OPERATIONS RESEARCH CENTER

working paper

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROBLEM FORMULATIONS AND NUMERICAL ANALYSIS
IN INTEGER PROGRAMMING
AND COMBINATORIAL OPTIMIZATION

by

Stephen C. Graves and Jeremy F. Shapiro*

OR 095-79 DECEMBER 1979

*Research supported, in part, by the National Science Foundation under Grant No. MCS77-24654.

Problem Formulations and Numerical Analysis in Integer Programming and Combinatorial Optimization

by

Stephen C. Graves and Jeremy F. Shapiro

1. Introduction

Experienced practitioners who use integer programming (IP) and other combinatorial optimization models have often observed that numerical problems to be optimized are sensitive, sometimes extremely so, to the specific problem formulations and data. Unlike linear programming (LP) and nonlinear programming, however, we have not seen the development of a coherent field of numerical analysis in IP and combinatorial optimization. By numerical analysis, we mean techniques for analyzing problem formulations and data that reveal the stability, or predictability, in optimizing these problems, and the expected degree of difficulty in doing so. Three apparent reasons for the lack of development are:

(1) Issues of numerical analysis in IP and combinatorial optimization are intermingled with the artistry of modeling. Thus, unlike LP and nonlinear programming, these issues cannot be related mainly to specific numerical problems that have already been generated by the practitioner.

(2) IP and combinatorial optimization problems possess a wide range of special structures that can sometimes be exploited by special purpose algorithmic methods. The relative merits of general versus special purpose approaches remains an open question, but the ambiguity has inhibited the development of general purpose numerical analytic methods.

(3) Many IP and combinatorial optimization problems can be very difficult to optimize exactly and sometimes even approximately. Systematic procedures for problem formulation and numerical analysis need to be related to approximate as well as exact methods, and the approximate methods are still under development.

Our purpose in this paper is to present a broad sampling of the issues in IP and combinatorial optimization problem formulation, and related questions of numerical analysis. First, primarily through illustrative examples, we discuss the importance of problem formulation. Second, we present briefly some formalisms that can facilitate both our understanding of the art of formulation, and our ability to perform numerical analysis. Finally, we suggest some areas of future research. We believe there is considerable room for the design and implementation of new numerical procedures for the practical solution of IP and combinatorial optimization problems that would greatly expand their usefulness.

2. Problem Formulations

Integer decision variables arise naturally in many applications, such as airline crew scheduling or investment problems where the items to be selected are expensive and indivisible. Integer variables are also used to model logical conditions such as the imposition of fixed charges. Fixed charge problems are among the most difficult IP problems to optimize, in large part because of the awkwardness of expressing logical conditions by inequalities.

Consider, for example, the problem

$$\begin{aligned} \min \ & \sum_{j=1}^{n} c_j x_j + \sum_{k=1}^{K} f_k y_k \\ \text{s.t.} \ & \sum_{j=1}^{n} a_{ij} x_j \geq b_i \qquad \text{for } i = 1,\ldots,m \\ & \sum_{j \varepsilon J_k} x_j - |J_k| y_k \leq 0 \qquad \text{for } k = 1,\ldots,K \\ & x_j = 0 \text{ or } 1,\ y_k = 0 \text{ or } 1, \end{aligned} \tag{1}$$

where each J_k is an arbitrary subset of $\{1,\ldots,n\}$ and $|J_k|$ denotes its size. The quantity f_k is the fixed charge associated with using the set of variables x_j for $j \varepsilon J_k$. The ordinary LP relaxation of (1) is the problem that results if we let the x_j and the y_k take on any values in the range zero to one.

The fixed charge problem (1), and others similar to it, is difficult to solve because the LP relaxations tend to be highly fractional. These LP's are the primary tool used in branch and bound, or other methods, for solving (1), hence there is a tendency for the branch and bound searches to be extensive. For example, suppose we let $\tilde{x}_j$ denote feasible LP or IP values for the x_j variables. If $f_k > 0$ for all k, then the corresponding y_k values for all k are given by

$$\tilde{y}_k = \frac{\sum_{j \varepsilon J_k} \tilde{x}_j}{|J_k|}$$

which in general are fractional numbers.

