FINITE-TERMINATION SCHEMES FOR SOLVING SEMI-INFINITE SATISFYCING PROBLEMS

by

E. Polak and L. He

Memorandum No. UCB/ERL M90/13

22 February 1990

COVER PAGE

FINITE-TERMINATION SCHEMES FOR SOLVING SEMI-INFINITE SATISFYCING PROBLEMS

by

E. Polak and L. He

Memorandum No. UCB/ERL M90/13

22 February 1990

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

TITLE PAGE

FINITE-TERMINATION SCHEMES FOR SOLVING SEMI-INFINITE SATISFYCING PROBLEMS

by

E. Polak and L. He

Memorandum No. UCB/ERL M90/13

22 February 1990

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

FINITE-TERMINATION SCHEMES FOR SOLVING SEMI-INFINITE SATISFYCING PROBLEMS*

by

E. Polak** and L. He**

ABSTRACT

The problem of finding a parameter which satisfies a set of specifications in inequality form is sometimes referred to as the satisfycing problem. We present a family of methods for solving, in a finite number of iterations, satisfycing problems stated in the form of semi-infinite inequalities. These methods range from adaptive uniform discretization methods to outer approximation methods.

KEY WORDS

Adaptive approximation, outer approximations, finite termination, semi-infinite programming, satisfycing problems.

* The research reported herein was sponsored in part by the National Science Foundation grant ECS-8713334, the Air Force Office of Scientific Research contract AFOSR-86-0116 and the State of California MICRO Program grant 532410-19900.

** Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory, University of California, Berkeley, CA 94720, U.S.A.

1. INTRODUCTION

Both engineering and economic system design problems often involve specifications in terms envelopes within which dynamic responses must be contained (see, e.g. [Pol.2],[Wie.1]). The mathematical expression for these specifications is likely to be of the form of a system of semi-infinite inequalities, such as

$$\max_{y_j \in \mathrm{Y}_j} \phi^j(x,y_j) \le b_j, \; j = 1,2,\dots,l, \tag{1.1a}$$

where $x \in \mathbb{R}^n$, the sets $\mathrm{Y}_j \triangleq [l_j,u_j]$ are compact intervals, all the functions are locally Lipschitz continuous, and the b_j express the desired *satisfycing* level. The satisfycing problem inequalities are obviously equivalent to the more compact form

$$\psi(x) \le 0, \tag{1.1b}$$

where, with $\mathbf{l} \triangleq \{ 1,2,\dots,l \}$,

$$\psi(x) \triangleq \max_{j \in \mathbf{l}} \max_{y_j \in \mathrm{Y}_j} [\, \phi^j(x,y_j) - b_j \,] . \tag{1.1c}$$

It is immediately clear from (1.1b) that whenever there exists an $\hat{x} \in \mathbb{R}^n$ such that $\psi(\hat{x}) < 0$, then any *conceptual* semi-infinite minimax descent algorithm, such as the Pironneau-Polak-Pshenichnyi algorithm ([Pir.1], [Pol.2], [Psh.1]) or *conceptual* outer approximations algorithm, such as the one in [Gon.1, May.1], are capable of finding a solution to (1.1b) in a finite number of iterations. The reason we insist that in the context of semi-infinite inequalities these algorithms are only *conceptual*, is that neither $\psi(x)$ nor any of the other quantities which these algorithms require at each iteration can be computed without discretization of the sets Y_j. Thus we see that the main issues are (i) *computational efficiency* of the implementation and (ii) the development of special features which will result in finite termination using finite precision.

In this paper we present a fairly general, implementable master algorithm for solving semi-infinite inequalities in a finite number of iterations under fairly mild assumptions. The master algorithm constructs approximating inequalities by discretizing the original inequalities and then invokes a descent method to try to solve these approximating inequalities. The finite termination property results from the fact that whenever possible, the master algorithm runs the descent method for one iteration over what is needed to solve the current approximating inequalities, and the use of Lipschitz constants which enable one to conclude that the original set of inequalities is satisfied from the fact that a discretized set of inequalities is satisfied. The required Lipschitz constants can be estimated in the course of algorithm execution. The discretizations are performed using either an outer approximations form of construction (see [Gon.1, May.1]), or an adaptive uniform discretization scheme (see [Pol.1, He.1]). Our numerical results show that even without the inclusion of constraint dropping schemes, such as those described in [Gon.1], the form of the algorithm based on outer approximations is considerably more efficient than the one based on adaptive uniform discretization scheme.

2. THE MASTER ALGORITHM

We will consider the following, normalized, semi-infinite satisfycing problems:

$$\textbf{SP}: \quad \text{find } x^* \in \mathbb{R}^n \text{ such that } \psi(x^*) \leq 0 \,. \tag{2.1a}$$

where

$$\psi(x) \triangleq \max_{y \in \mathbf{Y}} \phi(x,y) \,, \tag{2.1b}$$

$$\phi(x,y) \triangleq \max_{j \in \mathbf{l}} \phi^j(x,y) \,, \tag{2.1c}$$

where $\mathbf{l} \triangleq \{ 1,2,\dots,l \}$, $\mathbf{Y} = [0,1]$, and $\phi^j : \mathbb{R}^n \times \mathbf{Y} \to \mathbb{R}$. We will assume that the functions $\phi^j(\cdot,\cdot)$ and their gradients $\nabla_x\phi^j(\cdot,\cdot)$, are locally Lipschitz continuous.

