Attributes*

Diego Klabjan, Wojciech Olszewski, and Asher Wolinsky

May 26, 2010

Abstract

An agent makes the decision whether to acquire an object. Before making this decision, she can discover, at some cost, some attributes of the object (or equivalently, some signals about the object's value). We characterize the solution to the following problem of sequential discovering of attributes with the option of stopping at any point of time, and accepting or rejecting the object.

We also partially solve the problem of simultaneous choice of attributes which are to be discovered before making the decision regarding the object.

Preliminary and Incomplete

1 Introduction

Many economic decisions have the following form: An agent considers acquiring an object, which is characterized by several attributes. The agent knows the distribution of each attribute in the population, but not the particular realization of attributes for the object at hand. Before making the decision whether to acquire the object, the agent can at some cost discover the realization of each attribute; the cost of discovering can vary across different attributes. Examples abound, ranging from everyday goods to marriage. For instance, a potential spouse is characterized by intelligence, beauty, character, etc. Another example, from the labor markets, will be provided shortly.

Recent studies in decision theory provide additional support to the attribute approach. These studies argue that decision makers often use sequentially several rationales to discriminate among

*We would like to thank Hector Chade and Rakesh Vohra for helpful comments and suggestions.

the available alternatives. (See, for instance, Manzini and Mariotti (2007), although the basic ideas come yet from Tversky (1972).)

We study two versions of our decision problem: In one version of the model, the agent is allowed to discover attributes sequentially. In this case, we provide a complete solution for attributes which are distributed independently, and symmetrically around their means, and for the outside option, i.e., the value of rejecting the object, is equal to the sum of these means. The solution resembles Gittins' (1989) indices and Weitzman's (1979) Pandora rule. The attributes are attached indices, independent of other attributes, and it is determined by the largest index among the remaining attributes which attribute is to be discovered next.

In the other version of the model, the agent must decide up front which attributes to discover, and then, knowing the realizations of these attributes, decide to accept or reject the object. Here, we have only partial results. We provide a pseudo polynomial-time algorithms for solving some special cases, and we show that there is no polynomial-time algorithm solving the problem even in the binary case in which each attribute can take one of two possible values. The existence of pseudo polynomial-time algorithm, or another simple and easy to interpret solution in more general cases is an open problem.

Weitzman (1979) studied a problem similar to the sequential version of our multi-attribute model. Suppose an agent opens sequentially boxes containing a reward of unknown value. It is costly to open each box and discover the value of the reward contained in the box. And the agent can take only one reward. Weitzman shows that the optimal strategy (*Pandora rule*) assigns to each box a reservation price, independent of the other boxes, and prescribes to search next the remaining box with the highest reservation price. The agent terminates search when the highest reward from already sampled boxes exceeds the highest reservation price across remaining, closed boxes.

An important feature of the Weitzman's setting is that the rewards hidden in the remaining boxes are independent of the rewards discovered in the already sampled boxes. In our multi-attribute model, there is an inter-dependence between the rewards of the already sampled boxes and the still closed boxes. This makes the problem studied in the present paper more complicated.

The celebrated literature on multi-armed bandits is also closely related. This literature investigates situations in which an agent chooses which arm to pull in each single period. The reward obtained in this way depends on the state of the pulled arm, which then transits to another state. The state of all the other arms remain unaltered. Gittins and Jones (1974) (see Whittle (1980) and Weber (1992) for simpler proofs) showed that one can attach an index to each arm, which depends only on the state of that arm, and that pulling an arm with the largest index at any point in time is an optimal strategy.

Other related research includes Neeman (1995) and Chade and Smith (2005). Neeman is interested in an optimal strategy in the following stopping problem: An agent faces a sequence of i.i.d. multi-attribute products, and can observe only one attribute of each product. At each stage the agent has to decide whether to stop, taking the best product so far, or to continue by observing an attribute of the next product. Chade and Smith study the following choice problem: An agent selects a number of ranked stochastic options. The inclusion of each option to the selected set is costly. Only one option may be exercised from those that succeed. A leading example is a student applying to many colleges. They show that the celebrated greedy algorithm finds the optimal set.

1.1 Example

As an illustration, consider the following example. Suppose that an economics department considers hiring a new assistant professor, say an applied econometrician. The candidate may have, or may lack having, several types of skills, or attributes, which would be useful for the department We will consider only three of them: teaching skills, and two attributes related to research: breath and depth. Breath refers to wide interests and the ability of detecting relevant questions for analysis, while depth means the formal rigor and correctness of the analysis, as well as the ability of learning and applying of advanced econometric techniques.

Suppose that checking the candidate's breath, or teaching skills is relatively easy. It takes just a couple of professional conversations, or looking at teaching evaluations, syllabi and exam questions. Say the costs of checking these two attributes are 4 and 6, respectively. Suppose further that the dispersion of these attributes across the pool of all candidates is smaller than that for the third attribute, the candidate's depth. More precisely, normalize the expected value of each attribute to zero, and let breath take one of two equally likely values: $+108$ or -108; similarly, let teaching skills and the candidate's depth take values $+112$ or -112 and $+160$ or -160, respectively. Finally, assume that checking the candidate's depth is substantially more expensive, because it requires going into the technical details of the candidate's papers, and the cost of doing so is 30. These numbers are exhibited in Table 1.

Notice that if the department were restricted to looking at a single attribute, it would be indifferent between all three of them. The depth is more informative than the teaching skills or the breath, but the higher cost of discovering it exactly offsets the benefit of doing so. Consequently, the expected value is 50 in each case. It is of no help to discover two attributes, since the decision to accept or reject the object would disregard the signal that is less informative. Finally, checking all three attributes costs 40, and the optimal strategy in this case is to accept the object when (and only when) at least two out of three attributes take positive values, which yields the expected value (including the cost) of 55. Thus, the department should discover all three attributes.

Note here that one of our general results, Proposition 2, implies that in cases in which the agent restricted to looking at only one out of n attribute is indifferent between all of them, it is optimal to discover attributes $k, ..., n$ for some $k = 1, ..., n$, where the attributes are ordered from the most to the least informative. That is, in order to find out the optimal strategy in the present example we only need to compare the expected values of: discovering all attributes, learning about teaching skills and breath, and learning only about breath.

Under the sequential discovering of attributes, an optimal strategy is to discover teaching skills and breath first, which can be done by reviewing the candidate's packet and a sequence of short interviews; to accept the candidate if both these attributes take positive values, and to reject the candidate if both attributes take negative values; and to learn about the depth otherwise, that is, when one attribute is positive and the other one is negative, the department should study carefully the candidate's papers. This strategy is represented in Figure 1. The optimality of this strategy can be verified directly, and also follows from one of our general results.

2 Model

The problem we are studying in this paper can be formulated in the following two equivalent ways:

2.1 Attribute version of the model

An agent must decide whether to accept or reject an object that has been offered to her. The object has n attributes, represented by random variables $x_1, ..., x_n$, whose realizations are initially unknown to the agent. These random variables are distributed according to some cumulative density function (in abbreviation, *cdf*). The distribution is known to the agent. The agent's utility of having an object with attributes $x_1, ..., x_n$ is denoted by $u(x_1, ..., x_n)$, and the agent's reservation utility, i.e., the utility the agent obtains when she rejects the object, is denoted by V. Before making a decision, the agent can discover the realization of each attribute $x_1, ..., x_n$ of the object at the cost $c_1, ..., c_n$, respectively.

We consider two scenarios. Under one of them, the agent may discover the attributes sequentially. That is, she decides which attribute (if any) she wants to learn about first, and contingent on the realizations of the attributes she has already discovered, she decides which of the remaining attributes (if any) to learn about next. After the discovering of each attribute, the agent may stop the process of learning the realizations of attributes, in which case she may accept or reject the object. She may also decide to continue the process, and discover the realization of one of the remaining attributes. Every time she decides to discover an attribute, she pays the cost associated with this attribute.

Under the other scenario, the agent discovers the attributes simultaneously. That is, she decides which set of attributes $S \subset \{1, ..., n\}$ to discover. After learning the realizations x_i, $i \in S$ (at the cost $\sum_{i \in S} c_i$), she makes the decision whether to accept or reject the object.

2.2 Signal version of the model

Suppose that instead of having multiple attributes, the value of the object is uncertain. The agent has a prior over the value of the object offered to her, and before making the decision to accept or reject it, she can consult with several sources of information. Each available source $i = 1, ..., n$ provides a signal x_i at a cost c_i. The agent knows the value-dependent distributions of signals, and therefore for any set of signals x_i, $i \in S \subset \{1, ..., n\}$, she can compute the expected value of the object. Again, the agent may consult with the sources of information sequentially or simultaneously.

