1968

A decision making process for movement planning.

Regan, Thomas Joyce

Massachusetts Institute of Technology

A Decision Making Process for Movement Planning

Thomas J. Regan, Jr., Lieutenant, USN

August 19, 1968

A DECISION MAKING PROCESS FOR MOVEMENT PLANNING

BY

THOMAS JOYCE REGAN, JR.
BS, United States Naval Academy
(1965)

Sumitted in partial fulfillment
of the requirements for the degree of
Master of Science in Civil Engineering

at the

Massachusetts Institute of Technology
September, 1968

Signature of Author ______________________________
Department of Civil Engineering, August 19, 1968

Certified by ______________________________
Thesis Supervisor

Accepted by ______________________________
Chairman, Departmental Committee on Graduate Students

ABSTRACT

A DECISION MAKING PROCESS FOR MOVEMENT PLANNING

by

THOMAS JOYCE REGAN, JR.

Submitted to the Department of Civil Engineering on August 19, 1968,
in partial fulfillment of the requirements for the degree of
Master of Science in Civil Engineering

The problem of troop movement planning is greatly complicated by the fact that there is no single measure of effectiveness for a military deployment. The traditional approach has been to accept closure time as the critical goal of a deployment and to minimize it with little regard for the other variables surrounding the deployment.

This report develops a decision making process for movement planning that considers all of the elements of the goal vector. A linear programming model of the routing and scheduling problem is developed and the results of this model used to generate alternative movement plans. A choice process is described that chooses among the alternatives in light of the many goal variables of the deployment.

The results show that this process promises to be of much aid to the movement planner. The author concludes that the basic concepts of the process can be implemented at the present time.

Thesis Supervisor: Joseph H. Stafford

Title: Assistant Professor

ACKNOWLEDGMENTS

I wish to thank my thesis Advisor, Dr. Joseph H. Stafford, for his unfailing patience, guidance, and enthusiasm during the preparation of this report.

I am especially indebted to the United States Navy for giving me the opportunity to pursue graduate study at the Massachusetts Institute of Technology.

I also wish to acknowledge the support to M. I. T. from the Special Assistant to the Chairman of the Joint Chiefs of Staff for Strategic Mobility through the Defense Communications Agency (Contract Number 100-67-C-0003) which provided the problem area and furnished the clerical support necessary to prepare this report.

A special thanks to Mr. Kent L. Groninger, a graduate student at M. I. T., whose previous work in this field proved to be of great aid.

Of course, I can never fully thank my wife, Fedele, whose unfailing energy in the face of a home and two children inspired me to complete this work.

TABLE OF CONTENTS

TABLE OF CONTENTS, cont'd.

Chapter I

The Routing and Scheduling Problem

I. 1 Introduction

The movement of combat units and support elements from bases in the United States to a combat area overseas requires the utilization of scarce transportation resources. A movement plan is a tentative allocation of resources to each shipment in the deploying force. Its value lies in the fact that it provides decision makers with a guide for issuing movement orders for units, supplies, and transportation resources. Note that the movement plan is a simulation of a potential deployment. It does not have the authority of a movement order and it is used primarily to test the feasibility of providing the required transportation resources in support of an area contingency plan.[1]

The military effectiveness of a deploying force is dependent upon the individual units which comprise the force and the rate and sequence of arrivals of these units in the combat theater. The process of determining how the individual units of a movement plan are to be shipped is known as routing and scheduling. Thus, the military effectiveness of a deployment is dependent upon the routing and scheduling process.

Routing and scheduling of individual shipments is the core of all strategic mobility problems. At the most detailed level of analysis, a complete movement schedule is generated for each unit in the deploying force. This schedule follows the shipment from its origin through the ports of embarkation (POE) and debarkation (POD) to the combat zone. (See Fig. 1. 1) It includes such information as mode of transportation on each link, time over each link, time through each terminal, the shipment's routing, and the specific vehicle or vehicles assigned to each unit.[2]

In contrast, even the most general analysis of capabilities requires consideration of the routing and scheduling problem. For example, a basic single resource model might be described as having a 1000 short ton per day capability. The routing and scheduling process is implicit in this

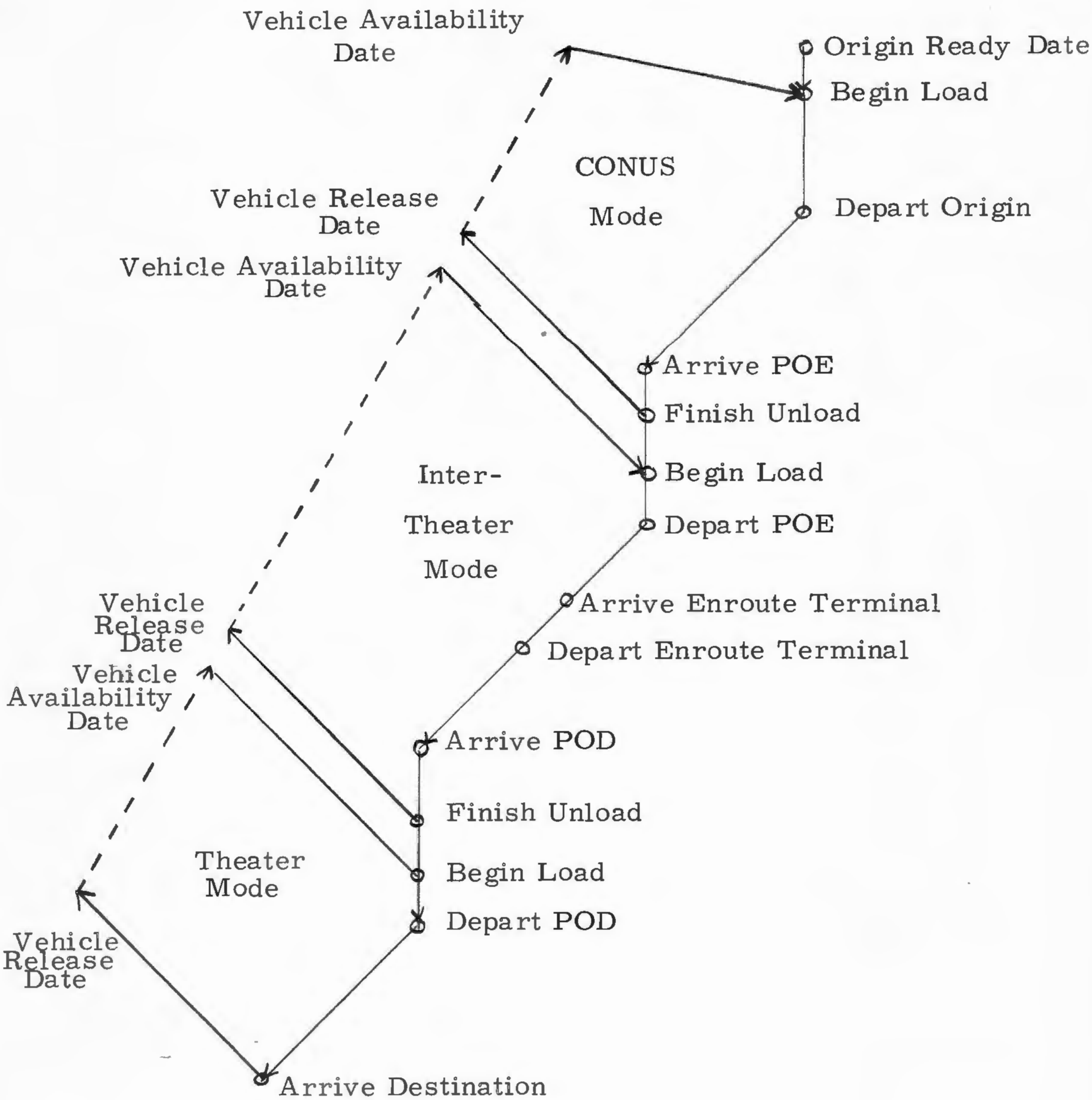


Figure 1.1 - Detailed Movement Schedule

figure because assumptions about types of aircraft, operating policies, turn-around times, and crew availabilities must be made.[3]

Routing and scheduling is traditionally done under an assumed set of operating conditions. The movement plan is generated assuming that air superiority will or will not be maintained, that either good or bad weather will be encountered, or that each port will have a specific capability. The assumptions may be useful in simulating what the planner can expect at the very least or, if his assumptions are optimistic, at the very best; but this procedure clearly limits the decision maker's ability to adjust to changes in the deployment variables. For instance, issues such as the political atmosphere surrounding a deployment, variations in warning time, and concurrent military operations make it unlikely that any set of assumptions will accurately describe a particular deployment.

The traditional procedure is convenient because it allows planners to model the problem and it leads to a feasible movement plan under the assumptions. This approach is inadequate for two reasons. First, closure times today are measured in hours or days rather than weeks and months. Any unforeseen delay represents a potential roadblock that cannot be detoured by three or four days of extra travel time. Such a delay could seriously decrease the effectiveness of the entire deployment. Every effort must be made to detect critical areas of the deployment so as not to stall the entire deployment.

The second reason why the traditional approach is not adequate by today's standards is that it limits the decision maker's flexibility. As soon as he assumes a condition under which the deployment will occur, he has committed himself to a set course of action. It would be much more realistic to give the planner a tool that allows him considerable freedom in the choice of movement plans and offers the needed flexibility to deal with unexpected occurrences. The objectives of this work are to develop a prototype operational tool which gives the planner the needed flexibility in movement planning, and to present a framework for choice among alternative movement plans in view of the complex goal structure of the deployment problem. Before we can consider these objectives in more detail, it is necessary to further define the routing and scheduling problem by identifying the criteria for movement planning.

I. 2 Criteria for Movement Planning

The primary objective of deploying a military force is to maximize the Commander-in-Chief's (CINC) ability to wage war. There is, however, no single, direct measure of the military effectiveness of a deploying force. In order to clarify this point, let us consider the routing and scheduling problem that precedes the movement of each unit of the deployment.

The overall goal, as stated above, is to maximize the military effectiveness of the deploying force. But military effectiveness is a many faceted measure. If we define "maximize military effectiveness" as the overall goal for each shipment of the deployment, then we can use the tree-like structure of Fig. 1. 2 to graphically present the relationships among the goal variables. [4]

Figure 1. 2 shows that the assignment effectiveness of transportation resources to a shipment is composed of two parts. First, the expected loss of effectiveness in other shipments due to the assignment of the resources, and second, the expected military effectiveness of the shipment. The loss of effectiveness can be expressed at a more detailed level as: 1) the added vulnerability due to concentration simply because the assignment of the resource may cause aggregation or concentration of units in a particular area, and 2) the loss of potential transportation capability due to the assignment of the resources for a fixed period of time.

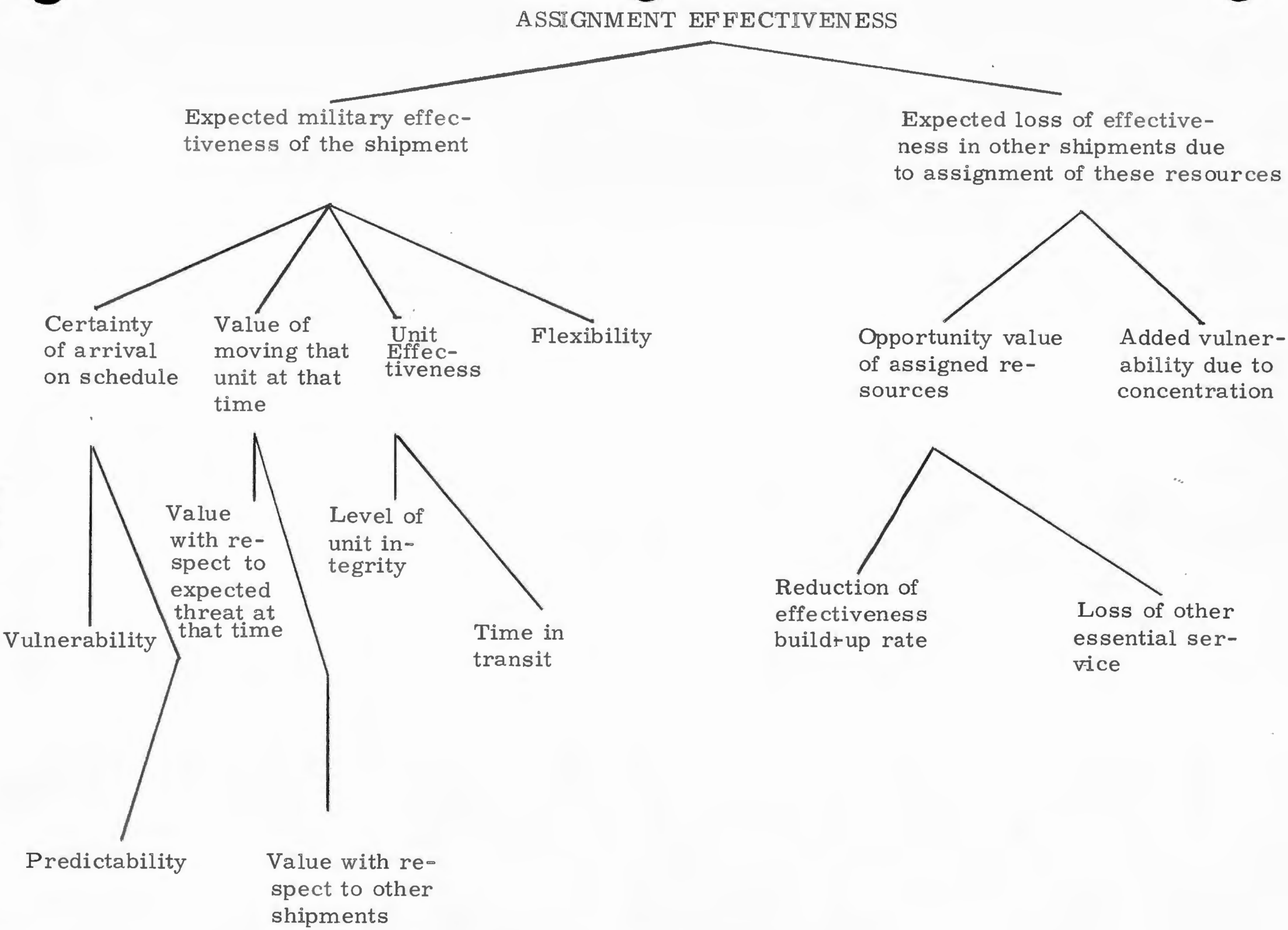


Figure 1. 2 - Routing and Scheduling Goals

The value of lost resources depends upon how those resources would be used. At level 4 in Figure 1.2 it is shown that the military effectiveness of the deploying force is sacrificed if the transportation resources could be used to move shipments which make up the force. Even if the resources were not used to move units of the deploying force, there would still be a loss of effectiveness due to the sacrifice of essential service such as the movement of retrograde cargo or evacuees.

Turning now to the benefit side of the goal structure, it is shown that the goal of expected military effectiveness of the shipment is composed of four goals at level 3. The goal certainty of arrival of the shipment on schedule is the predictability of such occurrences as weather delays and delays due to excessive handling in congested terminals, and the vulnerability of the shipment to enemy action, sabotage, or accident.

The second goal at level 3 of the benefit side is maximizing the value of moving a particular unit at a particular time. The maximum effectiveness would occur by placing every unit at its destination at exactly the time required by the CINC. Since this is seldom possible, the planner attempts to coordinate each shipment's arrival with that of the other units in the force. For instance, it would be improper planning to schedule an armored battalion for arrival if maintenance units and fuel shipments could not be scheduled as support elements.

Unit effectiveness is the next goal at level 3 and is composed of two goals at level 4. The first goal at level 4, unit integrity, is intended to reflect those conditions where the command structure of a unit is fragmented in shipment or where the matchup of personnel and cargo is not properly coordinated. The ideal situation would be to have each unit travel in one vehicle or group of vehicles with its command structure intact. Since this is not always possible, a small integral unit should be identified (such as a company) and every effort made to route it intact and coordinate the matchup between passengers and cargo. The second goal at level 4 is the time spent in transit. Each day a unit spends in transit away from its normal training routine reduces the effectiveness of the unit.

The fourth goal at level 3 is flexibility. Since a deployment occurs in a dynamic atmosphere, it is to the CINC's advantage to be able to alter the deployment at any time in response to a shift in the military threat. For instance, assume a particular unit is scheduled for arrival on D+20 and is routed via sea on D+12 for an eight day ocean transit. If the CINC should require the unit any time during those eight days in response to a shift in the military situation, the unit could not be delivered because it is at sea. However, if the unit had originally been scheduled to depart by air on D+18 to arrive on D+20, then the CINC's need for the unit to combat the new threat could have been met.

Based upon the above considerations, three conclusions can be made concerning the routing and scheduling problem. First, the assignment of scarce transportation resources to a unit of a deploying force is not determined by considering only a single measure of effectiveness. In fact, not only are there many variables to consider, but each is dependent to some extent upon the others. Second, the problem is further compounded because many of the goals are not easily quantified. Consider the problem of evaluating the effectiveness of a transportation resource in its next best assignment. In the commercial market, competitive pricing and rate regulation combine to determine market prices for transportation resources. However, in the deployment situation where values may shift dramatically as the military situation changes, no effective means of placing values on the resources used for a shipment exists.[5] The third conclusion we can draw concerns the effectiveness of the deployment as a whole. Since each shipment of the deploying force is subject to the many considerations of the goal structure of Figure 1. 2, then it follows that the effectiveness of the movement plan as a whole is dependent to some degree on these same variables. In other words, the military effectiveness of a movement plan cannot be expressed as a function of one variable or group of quantifiable variables. It is a vector of related variables. Each goal vector element and each combination of the elements must be considered before the military effectiveness of a movement plan can be determined.

I. 3 Alternative Measures of Effectiveness

The preceding section has shown that the great difficulty in establishing a measure of effectiveness for a deployment is our inability to accurately quantify the various goal elements and to clarify their interactions. In order to model the problem, planners have traditionally accepted the view that some measure of the closure time of a deployment is the most critical component of the goal vector.[6] This approach is comfortable and, through various analytic or hueristic techniques, produces feasible movement plans.

The objective in such an approach is most likely to be that of minimizing deployment closure time (delivering the entire force package to the CINC in the shortest possible time). There are other related measures of closure time. A very interesting approach is taken by Groninger[7] where he develops the concept of the "time weighted tonnage delivered" as a single measure of effectiveness. He uses the following weighting function to weigh early arrivals into the theater heavier than later arrivals:

$$v = \frac{T}{(1 + i)^t} \tag{1.1}$$

where

v is the weighted value of T tons delivered on day D+t.

T is the number of short tons delivered.

t is the time after D day that the shipment arrives in the theater.

i is the discounting factor, $0 < i < 1$, and normally, $.01 < i < .10$.

A weighting function of this type says that it is not only what units arrive in the theater but how soon they get there that determines the effectiveness of the deployment. An armored battalion arriving on D+ 3 is a greater asset to the CINC's ability to wage war than the same unit arriving on D+ 10.

Since the basis for the time weighted tonnage delivered factor (TWTD) is the time element of the goal vector, it is subject to the criticisms made earlier concerning the use of a single measure of effectiveness. In fact, if two movement plans have nearly equal TWTD's, there is no guarantee that the effectiveness of one plan will be greater than the effectiveness of the other plan. The value of this measurement is best appreciated when alternative plans have greatly different TWTD's. The most effective plan will have a TWTD that the experienced planner is able to identify as significantly greater than the alternative. In this case, the plan with the smaller TWTD probably need not be evaluated in terms of the other goals.

In summary, we can conclude that it is probably true that some form of closure time measure is the most significant element of the goal vector. Any plan with a closure time measure significantly less than optimal need not be evaluated in terms of the other variables. The difficulty arises when two or more feasible plans have nearly equal closure time measures that are nearly optimal. In this case, the entire range of the goal vector must be considered, each plan must be evaluated, and the trade-offs between the multiple criteria noted. Only after all of the goals have been identified can a rigorous choice procedure be applied.

I. 4 Movement Planning Considerations

The preceding sections have established the necessity for considering a complex set of goal variables when attempting to evaluate the military effectiveness of a deployment. The objectives of this section are to define each goal variable in more detail and to develop a greater appreciation of the interactions that must be considered.

The goal structure for movement planning is shown in Figure 1.3. (The actual mechanics of deciding exactly what variables should be included in the goal structure will be described in Chapter III.)

The goal of vulnerability reflects the delays a deployment can encounter because of congestion, enemy action, sabotage, accident, or weather. The planner must continually monitor terminal facilities and shipment flows so that capacities are not exceeded. If there is a possibility that a terminal or link capacity may be exceeded, then care must be exercised to ensure that not all of the units of a particular type are scheduled via the route of possible congestion. For instance, if two armored units are required by the CINC, it is preferable to schedule them over two different routes. If they are both scheduled through the same terminal, and that terminal becomes hopelessly congested, then the CINC will receive no armored support. If, however, one unit was scheduled through another terminal, then the CINC would receive half of the needed armored support on time. The important point is that the units that comprise a single phase of the CINC's capability (armored, fire support, infantry) should not be scheduled over a single network path. A delay anywhere along that path would delay an entire phase of the CINC's war making capability. This aspect of vulnerability is important because some units may require special terminal characteristics not common to all the terminal choices open to the planner. An example may be heavy duty cranes to speed the loading of an

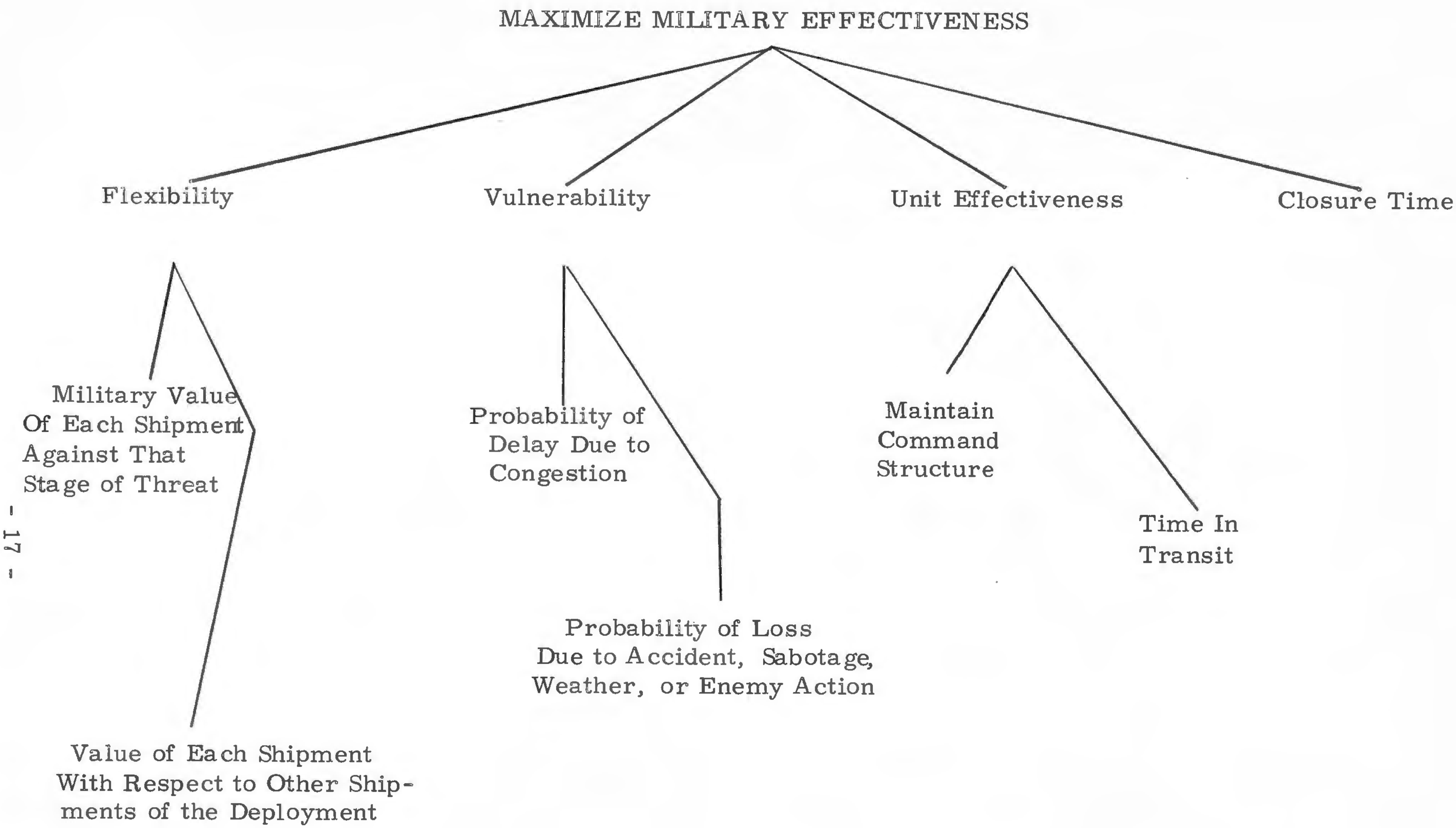


Figure 1. 3 - Deployment Goal Structure

armored battalion. If there are four possible POE's and only one of them has the heavy duty capacity, (but each POE could handle the armored unit with smaller equipment in a longer time period) then the unwary planner may schedule all of the needed armored units through the heavy duty POE in order to shorten deployment closure time. However, we have seen that any congestion of this POE will reduce the chances of these units arriving on time. The congestion could be caused by many factors such as poor scheduling practices, reduction of port capacity due to enemy action, adverse weather conditions, sabotage, or a combination of all of these factors. Regardless of the cause, the planner must consider the possibility of delay and attempt to determine how much it would cost in closure time measure to ship part of the armored units through one (or more) of the POE's less ideally equipped.

