Neighborhood Graphs and Symmetric Genetic Operators

Jonathan E. Rowe[1], Michael D. Vose[2], and Alden H. Wright[3]

[1] School of Computer Science
University of Birmingham
Birmingham B15 2TT
Great Britain
`J.E.Rowe@cs.bham.ac.uk`
[2] Computer Science Department
University of Tennessee
Knoxville, TN 37996
USA
`vose@cs.utk.edu`
[3] Dept. of Computer Science
University of Montana
Missoula, Montana 59812
USA
`wright@cs.umt.edu`

Abstract. In the case where the search space has a group structure, classical genetic operators (mutation and two-parent crossover) which respect the group action are completely characterized by formulas defining them in terms of the search space and its group operation. This provides a representation-free implementation for those operators, in the sense that the genotypic encoding of search space elements is irrelevant. The implementations are parameterized by distributions which may be chosen arbitrarily, and which are analogous to specifying distributions for mutation and crossover masks.

1 Introduction

This paper extends the theory developed in [RVW02, RVW04] concerning groups that act transitively on a search space. The special case where the search space itself has a group structure (so that it acts transitively on itself) is the primary focus.

One might legitimately wonder what mixing operators for classical Genetic Algorithms are possible (given reasonable restrictions).[1] We have answered such a question; the main results completely characterize classical genetic operators – mutation and two-parent crossover – which respect search space symmetries.

---

[1] Admittedly, such questions concerning *genetic algorithms* are *foundational* rather than applied, but this paper was intended for *"Foundations Of Genetic Algorithms"*...

C.R. Stephens et al. (Eds.): FOGA 2007, LNCS 4436, pp. 110–122, 2007.

The paper begins with motivation, describing how a search space might come to have a group acting upon it. This is done by way of first discussing neighborhood structures and their symmetries, next considering how neighborhood structures naturally arise as a consequence of neighborhood operators, and then describing conditions under which neighborhood operators have symmetries which may be ascribed to the search space itself. Consequently, the particular group ascribed to the search space can vary with the chosen neighborhood structure; more generally, it may be chosen arbitrarily. In any case, a transitive group action eliminates bias in the sense that the search space is made to look the same from every point. Results proved in sections 2 and 3 are probably not new; they are included as part of the motivation leading up to section 4. Additional motivation is provided by the fact that other authors have also considered group structures on the search space [Wei91, Sta96, RHKS02].

In the case where the search space has a group structure, the classical genetic operators (mutation and two-parent crossover) can be completely characterized when they commute with the group action. Moreover, they have representation-free implementations; the genotypic encoding of search space elements is irrelevant. Their implementations are parameterized by distributions which may be chosen arbitrarily, and which are analogous to specifying distributions for mutation and crossover masks.

1.1 Notation

Suppose the finite search space Ω is enumerated as $\{\omega_0, \ldots, \omega_{n-1}\}$. Without loss of generality Ω may be regarded as $\{0, \ldots, n-1\}$ through the association $i \mapsto \omega_i$.

The notation [*expression*] denotes 1 if *expression* is true, and 0 otherwise.

To maintain continuity with the thread of most relevant results [Vos99b], [RVW02], [RVW04], the "twist" of the matrix A (see section 4.1) is denoted by A^*.[2] A similar comment is apropos to how group operations are denoted. In the special case where a group can be expressed as a nontrivial direct sum of normal subgroups, the best choice would indubitably be $\oplus$ so as to be consistent with [Vos99b, RVW04] which are most relevant. This paper, however, concerns the general case; it uses $\circ$ to denote the group operation, which is consistent with [RVW02].

2 Neighborhood Structures

Assume the finite search space Ω has a *neighborhood structure*: every $x \in \Omega$ has a set $N(x) \subset \Omega$ of neighbors. A neighborhood structure N is equivalent to a *neighborhood graph* which has Ω as vertex set and which contains directed edge $x \to y$ iff $y \in N(x)$.

---

[2] We apologize for using superscript asterisk to denote something other than Kleene Closure...

Definition 1. *A neighborhood structure on a finite search space Ω is a function $N : \Omega \to 2^{\Omega}$ which to each $x \in \Omega$ assigns a set $N(x) \subset \Omega$ of neighbors.*

Definition 2. *The neighborhood graph corresponding to a neighborhood structure N has the domain Ω of N as vertex set and contains directed edge $x \to y$ iff $y \in N(x)$.*

In practice, a neighborhood structure often has symmetries. An *automorphism* of N is a bijection $\pi : \Omega \to \Omega$ such that if y is a neighbor of x, then $\pi(y)$ is a neighbor of $\pi(x)$. Equivalently, it is an invertible map (permutation) on the vertices of the neighborhood graph which preserves edges.

