Dynamic decision making without expected utility: an operational approach

Thomas D. Nielsen*

Department of Computer Science

Fredrik Bajers Vej 7E, 9220 Aalborg Øst

Denmark

Jean-Yves Jaffray

LIP6, Université Paris 6

8, rue du Capitaine Scott 75015 Paris

France

May 24, 2004

*Corresponding author.

Abstract

Non-expected utility theories, such as rank dependent utility (RDU) theory, have been proposed as alternative models to EU theory in decision making under risk. These models do not share the separability property of expected utility theory. This implies that, in a decision tree, if the reduction of compound lotteries assumption is made (so that preferences at each decision node reduce to RDU preferences among lotteries) and that preferences at different decision nodes are identical (same utility function and same weighting function), then the preferences are not dynamically consistent; in particular, the sophisticated strategy, i.e., the strategy generated by a standard rolling back of the decision tree, is likely to be dominated w.r.t. stochastic dominance. Dynamic consistency of choices remains feasible, and the decision maker can avoid dominated choices, by adopting a non-consequentialist behavior, with his choices in a subtree possibly depending on what happens in the rest of the tree. We propose a procedure which: i) although adopting a non-consequentialist behavior, involves a form of rolling back of the decision tree; and ii) selects a non-dominated strategy that realizes a compromise between the decision maker's discordant goals at the different decision nodes. Relative to the computations involved in the standard expected utility evaluation of a decision problem, the main computational increase is due to the identification of non-dominated strategies by linear programming. A simulation, using the rank dependent utility criterion, confirms the computational tractability of the model.

Keywords: Decision analysis, utility theory, uncertainty modeling.

Introduction

The standard normative model for decision making under risk, i.e., under situations of uncertainty in which events are endowed with probabilities, is *expected utility (*EU) *theory*: the value of a decision is the sum of the utilities of its potential consequences weighted by the probabilities of their occurrence. As a descriptive model however, EU theory performs poorly as demonstrated by numerous lab and field experiments (Machina 1987). These findings have the following implications: (i) the elicitation of the decision maker's (DM's) personal parameter in the EU model, his *von Neumann-Morgenstern* (vNM) *utility* function, poses a problem, since different methods are likely to produce different utility functions and there is no indisputable ground for preferring one to the others; (ii) in any case, whatever utility function is selected, there will be some simple decision situations where the DM is confident about his real preferences and disagrees with the choices advised by EU theory. This may make him reluctant to accept EU theory as a prescriptive model, and (iii) even if the DM is intellectually willing to use EU theory, he may feel unable to follow some of its prescriptions in real life problems.

For instance, take the following example inspired by the Allais paradox (Allais 1979). The DM (a firm), can carry out either one of two important 10 year projects in a politically unstable country, where a coup d'Etat is a permanent threat; in case of a coup, the project would be definitively stopped and the DM would sustain a very big loss L (possible bankruptcy). These projects differ in terms of investments and profits, as well as in the types of contracts. If he chooses the first project (decision Δ_1), the contract will specify that, after a 5-year period (and if no coup has taken place yet), he will have the opportunity to decide between continuing the project alone (decision D_1), selling his rights to the local state entirely (decision D_2) or offering the state a 50% level of participation (decision D_3). If he chooses the second project (decision Δ_2), the contract will not allow a mid-term revision. Thus the initially feasible strategies are (Δ_1, D_1), (Δ_1, D_2), (Δ_1, D_3) and Δ_2, but, if Δ_1 is initially chosen, the actual choice between D_1, D_2 and D_3 will only take place at midterm and conditionally on the absence of a coup.

Suppose that the DM ranks the strategies as follows: $(\Delta_1, D_1) \succ (\Delta_1, D_3) \succ \Delta_2 \succ (\Delta_1, D_2)$. Suppose moreover that the DM is subject to the *certainty effect*, i.e., he "overweights" the worst outcomes of the decisions in a way that EU theory cannot account for (see e.g. (Kahneman and Tversky 1979)). This effect is unlikely to play an important role in the preceding ranking since the DM incurs the risk of losing L for all the strategies (in case of a coup during the first 5 years). However, should he, after Δ_1, have to choose among D_1, D_2 and D_3, the fact that with D_2 he is sure to avoid losing L, while with D_1 and D_3 he risks losing L and $\frac{1}{2}\mathsf{L}$ respectively, may prove to be decisive in favor of D_2, e.g., $\mathsf{D}_2 \succ \mathsf{D}_3 \succ \mathsf{D}_1$.

EU theory, when applied with a vNM utility constructed by a "certainty effect bias free" method, e.g. the probability equivalent method (McCord and de Neufville 1984), would confirm the DM's initial ordering, and thus recommend strategy (Δ_1, D_1). Suppose the DM has started applying the strategy by taking decision Δ_1 and there has been no coup after 5 years; pursuing with D_1, when D_2 is so more attractive to him (no risk of bankruptcy), may require an amazing strength of resolution, a trait which we shall call *flexibility.* If the DM has such strength, he has no problem in applying the EU model.

On the other hand, consider another DM who has limited strength of resolution (limited flexibility) and is only able to deviate from D_2 to his second best choice D_3; aware of this, the DM will still initially choose Δ_1 against Δ_2 since (Δ_1, D_3) is the second best for the initial ordering and he is confident that he will later choose D_3 if this is his initial plan. A third DM, who knows he is unable to deviate at all from his preferences will choose Δ_2 which he prefers to (Δ_1, D_2). These are three different ways of coping with a situation where preferences (the same in the three cases) are dynamically inconsistent; on what basis should one be judged rational and another not? As a matter of fact, two different conceptions of rationality are recognizable in the normative arguments in favor of EU theory: The traditional line of argument for convincing a reluctant DM to conform to EU theory against his own judgment, consists of making him approve of a series of general "rationality principles" which his preferences should respect and which jointly are only satisfied by EU theory. In fact, it has been shown that it suffices to invoke the following three principles: *consequentialism*, *dynamic consistency* and *reduction of compound lotteries* (Hammond 1988). This elegant argument has a weak point though: the justification of each of these principles comes from the mere assertion of its natural or common sense character.

For this reason, the defense of EU theory has concurrently followed an alternative line of argument, first used by (de Finetti 1974) for justifying the universal use of subjective probabilities. The underlying idea on which de Finetti based his so-called *Dutch book argument* is that rationality is not a matter of preference but a matter of behavior: according to de Finetti, in the particular situation he considers, a DM behaves rationally if his actions do not make him risk sure loses. The extension of this idea to decision making under uncertainty in general, where the DM chooses among strategies (i.e., sequences of conditional decisions), suggests the following definition: *a DM is rational if his behavioral rule can never make him choose a dominated strategy.* In this definition, "dominated" refers to *point-wise* (or more properly: event-wise) *dominance*: a strategy dominates another when it leads to a better consequence whatever happens and to a strictly better consequence for some event. *Money-pump arguments* which have been advanced for justifying transitivity and independence axioms for preferences rest on this definition of rationality (Raiffa 1968).

In the special case of decision making under risk, however, events are endowed with probabilities, so that each strategy generates a lottery (a probability distribution) on the consequence set. Lotteries, hence also strategies, are partially ordered by *(first order) stochastic dominance* (FSD), an ordering which has a strong normative appeal since it is a direct extension of point-wise dominance: given two lotteries, $L_1$and L_2 such that L_1 FSD L_2, it is possible to exhibit a decision situation in which two decisions d_1 and d_2 generate $L_1$and L_2, respectively, and d_1 leads to a better consequence than d_2 on every event and to a strictly better one on some event.

Thus, rationality requirements under risk can be slightly reinforced with the following definition: *a DM is rational if his behavioral rule can never make him choose a FSD-dominated strategy.* Such a definition conforms to economic theory practice, which always limits rationality requirements to "no wasting" precepts; for instance, collective rationality only requires allocations among individuals to be Pareto optimal.

Thus, this definition captures an aspect of rationality –say, economic rationality– which arguably should be the only relevant aspect for decision aiding in managerial and economic applications. Henceforth we adopt this definition of rational behavior.

Coming back to our example, the choice of either strategy (Δ_1, D_3) or strategy Δ_2 by the DM can only be a proof of irrational behavior if the strategy is dominated by another feasible strategy. As a matter of fact, it suffices to suppose that Δ_2 is close to (Δ_1, D_1) but dominated by it, to show the irrationality of the totally inflexible non-EU maximizers. Observe that this argument does not extend to a flexible DM.

In this paper we present a decision model in which the DM has non-EU preferences (represented by a rank dependent utility (RDU) criterion (Quiggin 1982)) and behaves rationally, in the sense that in any decision tree he will select a non-dominated strategy, although he has non-EU preferences. The motivation for focusing on this particular class of non-EU preferences is to eliminate the main difficulties encountered in utility elicitation. Once the two parameters of the RDU model (the utility function and the probability weighting function) have been assessed, the strategy selection procedure is algorithmic and can be applied to any decision tree. The algorithm works by backward induction in the tree, but, due the fact that RDU preferences are not dynamically consistent, departs significantly from the standard dynamic programming algorithm; in particular, when a decision node is reached: instead of selecting one substrategy, a set of "acceptable" substrategies is selected, and the final choice among them is made at the next step. This feature of the procedure is given a psychological interpretation in terms of local preference flexibility. Poor flexibility may imply irrationality and the algorithm will then fail to produce a non-dominated strategy. Note that if the DM is an EU maximizer, the procedure always selects his optimal strategy, even if he has no flexibility at all.

