AIP Conference Proceedings

A variational discrete filled function approach in discrete global optimization

Woon Siew Fang, Sharmila Karim, and Mohd Saiful Adli Mohamad

Citation: AIP Conference Proceedings **1691**, 040027 (2015); doi: 10.1063/1.4937077
View online: http://dx.doi.org/10.1063/1.4937077
View Table of Contents: http://scitation.aip.org/content/aip/proceeding/aipcp/1691?ver=pdfcov
Published by the AIP Publishing

A Variational Discrete Filled Function Approach in Discrete Global Optimization

Woon Siew Fang[1,a)], Sharmila Karim[2,b)] and Mohd Saiful Adli Mohamad[3,c)]

[1,2,3] *School of Quantitative Sciences, UUM CAS,Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia*

[a)] Corresponding author: woonsiewfang@yahoo.com
[b)] mila@uum.edu.my
[c)]msadli@uum.edu.my

Abstract. Many real-life applications governed by discrete variables poss multiple local optimal solutions, which requires the utilization of global optimization tools find the best solution amongst them. The main difficulty in determining the best solution, or also known as the global solution, is to escape from the basins surrounding local minimums. To overcome this issue, an auxiliary function is introduced in discrete filled function method which turns the local minimizer of the original function become a maximizer. Then, an improved local minimum is found by minimizing the filled function, otherwise the edge of the feasible region is attained. Based on a discrete filled function method from the literature, we propose a modification particularly on the neighbourhood search to enhance its computational efficiency. Numerical results suggest that the proposed algorithm is efficient in solving large scale complex discrete optimization problems.

Keywords: *Global optimization, heuristics, discrete filled function.*

INTRODUCTION

The discrete filled function method (DFFM) [1,4,5,7,8,9] is a heuristics global optimization technique developed in the late 1990s to solve discrete optimization problems. An ordinary descent method is first applied to identify a local minimum. Once found, the DFFM is used to suggest sequential improvement of local optima strategy through an auxiliary function to escape from one local minima to a better one. Consequently, the local minimum of the original function becomes a local maximum of the auxiliary function. The search for the improved minimum in a lower basin continues until the parameter of the discrete filled function is satisfied. The final point found is expected to be the global minimum. Though DFFM have been a popular global optimization tool in recent years, there has been limited attention on investigating this method on discrete or mixed discrete optimization problems. A critical review on these discrete filled function techniques has been performed by [6]. The study also reveals that the DFFM developed in [5] turns out to be the most efficient method which yields the lowest number of original function evaluations in solving most benchmark test problems. The motivation of our work is inspired by the method developed in [5] to propose an improved algorithm in solving a general class of discrete optimization problems more competently by reducing the computational times.

CONCEPTS & APPROACH

A nonlinear discrete optimization problem is defined as follows:

$$\min f(\mathbf{x}), \quad s.t.\, \mathbf{x} \in X, \tag{1}$$

where $X = \{\mathbf{x} \in Z^n : \mathbf{x}_{i,\min} \le \mathbf{x}_i \le \mathbf{x}_{i,\max}, i = 1,\ldots,n,\}$, Z^n is the set of integer points in R^n, and $\mathbf{x}_{i,\min}, \mathbf{x}_{i,\max}, i = 1,\ldots,n,$ are given bounds. Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be any two distinct points in the box constrained set X. Since X is bounded, there exists a constant κ such that

$$1 \le \max_{\substack{\mathbf{x}_1,\mathbf{x}_2 \in X \\ \mathbf{x}_1 \ne \mathbf{x}_2}} \| \mathbf{x}_1 - \mathbf{x}_2 \| \le \kappa < \infty, \tag{2}$$

Innovation and Analytics Conference and Exhibition (IACE 2015)
AIP Conf. Proc. 1691, 040027-1–040027-6; doi: 10.1063/1.4937077

where $\|\cdot\|$ is the Euclidean norm. We make the following assumption.

Assumption 1: There exists a constant L , $0 < L < \infty$, such that

$$| f(\mathbf{x}_1) - f(\mathbf{x}_2) | \le L \| \mathbf{x}_1 - \mathbf{x}_2 \|, \quad (\mathbf{x}_1, \mathbf{x}_2) \in X \times X.$$

We now revise some familiar definitions in discrete optimization.