Definition 3. *An automorphism of a neighborhood structure N is a bijection $\pi : \Omega \to \Omega$ such that if y is a neighbor of x, then $\pi(y)$ is a neighbor of $\pi(x)$.*

The set $\mathcal{A}_N$ of all such automorphisms is the *symmetry group* of N (it is a group under function composition since Ω is finite); elements of $\mathcal{A}_N$ are called *symmetries* (of N).

Definition 4. *The symmetry group of a neighborhood structure N is the set $\mathcal{A}_N$ of all automorphisms of N (it is a group under function composition); elements of $\mathcal{A}_N$ are called symmetries of N.*

As a consequence of preserving edges of the neighborhood graph, the symmetry group of N *commutes* with the neighborhood structure; for all $x \in \Omega$ and all $\pi \in \mathcal{A}_N$

$$\begin{aligned} \pi \circ N(x) &= \{\pi(y) \,:\, y \in N(x)\} \\ &= \{z \,:\, z \in N(\pi(x))\} \\ &= N(\pi(x)) \end{aligned}$$

The theory developed in [RVW02, RVW04] concerns groups that act *transitively* on Ω : for every $x, y \in \Omega$ there exists a group element g such that $g(x) = y$. A direct consequence of $\mathcal{A}_N$ commuting with N is that a necessary (but not sufficient) condition for transitivity is that the neighborhood graph be regular (all vertices have the same degree).

Definition 5. ([Big71]) *A group $(G, \circ)$ is said to act on Ω if its elements act as permutations (of Ω) such that for all $g, g' \in G$, and all $x \in \Omega$,*

$$\begin{gathered} (g \circ g')(x) = g(g'(x)) \\ e(x) = x \end{gathered}$$

where $e \in G$ denotes the identity element. Moreover, G acts transitively (on Ω) if for every $x, y \in \Omega$ there exists a group element g such that $g(x) = y$.

2.1 Neighborhood Operators

A neighborhood structure might arise from a collection of *neighborhood operators* on Ω which could be used, for example, by a search algorithm; operator ν assigns to each x some particular neighbor $\nu(x)$. A collection $\mathcal{O}$ of neighborhood operators generates a neighborhood structure,

$$N(x) \,=\, \{\nu(x) \,:\, \nu \in \mathcal{O}\}$$

Definition 6. *A neighborhood operator is a function $\nu : \Omega \to \Omega$ which to each $x \in \Omega$ assigns a neighbor $\nu(x)$. The neighborhood structure N generated by a collection $\mathcal{O}$ of neighborhood operators maps $x \in \Omega$ to the set $N(x) = \{\nu(x) : \nu \in \mathcal{O}\}$.*

> **Example 1.** Let $\beta(n)$ denote the ℓ-bit binary expansion of the integer n, and let $\oplus$ denote bitwise exclusive-or. The Hamming neighborhood structure on the set $\mathcal{S}$ of binary strings of length ℓ is generated by
>
> $$\mathcal{O} \,=\, \{\nu_k \,:\, \forall x \in \mathcal{S} \,.\, \nu_k(x) = \beta(2^k) \oplus x \,,\; 0 \le k < \ell\}$$
>
> Here the operators of $\mathcal{O}$ are self-inverse, thus $x \in \mathcal{N}(y) \iff y \in \mathcal{N}(x)$. The corresponding neighborhood graph is the Hamming cube.

If, as in the example 1, every element of $\mathcal{O}$ is invertible, then $\mathcal{O}$ is said to be invertible.

Definition 7. *A collection $\mathcal{O}$ of neighborhood operators is said to be invertible if every element of $\mathcal{O}$ is invertible.*

The neighborhood structure generated by a collection of neighborhood operators can be *connected*—meaning that the neighborhood graph is connected (there is a directed path from x to y for all $x, y \in \Omega$)—yet its symmetry group can fail to be transitive (see example 2 below).

