Crossover Invariant Subsets of the Search Space for Evolutionary Algorithms

Boris Mitavskiy bmitavsk@umich.edu
Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109, USA

Abstract

This paper addresses the relationship between schemata and crossover operators. In Appendix A a general mathematical framework is developed which reveals an interesting correspondence between the families of reproduction transformations and the corresponding collections of invariant subsets of the search space. On the basis of this mathematical apparatus it is proved that the family of masked crossovers is, for all practical purposes, the largest family of transformations whose corresponding collection of invariant subsets is the family of Antonisse's schemata. In the process, a number of other interesting facts are shown. It is proved that the full dynastic span of a given subset of the search space under either one of the traditional families of crossover transformations (one-point crossovers or masked crossovers) is obtained after $\lceil \log_2 n \rceil$ iterations where n is the dimension of the search space. The generalized notion of invariance introduced in the current paper unifies Radcliffe's notions of "respect" and "gene transmission". Besides providing basic tools for the theoretical analysis carried out in the current paper, the general facts established in Appendix A provide a way to extend Radcliffe's notion of "genetic representation function" to compare various evolutionary computation techniques via their representation.

Keywords

Genetic algorithms, schema, forma, respect, invariance under crossover, dynastic span.

1 Introduction

One of the approaches to the study of the global behavior of genetic algorithms, initiated by John Holland, is to consider a certain family of subsets of the search space and to predict which ones of these subsets, say Q, possess the following target property:

Target Property: *The expected number of occurrences of elements of Q increases from one generation to the next.*

Holland, motivated by the binary genetic algorithm setting, defined this family of subsets in terms of the so-called schemata (see, for example, (Michalewicz, 1996)). To define all these notions explicitly, suppose we are dealing with a fixed length genetic algorithm, say all chromosomes are of fixed length n. Then our search space $S = \prod_{i=1}^{n} A_i$ where A_i denotes the set of all possible alleles which may occur at the i^{th} position in a chromosome. For example, in the case of a binary genetic algorithm, we have $A_i = \{0, 1\}$ for all i. In our setting, the set of all schemata $\mathcal{A} = \prod_{i=1}^{n}(A_i \cup \{*\})$ (here it is assumed, of course, that $* \notin A_i$ for all i). Every schema, say $(a_1, \ldots, a_n)$, represents a subset of the search space S of the form $\prod_{i=1}^{n} K_i$ where $K_i = \begin{cases} \{a_i\} & \text{if } a_i \neq * \\ A_i & \text{otherwise} \end{cases}$. For example, when $n = 4$ and $A_i = A = \{1, 2, 3, 4, 5\}$ for all i, a schema $(1, *, *, 5)$ would describe the subset $Q = \{1\} \times A \times A \times \{5\}$ of the search space $S = \prod_{i=1}^{n} A_i = \prod_{i=1}^{n} A$.

Holland derived the classical Schema Theorem, which suggests that "short, low order, above average schemata are expected to receive increasing trials in the subsequent generations of a genetic algorithm" (see (Michalewicz, 1996)). This means, in our terms, that this particular class of schemata tends to satisfy the Target Property.

Antonisse (1989) generalized Holland's definition of schemata to incorporate all subsets of the search space of the form $\prod_{i=1}^{n} T_i$ where T_i is an arbitrary nonempty subset of A_i. Notice that the two notions of a schema coincide in case of a binary genetic algorithm, namely when $A_i = \{0, 1\}$ for all i. Antonisse has shown that Holland's results naturally generalize to this extended family of subsets. Later, (Vose, 1991) considered the family of all subsets of the search space S and separated out the properties which are crucial for a subset to have in order to possess the Target Property as stated above. To define these properties, let $f : S \to (0, \infty)$ denote the fitness function.

Monotonicity: *A subset* $Q \subseteq S$ *is monotone increasing iff* $Q = f^{-1}([b, \infty))$ *for some* $b \in [0, \infty)$.

Stability: *A subset* $Q \subseteq S$ *is stable iff it is invariant under crossover.*

Vose (Vose, 1991) has shown that stable and monotone increasing subsets $Q \subseteq S$ possess the Target Property. In this paper we shall prove that stable subsets of the search space S are precisely of the form $\prod_{i=1}^{n} T_i$ where T_i is an arbitrary subset of A_i, which is precisely the family of subsets considered by Antonisse.

Radcliffe (Radcliffe, 1994) generalized the notion of a Holland schema for a representation-independent setting. To make the current paper self-contained, a brief list of definitions from (Radcliffe, 1994) is given in appendix C. Radcliffe's notion of forma (see definition 4, the discussion following definition 4 and definition 5 of (Radcliffe, 1994) or, alternatively, definition C.1 of the current paper) captures the essential properties of a Holland schema so that the operators which "transmit genes" (see definition 47 of (Radcliffe, 1994) or, alternatively, definition C.3 of the current paper) for a given set of forma resemble the usual crossover. Radcliffe also introduced a number of important concepts characterizing the behavior of the "mating" transformations. These include *purity* (two identical parents produce the child identical to both of them) (see remark 3.3), and *respect* (refers to the family of mating transformations preserving a given family of forma: see definition 41 of (Radcliffe, 1994) or, alternatively, definition C.2 of the current paper). In the current paper the notion of *respect* is generalized and is referred to as "invariance" under a given family of transformations. Incidentally, the notion of invariance also appears in a variety of contexts in other papers. (See, for instance (Rowe, Vose and Wright, 2002), (Vose and Wright, 2001).) Finally, Radcliffe's notion of the *dynastic span* (see definition A.5) of a given set under a given family of transformations is strongly related to the new, generalized notion of respect. It turns out, in fact, that the full dynastic span of a given set L under a given family of transformations Γ is simply the smallest invariant set under Γ containing L (see proposition A.4). In section 3 we find an exact bound on the number of dynastic spans of a given subset L under the two traditional families of crossover operators (the one-point crossover and the masked crossover) one needs to obtain the entire smallest invariant set containing L. It turns out that the bound is independent of the specific family of transformations one exploits (one-point crossover or masked crossover) and is logarithmic of the dimension of the search space (see theorem 3.2). Notice also that the notion of "gene transmission" introduced by Radcliffe is a particular case of the generalized notion of "respect" (invariance) introduced in the current paper (compare definition C.3 and the discussion below definition C.3 to proposition 4.1 and definition 4.1).

The notion of invariance appears in various contexts in current research (see, for instance, (Vose, 1999), (Vose and Wright, 2001)). In one of the most recent papers, (Rowe, Vose and Wright, 2002), Jonathan E. Rowe, Michael D. Vose and Alden H. Wright define schemata in terms of certain transitive group actions on the search space. Sufficient conditions for a given family of crossover operators which commute with the group action to respect this family of generalized schemata are established (see Theorem 13 of (Rowe, Vose and Wright, 2002)). Some results establishing connections between respect and purity are proved (see Theorems 12 and 15 of (Rowe, Vose and Wright, 2002)).

There is another reason why the notion of invariance is important. It turns out that the correspondence between the poset of m-fixable families of subsets of the search space and the poset of all families of m-ary mating transformations[1] established in appendix A provides a way to extend Radcliffe's notion of "genetic representation function" (see page 10 of (Radcliffe, 1994)) to compare various evolutionary computation techniques via their representations. A few theorems which illustrate this idea appear in the sequel papers (see (Mitavskiy, 2003a) and (Mitavskiy, 2003b)).

2 Notation and Summary of the Results from Appendix A

A less patient reader may want to postpone reading this section and proceed directly to section 3. The reader will then be referred to this section when necessary. Throughout the paper the following notation will be used:

Definition 2.1. Let Ω denote an arbitrary set. Let Γ denote a nonempty family of transformations from Ω^m to Ω for a fixed $m \geq 1$. We then denote by $\Lambda_\Gamma = \{S \mid S \subseteq \Omega,\ T(S^m) \subseteq S\ \forall\ T \in \Gamma\}$ the family of invariant subsets of Ω under the action of Γ.

In what follows, the family of transformations Γ usually represents a certain family of reproduction transformations. For instance, in section 3 these are the families of one-point and masked crossover transformations. There are two types of extensions that any family of reproduction transformations has. Both of them play an important role in the current paper. The first notion is the *composition closure* of a given family of transformations Γ, denoted by $\bar{\Gamma}$. Intuitively this is simply the family of all transformations obtained by repeatedly applying the transformations in Γ (see definition A.6). Another important extension of the family of transformations Γ is the family $\widetilde{\Gamma}$ (see definition A.7). $\widetilde{\Gamma}$ is actually defined to be an extension of the family $\bar{\Gamma}$ (hence, also an extension of Γ). One of the main results established in appendix A is that the family $\widetilde{\Gamma}$ is the largest family of transformations whose family of invariant subsets, $\Lambda_{\widetilde{\Gamma}} = \Lambda_\Gamma$. On the other hand, the family $\widetilde{\Gamma}$ is related to the family $\bar{\Gamma}$ in such a way that $\bar{\Gamma}$ is sufficiently large to model an evolutionary algorithm whose family of reproduction transformations is $\widetilde{\Gamma}$ (see discussion following definition A.8).

3 Comparing the Dynamics of the Subsets of the Search Space under Various Families of Crossover Operators.

Consider the limiting distribution obtained by repeatedly applying crossover to a given population. Clearly, only the elements consisting of the alleles present in the initial population have nonzero probability. Moreover, the exact probability distribution is known: alleles occur with a probability equal to their frequency in the initial population. This result is known as Geiringer's Theorem (see (Geiringer, 1944)). This theorem has also appeared in a number of publications in the theory of genetic algorithms (see

[1] Such a correspondence is known as a Galois Connection. See, for instance, (MacLane, 1971).

for example (Booker, 1993), (Vose and Wright, 1998), (Stephens and Waelbroeck, 1999) and (Stephens, 2002)).

Some of the modern research in genetic algorithm (GA) theory has been devoted to the study of crossover and mutation operators. For example, crossover defined by a probability distribution over binary masks is considered in detail by Michael Vose (see (Vose, 1999)). Some results related to the "probabilistic" invariance are proved (see, for instance, theorem 16.1 of (Vose, 1999)). Related properties of the crossover operator are considered in (Liepins and Vose, 1992).

In the current section we use the results from Appendix A to study the family of crossover operators alone and we make no assumptions about any probability distributions involved. This allows us to discover a tight bound related to the dynastic span of a given subset Q of the search space under the family of crossover operators (see theorem 3.2) and, also, to demonstrate that the family of masked crossover transformations is, for all practical purposes, the largest family of transformations all of whose invariant subsets are precisely the products of the subsets of the sets of alleles (see corollary 3.4, proposition 3.1 and the discussion following definition A.8).

First, suppose we deal with a fixed length GA, say all chromosomes are of length n, and recall that our search space is represented by $S = \prod_{i=1}^{n} A_i$ where A_i denotes the set of all possible alleles which may occur on the i^{th} position in a chromosome. For the crossover operators we shall fix the following notation: $L_i : S^2 \to S$, $R_i : S^2 \to S$, $L_M : S^2 \to S$ for $M \subseteq \{1, 2, ...n\}$ denote the following crossover operators:

$$L_i(\mathbf{a}, \mathbf{b}) = (a_1, a_2, \ldots, a_i, b_{i+1}, \ldots, b_n) \tag{1}$$

$$R_i(\mathbf{a}, \mathbf{b}) = (b_1, b_2, \ldots, b_i, a_{i+1}, \ldots, a_n) \tag{2}$$

$$L_M(\mathbf{a}, \mathbf{b}) = (x_1, x_2, \ldots, x_i, \ldots, x_n) \tag{3}$$

where $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n) \in S$ and $x_i = \begin{cases} a_i & \text{if } i \in M \\ b_i & \text{otherwise} \end{cases}$.

Remark 3.1. Notice that L_n and L_0 are the first and the second coordinate projections respectively.

Definition 3.1. We say that a subset $Q \subseteq S$ is invariant under one point crossover iff for all i in the range $0 \leq i \leq n$ we have $L_i(Q^2) \subseteq Q$.

Remark 3.2. Notice that, since $L_i(\mathbf{a}, \mathbf{b}) = R_i(\mathbf{b}, \mathbf{a})$ it follows immediately that $L_i(Q^2) \subseteq Q$ iff $R_i(Q^2) \subseteq Q$ so that in the definition above one can use L_i and R_i interchangeably.

In the notation of Appendix A, if $\Omega = S = \prod_{i=1}^{n} A_i$, $m = 2$, and $\mathcal{F}_1 = \{L_i \mid 1 \leq i \leq n\}$ denotes the family of one point crossover operators, then $Q \subseteq S$ is invariant under one point crossover if and only if $Q \in \Lambda_{\mathcal{F}_1}$ (see definition 2.1 in section 2 for the meaning of $\Lambda_{\mathcal{F}_1}$).

Remark 3.3. Notice also that crossover operators L_i and R_i are pure in the sense defined in (Radcliffe, 1994), meaning that $L_i(\mathbf{a}, \mathbf{a}) = R_i(\mathbf{a}, \mathbf{a}) = \mathbf{a}$ for any $\mathbf{a} \in S$, and, therefore, for any $Q \subseteq S$ we have $Q \subseteq L_i(Q^2)$ so that $L_i(Q^2) \subseteq Q$ iff $L_i(Q^2) = Q$ iff $R_i(Q^2) = Q$.

The following definition was introduced in (Radcliffe, 1994) under the name of "respect" for a given recombination operator. In the current paper we refer to this notion as invariance under a given family of transformations:

Definition 3.2. We say that a subset $Q \subseteq S$ is invariant under masked crossover iff for all $M \subseteq \{1, 2, \ldots n\}$ we have $L_M(Q^2) \subseteq Q$.

Likewise, in the notation of Appendix A, if $\Omega = S = \prod_{i=1}^{n} A_i$, $m = 2$, and $\mathcal{F}_M = \{L_M \mid M \subseteq \{1, 2, \ldots, n\}\}$ denotes the family of masked crossover operators, then $Q \subseteq S$ is invariant under masked crossover if and only if $Q \in \Lambda_{\mathcal{F}_M}$ (see definition 2.1 for the meaning of $\Lambda_{\mathcal{F}_M}$).

Remarks analogous to the ones above apply to the definition of invariance under masked crossover.

Both $\mathcal{F}_1 = \{L_i \mid 1 \leq i \leq n\}$, the family of one point crossover transformations, and $\mathcal{F}_M = \{L_M \mid M \subseteq \{1, 2, \ldots, n\}\}$, the family of masked crossover transformations, are families of operators $S^2 \to S$, so we can apply the general results from the appendix with $\Omega = S$, $m = 2$, and $\Gamma = \mathcal{F}_1$ or $\mathcal{F}_M$. It turns out that the family of masked crossover operators is the composition closure of the family of one point crossover operators in the sense of definition A.6:

Proposition 3.1. *The family of masked crossover transformations $\mathcal{F}_M$ is the composition closure of the family of one point crossover transformations. In the notation of Appendix A this says that $\mathcal{F}_M = \bar{\mathcal{F}}_1$ (see definition A.6 and discussion in section 2 for the meaning of $\bar{\mathcal{F}}_1$).*

Proof. The proof does not use any concepts which are of interest for the remainder of this paper and, hence, is given in Appendix B. □

Now, we investigate the dynamics of the subsets of the search space under the various crossover operators. The following definition was first introduced in (Radcliffe, 1994). In full generality the definition also appears in Appendix A (see definition A.5).

Definition 3.3. Given a subset $Q \subseteq S$, we define the dynastic span of Q under a given family $\mathcal{F}$ of transformations of Q^2 into Q to be the set $\mathcal{F}(Q) = \bigcup_{L \in \mathcal{F}} L(Q^2)$. We shall also write $\mathcal{F}^2(Q) = \mathcal{F}(\mathcal{F}(Q))$ and, in general, $\mathcal{F}^k(Q) = \underbrace{\mathcal{F}(\mathcal{F}(\mathcal{F}(\ldots \mathcal{F}}_{k \text{ times}}(Q) \ldots))$ for $k \geq 1$ and $\mathcal{F}^0(Q) = Q$. We say that $\mathcal{F}^\infty = \bigcup_{k=0}^{\infty} \mathcal{F}^k(Q)$ is the full dynastic span of Q under $\mathcal{F}$. We shall denote by $\mathcal{F}_1$ the family of one-point crossover transformations and by $\mathcal{F}_M$ the family of masked crossover transformations.