The normalized form (2.1a) is obtained from the general form (1.1b) by (a) absorbing the constants b^j into the definition of the functions $\phi^j(\cdot,\cdot)$, and (b) by changing the definition of $\phi^j(\cdot,\cdot)$ so that $\phi^j_{new}(x,y) = \phi^j_{old}(x, l_j + y/[u_j - l_j])$, which then requires that $y \in \mathbf{Y} \triangleq [0,1]$.

Since, in general, the exact calculation of the global maximum of $\phi(x,\cdot)$ over the interval $\mathbf{Y}$ is not a numerically feasible operation, numerical methods for solving the problem **SP** must involve a scheme for the discretization of the interval $\mathbf{Y}$. Hence, for any finite discrete set $\mathbf{D} \subset \mathbf{Y}$, we define a problem $\mathbf{SP_D}$, whose solutions approximate those of problem **SP**, by

$$\mathbf{SP_D}: \ \text{find } x^*_{\mathbf{D}} \in \mathbb{R}^n \text{ such that } \psi_{\mathbf{D}}(x^*_{\mathbf{D}}) \le 0, \tag{2.2a}$$

where

$$\psi_{\mathbf{D}}(x) \triangleq \max_{y \in \mathbf{D}} \phi(x,y). \tag{2.2b}$$

The master algorithm for solving the problem **SP** that we will shortly introduce, calls the Pironneau-Polak-Pshenichnyi (PPP) minimax algorithm [Pir.1, Pol.2, Psh.1] as a subroutine. The PPP algorithm computes search directions by evaluating an optimality function. Hence, proceeding as in [Pol.2], we define the *optimality functions* $\theta(x)$, $\theta_{\mathbf{D}}(x)$, for the problems **SP** and $\mathbf{SP_D}$, respectively, as follows:

$$\theta(x) \triangleq \min_{h \in \mathbb{R}^n} \max_{y \in \mathbf{Y}} \max_{j \in \mathbf{I}} \{ \phi^j(x,y) + \langle \nabla_x \phi^j(x,y), h \rangle + \tfrac{1}{2}\|h\|^2 - \psi(x) \}, \tag{2.3a}$$

$$\theta_{\mathbf{D}}(x) \triangleq \min_{h \in \mathbb{R}^n} \max_{y \in \mathbf{D}} \max_{j \in \mathbf{I}} \{ \phi^j(x,y) + \langle \nabla_x \phi^j(x,y), h \rangle + \tfrac{1}{2}\|h\|^2 - \psi_{\mathbf{D}}(x) \}. \tag{2.3b}$$

Note that because the discrete set **D** is finite, (2.3b) is an ordinary quadratic programming problem which can be solved finitely using standard quadratic programming subroutines.

Let $h, h_{\mathbf{D}}: \mathbb{R}^n \to \mathbb{R}$ be the PPP *search direction functions* defined by

$$h(x) \triangleq \arg \min_{h \in \mathbb{R}^n} \max_{y \in \mathbf{Y}} \max_{j \in \mathbf{I}} \{ \phi^j(x,y) + \langle \nabla_x \phi^j(x,y), h \rangle + \tfrac{1}{2}\|h\|^2 - \psi(x) \}, \tag{2.3c}$$

$$h_{\mathbf{D}}(x) \triangleq \arg \min_{h \in \mathbb{R}^n} \max_{y \in \mathbf{D}} \max_{j \in \mathbf{I}} \{ \phi^j(x,y) + \langle \nabla_x \phi^j(x,y), h \rangle + \tfrac{1}{2}\|h\|^2 - \psi_{\mathbf{D}}(x) \}. \tag{2.3d}$$

In the case of problem **SP**, the PPP minimax algorithm is only *conceptual* ; it uses the search direction function $h(\cdot)$. In the case of problem $\mathbf{SP_D}$ the PPP minimax algorithm is *implementable*; it uses the search direction function $h_{\mathbf{D}}(\cdot)$. In both cases, the PPP algorithm uses an Armijo type step size rule, which requires two parameters $\alpha, \beta \in (0,1)$.

The master algorithm, below, uses two sequences of discrete subsets of **Y**: one sequence, $\{ \mathbf{E}_i \}_{i=0}^{\infty}$, is used in conjunction with the PPP algorithm, while the second sequence, $\{ \mathbf{C}_i \}_{i=0}^{\infty}$, is used in the stopping criterion. We will consider two alternative schemes for constructing the sequence $\{ \mathbf{E}_i \}_{i=0}^{\infty}$. We choose the sequence $\{ \mathbf{C}_i \}_{i=0}^{\infty}$ to be a strictly increasing sequence of sets consisting of uniformly spaced $c_i \ge 2$ grid points in $[0,1]$, i.e., $\mathbf{C}_i \subset \mathbf{C}_{i+1}$ and $\mathbf{C}_i \triangleq \{ k/(c_i - 1) \mid k = 0,1,\dots,c_i - 1 \}$.