Definition 8. *A neighborhood structure N is said to be connected if there is a directed path from x to y for all $x, y \in \Omega$.*

> **Example 2.** Let $\mathcal{O}$ be the set of permutations $\{(120)(453), (210)(543), (012345), (543210)\}$ (in cycle notation) of $\Omega = \{0, 1, 2, 3, 4, 5\}$. The collection $\mathcal{O}$ of neighborhood operators is invertible (not because $\mathcal{O}$ is closed under inverse; what matters is that its elements are invertible). The neighborhood structure N generated by $\mathcal{O}$ is connected. The symmetry group of N is not transitive because the neighborhood graph is not regular (bi-directional edges are shown without arrow heads below)

Theorem 1. *Let N be the neighborhood structure generated by a collection $\mathcal{O}$ of neighborhood operators. The set G of all bijections of Ω which commute with all elements of $\mathcal{O}$ is a subgroup of $\mathcal{A}_N$.*

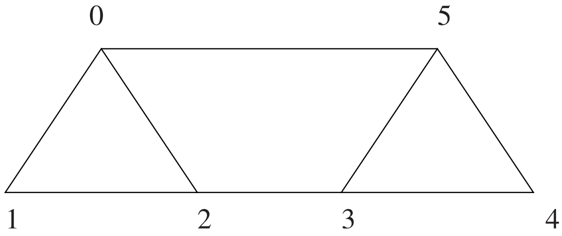

Fig. 1. Neighborhood graph of N

Proof. If $g, h \in G$ and $a \in \mathcal{O}$, then $g \circ h \circ a = g \circ a \circ h = a \circ g \circ h$, so $g \circ h \in G$. Moreover, $g \circ a = a \circ g$ implies $a \circ g^{-1} = g^{-1} \circ a$, and so $g \in G \Rightarrow g^{-1} \in G$. Hence G is a group.

Now suppose $y = a(x)$ for some $x, y \in \Omega$ and $a \in \mathcal{O}$, so that (x, y) is an edge of the neighborhood graph. If $g \in G$, then $g(y) = g \circ a(x) = a \circ g(x)$ so that $(g(x), g(y))$ is also an edge. Therefore $g \in \mathcal{A}_N$. □

Let N be the neighborhood structure generated by a collection $\mathcal{O}$ of neighborhood operators. Edge $x \to y$ in its neighborhood graph is said to have *color* ν if $\nu(x) = y$ and $\nu \in \mathcal{O}$ (an edge may have several colors). The group G (corresponding to $\mathcal{O}$ as in Theorem 1) not only preserves edges, it also preserves colors,

$$\nu(x) = y \Longleftrightarrow g \circ \nu(x) = g(y) \Longleftrightarrow \nu \circ g(x) = g(y)$$

The (edge-colored) neighborhood graph (of N) is said to correspond to $\mathcal{O}$. The group G is called the *symmetry group of* $\mathcal{O}$; its elements are called *symmetries* (of $\mathcal{O}$).

Definition 9. *The group of all bijections of Ω which commute with all elements of $\mathcal{O}$ is called the symmetry group of $\mathcal{O}$; its elements are called symmetries (of $\mathcal{O}$). The (edge-colored) neighborhood graph corresponding to the neighborhood structure generated by $\mathcal{O}$ is said to correspond to $\mathcal{O}$; edge $x \to y$ has color ν if $\nu(x) = y$ (for $\nu \in \mathcal{O}$; an edge may have several colors).*

Note that (by Theorem 1) the symmetry group of $\mathcal{O}$ is a subgroup of the symmetry group of the neighborhood graph corresponding to $\mathcal{O}$; the former preserves edges *and color*, whereas the latter need preserve only edges. For instance, the permutation (02)(35) is a symmetry of Figure 1 (i.e., a symmetry of N) but it is not a symmetry of the collection of neighborhood operators in example 1.

The next section deals with symmetries of $\mathcal{O}$.

3 Transitive Automorphism Groups

When using neighborhood operators as the basis of a search algorithm, the corresponding neighborhood graph is typically connected. As illustrated by example 2, connectivity is insufficient for the symmetry group of $\mathcal{O}$ to act transitively on Ω.

Nevertheless, when $\mathcal{O}$ acts on Ω the situation is quite different. If $\mathcal{O}$ is invertible, the group $\langle\mathcal{O}\rangle$ generated by $\mathcal{O}$ (under function composition) need not necessarily consist of symmetries of $\mathcal{O}$ (unless $\langle\mathcal{O}\rangle$ is Abelian), but it must act transitively; requiring that for all $x, y \in \Omega$ there exist $g \in \langle\mathcal{O}\rangle$ such that $g(x) = y$ is simply a restatement of connectivity.

