EXPERIMENT ON USEFULNESS OF FACETS OF THE 0–1 KNAPSACK PROBLEM AS CUTTING PLANES

YASUKI SEKIGUCHI
Associate Professor
Faculty of Economics
Hokkaido University

Abstract

The necessary and sufficient condition for a sequentially lifted facet to be a cutting plane is stated, and a simplified sufficient condition is induced and analyzed. Then, a procedure to construct a minimal cover and to determine a sequentially lifted facet which is expected to be a cutting plane is invented. The results can be combined with various algorithms to solve 0–1 programming problems. In order to assess the effect of the newly derived cutting planes, the procedure is combined to the Gomory's fractional cutting plane method and a new algorithm is proposed. Computational experiments on this algorithm clarify the following points:

The proposed algorithm solves the 0–1 knapsack problem better than the fractional method. The dominancy of the new algorithm is thought to be the result of prohibiting the increase of computational error by using facets with smaller coefficients than those of fractional cuts. It is difficult to see that primitive facets generated in new algorithm cut deeper than fractional cuts, but they complement each other. The experiments suggest strongly that the class of sequentially lifted facets generated by the proposed procedure does not include important classes of facets for cases with many variables.

1. Introduction

The continuous knapsack problem with a single constraint is an analytically solved problem. But, its integral version is one of the most difficult problems to solve. In fact, this is NP-complete (see for example [8]). On the other hand, this is simplest integer programming problem and has various applications (see for example [17]). This problem is interesting from both theoretical and practical viewpoints.

One of the current topics of theoretical interest is concerned with facets of the polytope which is the convex hull of the set of feasible solutions. Facets have been the main theoretical concern in recent years in the field of mathematical programming, and a fair amount of results have been ob-

tained (see for example [15]). The 0–1 knapsack problem is one of the well discussed problems. If all the facets of the convex hull of the set of feasible solutions could have been generated, an original integer programming problem would have been transformed to an equivalent continuous linear programming problem. Actually, generating all facets is also extremely difficult, and there was no confidence in the practical value of facet researches [4].

The author has been interested in the fact that an appropriately selected facet can be the deepest cutting plane, and has discussed a new algorithm, where facets are used for conventional cutting planes. In [14], an introductory discussion was given on an algorithm which utilizes the simplest facets, i. e. the canonical facets, of the knapsack polytopes. In the present paper, its improved version will be proposed and results of computational experiments will be shown. The improvements are, that facets other than the canonical ones are also utilized, that theoretical analysis on the use of facets are further developed, and that the algorithm is completed in detail.

During this research, reports [3] and [11] were published. These reports proposed a new algorithm of the symmetric travelling salesman problem, where facets are used as cutting planes. It was shown that facets work as deeper cutting planes than the conventional ones, and that the problems with 300 and some cities can be solved by combined usage of facet and an enumerative algorithm. But, they did not clarify the reason why facets work well. At present, we do not know a method for selectively generating facets which characterize the optimal integer solution, and need to acquire more experience on the use of facets.

Our new algorithm will be compared to the Gomory's fractional method and the dominancy of using facets will then be analyzed. It will be clarified through this analysis whether our knowledge on facets of the 0–1 knapsack problem is enough for constructing an effective algorithm.

In the next section, the construction of facets of the knapsack polytope is explained as long as it is concerned in the succeeding sections. In 3, conditions required for facets to be cutting planes, existency conditions of such facets and a method for constructing such facets are discussed. In 4, the new algorithm is proposed and examples are solved by the algorithm. Finiteness of this algorithm is easily shown. In the succeeding section, a FORTRAN program of the proposed algorithm is briefly introduced and the results of computational experiments are analyzed.

It is noteworthy that our objective here is not to invent a new algorithm better than the excellent existing ones. It is reported that problems of practical size can be solved by the existing ones. See [9] and [16] for the recent surveys of such algorithms. Our object here is to examine the possibility and effects of using facets in solution algorithms.

2. The Problem and Facets from Minimal Covers

The 0-1 knapsack problem is defined as follows.

$$\text{(KP)} \qquad \begin{aligned} &\text{Maximize} \quad \textstyle\sum_{i\in N} c_i x_i \\ &\textstyle\sum_{i\in N} a_i x_i \leqq a_0 \\ &x_i \in \{0, 1\} \qquad \forall i \in N = \{1, 2, \cdots, n\}\,. \end{aligned}$$

Here, $\quad a_{i-1} \geqq a_i > 0\,, \qquad c_i > 0 \qquad \forall i \in N\,.$

Besides, Assume the next to eliminate a trivial case.

$$\text{(2. 1)} \qquad \textstyle\sum_{i\in N} a_i > a_0\,.$$

We will denote as RKP the problem obtained by relaxing the integrality condition of KP. Denote also as F the set of feasible solutions of RKP, and as F_c the convex hull of the integral points in F. F_c is called a *knapsack polytope.* Let x be a column vector with n components x_i $(i=1, 2, \cdots, n)$ and R^n be the n-dimensional real space. Then, mathematically, we can write

$$\text{(2. 2)} \qquad \begin{aligned} &F = \left\{x \in R^n \,|\, \textstyle\sum_{i\in N} a_i x_i \leqq a_0,\ 0 \leqq x_i \leqq 1 \ \forall i \in N\right\}, \\ &F_c = \text{conv.}\left\{x \in F \,|\, x_i \in \{0, 1\} \ \forall i \in N\right\}. \end{aligned}$$

A subset S of N is called a *cover* if it satisfies

$$\text{(2. 3)} \qquad \textstyle\sum_{i\in S} a_i > a_0\,.$$

If S satisfies (2. 4) as well as (2. 3), it is called a *minimal cover.*

$$\text{(2. 4)} \qquad \textstyle\sum_{i\in S} a_i - a_j \leqq a_0 \qquad \forall j \in S\,.$$

Define a set S' for a minimal cover S by

$$\text{(2. 5)} \qquad S' = \left\{i \in N \backslash S \,|\, a_i \geqq \max\,(a_j \,|\, j \in S)\right\}.$$

Then, a set $E(S) = S \cup S'$ called an *extension of S.*
An inequality

$$\text{(2. 6)} \qquad \textstyle\sum_{i\in N} r_i x_i \leqq r_0$$

is said to be a *facet of* F_c, if and only if (2. 7) is satisfied.

(2. 7a) Every point in F_c satisfies (2. 6).

(2. 7b) There are n affinely independent points in F_c which satisfy (2. 6) with equality.