To state the following Theorem and the Lemma needed to prove it fix the following notation:

Definition 3.4. Given the search space $S = \prod_{i=1}^{n} A_i$, and a subset $M \subseteq \{1, 2, \ldots, n\}$, denote by $p_M : S \to \prod_{i \in M} A_i$ the usual coordinate projection. That is, if $M = \{i_1, \ldots, i_q\}$ with $i_1 < i_2 < \ldots < i_q$, then for all $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ we have $p_M(\mathbf{x}) = (x_{i_1}, x_{i_2}, \ldots, x_{i_q})$. Intuitively speaking, p_M simply forgets all the digits not indexed by the elements of M. We shall also write p_i instead of $p_{\{i\}}$ for the usual i^{th} coordinate projection.

Theorem 3.2. *Fix a subset $Q \subseteq S$. Then we have the following finite chains of proper inclusions: $Q = \mathcal{F}_1^0(Q) \subsetneq \mathcal{F}_1(Q) \subsetneq \mathcal{F}_1^2(Q) \subsetneq \ldots \subsetneq \mathcal{F}_1^k(Q) = \prod_{i=1}^{n} p_i(Q)$. Likewise, we have $Q = \mathcal{F}_M^0(Q) \subsetneq \mathcal{F}_M(Q) \subsetneq \mathcal{F}_M^2(Q) \subsetneq \ldots \subsetneq \mathcal{F}_M^l(Q) = \prod_{i=1}^{n} p_i(Q)$ where k is the smallest nonnegative integer such that $\mathcal{F}_1^k(Q) = \mathcal{F}_1^{k+1}(Q)$, and, likewise, l is the smallest nonnegative integer such that $\mathcal{F}_M^l(Q) = \mathcal{F}_M^{l+1}(Q)$.* [2]

Moreover, k and l satisfy the following:

1. $l \leq k \leq \lceil \log_2 n \rceil$

[2] In particular, the theorem also asserts that $\mathcal{F}_1^k(Q) = \mathcal{F}_M^l(Q) = \prod_{i=1}^{n} p_i(Q)$. Of course, if $Q = \prod_{i=1}^{n} p_i(Q)$ then $k = l = 0$ and both chains terminate right away.

2. *The upper bounds on the values of k and l cannot be improved in the sense that for every $S = \prod_{i=1}^{n} A_i$ with $|A_i| \geq 2$ for all $1 \leq i \leq n$ there exists a subset $Q \subseteq S$ with both, k and l equal to $\lceil \log_2 n \rceil$.*

Proof. First, observe that, by straightforward induction, we do have $Q = \mathcal{F}_1^0(Q) \subsetneq \mathcal{F}_1(Q) \subsetneq \mathcal{F}_1^2(Q) \subsetneq \ldots \subsetneq \mathcal{F}_1^k(Q) \subseteq \prod_{i=1}^{n} p_i(Q)$ (roughly speaking, because no i^{th} coordinate of any element of the dynastic span of Q can lie outside $p_i(Q)$ since crossover operators only mix the coordinates and don't introduce any new ones.) The fact that $\mathcal{F}_1^k(Q) = \prod_{i=1}^{n} p_i(Q)$ together with item (1) in the statement of theorem 3.2 follows from the following more general Lemma which we prove by induction on the integer q.

Lemma 3.3. *$p_M(\mathcal{F}_1^q(Q)) \supseteq \prod_{i \in M} p_i(Q)$ for every $M \subseteq \{1, 2, \ldots, n\}$ with $|M| \leq 2^q$ (see definition 3.4 for the notation).*

Proof. For $q = 0$ the statement follows immediately because then $|M| = 1$, and we have $p_M = p_i$ for some i. But then our statement simply reduces to the fact that $p_i(Q) \subseteq p_i(Q)$ and, thereby, the base case is established.

Now suppose the statement is true for $q = j \geq 0$. Fix any index set $M \subseteq \{1, 2, \ldots, n\}$ with $|M| \leq 2^{j+1}$. Say, $M = \{i_1, i_2, \ldots, i_k\}$ and consider any $\mathbf{x} = (x_{i_1}, x_{i_2}, \ldots, x_{i_k}) \in \prod_{i \in M} p_i(Q)$. Let $k_1 = \lceil k/2 \rceil$. Notice, then that both k_1 and $k - k_1 \leq 2^{j+1}/2 = 2^j$, which implies that $M_1 = \{i_1, i_2, \ldots, i_{k_1}\}$ and $M_2 = \{i_{k_1+1}, i_{k_1+2}, \ldots, i_k\}$ both have cardinalities less than or equal to 2^j. By inductive hypothesis there exists an element $\mathbf{a} = (a_1, a_2, \ldots, a_n) \in \mathcal{F}_1^j(Q)$ with $a_{i_1} = x_{i_1}, \ldots a_{i_{k_1}} = x_{i_{k_1}}$ and an element $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathcal{F}_1^j(Q)$ with $b_{i_{k_1+1}} = x_{i_{k_1+1}}, \ldots, b_{i_k} = x_{i_k}$. Notice that the element $\mathbf{u} = L_{k_1}(\mathbf{a}, \mathbf{b}) \in \mathcal{F}_1^{j+1}(Q)$ has the desired property that $p_M(\mathbf{u}) = \mathbf{x}$. This establishes the inductive step so that by the principle of induction our statement holds for all q. □

From the Lemma above it follows immediately that as soon as $2^q \geq n$ we have $\mathcal{F}_1^q(Q) = \prod_{i=1}^{n} p_i(Q)$ so that we must have $\mathcal{F}_1^k(Q) = \prod_{i=1}^{n} p_i(Q)$ for some k, and $k \leq \lceil \log_2 n \rceil$.

Observe now, that the fact that $l \leq k$ is a special case of Proposition A.5 together with Theorem A.9 since the composition closure of the family of one point crossovers is the family of masked crossovers so that by part 1 of Proposition A.5 together with Theorem A.9 we obtain that $\prod_{i=1}^{n} p_i(Q) = \mathcal{F}_1^k(Q) \subseteq \mathcal{F}_M^k(Q) \subseteq \prod_{i=1}^{n} p_i(Q)$ so that all the containments above become equalities which forces $l \leq k$.

To prove part 2, let $S = \prod_{i=1}^{n} \{0, 1\}$ and let
$Q = \{(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots 0, 1, 0, \ldots 0), \ldots, (0, 0, \ldots, 0, 1)\}$.

In other words, Q is the set of rows of an n by n identity matrix. Clearly, $\prod_{i=1}^{n} p_i(Q) = S$ so that the element $(1, 1, \ldots, 1) \in \prod_{i=1}^{n} p_i(Q)$. On the other hand, a straightforward induction on q shows that any element of $\mathcal{F}_1^q(Q)$ contains at most 2^q 1's among its digits, so that $\mathcal{F}_1^q(Q) = \prod_{i=1}^{n} p_i(Q)$ implies that $q \geq \lceil \log_2 n \rceil$ which proves the desired assertion. □

Remark 3.4. An interesting observation about the proof of theorem 3.2 is that it makes no reference to the sets involved being finite. The results go through even if we allow infinite sets A_i at each coordinate, and allow Q to be an infinite subset.

Theorem 3.2 basically says that the "pathlength" from $Q \subseteq S$ to any individual from the set $U_Q = \prod_{i=1}^{n} p_i(Q)$ is at most logarithmic of the "dimension" n of the search space even if we allow an infinite set at each coordinate, and allow Q to be an infinite subset (see remark 3.4).

Remark 3.5. Although Geiringer's theorem does imply that the full dynastic span of a given subset Q under any one of the traditional families of the crossover operators is $\prod_{i=1}^{n} p_i(Q)$, unlike theorem 3.2, it does not (at least not immediately) imply that the $\lceil \log n \rceil^{\text{th}}$ dynastic span of Q is its full dynastic span (meaning that $\mathcal{F}^{\lceil \log_2 n \rceil}(Q) = \mathcal{F}^{\infty}(Q)$ (see definition 3.3). In fact, when neither one of the sets involved is finite, Geirenger's theorem does not even say that the full dynastic span of the family of crossover transformations is obtained after finitely many iterations.

Remark 3.6. Interestingly enough, the exact logarithmic bound established in theorem 3.2 does not depend on whether one uses the family of one-point crossover transformations or the entire family of masked crossovers.

It is well known that all schemata are invariant under any reasonable family of crossover operators. The next corollary shows that if the notion of a schema is defined in the same way as done by Antonisse [2], then schemata are actually the only crossover invariant subsets of the search space no matter which one of the traditional crossover operators is used.

Corollary 3.4. *Let $Q \subseteq S = \prod_{i=1}^{n} A_i$. The following are equivalent:*

1. $Q = \prod_{i=1}^{n} T_i$ *for some* $T_i \subseteq A_i$
2. *Q is invariant under masked crossover.*
3. *Q is invariant under one point crossover.*

Proof. The equivalence of 2 and 3 follows at once from part 1 of Theorem A.9 together with the fact that the family of masked crossover operators is the composition closure of the family of one point crossover operators. Thereby, it suffices to prove that, say 3 is equivalent to 1. To see this, fix a subset $Q \subseteq S$ and observe that by proposition A.4 Q is invariant under one point crossover if and only if the degree of Q under the action of the family of one point crossover transformations is 0 (see proposition A.4), which, according to theorem 3.2, happens if and only if $Q = \prod_{i=1}^{n} p_i(Q)$ which happens if and only if $Q = \prod_{i=1}^{n} T_i$ for some $T_i \subseteq A_i$ and so the desired conclusion follows. □

Remark 3.7. Notice that corollary 3.4 together with proposition 3.1 and discussion following definition A.8 imply that the masked crossovers form, for all practical purposes, one of the largest families of transformations whose collection of invariant subsets is $\{\prod_{i=1}^{n} T_i \,|\, T_i \subseteq A_i\}$.

Corollary 3.5. *For the classical binary GA all the stable sets are represented by the Holland schemata. Precisely, a subset of the search space is represented by a schema $\Longleftrightarrow$ it is invariant under one point crossover $\Longleftrightarrow$ it is invariant under masked crossover*

Proof. An immediate consequence of Corollary 3.4 applied to the search space $S = \prod_{i=1}^{n} \{0, 1\}$. □

4 Generalizing a Genetic Algorithm with Respect to Holland Schemata

Radcliffe (Radcliffe, 1994) observed that the notion of "respect" of a given family of schemata (or formae) by itself is not sufficient to describe the family of traditional masked crossovers. From the point of view of the current paper, this simply means that the family of traditional masked crossovers is not the largest for all practical purposes (see the discussion following definition A.8) family of transformations which leaves a given set of schemata (or formae) invariant. This motivated Radcliffe to introduce two other interesting notions: "gene transmission" (see definition 47 of (Radcliffe, 1994)) and the "Random Respectful Recombination Operator". The notion of gene transmission is stronger than the notion of respect (see Lemma 48 of Radcliffe). In the case of a classical GA, the results of the previous section and Appendix A together with proposition 4.1 below show that a given family of binary transformations Γ on the search space $S = \prod_{i=1}^{n} A_i$ transmits genes if and only if $\Gamma \subseteq \widetilde{\mathcal{F}_M}$ (See definition A.7 and discussion in section 2 for the meaning of $\widetilde{\mathcal{F}_M}$. Here $\widetilde{\mathcal{F}_M}$ plays the role of Γ.) if and only if $\Lambda_\Gamma \supseteq \{\prod_{i=1}^{n} T_i \,|\, T_i \subseteq A_i\}$ (see definition 2.1 for the meaning of Λ_Γ).

Denote by $\mathcal{V} = \{\prod_{i=1}^{n} T_i \,|\, T_i \subseteq A_i\}$ the family of subsets of the search space determined by Antonisse's schemata. In the previous section it was shown that $\mathcal{V} = \Lambda_{\mathcal{F}_M}$ (see corollary 3.4 and definition 2.1) which, in turn, implies that $\mathcal{V}$ is a 2-fixable family of subsets (see theorem A.3). The proposition below classifies all of the 2-basic subsets of $\mathcal{V}$:

Proposition 4.1. *A given subset of the search space, $Q = \prod_{i=1}^{n} R_i \in \mathcal{V}$ is 2-basic (see definition A.4) for the 2-fixable family $\mathcal{V} = \{\prod_{i=1}^{n} T_i \,|\, T_i \subseteq A_i\}$ if and only if $1 \leq |R_i| \leq 2$ $\forall\, 1 \leq i \leq n$.*

Proof. Suppose $Q = \prod_{i=1}^{n} R_i \in \mathcal{V}$ is 2-basic, that is, $Q = S_{\vec{\mathbf{x}}}$ where $\vec{\mathbf{x}} = (\mathbf{x}, \mathbf{y}) \in S^2$ with $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$. It follows then that $R_i = \{x_i, y_i\}$. (Indeed, for any $K = \prod_{i=1}^{n} T_i$ containing both $\mathbf{x}$ and $\mathbf{y}$ we must have $T_i \supset \{x_i, y_i\}$. On the other hand, the set $U = \prod_{i=1}^{n}\{x_i, y_i\}$ does contain both $\mathbf{x}$ and $\mathbf{y}$ and belongs to $\mathcal{V}$, so that we must have $Q = U$ and the desired conclusion follows.)

Conversely, suppose $Q = \prod_{i=1}^{n} R_i \in \mathcal{V}$ and $1 \leq |R_i| \leq 2$. Let $M = \{i \,|\, |R_i| = 2\}$. For $i \in M$ let $R_i = \{a_i, b_i\}$ (with $a_i \neq b_i$) and, for $i \notin M$, let $R_i = \{c_i\}$. Define $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ as follows: for $i \in M$ $x_i = a_i$ and $y_i = b_i$, while for $i \notin M$, $x_i = y_i = c_i$. It is easily seen now that $Q = S_{(\mathbf{x},\, \mathbf{y})}$ (the argument is virtually the same as the one given above for the converse). □

In view of proposition 4.1, Radcliffe's definition of "gene transmission" can be restated as follows:

Definition 4.1. We say that a given family of binary transformations Γ on the search space $S = \prod_{i=1}^{n} A_i$ transmits genes if the 2-basic elements of $\mathcal{V} = \{\prod_{i=1}^{n} T_i \,|\, T_i \subseteq A_i\}$ are invariant under Γ (meaning that $\forall$ 2-basic $Q \in \mathcal{V}$ and $\forall\, T \in \Gamma$ we have $T(Q^2) \subseteq Q$).

