INFORMATICA, 1998, Vol. 9, No. 3, 279–296

Local Search Efficiency when Optimizing Unimodal Pseudoboolean Functions

Alexander ANTAMOSHKIN,
Eugene SEMENKIN
Siberian Aerospace Academy
P.O.Box 486, Krasnoyarsk, Russia-660014
e-mail: ant@iea.krasnoyarsk.su

Received: Februrary 1998

Abstract. This work is a continuation of our previous papers devoted to exploration of the regular search procedures efficiency in binary search spaces. Here we formulate the problem in a rather general form as a problem of optimization of an unimodal pseudoboolean function given implicitly and obtain analitical estimates of the expected time of a minimum point search for procedures of direct local search. These estimates are polinomial for the case of weakly nonmonotone functions and exponential for the general case of arbitrary unimodal functions. We hope that the proposed result will be usefull first of all for practical applications.

Key words: unimodal pseudoboolean functions, local search techniques, convergence analysis.

1. Introduction

In our previous papers (Antamoshkin, Saraev and Semenkin, 1990; Antamoshkin and Semenkin, 1991) we have considered unimprovable by time-complexity regular algorithms for unimodal and polimodal monotone pseudoboolean function optimization.

The objective of this work is to consider the local search potentialities as a method for arbitrary unimodal pseudoboolean function optimization. Deterministic local search methods in the case of binary search spaces were considered by Nemhauser *et al.* (1978) and Ausiello and Protasi (1995) only with the standpoint of a quality of solutions that can be found in polynomial time for the class of supermodular function.

A review of the latest achievements in the field of search efficiency in binary spaces has been given by Rudolf (1996). In this work he has shown a polynomial (by the expected number of trails) convergence of evolutionary algorithms for a number of special classes of unimodal functions, too. But as the problem of optimization of arbitrary unimodal pseudobbolean functions is rather unsolvable in deterministic polynomial time (by evidence of Rudolf (1996)) it makes sense to seek for exponential estimates of the expected time of search for the problem in general statement.

We consider the unconstrained pseudoboolean optimization problem which can be written as

$$f(X) \longrightarrow \min, \tag{1}$$

where

$$f : B_2^n \longrightarrow R^1, \tag{2}$$

$B_2^n = \{X = (x_1, ..., x_n) : x_j \in \{0, 1\},\ j = 1, ..., n\},\ R^1$ is the real line.

Proceeding from practical reasons, we will presuppose that the function f is given implicitly, i.e., it can not be presented with an evident analytical form. For instance, we use some computational algorithm (or a simulating model) to obtain its values or this function is an output of a real system.

In the rest of the paper, we inquire into ability of the local search method to solve the formulated problem. Some needed notions and theoretical results are given next. In Section 3, we consider concrete schemes of the local search method for different-valued and arbitrary unimodal functions, formulate and prove theorems giving estimates of the expected number trials for these schemes on different classes of the functions. In particular, it has been shown that the local search method realized in the terms of Papadimitriou and Steiglitz (1982) the information complexity of the class of weakly nonmonotone unimodal functions. Finally we close this paper with some concluding remarks in Section 4.

2. Some Theoretical Notions and Results

Following to Papadimitriou and Steiglitz (1982) the first that is needed to organize a local search is the "neighborhood search procedure".

DEFINITION 1. Points X^1, X^2 are called k-neighboring if they differ in the values of k coordinates $(k = 1, ..., n)$. 1-neighbouring points will be called simply neighbouring.

DEFINITION 2. The set $O_k(X)$ $(k = 1, ..., n)$ of points that are k-neighbouring to the point X will be called the kth level of the point X $(O_0(X) = X)$.

REMARK 1. It is clear, that $\operatorname{card} O_k(X) = C_n^k$, $k = 1, ..., n$. Here (and below) C_n^k is the number of combinations from n on k.

Lemma 1 (Antamoshkin, Saraev and Semenkin, 1990). $\forall X^k \in O_k(X) \subset B_n^2$:

$$\operatorname{card}\{O_1(X^k) \cap O_{k-1}(X)\} = k, \quad k = 1, ..., n,$$

$$\operatorname{card}\{O_1(X^k) \cap O_{k+1}(X)\} = n - k, \quad k = 0, .., n - 1.$$

Lemma 2 (Antamoshkin, Saraev and Semenkin, 1990). *Let* $X^k \in O_k(X) \subset B_n^2$. *Then for all* $m = 0, ..., n$:

$$O_m(X^k) \subset \bigcup_{j=0}^{M} O_{|k-m|+2j}(X), \quad M = (n - |n - m - k| - |k - m|)/2,$$

$$\operatorname{card}(O_m(X^k) \cap O_{|k-m|+2j}(X)) = C_k^{\min(m,k)-j} \cdot C_{n-k}^{n-\max(m,k)-j}.$$

COROLLARY 1. For any k-neighbouring points X^1, X^2 :

$$\operatorname{card}(O_l(X^2) \cap O_{k-l}(X^1)) = C_k^l, \quad l = 0, ..., k.$$

Lemma 3 (Antamoshkin, Saraev and Semenkin, 1990). $\forall X^k \in O_k(X) \subset B_n^2,\ k = 0, ..., n:$

$$O_k(X) \subset \bigcup_{l=0}^{N} O_{2l}(X^k), \quad N = \min\{k, n-k\},$$

and

$$\operatorname{card}(O_{2l}(X^k) \cap O_k(X)) = C_k^l \cdot C_{n-k}^l, \quad l = 0, ..., N.$$

DEFINITION 3. The point X^* for which $f(X^*) < f(X)\ \forall X \in O_1(X^*)$ will be called a local minimum point of the function f.

DEFINITION 4. A pseudoboolean function that has only local minimum on B_n^2 will be called unimodal.

