

# Formulating and Solving Sequential Decision Analysis Models with Continuous Variables

Jeffrey S. Stonebraker and Craig W. Kirkwood

**Abstract**— This paper presents a new decision analysis approach for modeling decision problems with continuous decision and/or random variables, and applies the approach to a research and development (R&D) planning problem. The approach allows for compact, natural formulation for classes of decision problems that are less appropriately addressed with standard discrete-variable decision analysis methods. Thus it provides a useful alternative analysis approach for problems that are often addressed in practice using simulation risk analysis methods. An illustrative application is presented to energy system R&D planning. The continuous-variable version of this model more directly represents the structure of the decision than a discrete approximation, and the resulting model can be efficiently solved using standard nonlinear optimization methods.

**Index Terms**—Continuous variables, decision analysis, government, model formulation, R&D planning, resource allocation, strategic planning.

## I. INTRODUCTION

THIS paper presents a decision analysis approach for formulating and solving sequential decision problems under uncertainty with continuous decision and/or random variables. An illustrative application to a large-scale decision problem in research and development (R&D) planning is presented. This approach provides a useful alternative to the simulation risk analysis methods that are often used to analyze such decisions in management and engineering practice.

Decisions with continuous variables are widespread in management and engineering. Examples include: 1) multiperiod resource allocation (such as R&D, oil lease bidding, or investment rollovers) where the intermediate results are uncertain, 2) civil litigation where the amount of the potential settlement is at question, 3) setting design parameters for new products, 4) corporate strategies and resource allocation across business portfolios, and 5) introducing a new product to the marketplace where the amount of capital to be invested and the price of the product need to be determined.

The standard decision analysis approach to such decisions is to discretize the continuous variables and then to solve the resulting discrete problem using decision tree or influence diagram methods [9]. This approach has disadvantages. First, it is not a natural way to think about decision problems with

continuous variables. Discrete models can appear cumbersome and inaccurate to managers or engineers who must implement the results of the analysis because these managers or engineers often think of the decision problem variables as continuous. A discrete approach seems to be distorting the “real” problem to fit the available analysis tools, a common criticism of management science methods.

This is of particular concern when there are a large number of variables because the discrete approximation must be relatively crude to keep the problem small enough to be solved. For example, in a decision problem with ten continuous variables, if each of these is approximated with a discrete variable having five possible levels, then the resulting decision tree has almost ten million endpoints. Unless there is a special structure to the problem, this is time consuming to solve. Yet, ten variables are not many for a practical management or engineering decision. If the number of variables increases to 20 and if five levels are used in the discrete approximation, then there are almost  $10^{14}$  endpoints in the decision tree. Without special structure, solving this is not possible with generally available computers.

Because of the solution difficulties, probability distributions for continuous random variables are often approximated in decision analysis practice with three-point discrete distributions. While research shows that a carefully chosen three-point discrete approximation can yield accurate results [14], [33], [15], [22], such an approximation appears crude to some potential users of decision analysis methods and can give inaccurate results in some situations. In practice, simulation risk analysis approaches to such problems are often preferred, and spreadsheet add-ins are available to implement these approaches [18]. Simulation approaches easily handle continuous random variables, but they generally cannot determine the optimal solution for a sequential decision problem. With these approaches, it is often necessary to heuristically determine what decision strategies will be analyzed. Thus the resulting analysis only determines the best of the strategies that were analyzed, and leaves open the possibility that another (unanalyzed) strategy might be better.

A continuous-variable decision analysis approach to this type of decision problem addresses these difficulties. For example, random variables can be formulated in terms of named distributions (normal, beta, triangular, etc.), as is often done in simulation risk analysis. While these distributions still must be discretized to be solved on a digital computer, this discretization is removed from the formulation of the problem and becomes part of the solution procedure, where it is merely

Manuscript received August 3, 1995; revised June 1996. Review of this manuscript was arranged by Editor-in-Chief D. F. Kocaoglu.

J. S. Stonebraker is with Headquarters Air Mobility Command, Quality Office, Scott Air Force Base, IL 62225-5014 USA (e-mail: stonebrj@hqamc.safb.af.mil).

C. W. Kirkwood is with the Department of Management, Arizona State University, Tempe, AZ 85287-4006 USA (e-mail: craig.kirkwood@asu.edu).

Publisher Item Identifier S 0018-9391(97)00692-2.

a numerical analysis issue. Thus, it does not raise the question of distorting a decision that involves continuous variables to fit (discrete) modeling tools. Similarly, with continuous decision variables, standard continuous mathematical optimization methods can be used to determine the optimum at each decision point in the model. As with continuous random variables, the solution procedure for continuous decision variables may require discretization, but this is once again only a solution issue, rather than being part of the problem's formulation.

Furthermore, continuous mathematical optimization solution methods can be much faster than the exhaustive search approach that is often necessary when a discrete sequential decision model is solved. Hence, while this may seem somewhat paradoxical, it may be possible to solve larger decision models with continuous decision variables than with discrete decision variables. Suppose, for example, that a decision problem includes five continuous decision variables, and it is desired to determine optimal values for each of these accurate to within 5%. With a discrete approximation, this requires using 19 different levels for each variable, and hence  $19^5 = 2\,476\,099$  different decision strategies must be analyzed. On the other hand, if the method of bisection [8] can be used with a continuous model to search for the optimal value of each continuous variable, it is only necessary to consider five levels for each variable to obtain a solution that is accurate to within 5%, or a total of  $5^5 = 3125$  different strategies, which is almost a factor of 800 fewer. The relative computational advantage of the method of bisection grows rapidly as the desired solution accuracy increases.

## II. BACKGROUND AND FORMULATION

Our approach to formulating decisions with continuous variables is a generalization of Kirkwood's approach [16] and [17] to formulating decisions with discrete variables. With his approach, information about the path which leads from the root node of a sequential decision model (decision tree) to a particular interior node or endpoint is summarized in a *state*  $s$ . Each node or endpoint is identified by a unique *node index*  $i$  and a *node type*  $t_i$ , which can be "decision," "chance," "endpoint," or "auxiliary." Each decision or chance node has  $b_i$  branches, where  $b_i \geq 1$  is an integer, and each branch is identified by a unique integer *branch index*  $j$ , with  $1 \leq j \leq b_i$ .

