72 5c

Handbook for Decision Analysis

DECISIONS AND DESIGNS INCORPORATED

Scott Barclay
Rex V. Brown
Clinton W. Kelly, III
Cameron R. Peterson
Lawrence D. Phillips
Judith Selvidge

ADVANCED DECISION TECHNOLOGY PROGRAM

CYBERNETICS TECHNOLOGY OFFICE
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY
Office of Naval Research • Engineering Psychology Programs

The objective of the Advanced Decision Technology Program is to develop and transfer to users in the Department of Defense advanced management technologies for decision making. These technologies are based upon research in the areas of decision analysis, the behavioral sciences and interactive computer graphics. The program is sponsored by the Cybernetics Technology Office of the Defense Advanced Research Projects Agency and technical progress is monitored by the Office of Naval Research – Engineering Psychology Programs. Participants in the program are:

Decisions and Designs, Incorporated
The Oregon Research Institute
Perceptronics, Incorporated
Stanford Research Institute
Stanford University
The University of Southern California

Inquiries and comments with regard to this report should be addressed to:

Dr. Martin A. Tolcott
Director, Engineering Psychology Programs
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

or

LT COL Roy M. Gulick, USMC
Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

(9)

TECHNICAL REPORT TR-77-6-30 (14)

(6)

HANDBOOK FOR DECISION ANALYSIS

(10)

Scott Barclay, Rex V. Brown, Clinton W. Kelly, III,
Cameron R. Peterson Lawrence D. Phillips, Judith Selvidge

Sponsored by

Defense Advanced Research Projects Agency
ARPA Order 3052

Monitored by

Office of Naval Research
under Contract N00014-76-C-0074

(15)

(11) Sep[redacted]77

(12) 287 p.

DECISIONS and DESIGNS, INC.
Suite 600, 8400 Westpark Drive
McLean, Virginia 22101
(703) 821-2828

TABLE OF CONTENTS

Page

PREFACE

To those who have grappled seriously with important problems of choice in their personal spheres (such as purchasing a home, deciding which automobile to buy, choosing among job options, etc.) the difficulties inherent in decision making are apparent. Even at the relatively uncomplicated level of personal choice, we often encounter more relevant decision dimensions than the intellect can cope with, many dimensions are difficult to value in an objective way, and a highly uncertain world is usually interposed between possible choices and their outcomes.

While most of us cope with personal decisions with more or less systematic consideration (and usually with unknown degrees of success or failure), it takes little extrapolation to realize that in national security decision contexts, the problems are often far more complex, the uncertainties greater, and the stakes involved of enormous magnitude. These factors, coupled with a research base that points to suboptimal human performance in complex decision tasks,[1] have served as the impetus for the development of formal methodologies to aid decision makers. During World War II, formal approaches to decision-making, especially in defense settings, began to be introduced under the name of operations research. They were typically applied to special types of clearcut, repetitive problems, such as those of systematic search and resource allocation.

Since the 1960's, however, a more general technology has emerged for imposing logical structure on the reasoning that underlies any specific decision. This technology is decision analysis. Since 1970, there has been a dramatic burgeoning of efforts by defense agencies to adapt this technology to their day-to-day decision making. Many have found it a way to make better, more defensible decisions.

Decision analysis is a quantitative method which permits the systematic evaluation of the costs or benefits accruing to courses of action that might be taken in a decision problem. It entails identification of the alternative choices involved, the assignment of values (costs/benefits) for possible outcomes,

---

[1]For excellent reviews of this research base, see Becher, G. M. and McClintock, C. G., "Value: Behavioral Decision Theory," Annual Review of Psychology 18:239-286 (1967); Rapoport, A. and Wallston, T. S., "Individual Decision Behavior," Annual Review of Psychology 23:131-175 (1972); Slovic, P., Fischhoff, B., and Lichtenstein, S., "Behavioral Decision Theory," Annual Review of Psychology 28:1-39 (1977).

and the expression of the probability of those outcomes being realized. With this information at hand, one can then systematically combine the values and probabilities to show the probable gain or loss that is associated with alternative choices.

In the application of decision analysis, a problem is decomposed into clearly defined components in which all options, outcomes, values, and probabilities are depicted. Quantification in the form of the value for each possible outcome and the probability of those values (or costs) being realized can be in terms of objective information or in the form of quantitative expressions of the subjective judgments of experts. In the latter case, the quantitative expression serves to make explicit those subjective qualities which would otherwise be weighed in the decision process, albeit in a more elusive, intuitive way.

Beyond its primary role of serving as a method for the logical solution of complex decision problems, decision analysis has additional advantages as well. The formal structure of decision analysis makes clear all the elements, their relationships, and their associated weights that have been considered in a decision problem. If only because the model is explicit, it can serve an important role in facilitating communication among those involved in the decision process. With a decision problem structured in a decision analytic framework, it is an easy matter to identify the location, extent, and importance of any areas of disagreement, and to determine whether such disagreements have any material impact on the indicated decision. In addition, should there be any change in the circumstances bearing upon a given decision problem, it is fairly straightforward to reenter the existing problem structure to change values or to add or remove problem dimensions as required.

It should be emphasized that in no sense does decision analysis replace decision makers with arithmetic or change the role of wise human judgment in decision making. Rather, it provides an orderly and more easily understood structure that helps to aggregate the wisdom of experts on the many topics that may be needed to make a decision, and it supports the skilled decision maker by providing him with logically sound techniques to support, supplement, and ensure the internal consistency of his judgments.

This handbook is intended to provide decision makers and their staffs (current or potential) with an introduction to the basic concepts and operations of decision analysis. As an introductory treatment, it is not expected that mastery of the material presented here will make an expert decision analyst of the reader. Rather, it is our hope that the material will provide readers a level of acquaintance with the methodology that will enable them to incorporate at least elements of decision analysis into their professional decision making activities, to be generally aware of the capabilities of decision analysis, to recognize decision contexts meriting formal decision analysis, and to be able to participate in and evaluate decision analysis applications which will assuredly be encountered with increasing frequency.

Chapters 1 through 5 of this handbook provide an overview of the major technical aspects of decision analysis: structuring models, assigning probabilities and values, concepts of personal probabilities, developing inferences from evidence, and information value. The concepts are developed in the context of defense or defense-related examples which we hope will enable our principal reading audience to better relate to the material. Chapter 6 provides a case study in the form of a dialogue between a decision analyst and a task force commander centered around the use of decision analysis in resolving a decision problem faced by the commander. The case study is designed to exemplify all the principles developed in Chapters 1 through 5. Thus, it can be used as a capstone for those who choose first to step through the development of the principles. Alternatively, those seeking the quickest possible acquaintance with what decision analysis is about and roughly how it works in practice may review Chapter 6 first. Cross references in Chapter 6 to earlier sections of the text which elaborate the principles being applied in the case study will guide the reader to fuller comprehension.

For the reader who elects to go into the subject matter in greater depth, excellent texts exist. Decision Analysis for the Manager by R. V. Brown, A. S. Kahr, and C. R. Peterson (New York: Holt, Rinehart, and Winston, 1974) is an up-to-date and comprehensive treatment of decision analysis and its applications. Replete with examples of applications of decision analysis in business contexts, the text is fairly easy reading with only very elementary mathematics used. Howard Raiffa's book, Decision Analysis: Introductory Lectures on Choices Under Uncertainty (Addison-Wesley, 1968) was the first comprehensive treatment of the theory and application of decision analysis. Somewhat more mathematical than the Brown, Kahr, Peterson book, Raiffa's text is still an excellent introduction to the subject matter. A non-mathematical treatment of decision analysis can be found in Analysis of Decisions Under Uncertainty by Robert Schlaiffer (McGraw-Hill, 1969). Schlaiffer's book develops all the key ideas of decision analysis logically and intuitively without recourse to the mathematical underpinnings. A much shorter and readily-comprehensible treatment of decision analysis is to be found in Dennis Lindley's book, Making Decisions (Wiley, 1973). For the student desiring a varied motivational approach to the subject matter, Howard Raiffa has prepared an excellent self-instructional program that uses lectures, discussions, readings, examples, and exercises. Presented in the form of audio cassettes and supporting printed materials, the course entitled Analysis for Decision Making: An Audio-graphic, Self-Instructional Course, is available from the Encyclopedia Britannica Educational Corporation.

ACKNOWLEDGMENTS

This handbook was prepared by decision analysts on the staff of Decisions and Designs, Incorporated. Contributing authors were Scott Barclay, Rex V. Brown, Clinton W. Kelly, III, Cameron R. Peterson, Lawrence D. Phillips, and Judith Selvidge. The manuscript benefitted enormously from the editorial attention of Michael L. Hays with splendid assistance from Theresa A. Cunningham.

Special acknowledgment is due the Defense Advanced Research Projects Agency (ARPA) and the Office of Naval Research whose funding support made possible the preparation of this handbook. These agencies have supported and managed an insightful program of research that has significantly advanced decision-aiding technology and planted the seeds of fundamental change in the way major national security decisions are made. For their role in initiating, nurturing, and guiding this research program, particular credit is due Col. Austin W. Kibler, Dr. Robert A. Young, and Lt. Col. Roy M. Gulick of the Defense Advanced Research Projects Agency, and Dr. Martin A. Tolcott, Director of Engineering Psychology Programs, Office of Naval Research.

CHAPTER 1

DIAGRAMING AND SOLVING DECISION PROBLEMS

Complex decision problems can be difficult to resolve for a variety of reasons. Frequently, options are not clearly d ed, the results that might be achieved by opting for one choice over another may be highly uncertain, and it is often difficult to determine relative preferences for the possible decision outcomes. Certainly, almost everyone has encountered decision problems characterized by such uncertainty. Usually, the reaction is either to devote more thought to the circumstances than would normally be afforded, or to resort to various devices to help sort out the decision such as listing pros and cons for each option, rank ordering preferences, listing the things that could go wrong, and so on. In either case, whether through extended contemplation of the problem or through recourse to more explicit written aids, the person with the problem, the decision maker, attempts to lend structure to the problem to reduce it to a more explicit, tractable form. In a much more systematic and formal way, that is exactly what decision analysis helps him do.

Although there are many alternative ways to structure decision problems, decision analysis builds upon four basic elements that are inherent in any decision problem. These elements lead to a rather natural way to conceptualize and resolve complex decisions.

The first element is a set of initial courses of action. You have a decision to make only if you face a choice among alternative possible acts. Each of the choices you want to consider should be made explicit.

Second, one needs to consider the possible consequences of each initial act. What are the important things that can happen that will make one act more valuable or worth more than another act? Relevant sequences of subsequent events and follow-up acts must be identified for each initial act.

The third element is concerned with how attractive or unattractive each possible consequence of each act is to you. What is its value to you? How undesirable is one outcome compared to others which might result from the same or another decision? This value could be measured in terms of money, utility, or some other carefully defined index.

Finally, how likely is it that a particular act will result in each of the consequences? This uncertainty may be measured either by a numerical probability from 0 to 1 or in the form of odds.

Decision-analytic methodology provides a way to quantify, organize, and trace the logical implications of the decision elements defined above. The primary object of decision analysis is to model the decision, or some part of it, so that at least some of the implications can be deduced. By the verb "model," we mean to represent in a quantified form. The resulting model may be complete in the sense that it incorporates all of the decision maker's relevant perceptions of alternatives of value or uncertainty. In that case, if it is an accurate model, it will determine exactly what course of action should be taken.

On the other hand, it may be, and usually is, the case that it is a partial model; it displays the logical implications of some, but not necessarily all, of the relevant inputs. It may, for example, simply display the probable monetary consequences for selected choices, leaving non-monetary consequences and other choices for later consideration. In that case, it is not necessary for the initial course of action that is favored by the model to be the one that is decided upon. For example, it may be found that the monetary consequences that favor one course of action are far outweighed by the non-monetary criteria that favor another action.

For the present we shall focus on complete decision analytic models because they are clearest for purposes of exposition. Later, we consider some of the shortcuts that are necessary to reduce the task of a decision analysis to manageable size.

The Structure of Decision Problems: Decision Diagrams

It is a central tenet of decision analysis that all relevant considerations in a decision can be assigned to one or another of the four components: initial options, possible consequences, values, and uncertainties. In addition, they can, in principle, be represented fully in a decision diagram. In other words, for every conceivable decision, it is theoretically possible to construct a decision diagram which captures everything a decision maker feels is relevant to the choice in question.

A decision diagram is a kind of road map. Like a road map, it visually displays possible destinations (outcomes) at which you may arrive if you take one or another of the routes (initial acts) immediately available to you. It also shows you what you may come across on the way (events) and what later choices (subsequent acts) you may have to make.

A decision diagram consists essentially of a network of branches corresponding to possible sequences of acts and events, fanning out from an origin at the left to a time horizon at the right. Acts are available choices. Events are possible occurrences which are partly or completely outside the decision maker's direct control, though the chance of one of them happening may be influenced by acts that were carried out earlier. The diagram is made up of a concatenation of forks which are either act forks or event forks. A path through the diagram corresponds to a possible sequence of acts and events characterized by a value assigned to the consequence (but we will deal with matters of value later).

The decision diagram graphically distinguishes acts from events. Act forks are represented by squares, and event forks are represented by circles, as shown in Figure 1-1.

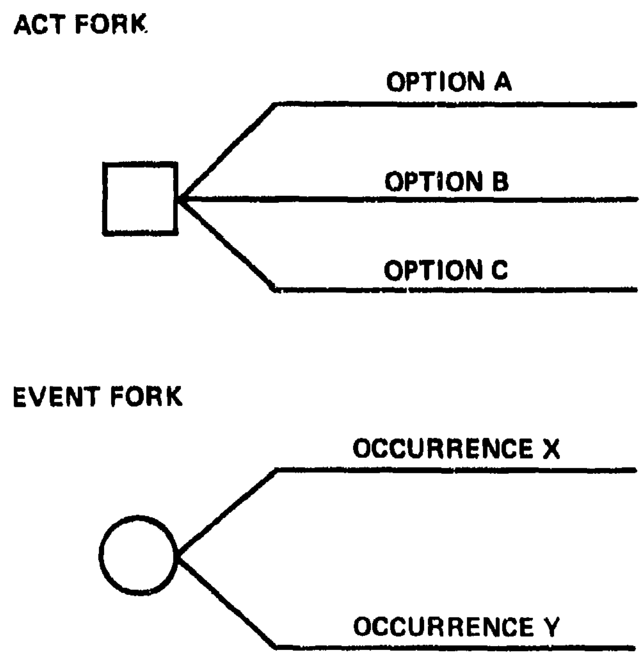


Figure 1-1
REPRESENTATIONS OF ACT AND EVENT FORKS

Although many examples throughout this text will alternate act forks and event forks, it is not necessary that acts and events alternate in sequence. By definition, however, the decision diagram must begin with an act fork at the point of origin at the left and end with an event circle at the termination of each branch at the time horizon on the right.

To see how a decision diagram is developed, let's consider a simple example. Suppose as a defense contractor, you must decide whether or not to bid on Contract A. To represent this situation, an initial act fork is required:

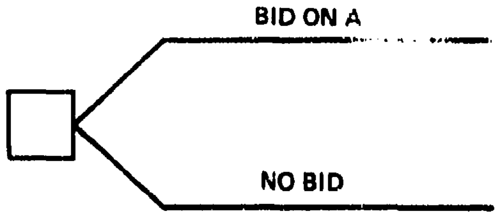


Figure 1-2
INITIAL ACT FORK

The possible outcomes of bidding on A are winning the bid and losing it. To represent this situation, an event fork is drawn at the right end of the bid-on-A line:

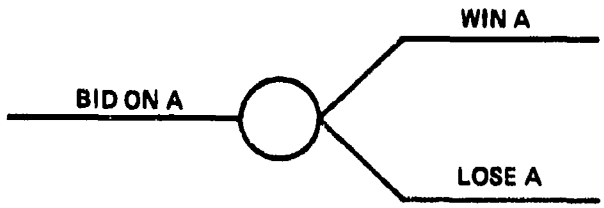


Figure 1-3
ACT A WITH EVENT FORK ADDED

At this point, the decision diagram is rather small and simple, as shown in Figure 1-4.

Suppose further that, after deciding whether or not to bid on A, you must decide whether or not to bid on Contract B. This situation is represented by an act fork like the initial act fork in Figure 1-2. However, since a bid on B is independent of your initial decision, this act fork is equally applicable to three situations: bidding on A and winning, bidding on A and losing, and not bidding on A at all. To represent this situation, a larger and more complex decision diagram is required, as shown in Figure 1-5. Needless to say, the possible outcomes of bidding on B are winning the bid and

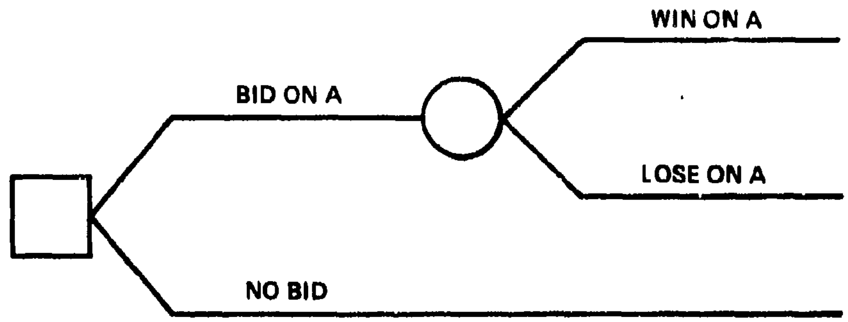


Figure 1-4

STRUCTURE OF BIDDING DECISION

losing it, which outcomes may be represented by an event fork analagous to that shown in Figure 1-3.

Clearly, we can continue this process of extending the decision diagram; we can add more contracts or vary the event forks of possible consequences of winning or losing additional contracts. Note that the branch representing a bid on A alternates act forks and event forks, but that the branch representing no bid on A does not alternate act and event forks. We do not show an event fork following an event fork, but it is easy to imagine events like merger, bankruptcy, and the like following the outcome of winning or losing a bid on B.

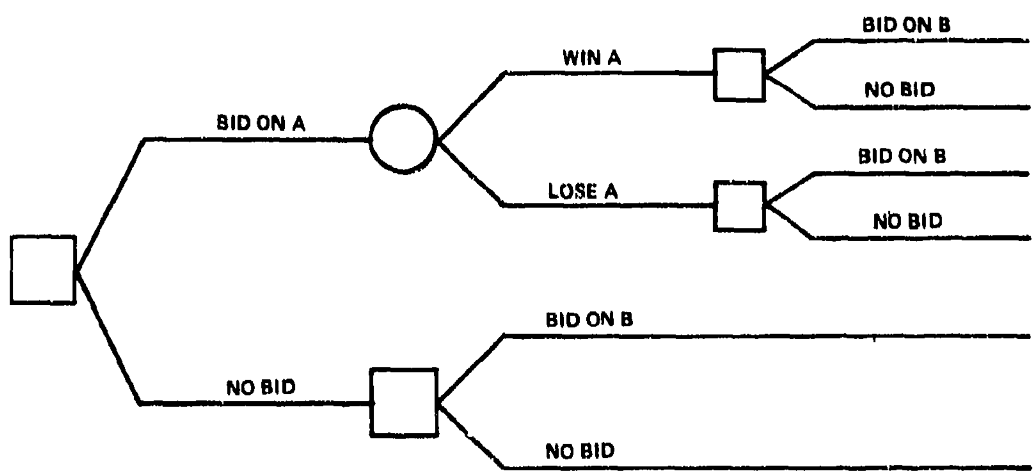


Figure 1-5

STRUCTURE OF BIDDING DECISION
WITH SECOND ACT FORK REPRESENTED

The decision diagram in Figure 1-6 reflects the structure of the bidding situation we originally described, in which the contracts were unrelated. But the technique of constructing decision diagrams is flexible enough to cover other, more

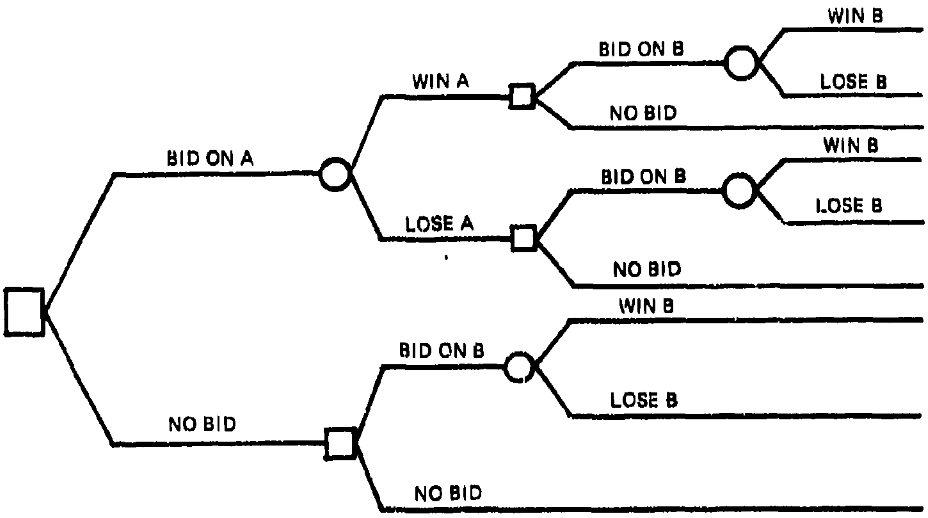


Figure 1-6

BIDDING DECISION STRUCTURE FOR INDEPENDENT BIDS

complex situations. Suppose, for example, that two different government policies establish different contract situations, especially in terms of the relationship between bids on contracts A and B. On the one hand, it may be stipulated that whoever wins A must bid on B. The stipulation may be based on the belief that the winner of A is in the best position to do the work required by B. This situation is represented in the decision diagram shown in Figure 1-7. Alternatively, it may be specified that whoever wins contract A may not bid on contract B, a stipulation based on a policy of distributing government contracts throughout an industry. This modified decision structure is shown in Figure 1-8.

In developing the decision diagrams for the situations represented in Figures 1-6 through 1-8, we followed two rules governing the construction of the diagrams. The two rules are:

RULE # 1: ACTS APPEAR ON A DECISION DIAGRAM AT THE MOMENT WHEN THE DECISION MAKER MUST IRREVOCABLY SELECT ONE OPTION AMONG SEVERAL OR COMMIT RESOURCES; and

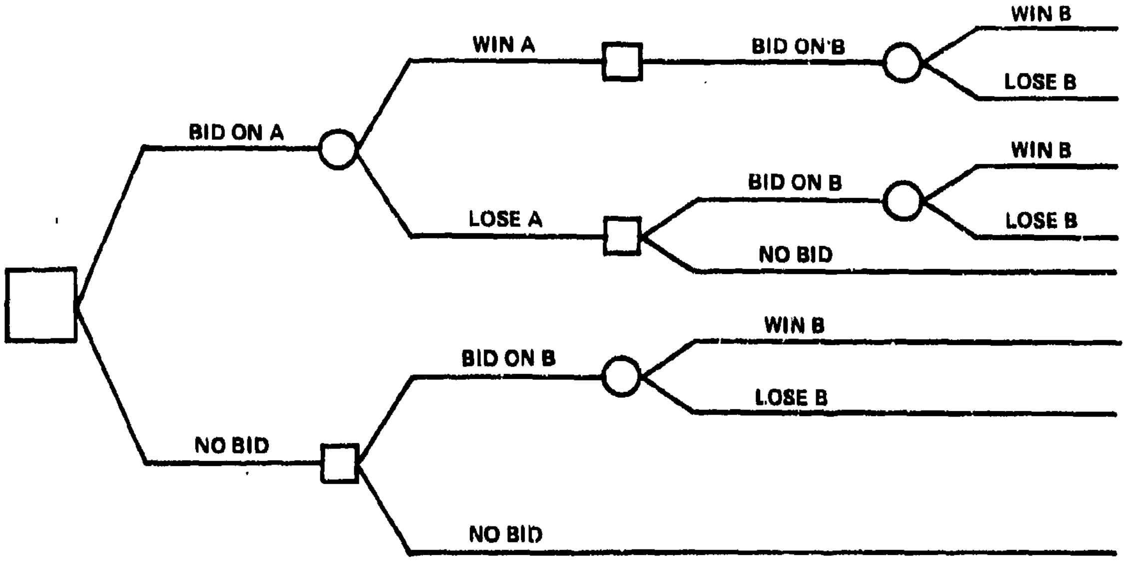


Figure 1-7

DECISION STRUCTURE WHEN WINNING A REQUIRES BID ON B

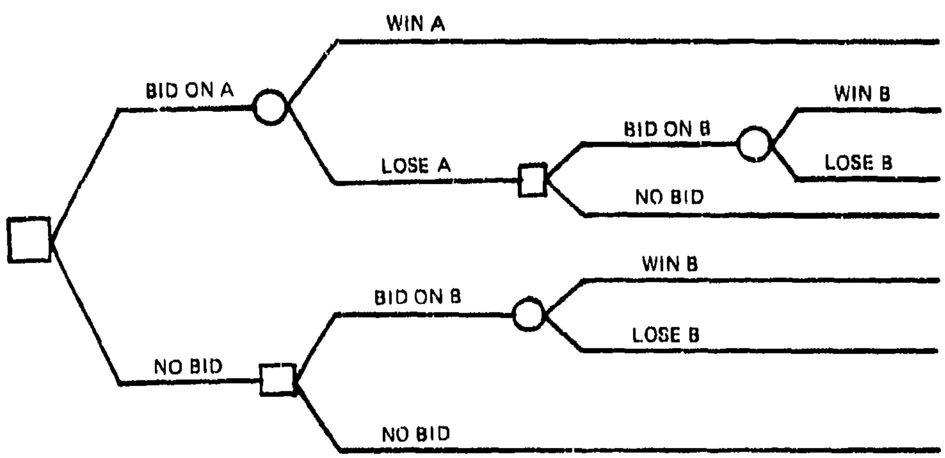


Figure 1-8

DECISION STRUCTURE WHEN WINNING A PRECLUDES BID ON B

RULE #2: EVENTS APPEAR ON A DECISION DIAGRAM AT THE MOMENT WHEN THE DECISION MAKER ACQUIRES KNOWLEDGE OF THEM.

Rule 1 directs us to represent actions in the decision making process when they are, in fact, made, not when they are contemplated. Rule 2 directs us to represent those events in the decision-making process when they become known to the decision maker. For only the decision-maker's knowledge of events, not the prior occurrence of the events themselves, can affect his decision-making process.

Thus far, the decision diagrams we have considered have depicted only the possible acts and events and shown how these relate to each other in a decision situation. Put another way, they have shown the structure of the sample decision problems. While the representations would help a decision maker to see at a glance his alternatives and identify those things that might affect any choice to be made, they do not yet address the central question: which choice should be made? That question cannot be answered from these skeletal diagrams since important information about the value of the possible outcomes and the likelihood of occurrence of events has not yet been considered. Once this information is incorporated into the decision diagram, a matter we turn to in the next section, we shall have completely represented the essential elements of the decision.

Before moving on, we should emphasize a number of points about the discussion thus far. First, it should be understood that for purposes of exposition we have used very simple examples. Real-world decisions of a size and complexity meriting analytic attention would be much more complex than the simple examples we have used. Since the diagrams of our simple examples already extend across a page, the thoughtful reader may well wonder about the tractability of more complex problems. It is enough to say at this point that such concern is appropriate. Fortunately, there are shortcuts that reduce very complex decision problems to manageable proportions. A second point to bear in mind is that the process of laying out the structure of a decision problem (diagraming the problem) is the most important and perhaps most difficult step in decision analysis. If the structure does not accurately represent the factors bearing on a decision, the resolution of that decision problem will be degraded.

Representing Values and Probabilities in Decision Diagrams

As was pointed out in the previous section, there are four essential elements in a decision: initial options, possible consequences, values, and uncertainties. In our discussion thus far, we have covered the way decisions are structured to depict two of the elements: initial options and the possible consequences of each initial option. Now we shall turn to the remaining elements, values and uncertainties, and show how these are incorporated into our decision representation. For convenience we will use the bidding example, variants of which we diagramed in the preceding section.

In considering the decision of whether or not to bid on contract A and subsequently to bid on contract B, let us assume that there are a number of relevant cost and value dimensions that can be calculated or estimated to help with the decision. Let's say you have calculated that it will cost $10,000 to prepare a bid on contract A and that if you win that bid, you will gain $50,000. Suppose further that to bid on contract B will cost $10,000, just as the bid on A did, but that since, if you win B after winning A, producing for both contracts will stretch your capacity to the point where you incur excessive costs, your profit on B alone would be only $20,000. Tracing through the uppermost path of the decision diagram shown in Figure 1-9, you can see these values reflected as follows: bid on A, -10; win A, +50; bid on B, -10; win B, + 20. If you do not bid on A, but do on B, your costs will be modest, leading to a $50,000 profit on B. The sequence of values is: no bid on A, 0; bid on B, -10; win B, +50. Corresponding values are shown on the other branches of the diagram as well. The diagram now reflects three of the elements essential to decision; initial options, possible consequences, and values.

The numbers shown at the right-hand side of the diagram are the path values. For any specific path through the diagram, the values and costs involved in achieving the indicated outcomes, if they occur for certain, will result in the path value shown. Tracing through the uppermost path of the diagram, we add or subtract values shown below each branch to obtain the path value:

-10[bid on A cost] + 50[win A] -10[bid on B cost] + 20[win B] =50.

The values for the remaining paths are calculated in a directly comparable manner. Note that in calculating these values, the probabilities of achieving the various outcomes did not enter into the calculation. Those will come into play in the next section when we discuss how to solve a decision diagram to determine the best course of action.

To complete the representation of the decision problem we need to reflect in the diagram the degree of uncertainty about the events which we cannot control, that you will win either or both of those contracts if you should elect to bid on them. From prior experience or perhaps from an intangible "sense of the situation," you judge that you are equally likely to win or lose the bid on contract A (probability .5 win, .5 lose).

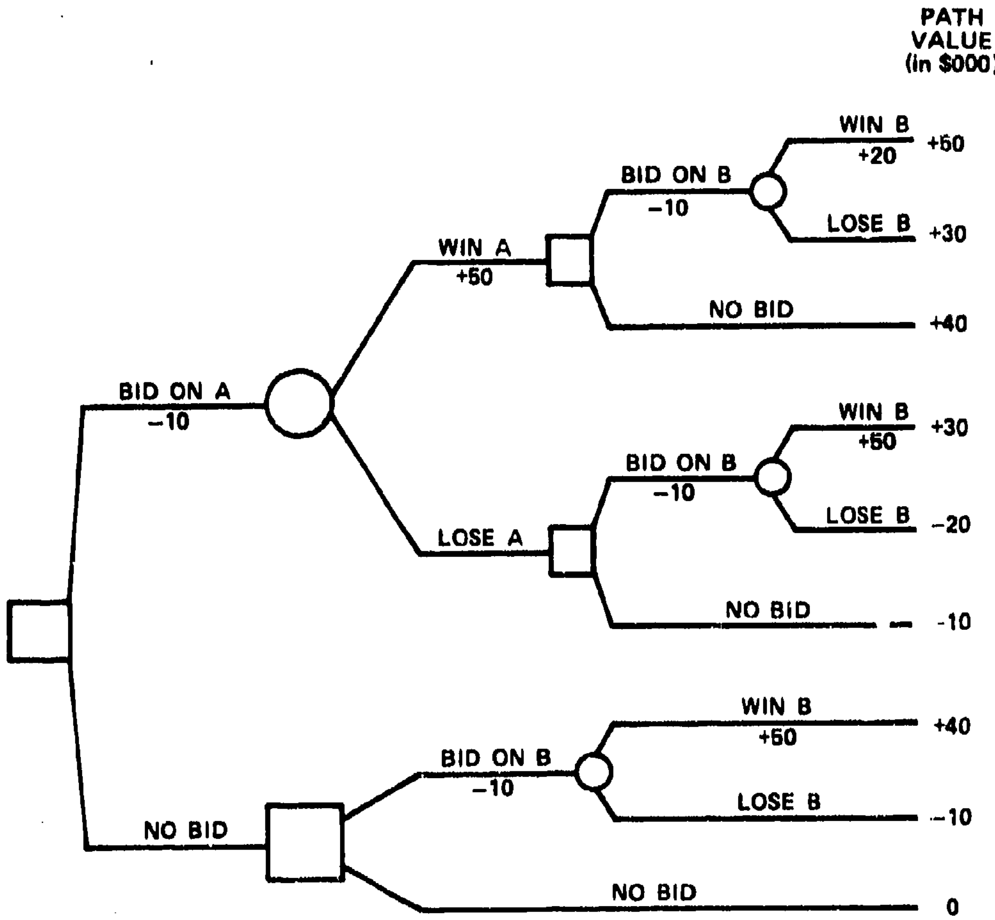


Figure 1-9
TREE FOR BIDDING EXAMPLE

If you win the bid on A, you judge your chances with respect to contract B as .4 win and .6 lose. On the other hand, if you lose the bid on contract A, you assess your probability of winning a bid on B to be .7 and the probability of losing the bid on B to be .3. These probability numbers are shown entered in our bidding diagram in Figure 1-10. Tracing through the "Bid on A" path of the diagram, you can see how these uncertainties (probabilities) are represented.

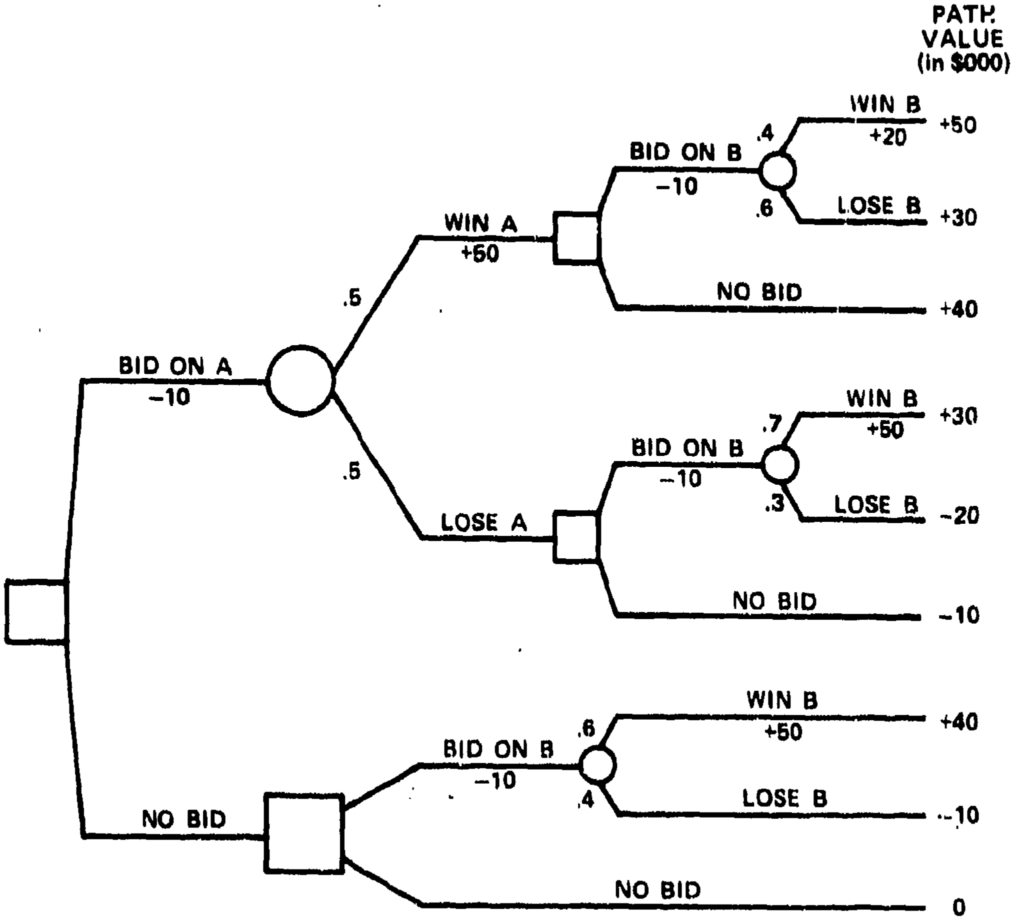


Figure 1-10
TREE FOR BIDDING EXAMPLE

Solving a Decision Diagram

We have seen how decision diagrams are constructed and shown how information relevant to the decision is represented in the diagram. In principle, such a diagram can be a virtually complete translation or model of your perception of the decision problem. If it is, you can be satisfied that the "best" decision in the problem model is also the best decision in your real world situation. Then you have only to determine which is the "best" decision. How is this determination to be made; that is, which immediate option is the most attractive?

The value of outcomes has been previously introduced as a measure of the attractiveness of a possible outcome of a decision option when that outcome is treated as certain. If all outcomes in decision situations were in fact certain, we would need go no further to solve our decision problem. We would need only to diagram the problem, calculate the value of each outcome and settle upon the choice that leads to the outcome with the highest value (or lowest cost). When uncertainty enters the picture, however, as we have noted in the event nodes of our bidding example in Figure 1-10, the value of an outcome is seldom the sole criterion for choosing among decision options. You know without thinking, for example, that you would not decide in favor of a gamble that offered you one chance in 100 of winning $10.00 over one offering you a 90% chance of winning $9.00. In these hypothetical gambles, the first, winning $10.00, has a higher outcome value than the second, winning $9.00. You intuitively would take the second choice, the lower outcome value, out of a sense of the likelihood of winning. What you have done implicitly is to weight the outcome values by the probability of achieving that value. As we shall see below, the solution of decision diagrams involves determining weighted values that accrue to alternative acts (decision options) in a decision problem. There are a variety of ways of making these determinations. In the remainder of this chapter we shall present but one method of solving decision diagrams. The method involves the calculation of weighted values, often called expected values, for each decision option. This technique is the simplest and statistically most straightforward method that can be used and is thus useful for demonstrating the general principles involved in solving a decision problem. The reader is cautioned, however, that frequently there are decision circumstances wherein an expected value solution to a decision problem may not be an optimal one. In these cases, alternative means of treating value are required. These alternative methods are covered in Chapter 2.

Inferring Act Values By Folding Back a Decision Diagram

Determining a value for each immediate act in a decision is done by a procedure called folding back the decision diagram. This is a process of substituting values for each decision fork or act fork, beginning at the right-hand side, or time horizon, of the diagram. The process of substitution is continued, selecting the better of higher-valued acts at each act fork subsequent to the initial decision, until the decision diagram is simplified to the initial act fork. The act with the highest value is the one that should be the "best" decision in any given situation.

One way of folding back a decision diagram, suitable for many decision situations, is to fold back using expected value. The expected value of an event fork is the sum of the values produced by multiplying the value of each possible outcome at that event fork by the probability of the occurrence of each outcome. The essential idea of folding back by expected value substitution can be illustrated by a simple hypothetical game (we shall return to our bidding example later).

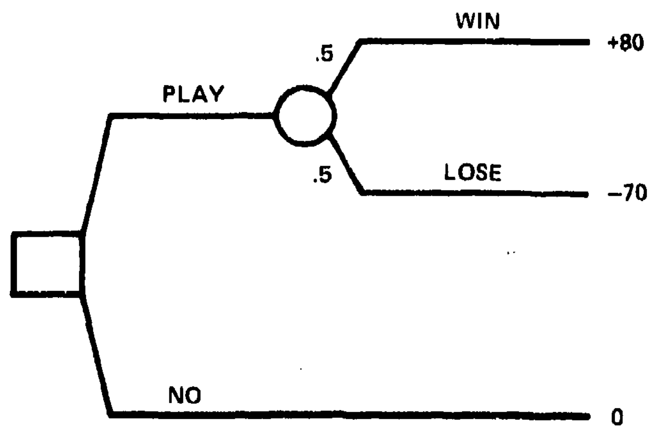
Suppose your decision is whether to pay $20 to get into a game where, if the coin comes down heads, you win $100, and if it comes down tails, you lose $50. If you play, the values of the possible outcomes are +$80 if you win, ($100 win - $20 cost of entry = $80) and -$70 if you lose (-$50 loss + -$20 cost of entry = -$70). A no-play decision has zero value for the outcome. The probabilities of heads or tails are of course even, or 50-50. Your decision options are fully represented by the decision diagram shown in Figure 1-11A.

The expected value of the decision to play is the value of winning times the probability of winning ($80 x .5 = $40) plus the value (or cost) of losing times the probability of losing (-$70 x .5 = -$35), or $5. Figure 1-11B shows the equivalent diagram after folding the original diagram back. It differs from Figure 1-11A only in that a value of +$5 appears in place of the win/lose fork and in the double bars across the no-play path. If you are satisfied that Figure 1-11B is indeed equivalent to Figure 1-11A in terms of the relative attractiveness of the play/no-play decision, you know that solving the equivalent diagram is the same as solving the original diagram. The solution says you would be indifferent about winning $5 for sure, and having a 50-50 chance of winning $80 or losing $70. Since winning $5 for sure is better than winning nothing, the solution to the equivalent diagram is to play; therefore, the preferred decision in the real problem is also to play. The double bars across the no-play path are a graphic convention indicating a rejected act branch.

Folding Back a More Complex Tree

Let's now return to a more complex example, the bidding problem which we developed in earlier sections and which is reproduced in Figure 1-12. To recapitulate, the contractor must decide whether or not to bid on contract A, which will cost $10,000 if he bids. He assesses his probability of winning the contract as 0.5, with a gain of $50,000. He has the option at a later time of bidding on contract B, at the same cost. If he has not bid on contract A, the probability of winning contract B is 0.6, and the profit is also $50,000. If he bids and wins contract A, then he assesses the probability of winning contract B to be 0.4. In this case, because of excessive production costs for both contracts, the profit

A. Original Diagram



B. Equivalent Diagram

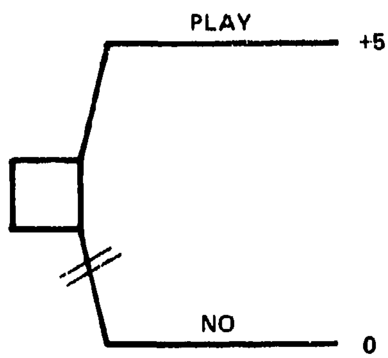


Figure 1-11
FOLDED-BACK DIAGRAM

on contract B will be only $20,000. Figure 1-12 is the complete decision diagram for solving the problem.

The first step in folding back is to eliminate the time-horizon event forks on the right of the diagram by substitution of expected values. Using the same technique as before, we can start at W, the top fork on the diagram, and compute the expected value of a 40% probability of making

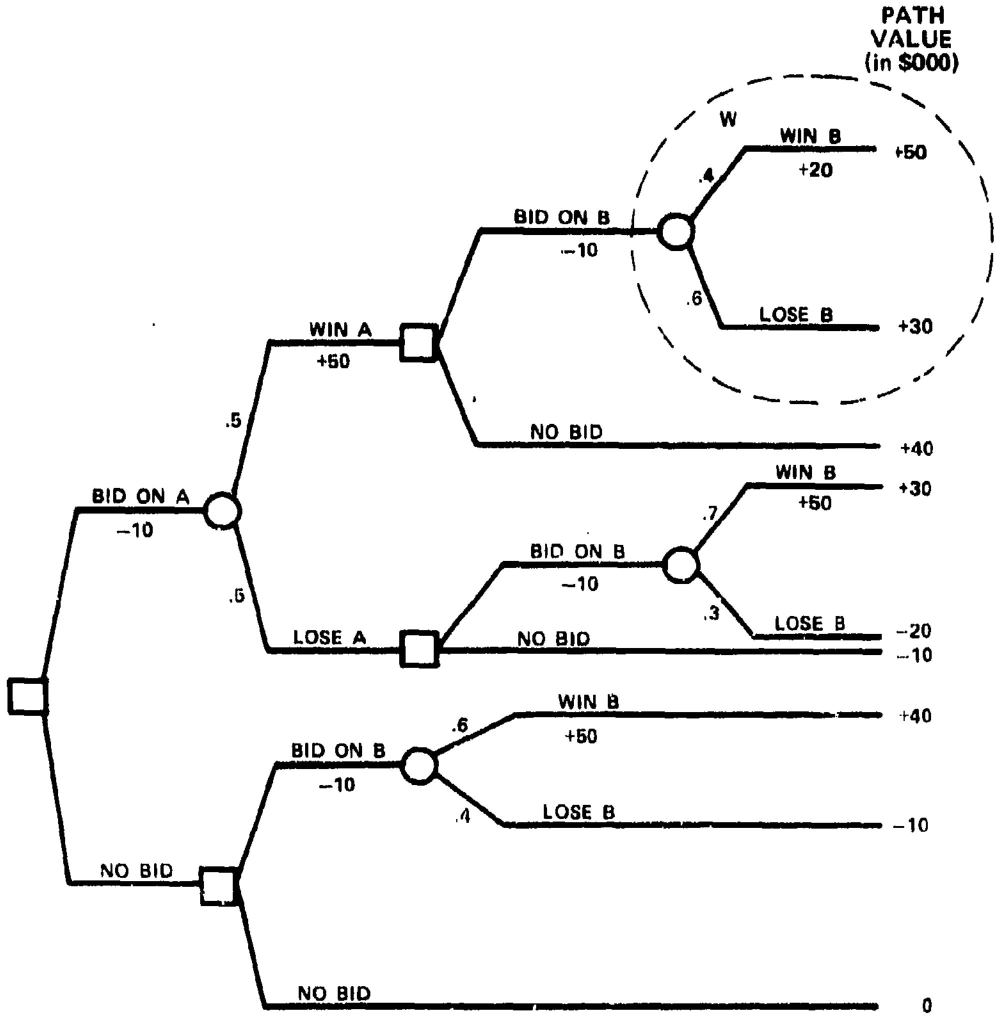


Figure 1-12

COMPLETED DIAGRAM FOR BIDDING DECISION

$50,000 (+50 on the diagram) and 60% probability of making $30,000 (+30). This turns out to be +38. Now we can eliminate the fork at W, and substitute the expected value of +38, as shown in Figure 1-13. The new diagram represents a situation where the attractiveness of the immediate acts is exactly the same as in the decision reflected by the original diagram (Figure 1-12). All we have done is substitute for the original gamble, depicted by the event fork at W, the expected value of that gamble.

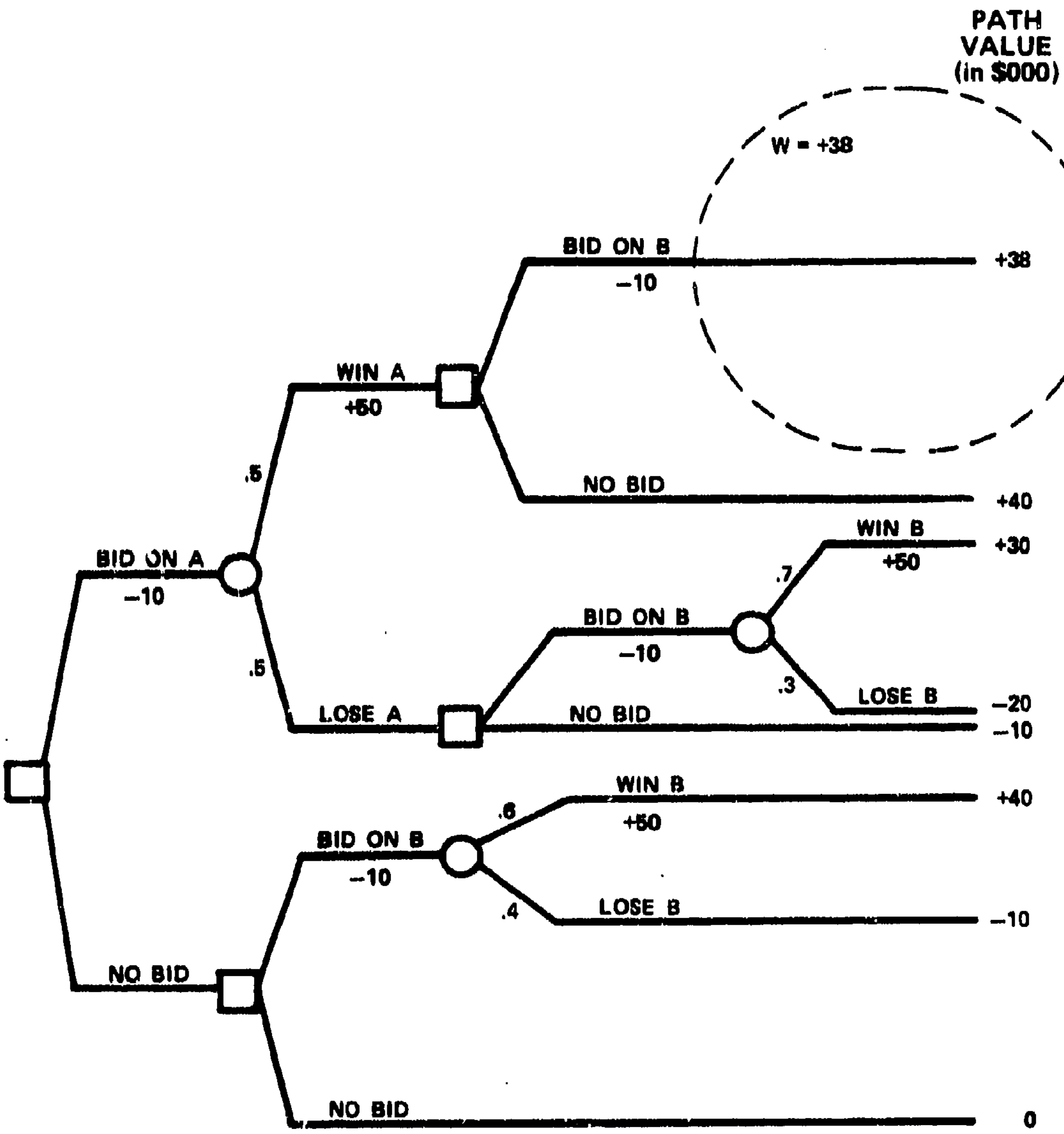


Figure 1-13

FIRST SUBSTITUTION ON BIDDING DECISION DIAGRAM

Proceeding in the same way, we can substitute expected values for each of the other two horizon forks, as shown in Figure 1-14A. It is apparent that the decision problem represented by Figure 1-14A should be simpler to analyze than the original problem of Figure 1-12. By continuing the process of substitution, we can make it simpler still.

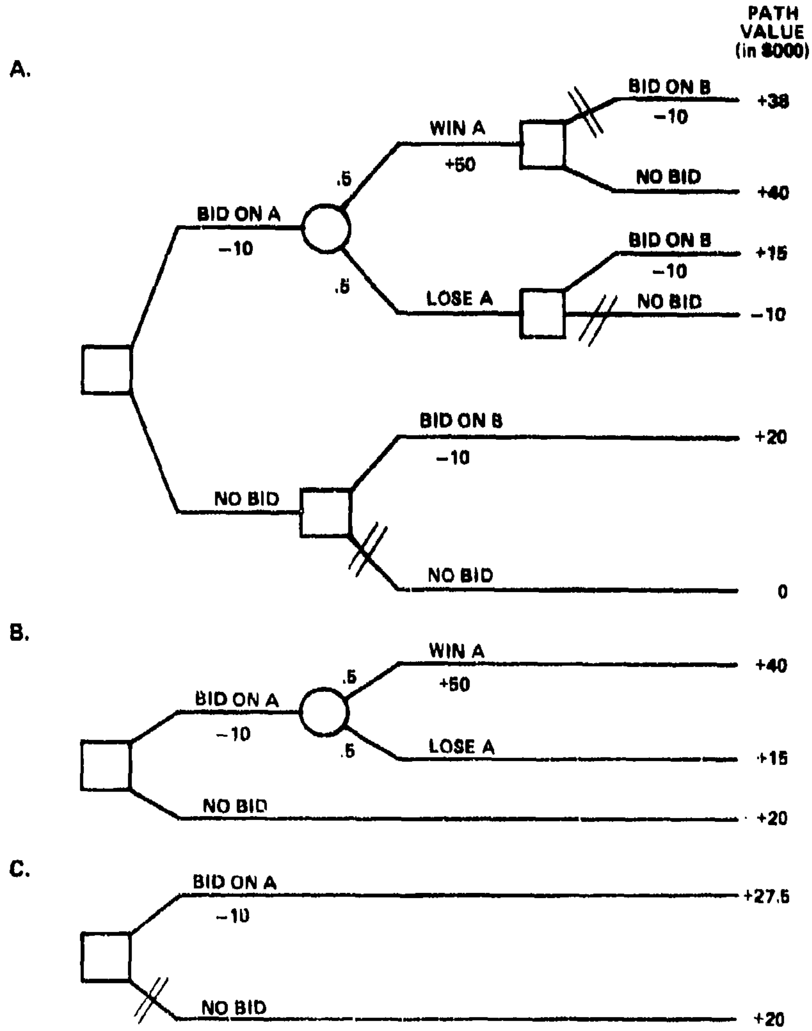


Figure 1-14

PROGRESSIVE SUBSTITUTION OF DIAGRAMS IN BIDDING DECISION

In Figure 1-14A, the forks on the right are now all act forks with the indicated expected values as shown. The next stage of the substitution is very straightforward since we should choose the act with the highest expected value. Taking the top fork of Figure 1-14A, the choices are to bid on B, with an expected value of +38, and not to bid, which is worth +40. Since we would pick +40, the other option is rejected, as shown by the double bars. Proceeding similarly with the other two act forks, we can now substitute Figure 1-14B for Figure 1-14A. This is the result of substituting for all act forks the value of the better act.

The final reduction is made by repeating the process of calculating an expected value. The event fork in Figure 1-14B is replaced by the expected value of the 50-50 chance of winning A(+40) or losing A(+15), which is +27.5. This results in the diagram of Figure 1-14C, and we are left with a single act fork, with one act valued at +27.5 and the other at +20. By a progression of substitutions, it is clear that we should bid on A, and the folding-back process is completed.

Folding back by substitution of expected values and choosing the set with the higher expected value is quite routine and can be easily computerized. As we have seen, the best initial decision is easily discernible from the simplified diagram resulting from the folding-back process. But what of the subsequent acts or decisions in our original diagram? It is often more useful, both from a compactness standpoint and for a complete picture of the problem, to retain the complete diagram and show the values computed for the event forks, those chosen for the act forks on the basis of the better act, and the rejected acts. Figure 1-15 shows the complete diagram with all the substitutions. The same immediate act is preferred, bid on A.

Solving a decision problem on the basis of maximizing expected value is often called a risk-neutral strategy. If you follow it, you will be indifferent between a decision to do nothing and one which gives you a 50-50 chance of making or losing, say, $1,000,000. A person controlling very large resources may well have such a risk-neutral attitude toward decisions which represent relatively small gambles for him. You or I, as individuals, would probably not be risk-neutral in the face of a gamble which involved an even chance of winning or losing $1,000,000. We would more than likely be averse to this kind of a gamble. This is known as being risk-averse, and our decision, if it is to reflect our personal preferences, would then require a risk-averse strategy.

The insensitivity of the expected-value approach to the decision maker's attitude toward risk constitutes a limitation on the usefulness of the technique. While the use of expected

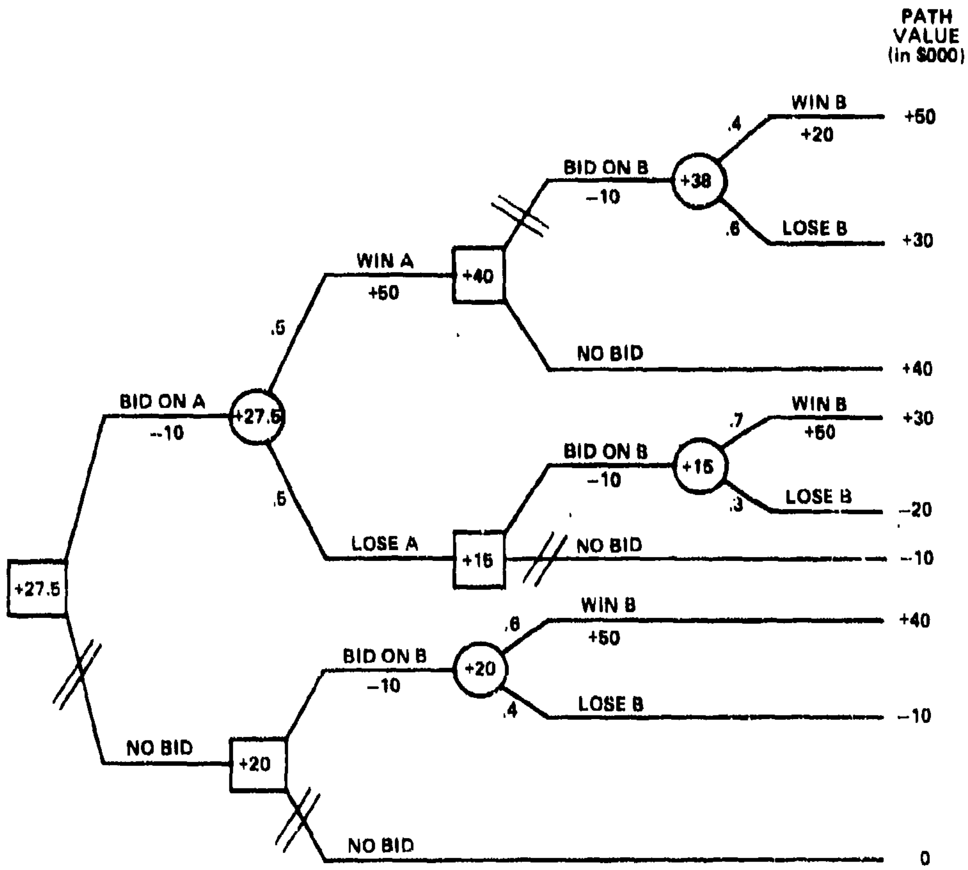


Figure 1-15
FOLDING BACK ON EXPECTED VALUE

value is quite appropriate and convenient in decision problems where factors of risk do not distort value considerations, the reader should be wary of the misapplication of this approach.

In circumstances where biasing effects of risks do come into play, there are other means of treating value which adjust for risk effects. There are techniques also for determining quantitative values for qualitative dimensions of worth and for combining multiple dimensions of outcome value into a single numerical expression of worth. These matters of value will be discussed in the next chapter.

CHAPTER 2

VALUE AND EVALUATION

In the preceding chapter, we discussed how to structure and solve simple decision problems. In the examples used, there was but a single dimension of worth assigned to the possible outcomes, their monetary value. In practice, it is rare to find a situation in which outcomes have only a single value dimension. Usually there are multiple dimensions of worth (or cost) associated with any outcome. Furthermore, as will be seen, many dimensions of value are qualitative; that is, there is no conventional metric to describe them. A military aid decision, for example, would be weighed not only in terms of quantifiable economic considerations but in terms of qualitative factors such as long- and short-term strategic advantage, the way in which economic or military aid would be perceived by allied nations, and domestic political considerations. And, as was discussed at the conclusion of the last chapter, there is a need to handle values in such a way that risk may be accommodated in the solution of a decision problem. Since decision-analytic methodology requires a single quantitative expression of worth for possible outcomes, how then does one handle situations having multiple dimensions of value, many of which are qualitative and all of which can be affected by the risk inherent in the decision? There are a number of ways of handling these complicating considerations, and all of them in one way or another use a notion of equivalent substitution. In this chapter, we shall present methods for taking account of risk in decision situations, for dealing with multiple dimensions of value, and for assigning values to qualitative dimensions of worth.

Folding Back by Substitution of Certainty Equivalents

Value for an option (or any event fork) can be thought of as the value (certain) which you would trade for the uncertain prospect that you face. In some circumstances, it is helpful to think of it as the selling price of the event fork treated as a gamble. The selling price of the gamble may be referred to as a "certainty equivalent," that is, an amount you would be just willing to accept for certain in lieu of the uncertain prospect represented by the gamble.

To illustrate how the use of certainty equivalents, which reflect your attitude toward risk, can affect your decision, let's look at the first example we used in Chapter 1 to illustrate the technique of folding back a decision diagram. Recall that you had an opportunity of paying $20 for a 50-50

Preceding Page Blank - Not Filmed

gamble between winning $100 and losing $50. As a result, either your net gain would be $80, or your net loss would be $70. The diagram in Figure 2-1A represents your decision option. Figure 2-1B differs from Figure 2-1A only in that a value of -$7 appears in the place of the win/lose event fork. What this difference means is that despite the expected value of $5 (and the chance of winning $80) you would be willing to give up, say, $7 rather than run the risk of losing $70. If the diagram is indeed equivalent as far as you are concerned, it says that you would be indifferent between losing $7 for sure, and having a 50-50 chance of winning $80, or losing $70.

Note that this certainty equivalent is not, and in general, does not have to be the same as the expected value of the gamble, which in this case, is +$5. If you are really averse to risk, you might be prepared to pay money to get out of a gamble like this one whose expected value is positive. Since losing $7 for sure is worse than losing nothing, the solution to the diagram in Figure 2-1B is not to play; accordingly, the preferred decision in the real problem is not to play either.

Substitution of certainty equivalents can be used in a more complex decision diagram such as our bidding problem shown again in Figure 2-2. Suppose that at W, you were indifferent between the gamble represented by the event fork and a value of +35 for sure. Substituting +35 for the event fork at W results in the simplified diagram shown in Figure 2-3. We have substituted for the prospect of the original gamble, the equivalent value (certainty equivalent) which is equally attractive, instead of the expected value of +38, which a risk-neutral strategy would dictate. Proceeding in the same way, you can substitute certainty equivalents for the other two horizon forks. You might be left with the diagram shown in Figure 2-4A. The values substituted would indicate that you are slightly risk-averse since the certainty equivalents values are somewhat lower than the expected values.

Since, in Figure 2-4A, the forks on the right are now all act forks with certain values, you can proceed as before by substituting for all act forks the value of the better act. The result is shown in the diagram in Figure 2-4B.

To make the final reduction, you simply repeat the process you started with. The act fork on the right is replaced by a certainty equivalent which reflects the attractiveness of a 50-50 gamble between +40 and +13. If +25 is a value which you find equivalent, that is the value substituted. In the simplified diagram of Figure 2-4C, you are left with a single act fork with one act valued at +25 and the other valued at +18; and it is clear which you prefer.

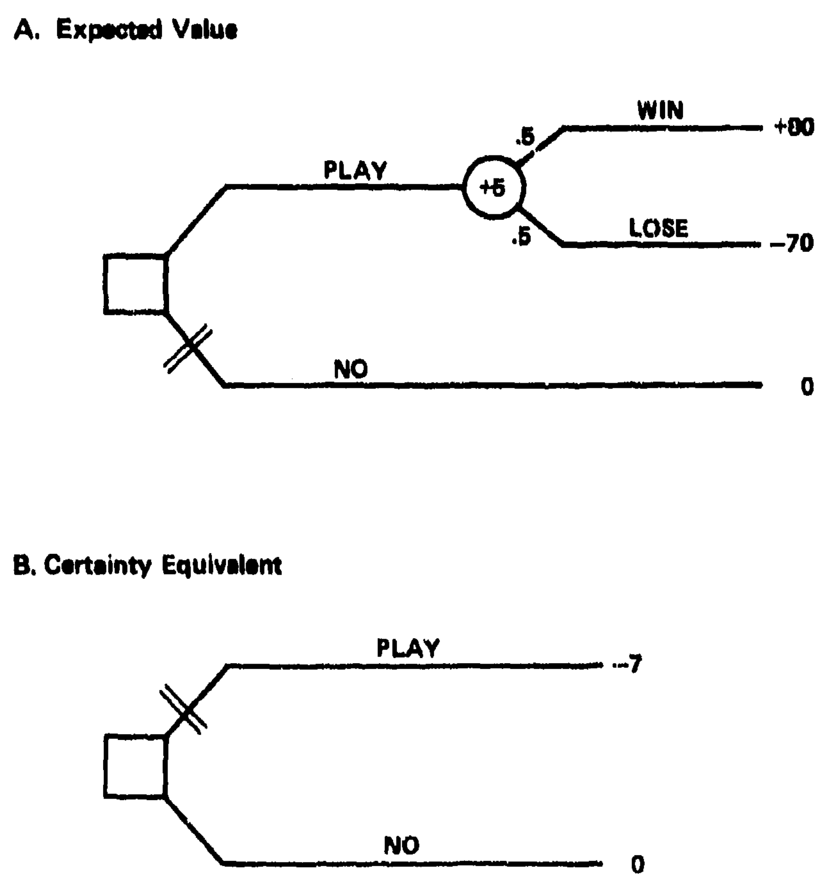


Figure 2-1
SUBSTITUTING A CERTAINTY EQUIVALENT FOR AN EXPECTED VALUE

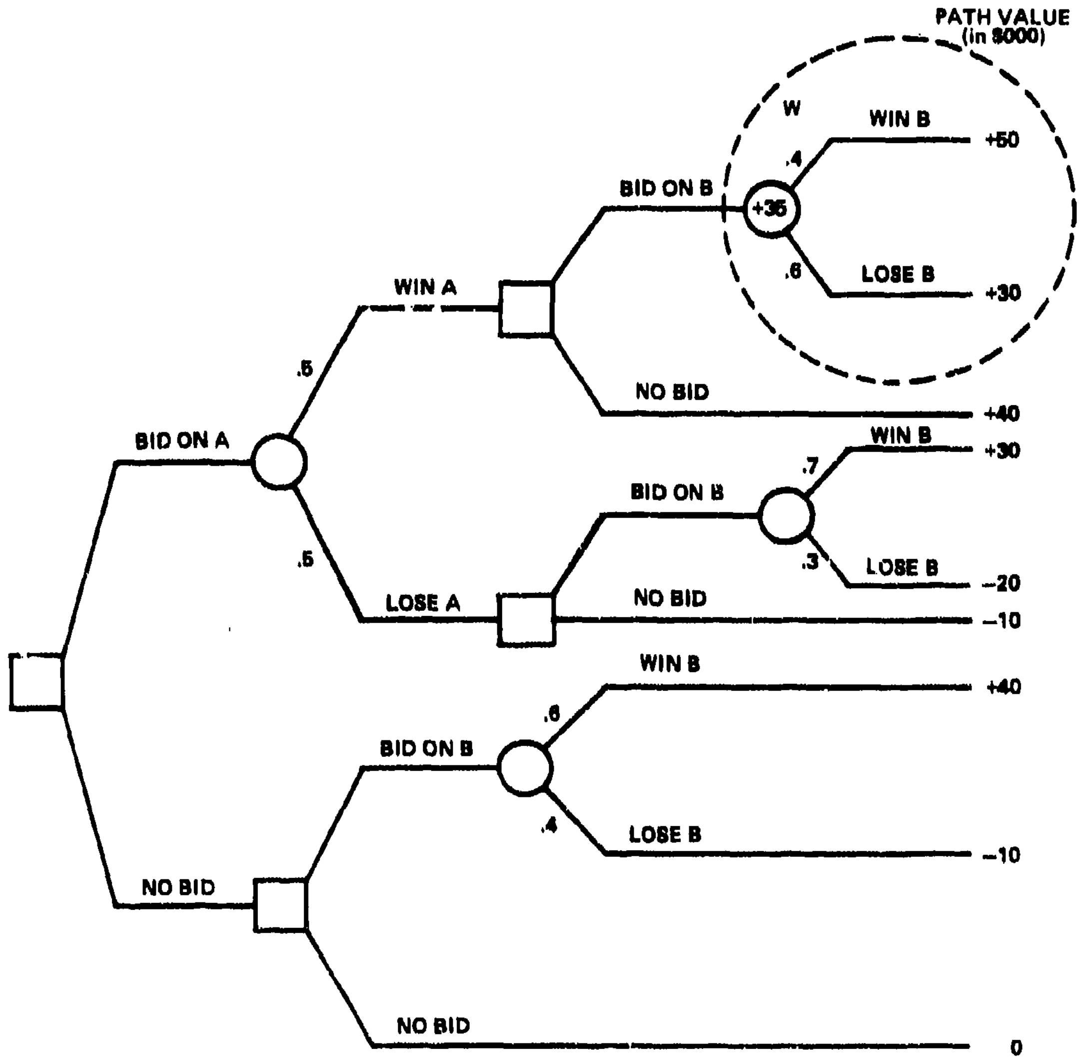


Figure 2-2
COMPLETED DIAGRAM FOR BIDDING DECISION

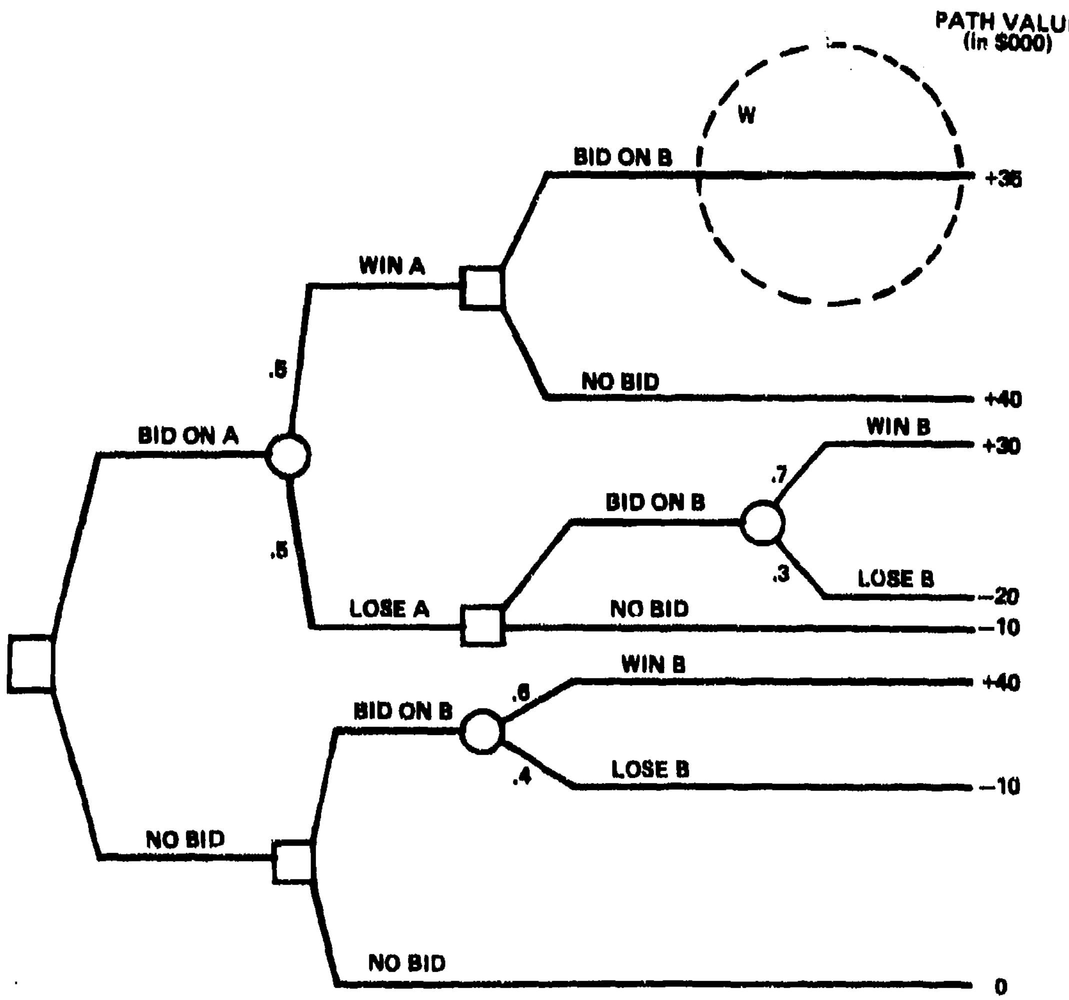


Figure 2-3
FIRST SUBSTITUTION ON BIDDING DECISION DIAGRAM

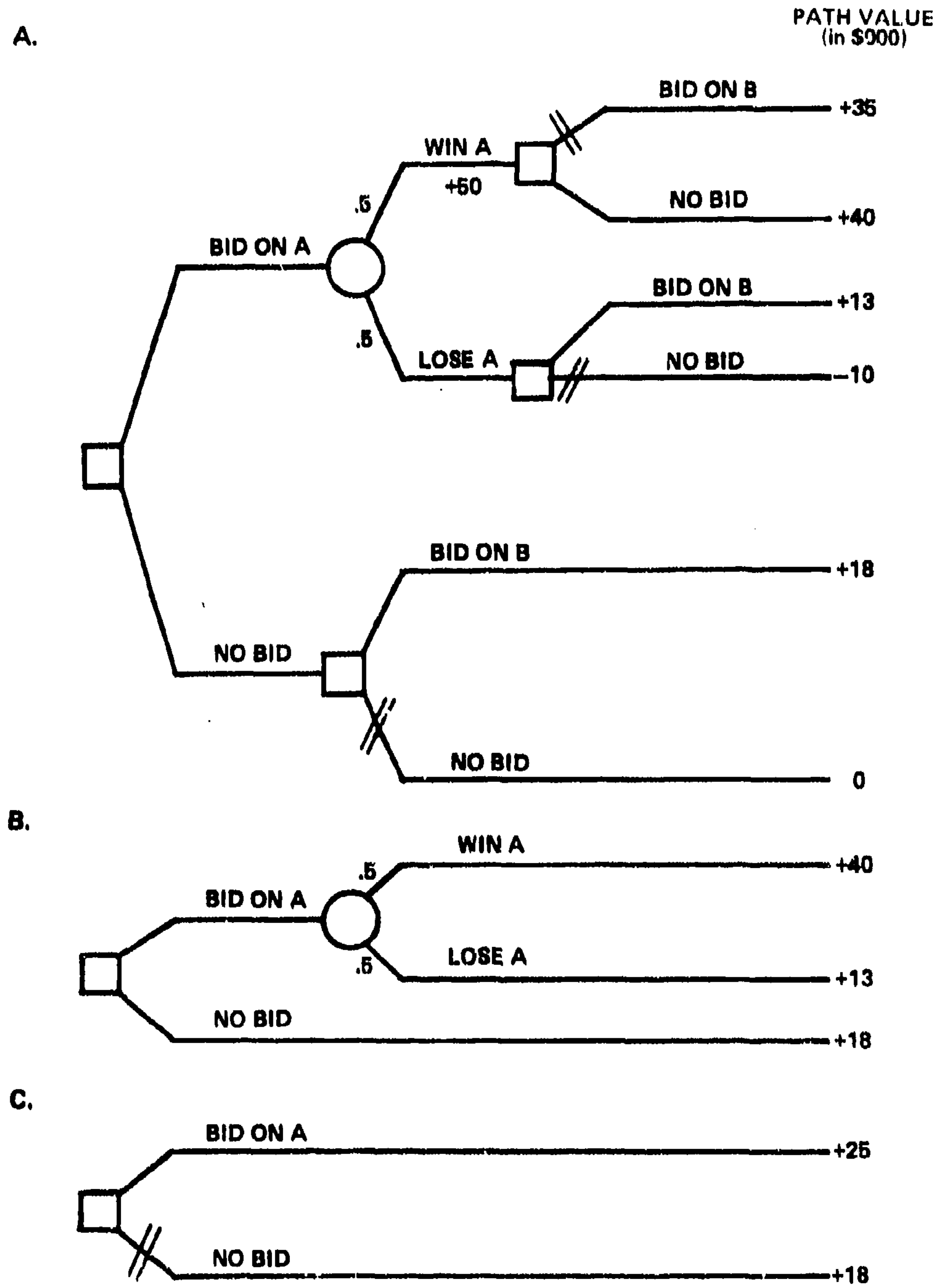


Figure 2-4

PROGRESSIVE SUBSTITUTION OF CERTAINTY EQUIVALENTS IN BIDDING DECISION DIAGRAM

Figure 2-5 shows the complete decision diagram, with the certainty equivalents you have used. The final valuations of the immediate acts (+25 and +18) are lower than those obtained by substitution of _expected_ values. The same immediate act is preferred (bid on A) and the contingent decisions are not changed. Although using certainty equivalents confirmed the preference obtained by using expected value in this case, the substitution of one for the other does not always yield the same preference. A very risky gamble might be preferred to a "sure thing" if you are risk-neutral, but a "sure thing" is preferred if you are risk-averse.

It may seem that the process whereby an equivalent value is substituted for an event fork is somewhat arbitrary, and, in a sense, it is. It is a measure of how _you_ personally feel about a gamble. It may be either higher or lower than the expected value that results from calculating a probability-weighted average of the possible values resulting from the gamble. If you are risk-averse, your certainty equivalent will be _lower_ than the expected value, as in the case above. On the other hand, if you are a risk seeker, the certainty equivalent will be higher.

While substitution of certainty equivalents has the handy property of taking into account the decision maker's attitude toward risk, there is a major limitation to its practical use. Requiring, as it does, node-by-node assessment of certainty equivalents, the method can become so laborious and time consuming in large decision problems as to be unmanageable. Fortunately, a method exists whereby one can achieve the same advantages offered by substituting certainty equivalents while avoiding the extensive labor of numerous assessments. It involves translating each of the original values on the diagram into a numerical expression of _utility_ and folding back the decision diagram based on _expected utility_ by the same operations previously presented for expected value.

Utility as an Index of Worth

It should be apparent from our discussion of certainty equivalents that the value of things is a personal, circumstantial matter and that it is these subjective values of things that are the basis for decision. In the case of money, for example, it is not necessarily the case that you would value a gain of $10,000 twice as highly as a gain of $5,000, nor is a loss of $1,000 necessarily twice as severe as a $500 loss. It depends on the decision maker's personal attitudes and circumstances, a dependency accommodated by the concept of utility.

Utility can be described as a subjective measure of "liking." It is a personal value reflecting how you subjectively value something. For example, for most families,

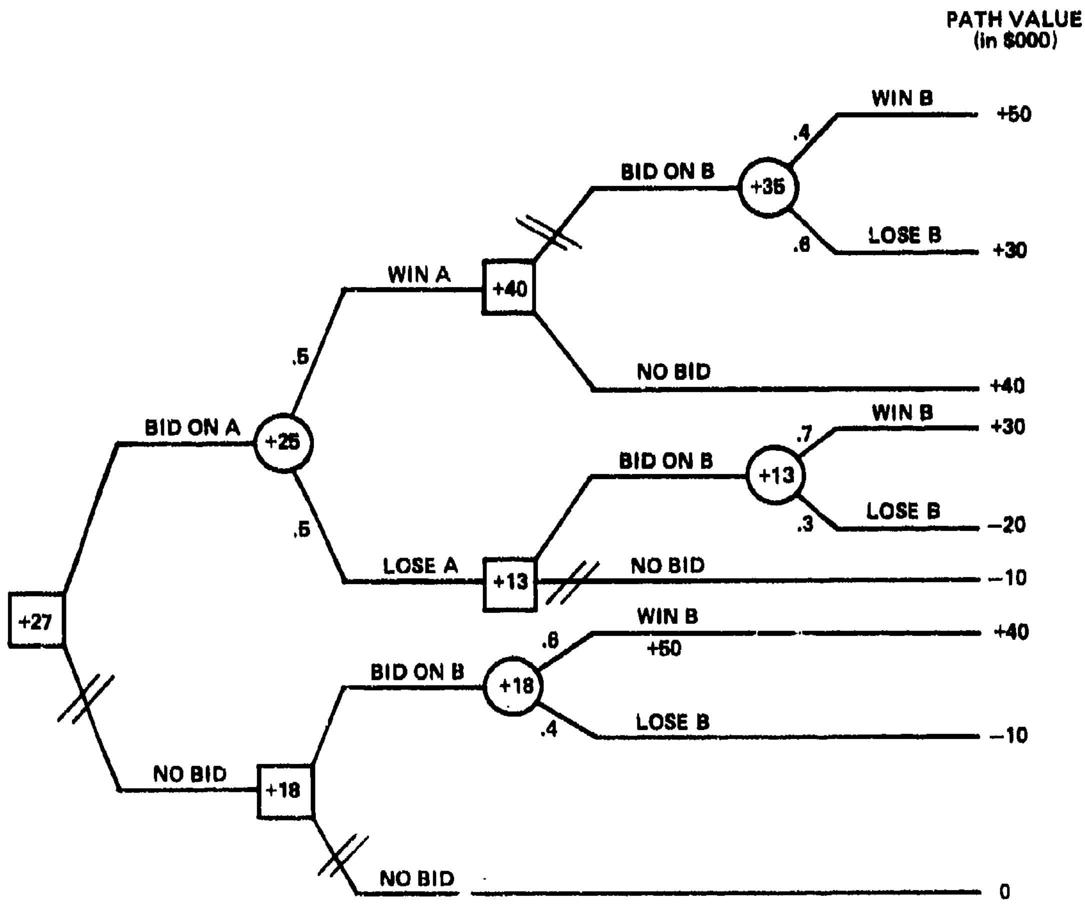


Figure 2-5

FOLDING BACK USING CERTAINTY EQUIVALENTS

owning one car would have high utility, a second car somewhat less utility, and a third car relatively little utility. Thus, there is a contextual and relative aspect of utility. If the three cars are identical, there is no difference between them in terms of objective value, but in the context of already owning two cars, the utility of the third car would be substantially less than either of the first two.

The contextual quality of the concept of utility is an interesting and valuable one for decision-analytic applications. It not only provides a means of quantifying subjective values, but also provides a convenient shorthand for accommodating attitudes toward risk in the solution of decision problems.

To illustrate the concept of utility and its application in a decision context, let's return to our now familiar bidding example (Figure 1-10). Note that the change in monetary assets at issue ranges from a loss of $20,000 (-$20) to a gain of $50,000 (+$50). In arriving at utilities in the form of a utility curve for a given value dimension, one typically restricts consideration to the range of values applicable to that problem. In this case we need to determine utility values over the range from -$20 to +$50. These values at the upper and lower end of the value dimensions are called "reference values," and are arbitrarily given utilities of 0 and 100, respectively.

Figure 2-6 presents a utility curve derived for the range of values at hand. The utility curve expresses the "degree of liking" for each of the range of payoffs shown along the bottom of the chart. This "degree of liking" is expressed in "utiles," shown on the arbitrary scale at the left side of the chart. To translate value to utility, one moves vertically from a value on the bottom of the chart to the utility curve, and from that point of intersection directly across to the utility scale. This operation is shown by the dotted line which links a value of $0 to a utility of 66. If you find it awkward to have a positive utility assigned to zero or, worse, to losses, you may rescale utility by assigning 0 utility to the 66-utile point with increasing negative utilities (0 to -200 at -$20,000) below it and increasing positive utilities (0 to 100 at +$50,000) above it. The scale can also be adjusted by adding or subtracting a constant or by changing the size of the unit. The utility scale, as was indicated before, is arbitrary, as is temperature measured on Celcius or Fahrenheit scales. A scientist, for example, would come to the same conclusion if he used a Celcius rather than a Fahrenheit scale. The choice of scale has no effect upon the relative utilities, only upon your convenience in assessing them.