DEFINITION 5. The set of points $W(X^0, X^l) = \{X^0, X^1, ..., X^i, ..., X^l\} \in B_n^2$ will be called a path between X^0 and X^l if points X^i and X^{i-1} are neighbouring for all $i = 1, ..., l$.

DEFINITION 6. Let $A \subset B_n^2$. Then A is a connected set if for all $X^0, X^l \in A$ there exists a path $W(X^0, X^l) \subset A$.

DEFINITION 7. A connected set $S_c \subset B_n^2, \quad \operatorname{card} S_c \geqslant 2$, such that $f(X) = c$ ($c =$ const) for all $X \in S_c$, is called the constancy set of the function f on B_n^2.

DEFINITION 8. An unimodal function f is monotone on B_n^2 if

$$f(X^{k-1}) \leqslant f(X^k)\ \forall X^{k-1} \in O_{k-1}(X^*)\ \wedge\ \forall X^k \in O_k(X^*), \qquad k = 1, ..., n, \tag{3}$$

or equivalently

$$\max_{X^{k-1} \in O_{k-1}(X^*)} f(X^{k-1}) \leqslant \min_{X^k \in O_k(X^*)} f(X^k) \quad \forall k = 1, ..., n. \tag{4}$$

If (3) and (4) are fulfilled with the sign of strict inequality then the function f is strictly monotone.

Lemma 4 (Antamoshkin, Saraev and Semenkin, 1990). *If f is an unimodal strictly monotone function then*

$$f(X^{k+1}) > f(X^k) \quad \forall X^{k+1} \in O_{k+1}(X^*) \cap O_1(X^k) \tag{5}$$

for all $X^k \in O_k(X^*)$, $k = 0, ..., n-1$.

DEFINITION 9. An unimodal function f, for which the condition (3) (or (4)) is violated at least in one point of B_n^2, will be called nonmonotone.

REMARK 2. It follows from the last definitions that constancy sets are possible for both monotone and nonmonotone on B_n^2 functions.

DEFINITION 10. A path $W(X^0, X^l) \subset B_n^2$ between k-neighbouring points X^0 and X^l will be called minimal if $l = k$.

DEFINITION 11. We will say that a function f is different-valued on B_n^2 if for all minimal paths $W(X^0, X^l)$, $X^0 \in B_n^2$, $X^l = X^*$:

$$f(X^i) \neq f(X^j) \quad \text{for } i \neq j, \quad X^i, X^j \in W(X^0, X^l), \quad i, j = 0, ..., l.$$

REMARK 3. One can see that constancy sets are possible for different-valued functions too.

DEFINITION 12. A path $W^f(X^0, X^l) \subset B_n^2$ will be called the path of non-increasing of a function f if for all $X^i, X^{i-1} \in W^f(X^0, X^l)$, $i = 1, ..., l$:

$$f(X^i) \leqslant f(X^{i-1}), \tag{6}$$

and the path of decreasing of the function f in the case of the strict inequality in (6).

REMARK 4. It directly follows from Definitions 8 and 12 that a function f is monotone on B_n^2 if for any $X^0 \in B_n^2$ all paths of non-increasing $W^f(X^0, X^*)$ are minimal.

DEFINITION 13. A path $W^f_{\min}(X^0, X^l) \subset B_n^2$ will be called the path of most decreasing of a function f if for all $X^i, X^{i-1} \in W^f_{\min}(X^0, X^l)$, $i = 1, ..., l$:

$$f(X^i) = \min_{X_j^1 \in O_1(X^{i-1})} f(X_j^1). \tag{7}$$

DEFINITION 14. An unimodal nonmonotone on B_n^2 function f will be called weakly nonmonotone if for all $X^k \in O_k(X^*)$, $k = 1, ..., n$, the point $X^1_{\min}$, such that

$$f(X^1_{\min}) = \min_{X_j^1 \in O_1(X^k)} f(X_j^1), \tag{8}$$

belongs to $O_{k-1}(X^*)$.

REMARK 5. As it follows from the last definition, the condition "for all $X^0 \in B_n^2$ the path $W^f_{\min}(X^0, X^*)$ is minimal" is necessary and sufficient for the weakly nonmonotonecity of an nonmonotone function f.

Lemma 5. *Let f is an unimodal different-valued function. Then for each point $X \in B_n^2 \setminus \{X^*\}$ among the points $X_j^1 \in O_1(X)$, $j = 1, ..., n$, there is at least one point $X_{j_i}^1$ such that $f(X_{j_i}^1) < f(X)$.*

Proof. Assuming inversely we have a minimum in the point X that contradicts to the presupposition $X \in B_n^2 \setminus \{X^*\}$.

COROLLARY 2. If f is an unimodal different-valued function then for all $X \in B_n^2 \setminus \{X^*\}$ there is at least one path $W^f(X, X^*)$ of the function f decreasing.

Proof. It follows from Definitions 3 and 11.

Lemma 6. *If f is an unimodal different-valued function then for all $X^i \in W^f_{\min}(X^0, X^*) = \{X^0, X^1, ..., X^i, ..., X^*\}$:*

$$\{X^0, X^1, ..., X^{i-2}, X^{i+2}, ..., X^*\} \subset B_n^2 \setminus O_1(X^i).$$

Proof. Presuppose inversely: $X^j \in O_1(X^i)$ and $j - i > 1$. Then by Definition 13 we have $f(X^i) > f(X^{i+1}) > ... > f(X^l) > ... > f(X^j)$, i.e., $f(X^j) < f(X^{i+1})$. Hence in accordance with (7) $X^j = X^{i+1}$ that contradicts to the supposition: $j - i > 1$. For the case $i - j > 1$ the proof is analogous.

