# On the Influence of the Representation Granularity in Heuristic Forma Recombination

Carlos Cotta
Dept. Lenguajes y CC.CC, Univ. of Málaga
Campus de Teatinos (2.2.A.6)
29071 – Málaga (SPAIN)
+34 952 13 3310
ccottap@lcc.uma.es

José Mª Troya
Dept. Lenguajes y CC.CC, Univ. of Málaga
Campus de Teatinos (2.2.A.6)
29071 – Málaga (SPAIN)
+34 952 13 1398
troya@lcc.uma.es

## ABSTRACT

This work studies a mechanism for reducing the computational cost of heuristic recombination operators. The mechanism is based on adjusting the size of the macro-formae processed during recombination. Using the Dynastically Optimal Forma recombination framework as a test model, it is shown that the computational cost can be dramatically reduced. Moreover, intermediate granularities seem to provide the best tradeoff between the computational cost and the quality of the results.

## Keywords

Genetic Algorithms, Heuristic Recombination, Scalability.

## 1. INTRODUCTION

Although evolutionary algorithms were originally devised as robust and general techniques, the necessity of incorporating problem dependent knowledge has been shown to be a strong requirement both in theory [6,14], and practice [4]. These problem-adapted algorithms are usually termed *hybrid* algorithms. As it can be seen in most application-oriented scientific papers, there exists a plethora of mechanisms for carrying out hybridization. One of the most popular techniques is the integration of a heuristic domain-dependent algorithm within the evolutionary algorithm, e.g., a local search procedure [8,9], a construction heuristic [13], an exhaustive search algorithm [1], etc.

One of the most common problems of using these hybrid approaches is the increased computational cost of the resulting algorithms (for example, consider the case of many *smart* recombination operators that build locally optimal offspring). This work studies a mechanism for alleviating this problem, based on tuning the so-called granularity of the representation. The DOR framework [3] is used as a test field for this mechanism. This choice has been made because of the strong heuristic behavior of DOR and its sensitiveness to the problem dimensionality. Nevertheless, the principles described can be generalized to other heuristic procedures.

The remainder of the article is organized as follows: first, some basic concepts and background information on the DOR operator are given in Sect. 2. Subsequently, the functioning of the operator is detailed in Sect. 3, introducing the concept of granularity of the representation and evaluating its influence in the computational cost of the algorithm. Next, experimental results regarding the effects of modifying the granularity are reported in Sect. 4. Finally, some conclusions are presented in Sect. 5, outlining future work as well.

## 2. DYNASTICALLY OPTIMAL FORMA RECOMBINATION

This section is intended to provide some background information on the functioning of the DOR operator. First of all, let $\Xi = \{\psi_1, \cdots, \psi_n\}$ be a set of $n$ independent equivalence relations defined over the search space $\mathcal{S}$ such that

$$\forall x \in \mathcal{S}\ \forall y \in \mathcal{S}\ \exists i\ (1 \leq i \leq n):\ \psi_i(x, y) = 0\ , \qquad (1)$$

i.e., no pair of solutions share membership to the same equivalence classes for all equivalence relations in $\Xi$. In this case, $\Xi$ is said to *cover* the search space, and each solution $x \in \mathcal{S}$ can be univocally represented as $x = \{\eta_1, \cdots \eta_n\}$ where $\eta_i$ is the equivalence class (for the sake of simplicity, the same symbol used to denote an equivalence class is utilized for labeling it) to which $x$ belongs under $\psi_i$. Each of these equivalence classes is termed a *basic forma* [10].

Now, let $x = \{\eta_1, \cdots \eta_n\}$ and $y = \{\zeta_1, \cdots \zeta_n\}$ be two feasible individuals. The *dynastic potential* of $x$ and $y$ is defined as

$$\Gamma(\{x, y\}) = \bigcap_{1 \leq i \leq n} (\eta_i \cup \zeta_i)\ , \qquad (2)$$

i.e., the set of individuals that can be constructed using nothing but the information contained in $x$ and $y$. Next, let $X$: $\mathcal{S} \times \mathcal{S} \times \mathcal{S} \rightarrow [0,1]$ be a stochastic recombination operator (where $X(x,y,z)$ is the probability of generating $z$ when recombining $x$ and $y$). The *immediate dynastic span* of $x$ and $y$ is

$$\Gamma_X^1(\{x, y\}) = \{\ z \mid X(x, y, z) > 0\ \}. \qquad (3)$$

If $\Gamma_X^1(\{x, y\}) \subseteq \Gamma(\{x, y\})$, $X$ is said to be a *transmitting* operator. An example of such an operator is the Random Transmitting Recombination operator [12] defined as

$$\mathrm{RTR}(x, y, z) = \begin{cases} 1/|\Gamma(\{x, y\})| & z \in \Gamma(\{x, y\}) \\ 0 & \text{otherwise} \end{cases}. \qquad (4)$$

Thus, RTR returns a random member of $\Gamma(\{x, y\})$, all individuals having the same probability of being selected, e.g., uniform crossover in binary representations. As stated in [3], this uniform selection is inappropriate if problem-dependent knowledge is available. For example, consider the case in which $\Xi$ is orthogonal, i.e., all combinations of formae induced by different equivalence relations are feasible. In this case, it is clear that

$$\Gamma(\{x, y\}) = \{\, z=\{\xi_1,\cdots\xi_n\} \mid (\xi_i\equiv\eta_i) \vee (\xi_i\equiv\zeta_i) \,\}. \quad (5)$$

Hence, RTR would simply select at random and independently whether $z\in\eta_i$ or $z\in\zeta_i$. However, if the formae are epistatic, the contribution of each forma $\eta$ to the fitness of an individual $x\in\eta$ depends on the other formae to which $x$ belongs. Hence, RTR may reduce to macromutation due to the fact that the context in which a certain forma is immersed is very likely to be disrupted.