Another closely related consideration is the possibility that units may be lost in transit due to weather or enemy action. Let us assume the CINC expresses a requirement for two infantry brigades, one of the two to be a regular infantry brigade for defense and occupation, and the other a mechanized brigade to be used as part of a strike force. If the planner schedules each unit to move as an entity, then the CINC will lose some phase of his capability if either unit is delayed or destroyed. That is, assume the mechanized infantry brigade is scheduled to move by convoy from port A to port B. If the convoy encounters heavy weather or enemy forces, the CINC's strike capability would be seriously weakened.

It is possible, though, to ship the units in combination, one-half brigade of infantry and one-half brigade of mechanized infantry to each shipment. Then, if one shipment is delayed, the CINC still retains both the occupation and strike capabilities (to a more limited extent). Once again, the vulnerability of the deployment has been decreased and the planner must decide what it cost in terms of the other goals. For instance, closure time would probably increase because one-half of each unit would be routed to a different POE from the other half of the unit. Also, unit integrity is obviously sacrificed as the brigades are physically separated during the deployment and unit integrity is maintained only at a lower level.

The second goal variable of Figure 1.3 is deployment flexibility. This goal is intended to model the uncertainty surrounding any military operation due to limited intelligence concerning the enemy's capabilities. As an example, assume that intelligence reports indicate a potential enemy to have no strong defensive lines in a potential crises area. On the strength of this information, the CINC may decide that he will not need armored units in the early stages of the deployment. Let us say the movement planner schedules the unit to depart on D+10 via ship and arrive in the theater on D+20.

If the CINC finds that on D+11 he has encountered a strong point in the enemy's lines, he would want the armored equipment immediately to carry the enemy's defenses. But since the unit is at sea, the CINC could expect to receive it no earlier than D+20. His campaign would be delayed nine days (from D+11 to D+20) in front of the enemy's lines. If, however, the armored unit was scheduled to depart via air on D+18 and arrive on D+20, then the CINC could have requested it when he encountered the unexpected defenses on D+11 and received the needed capability by D+14 or D+15. His attack would be stalled only three or four days. Clearly, flexibility in a movement plan is greatly desired. Here again, the movement planner must decide how much flexibility costs in terms of the other goals.

By scheduling a relatively heavy unit like an armored battalion to scarce air resources, the planner will require many aircraft simply to complete the one shipment. Other units may be forced to find other modes of tranpsortation with the end result that closure time will be affected. Vulnerability will also be affected to some degree because other units may have to be combined and shipped over routes that are not the most favorable. This increases the chances of delay due to terminal and link congestion. Only by considering these variables in their interdependent role can the planner decide if the goal of flexibility is worth achieving for any particular unit of the deployment package.

The third element of the goal tree of Figure 1.3 is unit effectiveness. Unit effectiveness is perhaps the most easily recognized of the goals considered. Military effectiveness is sacrificed each day a unit is away from its training base in a transit status. Also, the more fragmented a unit is in shipment, the more time will be required to re-organize and matchup personnel and equipment at its destination. Unit effectiveness is optimized by completing a shipment in as short a time as possible utilizing vehicles that maintain as nearly as possible the unit command structure.

As with all the elements of the goal vector, it is not always possible to achieve maximum unit effectiveness in routing and scheduling. For example, a division of troops would ideally be transported by air from its home base to its theater destination. However, it is often not possible to air-lift the logistic requirements of an entire division of men simply because the number of aircraft required exceeds the maximum available. Thus, it may be necessary to decrease unit effectiveness by routing some elements of the division by sea, and by doing so increase closure time. Also, by routing part of the division by air and part by sea, the planner has dispersed the unit over time and space so that the vulnerability of the division has been decreased. In this example, one begins to see the complex nature of the goal structure.

Just as it is in error to assume only closure time as a measure of effectiveness, it is equally so to assume merely that each goal of the goal vector must be considered in turn. The goal vector is more a goal fabric. Variations in each element can only be measured by observing how that element causes the others to vary. The fabric is inter-woven and a deformation at one location must produce a change at every location.

I. 5 The Concept of Partial Ordering

I. 5. 1 Introduction

In the preceding section the complexities of the goal vector were more fully developed. It was shown that when alternative movement plans have relatively equal closure times, the entire range of the goal vector must be evaluated before the choice process can be applied and the best plan selected.

In Chapters II and III, the evaluation and choice processes will be presented in some detail. The evaluation process is based upon a linear programming model of the routing and scheduling problem. The input to this model can be graphically represented by a network formulation of the movement plan; and the network is constructed by considering a partially ordered set of requirements furnished by the CINC. Since this procedure is a departure from the CINC's traditional method of presenting his requirements, it is necessary to consider the partially ordered sequence of arrivals concept before proceeding to the models based upon it.

I. 5. 2 Partial Ordering

Since the optimum sequence of arrival of units into the combat area is dependent upon the arrival rate, the CINC cannot specify a single best arrival sequence when stating his need for transportation resources unless the arrival rate is known. The current procedure is for the CINC to assume the arrival rate will be as predicted in certain planning documents and develop a sequenced troop list based upon the predicted build-up rate. This sequence of arrivals includes each requirement identified by the CINC as necessary to the operation and states the exact sequence in which the CINC desires these units to arrive in the combat theater. The transportation planners must meet the demand of this fully ordered sequence of arrivals. If the build-up rate predictions are accurate, and if unit availability is unrestricted, then the military effectiveness of the deployment will be a maximum.

Issues such as the political atmosphere surrounding the deployment, variations in warning time, and concurrent military operations make it extremely unlikely that the CINC will be able to accurately predict the arrival rate. As a result, the troop list may be improperly sequenced. Furthermore, specification of a fixed sequence of arrivals leaves the transportation planners little flexibility. Transportation and unit availability constraints can delay the arrival of specific units but should not delay the overall flow rate. In order to overcome these difficulties, the partially ordered sequence of arrivals has been proposed as a new approach to the routing and scheduling problem.[8]

Instead of being bound by the rigid demand of a fixed sequence of arrivals based upon an assumed build-up rate, the CINC is asked to take the more flexible approach of expressing a pairwise preference over key units of his requirements list. Thus, instead of forwarding a fixed list of arrivals to the planners, the CINC forwards a partially ordered list indicating the sequence in which he would prefer particular units to arrive. The combination of the partial ordering and unit availability results in a scheduling network such as that shown in Figure 1.4. Note that in the figure, (A < B) means that the CINC prefers A to arrive before B arrives and that (E < F) indicates the CINC's desire for Unit E before Unit F.

The preference orderings shown in Figure 1.4 might represent the following strategy to the CINC. At the outbreak of hostilities, he moves in the air-borne infantry because they are in a high state of readiness and can be deployed immediately. At the same time, the engineers and terminal service people are deployed to ready port and air facilities. After some delay, α_1, the facilities are ready to receive mechanized support groups. After ($\alpha_1 + \alpha_2$) days, the engineers have the facilities ready to receive the heavy equipment. At this point the CINC desires regular infantry to gain a strong position, and then requires armored and mechanized units to begin the offensive. The preceding example, as simple as it may be, is designed to bring out some of the kinds of strategic considerations which might be involved in determining the precedence relations among force units for a particular deployment.

Let us consider for a moment the set of seven elements (A, B, C, D, E, F, G) of Figure 1.4. There are 7! = 5040 different fixed orderings for this set. If we consider the preference orderings expressed by the CINC over this seven element set, then there are somewhat less than 5040 feasible orderings. In fact, the feasible orderings (those which satisfy the constraints of the pair-wise preferences made by the CINC) have been reduced to twelve:

A	A	A	A	A	A	A	A	A	A	A	A
B	C	B	B	B	B	B	C	C	C	C	C
C	B	C	C	C	C	C	B	B	B	B	B
E	E	D	D	E	D	E	E	D	E	D	D
D	D	E	E	F	G	D	D	G	F	E	E
F	F	F	G	D	E	G	G	E	D	G	F
G	G	G	F	G	F	F	F	F	G	F	G

Pair-Wise Preferences

A < B	D < G
A < C	E < F
B < E	C < E
B < D	C < D

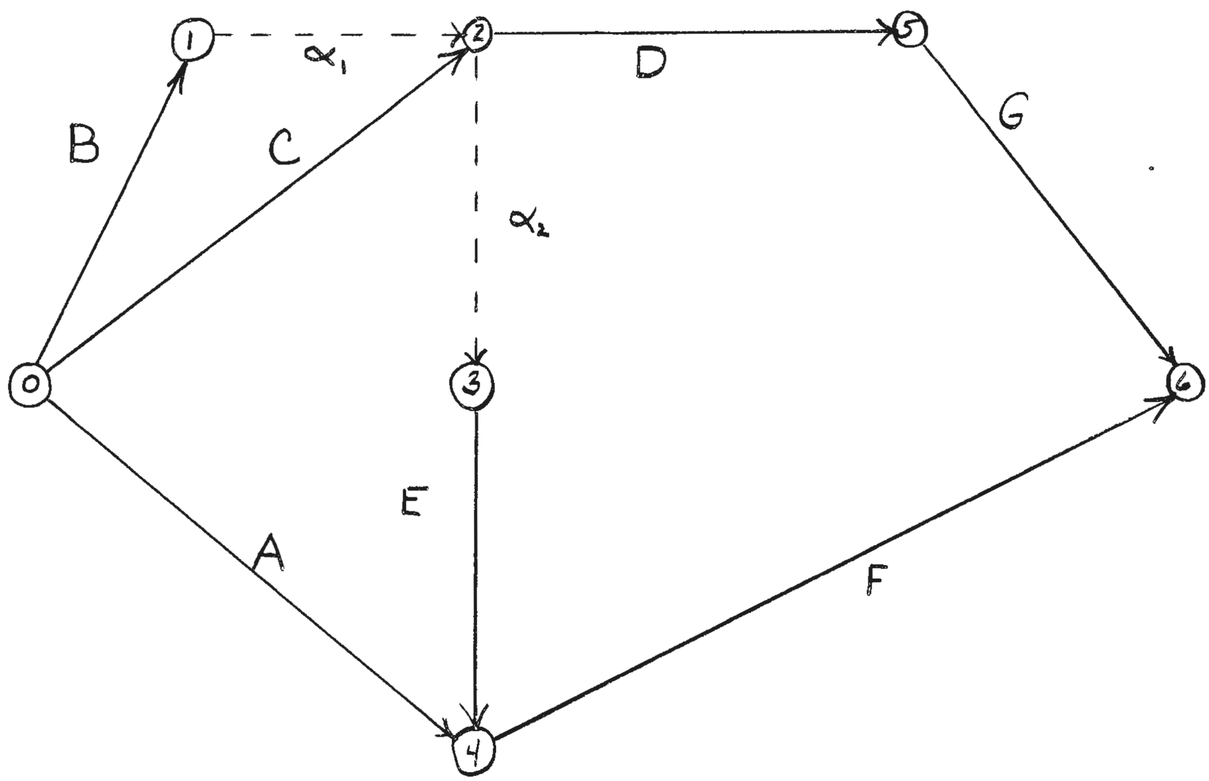


Figure 1.4. Scheduling Network

and these feasible orderings are best illustrated by the network of Figure 1. 4. Note that the directed arcs represent the elements of the set and that they interconnect in such a way that the CINC's precedence relations are satisfied at the nodes.[9]

Investigation of the network yields two important results. First, the transportation planners are not restricted to a single fixed sequence of arrivals. Second, the scheduling network aids in identifying the limited number of feasible movement plans that must be considered. The problem has now become one of evaluation and choice for the transportation planners. They must decide which of the alternative movement plans will yield the maximum military effectiveness.

Chapter II will present the evaluation process in some detail based upon the considerations of the complex goal structure developed in Chapter I. Chapter III will then be addressed to the problem of choice.

Chapter II

Movement Plan Evaluation

II. 1 Introduction

The network formulation of the routing and scheduling problem can be adapted to two general types of techniques, the hueristic models based upon critical path techniques and the analytic approaches such as linear programming. The objectives of Section II. 2 are to outline the various evaluation models considered for solution to the routing and scheduling problem, and to present the reasons why the linear programming model is considered to be the best approach to the problem.

In Section II. 3 the detailed linear programming model will be developed using minimum closure time as the objective function. In Section II. 4 the input data of the linear programming problem will be modified variously to reflect the complexities of the goal vector described in Chapter I. The objectives of Section II. 4 are to determine the trade-offs implied by evaluating a movement plan over the entire range of the goal vector and to produce alternative movement plans based upon those trade-offs.

II. 2 Critical Path Techniques

In the most common version (time-only) of the PERT/Critical Path Method scheduling problem, a project is represented by a network of precedence constrained activities and events. The critical path through the network is found solely from the component temporal relations, without regard to resource requirements, and an initial schedule is set for each activity to achieve a given project completion date.

At this point, the question of resources must be considered. If the resources available must be constrained to certain limits, it may be that the scheduling of the available resources to meet these constraints (and still maintain near minimum closure time) is a major problem. Consider the network problem shown in Figure 2. 1. The critical path has

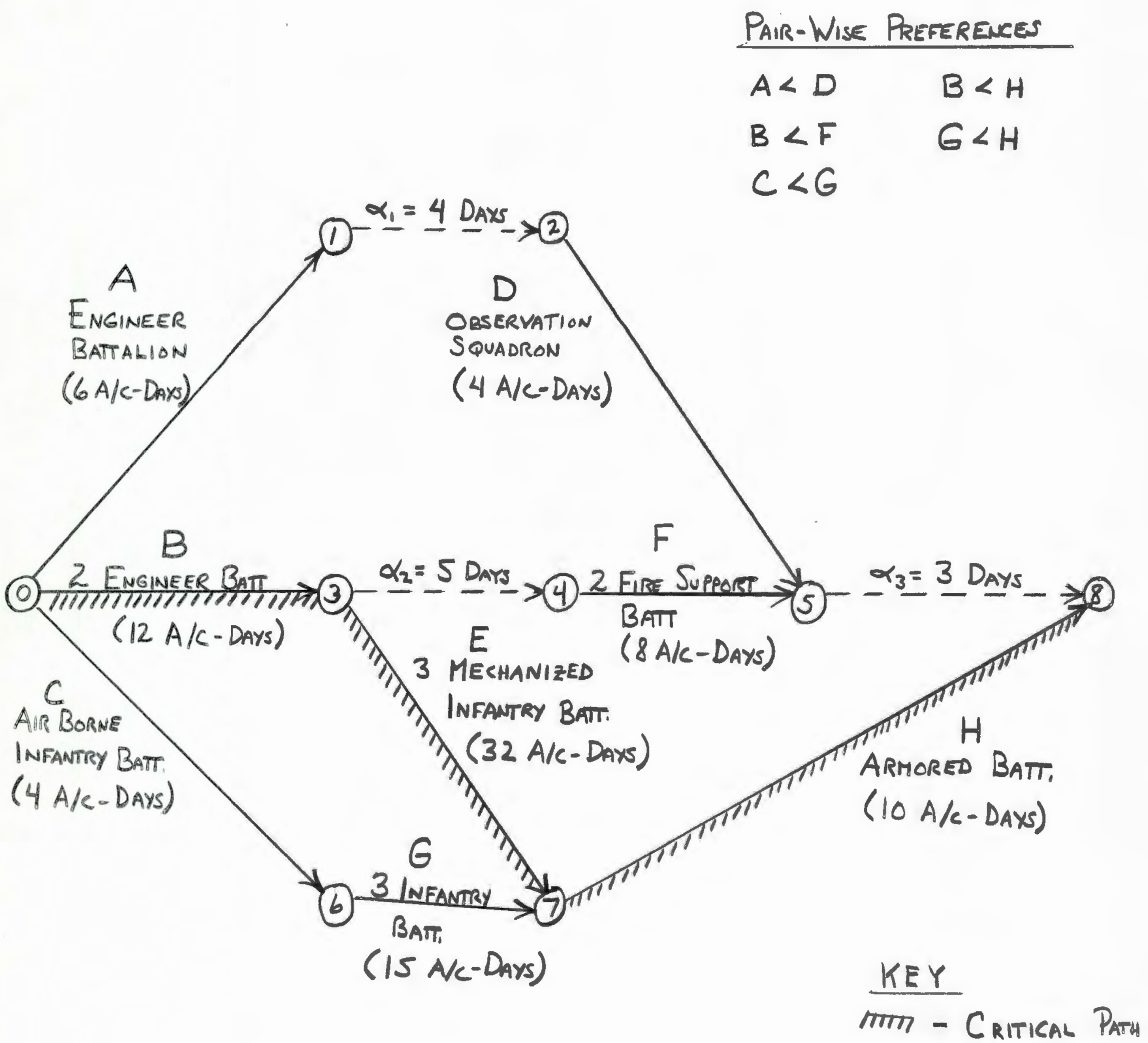


FIGURE 2.1. CRITICAL PATH NETWORK

been identified and labeled as such. The solution technique for this problem, suitable for hand computation and computer programming, was developed by G. H. Brooks of Purdue University, and is described in Moder and Phillips.[2] In order to use this problem as an example, we fix the maximum resource availability at seven.

The results of this example show that the project could be completed in nineteen days with the following aircraft utilization:

DAY:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
#A/C:	6	6	7	7	4	4	4	6	6	6	6	5	5	5	6	2	2	2	2

Note that only on days three and four is the maximum number of aircraft available actually used. On other days, as many as five aircraft sit idle. This observation leads one to suspect that this method has not led to an optimal solution. In fact, if we consider the force package, we see that there are a total of 91 aircraft-days that must be scheduled (Fig. 2. 2). If there are seven aircraft per day available, then the shipment could theoretically be moved in:

$$\frac{91 \ \text{A/C - day}}{7 \ \frac{\text{A/C - day}}{\text{day}}} = 13 \text{ days} \qquad (2.1)$$

This is the minimum time in which the deployment could occur. Clearly, the hueristic approach has produced a feasible, but far from optimal solution.

The reasons why there are no generalized procedures for producing an optimal schedule stem from the difficulties encountered in attempting a mathematical formulation of the problem.

These difficulties are described below.

TOTAL RESOURCE REQUIREMENT

ACTIVITY	UNIT	REQUIREMENT
(0,1)	ENGINEER BATTALION	6 A/C - DAYS
(0,3)	2 " "	12 " "
(0,6)	AIR BORNE INFANTRY	4 " "
(1,2)	DUMMY	0
(2,5)	OBSERVATION SQDN	4 A/C - DAYS
(3,4)	DUMMY	0
(3,7)	3 MECHANIZED INFANTRY BATT	32 A/C - DAYS
(4,5)	2 FIRE SUPPORT BATT.	8 " "
(5,8)	DUMMY	0
(6,7)	3 INFANTRY BATT.	15 A/C - DAYS
(7,8)	ARMORED BATT.	10 " "
		91 A/C - DAYS

FIGURE 2.2. TOTAL RESOURCE REQUIREMENT

First, it may be true that the activity time estimates are influenced subjectively by the estimator's knowledge of available resources. In this case, subdividing an activity may invalidate the original time estimates. In fact, the functional relationship between time estimates and resource availabilities is an unknown quantity. Even if the relationship were known, there is no provision in the hueristic approach for incorporating it into the model.

The second difficulty in obtaining an optimal solution is that the obvious approach when conflicts in resource usage occur is to shift the slack activities in order to remove the conflicts. In some cases, activities can also be divided. It is evident that the alternatives available for scheduling the various activities amount to a combinatorial problem of formidable magnitude for practical sized problems.

These difficulties were approached by Johnson[3] in his work on the resource constrained optimization problem. The problem he considered was to find a schedule for a known set of tasks that minimizes the time required to complete all work, while observing a complex set of logical task precedence requirements and stated limits on resource use in each time period. The basis for his work is the CPM/PERT techniques described above, but Johnson attacks the problem of resource constraints and optimality.

His approach is to view the problem as a non-stochastic sequential decision process where all schedules can be enumerated using a decision tree. However, since practical problems typically offer many feasible alternatives at any given decision point, and also contain a large number of such decision points, complete enumeration is difficult for real life problems.

The algorithm developed by Johnson is based on the "branch and bound" technique used by Little, et al[4], for the traveling salesman problem and employs the concept of "partial solution dominance" described by Weingartner and Ness.[5] The basic idea is to efficiently search the decision tree of all feasible schedules, avoiding investigation of branches that cannot possibly yield a shorter schedule than one already obtained. This is done by calculating a minimum bound for all complete schedules that can follow from a particular sequence of decisions, and terminating the search along this "route" if the minimum bound is greater than or equal to an existing complete schedule. In addition, the unbounded alternatives are explored in an efficient manner by hueristically selecting a likely alternative and proceeding on to the resulting new decision point (if it is not bounded). Only when a better complete schedule is found, or when all branches from a given decision point are bounded, does the algorithm return to consider branches from earlier decision points.

The results show that the extremely rapid growth of the decision tree with problem size often outstrips the algorithm's ability. Hence, Johnson concludes that the technique is not a reliable optimization procedure for projects of realistic size and complexity. However, optimal solutions were found for projects of fifty tasks with very short computation times.

The technique as it stands can be considered an economically reliable optimization procedure for less than fifty tasks, but may be used only as an approximate solution method for larger problems. Once again, for problems of practical size, the alternatives available for scheduling the various tasks amount to a combinatorial problem of formidable magnitude.

II. 3 The Linear Programming Model

In view of the limitations of the hueristic approaches, a linear programming formulation of the routing and scheduling problem was developed by Groninger.[6] This method is based upon a rather unique approach to the problem which lends itself to linear programming techniques quite easily. In this section, the linear programming model will be presented and an example problem solved with the objective of minimizing deployment closure time. The results of this example will then be used in the sensitivity analysis of Section II. 4.