The remainder of the paper is organized as follows. In Section 1 we first briefly review some basic concepts and notation relating to decision making based on both expected and non-expected utility; then we describe and discuss the model and outline the structure of the proposed algorithm. The algorithm is based on the identification of dominated strategies which is considered in Section 2. In Section 3 the proposed algorithm is specified and in Section 4 we present results brought by a simulation.

1 Preliminaries

1.1 Definitions and Notation

1.1.1 Decision trees

A *directed graph* $\mathsf{G} = (\mathcal{N}, \mathcal{E})$ consists of a finite set of nodes $\mathcal{N}$ and a finite set of arcs $\mathcal{E}$. An arc is a pair of nodes (X, Y), where $\mathsf{X}, \mathsf{Y} \in \mathcal{N}$ and $\mathsf{X} \neq \mathsf{Y}$; we say that X is the *parent* of Y and Y is the *child* of X.

A *directed path* from a node X to a node Y in a directed graph $\mathsf{G} = (\mathcal{N}, \mathcal{E})$ is a set of arcs $\mathcal{E}' = \{(\mathsf{X}_1, \mathsf{X}_2), (\mathsf{X}_2, \mathsf{X}_3), \ldots, (\mathsf{X}_{n-1}, \mathsf{X}_n)\}$, where $\mathcal{E}' \subseteq \mathcal{E}$ and $\mathsf{X} = \mathsf{X}_1$ and $\mathsf{Y} = \mathsf{X}_n$; the *length* of $\mathcal{E}'$ is the number of arcs in $\mathcal{E}'$.

A *subgraph* G' of a directed graph $\mathsf{G} = (\mathcal{N}, \mathcal{E})$ induced by a set of nodes $\mathcal{N}' \subseteq \mathcal{N}$ is the directed graph $\mathsf{G} = (\mathcal{N}', \mathcal{E}')$, where $\mathcal{E}' = \{(\mathsf{X}, \mathsf{Y}) | \{\mathsf{X}, \mathsf{Y}\} \subseteq \mathcal{N}' \text{ and } (\mathsf{X}, \mathsf{Y}) \in \mathcal{E}\}$. A *tree* is a directed graph, in which there is exactly one node (termed the *root*) from which there exists exactly one directed path to every other node. A node X is said to *precede* another node Y in a tree T if there exists a directed path from X to Y (analogously, Y is said to *succeed* X in T). The *subtree* of $\mathsf{T} = (\mathcal{N}, \mathcal{E})$ having a node X as root (denoted $\mathsf{T}(\mathsf{X}) = (\mathcal{N}_{\mathsf{T}(\mathsf{X})}, \mathcal{E}_{\mathsf{T}(\mathsf{X})})$) is the subgraph of T induced by X and the successors of X in T.

A *decision tree* T is a tree whose nodes $\mathcal{N}$ are partitioned into three disjoint subsets: *decision nodes* $\mathcal{N}_\mathsf{D}$ (drawn as rectangles), *chance nodes* $\mathcal{N}_\mathsf{C}$ (drawn as circles) and *value nodes* (or *consequences*) $\mathcal{C}$, where the nodes $\mathsf{X} \in \mathcal{C}$ have no children ($\mathcal{C}$ are the *leaves* in the decision tree). For notational convenience, we shall in the remainder of this paper assume that the root of the decision tree is a decision node.

An *immediate decision successor* for a decision node D is a succeeding decision node D', satisfying that all nodes on the path from D to D' are chance nodes. The set of decision nodes which constitute the immediate decision successors of a decision node D is denoted $\mathcal{D}_\mathsf{D}$. The *chance node successors* of a decision node D are the chance nodes which succeed

D but do not succeed any immediate decision successor of D. The set of chance node successors of a decision node D is denoted $\mathcal{C}_{\mathsf{D}}$. The *past* of a node X (denoted past(X)) is the path from the root to X. Finally, we denote by sp_{X} the set of all children of X in T. Semantically, the arcs emanating from a chance node X define *events* such that one and only one event may be achieved at X. Analogously, the arcs emanating from a decision node X define the *decision options* that are available to the DM at X.

Definition 1. A *strategy* in a decision tree $\mathsf{T} = (\mathcal{N}, \mathcal{E})$ is a set of arcs $\Delta = \{(\mathsf{Y}, \mathsf{Z}) | \mathsf{Y} \in \mathcal{N}^{\Delta}_{\mathsf{D}}, \mathsf{Z} \in \mathcal{N}^{\Delta}\} \subseteq \mathcal{E}$, where $\mathcal{N}^{\Delta} \subseteq \mathcal{N}$ is a set of nodes containing:

i) The root node.

ii) Exactly one child of each $\mathsf{X} \in \mathcal{N}^{\Delta}_{\mathsf{D}} = \mathcal{N}_{\mathsf{D}} \cap \mathcal{N}^{\Delta}$.

iii) All children of each $\mathsf{X} \in \mathcal{N}^{\Delta}_{\mathsf{C}} = \mathcal{N}_{\mathsf{C}} \cap \mathcal{N}^{\Delta}$.

The set of all strategies in T is denoted $\bar{\Delta}^{\mathsf{T}}$.

The restriction of a strategy to a subtree $\mathsf{T}(\mathsf{X})$, which is in fact a strategy in $\mathsf{T}(\mathsf{X})$, is called a *substrategy*. Notice that the definition of a strategy is not as the definition in game theory, where a strategy is a set of arcs consisting of exactly one arc for each $\mathsf{X} \in \mathcal{N}_{\mathsf{D}}$.

1.1.2 Decision making under risk

Decision making under risk deals with situations where events are endowed with probabilities: for every chance node X, we assume that a probability distribution $\{\mathsf{P}((\mathsf{X}, \mathsf{Y}) | \text{past}(\mathsf{X})), \mathsf{Y} \in \mathsf{sp}_{\mathsf{X}}\}$ specifies the probabilities of the achievable events, conditioned on the events and decisions belonging to the path from the root to X, past(X). For any substrategy with root X, these probabilities are sufficient (by repeated use of the theorem of compound probabilities) for determining the conditional probability of reaching a given leaf (consequence) from X given past(X). Thus, any (sub)strategy generates a probability distribution on the consequence set.

1.2 Single stage decision making

Consider a decision tree with a unique decision node, and where each decision option is associated with a (single stage) *lottery*, i.e., the discrete probability distribution it generates on the consequence set:

$$L = (c_1, p_1; ...; c_i, p_i; ...; c_n, p_n),$$

where $p_i \geq 0$ $(i = 1, ..., n)$ denotes the probability of obtaining the consequence c_i and $\sum_{i=1}^{n} p_i = 1$ (graphically, a lottery L can be represented as in Figure 1). With $L' = (c_1, p_1'; ...; c_i, p_i'; ...; c_n, p_n')$, $qL + (1-q)L'$ denotes a lottery, a *mixture* (convex linear combination) of L and L', giving consequence c_i with probability $qp_i + (1-q)p_i'$.

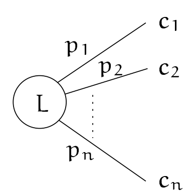

Figure 1: A graphical representation of the single stage lottery $L = (c_1, p_1; c_2, p_2; \ldots; c_n, p_n)$.

Models that have been proposed for single stage decision situations usually assume that the only relevant aspect of an event is its probability (*events anonymity*). Hence, the DM's preferences depend only on the lotteries generated by the decisions, and we can therefore directly define the DM's preferences by a weak order $\succsim$ on the lottery set: as usual, $L \succsim L'$ is read as "L is preferred or indifferent to L'". The derived relations, $\succ$, $\precsim$, $\prec$, $\sim$ have their standard meaning, and we shall assume that this preference ordering can be expressed by a real function V (a *preference function*) satisfying $V(L) \geq V(L') \Leftrightarrow L \succsim L'$, hence also $V(L) > V(L') \Leftrightarrow L \succ L'$ and $V(L) = V(L') \Leftrightarrow L \sim L'$.

1.2.1 Expected utility

Under EU theory, the preference function $V(\cdot) = EU(\cdot)$ on a lottery L takes on the form:

$$EU(L)_u = EU(c_1, p_1; ...; c_i, p_i; ...; c_n, p_n.)_u = \sum_{i=1}^{n} u(c_i)p_i, \tag{1}$$

where the function $u(\cdot)$ is a personal characteristic of the DM, his von Neumann-Morgenstern (vNM) utility function. The identification of a constant lottery with its unique consequence makes $u(\cdot)$ a utility representing the DM's preferences under certainty. Moreover, since $u(\cdot)$, like $EU(\cdot)$, is *cardinal* (unique up to affine transformations), its shape can receive an interpretation. In particular, when the consequence set $\mathcal{C}$ is a convex subset of a linear space, a concave $u(\cdot)$ characterizes *risk aversion*.

EU theory can be justified as an axiomatic model. The cornerstone of the axiom system is the following property which is clearly implied by the linearity of the preference function w.r.t. the probabilities.

Axiom 1 (Independence axiom). For all lotteries L, L', and L^* and for all q in $(0, 1]$, the lottery L is preferred to the lottery L' if and only if the mixture $qL + (1 - q)L^*$ is preferred to the mixture $qL' + (1 - q)L^*$.

Although the model allows lotteries to be ranked differently by different DMs, it is generally agreed that consistency with the following partial ordering is a natural rationality requirement.

