Proceedings of the 2007 IEEE Symposium on Computational Intelligence in Multicriteria Decision Making (MCDM 2007)

Multicriterion Decision Making with Dependent Preferences

Wynn C. Stirling, Richard L. Frost, Matthew S. Nokleby, and Yabing Luo
Electrical and Computer Engineering Department, Brigham Young University, Provo, Utah, 84602, USA
Email: {wynn, rickf, nokleby}@ee.byu.edu, ice@byu.net

***Abstract*—If preferential independence is assumed inappropriately when developing multicriterion search methods, biased results may occur. A new axiomatic approach to defining conditional preference orderings that naturally accounts for preferential dependencies is presented and illustrated. This approach applies both to scalar optimization techniques that identify a best solution and to evolutionary optimization approaches that approximate the Pareto frontier.**

I. Introduction

The classical approach to multicriterion decision making is *a priori optimization*, which is to define a scalar utility function as an aggregation of individual utility functions, and then search for the extrema of this function. Recent emphasis has also focused on *a posteriori optimization*, which involves searching for an entire family of solutions—the Pareto optimal set—by the application of evolutionary algorithms that exploit dominance to guide the search for the Pareto frontier. However, if the search algorithms employ the assumption of mutual preferential independence, then errors can be introduced that are difficult to quantify if that assumption is not appropriate.

As a simple example of preferential dependence, consider the following scenario introduced by [1, p. 232]. A farmer has preferences for various amounts of rain and sunshine because of the impact on his crops. However, his preferences for various amounts of sunshine will be different, depending on the amount of rain. Thus, the farmer cannot independently define his preference orderings for the two attributes.

This paper provides an alternative approach to both scalar optimization and dominance that is explicitly designed to account for preferential dependencies. We first provide a brief summary of preferential independence and then introduce a new axiomatic approach that accounts for such dependencies. We then define a scalar utility function and a concept of dominance, each of which accounts for preferential dependencies. We finish with a brief example and conclusions.

II. Background

A. Preferential Independence

Let A be a space of feasible alternatives, and suppose there are $n \geq 2$ distinct criteria that must be considered when choosing $a \in A$. In the parlance of multicriterion decision theory (see [1]), an *attribute* X_i is a function associated with the ith criterion that maps each $a \in A$ to the corresponding *consequence space* $\mathcal{X}_i$; that is, $X_i(a) = x_i \in \mathcal{X}_i$. Let the Cartesian product space $\boldsymbol{\mathcal{X}}_n = \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$ denote the n-dimensional consequence space, and let $[X_1(a), \ldots, X_n(a)] = (x_1, \ldots, x_n) \in \boldsymbol{\mathcal{X}}_n$ denote a *consequence vector* for the alternative a. It is convenient to denote elements of $\boldsymbol{\mathcal{X}}_n$ as vectors using the notation $\mathbf{x}_n = (x_1, \ldots, x_n)$.

A critical consideration when considering an alternative is to appreciate the significance in $\boldsymbol{\mathcal{X}}_n$ of its adoption. A *preference ordering* $\succeq_{\boldsymbol{\mathcal{X}}_n}$ is a binary relation over the vector consequences such that $\mathbf{x}_n \succeq_{\boldsymbol{\mathcal{X}}_n} \mathbf{x}'_n$ if and only $\mathbf{x}_n$ is considered, from the point of view of all consequences considered simultaneously, to be a least as good as $\mathbf{x}'_n$. Notationally, we write $\mathbf{x}_n \succ_{\boldsymbol{\mathcal{X}}_n} \mathbf{x}'_n$ if $\mathbf{x}_n$ is strictly better than $\mathbf{x}'_n$, and we write $\mathbf{x}_n \sim_{\boldsymbol{\mathcal{X}}_n} \mathbf{x}'_n$ if $\mathbf{x}_n$ and $\mathbf{x}'_n$ are equivalent. We shall assume that $\succeq_{\boldsymbol{\mathcal{X}}_n}$ is reflexive, antisymmetric, transitive, complete, and continuous.

The reason multicriterion theory is a challenging endeavor is that such an ordering can be difficult to comprehend and even more difficult to define unambiguously, due to conflicts. As Keeney and Raiffa observe, it is very difficult for a human decision maker to specify a preference order for more than two attributes under certainty or more than one attribute under risk because of the need to consider both preference orderings and probabilities simultaneously [1, p. 311]. Because of such conceptual difficulties, the decision maker is strongly motivated to formulate preference orderings over subspaces of $\boldsymbol{\mathcal{X}}_n$ and then aggregate them to form a decision. This approach is appropriate if it is possible to decompose the consequence space into subspaces that do not depend on each other. To this end, Keeney and Raiffa have employed the concept of preferential independence [1], which we briefly review.

Definition 2.1: Let $\boldsymbol{\mathcal{X}}_m = \{\mathcal{X}_{i_1} \times \cdots \times \mathcal{X}_{i_m}\}$ and $\boldsymbol{\mathcal{X}}_k = \{\mathcal{X}_{j_1} \times \cdots \times \mathcal{X}_{j_k}\}$ be m- and k-dimensional disjoint subspaces of $\boldsymbol{\mathcal{X}}_n$ (i.e., $\boldsymbol{\mathcal{X}}_m \cap \boldsymbol{\mathcal{X}}_k = \varnothing$) such that $m + k = n$. $\boldsymbol{\mathcal{X}}_k$ is the *complementary subspace* of $\boldsymbol{\mathcal{X}}_m$.

If $\boldsymbol{\mathcal{X}}_m$ and $\boldsymbol{\mathcal{X}}_k$ are complementary subspaces of $\boldsymbol{\mathcal{X}}_n$, we write $\boldsymbol{\mathcal{X}}_n = \boldsymbol{\mathcal{X}}_m \times \boldsymbol{\mathcal{X}}_k$, and elements of $\boldsymbol{\mathcal{X}}_n$ may be expressed as $\mathbf{x}_n = (\mathbf{x}_m, \mathbf{x}_k)$, where $\mathbf{x}_m = (x_{i_1}, \ldots, x_{i_m}) \in \boldsymbol{\mathcal{X}}_m$ and $\mathbf{x}_k = (x_{j_1}, \ldots x_{j_k}) \in \boldsymbol{\mathcal{X}}_k$. (Since the indexing of elements of $\boldsymbol{\mathcal{X}}_n$ is arbitrary, it may be necessary to permute the indices to obtain this decomposition.)

Definition 2.2: Let $\boldsymbol{\mathcal{X}}_m$ and $\boldsymbol{\mathcal{X}}_k$ be complementary subspaces of $\boldsymbol{\mathcal{X}}_n$. The consequence vector $\mathbf{x}_m \in \boldsymbol{\mathcal{X}}_m$ is *conditionally at least as good as* $\mathbf{x}'_m \in \boldsymbol{\mathcal{X}}_m$ given $\mathbf{x}_k \in \boldsymbol{\mathcal{X}}_k$ if $(\mathbf{x}_m, \mathbf{x}_k) \succeq_{\boldsymbol{\mathcal{X}}_n} (\mathbf{x}'_m, \mathbf{x}_k)$. The subspace $\boldsymbol{\mathcal{X}}_m$ is *preferentially independent* of $\boldsymbol{\mathcal{X}}_k$ if, for any two vectors $\mathbf{x}_m, \mathbf{x}'_m \in \boldsymbol{\mathcal{X}}_m$ such that $(\mathbf{x}_m, \mathbf{x}'_k) \succeq_{\boldsymbol{\mathcal{X}}_n} (\mathbf{x}'_m, \mathbf{x}'_k)$ for some $\mathbf{x}'_k \in \boldsymbol{\mathcal{X}}_k$, then

1-4244-0702-8/07/$20.00 ©2007 IEEE

$(\mathbf{x}_m, \mathbf{x}_k) \succeq_{\boldsymbol{\mathcal{X}}_n} (\mathbf{x}'_m, \mathbf{x}_k)$ for all $\mathbf{x}_k \in \boldsymbol{\mathcal{X}}_k$.

Preferential independence is not symmetric; that is, $\boldsymbol{\mathcal{X}}_m$ being preferentially independent of $\boldsymbol{\mathcal{X}}_k$ does not imply that $\boldsymbol{\mathcal{X}}_k$ is preferentially independent of $\boldsymbol{\mathcal{X}}_m$.