A number of researchers have observed that problem (1) is easier to solve if the fixed charge constraints are rewritten in the equivalent integer form as

$$x_j - y_k \leq 0 \qquad \text{for all } j \varepsilon J_k.$$

Note that we will have $y_k = \max_{j \varepsilon J_k} \{x_j\}$ if $f_k > 0$ implying $y_k = 1$ if any $x_j = 1$ for $j \varepsilon J_k$. In other words, we can expect the ordinary LP relaxation of the reformulation to be tighter in the sense that it produces solutions with fewer fractions. The reformulation has the effect of increasing the number of rows of the problem, but the improved formulation more than justifies the increase. In one application made by one of the authors, a difficult 80-row fixed charge problem was reformulated as indicated as a 200-row problem for which an optimal solution was easily computed. H. P. Williams (1974, 1978) has made a thorough study of fixed charge problem reformulations arising in a variety of applications. We will return again to the fixed charge problem in the next section when we discuss algorithmic methods.

Similar observations on the importance of obtaining tight LP formulations have been made for plant location problems by Spielberg (1969) and Davis and Roy (1969), for a distribution system design problem by Geoffrion and Graves (1974) and for a production allocation and distribution problem by Mairs, Wakefield, et al (1978). In all cases, the tightness of the LP relaxation was improved by the careful choice of model representation of the logical relationships.

Another class of IP reformulation "tricks" that has proven successful are procedures for reducing coefficients. For example, consider the inequality

$$114x_1 + 127x_2 + 184x_3 \leq 196$$

to be satisfied by integer values for the variables. The difficulty with this inequality arises again from the nature of LP relaxations to the IP problem with this constraint. If the constraint is binding, then we will surely have a fractional LP solution since no combination of the numbers 114, 127 and 184 will equal 196. However, an equivalent integer inequality is

$$x_1 + x_2 + x_3 \leq 1.$$

The inequality with smaller coefficients is less likely to produce fractional solutions in the LP relaxations.

In general, we will not be able to achieve an equivalent representation of an inequality in integer variables with coefficients reduced to the extreme degree of the above example. Gorry, Shapiro and Wolsey (1972) give the following general rule: Any non-negative solution satisfying

$$\sum_{j=1}^{n} a_{ij}x_j \leq b_i \tag{2}$$

will also satisfy

$$\sum_{j=1}^{n} [\frac{a_{ij}}{\lambda_i}]x_j \leq [\frac{b_i}{\lambda_i}] \tag{3}$$

where λ_i is any positive number and [a] denotes the largest integer smaller than a. In other words, the inequality (3) is a relaxation of the inequality (2). This is a useful property because a relaxed IP problem that is easier to solve can be used to provide lower bounds in a branch and bound scheme for solving the original IP problem from which the relaxation is derived. Of course, the critical issue is the strength of the lower bounds produced by the relaxation. Gorry, Shapiro and Wolsey (1972) report on some computational experience with relaxations of this type. Bradley, Hammer and Wolsey (1974) give more powerful coefficient reduction methods, but ones that can require significant computational effort.

There are a wide variety of other tricks that can be used to restructure IP problems to make them easier to solve. Some are very simple; for example, using a budget constraint $8x_1 + 10x_2 + 15x_3 \leq 29$ to deduce that the integer variables x_1, x_2, x_3 have upper bounds of 3, 2 and 1, respectively. Tight upper bounds on integer variables can be very important in reducing the search time required by branch and bound. Krabek (1979) reports good experience in solving IP and MIP problems that have been automatically re-formulated and simplified by tricks such as these. In the following section, we attempt to demonstrate that there is some underlying theory for under-standing and integrating many of these tricks.

A central issue in IP and combinatorial optimization problem formulation is the fact that most combinatorial optimization problems can be represented as IP problems, but sometimes these formulations do not permit special structures to be exploited. In this regard, an important class of problems are those that can be represented, perhaps by suitable transformations, as pure and generalized network optimization problems. For these problems, simplex-like network optimization algorithms will produce integer solutions, often in a very efficient manner.

As an example of this, consider the shift scheduling problem. Suppose each 24 hour day is broken into six four-hour periods, where the minimum staffing requirements for the i^{th} period are given as r_i, $i = 1,\ldots,6$. The problem is to determine the minimum work force to satisfy these requirements where each worker works a continuous two-period (eight hour) shift. Letting x_i be the number of people whose shift starts at period i, the problem can be stated as

$$\min \sum_{i=1}^{6} x_i$$

$$\begin{aligned} \text{s.t. } x_1 + \qquad\qquad x_6 &\geq r_1 \\ x_1 + x_2 \qquad\qquad\qquad &\geq r_2 \\ \ddots \qquad\qquad & \\ x_5 + x_6 &\geq r_6 \end{aligned}$$

$$x_i \geq 0, \text{ integer}$$

Figure 1 gives a reformulation of this problem as a network flow problem, where w_i is the actual staffing level for period i.

