JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 22, No. 2, JUNE 1977

Confidence Structures in Decision Making[1]

P. L. YU[2] AND G. LEITMANN[3]

Abstract. Decision making is defined in terms of four elements: the set of decisions, the set of outcomes for each decision, a set-valued criterion function, and the decision maker's value judgment for each outcome. Various confidence structures are defined, which give the decision maker's confidence of a given decision leading to a particular outcome. The relation of certain confidence structures to Bayesian decision making and to membership functions in fuzzy set theory is established. A number of schemes are discussed for arriving at *best* decisions, and some new types of domination structures are introduced.

Key Words. Confidence structures, domination structures, chance constraint formulation, multicriteria decision making, hierarchy of decision processes.

1. Introduction

We consider the process of decision making to be composed of four elements:

(i) the set of all feasible alternatives (decisions) X with elements denoted by x;

(ii) the set of all possible outcomes $Y(x) \subset R^m$ for each feasible alternative $x \in X$;

(iii) the criterion function $f(\cdot): x \mapsto Y(x)$, a set-valued function that measures the *value* of a decision;

(iv) the decision maker's *value judgment* or *preference* for each outcome.

[1] This research was partly supported by Project No. NR-047-021, ONR Contract No. N-00014-75-C-0569 with the Center for Cybernetic Studies, The University of Texas, Austin, Texas, and by ONR Contract No. N-00014-69-A-0200-1012 with the University of California, Berkeley, California.

[2] Associate Professor, Graduate School of Business, University of Texas, Austin, Texas.

[3] Professor, Department of Mechanical Engineering, University of California, Berkeley, California.

The totality of all possible outcomes is

$$Y = \cup\{Y(x)|x \in X\} \subset R^m.$$

The coordinates in R^m may be used for indexing quantitative or qualitative (linguistic) outcomes.

To illustrate these concepts, we consider a simple investment problem (SIP). The decision maker wishes to invest his savings $\$M$, so that he may be well off in the future. Here, X includes all possible stock purchases including deposits in banks. Of course, X may be not well defined. In fact, generating *good* alternatives (elements of X) is a very important ingredient of the decision process. Let us suppose that the decision maker uses two criteria, *growth rate of asset value* and *safety*, to measure the desirability of an investment (other criteria, such as liquidity, are important; but we shall not consider them here, for the sake of simplicity). Note that, depending on the economic situation, the outcomes of the decisions may be highly unpredictable. For instance, buying stock may yield a high growth rate of asset value and great safety in a bullish market, and quite the opposite in a bearish one. The set of all possible outcomes of a decision x to buy a certain stock (measured in terms of growth rate of asset value and safety) is a set $Y(x)$. Once each $Y(x)$ is specified, the set containing all possible outcomes Y is known.

The decision maker's value judgment of each element $y \in Y$ may be not simple. Various ways of forming such value judgments has been proposed; e.g., preference or utility construction, domination structures, etc. Subsequently, we shall classify value judgment in terms of single or multiple criteria.

Henceforth, we shall assume that the decision set X and the criterion function $f(\cdot)$ are specified. We shall focus out attention on two questions: How can one define the outcomes of a decision? What are methods of value judgment for arriving at a good decision?

In some of the literature (for instance, that quoted in Refs. 1 and 2), uncertainty of outcome is treated in terms of *a priori* probability distributions. Usually, a single criterion is employed to define the outcome of a decision. For many complex decision problems, uncertainty of outcome cannot be adequately represented by an *a priori* distribution. In the next section, we shall introduce the concept of *confidence structures* for the purpose of treating uncertainty. It will be shown that this concept is closely related to Bayesian *a priori* probability distribution and to Zadeh's membership function (Refs. 3 and 4). In Section 3, we shall exhibit some methods for attaining good decisions with a variety of confidence structures and a number of representations of value judgment or preference. In Section 4, we shall present a hierarchy of general decision processes.

2. Confidence Structures

Recall here that, given a decision $x \in X$, the set of possible outcomes[4] is denoted by $Y(x)$. Of course, $Y(x)$ depends on the decision maker's prior belief. Loosely stated, a *confidence structure* is a collection of information or prior beliefs which specifies, for each feasible decision $x \in X$, a set of prior probability measures for each $y \in Y(x)$ to be the outcome of x.

To help the intuitive understanding of confidence structure, we give first a definition for the case of $Y(x)$ consisting of discrete points only.

Definition 2.1. Suppose that, given any $x \in X$, $Y(x)$ consists of discrete points only. Let $\mathscr{I}$ denote the set of all nonempty subintervals, including isolated points, of the interval $[0, 1]$. A confidence structure over X (the set of all feasible decisions) and

$$Y = \cup\{Y(x) | x \in X\}$$

(a set that includes all possible outcomes of all feasible decisions) is a set-valued function[5]

$$c(\cdot, \cdot): X \times Y \to \mathscr{I}.$$

We interpret

$$c(x, y) = [a, b] \in \mathscr{I}, \qquad x \in X, \qquad y \in Y,$$

to mean that the decision maker has confidence in terms of prior probability from a to b that y will occur if he makes decision x. The interval $[a, b]$ is called the *confidence interval* for y to be the outcome of x.

Example 2.1. Suppose that the decision maker believes that his making a decision x will result in only two possible outcomes y^1 and y^2, with probabilities in $[0.2, 0.6]$ and $(0.4, 0.7]$, respectively. Then, one can specify $c(\cdot, \cdot)$, where

$$c(x, y) = \begin{cases} [0.2, 0.6] & \text{if } y = y^1 \\ (0.4, 0.7] & \text{if } y = y^2, \\ \{0\} & \text{otherwise.} \end{cases}$$

Here,

$$Y(x) = \{y^1, y^2\}.$$

[4] That is, outcomes whose probability of occurrence is nonzero.