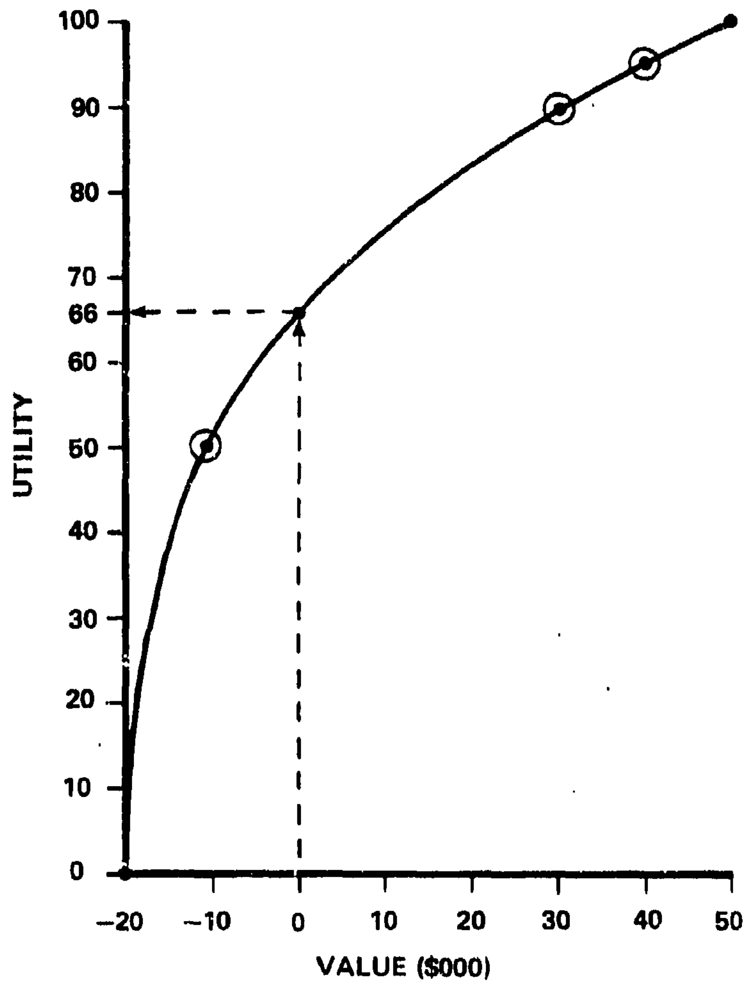


Figure 2-6

A TYPICAL RISK-AVERSE UTILITY CURVE FOR A CHANGE IN ASSETS FROM A LOSS OF $20,000 TO A GAIN OF $50,000

The utility curve shown in Figure 2-6 is arrived at by a series of hypothetical reference gambles presented to the decision maker. He might be asked, for example, what return for certain he would accept as being just equivalent to a gamble offering him a 66% chance of winning +$50,000 and a 34% chance of losing $20,000. This gamble is shown in Figure 2-7A. He judges $0 as his indifference point for that gamble; that is, if he were offered any positive amount of money for sure, he would rather take that money than risk the 34% chance of losing $20,000 even though he could win $50,000. But if he had to pay out any amount of money for sure, he would rather take the gamble. How is that judgment used to determine a point on the utility curve? By finding the utility of the gamble. We do this by calculating its expected utility, which is a weighted average of utilities just as expected value is a weighted average of values. First, we replace the values, +50 and -20, with their corresponding utilities, 100 and 0, as in Figure 2-7B. This gamble displays the decision maker's liking for the outcomes rather than the actual outcomes, 100 indicating most liking (for +50) and 0 indicating least liking (for -20). Next, we reduce this gamble to a single equivalent utility by determining the expected utility of the gamble. For this gamble, the expected utility is (100 x .66) + (0 x .34) = 66. Now we replace the gamble with its expected utility, 66, as in Figure 2-7C. Expected utilities are utilities, just as an average height of several people is itself a height, even if it does not represent the height of any one person. We have now found that the original hypothetical gamble in Figure 2-7A is worth 66 utiles on a scale from 0 to 100. Since the original gamble was judged to be equivalent to $0 for sure, it must follow that $0 and an expected a utility of 66 are equivalent. We say that, relative to $50 and -$10, a value of $0 has a utility of 66. The utility of $0 is 66, or $0 is worth 66 utiles. Now we can plot the point which is given by the intersection of the dashed lines in Figure 2-6. (The process of measuring utility in this manner seems backwards to many people: we chose a gamble with a particular utility and then asked for a judgment of an equivalent amount of money rather than starting with a particular amount of money and finding its utility.)

Next, we construct new gambles with different probabilities and judge their indifference points. For example, the decision maker might be asked what certain value he would accept as being equivalent to a gamble which offered him a 90% chance of winning $50,000 and a 10% chance of losing $20,000. In the case of our example, he declares a monetary value of about $30,000 as his indifference point for that gamble. Proceeding similarly for different gambles, we obtain utility

A.

B.

C.

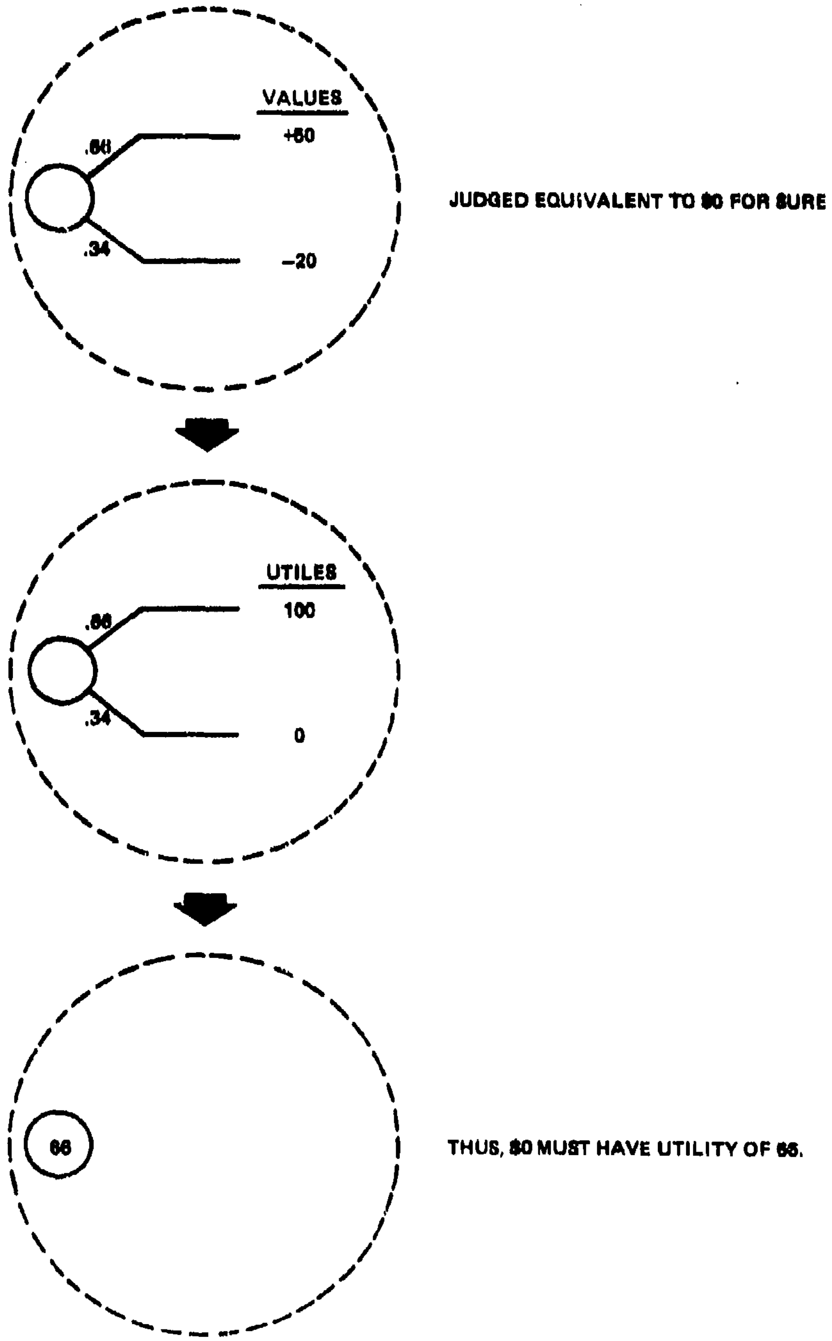


Figure 2-7

ONE HYPOTHETICAL GAMBLE USED IN DETERMINING A POINT ON THE UTILITY CURVE

points corresponding to a sample of the range of values in the bottom of the chart. With a sample of utility points in hand, we draw a smoothed curve through them to arrive at the decision maker's utility function over the whole range of values relevant to the decision problem. In effect, we have used the concept of certainty equivalents developed in the last section along with expected utility to derive the utility function.

It is worth noting that if the decision maker is risk-neutral, the certainty equivalent judged for any hypothetical gamble will equal the expected value of the gamble, with the result that his utility function may be represented by a straight line running diagonally across the chart. The fact that the utility function in Figure 2-6 bends as it does reflects risk aversion. For example, the risk-neutral value (expected value) for the implied gamble which our decision maker valued at $0 in Figure 2-7A is (.66 x $50K) + (.34 x -$20K) = $26.2. His aversion to the risk inherent in that gamble compelled him to assign a lower subjective value than the expected value.

A second useful and critical property of utility curves is that the certainty equivalent of any gamble within the value range can be determined very straightforwardly from its expected utility by reading across from the expected utility to the utility curve, and then down to the corresponding value.

At this point, we should warn the reader that there are several acceptable methods for assessing utility curves; the procedure we have shown here, while good for purposes of explaining the meaning of utility curves, is perhaps not the best way to proceed in a practical problem. To determine several points on the curve, we presented several gambles whose payoffs were always the same, +50 and -20, varied the probabilities and asked for judgments of the certainty equivalents. But because many people find it easier to think about 50-50 gambles, an alternative procedure is to present several gambles, all with 50-50 probabilities but with different payoffs, and ask for judgments of certainty equivalents. Another method that works well for some problems is to specify the certainty equivalent, the probabilities, and the payoff on the lower branch, and ask for a judgment of the payoff on the upper branch that would be needed for the decision maker to be indifferent between the certainty equivalent and the gamble. We would like to make the point that there is no one best procedure for assessing utility curves. The method you use in any particular situation will depend on the nature of the problem and the experience and background of the decision maker.

Once a utility curve has been defined for the range of values relevant to the decision problem under consideration, we are in a position to facilitate greatly the solution of the problem. The reason is that by using utilities as the outcome values, we can fold back the decision diagram using expected utility (calculated the same way as expected value). Expected utility is equal to the sum of products of the probabilities and utilities on the event branches of a particular fork.

Figure 2-8 shows a diagram of our bidding example where utilities from the curve in Figure 2-6 have been substituted for the original dollar values. Folding back on utilities, as in Figure 2-9, shows that bid on A has the higher expected utility (79) and should be the preferred act. Again, although this act is also the preferred act in the risk-neutral analysis previously conducted with expected value, it is important to realize that the two solutions need not have been the same. The difference in utility between bidding on A and not bidding is small, only 2 utiles (79-77). A slightly more risk-averse utility curve (more bowed upward) would result in a higher expected utility for not bidding and thus reverse the result of the expected value analysis.

Utilities Applied To Non-Monetary Values

The device of handling risk aversion through utilities is not restricted to problems where the values are monetary (or where non-monetary considerations have been translated into monetary units). Suppose, for example, that you are evaluating a Mideast foreign policy initiative whose primary purpose is to relieve potential U.S. energy shortages in the medium term.

Analysis has proceeded as follows: the main criterion chosen is an "energy shortage index," measured according to a complex formula which takes into account the pattern of supply and demand over time for oil (or its equivalents), and which is adjusted for any political side effects. Consideration of possible contingencies has been displayed in summary form so that a stripped-down decision diagram might look like that in Figure 2-10.

If the decision maker is risk-neutral in terms of the energy shortage index, then routine folding back based on expected value (with probabilities given in the diagram) shows the "treaty" decision to have a lower (i.e., better) expected shortage than "no treaty" (30 vs. 40). It would, on that basis, be preferred. On the other hand, if the result of a bad shortage is very serious, the decision maker would probably be risk-averse. He might then proceed by assessing a utility curve like the one shown in Figure 2-11. (Note that

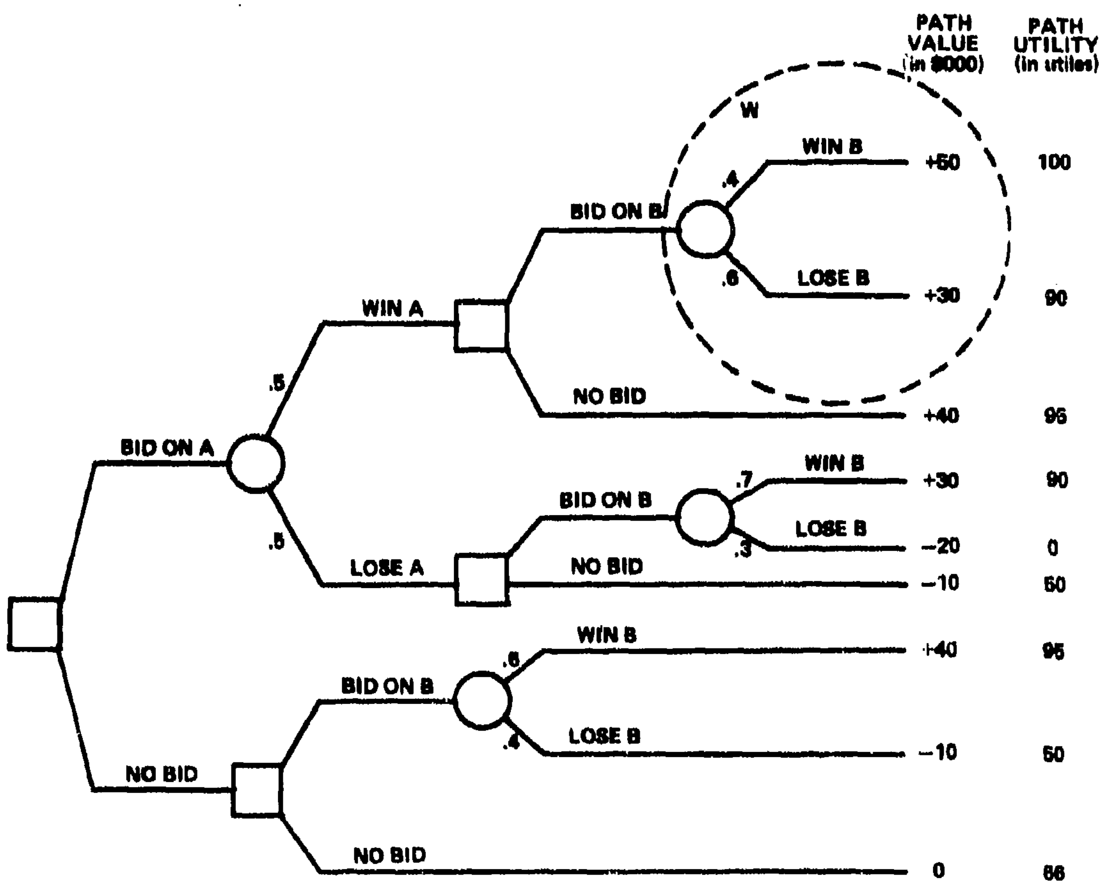


Figure 2-8

TRANSLATING VALUE TO UTILITY ON BIDDING DECISION

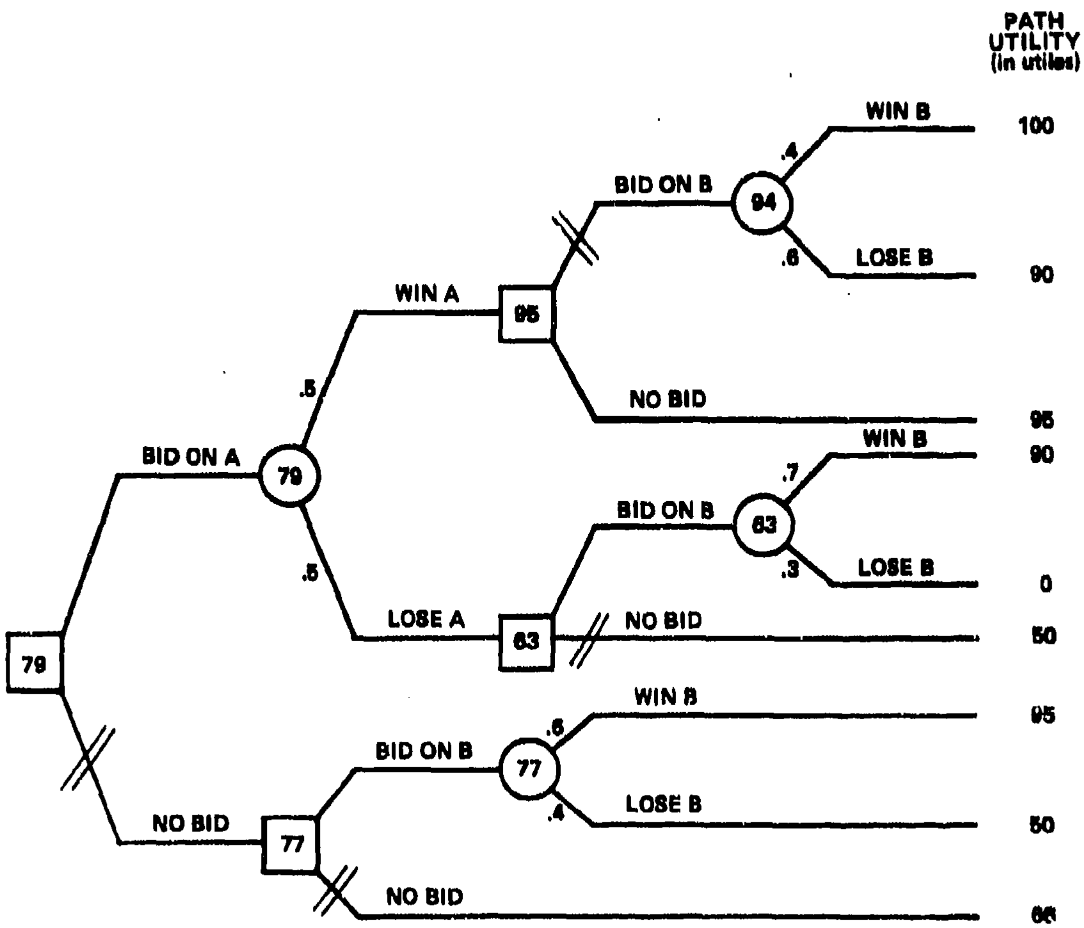


Figure 2-9
UTILITY ANALYSIS OF BIDDING DECISION

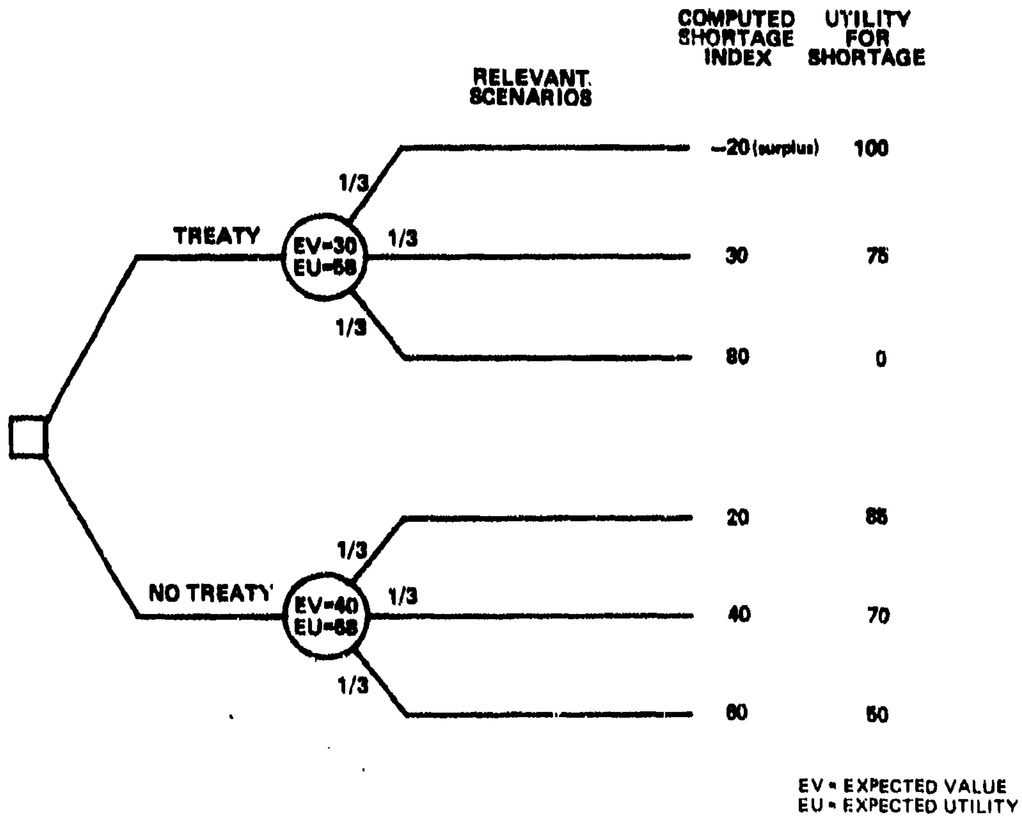


Figure 2-10

SUMMARY DIAGRAM FOR TREATY DECISION

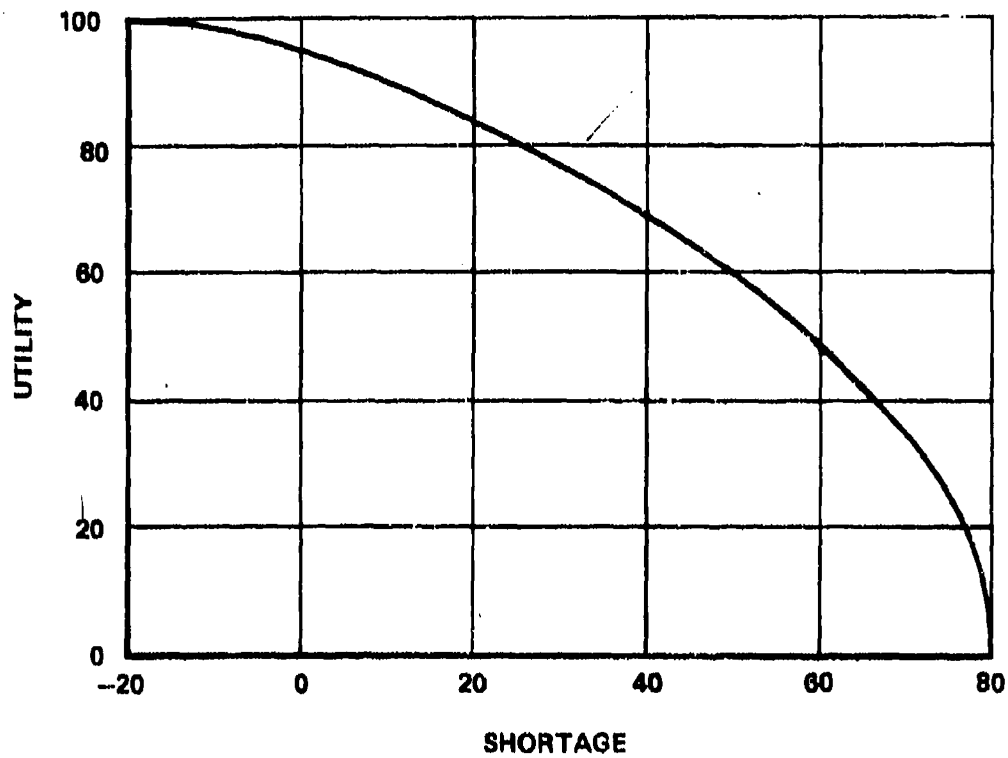


Figure 2-11
UTILITY FOR SHORTAGE

it slopes down from left to right because more shortage means less utility.) In this case, since the expected utility of "treaty" proves to be the less attractive (58 vs. 68), the "no treaty" option is now preferred. (Under risk neutrality as calculated by expected value, the decision went the other way.)

It can be demonstrated that in most decision situations, solving the decision problem in terms of expected utility instead of expected value leads to the best solution from the decision maker's perspective.

Why You Should Maximize Expected Utility

The following discussion is concerned with persuading you that the preceding procedure for using utilities (defined in terms of reference gambles) is logically sound. We shall use a concrete example to demonstrate that the act with the highest expected utility is the one that should be chosen. Assume that you have been offered the choice of the two gambles illustrated in Figure 2-12.

For Gamble A, a fair coin will be tossed; if it comes up heads, you will win $7,000 cash; and if it comes up tails, you will win nothing. For Gamble B, a fair coin will also be tossed; if it comes down heads, you will win $4,000; and if it comes down tails, you will win $1,000. The problem is to decide which of these two gambles you should take. You are offered either one of them, but only one. Gamble A is worth $1,000 more in terms of expected value ($3,500 vs $2,500), but that does not necessarily mean that you prefer Gamble A.

Let us further assume that you are quite risk-averse, that you do not like to take serious chances with money. In fact, assume that the degree to which you are risk-averse is described by the utility function in Figure 2-13. The reference values are $0 and $7,000, and amounts in between represent increases on your current scale.

We can now use this utility function and translate the values on the decision diagram as shown in Figure 2-14. For this diagram the end points are measured in utiles rather than dollars: the utility of $7,000 is 100; of $0, 0; of $4000, 84; and of $1000, 50. The expected utilities are 50 if you choose Gamble A, and 67 for choosing Gamble B.

Until now, we have been following the practice of solving a decision diagram by folding it back. Now, however, we are going to move in the opposite direction. We are going to expand the diagram rather than simplify it. We shall substitute a gamble, an event fork, for each of the end points of the diagram in Figure 2-14 in order to create the new decision diagram shown in Figure 2-15.

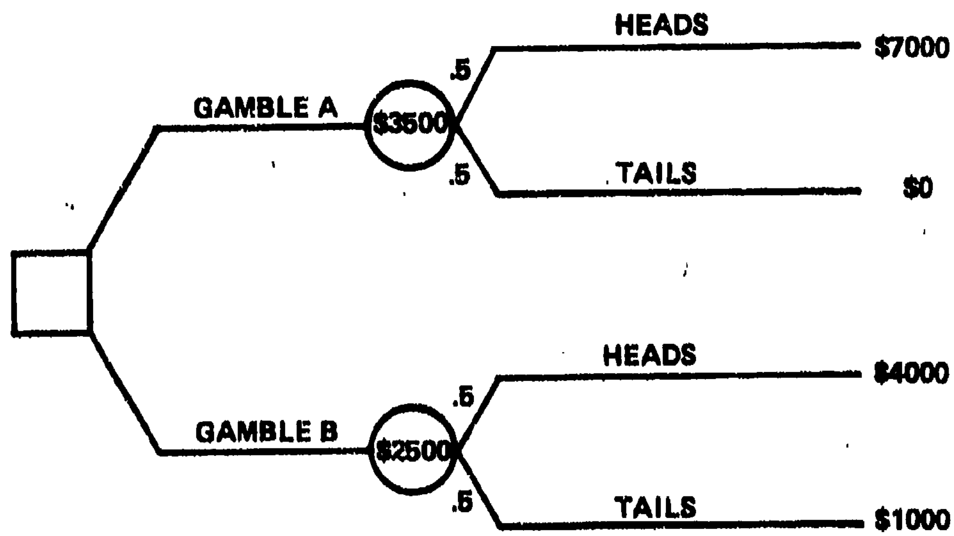


Figure 2-12
EXPECTED VALUE ANALYSIS OF COIN-TOSSING EXAMPLE

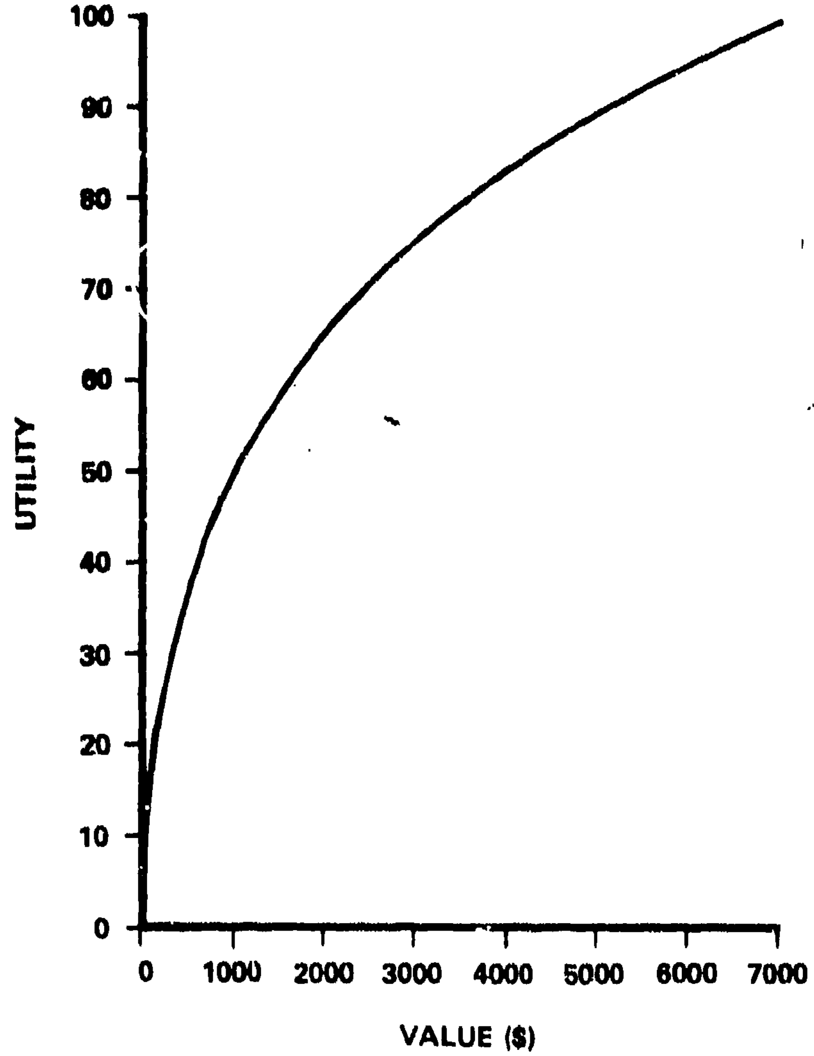


Figure 2-13
UTILITY CURVE FOR INCREASE OF $0 TO $7000 IN ASSETS

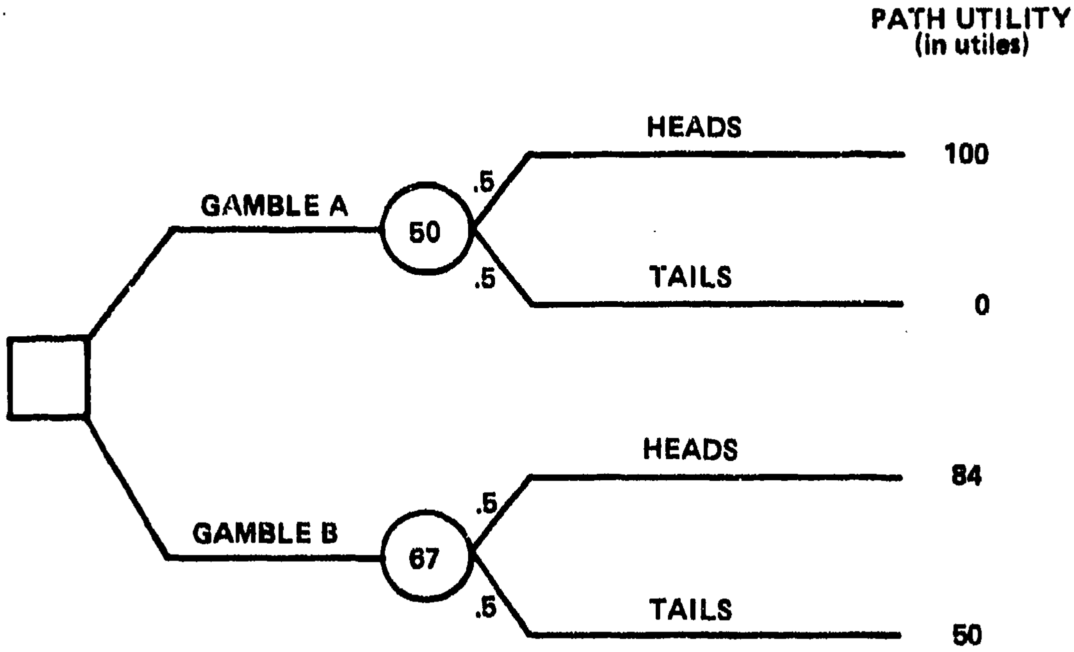


Figure 2-14
UTILITY ANALYSIS OF COIN-TOSSING EXAMPLE

There are four end points in Figure 2-14 and each has been converted into a corresponding event fork in Figure 2-15. The event fork is constructed so that \$7,000, the maximum gain, is associated with the probability (in percent) equal to the utility of the end point. Otherwise, \$0 is won. Accordingly, the utility of 100 on the top branch in Figure 2-14 is equivalent to the top event fork in Figure 2-15, where \$7,000 is won with probability 1.0 and \$0 is won with probability 0. Progressing down to the next end point, with utility 0, we substitute the event fork in which \$7,000 is won with probability 0 and \$0 is won with probability 1.0. The probabilities of the extended branches following Act A are always 1 or 0 because the utilities of the end point are 100 and 0, respectively.

Now consider Gamble B. Here a utility of 84 is equivalent to an event fork in which \$7,000 is won with probability 0.84 and \$0 with probability 0.16. Finally, for the bottom end point with the utility of 50, we have an equivalent gamble in which \$7,000 is won with probability 0.50 and \$0 with probability 0.50.

The numbers in the parentheses to the right of Figure 2-15 indicate the path probabilities. For the top branch, the path probability is .50 x 1.0, or is .50. In each case, the path probability is the product of the event probabilities which comprise the path.

We now have a rather complicated decision diagram, but one with only two possible outcomes: you either win \$7,000 or you win \$0. For each act, we now combine all branches with identical consequences and add the path probabilities for the identical payoffs. The result is a simpler but equivalent decision diagram with two possible consequences, win \$7,000 or win \$0, following each act. This simpler decision diagram is displayed in Figure 2-16. For Gamble A, you have simply ended up with the original gamble. You win \$7,000 with a probability 0.5, or you lose \$0. But for Act B, you have a very different gamble from the original one. Here, you win \$7,000 with probability 0.67, or you lose \$0 with probability .33. Note that the probability of winning \$7,000 is expressed as a percentage for each gamble, then the percentage is equal to the expected utility of that gamble. Since you presumably prefer a 67 percent chance of winning \$7,000 to a 50 percent chance of winning the same amount, Gamble B is preferred to Gamble A. Our previous analysis of Figure 2-14 showed that 50 is the expected utility of Gamble A, and 67 is the expected utility of Gamble B. The preferred gamble is the one with the higher expected utility.

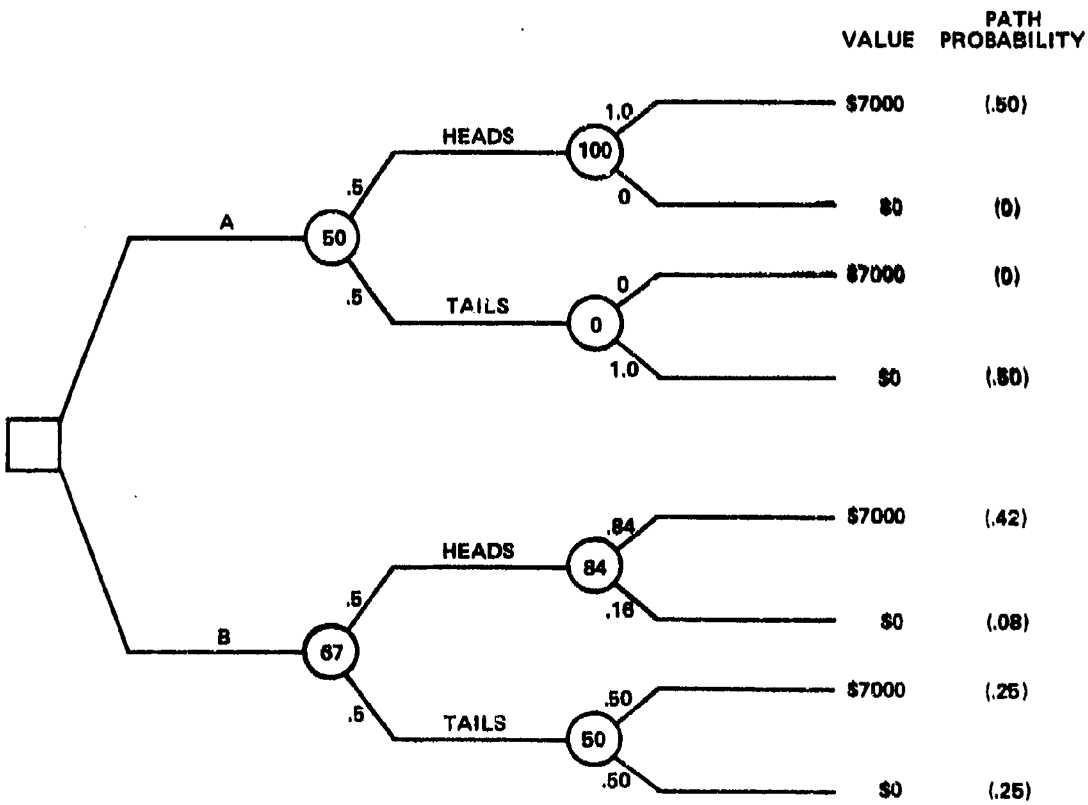


Figure 2-15

REFERENCE GAMBLE SUBSTITUTION IN COIN-TOSSING EXAMPLE

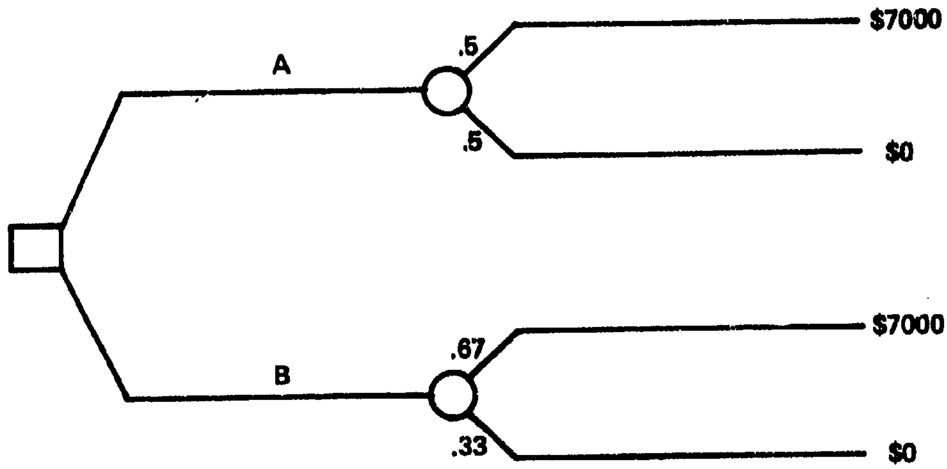


Figure 2-16

REFERENCE GAMBLE REDUCTION IN COIN-TOSSING EXAMPLE

This is the key point of the demonstration. It is always possible to substitute a new diagram that is equivalent to the original decision diagram in the manner described above. This new diagram will have the property that, for any initial act, you win the maximum amount with a probability (expressed as a percentage) equal to the expected utility of the initial act and you win the minimum amount otherwise. In this substitute diagram, we no longer find it necessary to be concerned with trade-offs between probabilities and amounts to be won, because the amounts to be won are held constant across all initial acts. You either win the maximum amount or you win the minimum amount. What varies from act to act is only the probability of winning, and that probability (expressed as a percentage) is equal to the expected utility associated with the act. It must be true that any decision maker faced with two probabilities of winning the maximum amount will select the gamble with the higher probability of winning. It follows that he should also select the initial act with the higher expected utility. In a somewhat anecdotal fashion, we have now demonstrated what we set out to show.

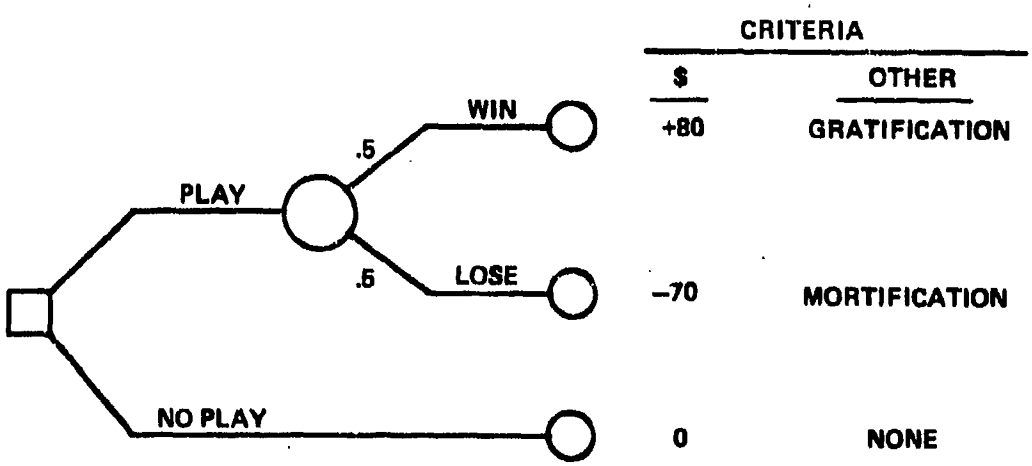
Multiple Attributes of Value

So far, we have considered decision situations wherein there has been but one attribute of value. In practice, it is rare to find decision situations where only one value attribute bears on a decision. A decision can often be dominated by one attribute but it is rare indeed to encounter a complex decision problem, one which would merit the effort of serious analytical effort, that has but one relevant value attribute. It is also the case that many value attributes which must be considered in decisions are qualitative, that is, are not expressed in any conventional metric. In the following sections, we shall present various methods to handle multiple attributes of value in a decision-analytic context and present additional means to handle qualitative value attributes.

Measuring Different Attributes by a Single Criterion

One approach is to select one readily quantified criterion, like cost in dollars, and express all other attributes in terms of that one. This procedure can best be explained in the context of an example. Suppose you are invited to play a coin tossing game which leaves you either $80 ahead or $70 behind with equal probability. Now further suppose that winning a game like this would give you some measure of gratification beyond the monetary reward but that getting into a game like this and <u>loning</u> carries a special cost to you, for example, embarrassment in explaining the loss to your spouse. Your real perception of the decision can be represented by the diagram in Figure 2-17A.

A. Multi-Attributed Problem and Multiple Criteria



B. Same Multi-Attributed Problem and a Single Criterion

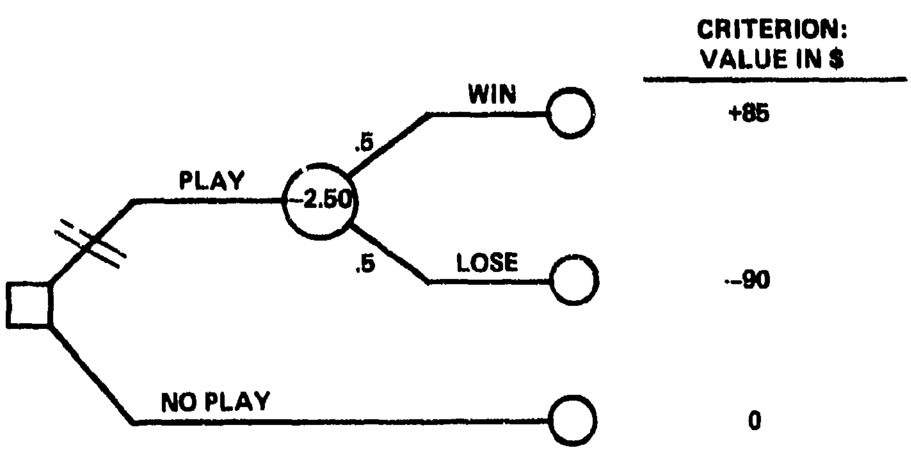


Figure 2-17

SIMPLE EXAMPLE OF MEASURING DIFFERENT ATTRIBUTES BY A SINGLE CRITERION

Suppose you would not exchange the gratification of winning for less than $5, but you would pay $20 to avoid the embarrassment of losing. These judgments allow you to substitute the problem shown in Figure 2-17B which has only one quantified criterion, and which therefore can be analyzed by folding back exactly as before.

In principle it is always possible to measure different attributes by a single quantified criterion whether or not all of them are numerically measurable. What is important is that the resulting values correctly reflect differences among outcomes in terms of their relative importance in the decision-making process. This principal is further illustrated in the following example.

A presidential aide is faced with the decision of whether or not to recommend signing a political and economic treaty with a Mideast country. The main motivation is to ensure a higher level of oil supplies over the next ten years. However, it is also important to the U.S. to take account of the price to be paid for the oil and the impact of the treaty on political goodwill vis-a-vis several groups of countries. Since the possible outcomes for either decision are very complex, the aide selects a limited number of possible scenarios to represent the real possibilities, assigns probabilities to them, and notes what their impact on each of the several attributes would be. Figure 2-18 shows a simplified version of the decision diagram. Note that the aide assigned a slightly higher probability to scenario A, leading to 5 million barrels a day, for the treaty than for no treaty.

He then evaluates the "trade-off" between the main attributes of oil supply and each other attribute in terms of oil supply, as shown in Figure 2-19. These measurements take into account not only his evaluation of the qualitative attributes, but also the importance of oil in the light of developments in the pattern of supply and demand of energy for the U.S. indicated in the scenario.

The aide nets out the values and folds back on equivalent value. The result, shown in Figure 2-19, shows that the aide should recommend rejecting the treaty for the particular evaluations and assessments he has made. It further shows that he would need to be guaranteed a bonus of 1.6 million barrels of oil a day (the difference between -0.5 bbl per calendar day and +1.1 bbl per calendar day) before he would reach an indifference point between having a treaty or not.

A Weighted Index of Attractiveness

Where no quantitative criterion dominates or the qualitative attributes are so critical as to swamp differences in

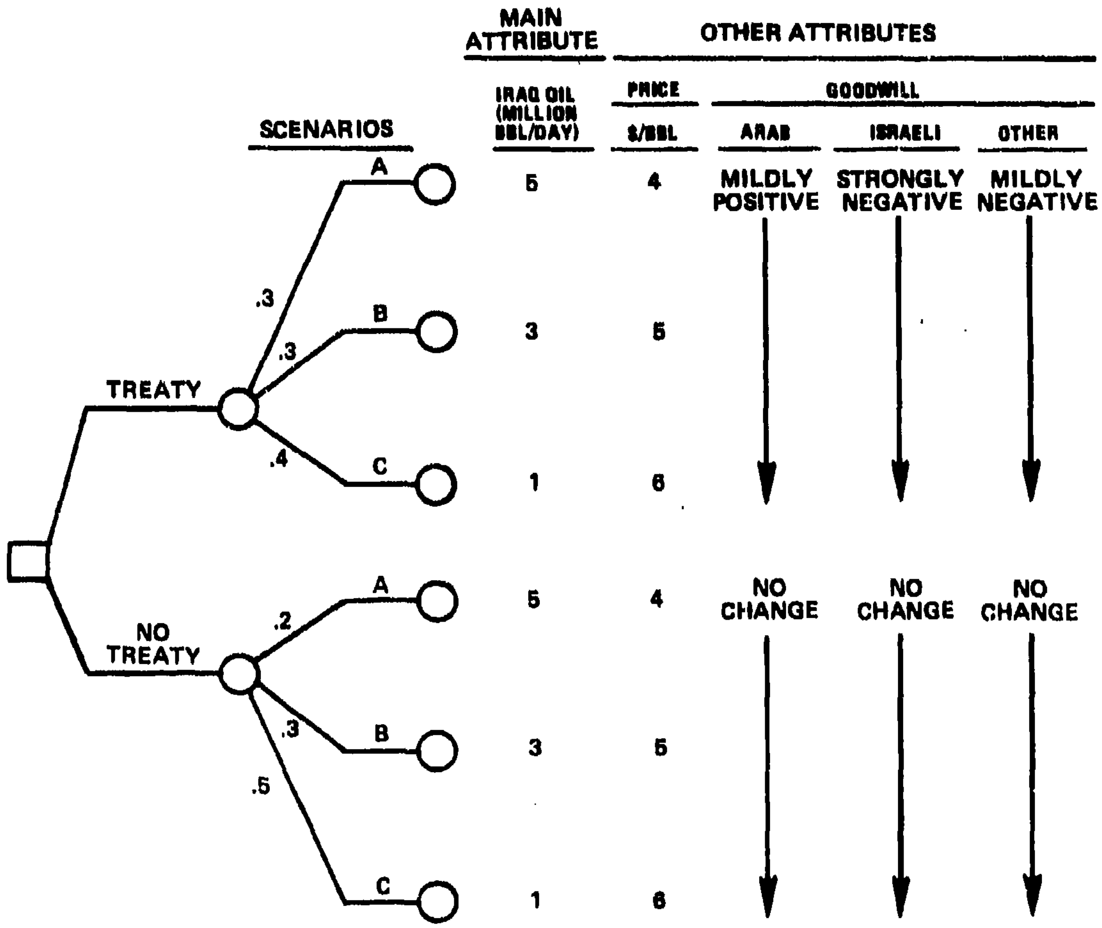


Figure 2-18

DECISION DIAGRAM FOR OIL TREATY WITH MULTIPLE CRITERIA

the main criterion, it may be difficult to make the required measurements with an adequate degree of precision. In such cases, as an alternative to the single criterion approach to multi-attributed decisions, an arbitrary weighted index of attractiveness can be used.

Suppose that the decision is whether or not to close down a military installation. In considering the possible outcome of his decision, the decision maker might take into account the following attributes: contribution of the installation to strategic objectives, goodwill in the local population, and, to a lesser degree, cost.

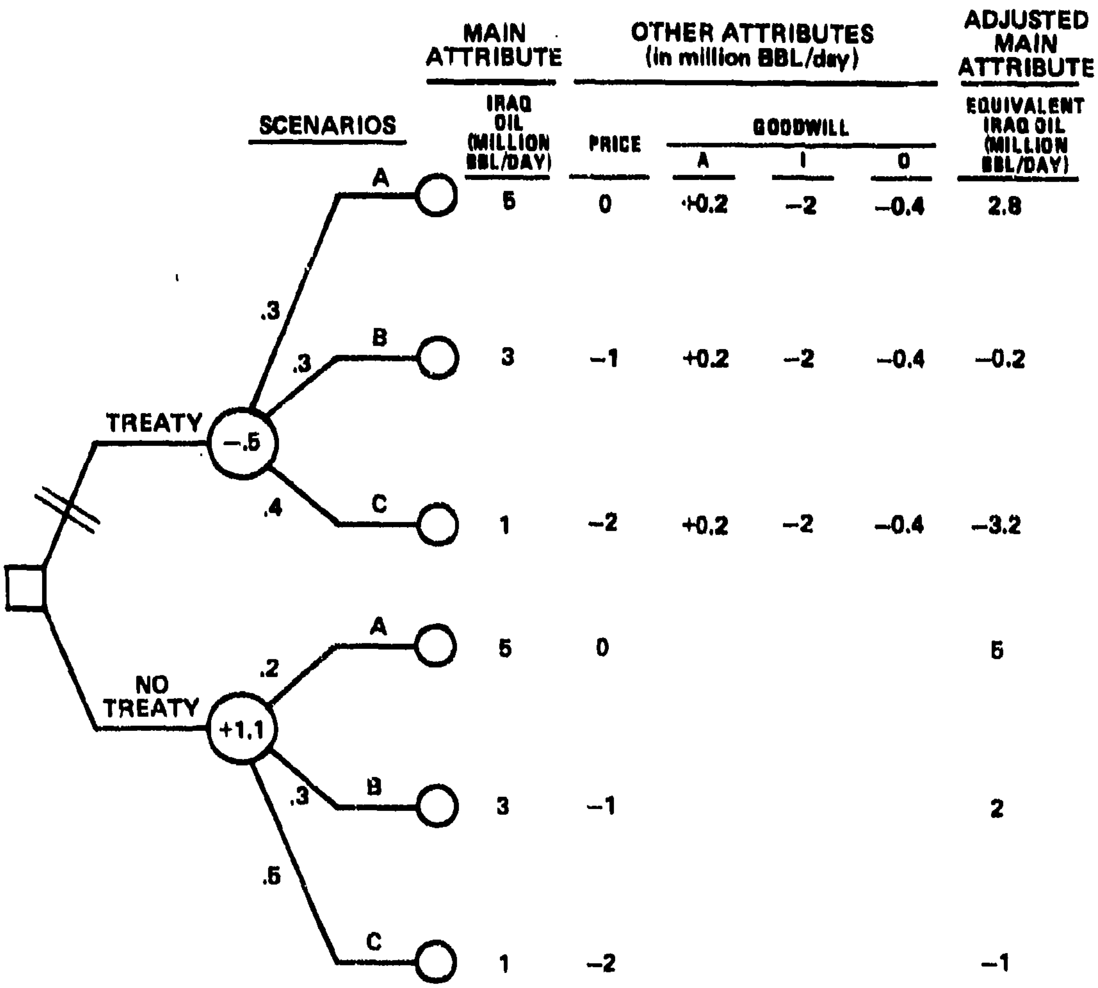


Figure 2-19

DECISION DIAGRAM FOR OIL TREATY WITH SINGLE, QUANTIFIED CRITERION

If uncertainty is not an issue, the choice might be represented as in Figure 2-20. This representation reflects the judgment that while closing the installation will represent an important sacrifice in terms of national strategic objectives, it will have a favorable impact on relations with the local population and will save $50 million a year.

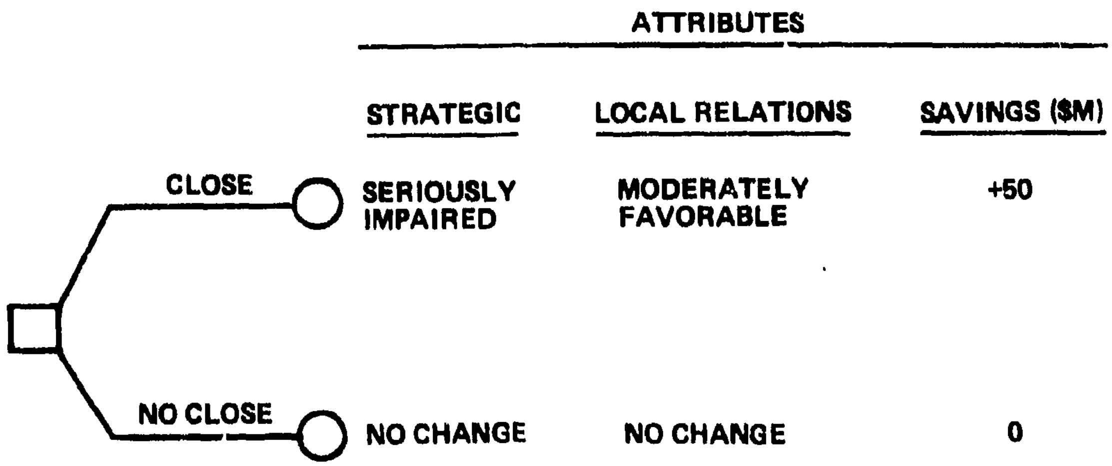


Figure 2-20

DECISION DIAGRAM FOR
CLOSING A MILITARY INSTALLATION

To construct a weighted index, we first assign an arbitrary scale for each attribute, say, from 0 to 100, defining carefully the ends of the scale. For example, 0 value of the strategic attribute might be a state of national emergency, and 100 might be the practical ultimate in preparedness. Next, each decision option is valued on each attribute scale, and a measure of relative importance between attributes is set up, as shown in Figure 2-21A.

Finally, the attribute values are weighted and summed, as shown in Figure 2-21B, to give a single index of attractiveness, which indicates that, on this set of value assessments, not closing the installation is preferred.

In this simplified hypothetical example, the formal diagram shown in Figure 2-21 may appear to add little to the initial statement of the problem in Figure 2-19 which, no

A. Scaling and Weighting Criteria

	ATTRIBUTES		
	STRATEGIC	LOCAL RELATIONS	SAVINGS ($M)
CLOSE	40	60	40
NO CLOSE	80	30	20
RELATIVE IMPORTANCE WEIGHTING FACTOR	10	5	3

B. Weighted Sum Index of Outcomes

	ATTRIBUTES (WEIGHTED)			
	STRATEGIC	LOCAL REL.	SAVINGS	WEIGHTED SUM
CLOSE	400	300	120	820
NO CLOSE	800	150	60	1010

Figure 2-21

DECISION DIAGRAM FOR CLOSING DOWN A MILITARY BASE USING WEIGHTED-SUM PROCEDURE doubt, could be readily digested intuitively. However, when the number of options, possible outcomes, and attributes is multiplied, the case for this variant of "divide and conquer" becomes more compelling.

It should be noted that this weighted-sum procedure, in contrast to the previous adjustment procedure, is not always logically sound, if, for example, there are important interactions among criteria. There are other procedures which are logically watertight, but they are typically less convenient to apply.

Multi-Attribute Utility (MAU) Models for Evaluating Complex Systems

The fact that utility values provide an index of merit for things, or multiples of things, and also provide a common metric for aggregating the measures into a single index of worth makes the concept of utility highly useful in evaluating complex systems. Determining a figure of merit for alternative complex system designs can be a formidable task largely because the competing system designs vary from each other in terms of many system attributes bearing on overall worth. Proposed radio set A, for example, may be lighter than B but have less transmission range. It may be more reliable but have fewer channels. It may require less power but be more readily jammed and so on. To further compound the problem, the system under consideration may face multiple roles (e.g. battle communication, rear-area use) and operating environments (urban area, open desert). Which design for which role for which environment? Which would be the best design for all?

A variety of procedures can be used to make the indicated evaluation. These could range from overall (global) judgments by experts as to which system is best to the use of very complex simulations. The former approach tends to be too simplistic and subject to many inherent difficulties, one of which is that the evaluation problem is usually far too complex for accurate, global judgments; the latter, though often used, suffers from the shortcoming that resulting performance measures may not be directly related to measures of worth or utility for mission performance. By using utility as a measure of merit and by applying this concept to a systematic disaggregation of the total problem, we can circumvent the shortfalls of global judgments and simulation approaches. This approach is known as multi-attribute utility (MAU) assessment. How it works in practice will be shown in the following example, which is a partial representation of an evaluation of alternative military radio systems that was actually conducted, but which is here greatly simplified for purposes of exposition.

The evaluation problem involves determining the military worth of four alternative field radio designs so that a rational selection could be made among them. In the MAU approach, the first step is to establish a model structured to relate physical system parameters to military value. This structure is a representation of how the physical and performance attributes of the system or systems being evaluated are translated into measures of utility, which in turn reflect how well each system accomplishes its military mission. In the case of the radio evaluation, the system attributes are first disaggregated into two major categories: technical system utility (how well each system functions as a radio) and operational acceptability (how attractive the system is from an operational point of view).[1] Each of these two major assessment categories is further disaggregated, as shown in Figures 2-22A and 2-22B. The disaggregation proceeds until, at the lower level of the model, the system is depicted in terms of its actual measurable attributes. These measurable attributes serve as the referent for the assessment of utility functions for those system attributes. These utility curves are elicited in the form of judgments by subject area experts and users.

After the utility functions have been developed, it is necessary to assign importance weights to each branch of the model. Two kinds of importance weights are required, depending upon the rules of combination. In the case of a set of attributes whose relationship is captured by an additive rule, importance weights simply provide a means of calculating a weighted average over the attributes being considered. Consequently, importance weights assume the form of coefficients that are multiplied by the utility of each of the attributes, and these coefficients sum to 1.0 across all related attributes within a given level of the hierarchical structure. The weight given to a particular attribute reflects the relative contribution of that factor vis-a-vis the contribution of the other related attributes to utility for that level of the model.

There are two sets of multiplicative factors in Figure 2-22A. If factors within a given level are multiplicative (that is, they are interactive in a value-wise sense), then the measure of utility of any factor may be considered as a measure of degradation. For example, technical performance and dependability are shown as interacting in Figure 2-22A. Systems capable of great technical performance are more complex than limited-capability systems, and so are less dependable since more things can go wrong. When variables interact in this manner, importance weights have the effect of rescaling the attributes. Prior to applying a weight, the

[1] The distinction is real. For example, a radio system may function well as a radio but be too heavy to carry in the field.

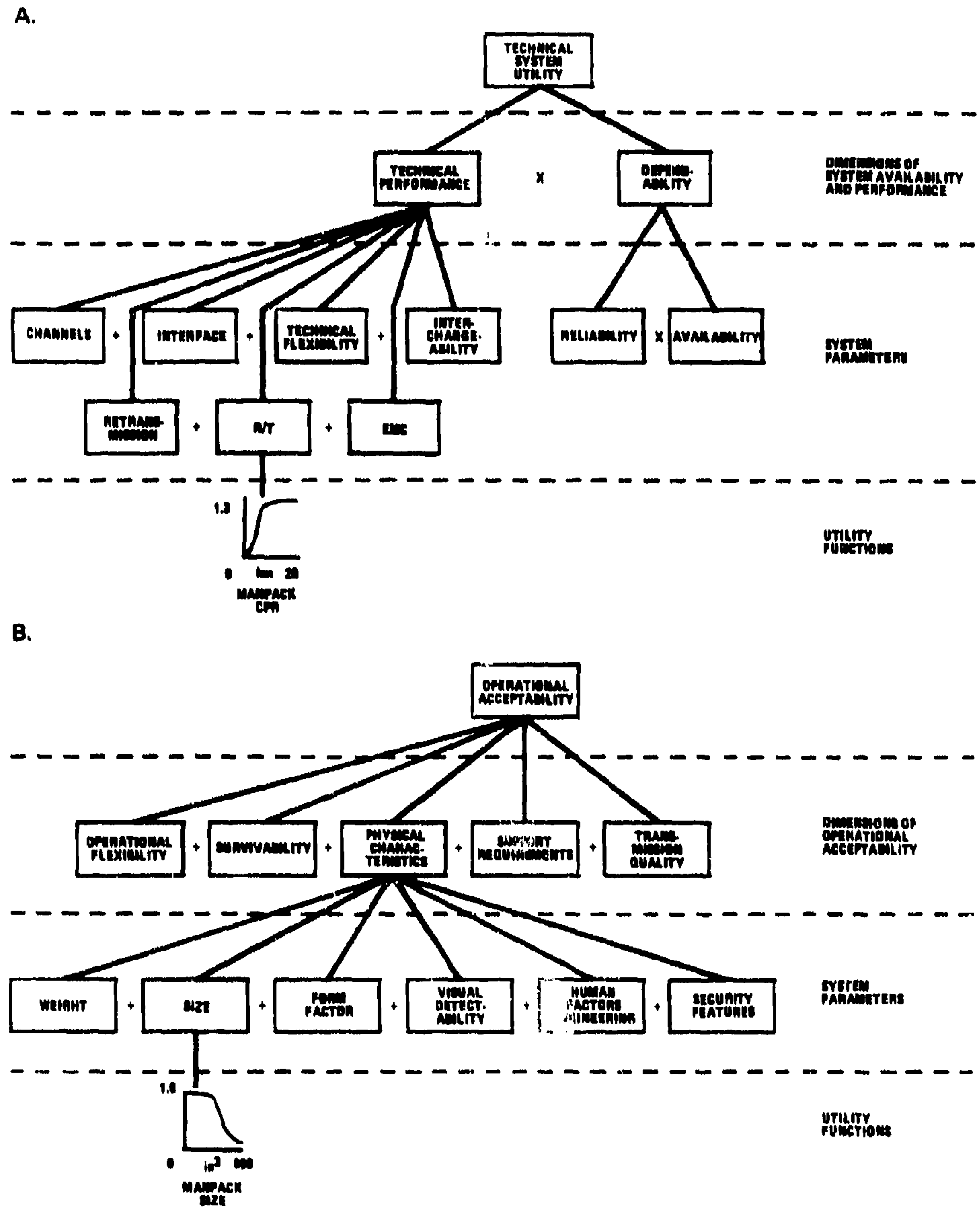


Figure 2-22

A SIMPLIFIED STRUCTURE OF A RADIO EVALUATION MODEL

utility over one of the attributes may have any value between zero and one. However, assume, for example, when this attribute takes on its worst value, that the effectiveness of a system is degraded to the order of 30%. In this case, the utility, originally scaled from 0 to 1.0, would be rescaled from 0.3 to 1.0.

Having now achieved weighted utilities for the elements of the model, the final step is to combine the values in accordance with the appropriate combination rule (in our case, a combination of additive and multiplicative rules) to arrive at utilities for successively higher levels of aggregation. The final utility value of a radio assessed in our sample model would be arrived at by adding the weighted utility value of technical system utility (Figure 2-22A) to that of operational acceptability (Figure 2-22B). Utility value thus derived for each of the system designs under consideration provides a systematically derived, numerical basis for comparison.

The multi-attribute utility approach described above has a number of advantages. First, it permits an individual who is an expert in a particular area to make judgments which involve his particular area of expertise rather than making an overall judgment of worth which may fall outside his area of expertise. Second, disaggregating the judgments of individual experts provides an explicit trail leading from measures of system performance to measures of benefit or utility. Thus, the judgments are public rather than private and are subject to screening.

Nonetheless, there are many critics who bridle at the use of subjective values no matter how systematically or logically they might be used. They contend that the subjective importance weights might be wrong, the assessed utilities inaccurate, or the logic of the model structure flawed. While it is true that such errors can occur, there are methods for determining what influence on final assessments such errors might have. These methods, called sensitivity analyses, can be used to check the sensitivity of the model to variances in subjective values, to compare the assessments made by the model to firm external reference criteria, and to permit further consideration of assessments crucial to the output of the model.

Sensitivity analyses within the model involve an evaluation of the degree to which changes in utility functions, importance weights, and combination rules influence the output of the model. For example, to what degree does the relative utility of Alternative System 1 vs Alternative System 2 depend upon the weights assigned to an attribute within the model or to different utility values? Sensitivity analyses external to the model provide a means of calibrating the model and involve the use of the model to compare known, well-understood alternative systems before using the model to evaluate less well-understood systems. This is to say that it is useful to test the model against educated judgments about the relative utility of existing systems, or components of existing systems, before applying the model to possible future systems. If the model appears to be blatantly at odds with expert judgments pertaining to the relative utility of known systems, analysis of the model should lead to identifying the source of the disagreement and thereby facilitate modification of the model in order to bring subsequent evaluations more into line with informed judgments. It is difficult to overemphasize the importance of these kinds of sensitivity analyses in terms of both improving the evaluation model itself and developing confidence and understanding on the part of the user who intends to make recommendations based upon the output of the model.

CHAPTER 3

MEASURING UNCERTAINTY

In this chapter we provide some useful background information about probabilities, describe ways of obtaining probability values as a numerical expression of uncertainty, discuss the biases that may result during the assessment of probabilities (and how to compensate for them), and finally, consider how to revise probability values on the basis of new information.

Qualitative Expressions of Uncertainty

Managers are accustomed to dealing with uncertainty informally. For instance, they may prepare contingency plans to be put in effect if some uncertain future event turns out differently from what is expected. The term "calculated risk" is sometimes used loosely in this context to indicate that uncertainties have to be considered. In the decision-analytic approach to problem solving, however, the amount of risk (equal to the sum of the outcomes times their probabilities) actually _is_ calculated. Many managers also express their feelings about future events by using semiquantitative terms such as "unlikely," "probable," or "rare." For example, a marketing person might discuss next year's projected sales for a product by citing three different possible levels of volume: optimistic, pessimistic, and best guess. Similarly, a military analyst might respond to a question about whether an attack will occur at a particular time and place by saying it is "highly unlikely." Figure 3.1 shows these two examples in decision tree format. Before these nodes can become part of a decision analysis, however, both the uncertainties and the values of the outcomes must be expressed numerically. Chapter 2 explained how the relative desirability of different outcomes can be translated into the measure called utility. Now we will discuss how to quantify _uncertainty_ in terms of a probability. This process, called "probability assessment," is a collection of techniques that help a person take feelings about uncertainty that may be vaguely but comfortably expressed in qualitative terms and convert these findings into numbers on a scale from zero to one. An outcome which is impossible will have a probability equal to zero. For example, the probability that the Soviets have a missile which exceeds the speed of light is zero. On the other hand, if an outcome is absolutely certain to occur, it will have a probability equal to one. For example, the probability that the temperature in downtown Nairobi tomorrow will be higher than absolute zero is 1.0. Intermediate values of probability represent intermediate degrees of certainty. A probability of 0.5 means that the event is just as likely to occur as not.

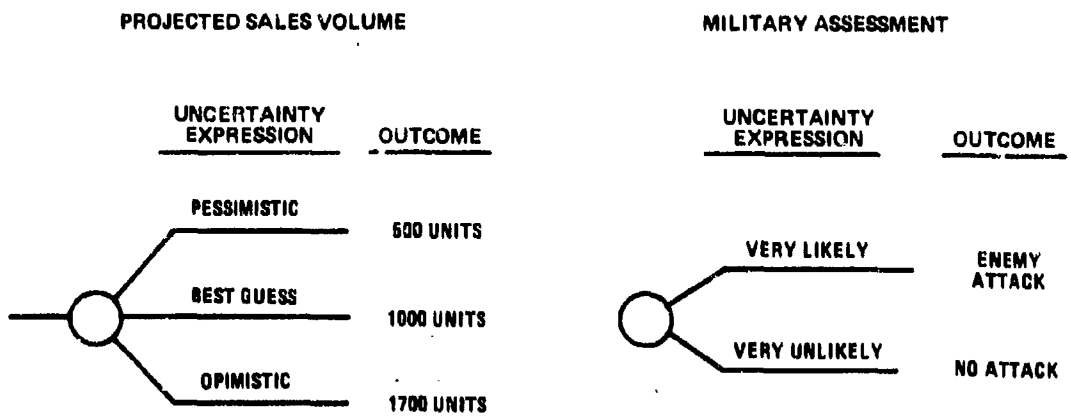


Figure 3-1

EXAMPLES OF EVENT NODES WHERE UNCERTAINTY IS EXPRESSED IN SEMI-QUANTITATIVE TERMS

Definition of Probability

The probability values we shall elicit represent a person's degree of belief that various events will occur. The source of this information may be the manager, an expert called in by the manager, or even a group of people. The idea of a probability as a degree of belief contrasts with some more narrow definitions of probability, for instance, probability as a relative frequency or probability derived from arguments of symmetry. Examples of the latter two types of probabilities would be:

1. Relative Frequency. A flight attendant asks passengers whether they want pan-fried chicken or Salisbury steak for dinner. By the time 50 passengers have been asked, the responses are chicken, 30 passengers; steak, 20 passengers. Therefore, he estimates that the probability that the next person will say "chicken" is 30/50 or 0.6.

2. Symmetry Argument. A carefully made "Las Vegas" die (perfectly symmetrical and made of homogeneous material) is rolled. Because of its physical characteristics, each of the six sides is equally likely to be on top at the end of the roll. The probability that a particular number, say, five, will show is equal to 1/6 or 0.1667.

Neither of these narrow methods for defining probability is very useful for the manager because a manager is rarely concerned with a situation where exactly the same event occurs over and over so that relative frequency can be used, or with a situation where a number of events being considered can be shown to be equally likely to occur. Consequently, we accept the broader definition of probability as a degree of belief and concentrate on ways of finding out the appropriate numerical values to express this degree of belief. These values express what are called personal probabilities.

Probability Rules

Probability theory has as its foundation a number of mathematical rules about the definition of probabilities and computations involving probabilities. In this introduction to decision analysis, however, only a few of these features of probability need to be discussed.

Range of Probability Values

We have already mentioned that probabilities must lie somewhere in the range from 0.00 through 1.00. According to this convention, it would be meaningless to say that some outcome had a negative probability or a probability greater than 1.0.

Mutually Exclusive Outcomes

The set of outcomes of an event should be chosen or defined so that different outcomes are mutually exclusive, that is, so that no more than one of the outcomes in the set can occur. In the military prospects example, two outcomes are considered: enemy attack and no attack. Since only one of these can occur, the events are mutually exclusive.

Exhaustive Outcomes

The outcomes of an event should be chosen or defined so that they are exhaustive in the sense that some one of the outcomes must occur. In the example above, "enemy attack" and "no attack" are obviously an exhaustive set of outcomes. Other exhaustive sets could be defined in this instance, for example, by breaking down each of those two outcomes to give four possibilities in all:

concerted attack,
skirmish,
stand fast but no attack,
retreat.

The choice of a particular exhaustive set of outcomes is an arbitrary one, depending on the judgment of the decision maker. However, no important possibilities should be omitted. To ensure the exhaustiveness of the set, the decision maker may choose to include a catch-all outcome labeled "Other."

Probabilities Sum to 1.00

Once a set of mutually exclusive, exhaustive outcomes is defined, probabilities are determined for each outcome, and these probabilities will all be in the range from 0.00 to 1.00. The last important feature of probabilities to be considered is that the total of all these probabilities must sum to 1.00. For example, in the first formulation of the military prospects illustration, where only two outcomes are considered, if the probability of "enemy attack" were equal to 0.15, then the probability of "no attack" would have to be equal to 0.85 to make their sum, 0.15 + 0.85, equal to 1.00.

Two Types of Outcomes

In assessing the probabilities of the different outcomes of an event, we distinguish between two cases: where there is a limited number of outcomes and where there is a whole range of outcomes. These situations require probabilities that are called discrete probabilities and continuous probabilities, respectively. The techniques used in their assessment are somewhat different.

Discrete Probabilities

Both the examples shown in Figure 3.1 are set up for discrete probabilities. In the sales volume example, only three levels of sales are here considered as possible outcomes. The probability of achieving each of these three levels would be assessed, and their total value would be 1.00, since, by definition, when all possible outcomes of an event have been considered, their probabilities must add up to 1.00. These can be displayed in a probability table as shown below.

Sales Level (Units Sold)		Probability
500		0.1
1000		0.7
1700		0.2
	Total:	1.0

In the military event node, only two outcomes are considered possible, attack or no attack. Two probabilities would be assessed and their total, as in the previous example, would be 1.00. If the analyst thought that the probability of an attack were 1/100 or 0.01, then the probability of "no attack" would have to be 1.00 - 0.01 or .99. These values are shown in the table below.

Outcome		Probability
Attack		0.01
No attack		0.99
	Total:	1.00

Another way to display the probabilities is graphically, as in Figure 3-2.

The probability for any outcome is read simply by moving across from the point above the outcome to the value on the vertical scale.

Continuous Probabilities

In the sales volume example, we know that considering only three sales levels is really a simplification of the problem. The sales volume could take on any of a great many values in a range from, say, 100 to 2500 units. If this is the case, and we wish to know the probabilities that the outcome will be anywhere in that range, then we require a continuous "probability distribution" over that range rather than discrete probability values. The result of assessing a probability distribution is shown in Figure 3-3. In this representation, the probability of any smaller range of values, such as sales from 1000-2000 units, is found by measuring

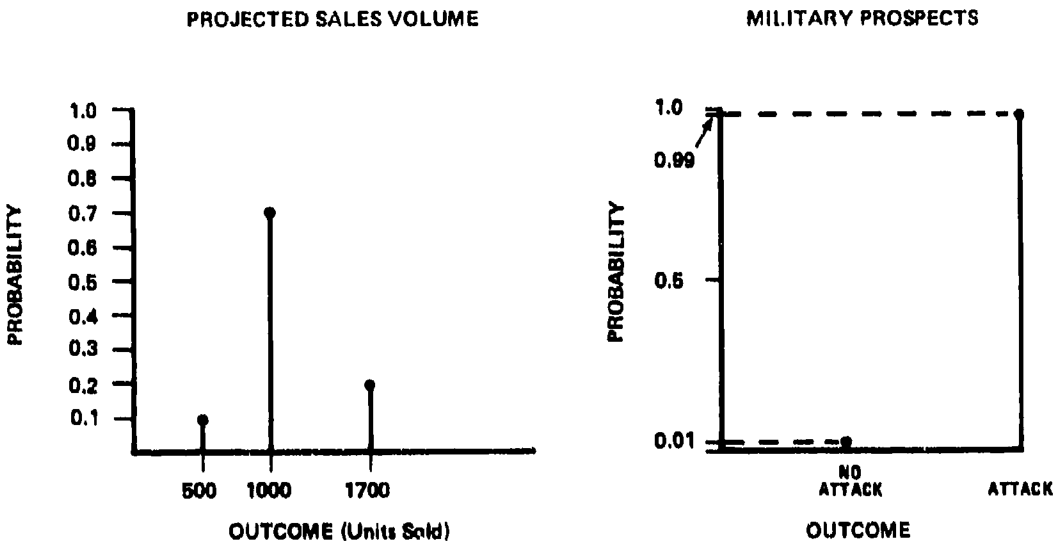


Figure 3-2
GRAPHIC DISPLAYS OF PROBABILITIES

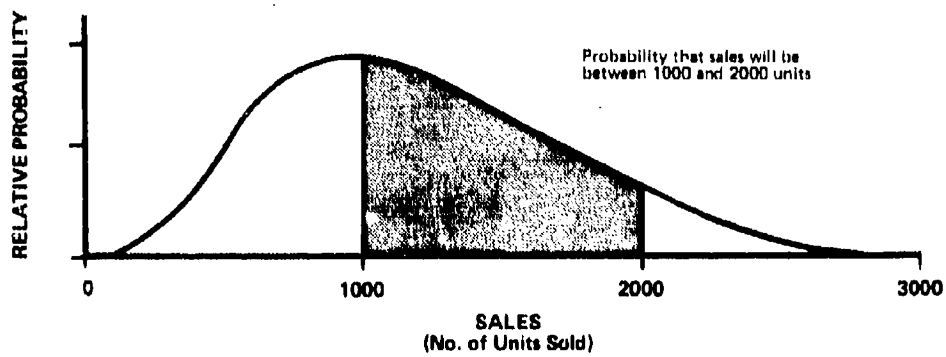


Figure 3-3
PROBABILITY DISTRIBUTION FOR PROJECTED SALES VOLUME

the area under the curve in that range. The total area under the curve is equal to 1.00.

Another way to represent a probability over a range is to draw what is called a "cumulative probability distribution." In this formulation, shown in Figure 3-4, the quantity plotted on the y (vertical) axis is the probability that the outcome will be equal to the corresponding value on the x (horizontal) axis or less. For instance, in the example shown, by reading up from the sales value of 1000 units to the curve and across to the cumulative probability scale (the y axis), we see that the probability that sales will be

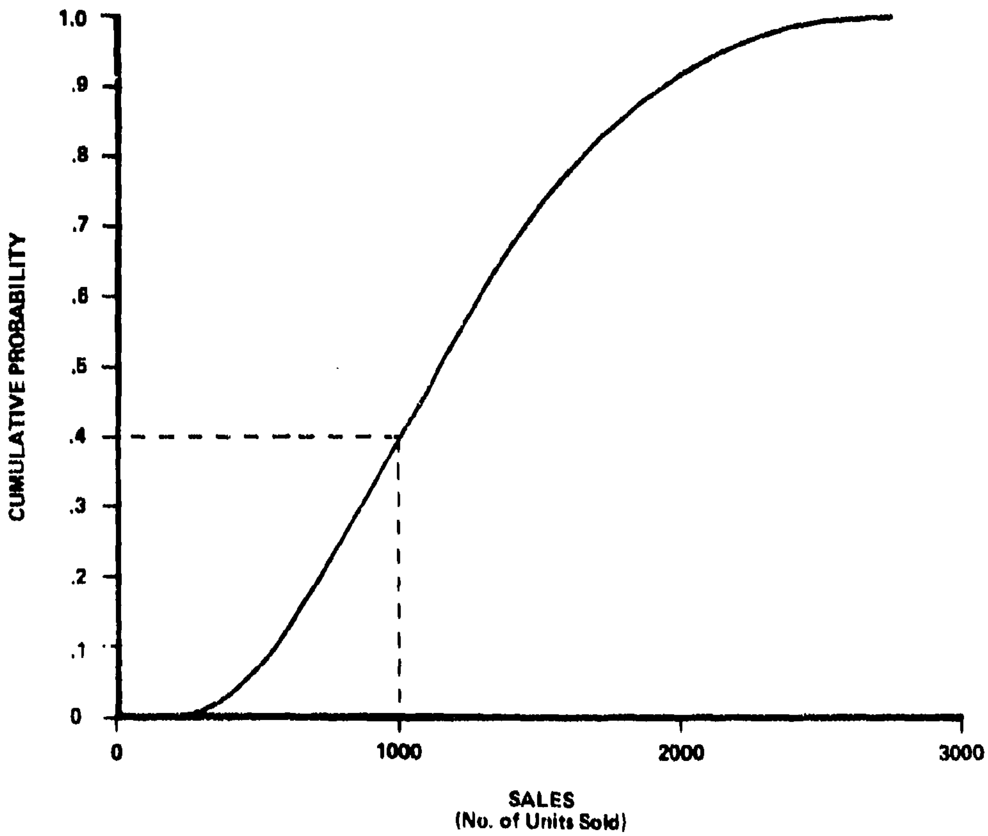


Figure 3-4

CUMULATIVE PROBABILITY DISTRIBUTION FOR PROJECTED SALES VOLUME

1000 units or less is equal to about 0.4. The cumulative probability distribution makes it easy to determine the probability that sales will be between 1000 and 2000 units. First, the cumulative probabilities up to those values, 0.40 and 0.92, respectively, are read off the graph. Then the smaller value is subtracted from the larger, 0.92-0.40=0.52, or a 52 percent chance that sales will be between 1000 and 2000 units.

These two ways of showing the continuous probability, the probability distribution and the cumulative probability distribution, are simply alternative ways of presenting the same information, and data presented in one form can be transformed into the other. For some purposes, one way is more convenient, for other purposes the other. It is easier to see the probability values from the cumulative curve since they are read directly from the y axis, but most people find the probability distribution in Figure 3-3 more intuitively understandable. Mathematically speaking, the cumulative probability distribution is the integral of the probability distribution.

When we assess discrete probabilities we are estimating probabilities for each of the outcomes; when we assess continuous probabilities, we are trying to establish the shape of the curve over the different outcomes.

Quantitative Expressions of Uncertainty

We cannot, unfortunately, directly translate semi-quantitative expressions of uncertainty such as "very likely" into probability values. The reason is that while most people would agree that "very likely" means a probability greater than 0.5, there would be no general agreement about how much more than 0.5 the probability is (0.8? 0.9? 0.99?). This lack of agreement in the use of such terms has been shown in a number of experiments (and can be easily verified by the reader by surveying a few friends).

For example, an intelligence analyst, a professional in the art of reasoning about the plausible, was asked to substitute probability estimates for some of the verbal qualifiers in an article he had written. The first statement was: "The cease-fire is holding but it could be broken within a week." The analyst said that he meant there was a 30% chance the cease-fire would be broken within a week. Later, an analyst who had helped the original analyst prepare the statement said she thought that there was an 80% chance that the cease fire would be broken. Yet, both analysts had previously believed that they were in agreement about what could happen.