An auxiliary node is analogous to a deterministic node in an influence diagram and can be used to clarify the structure of a decision analysis model. (Auxiliary nodes are equivalent to decision or chance nodes with one branch.) Each node (but not endpoint) has a *next node function*  $n_{ij}(s)$ , which yields a valid node index, and a *branch value function*  $f_{ij}(s)$ . Each chance node has a *branch probability function*  $p_{ij}(s)$ . Finally, each node has a *node variable*  $v_i$ , which specifies each algebraic variable in the decision analysis model. With this notation, the expected utility  $EU[i|s]$  for node  $i$  given state  $s$  is described by (see (1) at the bottom of the page) where  $u$  is an utility function, and  $\{v_i : f_{ij}(s), s\}$  is the state formed by adding to  $s$  the node variable  $v_i$  and its associated value as determined by  $f_{ij}(s)$ . For example, suppose that  $s$  is {COST : 500}, where "COST : 500" means that node variable COST has a value of 500. If the next node variable in the path is REVENUE and  $f_y(s)$  is 800, then this new information is summarized in  $\{v_i : f_y(s), s\}$  as {REVENUE : 800, COST : 500}. For a decision node  $i$ , a preferred alternative is a  $j$  that maximizes  $EU[n_{ij}(s)|\{v_i : f_{ij}(s), s\}]$ .

A number of other approaches have recently been developed for formulating and solving large-scale discrete sequential decision models [10], [26], [7], [25]. Kirkwood's approach emphasizes the use of algebraic variables and functions more strongly than these other approaches. Because of this emphasis on variables and functions, it is straightforward to generalize Kirkwood's approach to handle continuous decision and random variables.

The formulation can be generalized to include decisions with continuous decision and random variables by recognizing that the index  $j$  at each node  $i$  in (1) needs to be converted to a continuous variable, which we will call  $x_i$ . When this is done, (1) can be modified into (see (2) at the bottom of the page) where  $D_i(s)$  is the feasible region for decision variable  $x_i$ ,  $R_i(s)$  is the range of possible values for random variable  $x_i$ ,  $f_i(x_i, s)$  assigns a value to node variable  $v_i$ , and  $F_i(x_i|s)$  is the *cumulative distribution function* for random variable  $x_i$  given state  $s$ . (Note that the  $f_i(s)$  function at endpoint nodes, and the  $n_i(s)$  and  $f_i(s)$  functions at auxiliary nodes, do not depend on  $x_i$ .) As the notation indicates, all of these functions can depend on state  $s$ . A preferred alternative at decision node  $i$  is an  $x_i$  that maximizes  $EU[n_i(x_i, s)|\{v_i : f_i(x_i, s), s\}]$ .

$$EU[i|s] = \begin{cases} \max_j EU[n_{ij}(s)|\{v_i : f_{ij}(s), s\}], & t_i = \text{decision} \\ \sum_{j=1}^{b_i} p_{ij}(s) \times EU[n_{ij}(s)|\{v_i : f_{ij}(s), s\}], & t_i = \text{chance} \\ u[f_{i1}(s)], & t_i = \text{endpoint} \\ EU[n_{i1}(s)|\{v_i : f_{i1}(s), s\}] & t_i = \text{auxiliary} \end{cases} \quad (1)$$

$$EU[i|s] = \begin{cases} \max_{x_i \in D_i(s)} EU[n_i(x_i, s)|\{v_i : f_i(x_i, s), s\}], & t_i = \text{decision} \\ \int_{x_i \in R_i(s)} EU[n_i(x_i, s)|\{v_i : f_i(x_i, s), s\}] dF_i(x_i|s) & t_i = \text{chance} \\ u[f_i(s)], & t_i = \text{endpoint} \\ EU[n_i(s)|\{v_i : f_i(s), s\}] & t_i = \text{auxiliary} \end{cases} \quad (2)$$

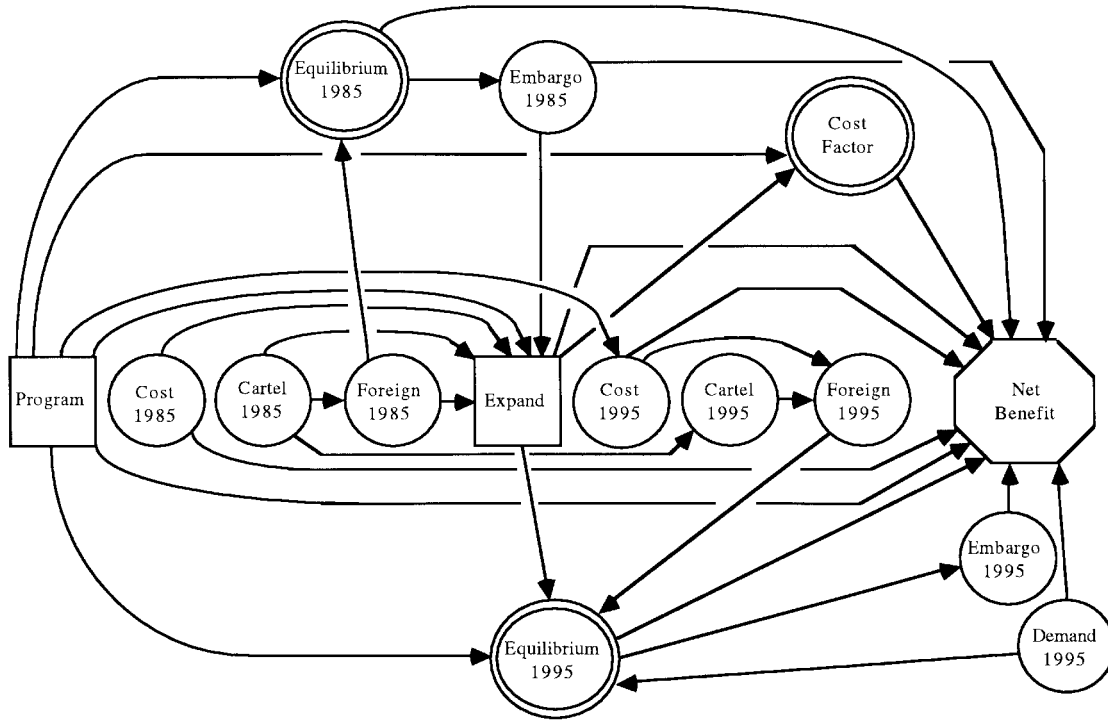


Fig. 1. Influence diagram for the Synfuels decision problem.

The discrete formulation in (1) can be viewed as a special case of (2) where  $D_i(s)$  consists of the integers  $1, 2, \dots, b_i$  and the probability distribution function at chance node  $i$  is discrete over the integers  $1, 2, \dots, b_i$ .

For many practical decisions, the feasible region  $D_i(s)$  will consist of a closed interval  $[L_i(s), U_i(s)]$ , where  $L_i(s)$  and  $U_i(s)$  will be called the *lower bound function* and *upper bound function*, respectively. For example, consider a three-period resource allocation problem where the budgets to be allocated for successive periods are  $x_1, x_2$ , and  $x_3$ , respectively, subject to a total budget constraint  $x_1 + x_2 + x_3 \leq B$ . Then the feasible regions of the three decision variables are

$$\begin{aligned} 0 &\leq x_1 \leq B \\ 0 &\leq x_2 \leq B - x_1 \\ 0 &\leq x_3 \leq B - x_1 - x_2 \end{aligned} \tag{3}$$

and therefore

$$\begin{aligned} L_i(s) &= 0 \quad \text{for } i = 1, 2, 3 \\ U_1(s) &= B \\ U_2(s) &= B - x_1 \\ U_3(s) &= B - x_1 - x_2. \end{aligned} \tag{4}$$

Also, for many practical decision problems with continuous decision variables, the node variable  $v_i$  for each decision node will be equal to the decision variable  $x_i$ , and thus  $f_i(x_i, s) = x_i$  when  $t_i = \text{decision}$ .