This model is formally equivalent with the model in which objects have attributes. However, some assumptions, or payoff functions that are easy to interpret in one version of the model may be less reasonable in the other version. In this paper, we will focus on the attribute version of the problem.

2.3 Assumptions

Throughout the paper, we assume that $x_1, ..., x_n$ are distributed *independently* with cdfs $F_1, ..., F_n$, respectively,

$$u(x_1, ..., x_n) = x_1 + ... + x_n$$

and

$$V = \int x_1 dF_1 = ... = \int x_n dF_n = 0.$$

We write that $y \precsim^{s.o.} z$ when random variable y that is *second-order stochastically dominated* by random variable z. We will often assume that $x_1 \precsim^{s.o.} ... \precsim^{s.o.} x_n$, i.e., that attributes are ordered by second-order stochastic dominance.

We will sometimes make even slightly stronger assumptions. A random variable y with the cdf G is *symmetric around the mean* (or briefly, symmetric) if for every $y \geq 0$,

$$G(-y) = 1 - G(y).$$

Consider two symmetric random variables y and z with cdfs G and H, respectively. We will say that y is *first-order stochastically dominated* by z on *positive realizations* (briefly, positive first-order stochastically dominated) if

$$G(x) \leq H(x)$$

for all $x \geq 0$. We will then write that $z \precsim^{p.f.o.} y$.

Notice that by symmetry, if y is positive first-order stochastically dominated by z, then z is negative first-order dominated by y, i.e.

$$H(x) \leq G(x)$$

for all $x \leq 0$. Intuitively, the idea is that variable y is obtained by moving the probability assigned by variable z from the center to the tails of the distribution. Notice that y is positive first-order stochastically dominated by z, then it is also second-ordered stochastically dominated.

2.4 Optimization problem

Under sequential discovering of attributes, we will seek simple characterizations of optimal strategies, i.e. rules that tell the agent the realization of which attribute to discover (or to stop the process of learning, and to accept or reject the object), contingent on any history of already discovered realizations.

Under simultaneous discovering of attributes, the agent who checked the values of attributes from set S will accept the object if $\sum_{i \in S} x_i > 0$, and reject the object if $\sum_{i \in S} x_i < 0$. The agent's tie-breaking rule when $\sum_{i \in S} x_i = 0$ will be inessential for the analysis, because we will consider only continuous distributions or discrete distributions that take values other than zero. Therefore, the agent's objective function is

$$U(S) = \int ... \int \max\left(0, \sum_{i \in S} x_i\right) d \prod_{i \in S} F_i(x_i) - \sum_{i \in S} c_i. \quad (1)$$

The question we are interested in is whether there exists a "nice" way of describing the solution of the agent's maximization problem:

$$\max_{S \subset \{1,...,n\}} U(S). \tag{2}$$

More rigorous versions of this question includes: Can the optimization problem (1) be solved in polynomial, or pseudo-polynomial time[1]? Or, can we find an optimal set S by computing $U(S)$ only at $P(n)$ sets S, where $P(n)$ is a polynomial function of the number of attributes n?

2.5 Comments, questions for future research

The results of this paper easily extend to the utilities of the form $u(x_1, ..., x_n) = \alpha_1 x_1 + ... + \alpha_n x_n$, where $\alpha_1, ..., \alpha_n$ are some positive real numbers; in this case, the assumptions regarding the distributions of $x_1, ..., x_n$ must be replaced with analogous assumptions regarding the distributions of $x_1/\alpha_1, ..., x_n/\alpha_n$. The research on related topics (e.g., Gittin's indices, Weitzman's Pandora rule) suggests that the analysis does not generalize, and is intractable without the independence assumption.

The assumption that the reservation utility V and the means of $x_1, ..., x_n$ are all equal is more problematic, because some of the results rely on the assumption that the distributions of $x_1, ..., x_n$ are symmetric around both their mean and the reservation utility. Extending the results to arbitrary values of V and the means is an open question. Moreover, the value of V is endogenous in some problems, related to one studied in the present paper, e.g., when the agent searches a sequence of objects, and decides whether to accept the current object and stop searching, or to reject the current object and search for another one.

Under sequential discovering (and a single object), one might also study the setting in which the agent decides first about the ordering in which she will discover attributes. She is allowed to interrupt this process, and accept or reject the object without learning the realizations of the remaining attributes, but she is not allowed to change the ordering. That is, the agent is allowed to discover attributes sequentially, but the ordering of attributes may not depend on the already discovered realizations.

General, even independent distributions of attributes may have the property that the value of signals about the object they provide depends on the exact subset of attributes which the agent is going to discover. This makes the optimization problem in the simultaneous case complicated. The assumptions that the attributes are ordered by stochastic dominance are attempts of capturing the

[1] Of course, we must restrict attention here to distributions with finite support.

idea that some distributions provide more valuable signals than others, independently of the exact subset of attributes which the agent is going to discover.

3 Sequential Discovering of Attributes

We will provide a complete solution of the problem in the case for two symmetrically distributed attributes x_1, x_2. We believe this result generalizes to any number n of symmetrically distributed attributes. The results generalize to possibly asymmetric distributions, but the solution is much less elegant and harder to interpret.

Let

$$x_i^* = -c_j + \int_{-x_i^*}^{+\infty} (x_i^* + x_j) dF_j(x_j) \tag{3}$$

for $i, j \in \{1, 2\}$ and $i \neq j$. This equation is interpreted as follows: Suppose that the agent discovers attribute i first and learn that its realization is x_i^*. Then, the agent is indifferent between accepting the object without learning the realization of the other attribute, and discovering the realization of the other attribute and making then the decision regarding the object.

Notice that equation (3) is equivalent to the equation

$$0 = -c_j + \int_{x_i^*}^{+\infty} (-x_i^* + x_j) dF_j(x_j). \tag{4}$$

Indeed, (3) can be rewritten as

$$x_i^* F_j(-x_i^*) = -c_j + \int_{-x_i^*}^{+\infty} x_j dF_j(x_j), \tag{5}$$

and (4) can be rewritten as

$$x_i^* [1 - F_j(x_i^*)] = -c_j + \int_{x_i^*}^{+\infty} x_j dF_j(x_j). \tag{6}$$

By the symmetry of x_j, the left-hand sides of these two equations coincide. The right-hand sides coincide as well; indeed, since the mean of x_j is zero,

$$\int_{-x_i^*}^{+x_i^*} x_j dF_j(x_j) = 0$$

by the symmetry of x_j.

Equation (4) says that when the agent learns that the realization of attribute i is $-x_i^*$, she is indifferent between rejecting the object without learning the realization of the other attribute, and discovering the realization of the other attribute and making then the decision regarding the object.

Notice that this equation has a unique solution $x_i^* > 0$. Indeed, for $x_i^* = 0$ the left-hand side of (3) falls below the right-hand side when

$$0 < -c_j + \int_0^{+\infty} x_j dF_j(x_j); \tag{7}$$

this condition means that the agent prefers discovering attribute j to accepting (or rejecting) the object without learning the realization of any attribute. For sufficiently large values of x_i^*, left-hand side of (3) exceeds the right-hand side. And the difference between the left-hand side and the right-hand side is increasing in x_i^*. Thus, condition (7) guarantees, that equation (3) has a unique solutions.

Theorem 1. According to every optimal strategy, the agent should discover attribute i first whenever $x_i^* < x_j^*$. After discovering the realization x_i of attribute i, the agent should accept the object whenever $x_i^* < x_i$; she should reject the object whenever $x_i < -x_i^*$; otherwise, she should discover the realization x_j of attribute j, accept the object when $x_i + x_j > 0$ and reject the object when $x_i + x_j < 0$.

The essence of Theorem 1, and the only part which requires a proof is the rule regarding which attribute to discover first. Suppose that $x_2^* < x_1^*$, i.e., $i = 2$ and $j = 1$. The payoffs contingent on any possible pair of realizations of the two attributes are exhibited in Figure 2 (a). The top row of each area is the payoff from discovering the realization of attribute 1 first, and and the bottom row is the payoff from discovering the realization of attribute 2 first. In each case, the agent plays an optimal continuation strategy contingent on the realization of the attribute she has learned first.

We obtain Figure 2 (b) from Figure 2 (a) by deleting the common components of corresponding top and bottom payoffs. The areas with no payoff appear when this deletion makes the top and bottom payoff equal to 0, i.e., the payoffs of the two rules are equal contingent on the realizations of attributes from that area.

