Continuous Variable Neighbourhood Search Algorithm
Based on Evolutionary Metaheuristic Components: A Scalability Test

Carlos García-Martínez
Dept. of Computing and Numerical Analysis
Univ. of Córdoba, 14071 Córdoba, Spain
cgarcia@uco.es

Manuel Lozano
Dept. of Computer Science and Artificial Intelligence
CITIC-UGR (Research Center on Information and Communications Technology)
Univ. of Granada, 18071 Granada, Spain
lozano@decsai.ugr.es

Abstract

Variable Neighbourhood Search is a metaheuristic combining three components: generation, improvement, and shaking components. In this paper, we describe a continuous Variable Neighbourhood Search algorithm based on three specialised Evolutionary Algorithms, which play the role of each aforementioned component: 1) an EA specialised in generating a good starting point as generation component, 2) an EA specialised in exploiting local information as improvement component, 3) and another EA specialised in providing local diversity as shaking component. We adopt the experimental framework proposed for the Special Session on Evolutionary Algorithms and other Metaheuristics for Continuous Optimization Problems - A Scalability Test, for the ISDA’09 conference, to test the ability of the model of being scalable for high-dimensional problems.

1. Introduction

Nowadays, high-dimensional optimisation problems arise as a very interesting field of research, because they appear in many important new real-world problems (bio-computing, data mining, etc.). Unfortunately, the performance of most available optimisation algorithms deteriorates rapidly as the dimensionality of the search space increases.

Metaheuristics (MHs) [11] are a family of search and optimisation algorithms based on extending basic heuristic methods by including them into an iterative framework augmenting their exploration capabilities. MHs coordinate subordinate components (such as probability distributions, tabu lists or genetic operators among others) with the aim of performing an effective and efficient process in search-

ing for the global optimum of a problem. Over the last years, a large number of search algorithms were reported that do not purely follow the concepts of one single classical MH, but they attempt to obtain the best from a set of MHs that perform together and complement each other to produce a profitable synergy from their combination. These approaches are commonly referred to as hybrid MHs [22].

Evolutionary Algorithms (EAs) [3] are stochastic search methods that mimic the metaphor of natural biological evolution. EAs rely on the concept of a population of individuals, which undergo probabilistic operators to evolve toward increasingly better fitness values of the individuals.

A novel method to build hybrid MHs, recently introduced in [19], concerns the incorporation of specialised EAs into classical MHs, replacing determinate components, but preserving the essence of the original MH as much as possible. The idea is to build customised EAs playing the same role as particular MH components, but more effectively, i.e., evolutionary MH components. In this way, a classical MH is transformed into an integrative hybrid MH (because one of its components is another MH). In the literature, we find some proposals that follow this idea: In [1] and [6], two EAs are applied to perform local search processes within a multi-start MH; and in [18, 19], an evolutionary ILS-perturbation technique is presented (ILS is for Iterated Local Search).

Variable Neighbourhood Search (VNS) [21] is a MH that exploits systematically the idea of neighbourhood change (from a given set of neighbourhood structures \(N_k, k = 1, \ldots, k_{\text{max}}\), both in the descent to local minima and in the escape from the valleys which contain them. It mainly consists of the following three components, which work on a single candidate solution, the current solution \(s^c\) (see Fig. 1):

1. Generation component: Firstly, a method is executed to generate \(s^c\) within the search space.
2. Improvement component: Secondly, \( s^c \) is refined, usually by a local search method.

3. Shaking component: Then, shaking is performed to escape from the valley where \( s^c \) lies. It selects a random solution in the \( k \)th neighbourhood of \( s^c \), which becomes a new starting solution for improvement component. At the beginning of the run, and every time last improvement process improved the best found solution, \( k \) is set to one; otherwise, \( k \) is set to \( k + 1 \).

In this work, we study a novel continuous VNS model, presented in [7, 8], that is based on three specialised EAs playing the role of each VNS component. Experiments are carried out to test the ability of the algorithm of being scalable for high-dimensional problems. With this aim, we adopt the experimental framework proposed for the Special Session on Evolutionary Algorithms and other Metaheuristics for Continuous Optimization Problems - A Scalability Test, for the ISDA’09 conference.