Let $\boldsymbol{\mathcal{X}}_m$ and $\boldsymbol{\mathcal{X}}_k$ be complementary subspaces of $\boldsymbol{\mathcal{X}}_n$. If $\boldsymbol{\mathcal{X}}_m$ is preferentially independent of $\boldsymbol{\mathcal{X}}_k$, then a preference ordering $\succeq_{\boldsymbol{\mathcal{X}}_m}$ over $\boldsymbol{\mathcal{X}}_m$ is induced by the equivalence $\mathbf{x}_m \succeq_{\boldsymbol{\mathcal{X}}_m} \mathbf{x}'_m$ if and only if $(\mathbf{x}_m, \mathbf{x}_k) \succeq_{\boldsymbol{\mathcal{X}}_n} (\mathbf{x}'_m, \mathbf{x}_k)$.

Definition 2.3: The members of $\boldsymbol{\mathcal{X}}_n$ are said to be *mutually preferentially independent* if every subspace of $\boldsymbol{\mathcal{X}}_n$ is preferentially independent of its complementary subspace.

The condition that each $\mathcal{X}_i$ is preferentially independent of its complementary subspace (a weaker condition than mutual preferential independence) induces an individual preference ordering $\succeq_{\mathcal{X}_i}$ over each $\mathcal{X}_i$. Given this condition, we may define the conventional notion of dominance.

Definition 2.4: Suppose individual orderings $\succeq_{\mathcal{X}_i}$ exist for $i = 1, \ldots, n$. The consequence vector $\mathbf{x}_n = (x_1, \ldots, x_n)$ *dominates* $\mathbf{x}'_n = (x'_1, \ldots, x'_n)$ if $x_i \succeq_{\mathcal{X}_i} x'_i$, $i = 1, \ldots, n$ and $x_j \succ_{\mathcal{X}_j} x'_j$, for some $j \in \{1, \ldots, n\}$.

Definition 2.5: The *Pareto frontier* is the set of all nondominated consequence vectors. The *Pareto optimal set* is the set of all $a \in A$ such that $(X_1(a), \ldots, X_n(a))$ is an element of the Pareto frontier.

B. Aggregation

The aggregation problem is to combine individual preference orderings to form a group preference ordering. A well-known result is Arrow's Impossibility Theorem [2], which states that it is generally impossible to combine a set of individual preference orderings to form a group preference ordering that satisfies a set of arguably reasonable or desirable properties (monotonicity, independence, unanimity, and non-dictatorship). Since Arrow's formulation involves only ordinal preference rankings, it does not permit interpersonal comparisons of preferences. By relaxing this prohibition through the introduction of numerical utilities, however, it becomes possible to form group preference orderings.

Definition 2.6: A function $p_{\boldsymbol{\mathcal{X}}_n} \colon \boldsymbol{\mathcal{X}}_n \to \mathbb{R}$ is a *utility* over $\boldsymbol{\mathcal{X}}_n$ corresponding to a preference ordering $\succeq_{\boldsymbol{\mathcal{X}}_n}$ if $\mathbf{x}_n \succ_{\boldsymbol{\mathcal{X}}_n} \mathbf{x}'_n$ if and only if $p_{\boldsymbol{\mathcal{X}}_n}(\mathbf{x}_n) > p_{\boldsymbol{\mathcal{X}}_n}(\mathbf{x}'_n)$ and $\mathbf{x}_n \sim_{\boldsymbol{\mathcal{X}}_n} \mathbf{x}'_n$ if and only if $p_{\boldsymbol{\mathcal{X}}_n}(\mathbf{x}_n) = p_{\boldsymbol{\mathcal{X}}_n}(\mathbf{x}'_n)$.

If $\boldsymbol{\mathcal{X}}_m$ and $\boldsymbol{\mathcal{X}}_k$ are preferentially independent of their complementary subspaces, then subspace orderings $\succeq_{\boldsymbol{\mathcal{X}}_m}$ and $\succeq_{\boldsymbol{\mathcal{X}}_k}$ can be defined over $\boldsymbol{\mathcal{X}}_m$ and $\boldsymbol{\mathcal{X}}_k$, respectively. Utilities $p_{\boldsymbol{\mathcal{X}}_m}$ over $\boldsymbol{\mathcal{X}}_m$ and $p_{\boldsymbol{\mathcal{X}}_k}$ over $\boldsymbol{\mathcal{X}}_k$ may then be defined that correspond to these preference orderings. Given these utilities, however, there is not a unique way to aggregate them to form a utility $p_{\boldsymbol{\mathcal{X}}_m \times \boldsymbol{\mathcal{X}}_k}$ over $\boldsymbol{\mathcal{X}}_m \times \boldsymbol{\mathcal{X}}_k$. Rather than there being too few ways (none) to aggregate the ordinal preference relations, there are now too many ways (infinite) to aggregate the utility functions. For example, any positively weighted sum of the individual utilities defines an ordering over $\boldsymbol{\mathcal{X}}_m \times \boldsymbol{\mathcal{X}}_k$ that satisfies all of Arrow's conditions except for the prohibition against interpersonal comparisons of utility. Thus, there exist many functions of utilities over $\boldsymbol{\mathcal{X}}_m$ and $\boldsymbol{\mathcal{X}}_k$ that define a preference ordering over $\boldsymbol{\mathcal{X}}_m \times \boldsymbol{\mathcal{X}}_k$. If each $\mathcal{X}_i$ is preferentially independent of its complement, then we may define a multiattribute utility function as

$$p_{\boldsymbol{\mathcal{X}}_n}(\mathbf{x}_n) = f[p_{\mathcal{X}_1}(x_1), \ldots, p_{\mathcal{X}_n}(x_n)], \tag{1}$$

where each $p_{\mathcal{X}_i}$ is a utility over the corresponding one-dimensional consequence subspace $\mathcal{X}_i$, $i = 1, \ldots, n$. A classical result of mutual preferential independence is the *additive value theorem* [3], which states that mutual preferential independence is necessary and sufficient to define a multiattribute utility function $p_{\boldsymbol{\mathcal{X}}_n}$ of the form

$$p_{\boldsymbol{\mathcal{X}}_n}(\mathbf{x}_n) = \sum_{i=1}^{n} w_i p_{\mathcal{X}_i}(x_i), \tag{2}$$

with weighting factors $w_i > 0$ for $1, \ldots, n$ and $\sum_{i=1}^{n} w_i = 1$.

A classical way to search for the Pareto frontier is to perform scalar optimization for various combinations of weighting factors. This approach, however, assumes that the consequence space is convex ([4, 5]). Recent research has focused on population-wide searches using evolutionary algorithms to generate an entire family of solutions that approximate the Pareto frontier without requiring the convexity assumption. The key feature of these search procedures is the use of dominance as the criterion for survival [6, 7]. Although these approaches do not require a scalar utility function to be defined, they do require a well-defined notion of dominance. With the conventional notion of dominance expressed in terms of individual preferences, each one-dimensional consequence subspace must be preferentially independent of its complementary subspace.

Preferential independence is a stringent assumption, and if it is not appropriate, then neither a scalar utility function of the form (1) nor dominance in terms of individual preference orderings can be defined, and the multicriterion decision problem becomes more difficult. Assuming preferential independence inappropriately can lead to errors that are difficult to quantify.

C. Conditional Preferential Independence

Keeney and Raiffa suggest a natural way to weaken the hypothesis of preferential independence by introducing the concept of conditional preferential independence.

Definition 2.7: Let $\boldsymbol{\mathcal{X}}_m$, $\boldsymbol{\mathcal{X}}_k$, and $\boldsymbol{\mathcal{X}}_\ell$ be disjoint subspaces of $\boldsymbol{\mathcal{X}}_n$ such that $\boldsymbol{\mathcal{X}}_m \times \boldsymbol{\mathcal{X}}_k \times \boldsymbol{\mathcal{X}}_\ell$ is preferentially independent of its complementary subspace. The subspace $\boldsymbol{\mathcal{X}}_m$ is *conditionally preferentially independent of* $\boldsymbol{\mathcal{X}}_k$ *given* $\mathbf{x}_\ell \in \boldsymbol{\mathcal{X}}_\ell$ if, for any two vectors $\mathbf{x}_m, \mathbf{x}'_m \in \boldsymbol{\mathcal{X}}_m$ and fixed $\mathbf{x}_\ell \in \boldsymbol{\mathcal{X}}_\ell$ such that $(\mathbf{x}_m, \mathbf{x}'_k, \mathbf{x}_\ell) \succeq_{\boldsymbol{\mathcal{X}}_m \times \boldsymbol{\mathcal{X}}_k \times \boldsymbol{\mathcal{X}}_\ell} (\mathbf{x}'_m, \mathbf{x}'_k, \mathbf{x}_\ell)$ for some $\mathbf{x}'_k \in \boldsymbol{\mathcal{X}}_k$, then $(\mathbf{x}_m, \mathbf{x}_k, \mathbf{x}_\ell) \succeq_{\boldsymbol{\mathcal{X}}_m \times \boldsymbol{\mathcal{X}}_k \times \boldsymbol{\mathcal{X}}_\ell} (\mathbf{x}'_m, \mathbf{x}_k, \mathbf{x}_\ell)$ for all $\mathbf{x}_k \in \boldsymbol{\mathcal{X}}_k$. If this relation holds for all $\mathbf{x}_\ell \in \boldsymbol{\mathcal{X}}_\ell$, then $\boldsymbol{\mathcal{X}}_m$ is *conditionally preferentially independent from* $\boldsymbol{\mathcal{X}}_k$ *given* $\boldsymbol{\mathcal{X}}_\ell$.