Finally, we obtain Figure 2 (c) from Figure 2 by means of (3) and (4). Notice that for any given value of x_2, the component $-c_1$ appears in the top row for every single value of x_1 or for no value of x_1. The component $-c_1$ appears for every single value of x_1 when $x_2 > x_2^*$ or when $x_2 < -x_2^*$. In the former case, we can replace $-c_1$ with $x_2^* + x_1$ for $x_1 < -x_2^*$ and 0 for $x_1 > -x_2^*$. This follows from (5), because

$$\int_{-x_2^*}^{+\infty} x_1 dF_1(x_1) = \int_{-x_2^*}^{+x_2^*} x_1 dF_1(x_1) + \int_{+x_2^*}^{+\infty} x_1 dF_1(x_1) = -\int_{-\infty}^{-x_2^*} x_1 dF_1(x_1)$$

by the symmetry of x_1. Similarly, we can replace $-c_1$ with $x_2^* - x_1$ for $x_1 > x_2^*$ and 0 for $x_1 < x_2^*$ in the latter case, by virtue of (6). And by analogous arguments, we dispense with $-c_2$.

Notice that it is essential for this last step of the proof that x_i^* solving (3) is equal to x_i^* solving (4). In order to conclude that the agent should discover attribute 2 before discovering attribute 1, notice that the entry in each top row of Figure 2 (c) is lower than the entry in the corresponding bottom row.

4 Simultaneous Discovering of Attributes

4.1 Some special cases

We begin with two propositions regarding some special cases in which we will be able to answer the questions raised in Section 2.4.

Proposition 1. Suppose the attributes $x_1, ..., x_n$ are ordered by the second-order stochastic dominance, i.e. $x_1 \precsim^{s.o.} ... \precsim^{s.o.} x_n$. Suppose further that all costs are equal, i.e., $c_1 \leq ... \leq c_n$. Then a set solving problem (2) has the form $S = \{1, ..., k\}$; that is, it consists of a k most <u>dominated</u> distributions.

This proposition generalizes the observation that stochastically dominated distributions provide more valuable signals about the object. Therefore, when the cost of discovering the realization is nondecreasing across attributes, the decision maker should learn first the realizations of most dominated attributes.

Proposition 2. Suppose the distributions $x_1, ..., x_n$ are symmetric and ordered by positive first-order dominance, i.e. $x_1 \precsim^{p.f.o.} ... \precsim^{p.f.o.} x_n$. Suppose further that

$$c_i - c_{i+1} \geq \int_0^{+\infty} x_i dF_i(x_i) - \int_0^{+\infty} x_{i+1} dF_{i+1}(x_{i+1}). \tag{8}$$

Then a set solving problem (2) has the form $S = \{k, ..., n\}$; that is, it consists of a k most <u>dominant</u> distributions.

Condition (8) says that the differences in the benefits from discovering single attributes correspond to the differences in the costs of discovering these attributes. That is, suppose the agent were restricted to discovering the realization of exactly one attribute. Then, she would be better off discovering a stochastically dominated attribute. Condition (8) says that discovering the stochastically dominant attribute is, however, less expensive, and this lower cost would make the agent at least indifferent across all single attributes. Then, if the attributes are ordered by positive first-order dominance, the decision maker should learn first the realizations of most dominant attributes.

Propositions 1 and 2 imply that under the assumptions of the two propositions, an optimal set S, solving the maximization problem (2), can be found by computing $U(S)$ at most n sets S. The propositions follow from the following two lemmas, which seem interesting by themselves, and will also be useful later, for other purposes.

Lemma 1. Let G_0, G_1 and G_2 be the cdfs of random variables y_0, y_1 and y_2. Suppose that all three random variables y_0, y_1 and y_2 have mean 0, and y_2 second-order stochastically dominates y_1. That is, $Ey_0 = Ey_1 = Ey_2 = 0$ and $y_1 \precsim^{s.o.} y_2$. Then

$$\int\int_{y_0+y_2\geq 0} (y_0 + y_2)\, dG_0(y_0) dG_2(y_2) \leq \int\int_{y_0+y_1\geq 0} (y_0 + y_1)\, dG_0(y_0) dG_1(y_1).$$

This lemma says that the benefit from learning the realizations of attributes will become higher when the decision maker replaces a second-order stochastically dominant attribute with a stochastically dominated attribute. The following lemma says, however, the marginal benefit from learning the realization of any additional attribute together with a positive first-order stochastically dominated attribute is lower than the marginal benefit of learning this additional attribute with a stochastically dominant attribute.

Lemma 2. Let G_0, G_1 and G_2 be the cdfs of random variables y_0, y_1 and y_2. Suppose that all three random variables y_0, y_1 and y_2 have mean 0, i.e., $Ey_0 = Ey_1 = Ey_2 = 0$, and are symmetric. Suppose further that y_1 is positive first-order stochastically dominated by y_2, i.e., $y_1 \precsim^{p.f.o.} y_2$. Then

$$\int\int_{y_0+y_1\geq 0} (y_0 + y_1)\, dG_0(y_0) dG_1(y_1) - \int_0^{+\infty} y_1 dG_1(y_1) \leq$$
$$\leq \int\int_{y_0+y_2\geq 0} (y_0 + y_2)\, dG_0(y_0) dG_2(y_2) - \int_0^{+\infty} y_2 dG_2(y_2)$$

The following example illustrates Lemmas 1 and 2.

Example 1. Suppose y_a with cdf G_a is distributed uniformly on the interval $[-a, a]$, and y_b with cdf G_b is distributed uniformly on the interval $[-b, b]$, where $b < a$. That is, y_a is positively first-order stochastically dominated, and so is second-order stochastically dominated by y_b. Then

$$\int_0^{+\infty} y_a dG_a(y_a) = \frac{a}{4}; \int_0^{+\infty} y_b dG_b(y_b) = \frac{b}{4},$$

and

$$\int\int_{y_a+y_b\geq 0}(y_a+y_b)\,dG_a(y_a)dG_b(y_b)=\frac{a}{4}+\frac{b^2}{12a}. \tag{9}$$

Suppose first that $y_0=y_a$, and $y_i=y_{b_i}$ for some $b_2<b_1<a$. That is, $y_1 \precsim^{s.o.} y_2$. Lemma 1 in this case says that (9) increases in b.

Suppose now that $y_0=y_b$, and $y_i=y_{a_i}$ for some $b<a_2<a_1$. That is, $y_1 \precsim^{p.f.o.} y_2$. Lemma 2 in this case says that

$$\int\int_{y_a+y_b\geq 0}(y_a+y_b)\,dG_a(y_a)dG_b(y_b)-\int_0^{+\infty}y_a dG_a(y_a)=\frac{b^2}{12a}$$

decreases in a.

Proposition 1 follows immediately from Lemma 1. Proposition 2 is in turn a consequence of Lemma 2. However, the proof is a little less immediate in this case, so we relegate it, together with the proofs of Lemmas 1 and 2, to Appendix.

Finally, it is worth mentioning that Propositions 1 and 2 generalize to the sequential discovering of attributes. That is, if all costs are nondecreasing, then it is an optimal strategy to discover second-order stochastically dominated attributes earlier; and if the costs satisfy condition (8), then positive first-order stochastically dominant attributes should be discovered first.

4.2 Greedy algorithm

We will now discuss the *greedy algorithm*, which is a standard method of solving optimization problems in the discrete convex analysis. Suppose the objective function $U: P(N)\to R$ is defined on the set $P(N)$ of all subsets of some finite set $N=\{1,...,n\}$. The greedy algorithm recursively defines a subset $S\subset N$ by the following procedure: Suppose we have already defined a subset S; we begin with $S=\varnothing$. We compare the value of U at the set S and all sets $S\cup\{k\}$ for all $k\notin S$. If

$$U(S)\geq U(S\cup\{k\})$$

for every $k\notin S$, then the algorithm stops and returns the set S as the output. Otherwise, we take an arbitrary

$$k\in\arg\max_{k\notin S}U(S\cup\{k\}),$$

replace S with $S\cup\{k\}$ and repeat the procedure.

Under some conditions, the output of the greedy algorithm is an optimum of the objective function. Since the greedy algorithm refers to the values of U at only $1+n(n+1)/2$ subsets, it is a pseudo-polynomial algorithm. It turns out that the greedy algorithm solves the problems raised in Section 2.4 in some special cases, but not in the general case.