Proposition 4.2. *A given family of binary transformations Γ on the search space $S = \prod_{i=1}^{n} A_i$ transmits genes if and only if $\Gamma \subseteq \widetilde{\mathcal{F}_M}$ (See definition A.7 and discussion in section 2 for the meaning of $\widetilde{\mathcal{F}_M}$. Here $\widetilde{\mathcal{F}_M}$ plays the role of Γ.) if and only if $\Lambda_\Gamma \supseteq \mathcal{V} = \{\prod_{i=1}^{n} T_i \,|\, T_i \subseteq A_i\}$.*

Proof. Indeed, in the language of the current paper, for a given family of transformations to transmit genes simply means to leave invariant all of the 2-basic members of the 2-fixable family of subsets $\mathcal{V}$ (see definition 4.1). According to remark A.3, this implies that $\Lambda_\Gamma \supseteq \mathcal{V}$. Proposition 3.1 together with proposition A.7 tell us that the family of masked crossovers is composition closed. Then from the discussion following definition A.8 it follows that $\Gamma \subseteq \widetilde{\mathcal{F}_M}$. On the other hand, if $\Gamma \subseteq \widetilde{\mathcal{F}_M}$, it follows immediately from part 3 of proposition A.5 combined with corollary 3.4 and theorem A.9 that $\Lambda_\Gamma \supseteq \mathcal{V}$ so that, in particular, Γ transmits genes. □

It turns out that the Random Respectful Recombination Operator introduced in (Radcliffe, 1994) and, incidentally, the so called "gene pool" recombination method (see (Syswerda, 1993), (Mühleinbeim and Mahnig, 2001) and (Wright, Rowe, Poli and Stephens, 2002)) are closely related to the following family of transformations:

Definition 4.2. For every given point $\mathbf{u} = (u_1, u_2, \ldots, u_n) \in S$ define a Holland transformation $T_{\mathbf{u}} : S^2 \to S$ as follows: for every $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in S$

$$T_{\mathbf{u}}(\mathbf{a}, \mathbf{b}) = (x_1, x_2, \ldots, x_n)$$

where

$$x_i = \begin{cases} a_i \text{ if } a_i = b_i \\ u_i \text{ otherwise} \end{cases}$$

In other words, if the i^{th} coordinates of $\mathbf{a}$ and $\mathbf{b}$ coincide, then the i^{th} coordinate of $T_{\mathbf{u}}(\mathbf{a}, \mathbf{b})$ also coincides with them. If, on the other hand, the i^{th} coordinates of $\mathbf{a}$ and $\mathbf{b}$ differ, then the i^{th} coordinate of $T_{\mathbf{u}}(\mathbf{a}, \mathbf{b})$ is that of $\mathbf{u}$, namely, u_i. Denote by $\mathcal{H} = \{T_{\mathbf{u}} \mid \mathbf{u} \in S\}$ the family of all Holland transformations.

For example, if $n = 5$, $A_i = A = \{1, 2, 3, 4\}$ for $1 \leq i \leq 5$, $\mathbf{u} = (1, 3, 4, 1, 2)$, then if, say, $\mathbf{a} = (3, 4, 2, 1, 3)$ and $\mathbf{b} = (3, 1, 2, 3, 4)$ we have $T_{\mathbf{u}}(\mathbf{a}, \mathbf{b}) = (3, 3, 2, 1, 2)$. Indeed, the alleles in the first and the third positions of $\mathbf{a}$ and $\mathbf{b}$ coincide: they are 3 and 2, respectively. On the other hand, the alleles in the second, fourth and fifth positions of $\mathbf{a}$ and $\mathbf{b}$ don't coincide, so that they are replaced with the corresponding alleles of $\mathbf{u}$, namely $3, 1$, and 2 respectively.

In a sense, the family of Holland transformations is more "destructive" than the family of crossover transformations: it only preserves the alleles which both individuals have in common, and replaces any other gene with any suitable randomly selected gene. We'll show that the family of Holland transformations is the largest for all practical purposes family of transformations which leaves the family of subsets of S determined by Holland schemata together with the empty set, call it $\mathcal{J}$ (i.e., $\mathcal{J} = \{\prod_{i=1}^{n} T_i \,|\, T_i = \{a_i\}$ for some $a_i \in A_i$ or $T_i = A_i\} \cup \{\emptyset\}$), invariant.

The following theorem, which is somewhat analogous to theorem 3.2, gives us a glance into the dynamics of the subsets of the search space under the family of Holland transformations. In order to state the theorem, let's recall the definition A.5 from the appendix. For the reader's convenience we restate the definition of the dynastic span under the family of Holland transformations below:

Definition 4.3. Given a subset $Q \subseteq S$, let $\mathcal{H}(Q) = \bigcup_{\mathbf{u} \in S} T_{\mathbf{u}}(Q^2)$. Just as before, we shall also write $\mathcal{H}^2(Q) = \mathcal{H}(\mathcal{H}(Q))$ and, in general,

$$\mathcal{H}^k(Q) = \underbrace{\mathcal{H}(\mathcal{H}(\mathcal{H}(\ldots \mathcal{H}}_{k \text{ times}}(Q) \ldots))$$

for $k \geq 1$ and $\mathcal{H}^0(Q) = Q$. Let the full dynastic span of Q under $\mathcal{H}$ be the set $\mathcal{H}^\infty(Q) = \bigcup_{k=1}^{\infty} \mathcal{H}^k(Q)$.

As before, let $p_i : S = \prod_{j=1}^{n} A_j \to A_i$ denote the i^{th} coordinate projection. More explicitly, $\forall\ \mathbf{x} = (x_1, x_2, \ldots, x_n) \in S,\ p_i(\mathbf{x}) = x_i$. We are finally ready to state the theorem:

Theorem 4.3. *For any nonempty subset $Q \subseteq S$ we have*

$$Q = \mathcal{H}^0(Q) \subsetneq \mathcal{H}(Q) \subsetneq \mathcal{H}^2(Q) \subsetneq \ldots \subsetneq \mathcal{H}^l(Q) = \prod_{i=1}^{n} T_i$$

$$T_i = \begin{cases} \{a_i\} & \text{if } p_i(Q) = \{a_i\} \\ A_i & \text{otherwise} \end{cases}$$

where l is the smallest nonnegative integer such that $\mathcal{H}^l(Q) = \mathcal{H}^{l+1}(Q)$. [3] *Moreover, the following is true:*

1. *Let k be the number of T_i's which are equal to the corresponding A_i's. In other words, $k = |\{i \mid T_i = A_i\}|$. Then, when $k > 1$ we have $l \leq \lceil \log_2 k \rceil$. If $k = 1$, then $l \leq 1$. (In the terminology of (Michalewicz, 1996), $k = n - o(D)$ where $o(D)$ is the order (number of fixed positions) of the schema D which determines the set $\prod_{i=1}^n T_i$.)*

2. *The upper bound on l cannot be improved: more precisely, for every n there exists a search space $S = \prod_{i=1}^n A_i$ such that for every element $\prod_{i=1}^n T_i \in \mathcal{J}$ there exists subset $Q \subseteq \prod_{i=1}^n T_i$ for which $l = \lceil \log_2 k \rceil$ where k, just as above, is the number of the asterisk (nonfixed) positions in the schema determining $\prod_{i=1}^n T_i$.*

Proof. The proof is similar to the proof of theorem 3.2. First, notice that every schema is invariant under the family of the Holland transformations. Indeed, when some two individuals, say $\mathbf{a}$ and $\mathbf{b}$, belong to a subset H, determined by a given schema $\vec{s} \in \mathcal{A} = \prod_{i=1}^n A_i \bigcup \{*\}$ (see the discussion in the introduction) whose non-asterisk coordinates (fixed positions) are $a_{i_1}, a_{i_2}, \ldots, a_{i_j}$, the $i_1^{\text{th}}, i_2^{\text{th}}, \ldots, i_j^{\text{th}}$ coordinates of both, $\mathbf{a}$ and $\mathbf{b}$ are $a_{i_1}, a_{i_2}, \ldots, a_{i_j}$ respectively, so that for any Holland transformation $T_{\mathbf{u}} \in \mathcal{H}$, the $i_1^{\text{th}}, i_2^{\text{th}}, \ldots, i_j^{\text{th}}$ coordinates of $T_{\mathbf{u}}(\mathbf{a}, \mathbf{b})$ are also $a_{i_1}, a_{i_2}, \ldots, a_{i_j}$, which, in turn, means that $T_{\mathbf{u}}(\mathbf{a}, \mathbf{b}) \in H$ as well.) But then, since the set $\prod_{i=1}^n T_i$, where T_i's are as in the statement of the theorem, is in $\mathcal{J}$ and contains Q, from proposition A.4 we get the following chain of inclusions:

$$Q = \mathcal{H}^0(Q) \subsetneq \mathcal{H}(Q) \subsetneq \mathcal{H}^2(Q) \subsetneq \ldots \subsetneq \mathcal{H}^l(Q) \subseteq \prod_{i=1}^n T_i$$

The fact that $\mathcal{H}^l(Q) = \prod_{i=1}^n T_i$ together with part 1 of the theorem follows from the following Lemma which is similar to Lemma 3.3. The statement of the Lemma uses the notation described in definition 3.4.

Lemma 4.4. *Let $J = \{i \mid p_i(Q) = \{a_i\}\}$ (in other words, J is the set of fixed positions of the smallest schema containing Q). Then for every $M \subseteq \{1, 2, \ldots, n\} - J$ with $|M| \leq 2^q, q \geq 1$ we have $p_M(\mathcal{H}^q(Q)) \supseteq \prod_{i \in M} A_i$.*

Proof. The proof is also similar to the proof of Lemma 3.3. We proceed by induction on q: For $q = 1$ we have $|M| \leq 2$. When $|M| = 1$, $p_M = p_i$ for some $i \in \{1, 2, \ldots, n\} - J$, and we want to show that given any $x_i \in A_i$ there exists an element $\mathbf{c} = (c_1, c_2, \ldots, c_n) \in \mathcal{H}(Q)$ with $c_i = x_i$. This can be seen as follows: Since $i \in \{1, 2, \ldots, n\} - J$, there must exist elements $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in S$ such that $a_i \neq b_i$. Now let $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ where

$$u_j = \begin{cases} x_j & \text{if } j = i \\ a_j & \text{otherwise} \end{cases}$$

and notice that, by definition 4.2, we have $p_i(T_{\mathbf{u}}(\mathbf{a}, \mathbf{b})) = x_i$. On the other hand, $T_{\mathbf{u}}(\mathbf{a}, \mathbf{b}) \in \mathcal{H}(Q)$, so that the statement is proved for $q = 1$ and $|M| = 1$. When

[3] In particular, the theorem also asserts that $H^l(Q) = \prod_{i=1}^n T_i$. Of course, if $Q = \prod_{i=1}^n T_i$ then $l = 0$ and the chain terminates right away.

$|M| = 2$, $M = \{i_1, i_2\}$ for some $i_1 \neq i_2$. Fix any element $\vec{x} = (x_{i_1}, x_{i_2}) \in A_{i_1} \times A_{i_2}$. Let $\mathbf{u} = (u_1, u_2, \ldots, u_n) \in S$ be defined as follows:

$$u_j = \begin{cases} x_j & \text{if } j = i_1 \text{ or } j = i_2 \\ a_j & \text{otherwise} \end{cases}$$

Since $i_1 \in M$, there exists an element $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in Q$ for which $a_{i_1} \neq b_{i_1}$. There are now exactly two possible cases: either $a_{i_2} \neq b_{i_2}$ or $a_{i_2} = b_{i_2}$

case 1: $a_{i_2} \neq b_{i_2}$. In this case notice that, according to definition 4.2, $p_M(T_{\mathbf{u}}(\mathbf{a}, \mathbf{b})) = \vec{x}$.

case 2: $a_{i_2} = b_{i_2}$. In this case there exists an element $\mathbf{c} = (c_1, c_2, \ldots, c_n) \in Q$ such that $c_{i_2} \neq a_{i_2} = b_{i_2}$. (otherwise $i_2 \notin M$) Notice that, no matter what, either $c_{i_1} \neq a_{i_1}$ or $c_{i_1} \neq b_{i_1}$ so that, according to definition 4.2, $p_M(T_{\mathbf{u}}(\mathbf{a}, \mathbf{c})) = \vec{x}$ or $p_M(T_{\mathbf{u}}(\mathbf{b}, \mathbf{c})) = \vec{x}$.

Thereby, in any case it follows that $\vec{x} \in \mathcal{H}(Q)$ which finishes the base case.

Now, suppose the statement holds for $t \geq 1$ and let $q = t + 1$. Let $M = \{i_1, i_2, \ldots, i_j\} \subseteq \{1, 2, \ldots, n\} - J$ with $j \leq 2^{t+1}$. Partition the set M into two subsets of size $\leq 2^{t+1}/2 = 2^t$: say, let $j_1 = \lceil j/2 \rceil$ and consider $M_1 = \{i_1, i_2, \ldots, i_{j_1}\}$ and $M_2 = \{i_{j_1+1}, \ldots, i_j\}$. Now, given any $\mathbf{x} = (x_{i_1}, x_{i_2}, \ldots, x_{i_n}) \in \prod_{i \in M} A_i$, by inductive hypothesis, there exist $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathcal{H}^t(Q)$ such that $a_{i_l} = x_{i_l}$ for $1 \leq l \leq j_1$ and $b_{i_l} = x_{i_l}$ for $j_1 + 1 \leq l \leq j$. Now let $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ where

$$u_i = \begin{cases} x_i & \text{if } i \in M \\ a_i & \text{otherwise} \end{cases}$$

and notice that for every $i \in M$ we have $p_i(T_{\mathbf{u}}(\mathbf{a}, \mathbf{b})) = x_i$. Indeed, for every $i \in M$ exactly two cases are possible: either $a_i = b_i$ or $a_i \neq b_i$

case 1: $a_i = b_i$. This case breaks down into two subcases: either $i \in M_1$ or $i \in M_2$. When $i \in M_1$, $b_i = a_i = x_i$, and, likewise, when $i \in M_2$, $a_i = b_i = x_i$ so that, in either of the two subcases, by definition 4.2 we obtain $p_i(T_{\mathbf{u}}(\mathbf{a}, \mathbf{b})) = x_i$.

case 2: $a_i \neq b_i$. In this case, by definition 4.2, we obtain right away that $p_i(T_{\mathbf{u}}(\mathbf{a}, \mathbf{b})) = u_i = x_i$.

Thereby, we deduce that $p_M(T_{\mathbf{u}}(\mathbf{a}, \mathbf{b})) = \mathbf{x}$, and $T_{\mathbf{u}}(\mathbf{a}, \mathbf{b}) \in \mathcal{H}^{t+1}(Q)$. Since the choice of $\mathbf{x} \in \prod_{i \in M} A_i$ is arbitrary, we deduce that $p_M(\mathcal{H}^{t+1}(Q)) \supseteq \prod_{i \in M} A_i$. The desired conclusion now follows by the principle of induction. □

The fact that $\mathcal{H}^l(Q) = \prod_{i=1}^n T_i$ and part 1 of the theorem now follow immediately from Lemma 4.4 applied to $q = \lceil \log_2 k \rceil$ so that M is the set of all nonfixed positions of the schema determining the set $\prod_{i=1}^n T_i$.

It still remains to prove part 2 of the theorem. So let's fix n and let $A_i = A = \{0, 1\}$. Suppose we are given a set $\prod_{i=1}^n T_i \in \mathcal{J}$. Let $i_1 \leq i_2 \leq \ldots \leq i_{n-k}$ be the indices of the fixed positions of the schema determining the set $\prod_{i=1}^n T_i$. Say these positions are $a_{i_1}, a_{i_2}, \ldots, a_{i_{n-k}}$ respectively. This means that

$$T_i = \begin{cases} \{a_{i_j}\} & \text{if } i = i_j \text{ for some } 1 \leq j \leq n - k \\ A_i & \text{otherwise} \end{cases}$$

Now, let $Q \subseteq \prod_{i=1}^n T_i$ denote the k-element subset of S, which is the set of rows of a k

by n matrix P pictured below:

$$\begin{pmatrix} 1 & 0 & \dots & a_{i_1} & 0 & 0 & \dots & a_{i_{n-k}} & 0\dots & 0 \\ 0 & 1 & \dots & a_{i_1} & 0 & 0 & \dots & a_{i_{n-k}} & 0\dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_{i_1} & 1 & 0 & \dots & a_{i_{n-k}} & 0\dots & 0 \\ 0 & 0 & \dots & a_{i_1} & 0 & 1 & \dots & a_{i_{n-k}} & 0\dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_{i_1} & 0 & 0 & \dots & a_{i_{n-k}} & 1\dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_{i_1} & 0 & 0 & \dots & a_{i_{n-k}} & 0\dots & 1 \end{pmatrix}$$

Every i_j^{th} column of P consists entirely of the entries a_{i_j} (this must be so because $Q \subseteq \prod_{i=1}^{n} T_i$) for $1 \leq j \leq n-k$. Upon deleting the columns i_j for $1 \leq j \leq n-k$ from P one should obtain the k by k identity matrix.

