

**CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION CFSTI
DOCUMENT MANAGEMENT BRANCH 410.11**

LIMITATIONS IN REPRODUCTION QUALITY

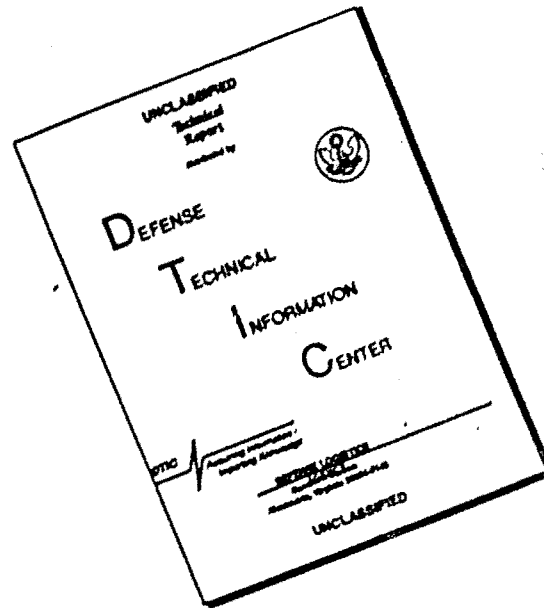
ACCESSION # 6004 6048

- ☒ 1. **WE REGRET THAT LEGIBILITY OF THIS DOCUMENT IS IN PART UNSATISFACTORY. REPRODUCTION HAS BEEN MADE FROM BEST AVAILABLE COPY.**
- ☐ 2. **A PORTION OF THE ORIGINAL DOCUMENT CONTAINS FINE DETAIL WHICH MAY MAKE READING OF PHOTOCOPY DIFFICULT.**
- ☐ 3. **THE ORIGINAL DOCUMENT CONTAINS COLOR, BUT DISTRIBUTION COPIES ARE AVAILABLE IN BLACK-AND-WHITE REPRODUCTION ONLY.**
- ☐ 4. **THE INITIAL DISTRIBUTION COPIES CONTAIN COLOR WHICH WILL BE SHOWN IN BLACK-AND-WHITE WHEN IT IS NECESSARY TO REPRINT.**
- ☐ 5. **LIMITED SUPPLY ON HAND: WHEN EXHAUSTED, DOCUMENT WILL BE AVAILABLE IN MICROFICHE ONLY.**
- ☐ 6. **LIMITED SUPPLY ON HAND: WHEN EXHAUSTED DOCUMENT WILL NOT BE AVAILABLE.**
- ☐ 7. **DOCUMENT IS AVAILABLE IN MICROFICHE ONLY.**
- ☐ 8. **DOCUMENT AVAILABLE ON LOAN FROM CFSTI (TT DOCUMENTS ONLY).**
- ☐ 9.

NBS 9/64

PROCESSOR: PM

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

604648

AD-604648

✓
①

RECENT ADVANCES
IN
LINEAR PROGRAMMING

by
George B. Dantzig

✓ P-652

12 April 1955

Approved for OTS release

32p

COPY	1	OF	1	jp
HARD COPY	\$.2.00			
MICROFICHE	\$.0.50			

DDC
RECEIVED
AUG 27 1964
DDC-IRA D

The RAND Corporation

1700 MAIN ST. • SANTA MONICA • CALIFORNIA

SUMMARY

As interest grows rapidly in industry on the potentialities of mathematical programming techniques, it appears worthwhile to have a paper devoted to some of the more promising developments which may speed up the transition from interest to use. Three topics; have been selected (in three sections that follow) which have recently come into prominence: uncertainty, combinatorial problems, and large scale systems. The reader will find in the course of their discussions that a survey — though perhaps not a systematic survey — has been made of current techniques* in the linear programming field.

*For consolidated sources on techniques in linear programming, see [1], [3], [10], and [21]; for introductory material see [3] and [13].

RECENT ADVANCES IN LINEAR PROGRAMMING

by

G. B. Dantzig*

I. UNCERTAINTY

In the past few months there have been important developments that point to the application of linear programming methods under uncertainty. By way of background let us recall that there are in common use two essentially different types of scheduling applications—one designed for the short run and those for the long run. In such applications the effect of probabilistic or chance events is reduced to a minimum. The usual technique for doing this is to provide plenty of fat in the system. For example, consumption rates, attrition rates, wear-out rates are all planned on the high side. Times to ship, time to travel, times to produce are always made well above actual needs. Indeed, the entire system is put together with plenty of slack and fat with the hope that they will be the shock absorbers which will permit the general objectives and timing of the plan to be executed in spite of unforeseen events.

