Pure Crossovers

Definition, Their Relation to Geiringer's Theorem for Finite Populations and Practical Value

J. Cervantes[1] and A. Moraglio[2]

[1] Universidad Autónoma Metropolitana
Departament of Applied Mathematics and Systems
Artificios s/n Del. Álvaro Obregón, México D.F. México01120
JorgeCervantesO@aim.com
[2] Centre for Informatics and Systems of the University of Coimbra
Polo II - Pinhal de Marrocos, 3030 Coimbra, Portugal
moraglio@dei.uc.pt

Abstract. Motivated by a practical interpretation of the Geiringer's theorem, we define the class of "Pure Crossovers" independently of the solution representation used and explain why they are important for both theorists and practitioners. We then give a geometric characterization of this class of operators and prove some general properties common to all pure crossovers.

Keywords. GA Theory, Pure Crossovers, Geometric Crossovers, Geiringer Theorem, Finite Populations

1 Introduction

Crossover is the most complicated operator to understand from a theoretical point of view compared to selection and mutation. What are the effects of crossover? What are the principles underlying a certain observed behaviour in an EA? What is the theory that can be used to unify criteria and analyze EAs?

One key factor that is needed is a clear classification of genetic operators and their associated characteristic effects on the evolutionary process. In this direction, the work of Moraglio and Poli [3] [4] sets a classification of crossovers in geometric and non-geometric. Interestingly, this classification is representation-independent and relies on the notion of inbreeding properties, which are properties that characterize recombination operators by stating what offspring may or may not be obtained as a result of recombining close relative solutions, such as, for example, parents and offspring. Many pre-existing recombination operators for the most frequently used representations belong to the class of geometric crossovers [5].

Motivated by a practical interpretation of the Geiringer's theorem for finite populations, in this work we extend this classification and define a new subclass of geometric crossovers, the Pure Crossovers. We analyze the inbreeding properties of crossovers of this kind and, more importantly, describe why it is good for theorists and practitioners to know these properties.

2 The Practical side of Geiringer's Theorem

According to Geiringer's theorem [1], (see also [2]), the effect on an infinite population with fixed length genomes of a crossover operator (without mutation or selection) that produces offspring at generation t by taking, for each position i of their genome, only one of the genes in position i of either of its parents, is to make the proportions P_I of every configuration I tend to P_I^* given by

$$P_I^* = \lim_{t \to \infty} P_I(t) = \prod_i P_{I_i}(0) \tag{1}$$

where P_{I_i} is the proportion of the population having the ith bit equal to the ith bit of I.

Given an initial distribution of an infinite population and applying only such a recombination operator, this formula provides the asymptotical proportions of every *schema* I in terms of its most basic components which are the proportions of order 1 schemata. For instance, if $I = **101$ then

$$P_I^* = P_{**1**}(0) P_{***0*}(0) P_{****1}(0) \tag{2}$$

considering that $P_{*****}(0) = 1$. Note that these asymptotic proportions depend only on the *initial* proportions of the order 1 schemata.

If one wants to use in a practical way (finite populations) the theoretical view of Geiringer's theorem then the crossover operator must keep every P_{I_i} constant over time because this theorem is expressed in terms of the initial values of each P_{I_i} and any change to them implies a change in the asymptotic distribution of the population which is contradictory with the asymptotic claim of Geiringer's theorem. This is possible, for instance, when every pair of parents produce two offspring by taking always, for each gene position, both parents' genes and placing them on both offspring and then both parents are replaced by both offspring If formula 1 does not hold it means that the crossover operator induces some effects that produce changes in at least one of the P_{I_i}. These effects can be compared to either selective or mutational effects. Selective effects would make the P_{I_i} tend to either 1 or 0, i.e., convergence. Mutational effects would make all P_{I_i} tend to 0.5 which is the value one would expect in a random population, i.e., divergence.

Above, we have seen an example of such an operator for binary strings. More in general, what kind of crossover operators have this non-selective and non-mutational property? It is clear that the specific representation used determines the form of the corresponding crossover. This imposes severe difficulties on identifying a general class of operators with this property because representations can vary from one application to another. Below, we give a first tentative definition of a general class of crossovers that have no selective or mutational effects independently of the representation one uses. In the remainder of the paper, we will make this definition exact. The advantage of having a representation independent definition is clear, one could, in principle, examine any particular operator on a particular representation and check if it belongs to this class. Then, if it belongs, it is reasonable to conjecture that a representation-independent generalization of the Geiringer's theorem would hold and one could analyze the effects of selection, mutation and crossover separately because the crossover operator contains no other effect than that stated in Geiringer's Theorem. We call this class of crossover operators "*Pure Crossovers*".