The Dynastically Optimal Forma Recombination Operator (DOR) tries to solve this problem by considering partial knowledge about the fitness function. To be precise, let $\phi: \mathcal{S} \rightarrow \mathbb{R}$ be the fitness function, using the notation $\phi[\mathcal{R}]$ to represent the image of $\mathcal{R}\subseteq\mathcal{S}$ under $\phi$. Assume that a partial order relation $\prec$ is defined over $\phi[\mathcal{S}]$, such that $x$ is better than $y$ if, and only if, $\phi(x) \prec \phi(y)$. Then, DOR is defined as a transmitting recombination operator for which

$$\mathrm{DOR}(x, y, z) > 0 \Rightarrow \phi(z) \in \mathbf{sup}_{\prec}(\phi[\Gamma(\{x,y\})]) \quad (6)$$

holds. Hence, DOR returns the best individual (or one of the best individuals) of the dynastic potential. Notice that, ideally, it would be desirable that DOR defined a uniform probability distribution over all members of the supreme set $\mathbf{sup}_{\prec}(\phi[\Gamma(\{x,y\})])$. Nevertheless, this is difficult to achieve in many situations for practical reasons. As a matter of fact, it may be hard even finding a single member of this set, as shown in next section.

# 3. REPRESENTATION GRANULARITY

According to the definition of the DOR operator given in Sect. 2, an exhaustive search must be performed in the dynastic potential $\Gamma(\{x,y\})$ in order to determine the returned child. For this purpose, the most efficient option is to use an A*-like mechanism for incrementally constructing solutions. This mechanism is described below in more detail, introducing the concept of granularity of the representation.

## 3.1 The Internal Functioning of DOR

Let $\Psi_i^j$, $0 \le i \le n$, $2^i \le j \le 2^{i+1}-1$, represent a forma of order $i$, i.e., a partially specified solution whose membership to $i$ equivalence classes is specified (or, alternatively, the set of all solutions that belong to the currently specified equivalence classes). Initially, $\Psi_0^1 = \mathcal{S}$; subsequently,

$$\Psi_{i+1}^{2j} = \Psi_i^j \cap \Sigma(\Psi_i^j, x, y), \quad \text{and} \quad (7)$$

$$\Psi_{i+1}^{2j+1} = \Psi_i^j \cap \Sigma(\Psi_i^j, y, x) \quad (8)$$

are considered, where the *construction units* $\Sigma(\Psi_i^j, w, u)$ are defined as

$$\Sigma(\Psi_i^j, w, u) = \bigcap_{1\le k\le g} \xi_k \,, w \in \xi_k, \quad (9)$$

i.e., the intersection of some of the basic formae $w$ belongs to (the purpose of the third parameter $u$ will be seen in the next subsection). The parameter $g$ is termed *granularity of the representation*. Now, it is necessary to use a monotonic function $\phi^*: \mathcal{P}(\mathcal{S}) \rightarrow \mathbb{R}$ verifying that

$$1.\ \forall x\in\mathcal{S}: \phi^*(\{x\}) = \phi(x), \quad (10)$$

$$2.\ \forall s\in\phi[\mathcal{S}]: s \prec \phi^*(\{\}), \quad (11)$$

$$3.\ \forall\mathcal{R}\subseteq\mathcal{S}\ \exists r\in\phi[\mathcal{R}]: r \prec \phi^*(\mathcal{R}), \text{ and} \quad (12)$$

$$4.\ \forall\mathcal{R}\subseteq\mathcal{S}\ \exists\mathcal{T}\subset\mathcal{R}: \phi^*(\mathcal{T}) \prec \phi^*(\mathcal{R}) \,. \quad (13)$$

Actually, $\phi^*$ need not be defined over all arbitrary sets of solutions but just on sets defined by the intersection of the basic formae induced by $\Xi$. This function comprises the available knowledge about the fitness function $\phi$, proportionating optimistic estimations of partially specified solutions. These estimations are used to determine the order in which the formae $\Psi_i^j$ are generated. They are also used to discard those formae for which $\phi_{\text{best}} \prec \phi^*(\Psi_i^j)$, where $\phi_{\text{best}}$ be the fitness of the best-so-far generated solution. This value is updated whenever a forma $\Psi_{i=n}^j$ such that $\phi^*(\Psi_{i=n}^j) \prec \phi_{\text{best}}$ is generated. Initially $\phi_{\text{best}} \in \mathbf{inf}_{\prec}\{\phi[\mathcal{S}]\}$).