A more systematic experiment along this line was conducted by NATO intelligence analysts who were concerned about the problem of communication. Several different sentences were constructed in the following manner. "It is highly likely that the Soviets will invade Czechoslovakia," or "It is almost certain that the Soviets will invade Czechoslovakia," or "We believe that the Soviets will invade Czechoslovakia." The basic structure of all sentences remained constant; only the verbal qualifiers changed.

Twenty-three officers, ranking from squadron leader to lieutenant general, served in the experiment. All participants were familiar with reading intelligence publications. They were asked to indicate the probability (in percent) that they would attribute to each message if they were to read it in that form in an intelligence article. The results of that experiment are displayed in Figure 3-5. Each dot represents one reader's probability assignment. While there were some phrases, particularly "better than even," which were interpreted in much the same way by most of the readers, it can be seen that, in general, there was tremendous variation.

In 1964, Sherman Kent attempted to mitigate this problem within the intelligence community by proposing a scale range of probabilities for various qualifying phrases. These ranges are indicated by the shaded areas in Figure 3-5. Clearly, the readers in this experiment were not using the Sherman Kent scale even though they were familiar with it. Even if the officers had been using the scale, the wide ranges for some of the phrases would still permit considerable variation in interpretation.

This and other demonstrations provide strong evidence that significant miscommunication can and does occur among analysts. Seeming agreement among analysts on an item being discussed could, therefore, be an artifact induced by the impreciseness of the verbal qualifiers. In reality, there could be considerable disagreement.

The fact that verbal qualifiers are imprecise descriptions of levels of certainty provides strong motivation for the use of a quantitative language of uncertainty. Appropriate use of numbers allows uncertainty to be expressed with precision. If a manager uses numbers, either percentages or odds, to convey a degree of belief about the likelihood of future events, and the numbers are carefully chosen to reflect the manager's uncertainty, then these numbers can be easily converted into probabilities.

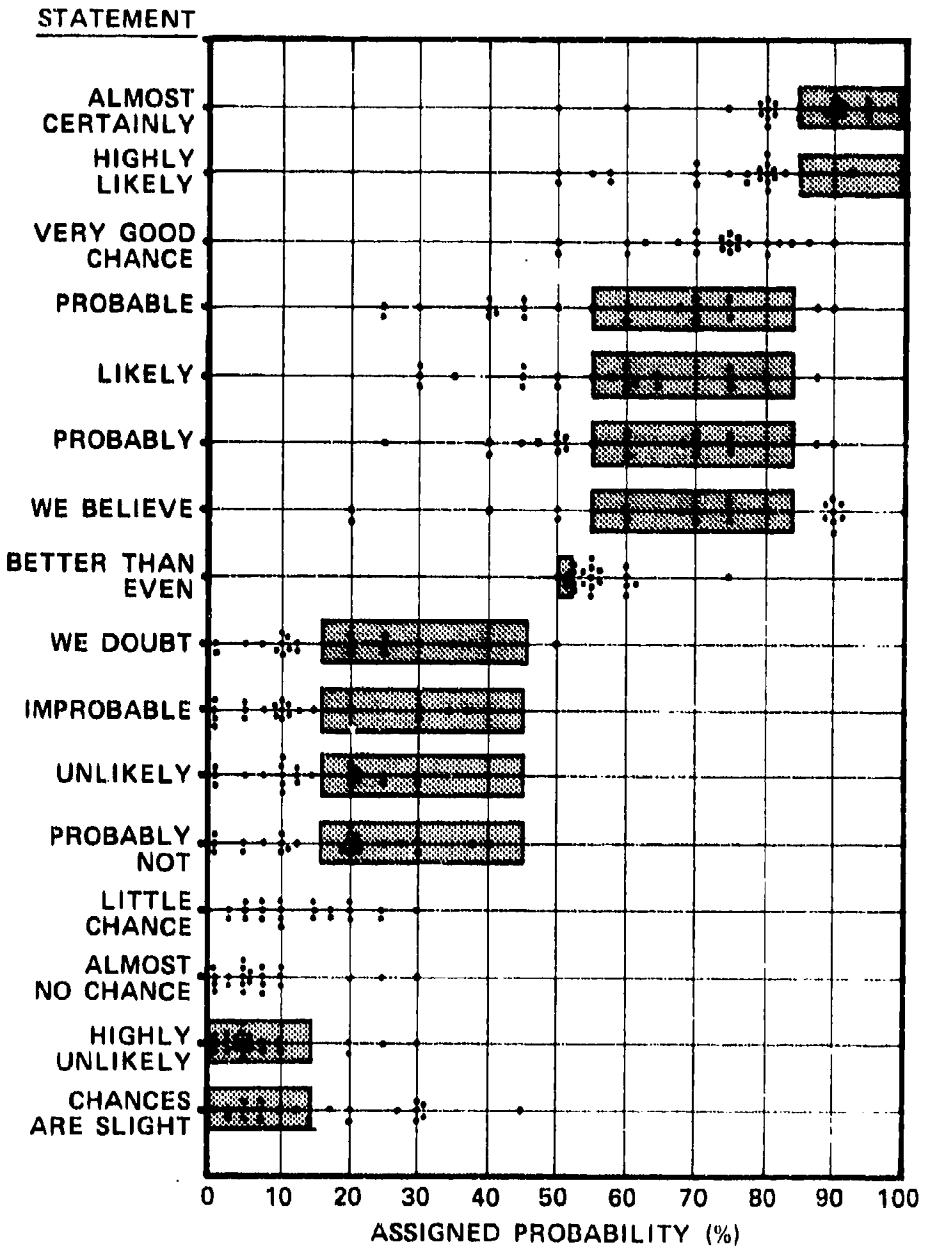


Figure 3-5

WHAT UNCERTAINTY STATEMENTS MEAN TO DIFFERENT READERS

Percentages

Percentages translate directly into probabilities simply by changing the scale for percentages running from 0 through 100 to a probability scale going from 0 through 1.00. Translation is easily accomplished by dividing the percentage figure by 100. For example, if a manager believes that there is an "80-percent chance" that the competition will respond to his new advertising campaign by increasing their own advertising expenditure, then he may be understood to mean that the probability of such a response is 0.80 (80 percent). We may then show those probabilities on the relevant branches of a fork in a decision tree, as in Figure 3-6.

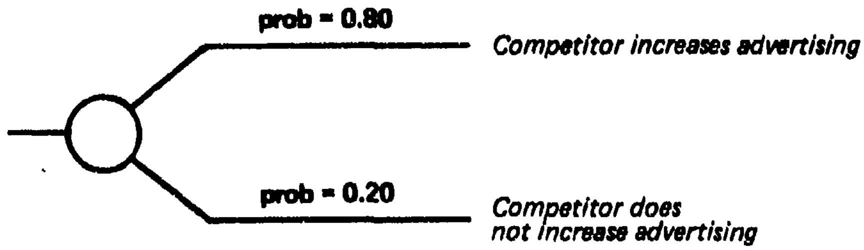


Figure 3-6

PROBABILITIES ON THE BRANCHES OF AN EVENT FORK

Odds

Some managers are accustomed to quoting odds to express their feelings of uncertainty. These values can also be directly converted to probabilities. Suppose, for instance, a Civil War military analyst had said before the battle at Gettysburg that the odds were 3 to 1 in favor of a victory by the Union troops. This statement of odds is equivalent to saying the probability of victory is 0.75. In other words, 3-to-1 odds implies that the probability of winning must be three times the probability of losing. Furthermore, the total of the probabilities of winning and losing must sum to 1.00. The only numbers between 0 and 1 satisfying these two conditions are 0.75 and 0.25. We can see that an odds statement is an expression of the ratio of the probabilities of the two outcomes being considered. In this example, the ratio is 0.75/0.25 or 3/1 or 3 to 1. Algebraically, the rule for converting odds to probabilities where there are two outcomes can be stated as follows:

If the odds are M to N, then the corresponding probabilities are:

$$\frac{M}{M + N} \quad + \quad \frac{N}{M + N} \quad = 1.00.$$

In the example above this becomes:

$$\frac{3}{3 + 1} \text{ (i.e., 3/4 or .75)} + \frac{1}{3 + 1} \text{ (i.e., 1/4 or .25)} = 1.00$$

Another common way of discussing uncertainty is to say that the chances are so many "in" or "out of" some total. For example, a manager may say "the chances are 1 in 100" of something occurring, or in another instance, "9 out of 10" that a particular outcome will occur. Rather than adding up the two figures in the statement to get the denominator for the probability, as was the case for the "odds" expressions, the second value is taken directly to be the denominator, as shown below:

Chances	Converts to	Probability
1 in 100		1/100
9 out of 10		9/10

Probability Assessment Techniques

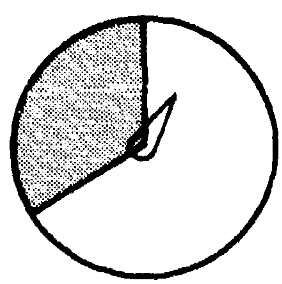
When uncertainties are not stated as odds or percentages, some more complicated, laborious procedures must be used to elicit feelings of uncertainty in numerical terms. A great deal of research in this area has been carried out to develop and test various probability assessment techniques. Among the more frequently used methods are:

Probability Wheels,
Reference Lotteries,
Relative Probability,
Quartile Assessment,
Trisection,
Other Fractile Assessment,
Cumulative Distribution Assessment.

The first three methods are used to assess discrete probabilities; the others are techniques for assessing continuous probabilities. The operation of the techniques is described below.

Probability Wheel

The probability wheel measures probabilities for people who are uncomfortable (usually because they object to the implication of precision) with assigning numbers to uncertainty. The wheel consists of a disk shaded in two colors, with a spinning pointer. The colors, blue and orange, can be adjusted so that the orange portion goes from being a tiny sliver, 1 percent, all the way up to 99 percent of the disk. The probability assessor who is trying to assess his feelings about, for example, the probability that the monarchy in Spain will be overthrown next year is asked to adjust the amount of orange (either decreasing or increasing it) until the wheel is configured so that the assessor believes that it is as likely that the monarchy in Spain will be overthrown as that the pointer on the wheel when spun will stop on the orange color. The assessor then looks at the back of the wheel where a scale shows what proportion of the wheel is orange. This value is taken to be the probability that the wheel will land on orange and is taken to be the assessor's probability that the Spanish monarchy will be overthrown.



Reference Lottery

The reference lottery is a technique that helps an assessor to focus fairly narrowly on his feelings about a discrete probability value, particularly when only two possible outcomes of the event are being considered. Suppose, for example, that a political analyst would like to estimate quantitatively the probability that a Republican candidate will be the winner of the next U.S. presidential election, as shown in Figure 3-7A. An estimate of that probability can be made by considering the following imaginary situation: First, the assessor is asked to think of a prize that he would like to have, such as a trip to Greece, series tickets for the local sports arena, a Picasso etching, and so on. Then a lottery is hypothesized where a single ball will be drawn blindly from an urn containing 100 balls, some red and

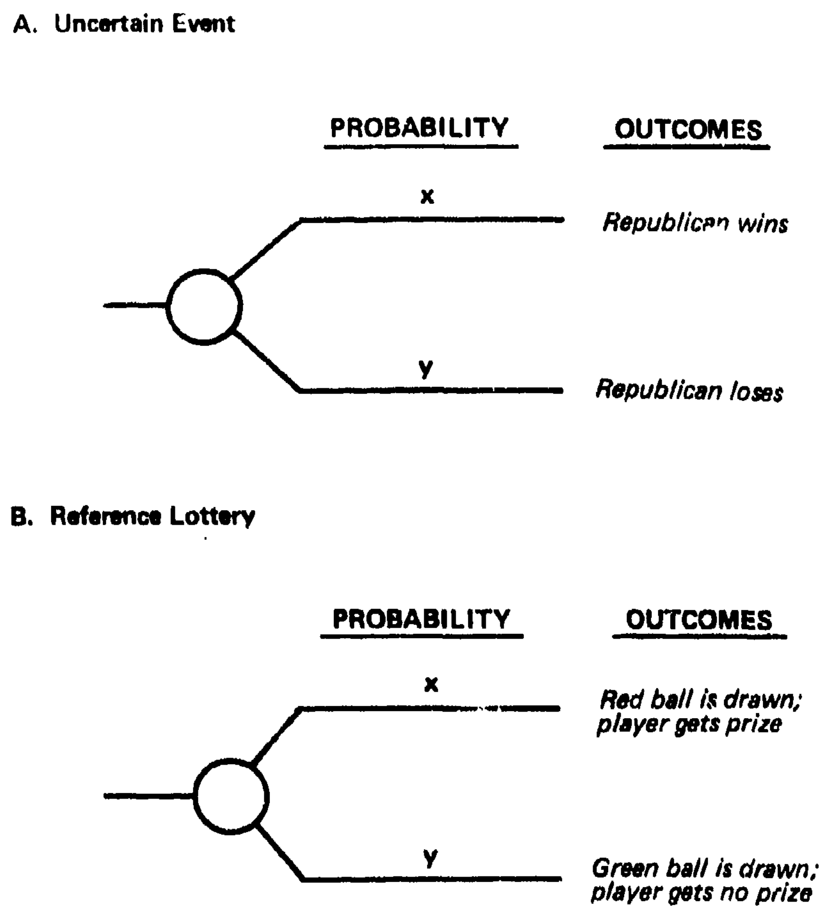


Figure 3-7

AN UNCERTAIN EVENT AND A REFERENCE LOTTERY USED TO ASSESS THE PROBABILITY OF THE UNCERTAIN EVENT

some green. The drawing will be made as soon as the winner of the election is known, but not before. Now suppose there are 60 red balls in the urn. The assessor is then asked if he would rather have the opportunity of letting his prize be contingent on drawing a red ball in the reference lottery <u>or</u> contingent on having a Republican winning the presidential election. If the assessor selects the reference lottery, where the probability of getting the prize is 60/100 or 0.6, then we conclude that his assessment of the probability of a Republican election winner is less than 0.6. By changing the composition of the lottery urn so that it contains, say, 40 red balls out of 100, we can find out if he still prefers the lottery to the election outcome. If so, his assessment of the probability of a Republican win is less than 0.4. We continue to change the composition of the urn until we have reached one such that the assessor is indifferent between getting the prize by winning the lottery or by having a Republican winner in the presidential election. The probability of winning the lottery (taken directly from the composition of the urn) is then equal to the assessor's personal probability or degree of belief that a Republican will win the election.

Relative Likelihood

When the discrete probabilities of an event which has several outcomes are considered, it is often helpful to ask the assessor to make a number of judgments about the relative likelihood between pairs of outcomes. In the simplified version of the "sales level" example where three outcomes (500 units, 1000 units, and 1700 units) were specified, the assessor could be asked the following questions:

> How much more likely is it that sales will be 1000 units rather than 500? Twice as likely? Five times as likely? Less likely?
>
> What about the likelihood of 1000 units compared to 1700 units?
>
> And 500 units compared to 1700 units?

After these relative relationships are elicited and then discussed and reassessed to eliminate any inconsistencies,[2] the values obtained (which are actually a form of odds) are converted to probabilities.

Quartile Assessments

The foregoing methods are designed to assess discrete probabilities. In the case of continuous probabilities, however, we need to develop an entire probability distribution or curve. Most techniques for assessing probability distributions consist of ways of obtaining a few points on the curve and then sketching in an approximation to the rest of the curve. The general name for one group of these techniques is "fractile assessment," and depending on the points being estimated, the specific techniques are called

[2]It would be inconsistent, for example, to say that A is twice as likely as B, B is three times as likely as C, and C is half as likely as A, since the latter would imply that A is twice as likely as C, which, by the first statement, makes B and C equally likely, thus contradicting the second statement.

quartile assessment, trisection, extremes assessment, and so on. How the probability information is elicited to get these points is most easily shown by an example.

Consider the case of a manager responsible for a research group who wants to assess a probability distribution for what dollar amount will be allocated to R&D in the next budget period. In other words, he would like to quantify the likelihood of various different possible budget amounts in some range. Suppose the range of possible values is from seven million dollars to fifteen million dollars. One way to elicit this probability information is by assessing three quartiles (Q_1, Q_2, Q_3) which are points that divide the range into four equally likely parts (or quarters). Notice the emphasis on the words "equally likely." These points will not normally divide the range into equally wide parts (see Figure 3-8A), but rather into segments such that the actual budget amount is just as likely to be between the minimum and point Q_1, or between Q_1 and Q_2, or between Q_2 and Q_3, or between Q_3 and the maximum.

Remember that for a continuous probability distribution the probability for any range of values is found by measuring the area under the curve between end points of the range. Since the quartiles divide the whole range into four equally likely parts and the total probability is 1.00 or 100%, the probability for each of the segments must be 0.25 or 25%. Information about what the quartiles are, therefore, is sufficient to allow us to sketch in (either by eye or by using a computer program) the approximate shape of the probability curve (see Figure 3-8B). The closer together the quartiles are, the more peaked the curve will be (Figure 3-8C). When the assessed quartiles are far apart, the curve will be flatter (Figure 3-8D). The location of these quartiles is determined by asking the decision maker to respond to questions like the following:

(1) What is the value such that you think the R&D budget is just as likely to be above as to be below this figure? (This number is Q_2, the second quartile because it divides the range into two equally likely parts, so each of these parts will have a probability of 0.5 or 50%. This is the probabilistic midpoint of the range.)

Suppose the decision maker has given "$10 million" as his estimate of Q_2. He is then asked:

(2) Now consider values that are less than $10 million. What single value divides this part of the range (the minimum to $10 million) into two equally likely parts? (This new value will be Q_1.)

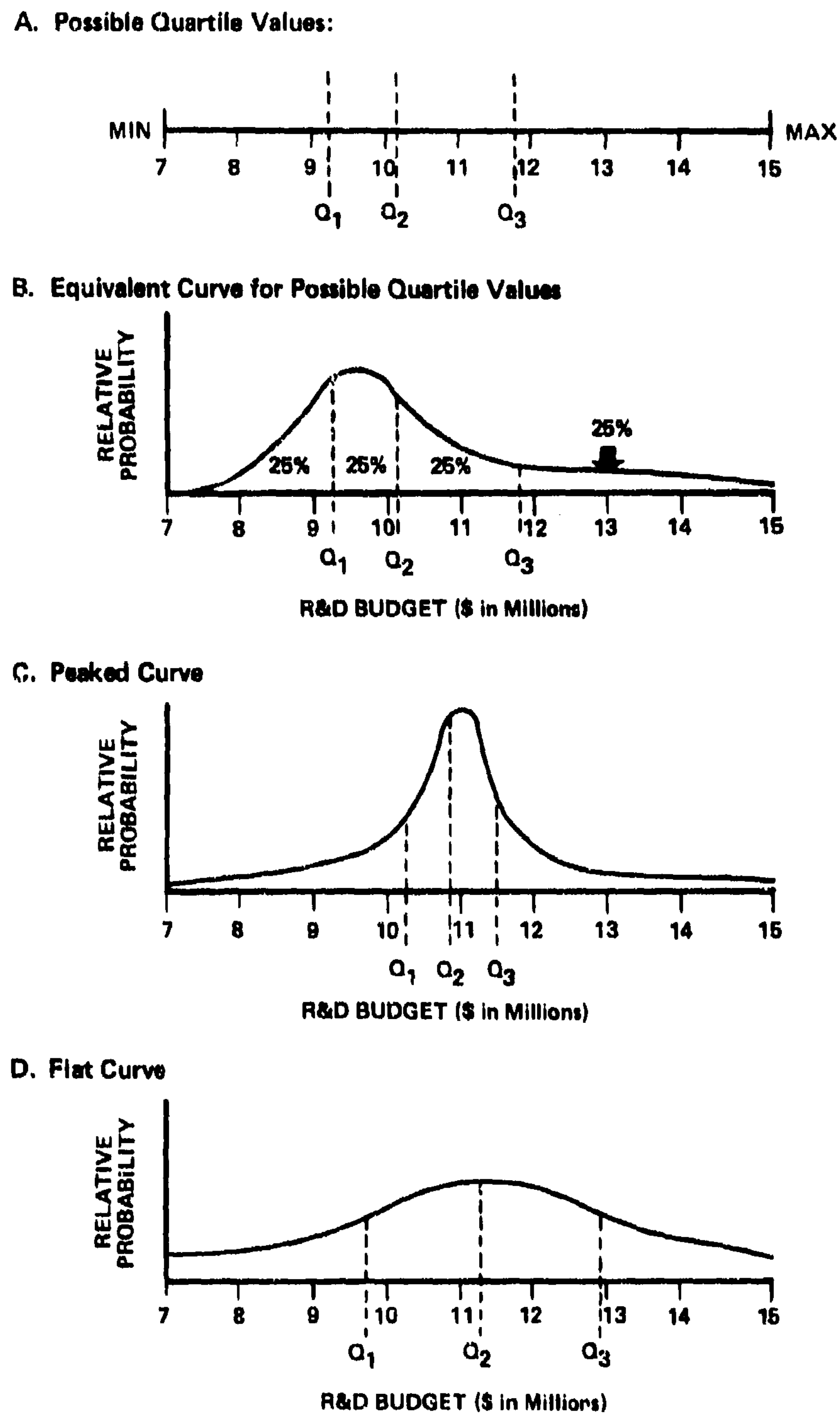


Figure 3-8
QUARTILE ASSESSMENTS

The value of Q_3 is found in a similar manner by looking at the range from Q_2 to the maximum. In the example shown in Figure 3-8C, these three quartile values assessed turned out to be approximately:

Q_2 = \$10.8 million,

Q_1 = \$10.2 million,

Q_3 = \$11.5 million.

Trisection

An alternative to assessing the quartiles of a probability distribution is to ask the decision maker to estimate the two points (T_1 and T_2) that divide the range into three equally likely parts. This process is called "trisection," and the resulting segments of the range each have probabilities equal to 33 1/3 percent of containing the actual value of the variable whose probability is assessed, in this example, the budget (see Figure 3-9).

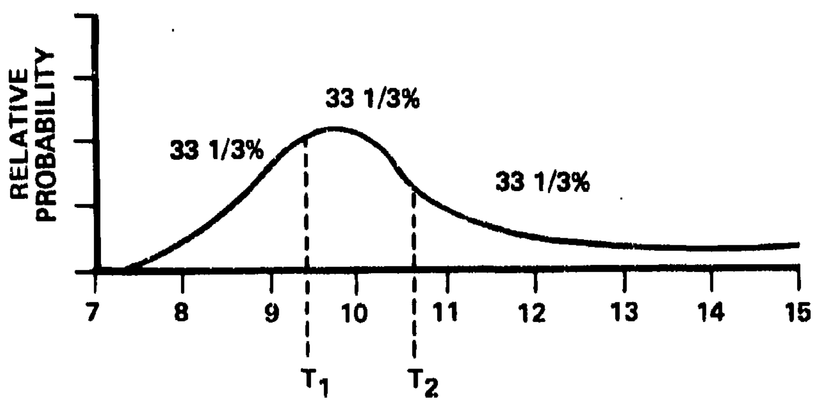


Figure 3-9
TRISECTION

Other Fractile Assessment

More than two or three points on the curve can be assessed if necessary for the problem being studied. The deciding factors here are how much detail or accuracy in the curve is desired and how much time and effort are to be spent on the probability assessment portion of the analysis.

For some problems, information about the extremes of the distribution may be relevant (in the budget example, for instance, an extreme would be the value such that there is only one chance in 100 of the budget being less than this value). There are special methods, not discussed here, that can be used to elicit this information.

Cumulative Distribution Assessment

A probability distribution can also be obtained by asking for estimates of points on the cumulative probability curve. Since the cumulative probability distribution is a graph of possible values of an uncertain quantity vs the probability of the occurrence of that value or something less, the questions asked to elicit points on the curve would be like those shown below. For the R&D budget example:

(1) What is the value such that you think there is a 0.5 probability that the budget will be this value or less?

(2) What is the value such that you think there is a 0.25 probability that the actual budget will turn out to be less than this?

(3) What is the value such that you think there is a 0.75 probability that the actual budget will be less than this?

Notice that the values elicited in this manner should be the same as the three quartiles (Q_2, Q_1, Q_3) assessed for the probability distribution. This procedure is simply another way of getting the same information about the assessor's uncertainty. For example, the value such that the assessor thinks there is a 0.5 probability that the budget will be this value or less (question 1 here) should be the same as the value such that the assessor thinks the R&D budget is just as likely to be above as to be below this figure (question 1 under Quartile Assessment). Figure 3-10 shows the probability distributions B, C, and D of Figure 3-8 redrawn in cumulative form. A comparison of these two figures shows that the more peaked the probability distribution is, the steeper will be the slope of the cumulative distribution curve. Despite the fact that the two types of curves are based on and display the same information, most people are more comfortable working with the probability information in the first form.

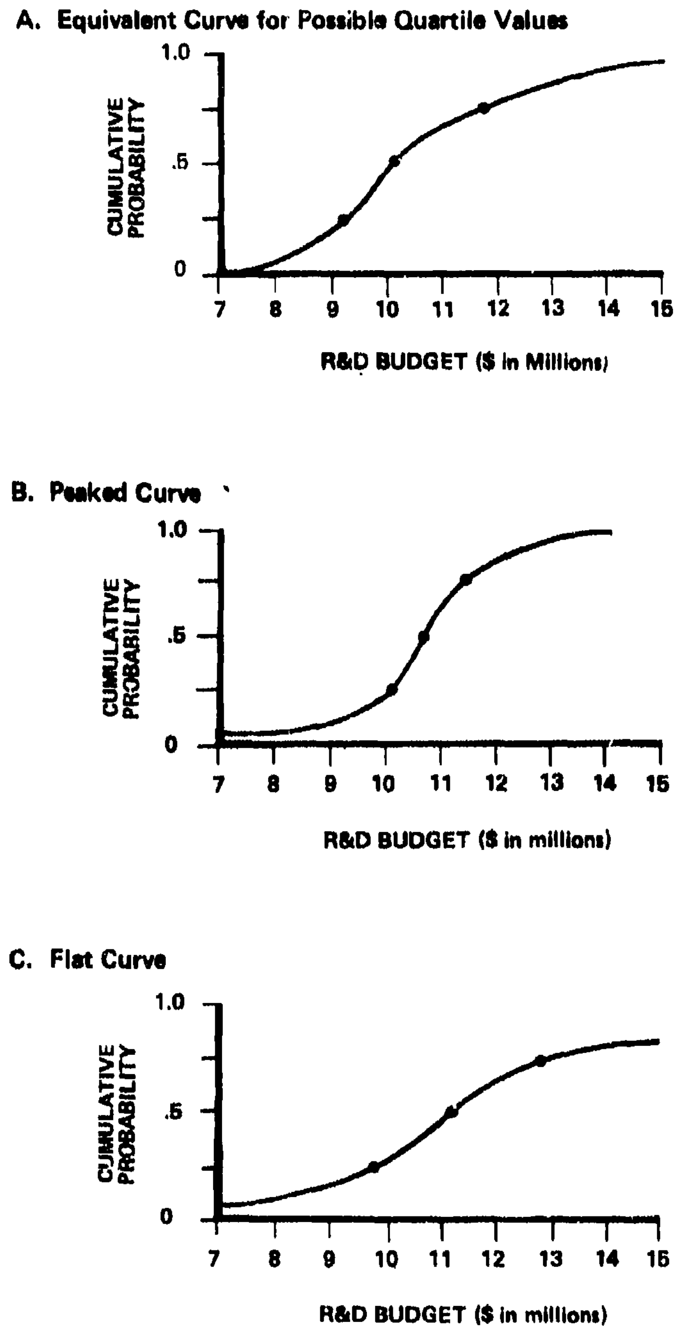


Figure 3-10
ASSESSED CUMULATIVE PROBABILITIES

Values of the Variable

In the methods discussed so far, the assessors are given specific probability statements and asked to estimate the value of the uncertain quantity which corresponds to the specified probability (e.g., "What is the value of the variable such that you think there is a 0.75 probability that the actual budget will be less than this?"). An alternative way to elicit the assessor's personal probabilities is to reverse the question by specifying the value or range of the variable and asking the assessor for a numerical estimate of the corresponding probability (e.g., "What do you think is the probability that the budget will be $11 million or less?").

Standard Mathematical Forms

The assessment of probability distributions can be simplified in some cases by assuming that the probability curve has some specific mathematical form. That is, the curve can be written according to a mathematical equation. When some points on the curve have been assessed, one can solve for the values of the unknowns in the equation (called the "parameters") and calculate mathematically the remaining probabilities of any points or ranges. The accuracy of the probabilities obtained in this way depends upon how good the initial assessments are and upon how realistic the assumption is that the desired distribution follows the mathematical form chosen.

Two examples of commonly used mathematical forms are the bell-shaped curve, known as the "normal" or "Gaussian" distribution, and the beta distribution. Typically, these distributions are assessed by specifying either two quartiles (the first and third) or the mean and variance of the distribution. If an interactive computer is used as an aid, the entire distribution can be computed immediately, and any selected ranges printed out.

Comparing Methods

Since the probability assessments obtained by different means often differ among themselves to varying degrees, depending upon which assessment methods are used, it is a good idea to assess the same probability in more than one way and then compare the values obtained. If they differ, some additional thought and discussion will usually enable the assessor to settle on a single value in which the assessor has more confidence than any of the initial values.

For example, a distribution assessed by trisection can be refined by means of the quartile technique. The two curves obtained this way are reconciled and then further tested by asking for probabilities associated with specific values of the variable to see whether these probabilities agree with those implied by the curve.

When time for eliciting probabilities by several procedures and for comparing results is limited, however, we must consider which one assessment procedure can be followed. The answer depends not only on the task (the initial decision in a tree may be so insensitive to the shape of the probability distribution that a very rough sketch based on trisection may be adequate), but also on the assessor. If he is unfamiliar with probabilities and reluctant to make numerical judgments of probability, then quartile or trisection assessment methods should be used. If he is comfortable with probability assessment, then the method of specifying values and asking for assessments of cumulative probability results in a distribution that is usually less biased than the distribution that results from quartile or trisection assessment. If the extreme ends of the range of the distribution are important to the decision, then interactive assessment with a computer[3] which uses standard mathematical functions allows their careful examination.

Updating Assessments

Probability assessments are designed to express quantitatively the assessor's current feelings of uncertainty about some outcome. In other words, they reflect the assessor's present state of knowledge. As the assessor gains additional information about the uncertain quantity or event, the probabilities may change. For example, suppose an intelligence analyst has, after careful consideration of many factors, assessed the probability of a coup in Freedonia as being equal to 0.60. Then, after making this estimate, the analyst learns that the children of the Freedonian President are traveling to London for a short vacation. The analyst would like to incorporate this new data into the probability estimate. It is easy to tell in a general way how such information would alter the probabilities. In the case of an event like the one above having two possible outcomes, it is natural to expect that if the additional information is in greater accord with the likelihood of one outcome than with that of the other, then the probability of the one will be increased, and vice versa.

---

[3]An extensive suite of programs has been developed at Harvard University. See Schlaifer, Robert, Computer Programs for Elementary Decision Analysis. (Cambridge, MA: Harvard Univ. Press, 1971).

To compute precisely the appropriate amount of change in the probability is a more complicated matter. An early and useful development in the theory of probability was Thomas Bayes' proposal of a formal, mathematical procedure for incorporating new information to revise a probability estimate. This procedure, called "Bayes' Theorem," is explained in detail and illustrated in Appendix A, and the application of that theorem in cases having multiple value dimensions is presented in Appendix B. In the case of discrete probabilities, Bayes' Theorem can easily be applied to revise the originally assessed probabilities, provided the new information is expressed in, or can be converted to, numerical form. For continuous uncertain quantities, on the other hand, the degree of difficulty in applying Bayes' Theorem depends upon the particular mathematical formulation of the curve.

When using Bayes' Theorem, the initially assessed probability is called a "prior" probability and the revised or updated probability is called a "posterior" probability (meaning that it is the revised probability after the new information is received). The form in which the new data is expressed is called a "likelihood ratio." In the example cited above, the analyst would have to express his feelings about the implications of the new data in this way:

> "What is the likelihood that the President's family would leave the country if a coup were imminent?"

and to determine the relevance of this new information to the other possibility,

> "What is the likelihood that the President's family would travel abroad in ordinary times?"

Bias in Assessments

A probability assessor must devote a certain amount of time and effort to learning how to use the appropriate assessment technique or techniques for his problems before useful probability values can be obtained. Even after the mechanics of the method are understood, however, research has shown that assessors typically are prey to a number of systematic errors or biases in making their assessments. Accordingly, to obtain the best assessments, some means must be developed for eliminating or correcting these biases. Ideally, the probability assessment should reflect as closely as possible an assessor's true feelings of uncertainty. Events given high probability values should occur frequently and those with low probability values less frequently. One way to check assessments for this characteristic is to collect a large number of assessments from the same assessor and see whether the assessed probability corresponds to the frequency of occurrence. For example, consider the case of a weather forecaster who has assessed precipitation probabilities for several months; looking back over his records, we would expect that if the assessments are good estimates of the forecaster's uncertainty, then it would have rained on about 60% of the days for which the forecaster said the probability of rain was 0.6. A probability assessor who exhibits this characteristic is said to be a "good" assessor or "well calibrated."

Assessors who are not well calibrated, on the other hand, may exhibit any of several types of biases. For example, they may be overconfident, that is, too sure of themselves and consequently assess probabilities that are always too extreme (too close to 1.00 or 0.00) for discrete events, or assess curves that are too narrow or tight for continuous probabilities. Some assessors, on the other hand, may be systematically underconfident in their assessments. Another bias observed is the phenomenon of availability, the tendency of people to assess probabilities according to the number of instances of the event that they remember and how easily they recall them. These factors may or may not be related to correct probability of occurrence. For example, the letter k is much more likely to occur as the third letter of a word in the English language than as the first letter, yet most people judge it as more likely to be a first letter. Presumably, this is because in making such a judgment, people try to think of words either starting with k or having k as a third letter. It is easier to think of words beginning with k, and, therefore, they are judged to be more frequent.

There are two ways of overcoming the bias in probability assessments. One is to observe an assessor long enough to discover which particular biases he exhibits and to what degree. Then a correction factor is computed and applied to any future assessments to eliminate the bias. The more usual practice, however, is to explain the kinds of bias and to train the assessor to avoid them so that he becomes well calibrated. This training consists of repeated assessments with feedback presented in the form of a "score" that tells the assessor how well he is doing compared to a perfectly calibrated assessor. Appendix C presents a discussion of scoring rules for training probability assessors.

Bias also appears when people are asked judgmentally to revise an initial probability estimate on the basis of new information. This bias is eliminated, however, by following the formal updating procedure of Bayes' Theorem.

CHAPTER 4

PROBABILITY DIAGRAMS AND HIERARCHICAL INFERENCE

When determining the probability of an event that is related to other events, an assessor will often find that it is easier to judge likelihoods for the related events than it is to assess the probability of the event of interest. Probability diagrams provide one of the most useful and straightforward techniques for decomposing the problem into its parts and assessing those probabilities. Instead of trying to assess the overall probability of an event or outcome, the analyst can use probability diagrams to assess the chance of that event by first assuming some of those related events have occurred and then by assuming those related events have not occurred.

Consider the following problems in which it is more effective to assess the components than it is to assess the total problem:

1. The probability of rain tomorrow will depend on several conditions, barometric pressure and prevailing winds.

2. The probability of new international monetary reform taking place this year will depend mainly on whether or not countries A and B both agree to support the proposed changes.

3. The probability of a coup in a particular country may depend jointly on three factors: the state of the economy, the country's support of a particular general, and threat of aggression from a neighboring country.

4. An analyst is interested in determining the probability that country X will launch a major attack across the border of country Y. He has come to the conclusion that X is more likely to attack if the leader of Y dies <u>and</u> if there is political upheaval.

Consider problem #1, determining the probability of rain occurring tomorrow morning. Since the possibility of rain will be dependent on several uncertain conditions like barometric pressure and prevailing winds, it will be difficult to assess the overall probability by looking at the conditions from a global viewpoint. A simpler method is to decompose the problem into separate components and to assess the implications of each section individually. Then it is possible to reconstruct the problem by applying the laws of probability to the decomposed assessments. The key is to attack the problem from its integral parts and to rebuild rather than to estimate an overall probability from a global viewpoint.

Structure

Once the analyst has decomposed the problem into those components he considers relevant, he can use a probability diagram to structure the problem. For example, if our weatherman were concerned with the probability of rain, he could structure a simple probability diagram (see Figure 4-1) with two possible outcomes, or target variables, that is, rain or no rain, with probabilities P_1 and P_2, respectively.

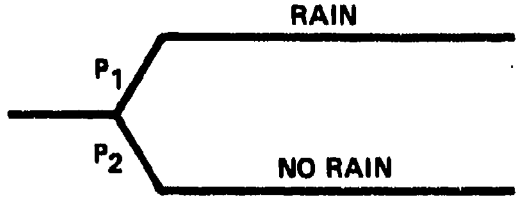


Figure 4-1
SIMPLE DIAGRAM

If the weatherman were now asked to assess the overall probabilities, P_1 and P_2, without knowing what other events will occur, he will have difficulty making the assessment. If he feels that the probability of rain will increase or decrease depending on whether the winds will be from the NW or from the SW, he can restructure his probability diagram as in Figure 4-2.

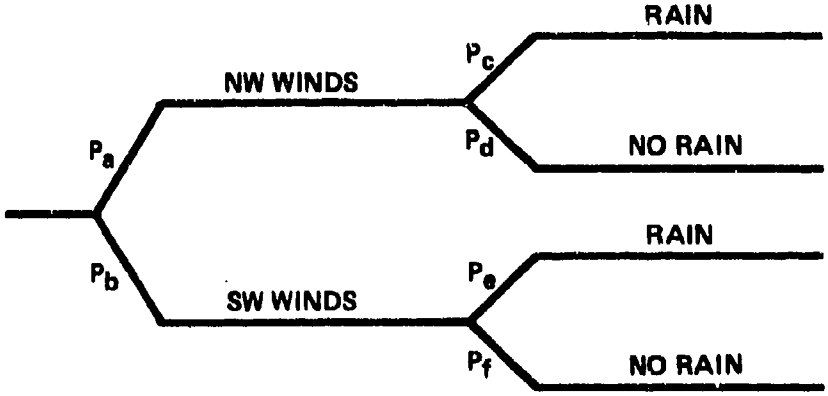


Figure 4-2
PROBABILITY DIAGRAM

Notice that the far right-hand branches of Figure 4-2 still have the same target variables, rain and no rain, as do the branches in Figure 4-1. He is still concerned with the same outcome; only now he bases his estimate of the likelihood of that outcome directly on wind condition. In order to assess the likelihood of rain for P_c, he can first assume that the winds will be from the NW, and assuming NW winds, he can more easily determine the probability of rain. Notice that it would be much easier to assess P_c in Figure 4-2 than it would be to assess P_1 in Figure 4-1.

In a similar manner, if the weatherman feels wind direction will be influenced by the existence of a low- or high-pressure area and could restructure his probability diagram accordingly, the result would appear as in Figure 4-3.

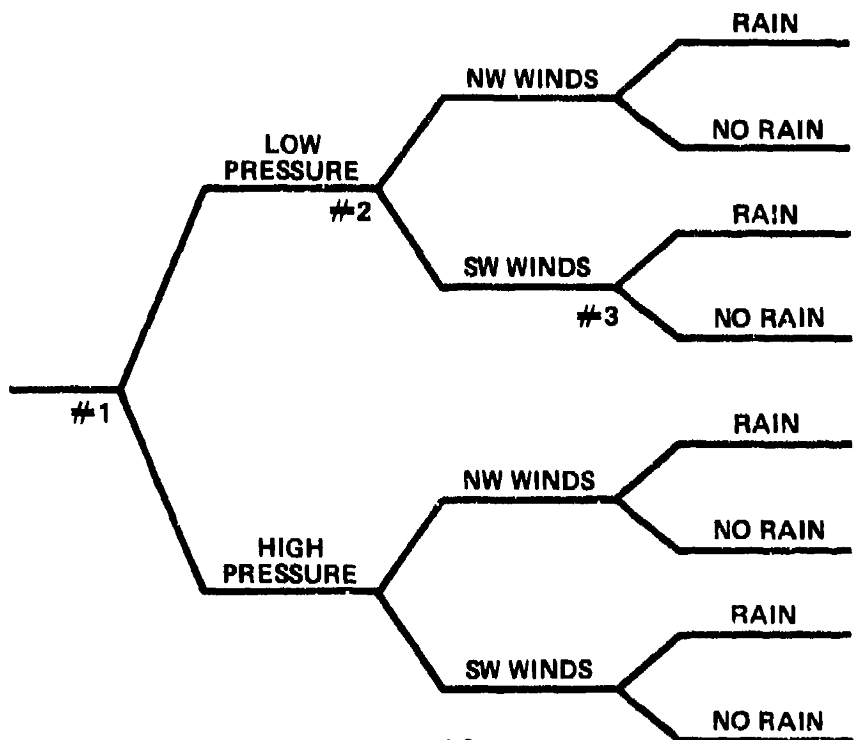


Figure 4-3
PROBABILITY DIAGRAM

If barometric pressure and wind direction were the only conditions affecting the likelihood of rain, then Figure 4-3 would be a complete probability diagram for this problem.

The overall diagram of this problem has been constructed in the shape of a tree with each of the possible sets of outcomes represented as a separate branch of the tree. Progressing to the right, an additional assumption has been made each time a point is reached from which branches emerge, this point being called a node. For example, if the analyst follows the numbered branches of Figure 4-3 at the upper branch of node #1, he assumes that a low-pressure area exists. At the lower branch of node #2, he assumes that in addition to a low-pressure area, the winds are from the SW. Finally, from the upper branch of node #3, he assumes that the likelihood of rain is determined by both a low-pressure area and SW winds. Notice that for this branch it would be much easier to assess the likelihood of rain tomorrow by first assuming a low-pressure area and SW winds than it would be to estimate the likelihood of rain tomorrow without making any previous assumptions. The same is true for each of the other branches. We call probabilities arrived at by combining probabilities estimated for relevant components <u>conditional probabilities</u>.

Once the probabilities have been estimated for each of the eight branches, the weatherman can total the probabilities for those branches resulting in rain and determine the original overall target variables, P_1 and P_2. We turn next to the rules that must be followed in constructing probability diagrams and assessing probabilities; we then discuss the rules used for combining the assessments.

Rules for Decomposition and Assessment

When a problem is decomposed into its relevant components and structured in the form of a probability diagram, it is necessary to assess the conditional probabilities for each of the branches. In order to do this correctly, it is first necessary to become familiar with the following rule:

Rule #1 All branches from a particular node must be <u>mutually exclusive</u>.

This means that a particular event cannot appear twice, as would be true of two different branches containing the same outcome. Consider a series of probability diagrams which represent the outcomes from the roll of a die. Figure 4-4 is incorrect because the branches are not mutually exclusive, but all three probability diagrams in Figure 4-5 show mutually exclusive branches and therefore are correct.

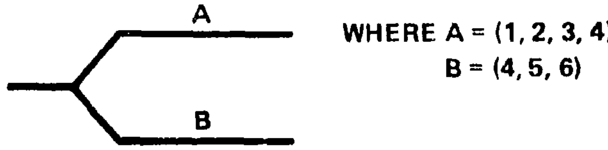


Figure 4-4

INCORRECT PROBABILITY DIAGRAM WITH
BRANCHES NOT MUTUALLY EXCLUSIVE

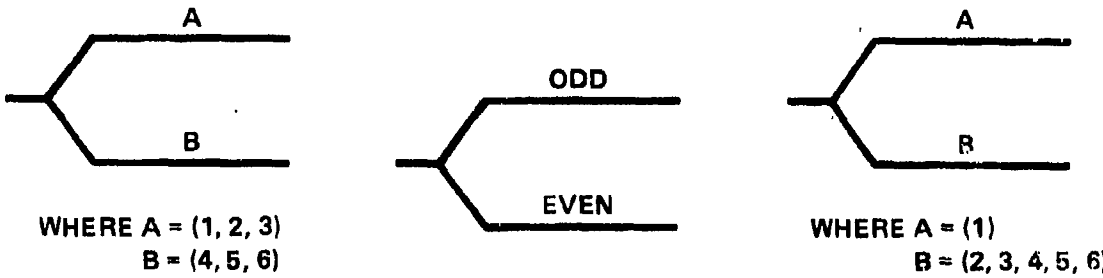


Figure 4-5

CORRECT PROBABILITY DIAGRAMS WITH BRANCHES MUTUALLY EXCLUSIVE

A second example represents a diagram by an analyst trying to determine the speed of a new Russian aircraft. The branches in Figure 4-6 are not mutually exclusive, and therefore are incorrect, while the branches in Figure 4-7 are mutually exclusive and therefore correct.

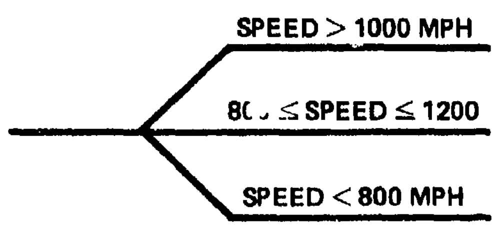


Figure 4-6

INCORRECT PROBABILITY DIAGRAM WITH
BRANCHES NOT MUTUALLY EXCLUSIVE

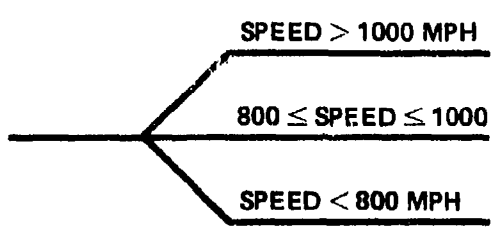


Figure 4-7

CORRECT PROBABILITY DIAGRAM WITH
BRANCHES MUTUALLY EXCLUSIVE

Rule #2 All branches stemming from a particular node must be exhaustive.

This means that every possible outcome must be included. In the case of the die problem, Figure 4-8 would be an incorrect representation because it does not include the case of rolling a 2.

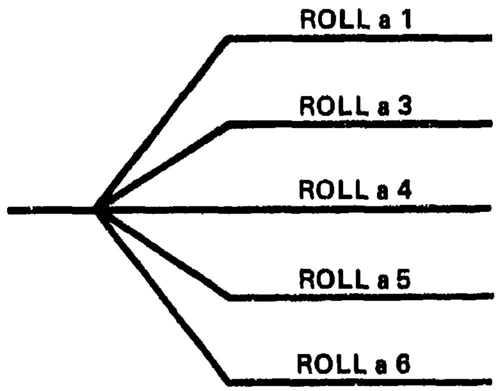


Figure 4-8

INCORRECT PROBABILITY DIAGRAM WITH BRANCHES NOT EXHAUSTIVE

Another example depicts the determination of the likelihood of the Soviet Union's targeting an ICBM at the U.S. or at Europe. Figure 4-9 is not exhaustive because it does not consider the possibility of the missiles also being aimed at another region of the world, as portrayed by the dotted line.

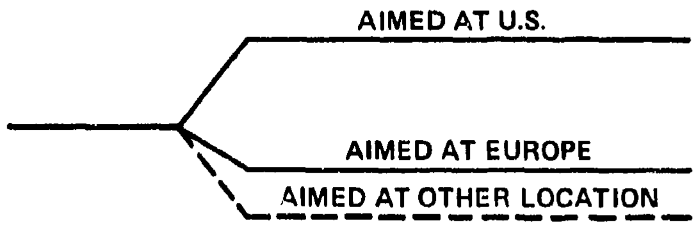


Figure 4-9

INCORRECT PROBABILITY DIAGRAM WITH BRANCHES NOT EXHAUSTIVE (CORRECT WHEN DOTTED LINE IS ADDED)

Rule #3 The probabilities branching from a particular node must sum to one.

This rule may serve as a check to verify whether the branches on the probability diagram are truly mutually exclusive and exhaustive. It also serves as a partial check on the consistency of the assessor's judgments, for if the diagram is correct, it is inconsistent to assign probabilities that do not sum to one.

Conditional Probabilities

Consider Figure 4-10, which refers to the original question concerning the likelihood of rain. When asked for estimates concerning the probability of rain from a global viewpoint, our weatherman was unsure, but estimated

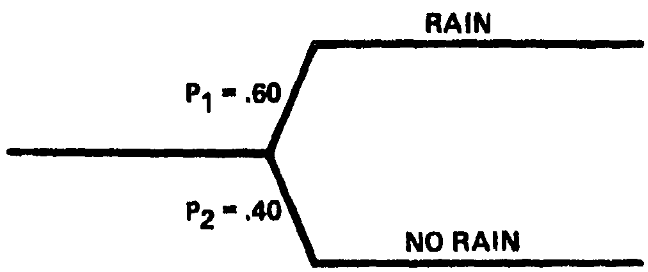


Figure 4-10
SIMPLE DIAGRAM

a 60% chance of rain. His estimate of $P_1 = .60$ is an unconditional probability, so called because it is not based on any previous assumptions, or branches, in the probability diagram.

Next, the weatherman considered the upper path of Figure 4-3. First, he assessed the probability of low pressure as 0.8; this is an unconditional probability. Then, he assessed the probability of NW winds after assuming that a low-pressure area occurs for sure. This is a conditional probability, which can be written:

$$p(\text{NW winds} \mid \text{low pressure}),$$

the vertical line indicating "given." The expression is read "the probability of NW winds given the occurrence of a low-pressure area." Turning to the upper right branch, the weatherman asked, "What is the probability of rain assuming NW winds and low pressure occur for certain?" This, too, is a conditional probability, and he assigned a value of 0.9:

p(Rain|low pressure and NW winds) = 0.9

Probabilities for the three events on the upper path are shown in Figure 4-11.

Strictly speaking, that first unconditional probability of 0.8 for low pressure is really a conditional probability; it is conditional on the person making the assessment, on that person's expertise, and on the information available to the assessor, like today's weather. In this sense, all probabilities are conditional. But to simplify communication, we rarely describe all the conditioning events, and when we do not explicitly describe any of them, we say that the probability is "unconditional."

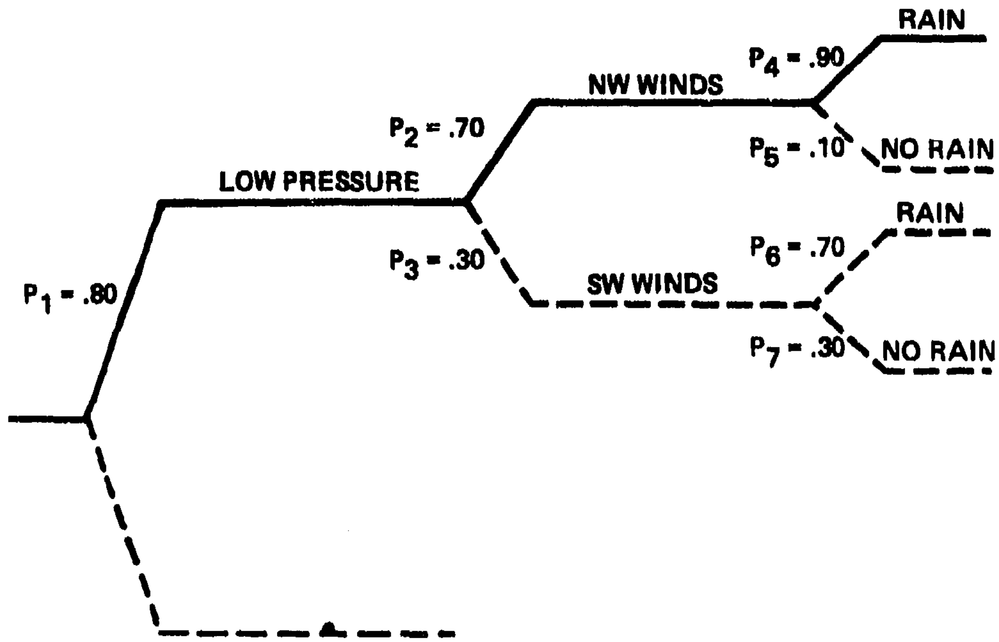


Figure 4-11

UPPER PATH OF FIGURE 4-3

In words, here is the description of each probability shown in Figure 4-11:

P_1 - Unconditional probability of a low-pressure area occurring tomorrow;

P_2 - Conditional probability of NW winds on the assumption there will definitely be a low-pressure area, or p(NW winds|low pressure);

P_3 - Conditional probability of SW winds on the assumption there will definitely be a low-pressure area, or p(SW winds|low pressure);

P_4 - Conditional probability of rain given low pressure and NW winds, or p(Rain|low pressure and NW winds);

P_5 - Conditional probability of no rain given low pressure and NW winds, or p(No rain|low pressure and NW winds);

P_6 - Conditional probability of rain given low pressure and SW winds, or p(Rain|low pressure and SW winds);

P_7 - Conditional probability of no rain given low pressure and SW winds, or p(No rain|low pressure and SW winds).

Next, the weatherman completes assessments for all the remaining branches. His probabilities are shown on the corresponding branches in Figure 4-12.

The assessment procedure can be summarized as a fourth rule:

Rule #4: In determining the probability for an event on a given branch of the probability diagram, the assessor must assume that all events on the path leading up to the given branch have definitely occurred.

Independence

No discussion of conditional probability is complete without considering the notion of independence. Since the definition of independence uses the idea of a conditional probability, we introduce the definition here even though we shall not use it until later in the chapter.

Definition: Two events are independent if the probability assigned to one event is unaffected by knowledge that the other event has occurred.

For example, to the event low pressure, we assigned a probability of 0.8. Presumably, that probability remains unchanged when the weatherman is told that his favorite football team has just lost to their arch rivals. If that is so, then the event "low pressure tomorrow" is independent of the event "Team X has lost to Team Y today." Here, independence is obvious because one event has no causal relationship to the other.

Lack of causality is not, however, an adequate basis for judging independence. It is quite possible for two events to have no causal connection, yet be judged non-independent. For example, without looking at my wristwatch, I judge there is a 0.6 probability that the time on my watch is within five minutes of 3:00 p.m. If you look at your watch and tell me it is 2:58 p.m., I now change my probability to 0.95 (your watch could be a little fast!). Since I have changed my probability for the event "3:00 p.m. plus or minus five minutes by my watch" when I heard about another event, the time shown on your watch, we conclude that the events are not independent. Yet, in no sense can one of our watches be said to "cause" the time shown on the other. The times shown on our watches are correlated, they "go together" in an orderly fashion, and it is this correlation that leads us to judge the events as non-independent.

In short, we cannot judge events to be independent because there is no causal relationship between them. Instead, we must ask if our <u>uncertainty</u> about an event is <u>changed</u> by knowledge about some other event. If that extra knowledge does not change our uncertainty, then the events are independent. If it does, then the events are dependent. This is the only adequate test of independence.

<u>Rules for Combining Probabilities</u>

We return now to the main problem at hand, that of finding the unconditional probability of rain. To calculate this from the assessments shown on the probability diagrams of Figure 4-12, we must apply two important rules of probability.

Multiplication Rule: To obtain the probability that all events on a path through the probability diagram will occur, multiply together all the individual probabilities on the path.

For example, the probability for the upper path is obtained by multiplying 0.8 by 0.7 by 0.9:

$$
\begin{aligned}
&p(\text{low pressure and NW winds and rain}) \\
&\quad = 0.8 \times 0.7 \times 0.9 \\
&\quad = 0.504
\end{aligned}
$$

That is, there is a 50.4% chance that all three events on the path will occur. This probability is called a path probability. Repeated applications of the multiplication rule gives all the path probabilities shown in the right column of Figure 4-12.

Now we are ready to determine the probability of rain. We need only one more rule.

Addition Rule: For mutually-exclusive events or paths, the probability that one or the other will occur is obtained by adding their individual probabilities.

For example, what is the probability that a fair die, when rolled, will show an odd number? Since there are six possible outcomes which are mutually exclusive, we can apply the addition rule. The probability of an odd number is the same as the sum of the probabilities of a 1, or of a 3, or of a 5 showing. The probability of each of those outcomes is 1/6. Thus, by the addition rule, the probability of an odd number must be 1/6 + 1/6 + 1/6 or 1/2.

$$\begin{aligned} p(\text{odd number}) &= p(1 \text{ or } 3 \text{ or } 5) \\ &= p(1) + p(3) + p(5) \\ &= 1/6 + 1/6 + 1/6 \\ &= 1/2 \end{aligned}$$

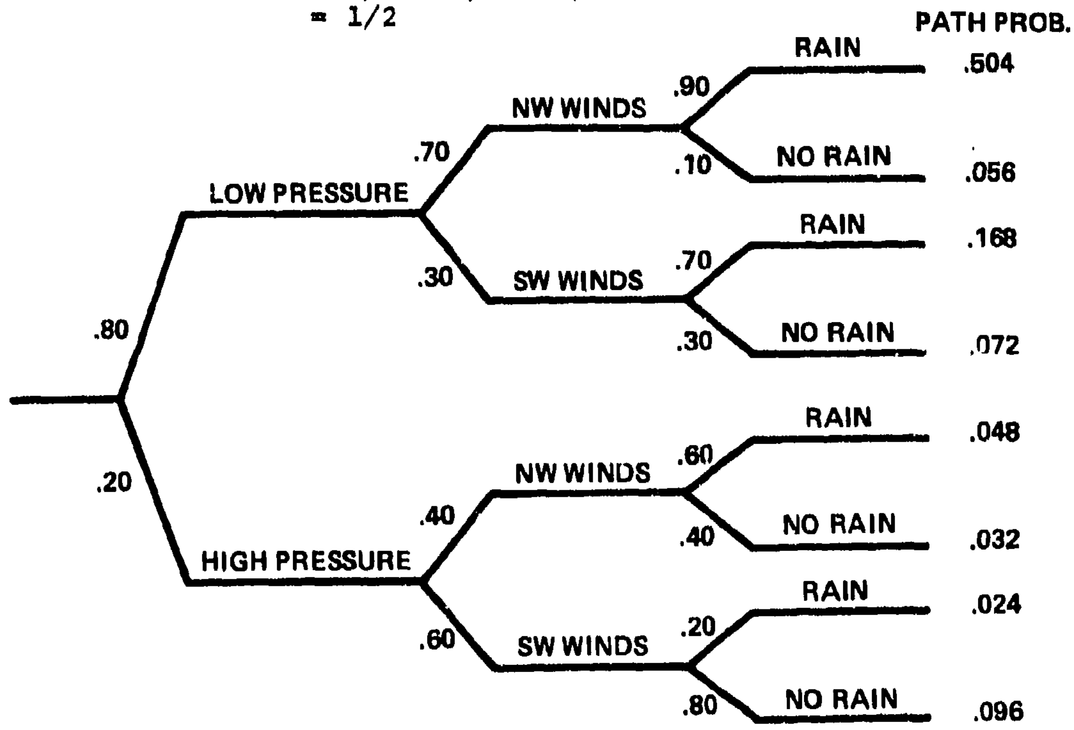


Figure 4-12

COMPLETE PROBABILITY DIAGRAM

Applying the addition rule to the weather problem, we find the probability of rain by adding the path probabilities for every path that leads to rain (the paths are mutually exclusive):

$$P\ (\text{rain}) = .504 + .168 + .048 + .024$$
$$= .74\ (\text{or } 74\%).$$

For no rain :

$$P\ (\text{no rain}) = .056 + .072 + .032 + .096$$
$$= .26\ (\text{or } 26\%).$$

Table 4-1 is a comparison of the estimates of the probability of rain made by the weatherman:

	"Off-the-Top-of-the-Head"	Probability Diagram
Rain	.60	.74
No Rain	.40	.26

Table 4-1

Initially, the weatherman estimated the first probability of rain as 60% from a global viewpoint by considering all factors "off-the-top-of-the-head." Finally, he assessed the probably more accurate figure of 74% by decomposing the problem into its more easily accessible component parts.

Overview

To help tie together the basics of probability diagrams, consider the following more complex example: What is the probability that the Soviet Union will launch a major attack across the border of the People's Republic of China during the next twelve months? Obviously, because of the complexity of this problem, it is very difficult to make an off-the-top-of-the-head estimate. One possible method of attacking the problem is to break it down into its component parts and derive separate probability assessments for each of those component parts.

The analyst working on this problem would have to decide which information concerning Sino-Soviet relations is most relevant and would have an effect on the target variables. Assume he came to the conclusion that Russia will decide to attack if there is political upheaval in China, as represented in Figure 4-13.

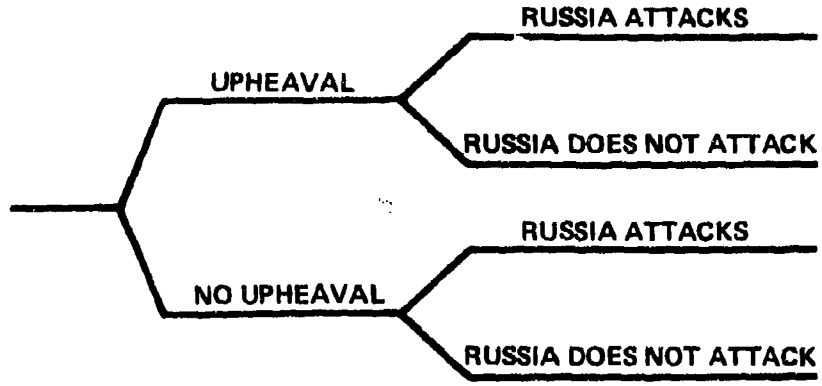


Figure 4-13

POLITICAL UPHEAVAL

The analyst might further feel that the likelihood of the occurrence of a political upheaval will depend on whether the Chinese political leader dies. The probability diagram in Figure 4-14 is one possible representation of these events.

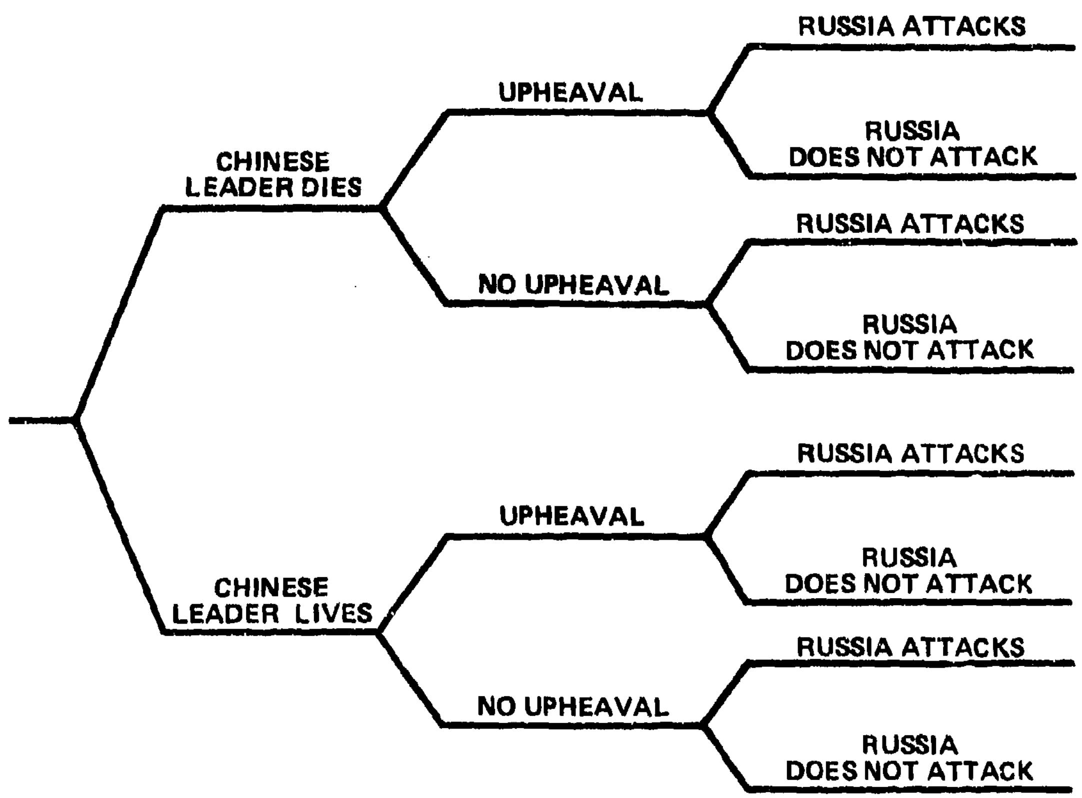


Figure 4-14

PROBABILITY OF RUSSIA ATTACKING CHINA

If this problem were going to be analyzed in depth, many more events might be helpful, and all of them would have to be very specifically defined. For our purpose, this depth, or to the exact probabilities, is not necessary; it is only the concepts with which we are involved.

Just considering the top path of the probability diagram, we derive each of the probabilities as follows:

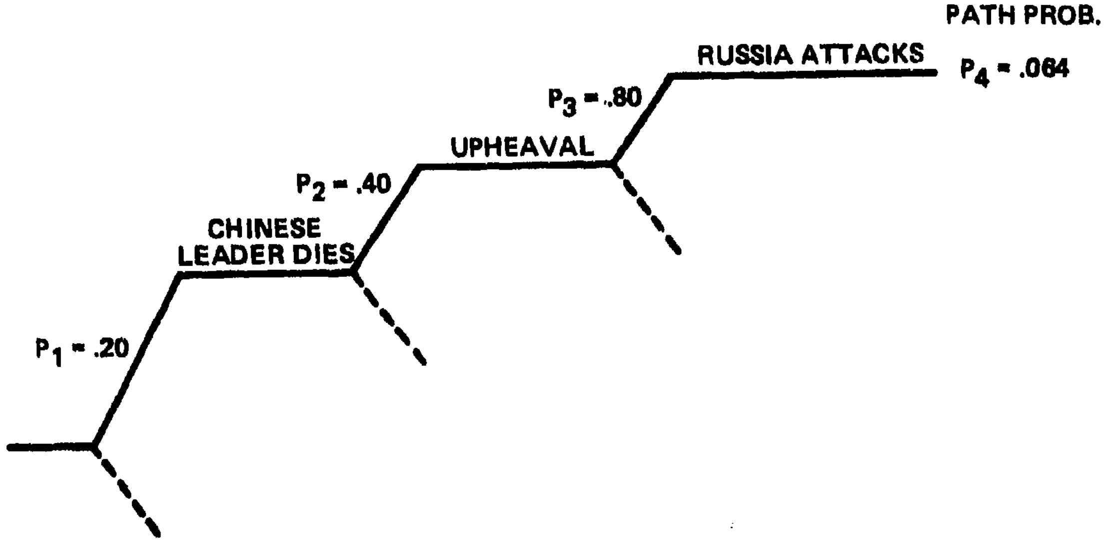


Figure 4-15

UPPER BRANCH OF FIGURE 4-14

P_1 - Assessed probability of Chinese leader dying.

P_2 - Conditional probability of upheaval given the leader dies - (upheaval|leader dies).

P_3 - Conditional probability of Russia attacking given that the leader dies and there is upheaval - (Russia attacks|leader dies and upheaval).

P4 - Path probability representing the likelihood of the whole branch occurring - (Leader dies and upheaval and Russia attacks).

$$P4 = P_1 \times P_2 \times P_3$$

$$= (.20) \times (.40) \times (.80)$$

P_4 = .064 (or, there is a 6.4% chance that all events will occur within the next twelve months).

Again, after considering some of the individual probabilities, let us go on to the complete probability diagram, Figure 4-16.

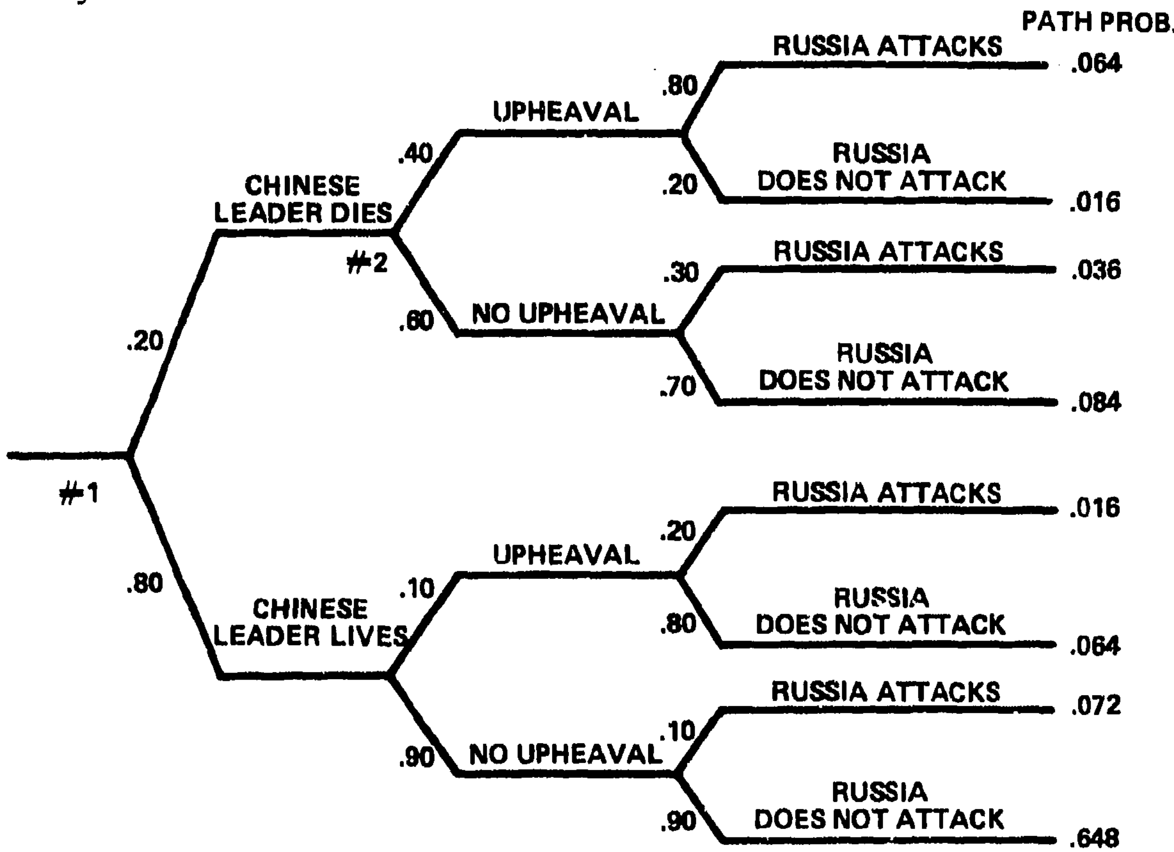


Figure 4-16

COMPLETE PROBABILITY DIAGRAM

Adding the path probabilities for the four paths which result in a Russian attack gives the overall probability of an attack,

$$P(\text{attack}) = .064 + .036 + .016 + .072$$
$$= .188$$

and

$$P(\text{no attack}) = .016 + .084 + .064 + .648$$
$$= .812$$

This is the solution to the original question, "What is the probability that the Soviets will launch a major attack across the border of the People's Republic of China during the next twelve months?" Given our assumptions about the relevant events and their likelihoods of occurrence, the overall probability of attack is 18.8% (from Figure 4-16). This final probability has been arrived at by breaking the overall problem into its small component events, events for which probabilities can be assessed easily and directly. Folding back these probabilities has made it possible to recombine them to determine their overall likelihood of attack, a likelihood which would have been extremely difficult to obtain from an overall estimate.

Folding Back

Frequently there is a need to know not only the likelihood of any one branch occurring, but also the likelihood of a series of branches occurring at one particular node. For example, the analyst may be interested in the probability Russia will attack assuming the Chinese leader dies. This could be determined by "folding back" to Node #2 the four right-hand branches of Figure 4-16, the result of which is shown in Figure 4-17.

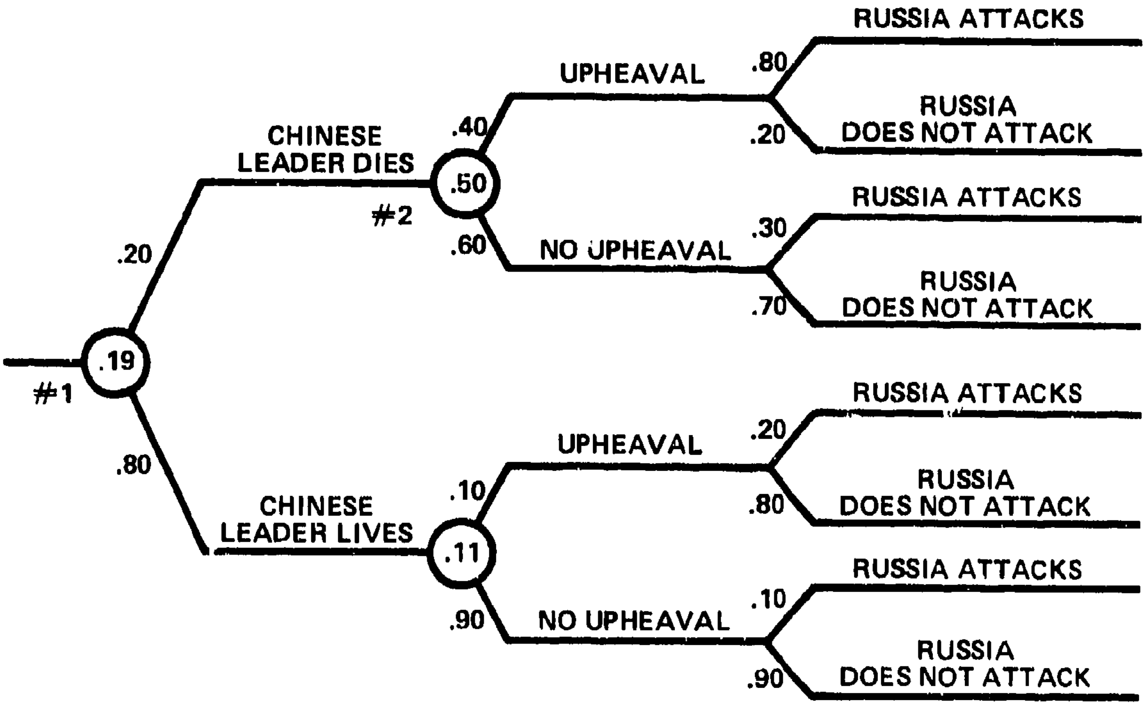


Figure 4-17

FOLDING BACK

At Node #2, the new probability, P(Russia attacks|Chinese leader dies), can be calculated by summing the appropriate branches of Figure 4-17.

P(Russia attacks|Leader dies)

= P(Upheaval|leader dies) x
P(Russia attacks| leader dies and upheaval) +
P(No upheaval|leader dies) x
P(Russia attacks|leader dies and no upheaval)

= (.40) x (.80) + (.60) x (.30)

= .50

In a similar manner, the whole probability diagram (Figure 4-17) can be folded back to Node #1 to determine the overall likelihood that Russia will attack, P = 19%.

Sensitivity Analysis

We are now ready to discuss the problem of subjective, or personal, probabilities. The reader will have noticed that throughout the analysis of both the weather and the Russian problems, very little explanation was given for the assessed conditional probabilities. In each case, the weatherman or analyst "assessed" there was a 70% probability of NW winds given there would be a low-pressure area, or "assessed" there was a 20% probability that the Chinese leader would die within twelve months. Although the analyst is using his expertise to provide his best assessment, other analysts may disagree; or the analyst may be unsure himself. The question naturally arises about the extent to which the probability of the target variable or event is influenced by any particular conditional probability. If the probability of NW winds, given low-pressure, is changed from 0.7 to, say, 0.5 or 0.9, what effect will this have on the overall probability of rain? If it has only a small effect, then disagreements between two experts, one of whom assigns 0.5 and the other 0.9, can be tolerated. Or, if an expert cannot decide what value between 0.5 and 0.9 to assign, he can be assured that it does not make any substantial difference so far as the overall probability of rain is concerned. Sensitivity analysis can resolve these kinds of questions.

First, let us reconsider the weather problem and the question of whether the likelihood of NW winds given there will be a low pressure area has a substantial effect on the overall probability of rain (see Figure 4-18).

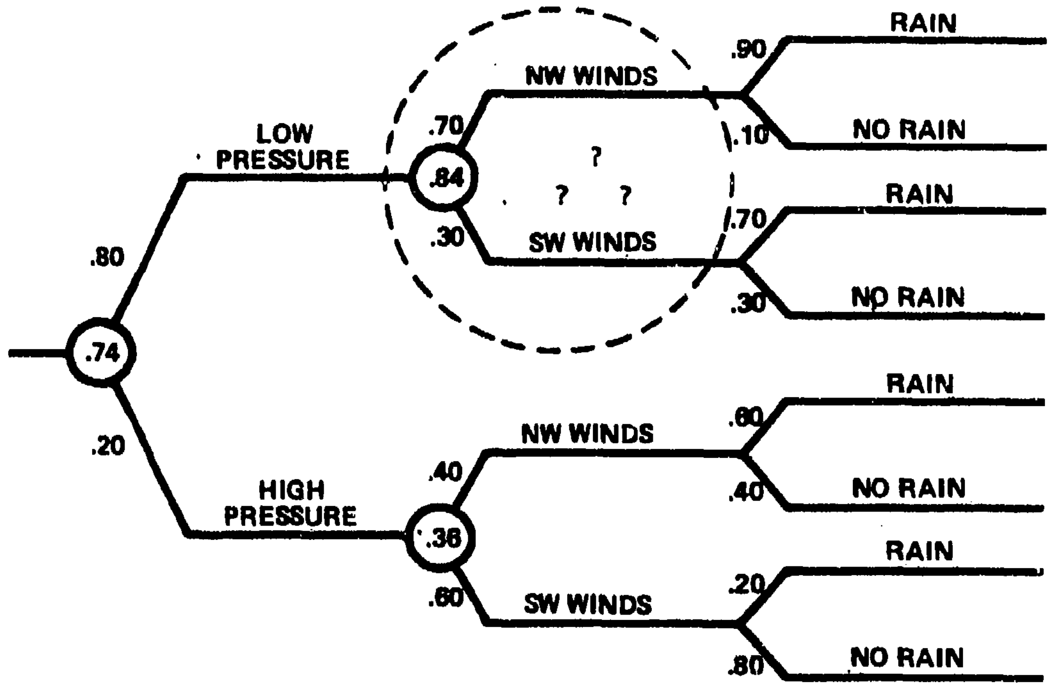


Figure 4-18

PROBABILITY OF RAIN

Possibly the weatherman is confident of all probabilities except the 70%. He has enough knowledge or expertise to feel comfortable about the other values, but not enough to feel sure about his assessment of 70%. Or perhaps he is discussing this analysis with two colleagues who are in total agreement with the analysis except for the 70% assessment. How crucial is 70% with regard to the overall probability of rain? Its degree of sensitivity can be determined by varying that probability over the range of possible values. If one colleague feels it is as low as 50% and the other colleague thinks it is as high as 100%, the weatherman can substitute those two extremes into the diagram and determine the overall effect on the P(Rain). Table 4-2 displays those results.

P (NW Winds Low Pressure)	P(Rain)
1.00 (Max. value)	.792
.70 (Original value)	.744
.50 (Min. value)	.712

Table 4-2

It is apparent from the small change in P(Rain), from .71 to .79, that this overall probability is relatively insensitive to the assigned probability. Because of this insensitivity, the weatherman and his colleague do not have to concern themselves with refining the probability of 70%. If this had been a sensitive branch and a fluctuation between the two extreme values did cause significant change in P(Rain), then more discussion or information would be needed to ascertain the best value for P(NW winds Low pressure).

Now reconsider the Russian problem where a 20% probability was assessed that the Chinese leader would die within the next twelve months (see Figure 4-19).

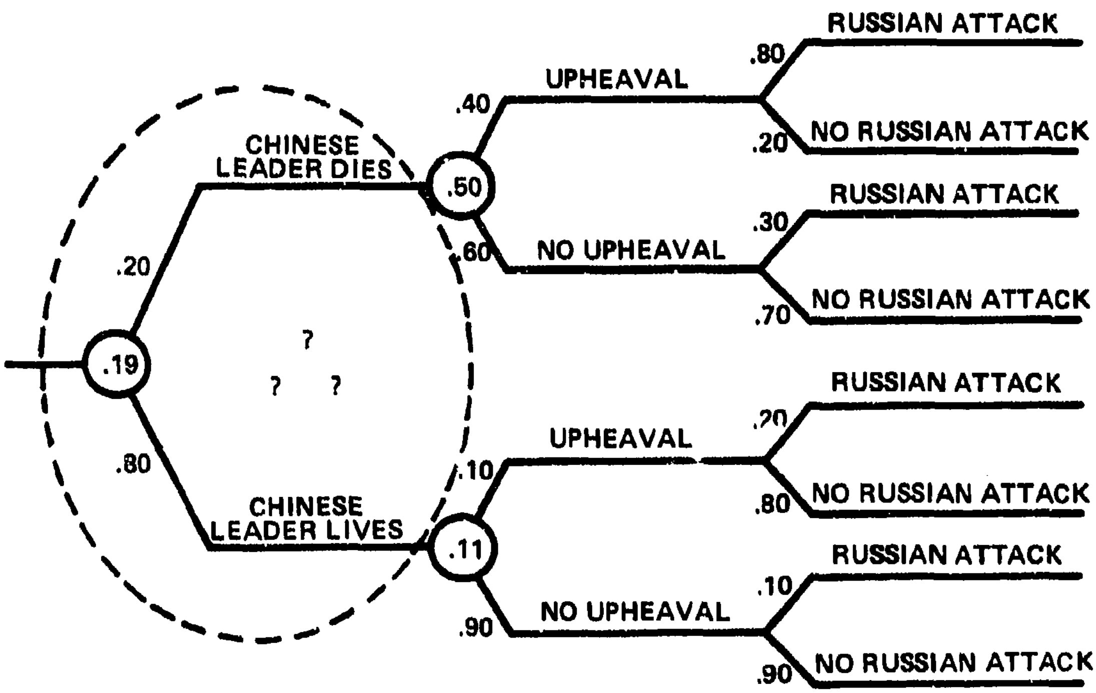


Figure 4-19

PROBABILITY OF RUSSIA ATTACKING

If there were a disagreement concerning this assessment, the analysts could determine the possible range of values. Depending on his age, the minimum actuarial probability of the Chinese leader's death might be 9%. The maximum might be 50%, an estimate supplied by an analyst who had recently received intelligence reports about the leader's deteriorating health. The effect of substituting these two extremes into the probability diagram on P(Russia attacking) is displayed in Table 4-3.

P(Leader dies)	P(Russian Attack)
.09 (Min. value)	.15
.20 (Original value)	.19
.50 (Max. value)	.30

Table 4-3

Unlike the outcome of the weather problem, the outcome of this problem is fairly sensitive to changes in probability. The probability of a Russian attack increases from 15% to 30% as the probability of the leader's dying increases from its minimum to its maximum value. In order to resolve this disparity among the different analysts, it would be necessary to obtain additional information and to refine the assessment of P(Leader dies). This could be done through intelligence work or possibly by a thorough analysis by persons knowledgeable in this area. If the target variable, a Russian attack, is sensitive to some of the component probabilities, then the analysts must pay more attention to those component probabilities than to those which are insensitive.

Sensitivity analysis is a very useful tool for helping the analyst to determine which components are crucial and which are relatively insensitive to the analysis. Refinement can be done on those decisive elements and less time can be spent on those which are secondary in importance.