Definition 1. *(Pure Crossovers (tentative definition)) A crossover operator is Pure if its effects on a population are exclusively recombinative, this is, without any selective (convergence) nor mutational (divergence) effects.*

Above, we have explained what selective and mutational effects are for the specific case of binary strings. However, for our tentative definition of pure crossovers to become exact and truly representation-independent, we need to characterize, in a representation-independent fashion, the notions of selective and mutational effects. Also, in order to relate them to a possible representation-independent generalization of the Geiringer's theorem, we need to define in a representation-independent way what we mean by "proportion of the population having the ith bit equal to the ith bit of a binary string I". Before proceeding with this, we explain, in the next section, why pure crossovers are valuable for both theoreticians and practitioners.

3 Why Pure Crossovers?

The definition of pure crossovers naturally allows for a clear-cut distinction of crossover (mixing), mutational (variation) and selective (fixation) effects on the basis of the convergence properties of the operator on the dynamic of the population: (i) pure mixing effects are equated with a population without convergence or divergence (ii) pure variation effects are equated with a divergent population, and (iii) pure fixation effects are equated with a convergent population.

Non-pure crossovers introduce a bias that can be rendered as a selective and/or mutational bias in the effects of crossovers. These induced effects are noisy, are not controlled by the designer of the algorithm and can hardly be measured separately. This makes the process much more unpredictable and does not allow for a precise analysis of the principles underlying a certain observed behaviour.

If a pure crossover is used, then any selection effect is due only to the selection operator being used. The same for mutation effects. This makes it easier to study and control each operator's effects separately with the immediate consequence that these effects would be better understood by theorists who would be able to create a better evolutionary theory and practitioners who would be able to better adapt their algorithms to their own needs.

Pure crossovers have another interesting and useful property. When a GA reaches a steady state means that selective and mutational effects cancel and there is no convergence or divergence anymore. This situation is not altered by the action of a Pure crossover. Under these circumstances, Geiringer's theorem still applies, even in the presence of the other operators. Thus one is able to calculate the (asymptotic) expected proportion of each genome generated by recombination for the next generation.

There are two ways to proceed in order to generalize the notion of pure crossover beyond the binary representation and in a representation-independent way. 1) The first one consists in generalizing the ability of pure crossover of keeping the proportions of some representation-independent generalized notion of order 1 schemata constant over time. 2) The alternative is to generalize the ability of pure crossover of keeping the population in a non-convergent non-divergent state. The first generalization can be naturally used as a starting point to generalize the Geiringer's theorem to any solution representation. The second generalization links naturally with the notion of geometric crossover. The two alternative generalizations are not mutually exclusive. They meet nicely as we will show in section 5. In the following we present a general, representation-independent definition of pure crossover based on the first alternative.

4 Fixed crossovers

Pure crossovers are inherently linked with the collection of sets they keep the proportions in the population constant over time. Therefore, we need to characterize formally this collection of sets.

Definition 2. *(Bipartition structure) Let S be a set of points (a search space of possible solutions). We say that the collection of sets $\mathcal{B} \subseteq 2^S$ is a bipartition structure on the set S if for any set $i \in \mathcal{B}$ also the complement of i with respect to S belongs to $\mathcal{B}$, which is $\bar{i} \in \mathcal{B}$.*

In other words, a bipartition structure $\mathcal{B}$ is a set of subsets i of S where these subsets of S are always accompanied by their complementary set $\bar{i}$ with respect to S. $\mathcal{B}$ can be seen as a list (exhaustive or not) of different ways to partition S in two subsets i and $\bar{i}$.

Definition 3. *(Proportion of a set) Let S be a set of points (a search space of possible solutions), let $P \subseteq S$ be a population of points and let i be any subset of S. The proportion of i in P is the number of points of P which are also in i over the number of points in P, which is $P_i = \frac{|i \cap P|}{|P|}$.*

The previous definition naturally extends to the case in which the population P is a multi-set. Notice that the proportion of the complement of i with respect to S is $P_{\bar{i}} = 1 - P_i$ because each point in the population belongs either to i or to $\bar{i}$.

Definition 4. *(Fixed-proportions population) A population has fixed-proportions with respect to a bipartition structure $\mathcal{B}$, if the proportion P_i of any set $i \in \mathcal{B}$ remains constant over time.*

Notice that a fixed-proportions population is not necessarily neither non-convergent nor non-divergent with respect to any metric. We make precise the notion of convergence in the next section. We will also see that when the bipartition structure is the set of all metric half-spaces for some metric, then the notion of non-divergent non-convergent population becomes equivalent to the notion of fixed-proportions population.