The set of vertices of F_c is the same as the set of feasible solutions of KP, and (2. 7) requires the hyperplane defined by (2. 6) to fit exactly on a face

of the polytope F_c. An inequality (2. 6) which satisfies (2. 7a) is said to be *valid.*

Suppose facets at the following form.

$$\sum_{i\in S} x_i + \sum_{i\in N\setminus S} b_i x_i \leqq |S| - 1 . \tag{2. 8}$$

If $N\setminus S=\phi$ or $b_i=0$ for all $i\in N\setminus S$, (2. 8) is called a minimal cover equation. Among this type of facets are the sequentially lifted* facets of Balas and Zemel [1]. Readers should refer to the other reports such as [2], [5], [10], [12], [13], [15] and [18] for details about facets of this form. In the present paper, results necessary for the succeeding discussion will be introduced.

Define h_i for each i in $N\setminus S$ by the following equation.

$$\sum_{j\in S_{h_i}} a_j \leqq a_i < \sum_{j\in S_{h_i+1}} a_j . \tag{2. 9}$$

Here, S_k denotes a subset of S with k elements taken in the increasing order of the values. By virtue of the assumption on the values of a_j's, h_i is the maximum number of elements taken from S in the decreasing order of a_j whose sum is less than or equal to a_i. It is evident from the definition that h_i depends on the minimal cover S.

Using h_i's, the set $N\setminus S$ can be partitioned into subsets I and J.

$$\begin{aligned} I &= \left\{ i\in N\setminus S \mid \sum_{j\in S\setminus S_{h_i+1}} a_j \leqq a_0 - a_i < \sum_{j\in S\setminus S_{h_i}} a_j \right\} . \\ J &= \left\{ i\in N\setminus S \mid \sum_{j\in S\setminus S_{h_i+2}} a_j < a_0 - a_i < \sum_{j\in S\setminus S_{h_i+1}} a_j \right\} . \end{aligned} \tag{2. 10}$$

Then, it can be shown (see [1], also [10] for the proof),

Theorem 2. 1. If (2. 8) is a facet of F_c, it must be satisfied that

$$\begin{aligned} b_i &= h_i && \forall i\in I , \\ h_i \leqq b_i &\leqq h_i+1 && \forall i\in J . \end{aligned} \tag{2. 11}$$

Theorem 2. 1 implies that the coefficients of facets of the type (2. 8) are uniquely determined, except those for $i\in J$. Moreover, the values of the coefficients for $i\in J$ are restricted within segments of length 1. If $b_i (i\in J)$ is either h_i or h_i+1, all coefficients of (2. 8) are integral. Every facet with integral coefficients of the form (2. 8) is known [1] to be generated by the sequential lifting below.

Determine an arbitrary order to the elements in $N\setminus S$, and denote this ordered set as $\bar{S}=\{j_1, j_2, \cdots, j_t\}$, $t=n-|S|$. Let $\bar{S}_u$ be a subset $\{j_1, j_2, \cdots, j_u\}$. The $KP(j_u)$ below is itself the 0–1 knapsack problem.

$$\begin{aligned} (KP(j_u)) \qquad & \text{Maximize} \quad \sum_{j\in S} x_j + \sum_{j\in \bar{S}_{u-1}} b_j x_j, \quad \sum_{j\in S\cup \bar{S}_{u-1}} a_j x_j \leqq a_0 - a_{j_u}, \\ & x_j \in \{0, 1\} \ \ \forall j\in S\cup \bar{S}_{u-1}, \ \ b_j = |S| - 1 - Z(j) \ \ \forall j\in \bar{S}_{u-1} . \end{aligned}$$

* In [12] on a general theory of facets, the lifting operation of Balas and Zemel is called an extension. But, the original term is adopted here.

$Z(j_u)$ denotes the optimal value of the objective function. b_{j_u} $(u=1, 2, \cdots, t)$ can be sequentially determined by solving $KP(j_u)$ for $u=1, 2, \cdots$ in turn. It is also known [1] that b_i's thus determined satisfy (2. 11) and determine a facet. This operation to find the values of b_i's is called a *sequential lifting*. About the optimal value of $KP(j_u)$, the following fact is known [1]. $\bar{Z}(j_u)$ is the optimal value of $RKP(j_u)$, which is a problem obtained from $KP(j_u)$ by relaxing the integrality condition.

Theorem 2. 2. Let $j_u \in J$, then $b_{j_u}=h_{j_u}+1$ if one of the following is true.

(2. 12a) $\quad j_u$ is the first in $\bar{S}$ which is included in J.

(2. 12b) $\quad \bar{Z}(j_u)<|S|-h_{j_u}-1$.

It is understood that through the theorems, facets (2. 8) in the following cases can be determined without solving combinatorial problems.

a. (direct partible) $J=\phi$.

b. (lift with $|J|=1$) $\quad |J|=1$.

c. (lift with $|J|>1$) $\quad |J|>1$ and every $j \in J$ satisfies (2. 12b).

It may happen in these cases that

(2. 13) $$b_i \in \{0, 1\} \qquad \forall i \in N \backslash S.$$

A facet (2. 8) is said *canonical* if (2. 13) is satisfied. Historically, canonical facets have been found earlier than the theorems above [2]. It is also known (see [5], [15] and [17] for details) that there is a subset S of $\{i \in N \mid b_i=1\}$ in which (2. 8) is exactly the same as

(2. 14) $$\sum_{j \in E(S)} x_j \leqq |S|-1,$$

for every canonical facet.

3. Facets and Cutting Planes

In this section, conditions under which sequentially lifted facets become cutting planes, or existency conditions under which minimal covers give facets as cutting planes are discussed.

Denote as $[RKP, k]$ a problem which is obtained by adding k valid inequalities to RKP. Let x^* be the optimal solution to $[RKP, k]$, and put

(3. 1) $$B_1=\{i \in N \mid x_i^*=1\},$$
$$B_f=\{i \in N \mid 0<x_i^*<1\}.$$

A facet is called a *cutting plane* (relative to x^*) if it cuts off the point x^*.

Theorem 3. 1. The inequality (2. 8) cuts off the point x^*, if and only if (3. 2) is satisfied.

