Official URL

DOI : https://doi.org/10.1007/978-3-319-91479-4_48

Sequential decision making under uncertainty: ordinal uninorms vs. the Hurwicz criterion

Hélène Fargier and Romain Guillaume

IRIT,CNRS and Université de Toulouse, France
{fargier, guillaum}@irit.fr

Abstract. This paper focuses on sequential decision problems under uncertainty, i.e. sequential problems where no probability distribution on the states that may follow an action is available. New qualitative criteria are proposed that are based on ordinal uninorms, namely R_* and R^*. Like the Hurwicz criterion, the R_* and R^* uninorms arbitrate between pure pessimism and pure optimism, and generalize the Maximin and Maximax criteria. But contrarily to the Hurwicz criterion they are associative, purely ordinal and compatible with *Dynamic Consistency* and *Consequentialism*. This latter important property allow the construction of an optimal strategy in polytime, following an algorithm of Dynamic Programming.

Keywords: qualitative decision making, uncertainty, sequential decision problems

1 Introduction

In a sequential decision problem under uncertainty, a decision maker faces a sequence of decisions, each decision possibly leading to several different states, where further decisions have to be made. A strategy is a conditional plan which assigns a (possibly non deterministic) action to each state were a decision has to be made (also called "decision node"), and each strategy leads to a compound lottery, following Von Neuman and Morgenstern's terminology [17] - roughly, a tree representing the different possible scenarios, and thus the different possible final states that the plan/strategy may reach. The optimal strategy is then the one which maximizes a criterion applied to the resulting compound lottery.

Three assumptions are desired to accept the optimal strategy without discussions on the meaning of optimal strategy. Those assumptions are:

- *Dynamic Consistency*: when reaching a decision node by following an optimal strategy, the best decision at this node is the one that had been considered so when computing this strategy, i.e. prior to applying it.
- *Consequentialism*: the best decision at each step of the problem only depends on potential consequences at this point.
- *Tree Reduction*: a compound lottery is equivalent to a simple one.

Those three assumptions are linked to the possibility to compute an optimal strategy using an algorithm of dynamic programming [13].

When the problem is pervaded with uncertainty the Hurwicz criterion [7] is often advocated since it generalizes the optimistic maximax and the pessimistic maximin approaches. It makes a "compromise" between these approaches, through the use of a coefficient α of optimism - the Hurwicz value being the linear combination, according to this coefficient, of the two criteria.

Unfortunately, this approach does not suit qualitative, ordinal, utilities: the Hurwicz criterion proceeds to an additive compensation of the min value by the max value. Moreover, the criterion turns out to be incompatible with the above assumptions: it can happen that none of the optimal strategies is dynamically consistent nor consequentialist - as a consequence the optimization of this criterion cannot be carried out using dynamic programming.

In such a situation, a decision maker using the Hurwicz criterion should adopt a resolute choice behavior [2], initially choosing a strategy and never deviating from it later. But many authors insist on the fact that Resolute Choice is not acceptable since a normally behaved decision maker is consequentialist. This leads some of them to use algorithmic approaches based on Veto-process [11] and Ego-dependent process [3] (see also [9],[8]).

In the present paper, rather than trying to "repair" the Hurwicz criterion in an algorithmic way, we are looking for new qualitative criteria which can take into account the level optimism/pessimism of the decision maker, like Hurwicz's criterion, and satisfies the three properties stated above (*Dynamic Consistency, Consequentialism* and *Tree Reduction*).

The paper is structured as follows. The next Section presents the Hurwicz criterion, the background on decision trees under pure uncertainty and the principle of dynamic programming. Section 3 then proposes the use of two qualitative uninorms, R^* and R_*, as alternatives to the Hurwicz criterion. Drowning them in the context of sequential decision making, we show in Section 4 that R^* and R_* are compatible with *Dynamic Consistency* and *Consequentialism* , and propose to apply an algorithm of dynamic programming to compute an optimal, consequentialist and dynamically consistent strategy. Section 5 eventually summarises the discussion between the two uninorms and the Hurwicz criterion [1].