Definition 2 (First order stochastic dominance: FSD). The lottery L *stochastically dominates lottery* L' *in the first order* when:

- $\sum_{\{i:u(c_i)\leq k\}} p_i \leq \sum_{\{i:u(c_i)\leq k\}} p'_i$, for all $k \in \mathfrak{R}$ and
- $\sum_{\{i:u(c_i)\leq k\}} p_i < \sum_{\{i:u(c_i)\leq k\}} p'_i$, for some $k \in \mathfrak{R}$.

This requirement imposes the same restrictions for different DMs having the same preferences under certainty, i.e., with $u(\cdot)$ in $\mathcal{U} = \{u(\cdot) : \mathcal{C} \to \mathfrak{R} | u(c) \geq u(c') \Leftrightarrow c \succsim c'\}$.

EU criteria are necessarily consistent with FSD. Actually, EU theory provides a characterization of FSD, as stated in the following theorem due to (Hadar and Russell 1969) and (Rothschild and Stiglitz 1970), which is exploited later in this paper.

Theorem 1. Let $\mathcal{U}$ be the set of utility functions representing the DM's ordering on the consequence set $\mathcal{C}$, and let L and L' be two lotteries. Then the following statements are equivalent:

i) L FSD L'.

ii) $\forall \mathsf{u} \in \mathcal{U} : \mathsf{EU}(\mathsf{L})_\mathsf{u} \geq \mathsf{EU}(\mathsf{L}')_\mathsf{u}$ and $\exists \mathsf{u} \in \mathcal{U} : \mathsf{EU}(\mathsf{L})_\mathsf{u} > \mathsf{EU}(\mathsf{L}')_\mathsf{u}$.

The theorem also implies that if there exists a utility function $\mathsf{u} \in \mathcal{U}$ s.t. $\mathsf{EU}(\mathsf{L})_\mathsf{u} > \mathsf{EU}(\mathsf{L}')_\mathsf{u}$, then L is not first order stochastically dominated by L'; this property also applies if for all utility functions $\mathsf{u} \in \mathcal{U}$ it holds that $\mathsf{EU}(\mathsf{L})_\mathsf{u} \geq \mathsf{EU}(\mathsf{L}')_\mathsf{u}$. Note that this utility function need not have any particular relation with the DM's preferences besides representing the same ordering of the consequences under certainty.

1.2.2 A non-EU model: Rank Dependent Utility (RDU)

Motivated by experimental observations of phenomena inconsistent with EU theory, in particular the *Allais Paradox* (Allais 1979) and (Kahneman and Tversky 1979), several authors have suggested an alternative to EU theory, now most often called *Rank Dependent Utility* (RDU) theory (see e.g. (Allais 1988) and (Yaari 1987)). This model was, however, first presented and axiomatized by (Quiggin 1982) under the name of *Anticipated Utility* theory; the preference function in the RDU model has the form of a *Choquet integral* (Choquet 1953).

In the RDU model, there are two parameters: i) a utility function $\mathsf{u}(\cdot)$ which, as the vNM utility function of EU theory, is cardinal and represents the DM's preferences under certainty, and ii) a *probability transformation* (or *weighting*) function $\mathsf{q}(\cdot)$, which is strictly increasing from $[0, 1]$ onto itself. The utility $V(L)_{u,q}$ of a lottery L is determined as follows. First, the components of L are re-indexed in increasing order of their utility:

$$L = (c_1, p_1; ...; c_i, p_i; ...; c_n, p_n) \text{ with } u(c_1) \leq ... \leq u(c_i) \leq ... \leq u(c_n).$$

Then, its utility is evaluated as:

$$V(L)_{u,q} = \sum_{i=1}^{n} u(c_i) \left[q\left(\sum_{j=i}^{n} p_j\right) - q\left(\sum_{j=i+1}^{n} p_j\right)\right]$$

or, equivalently as

$$V(L)_{u,q} = u(c_1) + \sum_{i=2}^{n} [u(c_i) - u(c_{i-1})]\, q\left(\sum_{j=i}^{n} p_j\right).$$

The crucial idea underlying the RDU model is that the utility of each consequence is given a variable weight, which depends on both its probability and on its ranking w.r.t. the other consequences of the lottery. The second expression suggests that the DM bases his evaluation on the probabilities of achieving at least such or such utility level. Note that RDU theory no longer satisfies Axiom 1.

From the second expression of $V(L)_{u,q}$ it is clear that, in RDU theory, preferences are consistent with FSD: the function $q(\cdot)$ is strictly increasing. It can be shown that consistency with *strong risk aversion* also holds if and only if $q(\cdot)$ is convex and $u(\cdot)$ is concave (Chew, Karni, and Safra 1987). Sufficient conditions for the preferences w.r.t. *weak risk aversion* ($EU(L) \succsim L$) are also known (Chateauneuf and Cohen 1984); these conditions also involve both $q(\cdot)$ and $u(\cdot)$ (see e.g. (Currim and Sarin 1989) and (Wu and Gonzalez 1996)). A frequently observed shape of $q(\cdot)$ is displayed in Figure 2.

Henceforth, we will focus on RDU when considering non-EU theory; however, we conjecture that the results presented in this paper also apply to other models.

1.3 Dynamic Decision Making

In the simple type of decision situation considered above, preferences can be inferred from behavior in a straightforward way: when choosing d in decision set sp_D, the DM reveals

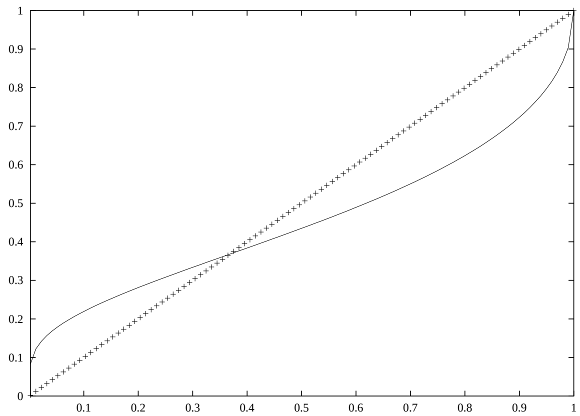

Figure 2: An instance of the probability transformation function $q(p) = \exp(-(-\ln(p))^{\omega}))$ with $\omega = 0.5$.

that he (weakly) prefers lottery L, generated by d, to any other lottery L' generated by some d' in sp_D.

Things are, however, not always that simple. For instance, consider a husband who, when dining alone, always eats in restaurant R_1, and, when dining with his wife, eats in restaurant R_2. Which is his preferred restaurant? A possible answer is that decisions that only concern him reveal his true intrinsic preferences, whereas decisions that also involve his wife are the result of a compromise between his and her intrinsic preferences. However, the choice of R_2 in some context (here a collective choice) and the claim that $R_1 \succ R_2$ are not in contradiction with the revealed preference paradigm: one only needs to realize that observations are sometimes best explained by a model in which the preferences have complex indirect links with the choices.

The dynamic decision model described below introduces intrinsic preferences which interact and determine the choice of a strategy. *The intrinsic preferences are by definition those revealed*, in the classical sense, *by the observation of choices in one stage decision situations*; they are assumed not to depend on the particular situation considered, and are therefore characterized by a single criterion defined on lotteries, e.g. $EU(L)_u$ or $V(L)_{u,q}$. The domain of definition of these intrinsic preferences can be extended by making the following assumption:

1.3.1 Reduction of compound lotteries (RCL)

With $L = (c_1, p_1; ...; c_i, p_i; ...; c_n, p_n)$ and $L' = (c_1, p'_1; ...; c_i, p'_i; ...; c_n, p'_n)$, $L^* = (L, q; L', 1-q)$ denotes a *compound lottery* which represents a two-step resolution of uncertainty, whereas $L'' = qL + (1-q)L'$ denotes a single stage lottery formed by the mixture of L and L' (see Figure 3).

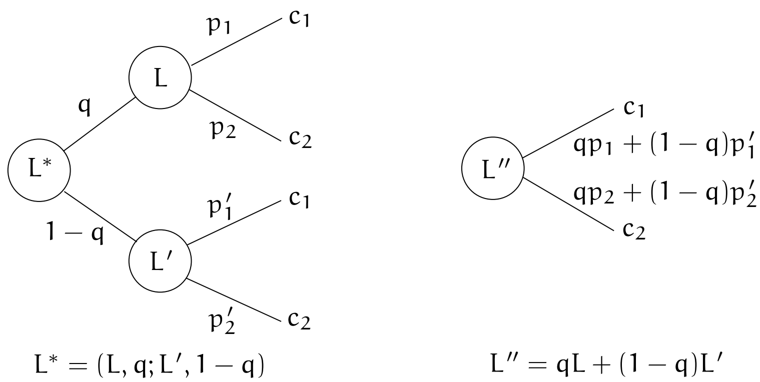

Figure 3: The figure illustrates a compound lottery and the lottery formed by the mixture of L and L'.

The RCL assumption implies that the intrinsic preferences of the DM make him indifferent between decisions respectively offering prospects of a compound lottery (e.g. L^*) and of the associated mixture (e.g. L''), i.e., the process of resolution of the uncertainty does not matter. Note that this principle can clearly be extended, by recursion, to processes involving n-stage lotteries.