Definition 1 A sequence $\{\mathbf{x}^{(i)}\}_{i=0}^{k+1}$ in X is a discrete path between two distinct points $\mathbf{x}^*$ and $\mathbf{x}^{**}$ in X if $\mathbf{x}^{(0)} = \mathbf{x}^*, \mathbf{x}^{(k+1)} = \mathbf{x}^{**}, \mathbf{x}^{(i)} \in X$ for all i, $\mathbf{x}^{(i)} \ne \mathbf{x}^{(j)}$ for $i \ne j$, and $\| \mathbf{x}^{(i+1)} - \mathbf{x}^{(i)} \| = 1$ for all i. Let A be a subset of X. If, for all $\mathbf{x}^*, \mathbf{x}^{**} \in A$, $\mathbf{x}^*$ and $\mathbf{x}^{**}$ are connected by a discrete path, then A is called a pathwise connected set.

Definition 2 For any $\mathbf{x} \in X$, the *neighbourhood* of $\mathbf{x}$ is defined by

$$N(\mathbf{x}) = \{\mathbf{w} \in X : \mathbf{w} = \mathbf{x} \pm e_i, i = 1, \ldots, n\},$$

where e_i denotes the i-th standard unit basis vector of R^n with the i-th component equal to one and all other components equal to zero.

Definition 3 The set of feasible directions at $\mathbf{x} \in X$ is defined by

$$D(\mathbf{x}) = \{\mathbf{d} \in R^n : \mathbf{x} + \mathbf{d} \in N(\mathbf{x})\} \subset E = \{\pm e_1, \ldots, \pm e_n\}.$$

Definition 4 $\mathbf{d} \in D(\mathbf{x})$ is a descent direction of f at $\mathbf{x}$ if $f(\mathbf{x} + \mathbf{d}) < f(\mathbf{x})$.

Definition 5 $\mathbf{d}^* \in D(\mathbf{x})$ is a steepest descent direction of f at $\mathbf{x}$ if it is a descent direction and $f(\mathbf{x} + \mathbf{d}^*) \le f(\mathbf{x} + \mathbf{d})$ for any $\mathbf{d} \in D(\mathbf{x})$.

Definition 6 $\mathbf{x}^* \in X$ is a *local minimizer* of X if $f(\mathbf{x}^*) \le f(\mathbf{x})$ for all $\mathbf{x} \in N(\mathbf{x}^*)$. If $f(\mathbf{x}^*) < f(\mathbf{x})$ for all $\mathbf{x} \in N(\mathbf{x}^*) \setminus \mathbf{x}^*$, then $\mathbf{x}^*$ is a *strict local minimizer* of f.

Definition 7 $\mathbf{x}^*$ is a *global minimizer* of f if $f(\mathbf{x}^*) \le f(\mathbf{x})$ for all $\mathbf{x} \in X$. If $f(\mathbf{x}^*) < f(\mathbf{x})$ for all $\mathbf{x} \in X \setminus \mathbf{x}^*$, then $\mathbf{x}^*$ is a *strict global minimizer* of f.

Definition 8 $\mathbf{x}$ is a vertex of X if for each $\mathbf{d} \in D(\mathbf{x})$, $\mathbf{x} + \mathbf{d} \in X$ and $\mathbf{x} - \mathbf{d} \notin X$. Let $\widetilde{X}$ denote the set of vertices of X.

Definition 9 $B^* \subset X$ is a discrete basin of f corresponding to the local minimizer $\mathbf{x}^*$ if it satisfies the following conditions:

- B^* is pathwise connected.
- B^* contains $\mathbf{x}^*$.
- For each $\mathbf{x} \in B^*$, any connected path starting at $\mathbf{x}$ and consisting of descent steps converges to $\mathbf{x}^*$.

Definition 10 Let $\mathbf{x}^*$ and $\mathbf{x}^{**}$ be two distinct local minimizers of f. If $f(\mathbf{x}^{**}) < f(\mathbf{x}^*)$, then the discrete basin B^{**} of f associated with $\mathbf{x}^{**}$ is said to be lower than the discrete basin B^* of f associated with $\mathbf{x}^*$.