Lemma 1. *Suppose the neighborhood graph corresponding to a collection $\mathcal{O}$ of neighborhood operators is connected. The only symmetry of $\mathcal{O}$ which has a fixed point is the identity (symmetry g has fixed point x iff $g(x) = x$).*

Proof. Let symmetry g and $x, y \in \Omega$ be such that $g(x) = x$. Since the neighborhood graph is connected, there exist $a_1, \ldots, a_k \in \mathcal{O}$ such that $y = a_1 \circ \cdots \circ a_k(x)$. Hence

$$g(y) = g \circ a_1 \circ \cdots \circ a_k(x) = a_1 \circ \cdots \circ a_k \circ g(x) = a_1 \circ \cdots \circ a_k(x) = y$$

Since y is arbitrary, g is the identity. □

Lemma 1, together with Theorem 17 from [RVW02], implies that if the neighborhood graph corresponding to $\mathcal{O}$ is connected, and if the symmetry group G of $\mathcal{O}$ acts transitivity, then there is a natural group structure isomorphic to G that can be ascribed to the search space itself. If the neighborhood graph corresponding to $\mathcal{O}$ is connected and if $\mathcal{O}$ is invertible, then the same conclusion holds with respect to $\langle\mathcal{O}\rangle$; there is a natural group structure isomorphic to $\langle\mathcal{O}\rangle$ that can be ascribed to Ω. These observations are recorded in the following theorem.

Theorem 2. *Suppose the neighborhood graph corresponding to a collection $\mathcal{O}$ of neighborhood operators is connected. If the symmetry group G of $\mathcal{O}$ acts transitively, then Ω has a group structure compatible with G (the search space can be given a group structure isomorphic to G such that the action of Ω on itself – via the group operation – is isomorphic to the action of G on Ω). If in addition $\mathcal{O}$ is invertible, then Ω has a group structure compatible with $\langle\mathcal{O}\rangle$.*

Proof. If G acts transitivity, then—*to use the language of* [RVW02]—it is a reduced group action on Ω (the only permutation fixing Ω is the identity), and Lemma 1 (above) implies that $\mathbf{Fix}(w) = \{0\}$, for all $w \in \Omega$ (only the identity fixes w); therefore Theorem 17 from [RVW02] applies to show that Ω has a group structure compatible with G.

If $\mathcal{O}$ is invertible, the comments preceding Lemma 1 imply the action of $\langle\mathcal{O}\rangle$ is reduced (the only permutation fixing Ω is the identity). Since $\nu(w) = w \Longrightarrow \nu(g(w)) = g(\nu(w)) = g(w)$ for all $g \in G$, it follows that if ν fixes w then ν is the identity (since G acts transitively); hence Theorem 17 from [RVW02] applies to show that G has a group structure compatible with $\langle\mathcal{O}\rangle$. □

Suppose that one starts with a connected undirected neighborhood graph. For each edge (x, y) of this graph, define an automorphism $a_{(x,y)}$ of Ω by $a_{(x,y)}(x) = y$, $a_{(x,y)}(y) = x$, and $a_{(x,y)}(z) = z$ for all $z \neq x, y$. (The permutation $a_{(x,y)}$ is the transposition that is denoted by (x, y) in cycle notation.) Let $\mathcal{O}$ be the collection of these automorphisms. We will call the elements of $\mathcal{O}$ the edge transpositions of the graph. Then the graph of the neighborhood structure generated by $\mathcal{O}$ is the graph that we started with, except that loops have been added at each vertex. One might wonder if the symmetry group corresponding to this $\mathcal{O}$ would give Ω as a group structure as in theorem 2. The following small example shows that this is not necessarily the case.

Let $\Omega = \{0, 1, 2, 3\}$ and let $\mathcal{O} = \{(0, 1), (1, 2), (2, 3), (0, 3)\}$, where the elements of $\mathcal{O}$ are permutations written in cycle notation. The corresponding neighborhood graph is the square with loops at each vertex which is connected, so $\langle\mathcal{O}\rangle$ acts transitively on Ω. But since $\langle\mathcal{O}\rangle$ consists of all permutations of Ω, the symmetry group G of $\mathcal{O}$ is identity, and theorem 2 does not apply.