A case against RCL has been made by some researchers, chiefly (Segal 1990), who pointed out that its non-observance might be responsible for the frequently observed deviation from EU theory, known as the common ratio effect. However, (Cubbit, Starmer, and Sugden 1998) in a recent experimental study have come to the conclusion that this effect could not be principally attributed to violations of RCL.

Making the RCL assumption we will not distinguish between a compound lottery and its probabilistically equivalent single stage lottery that is, we implicitly refer to single stage lotteries when making subsequent references to compound lotteries.

The preference function of EU theory, together with RCL, implies the notion of an *optimal strategy*: from RCL we have that any strategy $\Delta \in \bar{\Delta}^{\mathsf{T}}$ in a given decision tree T can be identified with a single stage lottery L_Δ; thus, we say that a strategy Δ is an optimal strategy w.r.t. the utility function u if and only if:

$$\mathsf{EU}(\mathsf{L}_\Delta)_u \geq \mathsf{EU}(\mathsf{L}_{\Delta'})_u, \forall \Delta' \in \bar{\Delta}^{\mathsf{T}} \backslash \{\Delta\}.$$

We now turn to non-EU theories, where the comparison of strategies is a more complex problem.

1.3.2 Dynamic consistency

Consider a decision problem represented by a decision tree, and suppose that at each decision node the DM has a preference ordering on the substrategies that are feasible at that node. We say that the DM is *dynamically consistent* if substrategies of optimal strategies are optimal substrategies. More precisely, assume that Δ_0 is an optimal strategy of the subtree rooted at decision node D_0 given the preference ordering at D_0. If decision node D_1 is reachable from D_0 with Δ_0, then substrategy Δ_1, the restriction of Δ_0 to the subtree rooted at D_1, should itself be optimal for the preference ordering at that node.

A DM who is an EU maximizer at each decision node (always with the same vNM utility function) is automatically dynamically consistent. This is a straightforward implication of the independence axiom. Consider now a DM who is an RDU maximizer at each decision node (always with the same vNM-like utility function and probability weighting function). The following example shows that this DM can be dynamically inconsistent; this problem was initially addressed by (Schick 1986).

Example 1. Consider the decision tree depicted in Figure 4 and assume that the DM's preferences are expressed by an RDU preference function, where the probability weighting function is given by:

$$q(p) = \exp(-(-\ln(p))^{0.5})$$

and where the utility function is linear.

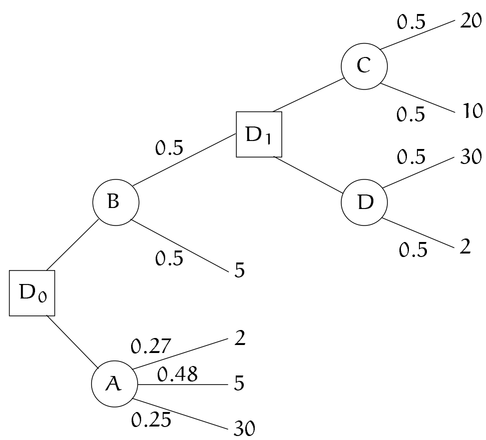

Figure 4: A decision tree representation of a dynamic decision problem with two decisions. The values associated with the leaf nodes correspond to the utilities of these consequences.

At D_0 the possible strategies are $\{(D_0, A)\}$, $\{(D_0, B), (D_1, C)\}$ and $\{(D_0, B), (D_1, D)\}$, where:

$$\begin{aligned}
V(\{(D_0, A)\}) &= 2 + (5-2) \cdot q(0.48 + 0.25) + (30-5) \cdot q(0.25) = 11.41 \\
V(\{(D_0, B), (D_1, C)\}) &= 5 + (10-5) \cdot q(0.25 + 0.25) + (20-10) \cdot q(0.25) = 10.26 \\
V(\{(D_0, B), (D_1, D)\}) &= 2 + (5-2) \cdot q(0.5 + 0.25) + (30-5) \cdot q(0.25) = 11.46
\end{aligned}$$

Thus, at D_0 the DM would prefer $\{(D_0, B), (D_1, D)\}$ over $\{(D_0, B), (D_1, C)\}$ and $\{(D_0, A)\}$ however, at D_1 we have:

$$\begin{aligned}
V(\{(D_1, C)\}) &= 10 + (20-10) \cdot q(0.5) = 14.35 \\
V(\{(D_1, D)\}) &= 2 + (30-2) \cdot q(0.5) = 14.18
\end{aligned}$$

That is, at D_1 the DM would prefer $\{(D_1, C)\}$ over $\{(D_1, D)\}$ meaning that the DM is not dynamically consistent. $\square$

1.3.3 Sophisticated behavior and consequentialism

Dynamic consistency is a relation between preference orderings intervening at various stages in a given decision problem. Let us now examine its implications concerning choice. At any given decision node, the DM ranks substrategies according to his preference ordering at that node but he can only control the immediate decision. It is generally admitted that, aware of that fact, a DM should at each moment anticipate his future choices and therefore value each immediate decision by the value of the lottery generated by the anticipated ensuing strategy. This is known as *sophisticated behavior*.

Going back to the preceding example, we see that a sophisticated behavior makes our dynamically inconsistent RDU maximizer choose an FSD-dominated strategy. This is generally considered as irrational.

Example 2. For the decision problem illustrated in Figure 4, the sophisticated strategy is $\{(D_0, A)\}$ which is FSD-dominated by $\{(D_0, B), (D_1, D)\}$. This can easily be seen by considering the corresponding cumulative distribution functions. □

Such an irrational choice cannot be observed with an EU maximizer. Consider then a dynamically consistent sophisticated DM whose preferences at the root node are expressible by a RDU criterion. What about his preferences at other decision nodes? Since the optimal strategy depends on the whole tree, so does its restriction to any subtree. Thus preferences at a decision node depend on elements outside the subtree rooted at that node. The same example as above shows it immediately.

Example 3. Assume that the DM is dynamically consistent, i.e., at D_0 the strategy $\{(D_0, B), (D_1, D)\}$ is preferred, which implies that $\{(D_1, D)\}$ is preferred at D_1. However, if the utility 5, in the subtree defined by (D_0, B), is replaced by the utility 9.5 we get:

$$\begin{aligned} V(\{(D_0, B), (D_1, C)\}) &= 9.5 + (10 - 9.5) \cdot q(0.25 + 0.25) + (20 - 10) \cdot q(0.25) = 12.80 \\ V(\{(D_0, B), (D_1, D)\}) &= 2 + (9.5 - 2) \cdot q(0.5 + 0.25) + (30 - 9.5) \cdot q(0.25) = 12.70 \end{aligned}$$

That is, in this modified decision problem the DM would prefer $\{(\mathsf{D}_0, \mathsf{B}), (\mathsf{D}_1, \mathsf{C})\}$ at D_0 and therefore $\{(\mathsf{D}_1, \mathsf{C})\}$ at D_1. Hence, the DM's choice at D_1 depends on a value outside the subtree having D_1 as root. $\square$

Consequentialism precisely forbids such an influence on choice. Thus the adaptation of RDU theory to a dynamic setting forces us to face the following dilemma when assuming RCL and sophisticated behavior: either accept possibly irrational behavior (choice of a dominated strategy) or renounce consequentialism.

In his seminal paper, (Hammond 1988) demonstrates the essential role of consequentialism in the justification of EU theory. A contrario, as first noted by (Machina 1989), renouncing consequentialism leaves the door open for other models. (Machina 1989) argues that it is inappropriate to impose consequentialism on non-EU behavior since the very notion of a non-EU criterion implies a type of non-separability which contradicts the consequentialist assumption. He also provides some examples of situations, where non-consequentialist behavior can indeed be deduced by underlying preferences which are influenced in a plausible way by the context. Thus non-consequentialist behavior can sometimes be justified by non-consequentialist preferences. Non-consequentialism can also appear through the modification of originally consequentialist preferences (see (McClennen 1990)). We shall develop and advocate this alternative approach.

1.3.4 Resolute choice

Resolute choice is a type of behavior introduced and discussed in (McClennen 1990): the resolute DM initially commits himself to a strategy, i.e., initially chooses a strategy and never deviates from it later. Conceptually, in this model each decision node is associated with a *Self*, who represents the DM at the time and state when the decision is made. The Selves, who at the outset may have inconsistent preferences, manage to depart from them and achieve a consensus. A restrictive interpretation of resolute choice is that the strategy to which the DM commits himself is necessarily the strategy judged optimal, from the very beginning, by the first Self. It has the advantage to straightforwardly endow the resolute strategy with all the qualities ensured by the first Self's criterion.

However the ability of the DM to commit himself, and especially to commit himself to this particular strategy, is a disputable matter. This makes it desirable to explicitly model the underlying process that leads to the determination of a consensual strategy (or fails to reach a consensus); notice that the model, being intended to be prescriptive rather than descriptive, does not necessarily mimic the underlying psychological process of the DM.