Definition 5. *(Fixed Operator) A genetic operator is Fixed if, acting alone, it makes the population remain fixed-proportions over time with respect to a bipartition structure $\mathcal{B}$.*

Notice that the property of being Fixed is relative to the structure $\mathcal{B}$. On two different bipartition structures, the same operator can be fixed with respect to one and non-fixed with respect to the other.

In the following, we show how the definitions above generalize traditional crossovers for binary strings.

Proposition 1. *The set of all order 1 schemata of length n is a bipartition structure on the set of all binary strings of length n.*

This is easy to see. For example, the order 1 schemata 1*** and 0*** are complementary and form a bipartition of all binary strings of size 4.

Proposition 2. *Any mask-based crossover for binary strings which replaces 2 parents with 2 complementary offspring is a fixed crossover with respect to the bipartition structure of order 1 schemata.*

This is because offspring match exactly the same order 1 schemata their parents match. Therefore, by replacing parents by their offspring, the proportions of strings in the population matching any order 1 schema remain constant.

Proposition 3. *For binary strings, Pure crossovers are fixed crossovers with respect to the bipartition structure of all order 1 schemata.*

5 Geometric Characterization of pure crossover

In the following we present a geometric characterization of pure crossover which links it to geometric crossover.

Definition 6. *(Metric segment) Given a metric space $M = (S, d)$, the metric segment between two points $a, b \in S$ is the set $[a, b]_d = \{c \in S | d(a, c) + d(c, b) = d(a, b)\}$. The length of the metric segment $[a, b]_d$ is $l([a, b]_d) = d(a, b)$.*

Proposition 4. *(Multiple end-points of a segment) The points a and b are called a pair of end-points of the metric segment $[a, b]_d$. In general, a metric segment can have more than a pair of end-points. So, for a segment $[a, b]_d$ two points $x, y \in S$ different from a, b may exist such that $[a, b]_d = [x, y]_d$.*

Definition 7. *(Metric convex set) Given a metric space $M = (S, d)$, a set $H \subseteq S$ is d-convex if $\forall a, b \in H : H \supseteq [a, b]_d$.*

Definition 8. *(Metric half-space) Given a metric space $M = (S, d)$, a d-convex set H is a d-half-space if its complement $S \setminus H$ is also a d-convex set.*

Definition 9. *(Geometric crossover) A recombination operator $OP : S \times S \to S$ is a geometric crossover with respect to metric d, if $\forall x, y \in S : OP(x, y) \subseteq [x, y]_d$.*

Definition 10. *(Segment-preserving crossover) A recombination operator $OP : S \times S \to (S \times S)$ is a Segment-preserving crossover with respect to metric d, if $\forall x, y \in S$ and $\forall w, z$ such that $OP(x, y) \mapsto (w, z) : [x, y]_d = [w, z]_d$.*

Proposition 5. *(Geometric schemata) (i) Convex sets are invariant under geometric crossover GX: for any convex set C if $x, y \in C$ then $GX(x, y) \subseteq C$. (ii) Metric convex sets for the Hamming distance coincide with traditional schemata.*

Proposition 6. *(Order 1 geometric schemata) Metric half-spaces for the Hamming distance coincide with the traditional order 1 schemata.*

Theorem 1. *(Segment-preserving crossover with parent replacement Fixed with respect to $\mathcal{B}$) For any distance d, any Segment-preserving crossover that replaces parents by their offspring is a Fixed crossover with respect to the bipartition structure $\mathcal{B}$ consisting of the set of all metric half-spaces i under d.*

Proof. We will prove this theorem by showing that a segment-preserving crossover that replaces parents by their offspring applied to any two parents in a population does not change the proportions P_i for any half-space i.

Case 1: both parents belong to the same half-space i. For any half-space i, if $a, b \in i$ then a, b do not belong to $\bar{i} = S \setminus i$, which is the complementary half-space to i. For the convexity of half-space i, since $a, b \in i$ then their offspring $x, y \in i$. Therefore, x, y do not belong to $\bar{i} = S \setminus i$.

Case 2: one parent belongs to the half-space i, the other parent belongs to the complementary half-space $\bar{i}$. For any half-space i, if $a \in i$ and $b \in \bar{i}$, then their offspring x, y under pure crossover must be in complementary half-spaces, say $x \in i$ and $y \in \bar{i}$. This can be shown by contradiction as follows. For the definition of pure crossover, we have $[a, b] = [x, y]$. From this we have that $a, b \in [x, y]$. If both offspring x, y are in the same half-space i then, for the convexity of i, we have $[x, y] \subseteq i$ and, therefore, $a, b \in i$ contradicting the hypothesis that parents belong to complementary half-spaces.