We assume that for x in a sufficiently large subset of $\mathbb{R}^n$, the Lipschitz constant of the functions $\phi^j(\cdot,\cdot)$ with respect to y is $L < \infty$. Note that in view of this fact, if $\psi_{C_i}(x) + L/(2(c_i - 1)) \leq 0$, then $\psi(x) \leq 0$ must also hold.

Assumption 2.1: We assume: (i) The functions $\phi^j(\cdot,\cdot)$, their gradients $\nabla_x\phi^j(\cdot,\cdot)$ and $\nabla\phi^j_{xx}(\cdot,\cdot)$ are continuous for $j \in \mathbf{l}$; (ii) For any $x \in \mathbb{R}^n$ such that $\psi(x) \geq 0$, $\theta(x) < 0$. ■

Note that Assumption 2.1 (ii) implies (i) that there exists an $x^* \in \mathbb{R}^n$ such that $\psi(x^*) < 0$ and hence, in turn that $\psi_{\mathbf{E}}(x^*) < 0$ for any discrete subset $\mathbf{E} \subset \mathbf{Y}$, and (ii) that the conceptual PPP algorithm is guaranteed to find such a point. However, it does not ensure that the implementable PPP algorithm is able to find such a point when applied to the problem $\min_{x \in \mathbb{R}^n} \psi_{\mathbf{E}}(x)$. That would require the much stronger assumption that for any discrete subset $\mathbf{E} \subset \mathbf{Y}$, for any $x \in \mathbb{R}^n$ such that $\psi_{\mathbf{E}}(x) \geq 0$, $\theta_{\mathbf{E}}(x) < 0$. This fact is reflected in the problem $\mathbf{SP}_{\mathbf{E}_i}$ abandonment test (2.5b) in the master algorithm below.

The master algorithm below, constructs a problem $\mathbf{SP}_{\mathbf{E}_i}$ and applies the PPP minimax algorithm to it until it finds a point $x_i^{n_i-1}$ such that $\psi_{\mathbf{E}_i}(x_i^{n_i-1}) \leq 0$. Then it performs one more iteration to construct a point $x_i^{n_i}$ such that $\psi_{\mathbf{E}_i}(x_i^{n_i}) < 0$. Although this is not transparent from our proofs, because we use contradiction arguments, the finite termination property of the master algorithm below depends on the fact that there exists a constant $\delta > 0$ such that, after a while, $\psi_{\mathbf{E}_i}(x_i^{n_i}) \leq -\delta$ must hold.

In the master algorithm, the index i counts the number of times the approximations to $\mathbf{Y}$ (i.e., $\mathbf{E}_i$) have been updated, the index j counts the number of iterations performed by the PPP minimax algorithm on the problem $\mathbf{SP}_{\mathbf{E}_i}$, and the index n_i counts the number of iterations performed in solving the problem $\mathbf{SP}_{\mathbf{E}_i}$.

Master Algorithm 2.1 (for problem SP):

Data: $x_0 \in \mathbb{R}^n$, a sequence of sets $\{ \mathbf{C}_i \}_{i=0}^{\infty} \subset \mathbf{Y}$, a finite discrete set $\mathbf{E}_0 \subset \mathbf{Y}$, α, $\beta \in (0,1)$, and a sequence $\{ \varepsilon_i \}_{i=0}^{\infty}$, such that $\varepsilon_i > 0$, $\varepsilon_i \to 0$ as $i \to \infty$.

Step 0: Set $x_0^0 = x_0$, $i = 0$ and $j = 0$.

Step 1: Compute $\psi_{\mathbf{E}_i}(x_i^j)$, $\theta_{\mathbf{E}_i}(x_i^j)$, and $h_{\mathbf{E}_i}(x_i^j)$.

Step 2: Compute the step size λ_i^j by the Armijo-like rule:

$$\lambda_i^j = \max\{\ \beta^k \mid k \in \mathbb{N}\ ,\ \psi_{\mathbf{E}_i}(x_i^j + \beta^k h_{\mathbf{E}_i}(x_i^j)) - \psi_{\mathbf{E}_i}(x_i^j) \le \alpha\beta^k\theta_{\mathbf{E}_i}(x_i^j)\ \}\ . \tag{2.4}$$

Set $x_i^{j+1} = x_i^j + \lambda_i^j h_{\mathbf{E}_i}(x_i^j)$.

Step 3: Consider three exclusive cases:

(a) If

$$\psi_{\mathbf{E}_i}(x_i^j) > 0 \text{ and } -\theta_{\mathbf{E}_i}(x_i^j) > \varepsilon_i\ , \tag{2.5a}$$

set $j = j+1$, and go to Step 1.

(b) If

$$\psi_{\mathbf{E}_i}(x_i^j) > 0 \text{ and } -\theta_{\mathbf{E}_i}(x_i^j) \le \varepsilon_i\ , \tag{2.5b}$$

go to Step 5.

(c) If

$$\psi_{\mathbf{E}_i}(x_i^j) \le 0, \tag{2.5c}$$

go to Step 4.

Step 4: Stopping criterion: If

$$\psi_{\mathbf{C}_i}(x_i^{j+1}) + L/(2(c_i - 1)) \le 0\ , \tag{2.6}$$

set $n_i = j+1$ and stop. Else, go to Step 5.

