
Curtin University
CURTIN UNIVERSITY LIBRARY
ESPACE@CURTIN
CURTIN'S INSTITUTIONAL RESEARCH REPOSITORY

Proceedings of the 5th International Conference on Optimization and Control with Applications (OCA2012), Beijing, China, December 4-8, 2012

109 A MODIFIED DISCRETE FILLED FUNCTION ALGORITHM FOR SOLVING NONLINEAR DISCRETE OPTIMIZATION PROBLEMS

Siew Fang Woon*[a], Volker Rehbock[b], Ryan Loxton[b]
[a]Department of Mathematics and Statistics
Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia
[b]Department of Mathematics and Statistics
Curtin University, Bentley, Western Australia 6102, Australia

Abstract: The discrete filled function method is a global optimization tool for searching for the best solution amongst multiple local optima. This method has proven useful for solving large-scale discrete optimization problems. In this paper, we consider a standard discrete filled function algorithm in the literature and then propose a modification to increase its efficiency.

Key words: Discrete filled function; Global optimization; Discrete optimization.

1 BACKGROUND

The discrete filled function method is one of the more recently developed global optimization tools for discrete optimization problems. Once a local minimum has been determined by an ordinary descent method, the discrete filled function approach involves the introduction of an auxiliary function, called a filled function, to avoid entrapment in the basin associated with this minimum. The local minimizer of the original function becomes a local maximizer of the filled function. By minimizing the filled function, the search moves away from the current local minimizer in the hope of escaping the basin associated with this minimizer and finding an improved solution.

The first filled function was introduced by Ge in the late 1980s Ge R. (1990) in the context of solving continuous global optimization problems. Zhu Zhu W. (1998) is believed to be the first researcher to introduce a discrete equivalent of the continuous filled function method in the late 1990s. This discrete filled function method overcomes the difficulties encountered in using a continuous approximation of the discrete optimization problem. However, the filled function proposed by Zhu contains an exponential term, which consequently makes it difficult to determine a point in a lower basin Ng C.K. (2007). Since the introduction of the original discrete filled function by Zhu, several new types of discrete filled functions with improved theoretical properties have been proposed, such as in Ng C.K. (2007); Yang Y. (2008); Shang Y. (2008), to enhance computational efficiency. A comprehensive survey of several discrete filled functions in the literature has been given in Woon S.F. (2010). The study showed that the discrete filled function developed in Ng C.K. (2007) seems to be the most reliable one since it guarantees that a local minimizer of the filled function is also a local minimizer of the original function, whereas other filled functions do not share this property. The goal of this paper is to propose an improved filled function algorithm based on the work in Ng C.K. (2007).

2 CONCEPTS & APPROACH

Consider the following nonlinear discrete optimization problem:

$$\min f(\mathbf{x}), \quad \text{s.t. } \mathbf{x} \in X, \tag{2.1}$$

where $X = \{\mathbf{x} \in \mathbb{Z}^n : \mathbf{x}_{i,\min} \leq \mathbf{x}_i \leq \mathbf{x}_{i,\max},\ i = 1, \ldots, n,\}$, $\mathbb{Z}^n$ is the set of integer points in $\mathbb{R}^n$, 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 $\mathcal{K}$ such that

$$1 \leq \max_{\substack{\mathbf{x}_1, \mathbf{x}_2 \in X \\ \mathbf{x}_1 \neq \mathbf{x}_2}} \| \mathbf{x}_1 - \mathbf{x}_2 \| \leq \mathcal{K} < \infty, \tag{2.2}$$

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

Assumption 2.1 *There exists a constant $\mathcal{L}$, $0 < \mathcal{L} < \infty$, such that*

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

We now recall some familiar definitions and concepts used in the discrete optimization area.

Definition 2.1 *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, $\mathbf{x}^{(i)} \neq \mathbf{x}^{(j)}$ for $i \neq j$, 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 2.2 *For any $\mathbf{x} \in X$, the neighbourhood of $\mathbf{x}$ is defined by*

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

where $\boldsymbol{e}_i$ denotes the i-th standard unit basis vector of $\mathbb{R}^n$ with the i-th component equal to one and all other components equal to zero.