The notion of conditional preferential independence has inspired several researchers to develop graphical models to represent multivariate utility functions and simplify calculations. Bacchus and Grove [8, 9] decompose the global utility function into sums of utility functions over conditionally independent subspaces of consequences. They extend this notion to define generalized additive independence, which permits the attributes to be expressed graphically. Boutilier, Brafman, Hoos, and Poole [10] extend this work to develop a Conditional Preference (CP) network involving directed graphs that represent the ordinal ranking of the attributes in terms of their importance. These relationships can be expressed graphically by associating a vertex with each attribute and associating an edge with each parent-child pair, thereby forming a directed graph. Boutilier, Bacchus, and Brafman [11] then combine CP nets with generalized additive independence to generate a directed graph that quantifies the preferential relationships with numerical values. The resulting UCP (Utility Conditional Preference) network is a directed acyclic graph with attributes as vertices and edges associated with the conditional utility functions for the children given the instantiation of the parents. Gonzales and Perny [12] have also applied generalized additive independence decompositions to construct graphical models to assist in the elicitation of preferences for risky multiattribute decision problems. Related work by Engel and Wellman [13] introduces the notion of conditional utility independence and shows this leads to the calculation of joint utility as the sum of conditional utility functions.

A somewhat different approach is taken by La Mura and Shoham [14], who introduce nonconventional notions of probabilistic independence and utility independence. Using these constructions, they define a hybrid graphical model called an Expected Utility Network (EUN) composed of attributes as vertices and with two sets of edges—one for probabilistic dependencies and one for utility dependencies.

III. An Axiomatic Approach to Aggregation

As is evident from the above discussion, much theoretical effort has focused on justifying and exploiting preferential independence and conditional preferential independence to simplify the construction of the multiattribute utility function and to define dominance. *This paper, however, reorients the problem by focusing directly on ways to accommodate preferential dependencies.* We present two fundamental axioms that we suggest should constrain the structure of preference orderings, and then we define preference orderings and utility functions that comply with these axioms.

A. Fundamental Axioms

As discussed above, classical multicriterion decision theory recognizes concepts of conditional preferences and preferential independence. These concepts are analogous to the corresponding concepts of probability theory. The main distinction is that dealing with attributes and criteria is a *praxeological*[1] consideration, whereas dealing with randomness and uncertainty is an *epistemological*[2] consideration. Two random phenomena are epistemologically (i.e., statistically) independent if the probability of the occurrence of one is not influenced by the occurrence of the other; two consequences are praxeologically independent if the utility of the instantiation of one is not influenced by the instantiation of the other.

The view of this paper is that the above analogy between the preferential and the statistical is more profound than is generally appreciated. Multicriterion decision theory considers aggregating relationships among praxeological attributes—efficiency and effectiveness, and probability theory considers aggregating relationships among epistemological attributes—knowledge and belief. The perhaps surprising thesis of this paper is that they can both be characterized by the same mathematical structure. To motivate this thesis, consider the following aggregation axioms.

Axiom 1 (Conditioning) *Preferences for consequences of a multicriterion decision problem may be conditioned on the preferences for other consequences.*

Axiom 2 (Endogeny) *If preference orderings exist for a space of consequences, they must be determined by relationships that exist among its subspaces.*

Axiom 1 represents an important shift in perspective from classical decision theory. With the classical approach, an attribute's preference ordering is with respect to the *instantiation of* consequences to itself and to other attributes [1, 10, 13]. This axiom, however, permits an attribute's preference ordering to be with respect to the *preference for* consequences to itself and to other attributes. If the consequences are mutually preferentially independent, then this ordering should be the same as with the conventional case. But if mutual independence does not apply, then the attribute's preferences can be dependent on the benefit to other attributes as well as on the benefit to the given attribute. For example, consider the farmer scenario introduced in Section I. Suppose, for health reasons and when viewed in isolation of his crop's success, he prefers much rain to little rain. Given that, he would naturally be inclined to consider his preference for much sun or little in the light of his preferences for rain.

Definition 3.1: Let $\boldsymbol{\mathcal{X}}_k$ be a subspace of $\boldsymbol{\mathcal{X}}_n$ that is preferentially independent of its complementary subspace, and let $\succeq_{\boldsymbol{\mathcal{X}}_k}$ be the preference ordering over $\boldsymbol{\mathcal{X}}_k$. A consequence vector $\mathbf{x}_k \in \boldsymbol{\mathcal{X}}_k$ is a *commitment* if $\mathbf{x}_k \succ_{\boldsymbol{\mathcal{X}}_k} \mathbf{x}'_k$ for all $\mathbf{x}'_k \in \boldsymbol{\mathcal{X}}_k \backslash \{\mathbf{x}_k\}$ and $\mathbf{x}'_k \sim_{\boldsymbol{\mathcal{X}}_k} \mathbf{x}''_k$ for all $\mathbf{x}'_k, \mathbf{x}''_k \in \boldsymbol{\mathcal{X}}_k \backslash \{\mathbf{x}_k\}$,where $\{\cdot\}$ denotes a singleton set.

Essentially, a commitment to $\mathbf{x}_k$ means that it is the *only* preferred consequence. All other consequences are equally inferior.

Definition 3.2: Let $\boldsymbol{\mathcal{X}}_m$ and $\boldsymbol{\mathcal{X}}_k$ be disjoint subspaces of consequences, with $m + k \le n$, such that $\boldsymbol{\mathcal{X}}_k$ is preferentially independent of its complementary subspace, and $\boldsymbol{\mathcal{X}}_m \times \boldsymbol{\mathcal{X}}_k$ is preferentially independent of its complementary subspace. A *conditional preference ordering* over $\boldsymbol{\mathcal{X}}_m$ given $\boldsymbol{\mathcal{X}}_k$ is a binary relation $\succeq_{\boldsymbol{\mathcal{X}}_m|\boldsymbol{\mathcal{X}}_k}$ such that $\mathbf{x}_m|\mathbf{x}_k \succeq_{\boldsymbol{\mathcal{X}}_m|\boldsymbol{\mathcal{X}}_k} \mathbf{x}'_m|\mathbf{x}'_k$ means $\boldsymbol{\mathcal{X}}_m$ considers $\mathbf{x}_m$ to be at least as good, given that $\boldsymbol{\mathcal{X}}_k$ is committed to $\mathbf{x}_k$, as $\mathbf{x}'_m$, given that $\boldsymbol{\mathcal{X}}_k$ is committed to $\mathbf{x}'_k$.

It is important to appreciate that a conditional ordering over $\boldsymbol{\mathcal{X}}_m$ is hypothetical; it does not depend on the actual

[1] Praxeology is the classification of choices on the basis of effectiveness and efficiency.

[2] Epistemology the classification of choices on the basis of knowledge and belief.

preference ordering over $\mathcal{X}_k$, nor does it mean that the ordering over $\mathcal{X}_k$ actually is a commitment to any consequence. Rather, it means that if $\mathbf{x}_k$ were a commitment, then the ordering over $\mathcal{X}_m$ would be as defined. To illustrate, suppose a committee is to buy a car, and committee members Y and Z have interest restricted to the attributes of model and color, respectively. A possible conditional preference would be that Y prefers convertibles given that Z is committed to red cars, to sedans given that Z is committed to green cars. A special case obtains if Y prefers convertibles to sedans given that Z is committed to red cars. The critical feature of this conditional preference structure is that it provides an ordering conditioned on preference, not on actual instantiation. We can view a conditional preference ordering in two ways. First, it may reflect Y's preference ordering out of its own selfish interest. Second, it may reflect Y's willingness to give deference to Z when defining its preferences (either benevolently or malevolently). Thus, this structure provides the opportunity for a social relationship to exist between Y and Z. It does not, however, require or impose a social relationship.

We now consider the problem of defining utility functions that correspond to conditional preference relations.

Definition 3.3: If a subspace ordering $\succeq_{\mathcal{X}_k}$ exists over $\mathcal{X}_k$, a *marginal utility* over $\mathcal{X}_k$ is a function $p_{\mathcal{X}_k}\colon \mathcal{X}_k \to \mathbb{R}$ such that $p_{\mathcal{X}_k}(\mathbf{x}_k) > p_{\mathcal{X}_k}(\mathbf{x}'_k)$ if and only if $\mathbf{x}_k \succ_{\mathcal{X}_m} \mathbf{x}'_k$ and $p_{\mathcal{X}_k}(\mathbf{x}_k) = p_{\mathcal{X}_k}(\mathbf{x}'_k)$ if and only if $\mathbf{x}_k \sim_{\mathcal{X}_m} \mathbf{x}'_k$.

