrix representing a. Then for any $k \in \Omega$,

$$
\begin{aligned}
(\sigma_a \pi_k)_{i,j} &= (\pi_k)_{a^{-1}(i),j} \\
&= [(a^{-1}(i), j) \sim (0, k)] \\
&= [(i, a(j)) \sim (0, k)] \\
&= (\pi_k)_{i,a(j)} \\
&= (\pi_k \sigma_a)_{i,j}
\end{aligned}
$$

So each π_k represents an element of $R(\Omega)$, and since there are n of each, then the set $\{\pi_k : k \in \Omega\}$ represents the group $R(\Omega)$ as required. □

Since in this case R is anti-isomorphic to L and therefore to the group structure on Ω, then this is equivalent to specifying a probability distribution over the elements of Ω, choosing an element using this distribution and applying it to the individual to be mutated *from the right*. We therefore have this corollary:

COROLLARY 21: *Assuming the conditions of Theorem 17 hold, a mutation matrix U commutes with $L(\Omega)$ if and only if there exists $\alpha \in \Lambda$ such that*

$$U_{i,j} = \sum_{k \in \Omega} \alpha_k [j \circ k = i] = \alpha_{j^{-1} \circ i}$$

5.8 Example

Assigning Jobs to Processors We are now in a position to determine all possible crossover and mutation operators that commute with the group action for the jobs and processors example. Let $\mathcal{X}_0, \ldots \mathcal{X}_7$ be probability distributions over Ω, and let

X be the matrix containing these distributions as rows. Then we can construct a mixing matrix by applying Theorem 19 to X. To proceed we need to index the group members (and equivalently the search space).

	A	B	C	D
0	A	B	C	D
1	D	A	B	C
2	C	D	A	B
3	B	C	D	A
4	C	B	A	D
5	A	D	C	B
6	B	A	D	C
7	D	C	B	A

We then set $M_{i,j} = X_{i^{-1} \circ j, i} = \mathcal{X}_{i^{-1} \circ j}(i)$. Following this rule, we get a matrix of the form

$$
\begin{bmatrix}
\mathcal{X}_0(0) & \mathcal{X}_1(0) & \mathcal{X}_2(0) & \mathcal{X}_3(0) & \mathcal{X}_4(0) & \mathcal{X}_5(0) & \mathcal{X}_6(0) & \mathcal{X}_7(0) \\
\mathcal{X}_3(1) & \mathcal{X}_0(1) & \mathcal{X}_1(1) & \mathcal{X}_2(1) & \mathcal{X}_7(1) & \mathcal{X}_6(1) & \mathcal{X}_4(1) & \mathcal{X}_5(1) \\
\mathcal{X}_2(2) & \mathcal{X}_3(2) & \mathcal{X}_0(2) & \mathcal{X}_1(2) & \mathcal{X}_5(2) & \mathcal{X}_4(2) & \mathcal{X}_7(2) & \mathcal{X}_6(2) \\
\mathcal{X}_1(3) & \mathcal{X}_2(3) & \mathcal{X}_3(3) & \mathcal{X}_0(3) & \mathcal{X}_6(3) & \mathcal{X}_7(3) & \mathcal{X}_5(3) & \mathcal{X}_4(3) \\
\mathcal{X}_4(4) & \mathcal{X}_7(4) & \mathcal{X}_5(4) & \mathcal{X}_6(4) & \mathcal{X}_0(4) & \mathcal{X}_2(4) & \mathcal{X}_3(4) & \mathcal{X}_1(4) \\
\mathcal{X}_5(5) & \mathcal{X}_6(5) & \mathcal{X}_4(5) & \mathcal{X}_7(5) & \mathcal{X}_2(5) & \mathcal{X}_0(5) & \mathcal{X}_1(5) & \mathcal{X}_3(5) \\
\mathcal{X}_6(6) & \mathcal{X}_4(6) & \mathcal{X}_7(6) & \mathcal{X}_5(6) & \mathcal{X}_1(6) & \mathcal{X}_3(6) & \mathcal{X}_0(6) & \mathcal{X}_2(6) \\
\mathcal{X}_7(7) & \mathcal{X}_5(7) & \mathcal{X}_6(7) & \mathcal{X}_4(7) & \mathcal{X}_3(7) & \mathcal{X}_1(7) & \mathcal{X}_2(7) & \mathcal{X}_0(7)
\end{bmatrix}
$$

One should then symmetrize this matrix to get the unique mixing matrix. Of course, while this tells us exactly what our mixing matrix must look like, it doesn't tell us how to implement a corresponding crossover operator in a natural way. This will be the subject of the sequel paper.