2 Background

2.1 The Hurwicz criterion [7]

Let us first consider simple, non sequential decision problems under uncertainty: each decision δ_i is characterized by the multi set of final states $E_{\delta_i} = \{s^i_1, ..., s^i_{m^i}\}$ it can lead to. Given a utility function u capturing the attractiveness of each of these final states, δ_i can be identified with a simple lottery over the utility levels that may be reached: in decision under uncertainty, where no probability distribution over the consequences of an act is available, a simple lottery is

[1] The proofs are omitted for the sake of brevity.

indeed the multiset of the utility levels of the s^i_j, i.e. $L_{\delta_i} = (u^i_1, ..., u^i_{m^i})$ (where $u^i_j = u(s^i_j)$).

A usual way to take the optimism of the decision maker (DM in the following) into account is to use the Hurwicz criterion. The worth of δ_i is then:

$$H(\delta_i) = H(L_{\delta_i}) = (1 - \alpha) \times \min(u^i_1, ..., u^i_{m^i}) + \alpha \times \max(u^i_1, ..., u^i_{m^i}).$$

where $\alpha \in [0, 1]$ is the degree of optimism. H indeed collapses with max aggregation when $\alpha = 1$ (and with the min aggregation when $\alpha = 0$).

2.2 Decision trees

A convenient language to introduce sequential decision problems is through decision trees [13]. This framework proposes an explicit modeling in a graphical way, representing each possible scenario by a path from the root to the leaves of a tree. Formally, a decision tree $\mathcal{T} = (\mathcal{N}, \mathcal{E})$ is such that $\mathcal{N}$ contains three kinds of nodes (see Figure 1 for an example):

- $\mathcal{D} = \{d_0, \ldots, d_m\}$ is the set of decision nodes (depicted by rectangles).
- $\mathcal{LN} = \{ln_1, \ldots, ln_k\}$ is the set of leaves, that represent final states in $\mathcal{S} = \{s_1, \ldots, s_k\}$; such states can be evaluated thanks to a utility function: $\forall s_i \in \mathcal{S}$, $u(s_i)$ is the degree of satisfaction of being eventually in state s_i (of reaching node ln_i). For the sake of simplicity we assume, without loss of generality, that only leaf nodes lead to utilities.
- $\mathcal{X} = \{x_1, \ldots, x_n\}$ is the set of chance nodes (depicted by circles).

For any node $n_i \in \mathcal{N}$, $Succ(n_i) \subseteq \mathcal{N}$ denotes the set of its children. In a decision tree, for any decision node $d_i, Succ(d_i) \subseteq \mathcal{X}$: $Succ(d_i)$ is the set of actions that can be chosen when d_i is reached. For any chance node $x_i, Succ(x_i) \subseteq \mathcal{LN} \cup \mathcal{D}$: $Succ(x_i)$ is the set of possible outcomes of action x_i - either a leaf node is observed, or a decision node is reached (and then a new action should be chosen).

The present paper is devoted to *qualitative* decision making under uncertainty; thus:

- the information at chance nodes is a list of potential outcomes - this suits situations of total ignorance, where no probabilistic distribution is available.
- the preference about the final states is purely qualitative (ordinal), i.e. we cannot assume more than a preference order on the consequences (on the leaves of the tree), captured by the satisfaction degrees. The scale $[0, 1]$ is chosen for these degrees, but any ordered set can be used.

Solving a decision tree amounts at building a *strategy*, i.e. a function δ that associates to each decision node d_i an action (i.e. a chance node) in $Succ(d_i)$: $\delta(d_i)$ is the action to be executed when decision node d_i is reached. Let Δ be the set of strategies that can be built for $\mathcal{T}$. We shall also consider the subtree $\mathcal{T}_n$ of $\mathcal{T}$ rooted at node $n \in \mathcal{T}$, and denote by Δ_n its strategies: they are subtrategies of the strategies of Δ.

Any strategy in Δ can be viewed as a connected subtree of $\mathcal{T}$ where there is exactly one edge (and thus one chance node) left at each decision node -

skipping the decision nodes, we get a chance tree or, using von Neuwman and Morgernstern's terminology, a compound lottery [2].

Simple lotteries indeed suit the representation of decisions made at the last step of the tree: $(u_1, ..., u_k)$ is the multiset of the utilities of the leaf nodes $(ln_1, ..., ln_k)$ that may be reached when some decision x is executed. Consider now a decision x made at the penultimate level: it may lead to any of the decision nodes d_i in $Succ(x)$, and thus to any of the simple lotteries $L_i = (u^i_1, ..., u^i_{m^i})$, $d_i \in Succ(x)$ - the substrategy rooted in x defines the compound lottery $(L_i,\ \text{s.t.}\ d_i \in Succ(x))$. The reasoning generalizes for decisions x at any level of the tree, hence the definition of the (possibly multi level) compound lottery L_δ associated to δ.