To develop the model, we start with the scheduling network portrayal of a movement plan. The network and the CINC's pair-wise preferences are shown in Figure 2. 3. We lable each node by a non-negative integer such that when node i precedes node j, then $i < j$. In particular, the source is labeled 0 and the sink is labeled n. Each arc is designated by the number of its predecessor and successor nodes. If node i is at the arc's tail and node j at the arc's head, then the arc will be denoted by the number pair (i, j).

Now we consider the dimension of time. The time required to perform any task in the movement plan is a function of the amount of transportation resource allocated to it. Let us call the time-resource trade-off a duration function, $g_i(x)$, where:

$g_i(x)$ = the minimum amount of time required to perform task i if an amount of resource x is allocated to it.

An important assumption of the model, as reflected in the above definition, is that the distribution of resource use over the duration of the task is uniform and equal to x. In general, the form of the duration function would be non-increasing. It is natural to assume that the

PAIR-WISE PREFERENCES

A < E	D < G	C < E
A < G	E < F	C < D
B < E	B < D	

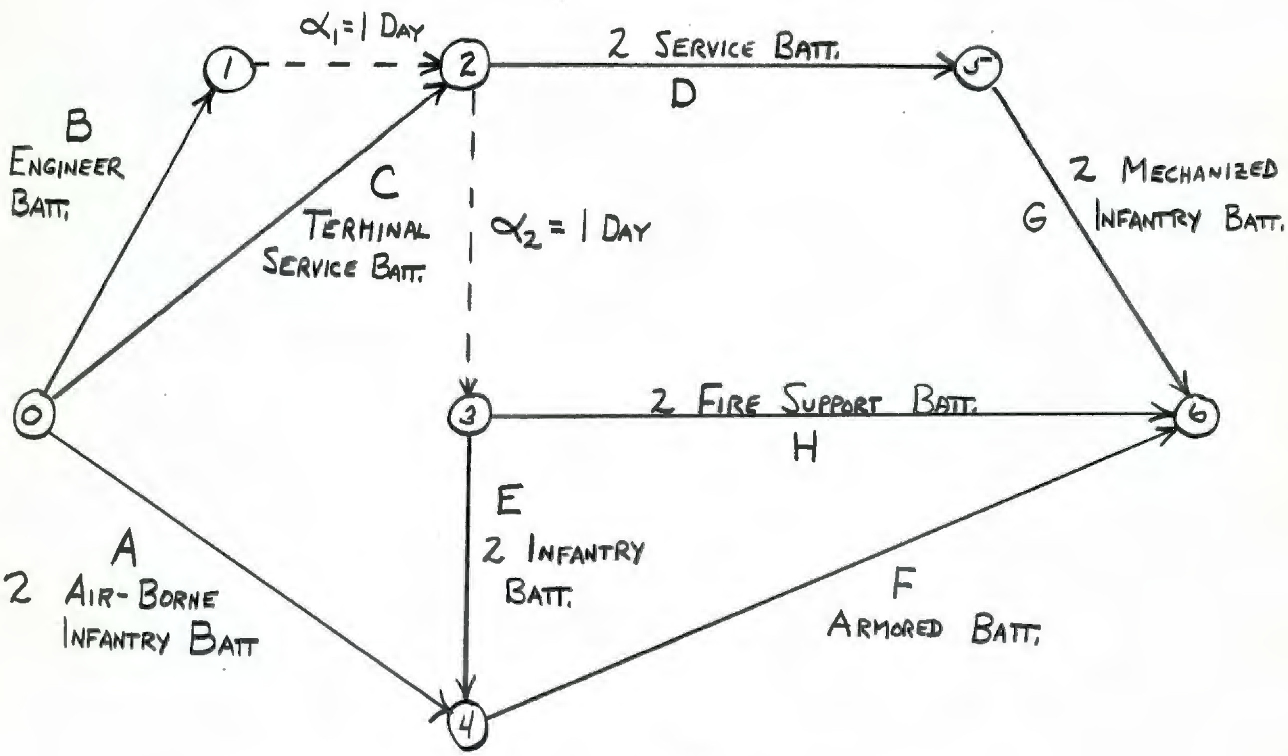


FIGURE 2.3. EXAMPLE DEPLOYMENT PROBLEM

greater amount of transportation resource assigned to a task the less time required to complete it. If we consider, for example, a single resource capability such as airlift, the time required to transport a movement package is inversely proportional to the number of aircraft employed:

$$g(x) = \frac{T \quad d}{x \quad p} \tag{2.2}$$

where:

T = size of movement package (short tons)

d = distance it must be transported (n - mi)

x = number of aircraft employed

p = productivity of a single aircraft (ST - n - mi/day)

The typical duration function for airlift is shown in Figure 2. 4 (a).

The shape of the duration function suggests that the shipment could be completed in the shortest time by assigning more and more transportation resources until congestion of terminal facilities occurs. However, this is not a practical approach to the problem because we must consider the transportation resource as a scarce commodity and assume the lift capability during a deployment is constrained by a fixed maximum.

In certain instances, the duration function for an airlift capability may have an entirely different shape. Consider the activity (1, 2) in the network shown in Figure 2. 3. Recall from Chapter I, Section I. 5. 2 that this activity represents a time delay of α_1 days required for the engineers to ready the facility to receive support elements. The shape of the duration curve for this dummy activity is shown in Figure 2. 4(b).

Note that regardless of the resources assigned to this activity, the duration time remains a constant.

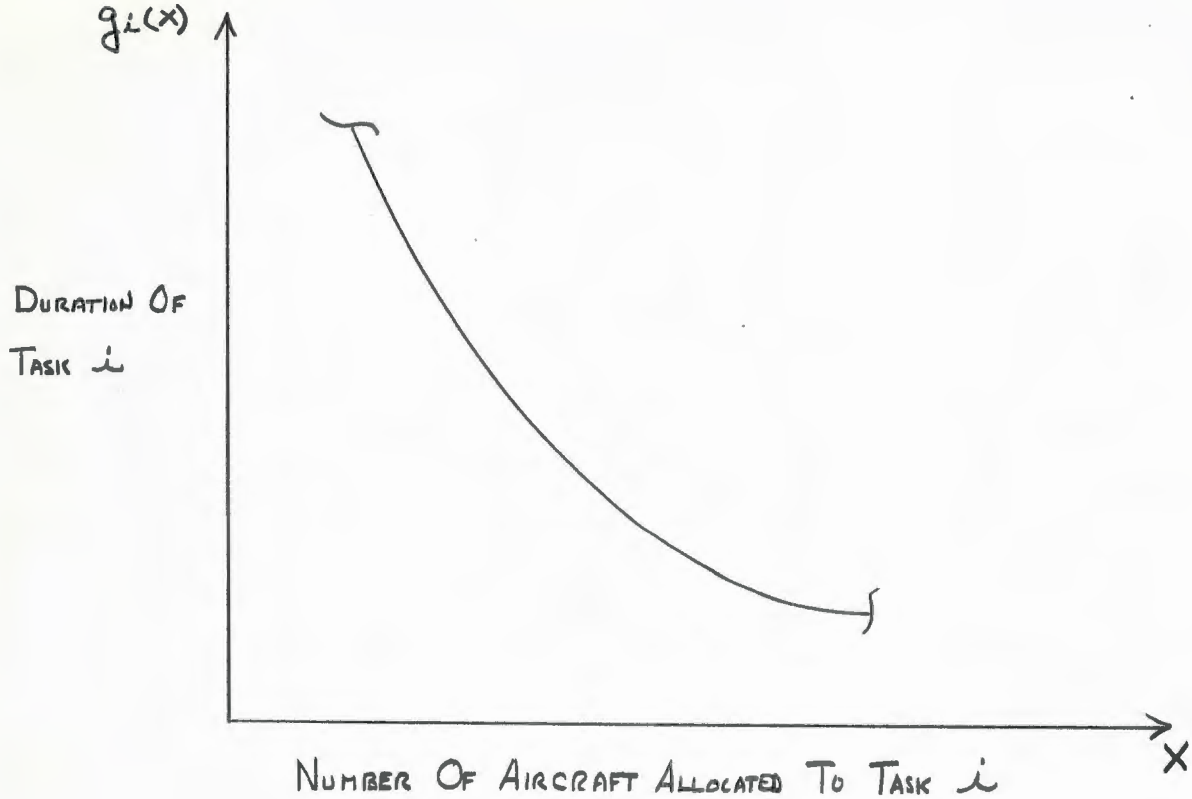


Figure 2.4(a). Typical Duration Function For Air Craft

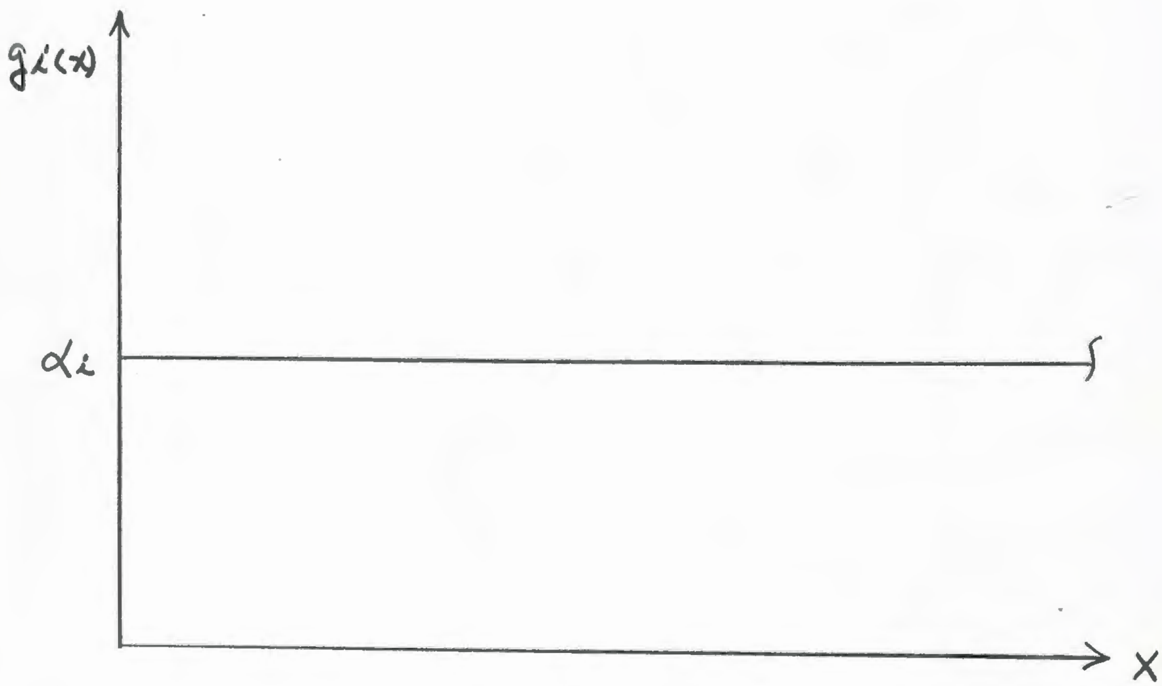


Figure 2.4(b). Duration Function For Dummy Activity

We see now that we have the makings of an optimization problem:

1. We would like to complete the movement as quickly as possible where
2. The tasks of the movement are to be performed in the sequence portrayed by the scheduling network and
3. Transportation resources do not exceed the stated availabilities.

The optimization problem can be concisely stated as follows:

> For a given scheduling network, determine the allocation of resource x to the arcs that minimizes the duration of the longest time chains through the network from source to sink where the time to perform each task is governed by its associated duration function, g(x), and the amount of resource employed does not at any time exceed a specified level.[7]

In order to formulate the model based upon the above considerations, we define:

gij(xij)	=	the duration function for arc (i, j)
xij	=	the amount of resource x allocated to arc (i, j)
X	=	the total amount of resource x available
ti	=	the time at which node i is reached (i. e., all arcs incident to node i have been traversed)

We set t_o, the time at which the movement begins, equal to zero; and t_n, the time at which the sink node is reached, is equal to the closure time of the movement.

Since each node represents an event in time (the completion of all tasks leading into it), then all tasks incident to a node must be completed before any tasks emanating from a node can begin. Thus, the operating rules of the scheduling network can be stated as follows:

1. All tasks in the network must be performed.
2. A task can be started only after its preceding node has been reached.
3. A node is considered reached only after all of its incident tasks are completed.

If we denote the set of arcs in a network by A and the set of nodes by N, then the following constraints guarantee that the operating rules of the network are satisfied:

$$t_i + g_{ij}(x_{ij}) - t_j \leq 0, \ \forall i,j \in A \tag{2.3}$$

Equation (2.3) says that the time of reaching node i, t_i, plus the duration of activity (i, j), $g_{ij}(x_{ij})$, must be less than or equal to the time at which node j is reached, t_j. There is one of these constraints for each arc in the network.

Now that the time constraints of the optimization problem have been modeled, we must ensure that the resource limit will not be exceeded. We can write the following constraint for each node:

$$\sum_k x_{ik} - \sum_i x_{ij} \quad \begin{matrix} \leq 0, \ j \neq 0, n \\ \leq X, \ j = 0 \\ \geq X, \ j = n \end{matrix} \tag{2.4}$$

These constraints say that we require the total amount of resource allocated to arcs emanating from node j (j not the source or sink) to be less than or equal to the total amount of resource allocated to arcs incident to node j. At the source and sink we require that the total amount of resource out and in, respectively, be less than or equal to X. These constraints (2.4) guarantee that the sum of the amount of resource committed to individual tasks at any one time during the deployment will be less than or equal to the resource limit.

Groninger points out a more general form of the resource constraints.[8] He states that the constraint on resource availability might be written in the following way

$$\sum_{(i,j) \in T(t)} x_{ij} \leq X, \; x_{ij} \geq 0, \; 0 \leq t \leq t_n \qquad (2.4b)$$

where

$T(t)$ = the set of all tasks (i, j) active at time t

This expression says that the sum of the amount of resource committed to individual tasks at any time during a deployment must be less than or equal to the resource limit. Of course, determining the members of the set $T(t)$ at any time is a difficult problem, since the concurrent tasks can be known only after their x_{ij} are determined. In fact, Groninger shows that for a network of reasonable size the number of constraints required to define (2.4b) is unmanageable. Thus, the need for the alternative approach defined in (2.4). Notice that the formulation (2.4) is more restricting than (2.4b) and that the space of feasible solutions to (2.4) is entirely contained in the space of feasible solutions to (2.4b).

A result of the more restrictive resource constraint equation is that the resource utilization rate for each activity is a constant. In other words, the model will allocate a resource level of x units to an activity and the allocation will remain constant throughout the activity duration.

As an example of the constraint equations (2.3) and (2.4), let us consider the scheduling network shown in Figure 2.3, and in particular, node 4 and arc (3, 4). The time constraint equation for arc (3.4) is written as follows

$$t_3 + g_{34}(x_{34}) - t_4 \leq 0 \tag{2.5}$$

and the resource constraint equation for node 4 would be:

$$-x_{04} - x_{34} + x_{46} \leq 0 \tag{2.6}$$

The complete set of constraints for this problem would consist of one equation of type (2.5) for each of the ten arcs and one equation of type (2.6) for each of the seven nodes.

There is one special case that must be considered involving the resource constraint equations. Let us write the resource constraint for node 3 of Figure 2.3. Note that node 3 is preceded by a dummy activity (2, 3) that requires no resource. It represents a delay of one day required by the engineer battalion to ready the area to receive the infantry units. The resource constraint for this node would be

$$-x_{23} + x_{34} + x_{36} \leq 0 \tag{2.7}$$

Activity (2, 3) is a dummy and requires no resource (see Figure 2.4b). If no resource is used, x_{23} is equal to zero; and since we cannot assign negative resources to activities (3, 4) and (3, 6), equation (2.7) is clearly infeasible.

In order to gain a feasible formulation, it has been necessary to sacrifice some degree of optimality. The approach to feasibility is to not require the dummy resource allocation to be equal to zero. Instead, the dummy activity would be allocated its share of the resource from the resource constraint equation of the <u>preceding</u> node, and these resources would sit idle for the duration of the dummy activity. In this way, the dummy resource allocation is not zero. However, in order to gain feasibility, we have required the resources assigned to the dummy to remain idle for some period of time. As a result, the solution will be somewhat less than optimal. The following example from Figure 2.3 is intended to further clarify this procedure.

Let us consider node 1 first. The resource constraint is:

$$- x_{01} + x_{12} \leq 0 \qquad (2.8)$$

Activity (1, 2) is a dummy with duration of one day. Its duration function (Fig. 2.4 b) is constant regardless of the amount of resource allocated to it. Therefore, we can assign to the dummy activity any resource allocation, $0 \leq x_{12} \leq X$, and still maintain a one day duration. In the specific case of (2.8), $x_{12} = x_{01}$.

We then proceed to node 2. The resource constraint equation is:

$$- x_{02} - x_{12} + x_{23} + x_{25} \leq 0 \qquad (2.9)$$

Once again, activity (2, 3) is a dummy with a duration of one day and a duration function like Figure 2.4(b). Instead of assigning a value of zero to x_{23}, it is allowed to take on any positive value between zero and X. In this particular case, activity (2, 5) would be allocated first and activity (2, 3) would be allocated the remaining resources:

$$x_{23} = x_{02} + x_{12} - x_{25} \qquad (2.10)$$

The resources, x_{23}, would then sit idle for one day in order to be available for allocation at node 3. The resource constraint at node 3 is:

$$- x_{23} + x_{34} + x_{36} \leq 0 \qquad (2.11)$$

However, x_{23} is no longer zero and the solution, although not optimal, is feasible.

The consequences of this aspect of the linear programming formulation have not been investigated to date and are not within the scope of this paper. It should be noted that dummy activities are encountered in many routing and scheduling problems. Since there are only two dummies in the example network of Figure 2. 3 and each has a duration of one day, the maximum deviation from the optimal is only two days. In a practical sized problem, the number and duration of the dummy activities would probably be greater. As a result, there could be a meaningful departure from the optimal solution.

It is especially important to realize that this inconsistency is not due to the linear programming routine but to the particular form of the resource constraints. This shortcoming will have to be answered before the linear programming formulation will accurately model the routing and scheduling problem.

Now let us return to the optimization problem. Figure 2. 5 shows the problem formulation written out for the sample network of Figure 2. 3. Notice that the problem is linear except for the duration functions, $g_{ij}(x_{ij})$. Groninger's [9] approach to the solution of a problem of this type is to approximate the duration functions by a piece-wise linear function and use the technique of separable programming presented in Hadley.[10]

Note that the problem shown in Figure 2. 5 is of the following form:

$$\text{MIN } t_n$$

$$\sum_{j=1}^{n} g_{ij}(x_j) \leq b_i \qquad i = 1, \ldots, m \qquad (2.12)$$

$$\sum_{k} x_{ik} - \sum_{j} x_{ij} \begin{cases} \leq 0 & , j \neq 0, n \\ \leq X & , j = 0 \\ \geq X & , j = n \end{cases}$$

$$x_j \geq 0 \qquad j = 1, \ldots, n$$

Minimize t_m

Subject To:

$$
\begin{array}{llr}
g_{01}(x_{01}) - t_1 & = 0 & (1) \\
g_{02}(x_{02}) - t_2 & \leq 0 & (2) \\
g_{04}(x_{04}) - t_4 & \leq 0 & (3) \\
g_{12}(x_{12}) + t_1 - t_2 & \leq 0 & (4) \\
g_{23}(x_{23}) + t_2 - t_3 & = 0 & (5) \\
g_{25}(x_{25}) + t_2 - t_5 & = 0 & (6) \\
g_{34}(x_{34}) + t_3 - t_4 & \leq 0 & (7) \\
g_{36}(x_{36}) + t_3 - t_6 & \leq 0 & (8) \\
g_{46}(x_{46}) + t_4 - t_6 & \leq 0 & (9) \\
g_{56}(x_{56}) + t_5 - t_6 & \leq 0 & (10)
\end{array}
$$

$$
\begin{array}{llr}
X_{01} + X_{02} + X_{04} & \leq X & (11) \\
-X_{01} + X_{12} & \leq 0 & (12) \\
-X_{02} - X_{12} + X_{23} + X_{25} & \leq 0 & (13) \\
-X_{23} + X_{34} + X_{36} & \leq 0 & (14) \\
-X_{04} - X_{34} + X_{46} & \leq 0 & (15) \\
-X_{25} + X_{56} & \leq 0 & (16) \\
-X_{36} - X_{46} - X_{56} & \leq X & (17)
\end{array}
$$

Figure 2.5. Example Problem Formulation

where the duration functions, $g_{ij}(x_j)$, are non-linear. The approximation technique replaces the $g_{ij}(x_j)$ functions by polygonal approximations, thereby reducing the problem to a form which can be solved by the simplex method.

The approximation functions are obtained in the following way (see Figure 2.6). Select $r+1$ points, x_k (they need not be equally spaced). Then compute the value of the ordinate at these points, $g_k = g(x_k)$, and connect (x_k, g_k) and (x_{k+1}, g_{k+1}) by a straight line. The dashed curve or polygonal approximation to $g(x)$ will be denoted by $\hat{g}(x)$. Notice that $\hat{g}(x)$ can be made arbitrarily good by selecting the x_k properly and subdividing the intervals.

We can now replace the original problem (2.12) by

$$\text{MIN } t_n$$

$$\sum_{j=1}^{n} \hat{g}_{ij}(x_j) \leq b_i \qquad i = 1, \ldots, m \tag{2.13}$$

$$\sum_k x_{ik} - \sum x_{ij} \begin{cases} \leq 0, & j \neq 0, n \\ \leq X, & j = 0 \\ \geq X, & j = n \end{cases}$$

$$x_j \geq 0, \quad j = 1, \ldots, n$$

which we shall call the approximation problem for (2.12).

Now that we have constructed the necessary approximation, we must show how to express $\hat{g}_{ij}(x_j)$ analytically before we can determine an optimal solution of the approximating problem.

