Dynamic Programming for Deterministic Discrete-Time Systems with Uncertain Gain

Gert de Cooman and Matthias C. M. Troffaes

Onderzoeksgroep SYSTeMS, Universiteit Gent, Technologiepark – Zwijnaarde 914, 9052 Zwijnaarde, Belgium

Abstract

We generalise the optimisation technique of dynamic programming for discrete-time systems with an uncertain gain function. We assume that uncertainty about the gain function is described by an imprecise probability model, which generalises the well-known Bayesian, or precise, models. We compare various optimality criteria that can be associated with such a model, and which coincide in the precise case: maximality, robust optimality and maximinity. We show that (only) for the first two an optimal feedback can be constructed by solving a Bellman-like equation.

Key words: optimal control, dynamic programming, uncertainty, imprecise probabilities, lower previsions, sets of probabilities

1 Introduction to the Problem

The main objective in optimal control is to find out how a system can be influenced, or controlled, in such a way that its behaviour satisfies certain requirements, while at the same time maximising a given gain function. A very efficient method for solving optimal control problems for discrete-time systems is the recursive *dynamic programming* technique, introduced by Richard Bellman [1].

To explain the ideas behind it, we refer to Figures 1 and 2. In Figure 1 we depict a situation where a system can go from state a to state c through state b in three ways: following the paths $\alpha\beta$, $\alpha\gamma$ and $\alpha\delta$. We denote the gains associated with these paths by $J_{\alpha\beta}$, $J_{\alpha\gamma}$ and $J_{\alpha\delta}$ respectively. Assume that path $\alpha\gamma$ is optimal, meaning that $J_{\alpha\gamma} > J_{\alpha\beta}$ and $J_{\alpha\gamma} > J_{\alpha\delta}$. Then it follows that path γ is the optimal way to go
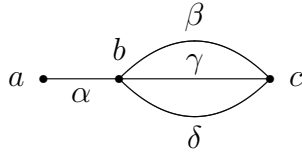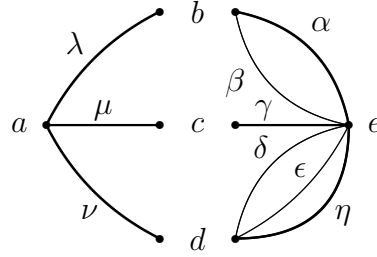
Fig. 1. Principle of Optimality   Fig. 2. Dynamic Programming

from b to c. To see this, observe that $J_{\alpha\nu} = J_\alpha + J_\nu$ for $\nu \in \{\beta, \gamma, \delta\}$ (we shall assume throughout that gains are additive along paths) and derive from the inequalities above that $J_\gamma > J_\beta$ and $J_\gamma > J_\delta$. This simple observation, which Bellman called the *principle of optimality*, forms the basis for the recursive technique of dynamic programming for solving an optimal control problem. To see how this is done in principle, consider the situation depicted in Figure 2. Suppose we want to find the optimal way to go from state a to state e. After one time step, we can reach the states b, c and d from state a, and the optimal paths from these states to the final state e are known to be α, γ and η, respectively. To find the optimal path from a to e, we only need to compare the costs $J_\lambda + J_\alpha$, $J_\mu + J_\gamma$ and $J_\nu + J_\eta$ of the respective candidate optimal paths $\lambda\alpha$, $\mu\gamma$ and $\nu\eta$, since the principle of optimality tells us that the paths $\lambda\beta$, $\nu\delta$ and $\nu\epsilon$ cannot be optimal: if they were, then so would be the paths β, δ and ϵ. This, written down in a more formal language, is what is essentially known as *Bellman's equation*. It allows us to solve an optimal control problem fairly efficiently through a recursive procedure, by calculating optimal paths backwards from the final state.

In applications, it may happen that the gain function, which associates a gain with every possible control action and the resulting behaviour of the system, is not well known. This problem is most often treated by modelling the uncertainty about the gain by means of a probability measure, and by maximising the *expected gain* under this probability measure. Due to the linearity of the expectation operator, this approach does not change the nature of the optimisation problem in any essential way, and the usual dynamic programming method can therefore still be applied.

As an example, consider the simple linear system described by

$$x_{k+1} = ax_k + bu_k, \quad k = 0, \ldots, N-1 \tag{1}$$

where $x_k \in \mathbb{R}$ denote the system state and $u_k \in \mathbb{R}$ the control at time k, and where a and b are non-zero real numbers. Given an initial state x_0 and a sequence $u.$ of successive controls $u_0, u_1, \ldots, u_{N-1}$, the systems goes through the successive states $x_1, x_2, \ldots, x_N$ determined by Eq. (1), and we assume that with this control there is associated a gain

$$J(x_0, u., \omega) = \sum_{k=0}^{N-1} [x_k^2 + \omega u_k^2],$$

where ω is some positive real constant. Solving the present optimal control problem

consists in finding a control $u.$ that brings the system at time N in a given final state x_f, while at the same time maximising the gain $J(x_0, u., \omega)$. The dynamic programming approach achieves this by reasoning backwards in time. First, the control u_{N-1} is determined that maximises the gain

$$x_{N-1}^2 + \omega u_{N-1}^2 = \left(\frac{x_f - b u_{N-1}}{a} \right)^2 + \omega u_{N-1}^2.$$