With this objective, (Jaffray 1999) elaborates on resolute choice (in the situation of decision making under uncertainty) and suggests a model which can be interpreted along the following lines: each Self is associated with a decision node and endowed with a consequentialist preference ordering on the substrategies at that node. Together, these preferences make them dynamically inconsistent. However the Selves are willing to cooperate. This assumption is quite the reverse of the prevailing one, which is to consider the Selves as non-cooperative players of a game (Karni and Safra 1989). As a justification, it can be put forward that the Selves are all part of the same mind and cannot hide intentions from each other; this rules out bluff, betrayal and manipulation and should therefore facilitate cooperation. The overall goal of the Selves is assumed to be the selection of a non-dominated strategy. For the Self associated with a decision node D, the selection of a non-dominated substrategy in the subtree rooted at D is an obvious goal. For the coalition of all the Selves associated with the immediate decision successors for D, one can argue that it is aware that one of the subcoalitions will in fact be selected by the decision option chosen at D, and therefore wants to avoid the selection of a subcoalition which offers a collectively disadvantageous prospect, i.e., a dominated substrategy. This common goal is the mutual incentive for cooperation between a Self and its immediate successors; since this cooperation requires a form of resolution, we shall also use the term *resolute choice* for this type of behavior.

Finally, analogously to Jaffray (1999) we will assume that the willingness of the Selves to cooperate is characterized by the maximum amount of value (denoted θ) they are ready to give up, when deviating from their best achievable choices in order to reach a satisfactory consensus. The fact that the Selves are clones makes it possible to assume a common value scale and a common acceptable deviation. A rough indication on the value of θ can clearly be obtained by presenting the DM with simple two-step decision trees. However, the design of the specific procedure is out of the scope of this paper. Measuring the flexibility of the Selves by θ, i.e., by a constant utility difference in a particular representation, makes it dependent on that representation; this should only be considered as a convenient simplifying assumption. Notice that an interactive version of the algorithm, presented in Section 3, would neither require an explicit knowledge of θ nor even the assumption that it exists. Notice, moreover, that the flexibility of the Selves might also be described using different models. For example, a Self might be willing to accept any of e.g. the three strategies with highest value.

The model by (Jaffray 1999) was originally presented for the situation of decision making under uncertainty, however, the general principles also apply to decision making under risk and in the remainder of this paper we shall consider the model (as described above) in this context.

2 Investigating strategies

A crucial aspect of the model is the identification (and subsequent proposal) of a non-dominated strategy that is acceptable to all the Selves. This means that in a given decision tree T, any proposed strategy $\Delta \in \bar{\Delta}^{\mathsf{T}}$ (represented by a lottery L_Δ) must satisfy:

$$\begin{aligned} \forall \Delta' \in \bar{\Delta}^{\mathsf{T}} \backslash \{\Delta\}, \text{ either } \exists u \in \mathcal{U} \text{ s.t. } \mathsf{EU}(\mathsf{L}_\Delta)_u > \mathsf{EU}(\mathsf{L}_{\Delta'})_u \\ \text{or } \forall u \in \mathcal{U} : \mathsf{EU}(\mathsf{L}_\Delta)_u \geq \mathsf{EU}(\mathsf{L}_{\Delta'})_u, \end{aligned}$$

where $\mathcal{U}$ is defined as in Theorem 1. If both of the expressions hold for all $\Delta' \in \bar{\Delta}^{\mathsf{T}} \backslash \{\Delta\}$, then Δ stochastically dominates all other strategies at the first order.

In order to investigate whether or not a strategy Δ satisfies one of the expressions above, we would in principle have to compare Δ with all the other strategies. To overcome this computational difficulty we shall instead look for strategies for which there exists a utility function s.t. the strategy in question is strictly EU-optimal:

$$\exists u \in \mathcal{U} \text{ s.t. } \mathsf{EU}(\mathsf{L}_\Delta)_u > \mathsf{EU}(\mathsf{L}_{\Delta'})_u, \forall \Delta' \in \bar{\Delta}^{\mathsf{T}} \setminus \{\Delta\}. \tag{2}$$

To say it in another way, if there exists a utility function s.t. Δ is an EU-optimal strategy, then Δ is non-dominated. The converse is, however, not necessarily true, i.e., the approach is sound but not complete since we cannot infer that Δ is dominated if no such utility function exists. Note again that this utility function need not have any particular relation with the DM's preferences besides representing the same ordering of the consequences under certainty.

2.1 Restating the problem

In this section we reduce the problem of determining whether a strategy Δ is non-dominated to the problem of finding a solution to a particular system of linear inequalities, i.e., the system of inequalities has a solution if and only if there exists a utility function resulting in Δ being a strictly optimal strategy (see Expression 2).

The linear inequalities are found by initially creating the variables which are subject to a constraint. Given a decision tree T these variables are constructed as follows:

- Create a variable y_c for each distinct consequence c in T .
- Create a variable y_{D} for each decision node D in T.

From these variables the system of inequalities is constructed incrementally by considering the consequences and the decision nodes separately.

The constraints on the consequences correspond to their preference ordering which can be read directly from the (original) utility function. That is, by assuming that the preference ordering over the consequences is given by their indexes, we require:

$$\forall 2 \leq i \leq m : y_{c_i} \geq y_{c_{i-1}} + \alpha, \tag{3}$$

where m is the number of distinct consequences and $\alpha > 0$ is chosen arbitrarily. Since a consequence can always be replaced by a privileged element of its indifference class, we assume, without loss of generality, that two different consequences c_i and c_j satisfy either $c_i \prec c_j$ or $c_j \prec c_i$. This set of inequalities ensures that any solution encodes a utility function $u \in \mathcal{U}$, defined by $u(c_i) = y_{c_i}$, which respects the preference ordering over all the consequences.

The remaining constraints are associated with the decision nodes and can be identified (locally) in the decision tree.

Assume that D is a decision node, and choose an arbitrary $\epsilon > 0$. Then, for each arc (D, X) which is not part of Δ, we require, if X is a chance node:

$$y_D \geq \sum_{l \in \mathcal{L}} p_l y_l + \epsilon, \tag{4}$$

where $\mathcal{L}$ is the set of all directed paths from X to either its immediate decision successors or to value nodes; p_l is the product of the probabilities associated with the path l and y_l is the variable associated with the "end" node of path l. In case X is a decision node or a value node, we simply have $y_D \geq y_X + \epsilon$. The constant $\epsilon > 0$ ensures strict optimality; note that α and ϵ need not be equal.

If (D, X) is part of the strategy Δ and if X is a chance node, we require:

$$y_D \leq \sum_{l \in \mathcal{L}} p_l y_l. \tag{5}$$

If X is a decision node or a value node we have $y_D \leq y_X$.

Example 4. Consider the case where we have to choose between the two lotteries $L_1 = (c_2, 0.4; c_3, 0.2; c_4, 0.4)$ and $L_2 = (c_1, 0.7; c_5, 0.1, c_6, 0.2)$, and assume that the preferences over the consequences are revealed by their index (see Figure 5).

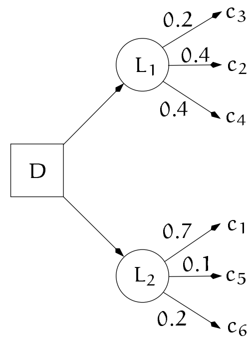

Figure 5: A decision tree with one decision node representing the choice between the two lotteries L_1 and L_2; the preference ordering over the consequences correspond to their index.

Consider the strategy $\Delta = \{(D, L_2)\}$, and assume that we want to determine whether there exists a utility function that respects the preference ordering over the consequences and yields Δ as an optimal strategy.

For e.g. c_4 we construct the inequality (the inequalities for the other consequences are constructed similarly):

$$y_{c_4} \geq y_{c_3} + \alpha.$$

For the decision node D we have:

$$\begin{aligned} y_D &\geq 0.2 \cdot y_{c_3} + 0.4 \cdot y_{c_2} + 0.4 \cdot y_{c_4} + \epsilon \\ y_D &\leq 0.7 \cdot y_{c_1} + 0.1 \cdot y_{c_5} + 0.2 \cdot y_{c_6} \end{aligned}$$

With $\alpha = 1$ and $\epsilon = 0.1$, an admissible solution to these inequalities encode the utility function $u(c_1) = 0$, $u(c_2) = 1$, $u(c_3) = 2$, $u(c_4) = 3$, $u(c_5) = 4$ and $u(c_6) = 8.5$. This utility function yields $\Delta = \{(D, L_2)\}$ as an optimal strategy ($EU(L_1)_u = 2$ and $EU(L_2)_u = 2.1$). □

Property 1. Let T be a decision tree with consequences $\mathcal{C}$ and let Δ be a strategy in T. Then there exists a utility function $u \in \mathcal{U}$ s.t. Δ is a strictly optimal strategy if and only if the corresponding system of linear inequalities has a solution for some $\alpha > 0$ and $\epsilon > 0$.

Proof. Follows from the construction of the set of inequalities. □

Property 2. Given a decision tree T and a strategy Δ, if the corresponding system of linear inequalities has a solution for some $\alpha > 0$ and $\epsilon > 0$ then the system of linear inequalities has a solution $\forall \alpha' > 0$ and $\forall \epsilon' > 0$.

Proof. The proof is straightforward by observing that a utility function is unique up to an affine transformation, and that the expected utility is linear in the utilities. □

Finding a solution for a system of inequalities can be done with "brute force" by considering all the inequalities simultaneously and then use stage one of the *simplex algorithm*; similarly, we may construct an artificial *objective function*, e.g. $c = 0$, and use an *interior point* algorithm. Notice that the structure of the inequalities does not directly support decomposition schemes like (Dantzig and Wolfe 1960), however, it should be noted that due to the relatively large size of manageable systems the brute force approach might be sufficient for some decision problems.

Alternatively to the brute force approach we can exploit the structure of the decision tree e.g. a solution may be found by recursively solving collections of smaller subsets of inequalities; this is a subject for further research.