Definition 11 Let $\mathbf{x}^*$ be a local minimizer of $-f$. The discrete basin of $-f$ at $\mathbf{x}^*$ is called a discrete hill of f at $\mathbf{x}^*$.

Definition 12 For a given local minimizer $\mathbf{x}^*$, define the discrete sets $S_L(\mathbf{x}^*) = \{\mathbf{x} \in X : f(\mathbf{x}) < f(\mathbf{x}^*)\}$ and $S_U(\mathbf{x}^*) = \{\mathbf{x} \in X : f(\mathbf{x}) \ge f(\mathbf{x}^*)\}$. Note that $S_L(\mathbf{x}^*)$ contains the points lower than $\mathbf{x}^*$, while $S_U(\mathbf{x}^*)$ contains the points higher than $\mathbf{x}^*$.

Let $\mathbf{x}^*$ be a local minimizer of f. In [5], the discrete filled function $G_{\mu,\rho,\mathbf{x}^*}$ at $\mathbf{x}^*$ is defined as follows:

$$G_{\mu,\rho,\mathbf{x}^*}(\mathbf{x}) = A_\mu(f(\mathbf{x}) - f(\mathbf{x}^*)) - \rho \| \mathbf{x} - \mathbf{x}^* \|, \tag{3}$$

$$A_\mu(y) = \mu y[(1-c)(\frac{1-c\mu}{\mu - c\mu})^{-y/\omega} + c],$$

where $\omega > 0$ is a sufficiently small number, $c \in (0,1)$ is a constant, $\rho > 0$, and $0 < \mu < 1$. It can be shown that the function $G_{\mu,\rho,\mathbf{x}^*}(\mathbf{x})$ is a discrete filled function when certain conditions on the parameters μ and ρ are satisfied, as detailed by the following properties proved in [5]:

- $\mathbf{x}^*$ is a strict local maximizer of $G_{\mu,\rho,\mathbf{x}^*}$ if $\rho > 0$ and $0 < \mu < \min\{1, \rho/L\}$.
- If $\mathbf{x}^*$ is a global minimizer of f, then $G_{\mu,\rho,\mathbf{x}^*}(\mathbf{x}) < 0$ for all $\mathbf{x} \in X \setminus \mathbf{x}^*$.
- Let $\bar{\mathbf{d}} \in D(\bar{\mathbf{x}})$ be a feasible direction at $\bar{\mathbf{x}} \in S_U(\mathbf{x}^*)$ such that $\| \bar{\mathbf{x}} + \bar{\mathbf{d}} - \mathbf{x}^* \| > \| \bar{\mathbf{x}} - \mathbf{x}^* \|$. If $\rho > 0$ and $0 < \mu < \min\{1, \frac{\rho}{2\kappa^2 L}\}$, then $G_{\mu,\rho,\mathbf{x}^*}(\bar{\mathbf{x}} + \bar{\mathbf{d}}) < G_{\mu,\rho,\mathbf{x}^*}(\bar{\mathbf{x}}) < 0 = G_{\mu,\rho,\mathbf{x}^*}(\mathbf{x}^*)$.
- Let $\mathbf{x}^{**}$ be a strict local minimizer of f with $f(\mathbf{x}^{**}) < f(\mathbf{x}^*)$. If $\rho > 0$ is sufficiently small and $0 < \mu < 1$, then $\mathbf{x}^{**}$ is a strict local minimizer of $G_{\mu,\rho,\mathbf{x}^*}$.
- Let $\mathbf{x}'$ be a local minimizer of $G_{\mu,\rho,\mathbf{x}^*}$ and suppose that there exists a feasible direction $\bar{\mathbf{d}} \in D(\mathbf{x}')$ such that $\| \mathbf{x}' + \bar{\mathbf{d}} - \mathbf{x}^* \| > \| \mathbf{x}' - \mathbf{x}^* \|$. If $\rho > 0$ is sufficiently small and $0 < \mu < \min\{1, \frac{\rho}{2\kappa^2 L}\}$, then $\mathbf{x}'$ is a local minimizer of f.