This control also determines a unique x_{N-1}, and the procedure is then repeated by finding a control u_{N-2} that maximises the gain $x_{N-2}^2 + \omega u_{N-2}^2$, and so on ... The principle of optimality then ensures that the control $u.$ found in this recursive manner indeed solves the optimal control problem. When ω is not well known, and only its probability distribution is given, the optimal control problem is solved by maximising the expected value of the gain, which can in this special example be done by replacing ω with its expectation.

It has however been argued by various scholars (see [2, Chapter 5] for a detailed discussion with many references) that uncertainty cannot always be modelled adequately by (precise) probability measures, because, roughly speaking, there may not be enough information available to identify a single probability measure. In those cases, it is more appropriate to represent the available knowledge through a so-called *imprecise probability model*, e.g., by a *coherent lower prevision*, or what is mathematically equivalent, by a *set of probability measures*. For applications of this approach, see for instance [3,4]. In the example above, it may for instance happen that the probability distribution for ω is only known to belong to a given set: e.g., ω is normally distributed with mean zero, but the variance is only known to belong to an interval $[\underline{\sigma}^2, \overline{\sigma}^2]$; or ω itself is only known to belong to an interval $[\underline{\omega}, \overline{\omega}]$. [2]

Two questions now arise naturally. First of all, how should we formulate the optimal control problem: what does it mean for a control to be optimal with respect to an *uncertain gain function*, where the uncertainty is represented through an imprecise probability model? In Section 2 we identify three different optimality criteria, each with a different interpretation (although they coincide for precise probability models), and we study the relations between them. Secondly, is it still possible to solve the corresponding optimal control problems using the ideas underlying Bellman's dynamic programming method? We show in Section 3 that this is the case for only two of the three optimality criteria we study: only for these a generalised principle of optimality holds, and the optimal controls are solutions of suitably generalised Bellman-like equations. In order to arrive at this conclusion, we study the properties that an abstract notion of optimality should satisfy for the Bellman approach to work. To illustrate how our ideas can be implemented, we present a numerical example in Section 4.

We recognise that other authors (see for instance [5,6,7,8,9]) have extended the

---

[2] This also covers the case where we want to find out how robust the optimal control solution is against variations of ω within a given interval.

dynamic programming algorithm to systems with uncertain gain and/or uncertain dynamics, where the uncertainty is modelled by an imprecise probability model. But none of them seem to have questioned under what assumptions their generalised dynamic programming method leads to optimal paths. Here we approach the problem from the opposite, and in our opinion, more logical side: one should *first* define a notion of optimality and investigate whether the dynamic programming argument holds for it, rather than blindly "generalise" Bellman's algorithm without showing that it actually yields optimal controls.

In the remainder of this section, we introduce the basic systems-theoretic concepts and notation used in the rest of the paper.

1.1 The System

For a and b in $\mathbb{N}$, the set of natural numbers c that satisfy $a \leq c \leq b$ is denoted by $[a, b]$. Let

$$x_{k+1} = f(x_k, u_k, k)$$

describe a discrete-time dynamical system with $k \in \mathbb{N}$, $x_k \in \mathcal{X}$ and $u_k \in \mathcal{U}$. The set $\mathcal{X}$ is the state space (e.g., $\mathbb{R}^n$, $n \in \mathbb{N} \setminus \{0\}$), and the set $\mathcal{U}$ is the control space (e.g., $\mathbb{R}^m$, $m \in \mathbb{N} \setminus \{0\}$). The map $f \colon \mathcal{X} \times \mathcal{U} \times \mathbb{N} \to \mathcal{X}$ describes the evolution of the state in time: given the state $x_k \in \mathcal{X}$ and the control $u_k \in \mathcal{U}$ at time $k \in \mathbb{N}$, it returns the next state x_{k+1} of the system. For practical reasons, we impose a final time N beyond which we are not interested in the dynamics of the system. Moreover, it may happen that not all states and controls are allowed at all times: we demand that x_k should belong to a set of *admissible states* $\mathcal{X}_k$ at every instant $k \in [0, N]$, and that u_k should belong to a set of *admissible controls* $\mathcal{U}_k$ at every instant $k \in [0, N-1]$, where $\mathcal{X}_k \subseteq \mathcal{X}$ and $\mathcal{U}_k \subseteq \mathcal{U}$ are given. The set $\mathcal{X}_N$ may be thought of as the set we want the state to end up in at time N.