3 Evaluation

In this section we present an algorithm for solving decision problems w.r.t. non-EU criteria such as RDU. The underlying idea of the algorithm is to assume consequentialist intrinsic preferences at each decision node and to accept non-consequentialist behavior,not induced solely by local preference optimization, as described in Section 1.3.3.

The consequentialist preference assumption is reflected in the use of backward induction, whereas the assumption about non-consequentialist behavior follows from a Self being willing to accept any of a set of strategies (whereas one strategy is enough when dealing with expected utility); the actual selection is made by a preceding Self and therefore depends on data outside the consequentially relevant subtree.

3.1 Algorithm

The algorithm works itself backwards from the leaves towards the root of the decision tree T. When a decision node (say D) is visited, a set of candidate (sub)-strategies Δ_{D}^{+} in $\mathsf{T}(\mathsf{D})$ is determined; since we work from the leaves towards the root, Δ_{D}^{+} is formed by taking the Cartesian product of the strategies acceptable by the Selves that are associated with the immediate decision successors of D in T. For each strategy in Δ_{D}^{+} we then determine whether or not it passes the non-dominance test; let $\Delta_{\mathsf{D}} \subseteq \Delta_{\mathsf{D}}^{+}$ be the set of strategies which are proven non-dominated. For each strategy Δ in Δ_{D} we calculate $\mathsf{V}(\Delta)$ and finally we remove all strategies with a value inferior to the best achievable one minus the amount (θ) the Self is willing to give up in order to achieve the goal of identifying a non-dominated strategy (see Section 1.3.4).

Obviously, if the value of θ is "sufficiently large" then no non-dominated strategies are removed, and at each decision node we may therefore get a set of strategies that is intractable to evaluate. In order to overcome this computational difficulty we impose an upper bound (denoted κ) on the number of non-dominated strategies acceptable by a Self. That is, if each Self is willing to accept at most κ strategies, then at any decision node D we only need to evaluate at most $\kappa^{\omega}|\mathsf{sp}_{\mathsf{D}}|$ strategies, where ω is the maximum number of decision nodes that immediately succeed D w.r.t. some decision option for D. Notice, that κ is only introduced to ensure the computational feasibility (as opposed to θ which is part of the underlying model), however, as the set of strategies still grows exponentially (in the number of immediate decision successors), one might be forced to only evaluate a subset of these strategies and then keep the κ best.

Finally, it should be emphasized that for each strategy in Δ_{D}^{+} we make sure that it is non-dominated before its value is calculated. This reflects the overall goal of finding a non-dominated strategy that is acceptable by all the Selves. Furthermore, this prevents a dominated (sub-)strategy to be included among the κ best strategies and thereby take the place of a non-dominated strategy. This "early" identification of dominated strategies is justified by the following property.

Property 3. Let T be a decision tree and let D be a decision node in T. If there does not exist a utility function $u \in \mathcal{U}$ s.t. Δ is a unique EU-optimal strategy w.r.t. u in $\mathsf{T}(\mathsf{D})$, then for any strategy Δ' that has Δ as a substrategy, there does not exist a utility function $u' \in \mathcal{U}$ s.t. Δ' is a unique EU-optimal strategy w.r.t. u'.

Proof. Follows immediately. □

Before presenting the algorithm we need the following definition.

Definition 3. Let $\mathsf{T}(\mathsf{D})$ be a decision tree with root D, and let $\mathcal{D}_{(\mathsf{D},\mathsf{X})}$ be the set of immediate decision successors for D that can be reached by a directed path which includes (D, X). The set of *candidate strategies* for D is then given by:

$$\Delta^{+}_{\mathsf{D}} = \bigcup_{\mathsf{X} \in \mathsf{sp}^{*}_{\mathsf{D}}} \left\{ \{(\mathsf{D}, \mathsf{X})\} \cup \mathsf{S} \mid \mathsf{S} \in \prod_{\mathsf{D}' \in \mathcal{D}_{(\mathsf{D},\mathsf{X})}} \Delta_{\mathsf{D}'} \right\},$$

where $\mathsf{sp}^{*}_{\mathsf{D}} \subseteq \mathsf{sp}_{\mathsf{D}}$ s.t. for all $\mathsf{X} \in \mathsf{sp}^{*}_{\mathsf{D}}$ it holds that every $\mathsf{D}' \in \mathcal{D}_{(\mathsf{D},\mathsf{X})}$ is associated with a non-empty set of strategies.

The algorithm outlined above is formalized as follows.

Algorithm 1. *Let θ denote a Self's willingness to cooperate, and let κ be the maximum number of strategies acceptable by a Self. To evaluate a decision tree T with root D, do:*

1. *Initialize T by associating the probability 1 with each consequence (leaf node) of T.*

2. *Invoke* Evaluation *(Algorithm 2) on T w.r.t. D, θ and κ, and let Δ_{D} be the set of strategies returned.*

3. **If** $\Delta_{\mathsf{D}} \neq \emptyset$ **then**

   Return $\arg\max_{\Delta \in \Delta_{\mathsf{D}}} \mathsf{V}(\Delta)$.

   else

   No solution exists for T *w.r.t.* κ *and* θ.

Algorithm 2 (Evaluation). *Given a decision tree* T*, let* θ *denote a Self's willingness to cooperate, and let* κ *denote the maximum number of strategies acceptable by a Self. To determine the strategies acceptable by decision node* D*, do:*

1. *Invoke* Evaluation *on* D', $\forall \mathsf{D}' \in \mathcal{D}_{\mathsf{D}}$.

2. *Let* Δ_{D}^{+} *be the set of candidate strategies for* D *in* $\mathsf{T}(\mathsf{D})$ *(see Definition 3).*

3. **For** *each chance node* $\mathsf{A} \in \mathcal{C}_{\mathsf{D}}$

   For *each* $\mathsf{B} \in \mathsf{sp}_{\mathsf{A}}$

   Multiply $\mathsf{P}((\mathsf{A}, \mathsf{B}) | \mathit{past}(\mathsf{A}))$ *with the probabilities associated with the consequences in* $\mathsf{T}(\mathsf{B})$.[1]

4. **For** *each strategy* $\Delta \in \Delta_{\mathsf{D}}^{+}$

   If Δ *is non-dominated in* $\mathsf{T}(\mathsf{D})$ *(see Section 2)* **then**

   a) Calculate $\mathsf{V}(\Delta)$.

   b) Set $\Delta_{\mathsf{D}} := \Delta_{\mathsf{D}} \cup \{\Delta\}$ *and let* $\mathsf{V}_{\mathsf{D}}^{\max} = \max_{\Delta \in \Delta_{\mathsf{D}}} \mathsf{V}(\Delta)$.

5. **For** *each strategy* $\Delta \in \Delta_{\mathsf{D}}$

   If $\mathsf{V}(\Delta) < \mathsf{V}_{\mathsf{D}}^{\max} - \theta$ **then**

   Set $\Delta_{\mathsf{D}} := \Delta_{\mathsf{D}} \backslash \{\Delta\}$

6. **If** $|\Delta_{\mathsf{D}}| > \kappa$ **then**

   repeat

   Set $\Delta_{\mathsf{D}} := \Delta_{\mathsf{D}} \backslash \{\Delta\}$*, where* $\Delta = \arg\min_{\Delta \in \Delta_{\mathsf{D}}} \mathsf{V}(\Delta)$.

   Until $|\Delta_{\mathsf{D}}| = \kappa$

[1] Note that this is basically an application of the RCL assumption w.r.t. the subtree in question.

7. Return Δ_{D} *as the set of strategies acceptable by* D*.*

During the evaluation the elements in Δ_{D} should, for reasons of efficiency, be organized in a data structure which is ordered (or accessible) w.r.t. $\mathsf{V}(\Delta)$. Furthermore, it should be noticed that step *3* ensures that all probabilities needed to calculate $\mathsf{V}(\Delta)$ in step *4.a* are available. Hence, to calculate $\mathsf{V}(\Delta)$ we basically only need to order the possible consequences of Δ w.r.t. their preference ordering and to calculate the cumulative distribution over these consequences; the latter can be done in linear time given the ordering over the consequences.

Note that the algorithm is also applicable in case we work with the *coalesced decision tree* by (Olmsted 1983), i.e., the graph formed from the decision tree by collapsing identical subtrees; two subtrees are identical if they agree on both structure and numeric information. Constructing the system of linear inequalities from the coalesced decision tree would significantly reduce the number of linear inequalities, however, we shall restrict our attention to the traditional decision tree.

Finally, notice that if no non-dominated strategy is found w.r.t. some values for κ and θ we would in principle need to perform the entire evaluation again, with another (larger) value for κ. However, by employing a form of *backtracking* we can re-use many of the calculations initially performed. This can be realized by storing all (or some) of the sets of substrategies acceptable by the decision successors for D. Given these sets of acceptable substrategies, a new set of candidate strategies for D can easily be generated.