Definition 2.3 *The set of feasible directions at $\mathbf{x} \in X$ is defined by*

$$\mathcal{D}(\mathbf{x}) = \{\mathbf{d} \in \mathbb{R}^n : \mathbf{x} + \mathbf{d} \in N(\mathbf{x})\} \subset E = \{\pm \boldsymbol{e}_1, \ldots, \pm \boldsymbol{e}_n\}.$$

Definition 2.4 $\mathbf{d} \in \mathcal{D}(\mathbf{x})$ *is a descent direction of f at $\mathbf{x}$ if $f(\mathbf{x} + \mathbf{d}) < f(\mathbf{x})$.*

Definition 2.5 $\mathbf{d}^* \in \mathcal{D}(\mathbf{x})$ *is a steepest descent direction of f at $\mathbf{x}$ if it is a descent direction and $f(\mathbf{x} + \mathbf{d}^*) \leq f(\mathbf{x} + \mathbf{d})$ for any $\mathbf{d} \in \mathcal{D}(\mathbf{x})$.*

Definition 2.6 $\mathbf{x}^* \in X$ *is a local minimizer of X if $f(\mathbf{x}^*) \leq f(\mathbf{x})$ 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 2.7 $\mathbf{x}^*$ *is a global minimizer of f if $f(\mathbf{x}^*) \leq f(\mathbf{x})$ 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 2.8 $\mathbf{x}$ *is a vertex of X if for each $\mathbf{d} \in \mathcal{D}(\mathbf{x})$, $\mathbf{x} + \mathbf{d} \in X$ and $\mathbf{x} - \mathbf{d} \notin X$. Let $\tilde{X}$ denote the set of vertices of X.*

Definition 2.9 $B^* \subset X$ *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 2.10 *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 2.11 *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 2.12 *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* $S_U(\mathbf{x}^*) = \{\mathbf{x} \in X : f(\mathbf{x}) \geq f(\mathbf{x}^*)\}$*. 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 Ng C.K. (2007), the discrete filled function $G_{\mu,\rho,\mathbf{x}^*}$ at $\mathbf{x}^*$ is defined as follows:

$$G_{\mu,\rho,\mathbf{x}^*}(\mathbf{x}) = A_\mu(f(\mathbf{x}) - f(\mathbf{x}^*)) - \rho \parallel \mathbf{x} - \mathbf{x}^* \parallel, \tag{2.3}$$

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

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 Ng C.K. (2007):

- $\mathbf{x}^*$ is a strict local maximizer of $G_{\mu,\rho,\mathbf{x}^*}$ if $\rho > 0$ and $0 < \mu < \min\{1, \rho/\mathcal{L}\}$.
- 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 $\bar{\mathbf{d}} \in \mathcal{D}(\bar{\mathbf{x}})$ be a feasible direction at $\bar{\mathbf{x}} \in S_U(\mathbf{x}^*)$ such that $\parallel \bar{\mathbf{x}} + \bar{\mathbf{d}} - \mathbf{x}^* \parallel > \parallel \bar{\mathbf{x}} - \mathbf{x}^* \parallel$. If $\rho > 0$ and $0 < \mu < \min\{1, \frac{\rho}{2\mathcal{K}^2\mathcal{L}}\}$, 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 $\acute{\mathbf{x}}$ be a local minimizer of $G_{\mu,\rho,\mathbf{x}^*}$ and suppose that there exists a feasible direction $\bar{\mathbf{d}} \in \mathcal{D}(\acute{\mathbf{x}})$ such that $\parallel \acute{\mathbf{x}} + \bar{\mathbf{d}} - \mathbf{x}^* \parallel > \parallel \acute{\mathbf{x}} - \mathbf{x}^* \parallel$. If $\rho > 0$ is sufficiently small and $0 < \mu < \min\{1, \frac{\rho}{2\mathcal{K}^2\mathcal{L}}\}$, then $\acute{\mathbf{x}}$ is a local minimizer of f.
- Assume that every local minimizer of f is strict. Suppose that $\rho > 0$ is sufficiently small and $0 < \mu < \min\{1, \frac{\rho}{2\mathcal{K}^2\mathcal{L}}\}$. 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 $G_{\mu,\rho,\mathbf{x}^*}$.