(3. 2) $$\sum_{j\in S\cap B_f} x_j^* + \sum_{j\in B_1\setminus S} b_j + \sum_{j\in B_f\setminus S} b_j x_j^* > |S\setminus B_1| - 1 .$$

Proof: $$\begin{aligned}\sum_{j\in S} x_j^* + \sum_{j\in N\setminus S} b_j x_j^* &= |S\cap B_1| + \sum_{j\in S\cap B_f} x_j^* + \sum_{j\in (B_1\cup B_f)\setminus S} b_j x_j^* \\ &= \sum_{j\in S\cap B_f} x_j^* + \sum_{j\in B_1\setminus S} b_j + \sum_{j\in B_f\setminus S} b_j x_j^* + |S\cap B_1| \\ &> |S| - 1 .\end{aligned}$$

Theorem 3. 1. suggests minimal covers which are convenient for obtaining facets of cutting planes. In Fig. 3. 1, S_D satisfies

$$|S_D| + |S_f| - 1 = |S\setminus B_1| - 1 .$$

Therefore, its contribution to the right hand side is $|S_D|$, but that to the left hand side is nothing. This implies that $S_D = \phi$ is desirable for a minimal cover in order to give a cutting plane. Effects of $B_1\cap S$ on both sides are equal and $B_1\cap S$ has no actual effect. $B_f\cap S$ contributes to the left hand side by at most less than 1, and to the right hand side by exactly 1. Thus this portion of S may be desirably as small as possible. Contributions by b_j's for j in $(B_1\cup B_f)\setminus S$ are made only to the left hand side, and it is desirable that they be as large as possible. Specifically, the effect of b_j's for j in $B_1\setminus S$ will be large if they have large values. Therefore, lift a minimal cover equation first with respect to the elements in $B_1\setminus S$ and with respect to those in $B_f\setminus S$. This sequential lifting procedure will be effective for obtaining a facet which becomes a cutting plane.

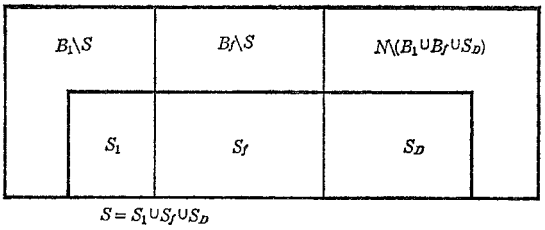

Fig. 3. 1. Relation of a Minimal Cover S to x^*.

From theorem 3. 1 the following can be obtained. This is convenient for understanding the desirable property of minimal covers which give facets.

Corollary 3. 2. Let $S\subset B_1\cup B_f$ be a minimal cover. Then, the facet (2. 8) obtained from S determines a cutting plane if (3. 3) is true.

(3. 3) $$\sum_{j\in S\cap B_f} x_j^* > |S\cap B_f| - 1 .$$

The proof of this corollary is immediate from theorem 3. 1.

The next problem is to see if it is possible to obtain a minimal cover with desirable properties given above from the optimal solution x^* to [*RKP*, *k*].

Lemma 3.3. For $[RKP, k]$ $(k=0, 1, 2, \cdots)$, $B_f \neq \phi$ is equivalent to (3.4).

(3.4) $\sum_{j \in B_1 \cup B_f} a_j > a_0$.

Proof: If (3.4) holds true, $B_f \neq \phi$ must also be true. Notice that x^* is feasible and $B_f = \phi$ contradicts this fact.

Assume $B_f \neq \phi$. x^* is feasible and (3.5) must be true.

(3.5) $\sum_{j \in B_1 \cup B_f} a_j x_j^* \leqq a_0$.

When this inequality holds with equality, (3.4) is evidently true. Assume that this inequality holds with an exact inequality and that (3.4) is not true. Then, $\bar{x}$ below is feasible with respect to $[RKP, 0]$.

$$\bar{x}_i = 1 \qquad \forall i \in B_1 \cup B_f.$$

$$\bar{x}_i = 0 \qquad \forall i \in N \backslash (B_1 \cup B_f).$$

Evidently, $\bar{x}$ is better than x^* (i. e. the objective function value of $\bar{x}$ is greater than that of x^*). Therefore, by definition, $\bar{x}$ must be infeasible with respect to $[RKP, k]$, and at least one, say V_q, of the additional inequalities, V_1, V_2, $\cdots$, V_k, is violated. This contradicts the validity of V_q.

From the lemma, $B_1 \cup B_f$ is a cover and it can be made minimal by dropping some elements. Every minimal cover made by doing this includes at least one element in B_f, because B_1 is not a cover. Thus, the next theorem follows immediately.

Theorem 3.4. For $[RKP, k]$ $(k=0, 1, 2, \cdots)$, if $B_f \neq \phi$ there exists a minimal cover S such that

$$S \cap B_f \neq \phi \quad \text{and} \quad S \subset B_1 \cup B_f.$$

Usually, there are plural minimal covers stated in the theorem, and specifically, the number will be large if $B_1 \cup B_f$ is large. Various procedures for constructing them can be devised. Two procedures stated below may be effective for constructing desirable minimal covers in conjunction with equation (3.3) and theorem 3.1.

Notice that minimal covers stated in theorem 3.4 can be constructed by the following procedure.

Step 1: Make a cover S by adding to B_1 an appropriate subset of B_f.

Step 2: Remove from S some elements in $S \cap B_f$ as long as the condition of covers (2.3) is not violated.

Step 3: Remove from S some elements in $S \cap B_1$ as long as the condition (2.3) is not violated.

If the sequence of addition in step 1 or those of removal in steps 2 and 3 are varied, resultant minimal covers will also vary. As stated previously, an important point which is induced from (3.3) or theorem 3.1 is which of

the elements in B_f are left in the minimal cover.

In order to obtain a minimal cover with a large difference between both sides of (3. 3), adding a subset of B_f corresponding to the optimal solution for the following 0-1 knapsack problem is quite sufficient.

$$\begin{aligned} \text{Maximize} \quad & \textstyle\sum_{j\in B_f}(x_j^*-1)\,y_j\,, \\ & \textstyle\sum_{j\in B_f}a_j y_j > a_0 - \sum_{j\in B_1} a_j\,, \\ & y_i \in \{0,1\} \qquad \forall j \in B_f\,. \end{aligned}$$

Because x_j^*-1 is negative, if the optimal solution to the above problem could be found, step 2 in the above procedure would be needless. But, when B_f is a large set it will be difficult to solve this problem. Therefore, the following approximate solution may be reasonable as steps 1 and 2.

Step 1: Arrange the elements in B_f in the nonincreasing order of

$$(x_j^*-1)/a_j\,.$$

Starting with $S=B_1$, add j in B_f in this order to S until S becomes a cover.

Step 2: Arrange the elements in $S_\cap B_f$ in the nondecreasing order of x_j^*. In this order, check each j in $S_\cap B_f$ if S remains a cover even when j is omitted from S, and if the answer is affirmative, omit j from S.

The three step procedure above is called algorithm AS if its step 1 and 2 are realized as step 1 and 2 just described.

Depth of a cut may be estimated through quantity of decrease of the optimal value of $[RKP, k]$. Assume that x_j^* $(j\in B_f)$ changes and becomes an integer. Then, the change in the objective function value is at least $c_j\cdot\min\{x_j^*,\ 1-x_j^*\}$ if changes in all the other variables are neglected. A possible way to find a good minimal cover is the following.