The illustrative application below to R&D planning shows that a formulation in terms of (2) can be compact and, furthermore, that such a formulation can show the quantitative nature of the interrelationships among the decision and random variables in a decision more clearly than a discrete formulation.

Cost 1985 (\$/barrel)

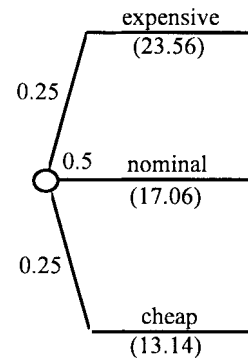


Fig. 2. Tree diagram for Cost 1985.

However, before turning to that application, we consider the solution of (2).

### III. SOLUTION APPROACHES

Kirkwood [16] presents a solution procedure for (1), and he notes that the recursive nature of the equation (that is, the function  $EU$  is defined in terms of itself) leads to a particularly simple solution algorithm when programmed in a computer language which supports recursive function calls. While the solution algorithm presented in [16] is “brute force” in the sense that it traverses all paths through the decision tree corresponding to (1), empirical results show that it is sufficiently fast to realistically analyze discrete decision models with up to a few million endpoints using widely available personal computers.

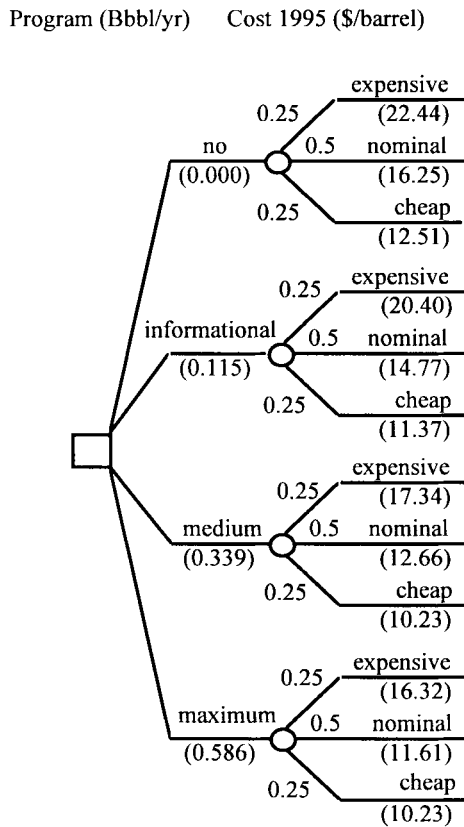


Fig. 3. Tree diagram for Cost 1995.

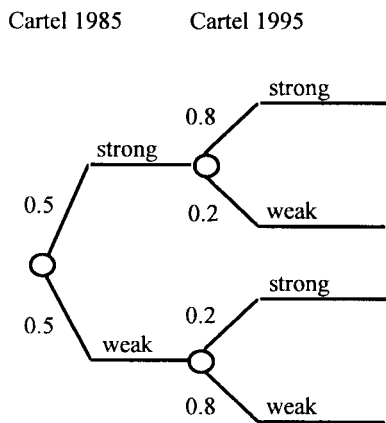


Fig. 4. Tree diagram for Cartel 1985 and Cartel 1995.

Developing an analogous solution procedure for (2) requires that an integration procedure be implemented for the chance nodes and that a maximization procedure be implemented for the decision nodes. The integration of probability density functions has been considered by Keefer [14], Smith [27], Merkhofer [21], Keefer and Bodily [15], and Miller and Rice [22] among others. Keefer [14] and Keefer and Bodily [15] show that in many cases accurate expected utilities and expected values can be obtained using a three-point discrete approximation consisting of the 0.05, 0.50, and 0.95 fractiles of the continuous distribution. Interested readers are referred to these papers for further details.

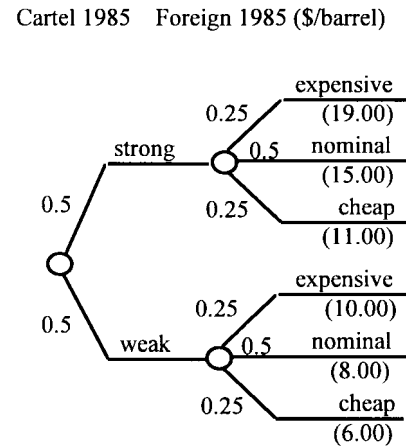


Fig. 5. Tree diagram for Foreign 1985.

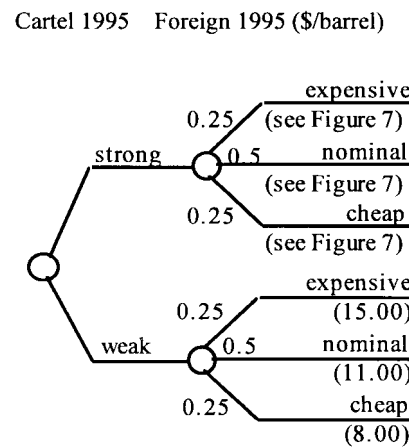


Fig. 6. Tree diagram for Foreign 1995.

The most common case in decision analysis practice is where the feasible region for each decision variable consists of an interval specified by lower and upper bound functions as discussed above. In general, the problem resulting from the analysis at each decision node is a nonlinear single-variable optimization problem subject to lower and upper bound constraints. In cases where there is a single local optimum at an interior point, the golden-section search algorithm [12] will efficiently determine the optimum. The Fibonacci search algorithm is also sometimes used. However, the golden-section search is often selected over the Fibonacci search because the total number of search intervals for the Fibonacci must be chosen in advance, whereas the golden-section search does not require this [8], which simplifies implementation. Other iterative approaches that use calculus, such as the Taylor series algorithm [8], are not practical to solve (2) because they require a closed-form specification of the objective function in order to differentiate it.