An alternative, explicit definition of the matrix P can be obtained in the following manner: first define a function $\gamma : \{1, 2, \dots, n\} \to \{0, 1, 2, \dots, n-k\}$ as follows: $\gamma(q) = |Z_q|$ where $Z_q = \{i_j \,|\, 1 \leq j \leq n-k \text{ and } i_j < q\}$ ($\gamma(q)$ is simply the number of i_j's which are smaller than q). Now P is the matrix with entries $p_{i\,l}$ for $1 \leq i \leq k$ and $1 \leq l \leq n$ where

$$p_{i\,l} = \begin{cases} a_{i_j} & \text{if } l = i_j \text{ for some } 1 \leq j \leq n-k \\ 1 & \text{if } i = l - \gamma(l) \text{ and } l \neq i_j \ \forall\, 1 \leq j \leq n-k \\ 0 & \text{otherwise} \end{cases}$$

Notice that any two distinct elements of Q differ in exactly two positions. The key to proving that Q has the property described in part 2 of the theorem is to generalize this observation by induction:

Claim: *Let Q be the subset of S described above. Let $M = \{1, 2, \dots, n\} - \{i_j \mid 1 \leq j \leq n\}$ (Notice that, by definition $|M| = k$). Then for any $q \leq \log_2 k$, given any element $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathcal{H}^q(Q)$ there exists a subset of indices $L_{\mathbf{x}} \subseteq M$ of size $|L_{\mathbf{x}}| = k - 2^q$ such that for every $i \in L$ $x_i = 0$*

Proof. We proceed by induction on q. For $q = 0$ the statement is clear from the definition of $Q = \mathcal{H}^0(Q)$. (Every element of Q, that is, every row of P has exactly one entry in a column indexed by an element of M equal to 1, and the other $k - 2^0 = k - 1$ entries of the row are 0's.) Now, suppose the statement holds for $q \leq j \geq 0$. Let $q = j + 1$. Fix an element $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathcal{H}^{j+1}(Q)$. Then, either $\mathbf{x} \in \mathcal{H}^j(Q)$ (in which case by inductive hypothesis there exists the desired subset $L_{\mathbf{x}}$ of size $k - 2^j \geq k - 2^{j+1}$) or, more of interest to us, $\mathbf{x} = T_{\mathbf{u}}(\mathbf{a}, \mathbf{b})$ for some $\mathbf{a}, \mathbf{b} \in \mathcal{H}^j(Q)$ and $\mathbf{u} \in S$. By inductive hypothesis, there are subsets $L_{\mathbf{a}}$ and $L_{\mathbf{b}} \subseteq M$ each of size $k - 2^j$ such that the coordinates of $\mathbf{a}$ and $\mathbf{b}$ indexed by elements of $L_{\mathbf{a}}$ and $L_{\mathbf{b}}$ respectively, are equal to 0. In particular, the set of indices in M of coordinates of $\mathbf{a}$ and $\mathbf{b}$ where $\mathbf{a}$ and $\mathbf{b}$ are not simultaneously equal to 0 must be a subset of $D_{\mathbf{x}} = M - (L_{\mathbf{a}} \cap L_{\mathbf{b}}) = (M - L_{\mathbf{a}}) \cup (M - L_{\mathbf{b}})$. But then the number of indices in $D_{\mathbf{x}}$ is $|D_{\mathbf{x}}| \leq |M - L_{\mathbf{a}}| + |M - L_{\mathbf{b}}| \leq 2^j + 2^j = 2^{j+1}$. The number of indices in the subset $L_{\mathbf{x}} = M - D_{\mathbf{x}} \subseteq M$ where the coordinates of $\mathbf{a}$ and $\mathbf{b}$ are simultaneously equal to 0, is, therefore, $k - |D_{\mathbf{x}}| \geq k - 2^{j+1}$. But according to definition 4.2 every coordinate of $\mathbf{x} = T_{\mathbf{u}}(\mathbf{a}, \mathbf{b})$ indexed by an element of $L_{\mathbf{x}}$ is equal 0. Thereby, the statement of the Claim holds for $q = j + 1$. By the principle of induction, the statement of the Claim is true. □

From the Claim above it follows immediately that, for the set Q constructed above, we must have $l = \lceil \log_2 k \rceil$. (For any $q < \lceil \log_2 k \rceil$ the element whose coordinates indexed by elements of M, are all equal to 1 is not in $\mathcal{H}^q(Q)$.) □

The following corollary follows immediately from theorem 4.3 in exactly the same manner as corollary 3.4 follows from theorem 3.2.

Corollary 4.5. *A subset $Q \subseteq S$ is invariant under the family of all Holland transformations if and only if $Q \in \mathcal{J}$ (which means that Q is determined by a schema or is empty).*

Proof. The argument is analogous to the proof of corollary 3.4: Given any nonempty subset $Q \subseteq S$, by proposition A.4 Q is invariant under the family of Holland transformations if and only if the degree of Q under the action of the family of Holland transformations (see proposition A.4) is 0 which, according to theorem 4.3, happens if and only if $Q = \prod_{i=1}^{n} T_i$ where

$$T_i = \begin{cases} \{a_i\} & \text{if } p_i(Q) = \{a_i\} \\ A_i & \text{otherwise} \end{cases}$$

which means that $Q \in \mathcal{J}$ □

The next theorem shows that the family of Holland transformations is composition closed, which, in turn, implies that it is large enough for all practical purposes in the sense of the discussion which follows definition A.8.

Theorem 4.6. *The family of Holland transformations, $\mathcal{H}$, is composition closed.*

Proof. According to Lemma A.6, it suffices to show that $\forall\, \mathbf{u}, \mathbf{v}, \mathbf{w} \in S\ \exists \mathbf{z} \in S$ such that $\forall\, \vec{\mathbf{x}} \in S^2$ we have $T_{\mathbf{z}}(\vec{\mathbf{x}}) = T_{\mathbf{w}}(T_{\mathbf{u}}(\vec{\mathbf{x}}), T_{\mathbf{v}}(\vec{\mathbf{x}}))$. Let $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ and $\mathbf{w} = (w_1, w_2, \ldots, w_n)$. We show that $\mathbf{z} = T_{\mathbf{w}}(\mathbf{u}, \mathbf{v})$ does the job for these $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$: Indeed, explicitly $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ where $z_i = \begin{cases} u_i & \text{if } u_i = v_i \\ w_i & \text{otherwise} \end{cases}$. Fix any $\vec{\mathbf{x}} = (\mathbf{a}, \mathbf{b}) \in S^2$. Now, simply notice that if the i^{th} coordinate of $\mathbf{a}$ is the same as that of $\mathbf{b}$, then, by definition 4.2, it is also the same as the i^{th} coordinate of $T_{\mathbf{u}}(\vec{\mathbf{x}})$ and the same as the i^{th} coordinate of $T_{\vec{\mathbf{v}}}(\mathbf{x})$, which, again by definition 4.2, implies that it is the same as the i^{th} coordinate of $T_{\mathbf{w}}(T_{\mathbf{u}}(\vec{\mathbf{x}}), T_{\mathbf{v}}(\vec{\mathbf{x}}))$. Now suppose the i^{th} coordinate of $\mathbf{a}$ is not the same as that of $\mathbf{b}$. Then, by definition 4.2, the i^{th} coordinate of $T_{\mathbf{u}}(\vec{\mathbf{x}})$ is u_i, and, similarly, the i^{th} coordinate of $T_{\mathbf{v}}(\vec{\mathbf{x}})$ is v_i. Then, if $u_i = v_i$, the i^{th} coordinate of $T_{\mathbf{w}}(T_{\mathbf{u}}(\vec{\mathbf{x}}), T_{\mathbf{v}}(\vec{\mathbf{x}}))$ is $u_i = v_i = z_i$. If, on the other hand, $u_i \neq v_i$, then the i^{th} coordinate of $T_{\mathbf{w}}(T_{\mathbf{u}}(\vec{\mathbf{x}}), T_{\mathbf{v}}(\vec{\mathbf{x}}))$ is $w_i = z_i$. Thereby, according to definition 4.2, $T_{\mathbf{z}}(\vec{\mathbf{x}}) = T_{\mathbf{w}}(T_{\mathbf{u}}(\vec{\mathbf{x}}), T_{\mathbf{v}}(\vec{\mathbf{x}}))$ and the desired conclusion follows. □

It is interesting to notice that the 2-fixable (see definition A.4 and theorem A.3) family of subsets determined by the Holland schemata together with the empty set, $\mathcal{J} = \{\prod_{i=1}^{n} T_i \,|\, T_i = \{a_i\}$ for some $a_i \in A_i$ or $T_i = A_i\} \cup \{\emptyset\}$, enjoys the following property:

Proposition 4.7. *Every nonempty $Q \in \mathcal{J}$ is 2-basic (see definition A.4). In other words, $\mathcal{J} = M_{\mathcal{J}} \cup \{\emptyset\}$ (see remark A.3 for the meaning of $M_{\mathcal{J}}$).*

Proof. Indeed, fix any nonempty $Q \in \mathcal{J}$, say $Q = \prod_{i=1}^{n} T_i$. Let $M = \{i \,|\, 1 \leq i \leq n\ T_i \neq \{a_i\}\ \forall\, a_i \in A_i\}$. Notice that $\forall\, i \in M$ we have $|A_i| \geq 2$ (if $|A_i| = 1$ then $T_i = A_i = \{a\}$ for the unique $a \in A_i$). For $i \in M$ let $b_i \neq c_i$ denote two distinct elements of A_i. Define

$\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n) \in S$ as follows: for $i \notin M$ let $x_i = y_i = a_i$ where $\{a_i\} = T_i$, and, for $i \in M$, let $x_i = b_i$ and $y_i = c_i$. It is easily seen now that Q is the smallest element of $\mathcal{J}$ containing $\mathbf{x}$ and $\mathbf{y}$. (Indeed, if $U = \prod_{i=1}^{n} R_i \in \mathcal{J}$ contains $\mathbf{x}$ and $\mathbf{y}$ then for $i \in M$, since $x_i = b_i \neq c_i = y_i$, we must have $|R_i| \geq 2$ which forces $R_i = A_i$. If, on the other hand, $i \notin M$ then, either $R_i = a_i$ or $R_i = A_i$. In either case, it follows that $R_i \supseteq T_i$, which implies that $U \supseteq Q$.) Since the choice of $Q \in \mathcal{J}$ was arbitrary, the desired conclusion follows. □

It may be worthwhile to notice that the property expressed in proposition 4.7 does not hold for the family of Antonisse's schemata, $\mathcal{V}$ (see proposition 4.1) unless the search space is binary, in which case the family of subsets determined by Antonisse's schemata coincides with the one determined by the Holland schemata.

One can generalize a genetic algorithm in a natural way using the family of Holland transformations as follows. Perform the first two steps, namely selection and pairing in exactly the same way as it is usually done in a genetic algorithm. Now, for every selected pair $(\mathbf{a}, \mathbf{b})$ select two probability distributions $D_l^{(\mathbf{a},\mathbf{b})}$ and $D_r^{(\mathbf{a},\mathbf{b})}$ on the search space S. (This is equivalent to selecting the probability distributions on the set of Holland transformations, $\mathcal{H}$.) We shall discuss some possible efficient ways of selecting the probability distributions later. Once the probability distributions are chosen, we randomly select elements $\mathbf{u}_{(\mathbf{a},\mathbf{b})}$ and $\mathbf{v}_{(\mathbf{a},\mathbf{b})} \in S$ with respect to the distributions $D_l^{(\mathbf{a},\mathbf{b})}$ and $D_r^{(\mathbf{a},\mathbf{b})}$ respectively, and replace each pair $(\mathbf{a}, \mathbf{b})$ with the pair $(T_{\mathbf{u}_{(\mathbf{a},\mathbf{b})}}(\mathbf{a}, \mathbf{b}),\, T_{\mathbf{v}_{(\mathbf{a},\mathbf{b})}}(\mathbf{a}, \mathbf{b}))$.

Choosing the probability distributions $D_l^{(\mathbf{a},\mathbf{b})}$ and $D_r^{(\mathbf{a},\mathbf{b})}$ can be done very efficiently by exploiting the fact that $S = \prod_{i=1}^{n} A_i$ so that one can choose probability distributions $D_i^{\mathbf{a}}$ and $D_i^{\mathbf{b}}$ on every A_i in parallel, and let $D_l^{(\mathbf{a},\mathbf{b})}$ and $D_r^{(\mathbf{a},\mathbf{b})}$ be the product measures with components $D_i^{\mathbf{a}}$'s and $D_i^{\mathbf{b}}$'s respectively. The simplest choice would be to make each $D_i^{\mathbf{a}}$ and each $D_i^{\mathbf{b}}$ a uniform distribution on A_i which results in a uniform probability distribution on $S = \prod_{i=1}^{n} A_i$ (by a uniform distribution on a finite probability space we mean a distribution with respect to which all the elements are equally likely to be chosen). Intuitively speaking, such a choice of a distribution results in preserving only the alleles which are common in both parents and replacing the rest of the alleles with randomly selected allowable alleles. This method was first introduced in (Radcliffe, 1994) under the name of "Random Respectful Recombination".

A more interesting choice can be made by letting each $D_i^{\mathbf{a}}$ and each $D_i^{\mathbf{b}}$ be the probability distribution on A_i such that the probability for a given allele (element of A_i) is given by its frequency within the current population. Another natural choice is to make each $D_i^{\mathbf{a}}$ and each $D_i^{\mathbf{b}}$ be the probability distributions which select a_i (the i^{th} coordinate of $\mathbf{a}$) with probability $\frac{f(\mathbf{a})}{f(\mathbf{a})+f(\mathbf{b})}$ and b_i (the i^{th} coordinate of $\mathbf{b}$) with probability $\frac{f(\mathbf{b})}{f(\mathbf{a})+f(\mathbf{b})}$. These methods may be viewed as possible modifications of what is sometimes referred to as "gene pool" recombination (see (Syswerda, 1993), (Mühleinbeim and Mahnig, 2001) and (Wright, Rowe, Poli and Stephens, 2002)) with respect to the family of Holland transformations.

Another interesting choice can be made as follows. For every pair $(\mathbf{a}, \mathbf{b}) \in S^2$, let $D_i^{\mathbf{a}} = D_i^{\mathbf{b}}$ be the following probability distribution on A_i: $D_i(\{a_i\}) = D_i(\{b_i\}) = 1/2$ and $D_i(c_i) = 0$ for any other $c_i \in A_i$. (Here, as usual, a_i and b_i are the i^{th} coordinates of $\mathbf{a}$ and $\mathbf{b}$ respectively.) Such a choice basically corresponds to applying the usual, randomly selected masked crossover operators to the pair $(\mathbf{a}, \mathbf{b})$.