*Material for this paper has been drawn from speeches delivered before The Institute of Management Sciences, Pittsburgh meeting, October, 1954, and the Symposium on Linear Programming, Washington, D.C., January 29, 1955.

In the general course of things, long-range plans are frequently revised because the stochastic elements of the problem have a nasty way of intruding. For this reason also the chief contribution, if any of the long-range plan, is to effect an immediate decision—such as the appropriation of funds or the initiation of an important development contract.

For short-run scheduling, many of the slack and fat techniques of its long-range brother are employed. The principal differences are attention to detail and the short time-horizon. As long as capabilities are well above requirements (or demands) or if the demands can be shifted in time, this approach presents no problems; i.e., it is feasible to implement the schedule in detail. However, where there are shortages, the projected plan based on such techniques may lead to actions far from optimal, whereas these new methods, where applicable, may result in considerable savings. I shall substantiate this later by reference to a problem of A. Ferguson on the routing of aircraft.

Having reviewed the need for techniques that more effectively account for uncertainty, let us turn our attention now to a sequence of simple ideas that has recently culminated in this extension of linear programming methods. To initiate the discussion, a special application of the "transportation" problem will be described, [1b], [1c], [2]. The RAND Corporation, as you know, has extensive computing facilities that are in

constant use by the research personnel. About two years ago the author was consulted with regard to finding an improved method for scheduling work requiring computation. The need arose because the computation laboratory's "customers," the research people, were unhappy with the long delays generally incurred between the time a request for computation was received and the time their work was completed. The principal cause of dissatisfaction was quite clear for there was one project that was both top priority and so large in volume that it completely absorbed the entire computing capacity for many weeks. The research people, being human, were no longer interested in the computed answers to their problems when the computing lab finally got around to them.

In this example we have a case where the priority method of scheduling is not necessarily the best.

In order to develop a more flexible decision method than priority scheduling, a model was devised* in which the value of a job fell off the longer its completion day was delayed. The final determination of the optimum schedule—where optimum

*The Model of Optimum Scheduling of Projects on Punch Card Equipment was developed by Clifford Shaw of RAND and the author, and reported jointly before the RAND-U.C.L.A. Seminar on Industrial Scheduling in the winter of 1952 (the latter, incidentally, being one of the forerunners of The Institute of Management Sciences).

meant that schedule which gave the greatest total "joy" to its customers—depended on the solution of a transportation problem whose array is given below:

Type Job	1st week	2nd week	3rd week		n-th week	Total Hours Req.
Job 1	x_{11}	x_{12}	x_{13}		x_{1n}	r_1
Job 2	x_{21}	x_{22}	x_{23}		x_{2n}	r_2
...						
Job m	x_{m1}	x_{m2}	x_{m3}		x_{mn}	r_m
Total Hours Available	h_1	h_2	h_3		h_n	

where x_{ij} is the hours (to be determined) assigned to the i^{th} job in the j^{th} week. Thus nonnegative x_{ij} were to be chosen such that

$$\sum_j x_{ij} = r_i, \text{ (hours assigned = hours required for } i^{\text{th}} \text{ job)}$$

$$\sum_i x_{ij} \leq h_j, \text{ (hours assigned } \leq \text{ hours available in } j^{\text{th}} \text{ week)}$$

and such that

$$\sum_i \sum_j c_{ij} x_{ij} = \text{Max}$$

where c_{ij} was the value to the customer per hour expended on his project in the j^{th} week. Typically one would assign less value per hour the longer the job is delayed,* i.e.,

$$c_{i1} > c_{i2} > \dots > c_{in}.$$

Now, in addition to these restrictions typical of a transportation problem this application had the added wrinkle that

$$0 \leq x_{ij} \leq a_{ij}$$

which states that the hours assigned to i^{th} project in the j^{th} week cannot exceed a_{ij} . Typically, $a_{ij} = 40$ hours meant that only one person could be assigned to job j . The adding of these upper bound restraints greatly enlarges the size of the problem if one proceeds in the usual manner of adding extra equations to represent these restraints, [1], [3], [4-I], [5].

*The reader will recognize that putting the value on a per hour basis rather than on a job completed basis is a dodge which an experienced formulator uses to get around indivisibilities that for the most part defy mathematical solution (see Section II on combinatorial problems).

To illustrate, a problem with 18 projects and 10 weeks has a total of 28 equations in 180 unknowns. However, after the upper bound conditions are added by means of the conditions

$$x_{1j} + y_{1j} = a_{1j}$$

(where y_{1j} is a "slack" variable, [3]) the number of equations is $28 + 180$ and the number of unknowns is 2×180 . The constants a_{1j} are referred to as "capacity restraints on routes."