In order to apply a criterion, e.g. Hurwicz's, a simple lottery is needed. To this extent the *Reduction* of the compound lottery relative to the strategy is computed, which is the simple lottery which gathers all the utilities reached by the inner lotteries. Formally, the reduction of a compound lottery $L = (L_1, ..., L_k)$ composed of lotteries L_i is defined by:

$$Reduction(L) = Reduction(L_1) \cup ... \cup Reduction(L_k) \quad (1)$$

where the reduction of a simple lottery is the simple lottery itself. For instance, if L composed of simple lotteries $(L_1, ..., L_k)$, with $L_i = (u^i_1, ..., u^i_{n^i})$:

$$Reduction(L) = (u^1_1, ..., u^1_{n^1}, ..., u^k_1, ..., u^k_{n^k}) \quad (2)$$

The principle of reduction make the comparison of compound lotteries (and thus of strategies) possible: to compare compound lotteries by some criteria O, simply apply it to their reductions:

$$O(L) = O(Reduction(L)) \quad (3)$$

For instance, considering the Hurwicz criterion, the preference relation over strategies is defined by:

$$\delta \preceq_H \delta' \text{ iff } H(Reduction(L_\delta) \preceq H(Reduction(L_{\delta'})) \quad (4)$$

In all the approaches that follow Equation (3), and in particular in the approach considered in this paper, *Tree Reduction* is thus obeyed by construction.

Optimality can now be soundly defined, at the global and the local levels:

- $\delta \in \Delta$ is optimal for T iff $\forall \delta' \in \Delta, O(Reduction(L_\delta)) \succeq O(Reduction(L_{\delta'}))$
- $\delta \in \Delta_n$ is optimal for T_n iff $\forall \delta' \in \Delta_n, O(Reduction(L_\delta)) \succeq O(Reduction(L_{\delta'}))$

Let us now consider *Dynamic Consistency*. An optimal strategy δ is said to be dynamically consistent iff for any decision node n, δ_n, the restriction of δ to node n and its descendent, is optimal for the subtree rooted in n. A criterion is

[2] Recall that a simple lottery $L = (u_1, ..., u_k)$ is a multiset of utilities; a compound Lottery $L = (L_1, \ldots, L_k)$ is a multiset of (simple or compound) lotteries.

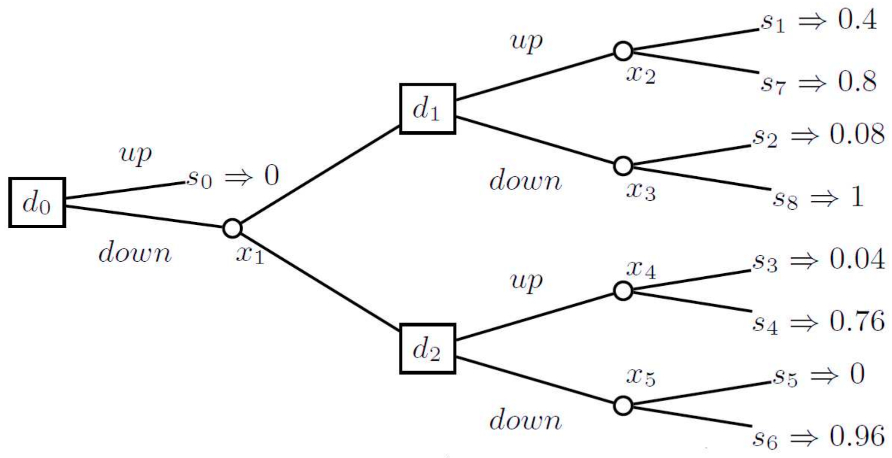

Fig. 1. A Decision Tree

said to be compatible with *Dynamic Consistency* if there is always an optimal strategy that is dynamically consistent.