[5] Since Y may include points in outcome space which are not possible due to a given decision x, one must allow zero probability.

Now, suppose that Y is an arbitrary, not necessarily countable, subset of R^m. Prior probability is not as readily specified as in Definition 2.1; however, the concept of probability measure appears to be useful, even though the intuitive meaning of confidence structure may not be as apparent.

Definition 2.2. Let $\mathscr{Y}$ be a collection of subsets of Y, and let $\mathscr{I}$ be the set of all nonempty subintervals, including isolated points, of the interval $[0, 1]$. A confidence structure over X and $\mathscr{Y}$ is a set to set-valued function

$$\mathscr{C}(\cdot, \cdot): X \times \mathscr{Y} \to \mathscr{I}.$$

We interpret

$$\mathscr{C}(x, U) = [a, b] \in \mathscr{I}, \qquad x \in X, \qquad U \in \mathscr{Y},$$

to mean that the decision maker has confidence in terms of prior probability from a to b that the outcome of decision x will be in set U.

Remark 2.1. Definition 2.2 is very general. To be mathematically manageable, the set $\mathscr{Y}$ may have to have structure such as a σ-algebra or Borel measurability. With this specification, for decision x fixed, a probability measure

$$\mathscr{P}(\cdot): \mathscr{Y} \to [0, 1]$$

satisfying

$$\mathscr{P}(U) \in \mathscr{C}(x, U)$$

for all $U \in \mathscr{Y}$ can represent a confidence structure for fixed x. Thus, a confidence structure induces, for each $x \in X$, a class of probability measures which describe the decision maker's belief in the outcomes of his decision; for further discussion see Example 2.4. From the point of view of information content, the smaller the confidence intervals $\mathscr{C}(x, U)$, $U \in \mathscr{Y}$, and the smaller the set $\mathscr{Y}$, the better.

Since we allow zero measure, we can extend Y to R^m, and we can use the concept of probability distribution function to define confidence structures.

Definition 2.3. Let $\mathscr{I}$ denote the set of all nonempty subintervals, including isolated points, of the interval $[0, 1]$. A confidence structure is a set-valued function

$$\mathscr{C}(\cdot, \cdot): X \times R^m \to \mathscr{I}.$$

We interpret

$$\mathscr{C}(x, y) = [a, b] \in \mathscr{I}, \qquad x \in X, y \in R^m,$$

to mean that the decision maker has confidence in terms of prior probability distribution from a to b that decision x will result in an outcome not exceeding y; that is,

$$\text{Prob}[\{y' \in R^m \mid y' \leq y\}] \in [a, b].$$

Remark 2.2. Definitions 2.2 and 2.3, while defining confidence structures in terms of probability measure, are cumbersome for purposes of application. To alleviate this we introduce the following convention.

Convention 2.1. Let $\mathscr{I}'$ be the set of all nonempty subintervals, including isolated points, of the nonnegative real half-line. A confidence structure over X and Y is a set-valued function

$$c(\cdot, \cdot): X \times Y \to \mathscr{I}'$$

such that, for each $x \in X$, if $y \in Y(x)$ is an isolated point[6] with respect to $Y(x)$, then $c(x, y)$ is a subinterval of $[0, 1]$ that specifies the range of the prior probability that y is the outcome of x, and such that if y is not an isolated point with respect to $Y(x)$, then $c(x, y)$ is an interval of $[0, \infty)$ that specifies the range of the probability density that y is the outcome of x.

Remark 2.3. Suppose that $Y \subset R^2$ and one of the coordinate axes is used for indexing qualitative (linguistic) outcome. Then, the probability density in Convention 2.1 is defined on a one-dimensional space. In general, if k axes $k \in \{0, 1, \ldots, m-1\}$, are used for indexing qualitative outcomes, then the density function is defined on $(m-k)$-dimensional space.

Example 2.2. In the SIP, $f_1(\cdot)$ and $f_2(\cdot)$ are the criterion functions for *growth rate of asset value* and *safety*, respectively. Thus, the higher their values, the better. Table 1 gives a set of choices X, outcomes $Y(x)$, and confidence intervals (in conformity with Convention 2.1). This is also illustrated in Fig. 1. This example will be used repeatedly hereafter.

An important special case of confidence structures, and the one usually considered, is the one for which $c(x, y)$ is exactly one point of $[0, 1]$, so that it maps $X \times Y \to [0, 1]$. To distinguish this from the general case, we state the following definition.

[6] That is, there is a neighborhood N of y such that $N \cap Y(x) = \{y\}$.

Table 1. Example 2.2.