The example can be generalized to $\Omega = \{0, 1, \ldots, n-1\}$ by defining $\mathcal{O}$ to be the set of edge transpositions of the edges of any connected graph whose vertices are Ω. We can show by induction on n that $\langle\mathcal{O}\rangle$ is the set S_n of all permutations of Ω. The base case is trivial. The induction hypothesis is that the edge transpositions of a connected graph on the vertices $\{0, 1, \ldots, n-2\}$ generates the set S_{n-1} of permutations of $\{0, 1, \ldots, n-2\}$. There must be an edge $(i, n-1)$ of the graph that contains $n-1$. We need to show that $S_{n-1} \cup \{(i, n-1)\}$ generates S_n. Let σ be any permutation of $S_n \setminus S_{n-1}$, and let $j = \sigma(n-1)$. Note that $(i, j) \cdot (i, n-1) \cdot (i, j) = (j, n-1)$. Let $\tau = (j, n-1) \cdot \sigma \in S_{n-1}$. Then $\sigma = (j, n-1) \cdot \tau = (i, j) \cdot (i, n-1) \cdot (i, j) \cdot \tau$.

In the case where the neighborhood operators mutually commute (which happens in example 1, for instance), then the following result holds.

Corollary 1. *Suppose the neighborhood graph corresponding to a collection $\mathcal{O}$ of neighborhood operators is connected. If $\mathcal{O}$ is invertible and if its elements commute, then its symmetry group is the group $\langle\mathcal{O}\rangle$ generated by $\mathcal{O}$ and Ω has a group structure compatible with $\langle\mathcal{O}\rangle$ (the search space can be given a group structure isomorphic to $\langle\mathcal{O}\rangle$ such that the action of Ω on itself – via the group operation – is isomorphic to the action of $\langle\mathcal{O}\rangle$ on Ω).*

Proof. If the bijections of $\mathcal{O}$ commute, then $\langle\mathcal{O}\rangle$ is a subgroup of the symmetry group G (of $\mathcal{O}$). Since the neighborhood graph is connected, the action of $\langle\mathcal{O}\rangle$ on Ω is transitive. Therefore, the action of G is also transitive and Theorem 2 applies. Note that $\mid G \mid = \mid \langle\mathcal{O}\rangle \mid$, since the search space has group structures isomorphic to G and to $\langle\mathcal{O}\rangle$. It follows that G is actually the same as $\langle\mathcal{O}\rangle$. □

In the case corresponding to corollary 1, the group structure ascribed to Ω is commutative. Each element of Ω is identified with a list of those operators that must be applied in order to reach it from 0. In this case, the search space is *structural*: a situation which is dealt with in detail in [RVW04].

4 Implementation

Having described how a search space might come to have a group acting upon it, and may in fact have a group structure ascribed to it (so that it acts transitively on itself via the group operation), consider the issue of designing mutation and crossover operators for such search spaces.

For the remainder of the paper, assume that Ω does indeed have its own group structure (the group operation is denoted by $\circ$). Let $m(x)$ denote the result of mutating x, and let $c(\{x, y\})$ denote the result of crossing parents x and y.

The formal requirement for mutation and crossover to commute with the group $(\Omega, \circ)$ is cast within the framework of *Random Heuristic Search* [Vos99a], [Vos99b, RVW02]. A population is represented by a distribution $p \in \Lambda$, where

$$\Lambda \;=\; \{p \in \mathrm{R}^n \;:\; p_k \geq 0,\;\; \sum p_k = 1\}$$

and where p_i is the proportion of i in the population. An element a of the group $(\Omega, \circ)$ corresponds to a permutation matrix $\sigma_a : \Lambda \to \Lambda$ defined by

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

Crossover acts on distribution p by mapping it to $\mathcal{C}(p)$ where

$$\mathcal{C}(p)_k \;=\; p^T M_k\, p$$

(superscript T denotes transpose) and where the matrix M_k is defined by

$$(M_k)_{i,j} \;=\; \mathrm{Prob}\{c(\{i, k\}) = k\}$$

Crossover is said to *commute with* $(\Omega, \circ)$ if for all $p \in \Lambda$ and all $a \in \Omega$,

$$\mathcal{C}(\sigma_a p) \;=\; \sigma_a \mathcal{C}(p)$$

An advantage of considering genetic operators (crossover, mutation, selection) at the distribution level is that some analysis can proceed with differentiable objects, and, finite population information is preserved [Vos99b, RVW05, RVW06].

A similar situation holds for mutation; it has a corresponding operator $\mathcal{U}$ at the distribution level, and, it is said to commute with $(\Omega, \circ)$ if for all $p \in \Lambda$ and all $a \in \Omega$,

$$\mathcal{U}(\sigma_a p) \;=\; \sigma_a \mathcal{U}(p)$$

The definitions above – for what it means for crossover and mutation to commute with the group $(\Omega, \circ)$ – are given to provide context and to be technically accurate. However, they are not displayed for subsequent use in this paper; working definitions are instead provided by the following theorem.