Let us return to Figure 2.6. When x lies in the in the interval $x_k \leq x \leq x_{k+1}$, we approximate $g(x)$ by $\hat{g}(x)$, where

$$\hat{g}(x) = g_k + \frac{g_{k+1} - g_k}{x_{k+1} - x_k}(x - x_k) \tag{2.14}$$

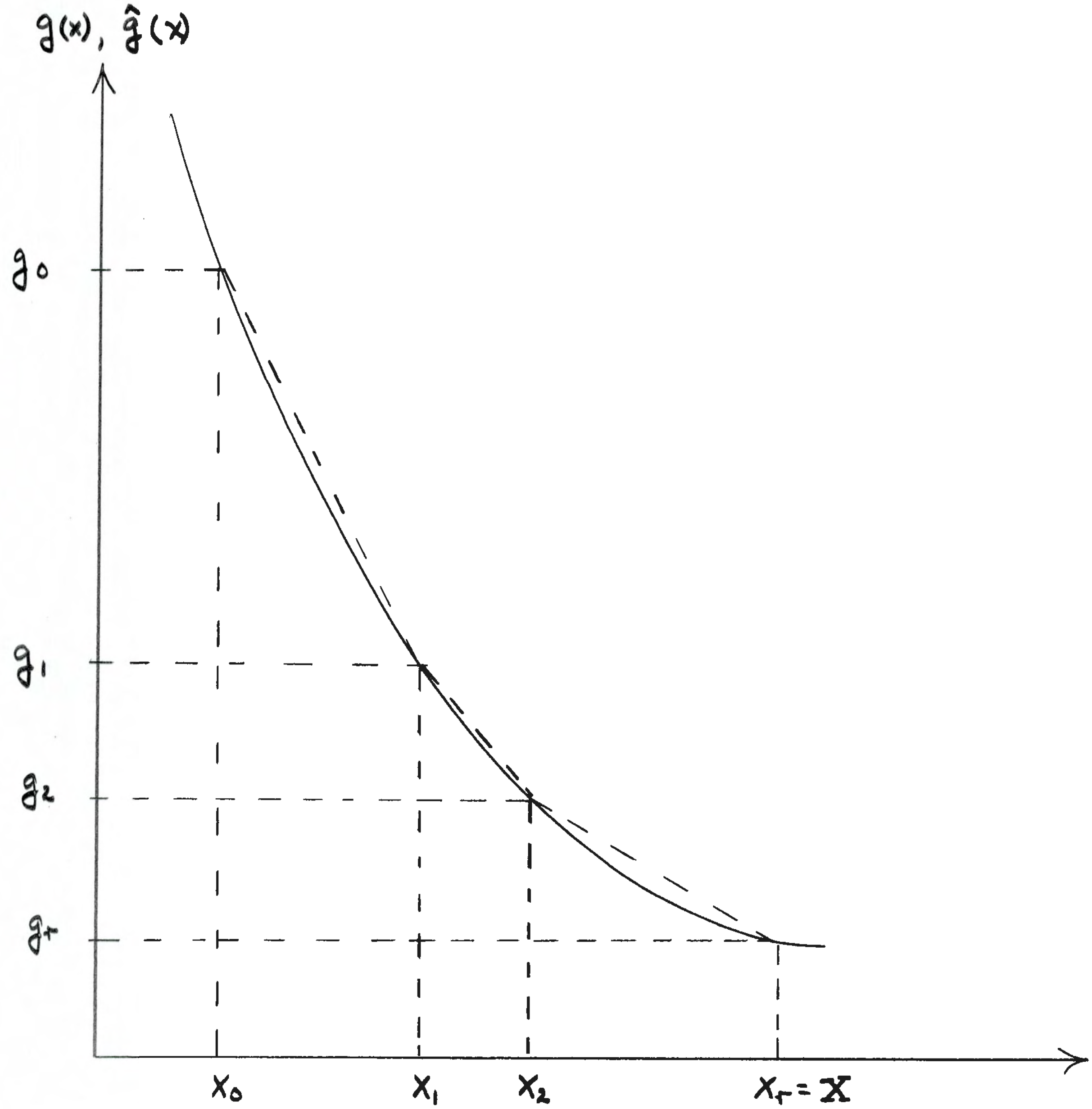


FIGURE 2.6. LINEAR APPROXIMATION OF A DURATION FUNCTION

Note that any x in the interval $x_k \leq x \leq x_{k+1}$ can be written $x = \lambda x_{k+1} + (1 - \lambda) x_k$ for some λ, $0 \leq \lambda \leq 1$. Then, $(x - x_k) = \lambda (x_{k+1} - x_k)$, so that (2.14) can be written $\hat{g}(x) = \lambda g_{k+1} + (1 - \lambda) g_k$. If we now write $\lambda = \lambda_{k+1}$, $(1 - \lambda) = \lambda_k$, it follows that when $x_k \leq x \leq x_{k+1}$, there exists a unique λ_k and λ_{k+1} such that

$$x = \lambda_k x_k + \lambda_{k+1} x_{k+1} \tag{2.15}$$

$$\hat{g}(x) = \lambda_k g_k + \lambda_{k+1} g_{k+1} \tag{2.16}$$

$$\lambda_k + \lambda_{k+1} = 1, \ \lambda_k, \lambda_{k+1} \geq 0. \tag{2.17}$$

For any x, $0 \leq x \leq X$, we can write

$$x = \sum_{k=0}^{r} \lambda_k x_k \tag{2.18}$$

$$\hat{g}(x) = \sum_{k=0}^{r} \lambda_k g_k \tag{2.19}$$

$$\sum_{k=0}^{r} \lambda_k = 1, \quad \lambda_k \geq 0, \quad k = 0, \ldots, r \tag{2.20}$$

provided that we require that no more than two of the λ_k shall be positive, and if two (say λ_s, λ_k) are positive, it must be true that $k = s + 1$, that is, they are adjacent λ's. With this description, the λ_k are uniquely determined, and for an x given by (2.18), $\hat{g}(x)$ determined by (2.19) will be on the dashed curve of Figure 2.6 and will be an analytic representation of $\hat{g}(x)$. By allowing no more than two of the λ_k to be positive, and then requiring that they be adjacent, we ensure that the points of (2.18) and (2.19) will be on the dashed curve.

We have now approximated the duration functions and obtained the mathematical representations (2.18) through (2.20) for the line. With this information we return to the approximating problem (2.13). Assume that the maximum value which the variable x_j can take on is α_j (this is our maximum airlift capability per day). We then subdivide the interval $0 \leq x_j \leq \alpha_j$ into r_j subintervals by the $r_j + 1$ points x_{kj}, $x_{oj} = 1$ and $x_{rj} = \alpha j$. Then, for all the functions $\hat{g}_{ij}(x_j)$ we can write

$$\hat{g}_{ij}(x_j) = \sum_{k=0}^{r_j} \lambda_{kj}\, g_{kij}\,, \quad g_{kij} = g_{ij}(x_k) \qquad i = 1, \ldots, m \tag{2.21}$$

where

$$x_j = \sum_{k=0}^{rj} \lambda_{kj}\, x_{kj} \tag{2.22}$$

$$\sum_{k=0}^{r_j} \lambda_{kj} = 1, \quad \lambda_{kj} \geq 0 \quad \text{all } k, j \tag{2.23}$$

and for a given j, no more than two λ_{kj} are allowed to be positive and these must be adjacent.

We can now use (2.21) to eliminate the function $\hat{g}_{ij}(x_j)$ in (2.13), to yield the following representation of the approximating problem in terms of the variables λ_{kj} rather than the variables k_j:

$$\sum_{j=1}^{n} \sum_{k=0}^{r_j} g_{kij}\, \lambda_{kj} \leq b i, \quad i = 1, \ldots, m$$

$$\sum_{k} x_{ik} - \sum x_{ij} \begin{cases} \leq 0, & j \neq 0, n \\ \leq X, & j = 0 \\ \geq X, & j = n \end{cases}$$

$$\sum_{k=0}^{r_j} \lambda_{k_j} = 1, \quad j = 1, \ldots, n \tag{2.24}$$

$$\lambda_{kj} \geq 0 \quad \text{all } k, j$$

$$\text{MIN } t_n$$

It should be noted that some of the variables x_j enter into the problem linearly. Let us consider the resource variable x_{01} in Figure 2. 5. Notice that it enters the first constraint equation non-linearly as $g_{01}(x_{01})$ and linearly in the eleventh and twelfth equations as x_{01} and $-x_{01}$ respectively. It is important to observe that if a particular variable x_j enters into the problem only linearly then it is unnecessary to write x_j in terms of the λ_{kj}. We simply use x_j as the variable. This observation is particularly important when we consider that the fairly small problem (2. 12) has given rise to the much larger approximating problem (2. 24). Since it will often be true that a sizable number of variables will enter the problem linearly, this observation may make it possible to reduce the size of the approximating problem.[11]

Figure 2. 7 shows the approximation problem (2. 24) written out for the case where each $g_{ij}(x_{ij})$ is approximated by three linear segments. This problem would be linear if we did not require that for each arc (i, j) no more than two λ_{ijk} be positive and then only if they are adjacent.

In this problem the duration functions are assumed hyperbolic and take the form

$$g(x) = \frac{T\,d}{x\,p} \tag{2. 25}$$

This is the same function defined in (2. 2). The function is approximated by three linear segments as shown in Figure 2. 6. The x_k (1, 1. 82, 4. 299, 20) have been chosen so as to minimize the vertical error between $g(x)$ and $\hat{g}(x)$ for each line segment.[12] Note that the limiting number of resource units (aircraft) for the example has been set at twenty. This figure (α_j) represents the deployment airlift capability per day.

Now that the duration function of an activity has been defined, it is necessary to clarify its use in the model. A duration function

λ_{010}	λ_{011}	λ_{012}	λ_{013}	λ_{020}	λ_{021}	λ_{022}	λ_{023}	λ_{040}	λ_{041}	λ_{042}	λ_{043}	λ_{120}	λ_{121}	λ_{230}	λ_{231}	λ_{250}	λ_{251}	λ_{252}	λ_{253}	λ_{340}	λ_{341}	λ_{342}	λ_{343}	λ_{360}	λ_{361}	λ_{362}	λ_{363}	λ_{460}	λ_{461}	λ_{462}	λ_{463}	λ_{560}	λ_{561}	λ_{562}	λ_{563}	t_1	t_2	t_3	t_4	t_5	t_6	
g_{01}	g_{01}	g_{01}	g_{01}																																	-1						= 0
				g_{02}	g_{02}	g_{02}	g_{02}																														-1					≤ 0
								g_{04}	g_{04}	g_{04}	g_{04}																												-1			≤ 0
												g_{12}	g_{12}																							+1	-1					≤ 0
														g_{23}	g_{23}																						+1	-1				= 0
																g_{25}	g_{25}	g_{25}	g_{25}																		+1			-1		= 0
																				g_{34}	g_{34}	g_{34}	g_{34}															+1	-1			≤ 0
																								g_{36}	g_{36}	g_{36}	g_{36}											+1			-1	≤ 0
																												g_{46}	g_{46}	g_{46}	g_{46}								+1		-1	≤ 0
																																g_{56}	g_{56}	g_{56}	g_{56}					+1	-1	≤ 0
X_{01}	X_{01}	X_{01}	X_{01}	X_{02}	X_{02}	X_{02}	X_{02}	X_{04}	X_{04}	X_{04}	X_{04}																															≤ X
$-X_{01}$	$-X_{01}$	$-X_{01}$	$-X_{01}$									X_{12}	X_{12}																													≤ 0
				$-X_{02}$	$-X_{02}$	$-X_{02}$	$-X_{02}$					$-X_{12}$	$-X_{12}$	X_{23}	X_{23}	X_{25}	X_{25}	X_{25}	X_{25}																							≤ 0
														$-X_{23}$	$-X_{23}$					X_{34}	X_{34}	X_{34}	X_{34}	X_{36}	X_{36}	X_{36}	X_{36}															≤ 0
								$-X_{04}$	$-X_{04}$	$-X_{04}$	$-X_{04}$									$-X_{34}$	$-X_{34}$	$-X_{34}$	$-X_{34}$					X_{46}	X_{46}	X_{46}	X_{46}											≤ 0
																$-X_{25}$	$-X_{25}$	$-X_{25}$	$-X_{25}$													X_{56}	X_{56}	X_{56}	X_{56}							≤ 0
																								$-X_{36}$	$-X_{36}$	$-X_{36}$	$-X_{36}$	$-X_{46}$	$-X_{46}$	$-X_{46}$	$-X_{46}$	$-X_{56}$	$-X_{56}$	$-X_{56}$	$-X_{56}$							≤ X
1	1	1	1																																							= 1
				1	1	1	1																																			= 1
								1	1	1	1																															= 1
												1	1																													= 1
														1	1																											= 1
																1	1	1	1																							= 1
																				1	1	1	1																			= 1
																								1	1	1	1															= 1
																												1	1	1	1											= 1
																																1	1	1	1							= 1

FIGURE 2.7. CONSTRAINT TABLEAU FOR APPROXIMATING PROBLEM

represents a family of curves:

$$g(x) = \frac{T\,d}{x\,p} = \frac{C}{x} \tag{2.26}$$

where, for a particular unit of the deployment and considering a single resource of one type of aircraft,

$$\frac{T\,d}{p} = \text{constant} = C \tag{2.27}$$

Thus, the family of curves is hyperbolic moving up with increasing C values (Figure 2. 8).

The method of applying the duration function is to define the C values as a measure of the work that must be done to transport a particular shipment to its destination. Since both the aircraft productivity, p, and the deployment distance, d, remain constant in this simple problem, the measurable variable from shipment to shipment is the weight of the unit expressed in short tons. In other words, (2. 27) can be written as:

$$T \sim C\ ;\ p\,,\ d \text{ constant} \tag{2.28}$$

In order to be consistent, unit total weight was taken to be that listed in STRATMAS TM 43.[13] These weights are shown in Figure 2. 9. As a relative measure, the infantry battalion of 550 short tons was arbitrarily assigned a C value of 10 and each unit scaled accordingly. Thus, for the armored battalion of 3800 short tons:

$$\frac{3800 \text{ ST}}{550 \text{ ST}} \cong 7 \tag{2.29}$$

and

$$7\,(10) = 70 = C \text{ value} \tag{2.30}$$

The complete list of C values is shown with the unit weights in Figure 2. 9.

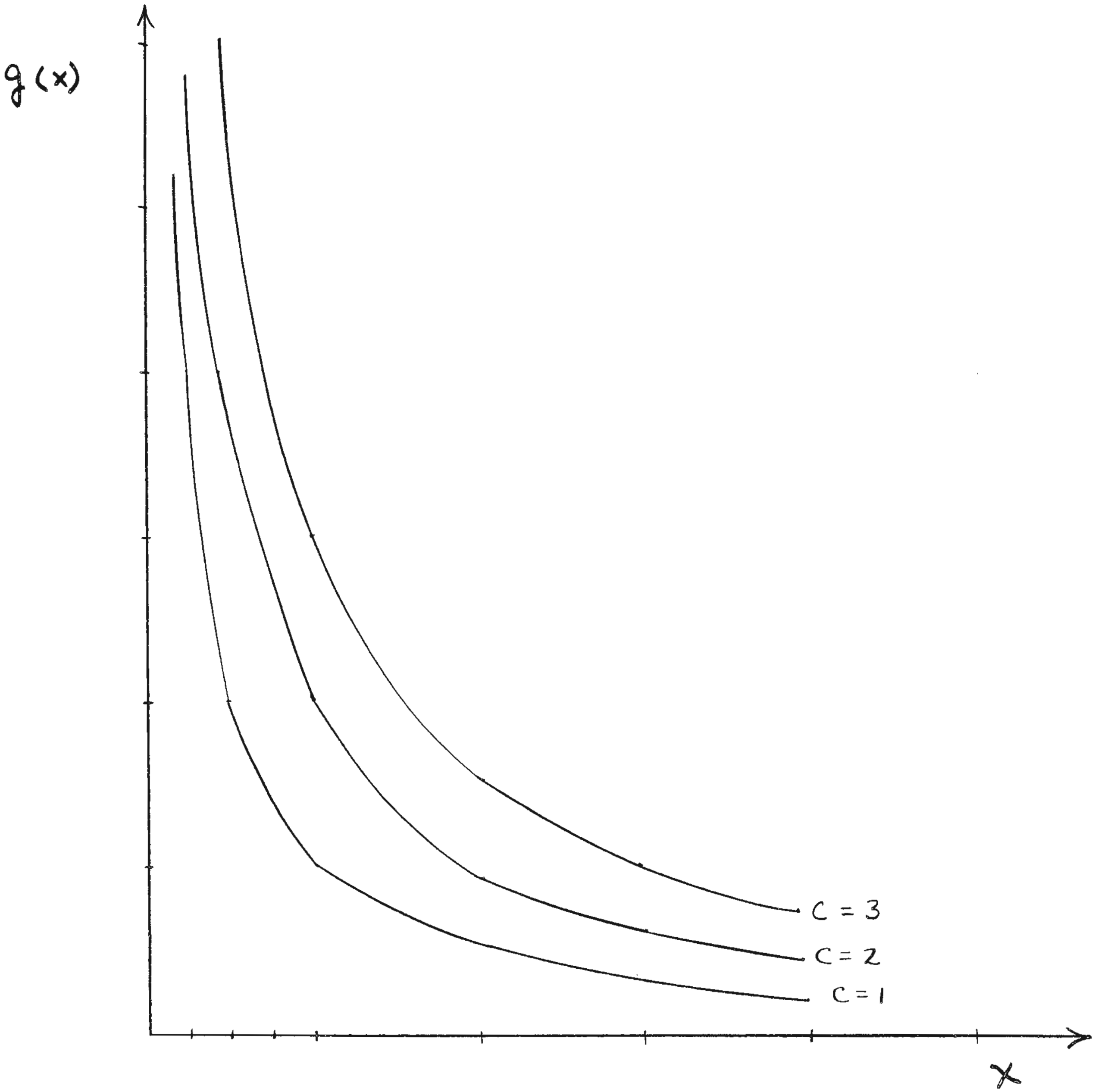


FIGURE 2.8. SET OF DURATION CURVES

Shipment	Unit	Number Units	Weight (Short Tons)	C Value
A Air Borne Infantry	Batt	2	900	20
B Engineers	Batt	1	400	8
C Terminal Service	Batt	1	200	4
D Support	Batt	2	1000	20
E Infantry	Batt	2	1000	20
F Armored	Batt	1	3800	70
G Mechanized Infantry	Batt	2	2200	40
H Fire Support	Batt	2	1800	36

(From STRATMAS TM 43)

Figure 2.9. Unit Weights And C Values

In general, the separable programming technique would lead us to obtain a local minimum of the approximating problem by applying the simplex method in the usual manner except that we would restrict entry into the basis in such a way that we would never allow more than two λ_{ijk} to be positive for a given (i, j). Furthermore, two λ_{ijk} could be positive only if they were adjacent. However, the constraints of Figure 2.7 have the convenient form that the set of feasible solutions which they define is convex. This would not necessarily be so if the duration functions had not been convex. The objective function, Min t_n, is also convex. For this special situation, it can be shown that the straight simplex technique yields the global optimum.[14] It should be noted that, in general, we are not guaranteed that the solution to the approximating problem will be a feasible solution to the original problem. However, since

$$g_{ij}(x) \leq \hat{g}_{ij} \quad \forall x, ij \tag{2.31}$$

by observing the constraint equations we can see that the solution to the approximating problem (2.24) $[\ldots, x^*_{ij}, \ldots, t^*_1, \ldots t^*_n]$, also satisfies the original non-linear constraint equations and is thus a feasible solution.[15]

It was stated above that the formulation of an approximating problem frequently gives rise to a very large linear programming problem. Figure 2.10 shows the number of constraints and variables of the linear programming problem as a function of network size and number of piece-wise linear sections of the duration function.[16]

As an example of the growth in problem size, consider the example problem of Figure 2.3 written out in the general form of (2.12) (Figure 2.5 shows the problem formulation in detail.). Notice in Figure 2.5 that the original problem has 17 rows, 16 columns, and 47 matrix elements. However, the approximating problem written out according

N = Number Of Nodes   A = Number Of Arcs

P = Number Of Piecewise Linear Sections Of Each Duration Function

	Number Of Constraints	Number Of Variables
The Original Nonlinear Programming Problem	1 Per Arc 1 Per Node $m = N + A$	1 Per Arc, x_{ij} 1 Per Node, t_i $n = N - 1 + A$
The Approximation Problem	$m_\lambda = N + 2A$	$n_\lambda = N - 1 + A(1-P)$

Examples

N	A	m_λ	n_λ P=3	n_λ P=5
4	5	14	23	33
11	16	43	74	106
100	150	400	699	999

Figure 2.10. The Relationships Between Network Size And Problem Size

to (2.24) and shown in Figure 2.7 has 28 rows, 56 columns, and 176 matrix entries.

A problem the size of the approximating problem can be solved by the linear programming routine used in this report in under two minutes. One of realistic size, say 200 rows, 300 columns, and more than 1000 entries, can be solved in five minutes.

The general form of the tableau presented in Figure 2.7 is shown in Figure 2.11. The activities consist of two major groups, the approximating variables and the node time variables. There is one time variable column for each node in the network, and the number of λ variable columns will be less than or equal to $(p+1)N$, where p equals the number of piece-wise linear sections and N is the number of arcs in the network. The number of λ variable columns can be less than $(p+1)N$ because it is convenient to represent dummy arcs by only one piece-wise section (see λ_{12} columns in Figure 2.7).

The constraints are divided into three major sections. A time constraint is written for each arc in the network and ensures that the operating rules of the network are followed. The resource constraints ensure that the sum of the resources utilized at any particular time does not exceed the maximum availability. There is generally one of these equations for each node. The third section of the constraints ensures that the sum of the λ variables is equal to 1 (see equation 2.24).

The right-hand side values are divided into the same three sections as the constraints. The right-hand side values opposite the time constraints are expressed in units of time (days in Figure 2.7). The values opposite the resource constraint equations are numbers representing resource availability (aircraft in the example). The bottom section of the right hand side values is composed of pure numbers.

Activities / Constraints	λ Variables	Time Variables	
Time Constraints	g_{kij} (Time)	t_λ (Time)	Time
Resource Constraints	Numbers		Resource Units
$\Sigma \lambda_s' = 1$	Numbers		1 1 . . . 1 . . .
Obj		t_m	Min

Figure 2.11. General Form Of The L P Tableau

The units of the various submatrices are as follows. The submatrix (λ variables, time constraints) contains elements, g_{kij} (k = 0, r_j where r_j is the number of piece-wise linear sections; i = 1, m where m equals the number of network arcs; j = 1, (p+1)N). The dimension of these elements is in units of time. This dimension satisfies the approximating problem constraint equation (2. 24)

$$\sum_{j=1}^{n} \sum_{k=0}^{r_j} g_{kij} \lambda_{kj} \leq b_i \qquad (2.32)$$

where λ_{kj} is a dimensionless number defined by (2. 22)

$$x_j = \sum_{k=0}^{r_j} \lambda_{kj} x_{kj} \qquad (2.33)$$

The dimension of the elements of the submatrix (time activities, time constraints) is also time. The elements of the other submatrices are pure numbers.

Notice that many of the elements of the tableau are zero values. This suggests the possibility of a computer program to interpret an input network and fill the tableau automatically. Since most of this work is routine, such a capability would be a great time saver. Also, we have suggested that the more piece-wise approximations are made to the duration functions, the more accurate the results will be. Thus, a loading capability that would allow the planner to vary the number of approximations without the burden of having to manually fill the increasing tableau would be of great value.

Granted, such a capability would require additional programs and added computer execution time. However, the input phase of a linear programming problem is very time-consuming. If this model were used extensively, such a capability would be an asset.

Now that we have developed the linear programming model in detail, we are in a position to utilize the model for analysis. The example problem of Figures 2.5 and 2.7 has been solved as a linear programming problem with the objective function of minimizing deployment closure time. This problem and the problems of the following section were run on the IBM 360/40 of the Civil Engineering Systems Laboratory at M. I. T. They were run under the control of the Integrated Civil Engineering System (ICES)[17], using the linear programming capabilities of the Optimization Techniques (OPTECH)[18] subsystem.

The input format for the example problem (Figure 2.5 with the C values of Figure 2.9) is shown in Appendix A. The results of the example run have been condensed and are shown in the network of Figure 2.12. Note that the number in the circle associated with each arc is the amount of resource assigned to that activity to achieve minimum closure time. The number in the rectangle associated with each node is the time at which the node is reached.

The results show that the minimum closure time for the example problem is 16.17 days. The objective of the next section will be to consider the example deployment in the light of the goal fabric considerations of Chapter I in order to clarify the relationships existing among the variables of the goal vector. The technique to be employed in the next section will be similar to a post-optimality analysis. The original network (Figure 2.12) will be evaluated in terms of vulnerability, flexibility, and unit effectiveness and network adjustments made variously to model these considerations. The modified problem will then be run under the model described in this section and the optimal network assignments compared.