In general, one could select the probability distribution $D_r^{(\mathbf{a},\mathbf{b})}$ depending on the element randomly selected according to the distribution $D_l^{(\mathbf{a},\mathbf{b})}$. In particular, one can make a deterministic choice. With such a setup it is easy to obtain both versions of the classical genetic algorithm: the one which uses the masked crossover operators and the one which uses only the one-point crossover operators. Indeed, let $D_l^{(\mathbf{a},\mathbf{b})}$ assign an equal probability to every element of the set $\mathcal{F}_M(\{\mathbf{a},\mathbf{b}\})$ (recall that $\mathcal{F}_M(\{\mathbf{a},\mathbf{b}\})$ is simply the set of all images of $(\mathbf{a},\mathbf{b})$ under the family of masked crossover transformations) and the probability 0 to every element of $\mathbf{S}$ which is not in $\mathcal{F}_M(\{\mathbf{a},\mathbf{b}\})$. This way one obtains the classical genetic algorithm with masked crossovers. To obtain the classical version with one-point crossovers, simply replace "$\mathcal{F}_M(\mathbf{a},\mathbf{b})$" with "$\mathcal{F}_1(\mathbf{a},\mathbf{b})$" everywhere in the previous sentence. Now, once we have chosen the element $\mathbf{u}_{(\mathbf{a},\mathbf{b})} \in S$ with respect to $D_l^{(\mathbf{a},\mathbf{b})}$, simply define the deterministic probability distribution, $D_r^{(\mathbf{a},\mathbf{b})}$ on S which assigns probability 1 to the element $\mathbf{v} \in S$ whose i^{th} coordinate is the i^{th} coordinate of $\mathbf{a}$ if the i^{th} coordinate of $\mathbf{u}_{(\mathbf{a},\mathbf{b})}$ is the i^{th} coordinate of $\mathbf{b}$, and vice versa: the i^{th} coordinate of $\mathbf{v}$ is the i^{th} coordinate of $\mathbf{b}$ if the i^{th} coordinate of $\mathbf{u}_{(\mathbf{a},\mathbf{b})}$ is the i^{th} coordinate of $\mathbf{a}$. (Since $\mathbf{u}_{(\mathbf{a},\mathbf{b})} \in \mathcal{F}_M(\{\mathbf{a},\mathbf{b}\})$ with probability 1, for every $1 \le i \le n$ the i^{th} coordinate of $\mathbf{u}_{(\mathbf{a},\mathbf{b})}$ is either the i^{th} coordinate of $\mathbf{a}$ or the i^{th} coordinate of $\mathbf{b}$ with probability 1.)

5 Conclusions and Future Work

In the current paper the following contributions have been made:

1. It was shown that the full dynastic span of a given subset Q (as defined by Radcliffe: see definitions 3.3 and A.5) of the search space $S = \prod_{i=1}^{n} A_i$ under either one of the traditional families of crossover transformations, $\mathcal{F}_1$ or $\mathcal{F}_M$, is obtained after $\lceil \log_2 n \rceil$ iterations, meaning that $\mathcal{F}^{\lceil \log_2 n \rceil}(Q) = \mathcal{F}^{\infty}(Q)\ \forall\, Q \subseteq S$ (see definition 3.3 and theorem 3.2). The result is independent of the cardinality of the sets involved (see remark 3.4).

2. It has been shown that the family of masked crossovers is a very natural extension (a composition closure: see definition A.6) of the family of one-point crossovers, which, in turn, implies that the family of Antonisse's schemata (the family $\{\prod_{i=1}^{n} T_i \,|\, T_i \subseteq A_i\}$) is exactly the family of all the subsets invariant under $\mathcal{F}_1$ or $\mathcal{F}_M$ (see corollary 3.4). Moreover, it also shows that the family of masked crossovers, $\mathcal{F}_M$, is, for all practical purposes, the largest family of transformations whose corresponding family of invariant subsets, $\Lambda_{\mathcal{F}_M}$, is the family of Antonisse's schemata, $\{\prod_{i=1}^{n} T_i \,|\, T_i \subseteq A_i\}$ (see remark 3.7).

3. It has been shown that Radcliffe's notion of "gene transmission" is a particular case of the notion of invariance introduced in the current paper. (This notion of invariance is simply a generalization of Radcliffe's notion of "respect".) This is described in detail in the discussion at the beginning of section 4.

4. The family of Holland transformations has been introduced. This family of transformations is convenient for modeling the Random Respectful Recombination Operator introduced in (Radcliffe, 1994). The discussion at the end of section 4 demonstrates how one can choose various probability distributions to obtain some variations of the usual GA. In particular, it is possible to use the family of Holland transformations to obtain the so-called "gene pool" recombination method which has been explored in some recent publications (see, for instance, (Syswerda, 1993), (Mühleinbeim and Mahnig, 2001) and (Wright, Rowe, Poli and Stephens, 2002)).

5. It is shown that the family of Holland transformations is, for all practical purposes, the largest family of transformations whose corresponding family of invariant

subsets, $\Lambda_{\mathcal{H}}$ (see the description of the notation following the introductory discussion in Appendix A), is precisely the family of Holland schemata.

6. It is shown that the full dynastic span of a given subset Q of the search space $S = \prod_{i=1}^{n} A_i$ under the family of Holland transformations is obtained after $\lceil \log_2 n \rceil$ iterations. In other words, $\mathcal{H}^{\lceil \log_2 n \rceil}(Q) = \mathcal{H}^{\infty}(Q)$ (see definition 4.3 and theorem 4.3). The result is very much analogous to the corresponding result for the families of crossover transformations (see theorem 3.2 and the remarks following theorem 3.2).

7. The results in Appendix A (see theorem A.9) establish an important correspondence between the poset of the m-fixable subsets of the search space and the poset of the families of m-ary transformations of the search space. In Category Theory such a correspondence is known as a Galois Connection (see (MacLane, 1971)). This correspondence provides a way to extend Radcliffe's notion of "genetic representation function" (see page 10 of (Radcliffe, 1994)) to compare various evolutionary computation techniques via their representations. A few theorems illustrating this idea will be proved in some forthcoming paper.

6 Acknowledgements

I want to thank Professors John Holland and Rick Riolo for the helpful discussions and for the encouragement I've received from them to write this paper. I also want to thank my thesis advisor, Professor Andreas Blass for the numerous helpful advisor meetings which have stimulated many ideas for this and for my future work. I am also very grateful to my fellow graduate student of mathematics, Ronald Walker for some very helpful discussions, and the University of Michigan Complex Systems Group for suggestions regarding the organization of this paper. Finally, I want to thank the anonymous referees for the numerous important remarks, suggestions and corrections which improved this paper.

A Appendix: General Mathematical Framework

The purpose of this section is to set up an abstract mathematical framework and to prove a few general facts which are exploited throughout the paper when we specialize them to study the behavior of genetic algorithms. The author's impression is that the reader should become familiar with the notation and the results of this section (skipping the proofs is perfectly fine) in order to understand the motivation and to appreciate the results of this paper.

Throughout this section, we shall fix the following notation:

Ω denotes an arbitrary set.

$\mathit{\Gamma}$ and Θ denote some nonempty families of transformations from Ω^m to Ω for a fixed $m \geq 1$. (i. e. $\mathit{\Gamma} \subseteq \{T \mid T : \Omega^m \to \Omega\}$ and so is Θ.)

$\mathit{\Lambda}_{\mathit{\Gamma}} = \{S \mid S \subseteq \Omega,\ T(S^m) \subseteq S\ \forall\ T \in \mathit{\Gamma}\}$ denotes the family of invariant subsets of Ω under the action of $\mathit{\Gamma}$. (This is simply a restatement of definition 2.)

For any set X, $\mathcal{P}(X)$ denotes the power set of X, that is, the set of all subsets of X.

To alleviate the level of abstraction, a simple example will be used throughout the discussion.

Example A.1. Let $\Omega = \prod_{i=1}^{3}\{0,\, 1\} =$

$$= \{(1,\, 0,\, 0),\, (0,\, 1,\, 0),\, (0,\, 0,\, 1),\, (1,\, 1,\, 0),\, (1,\, 0,\, 1),\, (0,\, 1,\, 1), (1,\, 1,\, 1),\, (0,\, 0,\, 0)\},$$

$m = 2$ and $\Gamma = \{L_i : \Omega^2 \to \Omega \,|\, 0 \le i \le 3\}$ where $L_i(\mathbf{a}, \mathbf{b}) = (a_1, a_2, \ldots, a_i, b_{i+1}, \ldots, b_n)$ for all $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3) \in \Omega$. (Notice that L_0 is the second coordinate projection, L_3 is the first coordinate projection, while L_1 and L_2 are the one-point crossover transformations) For instance, $L_1((1,\,0,\,0),(0,\,1,\,1)) = (1,\,1,\,1)$ and $L_2((1,\,0,\,0),(0,\,1,\,1)) = (1,\,0,\,1)$. Notice that for this example Λ_Γ is precisely the family $\mathcal{J} = \{\prod_{i=1}^3 T_i \,|\, T_i = \{a_i\}$ for some $a_i \in A_i$ or $T_i = A_i\} \cup \{\emptyset\}$ of all subsets determined by Holland schemata together with the empty set (see Corollary 3.5).

Definition A.1. We say that a subfamily, say $\mathcal{E} \subseteq \mathcal{P}(X)$, is a directed family iff given any two sets $A, B \in \mathcal{E}$, there exists another set $C \in \mathcal{E}$ with $A \subseteq C$ and $B \subseteq C$.

Definition A.2. We say that a subfamily $\mathcal{U} \subseteq \mathcal{P}(X)$ is closed under directed unions iff for any directed subfamily $\mathcal{E} \subseteq \mathcal{U}$ we have $\bigcup_{A \in \mathcal{E}} A \in \mathcal{U}$.

Proposition A.1. *For any family of transformations Γ, the family $\Lambda_\Gamma = \{S \mid S \subseteq \Omega,\ T(S^m) \subseteq S\ \forall\ T \in \Gamma\}$ is closed under arbitrary intersections and under directed unions. Also $\emptyset$ and Ω are in Λ_Γ.*

Proof. Indeed, for any indexed family $\{S_i\}_{i \in I} \subseteq \Lambda_\Gamma$ and for any $T \in \Gamma$ we have $T((\bigcap_{i \in I} S_i)^m) = T(\bigcap_{i \in I} S_i^m) \subseteq \bigcap_{i \in I} T(S_i^m) \subseteq \bigcap_{i \in I} S_i$. The first inclusion is a general fact that the image of the intersection is always contained in the intersection of the images, while the second inclusion is immediate from the assumption that every $S_i \in \Lambda_\Gamma$. Thereby, it follows that $\bigcap_{i \in I} S_i \in \Lambda_\Gamma$.

Now, suppose we are given any directed subfamily $\mathcal{E} \subseteq \Lambda_\Gamma$. We want to show that $\bigcup_{A \in \mathcal{E}} A \in \Lambda_\Gamma$, which means that for any $T \in \Gamma$ we have $T((\bigcup_{A \in \mathcal{E}} A)^m) \subseteq \bigcup_{A \in \mathcal{E}} A$. So, fix any point $\mathbf{x} = (x_1, \ldots, x_m) \in (\bigcup_{A \in \mathcal{E}} A)^m$. This means that $x_i \in A_i$ for some $A_i \in \mathcal{E}$. Since there are finitely many A_is, and the family $\mathcal{E}$ is directed, we conclude that there exists a set $B \in \mathcal{E} \subseteq \Lambda_\Gamma$ such that $A_i \subseteq B$ for all $1 \le i \le m$. But then we have $\mathbf{x} \in B^m$. So $T(\mathbf{x}) \in B \subseteq \bigcup_{A \in \mathcal{E}} A$. Thereby, we have shown that for every $\mathbf{x} \in (\bigcup_{A \in \mathcal{E}} A)^m$ we have $T(\mathbf{x}) \in \bigcup_{A \in \mathcal{E}} A$ which means that $T((\bigcup_{A \in \mathcal{E}} A)^m) \subseteq \bigcup_{A \in \mathcal{E}} A$ or, equivalently, $\bigcup_{A \in \mathcal{E}} A \in \Lambda_\Gamma$. The desired conclusion then follows.

Finally, the fact that $\emptyset$ and Ω are in Λ_Γ is completely trivial. □

Example A.2. Notice, indeed, that the intersection of any collection of sets represented by Holland schemata is either empty or a set represented by a Holland schema.

Remark A.1. Suppose a subset $\mathcal{A}$ of $\mathcal{P}$ is closed under arbitrary intersections and contains Ω. Then, given any $B \subseteq \Omega$, the set $S = \bigcap_{U \supseteq B,\ U \in \mathcal{A}} U$ is the smallest element of $\mathcal{A}$ containing B, in the sense that $B \subseteq S$, and for any $U \in \mathcal{A}$ such that $B \subseteq U$, we have $S \subseteq U$.

From Proposition A.1 together with remark A.1 it follows that for every subset of Ω, say K, there is a least element of Λ_Γ containing it. More precisely, there exists unique set $S \in \Lambda_\Gamma$ such that $K \subseteq S$ and for any $U \in \Lambda_\Gamma$ for which we also have $K \subseteq U$, it must be the case that $S \subseteq U$. This observation naturally leads us to the following definition:

Definition A.3. Fix a family of subsets $\mathcal{S} \subseteq \mathcal{P}(\Omega)$ which is closed under arbitrary intersections and contains Ω. For a given set $K \subseteq \Omega$, denote the smallest element of $\mathcal{S}$ containing K by $S_K^{\mathcal{S}}$. If $\mathcal{S} = \Lambda_\Gamma$ for some family of transformations Γ, we shall write S_K^Γ instead of $S_K^{\Lambda_\Gamma}$. When either $\mathcal{S}$ or the family of transformations Γ is clear from the context, we'll just write S_K.

Example A.3. Continuing with example A.1, when $\mathcal{S}$ is the family of subsets of Ω determined by Holland schemata, and $K = \{(0,\,1,\,0),\,(1,\,0,\,0)\}$ then the smallest invariant subset of Ω containing K is the subset

$$S_K = \{(0,\,1,\,0),\,(1,\,0,\,0),\,(1,\,1,\,0),\,(0,\,0,\,0)\}$$

which is determined by the schema $(*,\,*,\,0)$.

There is one more, slightly less obvious property which the family Λ_Γ has:

Proposition A.2. *For any $A \subseteq \Omega$ we have $A \in \Lambda_\Gamma \iff$ for every subset $K \subseteq A$ with $|K| \leq m$ we have $S_K \subseteq A$ (see definition A.3 for the meaning of S_K).*

Proof. The $\Rightarrow$ direction is clear from definition A.3. To prove the converse, fix $A \subseteq \Omega$ and suppose that for every subset $K \subseteq A$ with $|K| \leq m$ we have $S_K \subseteq A$. We want to show that $A \in \Lambda_\Gamma$, which means that for every $T \in \Gamma$ and for every $\mathbf{x} \in A^m$ we have $T(\mathbf{x}) \in A$. So fix $T \in \Gamma$ and $\mathbf{x} = (x_1, x_2, \ldots, x_m) \in A^m$. We then have $x_1, x_2, \ldots, x_m \in A$. Now consider the set $K = \{x \mid x = x_i \text{ for some } 1 \leq i \leq m\}$. In other words, K is the set of coordinates of $\mathbf{x}$. Notice that $\mathbf{x} \in K^m \subseteq (S_K)^m$. But then $T(\mathbf{x}) \in S_K \subseteq A$ by our assumption, since $|K| \leq m$. Thereby, we deduce that $A \in \Lambda_\Gamma$. □

Proposition A.2 says that a given subset $A \subseteq \Omega$ is in Λ_Γ if and only if A is a union of special elements of Λ_Γ, namely the sets S_K for $K \subseteq A$ and $|K| \leq m$. This motivates us to introduce the following definition:

Definition A.4. Fix a family $\mathcal{E} \subseteq \mathcal{P}(\Omega)$, such that $\mathcal{E}$ is closed under arbitrary intersections and contains $\emptyset$ and Ω. We say that a set $A \in \mathcal{E}$ is m-basic with base K, iff $A = S_K$ for some $K \subseteq \Omega$ and $|K| \leq m$. (see definition A.3) Given a point $\mathbf{x} = (x_1, x_2, \ldots x_m) \in \Omega^m$, we shall often write $S_{\mathbf{x}}$ to denote the m-basic set S_K with base $K = \{x \mid x = x_i \text{ for } 1 \leq i \leq m\}$. In other words, K is the set of coordinates of $\mathbf{x}$. If, in addition, for every subset $A \subseteq \Omega$ we have $A \in \mathcal{E}$ if and only if A is equal to the union of all m-basic subsets with base $K \subseteq A$, we say that $\mathcal{E}$ is an m-fixable family of subsets.

Example A.4. For instance, the set S_K of example A.3 can also be denoted by $S_{\vec{\mathbf{x}}}$ where $\vec{\mathbf{x}} = ((0,\,1,\,0),\,(1,\,0,\,0)) \in \Omega^2$.