In Sect. 2, we describe the general scheme of the VNS model based on evolutionary MH components. In Sect. 3, we introduce the generation component. In Sect. 4, we detail the improvement component. In Sect. 5, shaking component is described. In Sect. 6, we present the empirical study. Conclusions are presented in Sect. 7.

2. VNS based on Specialised EAs

In this study, we focus on a continuous VNS model based on three specialised EAs, which play the role of each VNS component (see Fig. 1):

1. Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [12, 13] generates initial \( s^c \) (Sect. 3).
2. Continuous Local EA improves \( s^c \) (Sect. 4).
3. Micro Cross-generational elitist sel., Heterogeneous recombination, and Cataclysmic mutation (\( \mu \)CHC) performs shaking to escape from the valley where \( s^c \) lies, according to the \( k \)th neighbourhood (Sect. 5).

Adaptation of parameter \( k \) is performed the same way original VNS algorithm does.

Neighbourhoods structures are defined by using metric \( \rho' \):

\[
N_k(s^c) = \{ s \mid r_{k-1} \leq \rho(s^c, s) \leq r_k \}
\]

where \( r_k \) is the radius of \( N_k(s^c) \) monotonically non-decreasing with \( k \). Hereafter, we will always consider the following distance metric because of computational reasons:

\[
\rho(x, y) = \sum_{i=1}^{n} |x_i - y_i|
\]

Bound constraints are checked every time a new solution is sampled. Every value that lies outside the ranges of its corresponding variable is recomputed to its symmetric value with regards to the violated bound. This process is repeated until no bound constraint is violated.

Stop condition is reached when \( k \) surpasses \( k_{\text{max}} \) (in our case, \( k_{\text{max}} \) is set to 20). At this point, the algorithm performs a restart. In this way, exploiting larger number of function evaluations may increase the chance to achieve better function values. In our case, a restart means that VNS is run once again from the CMA-ES execution, generating a new \( s^c \).

3. CMA-ES as Generator Component

The proposal applies CMA-ES [12, 13] to obtain a good point in the design space from which the algorithm carries out its search process. We have applied the C code, available at http://www.lri.fr/~hansen/cmaes_inmatlab.html with the suggested parameter values from a random solution.

4. Continuous Local EA as Improvement Component

Continuous Local EA plays the role of improvement component for the VNS model. It is an EA based on the principles of Binary Local Genetic Algorithm [10], which obtained promising results in efficacy and efficiency against classical improvement methods.

Continuous Local EA is a steady-state real-coded genetic algorithm [14, 15] that inserts one single new member into the population (\( \text{Pop} \)) in each iteration. It uses a replacement method in order to force a member of the current population to perish and to make room for the new offspring. It is important to know that the selected replacement method favours the acquisition of information about
the search space. Then, the algorithm exploits local information around \( s^c \), provided by the individuals in \( Pop \) close to it, to orientate the improvement process.

Let’s suppose that Continuous Local EA is to improve the given \( s^c \). Then, the following steps are carried out (see Fig. 2):

- **Mate selection**: A solution from \( Pop \) is selected as \( s^{mate} \) by means of nearest improving solution selection method (Sec. 4.1).

- **Crossover**: \( s^{mate} \) and \( s^c \) are crossed over by means of parent-centric BLX-\( \alpha \), creating offspring \( z \) (Sec. 4.2).

- **Acceptance and replacement**: If \( z \) is accepted (Sect. 4.3), it replaces \( s^c \) and is inserted into the population forcing the individual with lowest information contribution to perish (Sect. 4.4).

All these steps are repeated until a maximum number of iterations without improving the best visited \( s^c \) is reached (we stop the run after 100 iterations without improving best \( s^c \)). At last, the best found \( s^c \) is returned.

An important aspect when using Continuous Local EA in the proposed VNS model is that \( Pop \) undergoes initialisation only once, at the beginning of the run, and not at every invocation to improve \( s^c \). In this way, the algorithm may gather up valuable information about the search space and its optima, and employ accumulated search experience from previous refinements to enhance future ones.