3.2 Complexity

In what follows we consider the complexity of the proposed algorithm under the assumption that we have a decision tree T with a set $\mathcal{C}$ of distinct consequences; backtracking will not be considered. Since the model is proposed as an alternative to EU-theory, it is appropriate to consider the relative complexity. That is:

$$\text{Relative complexity} = \frac{\mathsf{O}(\text{non-EU})}{\mathsf{O}(\text{EU})}$$

In both Algorithm 1 and the "average-out and fold-back" algorithm, we work ourselves backwards from the leaves towards the root. When a chance node A is visited in the:

- EU model, it is assigned the sum of the values associated with its children weighted by the probabilities of their outcomes, i.e., $\sum_{\mathsf{X} \in \mathsf{sp}_\mathsf{A}} \mathsf{P}((\mathsf{A}, \mathsf{X})|\text{past}(\mathsf{A})) \cdot \mathsf{EU}(\mathsf{X})$.
- non-EU model, the probability $\mathsf{P}((\mathsf{A}, \mathsf{X})|\text{past}(\mathsf{A}))$ is multiplied by the probability (initially 1) associated with the consequences in the subtree $\mathsf{T}(\mathsf{X})$, for each $\mathsf{X} \in \mathsf{sp}_\mathsf{A}$.

When a decision node D is visited in the:

- EU model, it is associated with the decision option leading to the child having the highest expected utility; the decision node is assigned this value as well.
- non-EU model, it is associated with a set of substrategies. This set is formed by i) considering the set of substrategies associated with each of its immediate decision successors, and ii) iteratively removing substrategies s.t. the resulting set has cardinality no larger than κ, and consists of the non-dominated substrategies with highest values.

Compared to the "average-out and fold-back algorithm", the most computationally intensive step is when determining whether or not a strategy is non-dominated: when visiting a decision node D we consider the decision tree $\mathsf{T}(\mathsf{D})$ rooted at D; obviously, if some decision node D' in $\mathsf{T}(\mathsf{D})$ is not associated with any substrategies then we need not consider the part of $\mathsf{T}(\mathsf{D})$ corresponding to the subtree $\mathsf{T}(\mathsf{D}')$. For each of the candidate substrategies at D we construct a system of linear inequalities. Given that $\mathsf{T}(\mathsf{D})$ contains the decision nodes $\mathcal{N}_\mathsf{D}$ we obtain a system of inequalities where the:

- Number of variables $= |\mathcal{N}_\mathsf{D}| + |\mathcal{C}|$.

- Number of inequalities $\leq |\mathcal{N}_{\mathsf{D}}| \cdot \mathsf{sp}_{\max} + |\mathcal{C}|$; $\mathsf{sp}_{\max}$ is the maximum number of children for a decision node.

A solution to these inequalities can be found by e.g. employing a simplex algorithm or an interior point algorithm. Although the simplex algorithm can require an exponential running time for some difficult classes of linear problems, the running time has been shown to be polynomial on the average; this also agrees with the complexity observed in practice (see (Schrijver 1986)). On the other hand, the interior point algorithm is proven to have a polynomial running time (Karmarkar 1984).

Now, let $\#_{\mathsf{T(D)}}$ denote the time complexity of performing the non-dominance test in the subtree $\mathsf{T(D)}$; $\#_{\mathsf{T(D)}}$ is directly determined by the running time of the applied linear programming algorithm. If $\left|\mathcal{C}^{\mathsf{T(D)}}\right|$ denotes the maximum number of possible consequences for a strategy in $\mathsf{T(D)}$, then we have:

$$\text{Relative complexity} = \frac{\mathsf{O}\left(\sum_{\mathsf{D}\in\mathcal{N}_{\mathsf{D}}}\left[\kappa^{\omega}\cdot|\mathsf{sp}_{\mathsf{D}}|\cdot\left(\#_{\mathsf{T(D)}}+\left|\mathcal{C}^{\mathsf{T(D)}}\right|\right)\right]+\sum_{\mathsf{C}\in\mathcal{N}_{\mathsf{C}}}\left|\mathcal{C}^{\mathsf{T(D)}}\right|\right)}{\mathsf{O}\left(\sum_{\mathsf{D}\in\mathcal{N}_{\mathsf{D}}}|\mathsf{sp}_{\mathsf{D}}|+\sum_{\mathsf{C}\in\mathcal{N}_{\mathsf{C}}}|\mathsf{sp}_{\mathsf{C}}|\right)}.$$

It is important to notice that the complexity of investigating a strategy is strongly dependent on the number of distinct consequences in the decision tree. In real world decision problems it is reasonable to assume that the number of distinct consequences is significantly smaller than the number of possible consequences in the decision tree. For instance, most people are usually almost indifferent between receiving 1000\$ or 1005\$ thus, these consequences may be given equal utilities.

4 Empirical results

Two sets of tests have been performed w.r.t. the evaluation of decision trees based on non-expected utility. In the first set of tests, we have determined the average number of non-dominated strategies produced by rolling back the decision tree according to the sophisticated behavior of a RDU maximizer; the value of κ (the maximum number of strategies acceptable by a Self) was set to 1. In the second set of tests, the algorithm proposed in Section 3 was used to identify an average number for κ needed to produce a non-dominated strategy; the program *PCx* was used to solve the inequalities. It should be noted that considerations regarding the value of θ have been omitted from the tests since this value should be elicited in relation to the DM and is therefore specific to the decision problem (and the DM) in question. Generally, one would expect that, given two values θ_1 and θ_2 (where $\theta_1 < \theta_2$), the number of strategies acceptable by a Self under θ_1 is less than (or equal to) the number of strategies acceptable by the same Self under θ_2.

The first set of tests was performed on a collection of randomly generated decision trees with a branching factor between 2 and 4 (randomly chosen) and where the number of decision nodes and chance nodes varied from 15 to 7200; the probabilities associated with the chance nodes were also generated randomly. For a given number of nodes, 50 different decision trees were generated and for each of these decision trees a sophisticated strategy was determined w.r.t. three different utility functions; the range of these utility functions was either $\{0, 1, \ldots, 20\}$, $\{0, 1, \ldots, 100\}$ and $\{0, 1, \ldots, 1000\}$, i.e., each possible consequence was randomly assigned an integer utility level within these ranges. The average number of non-dominated sophisticated strategies is illustrated in Figures 6-8, which supports the prevailing assumption that a sophisticated strategy is likely to be dominated. This result also justifies the use of different models and evaluation schemes when considering dynamic decision problems with non-separable preference functions.

The second set of tests was also performed on a collection of randomly generated decision trees. The number of nodes varied between 15 and 4000, and for a given number of nodes 20 different decision trees were generated. For each of these decision trees the number of strategies acceptable by a Self was incremented (starting with $\kappa = 1$) until a non-dominated strategy was generated. The results of the tests are shown in Figures 9-11.

The second set of tests indicates that the average value for κ (needed to obtain a non-dominated strategy) increases w.r.t. both the size of the decision tree and the range of the

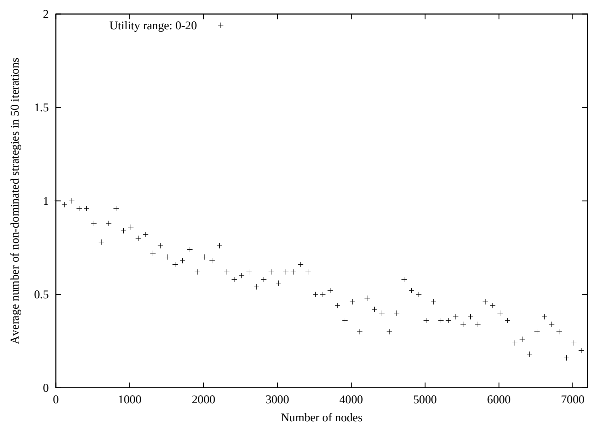

Figure 6: The average number of non-dominated sophisticated strategies as a function of the number of chance nodes and decision nodes in the decision tree. The utility levels varied between 0 and 20.

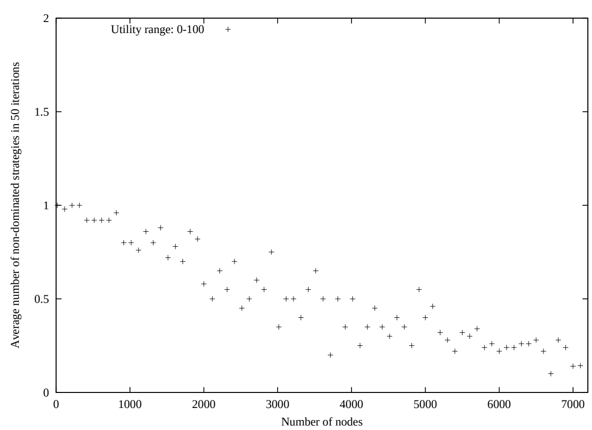

Figure 7: The average number of non-dominated sophisticated strategies as a function of the number of chance nodes and decision nodes in the decision tree. The utility levels varied between 0 and 100.

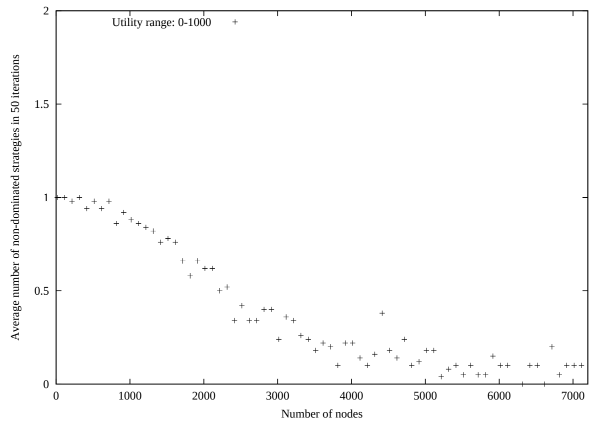

Figure 8: The average number of non-dominated sophisticated strategies as a function of the number of chance nodes and decision nodes in the decision tree. The utility levels varied between 0 and 1000.