1.2 Paths

A *path* is a triple $(x, k, u.)$, where $x \in \mathcal{X}$ is a state, $k \in [0, N]$ a time instant, and $u. \colon [k, N-1] \to \mathcal{U}$ a sequence of controls. Such a path fixes a unique state trajectory $x. \colon [k, N] \to \mathcal{X}$, which is defined recursively through $x_k = x$ and $x_{\ell+1} = f(x_\ell, u_\ell, \ell)$ for every $\ell \in [k, N-1]$. It is said to be *admissible* if $x_\ell \in \mathcal{X}_\ell$ for every $\ell \in [k, N]$ and $u_\ell \in \mathcal{U}_\ell$ for every $\ell \in [k, N-1]$. We denote the unique map from the empty set $\emptyset$ to $\mathcal{U}$ by $u_\emptyset$. If $k = N$, the control $u.$ does nothing: it is equal to $u_\emptyset$. The unique path starting and ending at time $k = N$ in $x \in \mathcal{X}$ is denoted by $(x, N, u_\emptyset)$.

The set of admissible paths starting in the state $x \in \mathcal{X}_k$ at time $k \in [0, N]$ is denoted by $\mathcal{U}(x, k)$, i.e.,

$$\mathcal{U}(x, k) = \{(x, k, u.) \colon (x, k, u.) \text{ admissible path}\}.$$

For example, $\mathcal{U}(x, N) = \{(x, N, u_\emptyset)\}$ whenever $x \in \mathcal{X}_N$ and $\mathcal{U}(x, N) = \emptyset$ otherwise.

If we consider a path with final time M different from N, then we write $(x, k, u_.)_M$ (assume $k \leq M \leq N$). Observe that $(x, k, u_.)_k$ can be identified with $(x, k, u_\emptyset)_k$; it is the unique path (of length zero) starting and ending at time k in x. Let $0 \leq k \leq \ell \leq m$. Two paths $(x, k, u_.)_\ell$ and $(y, \ell, v_.)_m$ can be concatenated if $y = x_\ell$. The concatenation is denoted by $(x, k, u_., \ell, v_.)_m$ or by $(x, k, u_.)_\ell \oplus (y, \ell, v_.)_m$. It represents the path that starts in state x at time k, and results from applying control u_i for times $i \in [k, \ell-1]$ and control v_i for times $i \in [\ell, m-1]$. In particular,

$$(x, k, u_.)_\ell = (x, k, u_.)_k \oplus (x, k, u_.)_\ell = (x, k, u_.)_\ell \oplus (x_\ell, \ell, u_.)_\ell.$$

The set of admissible paths starting in state $x \in \mathcal{X}_k$ at time $k \in [0, N]$ and ending at time $\ell \in [k, N]$ is denoted by $\mathcal{U}(x, k)_\ell$. In particular we have that $\mathcal{U}(x, k)_k = \{(x, k, u_\emptyset)_k\}$ if $x \in \mathcal{X}_k$, and $\mathcal{U}(x, k)_k = \emptyset$ otherwise. Moreover, for any $(x, k, u_.)_\ell \in \mathcal{U}(x, k)_\ell$ and any $\mathcal{V} \subseteq \mathcal{U}(x_\ell, \ell)$, we use the notation

$$(x, k, u_.)_\ell \oplus \mathcal{V} = \{(x, k, u_.)_\ell \oplus (x_\ell, \ell, v_.) \colon (x_\ell, \ell, v_.) \in \mathcal{V}\}.$$

1.3   The Gain Function

We assume that applying the control action $u \in \mathcal{U}$ to the system in state $x \in \mathcal{X}$ at time $k \in [0, N-1]$ yields a real-valued gain $g(x, u, k, \omega)$. Moreover, reaching the final state $x \in \mathcal{X}$ at time N also yields a gain $h(x, \omega)$. The parameter $\omega \in \Omega$ represents the (unknown) state of the world, and it is a device used to model that the gains are not well known. If we knew that the real state of the world was ω_o, we would know the gains to be $g(x, u, k, \omega_o)$ and $h(x, \omega_o)$. As it is, the real state of the world is uncertain, and so are the gains, which could be considered as random variables. It is important to note that the parameter ω only influences the gains; it has no effect on the system dynamics, which are assumed to be known perfectly well.

We shall only consider the important case where the gains are *additive along paths*, i.e., with a path $(x, k, u_.)$ we associate a gain $J(x, k, u_., \omega)$ given by:

$$J(x, k, u_., \omega) = \sum_{i=k}^{N-1} g(x_i, u_i, i, \omega) + h(x_N, \omega),$$

for any $\omega \in \Omega$ (gain additivity). If $M < N$, we also use the notation

$$J(x, k, u_., \omega)_M = \sum_{i=k}^{M-1} g(x_i, u_i, i, \omega).$$

It will be convenient to associate a zero gain with an empty control action: for $k \in [0, N]$ we let $J(x, k, u_., \omega)_k = 0$.