Step 1: Arrange the elements in B_f in the nonincreasing order of

$$c_j\cdot\min\{x_j^*,\ 1-x_j^*\}/a_j\,.$$

Starting with $S=B_1$, add j in B_f in this order to S until S becomes a cover.

Step 2: Do the same thing as step 2 of AS.

The three step procedure with step 1 and 2 above is called algorithm BS.

It is difficult to determine in advance whether it is possible to find a minimal cover which gives a facet satisfying theorem 3. 1, by AS or BS, or by some other adequate way. This is caused by difficulty in forecasting exact values of b_i's in (2. 8). But for some simple cases, it is possible to determine whether there exists a minimal cover which satisfies (3. 3).

Theorem 3.5. For $[RKP, k]$ $(k=0, 1, 2, \cdots)$, if $|B_f|=1$, then every minimal cover $S \subset B_1 \cup B_f$ satisfies (3.3).

Proof: By virtue of theorem 3.4 $|S \cap B_f|=1$, and $x_j^*(j \in B_f)$ are positive. Evidently, (3.3) must be satisfied.

Theorem 3.6. For $[RKP, k]$ $(k=0, 1, 2, \cdots)$, if $|B_f|=2$, then there exists a minimal cover $S \subset B_1 \cup B_f$ satisfying (3.3).

Proof: Let $B_f=\{j_1, j_2\}$, and $a_{j_1} \geqq a_{j_2}$. Also let $a_0'=a_0-\sum_{j \in B_1} a_j$. By lemma 3.3, $a_{j_1}+a_{j_2}>a_0'$ must be true. If $a_{j_1}>a_0'$, $B_1 \cup \{j_1\}$ is a cover and it is possible to construct a minimal cover so that $|S \cap B_f|=1$. In this case, such a cover will be shown to satisfy (3.3) by doing the same as in the proof of theorem 3.5. Therefore, assume $a_0' \geqq a_{j_1} \geqq a_{j_2}$. If $x_{j_1}^*+x_{j_2}^* \leqq 1$, the case can be reduced to the case of theorem 3.5 as follows. Let j_v ($v=1$ or 2) be the index of x_j^* whose coefficient in the objective function is greater than or equal to that of another. Make a solution by setting $x_{j_v}^*=1$ and the value of the other equal to zero in x^*. This solution is feasible and its objective function value is not less than that of x^*. It is another optimal solution for which theorem 3.5 is applicable. Assume $x_{j_1}^*+x_{j_2}^*>1$. Then $B_1 \cup \{j_v\}$ ($v=1$ or 2) can not be a cover, and a minimal cover $S \subset B_1 \cup B_f$ with $B_f \subset S$ must exist. For this S, the left hand side of (3.3) is greater than 1 and the right hand side is exactly equal to 1.

Unfortunately, an example can be constructed where there is no minimal cover satisfying (3.3) for $|B_f|=3$. A remaining problem is whether there is a minimal cover satisfying theorem 3.1 for this case.

4. A New Solution Algorithm and its Convergence Property

As discussed in the previous section, it is not garanteed that there exists a sequentially lifted facet which cuts off the nonintegral optimal solution to $[RKP, k]$. However, the results in the preceding section can be combined with various algorithms to solve 0-1 programming problems. For example, assume that a nonintegral optimal solution is obtained to a relaxed 0-1 programming problem and there is an active inequality of suitable form. Then, regarding the active inequality as a knapsack constraint, sequentially lifted facets in the preceding section can be determined and used as cutting planes related to the nonintegral solution. This will make it possible to improve the original solution algorithm. In order to assess effects of this new kind of cutting plane, a new algorithm is constructed and some computational experiments are performed below.

All coefficients in a sequentially lifted facet are always integer, and Gomory's fractional method can be applied to $[RKP, k]$ if its k additional constraints are sequentially lifted facets or Gomory's fractional cuts. Thus,

when a facet being a cutting plane is not found, it is always possible to add a cutting plane of the fractional method instead of a facet. We assumed in the preceding section only that additional inequalities $V_1, V_2, \cdots, V_k$ are valid, and the results can be applied even when all the additional inequalities are Gomory's cuts.

Essentially important is the fact that Gomory's fractional cuts are compatible with the facets which are discussed in the preceding section. In any stage of Gomory's fractional method, when a facet being a cutting plane is found, it can be utilized as an additional constraint instead of a Gomory's cut.

One of the possible algorithms where sequentially lifted facets are used as main additional constraints is the following cutting plane method.

[A New Solution Algorithm]

1. $k \leftarrow 0$.
2. Solve $[RKP, k]$. Terminate if the optimal solution is integral, else go to 3.
3. Find sequentially lifted facets.
4. Go to 6 if at least one of the facets is a cutting plane, else go to 5.
5. Find a Gomory's fractional cut.
6. $k \leftarrow k+$(the number of new additional constraints), and go to 2.

In 5 of the above algorithm, the row closest to 0.5 is selected as a source row among nonintegral rows. Nonintegrality of a row can be considered as maximum if its fractional part is just 0.5. On the other hand, a fractional cut is derived from a necessary condition for a basic nonintegral variable to have an integral value. So, it is natural to expect that a deep cut is realized by selecting a source row as described.

A procedure to find sequentially lifted facets can be partitioned into two stages. At the first stage, a minimal cover is sought, and at the second stage a sequential lifting is performed.

At the first stage, a minimal cover S_A is found by algorithm AS. If S_A satisfies (3. 3), the second stage is performed for S_A, after that, returning to the first stage, another minimal cover S_B is sought by algorithm BS. If S_B is determined, proceed again to the second stage.

The important thing in the second stage is the sequence $\bar{S}$ in which a sequential lifting is performed. Referring to the discussion just after theorem 3. 1, this sequence is determined as follows.

$$\bar{S} = (\underbrace{j_1, \cdots, j_a}_{B_1 \backslash S}, \underbrace{j_{a+1}, \cdots, j_b}_{B_f \backslash S}, \underbrace{j_{b+1}, \cdots, j_t}_{N \backslash (B_1 \cup B_f)})$$

where, $j_1, \cdots, j_a$; a sequence arranged in the nonincreasing order of c_j/a_j,

$j_{a+1}, \cdots, j_b$; a sequence where elements removed at step 2 of $AS(BS)$

are arranged in the nonincreasing order of x_j, followed by a sequence of elements which were not adopted at step 1 of $AS(BS)$ arranged in the order of that defined at the step,

$j_{b+1}, \cdots, j_t$; a sequence in the decreasing order of j's.

According to $\bar{S}$, for each $j_u (1 \leqq u \leqq t)$ the coefficient b_{j_u} in (2. 8) is determined by the following procedure.

