Symmetrical Building Blocks and the Simple Inversion Operator

Cristian Munteanu
Laseeb-ISR-Instituto Superior Técnico
Universidade Técnica de Lisboa
Av. Rovisco Pais, 1- Torre Norte 6.21,
1049-001 Lisboa, Portugal
cmunteanu@pop.isr.ist.utl.pt

Agostinho Rosa
Laseeb-ISR-Instituto Superior Técnico
Universidade Técnica de Lisboa
Av. Rovisco Pais, 1- Torre Norte 6.20,
1049-001 Lisboa, Portugal
acrosa@isr.ist.utl.pt

In this paper we analyze a Simple Inversion Operator (SIO), showing that a Genetic Algorithm (GA) with a reordering operation is capable of attaining very good performance on problems compliant with the Symmetrical Building Block Paradigm (SBB). Our approach is different from the classical view of inversion benefits, in that it is not concerned with increasing linkage between different genes, but rather aims at the direct discovery of building blocks with a specific symmetry. Inversion together with crossover works to properly discover and align the building blocks, to form the optimal solution. The main advantages of our approach are the application with good results of a fairly simple and unpretentious operator, and the clear identification of the class of problems where this operator has the highest potential.

Definition 1. The SBB of length δ, at position q along the chromosome, will be denoted as a string SBB_q^δ :

$$\mathrm{SBB}_q^\delta = \#_1 \#_2 \ldots \#_{q-1} a^1 a^2 \ldots a^i \ldots a^{\delta-1} a^\delta \#_{q+\delta-1} \ldots \#_l$$

$$q = j \cdot \delta + 1 \text{ for } j \in \left\{0, 1, \ldots, \frac{l-\delta}{\delta}\right\}, a^i \in \{0, 1, \#\} \text{ for } i \in \{1, \ldots, \delta\}$$

$$a^i = a^{\delta-i+1} \text{ for } i \in \left\{1, \ldots, \frac{\delta}{2}\right\} \text{ if } \delta = 2r \text{ or } i \in \left\{1, \ldots, \frac{\delta-1}{2}\right\} \text{ if } \delta = 2r+1,\ r \in \mathbf{N}^*$$

where l is the length of the chromosome, divisible by δ, # is a "don't care" value in {0, 1}. We call the sequence of all a^i with $i \in \{1, \ldots, \delta\}$ the kernel of the SBB.

Definition 2. The optimum string (SBB-compliant) is a string x_{opt} that is formed by juxtaposing SBB_q^δ with δ fixed, having the following structure $x_{\mathrm{opt}} = A_1 A_2 \ldots A_{l/\delta}$, $A_1 \equiv A_2 \equiv \ldots \equiv A_{l/\delta} = a^1 a^2 \ldots a^{\delta-1} a^\delta$. Under the assumption of single-string selection on a population containing a single instance of a SBB (SBB^δ_i) we derived a lower nonzero bound on the probability of constructing a *new* SBB (SBB^δ_j): $P(t)_{\mathrm{SBB}_j^\delta} \geq \frac{2 \cdot f_{\mathrm{SBB}_i^\delta}}{(1+l) \cdot l \cdot N \cdot \hat{\mu}(t)} P_{\mathrm{SIO}} \left(1 - P_{\mathrm{c}} \frac{\Delta}{l-1}\right) (1 - P_m)^o$ where N is the size of population, $f_{\mathrm{SBB}_i^\delta}(t)$ is the fitness of the existing SBB, Δ is the actual defining length of the SBB, o is the order of the SBB, $\hat{\mu}(t)$ is the mean fitness of the individuals in the population, while P_{SIO}, P_{c}, P_{m} are the genetic operators' rates. Furthermore, we have found an upper bound on the constructive capabilities of SIO in conjunction with the other genetic operators (selection, one-point crossover, mutation). That is, SIO can determine at most the doubling of SBBs present in a string. We analyzed all the potential inversions that can lead to this event. We carried out experimental results with a GA and SIO (SIOGA) on a set of test problems compliant with the SBB paradigm, as previously defined. The first function is a Royal Road Function (R1), as originally defined in (Mitchell, 1991). The second (SBBF1) and third test functions (SBBF2) are implemented directly, according to the definition of SBB-compliance. The functions differ in the order o of the SBBs. The fourth function (FUNC) is a problem that is not apparent to be SBB-compliant. The interesting part with this function, is the fact that the SBB-compliance becomes apparent only when a certain coding of the chromosome is used. The fifth and sixth functions, are also Royal Road functions, but this time we used the revised version defined by Holland (level 3, and 4). The last two functions tested are two versions of a 4-order fully deceptive problem. Results obtained on the whole test set clearly indicate the efficiency of applying SIO, comparison with GA showing for some functions an increase in speed and accuracy of optimum finding up to one order of magnitude bigger in the case of SIOGA. The sheer efficiency of SIO stems from the way it operates on SBBs. We experimentally shown that SIOGA performs very good, even on problems where the optimum consists of a partial sequence of identical SBB or the optimum does not have a perfect symmetry, and this was the case of the Royal Roads level 3 and 4. We have also shown how coding can turn a problem that was apparently not consistent with the SBB paradigm (i.e., the FUNC function), into a problem that was perfectly SBB-compliant.

References

M. Mitchell, S. Forrest, and J. Holland (1991). The Royal Road for Genetic Algorithms: Fitness Landscapes and GA Performance. In J. Varela, and P. Bourgine (eds.), *Proceedings of the First European Conference on Artificial Life,* 245-254. Cambridge, MA: MIT Press.