Remark A.2. Notice that saying that a subset $A \subseteq \Omega$ is equal to the union of all m-basic subsets with base $K \subseteq A$ is equivalent to saying that $A = \bigcup_{\mathbf{x} \in A^m} S_{\mathbf{x}}$, which, in turn is equivalent to saying that for every point $\mathbf{x} \in A^m$ we have $S_{\mathbf{x}} \subseteq A$.

The notions of m-fixable family and m-basic subset may be clarified by the following example:

Example A.5. Let $\mathcal{G} = (G,\,*)$ denote an arbitrary semigroup. (notice that $*$, the semigroup multiplication, is, by definition, a binary operation on the underlying set G (i. e. a function $* : G^2 \to G$). Let $\Gamma = \{*\}$. Then the family Λ_Γ is simply the collection of all sub-semigroups of $\mathcal{G}$. According to proposition A.2 Λ_Γ is an m-fixable family of subsets. The 2-basic elements of Λ_Γ are simply the sub-semigroups which are generated by 2 elements of G. For a more specific example, let $\mathcal{G} = \{\mathbb{N},\,\times\}$ where $\times$ is the usual multiplication of natural numbers. Fix any set of natural numbers, say $K = \{n_1,\,n_2,\ldots,n_m,...\}$ and notice that the set $S_K = \{\prod_{j=1}^{q} n_{i_j} \mid i_j \leq i_{j+1},\;\; i_j \leq |K| \text{ and } n_{i_j} \in K\}$ is the smallest element of Λ_Γ containing K. (It is, of course, the same thing as the sub-semigroup generated by K.) In particular, the 2-basic elements of Λ_Γ are of the form $S_{a,\,b} = \{a^m b^n \mid m,\,n \in \mathbb{N}\}$.

Corollary 3.4 states that the family of subsets determined by the Antonisse schemata is the same as Λ_Γ where Γ is either the family of all masked crossover transformations or the family of all one-point crossover transformations. According to proposition A.2, the family of subsets determined by the Antonisse schemata is 2-fixable. Proposition 4.1 explicitly describes the family of 2-basic subsets for the 2-fixable family of subsets determined by the Antonisse schemata. Thanks to proposition 4.1, definition 4.1 can be used to introduce Radcliffe's notion of "gene transmission" (see (Radcliffe, 1994)) in terms of leaving a certain family of 2-basic subsets invariant.

In all of the examples given above, a 2-fixable family of subsets arises as Λ_{Γ} for some family of m-ary transformations Γ. This is not in vain as the following theorem shows. Hence the reason for the name m-fixable:

Theorem A.3. *Fix a family of subsets $\mathcal{E} \subseteq \mathcal{P}(\Omega)$. Then $\mathcal{E}$ is an m-fixable family if and only if $\mathcal{E} = \Lambda_{\Gamma}$ for some family of transformations $\Gamma \subseteq \{T \mid T : \Omega^m \to \Omega\}$.*

Proof. From proposition A.1, proposition A.2 and definition A.4 it follows immediately that if $\mathcal{E} = \Lambda_{\Gamma}$, then $\mathcal{E}$ is m-fixable. Now, suppose that $\mathcal{E}$ is m-fixable. Let $\Gamma = \{T \mid T : \Omega^m \to \Omega,\ T(E^m) \subseteq E\ \forall\ E \in \mathcal{E}\}$. In other words, Γ is the family of all transformations from Ω^m to Ω under which all elements of $\mathcal{E}$ are invariant. By definition of Γ, we immediately deduce that $\Lambda_{\Gamma} \supseteq \mathcal{E}$. It remains to show the reverse inclusion, namely that $\Lambda_{\Gamma} \subseteq \mathcal{E}$. To do so, fix an element $K \in \Lambda_{\Gamma}$. We want to show that $K \in \mathcal{E}$. Without loss of generality, assume that $K \neq \emptyset$ and $K \neq \Omega$ (by assumption, $\emptyset$ and Ω are in $\mathcal{E}$). Now, fix any point $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in K^m$ and $b \notin K$. Consider the function $G_{\mathbf{a}}^b : \Omega^m \to \Omega$ defined as

$$G_{\mathbf{a}}^b(\mathbf{x}) = \begin{cases} x_1 & \text{if } \mathbf{x} \neq \mathbf{a} \\ b & \text{otherwise} \end{cases}, \text{ where } \mathbf{x} = (x_1, \ldots, x_m).$$

Notice that $b \in G_{\mathbf{a}}^b(K^m)$ so that $G_{\mathbf{a}}^b(K^m) \nsubseteq K$. Since $K \in \Lambda_{\Gamma}$, this implies that $G_{\mathbf{a}}^b \notin \Gamma$. But then, by definition of Γ above, this means that there exists a set $Q \in \mathcal{E}$ such that $G_{\mathbf{a}}^b(Q^m) \nsubseteq Q$. Notice, however, that the only way this may happen, is when $\mathbf{a} \in Q^m$ and $b \notin Q$. Thereby, we deduce that given any $\mathbf{a} \in K^m$, for any $b \notin K$ there exists a set $Q \in \mathcal{E}$ such that $\mathbf{a} \in Q^m$, and $b \notin Q$. Since the m-basic set $S_{\mathbf{a}} \subseteq Q$ for every such Q, we deduce that $S_{\mathbf{a}} \subseteq K$. Thereby, we have shown that for any $\mathbf{a} \in K^m$, the m-basic set $S_{\mathbf{a}} \subseteq K$. Since $\mathcal{E}$ is assumed to be m-fixable, according to remark A.2 and definition A.4 this implies that $K \in \mathcal{E}$. We now deduce that $\mathcal{E} \supseteq \Lambda_{\Gamma}$, and this finishes the argument. □

Remark A.3. Given an m-fixable family of subsets of Ω, $\mathcal{E}$, denote by $\mathcal{M}_{\mathcal{E}}$ the family of all m-basic subsets from $\mathcal{E}$. Notice now that for any family of transformations Γ from Ω^m into Ω, we have $\Lambda_{\Gamma} \supset \mathcal{E}$ if and only if $\Lambda_{\Gamma} \supset \mathcal{M}_{\mathcal{E}}$. (Indeed, the "only if" part is immediate. For the "if" part, assume that $\Lambda_{\Gamma} \supset \mathcal{M}_{\mathcal{E}}$. Now fix any $K \in \mathcal{E}$. We want to show that $\forall T \in \Gamma$ we have $T(K^m) \subseteq K$. So, fix any $T \in \Gamma$. Now, according to remark A.2, $\forall \mathbf{x} \in K^m$, $S_{\mathbf{x}} \subseteq K$. But, by our assumption, $S_{\mathbf{x}} \in \mathcal{M}_{\mathcal{E}} \subseteq \Lambda_{\Gamma}$, so that $T(\mathbf{x}) \in T(S_{\mathbf{x}}^m) \subseteq S_{\mathbf{x}} \subseteq K$. Since the choice of $\mathbf{x} \in K^m$ is arbitrary, it follows that $T(K^m) \subseteq K$ so that $K \in \Lambda_{\Gamma}$. The desired conclusion that $\Lambda_{\Gamma} \supset \mathcal{M}_{\mathcal{E}}$ now follows.)

It is natural to consider the forward images of a given set K under the elements of Γ. This idea was introduced in (Radcliffe, 1994).

Definition A.5. Define the dynastic span of $K \subseteq \Omega$ under a given family Γ of transformations of Ω^m into Ω to be the set $\Gamma(K) = \bigcup_{T \in \Gamma} T(K^m)$. We shall also write $\Gamma^2(K) = \Gamma(\Gamma(K))$ and, in general, $\Gamma^k(K) = \underbrace{\Gamma(\Gamma(\Gamma(\ldots\Gamma}_{k \text{ times}}(K)\ldots))$ for $k \geq 1$, $\Gamma^0(K) = K$, and $\Gamma^\infty(K) = \bigcup_{k=0}^{\infty} \Gamma^k(K)$. We say that $\Gamma^\infty(K)$ is the full dynastic span of Q under Γ.

Example A.6. Continuing with example A.1, if $K = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ then $\Gamma(K) = K \cup \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$ and $\Gamma^2(K) = \Gamma(\Gamma(K)) = \Omega$.

Now, suppose our family Γ of transformations satisfies the property that $\Gamma(K) \supseteq K$ for all $K \subseteq \Omega$. Then, if we start with an arbitrary subset $K \subseteq \Omega$ and keep taking dynastic spans we'll get $K \subseteq \Gamma(K)$ and $K = \Gamma(K)$ iff $K \in \Lambda_\Gamma$. Continuing in this manner we obtain the following chain of inclusions: $K = \Gamma^0(K) \subsetneq \Gamma(K) \ldots \subsetneq \Gamma^k(K) \ldots \subseteq \Gamma^\infty(K)$. The inclusions are proper only until the chain stabilizes, if it ever does. Of course, if Ω is a finite set, then the chain does stabilize after finitely many steps. Moreover, one can easily see by induction that $\Gamma^k(K) \subseteq \Gamma(S_K) = S_K$. Thereby, we have $\Gamma^\infty(K) = \bigcup_{k=0}^\infty \Gamma^k(K) \subseteq S_K$. On the other hand, for every $x \in \Gamma^\infty(K) = \bigcup_{k=0}^\infty \Gamma^k(K)$, and for any $T \in \Gamma$, we have $x \in \Gamma^k(K)$ for some k, so that $T(x) \in \Gamma^{k+1}(K) \subseteq \bigcup_{k=0}^\infty \Gamma^k(K) = \Gamma^\infty(K)$. Thereby, $T(\Gamma^\infty(K)) \subseteq \Gamma^\infty(K)$, that is, $\Gamma^\infty(K) \in \Lambda$, and, hence, we also have $S_K \subseteq \Gamma^\infty(K)$ by minimality of S_K, We then finally deduce that $S_K = \Gamma^\infty(K)$. All of the above is summarized in the following

Proposition A.4. *Fix a subset $K \subseteq \Omega$. Then we have the following chain of inclusions: $K = \Gamma^0(K) \subsetneq \Gamma(K) \ldots \subsetneq \Gamma^k(K) = S_K$. Notice that k here denotes the smallest element of the poset $\{0\} \cup \mathbb{N} \cup \{\infty\}$ such that $\Gamma^k(K) = S_K$. We then define $d_{\Gamma(K)} = k$ to be the degree of K under the action of Γ. Moreover, $d_{\Gamma(K)} = 0$ iff $K \in \Lambda_\Gamma$, and $d_{\Gamma(K)} < \infty$ if Ω is a finite set.*

We now investigate the relationship among different subfamilies of transformations from Ω^m into Ω. The first immediate fact is the following:

Proposition A.5. *Let $\Gamma \subseteq \Theta$. Then for every $K \subseteq \Omega$ we have*

1. $\Gamma^k(K) \subseteq \Theta^k(K)$ *for all* $k \leq \infty$.
2. $S_K^\Gamma \subseteq S_K^\Theta$.
3. $\Lambda_\Theta \subseteq \Lambda_\Gamma$.

Proof of 1: First, we prove the statement for finite k by induction: Indeed, the statement is true for $k = 0$ since, by definition, $\Gamma^0(K) = \Theta^0(K) = K$. Now, suppose the statement is true for some $n \geq 0$. Then we have $\Gamma^{n+1}(K) = \Gamma(\Gamma^n(K)) \subseteq \Gamma(\Theta^n(K)) \subseteq \Theta(\Theta^n(K)) = \Theta^{n+1}(K)$. The desired conclusion then follows by the principle of induction for $k < \infty$. The statement for $k = \infty$ follows from the statement for k finite because then we have $\Gamma^\infty(K) = \bigcup_{k=0}^\infty \Gamma^k(K) \subseteq \bigcup_{k=0}^\infty \Theta^k(K) = \Theta^\infty(K)$. □

Proof of 2: A particular case of 1 with $k = \infty$. □

Proof of 3: By proposition A.4 we have $K \in \Lambda_\Theta$ iff $d_{\Gamma(K)} = 0$ iff $K = S_K^\Theta$. But then we obtain from part 2 that $K = S_K^\Theta \supseteq S_K^\Gamma \supseteq K$ so that all containments above become equalities and we get $S_K^\Gamma = K$ iff $K \in \Lambda_\Gamma$. □

Definition A.6. Given a family of functions $\Gamma \subseteq \{T \mid T : \Omega^m \to \Omega\}$ define the composition closure of Γ by a standard recursive procedure: First define $\Gamma^{(1)} = \Gamma$. Now, let $\Gamma^{(k+1)} = \Gamma^{(k)} \cup \{T \mid T(\mathbf{x}) = T_0(T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_m(\mathbf{x})),\ T_0 \in \Gamma,\ T_i \in \Gamma^{(k)}\ \forall\ 1 \leq i \leq m\}$. Finally, let the composition closure, $\bar{\Gamma} = \bigcup_{k=1}^\infty \Gamma^{(k)}$. We shall say that Γ is composition closed iff $\Gamma = \bar{\Gamma}$.

Example A.7. Continuing with example A.1, notice that $\Gamma^{(2)}$ is precisely the family of all masked crossover transformations (see formula 3 of section 3 for the definition of masked crossover). Indeed, for $M = \{1\}, \{1, 2\}$ or $\{1, 2, 3\}$ the L_Ms are simply the one-point crossovers which are the elements of $\Gamma \subseteq \Gamma^{(2)}$. For $M = \{1, 3\}$ we have

$$L_M((a_1, a_2, a_3), (b_1, b_2, b_3)) =$$

$$= L_2(L_1((a_1,\, a_2,\, a_3),\, (b_1,\, b_2,\, b_3)), L_3((a_1,\, a_2,\, a_3),\, (b_1,\, b_2,\, b_3)))$$

For $M = \{2\}$ we have

$$L_M((a_1,\, a_2,\, a_3),\, (b_1,\, b_2,\, b_3)) =$$

$$= L_1(L_0((a_1,\, a_2,\, a_3),\, (b_1,\, b_2,\, b_3)), L_2((a_1,\, a_2,\, a_3),\, (b_1,\, b_2,\, b_3)))$$

The remaining $M \subseteq \{1,\, 2,\, 3\}$ are obtained as complements of $\{1\}, \{1,\, 2\}$ and $\{1,\, 2,\, 3\}$. The corresponding L_Ms are simply the R_is (see formula 2 in section 3 for the definition of R_is). In view of remark 3.2 we have

$$R_i((a_1,\, a_2,\, a_3),\, (b_1,\, b_2,\, b_3)) =$$

$$= L_i(L_0((a_1,\, a_2,\, a_3),\, (b_1,\, b_2,\, b_3)), L_3((a_1,\, a_2,\, a_3),\, (b_1,\, b_2,\, b_3)))$$

$\forall\, 0 \leq i \leq 3$. This shows that the family of masked crossovers is contained in $\mathit{\Gamma}^{(2)}$. On the other hand, if we exhaust all possible transformations of the form

$$T((a_1,\, a_2,\, a_3),\, (b_1,\, b_2,\, b_3)) =$$

$$= L_i(L_j((a_1,\, a_2,\, a_3),\, (b_1,\, b_2,\, b_3)), L_k((a_1,\, a_2,\, a_3),\, (b_1,\, b_2,\, b_3)))$$

by performing the computations similar to the ones above, we'll see that all of them are masked crossovers, so that $\mathit{\Gamma}^{(2)}$ is exactly the family of masked crossovers (see also proposition 3.1 and notice that $\mathit{\Gamma}^{(2)} \subseteq \bar{\mathit{\Gamma}}$).