Choice x	Outcome set $Y(x)$	Confidence interval $c(x, y)$
x^1	$\{y \mid \|y-y^1\|<0.6\}$	$[1/2(1+\|y-y^1\|), 1/(1+\|y-y^1\|)]$
	$\{y \mid \|y-y^1\|\geq 0.6\}$	$\{0\}$
	$y^1=(1.1, 0.9)$	
x^2	$y^{21}=(0.4, 0.5)$	$(0.9, 1]$
	$y^{22}=(1.6, 1.2)$	$[0, 0.05)$
	$y\notin\{y^{21}, y^{22}\}$	$\{0\}$
x^3	$y^3=(1.05, 1)$	$\{1\}$
	$y\neq y^3$	$\{0\}$

Definition 2.4. A confidence structure $c(\cdot, \cdot)$ is called point-valued iff, for each $(x, y)\in X\times Y$, it contains exactly one point of $[0, 1]$. It is denoted by

$$M(\cdot, \cdot): X\times Y\to[0, 1],$$

that is,

$$c(x, y)=\{M(x, y)\}, \qquad \forall(x, y)\in X\times Y.$$

Example 2.3. In the deterministic case, given any decision $x^0\in X$, there is one and only one outcome $y^0\in Y$. Then,

$$c(x^0, y)=\begin{cases}\{1\} & \text{if } y=y^0,\\ \{0\} & \text{otherwise,}\end{cases}$$

$$M(x^0, y)=\begin{cases}1 & \text{if } y=y^0,\\ 0 & \text{otherwise.}\end{cases}$$

Example 2.4. Suppose that each decision $x\in X$ results in outcomes which depend on the occurrence of mutually exclusive and collectively exhaustive events $\{E_1, E_2, \ldots, E_q\}$; e.g., in the SIP, E_1 may indicate a bullish market, E_2 a bearish one, etc. Let P_i be the prior probability for E_i, $i=1, 2, \ldots, q$, to occur. Let $y_i(x)$ denote the outcome of decision x when event E_i occurs. Then, we can give the confidence structure as

$$c(x, y)=\begin{cases}\{P_i\} & \text{if } y=y_i(x), \qquad i=1, 2, \ldots, q,\\ \{0\} & \text{otherwise;}\end{cases}$$

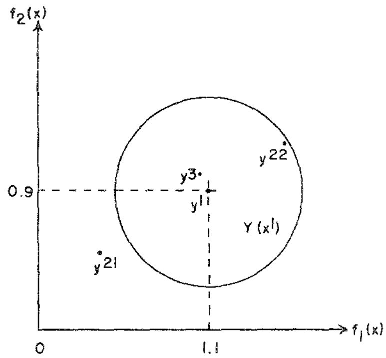

Fig. 1. Example 2.2.

or, in terms of $M(\cdot, \cdot)$,

$$M(x, y) = \begin{cases} P_i & \text{if } y = y_i(x), \quad i = 1, 2, \ldots, q, \\ 0 & \text{otherwise.} \end{cases}$$

More generally, $M(x, \cdot): Y \to [0, 1]$ can be a probability density function for $y \in Y$ to be the outcome of $x \in X$.

If one considers the value (or loss) of y to be a given function

$$V(\cdot): Y \to [0, \infty) \qquad (\text{or } L(\cdot): Y \to [0, \infty)),$$

then the *Bayes decision* is the one that maximizes

$$\sum_{i=1}^{q} P_i V(y_i(x)),$$

or minimizes

$$\sum_{i=1}^{q} P_i L(y_i(x)),$$

over X. For the continuous case, the summation is replaced by integration; see Section 3.2.

Example 2.5. Suppose that all possible outcomes are qualitatively (linguistically) described; e.g., in the SIP, the set of all possible outcomes can be:

1 ≡ very high return and very high risk,
2 ≡ very high return and medium risk,
3 ≡ high return and very high risk, etc.

Then, Y can be an index set of outcomes. That is, with each $y \in Y$, there is associated a qualitative (linguistic) outcome. Now, suppose that $c(\cdot, \cdot)$ is point-valued, so that the causal relation between x and y is represented by

$$M(\cdot, \cdot): X \times Y \to [0, 1];$$

that is, $M(x, y)$ represents a prior belief that y will be the outcome of x. Now, suppose that y is fixed, say

$$y = y^0 \in Y.$$

Then,

$$M(\cdot, y^0): X \to [0, 1]$$

can be versed as a *membership function* in the sense of Zadeh (Refs. 1–2); that is, $M(x, y^0)$ gives the degree of membership of decision x in the qualitatively (linguistically) described outcome y^0. Conversely, given a fuzzy set, one can construct a point-valued confidence structure.

In practice, it may not be easy to specify $c(x, y)$ as a point, but it may be less difficult to specify it as a subinterval of [0, 1]; see Sections 3 and 4.

Remark 2.4. In the Bayesian case (Example 2.4), the confidence structure is represented by a family of prior probability distributions, one for each decision $x \in X$. In the Zadeh case (Example 2.5), it is represented by a set of membership functions, one for each $y \in Y$. The membership functions are not probability distributions, so the sum of the degrees of membership of a decision x need not equal one.

Next, we consider a process of converting a general confidence structure to a point-valued one.[7]

Definition 2.5. The principle of insufficient reason is the following: If one is completely ignorant as to which event among a set of possible events will take place, then one can behave as if they are equally likely to occur.

If one applies the principle of insufficient reason to confidence structures, then the complexity of decision is greatly reduced, albeit at the possible risk of obtaining a poor decision. Then, each $\alpha \in c(x, y)$ is equally likely to be the true probability of x resulting in y. Thus, if

$$c(x, y) = [a, b],$$

then the expected value of one's confidence is $(a + b)/2$; that is, one uses the expected value of the probability as a representation for $c(x, y)$. More precisely, if

$$c(x, y) = [a, b],$$

then

$$M(x, y) = (a + b)/2$$

[7] See Ref. 5 for further discussion.

is used as a representation for $c(x, y)$; of course, $M(\cdot, \cdot)$ is a point-valued confidence structure. While

$$M(x, \cdot): Y \to [0, 1]$$

need not be a probability distribution, it can be normalized into one by dividing $M(x, y)$ by the sum (or integral) of $M(x, y)$ over $Y(x)$, provided it is defined (see Example 2.4).