Lemma 7. *If f is an unimodal different-valued function then*

$$\min_{X_j^k \in O_k(X^*)} f(X_j^k) < \min_{X_j^{k+1} \in O_{k+1}(X^*)} f(X_j^{k+1}) \quad \forall k = 0, ...n-1. \tag{9}$$

Proof. Suppose that for a certain $k > 0$ (when $k = 0$ the unequality (9) directly follows from the minimum definition) the condition (9) is violated, i.e.,

$$\min_{X_j^k \in O_k(X^*)} f(X_j^k) > \min_{X_j^{k+1} \in O_{k+1}(X^*)} f(X_j^{k+1}). \tag{10}$$

Consider the point $X^{k+1}_{\min}$, which is determined from the condition

$$f(X^{k+1}_{\min}) = \min_{X_j^{k+1} \in O_{k+1}(X^*)} f(X_j^{k+1}).$$

As it follows from (10) for all $X_j^1 \in O_1(X^{k+1}_{\min}) \cap O_k(X^*) : f(X_j^1) > f(X^{k+1}_{\min})$. Next, since $X^{k+1}_{\min} \in O_{k+1}(X^*)(k > 0)$ then $X^{k+1}_{\min} \neq X^*$ and hence there is the point $X^{k+2} \in O_1(X^{k+1}_{\min}) \cap O_{k+2}(X^*)$ such that

$$f(X^{k+2}) < f(X^{k+1}_{\min}). \tag{11}$$

Take the point $X^{k+2}_{\min} f(X^{k+2}_{\min}) = \min X^{k+2}_j \in O_{k+2}(X^*) f(X^{k+2}_j)$. Evidently, that $f(X^{k+2})_{\min}) \leqslant f(X^{k+2})$. Then the same as for the point $X^{k+1}_{\min}$ there is the point $X^{k+3} \in O_1(X^{k+2}_{\min}) \cap O_{k+3}(X^*)$ such that

$$f(X^{k+3}) < f(X^{k+2}_{\min}) \tag{12}$$

and so on. Finally, for $k = n - 1$ like (11) and (12) we have

$$f(X^n) < f(X^{n-1}_{\min}), \tag{13}$$

where $X^n \in O_n(X^*)$. As $\operatorname{card} O_n(X^*) = 1$ then it follows from (13) that X^n is a local minimum point and the last contradicts to the presupposition about the function f unimodality.

COROLLARY 3. Let $W^f_{\min}(X^0, X^*) = \{X^0, X^1, ..., X^*\}$ is the path of an unimodal different-valued function f most decreasing and let the original point of this path belongs to $O_1(X^k_{\min})$ $k = 1, ..., n-1$, where $X^k_{\min}$ is determined of the condition

$$f(X^k_{\min}) = \min_{X^k_j \in O_k(X^*)} f(X^k_j), \tag{14}$$

then

$$(W^f_{\min}(X^0, X^*) \setminus \{X^0, X^1\}) \subset \bigcup_{i=0}^{k-1} O_i(X^*). \tag{15}$$

Proof. By Definition 14: $f(X^0) > f(X^1) > ... > f(X^*)$ and in accordance with (9): $\forall X \in \bigcup_{i=k}^{n} O_i(X^*):\ f(X) > f(X^k_{\min})$ whence (15) appears.

DEFINITION 15. The points of the set $E_I(X, A) = O_I(X) \cap A, A \subset B^n_2, X \in B^n_2$, which answer the condition

$$O_I(X) \cap A \neq \emptyset \wedge \forall k = 0, ..., I-1:\ \ O_I(X) \cap A = \emptyset,$$

will be called the first points of the set A with respect to the point X.

DEFINITION 16. The points of the set $L_J(X, A) = O_J(X) \cap A, A \subset B^n_2, X \in B^n_2$, which answer the condition

$$O_J(X) \cap A \neq \emptyset \wedge \forall k = J+1, ..., n:\ O_J(X) \cap A = \emptyset,$$

will be called the last points of the set A with respect to the point X.

DEFINITION 17. A point $X \in B^n_2$ will be called k-neighbouring to a set $A \subset B^n_2$ if

$$O_k(X) \cap A \neq \emptyset \wedge \forall l = 0, ..., k-1:\ O_l(X) \cap A = \emptyset.$$

DEFINITION 18. The set $O_k(A), A \subset B_2^n$ of all points of B_2^n, which are k-neighbouring to the set A, we will call the kth level of the set A, $O_0(A) = A$.

DEFINITION 19. The function f constancy set $S_{c^*} \subset B_2^n$ such that $c^* < f(X) \forall X \in O_1(S_{c^*})$ will be called the extended local minimum of the function f.

DEFINITION 20. A set $A \subset B_2^n$ we will call the Φ set with respect to a point $X \in B_2^n$ if

$$A = E_I(X, A) \cup \Big(\bigcup_{k=I+1}^{J-1} O_k(X) \Big) \cup L_J(X, A),$$

where $E_I(X, A)$ and $L_J(X, A)$ are the sets of the first and the last points of the set A with respect to the point X.

REMARK 6. The Φ sets properties have been explored in detail by Antamoshkin, Saraev and Semenkin (1990).

Lemma 8 (Antamoshkin, Saraev and Semenkin, 1990). *Unimodal monotone on B_2^n functions admit existence of any Φ constancy sets with respect to a minimum point.*

COROLLARY 4. Unimodal nonmonotone on B_2^n functions admit existence of any Φ constancy sets with respect to a minimum point.

Proof. Evidently.