In situations where there are multiple local optima or the optimum occurs at a boundary, it is always possible to use a grid search solution procedure where the feasible region is sectioned into a grid of equally spaced points, and these points are then exhaustively searched to find the optimal point [12]. By using a fine enough grid, any desired degree of accuracy

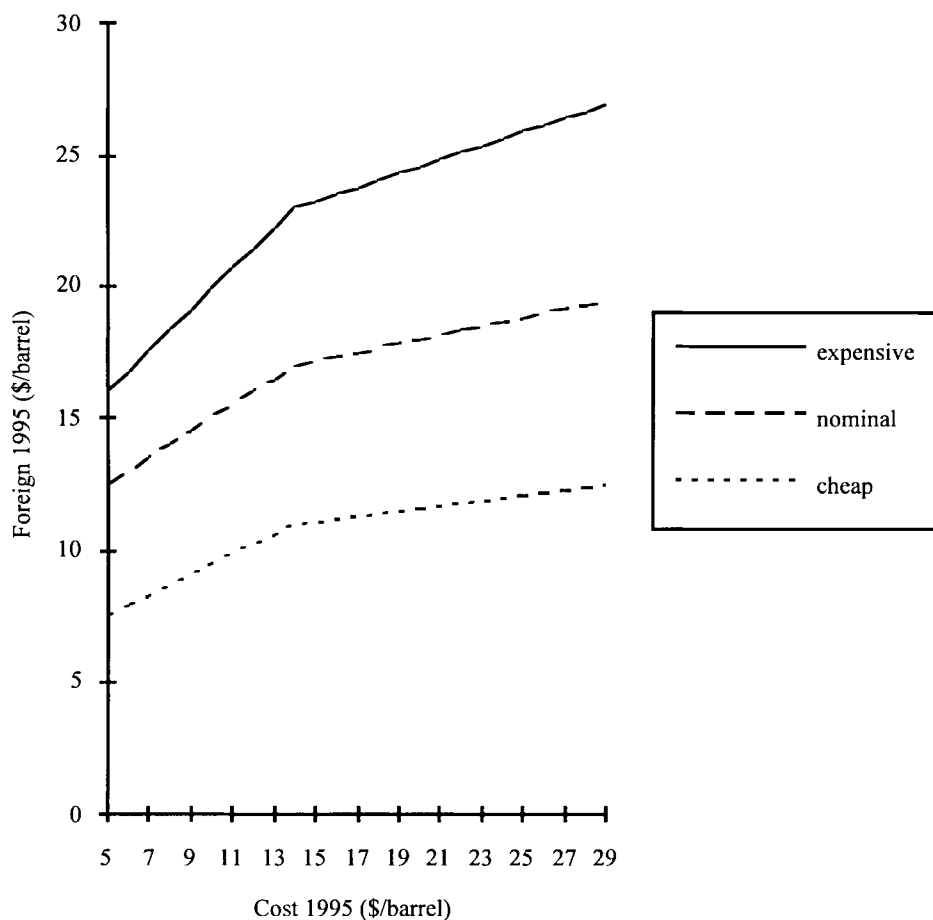


Fig. 7. Foreign 1995 when Cartel 1995 is strong.

can be obtained, but of course, considerable computation may be required if a fine grid is used.

We present empirical results related to the use of the golden-section search and grid search in the context of an illustrative application below. See [29] for further details of these solution procedures. Note that the specific solution procedure used does not impact the decision problem formulation, as given in (2). That is, the formulation of the decision model can be done using the notation of (2) regardless of the specific procedure used to solve this equation.

#### IV. ILLUSTRATIVE APPLICATION

In his January 1975 State of the Union Address, President Ford called for the accelerated development of U.S. energy technology and resources so as to achieve a low-cost alternative to imported oil. The Interagency Task Force on Synthetic Fuel (Synfuel) Commercialization was formed to evaluate alternatives and make recommendations to the President. The Decision Analysis Group at Stanford Research Institute was selected to assist the Task Force [31], [30]. This Group developed a large-scale discrete sequential decision analysis model (a 93 312 endpoint decision tree) to aid the Task Force. At the time of their modeling effort, no general purpose algebraic formulation method existed for decision analysis models, either discrete or continuous. Thus an *ad hoc* description was presented which covered 33 pages in an appen-

dix to [30], including a four-page, step-by-step calculation of the net benefit for a sample path through the 93 312 endpoint decision tree.

##### A. Original Discrete Model Formulation

The decision problem was formulated as a sequential resource allocation problem, with an initial decision in 1975 as to whether the United States should have a synfuels development program and, if so, at what funding. A subsequent decision in 1985 would determine if the current synfuels production capacity should be expanded and, if so, to what level of funding. These two decision variables were intrinsically continuous, as were a number of the random variables in the model (for example, the cost to produce the synfuel). However, following standard decision analysis practice, all of these continuous variables were approximated with discrete variables. Because of the complexity of the resulting discrete model, the analysis of the initial decision in the synfuels problem was limited to considering only four alternatives, namely: produce 0.000 billion barrels of synfuel per year (Bbbl/yr) ("no program"), produce 0.115 Bbbl/yr ("informational program"), produce 0.339 Bbbl/yr ("medium program"), and produce 0.586 Bbbl/yr ("maximum program"). The expansion decision in 1985 was approximated with six discrete levels, i.e., produce 0.000, 0.365, 0.730, 1.095, 1.460, and 1.825 Bbbl/yr.

TABLE I  
CAPACITY EXPANSION COST FACTORS [30]

| Program (Bbbl/yr) | Expand (Bbbl/yr) |       |       |       |       |       |
|-------------------|------------------|-------|-------|-------|-------|-------|
|                   | 0.000            | 0.365 | 0.730 | 1.095 | 1.460 | 1.825 |
| 0.000             | 0.94             | 1.00  | 1.13  | 1.32  | 3.00  | 10.00 |
| 0.115             | 0.90             | 0.96  | 1.05  | 1.16  | 1.30  | 1.61  |
| 0.339             | 0.89             | 0.94  | 1.01  | 1.11  | 1.23  | 1.46  |
| 0.586             | 0.88             | 0.93  | 0.99  | 1.06  | 1.16  | 1.36  |

These decisions are represented in the Fig. 1 influence diagram as *Program* and *Expand*. The *Program* decision is made before the *Expand* decision, hence the arrow from the former to the latter. In addition to these two decisions, there are a number of uncertain variables. For example, there are four uncertainties associated with the year 1985, namely: cost, cartel, foreign, and embargo; and five uncertainties for 1995: cost, cartel, foreign, embargo, and demand (where demand represents the structure of the market demand, that is, a parameter value for the demand curve). (The years 1985 and 1995 were used to typify the decades of the 1980's and 1990's, respectively. Each uncertainty associated with the year 1985 will be resolved prior to the *Expand* decision.)

*Cost 1985* and *Cost 1995* represent the uncertainty in the cost to produce synfuel in the years 1985 and 1995, respectively. While these are continuous variables, they are approximated as discrete variables. *Cost 1985* is independent of *Program*, and the probability distribution assumed for *Cost 1985* is shown in Fig. 2. *Cost 1995* depends on *Program*, and the assumptions made about the dependence are shown in Fig. 3. As shown in Fig. 3, the cost to produce a barrel of synthetic fuel in 1995 decreases as the production of such fuel increases.