## 3.2 Basic Granularity of the Representation

The structure of the construction units (and hence the granularity parameter $g$) is determined by the characteristics of the representation. The simplest scenario is that in which the representation is orthogonal. In this case, and according to Eq. (5), the dynastic potential $\Gamma(\{x, y\})$ is the Cartesian product of all pairs $\{\eta_i, \zeta_i\}$, $x\in\eta_i$, $y\in\zeta_i$, $1 \le i \le n$. For this reason, it is possible to extend any partially specified solution by considering a single basic forma at a time, i.e.,

$$\Sigma(\Psi_i^j, w, u) = \Sigma(\Psi_i^j, \{\xi_1, \cdots,\xi_n\}, u) = \xi_{i+1} \,. \quad (14)$$

This small granularity ($g$=1) is not always possible if the representation is non-orthogonal. As an example, consider the position-based representation of permutations [2,5]. In this representation $\Xi = \{\psi_1, \cdots, \psi_n\}$, where $\psi_i(x, y) = 1$ if, and only if, $x$ and $y$ have the same element in the $i$th position. Each of these equivalence relations $\psi_i$ induce $n$ equivalence classes $\eta_{ij}$, each one comprising all solutions in which $j$ occurs in the $i$th position. Clearly, using the construction units defined in Eq. (14) would be a waste of computational resources since many of the generated formae would consider repeated elements and hence would be empty (i.e., infeasible). Moreover, even when a particular forma $\Psi_i^j$ were not infeasible per se, it might happen that $\Psi_i^j \cap \Gamma(\{x, y\}) = \{\}$. This would be detected when at a further step both $\Sigma(\Psi_i^j, x, y)$ and $\Sigma(\Psi_i^j, y, x)$ were empty for all descendants of $\Psi_i^j$. All the computational effort needed to extend $\Psi_i^j$ to $\Psi_{i'}^{j'}$ would have been useless.

In the above-mentioned case, construction units must be more complex. To be precise, let the *compatibility set* of a basic forma $\eta$ ($x \in \eta$) be defined as

$$K(\Psi_i^j, \eta, x, y) = \bigcap \eta_k \,, \Psi_i^j \cap \Gamma(\{x, y\}) \cap \eta \cap \overline{\eta}_k = \{\}, \quad (15)$$

i.e., the intersection of all formae $\eta_k$ ($x\in\eta_k$) that must be included along with $\eta$ to preserve feasibility within the dynastic potential. Then, the construction units can be built as

$$\Sigma(\Psi_i^j, w, u) = K(\Psi_i^j, \xi, w, u) \,, \quad (16)$$

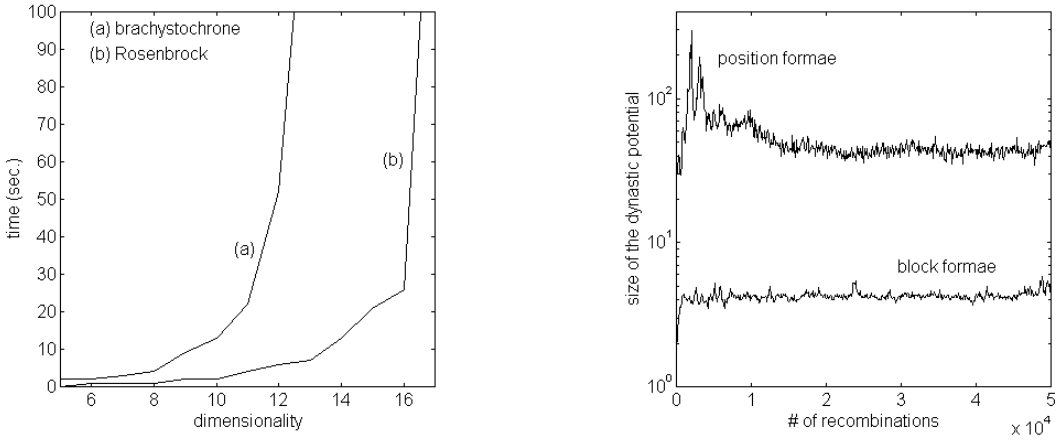


**Figure 1. (Left)** Time required for performing 100 recombinations of random individuals for increasing dimensionalities of two epistatic problems (the design of a brachystochrone and the Rosenbrock function). **(Right)** Evolution of the size of the dynastic potential for two different representations on a 1-EMP instance of 100 elements. Notice the use of a logarithmic scale.

where $\xi$ ($w \in \xi$) is a forma not yet considered in $\Psi_i^j$. Considering again the example of position formae, compatibility sets are recursively defined as follows:

$$\eta_{ab} \supseteq K(\Psi_i^j, \eta_{ab}, x, y) \quad (17)$$

$$\eta_{cd} \supseteq K(\Psi_i^j, \eta_{ab}, x, y),\ y \in \eta_{ce}, x \in \eta_{fe} \Rightarrow \eta_{fe} \supseteq K(\Psi_i^j, \eta_{ab}, x, y). \quad (18)$$

These compatibility sets are termed *cycles*, and constitute the basic units that can be transferred from parents to offspring in order to have a transmitting recombination of position formae. Notice that the user does not have explicit control on the size of these cycles.