If one wanted a crossover operator that respects a family of schemata, then the those entries of the mixing matrix that are incompatible relative to this family of schemata should be set to zero.

The search space group corresponds to the symmetries of the square (the dihedral group D_4). This group has five order 2 subgroups: the groups generated by the horizontal reflection (of the square), the vertical reflection, each of the two diagonal reflections, and the $180°$ rotation. This collection of subgroups defines a family of schemata. If one sets the incompatible entries of the above mixing matrix to zero, then the crossover matrix becomes:

$$
\begin{bmatrix}
\mathcal{X}_0(0) & \mathcal{X}_1(0) & \mathcal{X}_2(0) & \mathcal{X}_3(0) & \mathcal{X}_4(0) & \mathcal{X}_5(0) & \mathcal{X}_6(0) & \mathcal{X}_7(0) \\
\mathcal{X}_3(1) & 0 & \mathcal{X}_1(1) & 0 & 0 & 0 & 0 & 0 \\
\mathcal{X}_2(2) & \mathcal{X}_3(2) & 0 & \mathcal{X}_1(2) & 0 & 0 & 0 & 0 \\
\mathcal{X}_1(3) & 0 & \mathcal{X}_3(3) & 0 & 0 & 0 & 0 & 0 \\
\mathcal{X}_4(4) & 0 & 0 & 0 & 0 & 0 & \mathcal{X}_3(4) & \mathcal{X}_1(4) \\
\mathcal{X}_5(5) & 0 & 0 & 0 & 0 & 0 & \mathcal{X}_1(5) & \mathcal{X}_3(5) \\
\mathcal{X}_6(6) & 0 & 0 & 0 & \mathcal{X}_1(6) & \mathcal{X}_3(6) & 0 & 0 \\
\mathcal{X}_7(7) & 0 & 0 & 0 & \mathcal{X}_3(7) & \mathcal{X}_1(7) & 0 & 0
\end{bmatrix}
$$

The corresponding row stochastic matrix X is:

$$\begin{bmatrix} X_0(0) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X_1(0) & X_1(1) & X_1(2) & X_1(3) & X_1(4) & X_1(5) & X_1(6) & X_1(7) \\ X_2(0) & 0 & X_2(2) & 0 & 0 & 0 & 0 & 0 \\ X_3(0) & X_3(1) & X_3(2) & X_3(3) & X_3(4) & X_3(5) & X_3(6) & X_3(7) \\ X_4(0) & 0 & 0 & 0 & X_4(4) & 0 & 0 & 0 \\ X_5(0) & 0 & 0 & 0 & 0 & X_5(5) & 0 & 0 \\ X_6(0) & 0 & 0 & 0 & 0 & 0 & X_6(6) & 0 \\ X_7(0) & 0 & 0 & 0 & 0 & 0 & 0 & X_7(7) \end{bmatrix}$$

When there are only two non-zero entries in a row of the X matrix, these correspond to symmetric entries in the mixing matrix. They must therefore have equal values by symmetry. Because they also form a probability distribution, they must in fact take the value $1/2$. The original choice of 8 probability distributions has therefore been reduced to 2, namely X_1 and X_3.

If one wanted a mutation operator, then we have already seen that we can use a probability distribution over the center of the group $Z(L)$. More generally, we now see that we can use any probability distribution μ over Ω (which is now isomorphic to L) and apply the chosen element on the right. The mutation matrix constructed by the formula of Corollary 21 can be combined with a crossover matrix to give a single mixing matrix by using Theorem 1.