Indeed, suppose that the distributions $x_1, ..., x_n$ are symmetric and ordered by positive first-order dominance, i.e. $x_1 \precsim^{p.f.o.} ... \precsim^{p.f.o.} x_n$. To make the exposition more transparent, assume that the values of of U on different sets S sets are different. Let S denote the output set of the greedy algorithm, and let T denote the maximum of (1). Suppose that $U(T) > U(S)$. Consider the case that $S - T \neq \varnothing$ and $T - S \neq \varnothing$. (The other cases $T \subset S$ and $S \subset T$ are analogous and require no additional assumption.) Take $j := \min(S - T)$ and any $k \in T - S$. It cannot happen that $x_j \precsim^{p.f.o.} x_k$ and $c_j < c_k$, because then replacing k with j in the set T would increase the value of the objective function by Lemma 1; that is, we would have $U(T \cup \{j\} - \{k\}) > U(T)$. It cannot happen either that $x_k \precsim^{p.f.o.} x_j$ and $c_k < c_j$, because then, also by Lemma 1, the greedy algorithm would include k instead of j to the set S. Therefore, there are only two other cases:

(a) $x_k \precsim^{p.f.o.} x_j$ and $c_j \leq c_k$;

(b) $x_j \precsim^{p.f.o.} x_k$ and $c_j \geq c_k$.

Assume that

Condition 1. For any $S \subset \{1, ..., n\}$ and $k, j \notin S$ such that $k < j$,

$$x_k + \sum_{i \in S} x_i \precsim^{p.f.o.} x_j + \sum_{i \in S} x_i.$$

Notice that the fact that $x_1 \precsim^{s.o.} ... \precsim^{s.o.} x_n$ implies that $x_j + \sum_{i \in S} x_i \precsim^{s.o.} x_k + \sum_{i \in S} x_i$, but Condition 1 does not hold for some distributions $x_1, ..., x_n$ such that $x_1 \precsim^{p.f.o.} ... \precsim^{p.f.o.} x_n$. However, the condition seems to be a mild requirement, which is satisfied by many distributions of interest, including normal or uniform.

Under Condition 1, case (a) is impossible by Lemma 2. Indeed, the fact that the greedy algorithm included j not k implies that

$$U(T \cap S \cup \{j\}) > U(T \cap S \cup \{k\}). \tag{10}$$

Let $y_1 = \sum_{i \in T \cap S \cup \{k\}} x_i$, $y_2 = \sum_{i \in T \cap S \cup \{j\}} x_i$, and $y_0 = \sum_{i \in T - [S \cup \{k\}]} x_i$. Condition 1 implies that $y_1 \precsim^{p.f.o.} y_2$. Thus, by Lemma 2,

$$U([T - \{k\}] \cup \{j\}) - U(T \cap S \cup \{j\}) \geq U(T) - U(T \cap S \cup \{k\}). \tag{11}$$

Properties (10) and (11) together yield that

$$U([T - \{k\}] \cup \{j\}) > U(T).$$

Making case (b) impossible requires stronger assumptions regarding the distributions and costs. Intuitively, the greedy algorithm may include a stochastically dominated distribution x_j despite the fact that its cost is higher that the cost of a stochastically dominant distribution x_k, because it increases by more the integral component of the objective function. However, when we include other attributes, it may well happen that they are adding more to the integral component of the objective function if we include x_k instead of x_j. This is actually what Lemma 2 suggests. And this higher value added may offset the direct benefit from including x_j instead of x_k. Therefore, in order to make the greedy algorithm deliver a solution to problem (2), we must impose assumptions preventing this sort of scenario.

4.3 Polynomial-time algorithms

A naive algorithm solving problem (2) would be to compare the values of our objective function across all sets $S \subset \{1, ..., n\}$, and select the highest, but this would take 2^n time, and would be impractical except small values of n. We would prefer to have an algorithm whose time of running is polynomial in n.

The running time of an algorithm depends not only on the number of values $U(S)$ we have to compute, but also on the parameters of the model. These parameters include the number of values each random variable $x_1, ..., x_n$ is allowed to take, the size of these values, measured by the number of digits in their $0-1$ expansions, and the size the costs $c_1, ..., c_n$. Suppose the size of the parameters is commonly bounded by a number M. Then, an algorithm that is polynomial in M and n is called *pseudo polynomial*. To be truly polynomial, an algorithm must have running time that is independent in M and polynomial in n (*strongly polynomial*), or polynomial in $\log M$ and n (*weakly polynomial*). Polynomial-time algorithms (strongly or weakly) are considered practical in discrete convex analysis, while pseudo polynomial time algorithm are practical only for restricted sets of parameters.

We refer the reader to McCormick (2008) or Fujishige (2005) for a more-detailed discussion. These texts are to large extent devoted to the discussion of polynomial-time algorithms for minimization of submodular functions, or equivalently, the maximization of supermodular functions. In turn, there may not exist polynomial-time algorithms for maximization of submodular functions. Our objective function is (in the general case) neither supermodular nor submodular.

Theorem 2. There is no polynomial-time algorithm solving problem (2) even in the case in which each x_i is a binary variable taking value a_i with probability $1/2$ and value $-a_i$ with probability $1/2$.

Let $g(n)$ denote the expected value of

$$\max\left(0, \sum_{i=1}^{n} x_i\right).$$

The proof shows that if there exists a polynomial-time algorithm for evaluating $g(n)$ for every positive natural numbers $a_1, ..., a_n$, then P=NP. This result in turn follows from the fact that the following problem of partitioning is NP-complete (see (Garey and Johnson 1979)):

Given positive natural numbers $a_1, ..., a_n$, is there an $S \subset \{1, ..., n\}$ such that

$$\sum_{i \in S} a_i = \frac{1}{2} \sum_{i=1}^{n} a_i.$$

We relegate the details of the proof to Appendix.

Proposition 3. If there exists an algorithm solving problem (2) which requires computing $U(S)$ only for $P(n)$ sets S, where $P(n)$ is a polynomial function of n, then there also exists a pseudo polynomial-time algorithm solving problem (2).

This proposition follows immediately from the following lemma:

Lemma 3. There exists a pseudo polynomial-time algorithm for computing $g(n)$.

The proof of this lemma is relegated to Appendix.

Propositions 1-3 imply the existence of pseudo polynomial-time algorithm in some special cases. The existence of a pseudo polynomial-time algorithm in the general case is an open question, even when the attributes are ordered by positive first-order stochastic dominance. It is also an open question whether there exists an approximate polynomial-time algorithm solving problem (2). Recall that an *approximation algorithm* A with an *approximation ratio* r is a polynomial-time algorithm with the property that for any set of parameters of the model,

$$\frac{U^* - V}{U^*} \leq r,$$

where $U^* := \max_{S \subset \{1,...,n\}} U(S)$ and V denotes the returned value of algorithm A.

5 Appendix

5.1 Proofs of Lemmas 1 and 2, and Proposition 2

Proof of Lemma 1. Notice that

$$\int \int_{y_0 + y_i \geq 0} (y_0 + y_i)\, dG_0(y_0) dG_i(y_i) =$$

$$= \int_{-\infty}^{+\infty} \left[\int_{-y_i}^{+\infty} (y_0 + y_i)\, dG_0(y_0) \right] dG_i(y_i) =$$

$$= \int_{-\infty}^{+\infty} \left[\left(\int_{-y_i}^{+\infty} y_0 dG_0(y_0) \right) + y_i(1 - G_0(-y_i)) \right] dG_i(y_i) =$$

$$= \int_{-\infty}^{+\infty} \left[\left(\int_{-y_i}^{+\infty} y_0 dG_0(y_0) \right) - y_i G_0(-y_i) \right] dG_i(y_i) =$$

$$= \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} \max(y_0, -y_i) dG_0(y_0) \right] dG_i(y_i),$$

where the third equality follows from the assumption that $Ey_i = 0$.

So, it suffices to observe that

$$\int_{-\infty}^{+\infty} \max(y_0, -y_i) dG_0(y_0)$$

is a convex function of y_i. This, however, follows from the fact that

$$\alpha^{'} \max(y_0, -y_i^{'}) + \alpha^{''} \max(y_0, -y_i^{''}) \geq \max(y_0, -\alpha^{'} y_i^{'} - \alpha^{''} y_i^{''})$$

for every $y_0, y_i^{'}, y_i^{''}$, and $\alpha^{'}, \alpha^{''} \geq 0$ such that $\alpha^{'} + \alpha^{''} = 1$.