*Cartel 1985* and *Cartel 1995* represent the state (strong or weak) of the OPEC cartel in 1985 and 1995, respectively. *Cartel 1995* depends on the outcome of *Cartel 1985* as shown in Fig. 4. For example, if the cartel is strong in 1985, then it is more likely that the cartel in 1995 will be strong than if the cartel is weak in 1985.

*Foreign 1985* and *Foreign 1995* represent the price of foreign oil in 1985 and 1995, respectively. *Foreign 1985* depends on the state of the cartel in 1985 as shown in Fig. 5. For example, if the cartel is strong in 1985, then the price of foreign oil in 1985 will be higher than with a weak cartel. *Foreign 1995* depends on the state of the cartel in 1995 and the cost of synthetic fuel in 1995 as shown in Figs. 6 and 7. For example, if the cartel is strong in 1995, then the price of foreign oil in 1995 will be higher than if the cartel is weak.

*Equilibrium 1985* and *Equilibrium 1995* are deterministic or auxiliary variables that represent the market equilibrium prices of oil for 1985 and 1995, respectively. As shown in Fig. 1, *Equilibrium 1985* depends on *Program* and *Foreign 1985*; and *Equilibrium 1995* depends on *Program*, *Expand*, *Foreign*

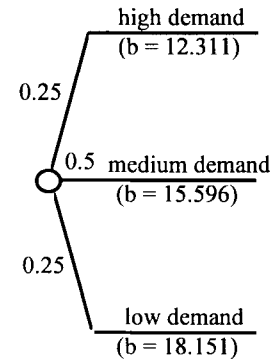


Fig. 8. Tree diagram for Demand 1995.

1995, and *Demand 1995*. As specified in [30], *Equilibrium 1985* and *Equilibrium 1995* are the minimum of the market-clearing price of synfuel and the price of foreign oil in the years 1985 and 1995, respectively. The market-clearing price is determined from the following demand curve:

$$p(q) = a/(b + q) + c \quad (5)$$

where  $p(q)$  is the price of synfuel,  $q$  is the amount of synfuel demanded, and  $a$ ,  $b$ , and  $c$  are parameter values for the demand curve. The parameter values used for the 1985 demand curve are  $a = 1888.875$ ,  $b = 19.893$ , and  $c = -69.000$ . The parameter values used for the 1995 demand curve are  $a = 809.910$  and  $c = -23.946$ , whereas the  $b$  parameter for the demand curve for 1995 (*Demand 1995*) is uncertain, and the assumed probability distribution is shown in Fig. 8.

*Embargo 1985* and *Embargo 1995* represent the likelihood of an embargo and its associated loss in 1985 and 1995, respectively. The probability of an embargo was judged to be 0.10 in 1985 and 0.05 in 1995. The duration of an embargo was assumed to be five months. As derived in [30], the economic effect of an embargo is a loss of consumer surplus, as given by

$$\text{embargo loss} = -(5/12) \cdot (2.5a) \cdot [q_0^2 - 0.25 \cdot (q - q_0)^2] / (b + q_0)^2 \quad (6)$$

where  $q_0$  is the amount of foreign oil and synfuel demanded at the market equilibrium price  $p_0$ ;  $q$  is the amount of synfuel demanded; and  $a$  and  $b$  are parameter values for the demand curve.

TABLE II  
CONTINUOUS-VARIABLE DECISION ANALYSIS MODEL FORMULATION OF THE SYNFUELS DECISION PROBLEM

| Node Index $i$ | Node Type $t_i$ | Node Variable $v_i$ | Lower Bound $L_i(s)$ | Upper Bound $U_i(s)$ | Branch Value $f_i(x_i, s)$ | Probability Distribution $F_i(x_i   s)$ | Next Node $n_i(x_i, s)$ |
|----------------|-----------------|---------------------|----------------------|----------------------|----------------------------|---|-------------------------|
| 1              | decision        | Program             | 0.000                | 0.586                |                            |   | 2                       |
| 2              | chance          | Cost 1985           |                      |                      |                            | $N[\mu_2(s), \sigma_2(s)]$              | 3                       |
| 3              | chance          | Cartel 1985         |                      |                      | $f_{3j}(s)$                | $p_{3j}(s)$                             | 4                       |
| 4              | chance          | Foreign 1985        |                      |                      |                            | $N[\mu_4(s), \sigma_4(s)]$              | 5                       |
| 5              | auxiliary       | Equilibrium 1985    |                      |                      | $f_5(s)$                   |   | 6                       |
| 6              | chance          | Embargo 1985        |                      |                      | $f_{6j}(s)$                | $p_{6j}(s)$                             | 7                       |
| 7              | decision        | Expand              | 0.000                | 1.825                |                            |   | 8                       |
| 8              | chance          | Cost 1995           |                      |                      |                            | $N[\mu_8(s), \sigma_8(s)]$              | 9                       |
| 9              | chance          | Cartel 1995         |                      |                      | $f_{9j}(s)$                | $p_{9j}(s)$                             | 10                      |
| 10             | chance          | Foreign 1995        |                      |                      |                            | $N[\mu_{10}(s), \sigma_{10}(s)]$        | 11                      |
| 11             | chance          | Demand 1995         |                      |                      |                            | $N[\mu_{11}(s), \sigma_{11}(s)]$        | 12                      |
| 12             | auxiliary       | Equilibrium 1995    |                      |                      | $f_{12}(s)$                |   | 13                      |
| 13             | chance          | Embargo 1995        |                      |                      | $f_{13j}(s)$               | $p_{13j}(s)$                            | 14                      |
| 14             | auxiliary       | Cost Factor         |                      |                      | $f_{14}(s)$                |   | 15                      |
| 15             | endpoint        |                     |                      |                      | $f_{15}(s)$                |   |                         |

Consumer surplus is the difference between the value of a good (i.e., energy) to a consumer and the amount paid for it. Given a demand curve, such as (5), and a market equilibrium price, the consumer surplus is determined from the area below the demand curve and above the market equilibrium price. Consumer surplus, as derived in [30], is given by

$$\text{consumer surplus} = a \cdot [\ln(a/b) - 1] - a \cdot \ln[p_0 - c] + b \cdot (p_0 - c) \tag{7}$$

where  $a$ ,  $b$ , and  $c$  are parameter values for the demand curve; and  $p_0$  is the market equilibrium price.

The supply curve used in [30] is the product of *Cost 1995* and the appropriate capacity expansion cost factor shown in Table I. Producer surplus is the difference between the amount received by a producer for a good and the cost to produce

it. Given a supply curve and a market equilibrium price, producer surplus is determined from the signed sum of the area above market equilibrium price and below the supply curve (thus including possible negative contributions to producer surplus), and the area below the market equilibrium price and above the supply curve (thus including possible positive contributions to producer surplus). The deterministic variable *Cost Factor* in Fig. 1 is a contribution to producer surplus and is determined by the area under the supply curve. As derived in [30], producer surplus for 1985 and 1995, respectively, are determined by and (8) and (9) at the bottom of the page.