The purely optimist (resp. pessimist) criterion, max (resp. min) is compatible with *Dynamic Consistency* - there always exist an optimal strategy whose substrategies are optimal. Unfortunately, the Hurwicz criterion is not compatible with *Dynamic Consistency.* Let us give a counter example:

Example 1. Consider the decision tree of Figure 1 and $\alpha = 0.1$; Strategy ($d_0 \leftarrow$ *down*, $d_1 \leftarrow$ *down*, $d_2 \leftarrow$ *up*) is optimal, with a Hurwicz value of $0.1{\cdot}0.04{+}0.9{\cdot}1 = 0.904$; as a matter of fact ($d_0 \leftarrow$ *down*, $d_1 \leftarrow$ *down*, $d_2 \leftarrow$ *down*) has a Hurwicz value of 0.9 and all the strategies with $d_0 \leftarrow$ *up* or $d_1 \leftarrow$ *up* have a lower value. Hence the (only) optimal strategy prescribes "*up*" for d_2. On the other hand, considering the tree rooted in d_2, "*up*" has a H value equal to 0.684, while "*down*" has a H value equal to 0.864 - *up* is not the optimal strategy in this subtree. This counter example shows that Hurwicz is not compatible with *Dynamic Consistency.*

2.3 Dynamic programming

Consequentialism prescribes that the decision maker selects a plan looking only at the possible futures (regardless of the past or counterfactual history) . This is the case when choosing, at each node n, the decision that maximizes O. Hence a consequentialist strategy can be built starting from the anticipated future decisions and rolling back to the present (see Algorithm 1). This is the idea implemented in the algorithm of dynamic programming, which simulates the behaviour of such a consequentialist decision maker: the algorithm builds the

Algorithm 1: Dynamic programming

Input: decision tree $\mathcal{T}$ of depth $p > 1$, criterion O
Output: A strategy δ which is optimal for O, its value $O(\delta)$

```
for ln ∈ LN do
    L(ln) = u(ln)
for t = p − 1 to 0 do
    for d ∈ D_t do
        // D_t denotes the decision nodes at depth t
        for n ∈ Succ(d) do
            L(n) = ⋃_{n' ∈ Succ(n)} L(n')
        δ(d) = argmax_{n ∈ Succ(d)} O(Reduction(L(n)))
        L(d) = L(δ(d))
Return (δ, O(Reduction(L(d_0))))
```

best strategy by a process of backward induction, optimizing the decisions from the leaves of the tree to its root. Roughly, one can say that a criterion is coherent with *Consequentialism* iff the strategy returned by the algorithm of dynamic programming is optimal according to this criterion.

Unfortunately this is not always the case when optimality is based on the principle of *Tree Reduction*: rolling back the Hurwicz optimization at each node of the tree of Figure 1 leads to strategy ($d_0 \leftarrow down$, $d_1 \leftarrow down$, $d_2 \leftarrow down$) which is *not* optimal according to equation (3).

The correctness of dynamic programming actually relies on an important property, called weak monotonicity:

Definition 1 *A preference criterion over lotteries is said to be weakly monotonic iff whatever* L, L' *and* L''*:*

$$O(L) \preceq_O O(L') \Rightarrow O((L, L'')) \preceq O((L', L'')) \tag{5}$$

Proposition 1 *If a criterion* O *satisfies weak monotonicity then the strategy returned by dynamic programming is optimal according to* O*.*

By construction, this strategy is dynamically consistent (any of its substrategies is optimal it its subtree), consequentialist and equivalent, according to O, to its reduction.

Corollary 1 *If a criterion* O *satisfies weak monotonicity then strategy returned by dynamic programming is consequentialist and dynamically consistent*

3 R_* and R^* as criteria for decision making under uncertainty

As we have seen in the previous Section, the Hurwicz criterion which is often advocated for decision making under uncertainty suffers from severe drawbacks, and in particular form its incapacity to satisfy *Dynamic Consistency*. This is regrettable from a prescriptive point of view: when optimizing this criterion, the decision planned for a node is not necessarily the one that would be the best one if the tree rooted at this node were be considered - when reaching this node, a Hurwicz maximizer would be tempted not to follow the plan. That is why we look for alternative qualitative generalizations of the maximax and maximin rules, which, like Hurwicz, allow a balance between pure pessimism and pure optimism.

3.1 The R_* and R^* uninorms

The uninorm aggregators [18] are generalization of t-norms and t-conorms. These operators allow the identity element (e) to lay anywhere in the unit interval - it is not necessarily equal to zero nor to one, as required by t-norms or t-conorms, respectively.