Example 2.6. *Continuation of Example 2.2.* We shall apply the principle of insufficient reason to the confidence structure given in Table 1. For $x = x^1$, we have then

$$M(x^1, y) = \begin{cases} 3/4(1 + \|y - y^1\|) & \text{if } \|y - y^1\| < 0.6, \\ 0 & \text{otherwise.} \end{cases}$$

Let

$$c_0 = \int_{Y(x^1)} [3/4(1 + \|y - y^1\|)]\, dy = (3\pi/2) \log(1.6),$$

so that

$$M(x^1, \cdot): Y(x^1) \to [0, 1],$$

given by

$$M(x^1, y) = 3/4c_0(1 + \|y - y^1\|)$$

is a probability distribution.

In similar fashion, upon applying the principle of insufficient reason, one obtains

$$M(x^2, y) = \begin{cases} 0.95 & \text{if } y = y^{21}, \\ 0.025 & \text{if } y = y^{22}, \\ 0 & \text{otherwise;} \end{cases}$$

and, after normalization,

$$M(x^2, y) = \begin{cases} 0.95/0.975 & \text{if } y = y^{21}, \\ 0.025/0.975 & \text{if } y = y^{22}, \\ 0 & \text{otherwise.} \end{cases}$$

To simplify our nomenclature we state the following definition.

Definition 2.6. The process of converting a general confidence structure into a point-valued one by means of the principle of insufficient reason, including normalization, will be called the reduction process.

3. Decision Making with Confidence Structures

In this section, we describe some methods for decision making with a variety of confidence structures. As mentioned before, the process of decision making depends not only on the confidence structures for the outcomes, but also on the *value judgment* or *preference* for the outcomes. We classify the situation according to these two elements and describe some methods for decision making. We begin with the simplest case.

3.1. Deterministic Confidence Structures. In the deterministic case, each decision results in a unique outcome (e.g., see Example 2.3). Thus, we can use a function

$$f(\cdot): X \to Y$$

to define the relation between decisions and outcomes.

Value judgment may involve a single (scalar) criterion or a multiple (vector) criterion (multicritical).

3.1.1. Deterministic Confidence Structure with a Single Criterion. Here, the possible outcome set Y is a subset of the one-dimensional Euclidean space. The points of Y may be arranged according to a preference ordering or, in the case of qualitative (linguistic) outcomes, Y may be merely an index set. In either case, we shall suppose there is a real-valued function

$$u(\cdot): Y \to R,$$

such that

$$u(y^1) > u(y^2) \quad \text{iff } y^1 \text{ is preferred to } y^2,$$
$$u(y^1) = u(y^2) \quad \text{iff } y^1 \text{ is indifferent to } y^2.$$

If Y is already arranged in accord with a preference ordering, then $u(\cdot)$ can be the identity map.

With the function $u(\cdot)$ so specified, decision making becomes a standard optimization problem:

$$\max_X [u \circ f(x)].$$

Generalizing $u(\cdot)$ to a total ordering over Y is feasible (e.g., see Refs. 6 and 7).

3.1.2. Deterministic Confidence Structure with a Multiple Criterion. Here, the possible outcome set Y is a subset of m-dimensional Euclidean space with criterion function

$$f(\cdot): X \to Y \subset R^m, \qquad m > 1.$$

Some of the coordinates axes, that is, some components of $f(\cdot)$, may be used as indices for qualitative (linguistic) outcomes.

Decision problems involving multicriteria are common in practice. For example, in the SIP, the decision maker is concerned with growth rate of asset value, safety, and probably other criteria, such as liquidity. In problems of national energy planning, decision makers are subject to considerations of self-sufficiency, cost of energy generation, unemployment, growth, etc.

Many solution concepts have been suggested for making decisions with multicriteria. Some of these are outlined below (for a survey, see Refs. 8 and 9). Except for (i) and (ii) of the listed methods, some monotonicity according to preference is assumed for each component $f_i(\cdot)$ of $f(\cdot)$, including the case for which an $f_i(\cdot)$ indexes a qualitative outcome. For example, in the SIP, $f_1(\cdot)$ may assume values in an index set $\{1, 2, \ldots, 5\}$, with

$1 \equiv$ very high growth rate,
$2 \equiv$ above average growth rate,
$3 \equiv$ average growth rate,
$4 \equiv$ below average growth rate,
$5 \equiv$ very low growth rate.

(i) *One-Dimensional Comparison.* Here, the multicriteria problem is converted into a single-criterion one. In this category are goal programming (Refs. 10 and 11), the additive weight method, compromise solutions (Refs. 12 and 13), utility construction (Ref. 14), and lexicographic ordering.

(ii) *Ordering and Ranking.* Instead of defining a real-valued function over Y as in (i), one defines a binary relation which may be a partial ordering over Y. Then, one seeks the maximum or minimum elements over Y, provided such elements exist. Among such methods are Pareto-optimality, efficient solutions, outranking relations (Ref. 15), and preference ordering (Refs. 6 and 7).

(iii) *Domination Structures and Nondominated Solutions.* Here, for each $y \in Y$, one defines a set of domination factors $D(y)$ such that, iff

$$y^0 \neq y \qquad \text{and } y^0 \in y + D(y),$$

then y^0 is dominated by y. An outcome that is not dominated by any other outcome is nondominated, and the final decision is to result in nondominated ones. For a detailed discussion, including the relation to (i) and (ii), see Refs. 8, 16, 17, and 20.