Pruning Probability Diagrams

When an analyst works with probability diagrams for any length of time, it becomes evident that this tool can frequently evolve into a complex network of branches unsuitable for practical analysis. How often has the analyst started with a small set of conditioning events and ended up with a diagram containing 16, 32, or 64 ending branches, or even a diagram which contains over 100,000 path outcomes? Since the complexity of probability diagrams increases geometrically with the number of factors considered, only 10 two-valued factors need to be examined before over 1000 paths result (assuming a symmetrical diagram), and 20 such factors will produce more than 1,000,000 ending branches.

This section is concerned with different procedures that will permit "pruning" of probability diagrams back to a manageable size without eliminating pertinent information.

If the analyst uses these procedures, he will greatly increase the applicability of probabilitity diagrams for probability assessment and permit many factors to be incorporated, while keeping the level of analytic effort within reasonable bounds.

Although the analyst probably will be able to develop his own simplification, the following two basic pruning methodologies will be discussed in detail: adjustment and relaxation of assumptions, and Markov chains.

Adjustment and Relaxation of Assumptions

The central idea of this method is to direct the primary thrust of the analysis to those branches of the diagram which possess very high probabilities. This amounts to devoting special care to the parts of the problem that matter the most.

Consider first an analysis of the likelihood of a U.S. dollar devaluation within the next twelve months. Figure 4-20 shows how the assessment of a possible devaluation might be conditioned on three factors: inflation, trade deficit, and European pressure.

To determine the overall probability of devaluation from the diagram, it would normally be necessary to assign conditional probabilities for each branch of the tree. As is evident from Figure 4-20, the upper half of the diagram is much more important than the lower half because a probability of 0.75 has been assigned to the event "inflation $\geq$ 6%/yr." Therefore, the required probabilities have been assigned by the analyst only to the upper half of the diagram, for reasons which will become clearer as the analysis progresses.

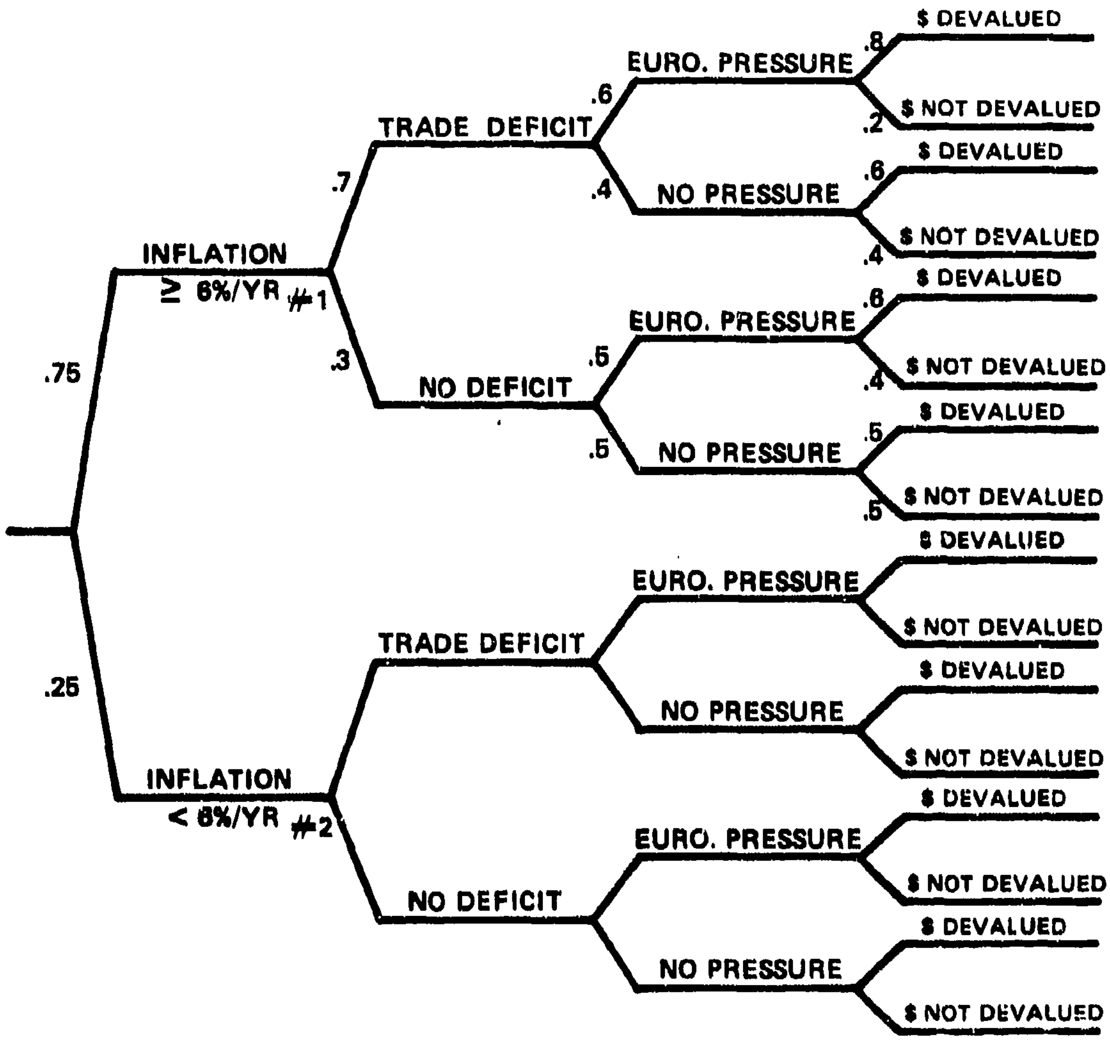
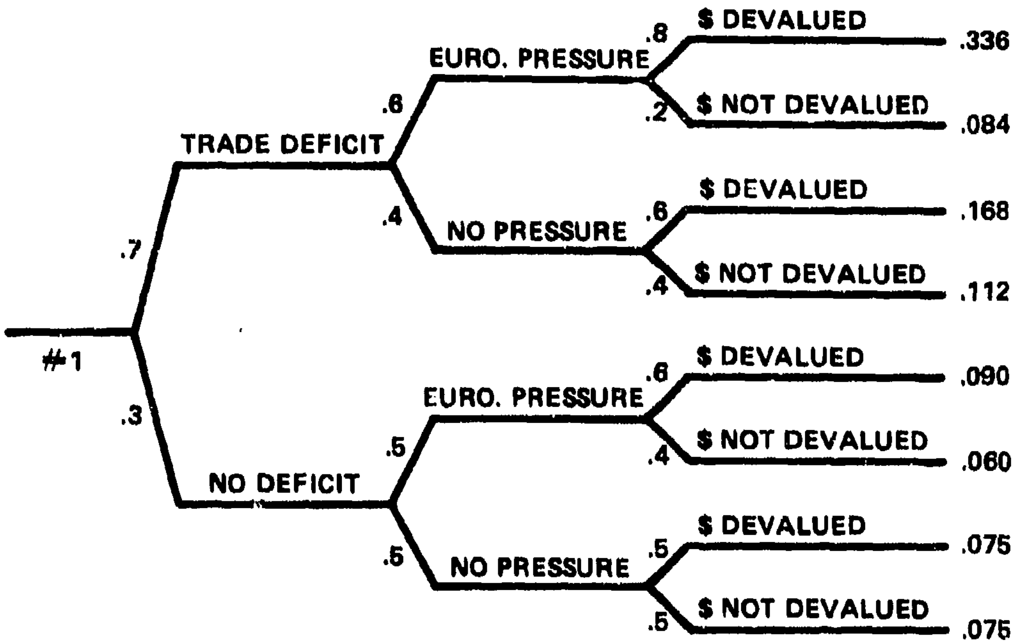


Figure 4-20

PROBABILITY DIAGRAM FOR U.S. DOLLAR DEVALUATION

At Node 1 in Figure 4-20, the effect of the factors subsequent to that node on the two possible outcomes are shown in Figure 4-21A. Folding this diagram back (by summing the path probabilities for each outcome) results in the simpler, but equivalent, diagram of Figure 4-21B.

A. Original Diagram Subsequent to Node 1



B. Simplified Diagram Subsequent to Node 1

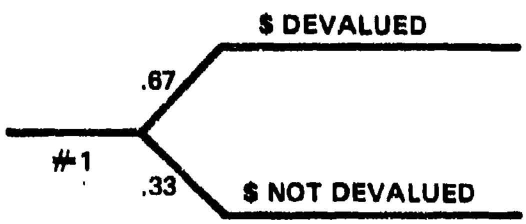


Figure 4-21

FOLDING BACK ORIGINAL DIAGRAM TO NODE 1

The next step would normally be the assignment of probabilities to each node of the lower half of the diagram, and the folding back of this half to Node 2, as was done above for the upper half of the diagram. Because the two halves of the original diagram (Figure 4-20) are identical in structure, (although the probabilities for each node are different), the diagram which results from folding back must resemble Figure 4-21B. This operation permits the simplification of the original diagram to the form shown in Figure 4-22, and we can approximate the probability of devaluation at Node 2 (given inflation is less than 6%/yr.) from the known probability of devaluation at Node 1 (with inflation equal to or greater than 6%/yr.).

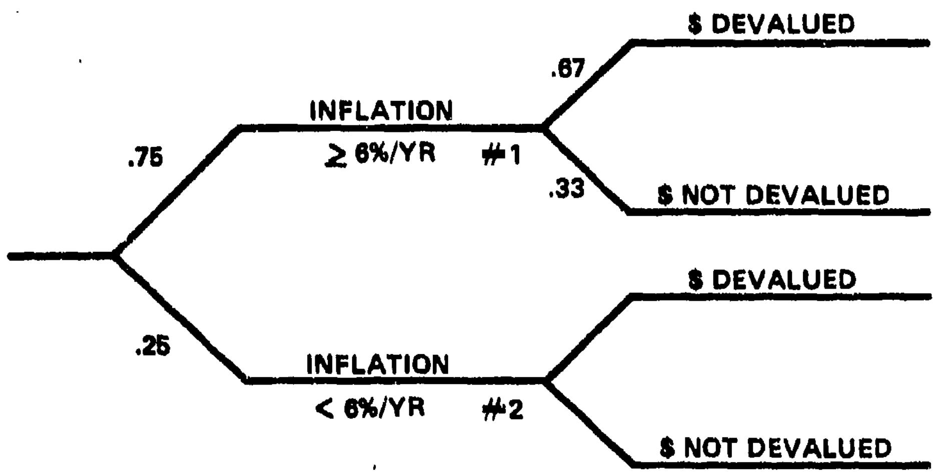


Figure 4-22

ORIGINAL DIAGRAM SIMPLIFIED BY FOLDING BACK

Assume the analyst feels the probability of devaluation at Node 2 could be as low as 25% or as high as 50% of that of Node 1. The effect of varying Node 2 from its highest to lowest plausible values is as follows:

P(Devaluation) Node 1	P(Devaluation) Node 2	P(Devaluation) Overall
0.67	0.33 (max.)	0.58
0.67	0.17 (min.)	0.54

It is apparent from the small change in the overall probability of devaluation that it is relatively insensitive to the assigned probability at Node 2. Therefore, for this analysis, there is no need for the analyst to refine his assessment of the probability at Node 2 or to complete the analysis by assessing individual probabilities for the branches to the right of that node.

It should be noted that the adjustment procedure has worked successfully on this problem because of 1) the identical structure of the two sections, and 2) the relatively low probability assigned to the likelihood of inflation being less than 6% per year. If this assessment had been much higher, it would have necessitated assessing probabilities for the entire remaining portion of the diagram, and laboriously completing the analysis in a manner similar to that employed for the upper half of the diagram.

The "adjustment" method represents one instance of relying on a precise analysis of highly probable branches of a probability diagram in order to produce, as simply as possible, a usefully less precise solution to the complete diagram. This principle can also be used in analyses where there is a sequence of highly probable assumptions. The effort is placed on analyzing the situation when the assumptions hold. Following this analysis, the assumptions are "relaxed" (i.e., assumed not to hold) in sequence and assessments relevant to the new branches of the diagram are made.

Consider two countries A and B and the possibility of conflict between them resulting in open hostility. If there are complex relations between A and B, the analysis will require a simplification of any probability diagram which attempts to interrelate major contributing factors leading to possible hostilities. In this analysis, we shall develop a probability diagram based upon several major assumptions and then relax those assumptions one at a time. Each of the assumptions to be relaxed must be a highly probable one.

In this scenario, A is an aggressive but militarily weak country, while B is powerful but politically unstable. To understand how this analysis will proceed, consider the following four incidents, which we shall call events, and the two possible outcomes, or hypotheses, which can result from those incidents:

Events

1. Drought in B (or no drought).
2. Coup in B (or no coup).
3. Arms shipments continue into A (or arms cease).
4. Acts of terrorism by A (or no acts of terrorism).

Hypotheses

Hostilities (H)

No Hostilities ($\bar{H}$)

The analysis of this problem can now proceed in two directions: the analyst can work with the four events simultaneously by constructing a 32-branch tree and assessing all of the necessary probabilities; or the analyst can simplify the analysis by assuming each event occurs, assessing the probability of the hypotheses occurring (given that each event has occurred), and then relaxing that assumption concerning the events one at a time. This latter method is useful only if some or all of the events are independent of each other. Accordingly, the next step is to make judgments about event independence.

Recall from the discussion of independence earlier in this chapter that two events can be judged to be independent if the probability assigned to one is unaffected by knowing of the occurrence of the other event. Looking down the list of events, the analyst decides that uncertainty about a coup in B would be affected by knowing there is a drought. Here the link is causal: drought can lead to political instability which might result in a coup; that is, a coup is more likely if there is a drought than if there is no drought. In short, the events in item 2 depend on the events in item 1.

Incidentally, we should point out to the interested reader that independence is a symmetrical relationship; if A is not independent of B, then B is not independent of A. This means that the probability assigned to "drought in B" depends on knowing whether or not there is a coup. That may sound odd: surely the occurrence of a drought is unaffected by the occurrence of a coup. True, as a statement of causality, but as we said earlier in the chapter, causality is not an adequate guide for judgments of independence. One must ask only whether knowledge of one event reduces uncertainty about some other event. In this case, the assessor should ask, "Is there any informtion in the event 'coup in B' that would affect my uncertainty about the event 'drought in B'?" If you had not heard about the weather and were assigning probabilities to it, would you change the probabilities if you were told that a coup occurred? The answer is yes, because droughts and coups have a slight tendency to go together, though, of course, not always. Knowing that either has occurred reduces our uncertainty about whether the other has happened. We are taking advantage of the association between these events, an association which is considerably less than perfect, yet useful for prediction. Knowing a person's height reduces your uncertainty, but does not completely eliminate it, about the person's weight. Height and weight are imperfectly associated: tall people are heavier, short people lighter, but there are exceptions. Knowing either height or weight reduces uncertainty about the other because height and weight co-vary. So with droughts and coups; they co-vary. Thus, knowing that one has occurred, I change my assessment of probability for the other.

Now we return to the task of identifying which events in this problem are independent. The analyst feels that the events in items 3 and 4 are independent of each other, that acts of terrorism, for example, are just as likely to occur whether or not arms shipments continue. Also, he decides that events in items 3 and 4 are independent of events in items 1 and 2.

Having determined that some events are independent, the analyst recognizes that he does not have to work with the full 32-branch tree. Instead, he adopts the alternative procedure, which starts with his assessing probabilities for the hypotheses assuming that all four events have occurred (See Figure 4-23). Then he will relax those assumptions.

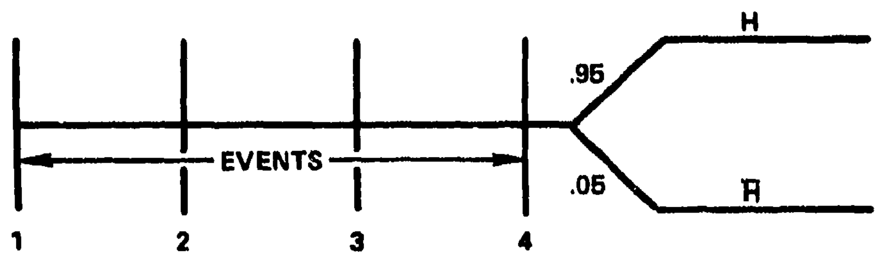


Figure 4-23

MODEL ASSUMING EVENTS 1–4 OCCUR

Assume the analyst has assessed the probability of hostilities, P(H), to be 0.95, assuming events 1-4 occur. The next step in the analysis is to relax the assumption concerning event 4 and construct the probability diagram shown in Figure 4-24.

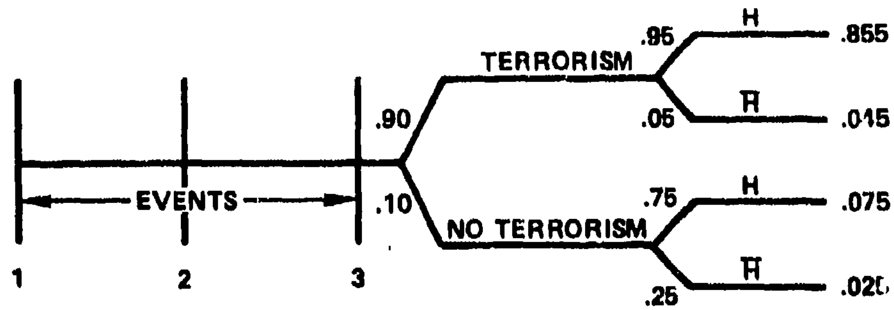


Figure 4-24

RELAXING ASSUMPTION CONCERNING EVENT 4

The analyst assesses two new probabilities, P(terrorism) = 0.90, and P(H|no terrorism) = 0.75, both given that events 1, 2, and 3 have occurred. He then folds the diagram back to the simpler form shown in Figure 4-25.

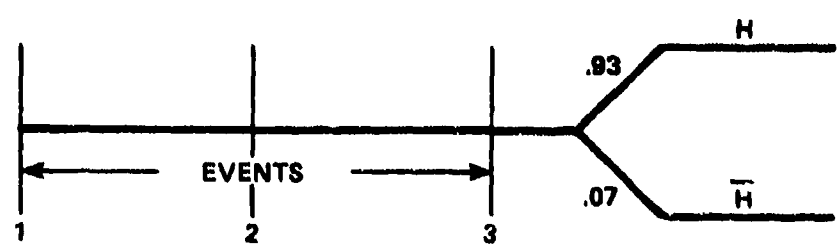


Figure 4-25

FOLDING BACK RELAXED ASSUMPTION FOR EVENT 4

The next step in the analysis is to relax the assumption concerning event 3 and construct a new probability diagram (Figure 4-26).

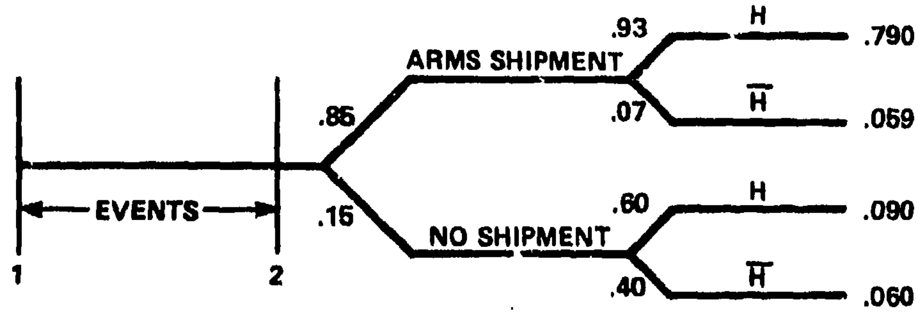


Figure 4-26

RELAXING ASSUMPTION CONCERNING EVENT 3

Again, two new probabilities, P(arms shipment) = 0.85, and P(H|no arms shipment) = 0.60, both given that events 1 and 2 have occurred, are assessed by the analyst. The resulting diagram can now be folded back to the diagram shown in Figure 4-27.

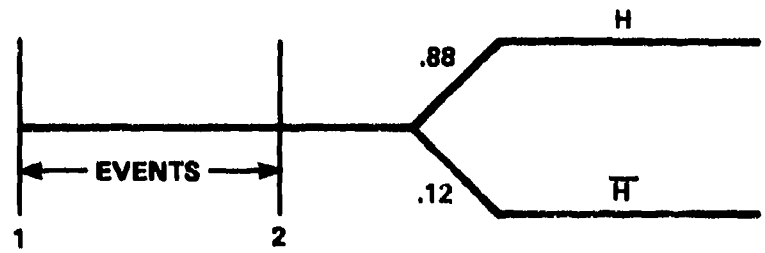


Figure 4-27

FOLDING BACK RELAXED ASSUMPTION FOR EVENT 3

At this point in the analysis, with events 1 and 2 still to be analyzed, the analysis needs to proceed with care. Up until this point the analysis has only been concerned with events 3 and 4, which are independent. Now, it must deal with events 1 and 2, which are not independent. If both dependent events are regarded as very highly likely to occur, the present method may be continued. Otherwise, the diagram for the dependent events must be constructed as shown in Figure 4-28.

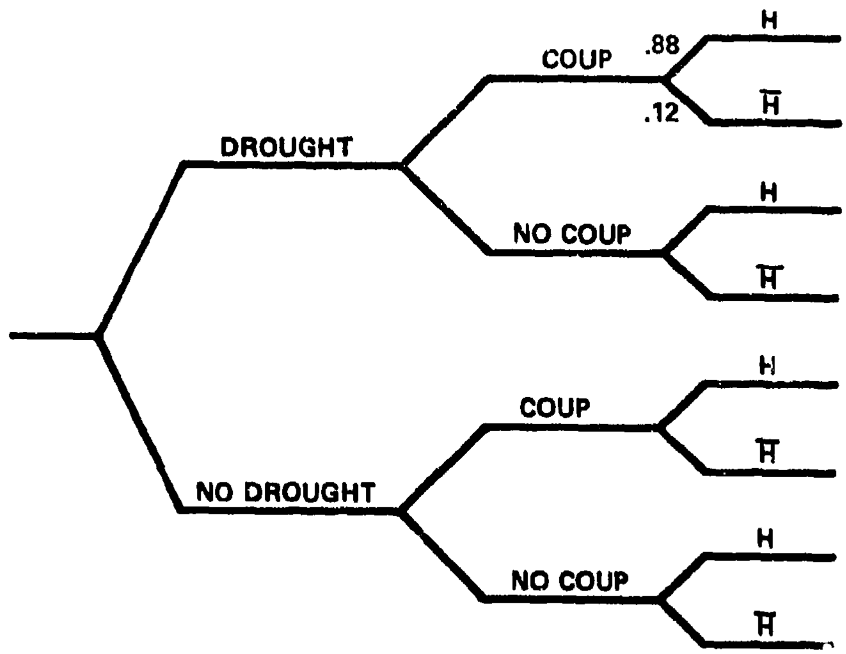


Figure 4-28

DIAGRAM FOR ANALYSIS OF DEPENDENT EVENTS 1 AND 2

Figure 4-28 now represents a standard probability diagram for which the standard method of analysis is possible. Although it still requires some work for completion, that work can be accomplished in less time with only an eight-branch diagram than with the 32-branch diagram which would originally have been generated.

Markov Chains

The second methodology for pruning probability diagrams, Markov Chains, is an extremely useful technique when the analysis involves a series of events which are repeated over a period of time. To illustrate its use, consider an analysis of Country C and its internal political unrest. The hypothesis in question is the possibility of serious hostilities developing within the next six months. If those hostilities are to develop, they can be considered to result from one of the following incidents:

1. Conflicts between the "Inpower" and the "Outpower" groups;

2. Internal conflict among factions of the "Inpower" group;

3. Other (all-inclusive factor).

First, consider the probability diagram in Figure 4-29, which traces out over a one-month period the sequence of events that can escalate the country from its normal state at the beginning of the one-month period into any one of the following three states at the end of that period: normal state (N), tension state (T), and hostilities (H). A one-month period was used because it was felt that the effect of the incident would be fully felt in this time frame.

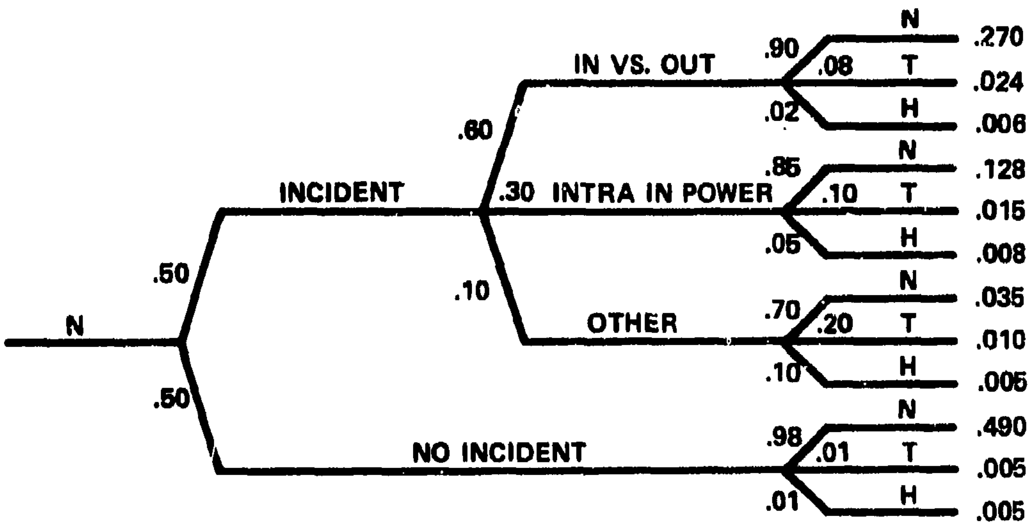


Figure 4-29

STARTING IN NORMAL STATE (N)

Now consider a one-month period which begins with Country C in a state of tension. The probability diagram for this situation is shown in Figure 4-30. Here, as before, the country may be in any one of the three states (N, T, or H) at the end of the one-month period. Note that none of the probabilities in either Figure 4-29 or Figure 4-30 are conditional on events occurring prior to the assumed beginning states. This quality is essential to the application of this method of analysis.

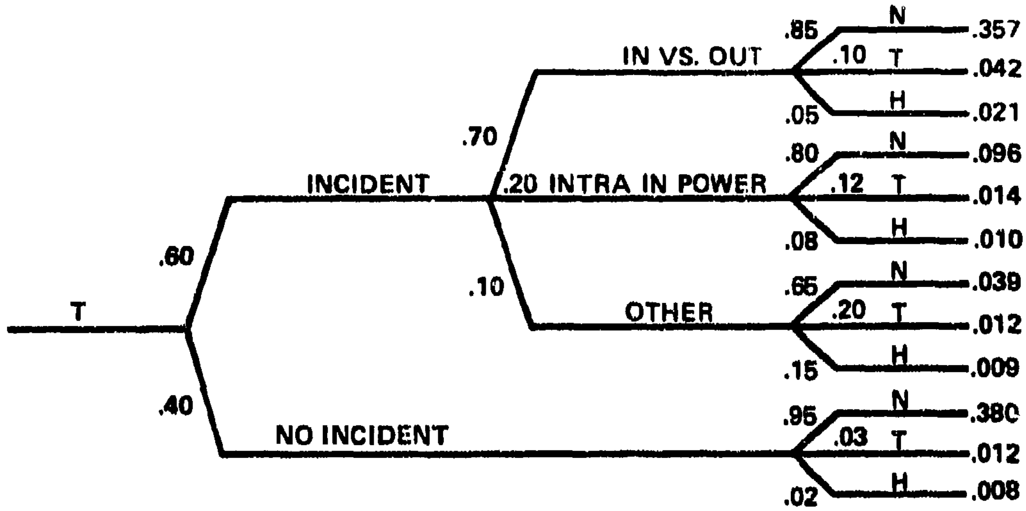
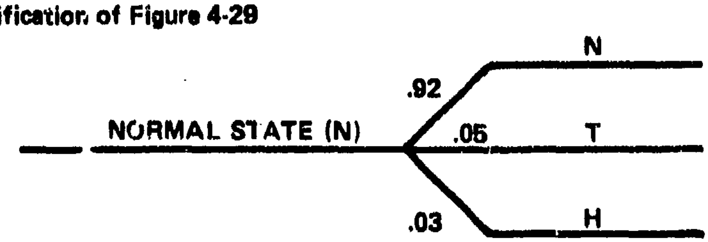


Figure 4-30

STARTING IN STATE OF TENSION (T)

The reader should note that although many different events can occur during the one-month interval, only three possible end states exist: normal, tension, and hostility. Because of this limited range, the large probability diagrams (Figures 4-29 and 4-30), which are useful for analysis of the many possible implications of an incident, can be pruned, or simplified, by using the diagrams of Figure 4-31. Both display the expected probabilities of terminating in each of the three possible end states.

A. Simplification of Figure 4-29



B. Simplification of Figure 1-30

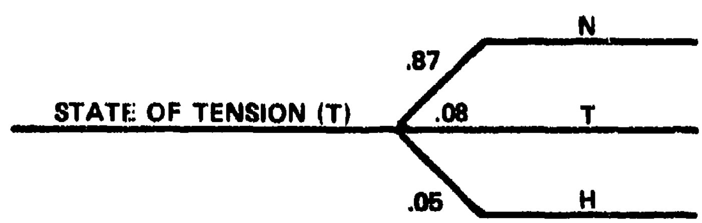


Figure 4-31

SIMPLIFICATION OF FIGURES 4-29 AND 4-30

Figures 4-31A and 4-31B are representations of one-month periods only. Since the hypothesis is concerned with the probability of hostilities over a six-month period, we must connect these diagrams together (link them into a "chain") so that if, at the end of one month, the H branch has not been reached, the process starts over at the beginning of the next one-month period. Figure 4-32, which assumes that Country C is in a normal state to start with and extends the analysis for a <u>second</u> one-month period, combines these two diagrams.

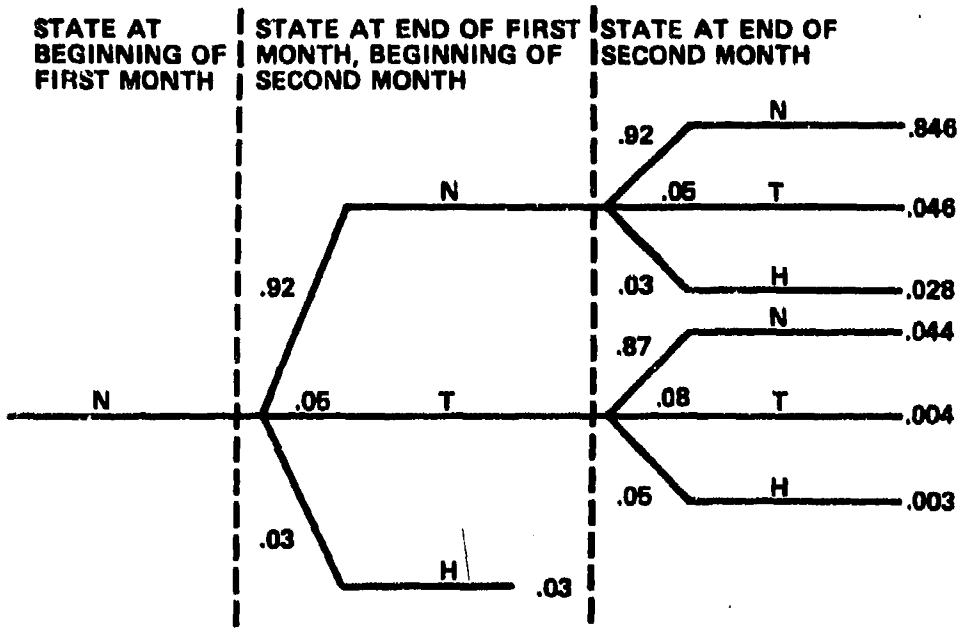


Figure 4-32

PROBABILITY DIAGRAM FOR A TWO-MONTH PERIOD

Although considerable pruning has occurred between Figures 4-29 and 4-30, and Figure 4-31, it is evident from Figure 4-32 that as new diagrams are added at the end of each one-month period, the number of end points will rapidly evolve into a complex network unsuitable for practical analysis. At the end of the first month there are 3 endpoints, 7 at the end of two months, and 123 endpoints at the end of six months.

Instead of drawing all 123 branches, the process can be simplified through the use of a Markov model, which chains together each of the one-month periods necessary to obtain six months. This simplification can be achieved most readily by constructing a transition probability matrix, such as that shown in Figure 4-33, which gives the probability of going from one state to another in a one-month period.

End State / Begin State	N	T	H
N	.92	.05	.03
T	.87	.08	.05
H	0	0	1

Figure 4-33

TRANSITION MATRIX FOR ONE-MONTH PERIOD

For example, assume you are in state N: the probability of still being in N at the end of one month is 92%; the probability of going into a state of tension is 5%; and the probability of going into hostilities during the one-month period is .03 or 3%. Note that while it is possible to enter state H from either N or T, once you have entered H, according to the purpose of the analysis, it is impossible to leave.

Implicit in this matrix is an important underlying assumption; namely, that the probability of moving from one state to another is dependent only on the state at the beginning of any period, not on how that beginning state was reached. Once the transition probabilities have been assessed, the matrix can be used to calculate the likelihood of hostilities having occurred at the end of any number of months, n. This is done by raising the matrix to the n^{th} power, which will give you not only the probability of hostilities, but all the other probabilities associated with the problem for the n^{th} month, in the form of a new matrix. Doing this requires a knowledge of matrix algebra and either a willingness to perform the laborious computations manually or a computer conveniently programmed to do it for you.

Figure 4-34 gives the results of such a computation for Country C, assuming that the starting state was normal (N). The probability of hostilities within the next six months, as read from the graph, is 17%.

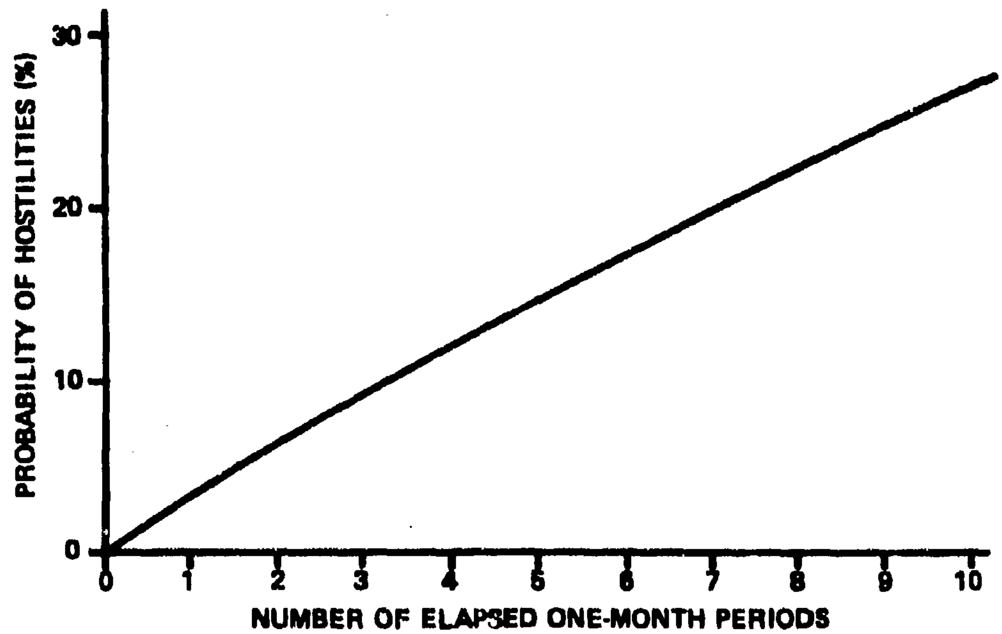


Figure 4-34

PROBABILITY OF HOSTILITIES AFTER SUCCESSIVE ONE-MONTH PERIODS

Hierarchical Inference

Up to this point, we have assumed that observed or observable data can be related directly to a given hypothesis, and that its impact on the probability of the hypothesis occurring can easily be determined by the application of Bayes' Theorem. However, the complexity of many real-world inference problems may require an amount or kind of knowledge necessary to assess probabilities or likelihood ratios directly linking all elements of data to the hypothesis beyond the capability of any one individual. In such cases, one approach is to decompose the complex problem by using a hierarchical structure. It may be easier to assess first those probabilities linking the data to intermediate variables (activities and/or indicators of activity) relevant to the hypothesis and then assess the probabilities linking these intermediate variables to the hypotheses.

One of the most compelling practical reasons for a hierarchical decomposition is that different individuals are experts in different areas; organizations are hierarchically structured to take advantage of this fact. For many inference problems, it is unlikely that any one individual has the necessary experience to relate the lower-level data or observables to the upper-level hypotheses. A problem decomposed along hierarchical lines makes it possible for the knowledge of each expert to be applied to the problem in a logically consistent manner.

To a great extent, the structure of a hierarchy reflects the conceptual model which the inference maker has of a particular problem. For this reason, any particular structure contains a degree of subjectiveness: the way a particular individual chooses to decompose a particular problem, and the way environmental constraints may make one choice more appealing than another.

Establishing a Hierarchical Structure

To illustrate how a hierarchical structure might be constructed, let us consider a substantive example. Suppose that an intelligence analyst is evaluating the intent of Country A to develop an independent nuclear weapons production capability within the next five years. This allows him to establish two hypotheses:

H_1 - Country A intends to develop a nuclear weapons production capability within 5 years, and

H_2 - Country A does not intend to develop a nuclear weapon production capability.

He then assembles a list of <u>activities</u> that would support these hypotheses, such as nuclear research and high-explosive research. The next step is to identify <u>indicators</u> related to these activities. One such indicator of increased nuclear research would be the construction or expansion of a centrifugal enrichment plant to increase the production capability of fissionable material. The last step in establishing the structure is to find <u>data</u> that would confirm or deny the activities or their related indicators. An example is a report from a photo interpreter describing the construction of several new cooling towers at the enrichment plant. This datum, which the analyst received from an expert in this technical area, and which increases the chances that the output of enriched uranium will increase over the next two years, lends support to H_1.

The analyst consults with experts in other areas which he considers relevant to his problem and, from the information obtained, constructs the hierarchical structure in Figure 4-35.

Quantitative Linkages Between Adjacent Levels

In order to use the observable information and to find the probabilities associated with the two hypotheses, it is necessary to assess quantitative linkages between levels. The linkages will take the form of conditional probabilities assessed by the analyst. The conditional probabilities refer to events at the level of interest given the occurrence of an event or the truth of a hypothesis at the next higher level.

The assessed conditional probabilities and likelihood ratios are measures of the strength of relations among variables. If, for example, the probability of an observed datum is about the same in the presence as in the absence of an indicator, then it provides very little information about whether or not the indicator is in fact present. If, on the other hand, the two probabilities are quite different, then the datum provides considerable information about the presence or absence of the indicator.

The assessment of quantitative linkages between levels begins at the top of the structure with the hypothesis to activity or data linkages and continues downward until a linkage is established among all of the elements in each level and the elements in the next higher level to which they directly relate. Accordingly, the first step is to assess conditional probabilities of data and activities given H_1 and H_2, for the row of boxes just below the top box, as shown in Figure 4-35.

For example, consider the activity, high-explosive R&D program. To assess the conditional probabilities associated with this activity given the truth of H_1 or H_2, the analyst begins by assuming that H_1 is true, that Country A intends to develop a nuclear weapons capability within five years. Under these circumstances, what is the likelihood that a high-explosive R&D program would be conducted? The analyst believes the chances to be high, about 95%. This leaves a 5% chance that such a program would not be carried out. Next, the analyst assumes H_2 to be true, that Country A does not intend to develop a nuclear weapons capability within five years. In this case, he feels there is only a 35% chance that the high-explosive R&D program would occur, and, therefore, a 65% chance that it would not occur. These assessments are shown in Figure 4-36.

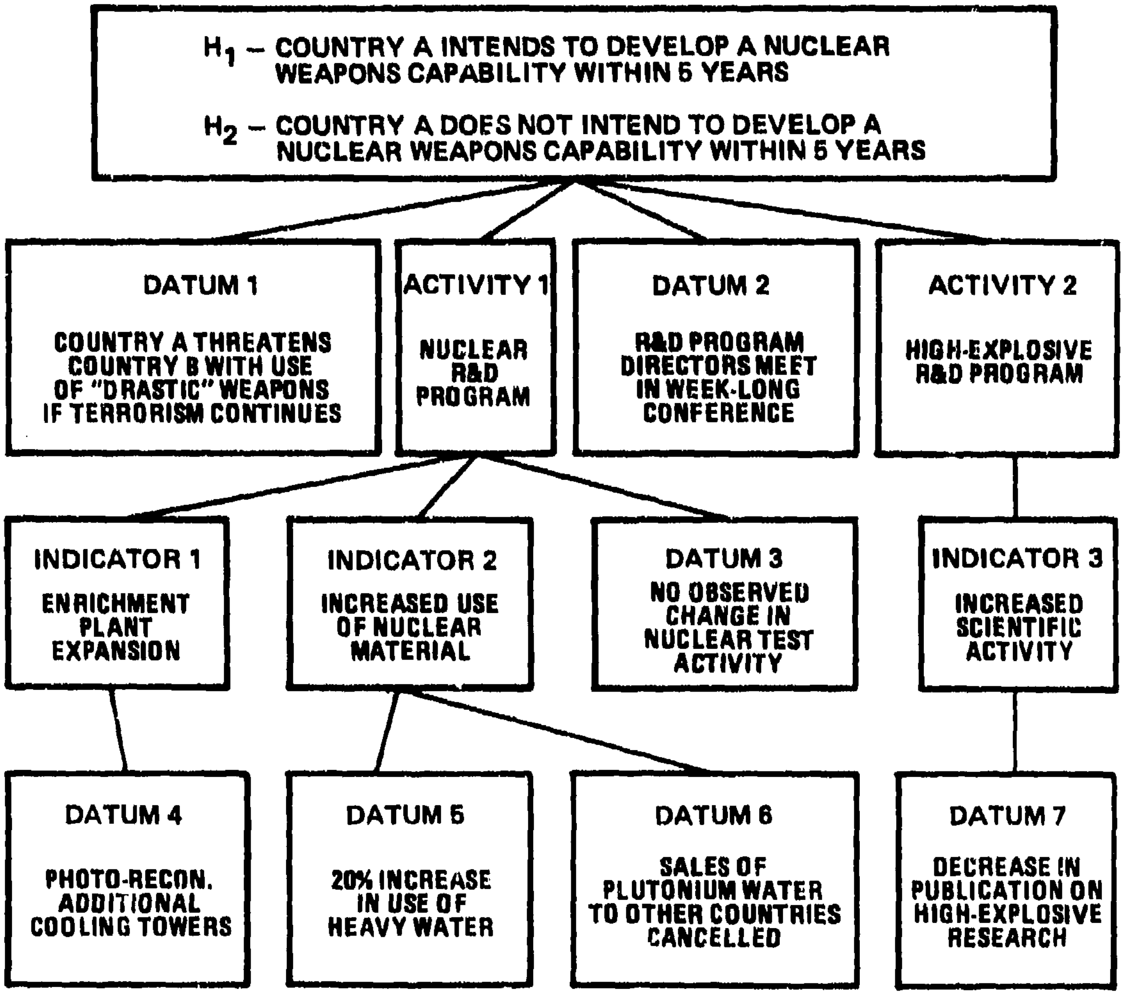


Figure 4-35

HIERARCHICAL STRUCTURE FOR
NUCLEAR WEAPONS DEVELOPMENT PROGRAM

	Nuclear Capability	No Nuclear Capability
High-Explosive R&D Program	0.95	0.35
No High-Explosive R&D Program	0.05	0.65

Figure 4-36

PROBABILITY MATRIX LINKING ACTIVITY 2 TO THE HYPOTHESES

In order to carry the analysis through to solution, similar assessments relating the second activity and the first two data to the hypotheses are made by the analyst.

The next step is to assess the conditional probabilities linking data and indicators in the third level to activities in the second level. For example, the analyst feels almost sure that <u>if</u> the nuclear R&D program were being conducted, then there would be an increase in the use of nuclear material. He assigns a likelihood of 0.80. On the other hand, he feels that there is a 60% chance that increased use would occur even without the R&D program. Figure 4-37 shows the matrix representing these assessments.

	Nuclear R&D	No Nuclear R&D
Increased Use of Nuclear Material	0.80	0.60
No Increase	0.20	0.40

Figure 4-37

PROBABILITY MATRIX LINKING INDICATOR 2 TO ACTIVITY 1

Again, similar assessments were made for indicators 1 and 3 and for datum 3.

The third step is the assessment of conditional probabilities linking data in the fourth level to indicators in the third level. As before, there will be a number of assessments required to link all the data to the hierarchical structure. Figure 4-38 shows the probabilities linking a 20% increase in the use of heavy water, to the indicator of increased use of nuclear material.

	Increase in Use of Nuclear Material	No Increase
20% Increase in Use of Heavy Water	0.99	0.50

Figure 4-38

CONDITIONAL PROBABILITIES LINKING DATUM 5 TO INDICATOR 2

The analyst believes there is a 99% chance that increased use of heavy water will be observed if there is a general increase in the use of nuclear material and only a 50% chance of observing this datum otherwise.

Although it is possible, as indicated in Figure 4-38, to assess conditional probabilities linking data to the appropriate level of the hierarchical structure, it is typically the case that any given datum is very unlikely given any hypothesized conditioning event. Analysts therefore often find it easier to assess a ratio of conditional probabilities for a datum than to assess the conditional probabilities themselves.

For example, if the analyst wishes to assess such a ratio for the link between datum 5 and indicator 2, conditional probabilities of which are given in Figure 4-38, the first step is to ask, "Would a 20% increase in the use of heavy water be more likely to occur if there was a general increase in the use of nuclear material or if there were no such general increase?" Clearly, that datum is the more likely result of a general increase in the use of nuclear material. Next, the analyst asks, "How much more likely?" In this case, he might say he thought the 20% increase in heavy water usage to be about twice as likely for a general increase as for no increase. This likelihood ratio is shown in Figure 4-39. Strictly speaking, because the likelihoods shown in Figure 4-38, 0.99

and 0.50, are in a ratio of 0.99/0.50 = 1.98, and the assessed likelihood ratio shown in Figure 4-39 is 2.0, the analyst has been inconsistent in his judgments. However, the inconsistency is very slight and is easily attributable to judgmental error that often arises when assessors round off probability judgments to the nearest 0.05, or ratios to the nearest unit.

Either likelihoods or likelihood ratios can be assessed in the inference problem since what is important is the ratio of the numbers rather than their absolute value. Whichever type of judgment seems easiest or most natural to the assessor is the one that should be used.

	Increase in Use of Nuclear Material	No Increase
20% Increase in Use of Heavy Water	2.00	1

Figure 4-39

NORMALIZED LIKELIHOOD RATIO SHOWING THE RELATIONSHIP OF DATUM 5 TO INDICATOR 2

There remains one other probability assessment necessary to the solution of the hierarchical inference problem. This is the analyst's assessment of the prior probabilities of the hypotheses, without regard for information included in the hierarchical structure. He believes that H_1 and H_2 are equally likely prior to determining the inference from the available data. His prior odds are, therefore,

$$\frac{P(H_1)}{P(H_2)} = \frac{0.50}{0.50} = \frac{1}{1}.$$

The purpose of the inference is to update or modify these prior probabilities by utilizing the explicit evidence available in the lower-level observables.

Once the analyst completed the task of assessing the conditional probabilities relating all the levels of the hierarchical structure, he assembled them in the form of a deductive hierarchical structure, shown in Figure 4-40.

Mathematical Solution of the Hierarchical Inference Problem

The analyst, having constructed the deductive structure from the top down, must now solve the problem starting at the bottom of the structure and working up, to determine the likelihood of all the data, D, given hypotheses H_1 and H_2.

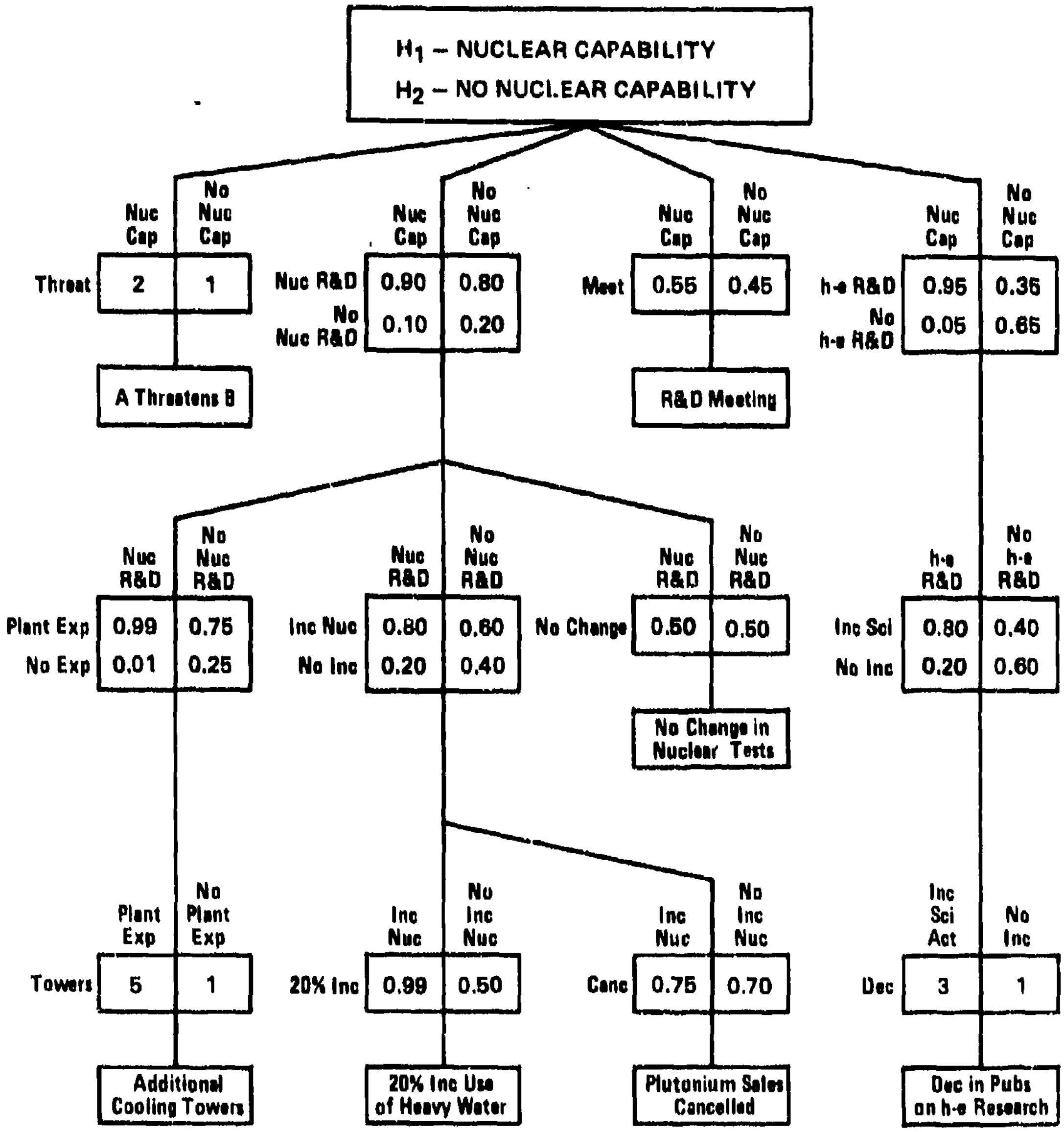


Figure 4-40

DEDUCTIVE HIERARCHICAL STRUCTURE FOR NUCLEAR WEAPONS DEVELOPMENT PROGRAM

The inductive hierarchical structure, represented in Figure 4-41, shows the steps and the computational results which the analyst went through to obtain a solution. (The principles behind those computations are given in Appendix D, along with a case study in hierarchical inference.) For each

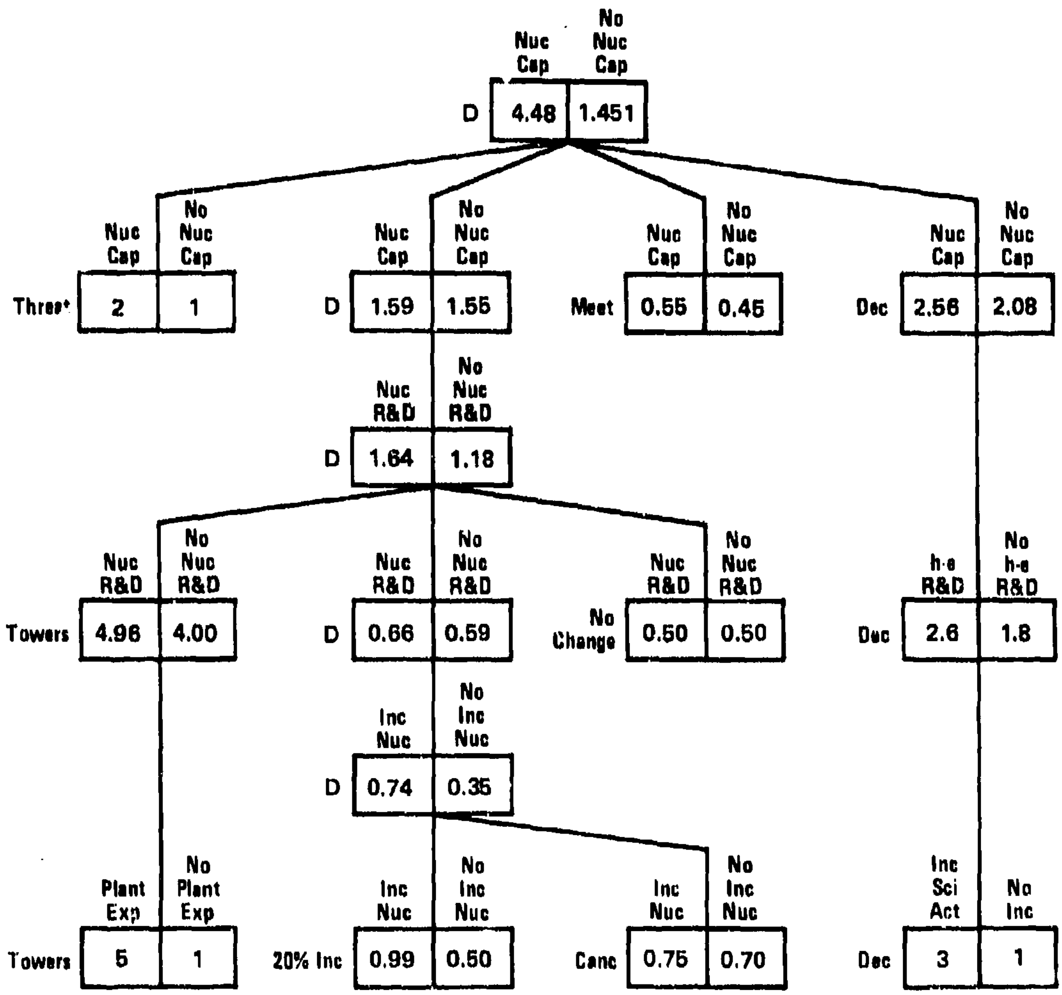


Figure 4-41

INDUCTIVE HIERARCHICAL STRUCTURE FOR NUCLEAR WEAPONS DEVELOPMENT PROGRAM

box, the two numbers show the relative likelihood of all the data D below that point in the structure, given that one of the next level events occurs. The top row vector shows that the final likelihood ratio was found to be

$$\frac{P(D|H_1)}{P(D|H_1)} = \frac{4.48}{1.45} = \frac{3.1}{1} .$$

Since the prior odds were assessed as 1:1, the posterior odds are obtained using Bayes' Theorem:

$$\frac{P(H_1|D)}{P(H_2|D)} = \frac{P(D|H_1)}{P(D|H_2)}\,\frac{P(H_1)}{P(H_2)} = \frac{3.1}{1} \times \frac{1}{1} = \frac{3.1}{1}\,.$$

This indicates that H_1 is now more than three times as likely as H_2, or expressed as probabilities.

$$P(H_1|D) = \frac{3.1}{4.1} = 0.76, \text{ and}$$

$$P(H_2|D) = \frac{1}{4.1} = 0.24.$$

Observation of those seven items of data leads the analyst to believe there is a 76% chance that Country A intends to develop a nuclear weapons capability within five years. That inference is the logical result of his assessments for the pieces of the decomposed problem.

CHAPTER 5

THE VALUE OF INFORMATION

Decisions on whether or not to gather new information, and if so, how much, can be analyzed in exactly the same way as any other decision by laying out a decision diagram with appropriate acts, events, probabilities, and outcome values, and then analyzing its implications, along the lines sketched in Chapters 1 and 2 of this Handbook. However, the type of linking which goes into such an analysis, and the kinds of problems which often arise, merit some special consideration.

Information decisions differ from what we might call primary decisions in that primary decisions lead directly to some outcome, whereas information decisions lead directly to the gathering of more data prior to making a subsequent decision. The assumption is that additional information will increase the chances of a desirable outcome. If the resulting expected improvement in the value of primary decisions exceeds the expected cost of the information gathering, then that difference is a measure of the net value of the information.

To illustrate how a specific analysis of the value of information might proceed, we shall take a quite realistic example from the foreign policy/defense field.

Suppose that in the course of a tense situation in Cyprus liable to lead to overt hostilities, the Commander-in-Chief for Europe (CINCEUR) has received a directive from the Joint Chiefs of Staff (JCS) instructing him to prepare to evacuate U.S. civilians from Cyprus only, or from Cyprus, Greece and Turkey, on three hours' notice. The evacuation notice, if it comes, will be a further JCS directive, presumably based on international developments focused on Cyprus and related to the possibility of conflict between Greek and Turkish Cypriots. There are three evacuation directive possibilities: major (E_M), involving evacuation from all three countries, Cyprus only (E_C), and none (E_N).

The primary decision facing CINCEUR is what posture of readiness to adopt in the anticipation of such an evacuation directive. The alternatives can be characterized by three basic postures: high, medium, and low, symbolized as P_H, P_M, P_L. The kinds of action each posture would entail are shown on the left of Figure 5-1, and an evaluation of some of their costs and other implications are shown on the right.

READINESS POSTURE	COST OF POSTURE	EXPECTED OUTCOME OF POSTURE
HIGH (P_H)		
STAND DOWN TRAINING	TO ATTAIN POSTURE	RAPID EVACUATION
DEPLOY ADDITIONAL AIRLIFT FROM U.S.	$950,000	MINIMUM U.S. CIVILIAN EXPOSURE
PREPARE EVACUATION BASES		MINIMUM U.S. CIVILIAN CASUALTIES
PLACE EVACUATION FORCES ON ALERT	TO MAINTAIN/DAY	LEAST COST IF TOTAL EVACUATION ORDERED
OBTAIN BASE AND OVERFLIGHT RIGHTS	$310,000	NEGATIVE REACTION FROM SOME NATO NATIONS TO U.S. ACTIONS -
DEPLOY AND POSITION RESCUE FORCES		HIGH COSTS IF NO EVACUATION
UPGRADE COMMUNICATIONS		DEGRADATION OF NORMAL MISSION
PROCURE REPLACEMENT AIRLIFT		
PREPARE ACCOMMODATIONS-FOREWARD TRANSPORTATION		
MEDIUM (P_M)		
PLACE FOUR THEATRE TRANSPORT AIRCRAFT ON ALERT	TO ATTAIN POSTURE $10,285	EVACUATION FROM ONE AREA ONLY COULD START IMMEDIATELY
PUT ONE FORWARD OPERATING BASE ON ALERT	TO MAINTAIN/DAY	INCREASED U.S. CIVILIAN EXPOSURE AND CASUALTIES
PREPARE ONE EVACUATION BASE	$3,350	LEAST COST TO EVACUATE CYPRUS ONLY
		INCREASED COST FOR TOTAL EVACUATION
		LITTLE MISSION DEGRADATION
		MINIMUM NATO REACTION
LOW (P_L)		
READINESS ORDERS ONLY	TO ATTAIN POSTURE	NO MISSION DEGRADATION
NO OTHER PREPAREDNESS ACTION	0	NO NATO REACTION
NO STANDOWN OF NORMAL OPERATIONS	TO MAINTAIN/DAY 0	FOR CYPRUS EVACUATION AND TOTAL EVACUATION HIGHEST EXPOSURE HIGHEST CASUALTIES HIGHEST COSTS

Figure 5-1

PRIMARY DECISION FACTORS

The value of these posture decisions is strongly influenced by whether there will be an evacuation directive and, if so, whether it is major or involves Cyprus only. In general, the higher the posture, the greater the cost (in money and non-monetary considerations), but also the greater the net benefit if a major evacuation is ordered.

CINCEUR could decide immediately on what readiness posture to adopt. Figure 5-2 shows a decision diagram that might reasonably reflect his best current judgments. The dollar values to the right incorporate judgmental adjustments for non-monetary factors such as U.S. civilian exposure. As we would expect, the values become more negative as the evacuation situation becomes more serious. Less obvious is the change in values as the posture becomes lower: although the cost of the posture is lower for less readiness, the exposure costs increase disproportionately, with the result that the overall values are higher for the lesser postures. On these judgments, we see that medium posture (E_M) has least expected cost ($-13.0M) and should therefore be preferred if the posture decision has to be made now.

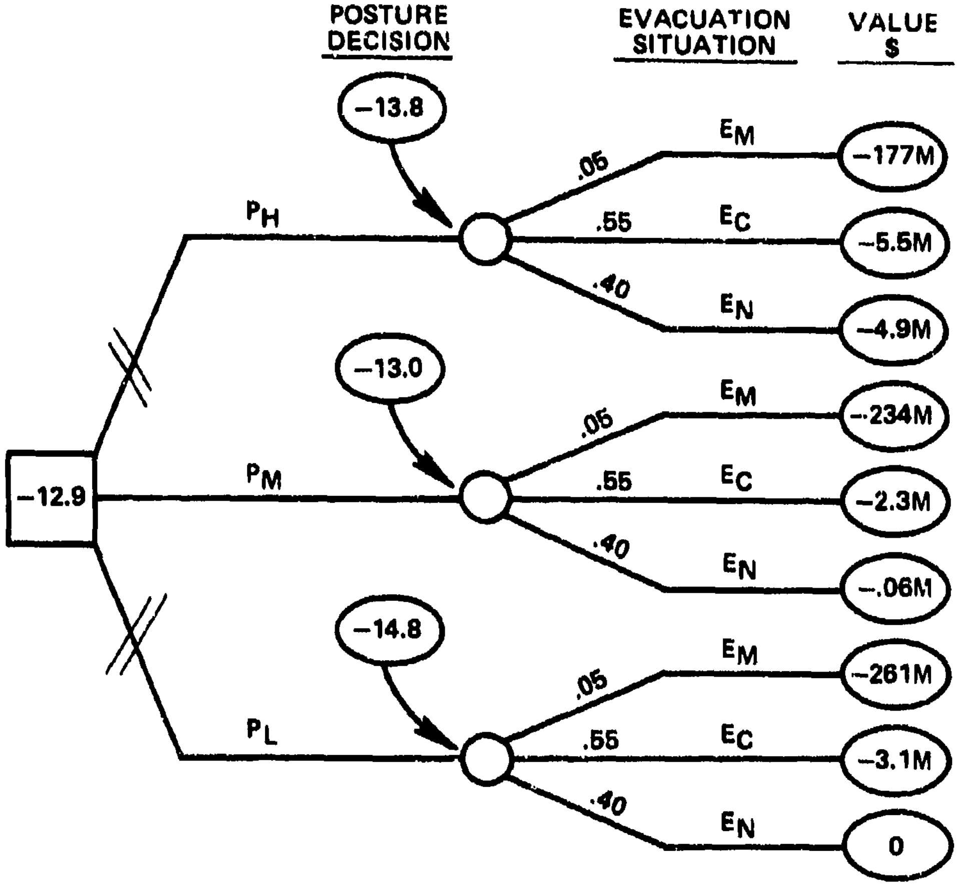


Figure 5-2

DIAGRAM OF PRIMARY DECISION

Intelligence Collection Decision

CINCEUR may have the option of gathering information, before making the posture decision, by varying the levels of his intelligence collection effort. He can direct a heavy collection effort (C_H), a moderate effort (C_M), or a low effort (C_L). Clearly, the more intelligence he collects, the more likely he is to receive information diagnostic of the prospects for evacuation and, therefore, the more likely it will be that his uncertainty about the proper course of action will be reduced. On the other hand, intelligence collection is costly in terms of money and resources diverted from other uses. The question is, at what level of intelligence collection is the best balance of cost and information value achieved?

These kinds of considerations, briefly and informally stated above, would be taken into account by CINCEUR in making his intelligence collection decision, whether or not he uses decision analysis. We shall now see how these considerations can be analyzed on a decision diagram. The general structure of a decision diagram corresponding to the above problem is given in schematic form in Figure 5-3. The information to be received is classified as being: highly indicative of impending hostilities (I_H), moderately indicative (I_M), or least indicative (I_L).

In conformity with the general rule for drawing decision diagrams, the order of acts and events is according to when the acts are taken and when the events are known to the decision maker. Thus, the immediate information collection (C) act fork is followed by the information received (I)

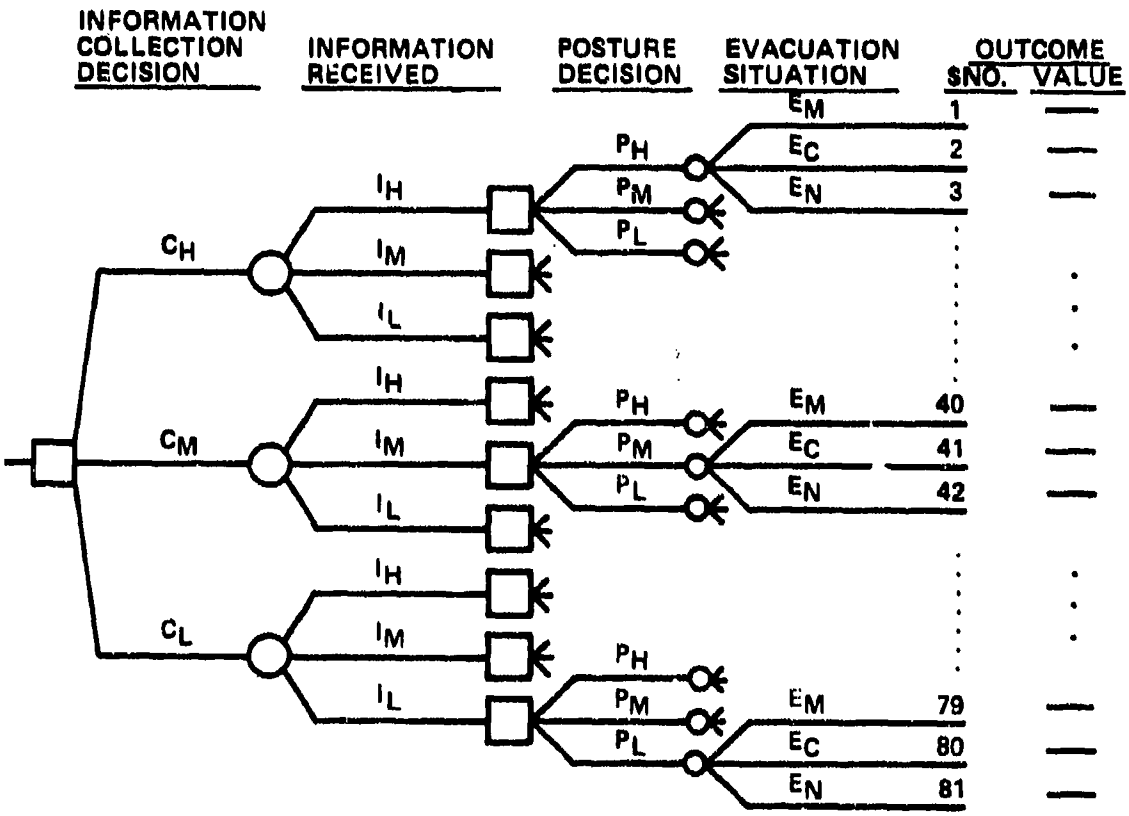


Figure 5-3

SCHEMATIC DIAGRAM FOR INFORMATION COLLECTION DECISION event forks, followed by the readiness posture (P) act forks, followed by the evacuation situation (E) event forks. Each fork happens to have three branches, corresponding to three act options or event possibilities, but that does not have to be the case and generally is not. Any path through the tree corresponds to one possible sequence of acts and events leading to outcomes, numbered from 1 to 81, whose values are to be assigned. For example, the path represented by the upper bounding branches of the tree corresponds to a sequence where a high level of collection effort has been undertaken (C_H), the information received turns out to be highly indicative of hostilities (I_H), a high readiness posture is taken by CINCEUR (P_H), and a situation develops which results in a directive from JCS to undertake major evacuation (E_M). This outcome is numbered 1, and its value will be recorded in the last column.

Quantifying the Decision Diagram

In principle, the problem is now simply one of assigning appropriate values and probabilities to the diagram and folding it back to find the immediate information collection decision with the highest expected value. The required probabilities for the two types of event forks are shown in the probability diagram in Figure 5-4A. The expert needs to know what the probabilities of receiving the three different levels of information are, regardless of the subsequent evacuation situation. He also needs to know the probabilities of the various evacuation situations given the particular level of information received. In principle, there is no reason why CINCEUR could not assess these probabilities directly. The decision diagram could then be folded back by using the techniques described in Chapter 1 to give the expected values of the various courses of action.

For this problem, however, it is much easier for the expert to assess the probability of receiving the various levels of information, given that the various evacuation situations occur, than it is for him to assess the probabilities of the evacuation situations, given that he has received the information. Similarly, it is easier for him to assess the probability of evacuation situations regardless of the information received. In other words, the probabilities which are most natural for the decision maker to assess are just the reverse of the probabilities required by the decision diagram. This suggests drawing the probability diagram shown in Figure 5-4B and then folding it back by using Bayes' Theorem and the techniques outlined in Chapter 4. This procedure will yield the required probabilities.

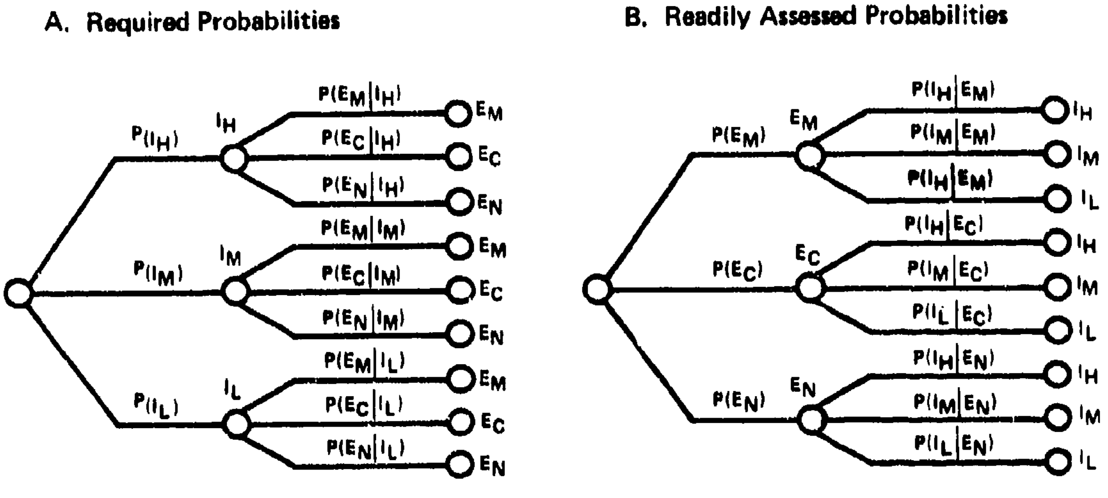


Figure 5-4

PROBABILITIES FOR INFORMATION DECISION DIAGRAM

There is a further complication in the present case, however. The evacuation situation, which is the state of affairs about which the collection system gathers information, is not observable directly. Instead, various levels of activity, such as fleet activities, military leaves, radio broadcasts, travel restrictions, navigation aids, and the like, are observed which provide some clue as to the true evacuation situation. Thus, the information system imperfectly reports about activity levels which, in turn, are probabilistically related to the true evacuation situation. This kind of an inference structure calls for the use of hierarchical inference (as discussed in Chapter 4) to obtain the required probabilities. Figure 5-5 shows the hierarchical structure of an information system for a particular level of collection effort. Using this structure, the expert assesses first the prior distribution, that is, the probability that each of the evacuation situations will occur; then, given a particular evacuation situation, he assesses the probability that each of the three levels of activity will occur; and finally, for each activity level, he assesses the probability that each of the different kinds of information will be received. The dotted line surrounding the activity level and information received forks indicates that the probability of various levels of information being received, given the activity levels, is a function of the particular level of collection employed. The other probabilistic connection between the evacuation situation and the activity level is not under the control of the information system. Because there are three levels of collection being

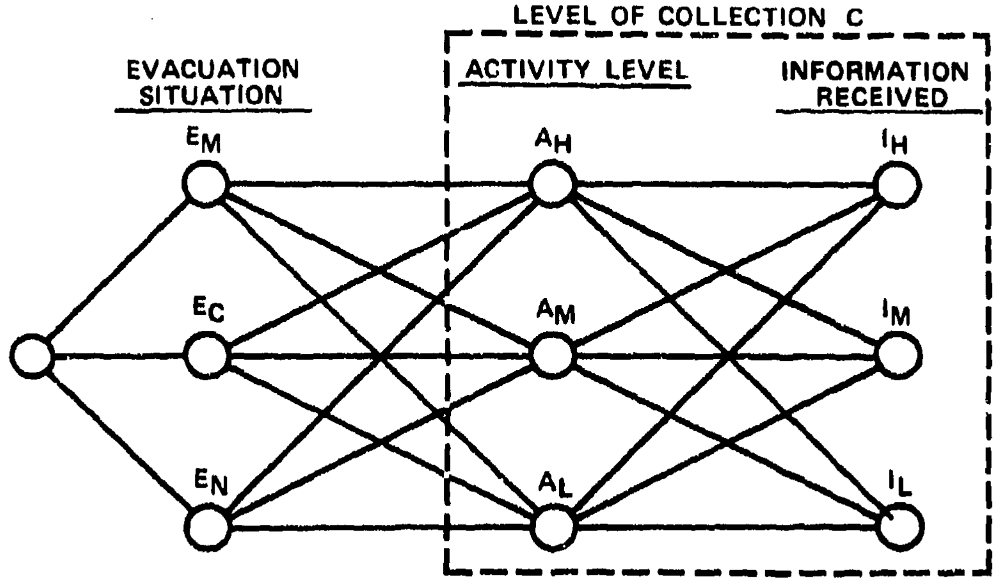


Figure 5-5

HIERARCHICAL STRUCTURE OF INFORMATION SYSTEM

considered in the present case, there will be three probability diagrams which have identical probabilities relating the activity levels to the evacuation situations, but which have different probabilities relating the observed information to the activity levels, which are dependent on the capability of the particular level of collection.

Figure 5-6 displays the assessed probabilities in matrix form for the present case. The expert has assessed prior probabilities of 5% that a major evacuation situation will occur, 55% that a "Cyprus only" evacuation situation will occur, and 40% that no evacuation will become necessary. The first column of the top matrix shows that, given the occurrence of a major evacuation situation, there is an 80% chance of a high level of activity, an 18% chance of a medium level of activity, and a 2% chance of a low level of activity. Similarly, the second column contains the probability of each of the activity levels given that a "Cyprus only" evacuation situation is occurring, and the final column contains the probabilities if the true situation is one in which evacuation is not required. The three matrices at the bottom reflect the extent to which each of the levels of collection reports accurately about the activity levels. For example, the high collection level,

P(E) PRIORS

E_M	E_C	E_N
0.05	0.55	0.40

P(A|E)

	E_M	E_C	E_N
A_H	0.80	0.60	0.10
A_M	0.18	0.30	0.40
A_L	0.02	0.10	0.50

P(I|A)

C_H

	A_H	A_M	A_L
I_H	0.90	0.18	0.005
I_M	0.09	0.74	0.072
I_L	0.01	0.08	0.923

C_M

	A_H	A_M	A_L
I_H	0.86	0.18	0.01
I_M	0.11	0.68	0.12
I_L	0.03	0.14	0.87

C_L

	A_H	A_M	A_L
I_H	0.81	0.19	0.02
I_M	0.13	0.63	0.16
I_L	0.06	0.18	0.82

Figure 5-6

ASSESSED PROBABILITIES FOR HIERARCHICAL INFERENCE

shown on the left, results in a 90% chance of reporting a high level of activity when in fact there is a high level of activity. In contrast, a medium collection effort has an 86% chance of reporting high activity when it occurs, and the status quo, or low collection effort, has an 81% chance.

Figure 5-7 shows how the assessed probabilities can be displayed in a probability diagram for hierarchical inference, in this case, for a high collection effort. The probability diagram in Figure 5-8 shows the required probabilities for the information decision diagram, which were obtained from the diagram in Figure 5-7 by using Bayesian hierarchical inference. Once this procedure has been followed for the medium and low collection efforts, the resulting probabilities may be inserted in the original information decision diagram, Figure 5-3, and the diagram may then be folded back to determine which level of collection effort has the highest expected value and is, therefore, the best course of action to take.

Folding Back the Decision Diagram

For convenience in presentation, the total decision diagram has been divided into three sections, each representing one of the three possible levels of collection.

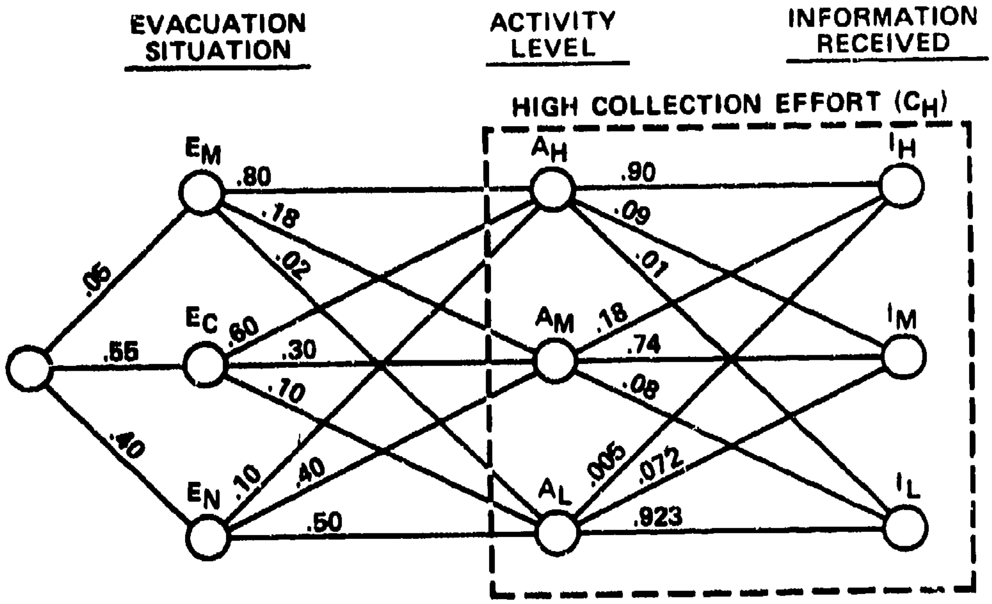


Figure 5-7

PROBABILITY DIAGRAM FOR HIERARCHICAL INFERENCE (HIGH COLLECTION EFFORT)

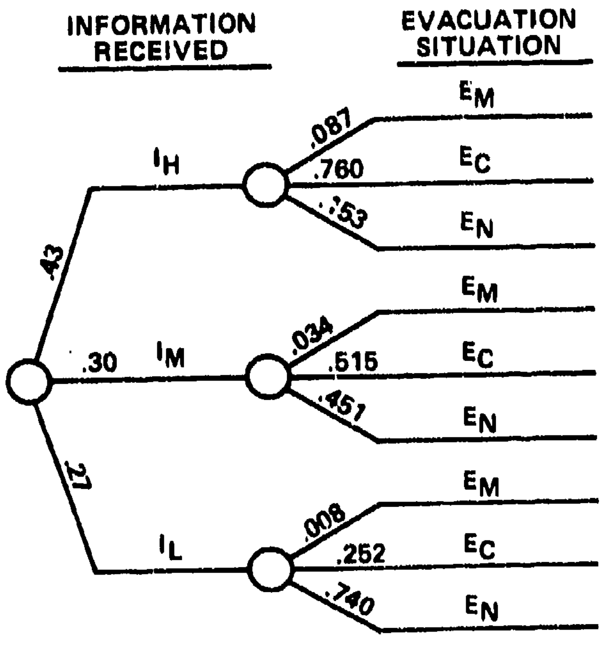


Figure 5-8

REQUIRED PROBABILITIES FOR HIGH COLLECTION EFFORT (C_H)

Figure 5-9 shows the branch leading from the left-most act fork and representing the high level of collection C_H. At the right are the posture and exposure costs for each outcome as well as the collection costs (which are higher for the less extensive evacuations because data would be gathered over longer periods of time). On the branches from the event forks are the probabilities obtained from hierarchical inference. Notice that the listing of the costs is not complete. The reason is that for a given level of collection, the costs are a function only of the posture assumed and the evacuation situation, and so repeat exactly for the branches representing medium and low indicators of activity. The only differences between these branches are the probabilities of the three evacuation situations. First, the costs are added at each end position and then, according to the principles described in Chapter 1, expected values are computed for the event forks leading to the evacuation situations. At each act fork, the posture which has the highest expected value is chosen, as indicated by blocking off those branches which are not chosen. Finally, the expected value for the high collection effort is obtained by multiplying the probabilities of receiving each of the three kinds of information by the expected values at the end of those three branches. In this case, the expected cost of the high collection effort is - \$13.64 million.

In exactly the same fashion, the expected value is calculated for mounting a medium collection effort or maintaining the status quo, as shown by the diagrams in Figures 5-10 and 5-11. The expected values for the two collection efforts are -12.69 million dollars and -12.26 million dollars, respectively. Thus, the best course of action is for CINCEUR to continue with the present level of collection. In addition to identifying which course of action has the highest expected value, the decision diagram also indicates which posture has the highest expected value given the chosen collection effort, namely, the status quo. Figure 5-11 shows clearly that the posture decision depends on the kind of information received. If the intelligence source indicates a high level of activity, then the course of action with the highest expected value is a high readiness posture, P_H. This conclusion is reached by following the I_H branch to the point at the right where the P_H branch was chosen. If the information indicates a medium or low level of activity, the best course is a medium posture, which is determined by following the I_M and the I_L branches to the point where the P_M branch was chosen.