## 3.3 Granularity and the Complexity of DOR

The granularity of the representation has an unquestionable impact on the time complexity of the algorithm. Consider that the size of the dynastic potential is $O(2^m)$, where $m$ is the number of construction units that can be identified in the selected individuals. Thus, in the case of orthogonal representations, the size of the dynastic potential is $|\Gamma(\{x, y\})| = 2^n$ since each construction unit comprises a single forma. Obviously, this size is only relevant when epistasis is involved. Otherwise, DOR must simply scan the recombined individuals, selecting the formae that individually optimize the fitness function. Hence it has linear-time complexity.

In the case of epistatic representations, the exponential growth of the dynastic potential becomes very important. As an example, Figure 1 (left) shows the time required for performing one hundred recombinations of randomly generated individuals as a function of the dimensionality of the problem. To be precise, the design of a brachystochrone [7] and the Rosenbrock function have been used[1]. As it can be seen, the computational cost grows extremely fast for dimensionalities above 10 in the first problem and 16 in the second one, thus confirming the influence of the number of construction units. Of course, there exist other factors affecting the time complexity of the operator, namely the amount of knowledge used in the $\phi^*$ function. If this function were defined to return a constant value for underspecified solutions (i.e., if no problem-knowledge were used), DOR would reduce to an exhaustive enumeration of all members of the dynastic potential and hence $2^n$ individuals would have to be evaluated. On the contrary, if $\phi^*$ provides good estimations of the fitness function this quantity can be dramatically reduced, so as to become affordable for a certain range of dimensionalities.

It is interesting to notice that, as the algorithm converges, individuals tend to be more similar and hence their dynastic potential is reduced. To be precise, suppose that the individuals to be recombined share membership to a set of basic formae $\Theta = \theta_{i1} \cap \cdots \cap \theta_{im}$. If the representation is *separable*[2] (i.e., it is possible to create a child $z \in \theta \cap \zeta \cap \xi$ provided that $x \in \theta \cap \zeta$, and $y \in \theta \cap \xi$, where $\theta$, $\zeta$ and $\xi$ are any three basic formae), it is clear that $\Gamma(\{x,y\}) \subseteq \Theta$. The search can be subsequently started from $\Psi_0^i = \Theta$, thus, reducing the search space to, at most, $2^{n-m}$ solutions.

Nevertheless, it must be taken into account that, even when the above consideration is true, it does not necessarily imply that the number of construction units in non-orthogonal representations decrease as well. As an example consider the permutations 12345678 and 23156784. From the point of view of position formae, they do not belong to any common basic forma, existing two compatibility sets in each permutation: {123■■■■■} and {■■■45678} from the first one and {231■■■■■} and {■■■56784} from the second

[1] These two problems are defined over real-valued variables. Different definitions of formae are possible in this case (e.g., see [11]. The equivalence relations $\psi_i(\vec{\mathbf{x}}, \vec{\mathbf{y}}) = 1 \Leftrightarrow (x_i = y_i)$ have been used in this work.

[2] Non-separable representations (e.g., the edge-based representation of permutations) are difficult to handle for several reasons (see [2]). Thus, they are deferred to a further work.

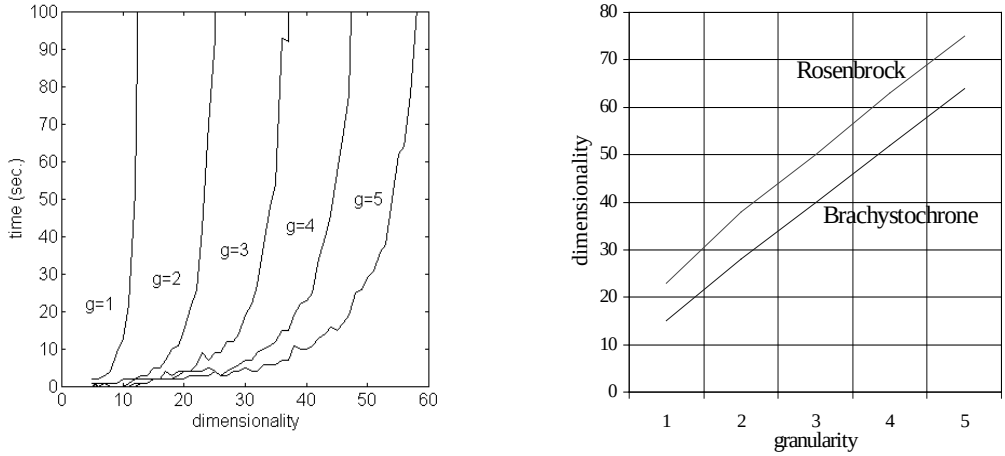


**Figure 2 (Left)** Times required for performing 100 recombinations of random individuals in the brachystochrone design problem. The results are shown for different granularities. **(Right)** Lowest dimensionalities for which the time required for performing 100 recombinations of random individuals is greater than 30 seconds.

one. Thus, there are $2^2$ solutions in their dynastic potential. On the other hand, the permutations 12345678 and 13246587 belong to two common position formae. However, there exist three compatibility sets in each of them: {■23■■■■■}, {■■■■56■■} and {■■■■■■78} in the first one and {■32■■■■■}, {■■■■65■■} and {■■■■■■87} in the second one. Hence their dynastic potential has $2^3$ solutions.