Consumer surplus, producer surplus, embargo loss, and environmental and socioeconomic cost specify the objective function, *Net Benefit*. The environmental and socioeconomic cost of producing synfuel is \$0.40 per barrel [30]. The net ben-

$$\text{1985 producer surplus} = (\text{Equilibrium 1985} - \text{Cost 1985}) \cdot \text{Program} \tag{8}$$

$$\text{1995 producer surplus} = (\text{Equilibrium 1995} - \text{Cost 1995}) \cdot \text{Program} + \text{Equilibrium 1995} \cdot (\text{Program} + \text{Expand}) - \text{Cost Factor} \tag{9}$$

TABLE III  
MEAN AND STANDARD DEVIATION FUNCTIONS FOR CONTINUOUS RANDOM VARIABLES IN THE SYNFUELS DECISION MODEL

| i  | $\mu_i(s)$   | $\sigma_i(s)$  | state s  |
|----|--|--|--|
| 2  | 17.7   | 3.74   |  |
| 4  | 15.0<br>8.0  | 2.83<br>1.41   | Cartel 1985 = 1<br>Cartel 1985 = 2   |
| 8  | 16.9 - 7.54 • Program  | 3.56 - 2.14 • Program  |  |
| 10 | 9.39 + 0.542 • Cost 1995<br>14.6 + 0.175 • Cost 1995<br>11.3 | 2.36 + 0.135 • Cost 1995<br>3.41 + 0.059 • Cost 1995<br>2.49 | Cartel 1995 = 1, 5 ≤ Cost 1995 < 14<br>Cartel 1995 = 1, 14 ≤ Cost 1995 ≤ 29<br>Cartel 1995 = 2 |
| 11 | 15.4   | 2.07   |  |

TABLE IV  
PROBABILITY DISTRIBUTIONS FOR DISCRETE RANDOM VARIABLES IN THE SYNFUELS DECISION MODEL

| Node Index i | Node Type $t_i$ | Node Variable $v_i$ | Number of Branches $b_i$ | Branch Index j | Branch Value $f_{ij}(s)$ | Branch Probability $p_{ij}(s)$ | Next Node $n_{ij}(s)$ |
|--------------|-----------------|---------------------|--------------------------|----------------|--------------------------|--------------------------------|-----------------------|
| 3            | chance          | Cartel 1985         | 2                        | 1              | 1                        | 0.50                           | 4                     |
|              |                 |                     |                          | 2              | 2                        | 0.50                           | 4                     |
| 6            | chance          | Embargo 1985        | 2                        | 1              | $f_{6j}(s)$              | 0.10                           | 7                     |
|              |                 |                     |                          | 2              | "                        | 0.90                           | 7                     |
| 9            | chance          | Cartel 1995         | 2                        | 1              | 1                        | $p_{9j}(s)$                    | 10                    |
|              |                 |                     |                          | 2              | 2                        | "                              | 10                    |
| 13           | chance          | Embargo 1995        | 2                        | 1              | $f_{13j}(s)$             | 0.05                           | 14                    |
|              |                 |                     |                          | 2              | "                        | 0.95                           | 14                    |

efit for 1985 is the sum of consumer surplus in 1985, producer surplus in 1985, embargo loss in 1985, and environmental and socioeconomic cost in 1985. Similarly, the net benefit for 1995 is the sum of consumer surplus in 1995, producer surplus in 1995, embargo loss in 1995, and environmental and socioeconomic cost in 1995. The total discounted net benefit is measured in constant 1975 dollars, as determined by multiplying the net benefits for 1985 and 1995 by the appropriate discount factors and summing. (A 10% discount rate is used in [30] and results in discount factors of 4.2 for 1985 and 1.62 for 1995.)

### B. Continuous Variable Model Formulation

As mentioned above, several of the variables in the synfuels decision problem are intrinsically continuous. These include the decisions *Program* and *Expand*, as well as the random variables *Cost 1985*, *Foreign 1985*, *Cost 1995*, *Foreign 1995*, and *Demand 1995*. As shown above, the original formulation approximated the possible levels for these variables with a small number of discrete levels.

To illustrate the usefulness of a continuous-variable decision analysis formulation, such a formulation is presented for the

synfuels decision problem in Tables II–V. Table II contains the basic structure of the decision model using the notation of (2). The two continuous decision variables *Program* and *Expand* have node indexes one and seven, and Table II shows that the feasible regions for these variables are specified by lower and upper bounds, which are simply fixed values.

The continuous random variables *Cost 1985*, *Foreign 1985*, *Cost 1995*, *Foreign 1995*, and *Demand 1995* have node indexes 2, 4, 8, 10, and 11, respectively. To illustrate the continuous formulation approach for these random variables, normal distributions specified by  $N[\mu, \sigma]$ , where  $\mu$  is the mean and  $\sigma$  is the standard deviation, are assumed for the five continuous random variables. (Since no continuous distributions were specified in the original analysis, these should be considered only illustrative.) The functional forms for the means and standard deviations of these normal distributions are specified in Table III.

The remaining (discrete) random variables are *Cartel 1985*, *Embargo 1985*, *Cartel 1995*, and *Embargo 1995*, with node indexes 3, 6, 9, and 13, respectively. The probability distributions for these random variables are given in Table IV, and some of the entries in this table are further defined in Table V.