6 Conclusion

This paper investigates the situation where there is a group of symmetries, or permutations, acting on the search space Ω of a genetic algorithm. If crossover and mutation commute with these symmetries, then these operators can be described by a mixing matrix and the group of permutation matrices that correspond to the group of symmetries. These considerations lead to a natural generalization of the concept of *schemata* and to conditions under which crossover and mutation respect schemata.

We have considered the possibility that the group action might induce a group structure on the search space itself. Necessary and sufficient conditions for this to happen have been given. Some of the implications of having a group structure on Ω are worked out in the sequel paper. This will lead to a general method for designing crossover and mutation operators in this case.

To summarize the key points of this paper:

1. Symmetries in a search space can be captured by defining a group action.

2. If crossover and mutation commute with the group action, then their effect on the population can be described by a single mixing matrix.

3. Necessary and sufficient conditions for this to happen have been given.

4. The structure of all possible mixing matrices has been characterized.

5. Crossover is *pure* if it fixes uniform populations.

6. Crossover is *respectful* if it is invariant on a given collection of subsets.

7. Crossover is pure if and only if the set $\{0\}$ is respected.

8. If crossover respects a subset A, then it also respects $\sigma(A)$, where σ is a group element.

9. The concept of *schemata* is generalized to capture the subgroup structure of the group action.

10. Necessary and sufficient conditions are given for crossover to respect all schemata. This is related back to the notion of purity.

11. The question of when a group action on the search space can induce a group structure on the search space was addressed. Necessary and sufficient conditions are found.

12. These conditions are always met when the group action is commutative.

13. When these conditions hold, it is possible to characterize all possible crossover and mutation operators.

The sequel paper will address the following issues:

1. Implications of the search space itself having a group structure.

2. The concept of a *structural* search space.

3. The definition of natural crossover and mutation operators for structural search spaces.

4. Proof that such genetic operators respect all schemata.

5. The existence of a Walsh Transform for Abelian groups (and only for Abelian groups).

6. Crossover, mutation, and proportional selection are analyzed with respect to the Walsh Transform in the Abelian case.

A third paper will discuss the relationship between *landscape graphs* and search space symmetries by studying groups of automorphisms of such graphs. This will lead to a generic method for designing crossover and mutation operators in relation to the neighborhood structure embodied in the landscape graph. A fourth paper will relate the construction of crossovers to the *twist* operator on matrices.

Acknowledgments

References

Battle, D. L. and Vose, M. D. (1993). Isomorphisms of genetic algorithms. *Artificial Intelligence*, 60:155–165.

Davis, L. (1985). Applying adaptive algorithms to epistatic domains. In Joshi, A., editor, *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 162–164, Morgan Kaufmann, San Francisco, California.

Radcliffe, N. J. (1992). The algebra of genetic algorithms. *Anals of Mathematics and Artificial Intelligence*, 10:339–384.

Rowe, J. E. (2001). A normed space of genetic operators with applications to scalability issues. *Evolutionary Computation*, 9(1):25–42.

Rowe, J. E., Vose, M. D., and Wright, A. H. (2001). Structural genetic operators. In preparation.

Vose, M. D. (1990). Formalizing genetic algorithms. In *Proceedings of Genetic Algorithms, Neural Nets, and Simulated Annealing Applied to Problems in Signal and Image Processing*, IEEE, UK and Ireland.

Vose, M. D. (1999). *The Simple Genetic Algorithm: Foundations and Theory*. MIT Press, Cambridge, Massachusetts.

Vose, M. and Whitley, D. (1999). A formal language for permutation recombination operators. In Banzhaf, W. and Reeves, C., editors, *Foundations of Genetic Algorithms (FOGA-5)*, Morgan Kaufmann, San Mateo, California.

Whitley, D. and Yoo, N.-W. (1995). Modeling permutation encodings in simple genetic algorithm. In Whitley, L. D. and Vose, M. D., editors, *Foundations of Genetic Algorithms 3*, pages 163–184, Morgan Kaufmann, San Mateo, California.