It is not difficult (as we shall show below) that the enlarged problem is still a transportation problem, although strangely enough x_{1j} now appears in three equations. If true, the values of x_{1j} in an optimal solution will still turn out to be integers if the h_j , r_1 and a_{1j} are integers. To show this, we will use Orden's transshipment device, [20]. The enlarged transportation array below illustrates the case for $m=2$, $n=3$, (the procedure is, of course, general):

This constitutes a formal proof that capacity restraints on routes do not alter the character of the problem. It is not recommended as it stands as a short cut computational device. Our purpose now is to show, however, that this problem can be solved with only slightly more effort than the original transportation problem without upper bounds. The procedure—referring back to our original array is to divide the variables x_{ij} into three classes:

- (a) "basic" variables ($m+n-1$ in number) whose determinant is nonvanishing;
- (b) nonbasic variables at lower bound value (0);
- (c) nonbasic variables at upper bound value (a_{ij}).

For example, in the original array, we may indicate basic variables—by putting a circle around them as is often done, [1], [2]; next, place an "upper bound" symbol for variables at their upper bounds. The prices u_i and v_j associated with the equations are computed in the usual manner, i.e., $u_i + v_j = c_{ij}$ for all basic variables x_{ij} . The criterion for optimality for such a system can be shown to be:

$u_i + v_j \geq c_{ij}$ for all nonbasic variables at lower bound value;

$u_i + v_j \leq c_{ij}$ for all nonbasic variables x_{ij} at upper bound value.

If any variable x_{ij} does not satisfy the above optimality criterion, an improved solution may be obtained by increasing or decreasing it until some variable (perhaps itself) either hits a lower or upper bound; it is this variable which is dropped out from the basis.

The reason why this device works—and my remarks now apply to the most general linear programming problem and not just the transportation type—is this: The simplex method divides the variables into two classes which are referred to as basic and nonbasic. The nonbasic variables are customarily set equal to zero and a rule is given as to when it pays to increase any one of the nonbasic variables from zero to the largest value possible which preserves feasibility. A little reflection, however, will

make it clear that the nonbasic variables could have been at any constant admissible value and the same rules would apply on whether it pays to increase their value and, of course, the opposite rule applies as to when it pays to decrease it. If the relative cost factors (e.g., $c_{1j} - u_1 - v_j$) are other than zero, then depending on sign it will always pay to increase or decrease the quantity of a variable. However, if the variable is at either end of its range of values it may not be possible to do this without losing feasibility. For a fuller account of this work see [5] and the Charnes-Lenke paper [4-I].

Let us now look at a second device that makes effective use of this upper bounding technique. Let us consider a situation in which the objective function, instead of being a linear form to be minimized, is of the form

$$(1) \quad \sum_{j=1}^n \phi_j(x_j) = \text{Min} \quad (x_j \geq 0)$$

where $\phi_j(x_j)$ is a convex function and x_j are subject to

$$(2) \quad \sum_{j=1}^n a_{1j}x_j = b_1 \quad (i=1,2,\dots,m) .$$

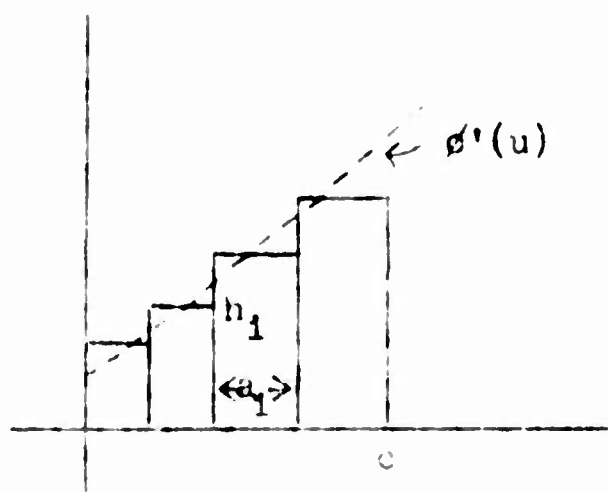
"Convex-separable" is the term used by Charnes to describe this class of objective forms, [4-II], also [5]. The trick here is

to represent $\phi_j(x)$ as indefinite (not definite) integral which is approximated by a sum over k fixed intervals.