In the proof of Lemma 2, we will need the following straightforward claim.

Claim 1. If a function $h : R \to R$ is increasing on $(-\infty, 0)$ and decreasing on $(0, +\infty)$, and y_1 with cdf G_1 is positive first-order stochastically dominated by y_2 with cdf G_2, where y_1 and y_2 are symmetric, then

$$\int_{-\infty}^{+\infty} h(y_1) dG_1(y_1) \leq \int_{-\infty}^{+\infty} h(y_2) dG_2(y_2).$$

The claim follows from the fact that a positively first-order stochastically dominated variable is obtained from a stochastically dominant variable by moving the probability from the center to the tails of the distribution, where the values of h are lower.

Proof of Lemma 2. Notice that

$$\int \int_{y_0 + y_i \geq 0} (y_0 + y_i)\, dG_0(y_0) dG_i(y_i) - \int_0^{+\infty} y_i dG_i(y_i) =$$

$$= \int_{-\infty}^{+\infty} \left[\int_{-y_i}^{+\infty} (y_0 + y_i)\, dG_0(y_0) - \max(0, y_i) \right] dG_i(y_i) =$$

$$= \int_{-\infty}^{+\infty} \left[\left(\int_{-y_i}^{+\infty} y_0 dG_0(y_0) \right) + y_i(1 - G_0(-y_i)) - \max(0, y_i) \right] dG_i(y_i) =$$

$$= \int_{-\infty}^{+\infty} \left[\left(\int_{-y_i}^{+\infty} y_0 dG_0(y_0) \right) + f(y_i) \right] dG_i(y_i),$$

where

$$f(y_i) = y_i(1 - G_0(-y_i)) - \max(0, y_i) = \begin{cases} -y_i G_0(-y_i) \text{ if } y_i \geq 0 \\ \\ y_i(1 - G_0(-y_i)) \text{ if } y_i \leq 0 \end{cases}.$$

Since the distribution of y_0 is symmetric around zero, $f(y_i) = -f(-y_i)$; and hence, since the distribution of y_i is also symmetric around zero,

$$\int_{-\infty}^{+\infty} f(y_i) dG_i(y_i) = 0.$$

And

$$\int_{-\infty}^{+\infty} \left(\int_{-y_i}^{+\infty} y_0 dG_0(y_0) \right) dG_i(y_i) = \int_{-\infty}^{0} \left(\int_{-y_i}^{+\infty} y_0 dG_0(y_0) \right) dG_i(y_i) +$$

$$+ \int_{0}^{+\infty} \left(\int_{-y_i}^{+y_i} y_0 dG_0(y_0) + \int_{+y_i}^{+\infty} y_0 dG_0(y_0) \right) dG_i(y_i) =$$

$$= \int_{-\infty}^{0} \left(\int_{-y_i}^{+\infty} y_0 dG_0(y_0) \right) dG_i(y_i) + \int_{0}^{+\infty} \left(\int_{+y_i}^{+\infty} y_0 dG_0(y_0) \right) dG_i(y_i) =$$

$$= \int_{-\infty}^{+\infty} h(y_i) dG_i(y_i),$$

where

$$h(y_i) = \begin{cases} \int_{+y_i}^{+\infty} y_0 dG_0(y_0) \text{ if } y_i \geq 0 \\ \\ \int_{-y_i}^{+\infty} y_0 dG_0(y_0) \text{ if } y_i \leq 0 \end{cases}.$$

The second equality follows from the symmetry of y_0, since the symmetry implies that

$$\int_{-y_i}^{+y_i} y_0 dG_0(y_0) = 0.$$

Notice that $h(y_i) = h(-y_i)$, is increasing in y_i for $y_i \leq 0$, and decreasing in y_i for $y_i \geq 0$. Thus, Lemma 2 now follows from Claim 1.

Proof of Proposition 2. Suppose that some set T solving problem (2) does not have the form $T = \{k, ..., n\}$. Consider the set R obtained from T by replacing any $i \in T$ with any $i + j \notin T$, where $j > 0$. By Lemma 2,

$$\int ... \int_{x_i, i \in R} \max\left(0, \sum_{i \in R} x_i\right) - \int ... \int_{x_i, i \in T} \max\left(0, \sum_{i \in T} x_i\right) \geq$$

$$\geq \int_{0}^{+\infty} x_{i+j} dF_{i+j}(x_{i+j}) - \int_{0}^{+\infty} x_i dF_i(x_i).$$

Apply Lemma 2 to the variable

$$x_0 = \sum_{i \in T - \{i\}} x_i = \sum_{i \in R - \{i+j\}} x_i.$$

By assumption,

$$\int_0^{+\infty} x_{i+j} dF_{i+j}(x_{i+j}) - \int_0^{+\infty} x_i dF_i(x_i) \geq c_{i+j} - c_i.$$

Therefore, since the cost of the attributes in R and the cost of the attributes in T differ only by $c_{i+j} - c_i$, it follows that the value of (1) for R is no lower than the value of (1) for T.

Replacing recursively the attributes in T by attributes that stochastically dominate them, we obtain a set of the form $S = \{k, ..., n\}$ for which the value of (1) is no lower than that for T.

5.2 Proof of Theorem 2

Suppose that $a_1, ..., a_n > 0$. Notice that for $S = \{i : x_i = -a_i\}$ and $T = \{i : x_i = a_i\}$, we have

$$\sum_{i=1}^{n} x_i = \sum_{i \in T} a_i - \sum_{i \in S} a_i = \left(\sum_{i=1}^{n} a - \sum_{i \in S} a_i \right) - \sum_{i \in S} a_i +_i = -2 \sum_{i \in S} a_i + \sum_{i=1}^{n} a_i,$$

and so

$$\sum_{i=1}^{n} x_i \geq 0 \;\Leftrightarrow\; \sum_{i \in S} a_i \leq \frac{1}{2} \sum_{i=1}^{n} a_i.$$

Thus,

$$g(n) = \frac{1}{2^n} \sum \left\{ \left(-2 \sum_{i \in S} a_i + \sum_{i=1}^{n} a_i \right) : S \subset \{1, ..., n\},\; \sum_{i \in S} a_i \leq \frac{1}{2} \sum_{i=1}^{n} a_i \right\} =$$

$$= \frac{1}{2^{n-1}} \sum \left\{ \left(\frac{1}{2} \sum_{i=1}^{n} a_i - \sum_{i \in S} a_i \right) : S \subset \{1, ..., n\},\; \sum_{i \in S} a_i \leq \frac{1}{2} \sum_{i=1}^{n} a_i \right\}$$

Define $a = (a_1, ..., a_n)$, $\overline{a} = (a_1, ..., a_n, 1)$,

$$M = M(a) := \frac{1}{2} \sum_{i=1}^{n} a_i,$$

and for any $\varnothing \neq S \subset \{1, ..., n\}$,

$$a(S) := \sum_{i \in S} a_i.$$

Define in a similar manner: $\overline{a}(\overline{S})$ and $\overline{M}$. Assume, without loss of generality, that M is an integer.

Further, define

$$h(a) := \sum_{S \subset \{1,...,n\} : a(S) \leq M} (M - a(S)),$$

$$f(a) := \sum_{S \subset \{1,\dots,n\} : a(S) \le M} 1,$$

$$h(a,-1) = \sum_{S \subset \{1,\dots,n\} : a(S) \le M-1} (M - 1 - a(S))$$

and

$$f(a,-1) := \sum_{S \subset \{1,\dots,n\} : a(S) \le M-1} 1.$$

Since all numbers $a(S)$ are integers, $M - a(S) = 0$ in the expression for $h(a)$ unless $a(S) \le M - 1$. This yields that

$$h(a) = h(a,-1) + f(a,-1). \tag{12}$$

Further, observe that

$$-1 = f(a) + f(a,-1) - f(\overline{a}). \tag{13}$$

Indeed, there are three cases: **(i)** If $n+1 \notin \overline{S} \subset \{1,...,n+1\}$, then $\overline{a}(\overline{S}) = a(\overline{S}) \le M + 1/2 \Leftrightarrow a(\overline{S}) \le M$; **(ii)** If $n+1 \in \overline{S} \neq \{n+1\}$, then for $S := \overline{S} - \{n+1\}$, we have $\overline{a}(\overline{S}) = a(S) + 1 \le M + 1/2 \Leftrightarrow a(S) \le M - 1$; (iii) If $\overline{S} = \{n+1\}$, then $\overline{a}(\overline{S}) = 1$. This yields that $f(\overline{a}) = f(a) + f(a,-1) + 1$.