Step 5: Construct a discrete set $\mathbf{E}_{i+1}$ satisfying $\mathbf{E}_i \subset \mathbf{E}_{i+1} \subset \mathbf{Y}$, $\mathbf{E}_{i+1} \ne \mathbf{E}_i$. Set $n_i = j+1$, $x_{i+1}^0 = x_i^{n_i}$, $i = i+1$ and $j = 0$, and go to Step 1. ■

Note that when (2.6) is satisfied, we must have that $\psi(x_i^{j+1}) \le 0$, and hence we have found a solution to our problem.

We will consider two schemes for constructing the sequence of discrete sets $\mathbf{E}_i$. The first is a uniform discretization scheme, while the second one is that used in outer approximations algorithms (see, e.g., [Gon.1]).

Adaptive Uniform Discretization Scheme 2.1:

In this scheme we set $\mathbf{E}_i$ to be a set of $2^i + 1$ uniformly spaced points in the interval $\mathbf{Y}$, i.e.,

$$\mathbf{E}_i \triangleq \{\, k/2^i \mid k = 0,1,\ldots,2^i \,\},\ i = 0,1,2,\ldots \tag{2.7}$$

■

Outer Approximation Scheme 2.2:

This scheme is recursive: Given $i \in \mathbb{N}$, a finite discrete set $\mathbf{E}_i \subset \mathbf{Y}$ and $x_i^{n_i} \in \mathbb{R}^n$, find a $y_i \in \mathbf{C}_i$ such that

$$\phi(x_i^{n_i}, y_i) = \psi_{\mathbf{C}_i}(x_i^{n_i}), \tag{2.8a}$$

and set

$$\mathbf{E}_{i+1} = \mathbf{E}_i \bigcup \{\, y_i \,\}\,. \tag{2.8b}$$

■

Lemma 2.1: Let $\{\, z_i^1 \,\}_{i=0}^{\infty}$ and $\{\, z_i^2 \,\}_{i=0}^{\infty}$ be any two converging sequences in $\mathbb{R}^n$ with $z_i^1 \to \hat{z}$ and $z_i^2 \to \hat{z}$ as $i \to \infty$. Let $\Omega_i \subset \mathbf{Y}$, $i = 0,1,\ldots$, be any sequence of compact sets contained in $\mathbf{Y}$. Then

$$|\psi_{\Omega_i}(z_i^1) - \psi_{\Omega_i}(z_i^2)| \to 0 \quad \text{as } i \to \infty\,. \tag{2.9}$$

Proof: Suppose that (2.9) does not hold. Then there exists a $\delta > 0$ and an infinite set $K \subset \mathbb{N}$ such that

$$|\psi_{\Omega_i}(z_i^1) - \psi_{\Omega_i}(z_i^2)| \geq \delta, \quad \text{for all } i \in K\,. \tag{2.10}$$

Now, $\psi_{\Omega_i}(z_i^1) = \phi(z_i^1, \omega_i^1)$ and $\psi_{\Omega_i}(z_i^2) = \phi(z_i^2, \omega_i^2)$ for some ω_i^1, $\omega_i^2 \in \Omega_i$. Without loss of generality, we may assume that

$$\phi(z_i^1, \omega_i^1) \geq \phi(z_i^2, \omega_i^2) + \delta\,. \tag{2.11}$$

Since $z_i^1 \to \hat{z}$, $z_i^2 \to \hat{z}$, and $\mathbf{Y}$ is a compact set, it follows from the continuity of $\phi(\cdot,\cdot)$, which is uniform in y, that there exists an i_0 such that

$$\phi(z_i^2, \omega_i^1) \geq \phi(z_i^1, \omega_i^1) - \delta/2 \quad \text{for all } i \in K,\ i \geq i_0\,. \tag{2.12}$$

But (2.11) and (2.12) imply that ω_i^2 is not a maximizer of $\phi(z_i^2, \cdot)$ over Ω_i, which is a contradiction.

Hence (2.9) is true. ■

Lemma 2.2: For any discrete set $\mathbf{D} \in \mathbf{Y}$ and any $x \in \mathbb{R}^n$,

(i) $\theta_{\mathbf{D}}(x) + \psi_{\mathbf{D}}(x) \le \theta(x) + \psi(x)$, (2.13a)

(ii) $\|h_{\mathbf{D}}(x)\|^2 \le -2\theta_{\mathbf{D}}(x)$. (2.13b)

Proof: (i) The inequality (2.13a) follows directly from (2.3a-b) and the fact that $\mathbf{D} \subset \mathbf{Y}$.

(ii) Referring to Theorem 5.6 in [Pol.2], we see that the application of the von Neumann minimax theorem to (2.3b) yields the fact that $-\theta_{\mathbf{D}}(x) = h_{\mathbf{D}}^0(x) + \frac{1}{2}\|h_{\mathbf{D}}(x)\|^2$, where $h_{\mathbf{D}}^0(x) \ge 0$. Hence (2.13b) follows immediately. ■

Theorem 2.1: Suppose that the sequence $\{ x_i^j \}_{i=0}^{\infty}{}_{j=0}^{n_i}$ is a sequence constructed by Algorithm 2.1 using either the Adaptive Uniform Discretization Scheme 2.1 or the Outer Approximation Scheme 2.2. Then

(i) n_i is finite for all i.