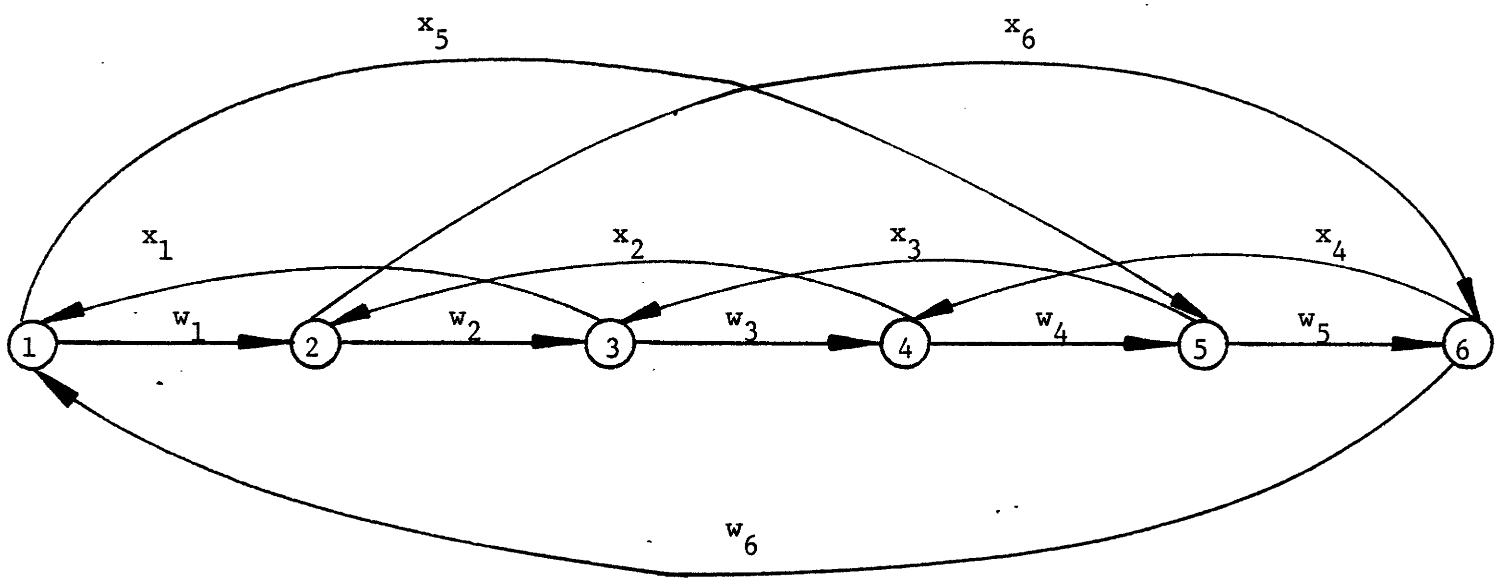

Figure 1

Here the arc flows w_i have a lower bound of r_i, while the arc flows x_i have a unit cost of one.

Since many IP and combinatorial optimization problems can be transformed into network optimization problems if one is willing to greatly expand problem size, an open empirical question is the extent of the useful class of such problems. Glover and Mulvey (1975) give details about these transformations. Additional discussion about the class of applications that can be formulated and solved as network optimization problems is given by Glover, Hultz and Klingman (1979).

In addition to imbedded network optimization problems, there are many other exploitable special structures that can arise in IP and combinatorial optimization problems. However, it sometimes takes considerable insight to identify the special structures. A prime example of this is the traveling salesman problem and related vehicle routing problems. Held and Karp (1970) show that a particular IP formulation of the traveling salesman problem can be recast as a simple graph optimization problem with a small number of side constraints. This reformulation led to great improvements in computation (Held and Karp (1971)). An alternative conceptualization of this problem is given by Miliotis (1978) who reports good computational experience with an IP formulation of the problem that is built up iteratively in the manner of the IP cutting plane method. A third approach is that of Picard and Queyranne (1978) who formulate the time-dependent traveling salesman problem as a shortest path problem on a multipartite graph, supplemented by a set of side constraints.

For the vehicle routing problem, Fisher and Jaikumar (1978) formulate the problem such that a natural decomposition arises in which a generalized assignment subproblem with side constraints is solved iteratively with a series of traveling salesman problems. The generalized assignment problem divides the cities amongst the vehicles, while each traveling salesman problem generates a route for a specific vehicle. Gavish and Graves (1978) give new formulations for the traveling salesman problem and for a variety of related transportation routing problems. These formulations suggest several approaches, which exploit the underlying network structure.