[Sequential Lifting]

1. Determine h_{j_u}.
2. If $j_u \in I$, let $b_{j_u} = h_{j_u}$.
3. If $j_u \in J$ and theorem 2. 2 is applicable, let $b_{j_u} = h_{j_u} + 1$.
4. If $j_u \in J$ and theorem 2. 2 is not applicable, let $b_{j_u} = h_{j_u}$ after all j's in $\bar{S}$ have been processed.

Let V_A be the inequality obtained in the second stage corresponding to S_A, and V_B be that to S_B. These equations are facets as long as cases of 4 do not happen. If some j's in J are of 4, they are not necessarily facets but are valid. They are also used in the experiment below and called non-facet cuts if they pass the test in step 4 of the new solution algorithm. Because S_B is sought only when S_A satisfies (3. 3), when V_B is generated, V_A is a cutting plane by virtue of corollary 3. 2. This is true even when there are some j's in J for which theorem 2. 2 is not applicable. In the new algorithm, only one or two constraints are augmented for every cycle.

Facets are the deepest cuts in the sense that there exists no valid inequality which is parallel to and deeper than it. But, it is noteworthy that facets are not the most effective cutting planes in the sense that the algorithm converges to an integral solution in the minimum number of augmented constraints.

The new algorithm, if no non-facet cut is used, is finite because there are at most a finite number of facets and because the Gomory's fractional method is proved to be finite.

Assume that a Gomory's cut is simultaneously added when only non-facet cuts are obtained for $[RKP, k]$ (this was not done in the experiment below). Then, the new algorithm is easily proved to be finite even when non-facet cuts are used. Therefore, it is understood that the new algorithm can easily be constructed to become finite if finiteness is a serious point.

[Example 1] (Taha [17])

$$\begin{aligned} \text{Maximize} \quad & 13x_1 + 14x_2 + x_3 + 5x_4 + 10x_5 + 2x_6 + 6x_7 + 2x_8 + 2x_9 \\ & 14x_1 + 12x_2 + 11x_3 + 10x_4 + 8x_5 + 3x_6 + 2x_7 + 2x_8 + x_9 \leqq 17 \\ & x_i \in \{0, 1\} \qquad \forall i \in \{1, 2, \cdots, 9\} . \end{aligned}$$

For this problem, the optimal solution to $[RKP, 0]$ is

$$x^* = (0, 1/2, 0, 0, 1, 0, 1, 0, 1) .$$

According to step 1 of AS,

$$S_A = \{2, 5, 7, 9\} .$$

By virtue of theorem 3. 5, step 2 of AS is skipped. In step 3 of AS, if j in B_1 is tested if it can be removed from S_A, in the nondecreasing order of c_j/a_j, the result is

$$S_A = \{2, 5\} .$$

Then, according to the definition,

$$\bar{S}_A = (7, 9, 8, 6, 4, 3, 1) .$$

By the sequential lifting procedure, b_j's are determined as follows.

$$h_7 = 0,\ 7 \in I,\ b_7 = 0 ;\ h_9 = 0\ \ 9 \in I,\ b_9 = 0 ;\ h_8 = 0,\ 8 \in I,\ b_8 = 0$$
$$h_6 = 0,\ 6 \in I,\ b_6 = 0 ;\ h_4 = 0,\ 4 \in J,\ b_4 = 1 ;$$
$$h_3 = 0,\ 3 \in J,\ \bar{Z}(3) = 3/4 < 1,\ b_3 = 1 ;\ h_1 = 1,\ 1 \in I,\ b_1 = 1 .$$

Therefore,

$$V_A :\quad x_1 + x_2 + x_3 + x_4 + x_5 \leqq 1 ,$$

is a facet corresponding to case c below theorem 2. 2. S_A satisfies (3. 3) and cuts off x^*. $|B_f| = 1$ and S_B is the same as S_A. Thus, V_B is identical with V_A. After addition of V_A to [RKP, 0], when the resultant problem [RPK, 1] is solved, its optimal solution is

$$x^* = (0, 1, 0, 0, 0, 0, 1, 1, 1) .$$

[Example 2] (Gomory's fractional cuts must be used)

$$\begin{aligned} \text{Maximize}\quad & 97x_1 + 11x_2 + 97x_3 + 44x_4 + 89x_5 + 91x_6 \\ & \quad + 62x_7 + 36x_8 + 89x_9 + 83x_{10} \\ & 94x_1 + 64x_2 + 61x_3 + 43x_4 + 40x_5 + 39x_6 \\ & \quad + 38x_7 + 31x_8 + 29x_9 + 11x_{10} \leqq 94 \\ & x_i \in \{0, 1\} \qquad \forall i \in \{1, 2, \cdots, 10\} . \end{aligned}$$

The optimal solution and additional constraints in each iteration are ;

$$x^* = (0, 0, 0, 0, 0.375, 1, 0, 0, 1, 1) ;\ S_A = \{5, 6, 9\} ,$$
$$V_A :\quad 2x_1 + x_2 + x_3 + x_5 + x_6 + x_7 + x_8 + x_9 \leqq 2 ;\ V_B = V_A ,$$
$$x^* = (0, 0, 0.6818, 0, 0, 0.3182, 0, 0, 1, 1) ;\ S_A = \{3, 9, 10\} ,$$
$$V_A :\quad 2x_1 + x_2 + x_3 + x_9 + x_{10} \leqq 2\ (J = \phi) ;\ V_B = V_A ,$$
$$x^* = (0, 0, 0.4688, 0, 0, 1, 0, 0, 0.5312, 1) ;\ S_A = \{3, 6\} ,$$

$$V_A: \quad x_1+x_2+x_3+x_6 \leqq 1 \ (J=\phi)\,; \ V_B = V_A\,,$$

$$x^*=(0, 0, 0.4545, 0, 0.4545, 0.5455, 0, 0.5455, 1)\,; \quad S_A=\{3, 6\}\,.$$

From the last S_A, an inequality identical with V_A in the preceding cycle is obtained. But, it can not be a cutting plane. Moreover, this S_A does not satisfy (3. 3) and S_B is not constructed. A Gomory's fractional cut is made.

$$5x_1+3x_2+3x_3+2x_4+2x_5+2x_6+x_7+x_8+2x_9+x_{10} \leqq 5\,,$$

$$x^*=(0, 0, 0, 0, 1, 1, 0, 0, 0, 1)\,.$$

Gomory's fractional method was used in the attempt to solve Example 2. In this case, computation was interrupted after six constraints were augmented, because a computational error greater than threshold value (5.0×10^{-5}) was accumulated at inverse matrix calculation (executed reinversion was five). Identical inspection for computational error was made in the new algorithm, too. In this case, increase of computational error was prohibited by using facets. The computer program used for this experiment will be dealt with in the next section.