Before preceding to the next section, it should be pointed out that there is an inconsistency in the results of Figure 2.12. Consider

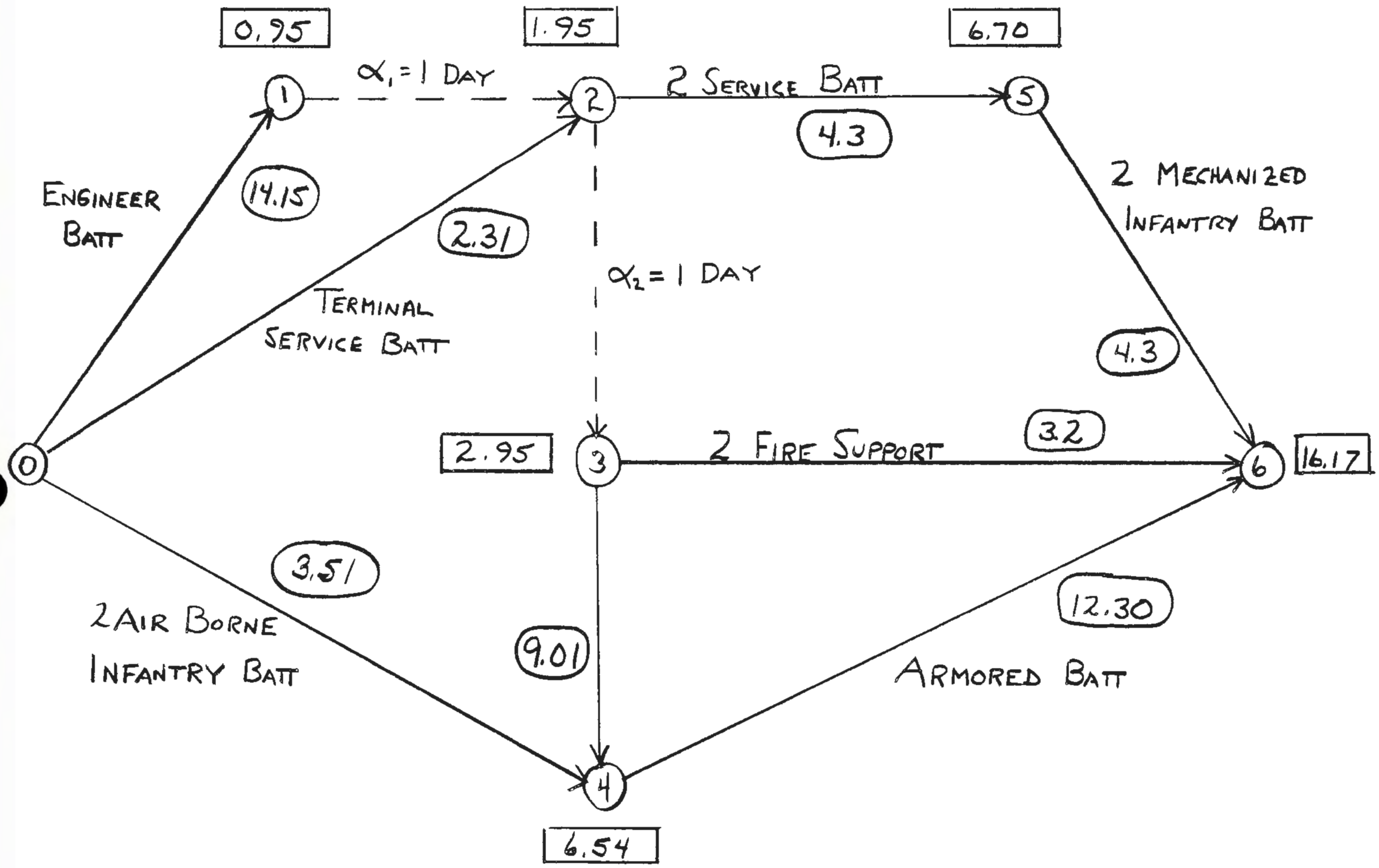


FIGURE 2.12. PROBLEM 1 SCHEDULING NETWORK

activity (0, 1). The results show that $\lambda_2 = .37$ and $\lambda_3 = .63$ (see Appendix A) so that x, the amount of resource allocated to (0, 1) is, from (2.15)

$$\begin{aligned} x &= \lambda_k x_k + \lambda_{k+1} x_{k+1} \\ &= .37(4.29) + .63(20) \\ &= 14.15 \text{ resource units} \end{aligned} \tag{2.34}$$

Then, from the duration function

$$g(x) = \frac{C}{x} \tag{2.35}$$

where C equals 8,

$$\begin{aligned} g(x) &= \frac{8}{14.15} \\ &= .566 \text{ days} \end{aligned} \tag{2.36}$$

However, TM01 (the time at which node 1 is reached) is .95 days in the optimal solution shown in Figure 2.12. The reason for the discrepancy is that the duration function used in the problem solution is an approximation defined by (2.16)

$$\hat{g}(x) = \lambda_k \cdot g_k + \lambda_{k+1} g_{k+1} \tag{2.37}$$

If (2.37) is used to compute TM01, the results are as follows (note that g_k in this case equals 8/4.29 or 1.86 and g_{k+1} equals 8/20 or 0.4)

$$\begin{aligned} g(x) &= .37(1.86) + .63(0.4) \\ &= 0.945 \text{ days} \end{aligned} \tag{2.38}$$

Thus, the closure times shown in the problem results of Figure 2.12 are based upon the approximating duration function and will always be greater than or equal to the actual values computed from (2.34) and (2.35). Of course, these results could be made more nearly equal throughout by subdividing the approximating intervals.

Throughout the following sections results will consistently be shown as those to the approximating problem.

II. 4 Post-optimality Considerations

II. 4. 1 Network Adjustments

Before we can implement any of the ideas concerning the various members of the goal vector developed in Section I. 4, we must first decide how these ideas can be physically modeled in the scheduling network.

Consider first the goal of vulnerability. This goal reflects the delays a deployment can encounter because of congestion, enemy action, sabotage, accident, or weather. It was pointed out in Section I. 4 that this goal is maximized when units comprising a single phase of the CINC's capability are split and scheduled over different routes. Using this scheduling approach, if one of the units is delayed or lost because of terminal congestion or enemy action, the CINC can still expect to receive some measure of the particular capability.

We can use the following techniques to model these considerations. First, the activity in question can be split. That is, instead of routing two infantry battalions as a unit, each could be routed separately. The network modification necessary would be to add one arc and possibly one node. Another approach to optimizing vulnerability is to re-define particular activities. For example, if one activity consists of two infantry battalions and another consists of two mechanized infantry battalions, then vulnerability considerations would be improved if these activities were re-defined so that one infantry and one mechanized infantry battalion comprised each activity. Of course, the only network modification necessary to model this approach is to re-define the new activity duration functions.

The next goal to consider is that of flexibility. This goal is intended to model the uncertainty surrounding any military operation due to limited intelligence concerning the enemy's capabilities. Flexibility of a deployment is optimized when the CINC or the movement planner maintains the ability to reschedule the movements of any particular shipment at any time during the planning process. Of course, the CINC can always demand a particular unit at any time --- and, if physically possible, it will be delivered. The important point is that the CINC and the planner must know how much it costs to modify the schedule; they must be able to determine what a particular degree of flexibility is worth in terms of the other goals.

There are two ways of modifying the scheduling network to model the flexibility problem. The use of dummy activities enables the planner to establish certain priorities among the units of a deploying force. For instance, in Figure 2.13 the mechanized infantry (5, 6) might arrive at a late date because it is competing for resources with the armored unit (7, 8). However, the addition of the dummy arc (6, 7) guarantees that (5, 6) must be completed before (7, 8) commences (Figure 2.14).

The second method of modeling flexibility is to maintain a capability to adjust the resource allocation of each activity. Thus, if a unit is originally allowed a long period of time in transit at a low resource level because its priority is considered small, the planner could adapt to a change in the military situation by directly assigning the activity in question to operate at a higher resource level, and, as a result of the shape of the duration curve,

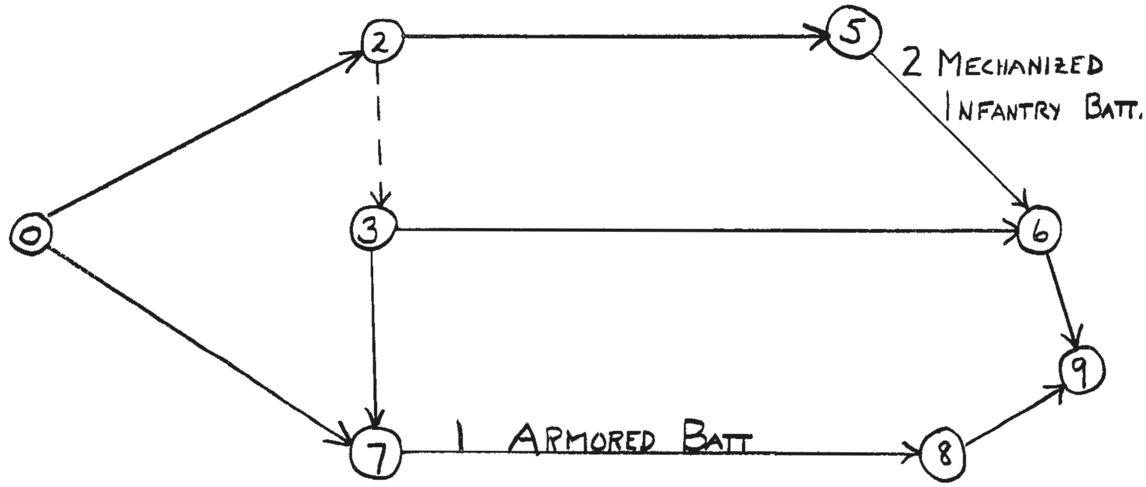


Figure 2.13. Example Network

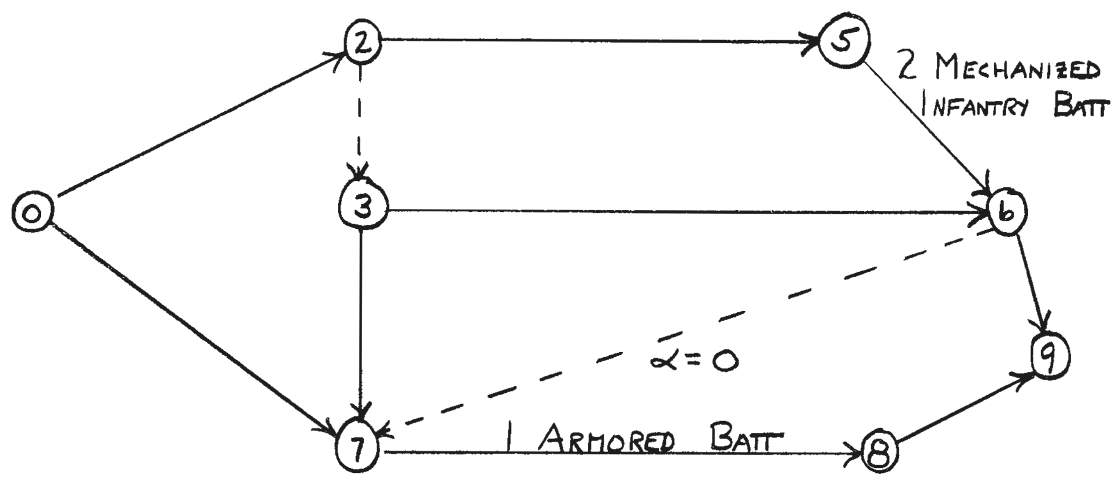


Figure 2.14. Establishing Unit Precedence

a lower duration time. This capability can be achieved by adding an extra constraint row to the linear programming problem in which the level of the resource is fixed (this procedure is applied in problem 4, below).

The third goal to consider is that of unit effectiveness. In order to maximize unit effectiveness, the planner should attempt to schedule a unit in as short a time as possible while utilizing transportation resources that maintain as nearly as possible the unit command structure.

The approach to modeling this goal is twofold. As with the flexibility considerations, a higher resource level can be assigned to each shipment by introducing additional constraint equations. In this way, the planner can assure that the shipment will be in transit a specified number of days.

Another approach is to introduce dummy activities with constant duration functions equal to zero. These activities will serve the purpose of diverting additional resources to the activity in question. For example, in Figure 2.15, the dummy arc (4. 5) is supplying additional resources to activity (5, 6) (see Problem 3, below).

Before proceeding to the application of these modeling techniques, a few comments are in order concerning the scheduling network concept. As the results of this chapter have shown, the scheduling network is a powerful tool in the analysis of the routing and scheduling problem. The network structure is the basis of the linear programming model formulated by Groninger. Also, we have seen that simple network modifications give the planner the capability to model the goal vector variables.

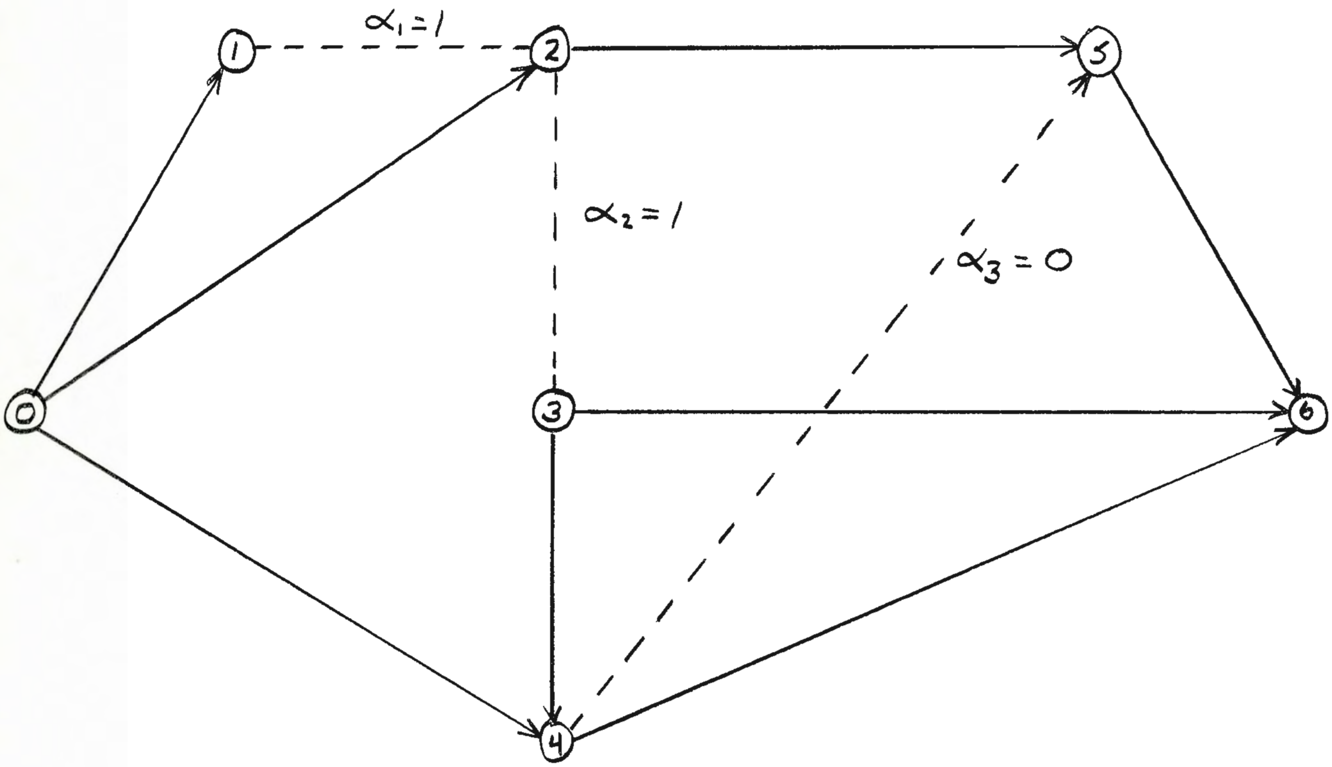


FIGURE 2.15. USE OF DUMMY ACTIVITIES

However, one must always bear in mind the fact that this is an elementary example of the single resource problem. Considerations of two or more resources greatly complicate the problem formulation and no solution method has been developed to date.[19]

Another problem area that must not be overlooked is that this formulation does not model the physical network. Great energy must be expended by the planners to detail the movement of each shipment from origin to destination (see Fig. 1.1, Chapter I). Then, the routing of the many shipments must be coordinated so as to minimize terminal and link congestion. Thus, the use of a model similar to this prototype will lead to the theoretical minimum closure time; but the post-optimality analysis that we have described above and will apply below, requires a great amount of detailed considerations. This point should be further clarified.

In order to utilize the linear programming model, the planner needs only the partially ordered sequence of arrivals from the CINC and the duration function for each activity. After solving the problem with this formulation, the planner has succeeded in optimizing deployment closure time. The next logical step would be to construct the detailed movement plan (Fig. 1.1) for each shipment utilizing the resource levels allocated and the time data from the optimal solution. (This information is shown in Figure 2.12 for the example problem.)

It is at this point that the planner begins to consider the physical constraints of the transportation network. After routing and scheduling each unit, he must coordinate the flows over links and through terminals. It is only after he has reached this stage of the analysis that he can determine what modifications are necessary to the network to reduce vulnerability due to congestion or accident. Then, after the adjustments have been made and a new optimal solution obtained, the entire procedure must be repeated.

The important point of this argument is that the linear programming model is not an end in itself. It is a single step in the analysis procedure, and the other elements of the procedure will probably demand as much, if not more, time and energy to accomplish.

II. 4. 2 Analysis of Adjustments

We begin this section by considering the results of the example problem in Section II. 3. This network and the results are shown in Figure 2. 12. Note that there are two air-borne (0, 4), two infantry (3, 4), two fire support (3, 6), and two mechanized infantry (5, 6) battalions scheduled to be shipped as single units. This violates the vulnerability criteria and suggests that these activities could be re-defined. Let us choose activities (3, 4), and (5, 6) and re-define them as shipments composed of one infantry battalion and one mechanized infantry battalion. This results in the scheduling network shown in Figure 2. 16. We will call this network Problem 2.

After re-defining the duration functions for these activities (the C value increases from 20 to 30 for (3, 4) and decreases from 40 to 30 for (5, 6)), the problem was solved with the results shown in Table 2. 1. The results of the original example (Problem 1) are also shown in the Table.

The new problem has decreased the vulnerability of the deployment by splitting the two activities and increasing the chances that the CINC will receive at least one-half of the particular capabilities. Now we must consider what the decreased vulnerability has cost. It is immediately apparent that the deploy-

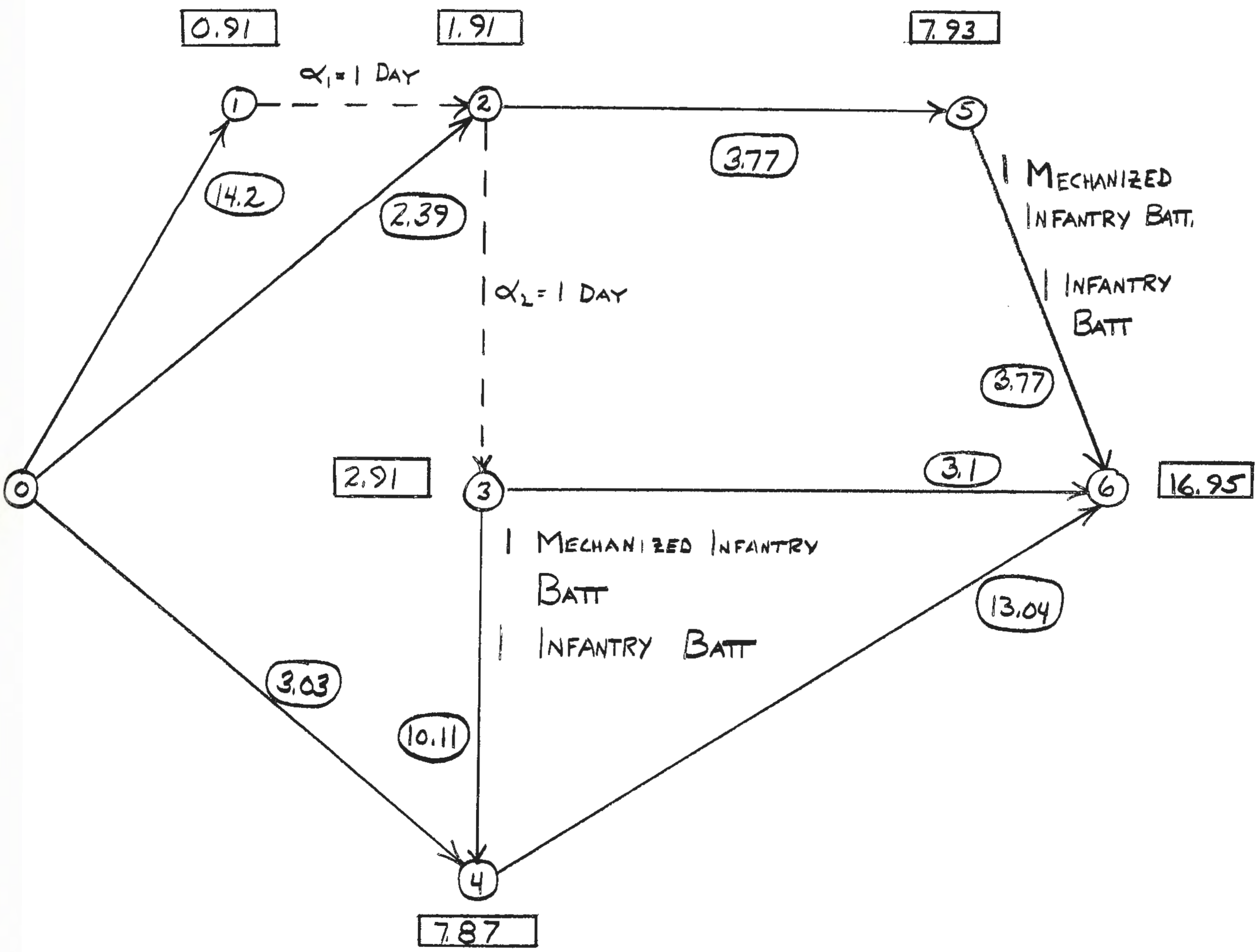


FIGURE 2.16. PROBLEM 2 SCHEDULING NETWORK

TABLE 2. 1

	PROBLEM 1		PROBLEM 2	
	Allocation (A/C)	Duration (Days)	Allocation (A/C)	Duration (Days)
0,1	14.15	.95	14.15	.91
0,2	2.31	1.94	2.39	1.90
0,4	3.51	6.54	3.03	7.9
2,5	4.3	4.75	3.77	6.01
3,4	9.01	3.62	10.11	4.96
3,6	3.2	13.026	3.1	13.14
4,6	12.3	9.64	13.04	9.13
5,6	4.3	9.42	3.77	9.00
CLTM	16.17		16.95	

	PROBLEM 3		PROBLEM 4	
0,1	14.2	.91	8.38	1.49
0,2	2.39	1.90	1.68	2.50
0,4	3.03	7.90	10.0	2.0
2,5	3.43	6.99	3.41	6.96
3,4	10.42	4.86	4.05	7.95
3,6	3.1	13.14	2.56	15.32
4,6	11.9	9.29	14.03	8.35
5,6	4.98	8.23	3.41	10.42
CLTM	17.00		19.85	

PROBLEM RESULTS

ment closure time has been increased (from 16. 17 to 16. 95 days). Also, the increased C value of activity (3, 4) of Figure 2. 16 has caused more resources to be allocated to the path 0-2-3-4 with the result that the engineer and terminal service battalions arrive in slightly less time (. 91 vs. . 94 and 1. 90 vs. 1. 94 days). However, the arrival date of the critical air-borne units has been increased from 6. 54 to 6. 90 days.

Also, since the C value of the activity (5, 6) was reduced, less resource is allocated to path 0-2-5-6. The result is that the activity (5, 6) spends 9. 00 days in transit. Table 2. 1 shows that the activity (5, 6) spends 9. 42 days in transit in Problem 1. Since these activities represent units that are vital to the military effort, their transit times should be reduced.