(ii) For any converging subsequence $\{ x_i^{n_i} \}_{i \in K}$, with $x_i^{n_i} \xrightarrow{K} \hat{x}$,

$$\psi_{\mathbf{E}_i}(x_i^{n_i}) \xrightarrow{K} \psi(\hat{x}) . \tag{2.14}$$

Proof: (i) Suppose that n_i is not finite for all i. Then, the tests (2.5b,c) were satisfied only a finite number of times. Hence there exists i_0 such that

$$\psi_{\mathbf{E}_{i_0}}(x_{i_0}^j) > 0 \, , \quad -\theta_{\mathbf{E}_{i_0}}(x_{i_0}^j) > \varepsilon_{i_0} \, , \quad \text{for all } j = 0, 1, ... \tag{2.15}$$

Making use of (2.4) and (2.15), we obtain that

$$-\infty < \sum_{j=0}^{\infty} [\psi_{\mathbf{E}_{i_0}}(x_{i_0}^{j+1}) - \psi_{\mathbf{E}_{i_0}}(x_{i_0}^j)] \le \sum_{j=0}^{\infty} \alpha\lambda_{i_0}^j \theta_{\mathbf{E}_{i_0}}(x_{i_0}^j) . \tag{2.16}$$

Referring to Lemma 2.2 (ii) and (2.15), we deduce that $\|h_{\mathbf{E}_{i_0}}(x_{i_0}^j)\|^2 \le 2\theta_{\mathbf{E}_{i_0}}(x_{i_0}^j)^2/\varepsilon_{i_0}$ for all $j \ge 0$. Hence, because of (2.16), for any positive integers $k > s \ge 0$,

$$\|x_{i_0}^k - x_{i_0}^s\| \le \sum_{j=s}^{k-1} \|x_{i_0}^{j+1} - x_{i_0}^j\| \le \sum_{j=s}^{k-1} \lambda_{i_0}^j \|h_{\mathbf{E}_{i_0}}(x_{i_0}^j)\| \le \sum_{j=s}^{\infty} (2/\varepsilon_{i_0})^{1/2} \lambda_{i_0}^j (-\theta_{\mathbf{E}_{i_0}}(x_{i_0}^j)) . \tag{2.17}$$

Therefore, $\{ x_{i_0}^j \}_{j=0}^{\infty}$ is a Cauchy sequence in $\mathbb{R}^n$, and hence it follows from Theorem 5.2b and Corollary 5.1 in [Pol.2] (which states that any accumulation point x^*, of the sequence $\{ x_{i_0}^j \}_{j=0}^{\infty}$, constructed by the PPP algorithm, satisfies $\theta_{\mathbf{E}_{i_0}}(x^*) = 0$) that $\theta_{\mathbf{E}_{i_0}}(x_{i_0}^j) \to 0$ as $j \to \infty$, contradicting (2.15).

(ii) We must consider the two discretization schemes separately.

(a) Suppose that the Adaptive Uniform Discretization Scheme 2.1 is used. Then $\psi_{\mathbf{E}_i}(\hat{x}) \to \psi(\hat{x})$ as $i \to \infty$ because the sequence of sets $\mathbf{E}_i$ is dense in $\mathbf{Y}$. Hence (2.14) follows immediately from Lemma 2.1.

(b) Suppose that the Outer Approximation Scheme 2.2 is used. Let K' be an infinite subset of K such that $y_i \xrightarrow{K'} \hat{y} \in \mathbf{Y}$, where y_i is chosen according to the Outer Approximation Scheme 2.2. Since the sequence of sets $\mathbf{C}_i$ is dense in $\mathbf{Y}$, $\psi_{\mathbf{C}_i}(\hat{x}) \to \psi(\hat{x})$ as $i \to \infty$. Hence, it follows from Lemma 2.1 and the fact that $\phi(x_i^{n_i}, y_i) = \psi_{\mathbf{C}_i}(x_i^{n_i})$, that $\phi(x_i^{n_i}, y_i) \xrightarrow{K'} \psi(\hat{x})$. Therefore $\phi(\hat{x}, \hat{y}) = \psi(\hat{x})$. Now, for $i \in K$, let $j(i) = \max\{ i' \in K' \mid i' < i \}$. Since $j(i) < i$ and $y_{j(i)} \in \mathbf{E}_i$, we have that $\psi(x_i^{n_i}) \geq \psi_{\mathbf{E}_i}(x_i^{n_i}) \geq \phi(x_i^{n_i}, y_{j(i)})$. Making use of the continuity of $\psi(\cdot)$ and $\phi(\cdot,\cdot)$, and the fact that $x_i^{n_i} \xrightarrow{K} \hat{x}$ and $y_{j(i)} \xrightarrow{K} \hat{y}$ as $i \to \infty$, we conclude that (i) $\psi(x_i^{n_i}) \xrightarrow{K} \psi(\hat{x})$ as $i \to \infty$ and (ii) $\phi(x_i^{n_i}, y_{j(i)}) \xrightarrow{K} \phi(\hat{x}, \hat{y}) = \psi(\hat{x})$ as $i \to \infty$. Hence (2.14) follows. ■