Lemma 9. *If the unimodal nonmonotone (weakly nonmonotone) function f constancy set:*

$$S_c = E_I(X^*, S_c) \cup \Big(\bigcup_{k=I+1}^{J-1} O_k(X^*) \Big) \cup L_J(X^*, S_c), \quad 0 \leqslant I \leqslant J \leqslant n-2, \tag{16}$$

includes at least one full level of X^* then

$$\forall X \in O_i(X^*), \quad i \geqslant J: \quad f(X) > c.$$

Proof. Let $L_J(X^*, S_c) = O_J(X^*)$ and $X \in O_{J+1}(X^*)$. Then by Corollary 3

$$f(X) \geqslant \min_{X^{J+1} \in O_{J+1}(X^*)} f(X^{J+1}) \geqslant \min_{X^J \in O_J(X^*)} f(X^J) = c,$$

but as we presuppose a connectedness of the considering constancy sets then

$$\forall X \in O_i(X^*), \quad i > J:$$

$$f(X) \geqslant \min_{X^i \in O_i(X^*)} f(X^i) \geqslant \ldots \geqslant \min_{X^{J+1} \in O_{J+1}(X^*)} f(X^{J+1}) \geqslant c.$$

Let now $O_J(X^*) \setminus L_J(X^*, S_c) \neq \emptyset$. Then, taking into account Corollary 3, $\forall X \in O_i(X^*),\ i \geqslant J$ we have

$$f(X) \geqslant \min_{X^i \in O_i(X^*)} f(X^i) \geqslant ... \geqslant \min_{X^J \in O_J(X^*)} f(X^J).$$

Assume the function minimum on the set $O_J(X^*)$ is reached in the point $X^J_{\min} \in O_J(X^*) \setminus L_J(X^*, S_c)$, i.e., $f(X^J_{\min}) < c$. Consider $O_1(X^J_{\min})$:

$$\forall X^{J-1} \in O_{J-1} \cap O_1(X^J_{\min}) : \quad f(X^{J-1}) = c > f(X^J_{\min}).$$

In accordance with the lemma conditions S_c has at least one full level and then by the Φ sets definition if $O_J(X^*) \setminus L_J(X^*, S_c) \neq \emptyset$ then $O_1(X^J_{\min}) \subset S_c$ and

$$\forall X^{J+1} \in O_{J+1}(X^*) \cap O_1(X^J_{\min}) :$$
$$f(X^{J+1}) \geqslant \min_{X^{J+1} \in O_{J+1}(X^*)} f(X^{J+1}) \geqslant f(X^J_{\min}),$$

i.e., either a local minimum is in the point $X^J_{\min}$ or the point $X^J_{\min}$ belongs to an extended local minimum that contradicts the lemma supposition on the function f unimodality. From that $\min_{X^J \in O_J(X^*)} f(X^J) = c$ and hence

$$\forall X \in O_i(X^*), \quad i > J :$$

$$f(X) \geqslant \min_{X^i \in O_i(X^*)} f(X^i) \geqslant ... \min_{X^J \in O_J(X^*)} f(X^J) = c,$$

or taking into consideration the requirement on the constancy sets connectness, finally, we have $f(X) > c$.

COROLLARY 5. Let $S_{c_q} \subset B_2^n,\ q = 1, ..., Q$, are the unimodal nonmonotone (weakly nonmonotone) function f Φ constancy sets with respect to the minimum point X^* and let each constancy set contains not less than one full level of X^*. Then for all q and g $(q, g, = 1, ..., Q;\ q \neq g)$ such that $I_q < I_g : \quad c_q < c_g$.

Here I_q and I_g are the numbers of the levels of the first point sets $E_{I_q}(X^*, S_{c_q})$ and $E_{I_g}(X^*, S_{c_g})$.

Proof. Let J_q is the number of the $L_{J_q}(X^*, S_{c_q})$ level. By the lemma $\forall X \in O_i(X^*) \cap S_{c_g}, i \geqslant J_q : \quad f(X) > c_q$. Whence (as $X \in S_{c_g}$) $c_g > c_q$ appears.

COROLLARY 6. Let the unimodal nonmonotone (weakly nonmonotone) function f Φ constancy set is determined by the lemma conditions. Then there is at least one point $X \in (O_{I-1}(X^*) \cup O_I(X^*)) \setminus S_c$ such that $f(X) < c$.

Proof. Otherwise we have that S_c is an extended minimum, that contradicts to the supposition on the function f unimodality.

3. Algorithms and Efficiency

By Papadimitriou and Steiglitz (1982) a local search concrete algorithm is determined with four points:

– the choice of the initial point for search;

– the determination of the neighborhood size for the search steps;

– the determination of strategy of moving to a new point ;

– the determination of the step size for moving to a new point.

These components of an algorithm are specified in depending on an available *a priori* information about an objective function features. In our case we assume the lack of *a priori* information on the objective function. Consequently, it will be reasonable to offer the following simple scheme to organize a local search.

Algorithm 1.

1. Suppose $r = 0$. The point $X^r \in B_2^n$ is chosen arbitrarily. Compute $f(X^r)$.

2. Inverting sequentially the point X^r components we find all points $X_j^1 \in O_1(X^r),\ j = 1, ..., n$.

3. For all $X_j^1 \in O_1(X^r) \setminus O_1(X^{r-2})$ ($O_1(X^r) = \emptyset$ if $r < 0$) $f(X_j^1)$ are computated. If $f(X^r) < f(X_j^1) \forall j = 1, ..., n$ then $X^* = X^r$, otherwise go to 4.

4. Suppose $r = r + 1$, determine the point X^r from the condition

$$f(X^r) = \min_{X_j^1 \in O_1(X^{r-1})} f(X_j^1)$$

and go to 2.