The main objective of optimal control can now be formulated as follows: given that the system is in the initial state $x \in \mathcal{X}$ at time $k \in [0, N]$, find a control sequence $u_. \colon [k, N-1] \to \mathcal{U}$ resulting in an admissible path $(x, k, u_.)$ such that the

corresponding gain $J(x, k, u., \omega)$ is maximal. Moreover, we would like this control sequence $u.$ to be such that its value u_k at time k is a function of x and k only, since in that case the control can be realised through state feedback.

If ω is known, then the problem reduces to the classical problem of dynamic programming, first studied and solved by Bellman [1]. We shall assume here that the available information about the true state of the world is modelled through a *coherent lower prevision* $\underline{P}$ defined on the set $\mathcal{L}(\Omega)$ of *gambles*, or bounded real-valued maps, on Ω. A special case of this obtains when $\underline{P}$ is a linear prevision P. Linear previsions are the precise probability models; they can be interpreted as expectation operators associated with (finitely additive) probability measures, and they are *previsions* or *fair prices* in the sense of de Finetti [10]. We assume that the reader is familiar with the basic ideas behind the theory of coherent lower (and linear) previsions (see [2] for more details).

For a given path $(x, k, u.)$, the corresponding gain $J(x, k, u., \omega)$ can be seen as a real-valued map on Ω, which is denoted by $J(x, k, u.)$ and is called the *gain gamble* associated with $(x, k, u.)$.[3] In the same way we define the gain gambles $g(x_k, u_k, k)$, $h(x_N)$ and $J(x, k, u.)_M$. There is gain additivity: $J(x, k, u., \ell, v.)_m = J(x, k, u.)_\ell + J(x_\ell, \ell, v.)_m$ for $k \le \ell \le m \le N$, and $J(x, k, u.)_k = 0$. We denote by $\mathcal{J}(x, k)$ the set of gain gambles for admissible paths from initial state $x \in \mathcal{X}_k$ at time $k \in [0, N]$:

$$\mathcal{J}(x, k) = \{J(x, k, u.) \colon (x, k, u.) \in \mathcal{U}(x, k)\}.$$

2 Optimality Criteria

2.1 $\underline{P}$-Maximality

The lower prevision $\underline{P}(X)$ of a gamble X has a behavioural interpretation as a subject's supremum acceptable price for buying the gamble X: it is the highest value of μ such that the subject accepts the gamble $X - x$ (i.e., accepts to buy X for a price x) for all $x < \mu$. The conjugate upper prevision $\overline{P}(X) = -\underline{P}(-X)$ of X is then the subject's infimum acceptable price for selling X. This way of looking at a coherent lower prevision $\underline{P}$ defined on the set $\mathcal{L}(\Omega)$ of all gambles allows us to define a strict partial order $>_{\underline{P}}$ on $\mathcal{L}(\Omega)$ whose interpretation is that of strict preference.

Definition 1 *For any gambles X and Y in $\mathcal{L}(\Omega)$ we say that X* strictly dominates *Y, or that X is strictly preferred to Y (with respect to $\underline{P}$), and we write $X >_{\underline{P}} Y$,*

---

[3] To simplify the discussion, we assume that this map is bounded. We have shown elsewhere how the boundedness requirement can be relaxed in the theory of coherent lower previsions [11].

if^4

$$\underline{P}(X - Y) > 0 \ or \ (X \geq Y \ and \ X \neq Y).$$

Indeed, if $X \geq Y$ and $X \neq Y$, then the subject should be willing to exchange Y for X, since this can only improve his gain. On the other hand, $\underline{P}(X - Y) > 0$ expresses that the subject is willing to pay some strictly positive price to exchange Y for X, which again means that he strictly prefers X to Y.

It is clear that we can also use the coherent lower prevision $\underline{P}$ to express a strict preference between any two *paths* $(x, k, u_\cdot)$ and $(x, k, v_\cdot)$, based on their gains: if $J(x, k, u_\cdot) >_{\underline{P}} J(x, k, v_\cdot)$ this means that the uncertain gain $J(x, k, u_\cdot)$ is strictly preferred to the uncertain gain $J(x, k, v_\cdot)$. We then say that the path $(x, k, u_\cdot)$ is strictly preferred to $(x, k, v_\cdot)$, and we use the notation $(x, k, u_\cdot) >_{\underline{P}} (x, k, v_\cdot)$.

The relation $>_{\underline{P}}$ is anti-reflexive and transitive.[5] It is therefore indeed a strict partial order on $\mathcal{L}(\Omega)$, and in particular also on $\mathcal{J}(x, k)$ and on $\mathcal{U}(x, k)$. But it is generally not linear: unless $\underline{P}$ is a linear prevision, there will typically be gambles X and Y such that $\underline{P}(X - Y) \leq 0 \leq \overline{P}(X - Y)$, and therefore $X \not>_{\underline{P}} Y$ and $Y \not>_{\underline{P}} X$. Two paths need not be comparable with respect to this order, and it does not always make sense to look for greatest elements, i.e., for paths that strictly dominate all the others. Rather, we should look for *maximal*, or undominated, elements: paths $(x, k, u_\cdot)$ that are not dominated by any other path, meaning that $(x, k, v_\cdot) \not>_{\underline{P}} (x, k, u_\cdot)$ for all paths $(x, k, v_\cdot)$ in $\mathcal{U}(x, k)$. Observe that a maximal gamble X in a set $\mathcal{K}$ with respect to $>_{\underline{P}}$ can be characterised as a maximal element of $\mathcal{K}$ with respect to $\geq$ (i.e., it is point-wise undominated) such that $\overline{P}(X - Y) \geq 0$ for all $Y \in \mathcal{K}$. In case $\underline{P}$ is a linear prevision P, maximal gambles with respect to $>_P$ are precisely the point-wise undominated gambles whose prevision is maximal; they maximise the expected gain. This motivates the following optimality definition.