Assume that every local minimizer of f is strict. Suppose that $\rho > 0$ is sufficiently small and $0 < \mu < \min\{1, \frac{\rho}{2\kappa^2 L}\}$. Then, $\mathbf{x}^{**} \in X \setminus \tilde{X}$ is a local minimizer of f with $f(\mathbf{x}^{**}) < f(\mathbf{x}^*)$ if and only if $\mathbf{x}^{**}$ is a local minimizer of $G_{\mu,\rho,\mathbf{x}^*}$.

The DFFM can be illustrated as follows. First, an initial point is chosen before applying a local search to find a discrete local minimizer. Then, the neighborhood of the discrete local minimizer is set up. Next, a discrete filled function is constructed and the local minimizer of the original function becomes the local maximizer of the discrete filled function. By minimizing the discrete filled function, a discrete local minimizer of the discrete filled function is found. The discrete local minimizer of the discrete filled function is examined whether it is an improved point or a corner point. If it is an improved point, the local minimizer of the discrete filled function transforms to be a new starting point to minimize the original function in such a way that an improved local minimizer can be found. In most situations, the local minimizer of discrete filled function is also a local minimizer of the original function when some standard assumption are satisfied. On the other hand, if the discrete local minimizer of the discrete filled function is a corner point, the next point in the neighbourhood in Step 3 is chosen to minimize the discrete filled function, until all points in the neighbourhood are tested. Then, the parameter ρ is reduced and the search of the discrete local minimizer of the discrete filled function is repeated. However, if the local minimizer of the discrete filled function is neither an improved point or corner point, the parameter μ is reduced and a new discrete filled function is constructed at the current point under the new parameter setting. The parameters reduction continues until they reached their presetting values, and the best solution found is treated as the global minimizer.

THE PROPOSED ALGORITHM

In the standard algorithm proposed in [5], the neighborhood points are set to be one unit away from the local minimizer of the original function. In our modified approach, we defined the neighborhood to be three units away from the local minimizer. The motivation for this modification is we believe that we can achieve faster computation by not choosing the immediate neighborhood of $\mathbf{x}^*$ but not too far away from the current local minimizer. An additional step is added to verify if any point in this outer neighbourhood is an improved point before starting to minimize the discrete filled function. The proposed method consists of two phases: the first phase involves finding a local minimum of original function; the second phase involves finding an improved point through discrete filled function using points in the outer neighbourhood from which the local search can be restarted as outlined below.

Algorithm 1 - Local search

1. Choose an initial point $\mathbf{x} \in X$.
2. If $\mathbf{x}$ is a local minimizer of f , then stop. Otherwise, find the steepest descent direction $\mathbf{d}^* \in D(\mathbf{x})$ of f at $\mathbf{x}$.
3. Set $\mathbf{x} := \mathbf{x} + \mathbf{d}^*$. Go to Step 2.

Algorithm 2 - Modified Algorithm with Outer Neighbourhood (MAON)