Value of Information

Based upon the expected values of each of the possible collection efforts, the best course of action to take under these conditions is to continue the present collection

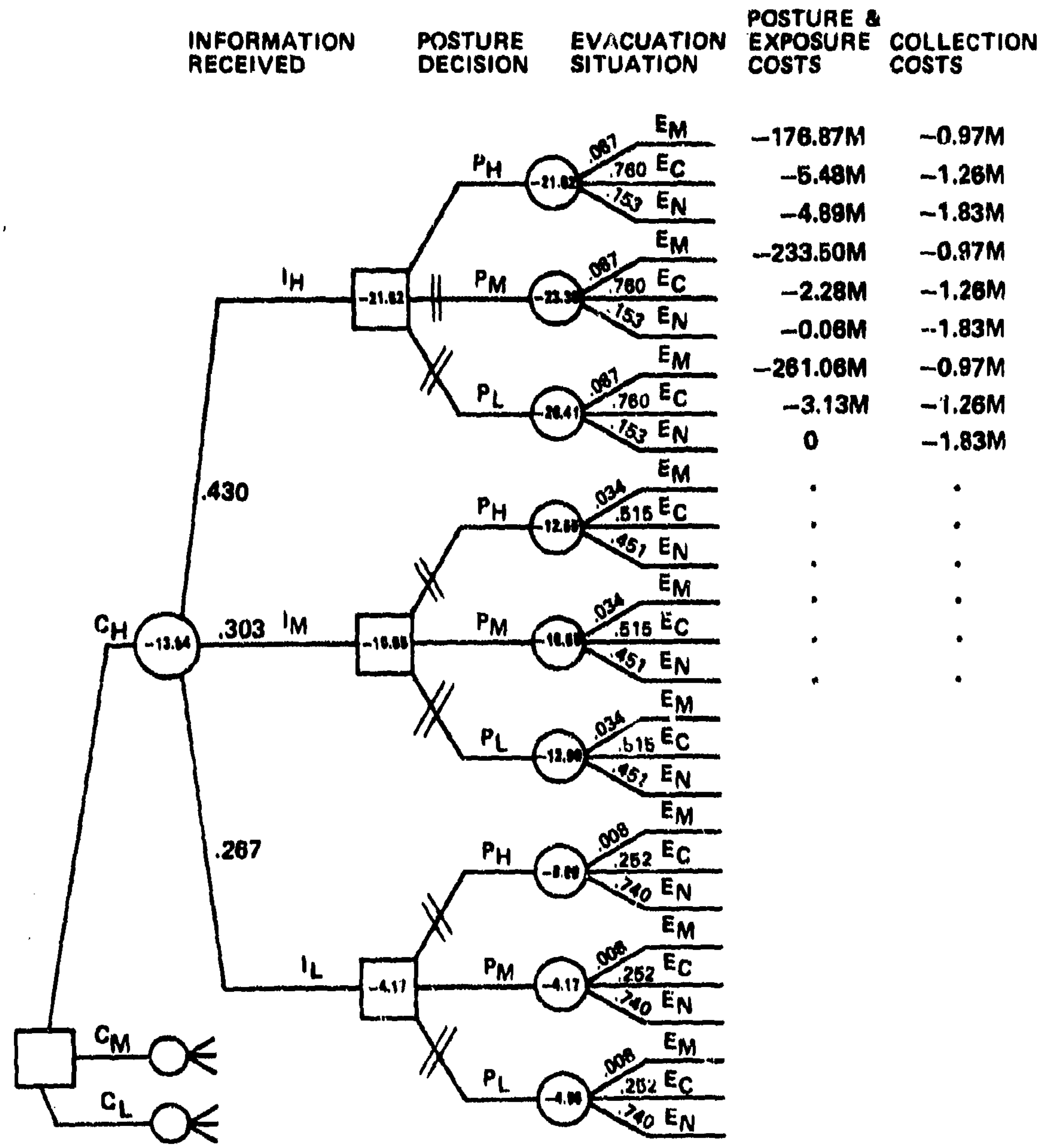


Figure 5-9

HIGH COLLECTION (C_H) BRANCH

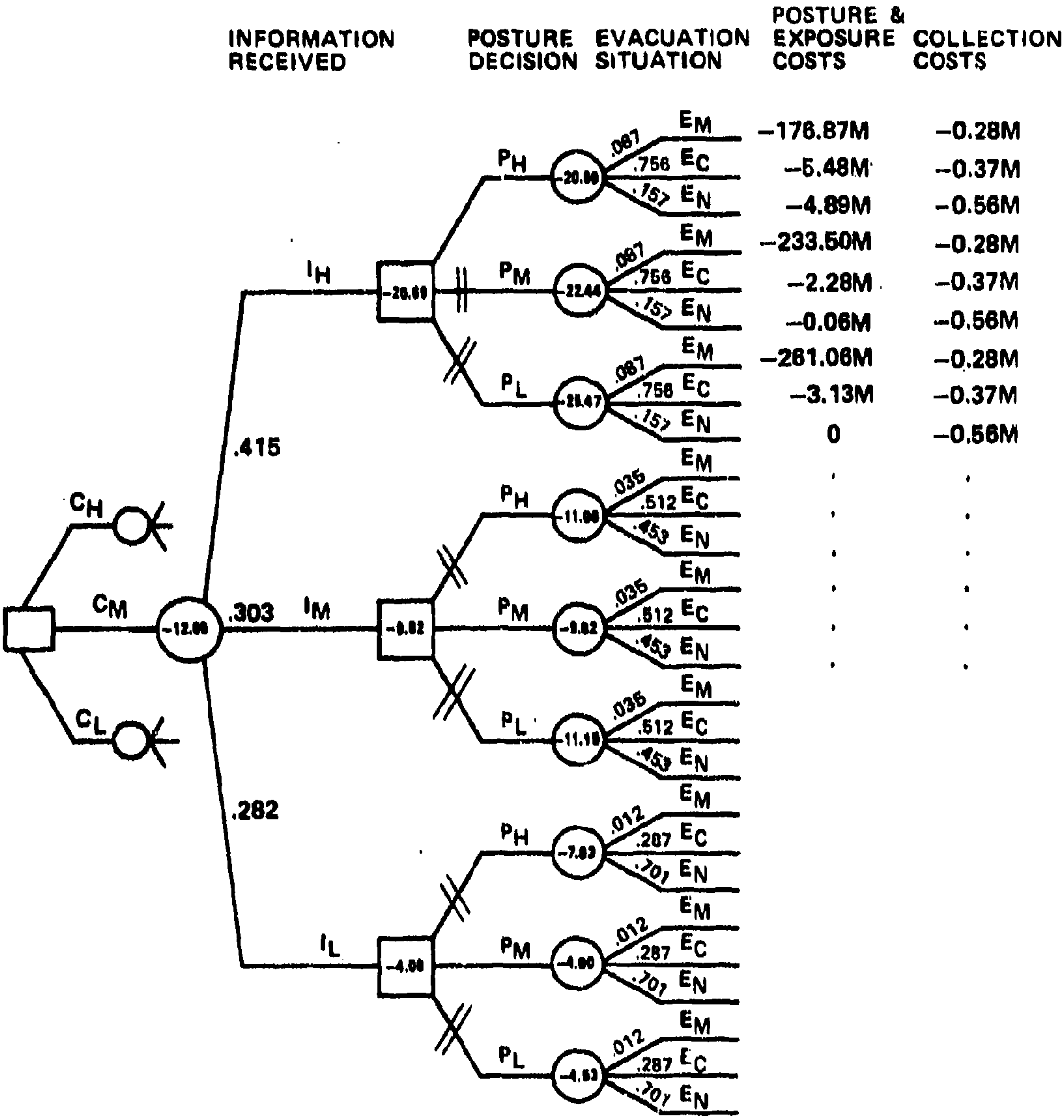


Figure 5-10

MEDIUM COLLECTION (C_M) BRANCH

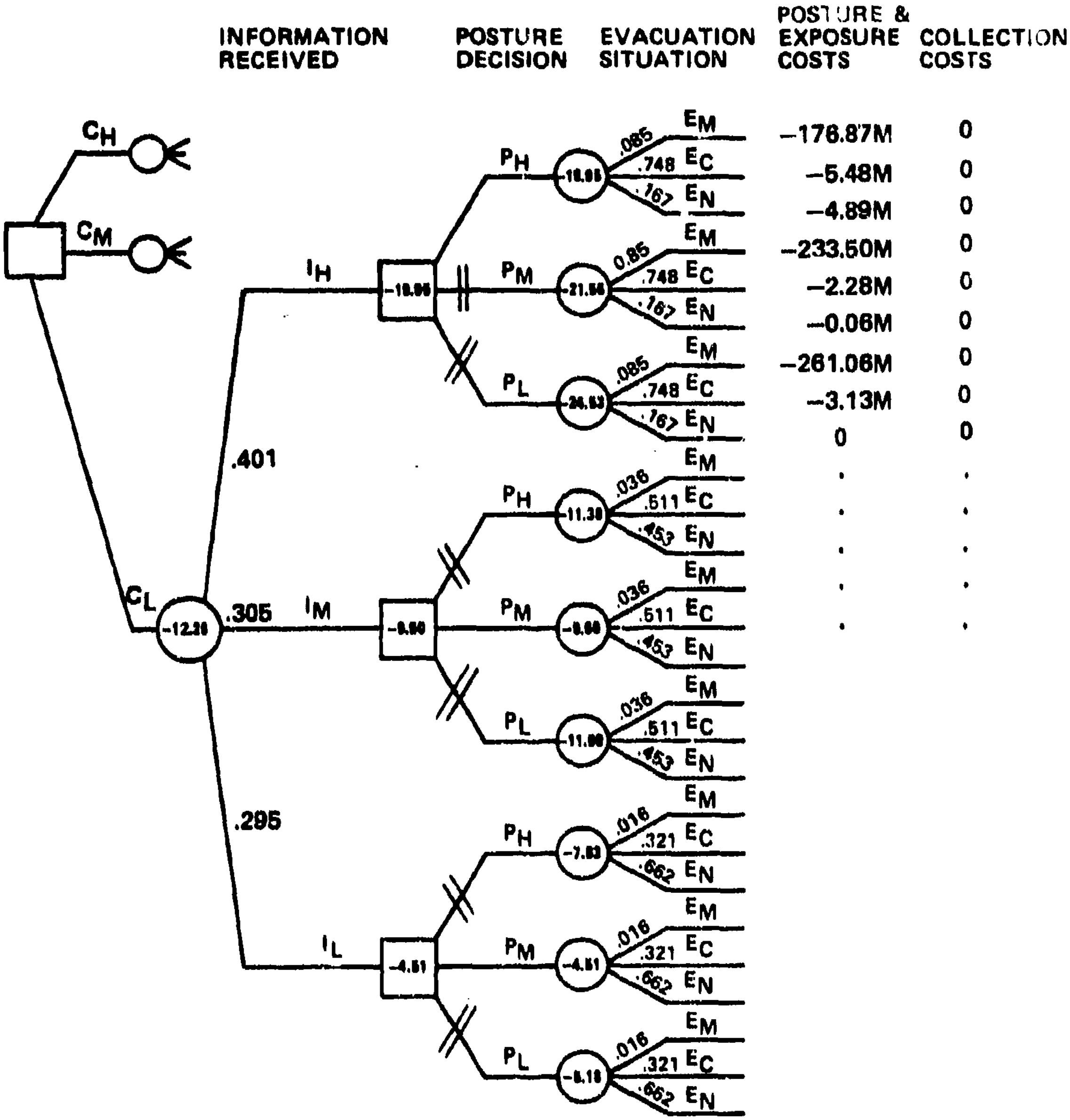


Figure 5-11

LOW COLLECTION (C_L) BRANCH

effort. But the expected values for alternative collection efforts differ for two reasons: first, because the information systems differ in their ability to report accurately the activity level, and second, because they differ in their costs. Compared to the status quo, the additional expected cost of a high collection effort is 13.64 minus 12.26, or 1.38 million dollars, while the extra expected cost of a medium collection effort is 12.69 minus 12.26, or 0.43 million dollars. The results of the decision analysis showed that one should continue the status quo collection effort and that the costs of the medium and high collection efforts are too large for the improved accuracy provided by these collection systems. However, the decision analysis does not tell how much too large these costs are.

An alternative way to analyze the information decision is to discover how much one should be willing to pay to obtain the information capabilities represented by the high and medium collection efforts. This determination can be made by computing the expected values for each alternative as before, without including the cost of collection in the diagram. Since the high and medium collection systems are more diagnostic than the status quo, they must yield a higher expected value than the status quo. In this case, the expected value of the high collection effort is -12.16 million dollars and the expected value of the medium collection effort is -12.25 million dollars. Therefore, since CINCEUR's expected savings are -12.16 minus -12.26 or $100,000 if he uses the high collection system, and -12.25 minus -12.26 or $10,000 if he uses the medium collection system, he should be willing to pay these amounts respectively for this information. Since the two systems actually cost more than this, he will save money by not buying them.

Notice that this method of analysis leads to the same decision as before, but it also gives additional information. It gives the decision maker a good idea of what is required of an information system in order for it to be a good buy under these circumstances. For example, suppose that someone offers a third information system which hadn't been previously considered. Rather than compute the entire decision diagram with the new probabilities and costs, it may be possible to determine that the system is uneconomical without making any calculations. That is, unless the system is substantially more diagnostic of the evacuation situation than those already considered or is considerably cheaper than either for the same level of diagnosticity, it will have a lower expected value than the status quo.

So far, for purposes of illustration, the example involving the Cyprus evacuation has been kept fairly simple. There are only four variables two decision variables, what level of collection level to pursue and what readiness posture to adopt; and two environmental variables, the level of information received and the true evacuation situation present. In this particular example, each of these variables can take on three values, yielding 34, or 81, branches of the diagram. However, in many realistic decision situations, the variables will often take on more than three values, with a resulting complexity that can make the decision diagram unmanageable. Consider what would happen if there were ten values for each of four variables: there would be 10^4, or 10,000, branches for the four variables. Even if the variables do not take on very many values, a realistic problem may contain many variables. If an analysis contained six variables with three branches leading from each, there would be 3^6, or 729, branches at the end of the decision diagram.

Because decision diagrams, especially those involving questions about the value of information, can become so complex, a number of useful analytical shortcuts have been developed to simplify the task of analyzing them. One shortcut is to compute the value of perfect information. That is, suppose you could obtain information from a clairvoyant who knew which evacuation situation was taking place. If you could compute the value of this information, then this value would place an upper bound on the value of any information system: since this is perfect information, you would never be willing to pay more for the information output of any real system. The reason for constructing and analyzing a perfect information model rather than the one involving the actual information options is that it is easier and more convenient to do if the actual information options cost more than perfect information is worth.

To calculate the value of perfect information, a decision diagram (Figure 5-12) which resembles the other decision diagrams (except that the chance fork representing the true evacuation situation can be eliminated) is used. This situation is omitted because once the decision maker has received information from the clairvoyant (that is, when a particular branch from the first chance fork has been taken), there is no remaining uncertainty about the true evacuation situation. Accordingly, no assessments of probabilities are necessary for the true evacuation situation. In addition, since the clairvoyant will provide the information that the situation requiring maximum evacuation is occurring whenever the situation is in fact occurring, the probability of the decision maker's receiving that information is equal to the probability of the situation occurring, or the prior probability. Figure 5-12 shows that the expected value of perfect information is -10.10 million dollars. Since the expected value of the status quo or low collection effort is -12.26 million dollars, the value of perfect information is -10.10 - (-12.26), or 2.16 million dollars.

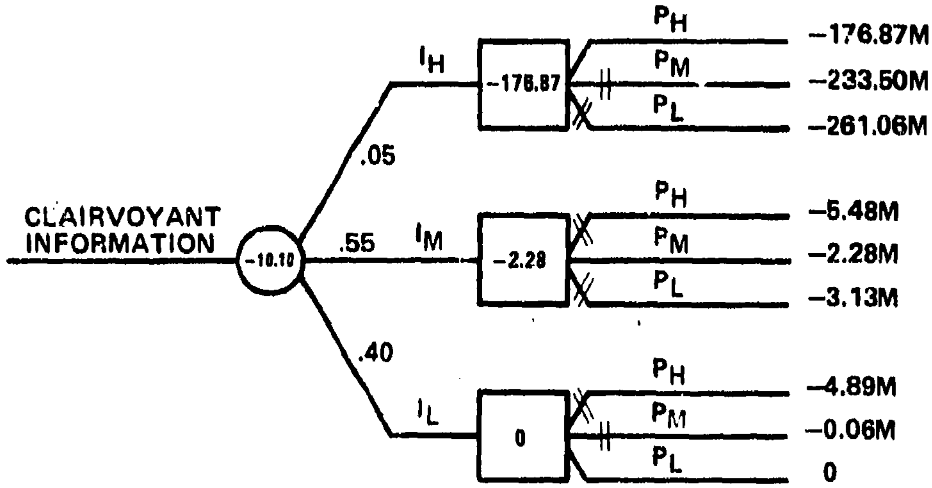


Figure 5-12

DECISION DIAGRAM FOR PERFECT (CLAIRVOYANT) INFORMATION

It is apparent that, in this example, perfect information puts only a weak upper bound on the value of information; the value of perfect information is 2.16 million dollars, which is considerably more than either the high or the medium collection effort costs. However, it is important to remember that the information system was part of a hierarchical inference. The information system was not designed to collect information directly about the evacuation situation but rather about activity level. Since activity level is not perfectly correlated with the evacuation situation, no information system which gathers information about activity level can be expected to predict the evacuation perfectly. This fact suggests a stronger upper bound, the value of a perfect report about activity level. Figure 5-13 shows the decision diagram for the calculation of the expected value of a perfect report about activity level. The only difference between this diagram and the decision diagrams for the high and the medium collection efforts is that the probabilities are derived from a simple single-stage Bayesian inference rather than from a hierarchical Bayesian inference. Specifically, the probabilities can be calculated by applying Bayes' Theorem to the probabilities in the upper matrix of Figure 5-6. The result of folding back this decision diagram is that the expected value of a perfect report about activity level is found to be -11.98 million dollars. Compared to the expected value of the status quo collection effort of -12.26 million dollars, the value of a perfect report about

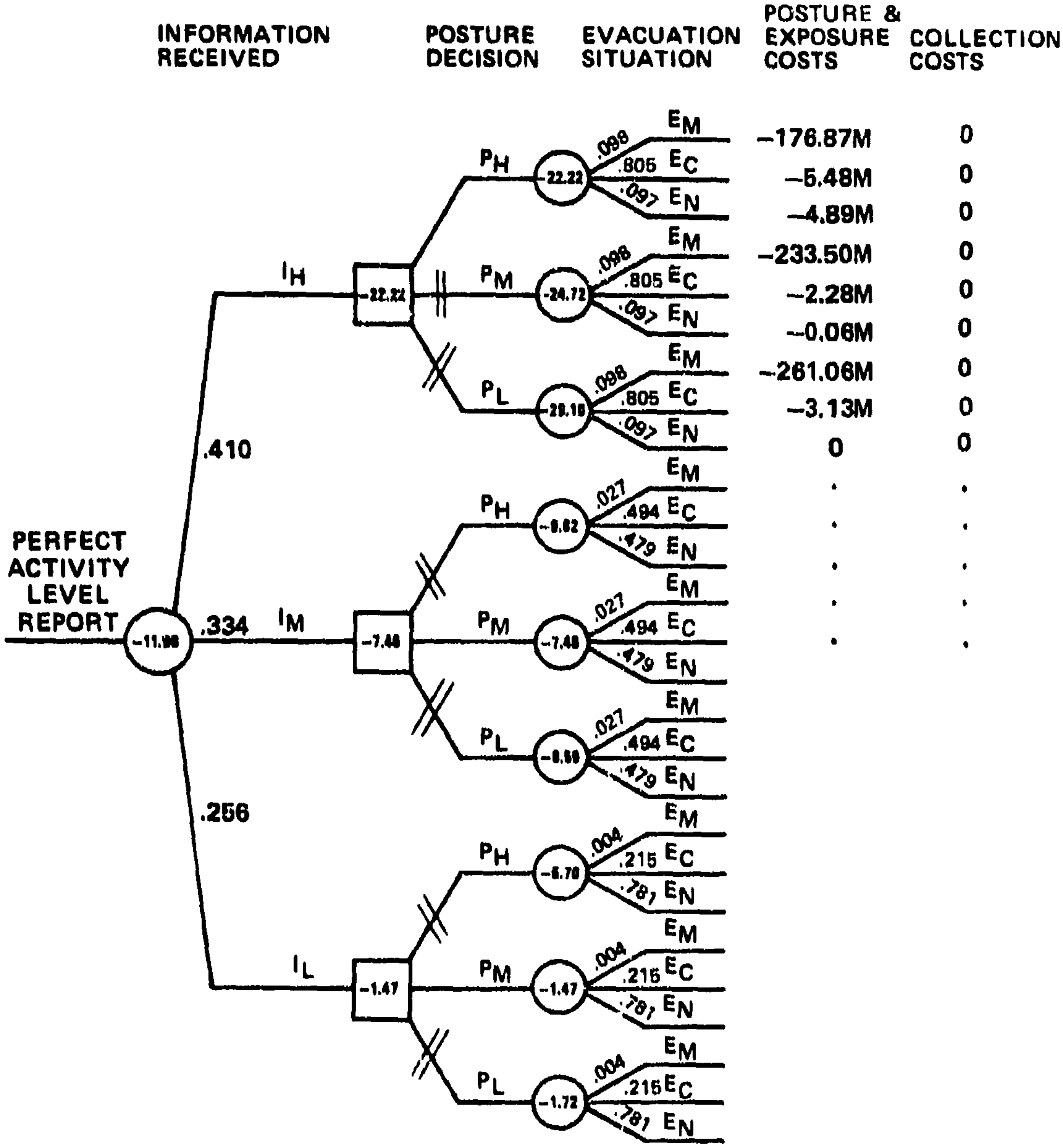


Figure 5-13

DECISION DIAGRAM FOR PERFECT ACTIVITY LEVEL REPORT activity level is $320,000. This means that no information system which reports about activity level, even if it gave a perfect report, is worth more than $320,000. Recall that the moderate collection effort has an expected additional cost of 0.43 million dollars over the low collection effort, and the expected increment for the high collection effort is 1.38 million dollars. Since both are considerably less than perfect in reporting about activity level, it is obvious without constructing decision diagrams for these two levels of information collection that they cost more than they are worth in terms of reducing the expected costs of the posture decision.

Figure 5-14 summarizes the value of information at low, medium, and high collection levels as well as the two upper bounds on information value, a perfect report about activity level and clairvoyance.

	VALUE OF DECISION	VALUE OF ADDITIONAL INFO.	COST OF INFO.	NET VALUE OF INFO.
	(MILLIONS OF DOLLARS)			
C_L	−12.26	0	0	0
C_M	−12.25	+0.01	−0.43	−0.42
C_H	−12.16	+0.10	−1.38	−1.28
PERFECT ACTIVITY LEVEL REPORT	−11.98	+0.28	—	—
CLAIRVOYANCE	−10.10	+2.16	—	—

Figure 5-14

VALUE OF INFORMATION SUMMARY

Value Diagrams

As was pointed out earlier, decisions to acquire information that has the potential of improving a primary decision are called "information decisions." Although the primary decision drives an information decision, the information decision must occur before the primary decision is actually made in order that the acquired information has a possibility of influencing the primary decision.

As described in the preceding section, decision-theoretic analysis of an information decision includes four distinguishable events. First, the information decision whether to purchase information and, if so, what kind and how much is made; second, the information is acquired;

third, based upon the outcome of the information acquisition, the primary decision is made; and finally, some state of nature occurs which, together with the decision made, results in a particular outcome which can be evaluated by the decision maker. The general principle is that information is purchased only if it has a potential for improving the expected value of a primary decision by an amount that is greater than the cost of collecting the information.

One important class of information decisions involves collecting information for intelligence agencies: what type and how many platforms should be used, should a new expensive collection system be purchased, how can the collection budget be cut substantially without doing serious damage to the amount of valuable information being collected, and the like? In these situations it is difficult if not impossible to apply the usual techniques of information decisions. One of the problems is that the primary decisions to be made in the future are not known presently, and even if they were known, it would be difficult to judge the impact that a particular item of information would have on them. Another problem is that the person who makes the information decision may not possess the necessary data to make the primary decision. For example, the manager in charge of allocating collection platforms may not be, and usually is not, responsible for the determination of U.S. foreign policy. Although this separation of functions may not be an ideal state of affairs from a conventional decision-theory standpoint, it is necessary for decision theory to provide a methodology that will operate within real-world constraints. Therefore, even though it is useful formally to consider the role of primary decisions as determiners of information decisions, it is also important to consider procedures for measuring the value of information without explicitly evaluating its impact on primary decisions.

Direct Assessment of Value - Value Diagrams

One theoretical model which is used to measure the value of information in situations in which it is inappropriate to consider primary decisions explicitly is called a value diagram (this approach has also been called assessment of goal-dependent utilities and goal-fabric analysis.)

The basic concept of a value diagram is the construction of a hierarchical structure in such a way that it is possible to measure or assess the relative contribution of different collection systems to the overall goals of collection. This structure is achieved by decomposing the problem so that it is possible to assess the value of information that each of the collection systems contributes to each of the several subgoals and then to assess the importance of the subgoals with respect to the major goals.

Figure 5-15 is a schematic diagram of an actual analysis illustrating how a value diagram works. The goal of the analysis was to assess the relative value of the information being collected by several different, separately fundable collection systems or platforms, with a particular emphasis on evaluating the air platforms. The ultimate purpose of the analysis was to reduce the current budget for air collection in such a manner that it would have a minimal effect on the value of information collected.

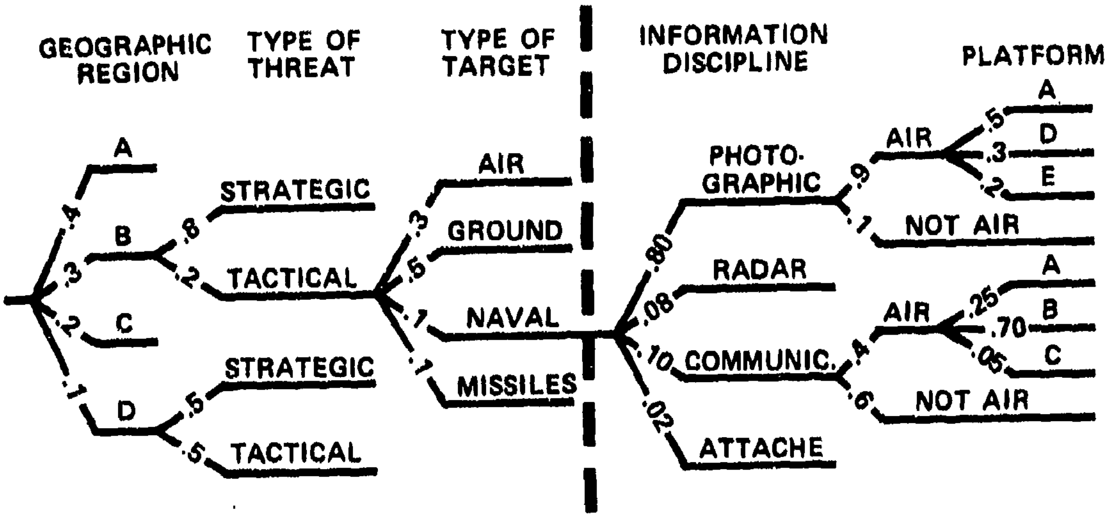


Figure 5-15

PARTIAL VALUE DIAGRAM
FOR INTELLIGENCE COLLECTION PLATFORMS

The left-hand side of Figure 5-15 indicates the goals and subgoals of the information collection and the right-hand side indicates the disciplines and platforms which are responsible for collecting information relevant to the goals and subgoals. The numbers on each branch are importance weights representing the relative importance of each branch subgoal to the subgoal or goal immediately to the left. At each fork the importance weights on all the branches are constrained to sum to 1.0. The importance weights on the left indicate the relative importance of information from different geographical regions, about differing threats, and about various types of targets. On the right side, the weights indicate the assessed relative values of information being collected by the different disciplines about each target type, and then the importance of each system within each discipline.

Consider first the geographic regions. The weights attached to the branches imply that it is twice as important to obtain information in region A as in region C and four times as important to obtain information in region A as in region D. The next branch indicates that, within region B, information about strategic threats is four times as important as information about tactical threats. By contrast, within geographic region D, information about tactical threats is assessed to be as important as information about strategic threats. Similarly, the relative importance of strategic and tactical threat information was also assessed for geographic regions A and C. However, rather than displaying the complete analysis, the diagram in Figure 5-15 illustrates the way the analyst proceeds in constructing the diagram.

Of the tactical information collected in region B, information about ground forces is the most important, and information about air forces is more important than that about naval and missile forces, which are equally important.

The collection part of the diagram is divided into disciplines and platforms. The figure shows that the value of photographic intelligence currently being collected against naval tactical forces in geographic region B is eight times as great as the value of information being collected through the intercept of communications, and ten times as important as information being collected from radar intercepts in that area. Finally, information which is provided by attaches has been assessed to have the least value for naval tactical forces in region B. The figure also shows that 90% of the value of photographic information is currently being obtained by airborne collection platforms and 10% by systems which are not airborne. Of the airborne platforms, aircraft A contributes 50% of the total value of the information collected, and the combination of aircraft D and E contributes the remaining 50%. For informtion obtained from communications intercepts, only 40% comes from airborne platforms, whereas 70% of the value of this information is obtained from airplane B. Notice that aircraft A contributes both photographic and communications information whereas B and C contribute only communication information and aircrafts D and E contribute only photographic information.

When all of the value assessments have been made for all of the branches of the complete value diagram, it can be folded back in exactly the same way as a probability diagram. The relative value of a given platform is a weighted sum equal to the value of the information contributed to a goal or subgoal weighted by the relative importance of those goals and summed across all goals. For example, using the values from Figure 5-15, we can compute the relative value of the photographic information collected by air platform A about tactical naval targets in geographic region B to be equal to the products of all the weights on that path (0.30 x 0.20 x 0.10 x 0.80 x 0.90 x 0.50 = 0.00216).

By proceeding in this manner for all paths through the complete diagram, we can obtain the simplified value diagram shown in Figure 5-16. This diagram shows that air platform A contributes 12% of the total value of all information from all collection systems about all threats in all geographic areas. It also shows that all the air platforms together contribute 30% of the value and "not air" sources contribute 70%.

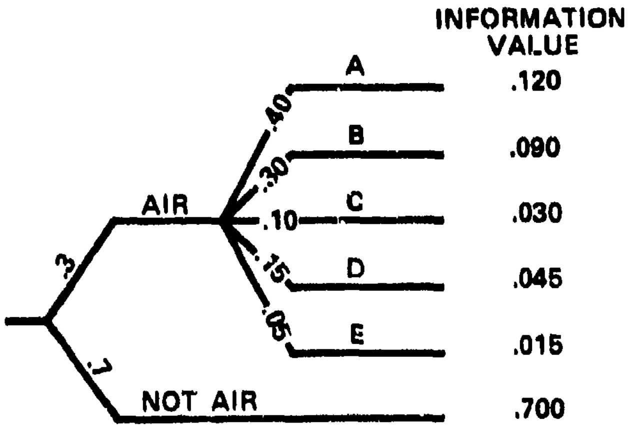


Figure 5-16

SOLUTION OF COMPLETE VALUE DIAGRAM

Methods of Assessing Importance Weights

There are three possible methods of assigning importance weights to the branches of the value diagram. While in theory there should be no difference in the result obtained by the three methods, in practice the three methods usually do differ. Unfortunately, research has not yet determined which of the three methods is most accurate. Therefore, it is a good practice to use all three methods and resolve any discrepancies that appear among them.

The first method is to distribute 100 points across all the alternatives. The advantage of this method is that it resembles the actual resource allocation problem. For example, the assessor, imagining that he had $100 million, might wish to spend $25 million on the Soviet Union, $10 million on East Germany, and so on. The disadvantage of the method is that the resulting distribution curve of money (points) is often too flat. This distribution often results if too many alternatives are considered, or if too great a discrepancy exists between the largest and smallest. In the above example, if the analyst assigns $28 million, or 28 of the 100 points, to Russia, it might be necessary for him to assign .003 points to Ghana in order that the relative sizes be kept accurate. Since analysts generally will not assign less than 1 point to any country, the consideration of many little countries can result in insufficient points remaining to give enough to the more important countries.

The second method is to use relative magnitudes. In this method the most important branch is assigned 100 points and the other branches are assigned points according to their importance relative to this branch. For example, if one branch has one-half the value of the most important, it would receive 50 points. Research has shown that this method also tends to result in a uniform distribution apparently because experts are reluctant to assign small values.

The third method of assessing importance weights is to use a ratio; typically, the most important factor is compared to each of the other factors and a direct ratio is assessed by the expert. For example, the expert might judge that the most important factor had five times the value of the least important but only three times the value of another factor and so on. Research has shown that this method generates a more extreme distribution, but it is not known whether the results are more accurate. However, it is a particularly valuable form of feedback when one of the other methods has been used initially. One possible disadvantage to this method is that a relatively small error in estimation of the ratio may result in a very large error in the final value of information. For example, if an expert assessed the importance of platform 1 to country A to be 3:1 rather than 2:1, and the importance of country A to the European area to be 5:1 instead of 4:1, the resulting relative importance of platform 1 to the European region would be 15:1 instead of 8:1. In contrast, the distribution of points would have been 75 and 25 instead of 67 and 33 in the first assessment, and 83 and 17 instead of 80 and 20 in the second assessment, for the same error to occur. Another possible problem is that the expert may jump from a ratio of 1:1 to 2:1 unless he is encouraged to use ratios such as 3:2, 5:4, and so on.

Once importance weights have been assigned to each of the forks in the value diagram, the diagram may be folded back in exactly the same manner as a probability diagram.

Thus, just as solving a probability diagram yields an average probability or an expected probability, folding back a value diagram yields a weighted value. If the ratio method has been used to check the values or to obtain them originally, the final scale should have the properties of a ratio scale. That is, it should have a meaningful zero (i.e., a zero on the information value scale means zero value, and a weighted value of 6 should be twice as important as a value of 3). Once such a scale has been obtained, it is possible to use the value of information to specify the allocation of resources.

Resource Allocation

One way in which a decision maker might allocate resources is to distribute money proportionally to the value of information; that is, a source of information which has 28% of the value would receive 28% of the resources. Another method is to use a lexicographic decision rule. The information sources are listed in decreasing order of value of information. The resources are allocated to the sources by starting at the top of the list and proceeding down it until all available resources have been allocated. The problem with both of these methods is that they make an implicit assumption that information is equally costly for any of the sources, an assumption which is practically never true. In the best of all possible worlds, the source which has the best information would be the cheapest to operate. In this case, one would purchase as much of that kind of information as possible, then as much of the next most valuable, and so on until the resources have been expended. In the worst case, which may occasionally occur, those sources that provide the most valuable information are the most costly. Accordingly, the decision maker must make a trade-off between the cost of information and the value of that information.

One way of proceeding in this situation is to describe each information source in terms of a value-to-cost ratio. As a first approximation, the optimum allocation policy can be approached by ordering these information sources lexicographically according to decreased value-to-cost ratio. The resulting decision rule says to start buying at the top and to continue to buy sources until the budget ceiling is reached. In some cases, however, this approach will be slightly suboptimum because it is possible that a combination of several information sources, each with a lower value-to-cost ratio, will give a slightly higher aggregate value than one information source with a slightly larger ratio. If the number of possible sources is small, it is possible to obtain the optimum solution by merely listing all combinations of sources and then selecting the combination which has the largest value for a total cost less than the budget ceiling. Unfortunately, since for even a modest number of sources, the total number of combinations is astronomical, a straightforward application of this algorithm is impractical. The optimum solution of the situation can be found by using integer programming techniques. Integer programming is a computational methodology which maximizes one variable, in this case, information value, subject to a constraint on another variable, in this case, cost.

In many applications, it is not merely a matter of deciding whether or not to use a particular source of information, but rather how much information to acquire from each source. In this case, it is not possible to describe the source of information in terms of a single ratio relating value to cost. In general, the value of a piece of information, i, is not the same as the value of the piece of information (i-1) that preceded it. It is often the case that the relationship of information value to cost is a monotonically increasing, negatively accelerated function. That is, the more information you receive, the more total value it has, but each increase in information, for a given increment of cost, is worth less than the previous increase for the same increment of cost. Under these conditions the optimum allocation of resources can be found by the following graphic technique:

First, a curve relating information value to cost for each source is drawn. Then, on the graphs relating value to cost for each source, a very high ratio of value to cost is set, and a straight line with that slope is drawn through the origin. The resulting cost of this allocation will be the sum of costs associated with each of the points at which this "criterion" line intersects each of the curves. (Because of the properties of the function, a straight line from the origin through any point on the curve to the left of the "criterion line" intersection will have a slope greater than the "criterion line" and, therefore, a greater value-to-cost ratio.) If the initial choice for the "criterion line" slope represents a sufficiently high value-to-cost ratio, the total cost obtained will be less than the cost ceiling. The process is then repeated, successively decreasing the slope of the criterion line. When the slope (value-to-cost ratio) of a chosen line results in costs for each source, derived from the intersection point on each curve, which add up to the cost ceiling, an optimal resource allocation will have been achieved for that cost constraint.

This technique works for any monotonically increasing function which has no inflection point. However, many actual collection systems exhibit an S-shaped function. This function occurs because, at a low level of collection, much information obtained is not identified or correlated with previous information and therefore has very low information value. Once a certain amount has been collected, however, the information begins to fit together in a pattern of significant interest. At this point, any subsequent increase in the level of collection results in a large increase in value until finally the system begins to approach the physical limits of information it can collect, at which point the value-to-cost ratio levels off again. In any function which has an inflection point, the "criterion line" may intersect the curve in more than one place; it is not, therefore, a simple matter to determine the optimum allocation of resources. In this case, the usual procedure is to try several alternative allocations in order to reduce the chance of a local maximum being found. A more systematic approach, however, employs a general, non-linear programming algorithm, which will nearly always produce the optimum solution.

CHAPTER 6

APPLICATION OF DECISION-ANALYTIC METHODS: A CASE STUDY

This chapter presents a case study of a decision analysis as an illustration of the way in which decision theory can be used to clarify a problem involving uncertainty and complex value judgments and thus help the decision maker to determine which of several courses of action is the best one to take. The example is simple, but it is extended in many directions which are often of interest to a decision maker. Very few real problems will be this simple, nor will they investigate so many avenues; typically, a decision maker is interested in a detailed application of one or two aspects of decision analysis rather than a more superficial application of many.

Within these restrictions, the analysis of the case study is developed in a logical fashion, much as it might be in a real decision problem. In this process, we use many of the concepts introduced in the previous chapters: decision tree, probability, subjective value, expected value, revision of probabilities, and the value of information. In addition to illustrating how these concepts can be applied, we also want to show how a decision analysis might proceed.

The case study involves a problem faced by the commander of a small naval task force in the scenario depicted in Figure 6-1. One of his ships has been attacked by an enemy airplane, and it has sustained damage and casualties. However, this is the first time such an attack has ever occurred, and the damaged ship, USS HERRINGBONE, did not fire at the attacking aircraft, although it had been given permission to fire at will. The task force commander, anticipating that such an attack might occur again in the future, wants to know whether the captain of the HERRINGBONE made the correct decision, given his uncertainty about the identity of the approaching aircraft. More generally, he would like to know if additional information would have resulted in a better decision, and how much that information would be worth. To help him answer these questions, he has called in a consultant who is expert in applied decision analysis.

What follows is a dialogue between the task force commander and the expert consultant, in which they try to arrive at a formal representation of the problem, and, from that, the solution. Notice that as the problem develops, it is necessary for the consultant to learn to understand the captain's view of the problem as well as to understand the problem itself. Unless the consultant can understand the

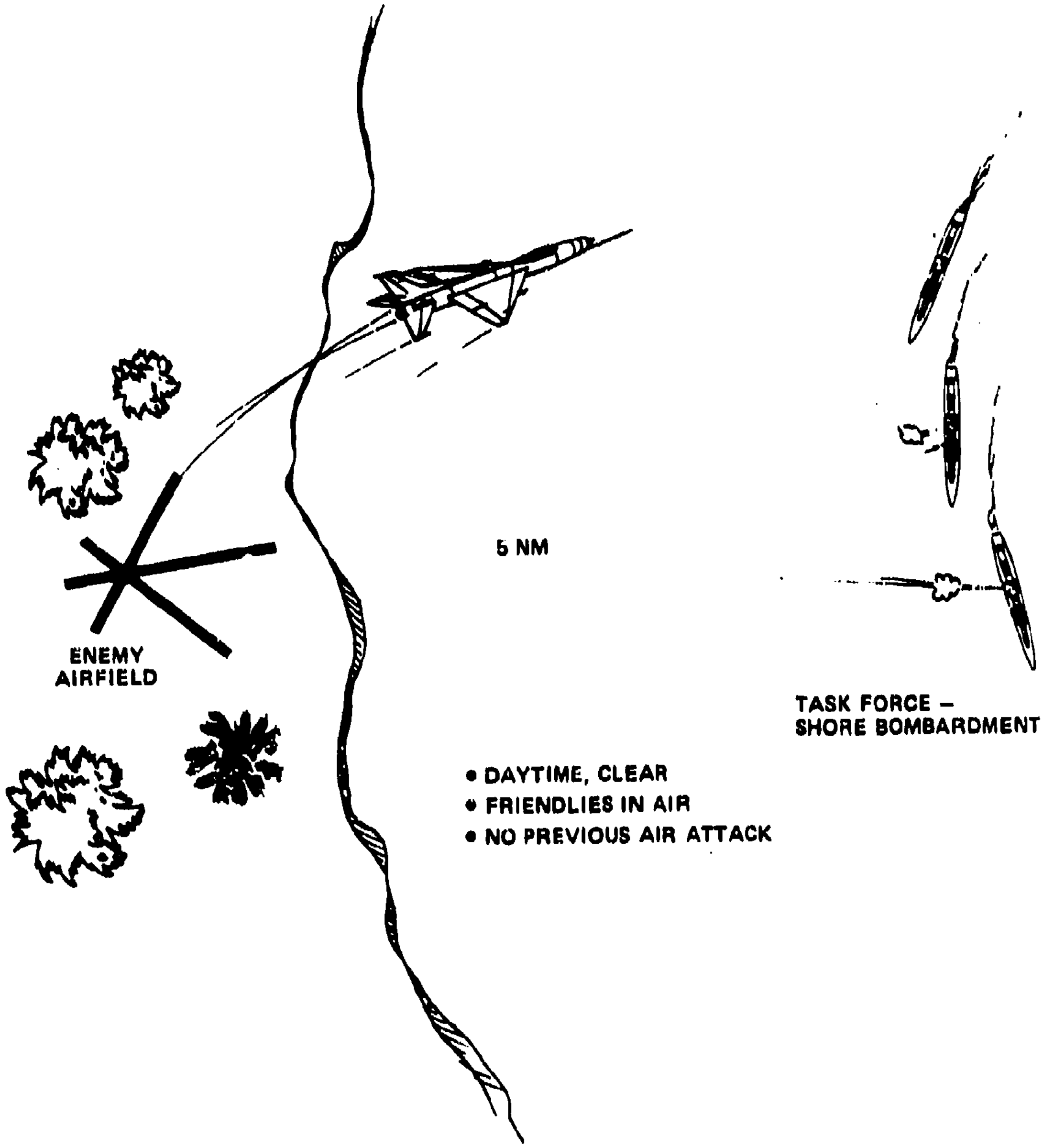

ENEMY
AIRFIELD
5 NM
TASK FORCE –
SHORE BOMBARDMENT
• DAYTIME, CLEAR
• FRIENDLIES IN AIR
• NO PREVIOUS AIR ATTACK

Figure 6-1

motivations of the decision maker as well as the objective characteristics of the problem, he cannot really make full use of the tools of decision theory.

Structure of the Decision

Commander: Let me describe the problem. Imagine that I have just picked up an unidentified aircraft on radar, and it is coming straight for my ship. It could be an enemy plane attacking, but it could also be one of our aircraft which has been shot up and is ditching in the water near my ship to be rescued. Now, of course, I have various kinds of information about whether it is friendly or not, but I can never be absolutely sure, and I may not have time to look at my information before I have to make a decision. I'd like to analyze this situation now so I can make a more informed decision if the problem ever occurs again. How do we start?

Consultant: I think the way to begin structuring the problem is to use a simple decision diagram (see Figure 6-2). We begin on the left, drawing lines for the options available to you and labeling them. This is called an act fork. In this case, there are only two available courses of action, SHOOT and NO SHOOT. Now, what will happen if you shoot at the airplane?

Commander: Well, you can either knock him out of the air, or you can miss him. Of course, our crews are trying their best to bring him down. And sometime during or after the action, we will find out that the airplane is a friend, or that it is an enemy.

Consultant: We'll represent the first set of possibilities as KILL or NO KILL, the second as FRIEND or ENEMY. In neither case are the possibilities entirely under your control, so we call them events and represent them by event forks. To distinguish the act forks from the event forks, I will use a box for the act forks, and a circle for the event forks. In drawing the event forks on the decision tree, I would normally show them in the order that the events will become known to the decision maker--shoot first, then learn the identity. But there are exceptions to that rule, and this may be one. It seems clear to me that the chances of the airplane being a friend or an enemy are the same whether you score a hit or not, but isn't it possible that the chances of a hit might be affected by the aircraft being a friend or enemy?

Commander: I suppose that could be so; some types of airplanes are slightly more difficult to hit then others.

Consultant: Then to be safe, we'd better show KILL or NO KILL as following FRIEND or ENEMY. So, we will put

FRIEND or ENEMY event forks at the end of each of the initial acts branches and, then, in the case of shooting, follow these with KILL or NO KILL event forks (see Figure 6-2). I assume that if an enemy gets through, he will try to bomb the ship.

Commander: That's right, but he may not score a hit.

Consultant: So, to complete the tree, we will add an event fork, with branches HIT and NO HIT, at the end of the path SHOOT, ENEMY, NO KILL, and the same fork after NO SHOOT, ENEMY (see Figure 6-2). Now, it is obvious that we can extend this structure in many ways, but at this point I think we have the basis for exploring the decision when we incorporate your opinions and values.

The first task is to decide what the outcomes are, and what attributes are important for assessing the value of those outcomes. The first outcome, A, is that you shoot at a friendly aircraft and hit it. What are the important things about this outcome?

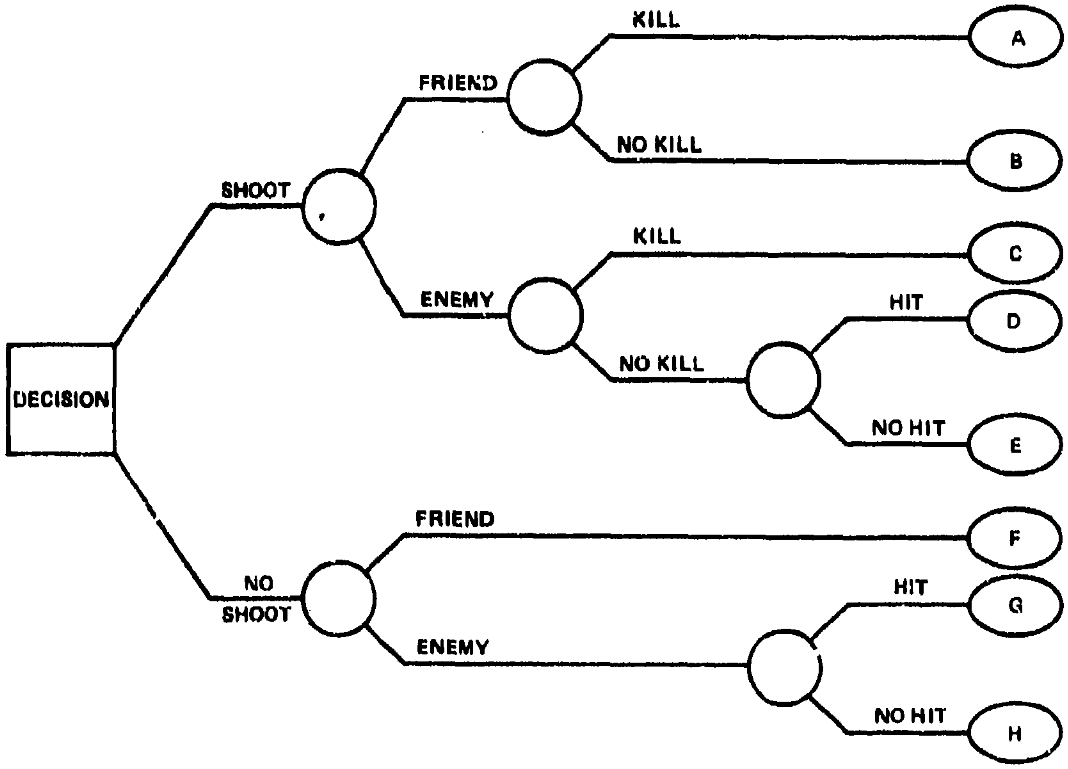


Figure 6-2

Commander: The most important thing is, of course, the life of the pilot. Second to that comes the value of the airplane.

Consultant: O.K., that seems reasonable, but when you say the life of the pilot, how do you mean that? Do you mean the effect that killing one of your own aviators has on the U.S. public, or on the morale of the men in the convoy, or do you mean the dollar value that the loss of an aviator means to the war effort?

Commander: Well, I hadn't really thought about it that way, but I guess I mean all three. I'm certain that all three of these enter into my decision to shoot or not shoot.

Consultant: Then we're going to have to make some assumptions or the decision diagram may become unmanageable. Let's assume that if you hit the aircraft that you kill the pilot. Also, we have to assume a particular type of U.S. airplane; let's say an A-4. What are the consequences if you miss him? That's outcome B.

Commander: Nothing happens. Well, you lose a missile, or a shell, but they're expendable, so they don't really count.

Consultant: At outcome C, you shoot down an enemy airplane. Again, we assume you kill the pilot, and again we will assume a particular enemy aircraft. What kind of plane did the enemy use when they attacked the HERRINGBONE?

Commander: They used an old MIG-17 that they got from the Soviets.

Consultant: What about outcome D? In this case, the enemy aircraft gets through, bombs the ship, and scores a hit. What damage would you say would be sustained?

Commander: Well, the HERRINGBONE lost eight men, but it was an unusually light hit. I would think that, in general, about 25 men, perhaps 3 officers and 22 enlisted men, would be lost in this kind of attack. And, of course, the ship would have to be dry-docked for repairs.

Consultant: Now let's move on to outcome E; the enemy fails to hit the ship. In this case, he quickly leaves the scene, so nothing further happens. At outcome F, you don't shoot at a friend, so nothing happens. Next, at outcome G, your ship is hit because you failed to shoot at the enemy aircraft and his bombing run was successful from his point of view. It looks like that outcome is almost the same as D. Finally, outcome H is like outcome E.

Value of Outcomes

In order to structure any decision problem, the person making the decision must assess how much each of the possible outcomes is worth to him. As we explained in Chapter 2, this worth is called the utility of the outcome, and when it is composed of several attributes or dimensions, it is called multi-attribute utility. One strategy here is to assess how much each outcome is worth for each attribute and how important each attribute is, then take the values from each of the attributes, weight them by the importance of the attribute, and combine them into an overall utility. This is the procedure followed by the consultant in this case.[1]

Consultant: Now we have to discover just how much each of these possibilities is worth to you, the captain of the attacked ship. To do this, we will use each of the four attributes we have talked about: men, equipment, political value, and morale. Let's start with the MEN attribute. Probably it is easiest to compare the value of one U.S. pilot to the value of one enemy pilot.

Commander: My first impulse is to say that U.S. pilots are more valuable because they are better pilots than the enemy, and it costs us more to train them than it does the enemy. On the other hand, we have a much larger supply of them than they do, and we can produce them in less time.

Consultant: O.K., I want you to think aloud and fill in the details of these considerations. The details themselves aren't too important, but as you think through the concrete facts, you should be more able to come up with a consistent assessment of the relative worth of the two aviators.

Commander: First of all, the U.S. produces 5,000 pilots a year, at a cost of about $1 million each, and it takes approximately 18 months to train them. The enemy, however, probably produces less than 100 pilots a year at a cost of perhaps $100 thousand each, and it takes about three years to train them. Also, it is important to remember that

---

[1]However, it should be pointed out that because direct assessments of worth are made, the commander's attitude toward risk is not accounted for, as would be the case if utilities were assessed using lotteries. Thus, the resulting assessments of worth here are equivalent to utilities only if the commander is assumed to be risk-neutral. To insure that this assumption is not forgotten, we will refer here to the assessments of worth as values rather than utilities.

the enemy pilots are generally less qualified than ours, so that the loss of a good pilot, like the one which attacked the HERRINGBONE, hurts them more than it does us. They have significantly fewer pilots than we do, but a higher percentage of them are fighter pilots. Of the fighter pilots, less than 10% are qualified enough to fly high-threat or night missions.

Consultant: Based on these facts, who stands to lose more from the loss of a qualified fighter pilot, the U.S. or the enemy? And how much more?

Commander: Things would be pretty even, except for the extreme shortage of enemy pilots. I would say that losing a qualified fighter pilot hurts them about three times as much as it does us. Of course, this is just in terms of the cost of replacement, how much they are worth to the war effort.

Consultant: Since you said that it cost about $1 million to replace a U.S. pilot, this would mean that an enemy pilot is worth $3 million in "equivalent" dollars. "Equivalent" means that they would just as soon lose $3 million as lose an aviator; it doesn't mean that they pay this much for one.

Now, on outcomes D and G, the U.S. loses three ship's officers and 22 ship's enlisted men. Since the Navy must train these men as well as the aviators, it should be relatively easy to assess how much they are worth.

Commander: Of course, I may want to check this out more exactly later, but offhand, I would say that one U.S. aviator is worth two ship's officers, or five enlisted men.

Consultant: Then the personnel loss due to the bombing of the ship is equal to 5.9 aviators, or 5.9 million "equivalent" dollars. To help you visualize all this, I'm going to make a scale of the values of the various outcomes (Figure 6-3). Outcomes B, E, F, and H are at zero, because in each of these outcomes, no men are lost. C is at +$3 million, because the enemy's loss is our gain, and outcome A is at -$1 million, which is the cost of the loss of our aviator. Finally, outcomes D and G are at -$5.9 million, because that is the loss of 25 men from the bombing of the ship. Does this scale look reasonable to you? Remember, we're considering only the men here, not equipment, political value, or morale.

Commander: Yes, that scale looks about right. Of course, since I came up with those values myself, it would have to look right, wouldn't it?

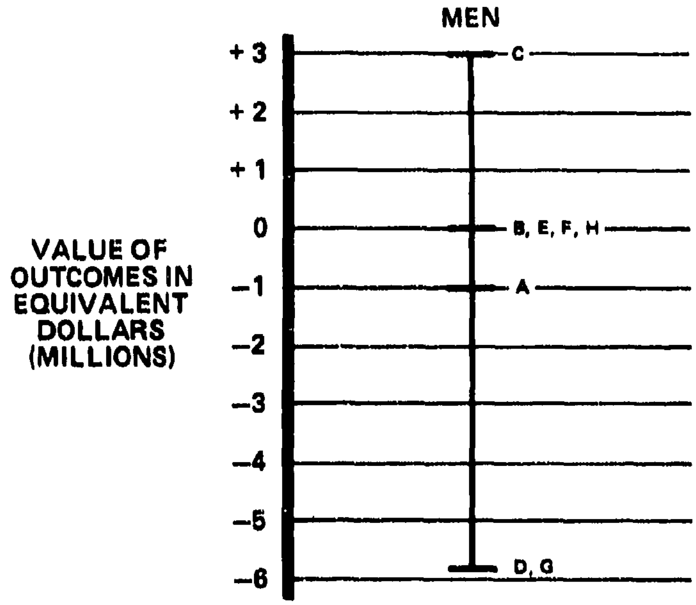


Figure 6-3

Consultant: Not necessarily. Sometimes in trying to break down a complex problem into its component parts we overlook something important, so it's best to check the formal procedure with your intuition every once in a while. When there is disagreement, this indicates that something has been overlooked. It may be that the formal procedure is being applied wrongly or your intuition is faulty or both. However, it is usually possible to resolve the difference without too much trouble.

Let's proceed to the EQUIPMENT attribute next. What is the cost of replacement for the U.S. aircraft?

Commander: Well, let's see. We said that it was an A-4, right? One of those costs about $2 million, I would guess. And I'll bet that next you'll want to know what the enemy airplane was worth, right?

Consultant: Right.

Commander: Well, Soviet MIG-17's aren't exactly the newest and best airplanes available, not as if they were MIG-21's. I would give a figure of about $1 million.

Consultant: In outcomes D and G, the enemy bombs the ship. What is the cost of that damage?

Commander: Whew! That's not so easy to figure out. First of all, there is the work required to repair the ship, which I would guess is around 200 man-months. Then, of course the ship's crew is idle for the time the ship is in dry dock, so we have to figure the down time for the ship's company--make that 300 man-months. That comes to a total of 500 man-months, which is 42 man-years; and figuring about 20 thousand dollars per man-year, that's $840,000.

What else is there? Oh yes, there's the cost of the materials to repair the ship, say $2 million, and of course, you have to bring another ship on the line to replace the damaged ship. This means transporting it across the ocean, which costs fuel and so on. Let's say this costs about $1 million. What does that come to?

Consultant: The total is approximately $4 million in costs. With this information, I think we have enough to make a scale (Figure 6-4) which is like the one for the men dimension. The value of shooting down the MIG is +$1 million (C), the value of either shooting at a friend and missing, or not shooting at him (B or F) is zero, the value of losing an A-4 is -$2 million (A), and the cost of the ship being bombed, because you failed to shoot at an enemy, or shot at him and missed (G or D) is -$4 million. Oh yes, an enemy who manages to make a bombing run but misses the ship (E or H), that's worth zero as far as equipment costs are concerned. Does the scale look about right? Notice that the length of the whole scale is $5 million, while the length of the MEN scale is nearly $9 million. This means that the EQUIPMENT attribute is about 5/9ths or slightly more than half as important as the MEN attribute. Do you agree?

Commander: Yes, I would consider the loss of life to be more important than the loss of equipment, and in about that ratio.

Consultant: Now, for the POLITICAL attribute. Here, there are no convenient units to measure things in, like dollars. However, I think we can still obtain the necessary values. To begin with, let's order the outcomes. What is the most favorable outcome politically? We're talking about the reaction back home, what the media might say, and potential action by Congress.

Commander: Of course, the best outcome is to shoot down an enemy airplane (C).

Consultant: What's the worst?

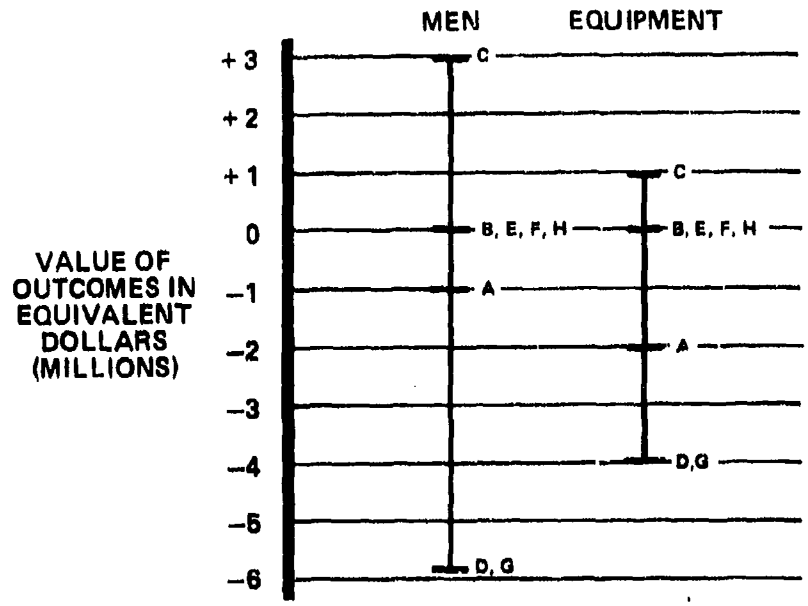


Figure 6-4
MEN AND EQUIPMENT

Commander: The worst possible thing is for the enemy to bomb your ship without your ever taking a shot at him (G).

Consultant: O.K., now order the other outcomes between C, shooting the enemy, and G, failing to shoot at the enemy and getting hit.

Commander: Well, the worst thing next to G is D, the enemy bombs your ship because you missed him when you shot at him. Next to that is A, shooting down one of your own aircraft and losing the pilot. After that is B, shooting at one of your own men and missing. I guess that H, not shooting at an enemy whose bomb misses the ship, is not quite as bad as B, and a little better than H is E, shooting at and missing an enemy who in turn misses the ship. Finally not shooting at one of your own aircraft, F, is neutral.

Consultant: Now that you have ordered the outcomes (Figure 6-5), the next step is to judge how much better or worse one outcome is compared to another. We can construct a relative scale of values using 0 for the worst outcome (G) and 100 for the best outcome (C). Now, relative to not shooting at one of your own aircraft (F), how much worse is it to have the ship bombed without a shot being fired (G), than it is to shoot down one of your own planes (A)? We're comparing G to A, relative to F.

1	C	SHOOT, ENEMY, KILL
2	F	DON'T SHOOT, FRIEND
3	E	SHOOT, ENEMY, DON'T KILL, NO HIT
4	H	DON'T SHOOT, ENEMY, NO HIT
5	B	SHOOT, FRIEND, DON'T KILL
6	A	SHOOT, FRIEND, KILL
7	D	SHOOT, ENEMY, DON'T KILL, HIT
8	G	DON'T SHOOT, ENEMY, HIT

Figure 6-5

POLITICAL VALUE: RANK ORDER OF OPTIONS

Commander: I would say that it is three times as bad to have the ship bombed. After all, we do occasionally shoot down one of our own planes, and while it is not a popular thing to do politically, it is certainly much better than allowing your ship to be bombed without even taking a shot.

Consultant: Well, since you said that F, not shooting at one of your own aircraft, is neutral, this means that the distance from F to G must be three times as great as the distance from F to A. That makes it three times as bad. What about D, shooting at the enemy but missing him so that he gets through to bomb the ship?

Commander: Well, that's almost as bad as G; put it down very close to G on the scale.

Consultant: O.K., what about shooting at a friend and missing him (B)?

Commander: Well, the pilot or one of his buddies might write someone, and that someone could write his congressman, and that would result in political pressure.

Consultant: How does it compare to shooting him down?

Commander: Of course, shooting him down is worse, I'd say almost twice as bad. Let's see, that must mean that the distance from F to B is a little more than halfway down the distance from F to A.

Consultant: That's good. Now, what about C, shooting down an enemy?

Commander: Well, politically, it's just not that big a deal to shoot down the enemy. After all, it's what you're supposed to do, so no one really pays much attention when you do it. It certainly doesn't have nearly the political impact of shooting at one of your own airplanes, even when you miss.

Consultant: How does the political impact of shooting down the enemy (C) compare to the impact of shooting and missing a friend (B)? Remember, we're making the comparison relative to F, not shooting at one of your own aircraft.

Commander: Well, I'd say that shooting at a friend and missing him has three times the political impact.

Consultant: O.K., so the distance from F to C is one-third the distance from F to B. What about outcomes E and H? You said they fall between F and B; can you say where?

Commander: That's difficult. E is a little worse than F, and H is a little worse than E but a little better than B. Make the differences all about the same.

Consultant: Now we have enough information to construct a scale (Figure 6-6) which gives us the relative values of each of the outcomes on the political attribute.

Commander: I must say that putting numbers to these political values lends spurious precision to my judgments. I feel pretty fuzzy about most of these values.

Consultant: Well, if we had fuzzy numbers we'd use them. But we don't, so for the moment, let's leave the over-precise numbers, and remember that each number represents an approximate location on the scale. Later, if we wish, we can move those numbers about to see the effects your fuzziness has on your decision. That would be one kind of sensitivity analysis.

POLITICAL VALUE

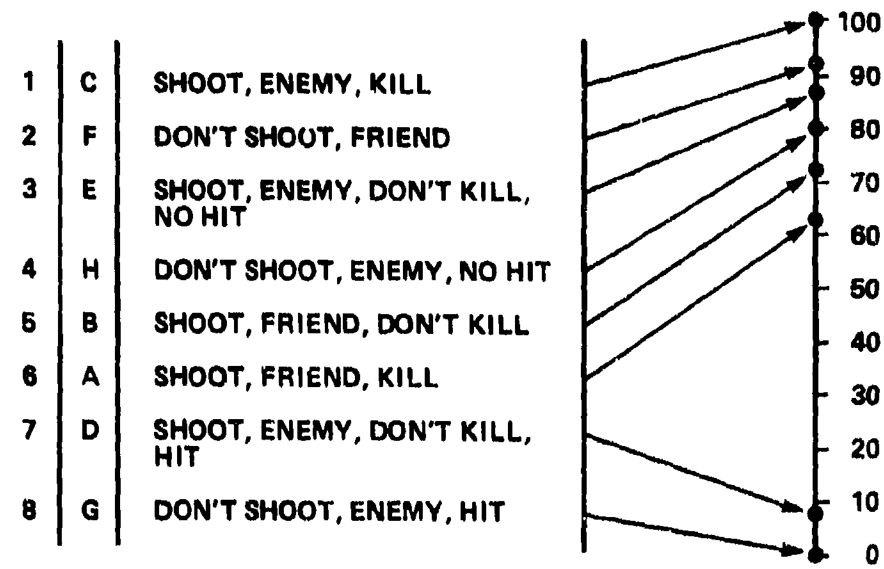


Figure 6-6

Commander: I see--we can incorporate my vagueness about these values by changing the numbers, but not so much that the ordering of the outcomes would be different. I'm reasonably confident about the order.

Consultant: That's right. Now our next job is to figure out some way to compare these political values, which are on a scale from zero to 100, with the MEN and EQUIPMENT values, which are expressed on a dollar scale. In other words, we need to get a total length for the political scale, and this length represents the importance of the political dimension. How important are political considerations, relative to either men or equipment?

Commander: Well, as you know, in this conflict, there has been a lot of public opinion, the people get stirred up, and their opinion has a lot more weight than it would have had back in World War II. I would say that in this situation, I consider political value to be about twice as important as men.

Consultant: So, we will now stretch the political value scale out to two times the length of the scale for men. Since the total length of the scale for men is about $9 million, the total length of the political scale must be 18, and we can call these 18 million "equivalent dollars." This makes the political scale between three and four times as long as the equipment scale, which is, recall, $5 million long. Does that seem about right?

Commander: Yes, political value is considerably more important than the value of equipment.

Consultant: Now let's determine the equivalent dollar values of the outcomes on the political scale. Outcome F is neutral, so we will place it at zero. Now, to keep all the relative distances about the same, as in Figure 6-6, I need to make the distance from F to A one-third the distance from F to G. A little over half the way from F to A, I'll put B. Then, I'll go up from F about one-third of the distance from F to B to locate C. Let's put C at 1.0. If I put G at -17, then A will be a little less than -5, and B will be about -3. You said D was next to G, where would you put it?

Commander: Put it at -16.

Consultant: That completes the scale (Figure 6-7). Now, does it agree with your intuition? Check both the location of the outcomes on the scale, and also check the scale against the other scales.

Commander: I think this is a pretty good picture of the situation. It looks as if the political value of shooting down the enemy is the same as the equipment value.

Consultant: That's right. Although the political value of C is small compared to other outcomes, the political attribute is so important that it makes up for it. Notice that we have done an interesting thing. We have taken a completely subjective scale, political value, and related it to a completely objective scale, dollars. Thus, we can talk about political value in terms of millions of equivalent dollars, and when we combine the different attributes to get an overall value for each of the outcomes, we can use units of millions of equivalent dollars for the total values.

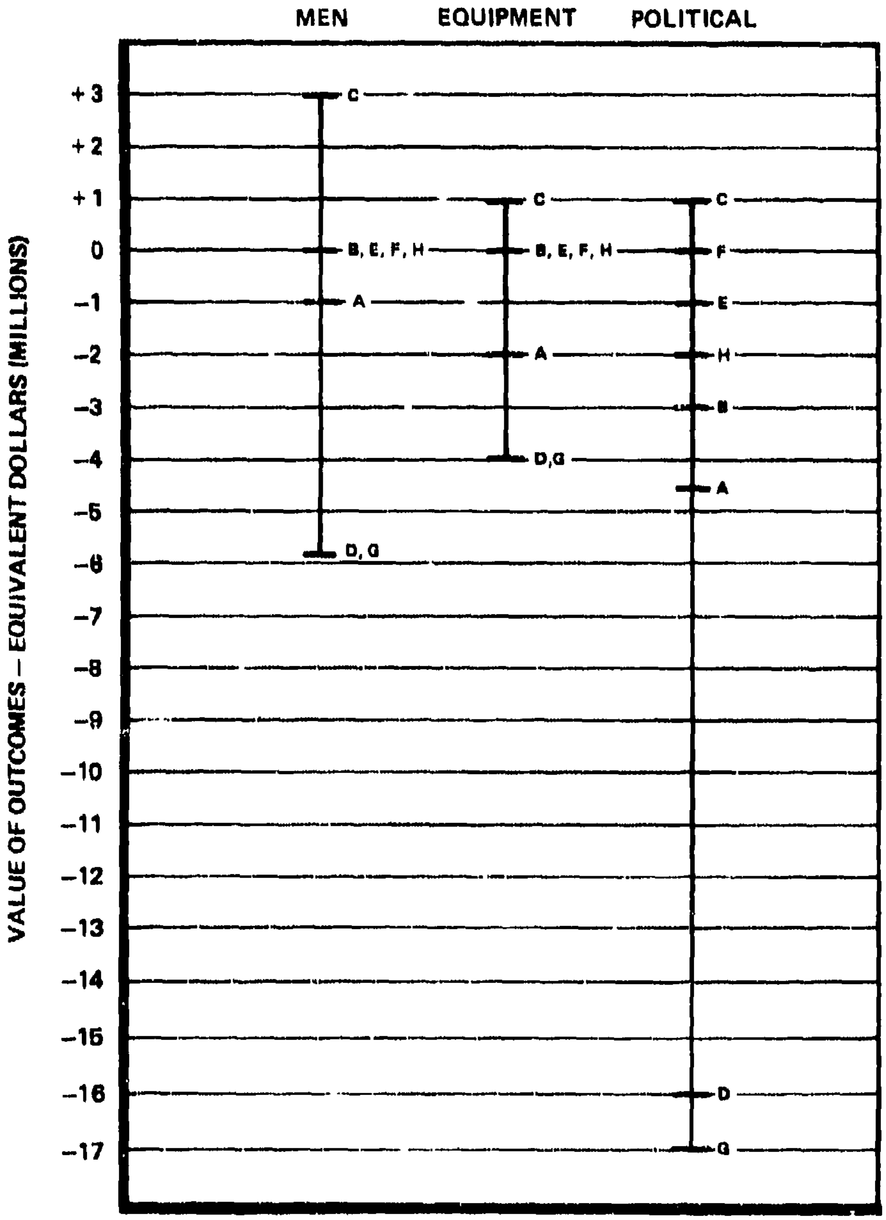


Figure 6-7

(At this point, the consultant and the Force Commander went through the morale attribute in a similar fashion. First, they ordered the outcomes, then they scaled them, and finally, they weighted the entire attribute. As with the other three attributes, outcome F was judged to have no positive or negative morale value, and thus, defined the zero point.)

Consultant: This is the final set of scales (Figure 6-8). The lengths correspond to their importances, and the relative position of each of the outcomes on any attribute indicates its relative value on that attribute. Does everything look O.K.?

Commander: Yes, I think those are the values I would use in making a decision in this situation. At least, I can't see anything which disagrees with my intuitions.

Consultant: Now that we have assessed your values on each attribute, we know what each of the outcomes would be worth to you if it occurred. The value of each outcome is simply the sum of the different values of men, equipment, political value, and morale which go with each one. This means that A is worth about -10 million equivalent dollars, B is worth -3.3 million, C is worth +5.8, and so on, as shown on our decision diagram (Figure 6-9).

Branch Probabilities

Consultant: We still don't know what decision to make, because we don't know which of the possible outcomes will be the one that occurs. In order to compare the two options, SHOOT and NO SHOOT, we need to put probabilities on each of the branches coming from an event fork; that is, we must assess your uncertainty about which of the branches will occur.

Working from the right to the left, the first event fork is the one which determines whether you hit the enemy or friendly aircraft when you shoot at it. How likely is it that the gun crews can hit an aircraft in this situation when they shoot at it? Remember that we have been assuming that a "hit" means that they destroy the aircraft and kill the pilot.

Commander: Well, let's see. They can't be absolutely certain of hitting it, nor is it impossible. The chances are definitely better than 50-50. I'd say that the probability is about 0.70 that the crews will hit the aircraft if they shoot at it.

Consultant: O.K., we'll put a probability of 0.70 on the "kill" branch, and since they either hit the plane or

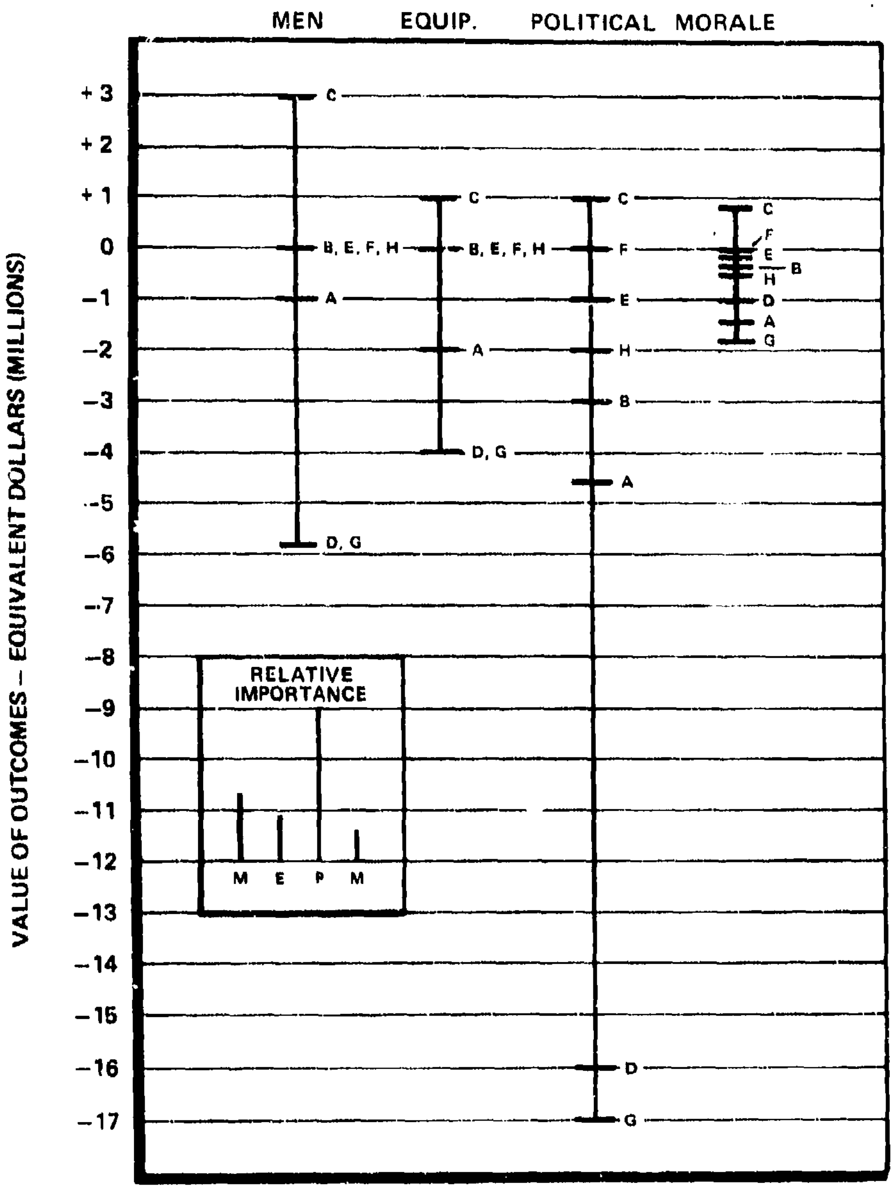


Figure 6-8

they don't, we'll put a 0.30 on the "don't kill" branch. Does it make any difference in the probability of hitting the aircraft whether it is friendly or an enemy?

Commander: I remember you asked me earlier if the chance of a hit is affected by the type of plane, but on

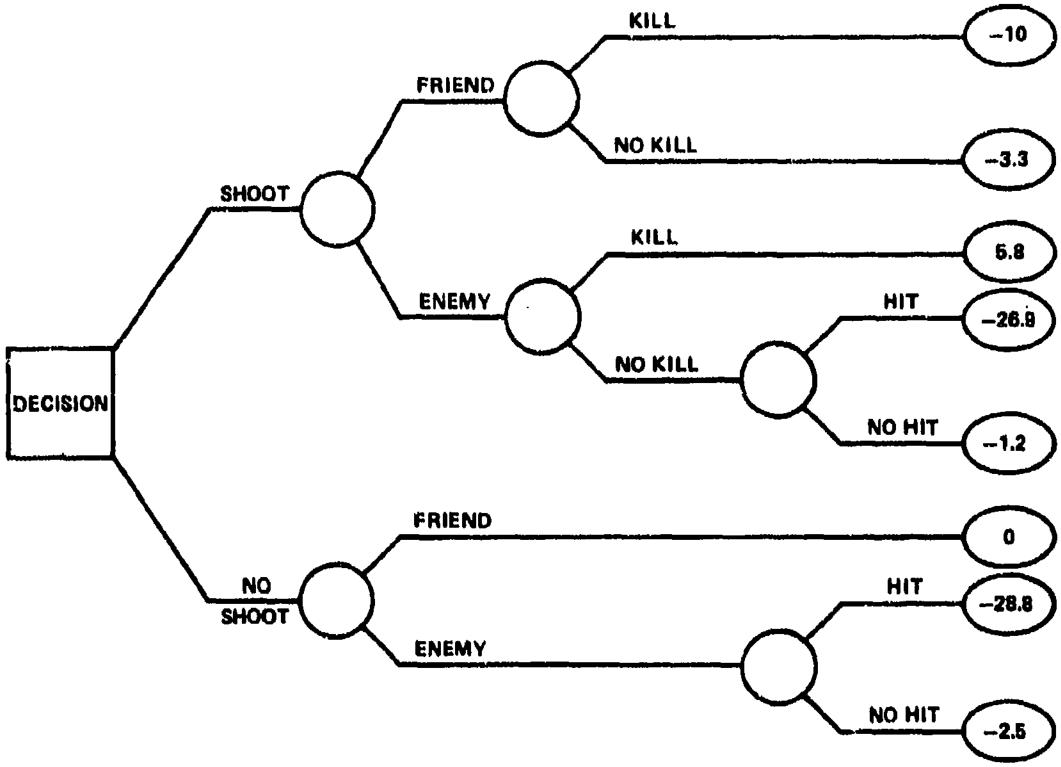


Figure 6-9

reflection, I'd have to say no. We would use a missile, and with this angle of attack, the probability stays the same whether it's a friend or an enemy.

Consultant: Then I'll put 0.70 and 0.30 on both sets of branches. What about the probability that the aircraft is a friend or an enemy?

Commander: Well, the ship was conducting shore bombardment, and about 80% of all this type of bombardment takes place south of Point Bunker. If the HERRINGBONE had been in that area, I would have said that there was only about a 1% chance of an enemy attack, mostly because there has never been an enemy attack before. It is true, however, that the enemy had been flying some risky missions at this time, trying to attack U.S. bombers. About 18% of the shore bombardments take place right around Point Bunker, and the remaining 2% occur significantly farther north, where the HERRINGBONE was. In this area, I would have to give a probability of about 15% of an enemy attack.

Consultant: That's the same as about one chance in six, and leaves five chances in six that the aircraft is friendly. Do you agree?

Commander: Yes, that's about right.

Consultant: Now look at the bottom right fork. What is the chance that the enemy plane will score a hit on the ship when you haven't shot at him?

Commander: I could give you a fairly informed assessment after consulting some tables giving hit probabilities for different types of targets, weapons, and so forth. But offhand, I'd say that for a target like the HERRINGBONE and the type of approach that would have to be taken by the pilot of the MIG-17--well, all things considered, about 35%.

Consultant: Does that probability change if you are shooting at him?

Commander: Oh yes. He's under pressure and will find it more difficult to make a clean run on the ship. He's also more likely to make a mistake. I'd say he has a 20% chance of hitting the ship.

Consultant: So, we put 35-65 on the branches of the lower right fork, but we assign 20-80 to the branches of the HIT - NO HIT fork just above. We now have all the values for all of the possible outcomes, and the branch probabilities for all of the event forks. Our decision diagram (Figure 6-10) now contains enough information to reach the best decision.

Maximization of Expected Value

Consultant: To reach the best decision, we apply this decision rule: Choose that act which has associated with it the highest expected value. You might think of an expected value as a sort of figure of merit, just as the values at the end of the tree indicate the relative worth or merit of the outcomes A through H. The outcomes range in value from -$28.8 million for G to +$5.8 million for C. Every outcome falls on that scale from -$28.8 to +$5.8. Now, by applying the expected value principle, we can put the acts on that same scale, thereby enabling us to compare easily the relative desirability of the acts. To find the expected value

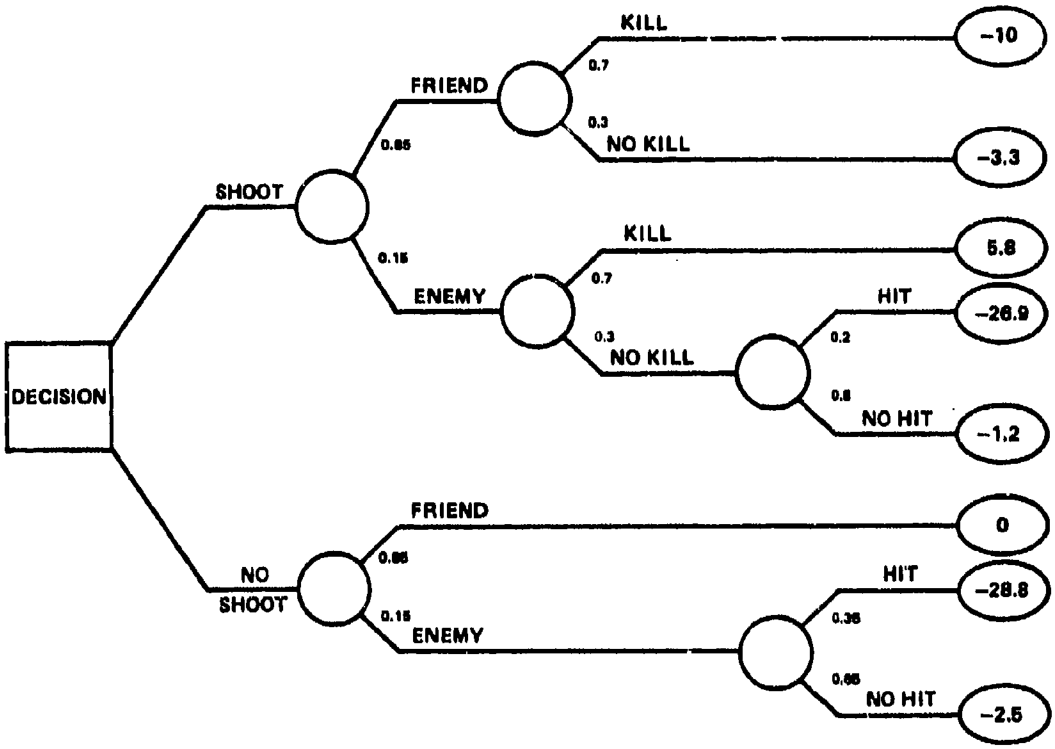


Figure 6-10

of the acts, we start at the right of the tree with the values for the outcomes and work back to the acts.