We also have

$$h(\overline{a}) = \sum_{\overline{S} \subset \{1,\dots,n+1\} : \overline{a}(\overline{S}) \le \overline{M}} (\overline{M} - \overline{a}(\overline{S})) =$$

$$= \sum_{\overline{S} \subset \{1,\dots,n\} : \overline{a}(\overline{S}) \le \overline{M}} (\overline{M} - \overline{a}(\overline{S})) + \sum_{\{n+1\} \in \overline{S} \subset \{1,\dots,n+1\} : \overline{a}(\overline{S}) \le \overline{M},\ \overline{S} \neq \{n+1\}} (\overline{M} - \overline{a}(\overline{S})) + \left(\overline{M} - 1\right) =$$

$$= \sum_{S \subset \{1,\dots,n\} : a(S) \le M} (M + 1/2 - a(S)) + \sum_{S \subset \{1,\dots,n\} : a(S) \le M-1} (M + 1/2 - a(S)) + (M - 1/2) =$$

$$= h(a) + \frac{1}{2} f(a) + h(a,-1) + \frac{1}{2} f(a,-1) + M - 1/2.$$

This is equivalent to

$$2\left(h(\overline{a}) - h(a) + 1/2 - M\right) = f(a) + f(a,-1) + 2h(a,-1). \tag{14}$$

Finally, define $u(a)$ as the number of sets S such that $a(S) = M(a)$. Clearly,

$$u(a) = \sum_{S \subset \{1,\dots,n\} : a(S) = M} 1 = \sum_{S \subset \{1,\dots,n\} : a(S) \le M} 1 - \sum_{S \subset \{1,\dots,n\} : a(S) \le M-1} 1 = f(a) - f(a,-1);$$

that is,

$$0 = -u(a) + f(a) - f(a,-1). \tag{15}$$

We also have

$$\sum_{S\subset\{1,...,n\}:a(S)\le M-1} 1+\sum_{S\subset\{1,...,n\}:a(S)=M} 1+\sum_{S\subset\{1,...,n\}:a(S)\ge M+1} 1=2^n-1.$$

Since $a(S)\le M-1 \Leftrightarrow a(T)\ge M+1$, where $T:=\{1,...,n\}-S$,

$$2\sum_{S\subset\{1,...,n\}:a(S)\le M-1} 1+\sum_{S\subset\{1,...,n\}:a(S)=M} 1=2^n-1;$$

that is,

$$2^n-1=2f(a,-1)+u(a) \tag{16}$$

If we were able to compute $g(n)$ in polynomial time, then we would able to compute $h(a)$. However, then we would also be able to compute $f(a)$, $f(\overline{a})$, $f(a,-1)$, $h(a,-1)$, $u(a)$ in polynomial time, by solving the system of equation (12)-(16). Indeed, the matrix

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 1 & -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 2 & 0 \\ 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 2 & 0 & 1 \end{bmatrix}$$

is nonsingular.

However, the condition $u(a)=0$ is equivalent to the existence of an S that solves the problem of partitioning.

5.3 Proof of Lemma 3

Suppose that variable x_i, $i=1,...,n$, takes values a_i from a finite set S_i.

For $t=1,...,n$ and $a=(a_1,...,a_t)$, define

$$p_t(a):=\prod_{i=1}^{t}\Pr\{x_i=a_i\},$$

and for any $t=1,...,n$, define

$$V_t(b)=\sum\left\{p_t(a)\cdot\left(\sum_{i=1}^{t}a_i\right):\sum_{i=1}^{t}a_i\ge b\right\}.$$

We will develop a recursive definition of $V_t(b)$:

$$V_t(b)=\sum\left\{p_t(a)\cdot\left(a_t+\sum_{i=1}^{t-1}a_i\right):\sum_{i=1}^{t}a_i\ge b\right\}=$$

$$= \sum \left\{ p_t(a) \cdot a_t + \Pr\{x_t = a_t\} p_{t-1}(a_1, ..., a_{t-1}) \cdot \left(\sum_{i=1}^{t-1} a_i \right) : \sum_{i=1}^{t} a_i \geq b \right\} =$$

$$= \sum \left\{ p_t(a) \cdot a_t : \sum_{i=1}^{t} a_i \geq b \right\} + \sum_{a_t} \Pr\{x_t = a_t\} V_{t-1}(b - a_t).$$

Let

$$U_t(b) := \sum \left\{ p_t(a) \cdot a_t : \sum_{i=1}^{t} a_i \geq b \right\}.$$

Then

$$U_t(b) = \sum_{a_t} a_t \Pr\{x_t = a_t\} V_{t-1}(b - a_t) \cdot \left(\sum \left\{ p_{t-1}(a_1, ..., a_{t-1}) : \sum_{i=1}^{t-1} a_i \geq b - a_t \right\} \right).$$

Let

$$z_t(b) := \sum \left\{ p_t(a) : \sum_{i=1}^{t} a_i \geq b \right\}.$$

Then

$$z_t(b) = \sum_{a_t} \Pr\{x_t = a_t\} \cdot \left(\sum \left\{ p_{t-1}(a_1, ..., a_{t-1}) : \sum_{i=1}^{t-1} a_i \geq b - a_t \right\} \right) =$$

$$= \sum_{a_t} \Pr\{x_t = a_t\} \cdot z_{t-1}(b - a_t).$$

Therefore $z_t(b)$ can be computed recursively, and so $U_t(b)$ and $V_t(b)$ can be computed recursively.

Let

$$\alpha := \sum_{i=1}^{n} \max_{a_i \in S_i} |a_i|$$

and

$$s := \max_{i=1,...,n} |S_i|.$$

The recursive definition of $V_t(b)$ provides an algorithm for computing $g(n) = V_n(0)$ which is polynomial-time in α, s, and n.

6 References

Chade H. and L. Smith (2005): "Simultaneous Search," *Econometrica* **74**, 1293–1307.

Fujishige S. (2005). *Submodular Functions and Optimization*. Elsevier, Amsterdam.

Garey M.R. and D.S. Johnson (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman & Co., New York.

Gittins, J.C. and D.M. Jones (1974): "A dynamic allocation index for the sequential design of experiments," in *Progress in Statistics (J. Gani et al., eds.)*, 241–266, North-Holland, Amsterdam.

Gittins, J.C. (1989). *Multi-armed bandit allocation indices.* John Wiley & Sons, Ltd., New York.

Manzini, P. and M. Mariotti (2007): "Sequentialy Rationalizable Choice," *American Economic Review*, **97**, 1824-1839.

McCormick S.T. (2008): "Submodular Function Optimization," Mimeo.

Neeman, Z. (1995): "On Determining the Importance of Attributes with a Stopping Problem," *Mathematical Social Sciences*, **29**, 195-212.

Tversky, A. (1972): "Elimination by Aspects: A Theory of Choice," *Psychological Review*, **79**, 281–99.

Weber, R. (1992): "On the Gittins Index for Multiarmed Bandits," *Annals of Applied Probability*, **2**, 1024-1033.

Weitzman, M.L. (1979): "Optimal Search for the Best Alternative," *Econometrica*, **47**, 641-654.

Whittle, P. (1980): "Multi-armed Bandits and the Gittins Index," *Journal of the Royal Statistical Society Ser. B (Methodology)*, **42**, 143–149.

	Breath	Teach. Skills	Depth
Cost of Check.	4	6	30
Value of Good Outcome	+108	+112	+160
Value of Bad Outcome	-108	-112	-160

Table 1.