Definition 2 *[18] A uninorm R is a mapping $R : [0,1] \times [0,1] \to [0,1]$ having the following properties:*

1. *$R(a,b) = R(b,a)$ (Commutativity)*
2. *$R(a,b) \geq R(c,d)$ if $a \geq c$ and $b \geq d$ (Monotonicity)*
3. *$R(a, R(b,c)) = R(R(a,b),c)$ (Associativity)*
4. *There exists some element $e \in [0,1]$, called the identity element, such that for all $x \in [0,1]$ $R(x,e) = x$*

In this paper we focus on two ordinal uninorms proposed by Yager [18]:

1. $R_* : [0,1]^n \to [0,1]$:
   - $R_*(a_1, ..., a_n) = Min(a_1, ..., a_n)$ if $Min(a_1, ..., a_n) < e$
   - $R_*(a_1, ..., a_n) = Max(a_1, ..., a_n)$ if $Min(a_1, ..., a_n) \geq e$
2. $R^* : [0,1]^n \to [0,1]$:
   - $R^*(a_1, ..., a_n) = Min(a_1, ..., a_n)$ if $Max(a_1, ..., a_n) < e$
   - $R^*(a_1, ..., a_n) = Max(a_1, ..., a_n)$ if $Max(a_1, ..., a_n) \geq e$

R_* specifies that if one of the a_i's is lower than e then the min operator is applied, otherwise max is applied. R^* specifies that if one of the a_i's is greater than e then the max operator is applied, otherwise min is applied. One can see that both R_* and R^* generalize the min and max uninorms, as Hurwicz does (min is recovered when $e = 1$, max when $e = 0$). The identity element e can represent the threshold of optimism (as α for Hurwicz).

R_* and R^* constitute two different ways of generalizing the maximin and maximax criterion, and capture different types of behaviours of the decision maker. In the context of decision making under uncertainty, we propose to interpret $[0, e[$ as an interval of hazards and $[e, 1]$ as interval of opportunities:

1. When all the possible utilities lay in the hazardous interval, both R_* and R^* behave in a pessimistic way and evaluate the lottery by its worst outcome.

2. When all the possible utilities lay in the interval of opportunity, both R_* and R^* behave in an optimistic way and evaluate the lottery by its best outcome.
3. When some possible utility belongs to the hazardous interval and others in interval of opportunities, R_* returns a pessimistic value (the worst one) while R^* returns the best, optimistic, one.

Hence, in the simultaneous presence of hazards and opportunities, R_* focuses on the hazards while R^* focuses on the opportunities. In other terms, the comparison of strategies is made as follows:

- R^*: if one of the two strategies may lead to (at least) one opportunity, the DM prefers the strategy with the greatest opportunity. If both lead surely into the interval of hazards, the DM prefers the more robust strategy.
- R_*: if one of the two strategies may lead to (at least) one hazardous utility, the DM prefers the more robust of the strategies. If both are exempt of hazards, the DM prefers the one with the greatest opportunity.

In robust decision making, where performance guarantees are looked for, one will obviously apply the R_* uninorm because of its cautiousness. R^* indeed appears as too adventurous: one single possible opportunity carries the final decision, and this even if all the other utilities lay in the hazard interval. On the contrary, R_* looks for opportunity only when the required level of satisfaction, e, is guaranteed for all the possible outcomes.

Example 2. Let us consider three decisions $\square = (0.55, 0.55)$, $\triangle = (0.7, 0.39)$ and $\bigcirc = (0.9, 0.2)$ (see Fig.2). In red, on the figure, is represented the zone containing decisions that the DM would like to avoid because too risky when e is set equal to 0.6 ((a).Figure 2 for R_* and (b). Figure 2 for R^*). One can see that if the DM uses R_*, all the solutions are in the red zone hence she/he will select $\square$. Conversely, if the DM uses R^*, decision $\square$ is the only decision in the red zone and $\bigcirc$ will be selected.

Depending on the value $e \in [0, 1]$, the optimal solutions are:

- $\forall e \in [0, 0.2]$ the optimal solution is $\bigcirc$ for both R_* and R^*.
- $\forall e \in]0.2, 0.39]$ for R_*: $\triangle$ and for R^*: $\bigcirc$
- $\forall e \in]0.39, 0.9]$ for R_*: $\square$ and for R^*: $\bigcirc$
- $\forall e \in]0.9, 1]$ the optimal solution is $\square$ for both uninorms.

Notice that $\triangle$ is favoured by R_*, when the degree of guaranteed performance, e, is moderate ($e \leq 0.39$). If a higher degree of performance must be ensured, R_* chooses $\square = (0.55, 0.55)$.