Theorem 1. *To locate the minimum point X^* of an unimodal weakly nonmonotone on B_2^n function f from the initial point $X^0 \in O_k(X^*), k = 0, ..., n$, Algorithm* 1 *requires T_1 computations of the function f values.*

$$T_1 = \begin{cases} 2n + (k-1)(n-2), & k > 0, \\ n+1, & k = 0. \end{cases}$$

Proof. For $k = 0$ the estimate T_1 directly follows from Definition 3. Let $k > 0$. In accordance with the algorithm the values of f in the points X^0 and $X_j^1 \in O_1(X^0),\ j = 1, ..., n$, have to be computeted, i.e., $(n + 1)$ computations of f are made. By Definition 13 the point $X^1 \in O_1(X^0) \cap O_{k-1}(X^*)$. The values of f in the point $X^1 (X^1 \in O_1(X^0))$ and in one neighbouring to it point are known, i.e., it is necessary to do $(n - 1)$ computations of f in the point X^1 for determining a point X^2. The point $X^2 \in O_1(X^1) \cap O_{k-2}(X^*)$ (in accordance with Definition 13) and besides by Corollary 1 the values of f in two points of $O_1(X^2)$) are known too, i.e., in the point X^2 for determining a point X^3 it is necessary to do $(n - 2)$ computations of f. An analogous situation takes place in the points $X^3, X^4, ..., X^k (X^k = X^*)$ and to establish in the point X^k the fact that $X^k = X^*$ will be required $(n - 2)$ computations of f in addition. Summarizing we have: $T_1 = (n+1)+(n-1)+(k-1)(n-2) = 2n+(k-1)(n-2)$.

COROLLARY 7.

$$\max_{0 \leqslant k \leqslant n} T_1 = 2n + (n-1)(n-2) = n^2 - n + 2. \tag{17}$$

Proof. Evidently.

Theorem 2. *To locate the minimum point X^* of an unimodal different-valued weakly nonmonotone on B_2^n function f Algorithm 1 on the average requires T_2 computations of the function values.*

$$T_2 = (n^2 + 4)/2 - 1/(2^n).$$

Proof. In accordance with the algorithm the original point X^0 is chosen arbitrarily hence it is possible to assume for all $X \in B_2^n : \Pr\{X^0 = X\} = 1/(\operatorname{card} B_2^n) = 1/(2^n)$. Then taking into consideration $\operatorname{card} O_k(X^*) = C_n^k,\ k = 0, ..., n$, we have

$$\Pr\{X^0 \in O_k(X^*)\} = C_n^k/(2^n).$$

Finally, from the last relation and Theorem 1 we obtain the following expression for the mathematical expectation of the number of calculations of f required for locating X^* :

$$\begin{aligned} T_2 &= \frac{n+1}{2^n} + \sum_{k=1}^{n} \frac{C_n^k}{2^n}[2n + (k-1)(n-2)] \\ &= \frac{n+1}{2^n} + \frac{1}{2^n}[2n \sum_{k=1}^{n} C_n^k + (n-2) \sum_{k=1}^{n} kC_n^k - \sum_{k=1}^{n} C_n^k] \\ &= \frac{1}{2^n}(n + 1 + 2n(2^n - 1) + n(n-2)2^{n-1} - (n-2)2^n). \end{aligned}$$

Hence T_2 follows after simple transformations.

REMARK 7. The estimates T_1 and T_2 are correct as well when applying this algorithm for optimization of unimodal strictly monotone functions. In fact, as it follows from (3) and (5) for all $X^k \in O_k(X^*),\ k = 1, ..., n$, the point $X^1_{\min}$ (determined by the condition (7)) belongs to $O_{k-1}(X^*)$.

Lemmas 5 – 6 and the corollaries to them permit to assert that the algorithm is acceptable for optimization of arbitrary unimodal on B_n^2 functions. However, in this case we can not eliminate the situation when Algorithm 1 degenerates in a total examination. By this reason to obtain some estimate "on the average" makes sense.

Theorem 3. *For locating the minimum point X^* of an unimodal different-valued on B_2^n function f Algorithm 1 requires less than T_3 computations of the function f values when $n \geqslant 5$.*

$$T_3 = 2^n - (n^2 + 3n - 2)/2 + (2n^3 + n^2 - 4n + 3)/(2^n). \tag{18}$$

Proof. By analogy with the previous theorem we have

$$\Pr\{X^0 = X^*\} = 1/(2^n), \tag{19}$$

$$\Pr\{X^0 \in O_1(X^*)\} = n/(2^n), \tag{20}$$

$$\Pr\{X^0 \in O_1(X^k_{\min})\} = n/(2^n), \quad k = 1, ..., n, \tag{21}$$

where $X^0_{\min} = X^*$ and for $k > 0 \quad X^k_{\min}$ is determined from the condition (14).

Let us introduce a random value $t(X^0)$. This is the number of function computations, which Algorithm 1 requires to locate the point X^* from the initial point $X^0 \in B^2_n$.

The points $O_1(X^*) \cup \{X^*\}$ answer the conditions of Theorem 2. Thus,

$$t(X^0) = \begin{cases} n+1 & \text{when } X^0 = X^*, \\ 2n & \text{when } X^0 \in O_1(X^*), \end{cases} \tag{22}$$

and in accordance with Corollaries 1, 3 and Lemma 1:

$$t(X^0) \leqslant (n+1) + (n-k+1-1) + \sum_{i=0}^{k-1} \operatorname{card} O_i(X^*) - k,$$

when $X^0 = X^0_{\min},\ k = 2, ..., n;$

$$t(X^0) \leqslant (n+1) + (n-1) + (n-k+1-2) + \sum_{i=0}^{k-1} \operatorname{card} O_i(X^*) - k,$$

when $X^0 \in O_1(X^k_{\min}) \setminus \{X^*\},\ k = 1, ..., n.$

Next, taking into account that $\operatorname{card} O_i(X^*) = C^i_n,\ i = 0, ..., n,$

$$t(X^0) \leqslant 2n - 2k + 1 + \sum_{i=0}^{k-1} C^i_n, \text{ when } X^0 = X^0_{\min}, \quad k = 2, ..., n;$$

$$t(X^0) \leqslant 3n - 2k - 1 + \sum_{i=0}^{k-1} C^i_n, \text{ when } X^0 \in O_1(X^k_{\min}) \setminus \{X^*\},\ k = 1, ..., n.$$