### 4.1. Nearest Improving Solution Selection

Assortative mating is the natural occurrence of mating between individuals of similar phenotype more or less often than expected by chance. Mating between individuals with similar phenotype more often is called positive assortative mating and less often is called negative assortative mating. Fernandes et al. [5] implement these ideas to design two mating selection mechanisms.

We introduce a new assortative mating mechanism to be used in Continuous Local EA, nearest improving solution selection. Its idea is to select the individual in \( Pop \) that provides valuable information of the search space (location and fitness value) for the task of locally improving \( s^c \). On the one hand, this mechanism dismisses individuals worse than \( s^c \), because they reduce chances for improving \( s^c \). On the other hand, nearest solution is selected because it provides more information about the search space region where \( s^c \) is located. We also include a distance threshold to avoid choosing a solution too similar to \( s^c \).

Therefore, first parent is always \( s^c \), then, second parent is the individual in \( Pop \) most similar to \( s^c \) (at phenotypic level) fulfilling two conditions: 1) it has a better fitness value than \( s^c \) and 2) similarity between both solutions is above a given threshold. Mating threshold is initially set to 1\( e - 2 \). Then, after every VNS restart, it is set to half its value to improve precision.

An important remark is that this selection mechanism may return no mate for \( s^c \). It occurs when the mechanism has dismissed all solutions in \( Pop \) because either they are worse than \( s^c \), or their distances to \( s^c \) are under mating threshold. Crossover is not performed in this case. Offspring is generated by perturbing \( s^c \) instead (Sect. 4.2).

### 4.2. PBX-\( \alpha \) Crossover Operator

PBX-\( \alpha \) [20] is an instance of parent-centric crossover operators, which have arisen as a meaningful and efficient way of solving real-parameter optimisation problems [2, 9].

Given \( s^c = (s^c_1 \cdots s^c_n) \) and \( s^{mate} = (s^{mate}_1 \cdots s^{mate}_n) \), PBX-\( \alpha \) generates offspring \( z = (z_1 \cdots z_n) \), where \( z_i \) is a random (uniform) number from interval \([s^c_i - I \cdot \alpha, s^c_i + I \cdot \alpha]\), with \( I = |s^c_i - s^{mate}_i| \). Parameter \( \alpha \) is initially set to 0.5. Then, after every VNS restart, it is set to 0.5/\( \log_{10}(\text{number of restarts} + 1) \) to improve precision.

As mentioned before, crossover operation occurs when mate selection returns a solution from \( Pop \). Otherwise, offspring is generated by adding a normal random vector to \( s^c \), whose standard deviation is the product of the current values of aforementioned parameters \( \alpha \) and mating threshold.

Bounds constraints are checked as in the general VNS framework (Sect. 2), either when PBX-\( \alpha \) is applied or when a vector of random values is added to \( s^c \).

### 4.3. Acceptance Criterion

Once offspring \( z \) has been generated either by PBX-\( \alpha \) or perturbation, Continuous Local EA decides which solution, between \( z \) and \( s^c \), becomes the new \( s^c \). It follows a similar idea to the one behind Simulated Annealing [16] to overcome rugged landscapes and local optima. In particular, it always accepts \( z \) if it is better than a specific computed bound. At the beginning, that bound is set equal to the fitness value of \( s^c \), only allowing improving offspring.
Then, every time a new best \( s^c \) is found, the bound is computed as the average of the fitness of the new best \( s^c \) and the old bound. This simple mechanism provides some selection pressure on the search process that also lets the algorithm to accept uphill moves, and overcome small local optima.

4.4. Lowest Information Contribution Replacement

One of the objectives of Continuous Local EA is to gather up valuable information about the search space and its optima, to enhance future refinements. To carry out this objective, \( P_{\text{op}} \) should maintain a set of good solutions properly spread over the search space. The problem of maintaining such a set of solutions recalls the problem of attaining a good set of non-dominated solutions in multi-objective problems. In [17], Kukkonen and Deb propose a method for pruning of non-dominated solutions in many-objective problems. The basic idea is to eliminate the most crowded members of a non-dominated set one by one, and update the crowding information of the remaining members after each removal.