1. Initialize $\mathbf{x}_0 \in X$, ρ_0 , μ_0 , $\rho_L > 0$, $0 < \hat{\rho} < 1$, and $0 < \hat{\mu} < 1$.
   Let $\rho := \rho_0$ and $\mu := \mu_0$.
   Choose an initial point $\mathbf{x}_0 \in X$.
2. Starting from $\mathbf{x}_0$, minimize $f(\mathbf{x})$ using Algorithm 1 to obtain a local minimizer $\mathbf{x}^*$ of f .
3. Define $\overline{N}(\mathbf{x}^*) = \{\mathbf{w}_l \in X \mid \mathbf{w}_l = \mathbf{x}^* \pm 3e_i : i = 1,2,\ldots,n\} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_q\}, q \le 2n$.
4. (a) Set $\ell := 1$.
   (b) If $f(\mathbf{w}_\ell) < f(\mathbf{x}^*)$, set $\mathbf{x}_0 := \mathbf{w}_\ell$ and go to Step 2. Otherwise, go to (c) below.
   (c) Set $\ell := \ell + 1$. If $\ell \le q$, go to (b) above. Otherwise, set $\ell := 1$ and go to (d) below.
   (d) Set the current point $\mathbf{x}_c := \mathbf{w}_\ell$.
5. (a) If there exists a direction $\mathbf{d} \in D(\mathbf{x}_c)$ such that $f(\mathbf{x}_c + \mathbf{d}) < f(\mathbf{x}^*)$, then set $\mathbf{x}_0 := \mathbf{x}_c + \mathbf{d}$ and go to Step 2. Otherwise, go to (b) below.
   (b) Let $D_1 = \{\mathbf{d} \in D(\mathbf{x}_c) : f(\mathbf{x}_c + \mathbf{d}) < f(\mathbf{x}_c) \textit{ and } G_{\mu,\rho,\mathbf{x}}(\mathbf{x}_c + \mathbf{d}) < G_{\mu,\rho,\mathbf{x}}(\mathbf{x}_c)\}$.
   If $D_1 \ne \varnothing$, set $\mathbf{d}^* := \arg\min_{\mathbf{d} \in D(\mathbf{x}_c)} \{f(\mathbf{x}_c + \mathbf{d}) + G_{\mu,\rho,\mathbf{x}^*}(\mathbf{x}_c + \mathbf{d})\}$.
   Then, set $\mathbf{x}_c := \mathbf{x}_c + \mathbf{d}^*$ and go to (a) above. Otherwise, go to (c) below.
   (c) Let $D_2 = \{\mathbf{d} \in D(\mathbf{x}_c) : G_{\mu,\rho,\mathbf{x}}(\mathbf{x}_c + \mathbf{d}) < G_{\mu,\rho,\mathbf{x}}(\mathbf{x}_c)\}$.
   If $D_2 \ne \varnothing$, set $\mathbf{d}^* := \arg\min_{\mathbf{d} \in D(\mathbf{x}_c)} \{G_{\mu,\rho,\mathbf{x}^*}(\mathbf{x}_c + \mathbf{d})\}$.
   Then, set $\mathbf{x}_c := \mathbf{x}_c + \mathbf{d}^*$ and go to (a) above. Otherwise, go to Step 6.
6. Let $\mathbf{x}' = \mathbf{x}_c$ be the local minimizer of $G_{\mu,\rho,\mathbf{x}^*}$ obtained from Step 5.
   (a) If $\mathbf{x}' \in \widetilde{X}$, set $\ell := \ell + 1$. If $\ell > q$, go to Step 7. Otherwise, go to Step 4(d).
   (b) If $\mathbf{x}' \notin \widetilde{X}$, reduce μ by setting $\mu := \hat{\mu}\mu$ and go to Step 5(b).
7. Reduce ρ by setting $\rho := \hat{\rho}\rho$. If $\rho < \rho_L$, terminate the algorithm. The current $\mathbf{x}^*$ is taken as a

global minimizer of the problem. Otherwise, set $\ell := 1$ and go to Step 4(d).

We test both proposed and standard algorithms as discussed above on Colville's function [2], a well-known test problem which has 1.94481×10^5 feasible points and a global minimum $\mathbf{x}^*_{\text{global}} = [1,1,1,1]^T$ with $f(\mathbf{x}^*_{\text{global}}) = 0$. The numerical results are shown in Table 1, where E_f is the total number of original function evaluations, E_G represents the total number of discrete filled function evaluations, and R_E denotes the ratio of the average number of original function evaluations to the total number of feasible points. We initialized both parameters μ and ρ as 0.1 and set $\rho_L = 0.001$. The parameter μ is reduced if $\mathbf{x}'$ is neither a vertex nor an improved point by setting $\mu := \mu/10$.

Problem: Colville's Function

$$\min f(\mathbf{x}) = 100(x_2 - x_1^2)^2 + (1-x_1)^2 + 90(x_4 - x_3^2)^2 + (1-x_3)^2$$
$$+10.1[(x_2-1)^2 + (x_4-1)^2] + 19.8(x_2-1)(x_4-1),$$
$$s.t. \quad -10 \le x_i \le 10, \quad x_i \; integer; \quad i = 1,2,3,4.$$