Remark A.4. Notice that a given family $\mathit{\Gamma}$ is composition closed, that is $\mathit{\Gamma} = \bar{\mathit{\Gamma}}$, if and only if $\mathit{\Gamma} = \mathit{\Gamma}^{(2)}$: Indeed, if $\mathit{\Gamma} = \mathit{\Gamma}^{(2)}$, then we have $\mathit{\Gamma} \subseteq \mathit{\Gamma}^{(1)} \subseteq \mathit{\Gamma}^{(2)} = \mathit{\Gamma}$ so that $\mathit{\Gamma}^{(2)} = \mathit{\Gamma}^{(1)}$. But then $\mathit{\Gamma}^{(3)} = \mathit{\Gamma}^{(2)} \cup \{T \mid T(\mathbf{x}) = T_0(T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_m(\mathbf{x})),\ T_0 \in \mathit{\Gamma},\ T_i \in \mathit{\Gamma}^{(2)}\ \forall\ 1 \leq i \leq m\} = \mathit{\Gamma}^{(1)} \cup \{T \mid T(\mathbf{x}) = T_0(T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_m(\mathbf{x})),\ T_0 \in \mathit{\Gamma},\ T_i \in \mathit{\Gamma}^{(1)}\ \forall\, 1 \leq i \leq m\} = \mathit{\Gamma}^{(2)} = \mathit{\Gamma}$. Likewise, by induction it follows that $\mathit{\Gamma} = \mathit{\Gamma}^{(k)}$ for all $k \geq 1$. Finally, we obtain $\bar{\mathit{\Gamma}} = \bigcup_{k=1}^{\infty} \mathit{\Gamma}^{(k)} = \bigcup_{k=1}^{\infty} \mathit{\Gamma} = \mathit{\Gamma}$. On the other hand, if $\mathit{\Gamma} = \bar{\mathit{\Gamma}}$, then we have $\mathit{\Gamma} \subseteq \mathit{\Gamma}^{(2)} \subseteq \bar{\mathit{\Gamma}} = \mathit{\Gamma}$ so that we get $\mathit{\Gamma} = \mathit{\Gamma}^{(2)}$.

It is not in vain that we call $\bar{\mathit{\Gamma}}$ the composition closure of $\mathit{\Gamma}$. Indeed, $\bar{\mathit{\Gamma}}$ is the smallest composition closed family containing $\mathit{\Gamma}$. In order to prove this fact it is convenient to prove the following Lemma first:

Lemma A.6. *A given family of transformations $\mathit{\Gamma}$ is composition closed if the following condition holds: The family of transformations*

$$\{T \mid T(\mathbf{x}) = T_0(T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_m(\mathbf{x})),\ T_i \in \mathit{\Gamma}\ \forall\, 0 \leq i \leq m\} \subseteq \mathit{\Gamma}$$

Proof. According to remark A.4, it suffices to show that $\mathit{\Gamma} = \mathit{\Gamma}^{(2)}$. To prove the last equality, simply notice that by definition A.6 $\mathit{\Gamma}^{(1)} = \mathit{\Gamma}$. But then $\mathit{\Gamma}^{(2)} = \{T \mid T(\mathbf{x}) = T_0(T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_m(\mathbf{x})),\ T_0 \in \mathit{\Gamma},\ T_i \in \mathit{\Gamma}^{(1)}\ \forall\ 1 \leq i \leq m\} \subseteq \{T \mid T(\mathbf{x}) = T_0(T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_m(\mathbf{x}))\ T_i \in \mathit{\Gamma}\ \forall\ 0 \leq i \leq m\} \subseteq \mathit{\Gamma}$ again, by assumption. That is, it follows that $\mathit{\Gamma}^{(2)} \subseteq \mathit{\Gamma}$. The reverse inclusion is automatic, so that the desired conclusion that $\mathit{\Gamma} = \mathit{\Gamma}^{(2)}$ follows at once. □

Proposition A.7. *The family $\bar{\mathit{\Gamma}}$ is composition closed. In other words, $\bar{\mathit{\Gamma}} = \bar{\bar{\mathit{\Gamma}}}$. Moreover, $\bar{\mathit{\Gamma}}$ is the smallest composition closed family containing $\mathit{\Gamma}$. In other words, given any other family of transformations Θ which is composition closed and contains $\mathit{\Gamma}$, we must have $\bar{\mathit{\Gamma}} \subseteq \Theta$.*

Proof. According to lemma A.6, in order to prove that $\bar{\Gamma} = \bar{\bar{\Gamma}}$ it suffices to show that $\bar{\Gamma}$ satisfies the condition of lemma A.6. To establish the validity of the condition, it suffices to show that if we are given any transformation $T : \Omega^m \to \Omega$ which is of the form $T(\mathbf{x}) = T_0(T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_m(\mathbf{x}))\ \forall\ \mathbf{x} \in \Omega^m$ where $T_i \in \bar{\Gamma}$ for $0 \leq i \leq m$, then $T \in \bar{\Gamma}$. Recall that $\bar{\Gamma} = \bigcup_{k=0}^{\infty} \Gamma^{(k)}$, which implies, in particular, that $T_0 \in \Gamma^{(k)}$ for some $0 \leq k$. Let q denote the least integer such that $T_0 \in \Gamma^{(q)}$. We shall prove that $T \in \bar{\Gamma}$ by induction on q: For $q = 1$, $T_0 \in \Gamma$. Notice also that $T_i \in \Gamma^{(k_i)}$ for $k_i \geq 0$. (recall, that $T_i \in \bar{\Gamma}$ and $\bar{\Gamma} = \bigcup_{k=0}^{\infty} \Gamma^{(k)}$) Since $\Gamma^{(k)}$'s are nested sets, (see definition A.6) $T_i \in \Gamma^{(k)}\ \forall\ 1 \leq i \leq m$ where $k = \max\{k_i \mid 1 \leq i \leq m\}$. But then, by definition A.6 we have $T \in \Gamma^{(k+1)}$. This establishes the base case. Now, suppose the statement holds for some $h \geq 0$, and suppose that the least integer k such that $T_0 \in \Gamma^{(k)}$ is $h + 1$. This implies that there exist transformations $U_1, U_2, \ldots, U_m \in \Gamma^{(h)}$ and $U_0 \in \Gamma$ such that for every $\mathbf{x} \in \Omega^m$ we have $T_0(\mathbf{x}) = U_0(U_1(\mathbf{x}), U_2(\mathbf{x}), \ldots, U_m(\mathbf{x}))$. In particular, we have

$$T(\mathbf{x}) = T_0(T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_m(\mathbf{x})) =$$

$$U_0(U_1((T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_m(\mathbf{x})), U_2((T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_m(\mathbf{x})), \ldots$$

$$\ldots, U_m((T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_m(\mathbf{x})))$$

Now consider the transformations $V_i : \Omega^m \to \Omega$ defined as follows: $\forall\ \mathbf{x} \in \Omega^m V_i(\mathbf{x}) = U_i((T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_m(\mathbf{x}))$. Since $U_i \in \Gamma^{(h)}$, from the inductive hypothesis it follows that $V_i \in \bar{\Gamma}$. Thereby, we have

$$\forall\ \mathbf{x} \in \Omega^m\ T(\mathbf{x}) = U_0(V_1(\mathbf{x}), V_2(\mathbf{x}), \ldots, V_m(\mathbf{x}))$$

where $U_0 \in \Gamma = \Gamma^{(1)}$ and $V_i \in \bar{\Gamma}$. Now, upon invoking the case with $q = 1$ which has been proved below, we immediately deduce that $T \in \bar{\Gamma}$. By principle of induction it follows now that $\bar{\Gamma}$ is composition closed.

The only remaining part to prove is that $\bar{\Gamma}$ is the smallest composition closed family containing Γ, but notice that if Θ is any other composition closed family such that $\Gamma \subseteq \Theta$, then, from definition A.6 together with remark A.4 it follows that $\Gamma^{(1)} \subseteq \Theta^{(1)} = \Theta$. Following the inductive chain of reasoning, we deduce that $\Gamma^{(k)} \subseteq \Theta$ for all $k \geq 0$, so that $\bar{\Gamma} = \bigcup_{k=0}^{\infty} \Gamma^{(k)} \subseteq \Theta$. □

Definition A.7. Given a family of functions $\Gamma \subseteq \{T \mid T : \Omega^m \to \Omega\}$ define the invariance closure of Γ, denoted by $\widetilde{\Gamma}$ as follows:
$\widetilde{\Gamma} = \{T \mid \forall\ \mathbf{x} \in \Omega^m\ \exists$ a transformation $T_{\mathbf{x}} \in \bar{\Gamma}$ such that $T(\mathbf{x}) = T_{\mathbf{x}}(\mathbf{x})\}$.

Example A.8. Continuing with example A.7, we have seen that $\Gamma^{(2)}$ is the family of all masked crossovers. Moreover, in section 3 we show that, in general, the family of masked crossovers is the composition closure of the family of one-point crossovers. The family $\widetilde{\Gamma}$ consists of many more transformations than $\bar{\Gamma}$ does. For instance, the following transformation, call it T, is in $\widetilde{\Gamma}$. T is defined as follows:

$$T((1, 0, 0), (1, 1, 1)) = T((1, 1, 1), (1, 0, 0)) = (1, 1, 1)$$

and $T(\mathbf{a}, \mathbf{b}) = L_2(\mathbf{a}, \mathbf{b})\ \forall\,(\mathbf{a}, \mathbf{b}) \in \Omega^2$ such that $\{\mathbf{a}, \mathbf{b}\} \neq \{(1, 1, 1), (1, 0, 0)\}$. Notice that, according to definition A.7 $T \in \widetilde{\Gamma}$. On the other hand, $T \notin \bar{\Gamma}$. Indeed, suppose $T = L_M$ for some $M \subseteq \{1, 2, 3\}$. Then, $L_M((1, 0, 0), (1, 1, 1)) = (1, 1, 1)$ which implies that $M \cap \{2, 3\} = \emptyset$. On the other hand, since $L_M((1, 1, 1), (1, 0, 0)) = (1, 1, 1)$, we must have $M \supseteq \{2, 3\}$ which contradicts the conclusion of the previous sentence.

Thereby, we must have $T \notin \bar{\Gamma}$. Below we prove that $\widetilde{\Gamma}$ is the largest family of transformations whose family of invariant subsets is Λ_Γ (see theorem A.9). It is, nevertheless, important to notice that although $\widetilde{\Gamma}$ is strictly larger than $\bar{\Gamma}$, when dealing with a genetic algorithm, $\bar{\Gamma}$ is large enough for all practical purposes, because performing crossover consists of selecting an appropriate transformation depending on the specific pair $(\mathbf{a}, \mathbf{b}) \in \Omega^2$, and $\forall T \in \widetilde{\Gamma}\ \exists\, T_{(\mathbf{a},\mathbf{b})} \in \bar{\Gamma}$ such that $T_{(\mathbf{a},\,\mathbf{b})}(\mathbf{a},\, \mathbf{b}) = T((\mathbf{a},\, \mathbf{b}))$ (see also the discussion following definition A.8).

It is clear from the definitions A.6 and A.7 that we have $\Gamma \subseteq \bar{\Gamma} \subseteq \widetilde{\Gamma}$. From Proposition A.5 it then follows that $\Lambda_\Gamma \supseteq \Lambda_{\bar{\Gamma}} \supseteq \Lambda_{\widetilde{\Gamma}}$.

On the other hand, it follows by a straightforward induction on k that $\Lambda_{\Gamma^{(k)}} = \Lambda_\Gamma$. But then, it easily follows that $\Lambda_{\bar{\Gamma} = \bigcup_{k=1}^{\infty} \Gamma^{(k)}} = \Lambda_\Gamma$.

We now show that, in fact, we also have $\Lambda_{\bar{\Gamma}} \subseteq \Lambda_{\widetilde{\Gamma}}$: This means that for any $K \in \Lambda_{\bar{\Gamma}}$ and for any $T \in \widetilde{\Gamma}$ we have $T(K^m) \subseteq K$. So, fix any $T \in \widetilde{\Gamma}$ and $K \in \Lambda_{\bar{\Gamma}}$. Fix any $\mathbf{x} \in K^m$. Our goal is to show that $T(\mathbf{x}) \in K$: Indeed, by definition A.7 there exists a transformation $T_{\mathbf{x}} \in \bar{\Gamma}$ such that $T(\mathbf{x}) = T_{\mathbf{x}}(\mathbf{x}) \in K$ because $K \in \Lambda_{\bar{\Gamma}}$. Thereby, $T(K^m) \subseteq K$ and the desired conclusion follows.

Thereby, we deduce that $\Lambda_\Gamma = \Lambda_{\bar{\Gamma}} = \Lambda_{\widetilde{\Gamma}}$. As a matter of fact, a lot more is true: Given any family of transformations $\Theta \subseteq \{T \mid T : \Omega^m \to \Omega\}$ such that $\Lambda_\Theta = \Lambda_\Gamma$, we must have $\Theta \subseteq \widetilde{\Gamma}$. To prove the last assertion we shall need the following Lemma:

Lemma A.8. *Fix any point $\mathbf{x} \in \Omega^m$. Then, for any $y \in S_{\mathbf{x}}$ (see definition A.4) there exists a transformation $F \in \bar{\Gamma}$ such that $F(\mathbf{x}) = y$.*

Proof. Fix a point $\mathbf{x} = (x_1, x_2, \ldots, x_m) \in \Omega^m$ and consider the set $K = \{x \mid x = x_i \text{ for } 1 \leq i \leq m\}$. In other words, K is the set of coordinates of the point $\mathbf{x}$. Then, we have $S_{\mathbf{x}} = S_K = \Gamma^\infty(K) = \bigcup_{k=1}^{\infty} \Gamma^k(K)$ (see definition A.4, definition A.3, definition A.5 and proposition A.4). Therefore, it suffices to show that for every $k \geq 1$ and $y \in \Gamma^k(K)$ (see definition A.5) there exists a transformation $F \in \Gamma^{(k)}$ such that $F(\mathbf{x}) = y$ (according to the definition A.6 this is a stronger statement). We proceed by induction on k: Indeed, for $k = 1$ we have $\Gamma^{(1)}(K) = \{y \,|\, y = T(\mathbf{x}),\ x \in K^m \text{ and } T \in \Gamma = \Gamma^{(1)}\}$ and the base case follows at once. Now, suppose the statement is true for some $k = j \geq 0$. We want to show the validity of the statement for $k = j + 1$: So, fix $y \in \Gamma^{j+1}(K) = \Gamma(\Gamma^j(K)) = \bigcup_{T \in \Gamma} T(\Gamma^j(K))$. But then we have $y = T_0(\mathbf{z})$ with $T_0 \in \Gamma$ and $\mathbf{z} = (z_1, z_2, \ldots z_m) \in (\Gamma^j(K))^m$. Thereby, $z_i \in \Gamma^j(K)$, so that, by inductive hypothesis, there exists a transformation $T_i \in \Gamma^{(j)}$ such that $T_i(\mathbf{x}) = z_i$ for every $1 \leq i \leq m$. Now, consider the transformation $F : \Omega^m \to \Omega$ defined as follows: $F(\mathbf{u}) = T_0(T_1(\mathbf{u}), T_2(\mathbf{u}), \ldots, T_m(\mathbf{u}))$ for all $\mathbf{u} \in \Omega^m$. Then, by definition A.6, we have $F \in \Gamma^{(j+1)}$ and $F(\mathbf{x}) = T_0(T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_m(\mathbf{x})) = T_0(z_1, z_2, \ldots, z_m) = y$. Now, the desired conclusion follows by the principle of induction. □

Theorem A.9. *Fix any family of transformations $\Gamma \subseteq \{T \mid T : \Omega^m \to \Omega\}$. The following are true:*

1. $\Lambda_\Gamma = \Lambda_{\bar{\Gamma}} = \Lambda_{\widetilde{\Gamma}}$
2. *For any family $\Theta \subseteq \{T \mid T : \Omega^m \to \Omega\}$ such that $\Lambda_\Theta = \Lambda_\Gamma$, we have $\Theta \subseteq \widetilde{\Gamma}$. In other words, $\widetilde{\Gamma}$ is the biggest family of mappings with the property that $\Lambda_{\widetilde{\Gamma}} = \Lambda_\Gamma$.*

proof of 1. See the discussion above. □

proof of 2. Fix any family of transformations Θ such that $\Lambda_\Theta = \Lambda_\Gamma$. We want to show that $\Theta \subseteq \widetilde{\Gamma}$. So, fix a transformation $T \in \Theta$. Our goal is to show that $T \in \widetilde{\Gamma}$, which means, according to the definition A.7, that for any given $\mathbf{x} \in \Omega^m$ there exists a transformation $T_\mathbf{x} \in \bar{\Gamma}$ such that $T(\mathbf{x}) = T_\mathbf{x}(\mathbf{x})$. Now consider the set $S_\mathbf{x}$. (see definition A.4) Since $S_\mathbf{x} \in \Lambda_\Gamma = \Lambda_\Theta$, we have $T(S_\mathbf{x}) \subseteq S_\mathbf{x}$. But then, since $\mathbf{x} \in S_\mathbf{x}$, we must have $T(\mathbf{x}) \in S_\mathbf{x}$, and the existence of the desired transformation $T_\mathbf{x} \in \bar{\Gamma}$ follows now from Lemma A.8. This finishes the argument. □

The following definition and the discussion which follows explain the significance of the composition closure:

Definition A.8. Given a family of functions $\Gamma \subseteq \{T \mid T : \Omega^m \to \Omega\}$ define the semi-invariance closure of Γ, denoted by $\widehat{\Gamma}$ as follows:

$$\widehat{\Gamma} = \{T \mid \forall\, \mathbf{x} \in \Omega^m\ \exists \text{ a transformation } T_\mathbf{x} \in \Gamma \text{ such that } T(\mathbf{x}) = T_\mathbf{x}(\mathbf{x})\}$$

Notice the similarity between the definition A.7 and the definition A.8. In fact, according to these definitions, the invariance closure of a given family of transformations Γ is the same thing as the semi-invariance closure of $\bar{\Gamma}$. One could say that the family $\bar{\Gamma}$ is one of the largest for all practical purposes families of transformations whose family of invariant subsets is Λ_Γ because, thanks to theorem A.9, the semi-invariance closure of $\bar{\Gamma}$ is the largest family of transformations fixing Λ_Γ.