(iv) *Satisficing Models.* In this approach, the decision maker establishes first either (a) a minimal *satisfaction* level for each criterion or (b) an upper *goal achievement* level for each criterion. In the first case, a decision resulting in any criterion not meeting or exceeding the minimal satisfaction level is unacceptable and ruled out as a candidate for a final decision. In the second case, any decision that results in all criteria meeting or exceeding the upper goal achievement level is acceptable for a final decision.

(v) *Iterative or Adaptive Procedures.* In these methods, a final decision is obtained in a sequence of steps. In one such method, at each step one considers the (remaining) feasible decisions and their outcomes, and eliminates the dominated ones from further consideration. A final decision is selected from among those which cannot be eliminated by this process (Ref. 8). In another method, one begins with a particular feasible decision and then finds a *better* one at each step until improvement becomes impossible. Such a technique, similar to a gradient search, is described in Ref. 18.

Some of the methods listed above can be combined in solving a particular decision problem. For instance, a combination of (i) and (iv) may result in a mathematical programming or an optimal control problem.

3.2. Point-Valued Confidence Structures. If the confidence structure is not deterministic, decision making is more complex. In this section, we consider decision making with point-valued confidence structures (e.g., see Examples 2.4 and 2.5).

3.2.1. Point-Valued Confidence Structure with a Single Criterion. As in Section 3.1.1, we consider a real-valued function

$$u(\cdot): Y \to R$$

such that

$$u(y^1) > u(y^2) \Leftrightarrow y^1 \quad \text{is preferred to } y^2.$$

We shall suppose that the following are defined for each $x \in X$ (recall Convention 2.1):

$$E(x) = \sum_{y \in Y_1(x)} u(y) M(x, y) + \int_{Y_2(x)} u(y) M(x, y)\, dy, \tag{1}$$

$$V(x) = \sum_{y \in Y_1(x)} [u(y) - E(x)]^2 M(x, y) + \int_{Y_2(x)} [u(y) - E(x)]^2 M(x, y)\, dy, \tag{2}$$

where

$$Y_1(x) = \{y \in Y(x) \mid y \text{ is an isolated point w.r.t. } Y(x)\}, \tag{3}$$

$$Y_2(x) = \{y \in Y(x) \mid y \text{ is not an isolated point w.r.t. } Y(x)\}. \tag{4}$$

If $M(x, \cdot)$ is a probability measure,[8] then $E(x)$ and $V(x)$ are the expected value and variance, respectively, of $u(\cdot)$ with respect to $M(x, \cdot)$.

The following decision optimization methods may be useful.

(i) *Maximization of Expected Value.* Here, one seeks $x^* \in X$ such that $E(x^*) \geq E(x)$ for all $x \in X$. Suppose that $u(\cdot)$ is *convex linear* in terms of lotteries. That is, if

$$y^0 = \alpha y^1 + (1-\alpha)y^2, \qquad \alpha \in [0, 1],$$

represents an outcome with chance α to have outcome y^1 and chance $(1-\alpha)$ to have outcome y^2, then

$$u(y^0) = \alpha u(y^1) + (1-\alpha)u(y^2).$$

In this event, maximizing the expected value seems logically sound. Unfortunately, such a utility function is difficult to find.

(ii) *Two Criteria for Value Judgment.* Here we treat $E(x)$ and $V(x)$ as two criteria for value judgment in decision making. Since $V(x)$ is the variance of $u(\cdot)$, it may be regarded as a measure of *fluctuation* or *risk*. Such a two-criteria formulation has been used extensively in portfolio analysis (Ref. 19). Of course, the methods listed in Section 3.1.2 are applicable here. Note that a nondeterministic single-criterion problem has been converted into a deterministic two-criteria problem.

(iii) *Chance Constraint Formulation.* Let

$$Y_\gamma = \{y \in Y \mid u(y) \geq \gamma\},$$

$$X(\beta, \gamma) = \{x \in X \mid \sum_{y \in Y_1(x) \cap Y_\gamma} M(x, y) + \int_{Y_2(x) \cap Y_\gamma} M(x, y)\, dy \geq \beta\},$$

where $Y_1(x)$ and $Y_2(x)$ are defined in (3) and (4), respectively.

Loosely speaking, $X(\beta, \gamma)$ is the set of feasible decisions whose final outcome in terms of $u(\cdot)$ has probability, of at least β, of exceeding a specified level γ. With γ and β specified, the chance constraint formulation is that of maximizing $E(x)$ over $X(\beta, \gamma)$ (e.g., see Ref. 21). This formulation combines the features of expected value maximization with those of a satisficing model. Again, combining two methods, such as (ii) and (iii), is possible.

3.2.2. Point-Valued Confidence with a Multiple Criterion. If outcomes are specified in terms of a multiple criterion with a point-valued confidence structure, decision making is more difficult than in the case of a multiple criterion with a deterministic confidence structure (Section 3.1.2), or in the case of a single criterion with point-valued confidence structure

[8] That is, the total measure of $M(x, \cdot)$ over $Y(x)$ is one. Recall that, by Convention 2.1, $M(x, \cdot)$ on $Y_1(x)$ is a prior probability, whereas on $Y_2(x)$ it is a probability density.

(Section 3.2.1). While it may be possible to combine the methods of Section 3.1.2 and 3.2.1, the success of so doing will depend on the skill of the decision maker. Here, we merely list some possibilities.

(i) Suppose that the problem can be converted into a single-criterion one as in (i) of Section 3.1.2. Then, it is a single-criterion problem with point-valued confidence structure, and the methods of Section 3.2.1 apply.