The unit effectiveness of activity (5, 6) can be increased by decreasing the time the unit spends in transit. As shown above, we can decrease the time a unit spends in transit by increasing its resource allocation. In Figure 2. 17, the dummy arc (4, 5) serves the purpose of diverting needed resources to activity (5, 6). Figure 2. 17 represents Problem 3.

Problem 3 has been solved for minimum closure time and the results shown in Table. 2. 1. Note that the duration time of activity (5, 6) has indeed been decreased to 8. 23 days. The unit's closure date has remained virtually unchanged at 17. 00 (up . 05 from 16. 95) if compared with Problem 2, but up . 83 days from the 16. 17 of Problem 1. Thus, we can conclude that it cost . 78 days (16. 95 - 16. 17) to decrease unit vulnerability, but the marginal price of increasing the unit effectiveness of activity (5, 6) is only . 05 days.

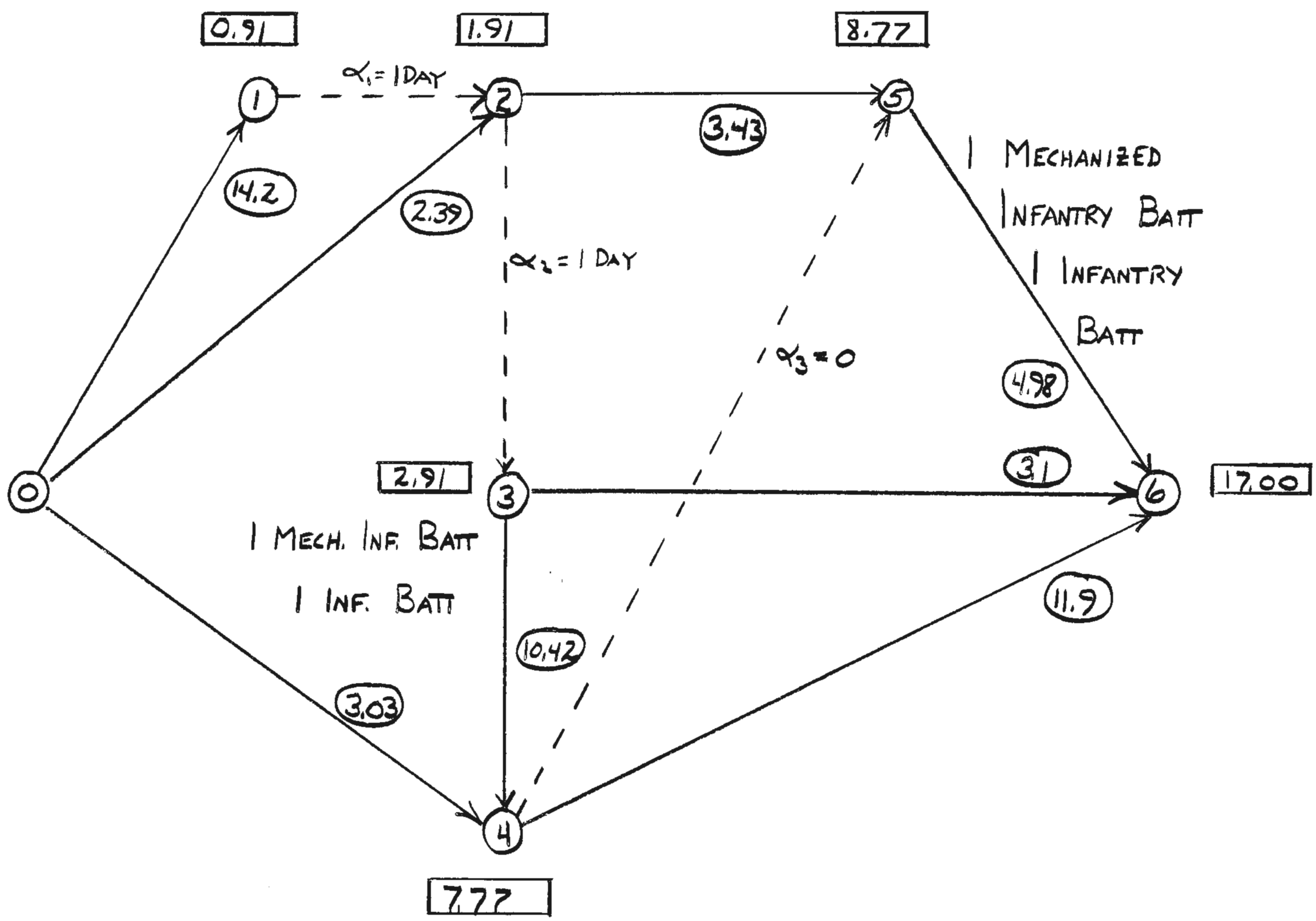


Figure 2.17. Problem 3 Scheduling Network

Throughout the three examples considered, the duration of the most critical activity, the air-borne units, has ranged from a low of 6. 54 days to a high of 7. 90 days. The CINC has expressed the urgency for these units in his partially ordered set of arrivals and it is apparent that these fighting forces should have immediate priority. In Problem 4 we have set the closure date for these units to day 2 by adding an additional row constraint setting X_{04} (the amount of resource allocated to activity (0, 4) equal to 10 units.

This problem is shown in Figure 2. 18 and the results tabulated in Table 2. 1.

Since a closure date of 2 is a radical change for the closure date of the same activity in Problem 2 (7. 9 days), we should expect other major changes in the results. Note, for example, that the closure times of (0, 1) and (0. 2) increase for the first time to 1. 49 and 2. 50 days respectively. This is because resources are diverted from these activities in order to allocate 10 units to the air-borne units.

It is especially important that the other combat shipments (3, 4) and (5, 6) have both been delayed because of the resource allocation. Activity (3, 4) will now arrive at day 11. 45 (up 3. 65 days from the 7. 8 of Problem 2) and activity (5, 6) does not arrive until day 19. 85 (up to 2. 90 days from the 16. 95 of Problem 2). Thus, the planner and CINC must decide if it is worth a delay of up to 3. 65 days in the main force combat units to gain 5. 9 days in the arrival of the initial fighting force.

There is another possible approach to the problem of the trade-off between a particular unit closure time and the deployment closure time. It might be feasible to include a second term in the

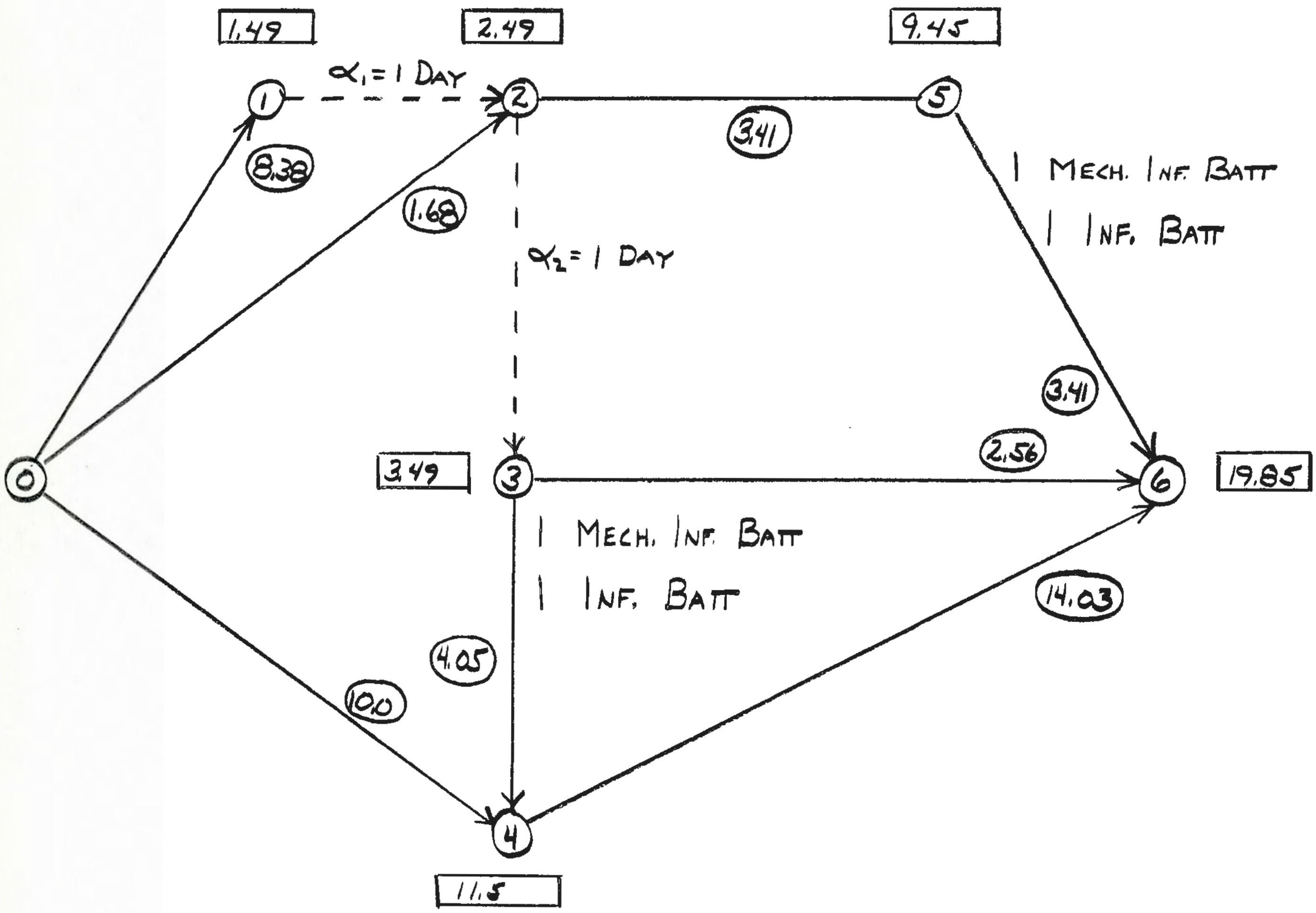


FIGURE 2.18. PROBLEM 4 SCHEDULING NETWORK

objective function to reflect the closure time of the particular shipment. Then, by varying the co-efficient of the additional term to reflect different closure times for the shipment, the range of the deployment closure time would trace out the trade-offs.

This approach was not used in this report, but is offered as an interesting modification technique.

II. 5 Summary

In Section II. 2 the hueristic approaches were outlined and their drawbacks noted. Then, in Section II. 3 the linear programming model was developed in detail and applied to an example problem with the objective of minimizing closure time. Section II. 4 showed the input data being modified variously to aid in determining the trade-offs implied by evaluating a movement plan over the entire range of the goal vector. Then a set of four alternative movement plans was developed based upon the trade-offs.

In Chapter III a choice procedure will be developed to aid the planner in selecting the optimal movement plan from among a set of alternatives. Since this procedure is designed to work over complex goals and to accept or delete goals at any time in the choice process, it is ideally suited for application to the routing and scheduling problem.

Chapter III

The Problem of Choice

III. 1 Introduction

In the preceding chapter we have developed several alternative movement plans based upon considerations over the entire range of the goal fabric. The principal problem left to overcome is that of making a choice among the alternatives. The problem of choice is greatly complicated by the fact that the transportation planner must consider many varied and complex goals. If he were concerned with one goal, say closure time, there would be little cause for concern. The planner would simply implement that plan with the smallest closure time. In the real world situation, however, the CINC wants not only a rapid deployment but also one that most guarantees unit arrival on time and permits some degree of freedom in the event of unexpected enemy operations. Some of the goals, such as closure time, are traditional and easily measured; but other goals are extremely difficult to quantify and there is no reliable technique for considering them in choice.

Manheim and Hall[1] have developed a method for choice that enables the planner to consider new goals as well as the traditional. This chapter will describe the method in some detail and then show its application to the choice problem faced by the movement planner. The objective will be to develop a framework for choosing the optimal plan from among several alternatives. The movement plans developed in Section II. 4 of Chapter II will be presented as examples to aid in clarifying the choice procedure.

III. 2 The Choice Procedure

III. 2. 1 Analysis Requirements

The process of choice used by movement planners must meet five specific requirements. First, it must deal systematically with multiple goals. Second, the procedure must be able to work with incomplete information about the relative values of the different goals to the movement planner. Third, the procedure used must clarify, not confuse or hide, the issues of choice and point out the trade-offs implied by certain decisions. Fourth, the procedure must be dynamic. Changes in the environment surrounding a deployment must be accepted as the rule and not the exception. Fifth, the method should provide an objective report of what is largely a subjective procedure, so that the logic of a decision can be understood by a second party. The method developed by Manheim and Hall meets these requirements to varying degrees and is described below in detail.

III. 2. 2 Goal Fabric Analysis

The procedure has two principal parts. The first is the goal fabric analysis and it consists of listing all the known goals for the project and then identifying the various relations among them. The second is the procedure utilizing the goal fabric analysis to rank the alternatives. This entails mapping each new alternative onto the goal fabric (i. e., predicting the performance of the alternative with respect to some of the goals) and then, using the mapped information and the structure of the goal fabric, compare the new alternative with one previously ranked. The method operates on only two alternatives at a time.

The list of goals is developed by considering the most general goal variables, such as maximize military effectiveness, and then asking such questions as, "What do we mean by that goal? How can we accomplish it? What does it have in common with the other goals?" Answering these questions will enable the planner to decide whether one goal is a more detailed specification of a higher level goal or a means to achieving the next higher level goal. For example, a means to achieving the end of increasing unit effectiveness is to reduce the time a shipment spends in transit. But a specification of the goal vulnerability is the probability of delay due to congestion.

After the list of goals has been expanded and the relationships between the goals determined, a hierarchical tree type structure can be produced with the general goals on top, proceeding down through the specification and means-end relationships to the lowest level goals, those for which we hope to predict and measure the performance of the alternatives. This step terminates the first part of the method.

It is worthwhile to stop and consider what the application of this procedure has accomplished to this point. Most important, it has forced the planner to ask himself exactly what it is he wishes to accomplish. He should not be content with vague goals that are meaningless in consequence, but should pursue each goal chain to its most elementary level. This process should help to clarify the overall objective.

Secondly, and perhaps just as important in an extremely complex problem, the method has helped the planner to systematically structure the problem. The hierarchical structure is indispensable in providing a physical framework for the analysis to follow.

III. 2. 3 Alternative Ranking

The second part of the analysis begins by mapping the alternative onto the goal fabric. This procedure entails predicting the performance of that alternative with respect to some of the goals. It is necessary to determine which goals can be predicted and measured with some accuracy. These goals will not always be the lowest level goals, but every branch of the tree must contain a predictable goal. If there is a branch of the tree that cannot be predicted, then the goals of that branch must be deleted from the decision process. Also, if there is a particular goal that cannot be accurately predicted, then that goal cannot enter the decision.

The predicted data is then converted into preference information on each goal. This entails deciding which alternative is preferred on that goal and the degree to which it is preferred. The measure of preference ranges from such absolute quantities as days for the measure of closure time, to relative measures like "good" or "poor to fair" for such goals as unit effectiveness and flexibility.

The last step is to condense all of the available data into a choice. The information available to the movement planner at this point consists of the following: 1) the CINC's partially ordered list of requirements representing his build-up strategy; 2) the goal relationships portrayed by the hierarchical tree-type structure; 3) preference information on each predictable goal; 4) data already accumulated about the CINC's preference over the various goals and among each goal combination; 5) any additional preferences asked of the CINC for the particular deployment.[2]

Note especially the last two sources of information. Consider the data already accumulated. This procedure is designed to build a data bank which grows each time it is used.

Subsequent choice procedures would then benefit from any pertinent information stored in the data bank. A simple example of the data storage might be as follows. Assume the planner makes the decision during a particular deployment that a closure time increase of one day is acceptable if the unit effectiveness of a shipment is increased by decreasing the time that shipment spends in transit by three days. If the procedure were automated, the routine would record that a one day deployment increase is worth a three day decrease in any particular shipment. This information would then be stored for future use.

Next, consider the possibility of additional preferences asked of the CINC. Many times it will be difficult, if not impossible, for the planner to maintain direct contact with an area CINC during the choice process. Thus, the need for a method that works despite incomplete information is apparent. Also, this source of information is a measure of the flexibility of the choice procedure. Even at this late date in the decision process the CINC and the movement planner maintain the capability to adapt to changes in the environment.

The last step in the method is to use the available information to move from one level to another in the goal tree, from the predictable goals to the next level goals. Manheim and Hall outline three techniques that can be used to condense the data.[3] All the techniques operate to give information on one higher level goal at a time, working with those goals which comprise the higher one. The three techniques are described below.

(1) Dominance: the same alternative is preferred on all the goals comprising the new one; hence, that same alternative is preferred on the new goal.

(2) Explicit choice by the planner: faced with a small subset of goals, the planner may be able to evaluate trade-offs and choices mentally, and give an answer.

(3) Comparison of intervals: find the interval between the alternatives on each goal, and then decide how these intervals compare with each other.

These techniques will be further clarified in Section III. 3 where the alternative movement plans developed in Chapter II will be analyzed by the choice procedure.

Before moving to the next section and the actual mechanics of applying the goal fabric concept described above, it is necessary to briefly summarize the benefits a movement planner might enjoy if he were to apply this method to the routing and scheduling problem. Its greatest advantage is that it provides a framework for the solution of complex problems. The degree of the framework can vary from a simple hand computation scheme conducted by a single movement planner, to an elaborate automated routine to allow for the information storage necessary to handle large, full scale problems. In either case, the method provides a rational approach which allows ample room for subjectivity in choice, but also points out the reasons for the choice.

Another advantage of the procedure is that it is a dynamic method that is designed to adapt to revision, addition, or deletion of goals. Its flexibility permits a high degree of direct planner participation; or it can use previously gathered preference data to indicate consistent choices without planner participation.

III. 3 Application of the Choice Procedure

III. 3. 1 Goal Analysis

Now that we have seen how the method operates, we can describe the particulars of the application of this method to the routing and scheduling problem. The first step is the goal fabric analysis.

We can say that the overall goal of deploying a force package is to maximize the military effectiveness of the deployment. But we must define what we mean by military effectiveness; and once we define the term we must decide how to maximize it. In fact, we should even become so basic in our approach that we ask ourselves who it is that decides the measure of military effectiveness. Clearly, the CINC is concerned with this ability to wage war. He would like to give little thought to considerations of transportation resource availability and terminal congestion. To the CINC, maximum military effectiveness means delivering each unit of the force package at the requested time and at the requested place.

On the other hand, the movement planner may find it impossible to meet the CINC's exacting requests. His conception of maximizing military effectiveness may be to deliver each unit as close to the desired delivery date as possible while observing the resource constraints imposed by limited movement capability. The planner may also attempt to route as many units through the desired ports of debarkation as possible while observing the capacity constraints of the terminals and links in the transportation network.

There are other individuals who view the problem from other points of view. For example, those responsible for transportation resource maintenance may see the need for periodic maintenance checks if they are to guarantee any resource availability. Thus, the planner and the CINC may have to accept a reduced capability in order to allow for the maintenance checks.

Also, a port captain may see the problem only as one of receiving each unit, processing it, and arranging for shipment. He may accept each unit as it arrives in the port area with little or no regard for priorities. The result being that a critical shipment could become delayed due to port congestion.

These are some of the considerations and the people that define the term military effectiveness. The important point is that no one measure of this variable could possibly please all those affected by the deployment. The approach in the choice process is to construct the goal tree out of all these considerations and viewpoints.

We have now decided that the objective of the choice procedure is to select that movement plan which maximizes military effectiveness. The next step, as presented in Section III. 2. 2, is to expand the list of goals from the general goal to the next level goals (shown as level 1 in Figure 3. 1). In order to expand the list we ask the questions, "What does maximize military effectiveness mean?" And, "How do we accomplish this goal?"

The first goal at level 1 is flexibility. This goal is important to both the movement planner and the CINC for the reasons detailed in Chapter I. After establishing flexibility as

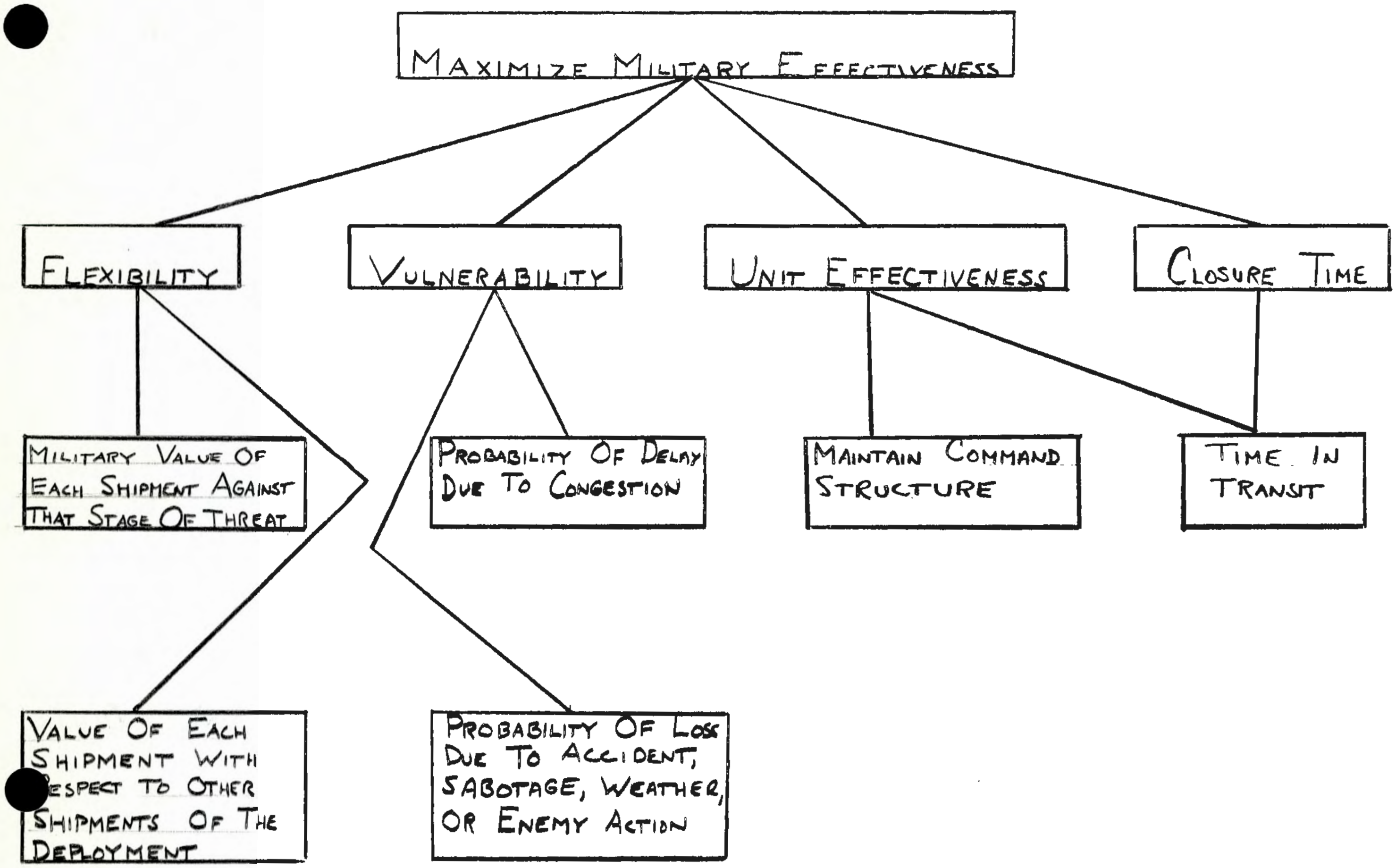


FIGURE 3.1. DEPLOYMENT GOAL STRUCTURE

a level 1 goal, the planner must ask once again what the goal means and how it can be achieved. The answers to these questions enable the planner to move to level of 2 of the goal tree (Fig. 3. 1). Notice that the level 2 goals under flexibility say that the effectiveness of a deployment is partially dependent upon, 1) the military value of each shipment against a particular stage of the enemy threat (this goal models the CINC's desire to retain the capability to adjust to unforeseen circumstances), and 2) the value of that shipment with respect to all the other shipments (here is where the planner decides what it "costs" to provide the requested flexibility). This example shows that constructing the goal tree does indeed define military effectiveness in terms of all those parties concerned with the deployment.