3 THE STANDARD ALGORITHM

The discrete filled function approach can be described as follows. First, an initial point is chosen and a local search is applied to find an initial discrete local minimizer. Then, the filled function is constructed at this local minimizer. By minimizing the filled function, either an improved discrete local minimizer is found or the boundary of the feasible region is reached. The discrete local minimizer of the filled function usually becomes a new starting point for minimizing the original objective with the hope of finding an improved point compared to the first local minimizer. A new filled function is constructed at this improved point. The process is repeated until no improved local minimizer of the original filled function can be found. The final discrete local minimizer is then taken as an approximation of the global minimizer. If a local minimizer of the filled function cannot be found after repeated searches terminate on the boundary of the box constrained feasible region, then the parameters defining the filled function are adjusted and the search is repeated. This adjustment of the parameters continues until the parameters reach their predetermined bounds; the best solution obtained so far is then taken as the global minimizer. The parameter μ is reduced if $\acute{\mathbf{x}}$ is neither a vertex nor an improved point and we return to Step 4(b). When all searches terminate at vertices ($\ell > q$), ρ is adjusted. The algorithm for minimizing $G_{\mu,\rho,\mathbf{x}^*}$ exits prematurely when an improved point $\mathbf{x}_k$ with $f(\mathbf{x}_k) < f(\mathbf{x}^*)$ is found in Step 4 of Algorithm **??**. The algorithm sets $\mathbf{x}_0 := \mathbf{x}_k$ and returns to Step 2 to minimize the original function f. Note that a direction yielding the greatest improvement of $f + G_{\mu,\rho,\mathbf{x}^*}$ is chosen when minimizing $G_{\mu,\rho,\mathbf{x}^*}$, assuming that a direction for improving f and $G_{\mu,\rho,\mathbf{x}^*}$ simultaneously exists. If such a direction does not exist, the algorithm chooses the steepest descent direction for $G_{\mu,\rho,\mathbf{x}^*}$ alone.

Table 4.1 Comparison of Algorithms - Colville's Function.

Types	$E_{f,avg}$	$E_{G,avg}$	$R_{E,avg}$
Standard Algorithm	1679.5	5247.2	0.008635805
Modified Algorithm	1143.2	2954.7	0.005878038

4 A MODIFIED APPROACH

We replace the neighbourhood $N(\mathbf{x}^*)$ in Step 3 of the standard Algorithm with a set of randomly chosen points from X. Then, an additional step is added to test whether any one of these random points happens to be an improved point. The motivation for this modification is to search for improved points more efficiently by choosing points which give a broader coverage of X, similar to the approaches proposed in Shang Y. (2005); Shang Y. (2008); Yang Y. (2007). In the standard approach, the initial points are chosen as the neighbouring points of the current local solution.

We tested both the original algorithm and our modified version on Colville's function. This function has 1.94481×10^5 feasible points and a global minimum $\mathbf{x}^*_{\text{global}} = [1,1,1,1]^\top$ with $f(\mathbf{x}^*_{\text{global}}) = 0$. We initialized both parameters μ and ρ as 0.1 and set $\rho_L = 0.001$. The parameter μ is reduced if $\acute{\mathbf{x}}$ is neither a vertex nor an improved point by setting $\mu := \mu/10$.

Computational results are shown in Table 4.1, 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.

Problem 1: Colville's Function Schittkowski K. (1987)

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

Six starting points are considered, namely $[1,1,0,0]^\top$, $[1,1,1,1]^\top$, $[-10,10,-10,10]^\top$, $[-10,-5,0,5]^\top$, $[-10,0,0,-10]^\top$, and $[0,0,0,0]^\top$. From Table 4.1, both algorithms succeeded in finding the global minimum from all starting points. Our algorithm succeeds in determining the global solution of Colville's function much more efficiently with an average $E_f = 1143.2$, compared with $E_f = 1679.5$ for the standard algorithm, which is a reduction of 31.9% in the average number of original function evaluations. However, the gain in efficiency for our Algorithm is offset somewhat by reduced reliability, since we sometimes needed to repeat the algorithm several times for each starting point before a global solution was attained.

References

Ge R. (1990), A filled function method for finding a global minimizer of a function of several variables, *Mathematical Programming*, Vol. 46, pp. 191-204.

Ng C.K., Li D., Zhang L.S. (2007), Discrete global descent method for discrete global optimization and nonlinear integer programming, *Journal of Global Optimization*, Vol. 37, pp. 357-379.

Shang Y., Zhang L. (2005), A filled function method for finding a global minimizer on global integer optimization, *Journal of Computational and Applied Mathematics*, Vol. 181, pp. 200-210.

Shang Y., Zhang L. (2008), Finding discrete global minima with a filled function for integer programming, *European Journal of Operational Research*, Vol. 189, pp. 31-40.

Schittkowski K. (1987), *More test examples for nonlinear programming codes*, Springer-Verlag, New York.

Woon S.F., Rehbock V. (2010), A critical review of discrete filled function methods in solving nonlinear discrete optimization problems, *Applied Mathematics and Computation*, Vol. 217, pp. 25-41.

Yang Y., Zhang L. (2007), A gradually descent method for discrete global optimization, *Journal of Shanghai University*, Vol. 11, pp. 39-44.

Yang Y., Wu Z., Bai F. (2008), A filled function method for constrained nonlinear integer programming, *Journal of Industrial and Management Optimization*, Vol. 4, pp. 353-362.

Zhu W. (1998), An approximate algorithm for nonlinear integer programming, *Applied Mathematics and Computation*, Vol. 93, pp. 183-183.