To see this clearly, let us note first that any convex function $\phi(x)$ may be written

$$\phi(x) = \int_0^x \phi'(u) du$$

where $\phi(x)$ convex means $\phi'(u)$ is nondecreasing and that $\phi'(u)$ may be approximated by a histogram over some interval $0 < u < c$ where c is some assumed very large upper bound for x . In order to avoid any discussion about how well the histogram fits $\phi'(u)$, which is not germane to what follows, we are actually assuming that $\phi'(u)$ has been replaced by a histogram.



$$h_1 \leq h_2 \leq h_3 \cdots \leq h_k$$

Here $h_1 \leq h_{i+1}$ follows from the convexity of ϕ . We now replace x by

$$x = \Delta_1 + \Delta_2 + \cdots + \Delta_k$$

$$\Delta_1 \geq 0$$

where Δ_1 are nonnegative variables and k is the fixed number of intervals in which the range of x was subdivided. It is easy to see that

$$\phi(x) = \text{Min} \sum_1^k h_1 \Delta_1 \quad 0 \leq \Delta_1 \leq a_1 .$$

Indeed, it is clear that the minimum is attained by choosing $\Delta_1 = a_1, \Delta_2 = a_2, \dots$ until the value of x is exceeded for some i in which case Δ_i is set equal to the residual. Thus the effect of the minimization is to represent $\phi(x)$ by the area under the histogram up to x .

We now employ this approach to solve (1) and (2); the procedure is to substitute for variable x_j and ϕ_j in the linear programming problem by

$$(3) \quad x_j = \sum_{i=1}^k \Delta_{ij} \quad j=1,2,\dots,n$$

and

$$(4) \quad \phi_j(x_j) = \sum_{i=1}^k h_{ij} \Delta_{ij} \quad 0 \leq \Delta_{ij} \leq a_{ij}$$

and noting that since a minimum for $\sum \phi_j(x_j)$ is sought in (1), this implies that the values of Δ_{ij} satisfying (3) must, at the minimum, satisfy $\sum h_{ij} \Delta_{ij} = \text{Min}$.

This manner of treating convex separable objective functions appears to greatly increase the number of variables without increasing the number of equations. However, it is the number of equations that, as a rule, determines the work in the simplex method. Moreover, it should be noted that there are numerous short cuts possible due to the appearance of several columns with identical coefficients (except for the cost row) so that, in fact, it is quite simple to rapidly solve cases involving a convex separable objective form.

Finally, let us turn to a problem involving uncertainty. For this purpose let us consider by way of an example the case of a cannery that has several factories as sources and a number of warehouses as outlets.* The typical formulation leads to a transportation problem.

$$\sum_{j=1}^n x_{1j} = a_1 \quad (a_1 = \text{availability at } 1^{\text{th}} \text{ source})$$

$$\sum_{i=1}^m x_{ij} = b_j \quad (b_j = \text{requirements at } j^{\text{th}} \text{ source})$$

$$\sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} = \text{Min} \quad (c_{ij} = \text{cost of shipping a unit from } i \text{ to } j)$$

*For example H. J. Heinz manufactures ketchup in a half dozen plants and distributes to some 70 warehouses scattered throughout the U.S.; a description of this application by Henderson and Schlaifly where the demands were assumed known, is given in [13].

Suppose, however, that the requirements at the j^{th} source are unknown but are given by some sort of a frequency distribution. Let us suppose that for some destination that

$$\sum x_{1j} = u_j$$

is the amount assigned and v_j is the actual amount demanded, which of course occurs later. Then the revenues which the company receives will be proportional to

$$\text{Min}(u_j, v_j).$$

It is clear that the expected revenues are a function of the amount assigned, u_j . In fact, if we let $\phi(u_j)$ be the expected costs (i.e., revenues, if negative) then

$$\phi(u) = -\alpha \left[\int_{v=0}^u v f(v) dv + u \int_{v=u}^{\infty} f(v) dv \right]$$

where α is the factor of proportionality and $f(v)dv$ is the probability density distribution of v . It is easy to see by differentiating that

$$(5) \quad \phi'(u) = - \int_u^{\infty} f(v) dv$$

and hence the expected cost as a function of u is convex.*

Let us suppose in the warehouse problem that all demands at destinations are subject to uncertainty, and that the objective is to arrange the shipments so that the total expected costs are minimized. The mathematical problem then becomes

$$\sum_{j=1}^n x_{1j} = a_1 \quad (i=1, 2, \dots, m)$$

$$\sum_{i=1}^m x_{1j} = u_j \quad (j=1, 2, \dots, m)$$

where now the expected costs are given by

$$\sum_i \sum_j c_{1j} x_{1j} + \sum_{j=1}^n \phi_j(u_j) = \text{Min}$$

and $\phi_j(u_j)$ are convex functions giving the expected revenues at j if u_j is assigned. From (5) it is clear that the derivative, $\phi'_j(u_j)$, may be approximated by a histogram obtained from the cumulative distribution of $f(v)$ starting with $+\infty$. Accordingly, we may approximate $\phi_j(u_j)$ by our second device and see that we can solve this problem as a regular transportation problem with

*Result due to H. Scarf, The RAND Corporation, see [6].

upper bounds on certain variables. From a practical point of view the work involved in introducing uncertainty into a problem of this type is only slightly greater than before. The resulting allocations are often quite different. In the hypothetical airline example considered by Ferguson and me, the savings in expected costs over the earlier procedure reported in [11] was about 10%.*

For a more complete account of the application of linear programming methods to a class of problems involving uncertainty the reader is referred to [6].