This effect is illustrated in Figure 1 (right), in which the evolution of the number of non-trivial construction units considered during recombination is shown for a 1-EMP instance of 100 elements (the *k*-EMP is a generalization of the min-permutation problem presented in [3], in which *k* indicates the degree of epistasis). As it can be seen, the size of the dynastic potential of recombined solutions grows during the initial stage of the search, reaching a maximum after a few thousands of recombinations, decreasing and stabilizing from that point.

## 3.4 Tuning the Representation Granularity

As discussed in the previous subsection, the size of the dynastic potential plays an important rôle in the computational cost of DOR. If this cost becomes prohibitive for the dimensionality of the problem under consideration, it may be appropriate to modify the granularity of the representation so as to reduce the number of solutions comprised in the dynastic potential.

First of all, consider the case of orthogonal representations. The granularity may be increased by making construction units contain more than one basic forma. These formae can be selected at random or using a priori knowledge about the epistatic relations of these formae. To be precise, in the two orthogonal problems considered, there exist epistatic relations between adjacent variables. For this reason, it is convenient to choose consecutive variables, thus preserving a larger part of their context. Hence,

$$\Sigma(\Psi_i^j, w, u) = \Sigma(\Psi_i^j, \{\xi_1, \cdots \xi_n\}, u) = \bigcap_{1 \le k \le \min(g, n-i)} \xi_{i+k} \quad . \quad (19)$$

Figure 2 (left) shows the results of the same benchmark depicted in Figure 1 (left). As it can be seen, when the granularity is increased, the algorithm reduces its computational cost, being capable of tackling larger problem instances. Moreover, and as it can be seen in Figure 2 (right), there exists a very good linear relation between the granularity of the representation and the highest affordable dimensionality of the problem when a constant computational cost is kept.

A similar approach can be taken when the representation is non-orthogonal although, in this case, compatibility sets rather than single basic formae must be joined. This is the underlying idea of the *block* representation [2]. A block is a macro-compatibility-set that can be defined as the closure of the following expressions:

$$x \in \eta_{ka} \Rightarrow [K(\Psi_i^j, \eta_{ka}, x, y) \supseteq B(\Psi_i^j, k, x, y)] \quad (20)$$

$$x \in \eta_{sb}, (\eta_{rc} \cap \eta_{td}) \supseteq B(\Psi_i^j, k, x, y), \; r < s < t \Rightarrow$$
$$\Rightarrow [K(\Psi_i^j, \eta_{sb}, x, y) \supseteq B(\Psi_i^j, k, x, y)]. \quad (21)$$

Figure 1 (right) illustrates the evolution of the size of the dynastic potential of recombined solutions using the block representation. Notice that, comparing with the canonical position representation, this size is more than one order of magnitude smaller. This implies that the computational cost of DOR is considerably reduced**.**

In any case, consider that reducing the dynastic potential also reduces the chances for combining valuable information taken from the parents. In a extreme situation, this could be as undesirable as the unaffordable computational cost of DOR for low (resp. high) values of *g* (resp. *n*). Next section provides experimental results that shed some light on the effects of tuning the representation granularity.

# 4. EXPERIMENTAL RESULTS

A set of experiments has been realized in order to assess the effects that a manual modification of the granularity of the representation has on the quality of the results. Except where otherwise noted, an elitist generational genetic algorithm

**Table 1.** Mean best fitness for the Rosenbrock function (averaged for 20 runs). The best results are shown in boldface.

| Operator | Dimensionality | | | | |
|---|---|---|---|---|---|
| | 8 | 12 | 16 | 24 | 32 |
| SPX | 14.59 | 20.88 | 21.18 | 103.02 | 442.59 |
| UX | 10.02 | 21.98 | 37.30 | 218.07 | 856.47 |
| AX | 5.91 | 14.38 | 29.80 | 127.29 | 319.26 |
| $R^3$ | 6.60 | 11.26 | 42.01 | 145.14 | 361.42 |
| $DOR_{g=1}$ | **3.44** | **6.36** | **15.53** | - | - |
| $DOR_{g=2}$ | 4.97 | 10.35 | 16.63 | 42.62 | 299.52 |
| $DOR_{g=4}$ | 10.74 | 22.78 | 27.91 | **39.26** | **69.86** |
| $DOR_{g=8}$ | 6.67 | 11.16 | 28.75 | 66.41 | 81.42 |

**Table 2.** Mean best fitness for the Brachystochrone problem (averaged for 20 runs). The best results are shown in boldface.