Continuous Local EA applies a similar method for updating \( P_{\text{op}} \) when solving mono-objective problems. The mechanism firstly inserts the new offspring in \( P_{\text{op}} \), then, it removes the most crowded solution, keeping population size constant (we use \(|P_{\text{op}}| = 100\)). To determine the solution to be removed, everyone in \( P_{\text{op}} \) holds a pointer to its nearest solution as well as the distance between them, at phenotypic level. Since, according to this definition, there are always, at least, two solutions holding the minimum distance, the worst of them is the one to be removed. Pointers and distances are updated every time \( P_{\text{op}} \) changes, i.e., when inserting the new solution and when removing the most crowded one. Notice that the best found solution is never removed.

5. \( \mu \text{CHC} \) as Shaking Component

The VNS model applies \( \mu \text{CHC} \) as the shaking component. \( \mu \text{CHC} \) was firstly presented as an evolutionary ILS-perturbation technique for binary-coded problems [18, 19]. The role of \( \mu \text{CHC} \) is to receive \( s^c \), provide local diversity, and generate a solution that is then considered as starting point for the next improvement process (see Fig. 3). The reason for choosing CHC, as the base model, is that it suitably combines powerful diversification mechanisms with an elitist selection strategy. The filtering of high diversity by means of high selective pressure favours the creation of useful diversity: many dissimilar solutions are produced during the run and only the best ones are conserved in the population, allowing diverse and promising solutions to be maintained. From our point of view, this behaviour is desirable for an EA assuming the work of a shaking operator.

5.1. \( \mu \text{CHC} \) Model

We have conceived \( \mu \text{CHC} \) to be an effective explorer in the neighbourhood of \( s^c \) to perform shaking process, because it provides local diversity. At the beginning, \( s^c \) is used to create its initial population (Sect. 5.2). Then, it is performed throughout a predetermined number of fitness function evaluations. The best reached individual is then considered as starting point for the next improvement process (see Fig. 3). Every time a new solution is sampled, bound constraints are checked as in the general VNS framework (Sect. 2).

The main components of the algorithm are:

- **Population size**: \( \mu \text{CHC} \) manages a population with few individuals (\(|P_{\text{op}}| = 5\)), and thus, it may be seen as micro EA. In standard VNS models, the number of evaluations required by the shaking mechanism is very low as compared with the one for the improvement method. With the aim of preserving, as far as possible, the essence of VNS, we have considered an EA with a low sized population; for being able to work adequately under the requirement of spending few evaluations.

- **Number of evaluations**: In particular, the number of evaluations assigned to \( \mu \text{CHC} \) for a particular invocation will be a fixed proportion \( (p_{\text{evals}}) \), of the number of evaluations consumed by the previously performed improvement method (Continuous Local EA). It is worth noting that \( p_{\text{evals}} \) should be set to a low value. We will use \( p_{\text{evals}} \) equals to 0.5.

- **Elitist selection**: Current population is merged with offspring population obtained from it and the best \(|P_{\text{op}}| \) individuals are selected for the new population.

- **Incest prevention mechanism**: Before mating, Hamming distance between the corresponding Gray-coding strings of paired individuals (with 20 bits per variable) is calculated. Only paired individuals whose distance exceed a difference threshold \( d \) are allowed to undergo
crossover operation. Aforementioned threshold is initialised to \( L/4 \) (with \( L \) being the number of bits coding a potential solution). If no offspring is obtained in one generation, difference threshold is decremented by one.

- **BLX-\( \alpha \) crossover operator**: Paired individuals allowed to produce offspring undergo BLX-\( \alpha \) crossover operator [4], producing one offspring. Given \( y^1 \) and \( y^2 \), BLX-\( \alpha \) generates offspring \( z \), where \( z_i \) is a random (uniform) number from \( [y^\min \cdot I - \alpha, y^\max + I \cdot \alpha] \) interval, with \( y^\min = \min(y^1_i, y^2_i), y^\max = \max(y^1_i, y^2_i) \), and \( I = |y^1_i - y^2_i| \). This parameter \( \alpha \) is set to 0.5.

- **Mating with \( s^c \)**: \( \mu \text{CHC} \) incorporates a strategy of recombining \( s^c \) with another solution. In addition to the typical recombination phase of CHC, our algorithm mates \( s^c \) with a member of \( \text{Pop} \) (selected at random) and, if they are finally crossed over (attending on the incest prevention mechanism), the resulting offspring will be introduced into the offspring population.