3. Algorithmic Methods and Numerical Analysis

We discussed in the previous section how a great deal of artistry is required to select efficient IP and combinatorial optimization problem formulations, and to apply the correct tricks to improve these formulations. Although it may never be possible to develop formalisms to automatically select efficient formulations for all problems, there remains considerable room for developing further the relevant numerical analytic techniques. We discuss how current and future research efforts in three important areas should facilitate these developments: Lagrangean techniques, elementary number theory, and approximation methods.

It is not within the intended scope of this paper to give an extensive survey of Lagrangean techniques applied to IP and combinatorial optimization problems (see, for example, Shapiro (1979b)). Instead we will review briefly their application to the family of (primal) IP problems, and relate them to the formulation issues discussed in the previous section. The family of IP problems is

$$\begin{aligned} v(y) = \min\ & cx \\ \text{s.t.}\quad & Ax \leq y \\ & x \in X \end{aligned} \qquad P(y)$$

where A is an m x n matrix of integers, y is an m x 1 vector of integers, and X is a discrete set in R^n in which the variables must lie. The function v(y) is called the <u>integer programming perturbation function</u>. We may be interested in v defined either for a specific or for a family of right hand sides y. The partition in P(y) is intended to separate the easy constraints $x \in X$ from the difficult ones $Ax \leq y$. We will discuss below the partition at greater length after we have discussed briefly the Lagrangean methods.

There is an entire family of Lagrangean functions and related dual problems that can be derived from P(y). The simplest is defined for $u \geq 0$ as

$$L^0(u;y) = -uy + \min_{x \in X}(c+uA)x. \tag{4}$$

The value $L^0(u;y)$ is a lower bound on $v(y)$, and the greatest lower bound is found by solving the dual problem

$$\begin{aligned} w(y) = \max\ & L^0(u;y) \\ \text{s.t. } & u \geq 0. \end{aligned} \qquad D(y)$$

In general, $w(y) \leq v(y)$ and if $w(y) < v(y)$, we say there is a <u>duality gap</u> between P(y) and D(y). For fixed $y = \bar{y}$, the lower bounds $L^0(u;\bar{y})$ can be used to fathom subproblems in a branch and bound scheme to solve $P(\bar{y})$ (see Shapiro (1979a,b)). Moreover, the Lagrangean sometimes provides optimal solutions to subproblems by appeal to the following global optimality conditions. These conditions also provide the rationale for selecting the m-vector of dual variables.

Global Optimality Conditions (Version One): For a given primal IP problem $P(\bar{y})$, the solutions $\bar{x} \in X$ and $\bar{u} \geq 0$ satisfy the global optimality conditions if

(i) $L(\bar{u};\bar{y}) = -\bar{u}\bar{y} + (c+\bar{u}A)\bar{x}$

(ii) $\bar{u}(A\bar{x} - \bar{y}) = 0$

(iii) $A\bar{x} \leq \bar{y}$

The following theorem establishes that these conditions provide (globally) optimal conditions to $P(\bar{y})$.

Theorem 1: If $\bar{x} \in X$, $\bar{u} \geq 0$ satisfy the global optimality conditions for the primal problem $P(\bar{y})$, then $\bar{x}$ is optimal in $P(\bar{y})$ and $\bar{u}$ is optimal in the dual problem $D(\bar{y})$. Moreover, $v(\bar{y}) = w(\bar{y})$.

Proof: See Shapiro (1979a, b).

The global optimality conditions are sufficient but not necessary for a given problem $P(\bar{y})$ in the sense that there may not be any $\bar{u} \geq 0$ such that an optimal solution $\bar{x}$ will be identified as optimal by the conditions. This is the case when there is a duality gap between $P(\bar{y})$ and $D(\bar{y})$. We will discuss below how duality gaps can be resolved, and how the resolution is related to the formulation of $P(\bar{y})$. As we shall see, methods for resolving duality gaps are derived from representation of the dual problem $D(\bar{y})$ as large scale LP problems. These LP problems are convexified relaxations of the primal problem $P(\bar{y})$, which lead Geoffrion (1974) to refer to the dual problems as Lagrangean relaxations.

The implicit assumption in our definition and proposed use of the Lagrangean $L^0(u;y)$ is that it can be easily computed, relative to the computation of an optimal solution to $P(y)$. The nature of the set X determines whether or not this assumption is valid. In some cases, the implicit constraints $x \in X$ contain the logical ones not involving data that may change or be parametrized; e.g., fixed charge constraints. Moreover, these constraints need not be stated as inequalities if they can be handled directly as logical conditions in the Lagrangean calculation. By contrast, the constraints $Ax \leq y$

might refer to scarce resources to be consumed or demand to be satisfied and do involve data that may change or be parametrized. In other cases, the constraints $x \varepsilon X$ may refer to network optimization substructures that can be solved very efficiently by special algorithms.