Theorem 3. *Crossover and mutation commute with* $(\Omega, \circ)$ *if and only if for all* $w, x, y, z \in \Omega$

$$Prob\{w = m(x)\} = Prob\{z \circ w = m(z \circ x)\}$$
$$Prob\{w = c(\{x, y\})\} = Prob\{z \circ w = c(\{z \circ x, z \circ y\})\}$$

Proof. Theorem 3 is a rephrasing of Theorems 5 and 6 from [RVW02]. □

4.1 Binary Crossover

Let $\mathcal{B}$ be the set of maps from Ω to itself. Given a probability distribution χ over $\mathcal{B}$, form an offspring from parents $u, v \in \Omega$ by

1. choosing an element $b \in \mathcal{B}$ according to χ,
2. returning (with equal probability) an element from

$$\{v \circ b(v^{-1} \circ u),\, u \circ b(u^{-1} \circ v)\}$$

This crossover method is called the *canonical crossover scheme.*

Theorem 4. *The canonical crossover scheme commutes with* $(\Omega, \circ)$*.*

Proof. Define the function r by

$$r(u, v, w) = \sum_{b \in \mathcal{B}} \chi_b [v \circ b(v^{-1} \circ u) = w]$$

The probability that u and v cross to form w is

$$s(u, v, w) = \frac{r(u, v, w) + r(v, u, w)}{2}$$

For every $z \in \Omega$

$$\begin{aligned} r(z \circ u, z \circ v, z \circ w) &= \sum_{b \in \mathcal{B}} \chi_b [z \circ v \circ b(v^{-1} \circ z^{-1} \circ z \circ u) = z \circ w] \\ &= \sum_{b \in \mathcal{B}} \chi_b [v \circ b(v^{-1} \circ u) = w] \\ &= r(u, v, w) \end{aligned}$$

It follows from Theorem 3 that crossover commutes with $(\Omega, \circ)$. □

> **Example 3.** Let Ω be the set of length ℓ binary strings under the group operation $\oplus$ of bitwise exclusive-or (in particular, $v^{-1} = v$). For every $k \in \Omega$ define the map $b_k \in \mathcal{B}$ by
>
> $$b_k(x) = k \otimes x$$
>
> where $\otimes$ denotes bitwise and. Let χ be a probability distribution that only assigns non-zero weight to such maps, and thus can be thought of as a probability distribution over Ω. Therefore,
>
> $$\begin{aligned} r(u, v, w) &= \sum_{b \in \mathcal{B}} \chi_b [v \oplus b(v \oplus u) = w] \\ &= \sum_{k \in \Omega} \chi_k [v \oplus (k \otimes (v \oplus u)) = w] \\ &= \sum_{k \in \Omega} \chi_k [v \oplus (k \otimes v) \oplus (k \otimes u) = w] \\ &= \sum_{k \in \Omega} \chi_k [((k \oplus \mathbf{1}) \otimes v) \oplus (k \otimes u) = w] \\ &= \sum_{k \in \Omega} \chi_k [(\bar{k} \otimes v) \oplus (k \otimes u) = w] \end{aligned}$$

where **1** is the string of all ones and $\bar{k}$ is the binary complement of k. It follows that the crossover scheme implements crossover by masks (see [Vos99b]).

According to Theorem 2 of [RVW02], every two-parent crossover commuting with $(\Omega, \circ)$ is completely determined by a mixing matrix M with the property that for all $i, j, k \in \Omega$,

$$\text{Prob}\{c(\{i,j\}) = k\} = M_{k^{-1}\circ i,\, k^{-1}\circ j}$$

Moreover, Theorem 19 of [RVW02][3] implies that $M = A^*$ for some row stochastic matrix A, where the *twist* A^* of A is defined by

$$A^*_{u,v} = A_{u^{-1}\circ v,\, u^{-1}}$$

In particular, $A = M^{**}$, since the twist operator has order three (i.e., $A = A^{***}$ for every matrix A). Another property of the twist is that $A^{*T} = A^{T**}$. Since mixing matrices are symmetric,

$$2M = A^* + A^{*T} \;=\; (A + A^{T*})^*$$

and therefore

$$\tfrac{1}{2}(A + A^{T*}) \;=\; M^{**}$$

Theorem 5. *Given any row stochastic matrix A, there exists a probability distribution α over $\mathcal{B}$ such that*

$$A_{i,j} = \sum_{f\in\mathcal{B}} \alpha_f [f(i) = j]$$

Moreover, given any such α, the formula above defines a row stochastic matrix.