Definition 3.4: If a conditional preference ordering exists over $\mathcal{X}_m$ given $\mathcal{X}_k$, a *conditional utility* over $\mathcal{X}_m$ given $\mathcal{X}_k$ is a function $p_{\mathcal{X}_m|\mathcal{X}_k}\colon \mathcal{X}_m \times \mathcal{X}_k \to \mathbb{R}$ such that $p_{\mathcal{X}_m|\mathcal{X}_k}(\mathbf{x}_m|\mathbf{x}_k) > p_{\mathcal{X}_m|\mathcal{X}_k}(\mathbf{x}'_m|\mathbf{x}'_k)$ if and only if $\mathbf{x}_m|\mathbf{x}_k \succ_{\mathcal{X}_m|\mathcal{X}_k} \mathbf{x}'_m|\mathbf{x}'_k$ and $p_{\mathcal{X}_m|\mathcal{X}_k}(\mathbf{x}_m|\mathbf{x}_k) = p_{\mathcal{X}_m|\mathcal{X}_k}(\mathbf{x}'_m|\mathbf{x}'_k)$ if and only if $\mathbf{x}_m|\mathbf{x}_k \sim_{\mathcal{X}_m|\mathcal{X}_k} \mathbf{x}'_m|\mathbf{x}'_k$.

The next step in our development is to impose requirements on the utility functions that ensure compliance with Axiom 2. Suppose we are given the preference ordering $\succeq_{\mathcal{X}_k}$ over $\mathcal{X}_k$ and the conditional preference ordering $\succeq_{\mathcal{X}_m|\mathcal{X}_k}$ over $\mathcal{X}_m$ given $\mathcal{X}_k$. The aggregation problem is to define a preference ordering $\succeq_{\mathcal{X}_m \times \mathcal{X}_k}$ over the product space of consequences $\mathcal{X}_m \times \mathcal{X}_k$. Axiom 2 prohibits an exogenous source from defining this ordering; rather, the aggregation must emerge from the relationships that exist between the two sets of consequences. Complying with this axiom thus requires that the utility function over $\mathcal{X}_m \times \mathcal{X}_k$ must be a function of the marginal utility function over $\mathcal{X}_k$ and the conditional utility function over $\mathcal{X}_m$ given $\mathcal{X}_k$.

Definition 3.5: Let $\mathcal{X}_m$ and $\mathcal{X}_k$ be two disjoint subspaces of consequences, with $m + k \leq n$, such that $\mathcal{X}_k$ is preferentially independent of its complementary subspace, and $\mathcal{X}_m \times \mathcal{X}_k$ is preferentially independent of its complementary subspace. The utility functions $p_{\mathcal{X}_k}$ and $p_{\mathcal{X}_m|\mathcal{X}_k}$ are *endogenously aggregated* if there exists a function F, nondecreasing in both arguments, such that

$$p_{\mathcal{X}_m \times \mathcal{X}_k}(\mathbf{x}_m, \mathbf{x}_k) = F[p_{\mathcal{X}_k}(\mathbf{x}_k), p_{\mathcal{X}_m|\mathcal{X}_k}(\mathbf{x}_m|\mathbf{x}_k)], \qquad (3)$$

is a utility over $\mathcal{X}_m \times \mathcal{X}_k$, and hence defines a preference ordering $\succeq_{\mathcal{X}_m \times \mathcal{X}_k}$.

A reasonable and intuitively important property of this type of aggregation is that, if $\mathbf{x}_m|\mathbf{x}_k \succeq_{\mathcal{X}_m|\mathcal{X}_k} \mathbf{x}'_m|\mathbf{x}'_k$ and $\mathbf{x}_k \sim_{\mathcal{X}_k} \mathbf{x}'_k$, then $(\mathbf{x}_m, \mathbf{x}_k) \succeq_{\mathcal{X}_m \times \mathcal{X}_k} (\mathbf{x}'_m, \mathbf{x}'_k)$, and, if $\mathbf{x}_k \succeq_{\mathcal{X}_k} \mathbf{x}'_k$ with $\mathbf{x}_m|\mathbf{x}_k \sim_{\mathcal{X}_m|\mathcal{X}_k} \mathbf{x}'_m|\mathbf{x}'_k$, then $(\mathbf{x}_m, \mathbf{x}_k) \succeq_{\mathcal{X}_m \times \mathcal{X}_k} (\mathbf{x}'_m, \mathbf{x}'_k)$. These relationships obtain if and only if F is nondecreasing in both arguments.

B. The Aggregation Theorem

Probability theory is traditionally concerned with epistemology and is used to express the degrees of belief regarding the truth of propositions. Mathematically, however, probability theory is a sub-topic of general measure theory [15]. We may also apply measure theory to the praxeological problem to express the degrees of suitability of choices. In the interest of brevity, we restrict our attention to the discrete case involving a finite alternative space and a finite consequence space.

Let $\mathcal{X}_n$ be a finite collection of n-dimensional consequence vectors. A *praxeological measure*[3] $P_{\mathcal{X}_n}$ is a function defined over a Boolean algebra $\mathcal{B}$ of subsets of $\mathcal{X}_n$ such that

1) $P_{\mathcal{X}_n}(A) \geq 0$ for all $A \in \mathcal{B}$.
2) $P_{\mathcal{X}_n}(\mathcal{X}_n) = 1$.
3) For any finite collection of pairwise disjoint elements $\{A_i\}$ of $\mathcal{B}$, $P_{\mathcal{X}_n}\left(\bigcup_i A_i\right) = \sum_i P_{\mathcal{X}_n}(A_i)$

Given a praxeological measure $P_{\mathcal{X}_n}$ over the power set of $\mathcal{X}_n$, we may construct a multivariate mass function $p_{\mathcal{X}_n}$, a family of marginal mass functions $\{p_{\mathcal{X}_m}, m = 1, 2, \ldots, n-1\}$, and a family of conditional mass functions $\{p_{\mathcal{X}_m|\mathcal{X}_k}, m, k = 1, \ldots, n-1\}, \mathcal{X}_m \cap \mathcal{X}_k = \varnothing, m + k \leq n\}$, such that, for all $\mathbf{x}_n \in \mathcal{X}_n$, $\mathbf{x}_m \in \mathcal{X}_m$, and $\mathbf{x}_k \in \mathcal{X}_k$,

$$\begin{aligned}
p_{\mathcal{X}_n}(\mathbf{x}_n) &= P_{\mathcal{X}_n}(\{\mathbf{x}_n\}) \\
p_{\mathcal{X}_m}(\mathbf{x}_m) &= P_{\mathcal{X}_n}(\{\mathbf{x}_m\} \times \mathcal{X}_k) \\
p_{\mathcal{X}_m|\mathcal{X}_k}(\mathbf{x}_m|\mathbf{x}_k) &= \frac{P_{\mathcal{X}_n}[\{\mathbf{x}_m\}\times\{\mathbf{x}_k\}\times(\mathcal{X}_n \setminus (\mathcal{X}_m \times \mathcal{X}_k))]}{P_{\mathcal{X}_n}[\{\mathbf{x}_k\} \times (\mathcal{X}_n \setminus \mathcal{X}_k)]},
\end{aligned}$$

where the last relation is defined if and only if the denominator is nonzero.

Theorem 3.6: **(The aggregation theorem)** *Let $\mathcal{X}_n$ be a finite collection of n-dimensional consequence vectors, and let $\mathcal{B}$ denote the power set of $\mathcal{X}_n$. Let $\{p_{\mathcal{X}_m}, \mathcal{X}_m \in \mathcal{B}, m = 1, \ldots, n\}$ be a family of utility functions associated with the subspaces of $\mathcal{X}_n$ and let $\{p_{\mathcal{X}_m|\mathcal{X}_k}, \mathcal{X}_m \cap \mathcal{X}_k = \varnothing, m+k \leq n\}$ be a family of conditional utility functions associated with all pairs of disjoint subspaces of $\mathcal{X}_n$. These utility functions are endogenously aggregated if and only if they are mass functions corresponding to a praxeological measure $P_{\mathcal{X}_n}$ over $\mathcal{B}$.*

A proof of this theorem is provided in the Appendix. This theorem formalizes the analogy between the epistemic and the praxeic. Indeed, we may view the alternative set A as analogous to a sample space, the attribute functions X_i as analogous to random variables, and the multiattribute utility function $p_{\mathcal{X}_n}$ as analogous to a multivariate probability mass function. It provides a complete valuation of all consequences, including

[3]Mathematically, this is same as a probability measure, but with different semantics.

the relationships between them. The aggregation theorem establishes that this function possess the same mathematical structure as multivariate probability mass functions, albeit with praxeological, rather than epistemological, semantics.