- **Cataclysmic mutation**: \( \mu \text{CHC} \), as CHC, uses no mutation in the classical sense of the concept, but instead, it goes through a process of cataclysmic mutation when the population has converged. The difference threshold is considered to measure the stagnation of the search, which happens when it has dropped to one. Then, the population is reinitialised with individuals generated by perturbing \( s^c \) attending to the \( k \)th neighbourhood structure (Sect. 5.2).

## 5.2. Initialisation and Cataclysmic Mutation

Every individual in the initial \( \mu \text{CHC} \) population is generated by perturbing \( s^c \) according to a relaxed idea of the given neighbourhood structure \( N_k \) (see equation 1). In particular, each individual is generated by adding a vector of random values from the following equation (see Fig. 4):

\[
N(0, 1) \cdot \text{range}_i / k_{\text{max}} \pm \varepsilon \cdot \text{range}_i / k_{\text{max}} \tag{3}
\]

where \( N(0, 1) \) is a normal random value with mean 0 and variance 1, \( \text{range}_i \) is the range of \( i \)th variable domain, \( k_{\text{max}} \) is the maximum number of neighbourhood structures \( \text{VNS} \) manages, and \( \varepsilon \) is the current neighbourhood index; \( \pm \) is chosen according to the sign of the random value.

<table>
<thead>
<tr>
<th>Table 1. Test problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
</tr>
<tr>
<td>( f_1 ) Shifting Sphere</td>
</tr>
<tr>
<td>( f_2 ) Schwefel</td>
</tr>
<tr>
<td>( f_3 ) Shifting Rosenbrock</td>
</tr>
<tr>
<td>( f_4 ) Shifted Rastrigin</td>
</tr>
<tr>
<td>( f_5 ) Shifted Griewank</td>
</tr>
<tr>
<td>( f_6 ) Shifted Ackley</td>
</tr>
<tr>
<td>( f_7 ) Schwefel 2.22</td>
</tr>
<tr>
<td>( f_8 ) Schwefel 1.2</td>
</tr>
<tr>
<td>( f_9 ) Extended F10</td>
</tr>
<tr>
<td>( f_{10} ) Bohachevsky</td>
</tr>
<tr>
<td>( f_{11} ) Schaffer</td>
</tr>
</tbody>
</table>

Cataclysmic mutation fills the population with individuals created by the same way as initial population is built, but preserving the best performing individual found in the previous generation. After applying cataclysmic mutation, the difference threshold is set to: \( \sigma \cdot (1 - \sigma) \cdot L \), with \( L \) being the number of bits coding a potential solution of the problem (20 bits per variable) and \( \sigma \) is the average of the minimal radius of neighbourhoods \( N_k \) and \( N_{k+1} \left( (2k + 1)/(2k_{\text{max}}) \right) \).

## 6. Results

The continuous VNS model with evolutionary MH components has been tested on the experimental framework proposed for the Special Session on Evolutionary Algorithms and other Metaheuristics for Continuous Optimization Problems - A Scalability Test, for the ISDA’09 conference. It consist of 11 scalable function optimisation problems. Table 1 shows their names, bounds, and optimum values. Functions \( f_1 \) to \( f_6 \) are defined in [23], and functions \( f_7 \) to \( f_{11} \), in [9]. Four dimensions (\( D = \{50, 100, 200, 500\} \)) are studied. The maximum number of fitness evaluations is 5,000 \( \cdot D \). Table 2 summarises the parameter setting (parameter values taken from [7]), where CLEA is for Continuous Local EA. Table 3 shows the error average, defined as the difference between the best attained fitness and the optimum value, and its standard deviation in brackets over 25 runs for each test function and dimension.