II. COMBINATORIAL PROBLEMS

Turning our attention now to the application of linear programming to combinatorial type problems — the future in this area appears to be less certain. Because of success with the "assignment" problem** and with large scale "traveling salesman"*** problems it does seem to be worthwhile to try to

*The paper on the extension of [11] to the case of uncertainty is under preparation.

**The Assignment Problem is one of assigning optimally m men to m jobs when the value of having the i -th man on the j -th job is known. It is a special case of a transportation problem discussed earlier, [1b], [1c], [2]; see also Votaw-Orden paper [23], and J. von Neumann [22].

***The Traveling Salesman Problem is one of finding the best order to tour m cities so as to minimize total distance covered; see [9].

find further use for such methods in this area.

It is interesting to note that mathematicians have been looking over other branches of mathematics as well with the hope that they may find significant applications of these new methods. Others have been busy, not only encouraging applications in different branches of mathematics, but in different fields such as chemistry, economics, engineering, etc. as well. Perhaps one reason for human progress arises from the curiosity of man to exploit every new tool.

An example that has recently received some attention is one which was discussed by the author briefly in a course in the Department of Agriculture Graduate School a number of years ago. It concerns a classical problem of Chemical Thermodynamics: Given any mixture of gases under constant pressure and temperature conditions, it will eventually reach an equilibrium position; problem — determine the amounts of various types of molecules in the mixture when the equilibrium is reached. This type of problem can be represented in mathematical terms in the form of a system of equations.

$$\sum_j a_{ij} x_j = b_i \quad (x_j \geq 0)$$

$$\sum_j x_j = x$$

$$\sum_j x_j \log (\gamma_j x_j / x) = \text{Min}$$

where x_j = number of molecules of the j^{th} type; a_{ij} = number of atoms of type i in a molecule of the j^{th} type; b_i = number of atoms of type i ; γ_j is a given constant. The form being minimized is called the free energy function. Chemists solve such systems in a fairly efficient way. They start out with that combination of m -type molecules which form the major mole fractions in the mixture. The first approximate solution can be obtained by setting, as first estimate, the minor components zero; from this first approximation of major components an improved estimate on minor components can be made and the process can then be repeated. Chemists are interested, however, in being able to solve such problems without the assumed prior knowledge of which combination constitutes the major components.

Recently, Selmer Johnson,* observing that the free energy function is a convex function in the number of molecules of different types in the mixture, was able from this to set up a procedure that extended many ideas found in the separable convex case discussed earlier to the case of a more general convex function. His procedure, of course, is free of any assumption about major components.

Now, my objective in discussing this problem in this section is simply to point out that there is a combinatorial aspect contained in every linear programming problem. Indeed, the basic problem is one of selecting from the class of extreme

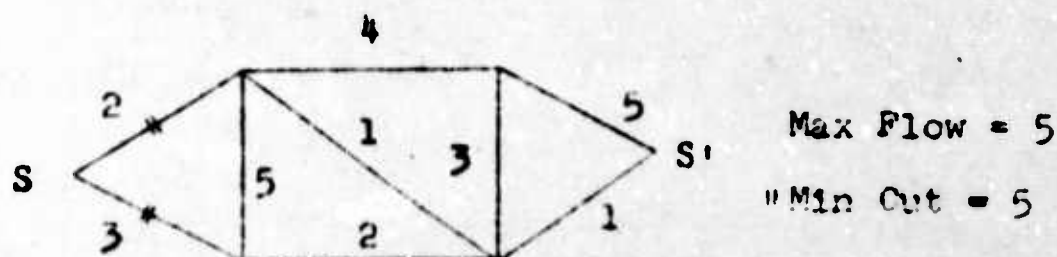
*RAND Corporation paper in preparation.

points of a polyhedral convex the one which maximizes a given linear form. The fact that there are procedures like the simplex method which are fairly efficient in selecting such combinations is the reason why it is tried for certain combinatorial problems. Indeed, it is just those problems where the extreme points of a convex can be identified with the combinations of interest where this approach has paid off. In the case of the traveling salesman problem it was necessary to go further and to find ways of removing extreme points of a convex which could not be identified with tours. In the case of the assignment problem this was not necessary.