- *Non-negativity*: $p_{\boldsymbol{\mathcal{X}}_n}(\mathbf{x}_n) \geq 0 \; \forall \mathbf{x}_n \in \boldsymbol{\mathcal{X}}_n$.
- *Normalization*: $\sum_{\mathbf{x}_n \in \boldsymbol{\mathcal{X}}_n} p_{\boldsymbol{\mathcal{X}}_n}(\mathbf{x}_n) = 1$.
- *Marginalization*: Let $\boldsymbol{\mathcal{X}}_m$ be an arbitrary m-dimensional subspace of $\boldsymbol{\mathcal{X}}_n$. Then the utility of $\mathbf{x}_m$ is obtained by summing $p_{\boldsymbol{\mathcal{X}}_n}(\mathbf{x}_n)$ over the complementary subspace $\boldsymbol{\mathcal{X}}_k$, where $k = n - m$, yielding

$$p_{\boldsymbol{\mathcal{X}}_m}(\mathbf{x}_m) = \sum_{\mathbf{x}_k \in \boldsymbol{\mathcal{X}}_k} p_{\boldsymbol{\mathcal{X}}_n}(\mathbf{x}_n). \tag{4}$$

- *Aggregation*: The marginal utility $p_{\boldsymbol{\mathcal{X}}_k}(\mathbf{x}_k)$ and the conditional utility $p_{\boldsymbol{\mathcal{X}}_m|\boldsymbol{\mathcal{X}}_k}(\mathbf{x}_m|\mathbf{x}_k)$ may be aggregated according to the chain rule. Let $\boldsymbol{\mathcal{X}}_m$ and $\boldsymbol{\mathcal{X}}_k$ be disjoint subspaces such that $\boldsymbol{\mathcal{X}}_k$ is preferentially independent of it complement. Then the function F in (3) has the form

$$\begin{aligned} F[p_{\boldsymbol{\mathcal{X}}_k}(\mathbf{x}_k), p_{\boldsymbol{\mathcal{X}}_m|\boldsymbol{\mathcal{X}}_k}(\mathbf{x}_m|\mathbf{x}_k)] &= p_{\boldsymbol{\mathcal{X}}_m \times \boldsymbol{\mathcal{X}}_k}(\mathbf{x}_m, \mathbf{x}_k) \\ &= p_{\boldsymbol{\mathcal{X}}_m|\boldsymbol{\mathcal{X}}_k}(\mathbf{x}_m|\mathbf{x}_k) p_{\boldsymbol{\mathcal{X}}_k}(\mathbf{x}_k). \end{aligned} \tag{5}$$

  More generally, if $\boldsymbol{\mathcal{X}}_m$, $\boldsymbol{\mathcal{X}}_k$, and $\boldsymbol{\mathcal{X}}_\ell$ are disjoint subspaces of $\boldsymbol{\mathcal{X}}_n$ such that $\boldsymbol{\mathcal{X}}_\ell$, $\boldsymbol{\mathcal{X}}_k \times \boldsymbol{\mathcal{X}}_\ell$, and $\boldsymbol{\mathcal{X}}_m \times \boldsymbol{\mathcal{X}}_k \times \boldsymbol{\mathcal{X}}\ell$ are all conditionally independent of their respective complementary subspaces. Then

$$\begin{aligned} & p_{\boldsymbol{\mathcal{X}}_n}(\mathbf{x}_m, \mathbf{x}_k, \mathbf{x}_\ell) = \\ & \quad p_{\boldsymbol{\mathcal{X}}_m|\boldsymbol{\mathcal{X}}_k \times \boldsymbol{\mathcal{X}}_\ell}(\mathbf{x}_m|\mathbf{x}_k, \mathbf{x}_\ell) p_{\boldsymbol{\mathcal{X}}_k|\boldsymbol{\mathcal{X}}_\ell}(\mathbf{x}_k|\mathbf{x}_\ell) p_{\boldsymbol{\mathcal{X}}_\ell}(\mathbf{x}_\ell). \end{aligned} \tag{6}$$

- *Praxeological Independence*: Let $\boldsymbol{\mathcal{X}}_m$ and $\boldsymbol{\mathcal{X}}_k$ be disjoint subspaces such that $\boldsymbol{\mathcal{X}}_m \times \boldsymbol{\mathcal{X}}_k$ is preferentially independent of its complementary subspace. Then $\boldsymbol{\mathcal{X}}_m$ and $\boldsymbol{\mathcal{X}}_k$ are praxeologically independent if $p_{\boldsymbol{\mathcal{X}}_m \times \boldsymbol{\mathcal{X}}_k}(\mathbf{x}_m, \mathbf{x}_k) = p_{\boldsymbol{\mathcal{X}}_m}(\mathbf{x}_m) p_{\boldsymbol{\mathcal{X}}_k}(\mathbf{x}_k)$. Praxeological independence implies preferential independence, but not vice versa.

IV. Multicriterion Decision Making

A. Constructing the Multiattribute Utility Function

The multivariate mass function $p_{\boldsymbol{\mathcal{X}}_n}$ provides a complete evaluation of the consequences that does not require assumptions of independence. It therefore is more general than the classical scalar utility function of the form (1). Generating this function requires the same kind of detailed knowledge about the problem as would the conventional approach. Fortunately, however, since $p_{\boldsymbol{\mathcal{X}}_n}$ is a multivariate mass function, we may apply methods that were originally designed for probabilistic applications to construct this function.

Directed acyclic graphs (DAGs) provide a convenient way to represent the influence one attribute may have on another. They have special significance to probability theory, and are used to construct *Bayesian networks* where each vertex represents a random variable and each path represents a conditional probability linking the parent to the child. Since our formulation of the aggregation problem employs the same mathematics as does probability theory, DAGs play a similar role. To distinguish between the two contexts, we will refer to such graphs as *praxeic networks*.

Consider the praxeic network displayed in Figure 1, whose vertices correspond to the attributes of a six-attribute decision problem. The edges consist of conditional utilities which express the degree of preferential influence one attribute has on another. The parents of a vertex X_i are denoted $\mathrm{pa}\,(X_i)$.

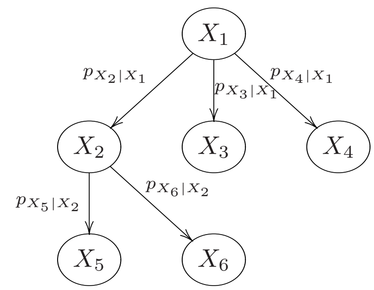

Fig. 1. (a) A praxeic network for a six-attribute decision problem.

The key property of Bayesian and praxeic networks is the *Markov property*: nondescendent nonparents of a vertex are conditionally independent of the vertex, given the state of its parent vertices (for a proof, see [16, 17]). Consequently, we may define conditional preference orderings for each X_i given the preferences of its parents.

Definition 4.1: Let $\mathrm{pa}\,(\mathcal{X}_i)$ denote the subspace corresponding to $\mathrm{pa}\,(X_i)$, that is, if $\mathrm{pa}\,(X_i) = \{X_{i_1}, \ldots, X_{i_{k_i}}\}$, were k_i is the number of parents of X_i, then $\mathrm{pa}\,(\mathcal{X}_i) = \mathcal{X}_{i_1} \times \cdots \times \mathcal{X}_{i_{k_i}}$. A conditional preference ordering $\succeq_{\mathcal{X}_i|\mathrm{pa}\,(\mathcal{X}_i)}$ may be associated with each X_i as follows. Let $\mathbf{x}_{k_i} \in \mathrm{pa}\,(\mathcal{X}_i)$ and $\mathbf{x}'_{k_i} \in \mathrm{pa}\,(\mathcal{X}_i)$ denote the k_i-dimensional sub-vectors of $\mathbf{x}_n$ and $\mathbf{x}'_n$, respectively, that lie in $\mathrm{pa}\,(\mathcal{X}_i)$. The expression $x_i|\mathbf{x}_{k_i} \succeq_{\mathcal{X}_i|\mathrm{pa}\,(\mathcal{X}_i)} x'_i|\mathbf{x}'_{k_i}$ means that X_i prefers x_i given that its parents are committed to $\mathbf{x}_{k_i}$, to x'_i given that its parents are committed to $\mathbf{x}'_{k_i}$. If $\mathrm{pa}\,(X_i) = \varnothing$ then $\mathcal{X}_i$ is preferentially independent of its complementary subspace and the conditional preference ordering $\succeq_{\mathcal{X}_i|\mathrm{pa}\,(\mathcal{X}_i)}$ reduces to the unconditional preference ordering $\succeq_{\mathcal{X}_i}$.