4 R_* and R^* in the sequential decision context

Let us now study the two uninorms in the context of sequential decision. Applying the principle of lottery reduction, we have:

$$\delta \succeq_{R_*} \delta' \text{ iff } R_*(Reduction(\delta)) \succeq R_*(Reduction(\delta')) \quad (6)$$

$$\delta \succeq_{R^*} \delta' \text{ iff } R^*(Reduction(\delta)) \succeq R^*(Reduction(\delta')) \quad (7)$$

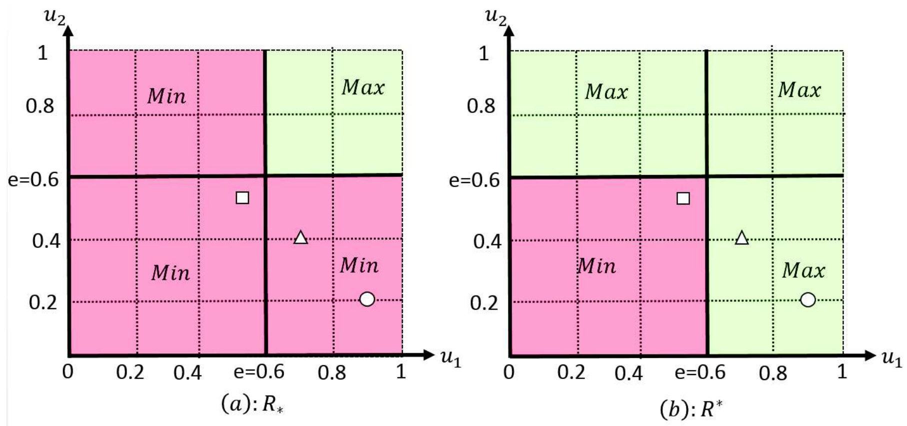

Fig. 2. Illustration of R_* and R^*

Example 3. Let us go back to the example of Figure 1 and focus first on criterion R_*. The strategies that decide *down* for d_2 are risky (may reach s_5, which have a utility of 0) and have a R_* equal to 0 whatever the value of e. This is also the case for all the strategies that decide *up* for d_0. Now,

– if $e \in]0, 0.04]$ $(d_0 \leftarrow down,\ d_1 \leftarrow down,\ d_2 \leftarrow up)$ is optimal, with a $R_* = 1$.
– if $e \in]0.04, 1]$ there are two optimal strategies, $(d_0 \leftarrow down,\ d_1 \leftarrow up,\ d_2 \leftarrow up)$ and $(d_0 \leftarrow down,\ d_1 \leftarrow down,\ d_2 \leftarrow up)$, both with $R_* = 0.04$.

It can be checked that any optimal strategy is dynamically consistent. For instance, $R_*(d_2 \leftarrow up)$, which is at least equal to 0.04 (whatever e), is always greater than $R_*(d_2 \leftarrow down)$,which is always equal to 0.

If we consider R^*, both $(d_0 \leftarrow down,\ d_1 \leftarrow down,\ d_2 \leftarrow down)$ and $(d_0 \leftarrow down,\ d_1 \leftarrow down,\ d_2 \leftarrow up)$ are optimal: their R^* is equal to 1, whatever the value e (and both are dynamically consistent)

Beyond this example, R_* and R^* behave well for sequential problems in the general case; indeed, both are compatible with *Dynamic Consistency* and *Consequentialism*. The reason is that, contrarily to the Hurwicz criterion, they satisfy weak monotonicity:

Proposition 2 *The R_* and R^* satisfies weak monotonicity*

A direct consequence of Propositions 1 and 2 is that both uninorms can be optimized by dynamic programming (see Algorithm 2)

Theorem 1 *Algorithm 2 computes a strategy optimal w.r.t R^* (resp. R_*) in time polynomial with respect to the size of the decision tree.*

This strategy is thus consequentialist and dynamically consistent; it follows from Theorem 1 that:

Algorithm 2: R^* and R_* under pure uncertainty

Input: decision tree $\mathcal{T}$ of depth $p > 1$, criterion $O \in \{R_*, R^*\}$, optimism coefficient e
Output: A strategy δ which is optimal for O, its value $O(\delta)$

for $ln \in \mathcal{LN}$ **do**
    $V(ln) = u(ln)$
for $t = p - 1$ *to* 0 **do**
    for $d \in \mathcal{D}_t$ **do**
        `//` $\mathcal{D}_t$ `denotes the decision nodes at depth` t
        for $n \in Succ(d)$ **do**
            $V(n) = O((V(n'), n' \in Succ(n)))$
        $\delta(d) = argmax_{n \in Succ(d)} V(n)$
        $V(d) = V(\delta(d))$
Return $(\delta, V(d_0))$