| Operator | Number of Pillars | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 8 | 12 | 16 | 24 | 32 | 40 | 48 | 56 | 64 |
| SPX | 1.1577 | 1.1788 | 1.2156 | 1.2713 | 1.3908 | 1.6262 | 1.7787 | 2.1049 | 2.5971 |
| UX | 1.1757 | 1.2300 | 1.2862 | 1.3499 | 1.5305 | 1.6771 | 1.9379 | 2.2172 | 2.6658 |
| AX | 1.1560 | 1.1668 | 1.1828 | 1.2219 | 1.2722 | 1.3534 | 1.4841 | 1.6327 | 1.8845 |
| $R^3$ | 1.1540 | 1.1627 | 1.1783 | 1.2149 | 1.2803 | 1.3683 | 1.5295 | 1.7357 | 1.9061 |
| $DOR_{g=1}$ | **1.1530** | **1.1549** | 1.1779 | - | - | - | - | - | - |
| $DOR_{g=2}$ | 1.1544 | 1.1628 | **1.1726** | **1.2110** | - | - | - | - | - |
| $DOR_{g=3}$ | 1.1548 | 1.1798 | 1.1815 | 1.2458 | **1.2660** | 1.3827 | **-** | **-** | **-** |
| $DOR_{g=4}$ | 1.1574 | 1.1901 | 1.1944 | 1.2485 | 1.2971 | **1.3413** | 1.4759 | 2.1635 | - |
| $DOR_{g=5}$ | 1.1601 | 1.1919 | 1.2111 | 1.2494 | 1.2977 | 1.3468 | 1.3886 | 1.5386 | 2.1009 |
| $DOR_{g=6}$ | 1.1649 | 1.2010 | 1.2053 | 1.2598 | 1.3001 | 1.3483 | **1.3695** | **1.4720** | 1.6750 |
| $DOR_{g=7}$ | 1.1729 | 1.1998 | 1.2112 | 1.2635 | 1.2981 | 1.3646 | 1.4073 | 1.4821 | **1.5894** |

(*popsize*=100, $p_c$=.9, *maxevals* = $10^5$) using ranking selection ($\eta^+$=2.0, $\eta^-$=0.0) has been utilized. In order to compare DOR with other classical operators, the internal partial evaluations have been considered. More precisely, computing the optimistic evaluation of a construction unit of granularity $g$ is accounted as $g/n$ evaluations.

The first results (shown in Table 1) correspond to the Rosenbrock function. In this problem, variables have been encoded using 64 bits, and $p_m$ has been set to 1/64. For comparison purposes, the experiments have been realized with some classical operators such as single-point crossover (SPX), uniform crossover (UX), arithmetic crossover (AX) and random respectful recombination ($R^3$) [11]. When using the DOR operator, mutation is always performed before recombination (thus, new material is still introduced in the population but the smart recombination performed by DOR is preserved).

As it can be seen, low values of $g$ are in general better for low dimensionalities of the function. It is interesting to see that, in the case $n$=32, $DOR_{g=1}$ has a prohibitive computational cost, and $DOR_{g=2}$ is less expensive but consumes very quickly the allocated number of epistatic calculations. For these instances, it seems that a ratio $7 \le n/g \le 9$ provides the best results.

The next results correspond to the brachystochrone design problem. As shown in Figure 2 (right), DOR has a comparatively higher computational cost in this problem than for the Rosenbrock function. For this reason, a wider range of granularities and dimensionalities has been tried. In this problem, variables are encoded with 16 bits, $p_m$ is set to 1/16 and *maxevals*=$5 \cdot 10^5$. The results are shown in Table 2. Again, low values of $g$ are better when the dimensionality of the problem is small, becoming the quality of the results slightly worse when the granularity is increased. Also, for large dimensionalities of the function (i.e., above 32), low values of $g$ are either prohibitive or provide worse results. As previously mentioned, this is due to the fact that a higher number of construction units are manipulated and hence the algorithm is prematurely terminated.

Finally, experiments have been done using $k$-EMP instances of different dimensionalities. The first results correspond to 1-EMP instances. As for the previous test problems, other recombination operators for permutations such as PMX, three variants of OX, RCX, UCX, BX and UBX (all described in [2]) have been also used. The results for these operators are shown in Table 3. Notice the poor results of the different variants of order crossover. This is due to the fact that this problem is defined on the basis of position formae rather than precedence formae. For that reason, UCX (*Uniform Cycle Crossover*) provides the best results. It is interesting to notice that this operator is based on a blind interchange of the compatibility sets defined in Eqs. (17) and (18).

Table 4 shows the results of the DOR operator. In this problem, the granularity is tuned by defining a maximum allowed number of construction units $f$. Whenever the number of construction units is greater than $f$, two of them are picked at random and joined. This process is repeated until at most $f$ construction units are available. Compacting cycles as indicated in Eqs. (20) and (21) has also been tried (*blocks*).

As it can be seen, the results for very coarse granularities (i.e., low values of $f$) are worse than for fine granularities. $DOR_{block}$ is included in the former category since, as shown in Figure 1 (left), its equivalent $f$ value is never greater than 4 on average. As a matter of fact, the performance of $DOR_{block}$ is

**Table 3.** Results of classical operators on 1-EMP problems (averaged for 20 runs). The best results are shown in boldface.

| | Operator | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| # of elements | OX#1 | OX#2 | OX#3 | PMX | RCX | UCX | BX | UBX |
| 100 | 2534 | 2809 | 2375 | 795 | 911 | **647** | 963 | 945 |
| 125 | 4412 | 4667 | 4046 | 1373 | 1598 | **1066** | 1715 | 1662 |
| 150 | 6803 | 7571 | 6502 | 2114 | 2537 | **1660** | 2637 | 2665 |
| 200 | 13607 | 14809 | 12845 | 4317 | 5244 | **3468** | 5364 | 5423 |

**Table 4.** Results of the DOR operator on 1-EMP problems (averaged for 20 runs). The best results are shown in boldface.