The Markov property may be used to prove the equivalence of a DAG whose edges are conditional mass functions with a joint mass function for all of the vertices. Thus, the multiattribute utility function may be constructed as

$$p_{\mathcal{X}_1 \times \cdots \times \mathcal{X}_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} p_{\mathcal{X}_i|\mathrm{pa}\,(\mathcal{X}_i)}\big(x_i|\mathbf{x}_{k_i}\big), \tag{7}$$

where, if $\mathrm{pa}\,(\mathcal{X}_i) = \varnothing$, then $p_{\mathcal{X}_i|\mathrm{pa}\,(\mathcal{X}_i)}(x_i|\mathbf{x}_{k_i}) = p_{X_i}(x_i)$, the marginal utility function for attribute X_i.

For the praxeic network illustrated in Figure 1, the multiattribute utility function is

$$\begin{aligned} p_{\boldsymbol{\mathcal{X}}_6}(x_1, x_2, x_3, x_4, x_5, x_6) = p_{\mathcal{X}_1}(x_1) p_{\mathcal{X}_2|\mathcal{X}_1}(x_2|x_1) \\ p_{\mathcal{X}_3|\mathcal{X}_1}(x_3|x_1) p_{\mathcal{X}_4|\mathcal{X}_1}(x_4|x_1) \\ p_{\mathcal{X}_5|\mathcal{X}_2}(x_5|x_2) p_{\mathcal{X}_6|\mathcal{X}_2}(x_6|x_2). \end{aligned} \tag{8}$$

The consequence of this property is that, if local influence relationships between attributes can be expressed with a DAG, then the influence relationships can be represented by conditional mass functions where the dependencies flow in only one way: from parents to children. Note that the DAGs associated with conditional preferences as defined in this paper are different from the DAGs that are associated with so-called CP (conditional preference) networks as defined by [10] and CUI (conditional utility independence) networks as defined by [13]. An essential distinction between these approaches and the approach we offer is that our approach conditions with respect to preferences for other consequences, whereas conditional independence is with respect to the instantiation of other consequences (see the discussion regarding Axiom 1). Although these previous approaches employ DAGs, they are *not* Bayesian networks.

The Markov property is used to define algorithms to construct the joint probability mass function from the local relationships that exist between vertices. Well known computationally tractable algorithms include Pearl's Belief Propagation Algorithm [16] and the Sum-Product rule [18].

B. Conditional Dominance

Once the joint utility function $p_{\boldsymbol{\mathcal{X}}_n}$ is defined, each $a \in A$ maps to utility space according to the relation $p_{\boldsymbol{\mathcal{X}}_n}(\mathbf{X}_n(a))$, where $\mathbf{X}_n(a) = (X_1(a), \ldots, X_n(a))$. The classical *a priori* optimal solution is then obtained as

$$a^* = \arg\max_{a \in A} p_{\boldsymbol{\mathcal{X}}_n}(\mathbf{X}_n(a)). \tag{9}$$

To follow the *a posteriori* approach, we must identify the Pareto frontier, which requires a notion of dominance. As conventionally defined (see Definition 2.4), dominance requires consequences to be preferentially independent. However, it is possible to extend this notion to the preferentially dependent case. To motivate, let us examine (8). We see that, since the utility function $p_{\mathcal{X}_1}$ appears in this product and praxeological independence implies preferential independence, $\mathcal{X}_1$ is preferentially independent of its complement. Thus $p_{\mathcal{X}_1}$ corresponds to an unconditional preference ordering $\succeq_{\mathcal{X}_1}$ over $\mathcal{X}_1$. Now consider the term $p_{\mathcal{X}_3|\mathcal{X}_1}(x_3|x_1)$. It is clear from the construction that $\mathcal{X}_1 \times \mathcal{X}_3$ is preferentially independent of its complement, and $p_{\mathcal{X}_3|\mathcal{X}_1}(x_3|x_1)$ corresponds to a conditional preference ordering $\succeq_{\mathcal{X}_3|\mathcal{X}_1}$ over $\mathcal{X}_3$ given $\mathcal{X}_1$, and so forth. We thus see that each conditional mass function in (8) corresponds to a conditional ordering over the associated subspace. These conditional preference orderings may be used to generalize the notion of dominance.

Definition 4.2: Conditional dominance and Pareto frontier. Let $\mathbf{x}_n = (x_1, \ldots, x_n)$ and $\mathbf{x}'_n = (x'_1, \ldots, x'_n)$ be arbitrary elements of $\boldsymbol{\mathcal{X}}_n$ and suppose X_i has k_i parents. Let $\mathbf{x}_{k_i} \in \mathrm{pa}(\mathcal{X}_i)$ and $\mathbf{x}'_{k_i} \in \mathrm{pa}(\mathcal{X}_i)$ be defined as in Definition 4.1. Then $\mathbf{x}_n$ *conditionally dominates* $\mathbf{x}'_n$ if $x_i|\mathbf{x}_{k_i} \succeq_{\mathcal{X}_i|\mathcal{X}_{\mathrm{pa}(\mathcal{X}_i)}} x'_i|\mathbf{x}'_{k_i}$, $i = 1, \ldots, n$, with strict inequality for at least one $j \in \{1, \ldots, n\}$. The *conditional Pareto frontier* is the set of all conditionally nondominated consequence vectors.

Notice that this modified definition coincides with the classical definition of dominance when each $\mathcal{X}_i$ is preferentially independent of its complementary subspace.

C. Example

To illustrate the effect of conditional preference orderings, consider the following simple example. Suppose a series of four different plays (P_1, P_2, P_3, P_4) are offered on each of four days (D_1, D_2, D_3, D_4). An individual has one ticket, and must choose one (play,day) combination out of the sixteen possible alternatives $a_{ij} = (P_i, D_j)$, $i, j = 1, \ldots, 4$. Let Criterion 1 correspond to enjoying the play, and let Criterion 2 correspond to desiring the day.

The operational concept of preference for the play is enjoyment; that is, $P_1 \succeq_{\mathcal{X}_1} P_2$ means that play P_1 will be enjoyed at least as much as play P_2. This ordering is illustrated on a normalized scale in Table I, yielding $p_{\mathcal{X}_1}$.

TABLE I
THE MARGINAL UTILITY FUNCTION $p_{\mathcal{X}_1}$.

x_1	P_1	P_2	P_3	P_4
$p_{\mathcal{X}_1}(x_1)$	0.3571	0.2857	0.2143	0.1429

Now suppose that the individual's preferences for the day depends on the individual's preference for the play (e.g., different actors may be performing on different days). Such a player's preference for the day would then be conditioned on her preferences for the play, yielding a conditional utility of the form $p_{\mathcal{X}_2|\mathcal{X}_1}(x_2|x_1)$. Table II displays this conditional utility function. For example, the conditional valuation for day D_3 given that P_2 is attended is $p_{\mathcal{X}_2|\mathcal{X}_1}(D_3|P_2) = 0.4348$, which is the highest valuation for that day. If, however, the individual were to attend on day D_3, play P_4 would result in the next-worst valuation. The multiattribute utility function thus assumes the form

$$p_{\mathcal{X}_1 \times \mathcal{X}_2}(x_1, x_2) = p_{\mathcal{X}_2|\mathcal{X}_1}(x_2|x_1) p_{\mathcal{X}_1}(x_1). \tag{10}$$

Maximizing (10) yields $a^* = a_{14}$.

TABLE II
THE CONDITIONAL UTILITY FUNCTION $p_{\mathcal{X}_1|\mathcal{X}_2}$.

	x_1			
x_2	P_1	P_2	P_3	P_4
D_1	0.1667	0.0870	0.1852	0.0909
D_2	0.1666	0.3043	0.3333	0.2727
D_3	0.2500	0.4348	0.4444	0.1818
D_4	0.4167	0.1739	0.0370	0.4545

It is instructive to compute the entire utility space, as displayed in Figure 2. By inspection, the Pareto optimal set is $\{a_{14}, a_{23}, a_{33}, a_{44}\}$. As expected, the globally optimal solution a^* is a member of this set.

Now let us assume preferential independence. Unfortunately, the problem specification does not include an unconditional ordering over $\mathcal{X}_2$, but it is possible to generate such

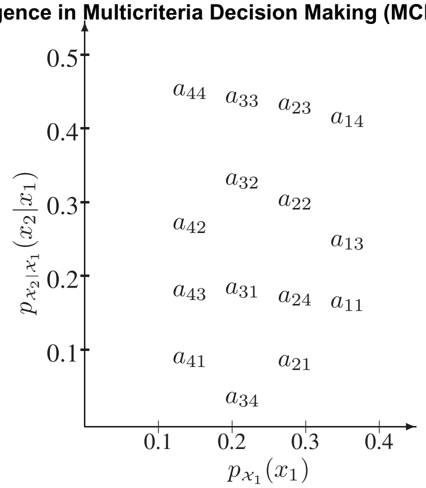

Fig. 2. The utility space illustrating the conditional Pareto optimal set.