Now we move to each level 1 goal in succession and ask what it means and how it can be achieved. (Each of these goals is shown in Figure 3. 1 and is defined in Chapter I.)

In considering the goal of vulnerability, we determine two areas of importance. First, we must consider the probability of individual unit delays due to congestion of terminals and links. The delay may be unavoidable in some cases, but at other times re-routing of critical units may relieve the congestion.

The second area of concern is the loss of some phase of the CINC's capability due to enemy action or weather. The planner must consider the possibility of splitting units in order to increase the probability that the CINC will retain at least some degree of capability in all the phases of his war-making effort.

Using this same procedure and considering the details of Chapter I, Section I. 4, the planner can move through the level 1 goals and establish the goal tree of Figure 3.1. Notice the description "completed goal tree" is not used. This method is designed to accept changes in the goal structure and the addition or deletion of goals at any time. The goal fabric tree presented here represents the author's conception of the routing and scheduling choice problem goal structure. The goals described and the interactions defined should be questioned and refined as more exacting techniques are developed to measure the variables.

It must be stressed that regardless of the shape or content of the goal fabric, the choice procedure can be applied to select the preferred alternative. The particulars of the example problem are offered primarily as a stimulus for further research in the problem area of complex goals.

III. 3. 2 Mapping the Alternatives

The next step in the choice process is to map the alternatives onto the goal fabric. This entails deciding which goals can be predicted and then listing the predictions (see Figure 3.2). Note that mapping an alternative onto the goal fabric requires a unit of measure over each goal. In some cases the measure is quantifiable and directly applied. For example, closure time is measured in days. Under the goal of unit effectiveness, the time a particular shipment spends in transit can be measured in days. Also, a measure of the fragmentation of the shipment is the number of vehicles required to transport it.

In many cases the measure over the goals cannot be quantified. When this occurs, the planner must exercise his subjective impression of the goal as a measure of the goal's value. The goal, value of that shipment with respect to other

GOAL	MEASURE			ALTERNATIVE 1	ALTERNATIVE 2
Probability Of Delay	Planner's	Subjective	Impression	High	Medium
Probability Of Loss	"	"	"	High	Med.
Military Value Of Each Shpmt	"	"	"	Med	Med
Value W/Rspct To Other Shpmts	"	"	"	Med	Med
Maintain Command Structure	"	"	"	Good	Good
Time In Transit		Days			
Deployment Closure Time		"		16.17	16.95

Figure 3.2. Predicted Performances Of Alternatives 1 And 2.

shipments, is difficult if not impossible to quantify. However, the planner can rank the alternative over that goal subjectively with rankings such as "good" or "poor to fair". In most instances, this type of ranking will be sufficient to move to the next higher level. If at any time the planner feels uncertain about the ranking, he should delete that goal from the decision process.

Before mapping the alternatives onto the goal fabric, we must note a few particulars of the single resource model. Notice, for example, that the goal flexibility in Figure 3.1 does not reflect quite the same meaning as it did in Chapter I. In that chapter, this goal was intended to reflect the costs incurred by shipping a unit by one mode vice another mode in order to guarantee that it would be available to the CINC at any time. In the single resource model, we consider only one mode (aircraft). Therefore, each shipment will always be available if needed unexpectedly.

Since we have only one mode in this model, the level 2 goal that reflects the cost incurred by shipping over one mode vice another mode has been deleted from the analysis. As a result, the goal of flexibility will model only the effectiveness of each shipment of the deployment with respect to the military threat and the value of each shipment to the shipments that compliment it.

We begin mapping the alternatives onto the goal fabric by considering first Problem 1 and Problem 2. The predictable goals have been listed in Figure 3.2 along with the measure over each goal. Notice that some measures can be quantified while others are subjective and express the particular experience of the planner.

In selecting the best plan between 1 and 2, we conduct a pair-wise comparison over all the goals. Consider first the goal of vulnerability. Over the level 2 goal of probability of delay due to congestion, the planner feels that Problem 2 is preferred. His reasoning may be as follows. In Problem 1 both of the heavy units (armor and mechanized infantry) are competing for those facilities that offer heavy duty equipment. However, in Problem 2, half of the mechanized units are shipped in activity (3, 4), before the armored shipment begins. As a result, there should be less congestion at the heavy duty terminals.

Moving to the next goal of Figure 3. 2, probability of capability loss, the planner prefers 2 over 1. This is because the two infantry capabilities have been re-defined and combined.

We now move across the level 2 goals to those considerations under flexibility. With respect to each of the two goals (value against military threat and value to other shipments), the planner feels the alternatives are equally effective.

In considering the goal of command structure under unit effectiveness, we see that the two alternatives are again considered equal. However, the time in transit goal has been affected and the planner prefers alternative 1 over this goal. Notice that the closure time of the air-borne units increases from 6. 54 to 7. 9 days for Problem 1 to Problem 2 (shown in Table 2. 1). Since these units are the first line of defense, the CINC would want them as quickly as possible.

Also, re-defining the infantry units has increased the closure time of activity (3,4), the first infantry-mechanized infantry combination, from day 6. 54 to day 7. 87. Similarly, activity (5, 6) does not close until day 16. 95 (vice 16. 17 for

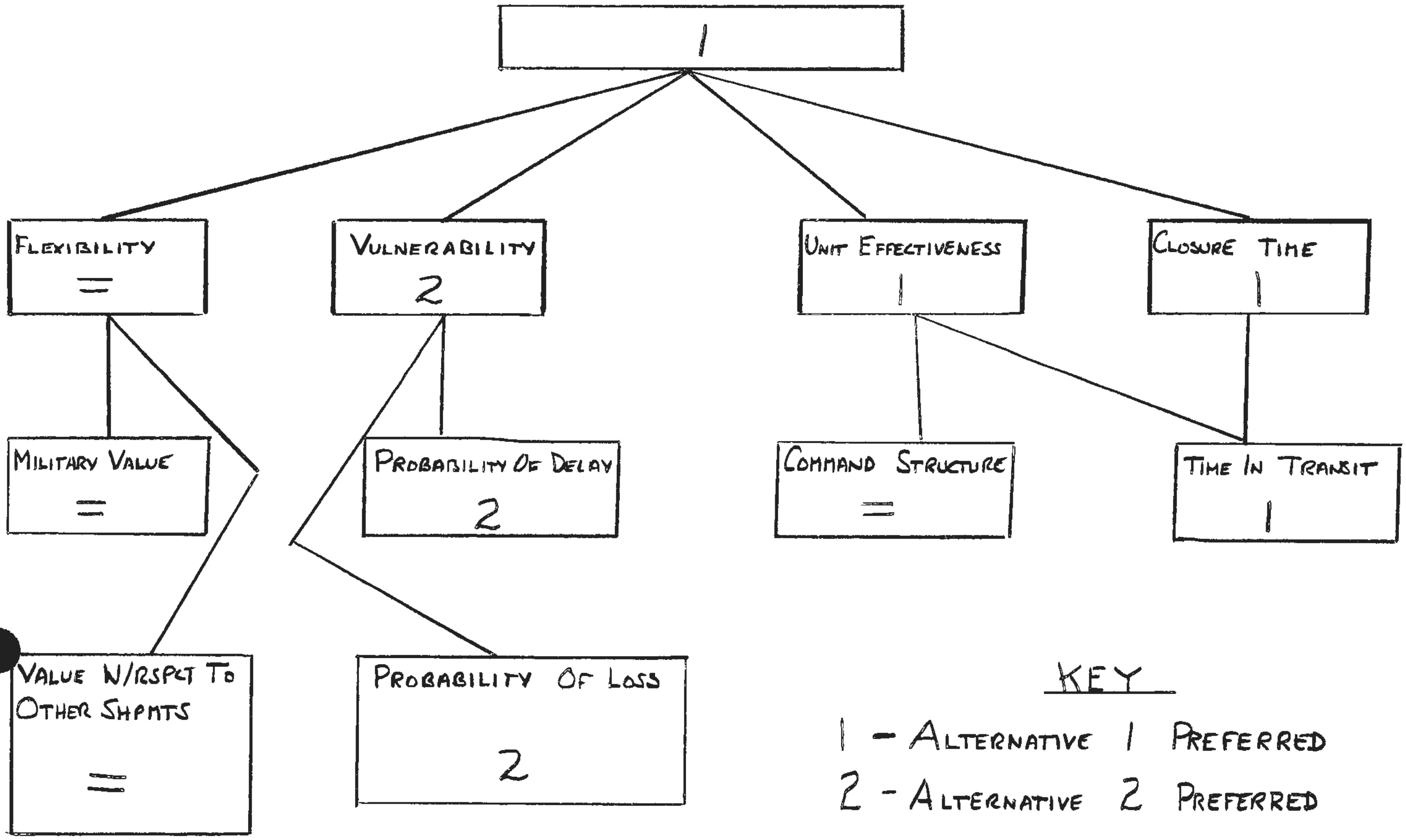
Problem 1). Thus, in regard to time in transit and closure time, the planner prefers Problem 1 to Problem 2.

After the pair-wise comparisons over each goal are completed, the next step is to move up one level in the goal structure from the predictable goals to the next level goals. In Figure 3.3, the preferences determined in Figure 3.2 have been recorded on the predictable goals of the goal fabric. Note that a 1 means that Problem 1 is preferred and a 2 means that Problem 2 is preferred. An = means that both plans are equally effective.

In moving to the next higher level we utilize the techniques presented in Section III. 2.3 for condensing the data into a choice. Referring to Figure 3.3, it is apparent that we can transfer the preferences directly to the level 1 goals because one alternative is dominant over each branch of the tree at level 2 (see Techniques (1) in Section III. 2.3, Dominance). For example, Problem 2 is preferred over both the probability of delay due to congestion and the probability of loss due to enemy action. Therefore, Problem 2 must be preferred over the next higher goal of vulnerability.

It is not quite so apparent what alternative should be transferred from level 1 to level 0. We see in Figure 3.3 that both are equally effective over flexibility, but Problem 2 is preferred over vulnerability. However, Alternative 1 is preferred over both unit effectiveness and closure time.

At this level, the planner may have to make an explicit choice as to which is the preferred alternative (Technique (2)). For instance, intelligence data may indicate that the enemy does not possess the capability to inflict losses on our units



KEY

1 - Alternative 1 Preferred

2 - Alternative 2 Preferred

= - The Two Are Equally Preferred

Figure 3.3. Goal Fabric Showing Alternative Preferences Over The Predictable Goals.

while they are in transit. Based upon this information, the planner may decide that the overriding consideration is "to get there fastest with the mostest". Let us assume this is indeed the case. Thus, alternative 1 is transferred to the level 0 goal and is the preferred movement plan.

Now that Alternative 1 has been identified as a likely choice for the optimal plan, we compare it in the same way with a new alternative, Problem 3. Figure 3.4 shows the predictable goals for this choice problem and the pair-wise comparisons.

Notice that alternative 3 is preferred to alternative 1 over the vulnerability goals for the same reasons that alternative 2 was preferred. Also, these two alternatives are no longer equally preferred over flexibility. The value of activity (5, 6) with respect to the military situation has decreased in alternative 3 because it does not close until day 17.00 (vs. day 16.17 in Problem 1). This occurs even though the unit effectiveness of the activity has increased because it spends less time in transit (8.23 vs. 9.42 days). This decrease in transit time causes an increase in the goal of command structure maintained for activity (5, 6).

However, before we can say that 3 is preferred over that goal, we must check the other critical units. In Alternative 1, activity (0, 4) closes at 6.54 days (vs. 7.90 for Alternative 3). Also, activity (3, 4) closes at day 6.54 with a duration of 3.62 days (vs. 7.77 and a duration of 4.86). Thus, even though 3 is preferred over activity (5, 6), Problem 1 is preferred over the goals of command structure and transit time because the primary consideration is still the delivery of the

GOAL	MEASURE	ALTERNATIVE 1	ALTERNATIVE 3
Probability Of Delay	Planners' Subjective Impression	High	Med
Probability Of Loss	" " "	High	Med
Military Value Of Each Shpmt	" " "	High	Med
Value W/Rspct To Other Shpmts	" " "	Med	Med
Maintain Command Structure	" " "	High	Med
Time In Transit	Days		
Deployment Closure Time	"	16.17	17.00

Figure 3.4. Predicted Performances Of Alternatives 1 And 3.

fighting units as quickly as possible.

Based upon Figure 3. 4, the preferences have been mapped onto the goal fabric of Figure 3. 5. Once again we move to the level 1 goals because all branches of the tree are dominated by one of the alternatives. Also, the planner must make the explicit choice of alternative 1 to move to the level 0 goal because the vulnerability advantages of Alternative 3 coupled with the decreased duration time of activity (5, 6) are not enough to offset the fact that the critical units actually reach the combat areas later than in Alternative 1.

Once again, Progirm 1 has proven to be the better plan. Now we conduct a pair-wise comparison of 1 and 4. Figure 3. 6 shows the predictable goals and the pair-wise comparison over the goals.

Over the goal of vulnerability we see that Alternative 4 is preferred for the same reasons 2 and 3 were preferred. Now we consider the two goals at level 2 under flexibility. First, the military value of that shipment with respect to the enemy's threat. Clearly, Alternative 4 is preferred to a great degree because the first combat units arrive at day 2 instead of day 6. 54 as in Alternative 1. However, the value of the shipment with respect to other shipments is less for Alternative 4 than Alternative 1. This is because allocating so much resource to (0, 4) has diverted resources from (0, 1), (0, 2), (3, 4), and (5, 6) and increased their duration times.

We move now to the goals at level 2 under unit effectiveness. Clearly, Alternative 1 is preferred over command structure and time in transit because, even though the duration of (0, 4) and (4, 6) have decreased, the duration of (0, 1), (0, 2),

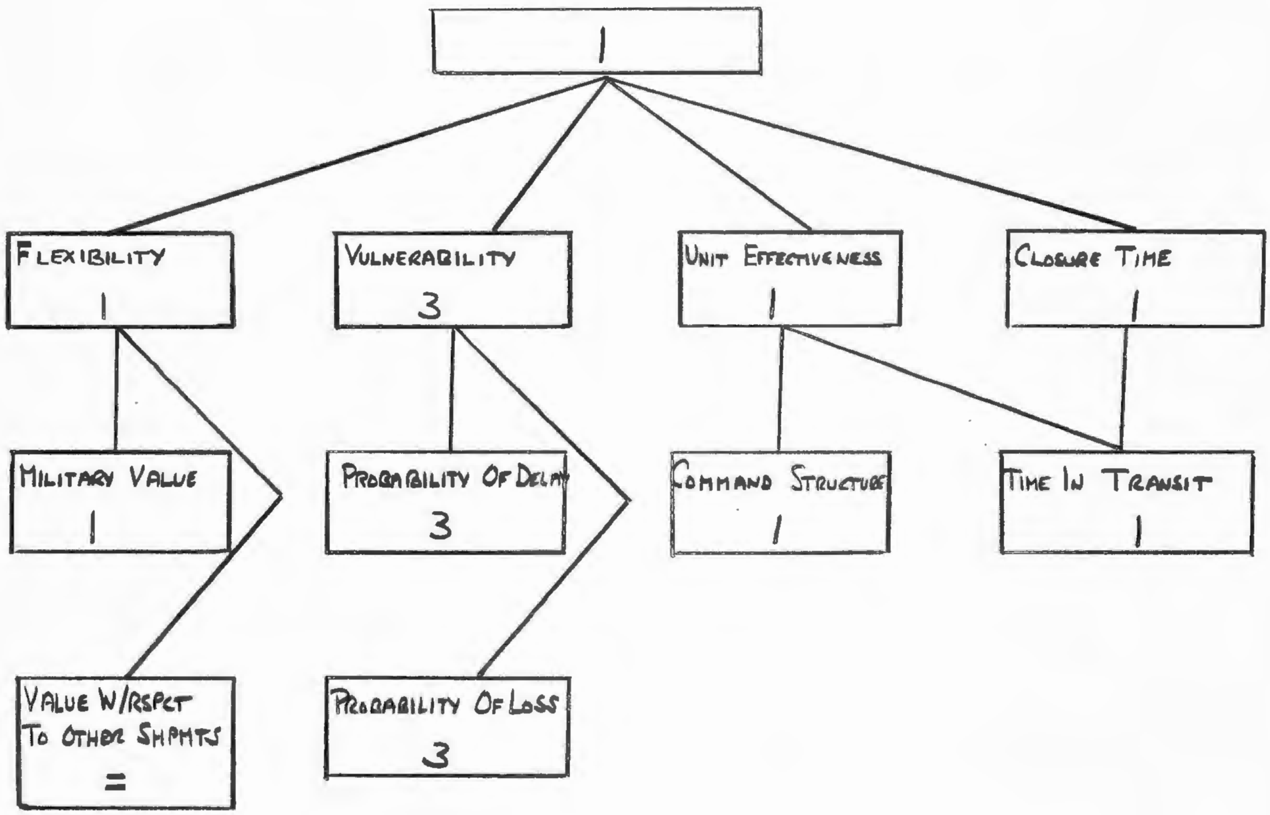


FIGURE 3.5. GOAL FABRIC SHOWING ALTERNATIVES 1 AND 3 PREFERENCES OVER THE PREDICTABLE GOALS.

GOAL	MEASURE	ALTERNATIVE 1	ALTERNATIVE 4
PROBABILITY OF DELAY	PLANNERS' SUBJECTIVE IMPRESSION	HIGH	MED
PROBABILITY OF LOSS	" " "	HIGH	MED
MILITARY VALUE OF EACH SHPMT	" " "	LOW	HIGH
VALUE W/RSPCT TO OTHER SHPMTS	" " "	HIGH	MED
MAINTAIN COMMAND STRUCTURE	" " "	HIGH	MED
TIME IN TRANSIT	DAYS		
DEPLOYMENT CLOSURE TIME	"	16.17	19.85

FIGURE 3.6. PREDICTED PERFORMANCES OF ALTERNATIVES 1 AND 4.

(3,4), (3,6), (2,5) and (5,6) have increased. Also, the closure time of 4 is 19.85 days (vs. 16.17 for 1).

These preferences have been mapped onto the goal fabric shown in Figure 3.7.

We can use the technique of dominance to move from level 2 to level 1 under the goals of vulnerability and unit effectiveness. Under flexibility, Alternative 4 is preferred over one goal and Alternative 1 over the other.

In order to move up to the level 1 goal of flexibility, the comparison of intervals technique must be used. We apply this technique by finding the interval between the alternatives on each goal and then decide how the intervals compare with each other. From Figure 3.6 the interval over the goal of military value against enemy threat is Low for Alternative 1 to High for Alternative 4. And over the goal of value with respect to the other shipments, the interval is High for Alternative 1 to Medium for Alternative 4.

The next step is to decide how the intervals compare with each other. First consider the interval Low to High over the first goal. This interval represents the fact that Alternative 4 delivers the first combat capability in two days versus 6.54 days in Alternative 1. The second interval of High to Medium over the goal of value of the shipment with respect to the other shipments represents the increase in the duration times of the critical activities (3,4) and (5,6) in Alternative 4 over the duration times of Alternative 1. The increases are 3.65 days for (3,4) and 2.9 days for (5,6). The planner must decide how these intervals compare with each other.

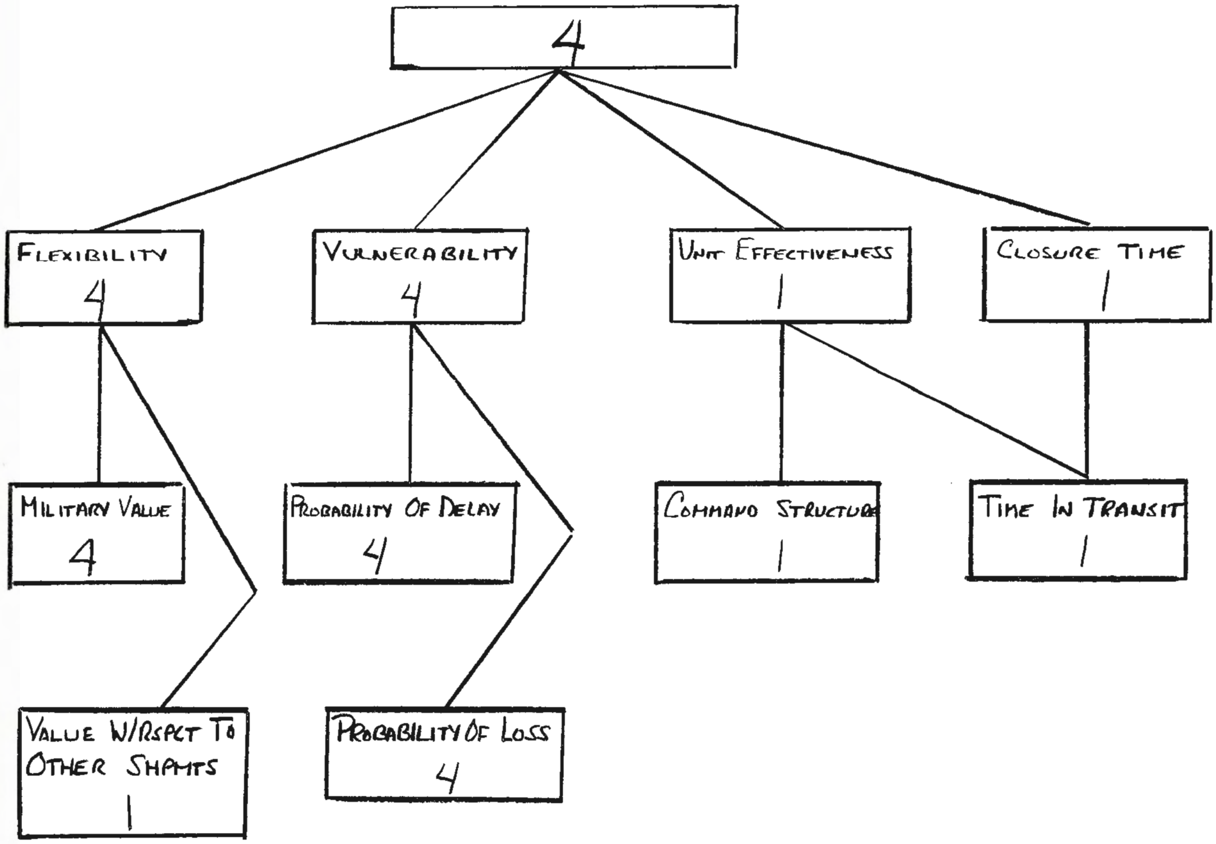


Figure 3.7. Goal Fabric Showing Alternatives 1 And 4 Preferences Over The Predictable Goals.

Let us assume once again that the overriding concern expressed by the CINC is to stop the enemy as quickly as possible. Based upon this criteria, the planner determines that the interval of 4.54 days between the arrival of activity (0, 4) in Alternative 1 and Alternative 4 should govern the decision. As a result, he accepts the delays in the combat units of (3, 4) and (5, 6) and chooses Alternative 4 over the goal of flexibility.

The final step is to move from level 1 to level 0. Note that Alternative 4 is preferred over vulnerability and flexibility, and Alternative 1 is preferred over unit effectiveness and closure time. Since he is now faced with a small subset of goals, the planner can evaluate the trade-offs and choices mentally and express a preference at the level 0.