| | | Number of construction units allowed (*f*) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| # of elements | *blocks* | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | ∞ |
| 100 | 572 | 773 | 346 | 319 | 316 | 315 | 302 | 312 | **297** | 319 | 317 |
| 125 | 1012 | 1318 | 533 | 482 | 478 | **454** | 481 | 494 | 466 | 463 | 478 |
| 150 | 1595 | 2151 | 773 | 691 | 668 | 670 | 657 | 675 | **648** | 679 | 672 |
| 200 | 3245 | 4293 | 1370 | 1193 | 1135 | **1116** | 1140 | 1142 | 1126 | 1175 | 1171 |

**Table 5.** Results of the DOR operator on *k*-EMP problems (averaged for 20 runs). The best results are shown in boldface.

| | *k* = 2 | | | | *k* = 5 | | | | *k* = 10 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *f* | 100 | 125 | 150 | 200 | 100 | 125 | 150 | 200 | 100 | 125 | 150 | 200 |
| 2 | 1114 | 1959 | 3099 | 6358 | 3193 | 5693 | 9007 | 18747 | 14847 | 27396 | 44099 | 94567 |
| 3 | 481 | 750 | 1075 | 2106 | 1295 | 2081 | 3261 | 6184 | 6397 | 10114 | 16183 | 33022 |
| 4 | 407 | 634 | 912 | 1614 | **994** | 1655 | 2452 | 4706 | 4878 | 8046 | 12622 | 23733 |
| 5 | 410 | **616** | 939 | 1537 | 1065 | 1625 | **2278** | 4195 | **4292** | **7721** | 11771 | 22875 |
| 6 | **406** | 633 | **898** | **1534** | 1023 | **1591** | 2415 | **4112** | 4754 | 8140 | **11717** | **21986** |
| 7 | 427 | 633 | 915 | 1566 | 1239 | 1710 | 2387 | 4395 | 4887 | 8381 | 12053 | 22266 |
| 8 | 444 | 653 | 961 | 1577 | 1074 | 1729 | 2557 | 4736 | 4667 | 8469 | 12339 | 23024 |
| 9 | 432 | 644 | 930 | 1622 | 1081 | 1747 | 2483 | 4819 | 4932 | 8632 | 13565 | 23958 |
| 10 | 416 | 634 | 955 | 1644 | 1181 | 1866 | 2661 | 4858 | 5252 | 9073 | 14210 | 25823 |
| ∞ | 431 | 681 | 1482 | 7282 | 1179 | 2416 | 13783 | 58132 | 7958 | 34085 | 139033 | 381955 |

intermediate between *f*=2 and *f*=3. Notice that the performance is stabilized around *f*=6, with a slight tendency to decrease for higher values. In any case, the quality of the results is always better than UCX (with the exception of the extreme situation *f*=2).

To confirm these results, further experiments with the DOR operator have been realized with higher degrees of epistasis (*k*=2, 5, and 10). The results are shown in Table 5. In these experiments, the behavior of the algorithm is clearer. First, notice that the quality of the results is improved when *f* is increased up from *f*=2 (this is also depicted in Figure 3, left). The best results are achieved in the intermediate values *f*=5 and *f*=6, existing a soft degradation of performance for higher values of *f*. The reason for this behavior has already been mentioned: low values of *f* reduce the chances for transferring information from parents to offspring, while high values of *f* quickly consume the allocated number of epistatic calculations. This bowl-like shape is illustrated in Figure 3 (right).

# 5. CONCLUSIONS

This work has studied the functioning of a heuristic recombination operator (DOR) on the basis of the construction units used for creating new solutions. To be precise, the size of these construction units (i.e., the so-called *granularity* of the representation) has been considered as a central factor for determining the computational cost of the operator and the quality of the results it provides.

It has been shown that there exists a basic (ground-level) granularity for each representation. This basic granularity is minimal in the case of orthogonal representations, in which basic formae can be freely combined. However, the basic granularity is variable when dealing with non-orthogonal separable representations. In this situation, it depends upon the size of the basic transference units (i.e., the *compatibility sets*), and it is not under direct control of the user.

There exists an empirically corroborated relationship between the granularity (and hence the size of the dynastic potential of the solutions to be recombined) and the computational cost of the DOR operator. Thus, it has been proposed to increase the granularity factor (i.e., use larger construction units) to reduce this cost.

An extensive experimental investigation has been done in order to assess the influence that this modification of the basic granularity has on the quality of the results provided by the algorithm. The results have been satisfactory, since it has been shown that the performance is significantly degraded only for extremely larger construction units. Moreover, intermediate granularities provide better results than very fine granularities since the former consume less computational resources (and hence the algorithm can be executed for a larger number of iterations).