The strict inequality in the last relations can be reached only for the situation when for all $k = 2, ..., n-2$:

$$O_1(X^{k-1}_{\min}) \cap \{X^k_{\min}\} = \emptyset, \tag{23}$$
$$O_1(X^{k+1}_{\min}) \cap \{X^k_{\min}\} = \emptyset, \tag{24}$$
$$O_1(X^{k-1}_{\min}) \cap \{X^{k+1}_{\min}\} = \emptyset. \tag{25}$$

Let $k = 2$. Then

$$\operatorname{Card}(O_1(X^1_{\min}) \cap O_2(X^*)) = n - 1, \quad \operatorname{card}(O_1(X^3_{\min}) \cap O_2(X^*)) = 3.$$

It leads to the condition

$$\operatorname{card} O_2(X^*) \geqslant n - 1 + 3 + 1 = n + 3$$

for fulfillment of (23) – (25) when $k = 2$. But $\operatorname{card} O_2(X^*) = C_n^2$ and we have the following inequality $n!/(2(n-2)!) \geqslant n + 3$ solving which with respect to n we will obtain: $n \geqslant 5$. The same is when $k = n - 2$.

From (19) – (25) in view of the fact that

$$\forall X^0 \in B_2^n \setminus \Big(\bigcup_{k=0}^{n} (O_1(X^k_{\min}) \cap \{X^k_{\min}\}) \Big) : \ t(X^0) \leqslant 2^n,$$

we have for the estimate of the mathematical expectation of the random value $t(X^0)$ "from above" when $n \geqslant 5$ the following relation:

$$T_3 = \frac{n+1}{2^n} + \frac{n2n}{2^n} + \frac{(n-1)(3n-2)}{2^n} + \frac{1}{2^n} \sum_{k=2}^{n} \Big[2n - 2k + 1 + \sum_{i=0}^{k-1} C_n^i \Big]$$
$$+ \frac{n}{2^n} \sum_{k=2}^{n} \Big[3n - 2k + \sum_{i=0}^{k-1} C_n^i - 1 \Big] + \frac{2^n - n^2 - 2n + 1}{2^n} 2^n.$$

As $\sum_{k=2}^{n} \sum_{i=0}^{k-1} C_n^i = n2^{n-1} - 1$ after simplest transformations from the last relation we obtain (18).

REMARK 8. It is easy to calculate that $T_3 = 6,26$ for $n = 3$ and $T_3 \approx 11,06$ for $n = 4$.

Next, we consider the case of pseudoboolean functions having constancy sets. First of all, note that Algorithm 1 is suitable if the condition

$$\forall X \in B_n^2 \setminus \{X^*\} \exists X^1 \in O_1(X) : \ f(X^1) < f(X) \tag{26}$$

is valid. This condition can be violated if the point X belongs to a certain constancy set. Consequently, to be able to optimize a pseudoboolean function admitting constancy sets we have to modify Algorithm 1 so as always to guarantee obtaining the situation (26) with least expenditures.

Algorithm 2.

1. Suppose that $r = 0$. The point $X^r \in B_2^n$ is chosen arbitrarily. Compute $f(X^r)$.
2. Inverting sequentially the point X^r components we find all points $X_j^1 \in O_1(X^r), \ j = 1, ..., n$. Computate the unknown values $f(X_j^r)$.

3. Form the set $V(X^r) = \{j \in \{1, ..., n\} : f(X^r) = f(X^r_j)\}$.
   If $\operatorname{card} V(X^r) = 0$ then go to 4, if $0 < \operatorname{card} V(X^r) < n$ then go to 6, if $\operatorname{card} V(X^r) = n$ then suppose $t = 2$ and go to 8.
4. If $f(X^r) < f(X^r_j)\ \forall j = 1, ..., n$ then $X^* = X^r$ and stop.
5. Determine $X^1_{\min}$ from the condition

$$f(X^1_{\min}) = \min_{X^1_j \in O_1(X^r)} f(X^1_j). \tag{27}$$

   Suppose $r = r + 1,\ X^r = X^1_{\min}$ and go to 2.
6. Determine $X^1_{\min}$ from the condition (27). If $f(X^1_{\min}) < f(X^r)$ then suppose $r = r + 1,\ X^r = X^1_{\min}$ and go to 2.
7. $\forall X^r_j, j \in V(X^r)$ invert sequentially the component values to find all points $X^{jr}_i \in O_1(X^r_j),\ j = 1, ..., n$. Computate the unknown values $f(X^{jr}_i)$. Determine $X^1_{\min}$ from the condition

$$f(X^1_{\min}) = \min_{j \in V(X^r)} \min_{X^{jr}_i \in O_1(X^r_j)} f(X^{jr}_i).$$

   If $f(X^1_{\min}) < f(X^r)$ then suppose $r = r + 1,\ X^r = X^1_{\min}$ and go to 2, otherwise suppose $t = 2$ and go to 8.
8. Invert t components of the point X^r to find all points $X^t_j \in O_t(X^r),\ j = 1, ..., C^t_n$. Sequentially calculate the unknown values $f(X^t_j)$: if $f(X^t_j) < f(X^r)$ then suppose $r = r + 1,\ X^r = X^t_j$ and go to 2, otherwise continue the calculations.
9. If $f(X^t_j) > f(X^r)\ \forall j = 1, ..., C^t_n$ then suppose $X^* = X^r$ and stop, if $f(X^t_j) = f(X^r)\ \forall j = 1, ..., C^t_n$ then go to 10.
10. Suppose $t = t + 1$, if $t > n$ then $X^* = X^r$ and stop, otherwise go to 8.

Let us give some explanation to this algorithm.

When realizing the point 3 the algorithm locates the search current point position with respect to possible constancy sets. If $\operatorname{card} V(X^r) = 0$ then the condition (26) is executed and the operation of Algorithm 1 and Algorithm 2 is identical. If $\operatorname{card} V(X^r) = n$ then the strategy of going out of a constancy set is realized (the points 8–10 of the algorithm). The case when $0 < \operatorname{card} V(X^r) < n$ corresponds to the situation when X^r is a boundary point of a certain constancy set S_c or $X^r \in O_1(S_c)$. For this case some additional test on the condition (26) fulfilling in the point X^r is realized (the point 6 of the algorithm). The points 4 and 9 permit to establish the fact that a current point of search is a minimum point.