Proof. Any matrix having the form above is row stochastic, since

$$\sum_j A_{i,j} \;=\; \sum_{f\in\mathcal{B}} \alpha_f \sum_j [f(i) = j] \;=\; \sum_{f\in\mathcal{B}} \alpha_f \;=\; 1$$

Conversely, if α is defined by

$$\alpha_f \;=\; \prod_{k\in\Omega} A_{k,f(k)}$$

then, using the identification $f \leftrightarrow \langle f_0, \ldots, f_{n-1}\rangle$ where $f_i = f(i)$,

$$\sum_{f\in\mathcal{B}} \alpha_f [f(i) = j] = \sum_{f_0\in\Omega} \sum_{f_1\in\Omega} \cdots \sum_{f_{n-1}\in\Omega} \Big(\prod_{k\in\Omega} A_{k,f_k}\Big)[f_i = j]$$

---

[3] The wording of Theorem 19 is directed towards *constructing* mixing matrices, hence the requirement to symmetrize (mixing matrices are symmetric). If one is representing a mixing matrix (see the discussion before Theorem 19), there is no need to symmetrize; mixing matrices are by definition symmetric.

$$
\begin{aligned}
&= \sum_{f_0 \in \Omega} A_{0,f_0} \sum_{f_1 \in \Omega} A_{1,f_1} \ldots \sum_{f_{n-1} \in \Omega} A_{n-1,f_{n-1}} [f_i = j] \\
&= \sum_{f_i \in \Omega} A_{i,f_i} [f_i = j] \\
&= A_{i,j}
\end{aligned}
\qquad \square
$$

For every $f \in \mathcal{B}$, define the function $\widehat{f} \in \mathcal{B}$ by

$$\widehat{f}(x) = x \circ f(x^{-1})$$

The transformation $\widehat{\ }: \mathcal{B} \to \mathcal{B}$ is a bijection (it is self-inverse).

Theorem 6. *Every two-parent crossover commuting with* $(\Omega, \circ)$ *is an instance of the canonical crossover scheme.*

Proof. By what has been explained above, it suffices to choose the distribution χ in the canonical crossover scheme such that the resulting probability of obtaining k by crossing parents i and j (via the canonical scheme) is

$$M_{k^{-1} \circ i,\, k^{-1} \circ j} \tag{1}$$

Define χ by

$$\chi_f \;=\; \tfrac{1}{2}(\alpha_f + \alpha_{\widehat{f}})$$

where α_f is the distribution referred to in Theorem 5 for $A = M^{**}$. Note that $\chi_f = \chi_{\widehat{f}}$, and the quantification $f \in \mathcal{B}$ is the same as $\widehat{f} \in \mathcal{B}$. The probability of obtaining k is

$$
\begin{aligned}
&\tfrac{1}{2} \sum_{f \in \mathcal{B}} \chi_f [j \circ f(j^{-1} \circ i) = k] + \tfrac{1}{2} \sum_{\widehat{f} \in \mathcal{B}} \chi_{\widehat{f}} [i \circ \widehat{f}(i^{-1} \circ j) = k] \\
&= \tfrac{1}{2} \sum_{f \in \mathcal{B}} \chi_f [j \circ f(j^{-1} \circ i) = k] + \tfrac{1}{2} \sum_{f \in \mathcal{B}} \chi_f [i \circ (i^{-1} \circ j) \circ f(j^{-1} \circ i) = k] \\
&= \sum_{f \in \mathcal{B}} \chi_f [j \circ f(j^{-1} \circ i) = k] \\
&= \tfrac{1}{2} \sum_{f \in \mathcal{B}} \alpha_f [f(j^{-1} \circ i) = j^{-1} \circ k] + \tfrac{1}{2} \sum_{f \in \mathcal{B}} \alpha_{\widehat{f}} [f(j^{-1} \circ i) = j^{-1} \circ k]
\end{aligned}
$$