If he were consistent in his approach to the choice process, he would recognize that the CINC must receive some combat units as quickly as possible in order to deny the enemy the initiative. As a result, he would reluctantly accept the increased closure time of Alternative 4 in order to put fighting units into the field at the end of day 2.

Plan number 4 is the optimal of the four alternatives considered.

III. 4 Summary

This chapter has presented in some detail a proposed approach to the problem of choice in light of the complex goal structure of the routing and scheduling problem. In Section III. 3 this procedure was applied to the alternatives developed in Chapter II and an optimal plan selected.

It is important to realize that the choice process does not occur in the quiet and order of the academic atmosphere. It is a dynamic procedure capable of assuming many forms as the priorities of a deployment change in response to many influences. Thus, it is especially important to note Section III. 2. 3 concerning the information available to the movement planner.

It is evident in the example of the preceding section that the problem of choice is simplified if the planner is continually aware of

the CINC's appraisal of the military situation. Since the military situation can be expected to change rapidly, the planner must monitor the information sources at all times in order to be aware of shifting priorities.

We can conclude, then, that the major importance of the choice process is its ability to effectively deal with changing goals. Also of importance is the fact that the process offers a framework for considering complex goal problems.

Chapter IV

The Movement Planning Environment

IV. 1 Introduction

The overall objective of this chapter is to present the necessary environment for the application of the results of Chapters II and III to the problem of movement planning. In particular, this chapter will focus upon such aspects as the logical processes, the capabilities and limitations of the detailed evaluation and choice procedures, and the alternatives for the movement planner.

In Section IV. 2 these aspects will be considered from the standpoint of today's movement planner. The objective will be to show how the described choice and evaluation procedures can be implemented using available hardware and software.

Section IV. 3 will consist of a presentation of the author's conception of the ideal movement planning process. Most of these concepts are available today in varying degrees but many require complex set-up phases. The objective of this section will be to indicate the potentialities of the system and to indicate a general direction of travel toward fulfilling the potentialities.

IV. 2 System Implementation

Consider the evaluation procedure presented in Chapter II. It is important to note that the mathematical formulation of the movement planning model could serve as input to any linear programming routine. The essential element of the formulation is the mathematical approach to the problem taken by Groninger and not the computer hardware used for a solution.

Although it may prove difficult to break standard procedure, the concept of partial ordering (which forms the foundation of the entire procedure) is available today. In fact, when a CINC prepares his fully ordered set of requirements, it is imperative that he make the pair-wise comparisons of the movement packages involved in the deployment. Therefore, it would be worthwhile to record these precedence relations before they get lost in the aggregation of a fixed ordering so that the routing and scheduling procedures can capitalize on their inherent flxibility. [1] Thus, the basic bools for implementation of this model are available today. These tools are the concept of partial ordering which leads to the scheduling network and the approximation techniques used by Groninger in the model formulation.

The model, in its present state, does have its limitations. It was emphasized in Chapter II that this is a single resource model. In a real world atmosphere, the planner could expect to encounter a multiple resource type problem where he has various lift capabilities at his disposal. The resource mix could range from a land, sea, and air capability with varying duration curve shapes to mixes of different aircraft types. In the latter case, the different aircraft productivities would cause the C constant to lose its effectiveness,

$$C = \frac{T\,d}{p} \tag{4.1}$$

If p, aircraft productivity, is allowed to vary, then we no longer have a constant C value to measure the work needed to move a shipment.

This question has been considered by Groninger and work to date indicates that a viable formulation is possible. [2]

Another limitation of the model has been pointed out in Chapter II and involves the scheduling of dummy activities. In order to achieve truly optimal results, there should be no idle resources per time unit. However, even with this inconsistency, this model produces highly acceptable results. In addition, as long as the planner and the

CINC are aware of the idle resources, they can be used to move retrograde cargo and evacuees; both of these tasks are necessary during an ongoing deployment.

In Section II. 4. 1, the linear programming model was shown not to be an end in itself but just one step in the evaluation process. It was shown that each movement plan must be evaluated for considerations of vulnerability, flexibility and unit effectiveness.

The output capability described in Phase IV of the proposed Batch Processing Mock-up computer program (BPM) would greatly aid in evaluating a movement plan.[3] This phase of the BPM would sort and process the results of earlier phases and produce a number of different reports. The planner could select the report he desires by means of an input card.

The types of reports that would be useful in determining the goal trade-offs might be as follows:

(1) Origin, POE, and destination tabulations---
For each of these locations, a table is produced, containing the units moving through each point with a total of tonnage and pax for each D-day.

(2) Subtotals for each origin, POE, and destination---
This report would be the same as in (1) except that only total tonnage and pax per D-day would be shown.

(3) Near capacity report----
This report would show only those figures in the (2) report where above 90% (or any chosen percentage) capacity is used.

(4) Rail cars, ships report---
This report would contain the number of rail cars and ships

by area and time.

As the model now stands, only the linear programming standard output is available. The above capabilities would be of great benefit to the planner.

Notice that the choice procedure of Chapter III is completely independent of the method for generating the alternatives. This procedure is not subject to the limitations of the model formulation. The only limitations the choice process encounters are those imposed by the decision maker. The procedure can only be as good as the goal fabric constructed by the planner. And the planner's choice of alternatives can only be as good as his use of the available information. Once again, the need to continually monitor the deployment atmosphere is imperative.

To date, no computer routine has been developed to model this procedure. However, the example of Chapter III has shown that the general framework of the procedure is quite adaptable to hand computation. In fact, this is a part of the procedure's strength. It encourages the planner to approach a complex goal problem from the most basic elements through the goal tree to the eventual choice. Any coding of the procedure should retain the planner interaction capability and remove only the burdensome calculations and provide convenient storage facilities.

In Section IV. 3 we shall see how these capabilities can be accomplished.

IV. 3 The Man-Machine Interaction

The objective of a military deployment is to maximize the military effectiveness of the deploying force and to obtain that effectiveness from the pool of available resources. Therefore, the basic planning problem is how to decide between alternative ways of using resources such that the maximum military effectiveness can be obtained.

These decisions are difficult to make because the complexities of a deployment overwhelm the logic of the human mind. Even the most experienced planner will find himself unable to cope with the complex projects created by advances in technology. The precedence constrained optimization approach and the resulting scheduling network, coupled with the use of computers to analyze the network, are answers to the planners need to extend his knowledge and more firmly exercise his control over the deployment. The logic of the planning network and the speed of the computer enable the planner to evaluate alternative plans before making a choice.

It is unwise, however, to expect a computer to produce a complete solution to a planning problem. The computer can predict according to the rules it has been taught, but some provision must be made so that the planner can guide the calculations along preferred paths. For this reason, the most effective way of maximizing the effectiveness of a deployment requires a combination of the planners judgment over the areas of uncertainty and the computer's logic to analyze the impact of the uncertainty on the factual data.

A flow chart of a possible dynamic planning model is shown in Figure 4.1. Notice that the evaluation and choice procedures of Chapters I and II are shown in the iteration process of Generation of Alternatives - Analysis - Implementation of Alternatives - Observations on the System. Note that the Observation of the System may also lead to a revision of the deployment objectives. This revision of objectives is an extremely important feature of the model because it allows

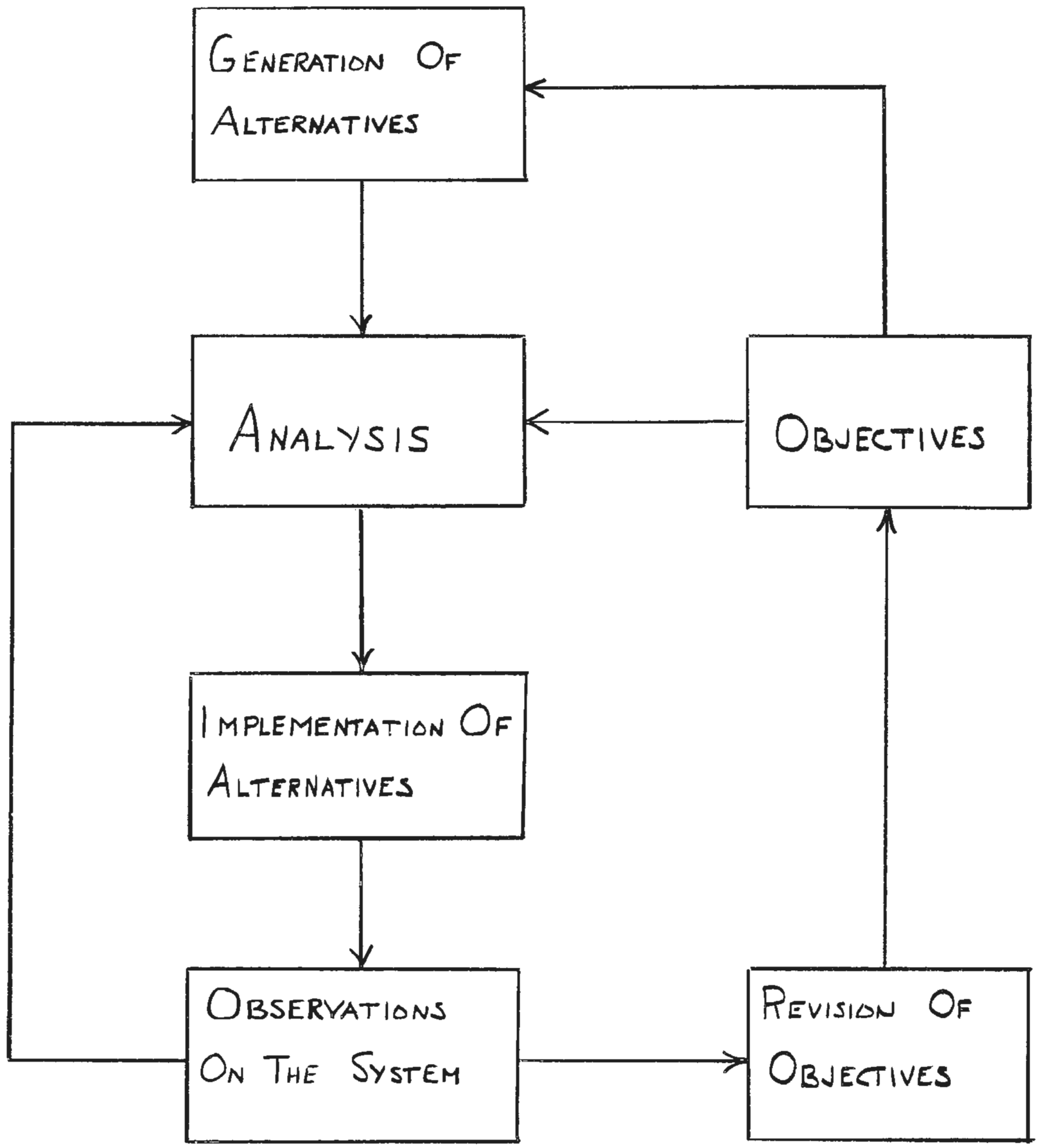


Figure 4.1. Dynamic Planning Model Flowchart

the planner the benefit of learning through experience. As a result, at any stage during the evaluation and choice process, the planner has the capability to re-define the objectives in terms of new information. The new objectives can then be used to generate new alternatives, to alter the criteria for selection in the choice process, or to do both.

It may be feasible to program a computer to make these decisions according to some fixed logic over a particular scheduling network. In fact, it is preferred that standard values be provided so that the planner need not delay the computer when the sequence proceeds as expected.[4] However, one cannot possibly foresee the many peculiarities of each deployment. The planner should be able to "manage by exception" and replace the standard values where necessary.[5]

This type of man-machine interaction requires a highly interactive time-sharing, remote access computer capability. The planner should have access to the computer at each stage of the problem, but need not always interrupt the program. By using a time-sharing system, he can do this most economically.

The ideal system would have a problem oriented language capability to allow the planner greater flexibility and more rapid utilization of the system. Also, the ideal system would include a graphic output capability that would present information that is difficult to convey in a printed line. Such a capability would enhance the understanding of the planner as he reads the output and also enable him to interpret its meaning much more rapidly.

It is important to note that the ideal capabilities described above would be desired only if the system were to be fully implemented and used extensively. It is fairly expensive and time consuming to develop graphic and problem-oriented language capabilities. Also, rather than purchase a time-sharing capability for a limited number of movement planning problems, it would be less expensive and more direct to use a dedicated machine and operator combination.

These trade-offs, of course, could not be identified in this report and must be decided by the individual planning agency.

Chapter V

Summary and Conclusions

V. Summary and Conclusions

The objective of this report was to present a new approach to the routing and scheduling problem. The need for such a departure from the traditional procedures has been created by the great strides achieved in the technology of deployment hardware. Because it is possible to move a division of troops in a matter of days today, instead of weeks or months as in the early 50's, closure time is no longer the sole measure of a deployment's effectiveness. The deployment atmosphere is dynamic, and the movement planner must consider the whole range of goal variables if he is to adapt to the constantly changing deployment environment. As a result, goals such as vulnerability, flexibility, and unit effectiveness have assumed new importance. Since these goals are difficult to quantify and measure, the new management techniques necessary to fully exploit advances in technology must accept their subjective nature and include them in the decision process.

In Chapter I the goal vector was developed in some detail and the concept of the goal fabric was introduced. The goal fabric is intended to show that each of the goal variables is dependent upon the others. A change at one location of the fabric will produce some distortion throughout the fabric.

In order to give the CINC and the movement planner more flexibility in the planning process, the concept of a partially ordered set of requirements was presented in Chapter I. It was shown that such an ordering is beneficial for several reasons. First, the CINC is not dependent upon any predicted build-up rate. He expresses pair-wise preferences over the critical elements of the deploying force and the planner meets these preferences with the available resource capability. The second benefit of partial ordering is that the planner is not tied to a single fixed sequence of arrivals which he must meet. He is free to work within the limits of resource and unit availability as long as he delivers the units in accordance with the CINC's partially ordered list of requirements.

A result of the partially ordered list of requirements and unit and resource availability is the scheduling network. In Chapter II, the network was presented as the basis for several evaluation methods including the hueristic approaches of the critical path techniques. After a study of these methods, it was decided that the analytic approach of the linear programming formulation best modeled the routing and scheduling problem. It was shown, however, that the model described must be refined before it can be considered operational. The major limitations of the model are that it does not satisfactorily handle dummy activities and a multiple resource capability is not fully developed.

Since the linear programming model is designed to minimize deployment closure time, it was necessary to apply certain post-optimal adjustments to the scheduling network in order to reflect considerations over the entire range of the goal fabric. In this way, alternative movement plans were developed that reflected the relationships and trade-offs among the goal variables.

The next step in the process is the choice procedure. In Chapter III, the method of choice among the alternative movement plans was presented. Throughout the chapter the importance of the dynamic nature of the procedure was emphasized. It was shown that the process was capable of adjusting to the addition, deletion, or revision of the goals at any time during the choice process.

The primary benefit that the planner gains from the choice procedure is that it gives him a framework for analysis in an otherwise complex problem. Hopefully, the procedure will cause the planner to go to the center of the problem and, as he works his way through the goal fabric, to develop new insights into the complexities of the goal variables.

The conclusions to be reached concerning the evaluation and choice procedure are presented in Chapter IV. It is important to note that the basic concepts are available and could be implemented today. Although the model is in need of further refinement, the idea of partial ordering and the construction of a scheduling network could be implemented immediately. Also, the author has no knowledge of any attempts to code the choice procedure. However, the basic structure remains at the planner's disposal to aid in ordering an otherwise overwhelming choice problem.

FOOTNOTES

CHAPTER I

1. J. H. Stafford, "Routing and Scheduling as a Problem in Systems Analysis and as an Optimization Problem", p. 1.
2. M. L. Manheim, "An Overview of Strategic Mobility and Its Implications for Design of Analysis Systems", p. 123.
3. Ibid., p. 124.
4. J. H. Stafford and L. N. Beckreck, "Decision Issues in Routing and Scheduling a Deploying Force", p. 6.
5. Ibid., p. 8.
6. Stafford, op. cit., p. 3.
7. K. L. Groninger, "An Analysis of Tandem Airlift-Sealift Operation for Strategic Mobility", p. 5.
8. K. L. Groninger, "Routing and Scheduling Procedures for Strategic Mobility", p. 3.
9. Ibid., p. 4.

CHAPTER II

1. E. W. Davis, "Survey of Critical Path Techniques", p. 177.
2. Moder, J. J., and C. R. Phillips, Project Management with CPM and PERT, p. 106.
3. T. J. R. Johnson, "An Algorithm for the Resource Constrained Project Scheduling Problem".
4. J. D. C. Little, K. G. Murtz, D. W. Sweeney, and C. Karel, "A Branch and Bound Algorithm for the Traveling Salesman Problem".

FOOTNOTES, cont'd.

CHAPTER II, cont'd.

5. H. M. Weingartner and D. N. Ness, "Methods for the Solution of the Multidimensional O/1 Knapsack Problem".
6. Groninger, K. L., "Routing and Scheduling Procedures for Strategic Mobility", p. 13.
7. Ibid., p. 17.
8. Ibid., p. 19.
9. Ibid., p. 21.
10. Hadley, G., Nonlinear and Dynamic Programming, p. 104.
11. Ibid., p. 121.
12. Groninger, K. L., op. cit., p. 23.
13. STRATMAS Tech. Memo #43, Headquarters, Department of the Army.
14. Groninger, K. L., op. cit., p. 26.
15. Ibid., p. 28.
16. Ibid., p. 29.
17. ICES : Programmers Reference Manual, Jan C. Jordan (ed.).
18. Ochoa-Rosso, F., D. K. Bivins, and H. Thiriez, OPTECH Users Guide.
19. Groninger, K. L., op. cit., p. 42.

CHAPTER III

1. M. L. Manheim, and F. G. Hall, "Abstract Representation of Goals".
2. Ibid., p. 11.

CHAPTER III, cont'd.

3. Ibid., p. 11.

CHAPTER IV

1. Groninger, K. L., op. cit., p. 6.

2. Ibid., p. 42.

3. STRATMAS Tech. Memo #30, Headquarters, Department of the Army.

4. M. L. Manheim, "An Overview of Strategic Mobility and Its Implications for Design of an Analysis System", p. 136.

5. Stafford and Beckreck, op. cit., p. 14.

BIBLIOGRAPHY

1. Davis, Edward W., "Survey of Critical Path Techniques", Journal of Industrial Engineering, April, 1966.

2. Groninger, Kent L., "Precedence Relations in Routing and Scheduling a Deploying Force", Discussion Paper T-18, Department of Civil Engineering, M.I.T., Cambridge, Mass., June, 1967.

3. Groninger, Kent L., "Routing and Scheduling Procedures for Strategic Mobility," Interim Technical Report, School of Engineering, M.I.T., Cambridge, Mass., May, 1968.

4. Groninger, Kent L., "An Analysis of a Tandem Airlift-Sealift Operation for Strategic Mobility", Discussion Paper, Department of Civil Engineering, M.I.T., Cambridge, Mass., January, 1967.

5. Hadley, G., Nonlinear and Dynamic Programming, Addison - Wesley Publishing Co., Inc., Reading, Mass., 1964.

6. ICES : Programmers Reference Manual, J. C. Jordan (ed.), Department of Civil Engineering, M.I.T., Cambridge, Mass., October, 1967.

7. Johnson, T.J.R., "An Algorithm for the Resource Constrained Project Scheduling Problem", Ph. D. Thesis, Sloan School of Management, M.I.T., August, 1967.

8. Little, J.D.C., K.G. Murtz, D.W. Sweeny, and C. Karel, "A Branch and Bound Algorithm for the Traveling Salesman Problem", Operations Research, November-December, 1963.

9. Manheim, Marvin L., and Frederick G. Hall, "Abstract Representation of Goals", Professional Paper P67-24, School of Engineering, M.I.T., Cambridge, Mass., January, 1968.

10. Manheim, Marvin, L., "An Overview of Strategic Mobility and Its Implications for Design of Analysis Systems", Professional Paper P67-25, Department of Civil Engineering, M. I. T., Cambridge, Mass., November, 1967.

11. Moder, J. J., and C. R. Phillips, Project Management with CPM and PERT, Reinhold Publishing Corp., London, 1964.

12. Ochoa-Rosso, F., D. K. Bivins, and H. Thiriez, OPTECH Users Guide, Department of Civil Engineering, M. I. T., Cambridge, Mass., October, 1967.

13. Stafford, Joseph H., and Laurence N. Beckreck, "Decision Issues in Routing and Scheduling a Deploying Force", Professional Paper P. 68-5, Department of Civil Engineering, M. I. T., Cambridge, Mass., August, 1967.

14. Stafford, Joseph H., "Routing and Scheduling as a Problem in Systems Analysis and as an Optimization Problem", Discussion Paper T-9, Department of Civil Engineering, M. I. T., Cambridge, Mass., March, 1967.

15. STRATMAS Tech. Memo #30, Headquarters, Department of the Army, Military Traffic Management and Terminal Service, Washington, D. C., March, 1966.

16. STRATMAS Tech. Memo #43, Headquarters, Department of the Army, Military Traffic Management and Terminal Service, Washington, D. C., March, 1966.

17. Weingartner, H. M., and D. N. Ness, "Methods for the Solution of the Multidimensional O/1 Knapsack Problem", Working Paper 177-66, Sloan School of Management, M. I. T., Cambridge, Mass., Apr. 1966.

LIST OF FIGURES AND TABLES

Appendix A

Input Tableau for Example Problem Number 1

λ_{00}	λ_{01}	λ_{02}	λ_{03}	λ_{020}	λ_{021}	λ_{022}	λ_{023}	λ_{040}	λ_{041}	λ_{042}	λ_{043}	λ_{120}	λ_{121}	λ_{220}	λ_{221}	λ_{250}	λ_{251}	λ_{252}	λ_{253}	λ_{340}	λ_{341}	λ_{342}	λ_{343}	λ_{320}	λ_{321}	λ_{322}	λ_{323}	λ_{460}	λ_{461}	λ_{462}	λ_{463}	λ_{560}	λ_{561}	λ_{562}	λ_{563}	t_1	t_2	t_3	t_4	t_5	t_6	
8	4.4	186	0.4																																	-1						=0
				4	2.2	.93	.42																														-1					≤0
								20	109	4.6	1																												-1			≤0
												1	1																							+1	-1					≤0
														1	1																						+1	-1				=0
																20	109	4.6	1																		+1			-1		=0
																				20	105	4.6	1															+1	-1			≤0
																								36	19.7	8.39	1.79											+1			-1	≤0
																												70	384	16.3	3.5								+1		-1	≤0
																																40	269	9.3	2.0					+	1	≤0
1	1.8	4.29	20	1	1.8	4.29	20	1	1.8	4.29	20																															≤20
-1	-1.8	-4.29	-20									0	20																													≤0
				-1	-1.8	-4.2	-20					-0	-20	0	20	1	1.82	4.29	20																							≤0
														-0	-20					1	1.82	4.29	20	1	1.82	4.29	20															≤0
								-1	-1.8	-4.29	-20									-1	-1.82	-4.29	-20					1	1.82	4.29	20											≤0
																-1	-1.82	-4.29	-20													+1	1.82	4.29	20							≤0
																								-1	-1.82	-4.29	-20	-1	-1.82	-4.29	-20	-1	-1.82	-4.29	-20							≤20
1	1	1	1																																							1
				1	1	1	1																																			1
								1	1	1	1																															1
												1	1																													1
														1	1																											1
																1	1	1	1																							1
																				1	1	1	1																			1
																								1	1	1	1															1
																												1	1	1	1											1
																																1	1	1	1							1

Appendix A.1. Input Tableau For Example Problem 1