Let us turn to topology. Recently, Alan Hoffman* was able to use linear programming to prove an interesting theorem of Dilworth on partially ordered sets—namely, that the maximum number of elements with the property that any pair of elements in the subset are unrelated is equal to the minimum number of disjoint chains covering the set. What he showed was that this theorem was our old friend the duality theorem of linear inequality theory in disguise. Stimulated by this, both Hoffman and independently Fulkerson (joint with the author), [8], were able to show that a theorem due to Fulkerson and Ford on capacities in networks was again our old friend the duality theorem in disguise. There is a close relation between this theorem and a well-known theorem of Menger on graphs. The

*National Bureau of Standards paper in preparation.

Fulkerson-Ford, Max Flow Min Cut Theorem, [12], states that the maximal flow between a fixed source and a fixed sink in a network is equal to minimum sum of capacities on arcs which separate or cut the source from the sink.



Almost all combinatorial applications that I am familiar with in this area depend on the integral character of the basic solutions of the transportation problem. Unless other types of matrices can be discovered with this property for their basic solutions or at least for the optimal basic solutions (as in Markowitz' Metal Process Model, [18]) the potential developments in the combinatorial field may be limited. I have found one case where the optimal values of variables were integers in a linear programming problem which was not of the transportation type in a puzzle which I recently ran across. Jacobs reports that his "Caterer Model"* has this property. This gives hope that the field may not soon be sterile.

*The Caterer Problem concerns the scheduling of purchases and laundering of fresh napkins for a known future schedule of meals; it is a paraphrased aircraft spare engine problem, [15].

III. SPECIAL METHODS FOR SOLVING LARGE SCALE SYSTEMS

The last section is devoted to a short plea that linear programmers pay greater attention to special methods for solving the larger matrices that are encountered in practice. With regard to the possibilities of solving large scale linear programming problems, one can sound both an optimistic and a pessimistic note. The pessimistic note concerns the ability of the problem formulator, either amateur or professional, to develop models that are large scale. The pessimistic note also concerns the inability of the problem solver to compute models by general techniques when they are large scale. If this is so, is not the great promise that the linear programming approach will solve scheduling and long range planning problems with substantial savings to the organizations adopting these methods but an illusion and a snare? Are the big problems going to be solved as they have always been solved—by a detailed system of on-the-spot somewhat natural set of priorities that resolve every possible alternative as it arises?

Let us consider a modest planner who is concerned with the expansion of motor production—let us say a special type motor that requires a special type of steel and must use tools fabricated from this steel and the tools which fabricate these tools also use this steel. The tools that fabricate steel we will call below steel capacity, those that fabricate tools —

tool capacity, and those that fabricate motors - motor capacity. The planner is quite modest because he is willing to consolidate all of the multitudinous activities and items into these simple terms. The initial inventory must satisfy the first 5 equations in detached coefficient form given in the tableau, while the outputs from the activities in first time period must satisfy the next 5 equations.

Variables	Activities (1st Period)										Activities (2nd Period)									
	Motors		Steel		Tool Prod.		Excess Cap.		Motors		Steel		Tool Prod.		Excess Cap.		Motors		Steel	
	Prod.	Storage	Prod.	Storage	Steel Cap.	Tool Cap.	Motor Cap.	Steel	Tools	Motors	Prod.	Storage	Prod.	Storage	Steel Cap.	Tool Cap.	Motor Cap.	Steel	Tools	Motors
	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}	x_{21}	x_{22}	x_{23}	x_{24}	x_{25}	x_{26}	x_{27}	x_{28}	x_{29}	x_{30}
Initial Inventories																				
Steel Cap. =			1					+1												
Tool Cap. =					λ	μ		+1												
Motor Cap. =	+1									+1										
Steel Stocks =	a		-1	1	λ'	μ'														
Motor Stocks =	-1	+1																		
Inventory Balance (2nd period)																				
$0 =$			-1	-1				-1				1						+1		
$0 =$						-1		-1							λ	μ			+1	
$0 =$	-1						-1			-1	+1									+1
$0 =$				-1							a	-1	1		λ'	μ'				
$-d_2 =$		-1									-1	+1								

one week using one shift per day. From the viewpoint of the computer the planner is no longer modest. However, for the planner it is clear that the so-called "detailed" model above is at best only useful as an over-all type of guide, but hardly detailed in a realistic sense.