TABLE V  
SUPPLEMENTAL FUNCTIONS IN THE SYNFUELS DECISION MODEL

|  |   |   |
|--|---|---|
| $f_5(s)$   | $\min \{ [1888.875 / (19.893 + \text{Program})] - 69, \text{Foreign 1985} \}$   |   |
| $f_{61}(s)$  | $-1967.58 \cdot [c^2 - 0.25 \cdot (c + \text{Program})^2] / (c + 19.893)^2$   |   |
| $f_{62}(s)$  | 0<br>where $c = 1888.875 / (\text{Equilibrium 1985} + 69) - 19.893$   |   |
|  | <u>Cartel 1985 = 1</u>  | <u>Cartel 1985 = 2</u>  |
| $p_{91}(s)$  | 0.80  | 0.20  |
| $p_{92}(s)$  | 0.20  | 0.80  |
| $f_{12}(s)$  | $\min \{ [809.91 / (\text{Demand 1995} + \text{Program} + \text{Expand})] - 23.946, \text{Foreign 1995} \}$   |   |
| $f_{13,1}(s)$  | $-843.66 \cdot [c^2 - 0.25 \cdot (c + \text{Program} + \text{Expand})^2] / (c + \text{Demand 1995})^2$  |   |
| $f_{13,2}(s)$  | 0<br>where $c = 809.91 / (\text{Equilibrium 1995} + 23.946) - \text{Demand 1995}$   |   |
| $f_{14}(s)$  | $\text{CF}(\text{Program}, 0.000) \cdot (\text{Program} + \text{Expand}),$  | $0.000 \leq \text{Expand} < 0.365$  |
|  | $0.1825 \cdot [\text{CF}(\text{Program}, 0.365) + \text{CF}(\text{Program}, 0.000)]$<br>$+ \text{CF}(\text{Program}, 0.365) \cdot (\text{Program} + \text{Expand} - 0.365),$  | $0.365 \leq \text{Expand} < 0.730$  |
|  | $0.1825 \cdot [\text{CF}(\text{Program}, 0.000) + 2 \cdot \text{CF}(\text{Program}, 0.365) + \text{CF}(\text{Program}, 0.730)]$<br>$+ \text{CF}(\text{Program}, 0.730) \cdot (\text{Program} + \text{Expand} - 0.730),$   | $0.730 \leq \text{Expand} < 1.095$  |
|  | $0.1825 \cdot [\text{CF}(\text{Program}, 0.000) + 2 \cdot \text{CF}(\text{Program}, 0.365) + 2 \cdot \text{CF}(\text{Program}, 0.730)]$<br>$+ \text{CF}(\text{Program}, 1.095) + \text{CF}(\text{Program}, 1.095) \cdot (\text{Program} + \text{Expand} - 1.095),$  | $1.095 \leq \text{Expand} < 1.460$  |
|  | $0.1825 \cdot [\text{CF}(\text{Program}, 0.000) + 2 \cdot \text{CF}(\text{Program}, 0.365) + 2 \cdot \text{CF}(\text{Program}, 0.730)]$<br>$+ 2 \cdot \text{CF}(\text{Program}, 1.095) + \text{CF}(\text{Program}, 1.460)]$<br>$+ \text{CF}(\text{Program}, 1.460) \cdot (\text{Program} + \text{Expand} - 1.460),$   | $1.460 \leq \text{Expand} \leq 1.825$   |
| The value for $\text{CF}(\text{Program}, \text{Expand})$ required for $f_{14}(s)$ is calculated by linear interpolation between adjacent Program entries and between adjacent Expand entries in the following table: |   |   |
|  |   | <u>Expand</u>   |
|  | <u>Program</u>  | <u>0.000</u> <u>0.365</u> <u>0.730</u> <u>1.095</u> <u>1.460</u> <u>1.825</u> |
| $\text{CF}(\text{Program}, \text{Expand})$   | 0.000   | 0.94   1.00   1.13   1.32   3.00   10.00                                      |
|  | 0.115   | 0.90   0.96   1.05   1.16   1.30   1.61                                       |
|  | 0.339   | 0.89   0.94   1.01   1.11   1.23   1.46                                       |
|  | 0.586   | 0.88   0.93   0.99   1.06   1.16   1.36                                       |
| $f_{15}(s)$  | $4.2 \cdot [8084.49 - 1888.875 \cdot \ln(\text{Equilibrium 1985} + 69) + 19.893 \cdot (\text{Equilibrium 1985} + 69) + \text{Embargo 1985} + (\text{Equilibrium 1985} - \text{Cost 1985}) \cdot \text{Program} - 0.4 \cdot \text{Program}] + 1.62 \cdot [809.91 \cdot \{\ln(809.91 / \text{Demand 1995}) - 1\} - 809.91 \cdot \ln(\text{Equilibrium 1995} + 23.946) + \text{Demand 1995} \cdot (\text{Equilibrium 1995} + 23.946) + \text{Embargo 1995} + \text{Equilibrium 1995} \cdot (\text{Program} + \text{Expand}) + (\text{Equilibrium 1995} - \text{Cost 1985}) \cdot \text{Program} - (\text{Cost 1995}) \cdot (\text{Cost Factor}) - 0.4 \cdot (\text{Program} + \text{Expand})]$ |   |

There are three auxiliary variables *Equilibrium 1985*, *Equilibrium 1995*, and *Cost Factor*, with node indexes 5, 12, and 14, respectively. The specific functions which determine the values of these auxiliary variables are given in Table V. Finally, node index 15 is the endpoint for the decision model, and the function which determines the endpoint values for the model is given in Table V.

The formulation in Tables II–V is significantly more compact than the original 33-page formulation shown in [30], but a greater advantage is that it shows the structure of the decision

problem more clearly. We quickly see that the two decision variables can lie anywhere within a bounded interval, and the specific quantitative nature of the interdependencies among the decision and random variables is more apparent. For example, the mean and standard deviation for *Cost 1995* both depend linearly on *Program*. In addition, *Embargo 1985* is a function of *Program* and *Equilibrium 1985*, and *Embargo 1995* is a function of *Program*, *Expand*, *Demand 1995*, and *Equilibrium 1995*, with the specific functional relationships shown in the table.

Having random variables as functions of decision variables as in this application is common in R&D decisions where the likelihood of success of a R&D project is characterized by the amount spent on funding the project. See, for example, [1]–[6], [11], [13], [19], [20], [23], [24], [28], and [32].

The formulation in Tables II–V was solved using (2). To solve this formulation, the continuous random variables were integrated using the extended Pearson–Tukey approximation [14]. Other integration procedures, such as the extended Swan–Megill approximation [15], moment methods [27], etc., could also be used. The expected net benefit was calculated as \$173.14 billion using the grid search solution procedure and \$173.23 billion using the golden-section search solution procedure, when the values for *Program* and *Expand* were determined to within 5%. At this accuracy level, the golden-section search was about seven times faster than the grid search (45 s versus 5 min 10 s on a 33-MHz 486 IBM-compatible personal computer with a math coprocessor.)

## V. CONCLUSIONS

The continuous-variable decision analysis approach offers a direct approach to formulating and solving sequential decision problems under uncertainty when the variables in such problems are continuous. It more explicitly shows the structure of such problems than a traditional discrete decision analysis model. In addition to possible computational advantages, this approach can result in better acceptance of the decision model by managers and engineers because the continuous variables in the decision problem are not distorted to fit discrete analysis tools.

## ACKNOWLEDGMENT

The reviewers provided helpful comments which substantially improved the presentation in this article.