To see how this works, let's start with the bottom right event fork, HIT or NO HIT following not shooting at the enemy. What is the value of not shooting at an enemy?

Commander: It's either -28.8 or -2.5 depending on whether he successfully bombs the ship.

Consultant: Yes, the value of not shooting an enemy if he hits you is -\$28.8, or <u>if</u> he doesn't hit you, -\$2.5. But what is the value just of <u>not</u> shooting at an enemy?

Commander: Well, since I'm not sure that he'll hit the ship, the value you want must be something less than -\$28.8. And, by the same line of reasoning, since I'm not sure he will miss the ship, the value must be worse than -\$2.5. Something between -\$28.8 and -\$2.5. But how much?

Consultant: That's where the probabilities come in. If you thought the chances of HIT or NO HIT were 50-50, then not shooting at an enemy could be valued as half way between -28.8 and -2.5, or -13.15. Here, the probabilities are 35-65, so we assign to NO SHOOT a value that is 65% of the distance from -28.8 to -2.5. That's -\$11.7. So, on our value scale for the outcomes, which extends from -28.8 to +5.8, we have placed not shooting at an enemy at -11.7.

Commander: Aren't you, in effect, using the probabilities as weighting factors? The more sure you are of a miss, the more heavily you weight that -\$2.5 outcome?

Consultant: Yes, in fact, that's what an expected value is: a weighted average. To calculate an expected value, you multiply each value by its probability and add up over all the branches of the fork. For the bottom fork:

Value	x	Probability				
-28.8	x	.35		=	-10.1	
- 2.5	x	.65		=	- 1.6	
			SUM	=	11.7	= Expected value.

We write that expected value in the circle at the node of the branch (Figure 6-11).

Commander: I see now how we arrive at an expected value for not shooting at an enemy, but how do we find the expected value of the act NO SHOOT?

Consultant: We apply the same approach again. Not shooting at a friend is valued at 0, and we've just found that not shooting at an enemy is valued at -\$11.7. To find the expected value of not shooting, we weight those two values with their respective probabilities, 0.85 and 0.15, and add:

Value	x	Probability			
0	x	0.85		=	0
-11.7	x	0.35		=	-1.8
			SUM	=	-1.8 .

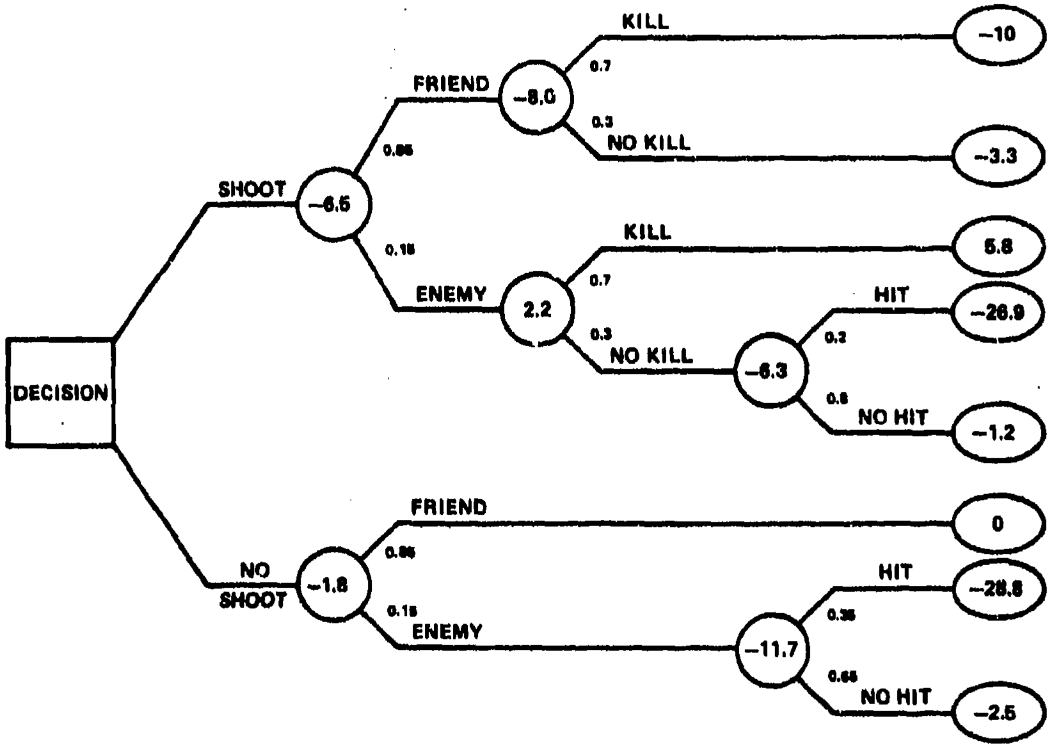


Figure 6-11

The expected value of not shooting is -\$1.8, and we write it in the circle following NO SHOOT. Now we know where the act NO SHOOT stands on the value scale from -28.8 to +5.8.

Commander: I see how you get the expected value, but what does it mean? It doesn't have any reality, does it?

Consultant: Remember that the -\$1.8 is not the value of any real <u>outcome</u>; it is the value of not shooting, taking into account the uncertainty about whether he is a friend or an enemy, and the uncertainty about whether an enemy would score a hit on the ship. The number is an index of the value of not shooting, relative to the best and worst outcomes valued at +5.8 and -28.8.

Commander: I see. Now, I suppose, we apply this same weighting process to get the expected value of shooting.

Consultant: That's right. You might start with the upper right fork. What is the value of shooting at a friend?

Commander: It's the weighted average of -10 and -3.3:

$$\begin{array}{rcrcl} -10 & \times & 0.7 & = & -7.0 \\ -3.3 & \times & 0.3 & = & \underline{-0.99} \\ & & \text{SUM} & = & -7.99\text{, say } -8.0. \end{array}$$

Shooting at a friend is valued at -$8.0, right?

Consultant: Yes. Now what about shooting at an enemy?

Commander: That should be a weighted average of the outcomes from KILL and NO KILL. But the latter involves another event fork, HIT and NO HIT, so I guess we should start there.

Consultant: Yes, the value of shooting at an enemy and missing him gives an outcome of -$26.9 if his bomb hits the ship, or -$1.2 otherwise. Weighting these outcome values with the probabilities 0.2 and 0.8 gives an expected value of $(-26.9) \times (0.2) + (-1.2)(0.8) = -6.3$. Shooting at an enemy and missing him is valued at -$6.3. Why don't you work out the value of shooting at an enemy?

Commander: It's 70% of +$5.8 plus 30% of -$6.3. I get $2.2. That's the first positive expected value we have obtained. Does that mean that there is a positive benefit in shooting at an enemy?

Consultant: Relative to not shooting at a friend, which is valued at 0, yes. But you could have obtained a negative expected value there if the chance of NO KILL had been high, for then the -6.3 would have been weighted more heavily, or if the chance of HIT had been higher, for then the -26.9 would have counted more.

Commander: Are you saying that I can't just choose a course of action on the basis of whether or not it has a positive or negative expected value associated with it?

Consultant: Right. The decision rule is to choose the act with the highest expected value. In some situations, all courses of action will have negative expected values;

then you choose the one that's least bad. Sometimes they may be all positive; in that happy circumstance, you choose the act with the highest positive expected value. Now let's finish the calculations. What is the expected value for shooting?

Commander: Well, shooting at a friend is valued at -$8.0, while shooting at an enemy is worth +$2.2. Weighting those values with the probabilities of FRIEND and of ENEMY gives (-8.0) x (0.85) + (+2.2) x (0.15) = -6.5. So, shooting is valued at -$6.5.

Consultant: Right. Now, if we show on the decision diagram (Figure 6-11) all the expected values we have calculated, we are clear to the left fork. What kind of fork is it?

Commander: It's an act fork.

Consultant: Which means that the way things go is under your control. You get to choose which branch to take. Which one will you take?

Commander: One branch is valued at -6.5, and the other one at -1.8. It looks like I lose, no matter which one I choose.

Consultant: Well, that's typical of this kind of tactical defense situation in which a relatively inexpensive enemy airplane can do a lot of damage to a relatively expensive ship. You're in a tough situation: if you shoot and it's a friend, then you've made a mistake. On the other hand, if you don't shoot, and it turns out to be an enemy, you've made another kind of mistake. Still, you must choose.

Commander: Since the NO SHOOT branch is -1.8, it's better than the SHOOT branch, which is -6.5. I choose to hold my fire.

Consultant: Our decision rule says that it's the best thing to do in this situation. If the captain of the HERRINGBONE had values and probabilities like those in the decision diagram, then he did the correct thing, even though the aircraft turned out to be an enemy and bombed him. Decision theory doesn't eliminate mistakes; it merely helps you avoid those which are avoidable.

Notice one thing: the probability that the aircraft is friendly or hostile is very important to the decision. After all, if you knew for certain that it was an enemy, then the expected values at the act fork would be +2.2, and -11.7, and you would shoot. If, on the other hand, you knew for certain that it was a friend, the expected values would be -8.0 and 0, you wouldn't shoot. Using these values as extremes, we can draw a graph (Figure 6-12) which shows how the expected value of shooting and not shooting changes as the probability that it's an enemy increases from 0.0 to 1.00. Where the two lines cross is where you should change your decision. In this case, you shouldn't shoot if the

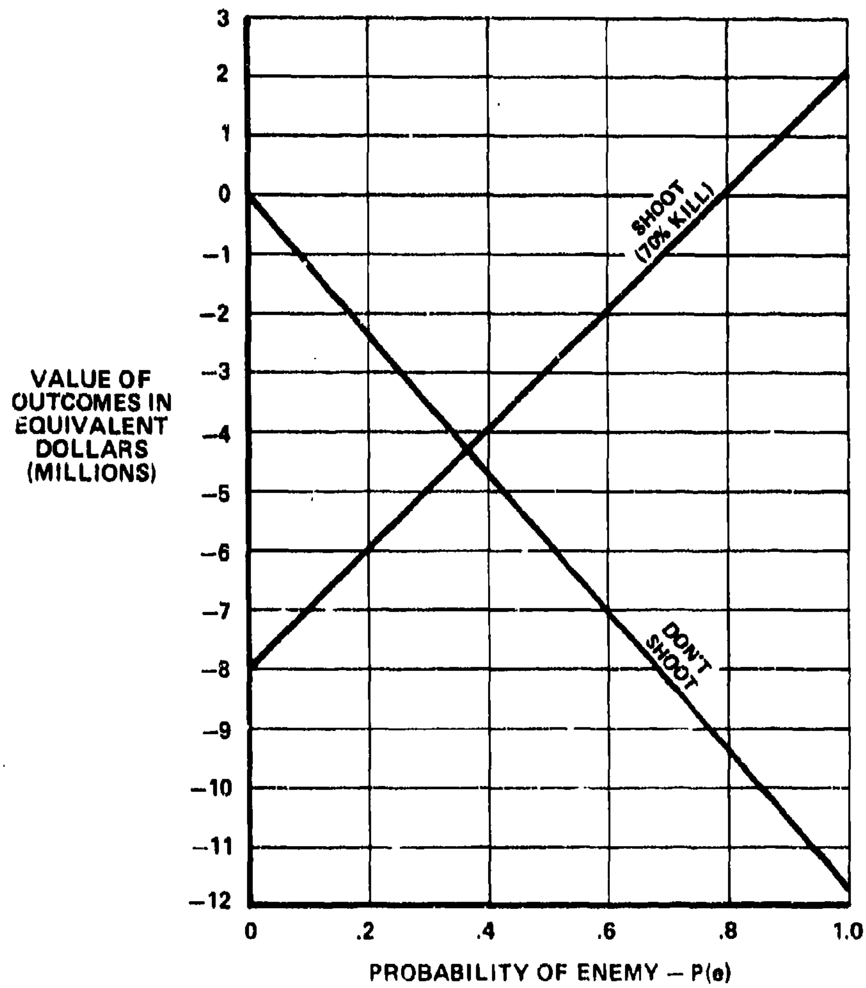


Figure 6-12

probability of enemy is less than 37%. But if the probability is greater than 37%, you should shoot. The decision analysis has enabled us to derive a useful rule of thumb that can be applied quickly if a similar situation arises.

Sensitivity Analysis

Commander: One of the problems I'm concerned with is the spurious precision of the numbers I have been giving. In many cases they are my personal judgments, and I am rather fuzzy about some of the assessments. If I use this type of analysis to make tactical decisions or to make recommendations about information systems or weapons systems, I want to know if my decisions or recommendations would be different if I had given different numbers for some of my fuzzier assessments. What should I do?

Consultant: The very first thing that you should do is a sensitivity analysis. This kind of analysis will tell you how important the fuzz in your assessments is. If a particular number can vary quite a bit without affecting the result, then the fuzz doesn't matter. On the other hand, if a small difference in one of the values makes a large difference in the result, it is important that you take time to make the best possible assessment. Which of the input values that we have used in this problem are you most unsure of?

Commander: Well, one of the numbers which seems to be very important in the analysis is the probability of kill for the current weapons system. I said there was a 70% probability, but I didn't get that figure by counting the number of times that it hit a target under these conditions, I merely gave my best judgment.

Consultant: Well, that may be good enough. How sure are you that the weapons system has a kill probability greater than 50%; that is, is it more likely to hit the aircraft than not hit it?

Commander: Oh, I am very sure of that.

Consultant: On the other hand, we know that the probability of kill cannot be greater than 100%. In order to show the sensitivity of our decision to shoot or not to shoot, let us take our original graph (Figure 6-12) which showed the expected value of each decision as a function of the probability of an enemy and try changing the probability of kill. Of course, if we decide not to shoot, the probability of kill is irrelevant so the NO SHOOT line will stay the same. However, the expected value of shooting certainly depends on the probability of kill. So let's work out a new SHOOT line by assuming 100% probability of kill, then another line by assuming the lowest possible probability of kill, 50%. Take a look at the decision diagram (Figure 6-11). First, we assume that the probability of kill is 100%. If we decide to shoot and the aircraft turns out to be an enemy, what is the expected value?

Commander: Well, with 100% chance of kill, the expected value will be +5.8.

Consultant: Right. What if we shoot and it turns out to be a friend?

Commander: Killing a friend was valued at -10.0.

Consultant: That's right. Let's put those two points on a graph and draw a straight line between them (Figure 6-13). This line represents the expected value for shooting, when you have 100% probability of kill, as a function of a probability that it is an enemy. Now, since you are sure that the probability of kill is at least 50%, let's use that as the other extreme. If there is 50% kill probability, you shoot and it turns out to be a friend, then the expected value must be halfway between the value for 100% probability of kill and the value for zero probability of kill from Figure 6-13. Halfway between -10.0 and -3.3 is -6.7. Again referring to Figure 6-11, if you shoot when you know for sure it is an enemy and you have a 50% kill probability, then the expected value would be halfway between -6.3 and +5.8, or -0.3. Now if we draw a line connecting -6.7 and -0.3 on our graph, it will represent the expected value of shooting, when there is a 50% probability of kill, as a function of the probability of an enemy (Figure 6-13). The important thing to notice on this graph is where each of the SHOOT lines crosses the NO SHOOT line. That is, at what threshold probability of an enemy will you change your decision from DON'T SHOOT to SHOOT. As we have already seen, at 70% probability of kill the change occurs when the probability of an enemy reaches 37%. Increasing the probability of kill to 100% or reducing it to 50% has hardly any effect on the threshold probability--it stays at 37%, give or take 1%.

Commander: So even if the kill probability of my weapons systems varies between hitting the target half the time and hitting it all the time, the effect on my decision is negligible.

Consultant: That's right, and therefore it is not too important that you be able to assess the probability of kill for your current weapon system to a high degree of accuracy, at least for this problem. Now let's try another sensitivity analysis. Which of the value dimensions are you most unsure of?

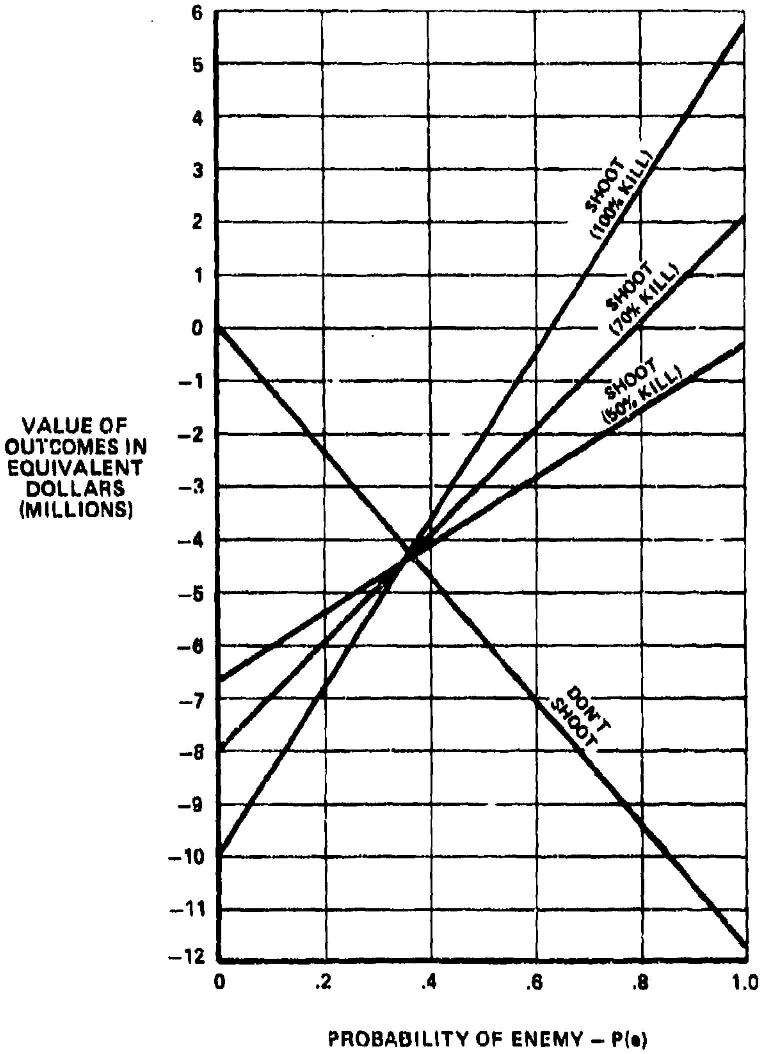


Figure 6-13

Commander: The thing that bothers me most about the value dimensions is the relative importance of the political factor. Until an incident of this kind happens, it is really difficult for me to say how important political factors should be in my decision.

Consultant: In the original diagram you gave the political value attribute twice the importance of the MEN attribute. What would be the effect if you assessed political value to be half the importance of the value of men? In order to find out, we go back to the original value assessments (Figure 6-8) and divide all the numbers on the political value scale by 4. Then we add up all the values as before and put them on the decision diagram. Assuming the probability of kill stays at 70%, we can now generate new SHOOT/NO SHOOT lines corresponding to the changed values.[1] When I do this, I come up with the two dotted lines shown on Figure 6-14. The solid lines are the results of the original analysis. As you can see, the effect of reducing the political value is to raise the expected value of any of the outcomes. This just means that there is less risk. However, notice what happens to the intersection of the two lines, the point at which one decides to start shooting. The threshold probability has changed from .37 to .34, a small difference. The reason for this is that the outcomes are located at approximately the same relative positions on each of the attribute scales. Outcome C is always at the top, followed by B, E, F and H all together or relatively close, then A, with D and G at the bottom. When this occurs, the scales are highly correlated. Therefore, changing the relative importance can change the expected values of all the outcomes but will have in general a very small effect on the decision. This insensitivity to value implies two things. First of all, it is not too critical that you be able to assess accurately the relative importance of the political dimension; and second, changing the importance weight for the political dimension (for example, if this situation were to occur in a different political climate) will not have much of an effect on the decision.

Impact of Information

Commander: Since the probability that the aircraft is a friend or an enemy is so important, is there a way to assess the impact additional information might have on the decision? In a situation like the one we have analyzed, there might be information available which, if utilized, could affect the decision.

Consultant: Yes, there is a way, and you are quite right in asking the question. Information which would change the probabilities, making you more certain, should be worth a great deal to you. If information, no matter how much or how precise, won't change your decision, then it is

---

[1]The reader is encouraged at this point to put the values obtained into the decision diagram, fold back the diagram and construct the expected value lines for the SHOOT/NO SHOOT decision as before.

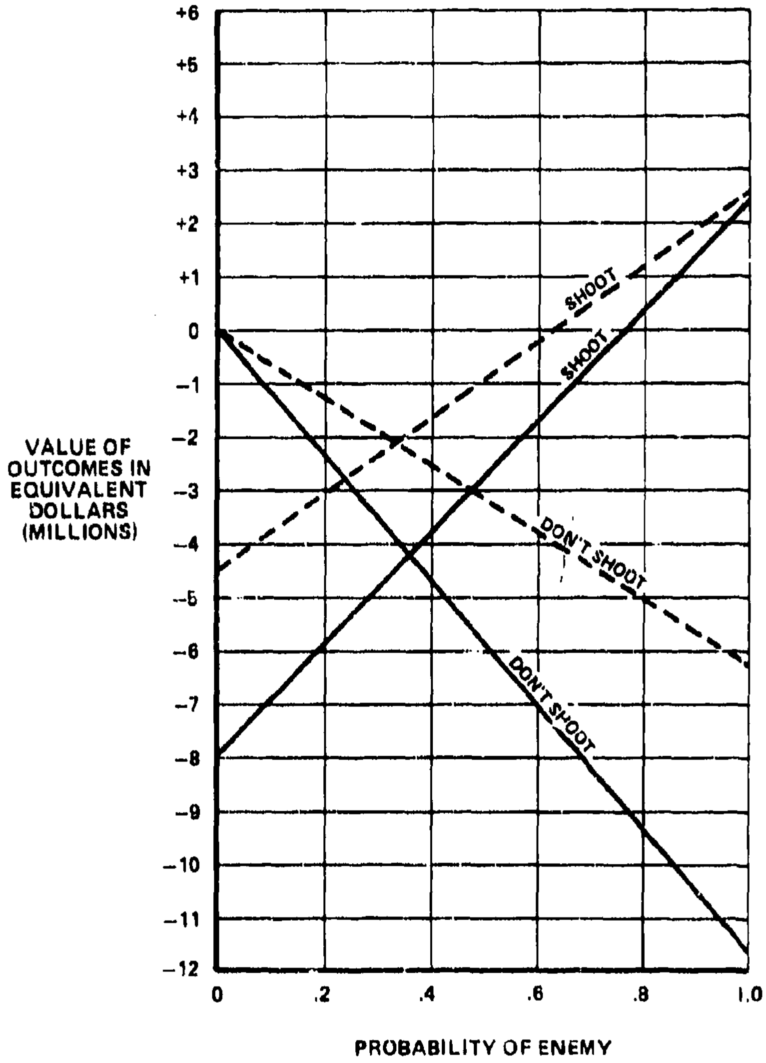


Figure 6-14

of no value to you. Let's discuss first how new information might change your prior opinion (your opinion prior to obtaining the information) about the probability of friend or enemy.

Bayes' Theorem: The Revision of Opinion

Consultant: The best known method for determining the impact of information on your opinion is called Bayes' Theorem:

$$\frac{P(H_1|D)}{P(H_2|D)} = \frac{P(D|H_1)}{P(D|H_2)} \times \frac{P(H_1)}{P(H_2)}$$

where H_1 means that hypothesis 1 is true, H_2 means that hypothesis 2 is true, and D represents the data you have observed. The theorem says that the odds in favor of one hypothesis over the other, after you observe the data, are equal to the odds in favor of the hypothesis before you observed the data, multiplied by the likelihood ratio. The likelihood ratio indicates how much more likely it would be to observe the data if one hypothesis were true, than if the other were true. So, if your odds in favor of one hypothesis were 3:1, that is, it was three times as likely that hypothesis 1 was true than hypothesis 2, and you saw some datum which was twice as likely to have been observed if hypothesis 2 were true than if hypothesis 1 were true, then your odds in favor of hypothesis 1 after you observed the datum are:

$$\frac{P(H_1|D)}{P(H_2|D)} = \frac{3}{1} \times \frac{1}{2} = \frac{3}{2}.$$

The original odds of 3 to 1 correspond to probabilities of 0.75 and 0.25 (the probabilities are in a ratio of 3 to 1), and the revised odds of 3 to 2 can be expressed as probabilities of 0.60 and 0.40 (they are in a ratio of 3 to 2) Thus, your probability in favor of hypothesis 1 has decreased from 0.75 to 0.60 because you have observed some datum which favors the other hypothesis--though not so much as to reverse your original belief that H_1 is more likely to be true than H_2.

Now, let's hypothesize possible information inputs to the captain of the USS HERRINGBONE, the probabilities associated with them, and apply Bayes' Theorem to see how they might affect our decision.

What is the first information input that comes to mind in this situation?

Commander: There might have been a tip-off that MIG's from a nearby airfield were in the air.

Consultant: O.K., there are two ways we can evaluate the effect of this information input on your uncertainty about whether the aircraft is a friend or an enemy. One approach is for you to revise your original assessment of

0.85-0.15 to take account of the new information, the tip-off. If we do this, there isn't any need to apply Bayes' Theorem, for you are assessing the impact of the information intuitively. Another approach is to assess likelihood ratios, and then apply Bayes' Theorem to combine the likelihood ratios with your original assessment to give the desired revised assessments. In the first approach, I will ask you to tell me how much more likely it is that the aircraft is a friend or an enemy assuming that we've received a tip-off. In the second approach, I will ask you how much more likely it is to receive a tip-off if the aircraft is a friend than if it is an enemy. Sometimes one way is easier than the other to think about in the situation. Which is easier for you?

Commander: Well, it's easier for me to imagine that it's a friend or an enemy and then assess the likelihood of receiving a tip-off. If it's a friend, I would say that you would get a tip-off only about 5% of the time. In the case of an enemy, the percentage goes up to about 20%.

Consultant: Well, assuming you received a tip-off, we can calculate the effect of this information by using the likelihoods in Bayes' Theorem:

$$\frac{P(E|T)}{P(F|T)} = \frac{P(T|E)}{P(T|F)} \times \frac{P(E)}{P(F)} = \frac{.20}{.05} \times \frac{.15}{.85} = \frac{.03}{.0425} = \frac{.41}{.59} .$$

Receiving a tip-off that MIG's were in the air changed your opinion from a 15% chance to a 41% chance that it was an enemy aircraft. Now, one nice property of Bayes' Theorem is that it allows you to use the posterior probabilities from the first piece of information as the prior probabilities for the next piece of information. This makes sense, since your opinion after the first information is your opinion before the second. What is the next information input likely in this situation?

Commander: The operator of the passive ECM would ordinarily pick up the enemy search radar next, if it were an enemy and if he had his radar switched on.

Consultant: What are the likelihoods?

Commander: Well, neither friendly nor enemy aircraft are very likely to be picked up by the ECM operator. First of all, it is extremely unlikely that any friendly airplanes have radar parameters which are sufficiently like the enemy's to be detected by the ECM operator. It's also extremely unlikely that the enemy would leave their radar on in this situation when they attack. It's kind of hard to say how much more unlikely one is than the other.

Consultant: All right. What if you don't get an ECM report? Is that more likely from a friend or an enemy?

Commander: It's about the same. A friend is just as likely as an enemy to generate no ECM report.

Consultant: O.K. Let's assume you don't get an ECM report. This means your posterior probabilities are the same as your prior probabilities:

$$\frac{P(E|T,\sim ECM)}{P(F|T,\sim ECM)} = \frac{.41}{.59} \times \frac{.50}{.50} = \frac{.205}{.295} = \frac{.41}{.59} \quad .$$

Now, what other information might be available to you?

Commander: The next thing would probably be IFF. (Note: IFF is a device in friendly aircraft, which, when turned on, returns a coded identifying signal in response to an interrogating signal.) Now, ordinarily, you would think that IFF would be a good indicator because all of our airplanes have it, and it's coded to prevent the enemy from imitating it. But, in a situation such as this, if it is a friendly airplane, he may be ditching, and therefore, he may have his IFF gear shot up or it might be turned off. The pilot would have turned it off before his bombing run, and he might have forgotten to turn it back on as he came back to the ship.

Consultant: What are the likelihoods?

Commander: If the airplane is friendly, I'd say there is about a 75% chance of getting an IFF signal back. If it's an enemy, there is about a 99% chance of no signal. The 1% is to take care of the fact that they may have duplicated the code and might mimic the signal.

Consultant: Let's assume that no IFF return signal has been received. Applying Bayes' Theorem again, we get:

$$\frac{P(E|T,\sim ECM,\sim IFF)}{P(F|T,\sim ECM,\sim IFF)} = \frac{.99}{.25} \times \frac{.41}{.59} = \frac{.4059}{.1475} = \frac{.73}{.27} \quad .$$

So, on the basis of getting a tip-off, not picking up enemy radar, and not getting an IFF signal, the probability of an enemy has moved up from 15% to 73%. Is there any additional information which might be available?

Commander: Yes, there might be a visual report. The members of the crew might confirm visually that it's an enemy airplane, or one of ours, or there might be no report at all.

Consultant: So really there are three possibilities: visual report of enemy, visual report of friend, and no report. What are the likelihoods?

Commander: Let me put them down on paper.

The Commander's final assessments looked like this:

	FRIEND	ENEMY
VISUAL (F)	0.80	0.02
NO REPORT	0.19	0.48
VISUAL (E)	0.01	0.50

Consultant: Assuming you received no report, and applying Bayes' Theorem, we have:

$$\frac{P(E|T,\sim ECM,\sim IFF,\sim R)}{P(F|T,\sim ECM,\sim IFF,\sim R} = \frac{.48}{.19} \times \frac{.73}{.27} = \frac{.3504}{.0513} = \frac{.87}{.13} \quad .$$

This means that when you don't get a report at all, you are more sure than you were before that it is an enemy. This is because the likelihood of getting no report is 48% when it's an enemy, and only 19% when it's a friend. Why is this?

Commander: Well, if it's an enemy, the observers on the ship aren't very familiar with what it is, so they would probably wait longer to be sure of identifying it. On the other hand, if it's one of our planes, they've seen so many that they are much more likely to identify it right away.

Consultant: That makes sense. Now, as a result of all these information inputs, the probability of the aircraft being an enemy has risen from the original 15% to 87%; therefore, you should shoot.

As you probably realized, we have picked out only one possible result for each information input for this analysis, while in reality, there are two possibilities for tip-off, ECM, and IFF, and three for the visual report input. To do a complete analysis of all combinations of possibilities and all orders of occurrence of the inputs, computing the posterior odds each time using Bayes' Theorem, would be extremely time consuming. Fortunately, some manipulation of Bayes' Theorem and the use of a "log odds" chart (see Appendix A) simplifies the problem. Recall that Bayes' Theorem states that the posterior odds (the revised odds in light of new evidence) equal the product of the likelihood ratio for that evidence and the prior odds (the odds prior to receiving the new evidence). This form can lead to scale problems if we attempt to display the impact of new data on the prior odds in a linear (uniform) graphical form. If the prior odds equal 1, and the likelihood ratio is 2:1, then the posterior odds equal 2; but with prior odds of 100 (i.e., 100:1) and the same likelihood ratio, the posterior odds equal 200. On an odds scale, the apparent degree of impact of the datum depends on the prior odds.

If we write Bayes' Theorem in logarithmic form, the log of the posterior odds equals the log of the likelihood ratio plus the log of the prior odds:

$$\log \frac{P(H_1|D)}{P(H_2|D)} = \log \frac{P(D|H_1)}{P(D|H_2)} + \log \frac{P(H_1)}{P(H_2)} .$$

For this form of Bayes' Theorem, the impact of a datum is to add on a fixed amount to the prior odds, whatever the value of the prior odds. Plotted on a log-odds chart, the log of the likelihood ratio produces the same increment of change independent of the value of the prior odds. As long as the likelihood ratio for each datum remains the same, the data can be considered in any order without affecting the final result.

Let's plot the results of our analysis on a log-odds chart (Figure 6-15) along with some information we have from Figure 6-12. That showed us that a P(e) of 0.37 was the crossover point between the SHOOT and NO SHOOT decisions.

Conveniently, a log-odds chart shows probability, in this case, P(e), on the left side and the corresponding odds, P(e):P(f), on the right side. We start with our prior probability of 0.15 (Point A) that the aircraft is an enemy, and plot the effect of each input in serial fashion as shown by the solid line. With the tip-off, the probability of enemy rises to 41% (Point B). The lack of an ECM contact doesn't change the probability (Point C) since the odds are 1:1 (log-odds = 0). No IFF return increases the probability to 73% (Point D), and no visual report raises it to 87% (Point E).

Suppose we had not had the tip-off at all. Recall that the likelihood of a tip-off if the aircraft is an enemy is

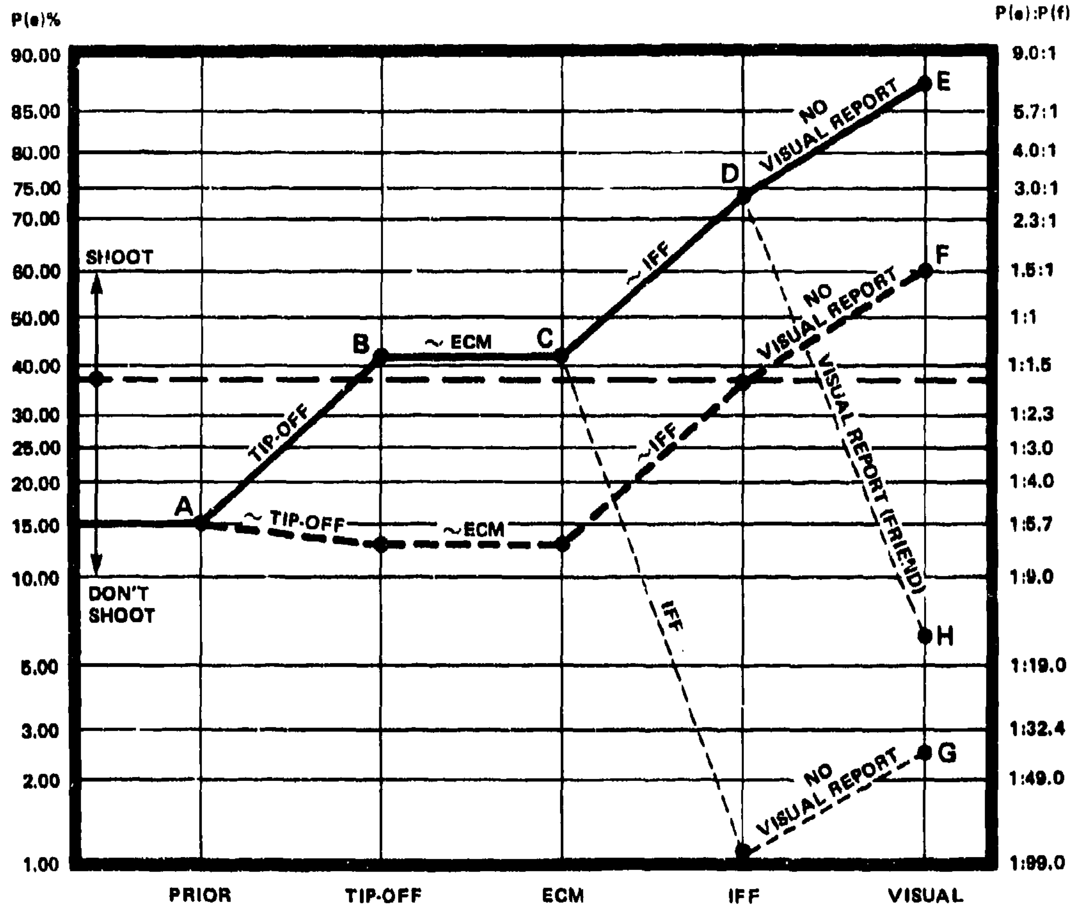


Figure 6-15

0.2; that leaves 0.8 for the likelihood of no tip-off given enemy. Also, you assessed the likelihood of a tip-off given friend as .05, so the likelihood of no tip-off if the aircraft is a friend must be 0.95. Thus, when no tip-off is received, the appropriate likelihood ratio is .80/.95 = .84, or 1:1.2. Starting with our prior probability of 0.15, or odds of 1:5.7, no tip-off changes the prior odds to

$$\frac{1}{5.7} \times \frac{1}{1.2} = \frac{1}{6.8}.$$

These odds give a posterior probability of 0.13. So, receiving no tip-off reduces the probability of enemy from 0.15 to 0.13.

Commander: This is an excellent way to display the analysis because it shows at a glance the impact of the various inputs. It's apparent that once the tip-off occurred, the odds immediately favor the decision to shoot, while without the tip-off, I have to wait for the IFF input before the odds favor shooting. In fact, if I had the tip-off that enemy aircraft were in the air prior to detecting the unknown aircraft on radar, I should have been ready to shoot at once.

Consultant: Yes, if you were sure that the remaining information inputs would remain the same. We didn't analyze all the possibilities, particularly the effect of a friendly IFF return or visual report that it was a friendly aircraft. With this plot, we can determine those impacts quite quickly.

From our earlier discussions, the probability of an IFF return if it's an enemy is 0.01 and 0.75 if it's a friend. This is a likelihood ratio of 1:75 against it being an enemy. Picking the log-odds increment off our scale, we can plot the results as shown on the dotted line from Point C down to the bottom of the chart and then on to point G in Figure 6-15.

Also, the probability of the aircraft being an enemy if the lookout reports it as friendly is 0.02, while there is an 80% probability it is friendly. This is a likelihood ratio of 1:40 against it being an enemy, and the results can be plotted as the dashed line from Point D to Point H, on Figure 6-15. In both cases, the impact has been sufficient to change your decision from SHOOT to NO SHOOT.

If you were at a different location, such as south of Point Bunker, where the prior probability of the aircraft being an enemy is 1%, the whole analysis would simply shift down the chart, starting at P(e) = 0.01, or 1%. This is true as long as the impact of the information inputs does not change as a result of the different location.

Commander: This whole session has been very enlightening, but also very time consuming. In the case of USS HERRINGBONE, there wouldn't have been enough time to do all this. How can decision analysis help in a timely manner in a tactical situation?

Consultant: You're right in what you say. Of course, the purpose of this particular analysis wasn't to help you make a decision in a real-time situation, but was rather to analyze a decision which had already been made. However, these tools should be applicable to a real-time situation if you can identify ahead of time the relevant acts and events, then develop procedures for assigning values and probabilities as the tactical situation changes. What you really need is a computer into which you can place a prestructured decision analysis in such a way that the expected values can be computed in real time as the inputs change. It is fairly clear that the values will change very slowly with time, while the probabilities will be changing quite rapidly in a tactical situation.

Commander: I have a computer available in my Tactical Data System, which is used to display the tactical situation.

(At this point, the Task Force Commander and the Consultant engage in a discussion of how the Tactical Data System works and how it might be adapted for computing and displaying the results of decision analysis in a tactical situation.)

Consultant: Figure 6-16 is a block diagram showing how the Tactical Data System could be used for decision-making in a given tactical situation. On the right are all the inputs that are either already available to you or that you would have to enter into the computer for the specific situation. The left-hand inputs are the dynamic information from the sensor systems available to you to determine the occurrence or non-occurrence of the events relevant to the decision. These inputs would cause the computer to update the probability associated with the decision in accordance with your prior analysis.

Values refer to outcome values for the decision diagram on which the computer, applying branch probabilities, would operate to generate expected values of decisions and the crossover point at which your decision changes. Your Plans Department would, in fact, provide the changes to the outcome values you have previously arrived at (Figure 6-8) by responding to direction from higher authority or from yourself. For example, the Seventh Fleet Commander, under whom you operate might send you a message stating, "aircraft losses must be reduced by 20%." Your weighting on the equipment dimension would change, and so might the outcome values.

The readiness inputs on the right-hand side would permit your Weapons Systems Department to make real-time changes to the probability of kill. Such changes could respond to either materiel or personnel values as they relate to readiness and could change from watch to watch, for instance. The last input on the right, which is the priors, would come from the Intelligence Department and is in fact the 15%/85% (enemy/friendly) assessment that you

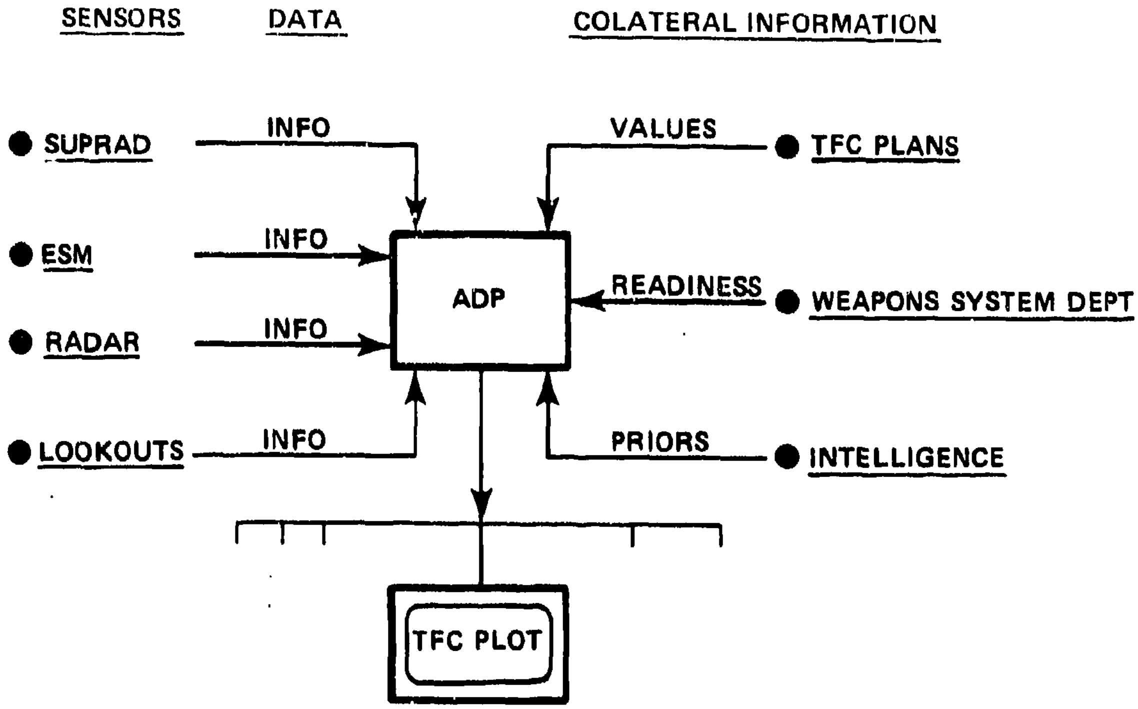


Figure 6-16

started off with as your prior probabilities. These prior probabilities are developed by your intelligence department by utilizing the all-source data base and applying it to the HERRINGBONE's actual tactical environment. Obviously, these inputs on the right-hand side vary, but probably rather infrequently.

Looking at the inputs on the left-hand side of Figure 6-16, we see all the real-time sensor information that will vary quite dynamically with the passage of time.

Using the HERRINGBONE analysis as an example, the detection of an aircraft by radar could generate a display like that shown in Figure 6-17. The target is given an identification number, Bogie 26, which is shown on the situation plot next to the target's present location. The time, 06, is shown on the outer ring. The target range to closest point of approach is 100 nautical miles, and the time at closest point of approach is 16. The decision

$t_1 = 06$

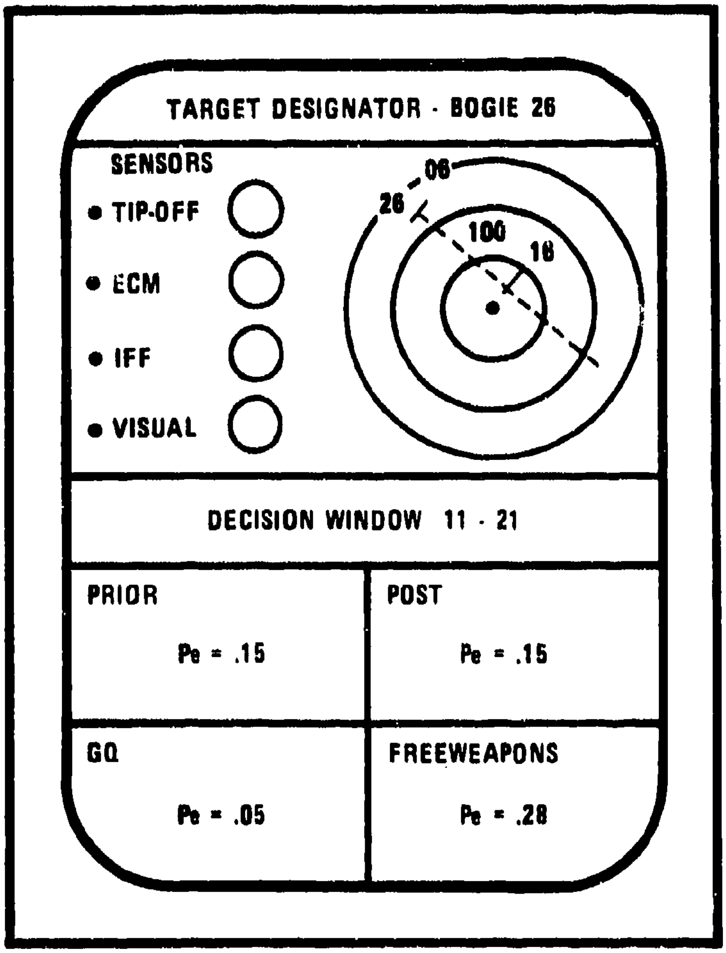


Figure 6-17

window shows the earliest time, 11, and the latest time, 21, that your weapons system can be used to engage the target, if the decision is to shoot. These window times will vary with maneuvering but are dynamically generated by the Tactical Data System. At time 06, there is no information available from the sensors about Bogie 26 except his present position, heading, and speed, from which the computer has generated the plot information. The probability, based on prior intelligence, that it is an enemy has not changed from the prior estimate already stored in the computer. Thus, the Task Force Commander's display says the prior odds are equal to the posterior odds, or a 15% probability of an enemy. Note also that the precomputed probabilities for going to general quarters and for freeing weapons are displayed for the convenience of the Commander and in this particular situation are 5% and 28%, respectively.

At time 10, you receive a tip-off that there is enemy aircraft activity in the area. The display (Figure 6-18) indicates that a tip-off has been received, and that the posterior probability of Bogie 26 being an enemy after accounting for the tip-off information has increased to 0.41.

$t_2 = 10$

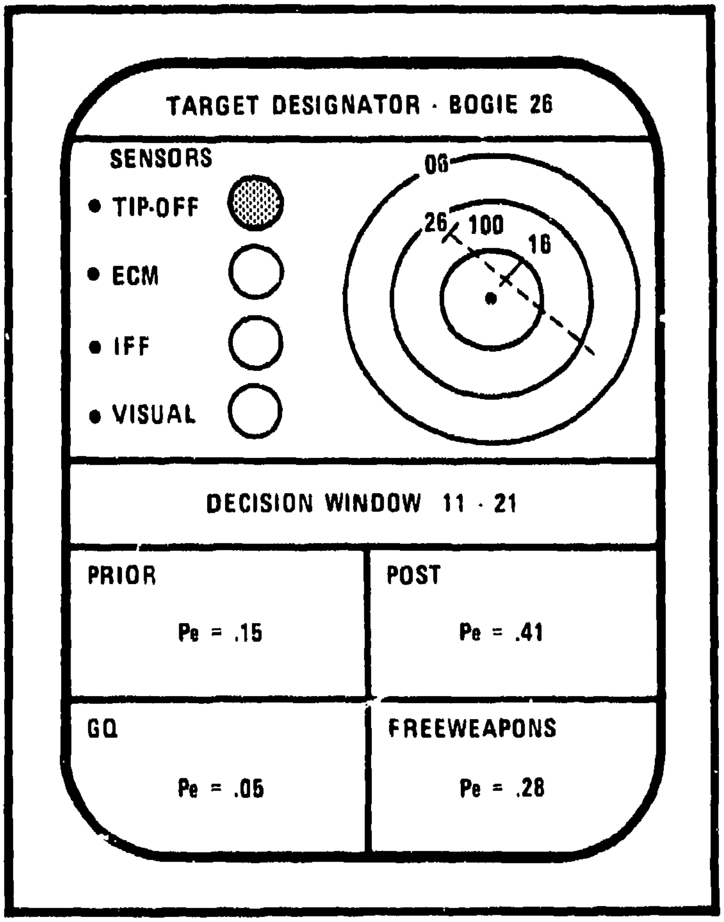


Figure 6-18

If you are uncertain about why such a large change in probability occurred, the computer could be programmed to display the likelihood used in the computation. If, in the present situation, you feel that your prior estimate of likelihood was too high, you could change it, enter the new ratio into the computer, and it would compute and display the revised probability.

Commander: Since I know in advance what my operations and operating areas are, it would be reasonable to perform an analysis like this for each one and store it in the computer. I would have to review each one periodically to be sure that my outcome values haven't changed or that my probabilities, based on previous use of the analyses, shouldn't be adjusted. But that is why you showed me Figure 6-16, to illustrate how these values and probabilities can be dynamically updated by the experts of my staff, from all-source data, and from the sensors organic to my task unit. Now I can see how decision analysis can be tactically useful.

Value of an Information Source

Consultant: Now that we have assessed the impact of the actual information on the probability of friend or enemy, we can use this same technique to calculate the potential value of a particular source of information. For example, suppose that you had the option of buying the IFF information in the present situation. How much should you be willing to pay for it? Should you be willing to buy it at all?

To calculate the value of the information, we will use the expected value of the basic decision as our reference. Any information we get should make the situation better than this, so we can take the expected value of the decision with the information and subtract the expected value for the decision without the information, and this difference is the expected value of the information. Now, since you get the IFF information after you have already received the tip-off, and failed to get the ECM signals, the basic decision to consider is the decision you would make at that point without any IFF. Let's take the basic decision diagram, and put in the probabilities. What are the probabilities?

Commarder: Let's see, we have received the tip-off, and no ECM report, so according to the chart, the probabilities at that point are .59 for friend and .41 for enemy.

Consultant: O.K., if we put those probabilities into the diagram, and compute expected values, we can find out what the expected value for the decision without information a out IFF is worth. (The consultant drew the diagram, put i. the probabilities, and calculated expected values; see Figure 6-19.)

Commander: This says that the expected value for shooting is -3.8, and for not shooting is -4.8.

Consultant: Which one would you choose?

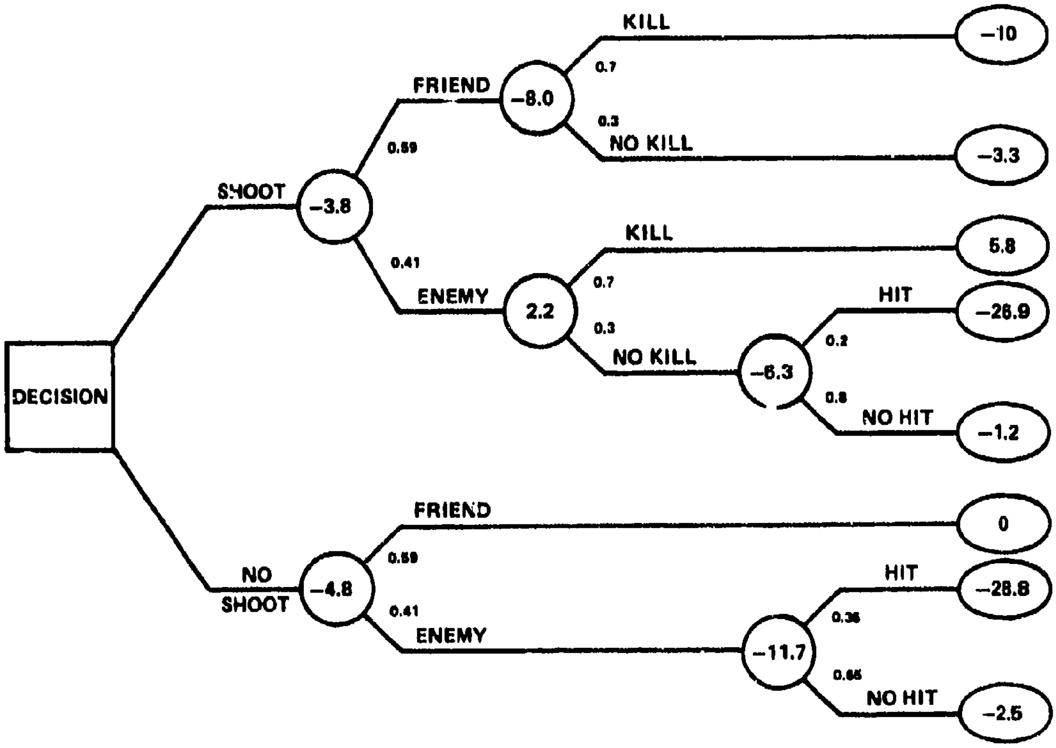


Figure 6-19

Commander: Naturally, I would choose to shoot, but I might not want to do that every time.

Consultant: Well, when would you prefer not to shoot?

Commander: When I was pretty sure it was a friend.

Consultant: Of course, that's right, but I don't see how you can know it's a friend without more information. Whenever you find yourself in the state of uncertainty shown by the probabilities in Figure 6-19, you should shoot. The point about receiving information is that it allows you to separate situations to some extent into two categories. In one category, the probability that it is an enemy is high enough that the expected value of shooting is greater than the expected value of not shooting, so you shoot. In the other category, the probability that it is an enemy is low enough that the expected value of not shooting is greater than the expected value of shooting, so you don't shoot. If there is no circumstance in which information could change your decision, then the information isn't worth anything, at least with respect to the decision. Now, we are going to do just that kind of analysis for one kind of information, in this case, IFF information.

Commander: O.K., how do we set it up?

Consultant: The first thing to realize is that we now have two decisions: the first is whether to buy information or not, and the second is whether to shoot at the airplane or not. I'll draw the diagram like this (Figure 6-20). The three act forks at the right are really condensed versions of the basic decision diagram. They are called the primary decision because they are what determines the information decision. The one at the bottom is the one we just calculated, the basic decision diagram, assuming we have received the tip-off and no ECM report, so we can put the values of -3.8 and -4.8 on the outcomes, and put -3.8 in the box as the expected value of the decision. Now, the other two decision diagrams are just the same, except that they incorporate the new information from IFF and, therefore, have new probabilities. Since we have already considered what happens when there is no IFF return signal, let's take the middle decision diagram first. Now everything is the same up to the point where the probabilities of friend and enemy occur. What are the probabilities?

Commander: According to our chart (Figure 6-15), the probabilities after the ship failed to get an IFF return signal were .27 for friend and .73 for enemy.

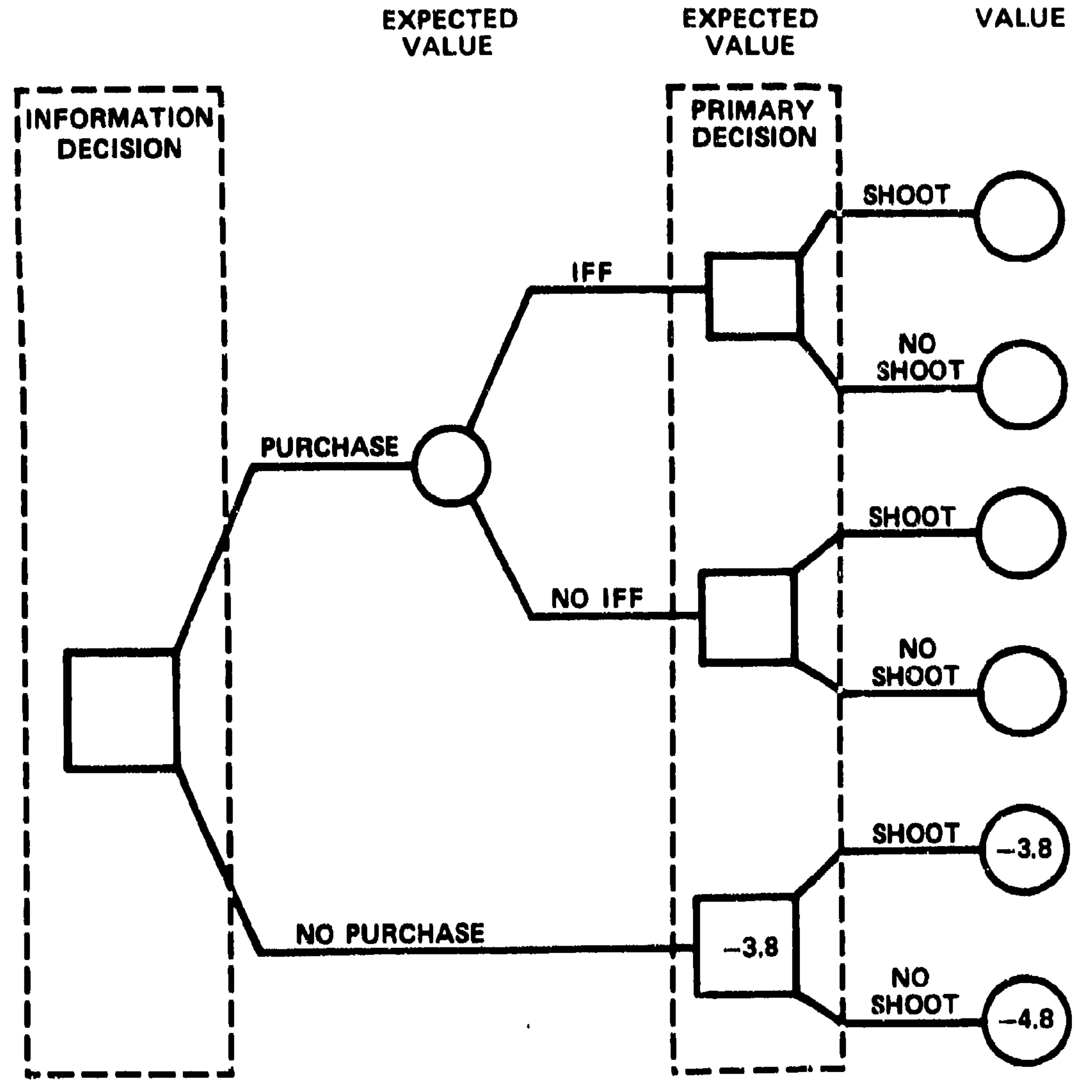


Figure 6-20

Consultant: Right. So, if we incorporate those probabilities into our decision digram (Figure 6-21), we get expected values of -0.55 for shooting and -8.5 for not shooting. Now what if the ship had received an IFF return signal? What is the probability of a friend in this case?

Commander: Since we didn't consider this before, I assume we have to use Bayes' Theorem again.

Consultant: That's right. What are the priors?

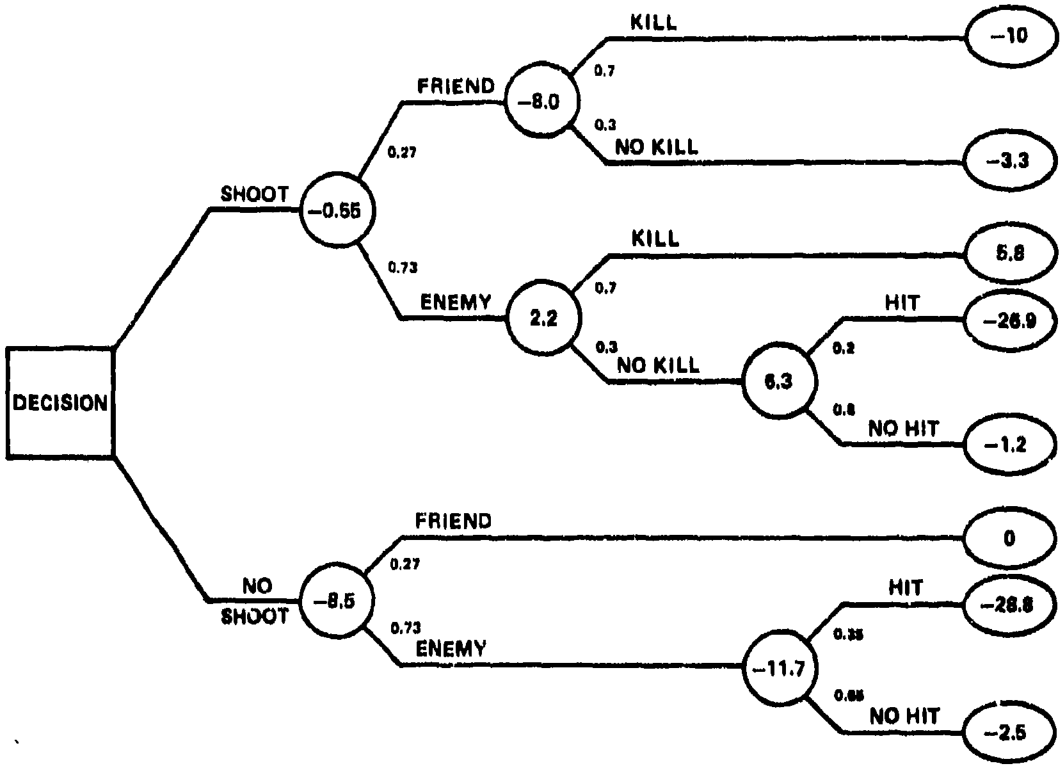


Figure 6-21
DECISION DIAGRAM

Commander: They are .85 for friend and .15 for enemy, right?

Consultant: No, remember that when I say priors, I mean the probabilities that you have prior to the current information. This means that they are probably posterior to some other information.

Commander: Since we have already received a tip-off and there was no ECM reception, the probabilities at this point are .59 for friend and .41 for enemy.

Consultant: That's right. Now, what are the likelihoods? What is the probability of getting an IFF return signal if the airplane is a friend?

Commander: I said before that it was .75, and that's about right.

Consultant: And if it's an enemy?

Commander: Then it's only 1%.

Consultant: O.K., then applying Bayes' Theorem, we have:

$$\frac{P(E|T,\sim ECM,IFF)}{P(F|T,\sim ECM,IFF)} = \frac{.41}{.59} \times \frac{.01}{.75} = \frac{.0041}{.4425} = \frac{.01}{.99} .$$

We can put these probabilities into the basic decision diagram as we did before (Figure 6-22), and the expected values are -7.9 for shooting, and -0.12 for not shooting. This makes sense, since we are very sure that it's a friend if we get an IFF return signal.

Notice that these last two decision diagrams make up the two categories, the SHOOT and NO SHOOT categories, so we know already that the information is worth something. The only thing left is to determine how much it is worth. If we put these last expected values into our information decision diagram, it looks like this (Figure 6-23).

Commander: It looks to me as though we could decide right now to purchase the IFF equipment. We don't need to assign probabilities to the IFF and NO IFF branches because the expected values for those branches, -0.12 and -0.55, are each smaller than the expected value for not purchasing the IFF, -3.8. No matter what probabilities we assign, the expected value of purchasing the IFF will be smaller than -3.8.

Consultant: That's true, but remember that we haven't included the cost of the IFF in our analysis. I doubt that you would purchase the equipment if it cost $10 million; the information wouldn't be worth the cost. Your observation is correct only if the cost of the IFF is less than the expected value of the information it provides. We have to complete the analysis so we can determine the expected value of the information. What do we need to know in order to finish the diagram?

Commander: There is an event fork here, and the two branches say IFF and no IFF.

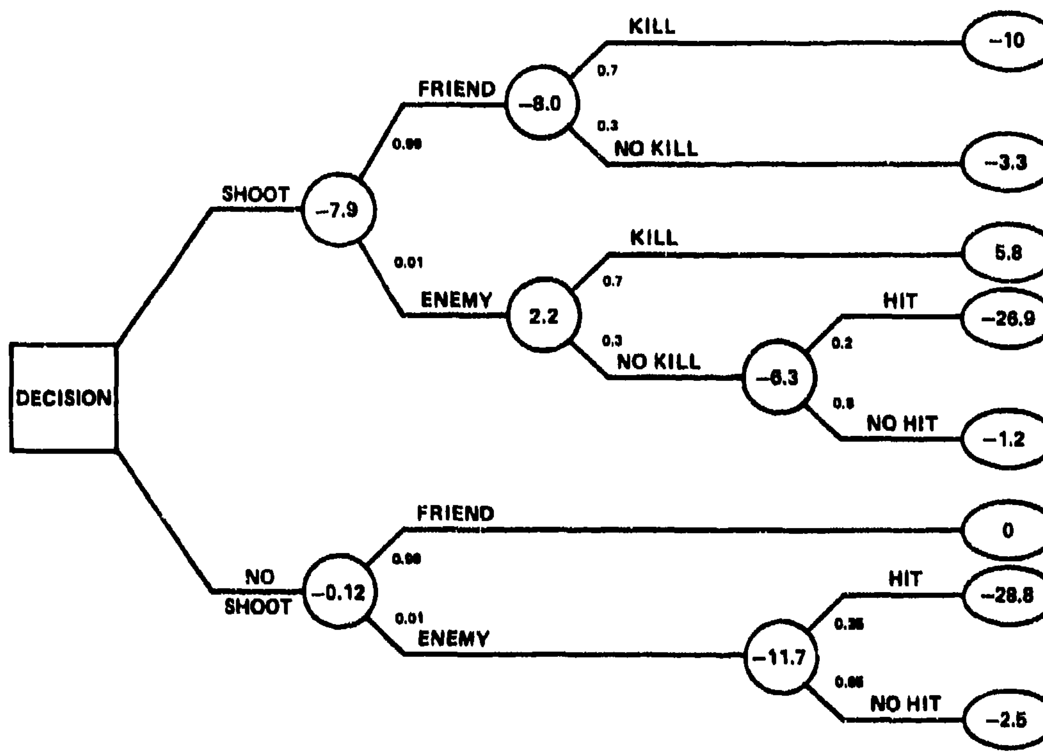


Figure 6-22

Consultant: Yes, there is some probability that you will get an IFF return signal. As you said earlier, the probability of getting an IFF return is high from a friend and low from an enemy. But we need to find the probability of getting an IFF return without having to specify whether the airplane is, in fact, a friend or an enemy. To do this,

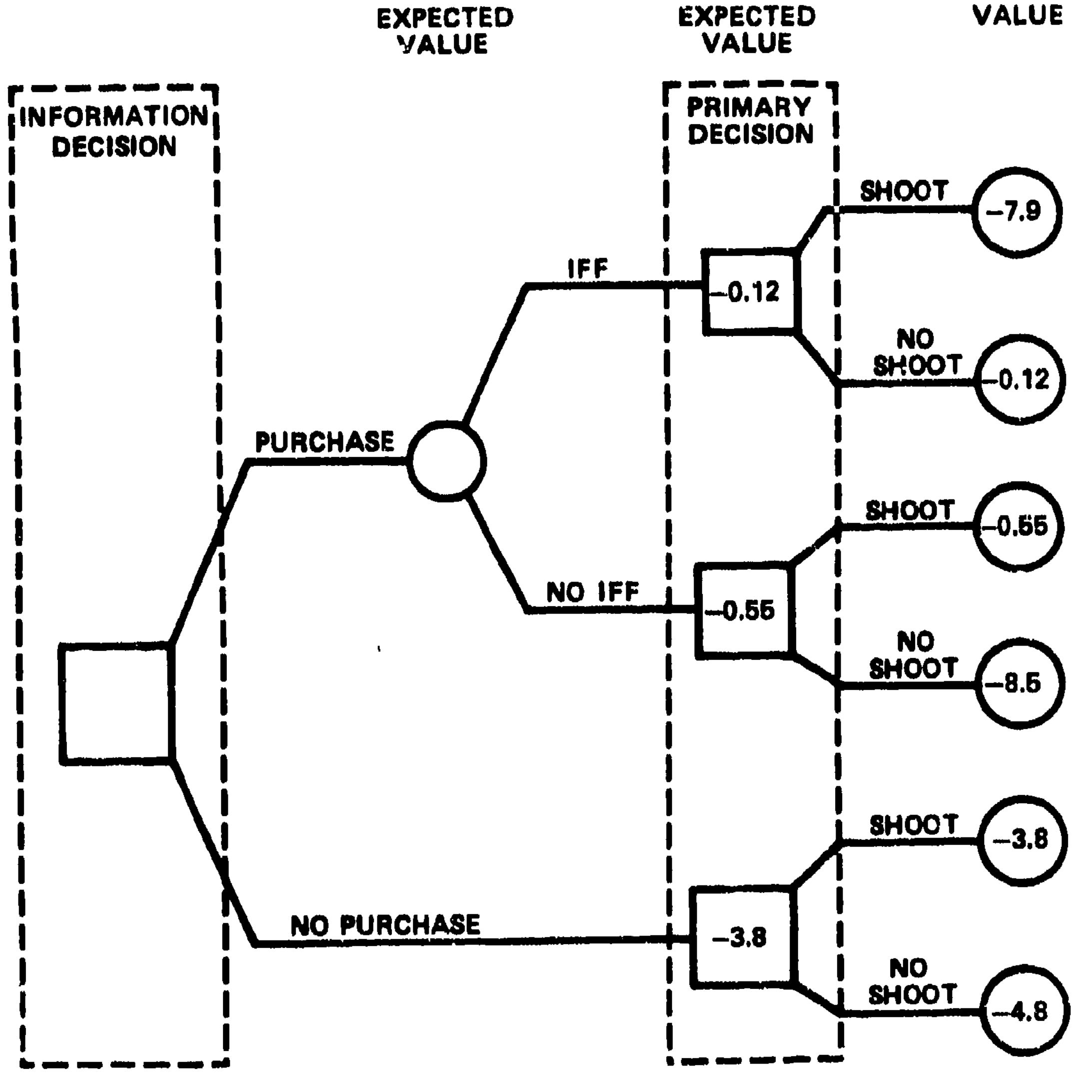


Figure 6-23

let's take a small detour, and draw a probability diagram (Figure 6-24) to represent the way you think about these uncertainties.

First, the airplane is either a friend or an enemy. If it's a friend, then there is either a return signal or there

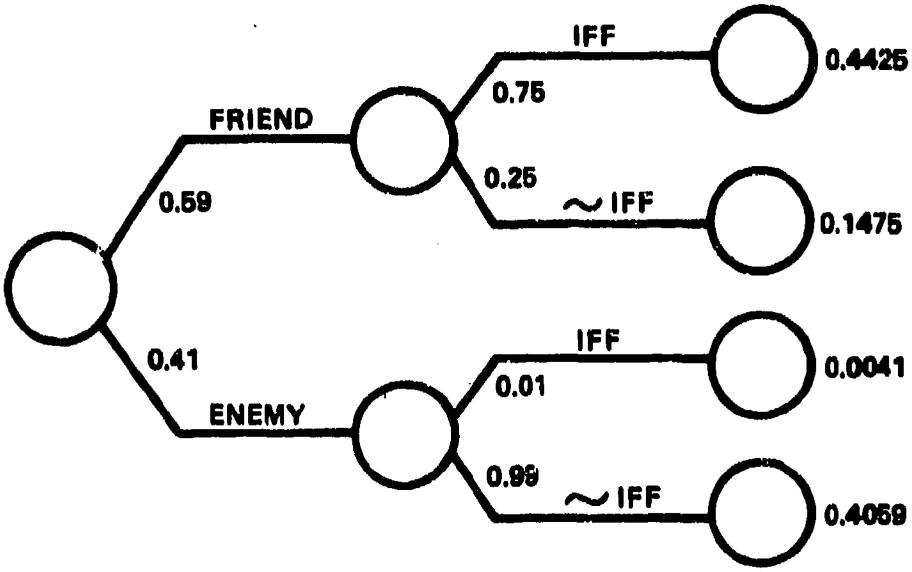


Figure 6-24

is not. Similarly, if it is an enemy, there is either a signal or not. Now, you have previously assessed the probabilities in this diagram relating to the IFF return, and we have used Bayes' Theorem to determine the probabilities on the branches of the left fork after having received a tip-off and no ECM, right?

Commander: Yes, the probability of a friend is 0.59, and the probability of enemy is 0.41. If it's a friend, the probability of IFF is 0.75, and the probability of no IFF is 0.25. If it's an enemy, the probability of IFF is .01, and the probability of no IFF is 0.99.

Consultant: O.K., the probability of IFF is just the probability of IFF if it's a friend times the probability that it is a friend, plus the probability of IFF if it's an enemy times the probability that it's an enemy.

$$
\begin{aligned}
P(IFF) &= P(IFF|F)\ P(F) + P(IFF|E)\ P(E) \\
&= (0.75)(0.59) + (0.01)(0.41) \\
&= 0.4425 + 0.0041 \\
&= 0.45
\end{aligned}
$$

(Note that we are applying both the multiplication and addition rules for probability as explained in Chapter 4.)

The probability of no IFF is (.25) x (.59) + (.99) x (.41) = .55. Now we can put these probabilities into our information diagram and calculate the expected value for buying IFF information (Figure 6-25). It comes out to be -0.4. So should you buy IFF information or not?

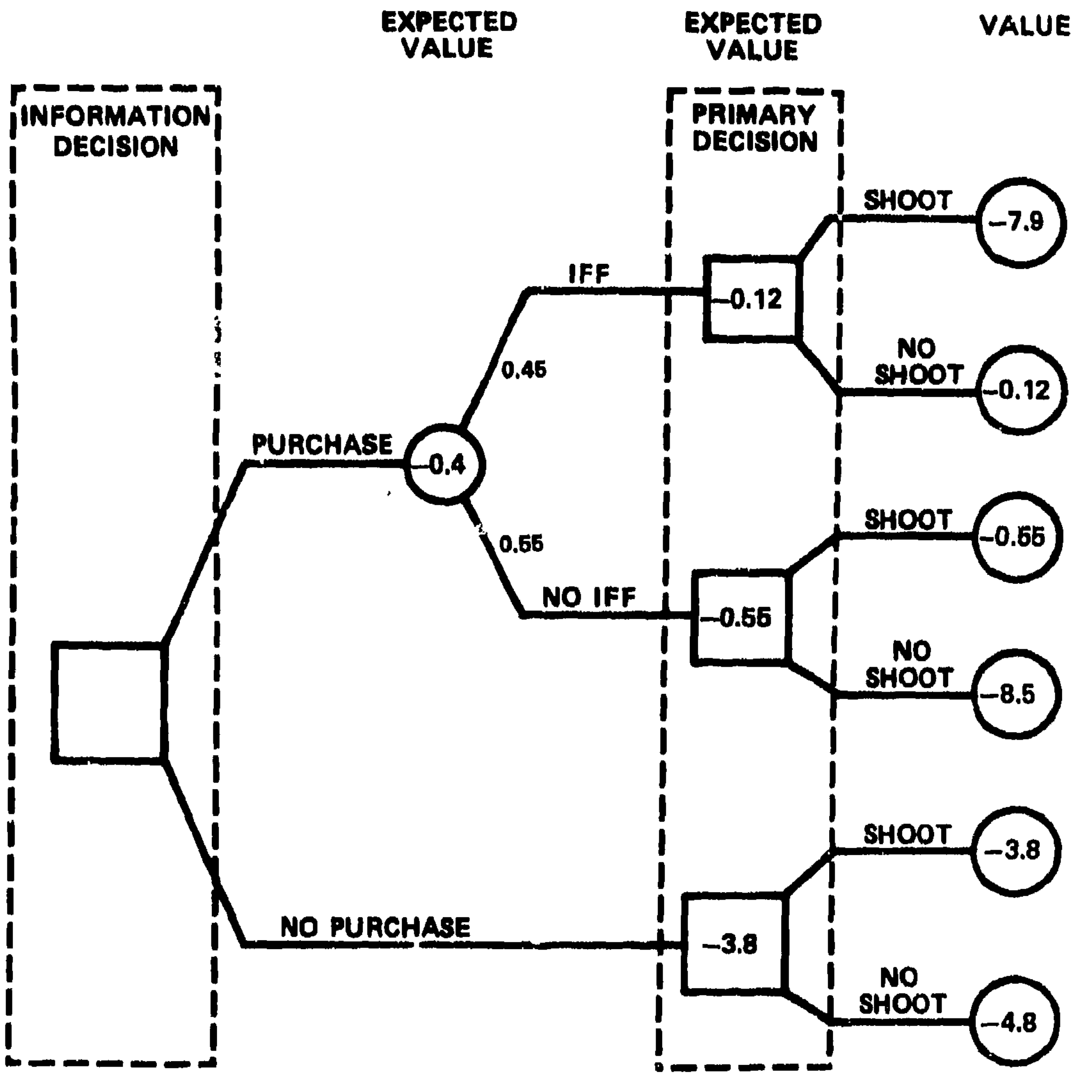


Figure 6-25

Commander: As you said before, that depends on how much the IFF costs.

Consultant: What if I offered to sell you the information at a cost of 3.3 million dollars. Should you buy it?

Commander: Let me think. If the information cost me 3.3 million, and the expected value of the decision with the information is worth -0.4 million, then the total expected cost of making the decision with the information is 0.4 + 3.3 = 3.7 million, right?

Consultant: That's right.

Commander: Then I would buy the information, because it still has a lower cost than the 3.8 million expected cost for deciding without the information.

Consultant: You have the whole concept of the value of information in a nutshell. What is the value of the information?

Commander: It must be 3.4 million dollars, because I would buy the information if it cost less than the amount, and I wouldn't buy it if it cost more.

Consultant: That's exactly right.

Commander: Does this mean that I should be willing to pay 3.4 million dollars to have IFF equipment installed on my ship?

Consultant: Not quite. This is only the value of IFF in this situation. In order to evaluate IFF in general, you would have to consider other scenarios, especially ones which are more likely to occur than this one. In addition, you would have to consider other uses of IFF, such as controlling traffic among your own aircraft, identification, and so forth.

Value of Information vs. Value of Weapons Systems

Consultant: Thus far, we have talked about the impact of information on the primary decision to shoot or not to shoot. We have demonstrated that information changes the expected value of the decision by changing the probability that the incoming unidentified aircraft is a friend or an enemy. Notice, however, that the probability that the incoming aircraft is a friend or an enemy is not the only probability which can affect the expected value of the primary decision. The likelihood that you will hit the incoming aircraft, given that you decide to shoot at it, will also cause the expected value to vary. To show this, let's fix the probability of enemy at 87% and the probability of a friend at 13% and vary the kill probability of the weapon system. In Figure 6-26, the kill probability is set at zero, the probability of not killing is set at 100%. Notice here that the expected value of shooting is a -5.9, the expected value of not shooting is a minus 10.2. Those choices are pretty unappealing, but the preferred decision is to shoot.

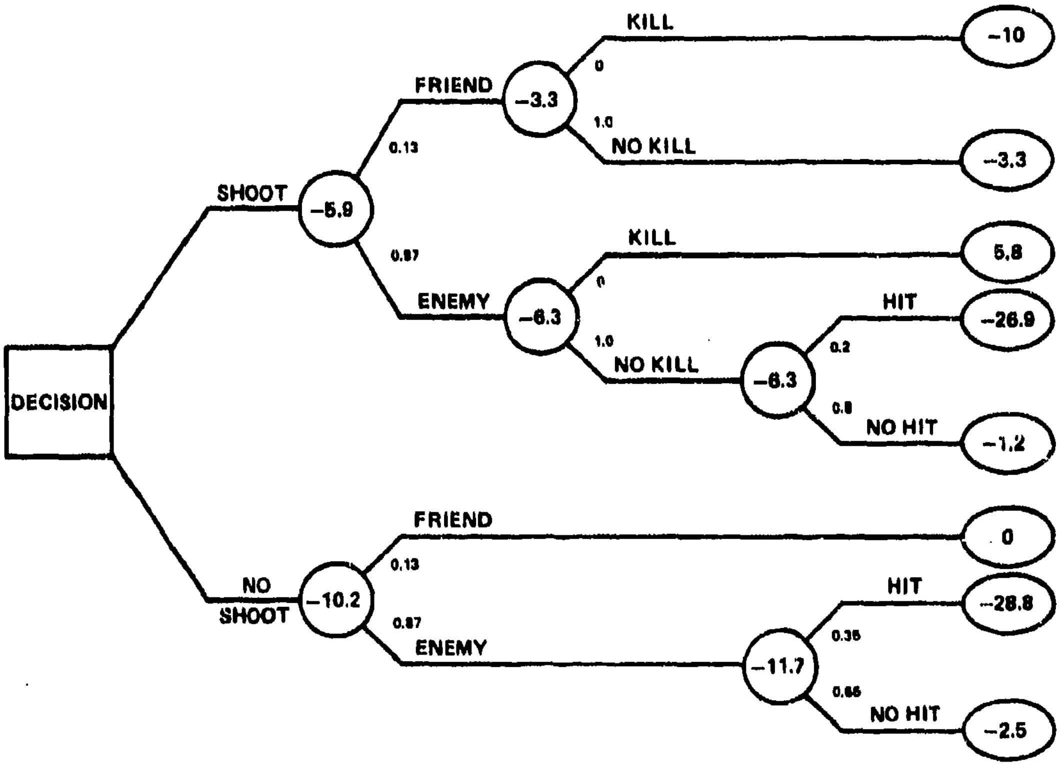


Figure 6-26

On the other hand, if the kill probability were moved to the other extreme, so that you could be 100% certain of hitting an airplane you shoot at, then, as Figure 6-27 shows, the expected value of shooting goes to +3.7, the expected value for not shooting remaining at -10.2. Thus, the value of increasing the kill probability from zero to 100% is the difference between the expected values of the decisions to shoot, or 3.7 -(-5.9) = 9.6 million dollars.