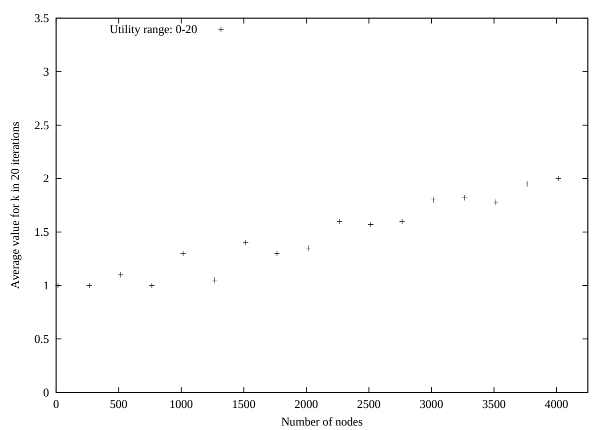

Figure 9: The average value of κ which yields a non-dominated strategy as a function of the number of chance nodes and decision nodes in the decision tree. The utility levels varied between 0 and 20.

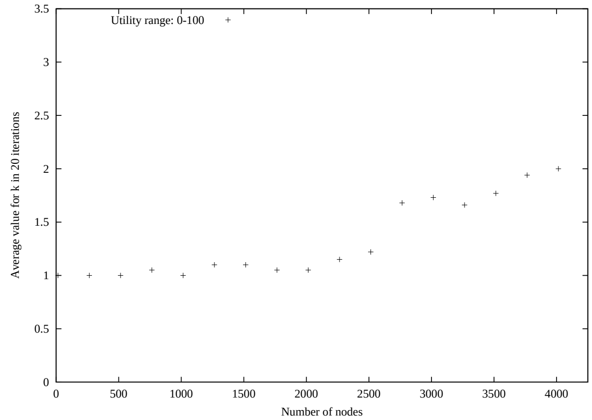

Figure 10: The average value of κ which yields a non-dominated strategy as a function of the number of chance nodes and decision nodes in the decision tree. The utility levels varied between 0 and 100.

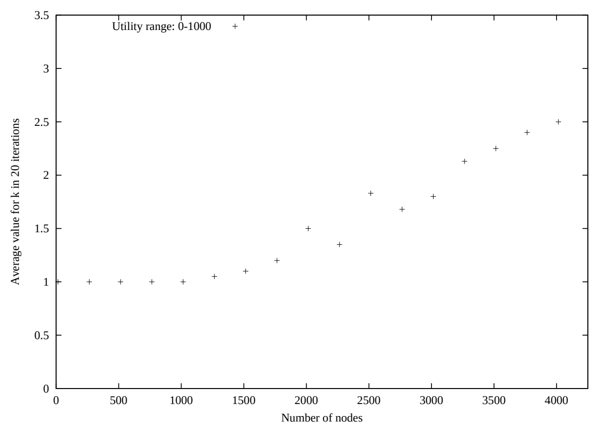

Figure 11: The average value of κ which yields a non-dominated strategy as a function of the number of chance nodes and decision nodes in the decision tree. The utility levels varied between 0 and 1000.

utility function: by increasing the size of the decision tree and the range of the utility function (i.e., the number of different consequences) we produce a larger system of inequalities, and, in general, we would therefore expect that fewer configurations satisfy this system.

As a final observation, the value assigned to κ can be guided by the empirical findings above, i.e., they support an initial evaluation of the decision tree (using backtracking) with an optimistic value for κ.

5 Conclusion

In this paper we have presented an operational approach to decision aiding based on non-EU theory. The approach was motivated by the fact that some DMs will not accept recommendations based on EU theory, on the ground that EU theory cannot account for certain factors, such as the certainty effect, that may have a strong influence on their preferences. In practice, the DA, not finding a specification of EU theory that fits sufficiently well with the DM's preferences, might consider using Rank Dependent Utility with an appropriate probability weighting function for selecting a strategy that the DM will accept.

Non-EU criteria such as Quiggin's Rank Dependent Utility are non-separable. Thus, the standard application of backward induction, known as sophisticated behavior, is likely to produce a dominated strategy. Instead we have proposed an algorithm which manages to simultaneously select a non-dominated strategy and comply (as closely as possible) with the DM's non-EU criterion, at each decision node. We conjecture that by making the RCL assumption the algorithm can straightforwardly be applied to other non-EU criteria besides RDU theory; basically the algorithm only relies on RDU theory for ranking substrategies and, as such, RDU theory could be substituted with another non-EU theory. Moreover, we have suggested a psychological interpretation which allows one to understand and justify the algorithm as a realization of a cooperative decision process between the DM's successive Selves.

Rational behavior is often identified with compliance to certain principles that do not need further justification other than one's intimate conviction that they should be abided by. However, a deeper justification has frequently been looked for in so-called Dutch Books or Money Pumps arguments, where the underlying idea is that a behavior rule is not acceptable if it can lead to some waste. Thus, in a limited sense, rational behavior under risk could be defined as behavior never leading to the selection of a dominated strategy. Such "no-wasting" requirements also conform with economic theory practice and in this sense, resolute choice, as presented in this paper, is an instance of rational behavior.

Acknowledgement

This research was done while Thomas D. Nielsen was associated with Université Paris 6, France as a visiting scholar; in the same period Thomas D. Nielsen was a Ph.D.-student at the Department of Computer Science, Aalborg University, Denmark. The research was partly supported by CIT project number 87.2 - "Automated Decision Support for Customer Support Operations - Phase 2". Finally, we would like to thank the anonymous referees for very helpful comments on this draft as well as on an earlier version of the paper.

References

Allais, M. (1979). The foundation of a positive theory of choice involving risk and a criticism of the postulate and axioms of the American school. In M. Allais and O. Hagen (Eds.), *Expected utility hypotheses and the Allais paradox*, pp. 27–145. Dordrecht, Holland. Original work published in 1952.

Allais, M. (1988). The general theory of random choices in relation to the invariant cardinal utility function and the specific probability function. In B. Munier (Ed.), *Risk, Decision and Rationality*, pp. 233–289. Dordrecht, Holland.

Chateauneuf, A. and M. Cohen (1984). Risk seeking with diminishing marginal utility in a non-expected utility model. *Journal of Risk and Uncertainty 9*, 77–91.

Chew, S., E. Karni, and Z. Safra (1987). Risk aversion in the theory of expected utility with rank dependent preferences. *Journal of Economic Theory 42*, 370–381.

Choquet, G. (1953). Theory of capacities. *Annales de l'Institut Fourier 5*, 131–295.

Cubbit, R. P., C. Starmer, and R. Sugden (1998). Dynamic choice and the common ratio effect: an experimental investigation. *The Economic Journal 108*, 1362–1380.

Currim, I. S. and R. K. Sarin (1989). Prospect versus utility. *Management Science 35*(1), 22–41.

Dantzig, G. B. and P. Wolfe (1960). Decomposition principle of linear programs. *Operations Research 8*(1), 101–111.

de Finetti, B. (1974). *Theory of Probability*, Volume 1. Wiley, New York.

Hadar, J. and W. Russell (1969). Rules for ordering uncertain prospects. *American Economic Review 59*(1), 97–122.

Hammond, J. (1988). Consequentialist foundations for Expected Utility. *Theory and Decision 25*, 25–78.

Jaffray, J.-Y. (1999). Rational decision making with imprecise probabilities. In 1*st International Symposium on Imprecise Probabilities and Their Applications*, pp. 183–188. Morgan Kaufmann Publishers.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*, 263–291.

Karmarkar, N. (1984). A new polynomial-time algorithm for linear progamming. *Combinatorica 4*, 373–395.

Karni, E. and Z. Safra (1989). Ascending bid auctions with behaviorally consistent bidders. *Annals of Operations Research 19*, 435–446.

Machina, M. J. (1987). Choice under uncertainty: problems solved and unsolved. *Journal of Economic Perspectives 1*(1), 121–154.

Machina, M. J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature 27*, 1622–1688.

McClennen, E. F. (1990). *Rationality and Dynamic choice: Foundational explorations*. Cambridge University Press, Cambridge.

McCord, M. R. and R. de Neufville (1984). Utility dependence on probability: An empirical demonstration. *Journal of Large Scale Systems 6*, 91–103.

Olmsted, S. M. (1983). *On representing an solving decision problems*. Ph. D. thesis, Department of Engineering-Economic Systems, Stanford University.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organisation 3*(4), 323–343.

Raiffa, H. (1968). *Decision Analysis, Introductory Lectures on Choices under Uncertainty*. Massachusetts: Addison-Wesley.

Rothschild, M. and J. Stiglitz (1970). Increasing Risk I: a definition. *Journal of Economic Theory 3*, 66–84.

Schick, F. (1986). Dutch bookies and money pumps. *Journal of Philosophy 83*, 112–119.

Schrijver, A. (1986). *Theory of Linear and Integer Programming*. Wiley, N-Y.

Segal, U. (1990). Two-stage Lotteries whithout the Reduction Axiom. *Econmetrica 58*, 349–377.

Wu, G. and R. Gonzalez (1996). Curvature of the probability weighting function. *Management Science 42*(12), 1676–1690.

Yaari, M. (1987). The dual theory of choice under risk. *Econometrica 55*(1), 95–115.