Thus, the artistry of IP problem formulation and analysis reduces in part to the scientific question of how best to partition the constraint set into easy and difficult subsets. As an illustration, consider the following formulation of the traveling salesman problem from Gavish and Graves (1978):

$$\min Z = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \tag{5a}$$

subject to:

$$\sum_{i=1}^{n} x_{ij} = 1 \qquad j = 1,2,\ldots,n \tag{5b}$$

$$\sum_{j=1}^{n} x_{ij} = 1 \qquad i = 1,2,\ldots,n \tag{5c}$$

$$\sum_{\substack{j=1 \\ j \neq i}}^{n} y_{ji} - \sum_{\substack{j=2 \\ j \neq i}}^{n} y_{ij} = 1 \qquad i = 2,\ldots,n \tag{5d}$$

$$y_{ij} \leq (n-1) x_{ij} \qquad \begin{array}{l} i = 1,\ldots,n \\ j = 2,\ldots,n \quad i \neq j \end{array} \tag{5e}$$

$$x_{ij} = 0,1, \; y_{ij} \geq 0 \tag{5f}$$

Here, x_{ij} denotes the inclusion of the arc connecting node i to j in the Hamiltonian circuit, while y_{ij} may be thought of as the flow along that arc. Three possible Lagrangean relaxations are suggested by this formulation. First, by dualizing on the constraint set (5b) or (5c), we obtain the relaxation given by Held and Karp (1970, 1971), a minimal cost 1-tree problem. An alternative relaxation is found by dualizing on the constraint set (5d); here, the resulting Lagrangean can be seen to be an assignment problem. Finally, by dualizing on the forcing constraints in (5e), the Lagrangean separates into an assignment problem in the $\{x_{ij}\}$ variables and a minimum cost network flow in the $\{y_{ij}\}$ variables. The best Lagrangean depends upon the particular problem specification.

Returning to the general IP problem P(y), we can see that some formulation tricks can be interpreted as manipulation of the set X to make it easier to optimize over; e.g., eliminating non-binding logical constraints, tightening upper bounds on the variables, etc. In addition, the form of $L^0(u;y)$ suggests that we may be able to ignore the magnitudes of the coefficients in the system $Ax \leq y$ since small round-off errors would tend to cancel and could be at least partially accounted for by corresponding variations in u. The validity of this observation is an open research question. As we shall see, the magnitude and accuracy of the coefficients in $Ax \leq y$ has an effect, probably an important one, on whether or not there is a duality gap between P(y) and D(y).

Resolution of duality gaps for P(y) is achieved by the application of elementary number theory to strengthen the Lagrangean. The number theoretic procedures are usefully formalized by abelian group theory. First, we need to reformulate the primal problem as

$$
\begin{aligned}
v(y) &= \min cx \\
\text{s.t.}\quad Ax + Is &= y \qquad\qquad P(y)\\
x \in X,\ s &\geq 0 \text{ and integer.}
\end{aligned}
$$

Let Z^m denote the abelian group consisting of integer m-vectors under ordinary addition. Let G denote a finite abelian group and let ϕ denote a homomorphism mapping Z^m onto G. For the moment, we ignore the rationale for G and ϕ. Methods for their selection will be clear after we show how they are used.

The homomorphism ϕ is used to aggregate the linear system $Ax + Is = y$ defined over the infinite abelian group Z^m to a group equation defined over the finite abelian group G. The group equation is added to the Lagrangean

$$L^{\phi}(u;y) = -uy + \min\{(c+uA)x + us\} \tag{6a}$$

$$\text{s.t.}\quad \sum_{j=1}^{n} \phi(a_j)x_j + \sum_{i=1}^{m} \phi(e_i)s_i = \phi(y) \tag{6b}$$

$$x \in X,\ s \geq 0 \text{ and integer} \tag{6c}$$

where $e_i = i^{th}$ unit vector in R^m, $a_j = j^{th}$ column of A.