Estimate the efficiency of Algorithm 2. In the case of a different-valued function or a function having only two-level Φ constancy sets, the efficiency estimates for Algorithm 2 will coincide with the corresponding estimates for Algorithm 1 (Theorems 1, 2 and Corollary 7). Similarly, these estimates will be correct and for the case when an unimodal weakly nonmonotone function is a different-valued one on the set $\bigcup_{i=0}^{I-1} O_i(X^*)$, where

I is the number of the first (with respect to X^*) points level of an arbitrary Φ constancy set, and

$$X^0 \in \Big(\bigcup_{i=0}^{I} O_i(X^*)\Big) \cup \big(B(S_c) \cap O_{I+1}(X^*)\big).$$

Here $B(S_c)$ is the set of the set S_c boundary points.

For the rest cases the following theorems give the estimates of efficiency.

Theorem 4. *Let*

$$S_c = \bigcup_{l=I}^{J} O_l(X^*), \quad I > 1,\ I+1 < J \leqslant n, \tag{28}$$

is a constancy set of a weakly nonmonotone function f and $X^0 \in O_k(X^*),\ I < k \leqslant J$ is the search initial point. Then Algorithm 2 requires to computate the f values not more than in T_4 points of B_2^n to locate the minimum point.

$$T_4 = \sum_{l=0}^{k-I+1} C_n^l - E[C_k^I/2] + (I-1)(n-1), \tag{29}$$

where $[\alpha]$ means the least integer greater than α.

Proof. As $I < k \leqslant J$ then

$$\bigcup_{l=0}^{k-I} O_l(X^0) \subset \Big(S_c \cup \big(\bigcup_{l=J+1}^{n} O_l(X^*)\big)\Big).$$

From Lemma 9 we have

$$\forall X \in \bigcup_{l=J+1}^{n} O_l(X^*):\ f(X) > c, \quad \text{i.e.,} \forall X \in \bigcup_{l=0}^{k-I} O_l(X^0):\ f(X) \leqslant c.$$

Thus, in accordance with of points 8 – 10 the algorithm

$$T_4^1 = \sum_{l=0}^{k-I} \operatorname{card} O_l(X^o) = \sum_{l=0}^{k-I} C_n^l, \tag{30}$$

computations of f will be done. Next, the values of f in the points of $O_{k-I+1}(X^0)$ are computated. In accordance with Corollary 1 $\operatorname{card}(O_{k-I}(X^0) \cap O_I(X^*)) = C_k^I$, and what is more, as it follows from Definition 14 $\forall X^I \in O_{k-I}(X^0) \cap O_I(X^*)$ at least one point $X^{I-1} \in O_1(X^I) \cap O_{k-I+1}(X^0) \cap O_{I-1}(X^*)$ such that $f(X^{I-1}) < f(X^I)$ has to exist. By Lemma 3 the set $O_{k-I}(X^0) \cap O_I(X^*)$ consists of 2-, 4-, 6- etc.

neigbouring points. Then taking into account Corollary 1 we can assert that not less than $E[C_k^I/2]$ points, in which the function values are less than c, exist among the points of $O_{k-I+1}(X^0)$. Thus, in accordance with the point 8 of the algorithm not more than

$$T_4^2 = C_n^{k-I+1} - E[C_k^I/2], \tag{31}$$

computations of f will be done additionally to determine the point $X^{I-1} \in O_{I-1}(X^*)$ such that $f(X^{I-1}) < c$. Further the work of Algorithm 1 and Algorithm 2 is identical and we can use the estimate T_1 (taking into consideration that the function values in $n-I+1$ points of the set $O_I(X^{I-1}) \cap O_I(X^*)$ and in the point X^{I-1} are known). Consequently, to locate X^* the algorithm has to do

$$T_4^3 = (I-1)(n-1), \tag{32}$$

computations of f in addition. Uniting (30) – (32) we have (29).

COROLLARY 8.

$$\max_{I<k\leqslant J} T_4 = \sum_{l=0}^{J-I+1} C_n^l - E[C_J^I/2] + (n-1)(I-1).$$

Proof. $\max_{I<k\leqslant J}(k-I+1) = J-I+1$.

COROLLARY 9.

$$\max_{1<I<J\leqslant n} \max_{I<k\leqslant J} T_4 = \sum_{l=0}^{n-1} C_n^l - E[C_n^2/2] + n - 1.$$

Proof. $\max_{1<I<J\leqslant n}(J-I+1) = n-1$.

COROLLARY 10. If $I=1$ then $T_4 = \sum_{l=0}^{k} C_n^l$.

Proof. When $I=1$ all points $X_j^1 \in O_1(X^*),\ j=1,...,n$, are 2-neighbouring points and

$$\Big(\bigcup_{j=1}^{n} O_1(X_j^1)\Big) \setminus \Big(\Big(\bigcup_{j=1}^{n} O_1(X_j^1)\Big) \cap S_c\Big) = \{X^*\},$$

i.e., in accordance with the algorithm the function values in all points of the set $O_{k-I+1}(X^0)$ are computated. Then the point X^* is determined without additional computations.

Finally we consider the case of an arbitrary function having Φ constancy sets.

Denote by T_5 the efficiency estimate "on the average" of the algorithm for this case. As the algorithm efficiency estimates for several Φ constancy sets will be less (it follows from Corollary 3 and the fact that if $s = s_1 + s_2, s_1 > 2, s_2 > 2$ then $\sum_{l=0}^{s} C_n^l > \sum_{l=0}^{s_1} C_n^l + \sum_{l=0}^{s_2} C_n^l,\ s \leqslant n-1$) we consider a function having only constancy set $S_c = \bigcup_{l=I}^{J} O_l(X^*),\ 1 \leqslant I < J \leqslant n$.