Definition 2 *Let $k \in [0, N]$, $x \in \mathcal{X}_k$ and $\mathcal{V} \subseteq \mathcal{U}(x, k)$. A path $(x, k, u_\cdot^*)$ in $\mathcal{V}$ is called $\underline{P}$-maximal, or $>_{\underline{P}}$-optimal, in $\mathcal{V}$ if no path in $\mathcal{V}$ is strictly preferred to $(x, k, u_\cdot^*)$, i.e., $(x, k, u_\cdot) \not>_{\underline{P}} (x, k, u_\cdot^*)$ for all $(x, k, u_\cdot) \in \mathcal{V}$. We denote the set of the $\underline{P}$-maximal paths in $\mathcal{V}$ by $\mathrm{opt}_{>_{\underline{P}}}(\mathcal{V})$. The operator $\mathrm{opt}_{>_{\underline{P}}}$ is called the* optimality *operator induced by $>_{\underline{P}}$, associated with $\mathcal{U}(x, k)$.*

The $\underline{P}$-maximal paths in $\mathcal{U}(x, k)$ are precisely those admissible paths starting at time k in state x for which the associated gain gamble is a maximal element of $\mathcal{J}(x, k)$ with respect to the strict partial order $>_{\underline{P}}$. If we denote the set of these $>_{\underline{P}}$-maximal gain gambles in $\mathcal{J}(x, k)$ by $\mathrm{opt}_{>_{\underline{P}}}(\mathcal{J}(x, k))$, then for all $(x, k, u_\cdot) \in \mathcal{U}(x, k)$:

$$(x, k, u_\cdot) \in \mathrm{opt}_{>_{\underline{P}}}(\mathcal{U}(x, k)) \iff J(x, k, u_\cdot) \in \mathrm{opt}_{>_{\underline{P}}}(\mathcal{J}(x, k)).$$

$\underline{P}$-maximal paths do not always exist: not every partially ordered set has maximal elements. A fairly general sufficient condition for the existence of $\underline{P}$-maximal ele-

---

[4] The symbol "$\geq$" denotes the point-wise order on gambles.
[5] Since $\underline{P}$ is coherent, we have $\underline{P}(X - X) = 0$ and $\underline{P}(Z - X) \geq \underline{P}(Z - Y) + \underline{P}(Y - X)$.

ments in $\mathcal{J}(x,k)$ (and hence in $\mathcal{U}(x,k)$) is that $\mathcal{J}(x,k)$ should be compact [6] (and of course non-empty). This follows from a general result mentioned in [2, Section 3.9.2], which is also proven in Lemma 3 below. In fact, we use this lemma to prove a stronger result in Theorem 4, whose Corollary 5 turns out to be very important in showing that the dynamic programming approach works for $\underline{P}$-maximality (see Section 3.2).

In order to prove Lemma 3 and Theorem 4, it is convenient to introduce the partial preorder (a reflexive and transitive relation) $\succcurlyeq_{\underline{P}}$ on $\mathcal{L}(\Omega)$ defined by:

$$X \succcurlyeq_{\underline{P}} Y \Leftrightarrow \underline{P}(X - Y) \geq 0.$$

Recall that an element X of a subset $\mathcal{K}$ of $\mathcal{L}(\Omega)$ is a maximal element of $\mathcal{K}$ with respect to $\succcurlyeq_{\underline{P}}$ if it is undominated, i.e., if and only if

$$(\forall Y \in \mathcal{K})(Y \succcurlyeq_{\underline{P}} X \Rightarrow X \succcurlyeq_{\underline{P}} Y). \tag{2}$$

Lemma 3 *For any non-empty compact subset $\mathcal{K}$ of $\mathcal{L}(\Omega)$ the following statements hold.*

(i) *If X is a maximal element of $\mathcal{K}$ with respect to $\succcurlyeq_{\underline{P}}$, then $\overline{P}(X - Y) \geq 0$ for all Y in $\mathcal{K}$.*

(ii) *For every X in $\mathcal{K}$ the subset*

$$\uparrow_{\underline{P}} X = \left\{ Y \in \mathcal{K} \colon Y \succcurlyeq_{\underline{P}} X \right\} = \{ Y \in \mathcal{K} \colon \underline{P}(Y - X) \geq 0 \}$$

of $\mathcal{K}$ is non-empty and compact.