Lemma 2.3: Suppose that a sequence $\{ x_i^j \}_{i=0}^{\infty}{}_{j=0}^{n_i}$, constructed by Algorithm 2.1, is in a bounded set S and that the Hessian matrices $\nabla \phi_{xx}^j(\cdot,\cdot)$, $j \in \mathbf{l}$, are continuous. Let $M = \max_{x \in S, y \in \mathbf{Y}} \max_{j \in \mathbf{l}} \|\nabla \phi_{xx}^j(x,y)\|$. Then for all i and $j = 0, \ldots, n_i - 1$,

$$\psi_{\mathbf{E}_i}(x_i^{j+1}) - \psi_{\mathbf{E}_i}(x_i^j) \leq \beta\alpha\theta_{\mathbf{E}_i}(x_i^j)/M \,. \tag{2.18}$$

Proof: Expanding some of the functions appearing in the test in (2.4) to second order, (with the indices suppressed to simplify notation) we obtain, for $\lambda \in [0, 1/M]$,

$$\psi_{\mathbf{E}}(x + \lambda h_{\mathbf{E}}(x)) - \psi_{\mathbf{E}}(x) - \alpha\lambda\theta_{\mathbf{E}}(x)$$

$$\leq \max_{y \in \mathbf{E}} \max_{j \in \mathbf{I}} \{ \phi^j(x,y) + \lambda \langle \nabla_x \phi^j(x,y), h \rangle + \tfrac{1}{2}\lambda^2 M \|h\|^2 \} - \psi_{\mathbf{E}}(x) - \alpha\lambda\theta_{\mathbf{E}}(x)$$

$$\leq \lambda(1-\alpha)\theta_{\mathbf{E}}(x) \,. \tag{2.19}$$

The desired result now follows directly from the fact that in view of (2.19) the actual step size used satisfies the inequality $\lambda \geq \beta/M$. ■

Theorem 2.2: Suppose that Assumption 2.1 holds and that Algorithm 2.1, using either the Adaptive Uniform Discretization Scheme 2.1 or the Outer Approximation Scheme 2.2, constructs a bounded sequence $\{ x_i^j \}_{i=0}^{\infty}{}_{j=0}^{n_i}$. Then this sequence must be finite, i.e., there exists an i_0 such that $\psi(x_{i_0}^{n_{i_0}}) \leq 0$.

Proof: Suppose that the sequence $\{ x_i^j \}_{i=0}^{\infty}{}_{j=0}^{n_i}$ is not finite. Since n_i is finite for all i, there are two possibilities: (i) The test (2.5b) was satisfied an infinite number of times, or (ii) the test (2.5c) was satisfied an infinite number of times, and, simultaneously, the stopping criterion (2.6) failed to be satisfied. Now, because the sequence $\{ x_i^j \}_{i=0}^{\infty}{}_{j=0}^{n_i}$ is bounded, it has a converging infinite sequence, indexed by $K \subset \mathbb{N}$, such that $x_i^{n_i} \xrightarrow{K} \hat{x}$, say, as $i \to \infty$, and either the test in (2.5b) is satisfied on this subsequence, i.e.,

$$\psi_{\mathbf{E}_i}(x_i^{n_i-1}) > 0 \quad \text{and} \quad -\theta_{\mathbf{E}_i}(x_i^{n_i-1}) \leq \varepsilon_i \,, \qquad \text{for all } i \in K \,, \tag{2.20a}$$

or the test in (2.5c) is satisfied and the test (2.6) fails on this subsequence, i.e.,

$$\psi_{\mathbf{E}_i}(x_i^{n_i-1}) \leq 0 \quad \text{and} \quad \psi_{\mathbf{C}_i}(x_i^{n_i}) + L/(2(c_i-1)) > 0 \,, \qquad \text{for all } i \in K \,. \tag{2.20b}$$

First, by considering these two cases separately, we will show that

$$\theta_{\mathbf{E}_i}(x_i^{n_i-1}) \xrightarrow{K} 0 \,, \quad x_i^{n_i-1} \xrightarrow{K} \hat{x} \text{ as } i \to \infty \,, \quad \text{and } \psi(\hat{x}) \geq 0. \tag{2.21}$$

(a) Suppose that (2.20a) holds. Since $\varepsilon_i \to 0$ as $i \to \infty$, $\theta_{\mathbf{E}_i}(x_i^{n_i-1}) \xrightarrow{K} 0$ as $i \to \infty$. Thus, we conclude from Lemma 2.2 (ii) that $h_{\mathbf{E}_i}(x_i^{n_i-1}) \xrightarrow{K} 0$ as $i \to \infty$. Hence, $x_i^{n_i} - x_i^{n_i-1} \xrightarrow{K} 0$ as $i \to \infty$. Since $x_i^{n_i} \xrightarrow{K} \hat{x}$ as $i \to \infty$ by assumption, we conclude that $x_i^{n_i-1} \xrightarrow{K} \hat{x}$ as $i \to \infty$. It follows from Lemma 2.1 and Theorem 2.1 (ii) that $\psi_{\mathbf{E}_i}(x_i^{n_i-1}) \xrightarrow{K} \psi(\hat{x})$. Since $\psi_{\mathbf{E}_i}(x_i^{n_i-1}) > 0$ for all $i \in \mathbb{N}$, we must have that $\psi(\hat{x}) \geq 0$.