With respect to the location of S_c in B_2^n the following essential for us situations are possible:

1. $I = 1,\ J = n$;
2. $I > 1,\ J = n$;
3. $I = 1,\ J < n$;
4. $I > 1,\ J < n$.

The Situation 1 conforms to the conditions of Theorem 3 and therefore in this case $T_5 = T_3$. For the Situation 2 with the probability $P_1 = (1 + n + I - 2 + (I-2)(n-2))/(2^n)$:

$$T_5 \leqslant 3n - 2I + \sum_{i=0}^{I-1} C_i^n - 1 < 2^n, \tag{33}$$

and with the probability $P_2 = 1 - P_1 : T_5 \leqslant T_3$, i.e., for the Situation 2 as well $T_5 < 2^n$. For the Situation 3, with the probability $P_1 = (\operatorname{card} S_c + 1)/(2^n) : T_5 \leqslant T_3$ and with the probability $P_2 = 1 - P_1 : T_5 \leqslant \operatorname{card}(\cup_{l=0}^{n} O_l(X^*)) = 2^n$, i.e., $T_5 \leqslant 2^n$. In the Situation 4, with the probability $P_1 = (1 + n + (I-2)(n-1))/(2^n)$ (33) is valid and with the probability $P_2 = (\operatorname{card} S_c)/(2^n) : T_5 \leqslant T_3$. Thus, "on the average" Algorithm 2 always requires less than 2^n computations of the function values.

4. Concluding Remarks

In this work we explored the possibilities of local search techniques for the exact determination of an arbitrary unimodal pseudoboolean function minimum. For the studied problem we assumed that the objective function is not known explicitly although usually other authors deal with objective functions given in an evident analytical form (see, e.g., Gramma *et al.*, 1990). The point is that such a kind of problems is important for authors' area of application, i.e., automatization of design processes in spacecrafts production. These design processes can be modeled in the form of pseudoboolean functions optimization with objectives given as simulated programs or real process outputs.

Direct local search algorithms were described and analytically investigated. These algorithms are applicable to any unimodal objective functions even if they have big regions of constancy what is usually a great problem for search discrete optimization methods. Convergence of the algorithms has been proved and analytical estimates of the rate of convergence have been obtained. The proposed local search algorithms will fulfill a total search in some cases. However, these cases are very special (e.g., if an objective function is constant in almost every point) and unimportant for practice. In any case, it is impossible to suggest any better algorithm for such a function. The average speed of convergence of the proposed algorithms is less than total search for any other forms of objective functions. In case of a monotone objective function, the proposed algorithms have a quadratic rate of convergence. For nonmonotone functions these algorithms have an exponential rate of convergence, that is they are of limited applicability in real problems with high dimension. Nevertheless, they are useful as they admit optimization in many cases when there are no other suitable algorithms.

The natural question can appear: if the explicit form of an objective function is unknown what are guarantees that it is an unimodal function? First of all, unimodality of an objective function fairly often can be based with some practical reasons. In the rest of cases the algorithms will supply a local minimum subject to the search original point location. Apperantly, estimates of the expected time of search for these cases will be less essentially than the estimates obtained above.

References

Antamoshkin, A.N., V.N. Saraev and E.S. Semenkin (1990). Optimization of unimodal monotone pseudoboolean function. *Kybernetika*, **26**(5), 432–442.

Antamoshkin, A., and E. Semenkin (1991). Optimization of polimodal locally monotone pseudoboolean functions. *Informatica*, **3**(2), 331–351.

Auselo, G., and M. Protasi (1995). Local search, reducibility and approximability of NP-optimization problems. *Information Processing Letters*, **54**(2), 73–79.

Gramma, Y., P. Hansen and B. Jaumard (1990). The basic algorithm for pseudo-boolean programming revisited. *Discrete Applied Mathematics*, 29, 171–185.

Nemhauser,G.L., L.A.Wolsey and M.L.Fisher (1978). An analysis of approximations for maximazing submodular set functions - I. *Mathematical Programming*, **14**, 265–294.

Papadimidriou, C.H., and K. Steiglitz (1982). *Combinatorial Optimization: Algorithms and Complexity*, Prentice-Hall, Englewood Cliffs, NJ.

Rudolf, G. (1996). *Convergence Properties of Evolutionary Algorithms*, Dissertation des Doktors Naturwissenschaften, Universität Dortmund, Dortmund.

A. Antamoshkin is a doctor of technical sciences, a professor of Siberian Aerospace Academy (Krasnoyarsk, Russia). His research interests are optimization techniques, modelling, control systems.

E. Semenkin is a doctor of technical sciences. He is a professor of Systems Analysis and Operations Research Department of Siberian Aerospace Academy (Krasnoyarsk, Russia). His research interests include adaptive search and artificial intellegence techniques in optimization of complex systems with application area in development of decision support systems for design of spacecrafts control systems and implementation of control systems for complex technological processes.

Lokalinės paieškos efektyvumas optimizuojant unimodalias psiaudo-bulines funkcijas

Alexander ANTAMOSHKIN, Eugene SEMENKIN

Šis darbas yra tęsinys straipsnių, skirtų reguliarios paieškos procedūrų efektyvumo tyrimui. Čia suformuluojamas bendresnis uždavinys, kuomet optimizuojama unimodali psiaudo-bulinė funkcija yra analitiškai išreikšta, ir gaunami analitiniai įverčiai laukiamų laiko sąnaudų minimumo paieškos, naudojantis tiesioginės lokalinės paieškos procedūromis. Parodyta, kuomet šie įverčiai yra polinominiai, o kuomet eksponentiniai. Tikimasi, kad gauti rezultatai bus ypač naudingi sprendžiant praktinius uždavinius.