(ii) Suppose that the problem cannot be converted into a single-criterion one. Then, one can introduce a simple utility function for each outcome component. Namely, if a higher outcome value is preferred to a lower one, let

$$u_i(y)=y_i=f_i(x), \qquad i=1,2,\ldots,m.$$

Then, one proceeds as in Section 3.2.1 by forming expected values and variances

$$E_i(x)=\sum_{y\in Y_1(x)} u_i(y)M(x,y)+\int_{Y_2(x)} u_i(y)M(x,y)\,dy, \tag{5}$$

$$V_i(x)=\sum_{y\in Y_1(x)} [u_i(y)-E_i(x)]^2M(x,y)+\int_{Y_2(x)} [u_i(y)-E_i(x)]^2M(x,y)\,dy, \tag{6}$$

with $i=1,2,\ldots,m$, where $Y_1(x)$ and $Y_2(x)$ are defined by (3) and (4), respectively.

Thus, by doubling the number of criteria, one converts a problem with point-valued confidence structure into one with deterministic structure as in (ii) of Section 3.2.1. Of course, one can also convert the problem into an m-criteria one, with m chance constraints as in (iii) of Section 3.2.1. In either case, the problem is reduced to one of multicriteria with deterministic confidence structure, so that the methods of Section 3.1.2 become applicable.

In the remainder of this section, we describe a new type of nondominated decisions (that is, decisions resulting in nondominated outcomes) for problems with point-valued confidence structures.

(iii) As discussed in (iii) of Section 3.1.2, given y^1 and y^2 in Y, iff

$$y^2\in y^1+D(y^1),$$

then y^2 is dominated by y^1. Now, given decisions x^1 and x^2 in X with possible outcome sets $Y(x^1)$ and $Y(x^2)$, respectively, iff

$$y^2\in y^1+D(y^1) \qquad \text{for all } y^1\in Y(x^1) \text{ and } y^2\in Y(x^2),$$

then x^2 is dominated by x^1. Loosely speaking, the decision x^2 is dominated by the decision x^1 if every possible outcome of x^2 is dominated by every possible outcome of x^1. In Fig. 2, $Y(x^1)$ and a *constant* domination cone $D(y)$ are given. The decision x^2 is dominated by x^1 if $Y(x^2)$ is contained in region B, whereas x^1 is dominated by x^2 if $Y(x^2)$ is contained in region A.

The above definition of domination may be too restrictive. To make the concept less restrictive, let us introduce the following kind of domination. For each

$$x \in X, \qquad \alpha \in [0, 1], \qquad \beta \in [0, \infty),$$

let

$$Y_1^\alpha(x) = \{y \in Y_1(x) | M(x, y) \geq \alpha\}, \tag{7}$$

$$Y_2^\beta(x) = \{y \in Y_2(x) | M(x, y) \geq \beta\}, \tag{8}$$

where $Y_1(x)$ and $Y_2(x)$ are given by (3) and (4), respectively.[9] Loosely speaking, $Y_1^\alpha(x)[Y_2^\beta(x)]$ is the set of all outcomes of decision x having probability [probability density] equal to or greater than $\alpha[\beta]$. Now, given α and β, and a domination structure

$$D(\cdot): y \mapsto D(y) \subset R^m,$$

$x^2 \in X$ is dominated by $x^1 \in X$ with respect to $(\alpha, \beta, D(\cdot))$ iff

$$y^2 \in y^1 + D(y^1)$$

for all

$$y^2 \in Y_1^\alpha(x^2) \cup Y_2^\beta(x^2) \qquad \text{and } y^1 \in Y_1^\alpha(x^1) \cup Y_2^\beta(x^1).$$

[9] Recall Convention 2.1.

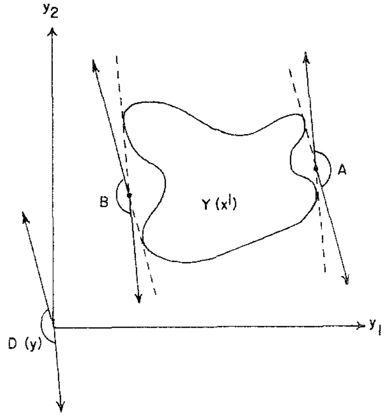

Fig. 2. Domination structure.

A nondominated decision is one which is not dominated by any other feasible decision. Roughly speaking, the domination relation is defined over those outcomes which have high enough probability of resulting from the decisions considered.

Example 3.1. Let

$$M(x^i, y), \qquad i = 1, 2,$$

be as given in Example 2.6. Let

$$\alpha = 0.05, \qquad \beta = 0.4.$$

Then,

$$Y_1^\alpha(x^1) = Y_2^\beta(x^2) = \varnothing,$$

and

$$Y_2^\beta(x^1) = \{y \mid \|y - y^1\| < 0.6\}, \qquad Y_1^\alpha(x^2) = \{y^{21}\}.$$

Figure 3 shows $Y_2^\beta(x^1)$ and $Y_1^\alpha(x^2)$ together with a constant domination cone $D(y)$.

For the given domination cone $D(y)$, the decision x^2 is dominated by x^1, so that x^1 is nondominated.