Although P(y) has been converted to an equality problem, the vectors u are still constrained to be non-negative because $L^{\phi}(u;y) = -\infty$ for all other dual vectors. The new dual problem is

$$
\begin{aligned}
w^{\phi}(y) &= \max L^{\phi}(u;y) \qquad\qquad D^{\phi}(y)\\
\text{s.t.}\ u &\geq 0
\end{aligned}
$$

We define for all $\beta \varepsilon G$ the sets

$$H(\beta) = \{(x,s) \mid (x,s) \text{ satisfy (6b) and (6c) with group right hand side } \beta \text{ in (6b)}\}.$$

Similarly, we define the function

$$\begin{aligned} Z^{\phi}(u;\beta) &= \min \{(c+uA)x + us\} \\ &\text{s.t. } (x,s) \; \varepsilon \; H \end{aligned} \tag{7}$$

For any $\beta \varepsilon G$ and any $y \; \varepsilon \; Z^m$ such that $\phi(y) = \beta$, problem (6) and (7) are connected by

$$L^{\phi}(u;y) = -uy + Z^{\phi}(u;\beta)$$

Problem (7) is a group optimization problem with side constraints determined by the set X. The size of G largely determines the relative computational effort to solve (6). If X simply constrains the x_j to be zero-one, or non-negative integer, then it can be efficiently solved for all $\beta \varepsilon G$ for groups of orders up to 5,000 or more (see Glover (1969), Gorry, Northup and Shapiro (1973), Shapiro (1979b)). If X contains fixed charge constraints defined over non-overlapping sets, efficient computation is still possible (Northup and Sempolinski (1979)). More generally, the effect of various side constraints in group optimization problems has not yet been fully explored (see also Denardo and Fox (1979)). Nevertheless, this is a future research direction of significant importance and promise.

The first version of global optimality conditions can now be adapted to the new Lagrangean.

Global Optimality Conditions (Version Two): For a given primal problem $P(\bar{y})$, the solutions $(\bar{x},\bar{s}) \in H$ and $\bar{u} \geq 0$ satisfy the global optimality conditions if

(i) $L^{\phi}(\bar{u};\bar{y}) = -\bar{u}\bar{y} + Z^{\phi}(\bar{u};\ \phi(\bar{y}))$

(ii) $A\bar{x} + I\bar{s} = \bar{y}$.

The complementary slackness condition of version one has been omitted because the inequalities were replaced by equations. As before, $(\bar{x},\bar{s})$ is optimal in $P(\bar{y})$ and $\bar{u}$ is optimal in $D^{\phi}(\bar{y})$ if the global optimality conditions hold.

With this background, we can make some observations about how the data of an IP problem affects its optimization. We define the <u>degree of difficulty</u> of the IP problem $P(y)$ as the size of the smallest group G for which there is a homomorphism ϕ mapping Z^m onto G such that the global optimality conditions hold relative to some u that is optimal in the induced dual problem. A degree of difficulty equal to one for some $P(y)$ means that the simple Lagrangean L^0 defined in (4) will yield an optimal solution to $P(y)$ via the first version of the global optimality conditions. Primal problems with small degrees of difficulty are not much more difficult to solve. These constructs permit us to study analytically the sensitivity of an IP problem to the data. For instance, there can be a great difference between the degree of difficulty of $P(y)$ and $P(y - e)$. Shapiro (1979c) presents results characterizing the degree of difficulty concept.

A more flexible approach to IP problem solving is achieved if we view the constraints $Ax \leq y$ as somewhat soft, in that we may tolerate some slight violations. If this is the case, we may be able to compute much more easily an optimal solution to $P(\tilde{y})$ for $\tilde{y}$ near y because the degree of difficulty of $P(\tilde{y})$ is much less than P(y). Moreover, each solution of the group optimization problem (7) for all $\beta \varepsilon G$ yields optimal solutions to $|G|$ primal IP problems. Letting $x(u;\beta)$ and $s(u;\beta)$ denote an optimal solution to (7) with group right hand side β, we can easily show by appeal to the second version of the global optimality conditions that this solution is optimal in $P(Ax(u) + Is(u))$. This is the principle of <u>inverse optimization</u> that has been studied in detail for the capacitated plant location problem by Bitran, Sempolinski and Shapiro (1979).

For studying the resolution of duality gaps, we find it convenient to give an LP representation of the dual problem $D^{\phi}(\bar{y})$. Letting (x^t, s^t) for $t = 1,\ldots,T$ denote the solutions in the set $H(\phi(\bar{y}))$, the dual problem $D^{\phi}(\bar{y})$ can be re-expressed as the large scale LP

$$\begin{aligned} w^{\phi}(y) = \min \quad & \sum_{t=1}^{T} (cx^t)\lambda_t \\ \text{s.t.} \quad & \sum_{t=1}^{T} (Ax^t + Is^t)\lambda_t = \bar{y} \\ & \sum_{t=1}^{T} \lambda_t = 1 \\ & \lambda_t \geq 0 \end{aligned} \tag{8}$$

The following theorem characterizes when duality gaps occur and provides the starting point for resolving them.