Let me cite an example from another area — the problem of routing cargo aircraft. Let the variable x_{ijk} represent the number of aircraft of type k routed between city i and j . Let us distinguish between six types of aircraft, ten time periods, and twenty cities. In addition, consider a second set of variables y_{ijl} which is the tons of cargo shipped between city i and j on the way to l . Our equations become

(2)

$$\text{Aircraft in} = \text{Aircraft out: } \sum_j x_{cjk} = \sum_i x_{ick} \quad (k=1, \dots, 6)(c=1, \dots, 20)$$

$$\text{Cargo in} = \text{Cargo out: } a_{cl} + \sum_j y_{cjk} = \sum_i y_{icl} + b_{cl} \quad (k=1, \dots, 20)(c=1, \dots, 20)$$

$$\text{Tonnage Cap.} \geq \text{Tonnage Req.: } \sum_k \lambda_{ij} x_{ijk} = \sum_l y_{ijl} \quad (i=1, \dots, 20)(j=1, \dots, 20)$$

$$\text{Plane Months Available: } \sum_i \sum_j \mu_{ij} x_{ijk} = p_k$$

As we see again such a system involving only a few cities, type aircraft, and cargo destinations generate easily a system in

1,000 equations in 10,000 unknowns. Superficially, a very discouraging situation.

Indeed, at the present time, it is possible to solve rapidly problems in the order of a hundred equations. The Orchard-Hays 701 Simplex Code, [19], has solved many problems of this size with as high as 1,500 unknowns and machine times of five to eight hours as a rule—all with excellent standards of accuracy. However, it is self-evident that no matter how much the general purpose codes are perfected, they will be unable to solve the next generation of problems which will be larger in size.

However, let us note there has been progress in this area: The excellent work of Jacobs on the caterer problem, [15], and the work of Jacobs, Hoffman, [14], Johnson, [16], on the production smoothing problem are examples of what may be done with certain dynamic models with a simple repetitive structure. Cooper and Charnes have employed in their work a number of shortcuts that have permitted resolution of certain large scale systems. At RAND we have found efficient ways to hand compute generalized transportation problems, and Markowitz has proposed a general procedure in this area that is promising. Many models exhibit a block triangular structure and certain partitioning methods have been proposed which take advantage of this type of structure, [5].

At RAND there is a special electronic computer code for a special type of problem called the "metal processing" model, [18]. Mathematically, this may be described as a slightly generalized type of transportation model of the form

$$\begin{aligned} \sum_j \lambda_{1j} x_{1j} &= a_1 & i=1,2,\dots,m \\ (3) \quad \sum x_{1j} &= b_j & j=1,2,\dots,n \\ \sum x_{1j} c_{1j} &= \text{Min} \end{aligned}$$

where λ_{1j} , a_1 , b_j , c_{1j} are known constants and $x_{1j} \geq 0$.

The special code (which only works for a particular choice of c_{1j}) can work out solutions to a 300-equation system of this type in about an hour.

Except in highly specialized models of the transportation type or others where unusual characteristics can be taken advantage of, it is suggested for industrial applications that every effort be made in the early stages to prepare a consolidated version of a model. There are several reasons for this. In the first place this effort results in a model which is often very useful in itself. Secondly, it provides an excellent dry run for methodology. In a word, it is better for administrative and technical reasons to keep the model initially small.

Inevitably when the size of models does increase, there are a number of devices that give promise of greatly reducing the amount of computation. However, considerably more research than has occurred to date is needed. In the first place there appears to be a number of important characteristics commonly found in practical models that need to be exploited:

- (1) Most factors in the coefficient matrix are zero.
- (2) In dynamic structures the coefficients are often the same from one time period to the next.
- (3) In dynamic solutions the activities employed often persist from one period to the next.
- (4) Transportation type submatrices are common.
- (5) Block triangular submatrices are common.

Block triangularity is one of the most promising characteristics to exploit, [5]. When a matrix is composed largely of zeros where the zeros are in no obvious pattern, it is often practical to solve directly for the prices and the representations of the vectors entering the basis rather than to solve for them by means of the inverses of successive bases. The transportation model is a classical case where this approach has paid off. When many variables have simple upper bounds, it is no longer necessary, as we have seen earlier, to add one more variable and equation for each such restraint. Instead, it is possible to slightly modify the original simplex algorithm and apply it to the system excluding the upper bounds.

In many problems there are equations that may be considered as forming a set of side conditions; for example, conditions that capacity of certain machines is never exceeded or the characteristics of a certain product (e.g., viscosity is within specifications). In most problems only a small subset of these "secondary" constraints are likely to be active, i.e., at their critical value—the others being well within capacities or specifications. In such cases it is recommended that the linear programming problem be first solved without regard to these secondary constraints; then the system is enlarged to include the secondary constraints and an initial basis is obtained by augmenting the final basis of the smaller problem. This will result in a basis in which not all variables associated with the secondary constraints are positive. However, in this form the dual simplex procedure of Lemke may be employed, [17]. Often in practical cases only a few iterations are needed to clean up the negative variables and obtain an optimum solution. For fuller discussion of these possibilities see [5], [7].