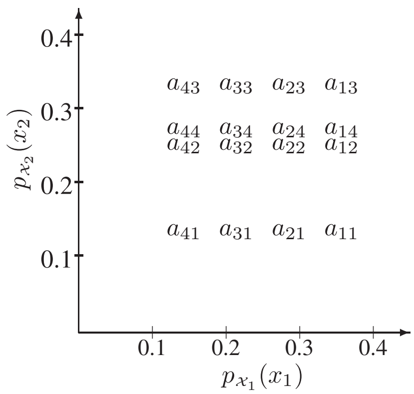

Fig. 3. The utility space assuming preferential independence.

an ordering by computing the marginal utility over $\mathcal{X}_2$ as

$$p_{x_2}(x_2) = \sum_{x_1 \in \mathcal{X}_1} p_{x_2|x_1}(x_2|x_1) p_{x_1}(x_1). \tag{11}$$

Using the marginal utility, however, does not result in the same solution as using the conditional information. Figure 3 displays the utility space assuming preferential independence. The Pareto optimal set is now the singleton $\{a_{13}\}$. This alternative, however, is not Pareto optimal when preferential dependence is taken into consideration. This example illustrates the fact, well known in the probability context, that the multiattribute utility cannot be reconstructed from the marginal utilities unless the consequences are praxeologically independent.

V. Conclusion

This paper provides a reorientation for multicriterion decision problems when the preferential independence assumption is not appropriate. Rather than focusing on ways to exploit preferential independence, as do most approaches in the literature, our approach is to account directly for preference dependencies. We do this by introducing a new utility structure that is explicitly designed to model dependencies according to two fundamental aggregation axioms. We show that compliance with these axioms requires the utilities to possess the syntactical structure of probability mass functions, albeit with praxeological, rather than epistemological, semantics. In a personal communication to Judea Pearl, Glenn Shafer observed that "probability is not really about numbers; it is about the structure of reasoning" [16, p 15]. We submit that this property applies not only to the epistemological activity of forming beliefs, but also to the praxeological activity of taking action.

An ancillary, but important benefit of the utility structure is that it is amenable to a powerful conceptual and synthesis tool: graph theory. Since the utility functions are structured mathematically as mass functions, the multiattribute utility function may be synthesized according to the Markov property of graph theory (i.e., the chain rule of probability), rather than relying upon *ad hoc* simplifications of the multiattribute function, such as additive decompositions.

This approach may result in more complexity than would be encountered with mutual preferential independence, but as Palmer has noted, "complexity is no argument against a theoretical approach if the complexity arises not out of the theory itself but out of the material which any theory ought to handle" [19, p. 176].

Appendix

Proof of the Aggregation Theorem (Preliminary versions of this theorem are discussed in [20–22].)

Sufficiency follows by the definition of marginalization and conditioning, yielding

$$F[p_{\boldsymbol{\mathcal{X}}_k}(\mathbf{x}_k), p_{\boldsymbol{\mathcal{X}}_m|\boldsymbol{\mathcal{X}}_k}(\mathbf{x}_m|\mathbf{x}_k)] = p_{\boldsymbol{\mathcal{X}}_m|\boldsymbol{\mathcal{X}}_k}(\mathbf{x}_m|\mathbf{x}_k) p_{\boldsymbol{\mathcal{X}}_k}(\mathbf{x}_k). \tag{12}$$

To prove necessity, let $\boldsymbol{\mathcal{X}}_i$, $\boldsymbol{\mathcal{X}}_j$, and $\boldsymbol{\mathcal{X}}_k$, where $i \neq j, i \neq k, j \neq k$, be arbitrary pairwise disjoint subspaces of $\boldsymbol{\mathcal{X}}_n$, and let $p_{\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j \times \boldsymbol{\mathcal{X}}_k}$, $p_{\boldsymbol{\mathcal{X}}_i | \boldsymbol{\mathcal{X}}_j \times \boldsymbol{\mathcal{X}}_k}$, $p_{\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j | \boldsymbol{\mathcal{X}}_k}$, $p_{\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j}$, $p_{\boldsymbol{\mathcal{X}}_i | \boldsymbol{\mathcal{X}}_j}$, and $p_{\boldsymbol{\mathcal{X}}_i}$ be endogenously aggregated non-negative utility functions. That is, there exists a function F such that $p_{\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j}(\mathbf{x}_i, \mathbf{x}_j) = F[p_{\boldsymbol{\mathcal{X}}_j}(\mathbf{x}_j), p_{\boldsymbol{\mathcal{X}}_i|\boldsymbol{\mathcal{X}}_j}(\mathbf{x}_i|\mathbf{x}_j)]$ $\forall (\mathbf{x}_i, \mathbf{x}_j) \in \boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j$.

Let $P_{\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j \times \boldsymbol{\mathcal{X}}_k}$, $P_{\boldsymbol{\mathcal{X}}_i | \boldsymbol{\mathcal{X}}_j \times \boldsymbol{\mathcal{X}}_k}$, and $P_{\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j | \boldsymbol{\mathcal{X}}_k}$ be non-negative functions over subsets of $\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j \times \boldsymbol{\mathcal{X}}_k$; let $P_{\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j}$, $P_{\boldsymbol{\mathcal{X}}_i | \boldsymbol{\mathcal{X}}_j}$ be non-negative functions over subsets of $\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j$, and let $P_{\boldsymbol{\mathcal{X}}_i}$ be non-negative functions over subsets of $\boldsymbol{\mathcal{X}}_i$ such that, for all singleton sets $\{x_i\} \in \boldsymbol{\mathcal{X}}_i$,

$$\begin{aligned}
P_{\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j \times \boldsymbol{\mathcal{X}}_k}(\{\mathbf{x}_i\} \times \{\mathbf{x}_j\} \times \{\mathbf{x}_k\}) &= p_{\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j \times \boldsymbol{\mathcal{X}}_k}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) \\
P_{\boldsymbol{\mathcal{X}}_i | \boldsymbol{\mathcal{X}}_j \times \boldsymbol{\mathcal{X}}_k}(\{\mathbf{x}_i\} | \{\mathbf{x}_j\} \times \{\mathbf{x}_k\}) &= p_{\boldsymbol{\mathcal{X}}_i | \boldsymbol{\mathcal{X}}_j \times \boldsymbol{\mathcal{X}}_k}(\mathbf{x}_i | \mathbf{x}_j, \mathbf{x}_k) \\
P_{\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j | \boldsymbol{\mathcal{X}}_k}(\{\mathbf{x}_i\} \times \{\mathbf{x}_j\} | \{\mathbf{x}_k\}) &= p_{\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j | \boldsymbol{\mathcal{X}}_k}(\mathbf{x}_i, \mathbf{x}_j | \mathbf{x}_k) \\
P_{\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j}(\{\mathbf{x}_i\} \times \{\mathbf{x}_j\}) &= p_{\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j}(\mathbf{x}_i, \mathbf{x}_j) \\
P_{\boldsymbol{\mathcal{X}}_i | \boldsymbol{\mathcal{X}}_j}(\{\mathbf{x}_i\} | \{\mathbf{x}_j\}) &= p_{\boldsymbol{\mathcal{X}}_i | \boldsymbol{\mathcal{X}}_j}(\mathbf{x}_i | \mathbf{x}_j) \\
P_{\boldsymbol{\mathcal{X}}_i}(\{\mathbf{x}_i\}) &= p_{\boldsymbol{\mathcal{X}}_i}(\mathbf{x}_i).
\end{aligned}$$

These functions, when restricted to singleton sets, are utility functions and thus, by hypothesis, are endogenously aggregated. We require these functions over sets to be endogenously aggregated when extended to arbitrary measurable rectangles

$A \times B \times C \subseteq \boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j \times \boldsymbol{\mathcal{X}}_k$. Thus, it must hold that

$$P_{\boldsymbol{x}_i \times \boldsymbol{x}_j}(A \times B) = F\left[P_{\boldsymbol{x}_j}(B), P_{\boldsymbol{x}_i|\boldsymbol{x}_j}(A|B)\right] = F\left[P_{\boldsymbol{x}_i}(A), P_{\boldsymbol{x}_j|\boldsymbol{x}_i}(B|A)\right]. \quad (13)$$