Theorem 2: Let $\bar{\lambda}_t > 0$ for $t = 1,\ldots,K$, $\bar{\lambda}_t = 0$ for $t = K+1,\ldots,T$, denote an optimal solution found by the simplex method to the dual problem $D^{\phi}(\bar{y})$ in the form (8). If $K = 1$, then the solution (x^1,s^1) is optimal in $P(\bar{y})$. If $K \geq 2$, then the solutions (x^t,s^t) for $t = 1,\ldots,K$ are infeasible in $P(\bar{y})$.

Proof: See Bell and Shapiro (1977).

In the latter case of Theorem 2, the column vectors $Ax^1 + Is^1,\ldots,Ax^K + Is^K$ are used to construct a super group G' of G and a homomorphism ϕ' mapping Z^m onto G' such that $Z^{\phi'}(u;\phi(\bar{y})) \geq Z^{\phi}(u;\phi(\bar{y}))$ for all $u \geq 0$. Specifically, the solutions (x^t,s^t) for $t = 1,\ldots,K$ are infeasible in (7) for the new group equation. Thus, the new dual problem $D^{\phi'}(\bar{y})$ is strictly stronger than its immediate predecessor $D^{\phi}(\bar{y})$. The procedure is repeated until a dual problem is obtained that provides an optimal solution to the primal problem.

Following the construction of Bell and Shapiro (1977), the size of the group G' depends on the size of G and the magnitude of the coefficients of the columns $Ax^t + Is^t$ for $t = 1,\ldots,K$. Clearly, the scale selected for the constraints $Ax \leq y$ plays an important role in determining the size of the group encountered and the degree of difficulty of the IP problems of interest. It is far better to measure resources y_i in tons rather than ounces, depending, of course, on the nature of the application. Recently Bell (1979) has derived some new procedures that permit the construction of a supergroup G' satisfying in many cases $|G'| = 2|G|$ or $|G'| = 3|G|$, regardless of the magnitude of the coefficients. Further theoretical and empirical research is needed to understand the importance of scaling to IP optimization.

Heuristics and approximation methods are the third research area that has an important bearing on understanding and achieving efficient IP problem formulations. The heuristic methods can be viewed as working backwards from IP and combinatorial problems with favorable problem structures towards more complex problems to extend the applicability of efficient solution methods. For example, Cornuejols, Fisher and Nemhauser (1977) have derived a "greedy" heuristic similar to an exact algorithm for the spanning tree problem for a class of uncapacitated location problems. The theoretical efficiency of the heuristic is evaluated using Lagrangean techniques. This approach suggests that, when appropriate, IP and combinatorial models should be selected so that they resemble as much as possible simpler models for which efficient solution methods are known.

The inverse optimization approach mentioned above is an exact approximate method providing optimal solutions to some primal problems that hopefully are in a close neighborhood of a given primal problem. This approach can be usefully combined with heuristics to modify the optimal solutions for the approximate problems to make them feasible in the given problem. More generally, specification by the user of a parametric range of interest for the problem specification might make analysis <u>easier</u> if it eliminated a slavish concern for optimizing a specific troublesome problem. A related point is that the branch and bound approach to IP and combinatorial optimization generally allows and even relies upon termination of the search of feasible solutions with a feasible solution whose objective function value is written ε of being optimal, where ε is a prespecified tolerance level.

4. Conclusions

We have discussed in this paper how efficient IP and combinatorial optimization is highly dependent on problem formulations and data. We also surveyed a number of seemingly unrelated formulations tricks that have been discovered to help achieve efficient formulations. Finally, we argued that formalisms based on Lagrangean techniques, elementary number theory and approximation methods provide a scientific basis for understanding the tricks. Moreover, the formalisms suggest a new research area devoted to numerical analysis of IP and combinatorial optimization problems.

The application of principles of efficient problem formulation is best achieved at the problem generation stage. One example is the cited approach of Krabek (1979) who developed automatic procedures to tighten up MIP problem formulations prior to optimization by a commercial code. Northup and Shapiro (1979) report on a general purpose logistics planning system, called LOGS, that generates and optimizes large scale MIP problems from basic decision elements specified by the user. Many of the MIP problem formulation tricks can be brought to bear on the specific MIP model generated from the decision elements.

References:

Bell, D. E. (1979), "Efficient group cuts for integer programs," Math. Programming, 17, pp. 176-183.