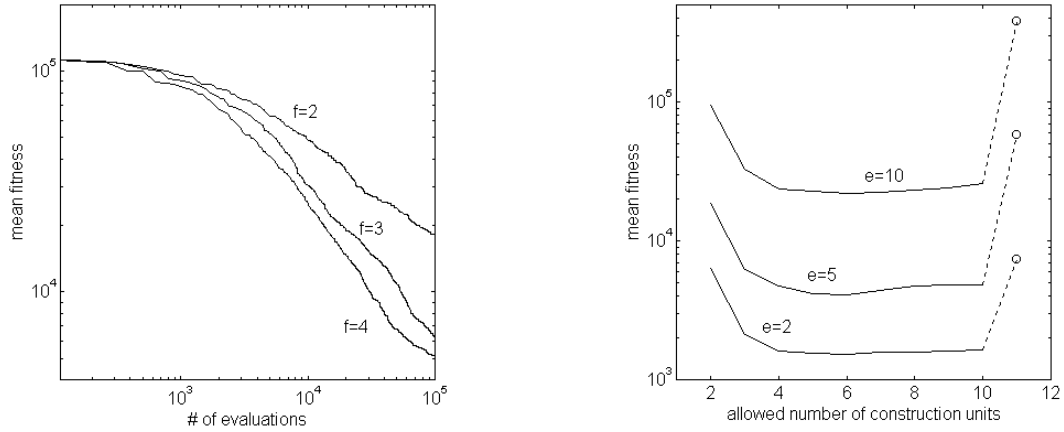


**Figure 3 (Left)** Comparison of the evolution of fitness in a 5-EMP instance of 200 elements for different numbers of construction units. **(Right)** Results of DOR for 2-EMP, 5-EMP and 10-EMP instances of 200 elements as a function of the number of allowed construction units. The rounded point corresponds to the case $f = \infty$. Notice the use of a logarithmic scale in both graphs.

This work admits several extensions. First of all, and as mentioned in Sect. 1, DOR has been chosen as test model because it is a strongly heuristic operator and it is sensitive to the dimensionality of the problem considered. However, the idea of granularity adjustment can be applied to other heuristic techniques, e.g., embedded hill-climbers. In this sense, validating the conclusions of this work in the context of other hybrid models constitutes a very interesting line of research in which work is already in progress.

On the other hand, extending these results to non-separable representations is a line of future work too. Among other problems, it is difficult to determine the compatibility sets for this kind of representations. Furthermore, common formae cannot be imposed to the child. For these reasons, it is not completely clear how to increase the granularity in this case.

## 6. ACKNOWLEDGMENTS

This work is supported by the *Comisión Interministerial de Ciencia y Tecnología* under grant TIC99-0754-C03-03.

## 7. REFERENCES

[1] Cotta C., Aldana J.F., Nebro A.J., Troya J.M., "Hybridizing Genetic Algorithms with Branch and Bound Techniques for the Resolution of the TSP", *Artificial Neural Nets and Genetic Algorithms 2* (pp. 277-280), Pearson D.W., Steele N.C., Albrecht R.F. (eds.), Springer-Verlag, Wien, 1995

[2] Cotta C., Troya J.M., "Genetic Forma Recombination in Permutation Flowshop Problems", *Evolutionary Computation* 6(1):25-44, 1998

[3] Cotta C., Alba E., Troya J.M., "Utilizing Dynastically Optimal Forma Recombination in Hybrid Genetic Algorithms", *Parallel Problem Solving From Nature - PPSN V* (pp. 305-314), Eiben A.E., Bäck Th., Schoenauer M., Schwefel H.-P. (eds.), Lecture Notes in Computer Science 1498, Springer-Verlag Berlin, 1998

[4] Davis L., *Handbook of Genetic Algorithms*, Van Nostrand Reinhold Computer Library, New York , 1991

[5] Goldberg D.E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading MA, 1989

[6] Hart W.E., Belew R.K., "Optimizing an arbitrary function is hard for the genetic algorithm", *Proceedings of the Fourth International Conference on Genetic Algorithms* (pp. 190-195), Belew R.K., Booker L.B. (eds.), Morgan Kaufmann, San Mateo CA, 1991

[7] Herdy M., Patone G., "Evolution Strategy in Action: 10 ES-Demonstrations", Technical Report TR-94-05, Technische Universität Berlin, 1994

[8] Moscato P., "On Evolution, Search, Optimization, Genetic Algorithms and Martial Arts: Towards Memetic Algorithms", Caltech Concurrent Computation Program, C3P Report 826, 1989

[9] Mühlenbein H., "Parallel genetic algorithm, population dynamics and combinatorial optimization", *Proceedings of the Third International Conference on Genetic Algorithms* (pp. 416-421), Schaffer J.D. (ed.), Morgan Kaufmann, San Mateo CA, 1989

[10] Radcliffe N.J., "Equivalence Class Analysis of Genetic Algorithms", *Complex Systems* 5:183-205, 1991

[11] Radcliffe N.J., "Forma Analysis and Random Respectful Recombination", *Proceedings of the Fourth International Conference on Genetic Algorithms* (pp. 222-229), Belew R.K., Booker L.B. (eds.), Morgan Kaufmann, San Mateo CA, 1991

[12] Radcliffe N.J., "The Algebra of Genetic Algorithms", *Annals of Mathematics and Artificial Intelligence* 10: 339-384, 1994

[13] Storer R.H., Wu S.D., Vaccari R., "New search spaces for sequencing problems with application to job shop scheduling", *Management Science* 38:1495-1509, 1992

[14] Wolpert D.H., Macready W.G., "No Free Lunch Theorems for Search", Sante Fe Institute Technical Report SFI-TR-95-02-010, 1995