Corollary 2 *The uninorm R^* and R_* are compatible with* Dynamic Consistency*,* Consequentialism *and* Tree Reduction*.*

As already outlined compatibility with *Dynamic Consistency* guarantees that the DM cannot be tempted to deviate from the plan during its execution. Because R_* is consequentialist, the evaluation of a decision can be conservative at some node in the tree (because hazard cannot be excluded) and become optimistic when some safer point is reached (e.g. at node d_1 when $e \leq 0.08$). On the example of Figure 1, with $e = 0.05$, R_* compares the min values of the two candidate decisions at node d_2, but is optimistic at node d_1: all the outputs that can be reached from d_1 are greater than 0.05, i.e. all the decision are safe when d_1 is reached. Similar examples can be built for R^* (which is nevertheless less in accordance with the intuition, since pessimism is taken into account only when no opportunity is available).

A last algorithmic advantage of R^* and R_* over Hurwicz is that they are associative (like any uninorm). This allows the algorithm of dynamic programming to memorize, for each node, the *value* of the corresponding reduced lottery rather than the lottery itself

Definition 3 *A criterion O satisfies the decomposition principle iff whatever L, L', $O(L \cup L') = O(O(L), O(L'))$.*

Proposition 3 *R^* and R_* satisfy the decomposition principle*

Hurwicz, which is not associative, does not satisfy this principle - for instance $H((1,0),(0)) = \alpha^2$ while $H((1,0,0)) = \alpha$.

5 R_* and R^* vs. Hurwicz

Let us now focus on the comparison between the uninorms (and especially of R_*, which has a well founded interpretation in terms of robustness) and Hurwicz's criterion. All are generalization of the maximax and maximin criteria, allow a tuning between optimism and pessimism, and extend to sequential problems through the application of the principle of lottery reduction.

The first remark is that R_* can capture the desiderata of a decision maker who looking for guarantees of performance, the level of performance being represented by e. This kind of requirement cannot be captured by the Hurwicz criterion, unless $\alpha = 0$, i.e. unless Hurwicz collapses with the min (and also collapses with R_* and with R^*, setting $e = 0$).

Moreover, the behaviour of Hurwicz's approach may appear chaotic in its way to move from pessimism to optimism. Consider again Example 2: $\Box = (0.55, 0.55)$ and $\bigcirc = (0.2, 0.9)$ are the min optimal and max optimal solutions, respectively. The max (resp. the min) value of $\triangle$ lays between the ones of $\Box$ and $\bigcirc$, so $\triangle = (0.39, 0.7)$ appears as an intermediate solution between $\Box$ and $\bigcirc$ (see Figure 2). Nevertheless, $\triangle$ is never optimal for Hurwicz. It can indeed be checked than $H(\Box) = 0.55$ whatever α. $H(\triangle) = 0.545$ at $\alpha = 0.5$. When $\alpha \leq 0.5$, $H(\triangle) < 0.55 = H(\Box)$; when $\alpha \geq 0.5$ $H(\bigcirc) \geq H(\triangle)$, because $H(\bigcirc)$ increases faster than $H(\triangle)$. Hence a slight variation of α makes Hurwicz jump directly from the pessimistic solution $\Box$ to the very optimistic solution $\bigcirc$, without considering $\triangle$, which is Pareto optimal and intermediate between $\Box$ and $\bigcirc$.

If we look at the formal properties that may be looked for, the first difference is that the uninorms are purely ordinal. They do not need to assume that the utilities are additive to some extent, while Hurwicz is basically an additive criterion. The second one is their associativity - a basic property that is not satisfied by the Hurwicz's aggregation. Last but not least, R_* and R^* are compatible with *Dynamic Consistency* and *Consequentialism*, while Hurwicz is not.

A first, practical consequence is that a polynomial algorithm of dynamic programming can be designed to find consequentialist and dynamically consistent optimal solutions. *Dynamic Consistency* and *Consequentialism* are also important from a prescriptive point of view. Because the R_* and R^* optimal strategies are dynamically consistent, the DM will never be tempted to deviate from it - we have seen on Example 1 that Hurwicz does not prevent for such deviations.