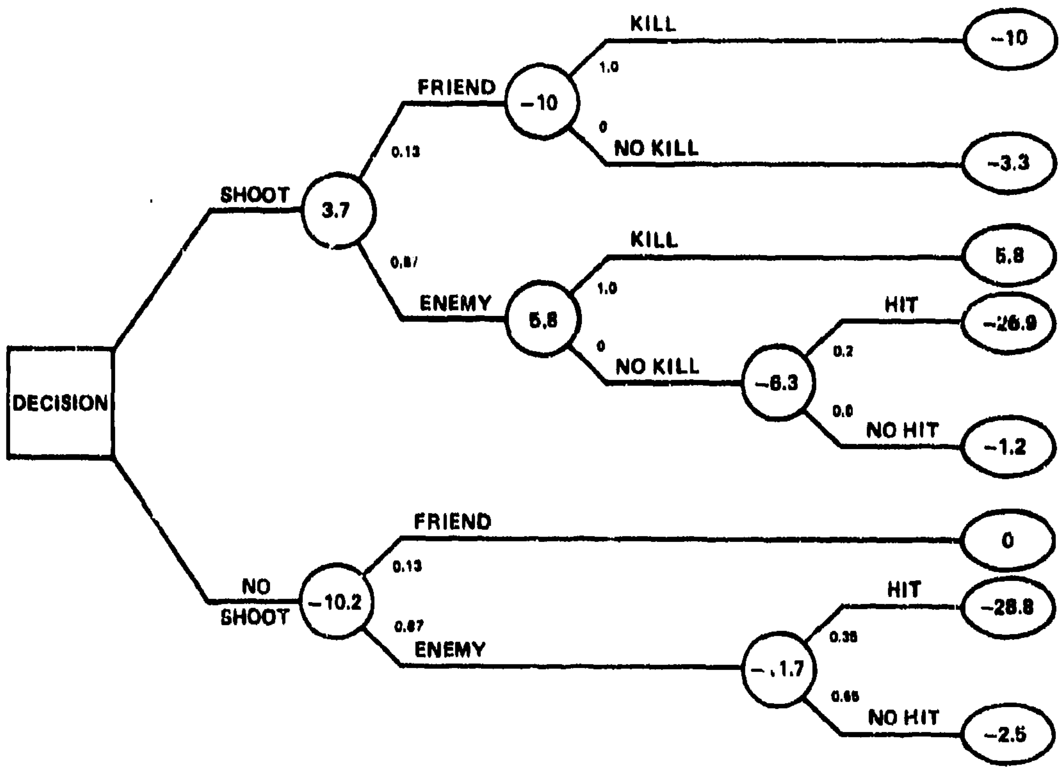


Figure 6-27

Commander: As I understand what you've just done, you've said that the value of the weapon system is equal to the difference in expected values of the decisions which would be made, given that the kill probability characterizing that weapon system has been changed. This seems to be the same procedure we used to calculate the value of information.

Consultant: That's right. Just as the value of information depended upon the assigned kill probability, so the value of changing the kill probability (the value of the weapon system) depends upon the probabilities assigned to friend and enemy. Obviously, if you were completely certain that the airplane was friendly, then you wouldn't pay anything to increase your kill probability since you wouldn't be shooting.

Commander: I get the feeling that there must be a trade-off between purchasing information and improving the weapon system. It looks to me like different combinations of probability of friend or enemy and probability of kill or no kill, would result in the same expected value. It would seem to me that one way to display this would be by plotting those combinations of P(e) and P(k) which gave the same expected value. I can imagine doing this for different expected values so I'd end up with something like a series of contour lines. Then if someone offered me the opportunity to purchase information which would change my prior probability a certain amount or to improve my kill probability, I could determine exactly what the trade-off was.

Consultant: That's almost right. I think from the way you described your approach, however, you've overlooked a subtle point. That is, you won't know in a given situation what kind of information you're going to get. As I understand what you were saying, you envisage making a trade-off between, say, receiving a positive IFF report, and increasing P(k) from 70 to 75%. The problem with this is that you won't know in advance exactly what information you're going to receive. You might get a positive IFF report which would decrease the probability of enemy, but on the other hand, you might get a negative IFF report which would increase that probability.

Commander: You're right. I'd forgotten about that. I remember now that when we were doing the value of information analysis, we had to consider the possibility of getting both kinds of reports and then weight the impact of each report by the probability that we'd get that report.

Consultant: That's right and that's almost exactly the procedure we could use here. The calculations involve no new concepts, and since they are a little involved, we won't do them just now. However, eventually we could draw curves that would enable you to compare the relative values of better information and increased weapons capability.

Commander: That's really a very clever trade-off. It seems to me that one of the biggest problems we have had in the past is, in fact, trying to show the value of information versus the value of weapons systems, and this looks like a very good way to do it.

Consultant: Now we've completed the analysis of the decision problem faced by the commander of USS HERRINGBONE. I think the major benefit of any decision analysis comes through the increased understanding of the problem afforded by the analysis. When this occurs, the numbers that went into the analysis become less important. You might say that we've taken a quantitative approach in order to get a better qualitative feel for the problem.

Commander: The sensitivity analyses were particularly informative. You quickly get a sense of what is important and what isn't. I only wish we could have done more of them.

Consultant: We could have done a great many more, very quickly, if we had used a computer. That's a major role that is emerging for the new generation of desk-top computers. Using an interactive program for computer-aided decision analysis, you could do on the computer everything we did here, and more. You use the computer as an extension of your intellect; it does the messy, complex, and tedious chores while you supply the judgments. In addition, it can check on the consistency of your judgments, inform you of inconsistencies, and give you the opportunity to correct errors. Because judgment is an important element in most important or difficult decisions, it is neither possible nor desirable to hand over the decision-making process to the computer. But the computer can be a useful tool for the decision maker; that's why we talk about computer-aided decision-making.

Commander: We've covered a lot of ground, and I'm not sure I'll remember all this.

Consultant: Because decision analysis is as much an approach as it is a set of techniques, it usually takes further work and attempts to apply the approach before the ideas settle comfortably. Eventually, however, you will find that decision analysis provides a useful and natural way of thinking about problems involving risk and uncertainty.

APPENDIX A

INFERENCE FROM EVIDENCE: BAYES' THEOREM

An analyst, especially one in the field of intelligence, is often required to work with data that are incomplete or possibly unreliable. Since he must use his judgment to evaluate the worth of that material, he needs to develop a technique that will help him to give his best judgment. This appendix describes such a method; it is based upon Bayes' Theorem and allows the analyst to combine new information with his prior assessment of the probability of an occurrence.

Consider the following problems. In each there is a need to revise the judgments of the probabilities of events:

1. A Russian submarine has been sighted leaving the Mediterranean Sea, and the analyst needs to infer whether or not it was a uclear submarine. Previously, the analyst felt there was a 60% probability it was nuclear. How can he revise this probability in light of the fact that the submarine seems to be following a straight track rather than taking evasive action?

2. The flag officer of an aircraft carrier is concerned with an unidentified plane on a straight course to his ship. Before attempted communication, there is a 50:50 chance it is friendly. If the plane refuses radio acknowledgment, is the probability likelihood of its being an enemy airplane now high enough to warrant attempting to shoot it down?

3. During the eight months leading up to the Soviet Union's invasion of Czechoslovakia, how did the occurrence of new intelligence data affect the probability of an invasion? (This problem will be examined in depth later in the appendix.)

An analyst can of course make direct assessments of any of the probabilities called for in the above examples. It will often happen, however, that he feels more comfortable or more confident with a somewhat different type of assessment. For example, it may be more natural for him to address the question "What is the likelihood that a nuclear submarine would travel in a straight track rather than take evasive action?" instead of the required probability which, in a sense, is the other way around: "What is the probability it is nuclear if we know it is following a straight track?"

One of the very important roles of probability theory is that it often allows the analyst to deduce in a logical fashion the probabilities he wants from probabilities and likelihoods already at hand. For example, probabilities "posterior" to a new piece of evidence can be deduced from "prior" probabilities and from the likelihood of observing the evidence given the hypothesis is or is not true.

We ask the reader to bear with us for a few paragraphs while we introduce some basic concepts and notations which underlie a particularly useful procedure for updating uncertainty as new information is received.

Bayes' Theorem

All of the problems cited above can be solved through the use of Bayes' Theorem, a formal, optimal rule for the revision of probabilities in the light of new evidence. The actual formula is fairly simple and can be derived from the rules of probability discussed in Chapter 3.

First, it is necessary to define three inputs to Bayes' Theorem. (1) A prior (or unconditional) probability of a hypothesis, H, is the probability an assessor assigns to it prior to receiving any information and is represented as P(H). (2) A likelihood of a datum is assigned given knowledge of some hypothesis, H, and is represented as P(D|H). (3) A joint probability is the probability of two events both occurring, that is, the probability that both the hypothesis is true and that a datum is observed. This probability is represented as P(H&D).

These three terms can be represented conveniently in a probability diagram as shown in Figure A-1: P(H) = .2 represents an assigned probability of .2 that the hypothesis is true, unconditional on the data; P(D|H) = .6 represents an assigned likelihood of .6 that a particular datum would be observed if the hypothesis were true; and P(D&H) = .12 represents a probability of .12 that both the hypothesis is true and the datum is observed. [Note: P(~D) is the probability of the datum, D, not occurring.]

When these three quantities are defined in general terms, they can be used to construct Bayes' Theorem. First, the definition of the likelihood of an event D, or datum, occurring, given hypothesis H is,

$$P(D|H) = \frac{P(D\&H)}{P(H)} \, . \qquad (1)$$

Equation (1) also holds for the reverse case of the probability of the hypothesis given the datum,

$$P(H|D) = \frac{P(D\&H)}{P(D)} \, . \qquad (2)$$

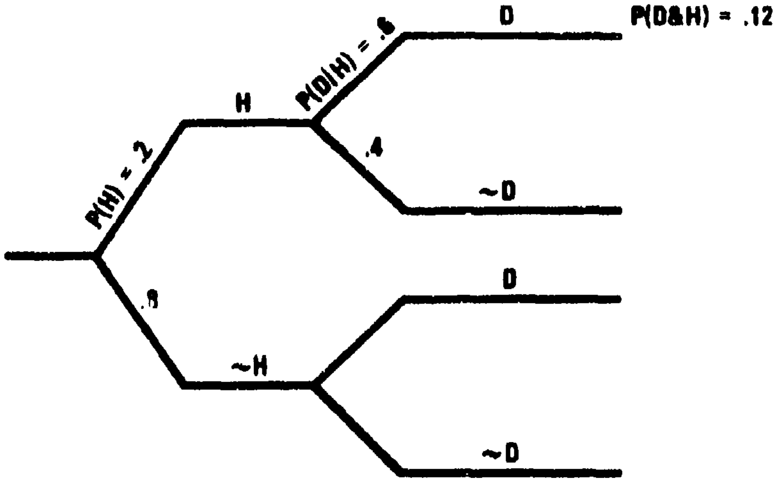


Figure A-1
TYPES OF PROBABILITIES

By combining equations (1) and (2), we derive the basic form of Bayes' Theorem,

$$P(H|D) = \frac{P(D|H)\ P(H)}{P(D)} \quad . \tag{3}$$

In review,

$P(D)$	-	Probability of the datum occurring.
$P(D\|H)$	-	Likelihood of the datum, given the hypothesis is true.
$P(H)$	-	<u>Prior</u> probability of hypothesis H, based on all information about H which is available <u>before</u> the new datum D is learned.
$P(H\|D)$	-	<u>Posterior</u>, or revised, probability of the hypothesis occurring, given that the datum has been observed.

For the case of two hypotheses, H_1 and H_2 (i.e., H_1 = the Russian submarine is nuclear, and H_2 = the Russian submarine is non-nuclear), Bayes' Theorem can be written twice:

$$P(H_1|D) = \frac{P(D|H_1)\ P(H_1)}{P(D)} \tag{4}$$

$$P(H_2|D) = \frac{P(D|H_2)\ P(H_2)}{P(D)} \quad . \tag{5}$$

By dividing one equation by the other, the P(D)'s in the denominators cancel out and leave the following:

$$\frac{P(H_1|D)}{P(H_2|D)} = \frac{P(D|H_1)\ P(H_1)}{P(D|H_2)\ P(H_2)} \quad . \tag{6}$$

The ratio far to the right of the equal sign in equation (6) is the

$$\text{Prior Odds} = \frac{P(H_1)}{P(H_2)} \quad . \tag{7}$$

It is equal to the probability that H_1 is true divided by the probability that H_2 is true <u>prior</u> to observing the evidence. The ratio immediately to the right of the equal sign in equation (6) is called the <u>likelihood</u> <u>ratio</u> (LR), or

$$LR = \frac{P(D|H_1)}{P(D|H_2)} \quad . \tag{8}$$

It is equal to the probability that the datum would have occurred if H_1 is true divided by the probability that it would have occurred if H_2 is true. The ratio to the left of the equation is the resulting <u>posterior</u> odds, the odds in favor of H_1 over H_2 after the datum has been observed:

$$\text{Posterior Odds} = \frac{P(H_1|D)}{P(H_2|D)}, \text{ or} \tag{9}$$

$$\text{Posterior odds} = LR \times \text{Prior Odds.} \tag{10}$$

Equation (10) can now be used to revise the odds from prior to posterior as a function of the information in the observed datum. Because of the kind of thinking required to use this method, Bayes' Theorem is often called the theorem of inverse probabilities. To go from prior odds to the posterior odds, it is necessary to assume first that H_1 is true and then assume that H_2 is true, and to ask how likely it is that the datum would be observed under each assumption. The answers to that question are specified in the likelihood ratio. To the extent that the datum is more likely if H_1 is true, the posterior odds are greater than the prior odds. To the extent that the datum is more likely under H_2, the reverse is true.

Czechoslovakia Case

Since Bayes' Theorem can be used to update existing information, consider the following problem in which new data are periodically received: Assume that today is January 6, 1968, and the following article appeared in yesterday's New York Times newspaper:

> Vienna, Jan. 5, 1968. Antonin Novotny has been ousted as the Czechoslovak Communist Party chief, the Czecn press agency CTK confirmed tonight. However, he will remain as President of the country. The Yugoslav press agency Tanyug first reported the ouster of Mr. Novotny as First Secretary of the Party and the Czechoslovak government confirmed it by distributing photographs of Alexander Dubcek and identifying him as the new Party leader.

Now that the Soviet favorite, Antonin Novotny, has been replaced by a liberal, Alexander Dubcek, the chance for a Soviet invasion has become high enough that the situation is worth considering. (Note: Although the invasion of August 20, 1968, is now history, assume this is a future event that may or may not occur.) The job of the U.S. analyst is to follow the situation for the next year and to update periodically the likelihood of an invasion as new intelligence data are received. For simplification, the only source of information will be actual newspaper articles from the New York Times, i.e., public information.

Consider the two hypotheses: The first hypothesis, H_1, is that the Soviet Union will invade Czechoslovakia within one year; the second hypothesis, H_2, is that it will not invade Czechoslovakia within one year. The analyst's first step is to estimate today, January 6, 1968, the likelihood of an invasion occurring within one year. Although a new Communist Party chief is now in power in Czechoslovakia, it is unlikely that the USSR will consider an invasion unless Dubcek starts to make radical changes. In this case, the analyst may determine that there is a one to 20 chance of H_1 being true, that is, the prior odds are given by

$$\frac{P(H_1)}{P(H_2)} = \frac{1}{20} .$$

The easiest way of displaying the impact of new data on the prior odds is through the use of a "log odds" chart. Consider once again Bayes' Theorem as written in equation (6). The posterior odds equal the product of the likelihood ratio and the prior odds. If the prior odds equal 1 and the likelihood ratio is 2:1, the posterior odds equal 2. But if the prior odds equal 100 and the likelihood ratio is 2:1, the posterior odds equal 200. Thus, on an odds scale, the apparent degree of impact of the datum depends upon the value of the prior odds. Any given likelihood ratio will result in greater movement of the odds when the prior odds are already high than when the prior odds are low.

This problem does not occur when Bayes' Theorem is written in logarithmic form. When both sides of equation (6) are transformed into logarithms, the log of the posterior odds equals the log of the prior odds plus the log of the likelihood ratio:

$$\log \frac{P(H_1|D)}{P(H_2|D)} = \log \frac{P(D|H_1)}{P(D|H_2)} + \log \frac{P(H_1)}{P(H_2)} . \tag{11}$$

Whatever the prior odds, the impact of a datum is to add on a fixed amount to the prior odds. For example, assume that the log of the likelihood ratio is 0.5. If the prior log odds equal 0.0, then the posterior log odds equal +0.5. This additive version of Bayes' Theorem yields a simple graphical display of posterior odds.

Consider the chart in Figure A-2. The Y-axis is scaled in log odds, but the right vertical axis is marked in odds, the left vertical in probabilities. The horizontal axis is scaled in time. Both hypotheses are equally likely in the middle of the Y-scale where the odds are 1:1. H_1 becomes increasingly likely as the scale goes up, and H_2 becomes increasingly likely as the scale moves down. According to the previous estimate, the prior odds are 1:20 and are represented on the following log odds chart,[1] Figure A-2.

Datum #1: The first datum concerning the hypothesis occurs on February 22:

> Prague, Feb. 22. The Czechoslovak Communist Party began celebrating today its 20th Anniversary in power amid indications that some members were uneasy about the presence of their most prestigious foreign guest, Leonid I. Brezhnev, the Soviet Party Leader . . . Some Czechoslovaks are concerned that Mr. Brezhnev may be here to put pressure on Alexander Dubcek, the 47-year-old Slovak who replaced the Czech party leader, Antonin Novotny, less than two months ago.

[1]The log odds charts used in this Appendix were drawn up to facilitate the use of the odds scale on the right-hand side of the chart. The probabilities on the left-hand scale have been approximated to the nearest integral 1%, and were included only for the reader's reference.

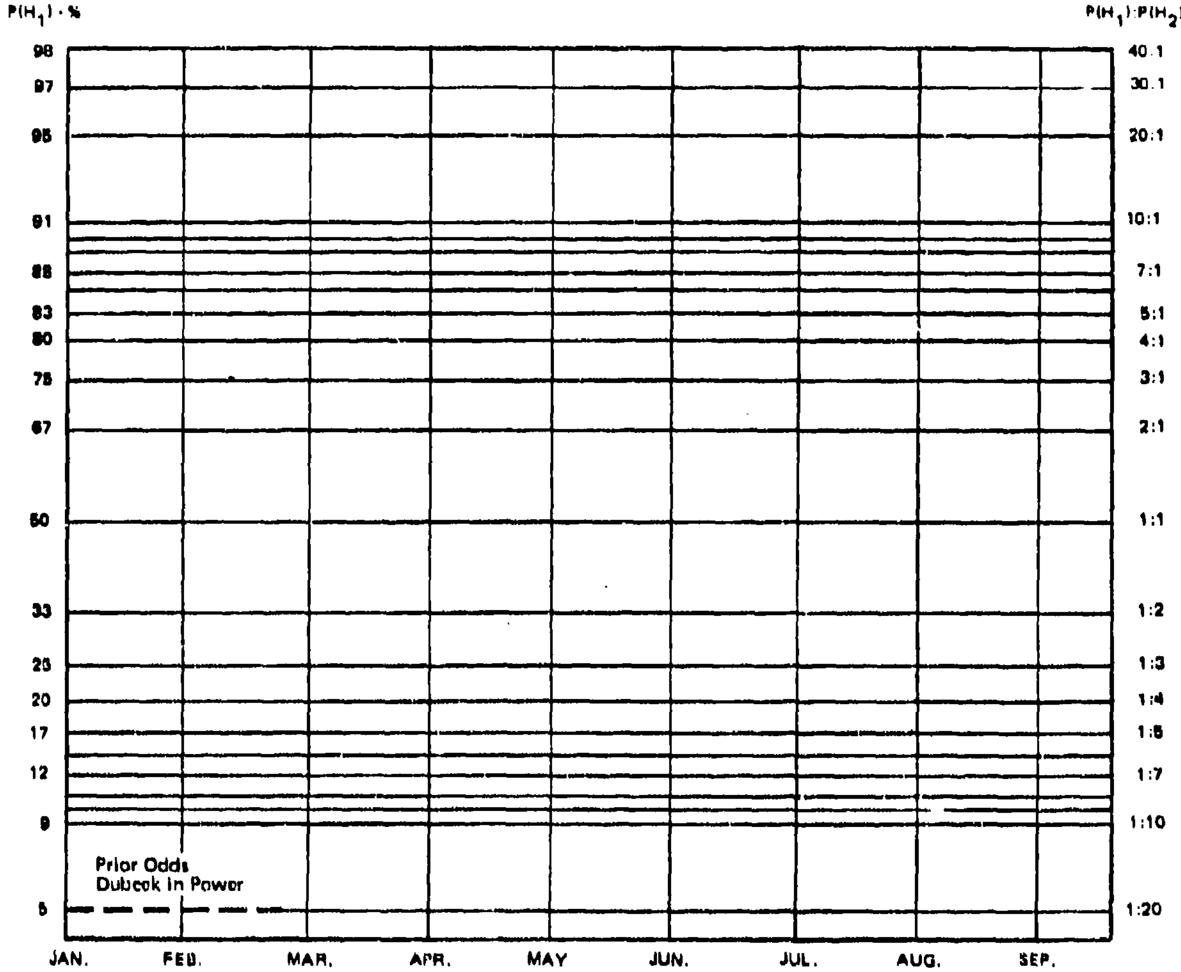


Figure A-2
PRIOR ODDS

Now that a new piece of datum is available, the analyst needs to update his earlier prediction from the prior odds of 1:20. He must decide if this datum has any bearing on H_1, and if it does, whether it favors H_1 or H_2, and to what extent. Since the new posterior odds are equal to the LR times the prior odds, we need only to determine the LR to calculate the revised odds. Because Brezhnev may be in Czechoslovakia to put pressure on Alexander Dubcek, this datum slightly favors the hypothesis of an invasion. In this case, the analyst assesses a probability of .80 that Brezhnev would be there if an invasion were planned, and a .50 probability that he would be there if there were to be no invasion. Therefore the likelihood ratio of the first datum is .80/.50 = 1.6:1. With this value and the prior odds, we can compute the posterior odds:

$$\text{Posterior} = \text{LR} \times \text{Prior}$$

$$= \frac{1.6}{1} \times \frac{1}{20}$$

$$= \frac{1}{12.5} \quad .$$

The updated posterior odds, 1:12.5, are shown on the chart in Figure A-3, the dashed line from the prior odds to the posterior odds indicating that the odds have been revised.

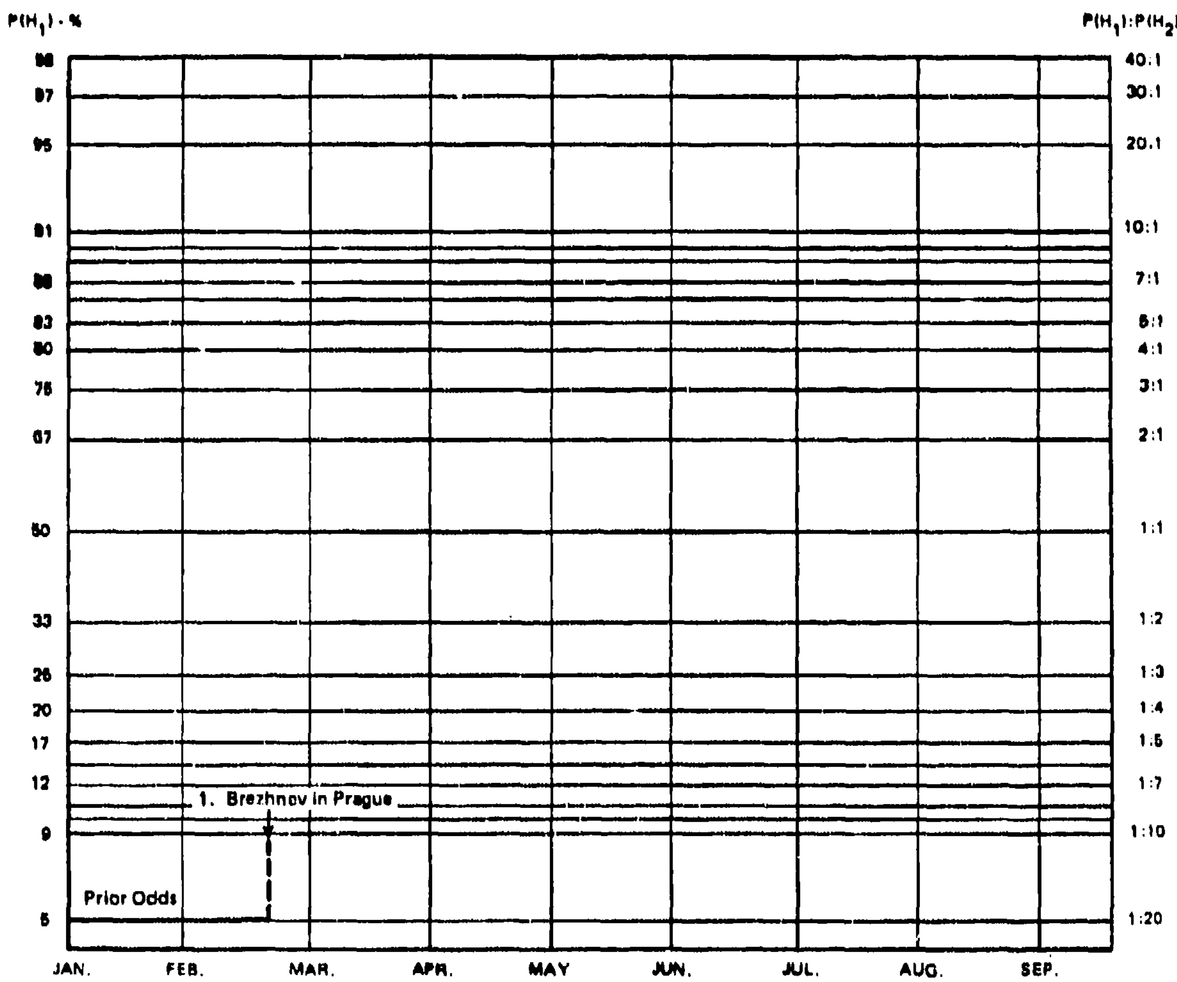


Figure A-3

DATUM No. 1

Note that it is only the likelihood ratio which affects the posterior odds. It does not matter what value the likelihoods have as long as their ratio is unaffected.

This has an important practical implication for work, especially in the field of intelligence, because it permits the analyst to avoid awkwardness in specifying exactly what the datum is and, therefore, its precise probability conditional on the alternative hypotheses. Data can be defined loosely or precisely. Datum #1, for example, could be interpreted as meaning the exact wording of the newspaper report. In that case, the probabilities appearing in the likelihood ratio would be infinitesimally small, but it might still be reasonable to assess their ratio to each other, i.e., the ratio of 1.6 to 1. On the other hand, the datum might be more loosely defined in a way that simply captured the flavor of the communique. In this case the likelihoods would be larger, but in all probability, the ratio would still be very close to 1.6 to 1. In circumstances like this--which are very common--when the likelihood ratio is not very sensitive to how the datum is defined, there is no need to be overly precise. This is a marked practical convenience.

Datum #2:

> Berlin, April 21. Marshal Ivan I. Yakubovsky of the Soviet Union, the commander of the Warsaw Pact Military Alliance, arrived in East Berlin today for talks with Walter Ulbricht, First Secretary of the Communist Party, on measures to reinforce the Alliance's defense system. The marshall came from Warsaw, where he had conferred with Polish leaders.

In this case, if the Soviet Union were planning an invasion, there is an assessed 95% probability that Marshal Yakubovsky would meet with the different members of the Warsaw Pact; and there is an assessed probability of 40% that he might be meeting with them on other routine matters if no invasion were planned. The likelihood ratio, 2.4:1, for the second datum also favors H_1, giving new posterior odds that can be drawn on the chart, Figure A-4.

$$\text{Posterior} = \text{LR} \times \text{Prior}$$

$$= \frac{2.4}{1} \times \frac{1}{12.5}$$

$$= \frac{1}{5.2}$$

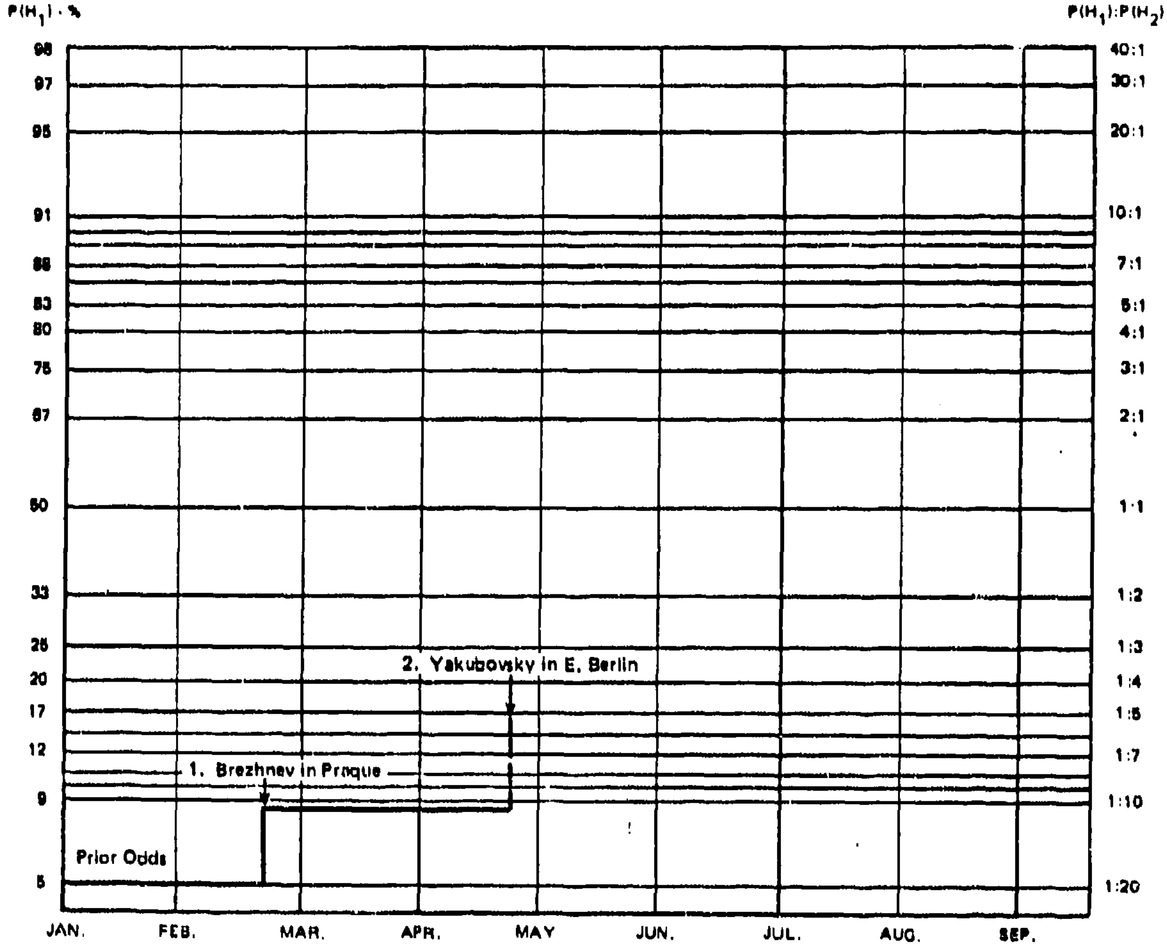


Figure A-4
DATUM No. 2

Datum #3:

> Prague, May 3. It was disclosed tonight that Alexander Dubcek, the new Communist Party chief, had flown to Moscow for meetings with the Soviet Government.

Datum #3 only informs the analyst that Dubcek is in Moscow meeting with the Soviet Government. It is up to the analyst to use his knowledge of the situation to determine if this is just a routine meeting to discuss trade or economic matters, or if it is an indication of a serious rift concerning communist ideology. The analyst may feel there is an 80% probability Datum #3 would occur if there is to be an invasion and a 40% probability Datum #3 would occur without an invasion being planned. The effect of the new likelihood ratio, 2.0:1, is displayed in Figure A-5.

Note that when using Bayes' Theorem and log odds charts, it is not necessary to take the data in order. The present posterior odds of 1:2.6 would have been determined regardless of the order in which Data #1, #2 and #3 had been considered as long as the three likelihood ratios could remain the same. The order must be considered only if the component likelihood ratios change for different orders.

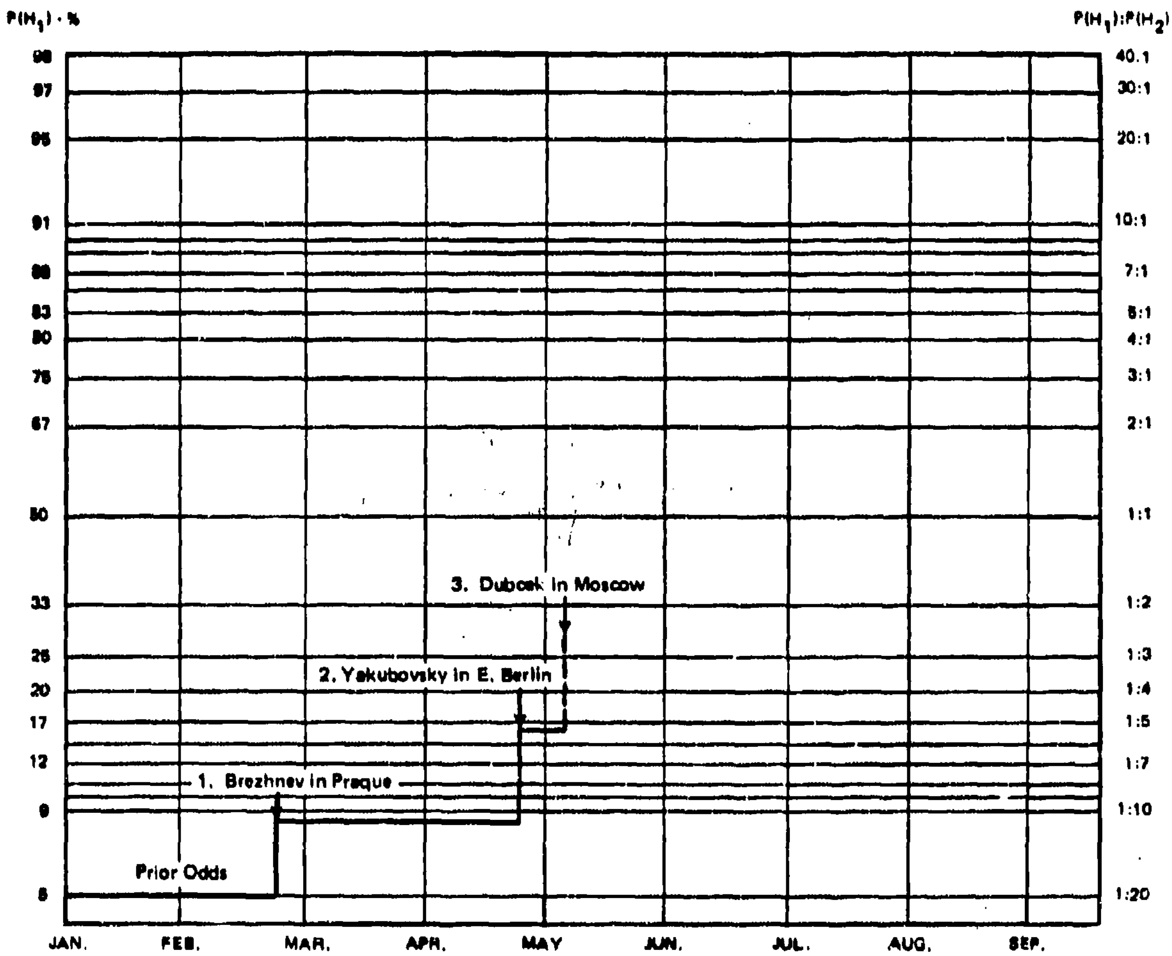


Figure A-5
DATUM No. 3

Datum #4:

Prague, May 24. The armed forces of the Warsaw Pact will hold exercises on Czechoslovak and Polish territory next month, Czechoslovakia announced tonight. The statement distributed by the Ceteka press agency, said that the exercises would be under the command of Marshal Ivan F. Yakubovsky of the Soviet Union, Commander-in-Chief of the forces of the Warsaw Pact. The military alliance of the Eastern Bloc consists of the Soviet Union, Bulgaria, Czechoslovakia, East Germany, Poland and Rumania.

This new datum, considered after the occurrence of the other previous events, is a very strong one and now clarifies Marshal Yakubovsky's actions in East Berlin on April 21 (Datum #2). The analyst assessed that the event was certain to occur if an invasion were imminent, but the probability was only 10% if no invasion were planned. The resulting LR of 10:1 is now added to the log odds chart in Figure A-6. With

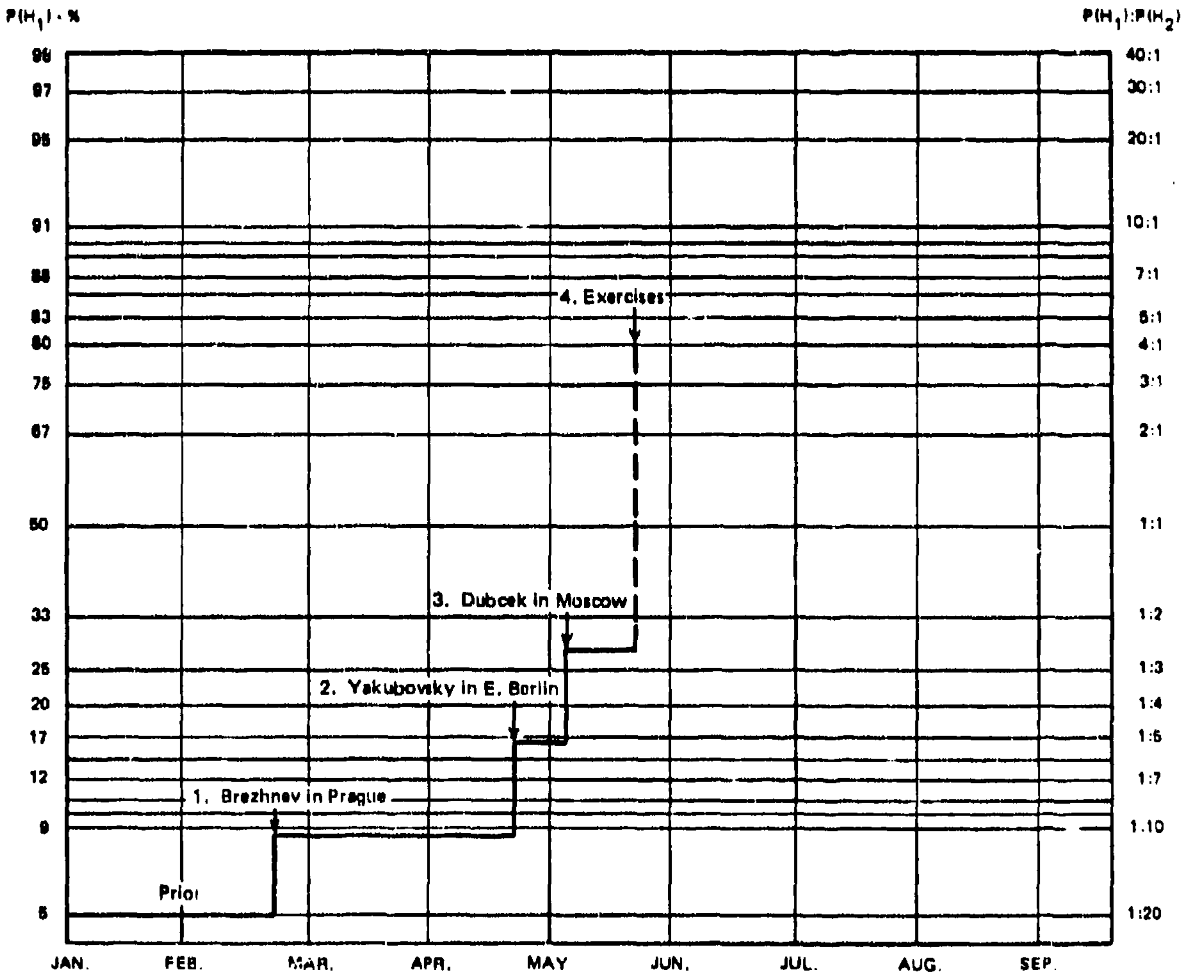


Figure A-6

DATUM No. 4

this last datum, the posterior odds of an invasion have gone over the 50:50 chance for the first time. In fact, the odds are now 3.9:1 in favor of H_1, or almost an 80% likelihood of an invasion occurring.

Datum #5:

> Prague, June 18. The Soviet commander of Warsaw Pact forces has said that maneuvers this month in Czechoslovakia would involve only command staffs of the various services . . . The Soviet commander's remarks were apparently intended to forestall rumors that the maneuvers might be aimed at overturning Czechoslovakia's liberal regime.

Until this point, the new data have favored H_1. Datum #5 now seems to indicate that the implied meaning of Datum #4 is incorrect. Inasmuch as there is no way of determining whether or not this statement is only propaganda, the analyst might assess a 60% likelihood of this statement being released given there is going to be an invasion and 80% probability if there is to be no invasion; a likelihood ratio equal to 1:1.3 against H_1 giving posterior odds of 3:1 (Figure A-7).

Data #6, #7, #8, #9:

The next four pieces of data, from June 30 to August 5, all indicate a decreasing amount of tension in Czechoslovakia, favoring H_1, no invasion. Figure A-8 represents the NY Times articles and the assessed LR's, and their impacts.

Datum #6:

> Prague, June 30. The military staff exercises of Warsaw Pact forces that created anxiety in Czechoslovakia for about a month ended today.

$$LR_6 = \frac{P(D_6|H_1)}{P(D_6|H_2)} = \frac{.30}{.80} = \frac{1}{2.7}$$

Datum #7:

> Prague, July 13. Some of the Soviet troops still in Czechoslovakia two weeks after the end of the joint Warsaw Pact maneuvers started for home today... At the same time, leaders of the Soviet Union and four East European allies gathered in Warsaw without the Czechoslovaks to discuss Prague's determination to make the Communist regime more democratic.

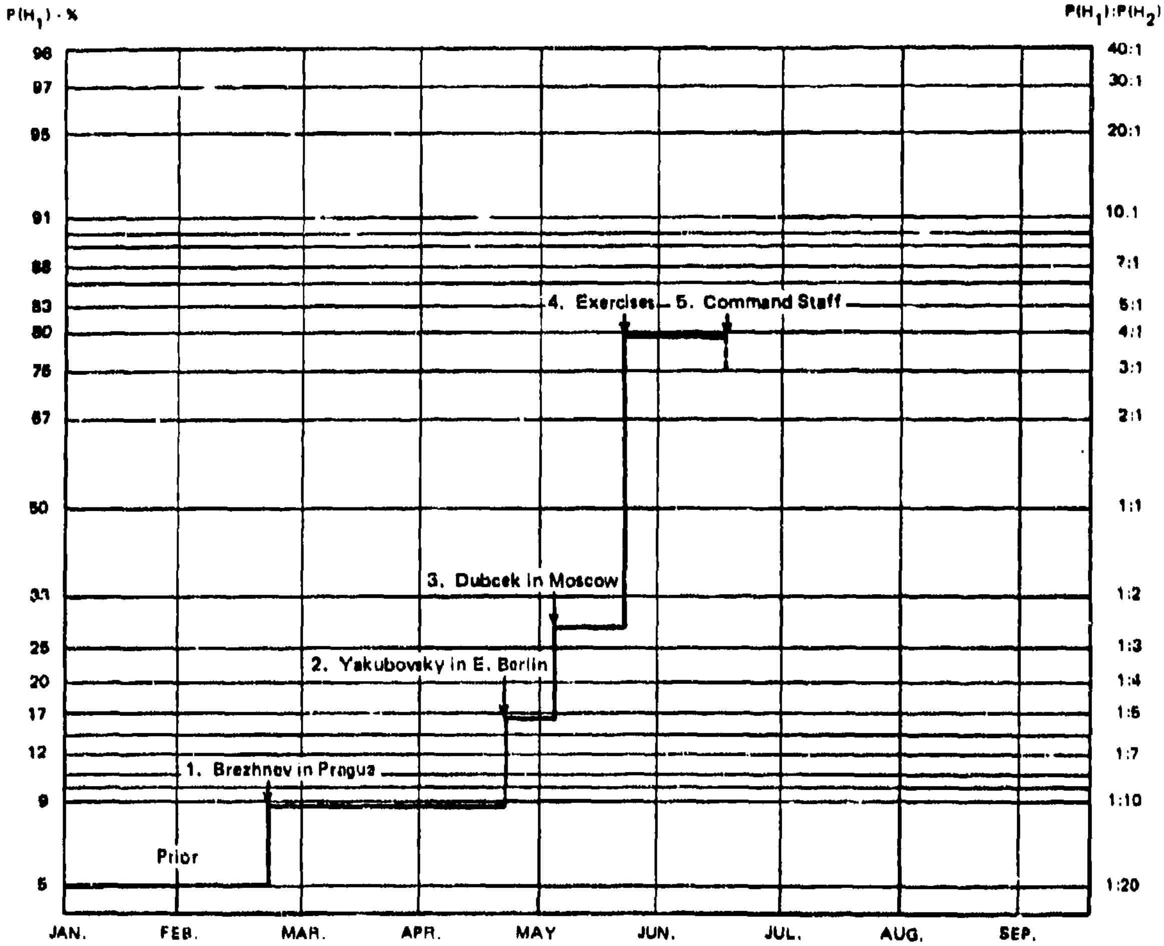


Figure A-7

DATUM No. 5

$$LR_7 = \frac{P(D_7|H_1)}{P(D_8|H_2)} = \frac{.30}{.60} = \frac{1}{2}$$

<u>Datum #8:</u>

Warsaw, July 30. Informed sources said today that the Soviet Army had concentrated combat units in Poland capable of armed intervention in Czechoslovakia and that the Soviet troop movements in the last three days had assumed dimensions in excess of a simple show of force.

Prague, July 30. Well-informed Czechoslovak sources believe that the meeting between the leaders of this country and those of the Soviet Union has served to push the possibility of Soviet military intervention further into the background. Their belief is based on authoritative reports that the first day of negotiations at Cierna, Czechoslovakia, yesterday brought no new Soviet demands or threats.

$$LR = \frac{P(D_8|H_1)}{P(D_8|H_2)} = \frac{.50}{.55} = \frac{1}{1.1}$$

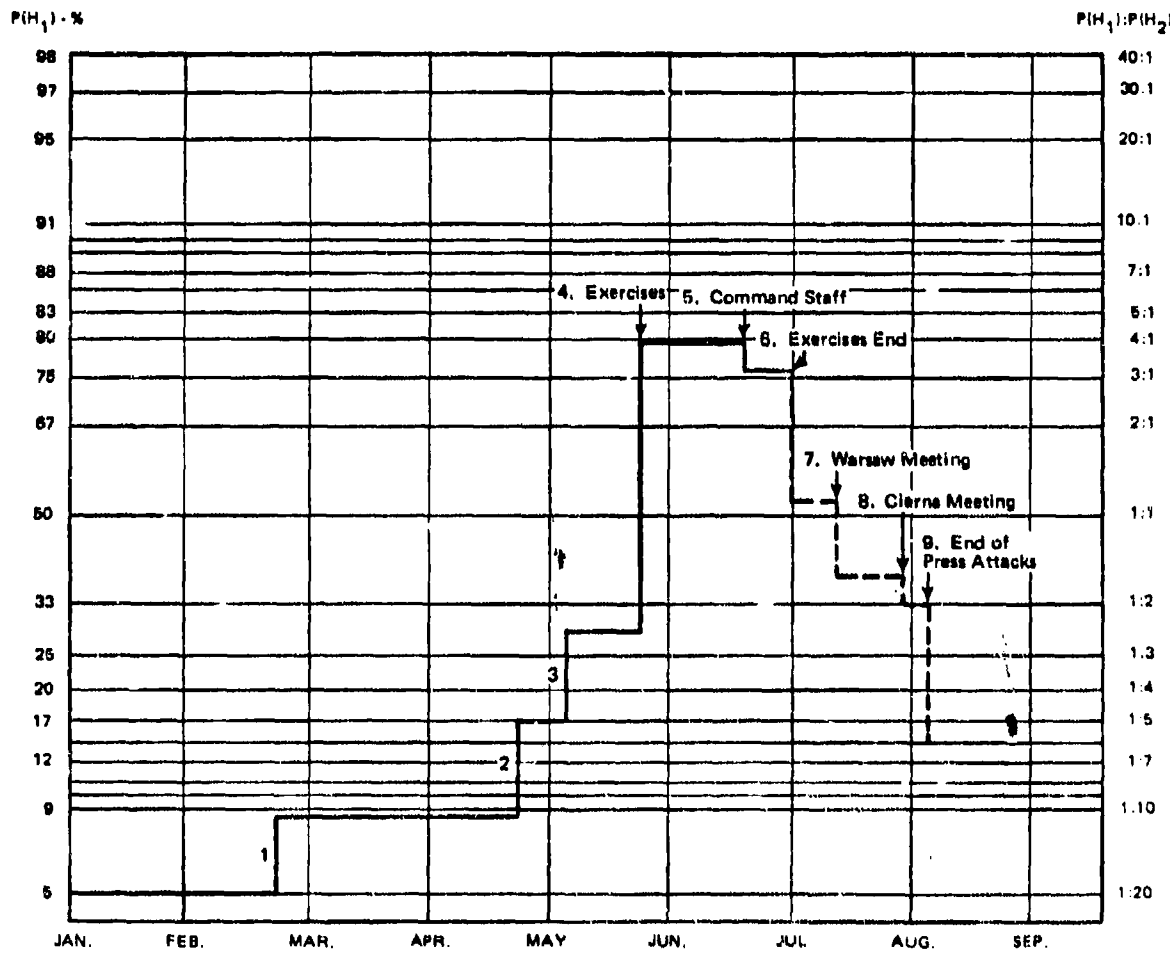


Figure A-8

DATA No. 6 THROUGH No. 9

Datum #9:

Moscow, Aug. 5. The specter of "counterrevolution" in Czechoslovakia that evoked cries of outrage and alarm from the Soviet Union seems in Moscow to have disappeared abruptly into the mists of unhistory. Scornful denunciations of democratization, allegations of anti-Communist conspiracies and similar ominous accusations aimed at the Prague liberals have vanished from the Soviet press. Mass propaganda rallies throughout the country to stir up hostility against the Czechoslovak liberation program have ended.

$$LR_9 = \frac{P(D_9|H_1)}{P(D_9|H_2)} = \frac{.30}{.90} = \frac{1}{3}$$

Datum #10:

Moscow, Aug. 9. The Soviet Communist Party warned its members today that if any of them had any ideas about liberalization in the party along the line of the recent Czechoslovak reforms, they had best forget them. Pravda, the party organ, printed a long article defending the principle of "democratic centralism" and declaring that an abandonment of this system of unquestioning obedience to decisions from the top would lead to a collapse of Communist Party rule.

From June 18 (Datum #5) to August 5 (Datum #9), the differences between the Soviet Union and Czechoslovakia seemed to have been resolved, and general relationships appeared to have improved; Datum #10 is the first real piece of evidence that those differences may not have been completely resolved. With this new datum, the analyst might assess that LR = .80/.45 = 1.8/1. Figure A-9 shows the effect of the addition of this new information.

Datum #11:

Moscow, Aug. 16. The Soviet Union has resumed polemics against the Czechoslovak press after a lapse of three weeks.

The Soviet Union has now renewed direct attacks against Czechoslovakia. Apparently all private efforts at a reconciliation have failed, resulting in an increasing LR = .85/.25 = 3.4/1 (see Figure A-10).

Datum #12:

Moscow, Aug. 18. The Soviet Communist Party expressed apprehension today that the Czech leadership appeared

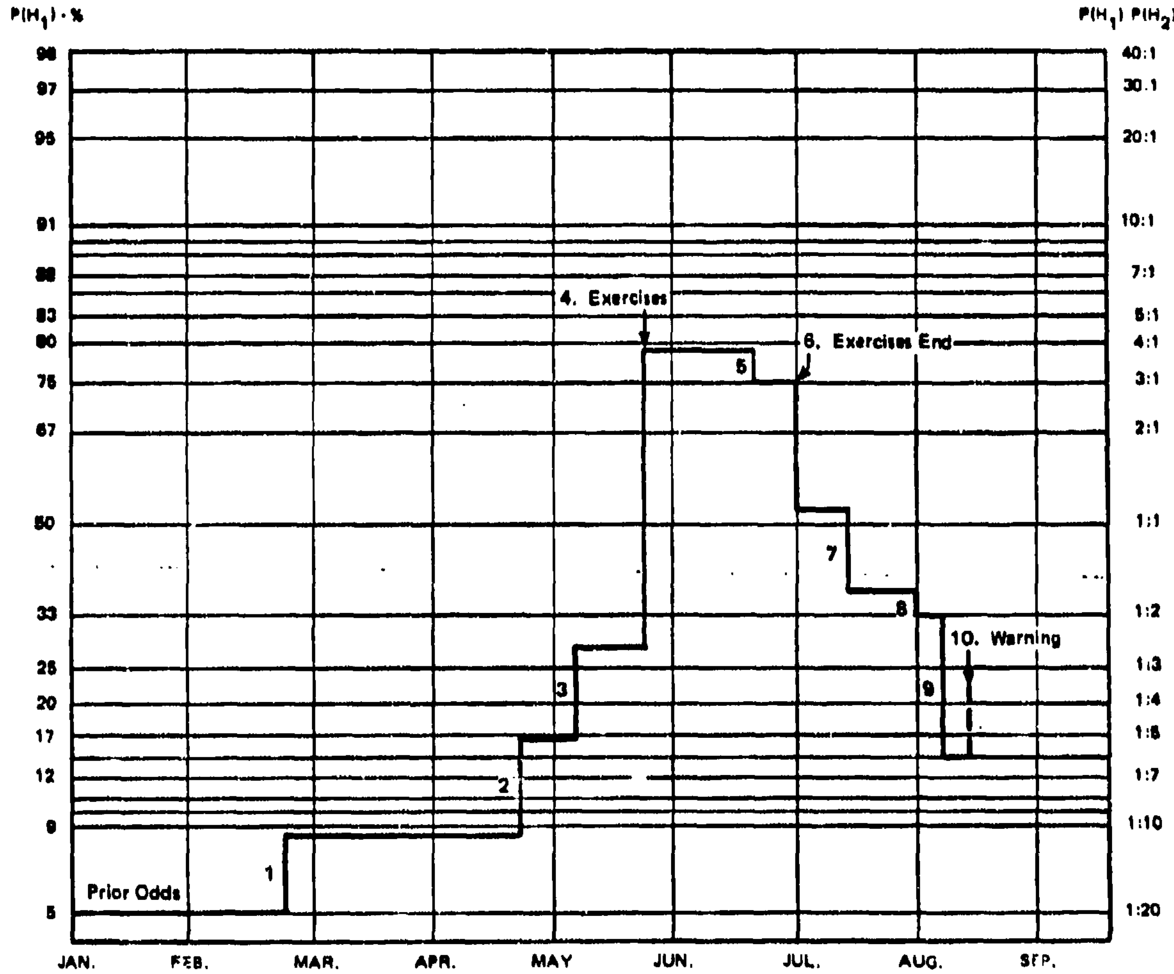


Figure A-9

DATUM No. 10

to be losing control in the country. "Subversive activities by antisocialist forces" have resumed in Prague, Pravda, the Soviet Party's newspaper asserted. It suggested that the Czechoslovak leaders were unable to cope with the threatening situations.

Soviet criticism of Czech leaders is now greatly increasing and seems to imply an invasion is imminent. The analyst assesses an 80% probability that these remarks would

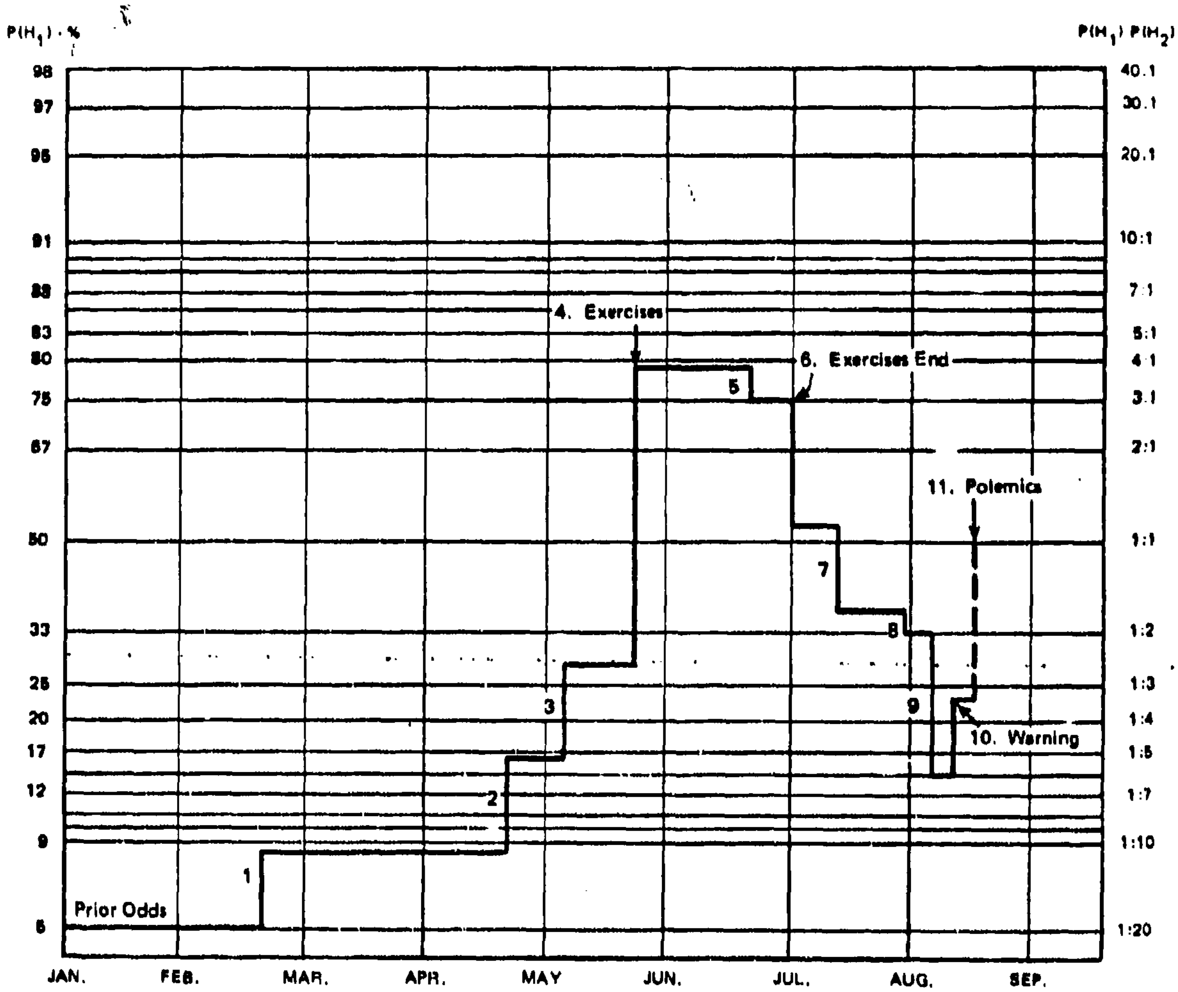


Figure A-10

DATUM No. 11

be in Pravda if an invasion were planned, and only a 20% probability if an invasion were not planned. This twelfth datum is a very strong one with an LR = 4.0/1 (see Figure A-11).

Datum #13:

> Prague, Aug. 19. The press agency CTK announced that Czechoslovak Army divisions will exercise Wednesday and Thursday in Bohemia, with observers from other Warsaw Pact countries looking on.

Datum #13 is the last piece of data to be considered. It is obvious now that there is an intention to invade Czechoslovakia. The LR = 1.0/.12 = 8.3/1 and raises the total posterior odds to 33:1 or about a 97% probability in

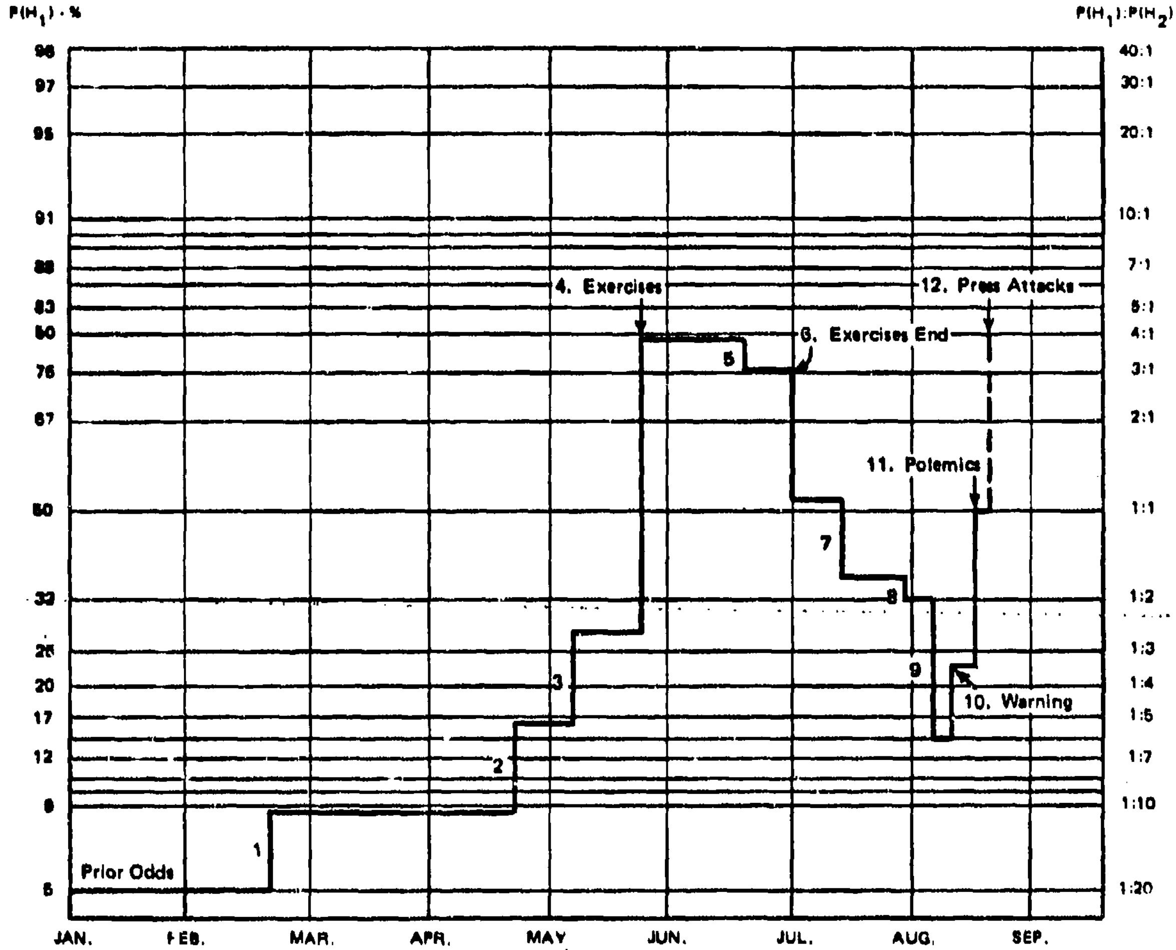


Figure A-11

DATUM No. 12

favor of H_1, the Soviet Union is planning to invade Czechoslovakia (see Figure A-12).

Two days later the following headlines and article appeared in the <u>NY Times</u>:

> CZECHOSLOVAKIA INVADED BY RUSSIANS AND FOUR OTHER WARSAW PACT FORCES
>
> Moscow, Wed., Aug. 21. Moscow announced this morning that troops from the Soviet Union and four other Communist countries [on Aug. 20] had invaded Czechoslovakia at the request of the "party-Government leaders of the Czechoslovak Socialist Republic."

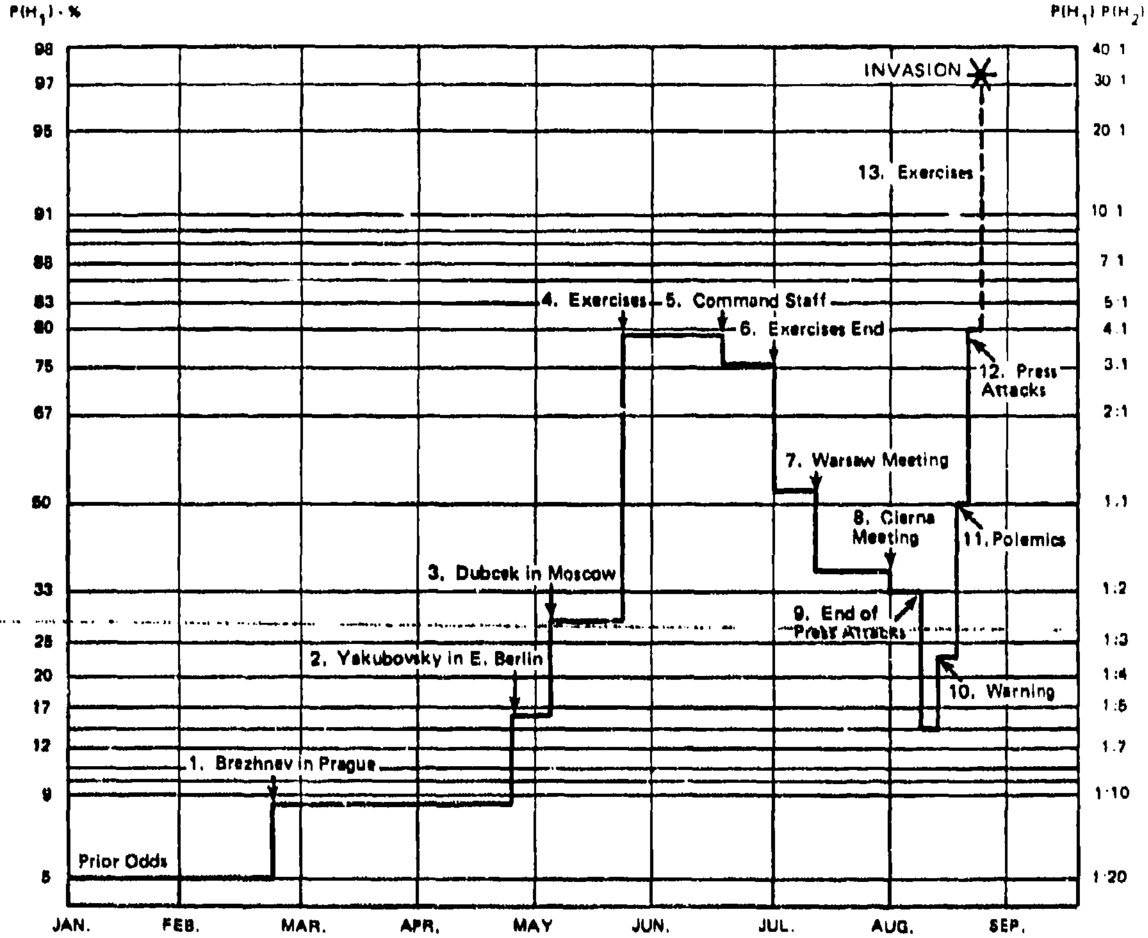


Figure A-12

DATUM No. 13

In this appendix the events leading up to the August 20, 1968, invasion of Czechoslovakia have been studied. The initial assessment of prior odds 1:20 against the invasion was slowly increased until May 24 when it was announced that the armed forces of the Warsaw Pact would hold exercises on Czech territory. At that time the odds were 3.9:1 in favor of H_1, or almost an 80% likelihood of an invasion occurring. After that point, four consecutive sets of data (#6-#9) all indicated a resolution of differences and a decreasing amount of tension between the USSR and Czechoslovakia. On August 5 (Datum #9) the likelihood of an invasion dropped to 1:6 or only a 14% chance. But on August 9, eleven days before the actual invasion, the Soviet Union resumed press attacks against Czech leadership. Those attacks continued until August 19 (Datum #13) and the press agency CTK announcement of exercises in Czechoslovakia. That announcement raised the total posterior odds to 33.1, or a 97% probability of invasion. In each case, Bayes' Theorem has enabled the analyst to revise the probabilities as new information was received. It permitted him to substitute quantity for quality, with large amounts of data sufficing in situations where the information was incomplete or not entirely reliable. Finally, Bayes' Theorem allowed the analyst to combine data from many different sources, to weigh each datum as it was received, and to visualize the effect of a new datum graphically on a log odds chart.

APPENDIX B

BAYESIAN INFERENCE ABOUT MANY-VALUED UNCERTAINTIES

In Appendix A, "Inference from Evidence: Bayes' Theorem," we saw how "prior" uncertainty about a two-valued event (an invasion occurs or not) can be updated after the receipt of evidence (e.g., a certain intelligence report), by means of Bayes' Theorem. The prior probabilities are combined with a "likelihood ratio," which measures the diagnostic value of the evidence, to produce "posterior" probabilities. Essentially the same prior-posterior logic can be used to update uncertainty about a many-valued event, such as which of several possibilities (e.g., type of contract to be awarded) or what value on a continuum (e.g., size of an R&D budget) will occur. Instead of just two alternative hypotheses (H_1 and H_2) to consider, there are any number of hypotheses, from three to infinity (H_1, H_2, H_3, ..., and so on, or, more briefly, H_i).

As with two-valued uncertainties, the key inputs are prior probabilities and likelihoods reflecting the diagnostic value of the evidence. Prior probabilities, $p(H_1)$, $p(H_2)$, $p(H_3)$, and the like are written as $p(H_i)$. Likelihoods associated with a datum of evidence are written as $p(D|H_i)$. In the special two-value case discussed in Appendix A, the prior-posterior calculation is rather simple. The posterior odds are calculated as the product of the prior odds and the likelihood ratio:

$$\frac{P(H_1|D)}{P(H_2|D)} = \frac{P(H_1)}{P(H_2)} \times \frac{P(D|H_1)}{P(D|H_2)} . \qquad (1)$$

When there are more than two hypotheses, the procedure is somewhat more complicated.

Bayes' Theorem tells us that the posterior probability of any one hypothesis H_i is related to its prior probability and its likelihood, given the evidence, according to the following formula:

$$P(H_i|D) = \frac{P(H_i)P(D|H_i)}{P(D)} . \qquad (2)$$

This formula is repeated for each of the possible hypotheses, the number of repetitions indicated by the largest value of i. All of these posterior probabilities must, of course, sum to one on the assumption that one of the hypotheses is true. Therefore, since the unconditional probability of the datum, P (D), appears in the denominator of the posterior probability expressions for all H_i, equation (2), can be restated as a proportionality indicated by the symbol $\propto$:

$$P(H_i|D) \propto P(H_i)P(D|H_i). \tag{3}$$

Equation (3) states that the posterior probability of a hypothesis is proportional to its prior probability times the likelihood of the datum and thereby eliminates the need to consider P(D) specifically. Since p(D) serves as a constant of proportionality, its actual value can be determined by summing the right-hand products of the equations that result from applying Equation (3) for each hypothesis. That is,

$$P(D) = \sum_{i=1}^{n} P(H_i)P(D|H_i). \tag{4}$$

A simple example may be used to illustrate the application of this algebra.

A Three-Valued Example

Suppose that the Commanding Officer of an aircraft carrier is concerned with an unidentified plane flying on a straight course to his ship. It may be American or non-American and either friend or foe. He assigns prior probabilities of 0.5, 0.2, and 0.3, respectively. When the plane refuses radio acknowledgment, the Commanding Officer must decide whether his posterior probabilities would justify shooting the plane down. His reading of the evidence is: if the plane were American, the likelihood of refused radio communication would be 0.1; if other friend, 0.3; and if foe, 0.8. (Note that these are likelihoods and do not need to add to one).

Figure B-1 shows the calculation of the posterior probabilities (Column 5) from the prior probabilities (Column 2) and the likelihoods (Column 3). The joint probabilities in Column 4 are the products of Columns 2 and 3 for each hypothesis, and the sum of Column 4 is P(D), in this case, 0.35. The posterior probability for each hypothesis, Column 5, is the joint probability in Column 4 divided by the sum of the joint probabilities, P(D). These computations are all implied by equations (2), (3) and (4).

HYPOTHESIS H_i	PRIOR PROBABILITY $P(H_i)$	LIKELIHOOD $P(D\|H_i)$	JOINT PROBABILITY $P(H_i)\,P(D\|H_i)$	POSTERIOR PROBABILITY $P(H_i\|D)$
AMERICAN	0.5	0.1	0.05	0.14
FRIEND	0.2	0.3	0.06	0.17
FOE	0.3	0.8	0.24	0.69
	1.0	—	P(D) = 0.35	1.00

Figure B-1

COMPUTING POSTERIOR PROBABILITIES FOR A THREE-VALUED UNCERTAINTY

Generalization of the Prior-Posterior Worksheet

The same worksheet can be used for any prior-posterior analysis, regardless of how many hypotheses there are, provided that, for each hypothesis, there is a prior probability and a likelihood. Each hypothesis will correspond to one row on the worksheet as in Figure B-1.

Where the Likelihoods Come From

It makes sense to use this kind of prior-posterior analysis to update uncertainty only if the inputs required, in this case, prior probabilities and likelihoods, are more conveniently or more reliably obtained than the direct assessment of a posterior distribution would be. In fact, since likelihoods are often quite difficult to assess subjectively, this condition will not necessarily hold. It may be quite difficult to visualize how likely it is that you would have observed the evidence that you have observed, given alternative possible hypotheses.

However, the likelihoods can sometimes be determined quite routinely, notably when the process whereby the data generated is clear-cut and well-understood. Suppose, for example that the "target" uncertainty is the proportion of red chips in a bag of red and white chips, and suppose that the datum observed is that a red chip has been drawn randomly from the bag. In this case, the assessment of the likelihoods presents no problem. It is obvious that if 20% of the chips are red, there is a 0.2 chance that a red chip will be drawn; if 40% of the chips are red, a 0.4 chance; and so on. The only subjective judgment that has to be supplied, therefore, is the prior probability distribution. If only four hypotheses were considered (20% or 40% or 80% or 100% of the chips are red), and each of these were given equal probability, Figure B-2 shows how the posterior probability distribution would be computed. The last column shows that the posterior probabilites are now .08, .17, .33, and .42, respectively.

HYPOTHESIS % RED CHIPS	PRIOR PROBABILITY	LIKELIHOOD	JOINT PROBABILITY	POSTERIOR PROBABILITY
20	0.25	0.2	0.05	0.08
40	0.25	0.4	0.10	0.17
80	0.25	0.8	0.20	0.33
100	0.25	1.00	0.25	0.42
	1.00		0.60	1.00

Figure B-2

PRIOR-POSTERIOR ANALYSIS WITH OBVIOUS LIKELIHOODS

It sometimes happens that the likelihoods are not obvious but can be determined in some manner, such as by using the laws of probability to calculate them from simple assessments about which the assessor feels relatively confident. Suppose that instead of one, ten chips had been drawn from the bag (and replaced after each drawing), and six of them were red. It would no longer be obvious, at least to a non-statistician, what the likelihoods would be in this case. Intuitively, it is difficult to say what the probability is of drawing six red chips out of ten, if the bag they are drawn from has, for instance, 20% red chips. However, a statistician can calculate these probabilities if we tell him the likelihood for just one red chip (he also has to know that we put each chip back in the bag after it is drawn, and that the red and white chips are identical except for color). The results of his calculations are shown in Column 3 in Figure B-3, and the prior-posterior analysis proceeds as before. Figure B-3 shows the posterior probabilities for the red fraction being 20%, 40%, 80%, 100% as being .03, .54, .43, and 0, respectively.

HYPOTHESIS % RED CHIPS	PRIOR PROBABILITY	LIKELIHOOD	JOINT PROBABILITY	POSTERIOR PROBABILITY
20	0.25	0.000262	0.0000066	0.03
40	0.25	0.0005308	0.0001327	0.54
80	0.25	0.0004194	0.0001048	0.43
100	0.25	0.000	0.0000000	0
	1.00		0.0002441	1.00

Figure B-3

PRIOR-POSTERIOR ANALYSIS WITH LIKELIHOODS CALCULATED FROM SIMPLE ASSESSMENTS

The situation in which likelihoods can be determined easily or routinely tends to occur where carefully controlled samples have been taken, not where evidence "just turns up."

When There Are Many Possible Values of the Target Uncertainty

Although real-world quantities are very rarely strictly continuous (even the GNP must be measured to the nearest cent!), their possible values are frequently very numerous, like the size of an R&D budget or the maximum range of a missile. In such cases, it would obviously be impractical to list all possible values and assign prior probabilities and likelihoods to each. Three common shortcuts are used to avoid this difficulty.

In very special cases, there are mathematical shortcuts. Take, for example, the above "chips-in-a-bag" example, in which six out of ten chips prove to be red. Suppose that instead of only four possible values for the proportion of red chips, there had been a large number of values (101 if there were 100 chips in the bag). If the prior distribution over this large number of values had been described in a special mathematical form, called a "beta" function, the required posterior distribution could have been obtained routinely by using statistical theory.[1] Since this condition is rarely found outside certain types of sample survey situations, we shall not pursue it here.

An approximate graphic solution can often be obtained with a little practice. Equation (3) states that the posterior probability of any particular value is proportional to the product of its prior probability and its likelihood. If the probabilities and likelihoods are plotted as a function of the variable, a graph of the posterior distribution can be sketched so that its height is proportional to the heights of the two input graphs.

To extend an earlier example, suppose that a red chip has been drawn from a bag with possible percentages of red chips ranging from zero to 100, and that the prior probabilities are as indicated in Figure B-4. The likelihood

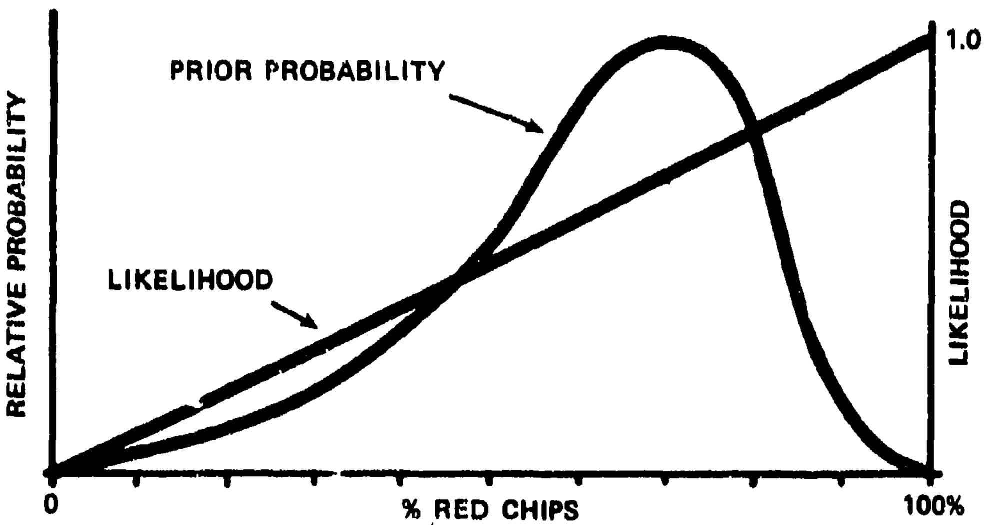


Figure B-4

GRAPHIC REPRESENTATION OF PRIOR PROBABILITY AND LIKELIHOOD

---

[1]In fact, the posterior distribution is a beta distribution whose two parameters are those of the prior beta distribution with the size of the sample added to one parameter and the number of "successes" added to the other. The reader interested in getting more detailed information on this special case and others like it is referred to Howard Raiffa and Robert Schlaifer, Applied Statistical Decision Theory (Cambridge, MA: Harvard University Press, 1961).

function will be a straight, upward-sloping line reflecting the fact that the probability of observing a red chip is exactly equal to the fraction of red chips in the bag, as in Figure B-4. The height of the posterior distribution for a given percentage of red chips is proportional to the product of the heights of the two curves for the same percentage in Figure B-4. A first approximation of the posterior distribution can be drawn by eye, as in Figure B-5.

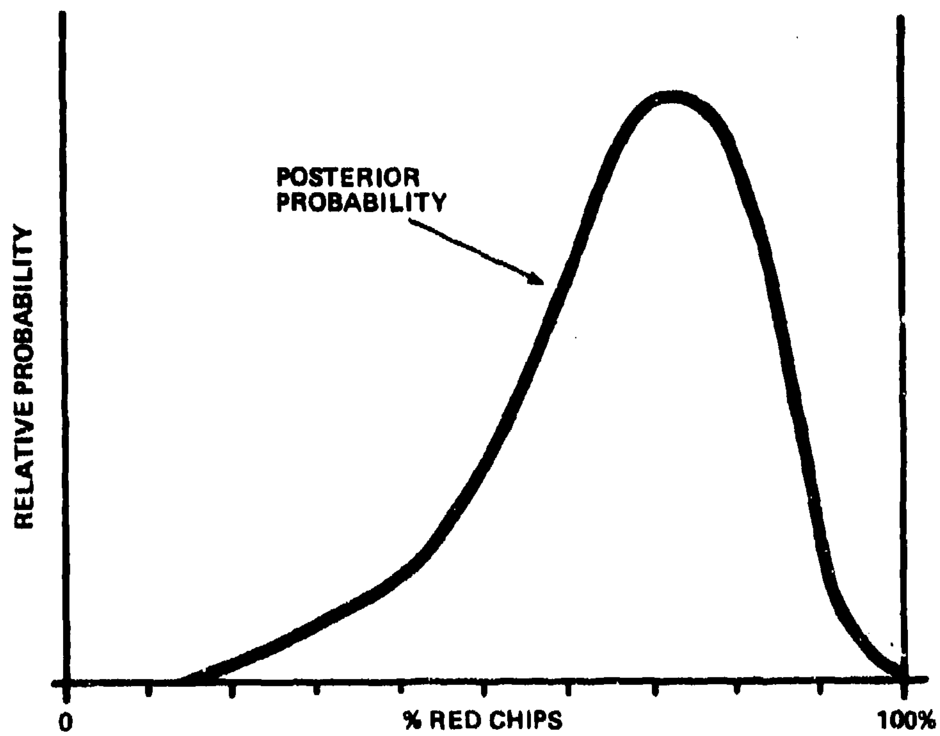


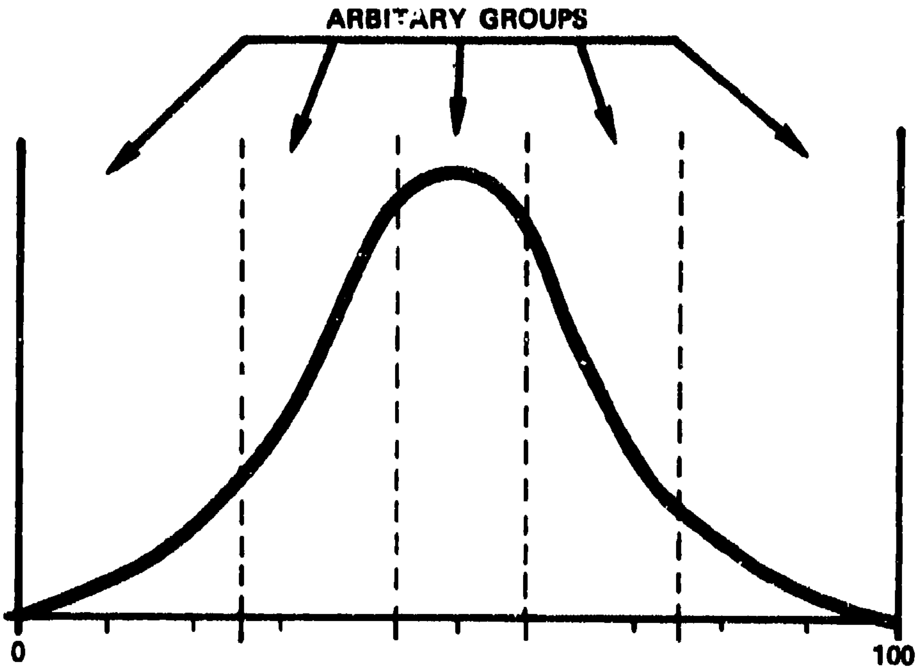
Figure B-5

VISUAL APPROXIMATION OF
POSTERIOR PROBABILITY FROM CURVES OF FIGURE B-4

The third shortcut for prior-posterior analysis on many-valued uncertainties is probably the most widely applicable. Called "grouping," this shortcut is accomplished by substituting a few-valued probability distribution for a many-valued probability distribution. The procedure essentially consists of dividing the original scale of values into any convenient number and size of segments, or "groups," of values and representing each group by a single value.