5. Experiment and Results

Computer programs for the new algorithm and for the fractional method were made and experiments on the programs were performed using numerical examples generated by a uniform random number generator.

The new algorithm can be regarded as a version of the fractional algorithm where a routine for generating sequentially lifted facets (steps 3 and 4 in the new algorithm) is inserted just before fractional cut generation. A program MIF (Method of Integer Form) of Land and Powell [7] was used with slight modifications as a code for the fractional method. In MIF, every time an optimal solution to a relaxed problem is found, cutting planes corresponding to all non-integer rows are generated and augmented (i. e. subroutine INTCON). But in the modified program, only one cutting plane is generated, corresponding to a source row selected by the rule in step 5 of the new algorithm. This is the main modification and it was necessary in order to confirm the number of augmented cutting planes until an integral optimal solution could be obtained. Besides, some routines such as a systematic numerical example generator, an output routine of experimental results, etc. were added, but they are independent of the solution algorithm itself. A routine to remove additional constraints which are not active at a current optimal point were activated (PURGE). The fractional method can be applied only to problems defined by integral constants and coefficients, therefore, normalization of coefficients was not performed.

A program of the new algorithm was made by inserting steps 3 and 4

into the modified MIF. MIF was used here without further modification.

The problem instances used in the experiment are those with 10, 20, 30 and 50 variables. For each number of variables, 20 instances were tried. The random number generator supported by Hitachi Ltd. was used to generate problems with n variables in the following way. Starting with an initial value (to the generator) 3457, random numbers are continuously generated and multiplied by an upper bound AMAX. The resultant numbers were rounded off. Doing this, integral numbers between 1 and AMAX were generated. The number of generated random numbers were $(2n+1)\times 20$. They were divided into 20 sets of $2n+1$ numbers. The first n numbers in

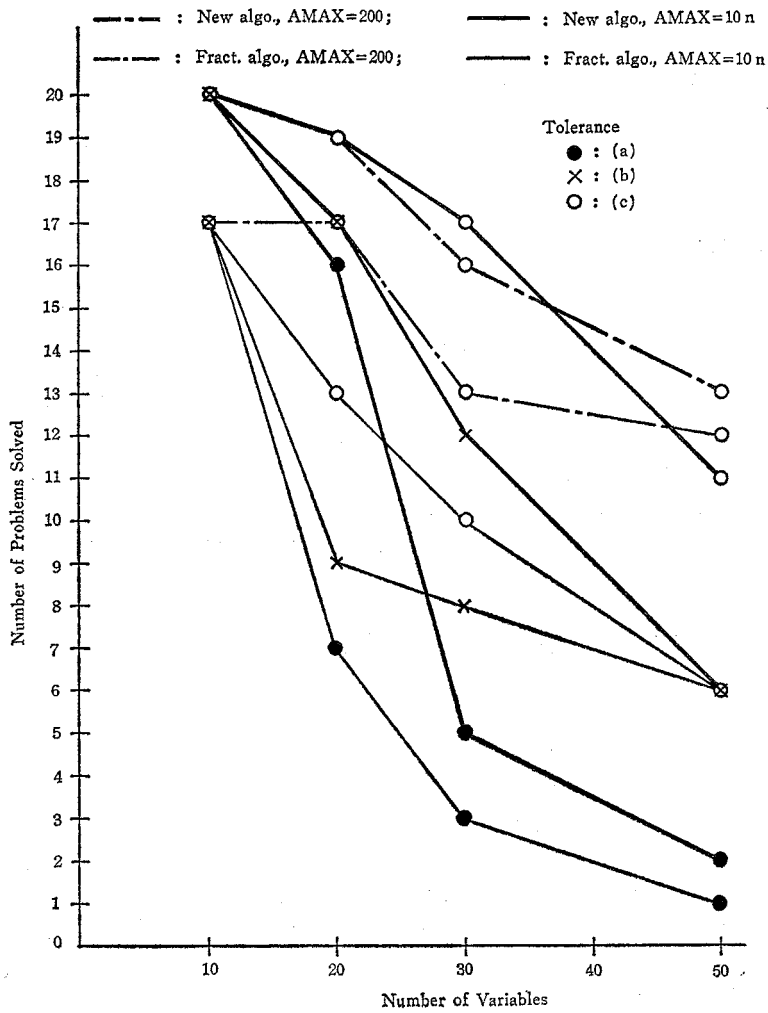

Fig. 5.1. Number of Solved Problems.

each set were used as coefficients of the objective function. The last $n+1$ numbers were arranged in the non-increasing order of the values, the largest was used as the constant and the others were used as coefficient of the constraint. Example 2 in the previous section was generated in this way.

Coefficients were not normalized, and AMAX will affect accuracy of inverse matrix calculations. On the other hand, if AMAX are fixed, the tendency of generating identical values will increase with an increase of n, and the numerical property of generated problems will change. Experiments should be done both for variable AMAX and for fixed AMAX.

When AMAX is changed with n, determining an upper bound of admissible error (tolerance) for inverse matrix calculations is a difficult problem. In our experiments, the following three tolerances were used.

(a) 0.5×10^{-4}, (b) $0.5\times10^{-4}\times(\mathrm{AMAX}/100)$, (c) $0.5\times10^{-4}\times(\mathrm{AMAX}/100)^2$.

AMAX/100 was used because AMAX was set equal to $100(n/10)$ and because all examples were solved by the new algorithm with $n=10$.

Fig. 5. 1 shows the number of problems solved exactly among twenty examples via the number of variables. The heavy lines are for the new algorithm and the light lines for the fractional method. The solid lines are for the cases of $\mathrm{AMAX}=10n$ and the chain lines for $\mathrm{AMAX}=200$. In any case, the new algorithm solved more problems than the fractional one. But, a few problems were solved by the fractional algorithm but not

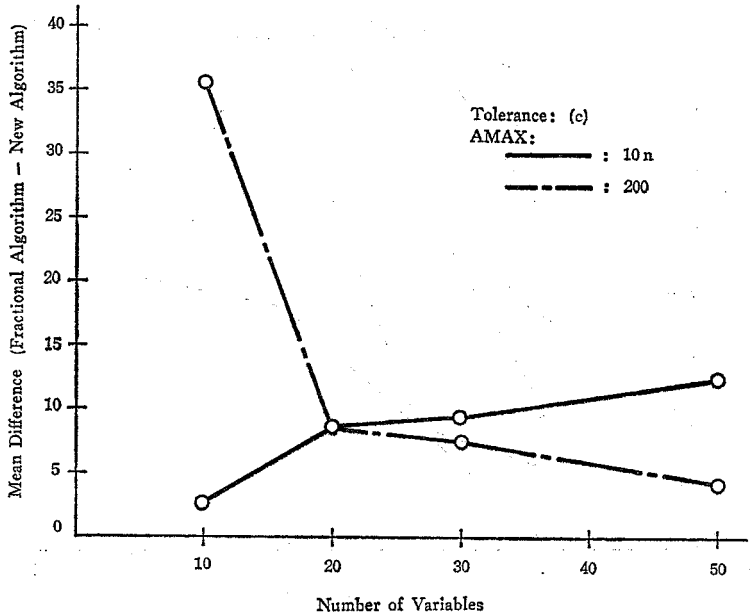

Fig. 5. 2. Mean Difference of Objective Function Values at Termination.