<table>
<thead>
<tr>
<th>Table 2. Parameter setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>( k_{\text{max}} )</td>
</tr>
<tr>
<td>( \text{CLEA’s Pop} )</td>
</tr>
<tr>
<td>( \mu \text{CHC’s Pop} )</td>
</tr>
<tr>
<td>( \text{BLX-( \alpha )} )</td>
</tr>
<tr>
<td>( p_{\text{evals}} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMA-ES parameters</td>
<td>Suggested in code</td>
</tr>
<tr>
<td>( \text{CLEA stop cond.} )</td>
<td>100 its without improv.</td>
</tr>
<tr>
<td>( \text{CLEA mating th.} )</td>
<td>1e-2 initially</td>
</tr>
<tr>
<td>( \mu \text{CHC bits/var} )</td>
<td>20</td>
</tr>
</tbody>
</table>
Table 3. Results

<table>
<thead>
<tr>
<th></th>
<th>50</th>
<th>100</th>
<th>200</th>
<th>500</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_1$</td>
<td>5.7e-14(3e-29)</td>
<td>1.3e-13(3e-14)</td>
<td>2.6e-13(3e-14)</td>
<td>6.0e-13(3e-14)</td>
</tr>
<tr>
<td>$f_2$</td>
<td>2.7e-11(6e-12)</td>
<td>5.8e-9(1e-8)</td>
<td>1.1e-9(6e-10)</td>
<td>3.6e-4(1e-4)</td>
</tr>
<tr>
<td>$f_3$</td>
<td>1.5e+0(2.8e+0)</td>
<td>1.8e+1(8.4e+0)</td>
<td>1.0e+2(1.5e+1)</td>
<td>6.1e+4(3.0e+5)</td>
</tr>
<tr>
<td>$f_4$</td>
<td>1.3e+4(4.5e+0)</td>
<td>4.4e+4(1.3e+1)</td>
<td>1.9e+2(1.1e+2)</td>
<td>2.5e+3(4.3e+2)</td>
</tr>
<tr>
<td>$f_5$</td>
<td>3.3e-14(1e-14)</td>
<td>7.7e-14(2e-14)</td>
<td>1.6e-13(1e-14)</td>
<td>3.5e-13(1e-14)</td>
</tr>
<tr>
<td>$f_6$</td>
<td>1.8e+7(5.7e+0)</td>
<td>2.1e+1(5.8e-1)</td>
<td>2.1e+1(1.9e-1)</td>
<td>2.2e+1(4.3e-2)</td>
</tr>
<tr>
<td>$f_7$</td>
<td>1.5e-9(4.0e-9)</td>
<td>7.7e-5(2.6e-4)</td>
<td>4.1e-2(7.4e-2)</td>
<td>5.0e-2(5.1e-2)</td>
</tr>
<tr>
<td>$f_8$</td>
<td>1.2e-14(1e-14)</td>
<td>5.8e-14(1e-14)</td>
<td>2.6e-13(4e-14)</td>
<td>4.9e-4(2.3e-4)</td>
</tr>
<tr>
<td>$f_9$</td>
<td>4.8e+0(9.0e+0)</td>
<td>2.2e+1(1.5e+1)</td>
<td>1.5e+2(8.0e+1)</td>
<td>2.5e+3(6.6e+2)</td>
</tr>
<tr>
<td>$f_{10}$</td>
<td>1.7e-4(1.5e-4)</td>
<td>4.5e-4(2.7e-4)</td>
<td>1.3e-3(5.7e-4)</td>
<td>4.7e-3(6.5e-4)</td>
</tr>
<tr>
<td>$f_{11}$</td>
<td>3.6e+0(6.5e+0)</td>
<td>2.2e+1(1.7e+1)</td>
<td>1.7e+2(7.2e+1)</td>
<td>2.4e+3(6.1e+2)</td>
</tr>
</tbody>
</table>

7. Conclusions

We have presented the continuous VNS model based on three evolutionary MH components: 1) CMA-ES as an EA specialised in generating a good starting point, as generation component, 2) Continuous Local EA, specialised in exploiting local information, as improvement component, and 3) $\mu$CHC, which provides local diversity, as shaking component. We have tested the continuous VNS model with evolutionary MH components on the experimental framework proposed for the Special Session on Evolutionary Algorithms and other Metaheuristics for Continuous Optimization Problems - A Scalability Test, for the ISDA’09 conference, to study the ability of the algorithm of being scalable for high-dimensional problems.

8 Acknowledgments

This work was supported by Research Projects TIN2008-05854 and P08-TIC-4173.

References