We used six starting points in our simulations, which were $[1,1,0,0]^T$, $[1,1,1,1]^T$, $[-10,10,-10,10]^T$, $[-10,-5,0,5]^T$, $[-10,0,0,-10]^T$, and $[0,0,0,0]^T$. Both discrete filled function algorithms found the global minimum from all starting points. Table 1 shows the comparison of the results obtained from the proposed and standard algorithm, where the proposed algorithm has a slightly lower average number of original function evaluations with 1645.67, compared with the standard algorithm with 1679.5, which is an improvement of 2%. In addition, the average E_G observed are 4898.67 and 5247.17 for the proposed and standard algorithms, respectively. The results shown that the proposed algorithm is more efficient in finding the minimizer of the filled function compared to the standard algorithm with a reduction of 6.06%. Therefore, the proposed algorithm outperforms the standard one in the overall numerical performance with the average total number of both original and filled functions evaluations at 6544.33 compared to 6926.67 obtained by the standard algorithm, which is an improvement of 5.52%.

TABLE 1. Numerical Results of Colville's Function

Algorithms	$\mathbf{x}_0$	$\mathbf{x}^*_{\text{final}}$	$f(\mathbf{x}^*_{\text{final}})$	E_f	E_G	$E_f + E_G$
Proposed	$[1,1,0,0]^T$	$[1,1,1,1]^T$	0	1550	4851	6401
Algorithm	$[1,1,1,1]^T$	$[1,1,1,1]^T$	0	1525	4812	6337
	$[-10,10,-10,10]^T$	$[1,1,1,1]^T$	0	1828	5096	6924
	$[-10,-5,0,5]^T$	$[1,1,1,1]^T$	0	1709	4906	6615
	$[-10,0,0,-10]^T$	$[1,1,1,1]^T$	0	1688	4856	6544
	$[0,0,0,0]^T$	$[1,1,1,1]^T$	0	1574	4871	6445
[5]	$[1,1,0,0]^T$	$[1,1,1,1]^T$	0	1426	5097	6523
	$[1,1,1,1]^T$	$[1,1,1,1]^T$	0	1422	5076	6498
	$[-10,10,-10,10]^T$	$[1,1,1,1]^T$	0	2674	5979	8653
	$[-10,-5,0,5]^T$	$[1,1,1,1]^T$	0	1567	5134	6701
	$[-10,0,0,-10]^T$	$[1,1,1,1]^T$	0	1557	5098	6655
	$[0,0,0,0]^T$	$[1,1,1,1]^T$	0	1431	5099	6530

In addition, we also compare the proposed method with some other DFFM available in the literature in solving Colville's function and summarized the results in Table 2. The results indicate that the proposed algorithm has the lowest total average number of the original and filled functions evaluations, hence turns out to be the most efficient method amongst these heuristics approaches.

TABLE 2. Comparison with other DFFM

Algorithms	$E_f + E_G$
Proposed algorithm	6543.33
[4]	10085.67
[5]	6926.67
[7]	39649.83
[8]	19258.33

CONCLUDING REMARK

A modified algorithm based on the existing DFFM is proposed in solving discrete optimization problem to enhance the numerical efficiency. Further study in solving applications of sciences and engineering problems, such as in biotechnology, process control, and financial planning will be performed.

ACKNOWLEDGMENTS

We would like to thank Universiti Utara Malaysia for providing the research funding (S/O Code: 12408) in completing this research project.

REFERENCES

1. R. Ge, *Mathematical Programming* **46**, 191-204 (1990).
2. A. A. Goldstein and J. F. Price, *Mathematics of Computation* **25**, 569-574 (1971).
3. H. Lin, Y. Wang, L. Fan, and Y. Gao, *Applied Mathematics and Computation* ***219***, 4371-4378 (2013).
4. C. K. Ng, L. S. Zhang, D. Li, and W. W. Tian, *Computational Optimization and Applications* **31**, 87-115 (2005).
5. C. K. Ng, D. Li, and L. S. Zhang, *Journal of Global Optimization* **37**, 357-379 (2007).
6. S. F. Woon and V. Rehbock, *Applied Mathematics and Computation* **217**, 25-41 (2010).
7. Y. Yang and Y. Liang, *Journal of Computational and Applied Mathematics* **202**, 280-291 (2007).
8. Y. Yang, Z. Wu, and F. Bai, *Journal of Industrial and Management Optimization* **4**, 353-362 (2008).
9. W. Zhu, *Applied Mathematics and Computation* **93**, 183-183 (1998).