by the new. One of the purposes of these tests was to determine which algorithm stopped with approximate solutions closer the exact optimal solutions. The result is shown in Fig. 5. 2. (Hereafter, we will treat the cases of tolerance (c) because Fig. 5. 1 seems to tell us that tolerance (c) is appropriate for variable AMAX). The vertical axis is the average difference of objective function values between the fractional method and the new algorithm, over the problems which were solved only by one of the two or which were not solved by either. This figure, too, tells us that the new algorithm is excellent.

From the viewpoint of our original objective, it is more interesting to know the causes of this difference. Fig. 5. 3 shows the average number of

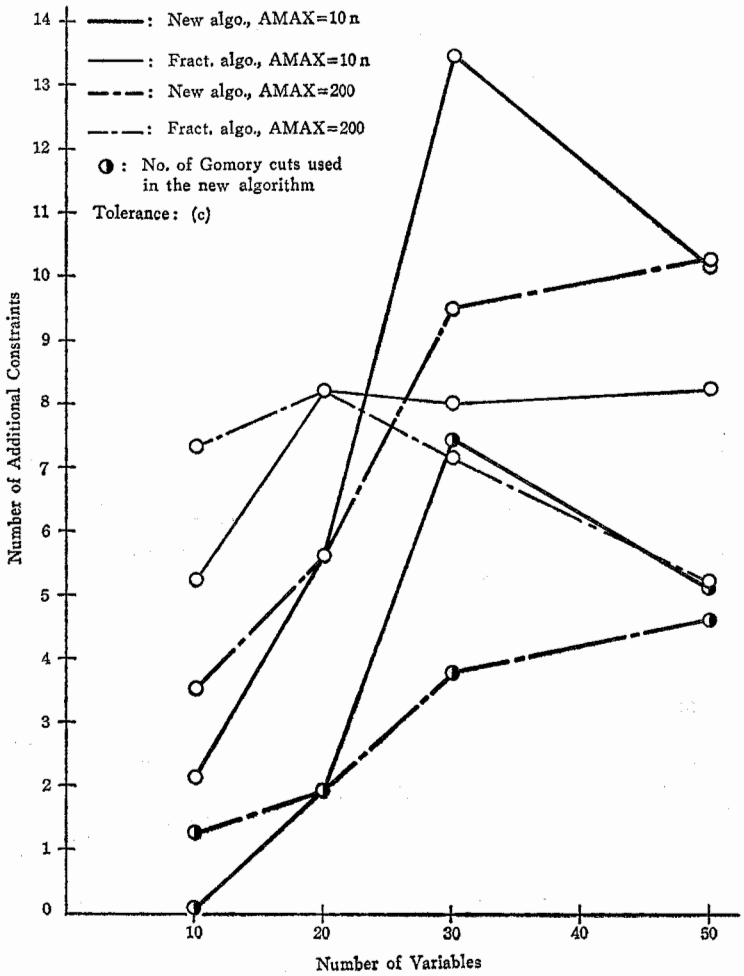

Fig. 5. 3. Number of Additional Constraints.

augumented facets and cutting planes over the problems exactly solved by the new algorithm and by the fractional. First, the number of augmented constraints is smaller in the new algorithm than that in the fractional if the number of variables is small, and larger if the number of variables is large. But, the number of fractional cuts is smaller in the new algorithm than that in the fractional method for any value of n. From this fact, using facets seems to be effective in prohibiting an increase of computational error in matrix inversion. In other words, we could not utilize the fact that facets are, in a sense, the deepest cuts.

Fig. 5.4 shows the proportion of augmented facts when AMAX$=10\,n$ and tolerance of computational error is (c). Numbers above bars are the total numbers of facets for twenty problems. It is clear that the proportion of sophisticated facets becomes high with an increase of n, and that at the same time the ratio exactly solvable problems decreases. Sequentially lifted facets generated by the proposed procedure seem not to cover an important class of facets for cases involving many variables. Validity of this inference is strengthened also by the fact that facets generated from minimal covers

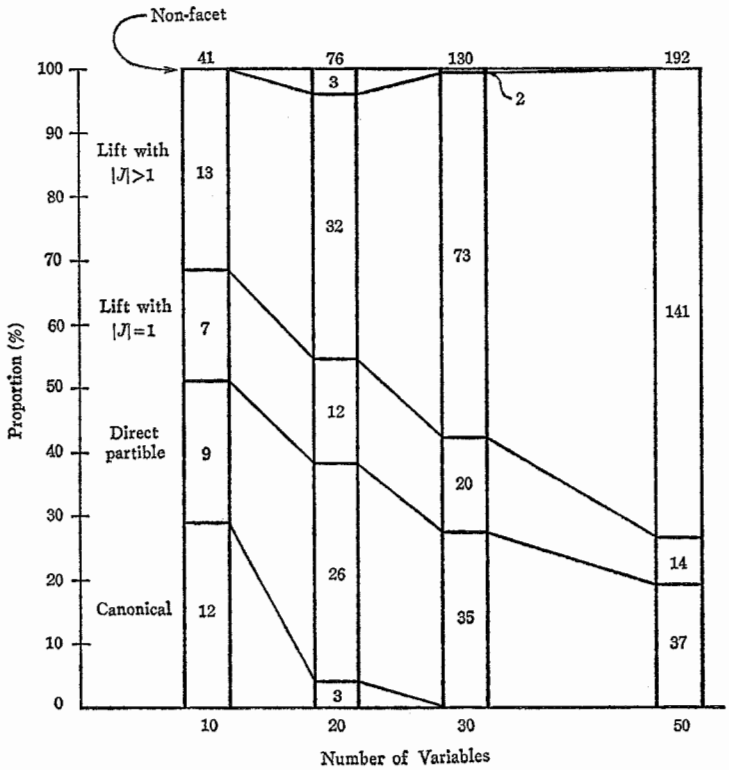

Fig. 5.4. Proportion of Various Facets.

do not cut the optimal solution (i. e. fractional cutting planes must be generated) increasingly with an increase of n.