(iii) *There is a maximal element of $\mathcal{K}$ with respect to $\succcurlyeq_{\underline{P}}$.*

(iv) *For every X in $\mathcal{K}$ there is a maximal element Y of $\mathcal{K}$ with respect to $\succcurlyeq_{\underline{P}}$ such that $Y \succcurlyeq_{\underline{P}} X$.*

(v) *For every X in $\mathcal{K}$ there is a maximal element Y of $\mathcal{K}$ with respect to the pointwise order $\geq$ such that $Y \geq X$.*

(vi) *There is maximal element of $\mathcal{K}$ with respect to $>_{\underline{P}}$.*

PROOF. Assume that X is a maximal element of $\mathcal{K}$ with respect to $\succcurlyeq_{\underline{P}}$. Consider Y in $\mathcal{K}$, then it follows from Condition (2) that $\underline{P}(Y - X) < 0$ or $\underline{P}(X - Y) \geq 0$. In both cases it follows that $\overline{P}(X - Y) \geq 0$. This proves (i).

It is obvious that $X \succcurlyeq_{\underline{P}} X$, so $\uparrow_{\underline{P}} X$ is non-empty. Consider a sequence (X_n) in $\uparrow_{\underline{P}} X$ that converges to some gamble X_∞: $\sup_{\omega \in \Omega} |X_\infty(\omega) - X_n(\omega)| \to 0$. Since $\mathcal{K}$ is compact and therefore closed, we know that $X_\infty \in \mathcal{K}$. It now follows from the coherence of $\underline{P}$ (see [2, Theorem 2.6.1]) and $\underline{P}(X_n - X) \geq 0$ that

$$\underline{P}(X_\infty - X) = \underline{P}(X_\infty - X_n + X_n - X)$$
$$\geq \underline{P}(X_\infty - X_n) + \underline{P}(X_n - X) \geq \underline{P}(X_\infty - X_n),$$

---

[6] In this paper, we always assume that $\mathcal{L}(\Omega)$ is provided with the supremum-norm topology.

for all n. Since $-|X_\infty - X_n| \leq X_\infty - X_n \leq |X_\infty - X_n|$, the coherence of $\underline{P}$ (again see [2, Theorem 2.6.1]) also tells us that

$$- \overline{P}(|X_\infty - X_n|) = \underline{P}(-|X_\infty - X_n|)$$
$$\leq \underline{P}(X_\infty - X_n) \leq \underline{P}(|X_\infty - X_n|) \leq \overline{P}(|X_\infty - X_n|),$$

and that $0 \leq \overline{P}(|X_\infty - X_n|) \leq \sup|X_\infty - X_n|$. This implies that $\underline{P}(X_\infty - X_n) \to 0$, and therefore $\underline{P}(X_\infty - X) \geq 0$, whence $X_\infty \in \uparrow_{\underline{P}} X$. This tells us that $\uparrow_{\underline{P}} X$ is a closed subset of the compact $\mathcal{K}$ and therefore also compact, proving (ii).

To prove (iii), let $\mathcal{K}'$ be any subset of the non-empty compact set $\mathcal{K}$ that is linearly ordered with respect to $\succcurlyeq_{\underline{P}}$. If we can show that $\mathcal{K}'$ has an upper bound in $\mathcal{K}$ with respect to $\succcurlyeq_{\underline{P}}$, then we can infer from Zorn's lemma that $\mathcal{K}$ has a $\succcurlyeq_{\underline{P}}$-maximal element. Let then $\{X_1, X_2, \dots, X_n\}$ be an arbitrary finite subset of $\mathcal{K}'$. We can assume without loss of generality that $X_1 \succcurlyeq_{\underline{P}} X_2 \succcurlyeq_{\underline{P}} \dots \succcurlyeq_{\underline{P}} X_n$, and consequently $\uparrow_{\underline{P}} X_1 \subseteq \uparrow_{\underline{P}} X_2 \subseteq \dots \subseteq \uparrow_{\underline{P}} X_n$. This implies that the intersection $\bigcap_{k=1}^{n} \uparrow_{\underline{P}} X_k = \uparrow_{\underline{P}} X_1$ of these up-sets is non-empty. We see that the collection $\left\{ \uparrow_{\underline{P}} X \colon X \in \mathcal{K}' \right\}$ of compact and therefore closed subsets of the compact set $\mathcal{K}$ has the finite intersection property. Consequently, the intersection $\bigcap_{X \in \mathcal{K}'} \uparrow_{\underline{P}} X$ is non-empty as well, and this is the set of upper bounds of $\mathcal{K}'$ in $\mathcal{K}$ with respect to $\succcurlyeq_{\underline{P}}$.

To prove (iv), combine (ii) and (iii) to show that the non-empty compact set $\uparrow_{\underline{P}} X$ has a maximal element Y with respect to $\succcurlyeq_{\underline{P}}$. It is then a trivial step to prove that Y is also $\succcurlyeq_{\underline{P}}$-maximal in $\mathcal{K}$.