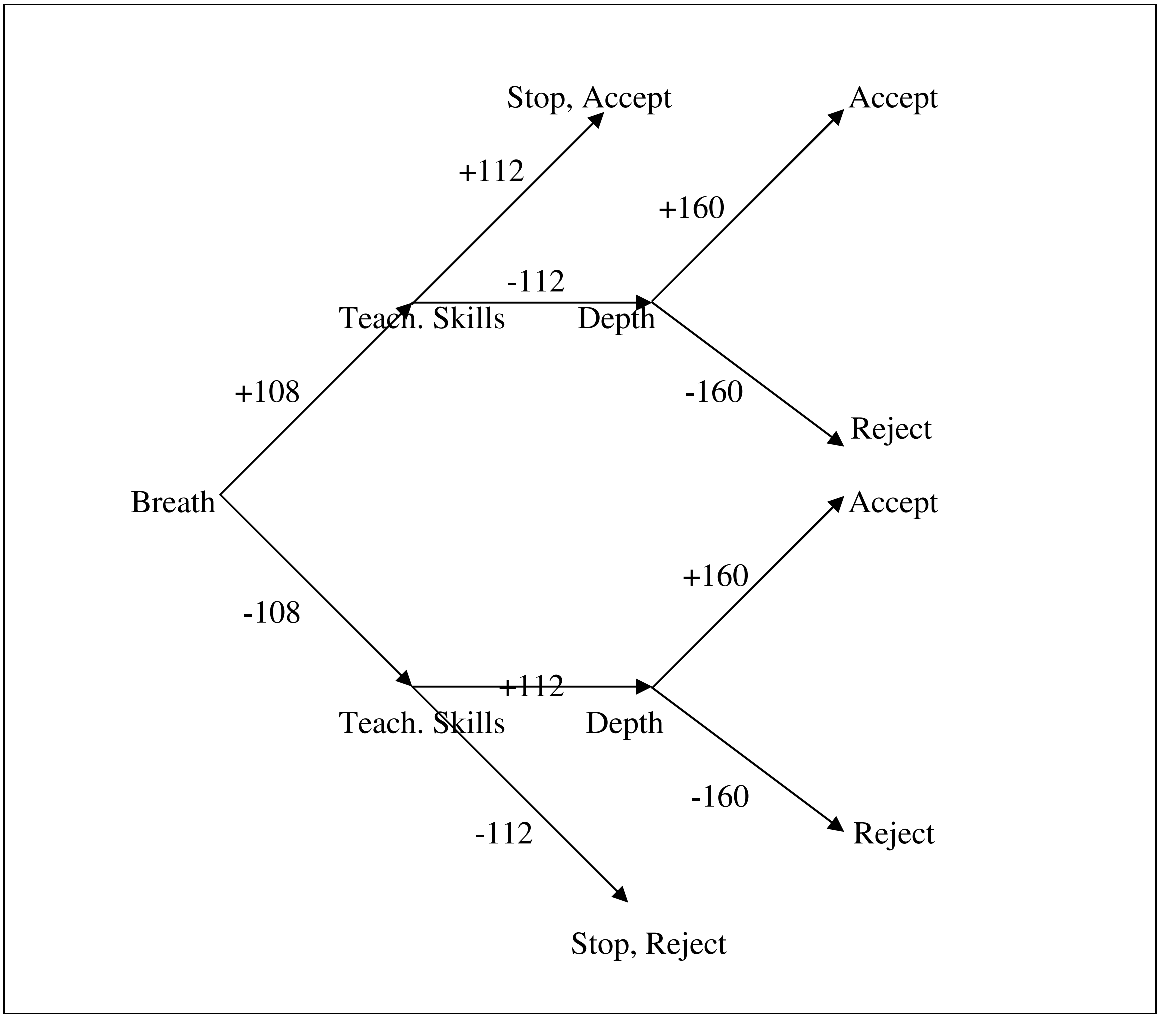

Stop, Accept
Accept
+112
+160
-112
Teach. Skills
Depth
+108
-160
Reject
Breath
Accept
+160
-108
+112
Teach. Skills
Depth
-160
-112
Reject
Stop, Reject

Figure 1.

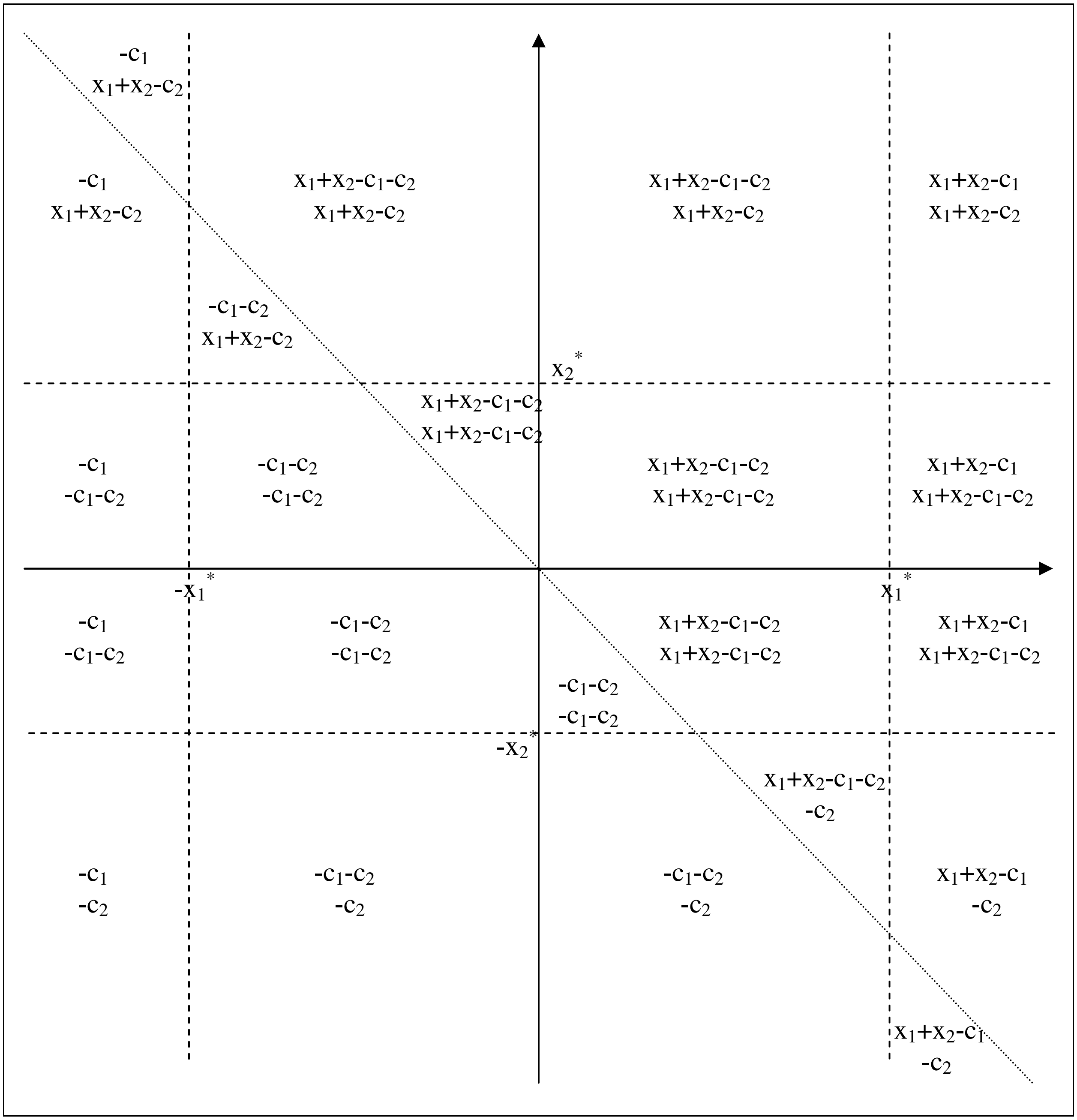

$-c_1$
$x_1+x_2-c_2$
$-c_1$
$x_1+x_2-c_2$
$x_1+x_2-c_1-c_2$
$x_1+x_2-c_2$
$x_1+x_2-c_1-c_2$
$x_1+x_2-c_2$
$x_1+x_2-c_1$
$x_1+x_2-c_2$
$-c_1-c_2$
$x_1+x_2-c_2$
x_2^*
$x_1+x_2-c_1-c_2$
$x_1+x_2-c_1-c_2$
$-c_1$
$-c_1-c_2$
$-c_1-c_2$
$-c_1-c_2$
$x_1+x_2-c_1-c_2$
$x_1+x_2-c_1-c_2$
$x_1+x_2-c_1$
$x_1+x_2-c_1-c_2$
$-x_1^*$
x_1^*
$-c_1$
$-c_1-c_2$
$-c_1-c_2$
$-c_1-c_2$
$x_1+x_2-c_1-c_2$
$x_1+x_2-c_1-c_2$
$x_1+x_2-c_1$
$x_1+x_2-c_1-c_2$
$-c_1-c_2$
$-c_1-c_2$
$-x_2^*$
$x_1+x_2-c_1-c_2$
$-c_2$
$-c_1$
$-c_2$
$-c_1-c_2$
$-c_2$
$-c_1-c_2$
$-c_2$
$x_1+x_2-c_1$
$-c_2$
$x_1+x_2-c_1$
$-c_2$

Figure 2 (a).