Figure B-6 illustrates this procedure. The probabilities corresponding to areas under different parts of the

A. Original Many — Valued Prior Probability Distribution



B. Substitute Few — Valued Prior Probability Distribution

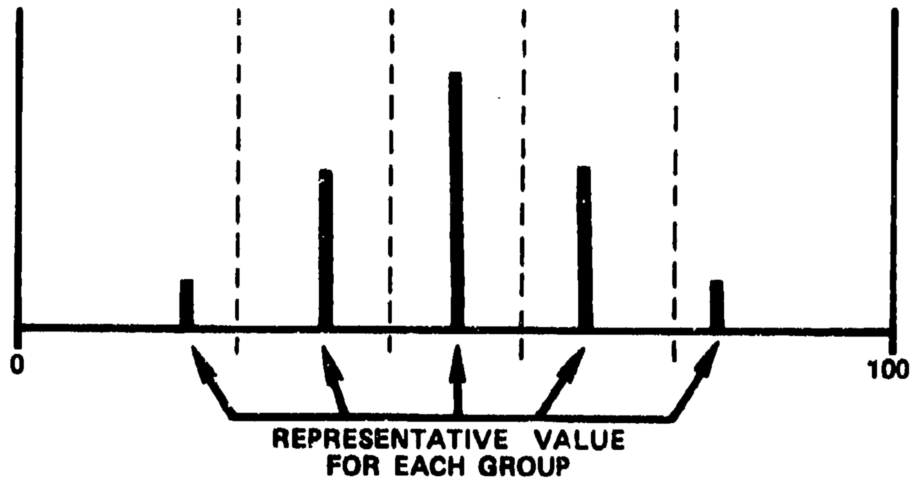


Figure B-6

GROUPING MANY-VALUED PROBABILITY DISTRIBUTION original curve (Figure B-6A) are concentrated at representative values in the substitute few-valued curve (Figure B-6B). If you are satisfied that, for your purposes, the substitute distribution is similar to your real prior distribution, you can proceed as if it were your prior distribution. In the particular example of Figure B-6 where there are only five groups, there would only be five rows in the prior-posterior worksheet, instead of the very large numbers of rows that would have been necessary had the ungrouped prior distribution been used.

Certain methods of effecting this grouping are moderately convenient and reliable. One of them is called the bracket median method. To use this method, you draw the prior probability distribution in the form of a cumulative curve, pick up the .05, 0.15, 0.25, through to 0.95 fractiles, and assign 0.1 probability to each. Figure B-7 illustrates the

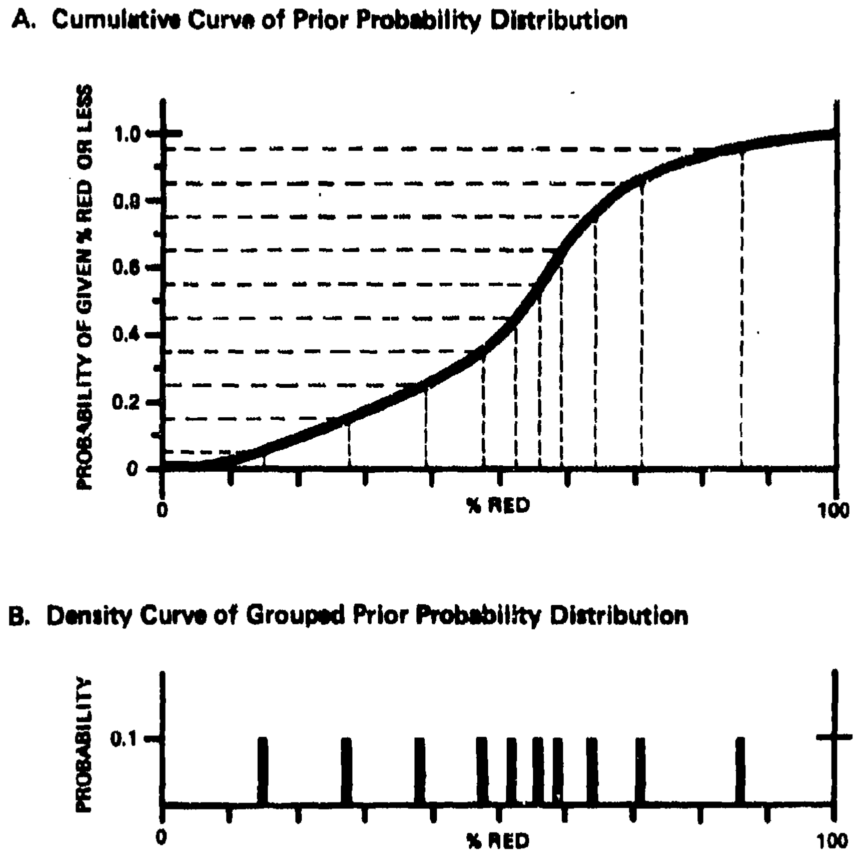


Figure B-7
GROUPING BY BRACKET MEDIAN METHOD

bracket median method in the context of the "red fraction" example used before. The original prior distribution shown in cumulative form in Figure B-7A is reduced to a ten-point probability distribution with equal probabilities in Figure B7-B. Figure B-8 shows the worksheet calculations for a prior-posterior analysis using the grouped prior distribution derived in Figure B-7.

HYPOTHESIS % RED	PRIOR PROBABILITY	LIKELIHOOD	JOINT PROBABILITY	POSTERIOR PROBABILITY	CUMULATIVE POSTERIOR PROBABILITY
15	0.1	0.15	0.015	0.029	0.029
28	0.1	0.28	0.028	0.054	0.083
39	0.1	0.39	0.039	0.076	0.159
46	0.1	0.46	0.046	0.089	0.248
52	0.1	0.52	0.052	0.101	0.349
56	0.1	0.56	0.056	0.109	0.458
59	0.1	0.59	0.059	0.114	0.572
64	0.1	0.64	0.064	0.124	0.696
71	0.1	0.71	0.071	0.138	0.834
86	0.1	0.86	0.086	0.166	1.000
	1.0		0.516	1.000	

Figure B-8

PRIOR-POSTERIOR ANALYSIS
USING GROUPED PRIOR PROBABILITY DISTRIBUTION

Note that the output of this analysis is itself grouped; that is, the posterior distribution has ten possible values in this case. The possible values are the same as those in the grouped prior distribution, but the probabilities are different. A many-valued cumulative curve can be plotted by what, in effect, is the reverse of grouping. The cumulative distribution for the ten-value posterior distribution is given in Column 6 of Figure B-8. These values can be plotted and a smooth curve drawn through them, as in Figure B-9. If all that is needed from the posterior distribution is some summary measure like a mean or a variance, this measure can be calculated directly from the grouped posterior.

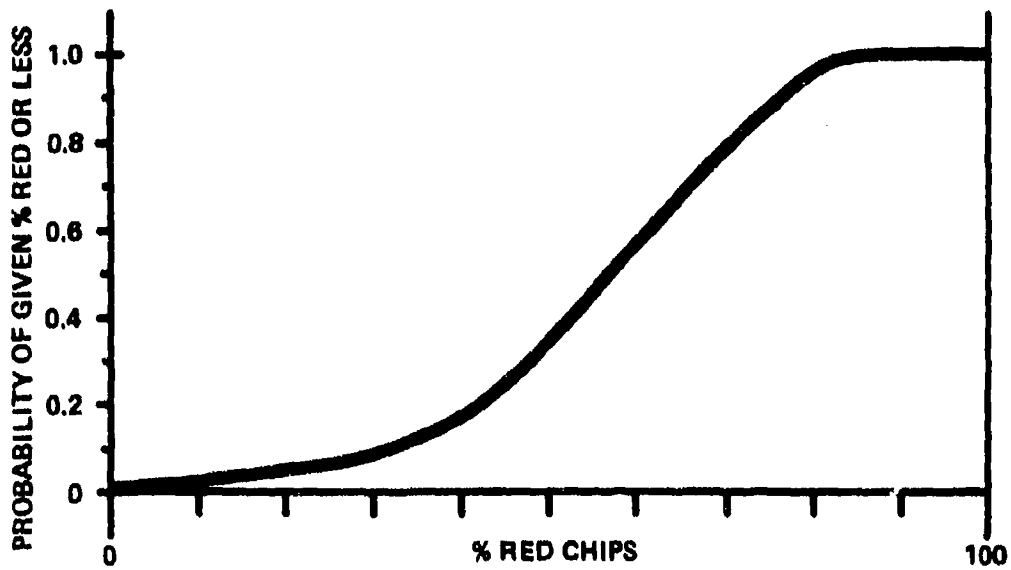


Figure B-9
DEGROUPING THE POSTERIOR DISTRIBUTION

A SCORING RULE FOR PROBABILITY ASSESSMENT

We have argued that probability assessments should be coherent and that they should obey the rules of probability theory. But it is not enough that they obey the rules; they should also be related to reality. A decision maker would soon lose faith in probability assessments if it turned out that the events that occurred were associated with low probabilities and those that did not occur were associated with high probabilities.

The ideal in probability assessments is clairvoyance, that is, the assignment of a probability of 100% to the events that turn out to be true, and 0% to the events that are not true. But clairvoyance is, of course, not possible. Since human beings are unfortunately not blessed with powers of omniscience and future events are only partially accessible to their foresight, the ideal standard of clairvoyance is not obtainable.

Proper Scoring Rules

Clairvoyance is, however, a useful standard against which actual probability assessments can be evaluated. They can be measured against this standard for accuracy (or error) by scoring rules developed for that purpose. Since much of the development of proper scoring rules has occurred in the area of weather forecasting, where scoring rules are used to evaluate precipitation probability forecasts, we shall use the following example of precipitation probabilities to explain the concept of a proper scoring rule. A weather forecaster assesses a precipitation probability between 0% and 100% and he later measures whether or not it rained. He decides that it has rained if his rain gauge gathers more than .01 inch of moisture; otherwise, it has not rained. Given his probability forecast about rain, and later information about whether it does or does not rain, how good is his forecast?

Consider first an intuitively appealing but improper scoring rule, called a linear scoring rule. The first step is to measure the actual distance, in probability units, between a clairvoyant assessment and the actual assessment. For example, if a 60% precipitation probability is assessed, and is followed by rain, then the error score would be 40 points, the distance between 100% and 60%. If, on the other hand, a precipitation probability of 60% is assessed but is not followed by rain, then the error score would be 60 points, the difference between the clairvoyant 0% and the assessed 60%. Figure C-1 displays this scoring rule. The error score is plotted on the vertical axis as a function of the percentage probability assigned to the true event which is displayed on the horizontal axis. Accordingly, if 100% is assigned to the true event, there is an error of 0; if 50% is assigned to the true event, the error score is 50 points; and if 0% is assigned to the true event, the error score is 100 points.

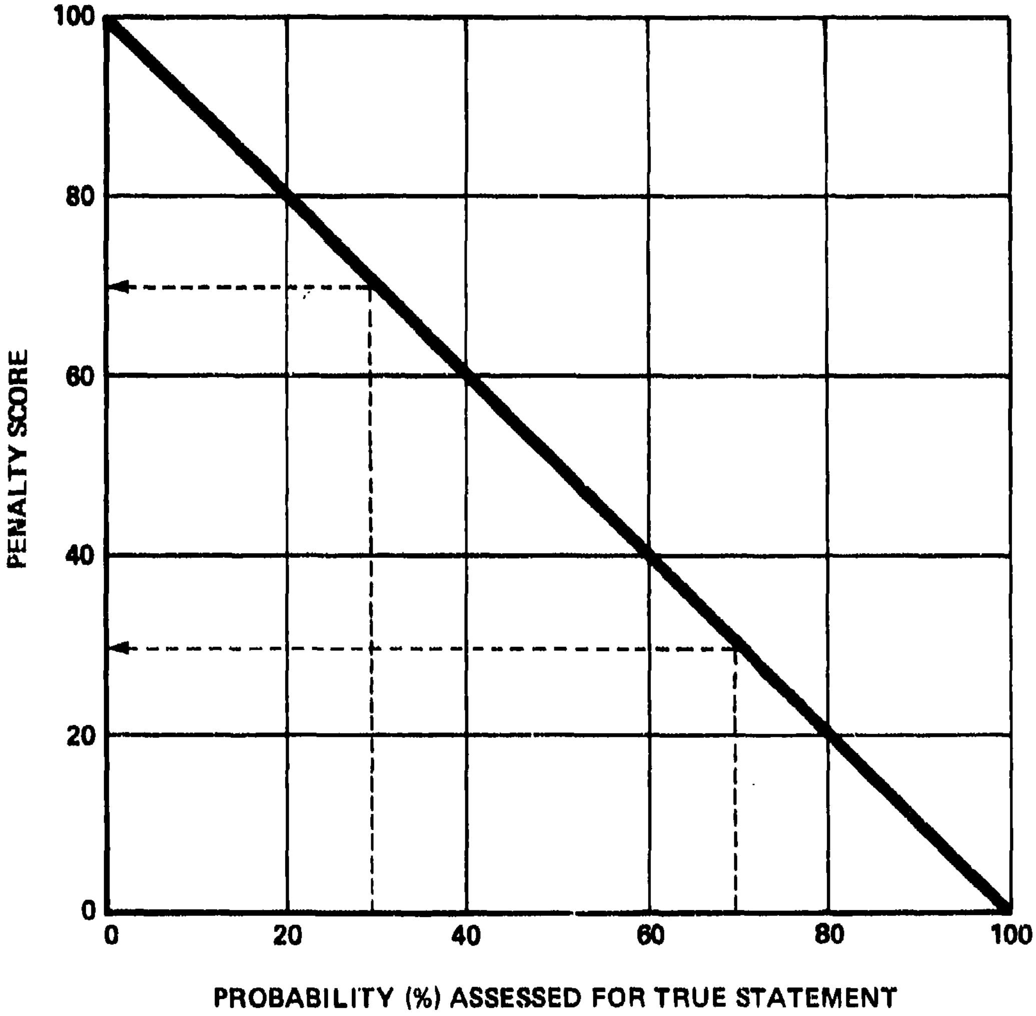


Figure C-1
LINEAR SCORING RULE

In order to understand why this apparently reasonable scoring rule in Figure C-1 is inappropriate, consider the following example. After careful consideration of all relevant information, an expert forecaster assesses a 70% precipitation probability. First, suppose that he tells the truth when reporting his assessment. In other words, he reports a precipitation probability of 70%. What is his expected score for telling the truth? If it rains, he will receive an error score of 30 points since he has assessed a 70% probability that it will rain. If it does not rain, he will receive an error score of 70 points since he has assessed a 30% probability of that event. Therefore, his expected error score is equal to 70% of 30 points plus 30% of 70 points, which is 21 + 21, or 42 points.

Now suppose that our forecaster decides not to tell the truth when reporting his assessment. Specifically, suppose that he reports a precipitation probability of 100%. In this case, what is his expected score? If it rains, he will receive an error score of 0 points. If it does not rain, he will receive an error score of 100 points. His expected score is equal to 70% of 0 points plus 30% of 100 points, which is 0 plus 30, or 30 points. Since the error of 30 points is a much smaller penalty score than is 42 points, our weather forecaster is well advised not to tell the truth but rather to report a 100% precipitation probability.

Clearly, this is an unsatisfactory state of affairs. A scoring rule, to be useful in evaluating probability assessments, should have the property that the probability assessor not only need not "play games" in order to do well, but also is rewarded for telling the truth. A class of scoring rules, called proper scoring rules, has been developed whereby the probability assessor can minimize his expected penalty score only by reporting his true assessed probability.

The Brier score, which is used for weather forecasting, is an example of a proper scoring rule. In the two-event case, of which the precipitation problem is an example, the Brier rule assigns an error score which is equal to .01 times the square of the distance of the assessed probability from clairvoyance. It is calculated by finding the true event and then squaring the distance of the probability assessed for that event from 100% or 0%, depending on whether the event confirms or refutes the assessment. For example, consider an assessed precipitation probability of 60%. If it rains, the assessment is assigned an error score of $.01 \times (100-60)^2$, or 16 points. If it does not rain, the assessment is assigned an error score of $.01 \times (60-0)^2$, or 36 points. The proper Brier score is thus a very simple modification of the improper linear scoring rule.

Figure C-2 displays the Brier rule. The error score is shown on the vertical axis as a function of the percentage probability assigned to the true event, which is shown on the horizontal axis. Now apply the Brier score to the weather forecaster whose assessed precipitation probability is 70%. Suppose he reports this assessment as it stands. If it rains, he will receive an error score of 9; if it does not rain, he will receive an error score of 49. His expected score, then, is 70% of 9 (6.3) plus 30% of 49 (14.7), or 21 points. Next, consider the case in which he "plays games" and reports a precipitation probability of 100%. If it rains, he will receive an error score of 0; if

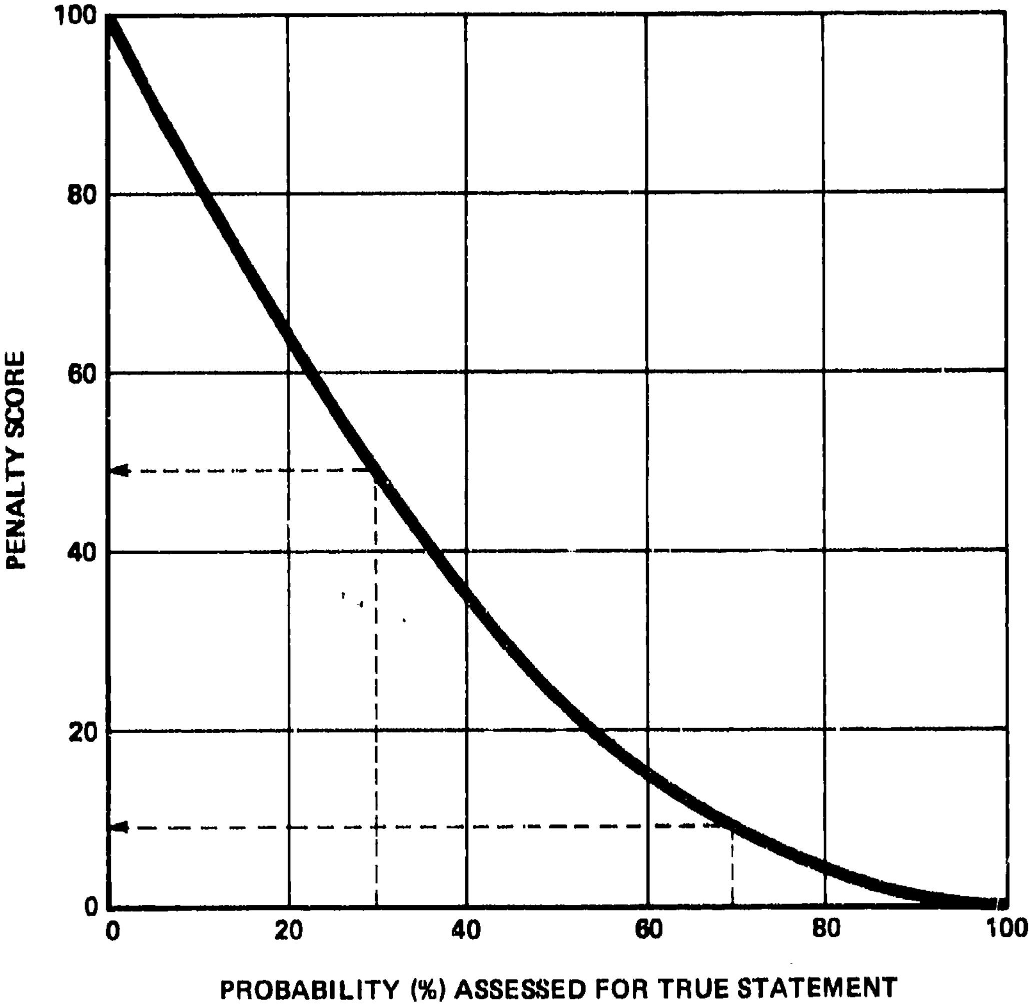


Figure C-2

BRIER SCORING RULE

it does not rain, he will receive an error score of 100 points. Against the standard of a truthful report of his true assessment of the probabilities, his expected error score is 70% of 0 (0) plus 30% of 100 (30), or 30 points. Since an error score of 30 is substantially greater than a score of 21, the forecaster is well advised to report truly his best assessment of his probability of precipitation. Indeed, it can be shown mathematically that a weather forecaster can earn the highest expected Brier score only by reporting his true precipitation probability.

Sorting Errors Versus Labeling Errors

Even the most expert forecaster, however truthful his reports of assessed probability are, will be wrong some of the time. His assessments have no bearing on the outcomes which may confirm or refute them. However, it is possible to help the forecaster make statements of assessed probability which have the same probability of being confirmed as he assesses the probabilities of the outcomes. In short, 70% of all of his statements assessing a probability of 70% ought to be confirmed.

If not, the forecaster is making one of two kinds of errors. These are a sorting error or a labeling error. We may think of assessing probabilities as a two-step procedure. First, we consider whether a certain event will happen or not; then, to what degree we have confidence in that determination. (The manner in which we computed a proper scoring rule depends on a comparable procedure, first, what happens in the outcome, then a computation of a score based on the differences between the probability and the certainties.) A sorting error, then, is deciding wrongly that an event will or will not happen; a labeling error, an improper degree of confidence in that decision.

The following example illustrates these two error components. Intelligence analysts in a Washington intelligence agency made weekly forecasts of many different kinds of events such as whether a military coup would occur within a particular time interval, whether a reconnaissance plane would be shot down, or whether an arms shipment would be made to a particular country within a specified time interval. In each case, it was possible to determine some time after the forecast whether or not the event in question occurred, that is, whether or not the statement for which the probability was assessed turned out to be true.

The probability assessments were evaluated in the following manner. First, the assessments were categorized into common probabilities. Thus, all assessments of 70% were placed into one category, assessments of 40% were placed into another category, assessments of 10% into yet another category, and so for all different probability assessments that the analysts used. The goal of this analysis was to calculate the percentage of true statements, known as the hit rate, associated with each category. In an effort to obtain dependable hit rates, adjacent categories were combined so that the combined categories contained approximately 100 different assessments. Then the percentage of true statements, or the hit rate, was calculated within each category by dividing the number of true statements by the total number of statements in that category. Figure C-3 displays the results of this analysis. The

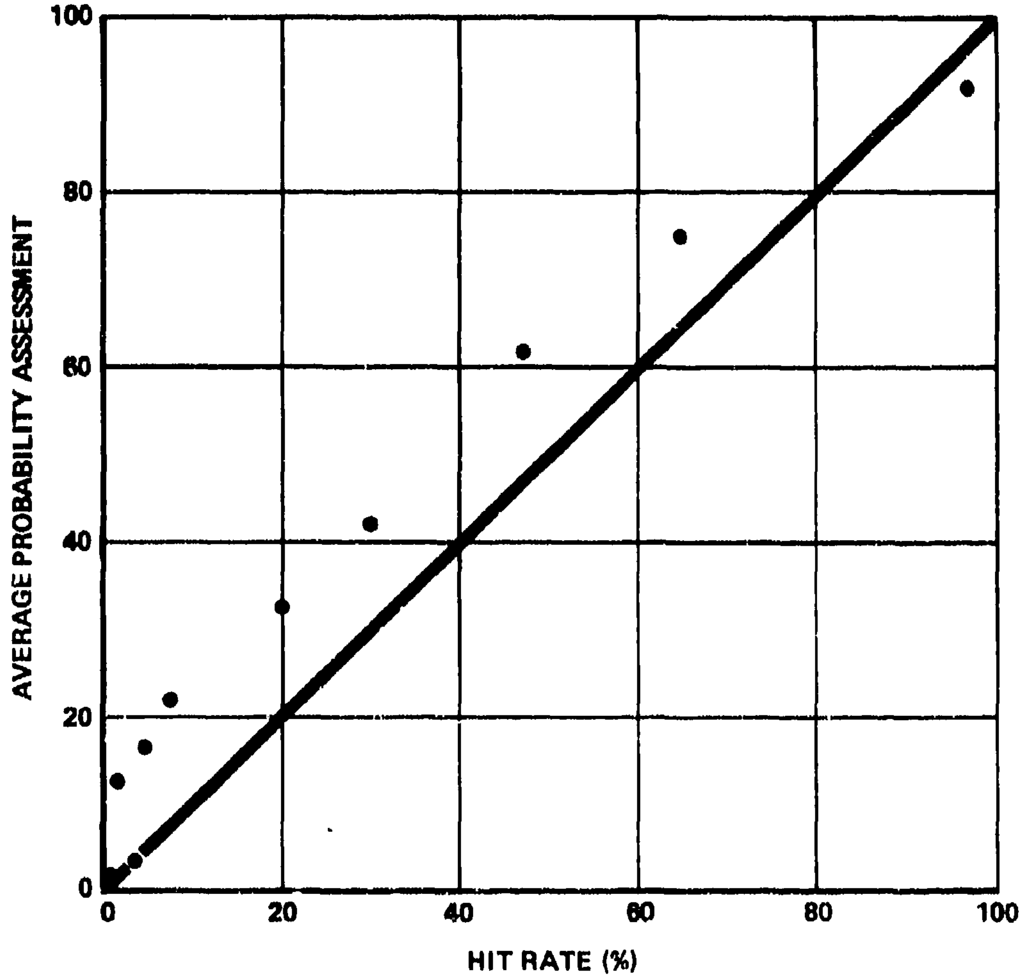


Figure C-3

vertical axis refers to the average assessed probability for each category and the horizontal axis refers to the corresponding hit rate in the category. These results seem to indicate good performance on the part of the analysts. As the average assessed probability increased, there was an accompanying increase in the percentage of correct statements. An exact measure of performance can be obtained by using the Brier rule to calculate the average error score for all assessments that contributed to each data point in Figure C-3.

Since the data points do not all fall on the line indicating a perfect correlation between the assessed probabilities and the percentage of true statements, one or both kinds of errors are affecting the hit rate. Let us take an example that measures the contribution each makes to the error score.

Consider a hypothetical data point at which the average assessed probability is equal to 70% and the hit rate is equal to 60%. The sorting error is a measure of the degree to which the analyst failed according to the standard of clairvoyance to sort perfectly all true statements from all false statements. Specifically, the sorting component of the average error for a data point is equal to (P)(1-P)/100, where P is equal to the hit rate. In this case, the average error attributable to sorting is (60 x 40)/100, or 24. Figure C-4 illustrates the manner in which the sorting component of the error changes with P. The sorting component is at 0 for perfect sorting, when P is equal to either 0% or 100%. It is at a maximum, 25 points, when the hit rate is equal to 50%. Intuitively, the sorting error can be thought of as a measure of the degree to which the probability assessor is knowledgeable about the subject matter. To the degree that he is knowledgeable and that his probability assessments serve to sort the statements into almost all true statements and almost all false statements, the sorting component of his error score will be low. If he is not very knowledgeable, and his probability assessments do not discriminate very well between true and false statements, then the sorting component of his error score will be high.

The second component of the error score results from errors in the labeling process. Having decided which event is likely to occur, that is, having decided how much he knows about the likelihood of a certain outcome, he then determines his degree of confidence in his decision. Ideally, the label should be equal to the hit rate. For example, if a weather forecaster is doing a good job, then he should use the probability label in such a manner that it will turn out to rain on about 70% of the days for which he issues a 70% precipitation probability; and it should turn out to rain on about 20% of the days for which he has assessed a 20% probability. In general, the assessed probabilities should be equal to the corresponding hit rates. To the degree that the data points are either above or below the identity line, we infer that the probability assessor has erred in the process of labeling. In terms of Figure C-3, an error-free labeling process would imply that the data points would lie along the 45-degree line.

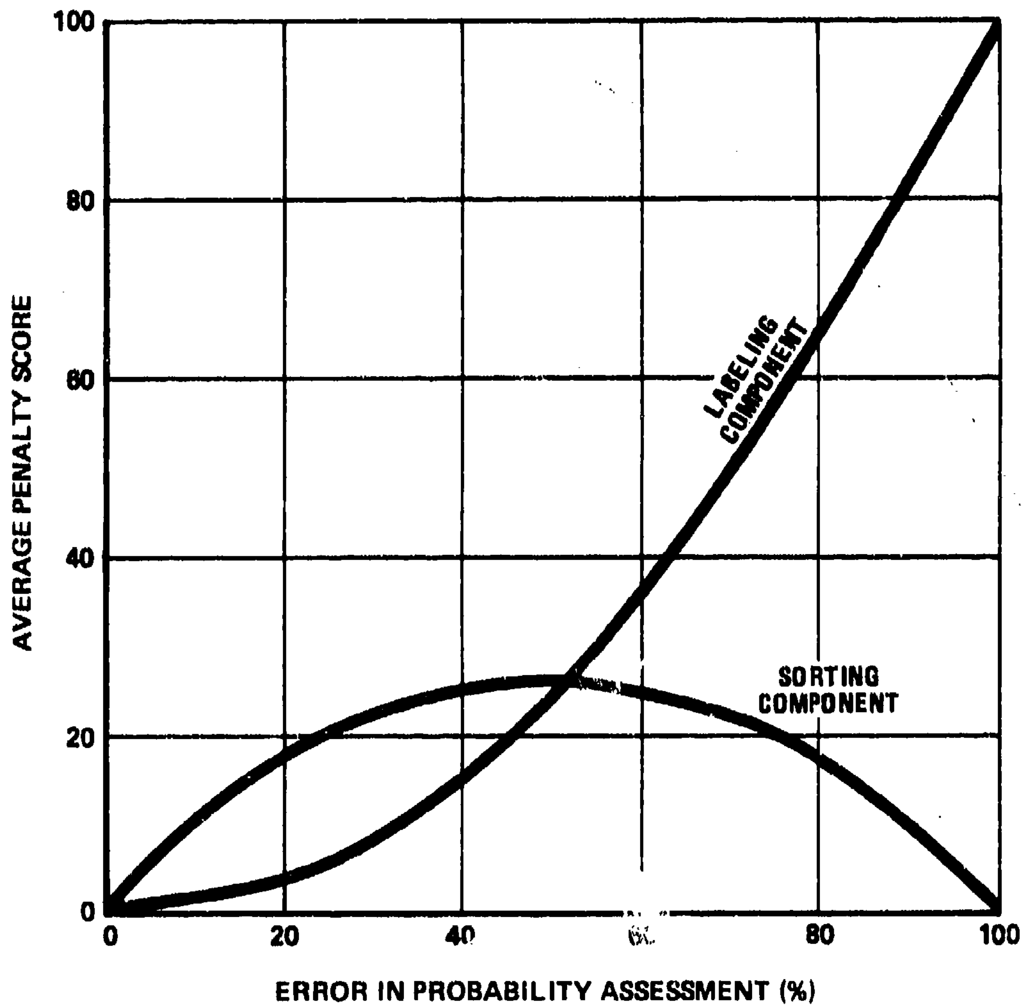


Figure C-4

RELATIVE CONTRIBUTION OF COMPONENT ERRORS TO ERROR SCORE

The exact contribution of the error in labeling to the error score is equal to the square of the difference between the assessed probability and the hit rate, divided by 100. Accordingly, in our example in which the assessed probability is 70% and the hit rate is 60%, the labeling component of the error score is equal to the square of the difference between 70 and 60, all divided by 100: it is only 1. Figure C-4 shows the manner in which the labeling component contributes to the error score as the distance between the assessed probability and the hit rate increases. The labeling component contributes 0 points when the assessed probability is equal to the hit rate, and it contributes 100 points when there are 100 percentage points between the two.

These calculations for our example imply that the average error for all assessments is equal to 25 points; 24 are attributable to sorting error, and 1 is attributable to labeling error. Thus, for that data point, 96% (.24/.25) of the error can be attributed to improper sorting whereas only 4% (.01/ .25) can be attributed to mislabeling. Thus, this particular error in labeling is relatively unimportant as compared with the inability to sort appropriately or discriminate between true and false events. The relative degree of importance will, of course, vary with the data point analyzed. Data points toward the extreme, where the hit rate is near 0% or 100%, will have relatively small amounts of sorting error where data points in the center will have large amounts. The degree of labeling error depends only upon the vertical distance of a data point from a straight line. The degree of labeling error reflected in Figure C-3 is, for the most part, relatively small because few data points fall very far from the straight line.

Notice that a large labeling error, say, above 50%, will generally occur only when a probability assessor is being deliberately misleading. That is, it will occur when a probability assessor assigns high probabilities to statements that usually turn out to be false and low probabilities to statements that usually turn out to be true. Even if the probability assessor is uncommonly benighted, this labeling error can be extreme only when the hit rate is extreme and thus permits a substantial degree of error.

Sorting errors can be reduced primarily by improving the analytic process, either by making more information available or by providing a better means for processing the available information. Labeling errors, on the other hand, can be reduced by improving the process by which the probability assessor assigns numbers to reflect his degree of knowledge. A better assignment of numbers requires an intuitive appreciation for the meaning of a quantitative probability scale.

A computer-based scoring rule training procedure has been developed to calibrate and improve the accuracy of probability assessors. In application, the computer poses a series of multiple-choice questions to the trainee. The trainee is required to indicate the correct answer along with a probabilistic assessment of his degree of certainty about the designated answer. Automatic feedback as to the accuracy of the trainee's response is provided, and the computer maintains a running calculation of the implication of cumulative response accuracy and uncertainty levels to determine and display the degree of sorting error and labeling error in the trainee's performance. In experimental situations, use of this computer-based procedure has yielded gains in assessment accuracy of from 10 to 40%.

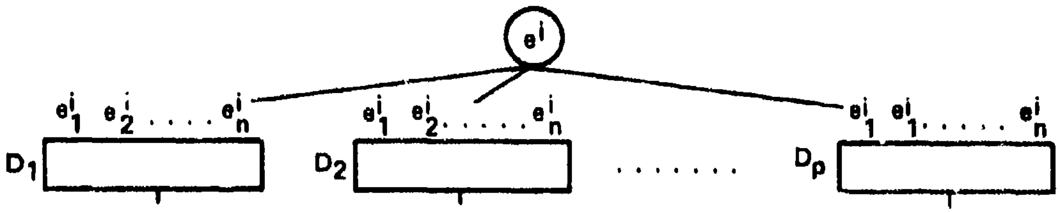
APPENDIX D

PROBLEMS IN HIERARCHICAL INFERENCE AND A CASE STUDY

In this appendix we first discuss the general principles to be followed in making the calculations whose results appear on the inductive structure of a hierarchical inference problem. Next, we mention an error that is easily made when modifying the calculations upon receipt of a new datum. Finally, we explain in some detail a case study in hierarchical inference.

Principles

In the process of proceeding inductively up through the hierarchical structure for the nuclear weapons development program discussed in Chapter 4, the analyst was confronted with one of two situations at each branch point. The first situation occurs whenever it is necessary to determine the likelihood of two or more data for a given indicator or activity, that is, whenever two or more data branches merge on the diagram of the hierarchical structure. A diagram illustrating this situation and the general equation (1) used in the solution is shown below:



$$P[D \mid e^i_k] = P[D_1 D_2 \ldots\ldots D_p \mid e^i_k] = \prod_{l=1}^{p} P(D_l \mid e^i_k) \qquad (1)$$

FOR $k = 1, 2 \ldots\ldots n$

This equation says that the probability of all the data given each state of the variable is equal to the product of the probability for each datum for a given state, provided that the data are independent. This equation is an application of the multiplication rule given in the section "Rules for Combining Probabilities," in Chapter 4.

For example, this situation occurred in the nuclear development problem when the analyst computed likelihoods for the combination of datum 5 (D_5) and datum 6 (D_6), on the assumptions, first, that indicator 2 has occurred (I_2), then that it has not occurred ($\bar{I}_2$). (Figures 4-35 and 4-40 refer.) Using equation 1, he proceeded as follows:

$$P(D|I_2) = P(D_5|I_2)\ P(D_6|I_2)$$

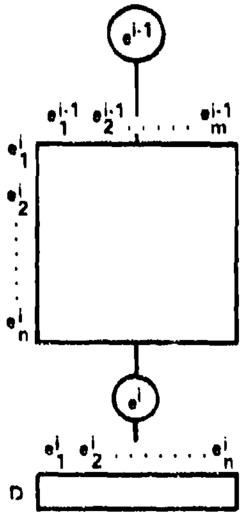
$$= (0.99)\ (0.75) = 0.74$$

$$P(D|\bar{I}_2) = P(D_5|\bar{I}_2)\ P(D_6|\bar{I}_2)$$

$$= (0.50)\ (0.70) = 0.35\ .$$

D refers, in this case, to the combination of D_5 and D_6. These computed likelihoods, 0.74 and 0.35, can be seen in their appropriate places on Figure D-1.

The second situation occurs whenever there is no merging of branches and it is necessary to determine the likelihood of the data given the activity or indicator at the next higher level. A diagram of this situation along with the general equation (2) used in the solution is shown below:



$$P(D \mid e_j^{i-1}) = \sum_{k=1}^{n} P(D \mid e_k^i)\ P(e_k^i \mid e_j^{i-1}) \qquad (2)$$

FOR $j = 1, 2, \ldots, m$

This equation, which applies the addition rule for probabilities given in Chapter 4, says that the probability of the data occurring for a given state of the upper-level event is found by multiplying each probability in the lower box by each probability in the first column of the upper box (i.e., the left-most number in the row by the top number in the column, the second from left in the row by the second from top in the column etc.) and by summing those products, and then by repeating this procedure for each remaining column.

For example, this situation occurred in the current problem just after the analyst computed the likelihoods given on the previous page, at the point where he wanted to calculate the likelihoods for the combined effects of D_5 and D_6 with activity 1 as the conditioning event rather than, as previously, indicator 2 as the conditioning event. Using the previously-computed likelihoods along with the assessed probabilities relating activity 1 to indicator 2, he applied equation 2 as follows:

$$P(D|A_1) = P(D|I_2)\ P(I_2|A_1) + P(D|\bar{I}_2)\ P(\bar{I}_2|A_1)$$

$$= (0.74)(0.80) + (0.35)(0.20)$$

$$= 0.59 + 0.07 = 0.66$$

$$P(D|\bar{A}_1) = P(D|I_2)\ P(I_2|\bar{A}_1) + P(D|\bar{I}_2)\ P(\bar{I}_2|\bar{A}_1)$$

$$= (0.74)(0.60) + (0.35)(0.40)$$

$$= 0.45 + 0.14 = 0.59 \quad .$$

These likelihoods, 0.66 and 0.59, are also shown on Figure D-1.

(At this point, the reader is encouraged to examine both the deductive and inductive structures, to determine which rule applies to the generation of the remaining row vectors, and to perform the indicated calculations, in order to reinforce understanding of the computational procedure.)

Suppose the analyst is now given a new datum, D_8, equally as diagnostic about increased scientific activity as datum 7, decrease in publications on high-explosives research. If the analyst is careless or does not understand how to proceed, he may attempt to incorporate this datum in the following way: To accommodate the new datum, he adds to the diagram of Figure D-1 a fifth branch from the top box. Since the impact of the new datum is the same as datum 7, he simply duplicates on the new branch all the boxes and their associated numbers shown in the extreme right branch of Figure D-1. Recall that the likelihood ratio of 4.48 shown in the top box is calculated by multiplying together the four likelihood ratios on the left sides of the four boxes

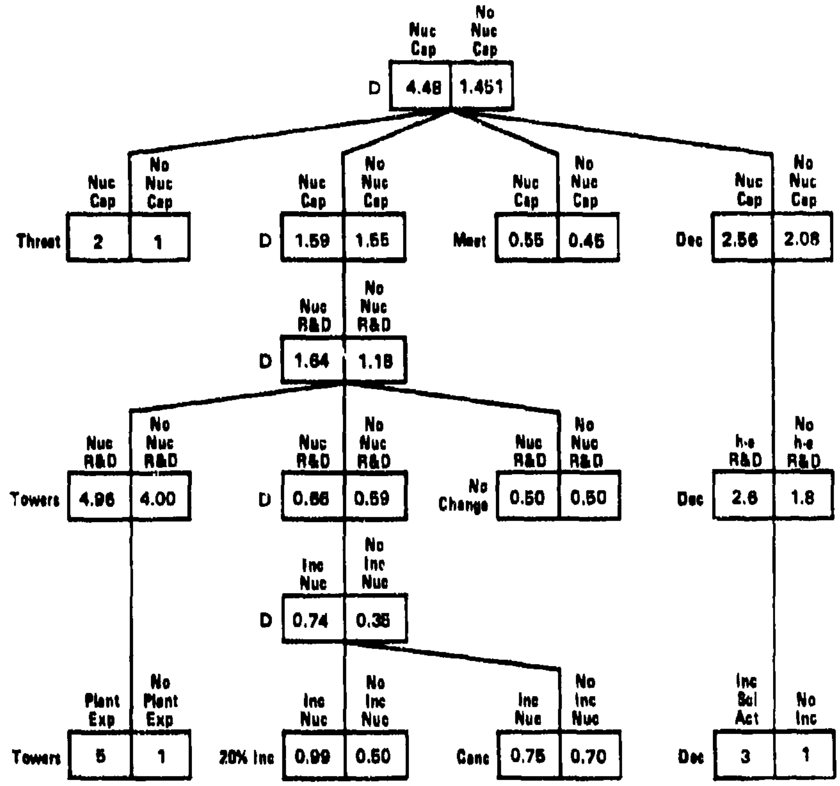


Figure D-1

INDUCTIVE HIERARCHICAL STRUCTURE FOR NUCLEAR WEAPONS DEVELOPMENT PROGRAM

on the next level below. Thus, to take account of the new datum, another 2.56 must be included in the product. This same line of reasoning applies, of course, to the calculation of a new likelihood ratio for no nuclear capability, the right figure in the top box. So, the analyst combines the likelihood for the new datum with the previous likelihood at the top of Figure D-1:

$$\begin{array}{cc} \text{nuc cap} & \text{no nuc cap} \\ D_8 \; \boxed{2.56} & \boxed{2.08} \end{array} \times \begin{array}{cc} \text{nuc cap} & \text{no nuc cap} \\ D \; \boxed{4.48} & \boxed{1.45} \end{array} = \begin{array}{cc} \text{nuc cap} & \text{no nuc cap} \\ (D_8 \,\&\, D) \; \boxed{11.5} & \boxed{3.02} \end{array}$$

$$= D' \; \boxed{3.81} \; \boxed{1}$$

But is this the correct inference, given the additional datum?

It is easy to show that this is an incorrect inference. Assume, for the moment, that some datum which tells us for certain that increased scientific activity is in progress has been obtained, that is, that I_3 has definitely occurred. Then the likelihood of this datum, given that a high-explosive R&D program either is or is not in progress, is as shown in the top row of the middle box in the right branch of Figure D-2. It is 0.8 if the h-e R&D program is underway, 0.4 otherwise.

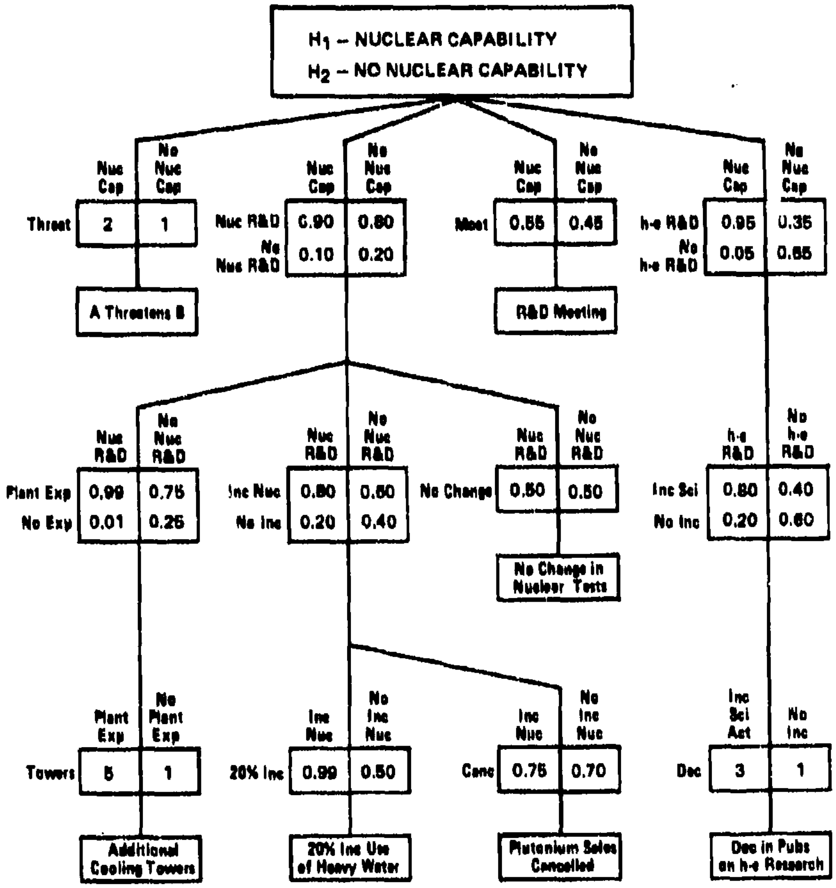


Figure D-2

DEDUCTIVE HIERARCHICAL STRUCTURE
FOR NUCLEAR WEAPONS DEVELOPMENT PROGRAM

Carrying out calculations as before, but with this modification of the middle box, gives the new values shown in Figure D-3. Note that the likelihood at the top is 3.62 to 1, based on the assumption of perfect information for the right branch. But in adding a new piece of imperfect information, the analyst arrived at a likelihood of 3.81 to 1. Clearly, additional imperfect information in one branch should not make us more certain than when perfect information is received for that same branch. So, the procedure he adopted in arriving at the figure of 3.81 to 1 must be incorrect. His mistake was to add an entire fifth branch when the new datum was received. This mistake is equivalent to an assumption that the seventh and eighth data were independent with respect to the hypotheses H_1 and H_2. In fact, they are independent only with respect to the indicator I_3. Thus, the new datum must be shown entering the diagram just below I_3.

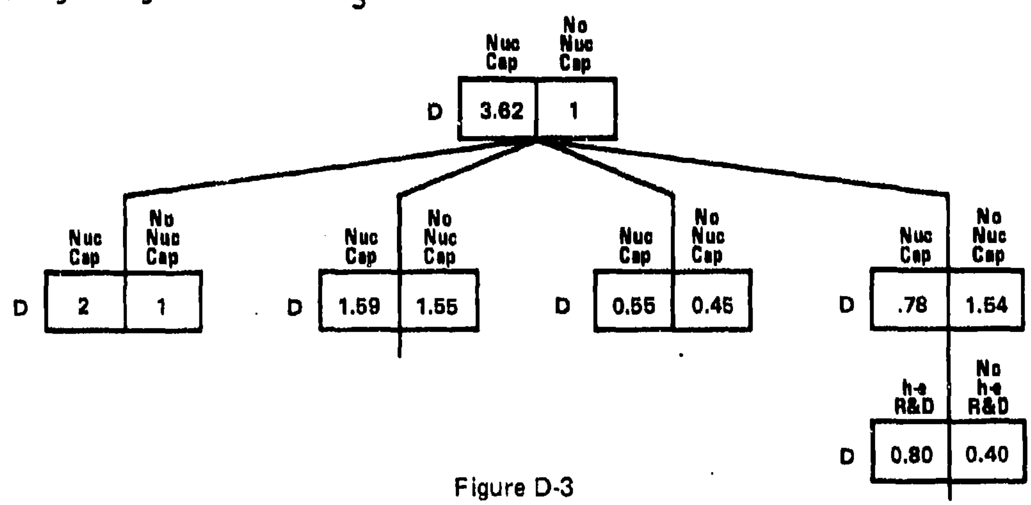


Figure D-3

INFERENCE STRUCTURE ASSUMING
PERFECT INFORMATION ABOUT INDICATOR I_{41}

Figure D-4 shows the inference structure which correctly relates D_8 to the hierarchy, and the correct likelihood of D given H.

Figure D-4 demonstrates the following rule: new information about indicators, activities, or hypotheses must be aggregated with the existing information at the point at which they enter the hierarchical structure.

The Generic Hierarchical Structure

We would like to present now a complete case study of hierarchical inference as a technique for the solution of complex inference problems, but first it will be useful to introduce notational conventions that will apply throughout

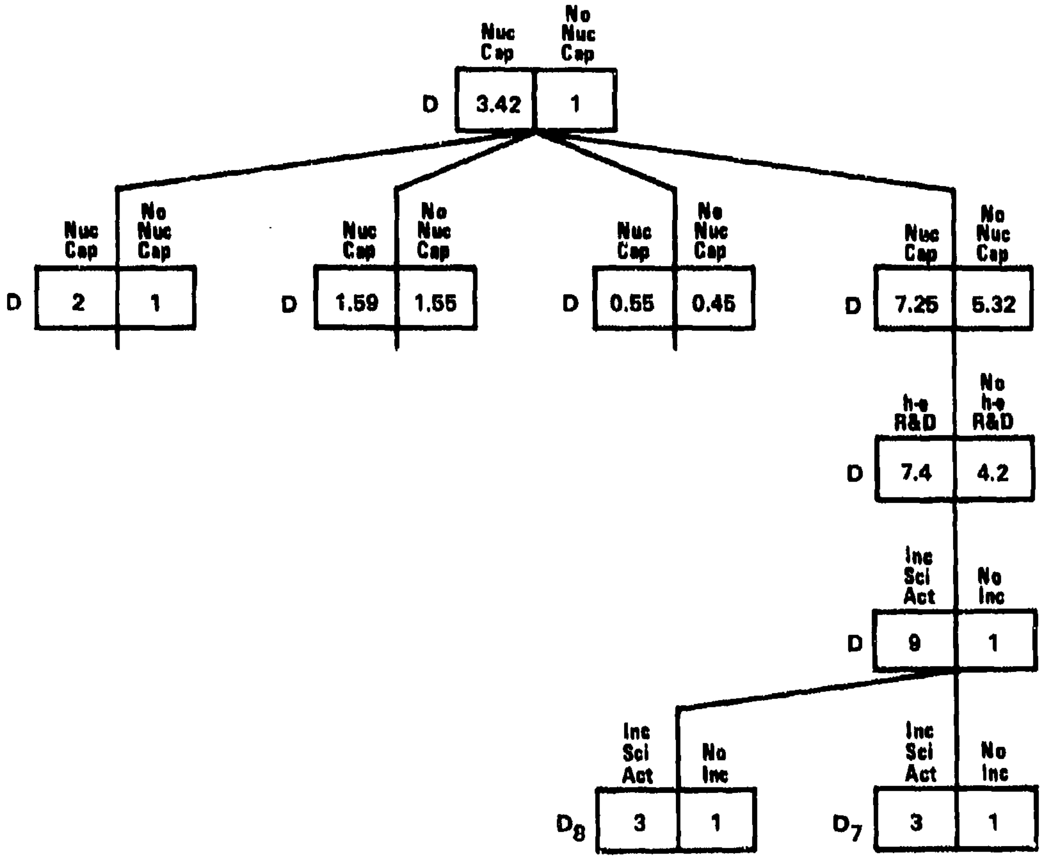


Figure D-4

INFERENCE STRUCTURE GIVEN DATUM D^{412}

the remainder of this appendix. Figure D-5 shows the structure formulated for a hypothetical problem, which contains four levels.

The first level contains the exhaustive set of mutually exclusive hypotheses, H_i, required to describe the central issue.

The second level contains the data, D^i, and the activities, A^i, which impact directly on the hypotheses. These elements are numbered with a single superscript, in consecutive order from left to right on the diagram. Each category of activity, A^i, may be described by a number of mutually exclusive and exhaustive states A^i_j which are diagnostic of the hypotheses. For example, activity A^2 might have three states, which would be indicated A^2_1, A^2_2, and A^2_3.

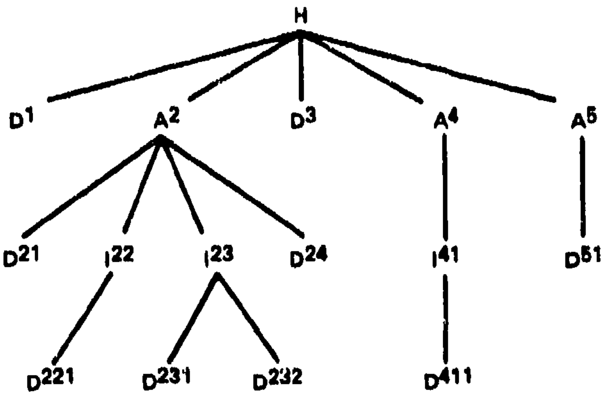


Figure D-5
ILLUSTRATIVE HIERARCHICAL STRUCTURE

The third level of the hierarchy contains those data and indicators which impact on the hypotheses through activities in the second level. The first superscript is their consecutive order from left to right on the diagram. Each indicator may consist of a number of mutually exclusive and exhaustive states which are diagnostic of the related activity. These states are indicated by consecutive subscripts (I_1^{23}, I_2^{23}, ...).

The fourth level of the hierarchy contains those data which impact on the hypotheses through indicators in the third level and activities in the second level. Each datum is identified by the superscript of the indicators to which it is related, with the third superscript being the consecutive order of occurrence, in relation to that indicator, from left to right in the diagram.

Complete Case Study

Now we turn to the case study. We shall use the methodology described in Chapter 4, with the technical refinements described earlier in this appendix.

Imagine an analyst who has been given the assignment of determining if research which could lead to the capability of detecting a moving submarine by infrared or radar observation of the surface effects is being conducted by Country B.

This analysis would be important because the development of such a capability would allow the use of high-altitude aircraft in detecting and tracking submarines and would eliminate the need for the sensing vehicle to be in close proximity to a submarine. The conventional, localized method of detection using sonar would be replaced by a more wide-ranging system.

Hierarchy Construction

The first step in the analysis was to develop a set of mutually exclusive and exhaustive hypotheses. The analyst felt that if such a program existed in Country B, it might take one of two possible forms. It could be a pure research program aimed at obtaining an understanding of the physical processes involved, or it could be a program aimed at observing and measuring possible surface effects. The third possibility to be considered is that there is no program. He therefore developed the following hypotheses:

H_1. There is a research program to investigate surface effects of a submerged submarine for potential military use. The emphasis is on obtaining a thorough understanding of the physical processes involved in the generation of surface effects.

H_2. There is a research program to investigate surface effects of a submerged submarine for potential military use. The emphasis is on observation and measurement of the surface effects, with little attention given to the physical processes involved.

H_3. There is no program to investigate surface effects of a submerged submarine.

The analyst next developed a list of six R&D program activities which would be likely to occur if the hypothesized research programs existed. These activities would be necessary to produce solutions to critical technological problems and would probably produce observable data indicating their presence. These six activities, comprising the second level of the hierarchy, are:

A^1. Wake Formation

A^2. Wake Observables

A^3. Ambient Ocean Conditions

A^4. Infrared Radiometry of the Ocean Surface

A^5. Radar Scattering by Surface Waves

A^6. Pattern Distortion by Ocean Waves

From his own knowledge of the problem and with the assistance of experts in the technical areas encompassed by the six activities, the analyst was able to develop a number of indicators which would be diagnostic about the existence of the R&D activities. These the analyst related to the activities in the following manner:

A^1. Wake Formation
- I^{11} - Turbulent Wake
- I^{12} - Fluid Displacement
- I^{13} - Wake Measurement
- I^{14} - Internal Waves

A^2. Wake Observation
- I^{21} - Modeling
- I^{22} - Surface Roughness
- I^{23} - Slick Formation
- I^{24} - Temperature Measurement

A^3. Ambient Conditions
- I^{31} - Wave Generation
- I^{32} - Heat Transport
- I^{33} - Temperature Fluctuations
- I^{34} - Cloud Reflections
- I^{35} - Surface Slicks

A^4. Infrared Detection
- I^{41} - Detector Development
- I^{42} - IR Systems
- I^{43} - Atmospheric Effects
- I^{44} - Signal Processing

A^5. Radar Detection
- I^{51} - Theoretical Models
- I^{52} - Differential Roughness
- I^{53} - Radar Measurements

A^6. Pattern Distortion
- I^{61} - Weather Conditions
- I^{62} - Wind Conditions
- I^{63} - Wave Conditions

At this point the analyst had sufficiently decomposed the problem so that he could lay out the hierarchical structure shown in Figure D-6, which linked all the available data to the top-level hypotheses. Note that two activities, A^2 and A^5, are not in this diagram (Figure D-6), indicating that there were no available data related to these activities.

It is evident from Figure D-6 that most of the information in the observed data is related to the hypotheses under analysis indirectly through a chain of indicators and activities. As an example, datum D^{111} is the information content of a paper published recently by Dr. X. This datum is diagnostic in discriminating between the two hypothesized

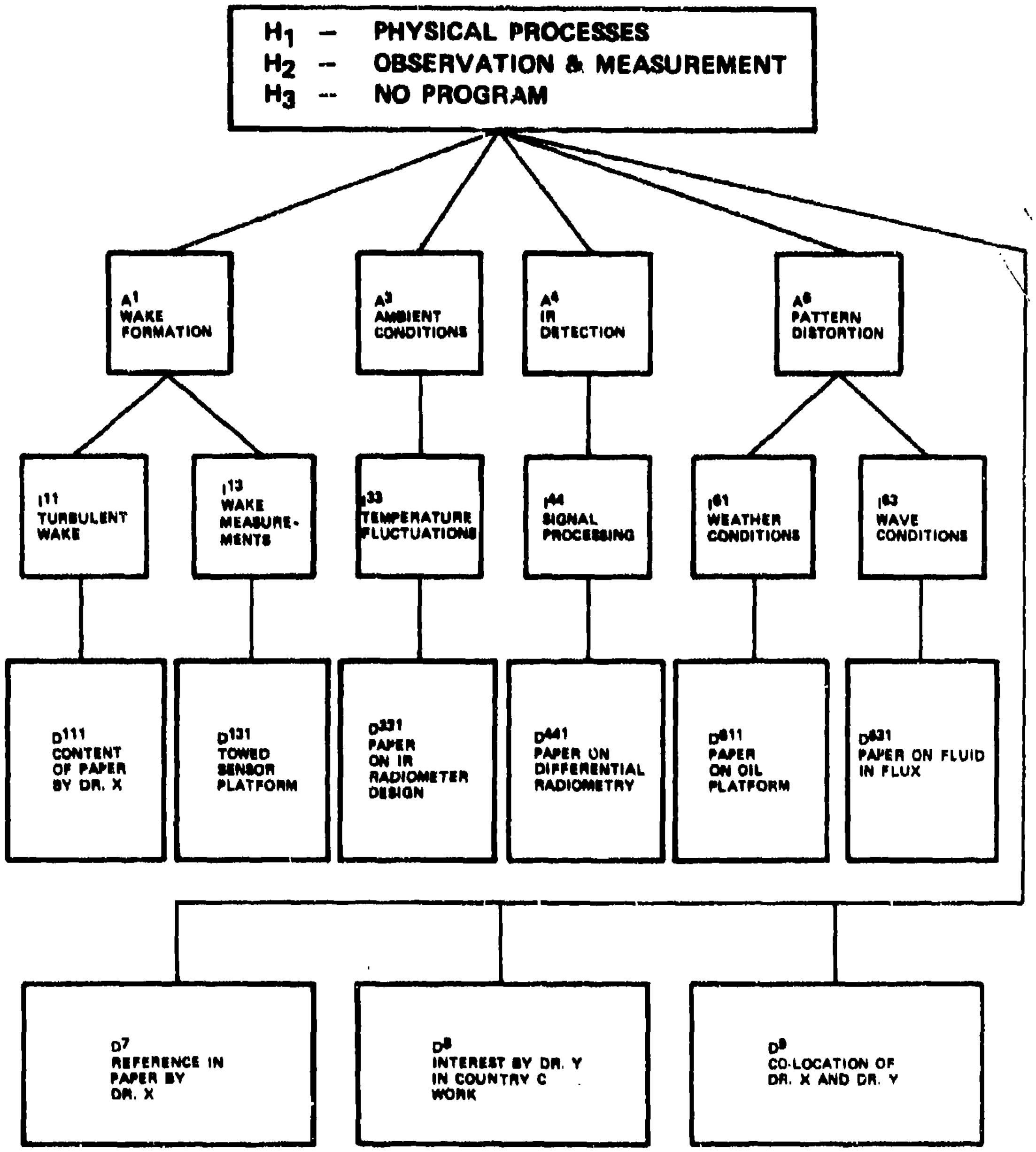


Figure D-6

HIERARCHICAL STRUCTURE FOR
SUBMARINE DETECTION PROGRAM, COUNTRY B

states of indicator I^{11}, which is in turn diagnostic between the states of A^1, and therefore diagnostic among the hypotheses. It is important to note that the impact of data, linked to the hypotheses through the intermediate variables, will be filtered by each level through which they pass. For this eason the intermediate variables must be defined carefully, and the probabilities linking the variables must be assessed as accurately as possible. On the other hand, data D^7, D^8, and D^9 are linked directly to the hypotheses.

Assessments and Mathematical Solution

The analyst first assessed the prior probabilities of different hypotheses occurring. He felt that the odds were 4:1 in favor of H_3, "no program." He also felt there was an equal probability that the programs represented by H_1 and H_2 existed. Thus his prior probability assessments were $P(H_1) = 0.10$, $P(H_2) = 0.10$, and $P(H_3) = 0.80$.

The analyst next assessed the linkages of data D^7, D^8, and D^9 to the hypotheses, and the linkages of the lower-level data to the indicators about which they were diagnostic. He found it more convenient to express these relationships in terms of likelihood ratios as follows:

D^7 - Reference to Dr. Y in paper by Dr. X. The analyst felt that this datum was least likely given H_3. He then assessed the relative likelihoods of the datum, given the other hypotheses in comparison to H_3.

$$\frac{P(D^7|H_1}{P(D^7|H_3} = \frac{6}{5} = 1.2$$

$$\frac{P(D^7|H_2}{P(D^7|H_3} = \frac{8}{7} = 1.14$$

He then expressed these relationships in ratio form:

$$P(D^7|H_1):P(D^7|H_2):P(D^7|H_3)$$

$$1.2 \quad : \quad 1.14 \quad : \quad 1$$

D^8 - Interest by Dr. Y in Country C work during visit to Country C facilities:

$$P(D^8|H_1):P(D^8|H_2):P(D^8|H_3)$$

$$2 \quad : \quad 2 \quad : \quad 1$$

D^9 - Co-location of Dr. X and Dr. Y:

$$P(D^9|H_1):P(D^9|H_2):P(D^9|H_3)$$

$$4 \quad : \quad 2 \quad : \quad 1$$

D^{111}- Content of paper by Dr. X:

$$\frac{P(D^{111}|I_1^{11})}{P(D^{111}|I_2^{11})} = \frac{5}{1}$$

D^{131}- Existence of towed, controllable sensor platform suitable for probing submarine wakes:

$$\frac{P(D^{131}|I_1^{13})}{P(D^{131}|I_2^{13})} = \frac{5}{2}$$

D^{331}- Paper describing design characteristics of an Ir radiometer:

$$\frac{P(D^{331}|I_1^{33})}{P(D^{331}|I_2^{33}} = \frac{6}{1}$$

D^{441}- Paper discussing differential radiometry:

$$\frac{P(D^{441}|I_1^{44})}{P(D^{441}|I_2^{44})} = \frac{20}{7}$$

D^{661}- Paper by Dr. X describing effects of turbulent seas on oil platform operations:

$$\frac{P(D^{661}|I_1^{61})}{P(D^{661}|I_2^{61})} = \frac{2}{3}$$

D^{631}- Paper by Dr. X discussing turbulent wake of a fluid in flux:

$$\frac{P(D^{631}|I_1^{63})}{P(D^{631}|I_2^{63})} = \frac{7}{5}$$

Next, the analyst needed to assess the conditional probabilities relating the indicators I^{11} and I^{13} to activity A^1, and the activity to the hypotheses. After making these judgments, he set out assessments pertaining to the A^1 branch of the hierarchy in the form of a deductive structure, Figure D-7A.

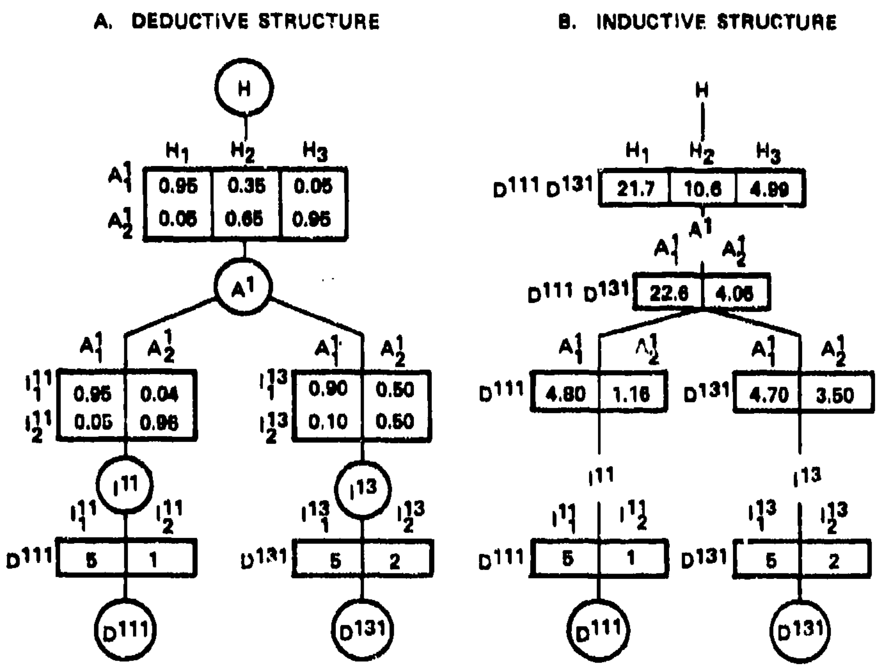


Figure D-7

HIERARCHICAL STRUCTURE FOR BRANCH A^1

The results of working up the hierarchy from the observation of data D^{111} and D^{131} are given in Figure D-7B. Calculations are shown in the steps below.

$$P(D^{111}|A_1^1) = P(D^{111}|I_1^{11})P(I_1^{11}|A_1^1) + P(D^{111}|I_2^{11})P(I_2^{11}|A_1^1)$$

$$= (5)(0.95) + (1)(0.05)$$

$$= 4.75 + 0.05 = \underline{4.80}$$

$$P(D^{111}|A_2^1) = P(D^{111}|I_1^{11})P(I_1^{11}|A_2^1) + P(D^{111}|I_2^{11})P(I_2^{11}|A_2^1)$$
$$= (5)(0.04) + (1)(0.96)$$
$$= 0.20 + 0.96 = \underline{1.16}$$

$$P(D^{131}|A_1^1) = P(D^{131}|I_1^{13})P(I_1^{13}|A_1^1) + P(D^{131}|I_2^{13})P(I_2^{13}|A_1^1)$$
$$= (5)(0.90) + (2)(0.10)$$
$$= 4.50 + 0.20 = \underline{4.70}$$

$$P(D^{131}|A_2^1) = P(D^{131}|I_1^{13})P(I_1^{13}|A_2^1) + P(D^{131}|I_2^{13})P(I_2^{13}|A_2^1)$$
$$= (5)(0.50) + (2)(0.50)$$
$$= 2.50 + 1.00 = \underline{3.50}$$

$$P(D^{111}D^{131}|A_1^1) = P(D^{111}|A_1^1)P(D^{131}|A_1^1)$$
$$= (4.80)(4.70) = \underline{22.6}$$

$$P(D^{111}D^{131}|A_2^1) = P(D^{111}|A_2^1)P(D^{131}|A_2^1)$$
$$= (1.16)(3.50) = \underline{4.06}$$

$$P(D^{111}D^{131}|H_1) = P(D^{111}D^{131}|A_1^1)P(A_1^1|H_1) + P(D^{111}D^{131}|A_2^1)P(A_2^1|H_1)$$
$$= (22.6)(0.95) + (4.06)(0.05)$$
$$= 21.5 + .20 = \underline{21.7}$$

$$P(D^{111}D^{131}|H_2) = P(D^{111}D^{131}|A_1^1)P(A_1^1|H_2) + P(D^{111}D^{131}|A_2^1)P(A_2^1|H_2)$$
$$= (22.6)(0.35) + (4.06)(0.65)$$
$$= 7.92 + 2.64 = \underline{10.6}$$

$$P(D^{111}D^{131}|H_3) = P(D^{111}D^{131}|A_1^1)P(A_1^1|H_3) + P(D^{111}D^{131}|A_2^1)P(A_2^1|H_3)$$
$$= (22.6)(0.05) + (4.06)(0.95)$$
$$= 1.13 + 3.86 = \underline{4.99}$$

The analyst next assessed the necessary conditional probabilities relating the indicators to the activity A^3, and the activity to the hypotheses, as shown in Figure D-8A.

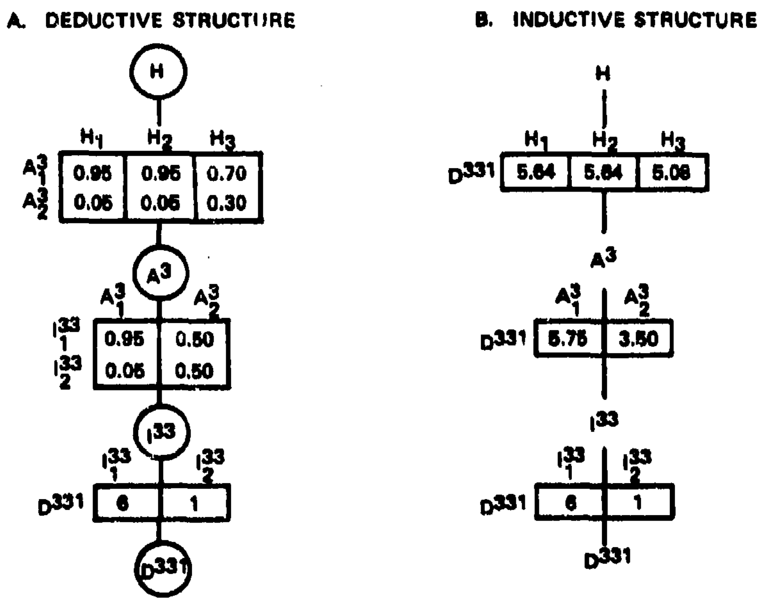


Figure D-8

HIERARCHICAL STRUCTURES FOR BRANCH A^3

He then performed the calculations, as shown below, to obtain the values for the inductive structure, as shown in Figure D-8B.

$$P(D^{331}|A_1^3) = (6)(0.95) + (1)(0.05)$$

$$= 5.70 + 0.05 = \underline{5.75}$$

$$P(D^{331}|A_2^3) = (6)(0.50) + (1)(0.50)$$

$$= 3.00 + 0.50 = \underline{3.50}$$

$$P(D^{331}|H_1) = (5.75)(0.95) + (3.50)(0.05)$$

$$= 5.46 + 0.18 = \underline{5.64}$$

$$P(D^{331}|H_2) = (5.75)(0.95) + (3.50)(0.05)$$
$$= 5.46 + 0.18 = \underline{5.64}$$

$$P(D^{331}|H_3) = (5.75)(0.70) + (3.50)(0.30)$$
$$= 4.03 + 1.05 = \underline{5.08}$$

Figures D-9 and D-10 show the deductive structures assessed by the analyst for branches A^4 and A^6, and the results of solving the inductive structure.

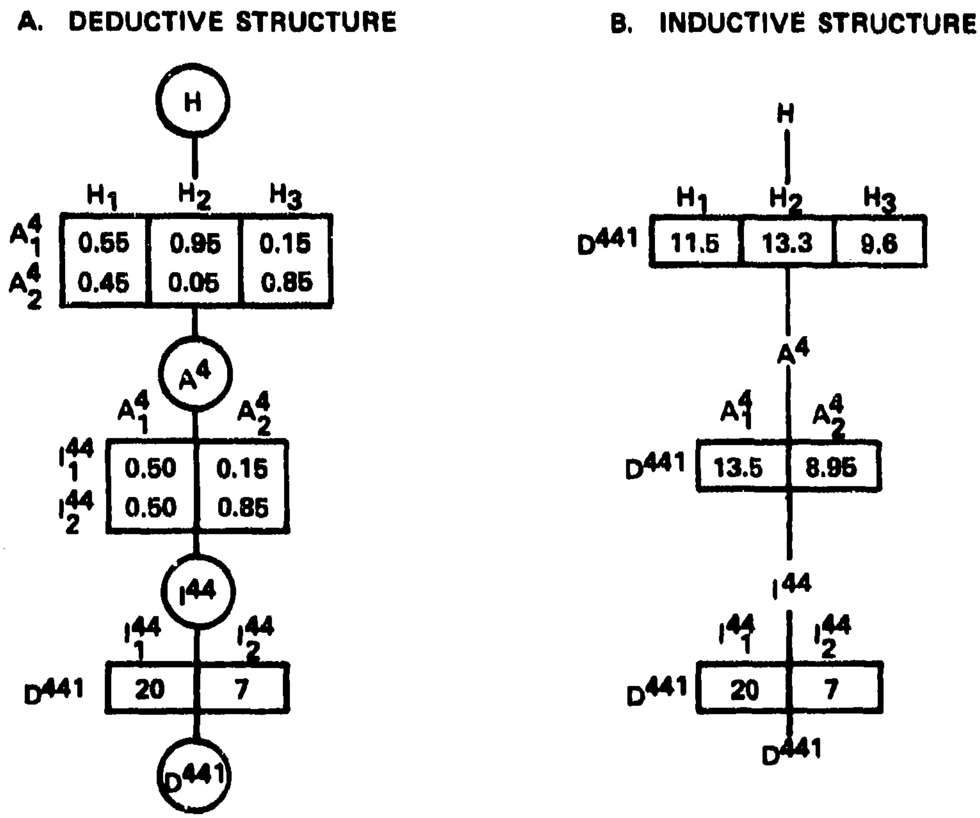


Figure D-9

HIERARCHICAL STRUCTURES FOR BRANCH A^4

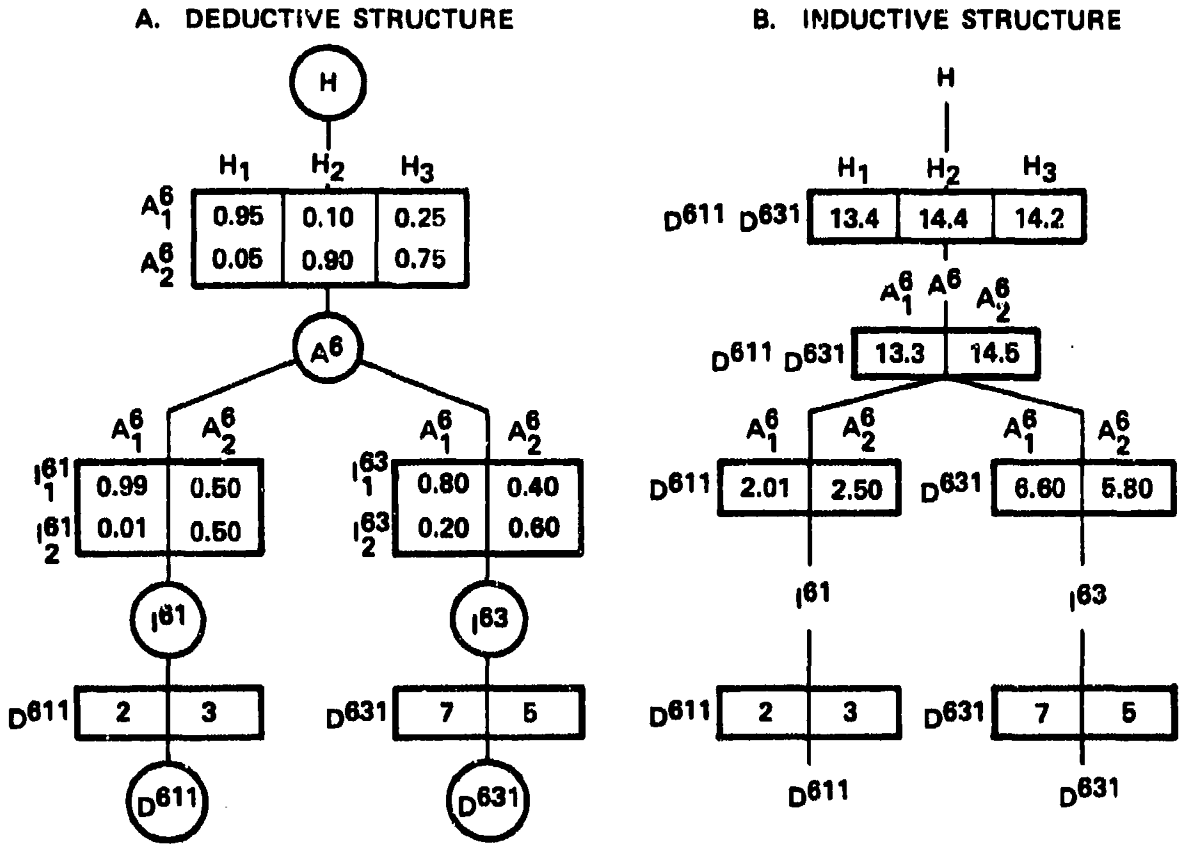


Figure D-10
HIERARCHICAL STRUCTURE FOR BRANCH A^6

Figure D-11 shows the final step in solving the inference; it uses the computed likelihoods for the activity branches and the given likelihoods for the data impacting directly on the hypotheses. Each likelihood has been normalized so that its impact on the final likelihood is easily visualized.

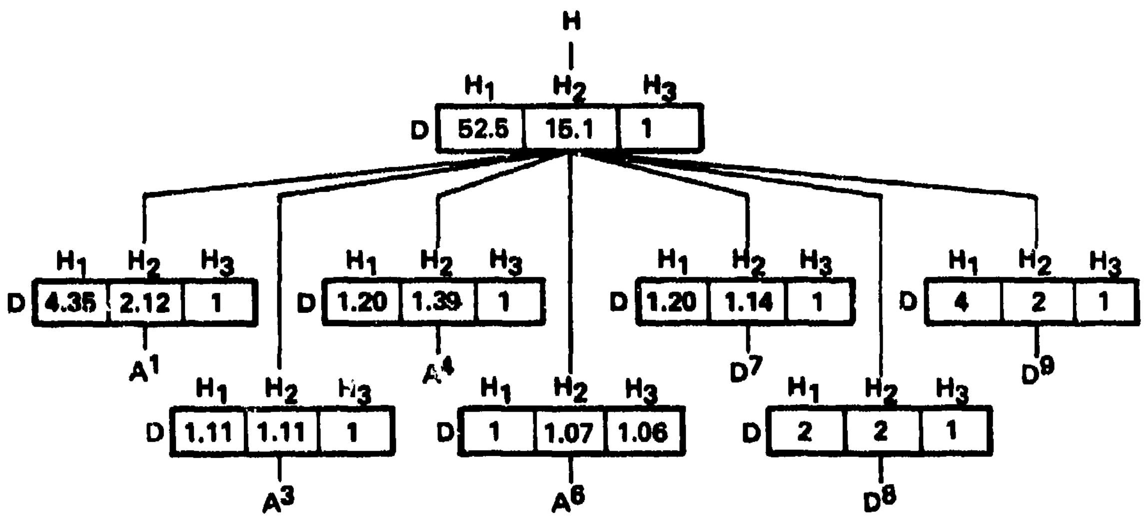


Figure D-11

INDUCTIVE STRUCTURE FOR FINAL STEP IN INFERENCE SOLUTION

The final step in the calculations is as follows:

$$\begin{aligned} P(D|H_1) &= P(D^{111}D^{131}|H_1) \times P(D^{331}|H_1) \times P(D^{441}|H_1) \\ &\quad \times P(D^{611}D^{631}|H_1) \\ &= (4.35)(1.11)(1.20)(1)(1.20)(2)(4) = \underline{55.62} . \end{aligned}$$

Similarly,

$$P(D|H_2) = (2.12)(1.11)(1.39)(1.07)(1.14)(2)(2) = \underline{15.96}$$

$$P(D|H_3) = (1)(1)(1)(1.06)(1)(1)(1) = \underline{1.06} .$$

Normalizing these gives likelihoods of 52.5, 15.1 and 1.
Using Bayes's Theorem,

$$\begin{array}{c} P(H_1|D) \\ \hline P(H_2|D) \\ \hline P(H_3|D) \end{array} = \begin{array}{c} 52.5 \\ \hline 15.1 \\ \hline 1 \end{array} \times \begin{array}{c} 0.1 \\ \hline 0.1 \\ \hline 0.8 \end{array} = \begin{array}{c} 6.56 \\ \hline 1.89 \\ \hline 1 \end{array} .$$

And from this the posterior probabilities are:

$$P(H_1|D) = \frac{6.56}{9.45} = 0.694,$$

$$P(H_2|D) = \frac{1.89}{9.45} = 0.200,$$

$$P(H_3|D) = \frac{1}{9.45} = 0.106.$$

INDEX

INDEX

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TR 77-6-30	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) HANDBOOK FOR DECISION ANALYSIS		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) S. Barclay, R.V. Brown, C.W. Kelly, III, C.R. Peterson, L.D. Phillips, J. Selvidge		8. CONTRACT OR GRANT NUMBER(s) N00014-67-A-0181-0034 N00014-76-C-0074
9. PERFORMING ORGANIZATION NAME AND ADDRESS Decisions and Designs, Incorporated 8400 Westpark Drive, Suite 600 McLean, Virginia 22101		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, Virginia 22209		12. REPORT DATE September 1977 13. NUMBER OF PAGES 285
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Office of Naval Research 800 North Quincy Street Arlington, Virginia 22217		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release: Distribution Unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Best Available Copy

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Decision analysis
Probability assessment
Multi-attribute utility
Hierarchical inference
Decision making
Decision theory
Decision aids

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This handbook is intended to provide decision makers and their staffs (current or potential) with an introduction to the basic concepts and operations of decision analysis.

Decision analysis is a quantitative method which permits the systematic evaluation of the costs or benefits accruing to courses of action that might be taken in a decision problem. It entails identification of the alternative choices involved, the assignment of values over

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) (costs/benefits) for possible outcomes, and the expression of the probability of those outcomes being realized. With this information at hand, one can then systematically combine the values and probabilities to show the probable gain or loss that is associated with alternative choices.

Chapters 1 through 5 of this handbook provide an overview of the major technical aspects of decision analysis: structuring models, assigning probabilities and values, concepts of personal probabilities, developing inferences from evidence, and information value. The concepts are developed in the context of defense or defense-related examples which will enable the principal reading audience to better relate to the material. Chapter 6 provides a case study in the form of a dialogue between a decision analyst and a task force commander centered around the use of decision analysis in resolving a decision problem faced by the commander. The case study is designed to exemplify all the principles developed in Chapters 1 through 5.