The fifth statement follows from the fourth: let $\underline{P}$ be the (coherent) so-called vacuous lower prevision, defined by $\underline{P}(X) = \inf \{X(\omega) \colon \omega \in \Omega\}$. Then the order $\succcurlyeq_{\underline{P}}$ is nothing but the pointwise order $\geq$.

We now come to the last statement. By combining (i) and (iii), we know that there is some Y_o in $\mathcal{K}$ such that $\overline{P}(Y_o - X) \geq 0$ for all $X \in \mathcal{K}$. From (v) we infer that there is some $\geq$-maximal Y in $\mathcal{K}$ such that $Y \geq Y_o$, and therefore (by coherence) $\overline{P}(Y - X) \geq \overline{P}(Y_o - X) \geq 0$ for all $X \in \mathcal{K}$. This means that Y is a maximal element of $\mathcal{K}$ with respect to $>_{\underline{P}}$.

Theorem 4 *For every element X of a compact subset $\mathcal{K}$ of $\mathcal{L}(\Omega)$ that is not a maximal element of $\mathcal{K}$ with respect to $>_{\underline{P}}$, there is some maximal element Y of $\mathcal{K}$ with respect to $>_{\underline{P}}$ such that $Y >_{\underline{P}} X$.*

PROOF. Consider an element X of $\mathcal{K}$ that is not $>_{\underline{P}}$-maximal in $\mathcal{K}$. We may assume that X is $\geq$-maximal in $\mathcal{K}$. Indeed, if X is not $\geq$-maximal then by Lemma 3(v) there is some $\geq$-maximal Z in $\mathcal{K}$ such that $Z \geq X$ and $Z \neq X$, whence $Z >_{\underline{P}} X$. If Z is $>_{\underline{P}}$-maximal in $\mathcal{K}$ then there is nothing left to prove. So we are left with the case that Z is not $>_{\underline{P}}$-maximal. If we can prove for this $\geq$-maximal Z that there is some $>_{\underline{P}}$-maximal Y in $\mathcal{K}$ such that $Y >_{\underline{P}} Z$ then also $Y >_{\underline{P}} X$ and the proof is complete.

Since X is $\geq$-maximal in $\mathcal{K}$, there is some U in $\mathcal{K}$ such that $\underline{P}(U - X) > 0$. By Lemma 3(ii) and (vi) there is a $>_{\underline{P}}$-maximal element Y in $\uparrow_{\underline{P}}U$. Since $\underline{P}(Y - U) \geq 0$ we infer from the coherence of $\underline{P}$ (see [2, Theorem 2.6.1(e)]) that $\underline{P}(Y - X) = \underline{P}(Y - U + U - X) \geq \underline{P}(Y - U) + \underline{P}(U - X) > 0$, whence $Y >_{\underline{P}} X$. It remains to prove that Y is also $>_{\underline{P}}$-maximal in $\mathcal{K}$. Assume *ex absurdo* that there is some V in $\mathcal{K}$ such that $V >_{\underline{P}} Y$. Then there are two possibilities. If $V \geq Y$ and $V \neq Y$ then it follows from the coherence of $\underline{P}$ (see [2, Theorem 2.6.1(d)]) that $\underline{P}(V-U) \geq \underline{P}(Y-U) \geq 0$, whence $V \in \uparrow_{\underline{P}}U$, a contradiction. If $\underline{P}(V-Y) > 0$ then it follows from the coherence of $\underline{P}$ (see [2, Theorem 2.6.1(e)]) that $\underline{P}(V - U) = \underline{P}(V - Y + Y - U) \geq \underline{P}(V - Y) + \underline{P}(Y - U) > 0$, whence $V \in \uparrow_{\underline{P}}U$, again a contradiction.

Corollary 5 *Let $k \in [0, N]$ and let $x \in \mathcal{X}_k$. If $\mathcal{J}(x, k)$ is compact then for every admissible, non-$\underline{P}$-maximal path $(x, k, u_\cdot)$ in $\mathcal{U}(x, k)$ there is a $\underline{P}$-maximal path $(x, k, u_\cdot^*)$ in $\mathcal{U}(x, k)$ that is strictly preferred to it.*

2.2 $\underline{P}$-Maximinity

We now turn to a different optimality criterion that can be associated with a lower prevision $\underline{P}$. We use $\underline{P}$ to define another strict order on $\mathcal{L}(\Omega)$:

Definition 6 *For any gambles X and Y in $\mathcal{L}(\Omega)$ we write $X \sqsupset_{\underline{P}} Y$ if*

$$\underline{P}(X) > \underline{P}(Y) \text{ or } (X \geq Y \text{ and } X \neq Y).$$

$\sqsupset_{\underline{P}}$ induces a strict partial order on $\mathcal{U}(x, k)$, since it is anti-reflexive and transitive on $\mathcal{L}(\Omega)$. A maximal element X of a subset $\mathcal{K}$ of $\mathcal{L}(\Omega)$ with respect to $\sqsupset_{\underline{P}}$ is easily seen to be a point-wise undominated element of $\mathcal{K}$ that maximises the lower prevision: $\underline{P}(X) \geq \underline{P}(Y)$ for all $Y \in \mathcal{K}$.