So, for any half-space i, by applying a distance preserving crossover and replacing parents by their offspring does not change the proportions P_i. Thus the population is non-convergent and non-divergent and consequently this crossover is Fixed.

Notice that a segment-preserving crossover can be fixed also with respect to bipartition structures other than the one consisting of all half-spaces. For example, a segment-preserving crossover is fixed with respect to all bipartition structures consisting of only some half-spaces.

6 Inbreeding properties of Fixed crossover

Inbreeding properties characterize recombination operators in a representation-independent fashion by stating what offspring may or may not be obtained as a result of recombining close relative solutions, such as, for example, parents and offspring. In the following, we present the inbreeding properties that completely characterize Fixed crossovers.

Definition 11. *(Cycle inbreeding property) A recombination operator has the cycle inbreeding property if for any choice of parents it is possible to obtain both of them as a result of recombining their offspring.*

```
X --- Y
   |
-------
|     |
W --- Z
   |
-------
|     |
X     Y
```

Fig. 1. Simple cycle property

```
X --- Y
   |
-------
|     |
W1---Z1
   |
-------
|     |
W2---Z2
   |
   :
   :
-------
|     |
X     Y
```

Fig. 2. General cycle property

The inbreeding diagram in Figure 1 illustrates the cycle inbreeding property: the parents X and Y are recombined obtaining the offspring W and Z, which when recombined can produce as their offspring X and Y.

The cycle inbreeding property can be generalized by requiring only that by successive recombinations of offspring among themselves, eventually both their parents can be obtained. This is illustrated in Figure 2: the parents X and Y are recombined obtaining the offspring $W1$ and $Z1$, which when recombined produce as their offspring $W2$ and $Z2$, which when recombined produce $W3$ and $Z3$, and so on and so forth, until at a certain generation when X and Y can be generated back. This property implies that at each generation the genetic material needed to generate the original parents is not altered, while keeping the population size constant too, so that any pair of parents can be eventually recovered.

Definition 12. *(Invariant set) A set H is said to be invariant for a recombination operator OP if all offspring obtained by recombining elements of H are also in H.*

Theorem 2. *(Cycle inbreeding property and Fixed crossover) Every recombination operator OP with the general cycle inbreeding property is a fixed crossover with respect to any bipartition structure made of invariant sets for the operator OP.*

Proof. A simple generalization of the proof of theorem 1 proves it.

7 Selection and mutation as non-fixed operators

Theorem 2 can be generalized to the case of a recombination operator of input arity n and output arity n, which is for a recombination operator that takes n parents in input and produces n offspring. The interesting aspect of this generalization is that it allows us to relate fixed crossover with selection and mutation. We illustrate this in the following.

Firstly, let us consider selection. Selection can be seen as an operator which acts at a population level: it takes all the population in input and produces a new population in output of the same size, with the same members or with less members, and with possibly changed frequency of each member. If the population has size n, the selection operator has input arity n and output arity n.

Let us consider if the selection operator meets the conditions of theorem 2. Interestingly, the selection operator is invariant with respect to any set H, because if its input members (parents) belong to H, clearly its output members (offspring) belong to H as well. Therefore, this condition holds. However, the selection operator does not have the general cycle inbreeding property. This is because by successive applications of selection to a population we may never obtain the original population back. So, selection seen as an operator fails to be fixed because of the lack of cycle property.

Let us now consider mutation. Mutation is a unary operator both in input and in output: it takes one parent and it returns one offspring. By repeated applications of mutation, we can obtain back the original parent from one of its successors for any choice of the parent. So, the cycle property holds. However, there are no invariant sets for (ergodic) mutation, because for any set H not including all solutions, the mutation operator can with probability greater than zero generate an offspring not in H. So, mutation fails to be fixed because of the lack of invariant sets to it.

The interesting aspect of analyzing selection and mutation as fixed operators is that they fail to be so for opposite reasons. Both, mutation and selection, are not fixed operators since they necessarily change the proportions in the initial population with respect to any proper bipartition structure.