B Appendix: Proof of Proposition 3.1

Proof of Proposition 3.1. In order to prove proposition 3.1 it is convenient to introduce the following definitions:

Definition B.1. Given a subset $M \subseteq \{1, 2, \ldots, n\}$, the first line segment of M is simply the ordered pair (i_1, i_2) where $i_1 = \min(M) - 1$ and $i_2 = \min\{i \mid i \notin M,\ i - 1 \in M\} - 1$. The 1^{st} successor of M is the set $M_1 = \{i \mid i \in M\ i > i_2\}$. For $q > 1$, if the $(q-1)^{\text{st}}$ successor of M, M_{q-1}, is not empty, the q^{th} line segment of M is the first line segment of the $(q-1)^{\text{st}}$ successor of M. Finally, if $M_{q-1} \neq \emptyset$, then the q^{th} successor of M is the 1^{st} successor of the $(q-1)^{\text{st}}$ successor of M.

For example, if $n = 10$ and $M = \{1, 5, 6, 7, 9, 10\}$ then the first line segment of M is the pair $(0, 1)$, the first successor of M is the set $M_1 = \{5, 6, 7, 9, 10\}$, the second line segment of M is the pair $(4, 7)$ and the second successor of M is the first successor of M_1, which is $(9, 10)$. Finally, the third line segment of M is the set $(8, 10)$ and the third successor of M is $\emptyset$.

Definition B.2. Given a subset $M \subseteq \{1, 2, \ldots, n\}$, we say that M is of degree l if its l^{th} line segment is defined. (This means that the $(l-1)^{\text{st}}$ successor of M, M_{l-1}, is nonempty. See definition B.1.)

We now prove the following claim by induction on the degree of the set M:

Claim: The set of transformations $\{L_M \mid M \text{ is of degree } l\} \subseteq \mathcal{F}_1^{(l+1)}$. (For the meaning of $\mathcal{F}_1^{(q)}$ apply definition A.6 to the family of one point crossover transformations, $\mathcal{F}_1$.)

Proof. First, notice that $\forall\, \vec{\mathbf{x}} \in S^2\ R_i = L_i(L_0(\vec{\mathbf{x}}), L_n(\vec{\mathbf{x}}))$. (Recall that L_n and L_0 are the first and the second coordinate projections respectively so that the formula follows from remark 3.1 and remark 3.2.) Thereby, according to definition A.6 $R_i \in \mathcal{F}_1^{(2)}$. Now fix any subset $M \subseteq \{1, 2, \dots, n\}$ of degree 1. Let (i_1, i_2) denote the first line segment of M. Then simply notice that $\forall\, \vec{\mathbf{x}} \in S^2$ we have $L_M(\vec{\mathbf{x}}) = L_{i_2}(R_{i_1}(\vec{\mathbf{x}}), L_{i_2}(\vec{\mathbf{x}}))$. (Indeed, since the degree of M is 1, we must have $M = \{i \,|\, i_1 < i \le i_2\}$. Let $\vec{\mathbf{x}} = (\mathbf{a}, \mathbf{b})$ with $\mathbf{a} = (a_1, a_2, \dots, a_n)$ and $\mathbf{b} = (b_1, b_2, \dots, b_n)$. Now according to the formulas (1), (2) and (3) in section 3 which define the various crossover operators involved here, we simply have $L_M(\vec{\mathbf{x}}) = (x_1, x_2, \dots, x_n)$ where $x_i = \begin{cases} a_i & \text{if } i \in M \\ b_i & \text{otherwise} \end{cases}$, or, since $M = \{i \,|\, i_1 < i \le i_2\}$, we can write $x_i = \begin{cases} a_i & \text{if } i_1 < i \le i_2 \\ b_i & \text{otherwise} \end{cases}$ Likewise, we have $R_{i_1}(\vec{\mathbf{x}}) = (y_1, y_2, \dots, y_n)$ where $y_i = \begin{cases} a_i & \text{if } i > i_1 \\ b_i & \text{otherwise} \end{cases}$ and $L_{i_2}(\vec{\mathbf{x}}) = (z_1, z_2, \dots, z_n)$ where $z_i = \begin{cases} a_i & \text{if } i \le i_2 \\ b_i & \text{otherwise} \end{cases}$. But then $L_{i_2}(R_{i_1}(\vec{\mathbf{x}}), L_{i_2}(\vec{\mathbf{x}})) = (u_1, u_2, \dots, u_n)$ where $u_i = \begin{cases} a_i & \text{if } i_1 < i \le i_2 \\ b_i & \text{otherwise} \end{cases}$ which coincides with $L_M(\vec{\mathbf{x}})$.) This shows that $L_M \in \mathcal{F}_1^{(2)}$ and establishes the base case of induction.

For the inductive step suppose the Claim holds for $q = j$. Fix any $M \subseteq \{1, 2, \dots, n\}$ of degree $j + 1$ and notice that the set M_1, the first successor of M is of degree j. (This follows by direct verification of definitions B.1 and B.2.) But then, by inductive hypothesis, $L_{M_1} \in \mathcal{F}_1^{(j+1)}$ and also $\forall\, \vec{\mathbf{x}} \in S^2$ we have $L_M(\vec{\mathbf{x}}) = L_{i_2}(R_{i_1}(\vec{\mathbf{x}}), L_{M_1}(\vec{\mathbf{x}}))$ where (i_1, i_2) is the first line segment of M. (this is seen by direct verification completely analogous to the one carried out in the base case.) So, according to definition A.6, $L_M \in \mathcal{F}_1^{((j+1)+1)}$. The desired conclusion now follows by principle of induction. □

It is easily seen that the degree of any subset $M \subseteq \{1, 2, \dots, n\}$ is at most $\lceil n/2 \rceil$, so that, according to the Claim above, we immediately obtain $\mathcal{F}_M \subseteq \bar{\mathcal{F}}_1$ (see definition A.6).

On the other hand, one can easily see by direct verification, that the family of masked crossovers $\mathcal{F}_M$ is composition closed, which, roughly speaking, means that performing masked crossover several times repeatedly gives us a masked crossover again. But then, from the minimality part of proposition A.7, the reverse inclusion follows. □

C Appendix: Summary of Radcliffe's Work

This section is devoted to summarizing the definitions and results of (Radcliffe, 1994) which are highly relevant to the current paper. As mentioned in the introduction, Radcliffe's notion of a forma is an extension of Holland schema for a representation independent setting.

Definition C.1. Let Ξ denote a partition of a set (search space) Ω into equivalence classes. A forma, by definition, is an element of Ξ. Given a sequence (usually finite) of partitions $\Psi = \Xi_1, \Xi_2, \dots, \Xi_n$, let $\mathbf{\Xi}(\Psi) = \prod_{i=1}^{n} \Xi_i$.

A classical example of forma is Holland schemata: Let $\Omega = \prod_{i=1}^{n} A_i$. Let $\Xi_i = \{(*, \dots, *, a_i, *, \dots, *) \mid a_i \in A_i\}$ [4]. In words, Ξ_i is simply a collection of subsets of Ω

[4] To simplify notation we shall denote by $(*, \dots, *, a_i, *, \dots, *)$ not only the schema itself, but also the subset of Ω determined by the schema $(*, \dots, *, a_i, *, \dots, *)$

determined by Holland schemata with i^{th} position fixed.

Definition C.2. A given family of m-ary reproduction transformations, Γ (Γ is simply a family of functions $\Omega^m \to \Omega$, see section 2), respects a given sequence of partitions $\Psi = \Xi_1,\ \Xi_2, \ldots, \Xi_n$ of Ω if and only if $\forall\, F \in \Gamma\ \forall\, 1 \leq i \leq n$ and $\forall\, \mathbf{x_1},\ \mathbf{x_2}, \ldots, \mathbf{x_m} \in \xi$ where $\xi \in \Xi_i$, we have $F(\mathbf{x_1},\ \mathbf{x_2}, \ldots, \mathbf{x_m}) \in \xi$.

In our notation definition C.2 simply says that every one of the forma from a given collection of partitions of Ω is invariant under Γ (i. e. everyone of the forma from a given collection of partitions of Ω is in Λ_Γ: see definition 2.1). In the case of the classical Holland schemata a given family of reproduction transformations, Γ, respects the family of Holland schemata if and only if whenever the i^{th} allele of both parents $\mathbf{a}$ and $\mathbf{b}$ is equal to x_i, then so is the i^{th} allele of their child. (In the notation of the current paper this simply means that Γ is a subfamily of the family of Holland transformations: see definition 4.2.) The family of the traditional one-point or masked crossover transformations enjoys more restrictive properties. This led Radcliffe to define the notion of *gene transmission*.

Definition C.3. We say that a given family of m-ary reproduction transformations, Γ, transmits genes if and only if $\forall\, F \in \Gamma\ \forall \xi_1 \in \Xi_1, \ldots, \xi_j \in \Xi_j, \ldots \xi_n \in \Xi_n$ and $\mathbf{x}_1 \in \xi_1, \ldots, \mathbf{x}_j \in \xi_j, \ldots, \mathbf{x}_m \in \xi_n$ we have $F(\mathbf{x_1},\ \mathbf{x_2}, \ldots, \mathbf{x_m}) \in \xi$.

When specialized to Holland schemata definition C.3 simply says that if the i^{th} allele of the parent $\mathbf{a}$ is a_i and the i^{th} allele of the parent $\mathbf{b}$ is b_i then the i^{th} allele of their child can only be either a_i or b_i. It turns out that the notion of *gene transmission* is also a special case of the notion of invariance which is studied in detail in the current paper (see proposition 4.1 and definition 4.1). The relationship between the traditional families of crossover transformations is studied in detail in section 3 and in the beginning of section 4.

Although the results of sections 3 and 4 can be formulated in a representation independent setting using the language of Radcliffe's forma, thanks to theorem 20 of (Mitavskiy, 2003a) and theorem 3.7 of (Mitavskiy, 2003b) greater generality would not be attained, yet the notation would become more complicated.

References

Antonisse, J. (1989). A new interpretation of schema notation that overturns the binary encoding constraint. In Schaffer, J., editor, *Proceedings of the Third International Conference on Genetic Algorithms*, pages 86–97, Morgan Kaufmann.

Booker, L. (1993). Recombination distributions for genetic algorithms. In Whitley, D., editor, *Foundations of Genetic Algorithms 2*, pages 29–44, Morgan Kaufmann.

Geiringer, H. (1944). On the probability of linkage in Mendelian heredity. *Annals of Mathematical Statistics*, 15:25–57.

Liepins, G. and Vose, M. (1992). Characterizing crossover in genetic algorithms. *Annals of Mathematics and Artificial Intelligence*, 5:27–34.

MacLane, S. (1971). *Categories for the Working Mathematician.* Springer-Verlag.

Michalewicz, Z. (1996). *Genetic Algorithms + Data Structures = Evolution Programs.* Springer-Verlag.

Mühleinbeim, H. and Mahnig, T. (2001). Evolutionary computation and beyond. In Uesaka, Y., Kanerva, P. and Asoh, H., editors, *Foundations of Real-World Intelligence*, pages 123–188, CSLI Publications.

Mitavskiy, B. (2003a). Comparing evolutionary computation techniques via their representation. In Cantú-Paz et al., editors, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*, pages 1196–1209, Springer Verlag.

Mitavskiy, B. (2003b). A category theoretic method for comparing evolutionary computation techniques via their representation. In ten Cate, B., editor, *Proceedings of the Eighth ESSLLI Student Session*, pages 201–210.

Radcliffe, N. (1994). The algebra of genetic algorithms. *Annals of Mathematics and Artificial Intelligence*, 10:339–384. http://users.breathemail.net/njr/papers/amai94.pdf

Rowe, J., Vose, M., and Wright, A. (2002). Group properties of crossover and mutation. *Evolutionary Computation*, 10(2):151–184.

Stadler, B. and Stadler, P. (2002). Generalized topological spaces in evolutionary theory and combinatorial chemistry. *Journal of Chemical Information and Computer Sciences*, 42:577–585.

Stadler, P. and Wagner, G. (1997). Algebraic theory of recombination spaces. *Evolutionary Computation*, 5(3):241–275.

Stephens, C., Poli, R., Wright, A., and Rowe, J. (2002). Exact results from a coarse grained formulation of the dynamics of variable-length genetic algorithms. In Langdon, W., and Cantu-Paz, E., editors, *GECCO 2002: Proceedings of the Genetic and Evolutionary Computation Conference*, pages 578–585, Morgan Kaufmann Publishers.

Stephens, C. (2001). Some exact results from a coarse grained formulation of genetic dynamics. In Cantu-Paz, E., Goldberg, D., *Proceedings of the Genetic and Evolutionary Computation (GECCO) conference*, pages 631–638, Morgan Kaufmann.

Stephens, C. and Waelbroeck, H. (1999). Schemata evolution and building blocks. *Evolutionary Computation*, 7(2):109–124.

Stephens, C., Waelbroeck, H., and Aguirre, R. (1997). Schemata as building blocks: does the size matter. In Belew, R. and Vose, M., editors, *Foundations of Genetic Algorithms 5*, pages 117–133, Morgan Kaufmann.

Stephens, C. (2002). The renormalization group and the dynamics of genetic systems, to be published in *Acta Physica Slovaka*. http://arXiv.org/abs/cond-mat/0210271/.

Syswerda, G. (1993). Simulated crossover in genetic algorithms. In Whitley, D. editor, *Foundations of Genetic Algorithms 2*, pages 239–255, Morgan Kaufmann.

Vose, M. and Wright, A. (1998). The simple genetic algorithm and the Walsh transform: Part II, the inverse. *Evolutionary Computation*, 6(3):275–289.

Vose, M. and Wright, A. (2001). Form invariance and implicit parallelism. *Evolutionary Computation*, 9(3):355–370.

Vose, M. (1991). Generalizing the notion of a schema in genetic algorithms. *Artificial Intelligence*, 50(3): 385–396.

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

Wright, A., Rowe, J., Poli, R., and Stephens C. (2002). A fixed point analysis of a gene pool GA with mutation. *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)* Morgan Kaufmann. http://www.cs.umt.edu/u/wright/.