It is convenient to follow the development that Jaynes applied to the epistemological case [23, Chapter 2]. Let $A \subseteq \boldsymbol{\mathcal{X}}_i$, $B \subseteq \boldsymbol{\mathcal{X}}_j$, and $C \subseteq \boldsymbol{\mathcal{X}}_k$ be arbitrary. Since product space operations are associative, $A \times B \times C = (A \times B) \times C = A \times (B \times C)$, it must hold that

$$\begin{aligned} P_{\boldsymbol{x}_i \times \boldsymbol{x}_j \times \boldsymbol{x}_k}(A \times B \times C) &= F\left[P_{\boldsymbol{x}_j \times \boldsymbol{x}_k}(B \times C), P_{\boldsymbol{x}_i|\boldsymbol{x}_j \times \boldsymbol{x}_k}(A|B \times C)\right] \\ &= F\left[P_{\boldsymbol{x}_j}(C), P_{\boldsymbol{x}_i \times \boldsymbol{x}_j|\boldsymbol{x}_k}(A \times B|C)\right]. \quad (14) \end{aligned}$$

By the endogenous aggregation hypothesis, we have

$$P_{\boldsymbol{x}_i \times \boldsymbol{x}_j|\boldsymbol{x}_k}(A \times B|C) = F\left[P_{\boldsymbol{x}_j|\boldsymbol{x}_k}(B|C), P_{\boldsymbol{x}_i|\boldsymbol{x}_j \times \boldsymbol{x}_k}(A|B \times C)\right]. \quad (15)$$

Applying (13) to $P_{\boldsymbol{x}_j \times \boldsymbol{x}_k}(B \times C)$ and (15) to $P_{\boldsymbol{x}_i \times \boldsymbol{x}_j|\boldsymbol{x}_k}(A \times B|C)$ yields, upon substitution into (14),

$$F\{F[P_{\boldsymbol{x}_k}(C), P_{\boldsymbol{x}_j|\boldsymbol{x}_k}(B|C)], P_{\boldsymbol{x}_i|\boldsymbol{x}_j \times \boldsymbol{x}_k}(A|B \times C)\} = F\{P_{\boldsymbol{x}_k}(C), F[P_{\boldsymbol{x}_j|\boldsymbol{x}_k}(B|C), P_{\boldsymbol{x}_i|\boldsymbol{x}_j \times \boldsymbol{x}_k}(A|B \times C)]\}. \quad (16)$$

In terms of general arguments, this equation becomes

$$F\left[F(x,y), z\right] = F\left[x, F(y,z)\right], \quad (17)$$

which is called the *associativity equation* [23, 24]. By direct substitution it is easy to see that (17) is satisfied if

$$f\left[F(x,y)\right] = f(x)f(y) \quad (18)$$

for any function f.

It has been shown by Cox [25] that if F is differentiable in both arguments, then

- (18) is the general solution to (17) for any function f. Taking f as the identity function, (13) becomes

$$P_{\boldsymbol{x}_i \times \boldsymbol{x}_j}(A \times B) = P_{\boldsymbol{x}_i|\boldsymbol{x}_j}(A|B)P_{\boldsymbol{x}_j}(B) = P_{\boldsymbol{x}_j|\boldsymbol{x}_i}(B|A)P_{\boldsymbol{x}_i}(A) \quad (19)$$

- For any $A_1, A_2 \subseteq \boldsymbol{\mathcal{X}}_i$ and any $B_1, B_2 \subseteq \boldsymbol{\mathcal{X}}_j$ such that $(A_1 \times B_1) \cap (A_2 \times B_2) = \varnothing$,

$$P_{\boldsymbol{x}_i \times \boldsymbol{x}_j}[(A_1 \times B_1) \cup (A_2 \times B_2)] = P_{\boldsymbol{x}_i \times \boldsymbol{x}_j}(A_1 \times B_1) + P_{\boldsymbol{x}_i \times \boldsymbol{x}_j}(A_2 \times B_2). \quad (20)$$

- $P_{\boldsymbol{x}_i \times \boldsymbol{x}_j}[\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j] = 1$.

In particular, if $\boldsymbol{\mathcal{X}}_i \times \boldsymbol{\mathcal{X}}_j = \boldsymbol{\mathcal{X}}_n$, then $P_{\boldsymbol{x}_n}$ is a praxeological measure over $\mathcal{B}$ and the utility function $p_{\boldsymbol{x}_n}$ is the associated multivariate mass function over $\boldsymbol{\mathcal{X}}_n$.

REFERENCES

[1] R. L. Keeney and H. Raiffa, *Decisions with Multiple Objectives*. Cambridge, UK: Cambridge University Press, 1993.
[2] K. J. Arrow, *Social Choice and Individual Values*. New York, NY: John Wiley, 1951, 2nd ed. 1963.
[3] G. Debreu, "Topological methods in cardinal utility theory," in *Mathematical Methods in the Social Sciences*, K. J. Arrow, S. Karlin, and P. Suppes, Eds. Cambridge, MA: Harvard Univ. Press, 1960.
[4] K. Miettinen, *Nonlinear Multiobjective Optimization*. Boston, MA: Kluwer, 1999.
[5] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*. New York, NY: John Wiley & Sons, 2001.
[6] D. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, Massachusetts: Addison-Wesley Publishing Company, 1989.
[7] C. M. Fonseca and P. J. Fleming, "Genetic algorithms for multiobjective optimization: Formulation, Discussion and Generalization," in *Proceedings of the Fifth International Conference on Genetic Algorithms*, 1993, pp. 416–423.
[8] F. Bacchus and A. Grove, "Graphical models for preference and utility," in *Uncertainty in Artificial Intelligence. Proceedings of the Eleventh Conference (1995)*. San Francisco: Morgan Kaufmann Publishers, 1995, pp. 3–10.
[9] ——, "Utility independence in a qualitative decision theory," in *Principles of Knowledge Representation and Reasoning (KR-96)*, 1996, pp. 542–552.
[10] C. Boutilier, R. Brafman, H. H. Hoos, and D. Poole, "Reasoning with conditional ceteris paribus preference statements," in *Uncertainty in Artificial Intelligence. Proceedings of the Fifteenth Conference (1999)*. San Francisco: Morgan Kaufmann Publishers, 1999, pp. 71–80.
[11] C. Boutilier, F. Bacchus, and R. Brafman, "UCP-networks: A directed graphical representation of conditional utilities," in *Uncertainty in Artificial Intelligence. Proceedings of the Seventeenth Conference (2001)*. San Francisco: Morgan Kaufmann Publishers, 2001, pp. 56–64.
[12] C. Gonzales and P. Perny, "GAI networks for utility elicitation," in *Proceedings of the Ninth International Conference on the Principles of Knowledge Representation and Reasoning*, 2006, pp. 224–234.
[13] Y. Engel and M. P. Wellman, "CUI networks: A graphical representation for conditional utility independence," in *Proceedings of the Twenty-First National Conference on Artificial Inelligence*, 2006, pp. 1137–1142.
[14] P. L. Mura and Y. Shoham, "Expected utility networks," in *Uncertainty in Artificial Intelligence. Proceedings of the Fifteenth Conference (1999)*. San Francisco: Morgan Kaufmann Publishers, 1999, pp. 366–373.
[15] J. Neveu, *Mathematical Foundations of the Calculus of Probability*. San Francisco: Holden Day, 1965.
[16] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*. San Mateo, CA: Morgan Kaufmann, 1988.
[17] R. G. Cowell, A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter, *Probabilistic Networks and Expert Systems*. New York, NY: Springer Verlag, 1999.
[18] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Info. Theory*, vol. 47, no. 2, pp. 498–519, 2001.
[19] F. R. Palmer, *Grammar*. Harmondsworth, Middlesex, UK: Harmondsworth, Penguin, 1971.
[20] W. C. Stirling, *Satisficing Games and Decision Making: with applications to engineering and computer science*. Cambridge, UK: Cambridge Univ. Press, 2003.
[21] ——, "Social Utility Functions, Part 1 — Theory," *IEEE Transactions on Systems, Man, and Cybernetics (Part C)*, vol. 35, no. 4, pp. 522–532, 2005.
[22] W. C. Stirling and R. L. Frost, "Social Utility Functions, Part 2 — Applications," *IEEE Transactions on Systems, Man, and Cybernetics (Part C)*, vol. 35, no. 4, pp. 533–543, 2005.
[23] E. T. Jaynes, *Probability Theory: The Logic of Science*. Cambridge, UK: Cambridge Univ. Press, 2003.
[24] J. Aczél, *Lectures on Functional Equations and Their Applications*. New York, NY: Academic Press, 1966.
[25] R. T. Cox, "Probability, frequency, and reasonable expectation," *American Journal of Physics*, vol. 14, pp. 1–13, 1946.