Above we have seen how selection and mutation can't be fixed operators, only crossovers can. This suggests that any crossover that fails to be fixed must have an intrinsic effect that equals either selection or mutation effects and, consequently, can't be Pure. (According to definition 1) Under this view, a non-pure crossover can be modeled as a Pure crossover plus some amount of either selection or mutation accordingly. For the inbreeding properties of fixed crossover, a crossover operator must have the cycle property in order to be Fixed. The family of sets this operator is fixed for is the family of sets for which it is invariant.

Proposition 7. *Pure crossovers have the cycle inbreeding property.*

Proposition 8. *A class of crossovers which can concretely be characterized as Pure is the class of segment-preserving crossovers with parent replacement, which are both fixed with respect to the set of half-spaces associated with them and (hence) geometric.*

8 Convergence and pure crossover

Fixed crossovers, by definition, keep the proportions of some sets covering the population constant over time. Therefore, intuitively, fixed crossovers may keep the population in an orbit which is non-convergent and non-divergent. In the following, we analyze formally how Pure crossovers relate with population convergence.

Definition 13. *(Non-divergence) A population P is said to be non-diverging under the operator OP acting at a population level, if there is a metric d such that the metric convex hull of P includes the metric convex hull of $OP(P)$, which is if $co(P) \supseteq co(OP(P))$.*

When the inclusion between convex hulls is strict, the population is said to be converging.

Definition 14. *(Non-convergence non-divergence) A population P is said to be non-convergent non-divergent under the operator OP acting at a population level, if there is a metric d such that the metric convex hull of P equals the metric convex hull of $OP(P)$, which is if $co(P) = co(OP(P))$.*

Theorem 3. *(Segment preserving and non-convergence non-divergence) A population under Segment-preserving crossover replacing n parents by n offspring is non-convergent non-divergent.*

Proof. $co(P)$ is the union of all segments $[a, b]$ with $a, b \in P$. Since the crossover operator preserves all segments, co(P) $= co(OP(P))$.

Proposition 9. *A population under fixed crossover can be diverging or converging depending with respect to which family of sets it is fixed.*

Proposition 10. *A population evolving under a crossover operator that is fixed with respect to the bipartition structure of all metric half spaces is non-convergent non-divergent.*

Hence, a population under the Pure crossover in proposition 8 is non convergent non-divergent.

Conjecture 1. A population evolving under *any* Pure crossover is non-convergent non-divergent.

9 Conclusions and future work

Crossovers have an intrinsic effect on populations that is not present in either selection or mutation but it is possible that a crossover operator contains, besides its recombinative effect, additional selective and/or mutational effects. So we have defined the class of Pure crossovers which have exclusively recombinative effects. This classification is important because with a non-pure crossover one can not understand and effectively control how much they influence in the evolutionary process. Hence, one has to be careful on the choice of a crossover operator and, preferably, one should choose a Pure Crossover over non-Pure ones. On the theoretical side, Pure Crossovers precisely characterize the class of crossovers that will act exclusively as a recombinative force which must be distinguished and studied separately from other forces such as selection and mutation. This is very important given that the study of the dynamics of selection and/or mutation is relatively easy compared to recombination. If one gets good knowledge about them separately, then when used together, one should be able to distinguish the effects that must be caused by recombination or by its interactions with other operators. If a non-Pure Crossover is used one would see effects of this operator that are similar to either selection or mutation besides the effects of the selection or mutation operators and that would be confusing. With the proposed definition of Pure crossovers one is able to classify existing implementations of crossovers and design new ones taking into account their very true purpose, mixing genetic material, and also to get rid of any effect that could be accomplished by either selection or mutation in a controlled way.

In future work we are planning to formally generalize a form of Geiringer's theorem for finite populations that is representation-independent. We already have all the necessary formal definitions to do it but further assumptions on the bipartition structure used may need to be taken into account. Although intuition leads to think it is possible, the proof may not be trivial.

References

1. Geiringer, H. On the Probability Theory of Linkage in Mendelian Heredity. Annals of Mathematical Statistics, 15:25-27. 1994.
2. C.R. Stephens. Some exact results from a coarsed grained formulation of genetic dynamics. In Proceedings of the genetic and evolutionary computation conference (GECCO 2001), pages 631-638, San Francisco California, USA, 7-11 July 2001.
3. Moraglio, A. and Poli, R. Topological interpretation of crossover. Proceedings of the Genetic and Evolutionary Computation Conference, 1377-1388. 2004.
4. Moraglio, A. and Poli, R. Inbreeding properties of geometric crossover and non-geometric recombinations. Proceedings of the workshop on the Foundations of Genetic Algorithms, 1-14. 2007.
5. Moraglio, A. Towards a geometric unification of evolutionary algorithms. PhD thesis, University of Essex. 2007.