REFERENCES

- [1] Activity Analysis of Production and Allocation, T. C. Koopmans, Editor, John Wiley and Sons, 1951.
 - (a) Dantzig, George B., "Maximization of a Linear Function of Variables Subject to Linear Inequalities," Chapter XXI, pp. 339-347.
 - (b) _____, "Application of the Simplex Method to a Transportation Problem," Chapter XXIII, pp. 359-373.
 - (c) Koopmans, Tjalling C., and Stanley Reiter, "A Model of Transportation," Chapter XIV, pp. 222-259.
- [2] Charnes, A., and W. W. Cooper, "The Stepping Stone Method of Explaining Linear Programming Calculations in Linear Programming Problems," Management Science, Vol. 1, No. 1, October, 1954, pp. 49-69.
- [3] Charnes, A., W. W. Cooper, and A. Henderson, An Introduction to Linear Programming, John Wiley and Sons, Inc., New York, 1953.
- [4] Charnes, A., and C. E. Lemke, Computational Theory of Linear Programming.
 - I. "The Bounded Variable Problem," O.N.R. Research Memorandum No. 10, Carnegie Institute of Technology, January 7, 1954.
 - II. "Minimization of Non-Linear Separable Convex Functions," O.N.R. Research Memorandum No. 16, May 27, 1954.
- [5] Dantzig, G. B., "Upper Bounds, Secondary Constraints, and Block Triangularity in Linear Programming," (forthcoming Econometrica).
- [6] _____, Linear Programming under Uncertainty, The RAND Corporation, Research Memorandum RM-1374-1, Rev. 8 March 1955.
- [7] _____, Status of Solution of Large Scale Linear Programming Problems, The RAND Corporation, Research Memorandum RM-1375, 30 November 1954.
- [8] Dantzig, G. B., and D. R. Fulkerson, On the Min Cut Max Flow Theorem of Networks, The RAND Corporation, Research Memorandum RM-1418, 1 January 1955.

- [9] Dantzig, G. B., D. R. Fulkerson, and S. M. Johnson, "Solution of a Large-Scale Traveling-Salesman Problem," Journal Operations Research Society, November, 1954, pp. 393-410.
- [10] Dorfman, Robert, Application of Linear Programming to the Theory of the Firm, Univ. of Calif. Press, Berkeley and Los Angeles, 1951.
- [11] Ferguson, A. R., and G. B. Dantzig, "The Problem of Routing Aircraft—A Mathematical Solution," Aeronautical Engineering Review, June, 1955.
- [12] Ford, L. R., and D. R. Fulkerson, Maximal Flow through a Network, The RAND Corporation, Research Memorandum RM-1400, 19 November 1954.
- [13] Henderson, Alexander, and Robert Schlaifer, "Mathematical Programming—Better Information for Better Decision Making," Harvard Business Review, Vol. 32, No. 3, May-June, 1954.
- [14] Hoffman, A. J., and Walter Jacobs, "Smooth Patterns of Production," Management Science, Vol. 1, No. 1, October, 1954, pp. 86-91.
- [15] Jacobs, Walter, "The Caterer Problem," Naval Research Logistics Quarterly, Vol. 1, No. 2, June, 1954, pp. 154-165.
- [16] Johnson, Selmer, and George Dantzig, A Production Smoothing Problem, The RAND Corporation, Research Memorandum RM-1432, 6 January 1955.
- [17] Lemke, C. E., "The Dual Method of Solving the Linear Programming Problem," Naval Logistics Quarterly, Vol. 1, No. 1, March, 1954.
- [18] Markowitz, H., Concepts and Computing Procedures for Certain X_1 Programming Problems, The RAND Corporation, Paper P-602, 19 November 1954.
- [19] Orchard-Hays, William, The RAND Code for the Simplex Method (SXA) (for the IBM 701 Electronic Computer), The RAND Corporation, Research Memorandum RM-1440, 7 February 1955.
- [20] Orden, A., "Transshipment and Storage in the Hitchcock-Koopmans Transportation Problem," DCS/Comptroller, Hq. USAF, March 15, 1951.

- [21] "Symposium on Linear Inequalities and Programming," Wash., D.C., June 14-16, 1951, Project SCOOP, No. 10, DCS/Comptroller, Hq. USAF, 1 April 1952.
- [22] von Neumann, John, "A Certain Zero-sum Two-person Game Equivalent to the Optimal Assignment Problem," Contributions to the Theory of Games, H. Kuhn and A. Tucker (Editors), Vol. 2, Paper No. 1, Princeton Univ. Press, 1953.
- [23] Votaw, D. F., and A. Orden, "The Personnel Assignment Problem," Project SCOOP, No. 10, DCS/Comptroller, Hq. USAF, 1 April 1952, pp. 155-163.