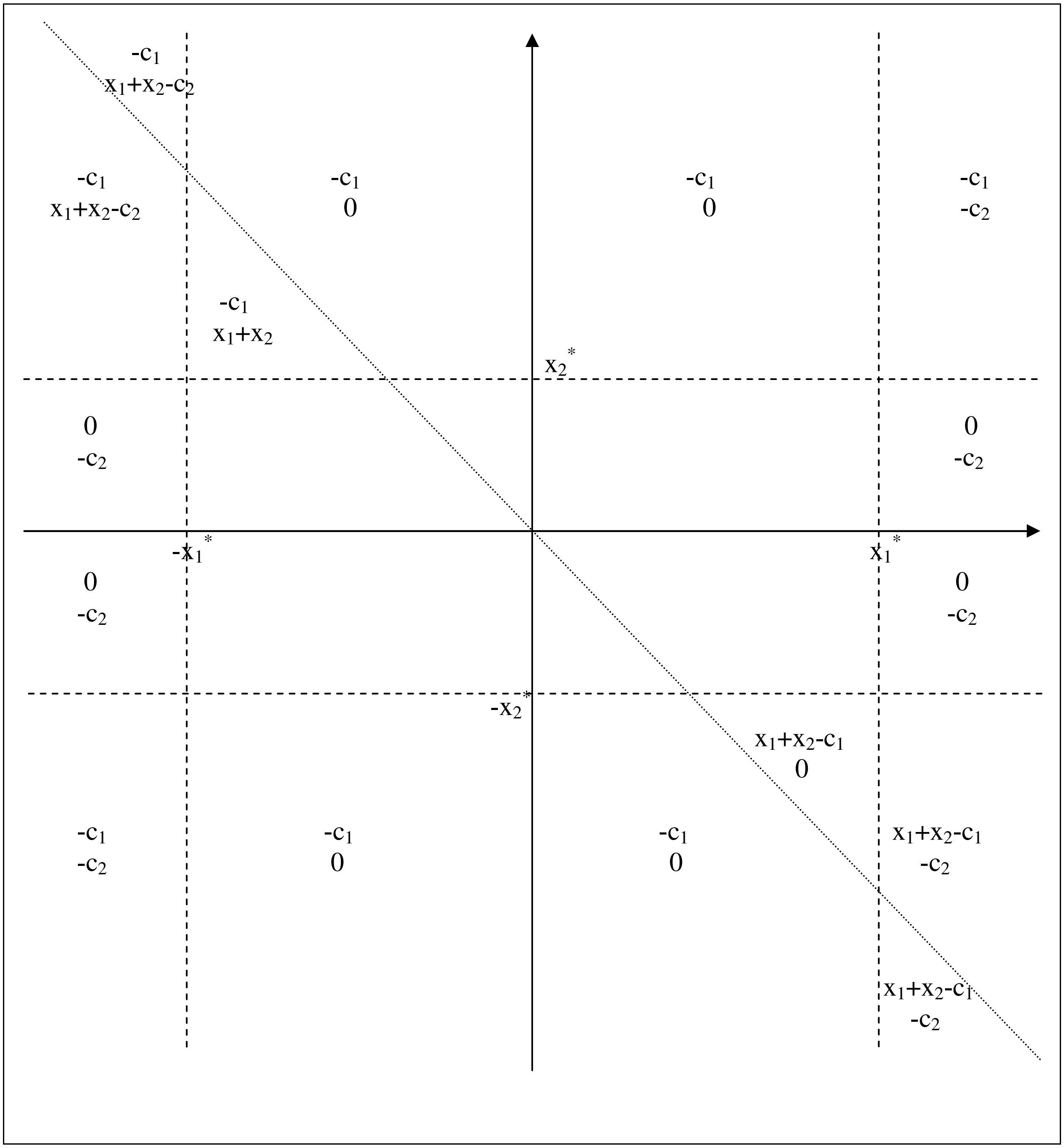
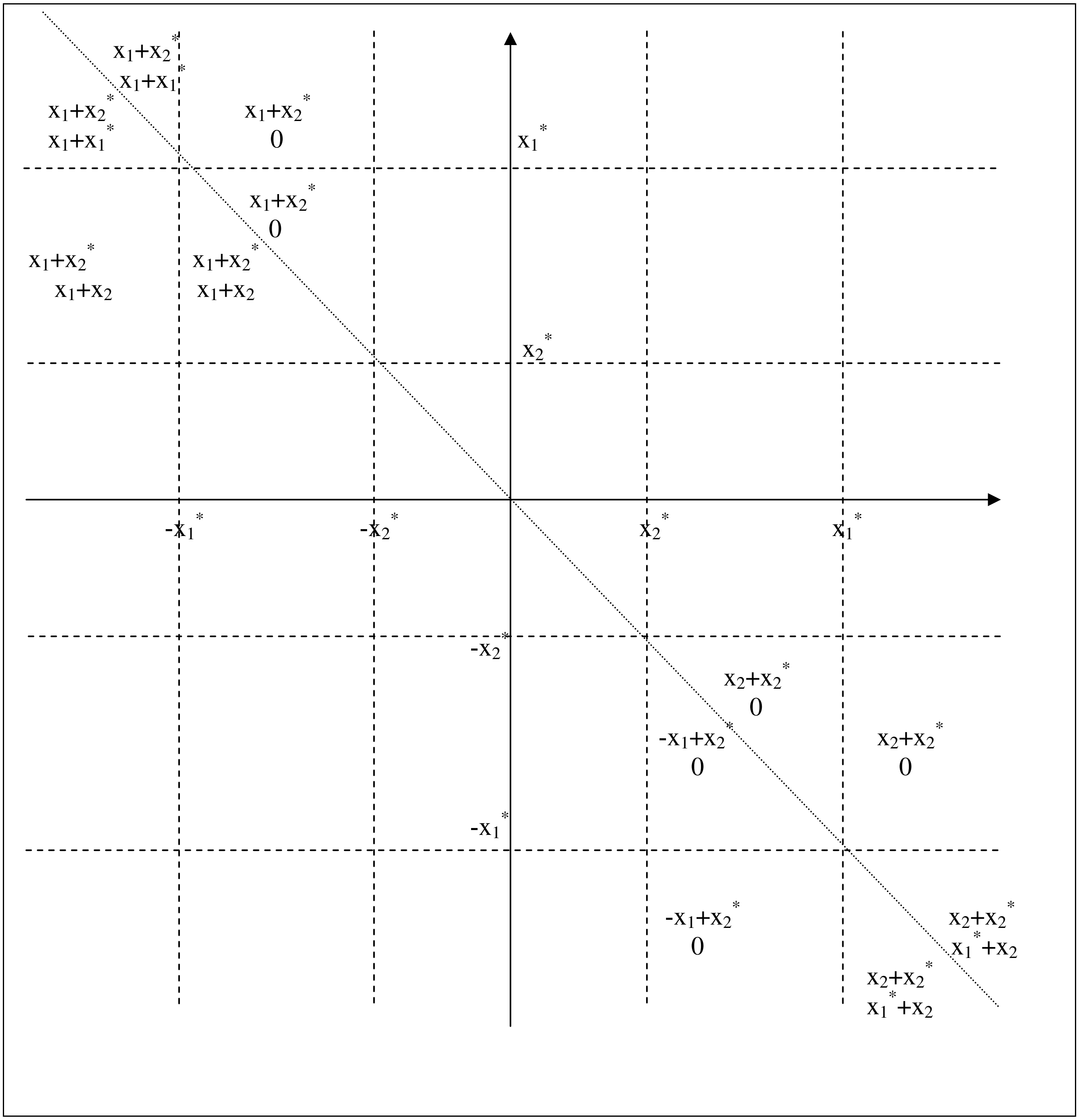

-c1
x1+x2-c2
-c1
x1+x2-c2
-c1
0
-c1
0
-c1
-c2
-c1
x1+x2
x2*
0
-c2
0
-c2
-x1*
x1*
0
-c2
0
-c2
-x2*
x1+x2-c1
0
-c1
-c2
-c1
0
-c1
0
x1+x2-c1
-c2
x1+x2-c1
-c2

Figure 2 (b).



x1+x2*
x1+x1*
x1+x2*
x1+x1*
x1+x2*
0
x1*
x1+x2*
0
x1+x2*
x1+x2
x1+x2*
x1+x2
x2*
-x1*
-x2*
x2*
x1*
-x2*
x2+x2*
0
-x1+x2*
0
x2+x2*
0
-x1*
-x1+x2*
0
x2+x2*
x1*+x2
x2+x2*
x1*+x2

Figure 2 (c).