If

$$\alpha = 0.01, \qquad \beta = 0.4,$$

then

$$Y_1^\alpha(x^2) = \{y^{21}, y^{22}\}.$$

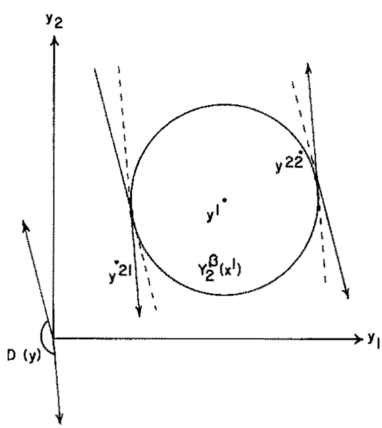

Fig. 3. Example 3.1.

x^2 is not dominated by x^1, nor is x^1 dominated by x^2. Thus, both decisions are nondominated in this case.

In general, the larger α and β are, the smaller is the set of nondominated decisions. If α and β are too large, $Y_1^\alpha(x)$ and $Y_2^\beta(x)$ may be empty, and the domination relation ceases to be meaningful.

3.3. General Confidence Structures. Recall Convention 2.1 for general confidence structures and Definition 2.6 for the reduction process of converting a general confidence structure into a point-valued one.

3.3.1. General Confidence Structure with a Single Criterion. In this case, one can convert the general confidence structure into a point-valued one by the reduction process discussed in Section 2. Thereafter, the methods of Section 3.2.1 are applicable.

There is one special case that is somewhat *equivalent* to a multicriteria problem with deterministic confidence structure. This case is characterized by the fact that the number of possible outcomes of each decision is finite and fixed. For example, in the SIP, if asset value is the only concern, the outcomes of each decision may depend either on a bearish market or a bullish one. Thus, the possible outcomes of each decision may be represented by a pair of real numbers. In general, if the outcomes of each decision x depend on m possible situations, they then may be represented by m real numbers $(f_1(x), \ldots, f_m(x))$. Now, let I_i be an interval of $[0, 1]$ indicating the confidence that the ith situation will occur. More precisely, let the confidence structure be given by

$$c(x, y) = \begin{cases} I_i, & \text{if } y = f_i(x), \\ \{0\}, & \text{otherwise.} \end{cases}$$

Now, let

$$\lambda = (\lambda_l, \ldots, \lambda_m)$$

and

$$\Omega = \left\{ \lambda \in R^m \,|\, \lambda_i \in I_i, \sum_{i=1}^{m} \lambda_i = 1 \right\}.$$

Thus, each $\lambda \in \Omega$ represents a prior probability that is consistent with the confidence structure, because each $\lambda_i \in I_i$. Note that, given a $\lambda \in \Omega$, the expected value of outcome for a given decision x is

$$\sum_{i=1}^{m} \lambda_i f_i(x).$$

Maximization of the expected value over X is equivalent to maximizing the value of the additive weight function $\lambda \cdot f(x)$ with $\lambda \in \Omega$. Let

$$\Lambda = \{\alpha\lambda \mid \alpha \geqq 0, \lambda \in \Omega\},$$

$$\Lambda^* = \{d \in R^m \mid d \cdot \lambda \leqq 0, \forall \lambda \in \Omega\};$$

hence, Λ^* is the polar cone of Λ. From Refs. 8 and 16, it is seen that the confidence structure induces a domination structure such that

$$\text{int}\, \Lambda^* \subset D(y)$$

for all $y \in Y$, where

$$Y = \{f(x) \mid x \in X\}.$$

As shown in Refs. 8 and 16, the advantage of using domination structures is that good candidates are not disregarded when Y does not possess suitable cone-convexity.

3.3.2. General Confidence Structure with a Multiple Criterion. This is the most general as well as the most common decision problem. It may be possible first to convert the multicriteria problem into one with a single criterion, and then apply the methods of Section 3.3.1. It may also be possible to use the reduction process of converting a general confidence structure into a point-valued one, and then to apply the methods of Section 3.2.2. A combination of these steps may be possible depending on the particular problem.

Another method may be an extension of that of (iii) of Section 3.2.2. Let $a(x, y)$ and $b(x, y)$ denote the greatest lower bound and the least upper bound, respectively, of the confidence interval $c(x, y)$. Analogously to (7) and (8), define

$$\tilde{Y}_1^{\alpha}(x) = \{y \in Y_1(x) \mid P(a(x, y), b(x, y)) \geqslant \alpha\}, \tag{9}$$

$$\tilde{Y}_2^{\beta}(x) = \{y \in Y_2(x) \mid P(a(x, y), b(x, y)) \geqslant \beta\}, \tag{10}$$

where

$$P(\cdot, \cdot): R_+^2 \to R_+.$$

For instance, if one applies the principle of insufficient reason,

$$P(a(x, y), b(x, y)) = \tfrac{1}{2}[a(x, y) + b(x, y)],$$

or one may let

$$P(a(x, y), b(x, y)) = b(x, y).$$

Now, we introduce the following value judgment. Given $P(\cdot,\cdot)$, α, β, and a domination structure $D(\cdot)$, $x^2 \in X$ is dominated by $x^1 \in X$ with respect to $(P(\cdot,\cdot), \alpha, \beta, D(\cdot))$ iff

$$y^2 \in y^1 + D(y^1)$$

for all

$$y^2 \in \tilde{Y}_1^{\alpha}(x^2) \cup \tilde{Y}_2^{\beta}(x^2) \qquad \text{and } y^1 \in \tilde{Y}_1^{\alpha}(x^1) \cup \tilde{Y}_2^{\beta}(x^1).$$

The proper specification of $P(\cdot,\cdot)$, α, β, and $D(\cdot)$ is clearly of great importance and remains a subject for further investigation.