It is evident that the larger the number of variables, the more the augmented constraints in the new algorithm, compared to the fractional method. But, this does not necessarily mean that fractional cuts shear deeper than the new algorithm when the number of variables is large. Of course, Fig. 5.3 tells us that there are many cases where fractional cuts shear deeper. But, notice also that Fig. 5.2 shows that using facets makes it possible to arrive close to the optimal solution. Besides, there are some cases where the new algorithm solves with fewer augmented constraints. For a problem instance with AMAX $=300$ and $n=30$, the fractional algorithm does not solve it even after twenty-four constraints have been added, but the new algorithm does it after only eleven constraints (five facets and 6 fractional cuts) have been augmented. For a problem with AMAX$=500$ and $n=50$, the fractional algorithm does not solve it with twenty-eight additional constraints, but the new algorithm does it with only ten constraints (nine facets and 1 fractional cut).

The author concludes from the experiment that facets and fractional cuts complement each other. Many more experiments are needed in order to be able to state a concrete result, because the combination of generated Gomory's cuts and facets varies with changes of rules for source row selection, minimal cover construction sequential lifting, etc.

Finally, it is noteworthy that there is a distinguishable difference between the two cutting planes, i. e., that the coefficients and the constant in a facet are relatively small and stable but those of a fractional cut sometimes become very large and instable. This may be a cause for the steep decline in computational accuracy in the fractional method.

6. Concluding Remarks

Conditions under which sequentially lifted facets of the 0–1 knapsack polytope become cutting planes were clarified. If the optimal solution to a relaxed problem has only one or two non-integral variables, the existence of sequentially lifted facets which are of integer coefficients and cutting planes for the optimal solution were guaranteed. Besides, a procedure for minimal cover construction and for sequential lifting, which is expected to generate sequentially lifted facets that can be used as cutting planes, were discussed; and using facets generated by this procedure instead of Gomory's fractional cuts, a new cutting plane algorithm was constructed.

Experiments on this algorithm were performed and the results have clarified the following facts.

The proposed algorithm solves the 0–1 knapsack problem better than

the traditional fractional method. This is hardly thought to result from utilizing the property of facets with the deepest cutting planes. Rather, it is natural to understand that this improvement is accomplished by prohibiting an increase of computational error by using facets with smaller coefficients than those of fractional cuts. On the other hand, the experiments suggest strongly that the class of sequentially lifted facets does not include an important class of facets for cases with many variables.

Although the present research has proved to be an effective approach in using facets to solve the 0-1 knapsack problem, many difficulties remain unsolved. Among them, the improvement of a procedure for constructing minimal covers or for lifting sequentially, the combining of our idea with Gomory's all-integer method, relationships between the change of numerical property in problems to effectiveness of facets, etc. An even more important question is how to improve existing practical algorithms by the combined use of facets.

We have discussed practical implications of facets through integer programming problems with the simplest structure. Facets have been shown to be useful for one of the most complicated problems [3, 11]. Research of this kind has just started and many problems are expected to arise in the future.

Acknowledgement

The author wishes to thank Ms. S. Mineno, Assistant in the Faculty of Economics, Hokkaido University for her kind support on the computational experiments. The experiments were performed by using facilities at the Hokkaido University Computing Center.

References

[1] Balas, E. and Zemel, E.: Facets of the Knapsack Polytope from Minimal Covers. *MSRR#352 GSIA Carnegie-Mellon University*, (Dec. 1974). Also in *SIAM Journal on Applied Mathematics*, Vol. 34, No. 1 (1978), 119-148.

[2] Balas, E.: Facets of the Knapsack Polytope. *Mathematical Programming*, Vol. 8, No. 2 (1975), 146-164.

[3] Crowder, H. and Padberg, M. W.: Solving Large-scale Symmetric Travelling Salesman Problems to Optimality. *Management Science*, Vol. 26, No. 5 (1980), 495-509.

[4] Geoffrion, A. M. and Marsten, R. E.: Integer Programming Algorithms: A Framework and State-of-the-Art Survey. *Management Science*, Vol. 18, No. 7 (1972), 465-491.

[5] Hammer, P. L., Johnson, E. L. and Peled, U. N.: Facets for Regular 0-1 Polytopes. *Mathematical Programming*, Vol. 8, No. 2 (1975), 179-206.

[6] Konno, H.: Structural Approaches. *Communications of the Operations Research Society of Japan*, Vol. 23, No. 12 (1978), 780-787 (in Japanese).

[7] Land, A. H. and Powell, S.: *Fortran Codes for Mathematical Programming: Linear, Quadratic and Discrete.* John Wiley and Sons, 1973.

[8] Maffioli, F.: The Complexity of Combinatorial Optimization Algorithms and the Challenge of Heuristics. Chapter 5 in *Combinatorial Optimization*, Christofides, N. et al. (eds.), John Wiley and Sones, 1979.

[9] Martello, S. and Toth, P.: The 0-1 Knapsack Problem. Chapter 9 in *Combinatorial Optimization*, Christofides, N. et al. (eds.), John Wiley and Sons, 1979.

[10] Mineno, S.: Lifted Facets of the Knapsack Polytope. Manuscript (Jan. 1979).

[11] Padberg, M. W. and Hong, S.: On the Symmetric Travelling Salesman Problem: A Computational Study. *Mathematical Programming Study 12* (April 1980), 78-107.

[12] Peled, U. N.: Properties of Facets of Binary Polytopes. pp. 435-456 in *Annals of Discrete Mathematics*, Hammer, P. L. et al. (eds.), North Holland, 1977.

[13] Sekiguchi, Y.: On the Facets of 0-1 Integer Programming. *Hokkaido Daigaku [Keizaigaku Kenkyu]*, Vol. 30, No. 1 (1980), 65-81 (in Japanese).

[14] Sekiguchi, Y.: A Cutting Plane Method Using Primitive Facets for the 0-1 Knapsack Problem. *Abstract for the Autumn Meeting of the Operations Research Society of Japan*, (Oct. 1980), 136-137 (in Japanese).

[15] Sekiguchi, Y.: Facets of 0-1 Knapsack Polytopes and Applications. *Proceedings of the 1st Mathematical Programming Symposium, Japan* (Nov. 1980), 65-74 (in Japanese).

[16] Suzuki, H. and Iwamura, K.: The Knapsack Problem and the Set Covering (Partitioning) Problem. *Communications of the Operations Research Society of Japan*, Vol. 24, No. 6 (1979), 359-368 (in Japanese).

[17] Taha, H. A.: *Integer Programming: Theory, Applications and Computations.* Academic Press, 1975.

[18] Wolsey, L. A.: Facets for a Linear Inequality in 0-1 Variables. *Mathematical Programming*, Vol. 8, No. 2 (1975), 165-178.