(b) Suppose that (2.20b) holds. Since the sequence of sets $\{ C_i \}_{i=0}^{\infty}$ is dense in $\mathbf{Y}$, it follows from the reasoning used in proving (2.14) that $\psi_{C_i}(x_i^{n_i}) \xrightarrow{K} \psi(\hat{x})$. Since by (2.20b), for all $i \in K$, $\psi_{C_i}(x_i^{n_i}) + L/(2(c_i-1)) > 0$ and $c_i \to \infty$, we conclude that $\psi(\hat{x}) \geq 0$. Making use of Lemma 2.3 and the fact that by (2.20b), for all $i \in K$, $\psi_{E_i}(x_i^{n_i-1}) \leq 0$, we obtain that

$$\psi_{E_i}(x_i^{n_i}) \leq \psi_{E_i}(x_i^{n_i}) - \psi_{E_i}(x_i^{n_i-1}) \leq \alpha\beta\theta_{E_i}(x_i^{n_i-1})/M \ , \quad \text{for all } i \in K \, . \tag{2.22}$$

It now follows from Theorem 2.1 that

$$\liminf_{i \in K, i \to \infty} \quad \theta_{E_i}(x_i^{n_i-1}) \geq M\psi(\hat{x})/\alpha\beta \, . \tag{2.23}$$

Making use of (2.23) and the fact that (i) $\psi(\hat{x}) \geq 0$ and (ii) $\theta_{E_i}(x) \leq 0$ for all $x \in \mathbb{R}^n$, we conclude that $\theta_{E_i}(x_i^{n_i-1}) \xrightarrow{K} 0$ as $i \to \infty$. Now, from Lemma 2.2 (ii), we conclude that $h_{E_i}(x_i^{n_i-1}) \xrightarrow{K} 0$ as $i \to \infty$. Hence, $x_i^{n_i} - x_i^{n_i-1} \xrightarrow{K} 0$ as $i \to \infty$, and, since $x_i^{n_i} \xrightarrow{K} \hat{x}$ as $i \to \infty$, this leads to the conclusion that $x_i^{n_i-1} \xrightarrow{K} \hat{x}$ as $i \to \infty$, which completes our demonstration that (2.21) must hold.

Now, since $x_i^{n_i-1} \xrightarrow{K} \hat{x}$ as $i \to \infty$, It follows from Lemma 2.1 and Theorem 2.1 (ii) that $\psi_{E_i}(x_i^{n_i-1}) \xrightarrow{K} \psi(\hat{x})$. By Lemma 2.2 (i), we have that for all $i \in \mathbb{N}$,

$$\theta_{E_i}(x_i^{n_i-1}) \leq \theta(x_i^{n_i-1}) + \psi(x_i^{n_i-1}) - \psi_{E_i}(x_i^{n_i-1}) \, . \tag{2.24}$$

Letting $i \to \infty$ in K, and making use of (2.21) and the fact that $\psi(x_i^{n_i-1}) - \psi_{E_i}(x_i^{n_i-1}) \xrightarrow{K} 0$, we obtain that $\theta(\hat{x}) \geq 0$. Hence, it follows from the fact that $\theta(x) \leq 0$ for all $x \in \mathbb{R}^n$, that $\theta(\hat{x}) = 0$, which, together with the fact that $\psi(\hat{x}) \geq 0$ in (2.21), contradicts Assumption 2.1(ii). ■

3. NUMERICAL RESULTS

Algorithm 2.1 was coded in C and was executed on a SUN 3/140 Workstation. In the experiments below, the algorithm parameters were set as follows: $\alpha = 0.9$, $\beta = 0.9$, $\gamma = 1.0$, $\varepsilon_i = 0.1^{i+1}$. The sets C_i were chosen to consist of $2^{i+4} + 1$ uniformly spaced points in $[0,1]$. In the Adaptive Uniform Discretization Scheme 2.1, the sets E_i, $i = 0,1,...,$ were chosen to consist of $2^{i+2} + 1$ uniformly spaced points in $[0,1]$. The Outer Approximations Scheme 2.2 was initialized by setting E_0 to consist of 5 uniformly spaced points in $[0,1]$. In our experiments, the Lipschitz constant L was calculated by averaging the estimates of local Lipschitz constants, L_i^j, around the points x_i^j, with respect to the set C_i, where L_i^j is obtained as follows

$$L_i^j = \max_{p \in I} \max_{k = 0,1,\ldots,c_i-1} |\phi^p(x_i^j, k/(c_i-1)) - \phi^p(x_i^j, (k+1)/(c_i-1))|(c_i - 1) . \tag{3.1}$$