4. Hierarchy of Decision Processes

In view of the discussion of Section 3, it appears reasonable to set up a hierarchy of decision processes. After the feasible decision set X is specified, a decision problem is characterized by its confidence structure and by the value judgment of the outcomes. The process begins at the most common starting point, multicriteria with a general confidence structure, and ends with adoption of a final decision. During the process, consecutive simplification of confidence structure and value judgment takes place. Figure 4 shows the direction of simplification as indicated by the arrow.

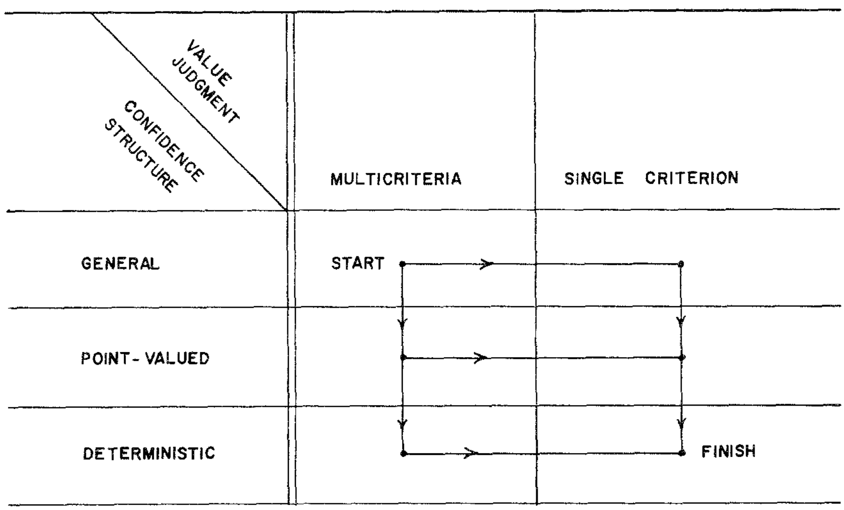

Fig. 4. Decision process.

5. Conclusions

The concept of confidence structure has been introduced into decision-making problems. Various concepts and techniques for simplifying and solving such problems have been discussed, and a hierarchy of decision processes has been outlined.

References

1. FERGUSON, T. S., *Mathematical Statistics, a Decision Theoretic Approach*, Academic Press, New York, New York, 1967.
2. RAIFFA, H., *Decision Analysis*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1970.
3. ZADEH, L. A., *Fuzzy Sets*, Information and Control, Vol. 8, pp. 338–353, 1965.
4. BELLMAN, R. E., and ZADEH, L. A., *Decision-Making in a Fuzzy Environment*, Management Science, Vol. 17, No. 4, 1970.
5. LUCE, R. D., and RAIFFA, H., *Games and Decisions*, John Wiley and Sons, New York, New York, 1967.
6. DEBREU, G., *Smooth Preferences*, Université Catholique de Louvain, CORE Discussion Paper No. 7203, Heverlee, Belgium, 1972.
7. STADLER, W., *Preference Optimality*, Multicriteria Decision Making, Edited by G. Leitmann and A. Marzollo, Springer-Verlag, New York, New York, 1975.
8. YU, P. L., *Domination Structures and Nondominated Solutions*, Multicriteria Decision Making, Edited by G. Leitmann and A. Marzollo, Springer-Verlag, New York, New York, 1975.
9. MACCRIMON, K. R., *An Overview of Multiple Objective Decision Making*, Multiple Criteria Decision Making, Edited by J. L. Cochrane and M. Zeleny, University of South Carolina Press, Columbia, South Carolina, 1973.
10. CHARNES, A., and COOPER, W. W., *Management Models and Industrial Applications of Linear Programming, Vol. 1*, John Wiley and Sons, New York, New York, 1961.
11. IJIRI, Y., *Management Goals and Accounting for Control*, North-Holland Publishing Company, Amsterdam, Holland, 1965.
12. YU, P. L., *A Class of Solutions for Group Decision Problems*, Management Science, Vol. 19, No. 8, 1973.
13. FREIMER, M., and YU, P. L., *Some New Results on Compromise Solutions*, Management Science, Vol. 22, No. 6, 1976.
14. FISHBURN, P. C., *Utility Theory for Decision Making*, John Wiley and Sons, New York, New York, 1970.
15. ROY, B., *How Outranking Relation Helps Multiple Criteria Decision Making*, Multiple Criteria Decision Making, Edited by J. L. Cochrane and M. Zeleny, University of South Carolina Press, Columbia, South Carolina, 1973.

16. YU, P. L., *Cone Convexity, Cone Extreme Points, and Nondominated Solutions in Decision Problems with Multiobjectives*, Journal of Optimization Theory and Applications, Vol. 14, No. 3, 1974.
17. YU, P. L., and LEITMANN, G., *Compromise Solutions, Domination Structures, and Salukvadze's Solution*, Journal of Optimization Theory and Applications, Vol. 13, No. 3, 1974.
18. GEOFFRION, A. M., DYER, J. S., and FEINBERG, A., *An Interactive Approach for Multicriterion Optimization with an Application to Operation of an Academic Department*, Management Science, Vol. 19, pp. 357–368, 1972.
19. SHARPE, W., *Portfolio Theory and Capital Markets*, McGraw-Hill Book Company, New York, New York, 1970.
20. BERGSTRESSER, K., CHARNES, A., and YU, P. L., *Generalization of Domination Structures and Nondominated Solutions in Multicriteria Decision Making*, Journal of Optimization Theory and Applications, Vol. 18, No. 1, 1976.
21. CHARNES, A., and COOPER, W. W., *Chance-Constrained Programming*, Management Science, Vol. 6, No. 1, 1959.