Setting $u = j^{-1} \circ i$, $v = j^{-1} \circ k$, and re-indexing the second sum in the last line above yields

$$
\begin{aligned}
&\tfrac{1}{2} \sum_{f \in \mathcal{B}} \alpha_f [f(u) = v] + \tfrac{1}{2} \sum_{\widehat{f} \in \mathcal{B}} \alpha_f [\widehat{f}(u) = v] \\
&= \tfrac{1}{2} A_{u,v} + \tfrac{1}{2} \sum_{f \in \mathcal{B}} \alpha_f [u \circ f(u^{-1}) = v] \\
&= \tfrac{1}{2} A_{u,v} + \tfrac{1}{2} A_{u^{-1},\, u^{-1} \circ v}
\end{aligned}
$$

$$
\begin{aligned}
&= \tfrac{1}{2}(A + A^{T*})_{u,v} \\
&= M^{**}_{u,v} \\
&= M_{v^{-1},v^{-1}\circ u} \\
&= M_{k^{-1}\circ j,k^{-1}\circ i}
\end{aligned}
$$

This agrees with (1) since M is symmetric. $\square$

4.2 Mutation

Theorem 20 of [RVW02] and the Corollary following it describe how to implement mutation: given a probability distribution μ over Ω, mutate $j \in \Omega$ by

1. choosing an element $k \in \Omega$ according to μ,
2. returning the element $j \circ k$

According to the Theorem, all possible mutation operators which commute with $(\Omega, \circ)$ are of this form.

> **Example 4** Let Ω be the set of length ℓ binary strings under the group operation of bitwise exclusive-or. Then the mutation scheme above implements mutation by masks.

Mutation can be seen as a special case of crossover, in the sense that the resulting child will be the mutation of some parent. For each element $k \in \Omega$ define the map $b_k \in \mathcal{B}$ by

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

Using these maps to implement crossover (nonzero probability is assigned by χ only to such maps), the set of possible children resulting from parents u, v and map b_k is

$$\{v \circ b_k(v^{-1} \circ u),\, u \circ b_k(u^{-1} \circ v)\} = \{u \circ k,\, v \circ k\}$$

Moreover, choosing χ to satisfy

$$\chi_{b_k} = \mu_k$$

arranges for the resulting child (of crossover) to not only be the mutation of a parent, but to occur according to the probabilities specified by μ.

5 Conclusion

This paper introduces neighborhood structures and their symmetries, and describes how neighborhood structures naturally arise as a consequence of neighborhood operators. Conditions are given under which neighborhood operators have symmetries which may be ascribed to the search space itself. This motivates—by providing a concrete example—the situation with which the paper is primarily concerned; the search space itself has a group structure.

Irrespective of how or why the search space may have a group structure, the main result is that those classical genetic operators (mutation and two-parent crossover) which respect the group action are completely characterized.

Formulas are given which define such genetic operators in terms of the search space and its group operation. This provides a representation-free implementation for those operators, in the sense that the genotypic encoding of search space elements is irrelevant. The implementations are parameterized by distributions which may be chosen arbitrarily, and which are analogous to specifying distributions for mutation and crossover masks (when specialized to a classical fixed-length binary GA, the standard crossover and mutation operators defined by masks result).

References

[Big71] Biggs, N.L.: Finite groups of automorphisms. Cambridge University Press, Cambridge, UK (1971)

[RHKS02] Rockmore, D., Hordijk, W., Kostelec, P., Stadler, P.: Fast fourier transforms for fitness landscapes. Applied and Computational Harmonic Analysis 12(1), 57–76 (2002)

[RVW02] Rowe, J.E., Vose, M.D., Wright, A.H.: Group properties of crossover and mutation. Evolutionary Computation 10, 151–184 (2002)

[RVW04] Rowe, J.E., Vose, M.D., Wright, A.H.: Structural search spaces and genetic operators. Evolutionary Computation 12, 461–494 (2004)

[RVW05] Rowe, J.E., Vose, M.D., Wright, A.H.: Coarse graining selection and mutation. In: Wright, A.H., Vose, M.D., De Jong, K.A., Schmitt, L.M. (eds.) FOGA 2005. LNCS, vol. 3469, pp. 176–191. Springer, Heidelberg (2005)

[RVW06] Rowe, J.E., Vose, M.D., Wright, A.H.: Differentiable coarse graining. Theoretical Computer Science 361, 111–129 (2006)

[Sta96] Stadler, P.F.: Landscapes and their correlation functions. Journal of Mathematical Chemistry 20(1), 1–45 (1996)

[Vos99a] Vose, M.D.: Random heuristic search. Theoretical Computer Science 229, 103–142 (1999)

[Vos99b] Vose, M.D.: The Simple Genetic Algorithm: Foundations and Theory. MIT Press, Cambridge, MA (1999)

[Wei91] Weinberger, E.D.: Fourier and taylor series on fitness landscapes. Biological Cybernetics 65(5), 321–330 (1991)