Because there are few semi-infinite satisfycing test problems in the literature, we have constructed four semi-infinite minimax problems by converting four constrained problems in [Tan.1] into semi-infinite satisfycing problems using L_∞ exact penalty functions. Two versions of each problem were tested, both with the Adaptive Uniform Discretization Scheme 2.1 and with the Outer Approximation Scheme 2.2. The first version used a large penalty coefficient (PC = 100.0), while in the second version used a smaller penalty coefficient (PC = 10.0). Table 3.1 summarizes the performance of the two discretization schemes on these problems. We evaluate the performance of the two discretization schemes by comparing the number of function evaluations (denoted as NF), the number of gradient evaluations (denoted as NG), and the total number of evaluations (NT = NF + n*NG). It is clear from our experimental results that with the exception of one case, the Outer Approximation Scheme 2.2 outperformed the Adaptive Uniform Discretization Scheme 2.1. The test problems were as follows:

TFI1 Problem[Tan.1]:

$$\psi(x) = (x_1)^2 + (x_2)^2 + (x_3)^2 + PC \max\{ \max_{y \in [0,1]} (x_1 + x_2\exp(x_3 y) + \exp(2y) - 2\sin(4y)) \, , 0 \} - 5.5 .$$

Initial point $x_0 = (1.0, 1.0, 1.0)$.

TFI2.1 Problem[Tan.1]:

$$\psi(x) = x_1 + x_2/2 + x_3/3 + PC \max\{ \max_{y \in [0,1]} (\tan(y) - x_1 - x_2 y - x_3 y^2), 0 \} - 0.66 .$$

Initial point $x_0 = (0.0, 0.0, 0.0)$.

TFI2.2 Problem[Tan.1]:

$$\psi(x) = x_1 + x_2/2 + x_3/3 + x_4/4 + x_5/5 + x_6/6 +$$

$$PC \max\{ \max_{y \in [0,1]} (\tan(y) - x_1 - x_2 y - x_3 y^2 - x_4 y^3 - x_5 y^4 - x_6 y^5), 0 \} - 0.63 .$$

Initial point $x_0 = (0.0, 0.0, 0.0, 0.0, 0.0, 0.0)$.

TFI3 Problem[Tan.1]:

$$\psi(x) = \exp(x_1) + \exp(x_2) + \exp(x_3) + PC \max\{ \max_{y \in [0,1]} (1/(1 + y^2) - x_1 - x_2 y - x_3 y^2) , 0 \} - 4.45 .$$

Initial point $x_0 = (1.0, 0.5, 0.0)$.

Problems	Adaptive Uniform Discretization			Outer Approximation		
	NF	NG	NT	NF	NG	NT
TFI1.a	56808	2476	64236	11332	468	12736
TFI1.b	10338	648	12282	4494	168	4998
TFI2.1.a	100884	4394	114066	20546	278	21380
TFI2.1.b	6014	344	7046	4338	158	4812
TFI2.2.a	9454	750	13954	11669	346	13745
TFI2.2.b	1890	248	3378	1600	180	2600
TFI3.a	2160	168	2664	1376	71	1589
TFI3.b	286	22	352	232	19	289

Table 3.1 Summary of Numerical Results

4. CONCLUSION

We have presented an efficient method for solving semi-infinite inequalities in a finite number of operations, which can incorporate either outer approximations type discretization or adaptive uniform discretization. It is most interesting to observe that even without the inclusion of constraint dropping schemes, such as those described in [Gon.1], the version of our method based on outer approximations is considerably faster than the version based on an adaptive uniform discretization scheme.

5. REFERENCES

[Gon.1] C. Gonzaga and E. Polak, "On Constraint Dropping Schemes and Optimality Functions for a Class of Outer Approximations Algorithms", *SIAM J. Control Optim.* Vol. 17, pp. 477-493, 1979.

[He.1] L. He and E. Polak, "Optimal Diagonalization Strategies for the Solution of a Class of Optimal Design Problems", *Electronics Research Laboratory, University of California, Berkeley,* Memorandum No. UCB/ERL M88/41, 1988, *IEEE Trans. Automatic Control,* in press.

[May.1] D. Q. Mayne and E. Polak, "Outer Approximation Algorithm for Nondifferentiable Optimization Problems" *J. Optimization Theory and Applications,* Vol. 42, No. 1, pp. 19-30, 1984.

[Pir.1] O. Pironneau and E. Polak, "On the Rate of Convergence of Certain Methods of Centers", *Mathematical Programming,* Vol. 2, No. 2, pp. 230-258, 1972.

[Pol.1] E. Polak, *Computational Methods in Optimization: A Unified Approach,* Academic Press, New York, 1971.

[Pol.2] E. Polak, "On the Mathematical Foundations of Nondifferentiable Optimization in Engineering Design", *SIAM Review,* Vol. 29, No. 1, pp. 21-91, March 1987.

[Psh.1] B. N. Pshenichnyi and Yu. M. Danilin, *Numerical methods in extremal problems (Chislennye metody v ekstremal'nykh zadachakh),* Nauka, Moscow, 1975.

[Tan.1] Y. Tanaka, M. Fukushima and T. Ibaraki, "A comparative study of several semi-infinite nonlinear programming algorithms", *European Journal of Operational Research,* Vol. 36, pp. 92-100, 1988.

[Wie.1] Wierzbicki, A. P., "A Mathematical Basis for Satisfycing Decision-Making", *Mathematical Modeling,* Vol. 3, No. 5, pp. 391-405, 1982.