Consequentialism says that the value of a (sub)strategies only depends on the future consequences - R_* and R_* never care of "parallel", counter factual worlds. As we have seen, Hurwicz is not compatible with this principle: what happens in a world (e.g., in Example 1 in d_2 when *up* is chosen for d_2) may influence the decision in an independent, parallel world (here, in d_1). Indeed, Hurwicz will always prefer $d_1 \leftarrow down$ to $d_1 \leftarrow up$ even in case of a very low - but positive - degree of optimism. This is due to the fact the low value (0.04) for s_3, which is not a descendent of d_1 but of d_2, masks the 0.08 utility of s_2.

Our running example also shows that Hurwicz can be very adventurous even for small positive α's: ($d_0 \leftarrow down$, $d_1 \leftarrow down$, $d_2 \leftarrow up$) might reach a very low utility (0.08) is indeed optimal for Hurwicz as soon as $\alpha > 0$. This strategy will on the contrary be considered as too risky for R_*, unless a low level ($e < 0.08$) of guaranteed performance is looked for .

6 Conclusion

In this paper, we have shown how the R_* and R^* uninorms can be used for decision under uncertainty. They constitute an appealing alternative to Hurwicz's criterion to model the behavior ofa DM who is not purely optimistic nor purely pessimistic: an optimal strategy can be computed in polytime, which satisfies the three natural assumptions of sequential decision making. Moreover, these utilities are purely qualitative; as a perspective, it would be natural to extend them to possibilistic (qualitative) decision trees [14], that allow the expression of some knowledge about the more or less possible consequences of the decisions.

References

1. N.BenAmor, H.Fargier, W.Guezguez:Possibilistic sequential decision making, International Journal of Approximate Reasonning,Elsevier,55(5): 1269–1300, 2014.
2. E.F.McClennen: Rationality and Dynamic choice: Foundational Explorations. Cambridge University Press, 1990.
3. D.Dubois, H.Fargier, R.Guillaume, C.Thierry: Deciding under ignorance: in search of meaningful extensions of the Hurwicz criterion to decision trees, In Strengthening Links Between Data Analysis and Soft Computing 3-11. Springer, 2014.
4. I. Gilboa: A Combination of Expected Utility and Maxmin Decision Criteria. Journal of Mathematical Psychology, 32:405-420, 1988.
5. S. Grant, A. Kajii, B. Polak: Decomposable Choice under Uncertainty. Journal of Economic Theory 92(2): 169-197, 2000.
6. P. Hammond: Consequentialist Foundations of Expected Utility. Theory and Decision, 25:25-78, 1988.
7. L. Hurwicz, 1951. Optimality Criteria for Decision Making under Ignorance. Cowles commission papers 370.
8. D.Kikuti, F.Gagliardi Cozman and R.Shirota Filho: Sequential decision making with partially ordered preferences. Artificial Intelligence 175: 13461365,2011.
9. N.Huntley and M.Troffaes: An efficient normal form solution to decision trees with lower previsions. Soft Methods for Handling Variability and Imprecision, 419-426,2008.
10. J.-Y. Jaffray: Linear Utility Theory for Belief Functions. Operational Research Letters, 82:107-112, 1989.
11. J.-Y. Jaffray: Rational Decision Making With Imprecise Probabilities. Proc. Int. Sympos. on Imprecise Probabilities (ISIPTA 1999), 183-188, 1999.
12. G. Jeantet: Algorithmes pour la décision séquentielle dans l'incertain: optimisation de l'utilité espérée dépendant du rang et du critère de Hurwicz, Ph. D. Thesis, Université Paris VI, 2010 (in French)
13. H. Raiffa: Decision Analysis: Introductory Lectures on Choices Under Uncertainty. Addison-Wesley, Reading, MA, 1968.
14. R. Sabbadin, H. Fargier, J. Lang: Towards qualitative approaches to multi-stage decision making. International Journal of Approximate Reasoning, 19, 441471, (1998).
15. L. J. Savage: The Foundations of Statistics. New York: Dover Publications 1972.
16. G. Shafer: The Art of Causal Conjecture, MIT Press, 1996.
17. J. Von Neumann, O. Morgenstern: Theory of games and economic behavior. Princeton University Press, Princeton (1947)
18. R. Yager, A. Rybalov: Uninorm aggregation operators. Fuzzy Sets and Systems,80(1):111-120, 1996.