Bell, D. E. and J. F. Shapiro (1977), "A convergent duality theory for integer programming," Operations Research, 25, pp. 419-434.

Bitran, G. R., D. E. Sempolinski and J. F. Shapiro (1979), "Inverse optimization: An application to the capacitated plant location problem," Working Paper OR 091-79, MIT Operations Research Center, Cambridge, MA.

Bradley, G., P. L. Hammer and L. A. Wolsey (1974), "Coefficient reduction for inequalities in 0-1 variables," Math. Programming, 7, pp. 263-282.

Cornuejols, G., M. L. Fisher and G. L. Nemhauser (1977), "Location of bank accounts to optimize float: An analytic study of exact and approximate algorithms," Management Science, 23, pp. 789-810.

Davis, P. S. and Ray, T. L. (1969), "A branch-bound algorithm for the capacitated facilities location problem," Nav. Research Logistics Quarterly, pp. 331-344.

Denardo, E. V. and B. L. Fox (1979), "Shortest route methods II: group knapsacks, expanded network, and branch-and-bound," Operations Research, 27, pp. 548-566.

Fisher, M. L. and R. Jaikumar (1978), "Solution of large vehicle routing problems," paper presented at the TIMS/ORSA meeting, New York City, NY, May.

Gavish, B. and S. C. Graves (1978), "The traveling salesman problem and related problems," Working Paper OR 078-78, MIT Operations Research Center, Cambridge, MA.

Geoffrion, A. M. and G. W. Graves (1974), "Multicommodity distribution system design by Benders decomposition," Management Science, 20, pp. 822-844.

Geoffrion, A. M. (1974), "Lagrangean relaxation for integer programming," Math. Programming Study 2, pp. 82-114.

Glover, F. (1969), "Integer programming over a finite additive group," SIAM J. Control, 7, pp. 213-231.

Glover, F., J. Hultz and D. Klingman (1979), "Improved computer-based planning techniques, part II," Interfaces, 9, pp. 12-20.

Glover, F. and J. Mulvey (1975), "Equivalence of the 0-1 integer programming problem to discrete generalized and pure networks," to appear in Operations Research.

Gorry, G. A., W. D. Northup and J. F. Shapiro (1973), "Computational experience with a group theoretic integer programming algorithm," Math. Programming, 4, pp. 171-192.

Gorry, G. A., J. F. Shapiro and L. A. Wolsey (1972), "Relaxation methods for pure and mixed integer programming problems," Management Science, 18, pp. 229-239.

Held, M. and R. M. Karp (1970), "The traveling salesman problem and minimum spanning trees," Operations Research, 18, pp. 1138-1162.

Held, M. and R. M. Karp (1971), "The traveling salesman problem and minimum spanning trees: part II," Math. Programming, 1, pp. 6-25.

Krabek, B. (1979), "Experiments in mixed integer matrix reduction," paper presented at TIMS XXIV International Meeting, Hawaii, June.

Mairs, T. G., G. W. Wakefield, E. L. Johnson and K. Spielberg (1978), "On a production allocation and distribution problem," Management Science, 24, pp. 1622-1630.

Miliotis, P. (1978), "Using cutting planes to solve the symmetric traveling salesman problem," Math. Programming, 15, pp. 177-188.

Northup, W. D. and D. E. Sempolinski (1979), "Fixed-charge integer programming and group optimization problems," (in preparation).

Northup, W. D. and J. F. Shapiro (1979), "LOGS: An optimization system for logistics planning," Working Paper OR 089-79, MIT Operations Research Center, Cambridge, MA.

Picard, J. and M. Queyranne (1978), "The time-dependent traveling salesman problem and its application to the tardiness problem in one-machine scheduling," Operations Research, 26, pp. 86-110.

Shapiro, J. F. (1979a), Mathematical Programming: Structures and Algorithms, Wiley Publishing Co., New York, NY.

Shapiro, J. F. (1979b), "A survey of Lagrangean techniques for discrete optimization," in Annals of Discrete Mathematics 5: Discrete Optimization, pp. 113-138, edited by P. L. Hammer, E. L. Johnson and B. H. Kortz, North-Holland Publishing Co.

Shapiro, J. F. (1979c), "The integer programming perturbation function," (in progress).

Spielberg, K. (1969), "Algorithms for the simple plant-location problem with some side conditions," Operations Research, 17, pp. 85-111.

Williams, H. P. (1974), "Experiments in the formulation of integer programming problems," Math. Programming Study 2, pp. 180-197.

Williams, H. P. (1978), "The reformulation of two mixed integer programming problems," Math. Programming, 14, pp. 325-331.