## REFERENCES

- [1] A. C. Atkinson and A. H. Bobis, "A mathematical basis for the selection of research projects," *IEEE Trans. Eng. Manage.*, vol. EM-16, pp. 2–8, 1969.
- [2] N. Baker and J. Freeland, "Recent advances in R&D benefit measurement and project selection methods," *Manage. Sci.*, vol. 21, pp. 1164–1175, 1975.
- [3] J. F. Bard, R. Balachandra, and P. E. Kaufmann, "An interactive approach to R&D project selection and termination," *IEEE Trans. Eng. Manage.*, vol. 35, pp. 139–146, 1988.
- [4] A. H. Bobis and A. C. Atkinson, "Analyzing R&D investments via dynamic modeling," *Chem. Eng.*, vol. 77, no. 5, pp. 133–141, 1970.
- [5] ———, "Analyzing potential research projects," *Chem. Eng.*, vol. 77, no. 4, pp. 95–100, 1970.
- [6] A. H. Bobis, T. F. Cooke, and J. H. Paden, "A funds allocation method to improve the odds for research successes," *Res. Manage.*, vol. 14, no. 2, pp. 34–49, 1971.
- [7] H. J. Call and W. A. Miller, "A comparison of approaches and implementations for automating decision analysis," *Reliability Eng. Syst. Safety*, vol. 30, pp. 115–162, 1990.
- [8] E. W. Cheney and D. R. Kincaid, *Numerical Mathematics and Computing*, 2nd ed. Pacific Grove, CA: Brooks/Cole, 1985.
- [9] R. T. Clemen, *Making Hard Decisions: An Introduction to Decision Analysis*, 2nd ed. Belmont, CA: Duxbury, 1996.
- [10] Z. Covaliu and R. M. Oliver, "Representation and solution of decision problems using sequential decision diagrams," *Manage. Sci.*, vol. 41, pp. 1860–1881, 1995.
- [11] B. V. Dean and L. E. Hauser, "Advanced materiel systems planning," *IEEE Trans. Eng. Manage.*, vol. EM-14, pp. 21–43, 1967.
- [12] B. S. Gottfried and J. Weisman, *Introduction to Optimization Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1973.
- [13] S. W. Hess, "A dynamic programming approach to R and D budgeting and project selection," *IRE Trans. Eng. Manage.*, vol. EM-9, pp. 170–179, 1962.
- [14] D. L. Keefer, "Certainty equivalents for three-point discrete distribution approximations," *Manage. Sci.*, vol. 40, pp. 760–773, 1994.
- [15] D. L. Keefer and S. E. Bodily, "Three-point approximations for continuous random variables," *Manage. Sci.*, vol. 29, pp. 595–609, 1983.
- [16] C. W. Kirkwood, "Implementing an algorithm to solve large sequential decision analysis models," *IEEE Trans. Syst., Man Cybern.*, vol. 24, pp. 1425–1432, 1994.
- [17] ———, "An algebraic approach to formulating and solving large models for sequential decisions under uncertainty," *Manage. Sci.*, vol. 39, pp. 900–913, 1993.
- [18] ———, "An overview of methods for applied decision analysis," *Interfaces*, vol. 22, no. 6, pp. 28–39, 1992.
- [19] G. R. Madley and B. V. Dean, "Strategic planning for investment in R&D using decision analysis and mathematical programming," *IEEE Trans. Eng. Manage.*, vol. EM-32, pp. 84–90, 1985.
- [20] J. M. Mellichamp, W. L. Dixon Jr., and S. L. Mitchell, "Ballistic missile defense technology management with goal programming," *Interfaces*, vol. 10, no. 5, pp. 68–75, 1980.
- [21] M. W. Merkhofer, "Quantifying judgmental uncertainty: methodology, experiences, and insights," *IEEE Trans. Syst., Man Cybern.*, vol. SMC-17, pp. 741–752, 1987.
- [22] A. C. Miller and T. R. Rice, "Discrete approximations of probability distributions," *Manage. Sci.*, vol. 29, pp. 352–362, 1983.
- [23] D. L. Owen, "Selecting projects to obtain a balanced research portfolio," in *Readings on the Principles and Applications of Decision Analysis*, vol. I, R. A. Howard and J. E. Matheson, Eds. Menlo Park, CA: Strategic Decisions Group, 1983, pp. 337–361.
- [24] E. M. Rosen and W. E. Souder, "A method for allocating R&D expenditures," *IEEE Trans. Eng. Manage.*, vol. EM-12, pp. 87–93, 1965.
- [25] R. D. Shachter, "Evaluating influence diagrams," *Operat. Res.*, vol. 34, pp. 871–882, 1986.
- [26] P. P. Shenoy, "Valuation-based systems for bayesian decision analysis," *Operat. Res.*, vol. 40, pp. 463–484, 1992.
- [27] J. E. Smith, "Moment methods for decision analysis," *Manage. Sci.*, vol. 39, pp. 340–358, 1993.
- [28] W. E. Souder, "Analytical effectiveness of mathematical models for R&D project selection," *Manage. Sci.*, vol. 19, pp. 907–923, 1973.
- [29] J. S. Stonebraker, "Applying decision analysis to sequential decisions under uncertainty with continuous decision variables," Ph.D. dissertation, Dept. Decision and Inform. Syst., Arizona State Univ., 1993.
- [30] Synfuels Interagency Task Force, "Recommendations for a synthetic fuels commercialization program," vol. II, Cost/Benefit Analysis of Alternative Production Levels, 1975.
- [31] S. N. Tani, "Decision analysis of the synthetic fuels commercialization program," *Readings on the Principles and Applications of Decision Analysis*, vol. I, R. A. Howard and J. E. Matheson, Eds. Menlo Park, CA: Strategic Decisions Group, 1978, pp. 435–443.
- [32] B. W. Taylor III, L. J. Moore, and E. R. Clayton, "R&D project selection and manpower allocation with integer nonlinear goal programming," *Manage. Sci.*, vol. 28, pp. 1149–1158, 1982.
- [33] N. A. Zaino, Jr. and J. D'Errico, "Optimal discrete approximations for continuous outcomes with applications in decision and risk analysis," *J. Operat. Res. Soc.*, vol. 40, pp. 379–388, 1989.



**Jeffrey S. Stonebraker** received the B.S. degree in electrical engineering from the University of South Florida, Tampa, the M.S. degree in engineering-economic systems from Stanford University, Stanford, CA, and the Ph.D. degree in business administration from Arizona State University, Tempe.

He is a Major in the United States Air Force. His work experience includes four years as a ground-based radar test engineer at Hill Air Force Base, and five years as an Assistant Professor at the United States Air Force Academy's Department of Mathematical Sciences. He is currently a Quality Consultant at Scott Air Force Base, and has consulted for numerous companies in the private sector, as well as the Departments of Defense and Energy. His research interest is in decision analysis.

Maj. Stonebraker is a member of the Institute for Operations Research and the Management Sciences and its Decision Analysis Society, Tau Beta Pi, Phi Kappa Phi, Alpha Iota Delta, and Omega Rho.



**Craig W. Kirkwood** received the S.B., S.M., E.E., and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, MA.

He is currently a Professor at Arizona State University, Tempe, and has served on the faculties of the University of Colorado and the University of Michigan. He worked with the decision analysis group at Woodward-Clyde Consultants, San Francisco. His research interests include large scale decision analysis models and decision analysis using partial or approximate information.

Dr. Kirkwood is a member of the Institute for Operations Research and the Management Sciences and has served on the Advisory Panel for the National Science Foundation's Decision, Risk, and Management Science Program.