We can consider as optimal in $\mathcal{U}(x, k)$ those admissible paths $(x, k, u_\cdot)$ for which the associated gain gamble $J(x, k, u_\cdot)$ is a maximal element of $\mathcal{J}(x, k)$ with respect to $\sqsupset_{\underline{P}}$; they are the paths $(x, k, u_\cdot)$ that maximise the 'lower expected gain' $\underline{P}(J(x, k, u_\cdot))$ and whose gain gambles $J(x, k, u_\cdot)$ are point-wise undominated.

Definition 7 *Let $k \in [0, N]$, $x \in \mathcal{X}_k$ and $\mathcal{V} \subseteq \mathcal{U}(x, k)$. A path $(x, k, u_\cdot^*)$ in $\mathcal{V}$ is called $\underline{P}$-maximin, or $\sqsupset_{\underline{P}}$-optimal, in $\mathcal{V}$ if no path in $\mathcal{V}$ is strictly preferred to $(x, k, u_\cdot^*)$, i.e., $(x, k, u_\cdot) \not\sqsupset_{\underline{P}} (x, k, u_\cdot^*)$ for all $(x, k, u_\cdot) \in \mathcal{V}$. We denote the set of the $\underline{P}$-maximin paths in $\mathcal{V}$ by $\mathrm{opt}_{\sqsupset_{\underline{P}}}(\mathcal{V})$. The operator $\mathrm{opt}_{\sqsupset_{\underline{P}}}$ is called the optimality operator induced by $\sqsupset_{\underline{P}}$, associated with $\mathcal{U}(x, k)$.*

Proposition 8 *$\underline{P}$-maximinity implies $\underline{P}$-maximality. For a linear prevision P, P-maximinity is equivalent to P-maximality.*

PROOF. Consider a set of gambles $\mathcal{K}$ and assume that X is a maximal element of $\mathcal{K}$ with respect to $\sqsupset_{\underline{P}}$. In order to prove that X is also a maximal element of $\mathcal{K}$ with respect to $>_{\underline{P}}$, it obviously suffices to show that $\overline{P}(X - Y) \geq 0$ for all $Y \in \mathcal{K}$. We know that $\underline{P}(X) \geq \underline{P}(Y)$ for any $Y \in \mathcal{K}$, and consequently, taking into account coherence (see [2, Section 2.6.1(e)]):

$$\overline{P}(X - Y) \geq \underline{P}(X) + \overline{P}(-Y) = \underline{P}(X) - \underline{P}(Y) \geq 0.$$

If $\underline{P}$ is a linear prevision P, assume that X is a maximal element of $\mathcal{K}$ with respect to $>_P$. In order to prove that X is also a maximal element of $\mathcal{K}$ with respect to $\sqsupset_{\underline{P}}$, it suffices to show that $P(X) \geq P(Y)$ for all $Y \in \mathcal{K}$. Since we know that for any $Y \in \mathcal{K}$, $P(X - Y) \geq 0$, and that $P(X - Y) = P(X) - P(Y)$, the desired result follows at once.

The existence of maximal elements with respect to $\sqsupset_{\underline{P}}$ in an arbitrary set of gambles $\mathcal{K}$ is obviously not guaranteed. But if $\mathcal{K}$ is compact, then we may easily infer from the continuity of any coherent lower prevision $\underline{P}$, that the counterparts of Theorem 4 and Corollary 5 hold for $\sqsupset_{\underline{P}}$.

2.3 $\mathcal{M}$-Maximality

There is a tendency, especially among robust Bayesians, to consider an imprecise probability model as a compact convex set of linear previsions $\mathcal{M} \subseteq \mathcal{P}(\Omega)$, where $\mathcal{P}(\Omega)$ is the set of all linear previsions on $\mathcal{L}(\Omega)$. $\mathcal{M}$ is assumed to contain the true, but unknown, linear prevision P_T that models the available information [12,13].

A gamble X is then certain to be strictly preferred to a gamble Y under the true linear prevision P_T if and only if it is strictly preferred under all candidate models $P \in \mathcal{M}$. This observation leads to the definition of a 'robustified' strict partial order $>_{\mathcal{M}}$ on $\mathcal{L}(\Omega)$.

Definition 9 $X >_{\mathcal{M}} Y$ *if* $X >_P Y$ *for all* $P \in \mathcal{M}$.

Since $\mathcal{M}$ is assumed to be compact and convex, it is not difficult to show that the strict partial orders $>_{\mathcal{M}}$ and $>_{\underline{P}}$ are one and the same, where the coherent lower prevision $\underline{P}$ is the so-called *lower envelope* of $\mathcal{M}$, defined by $\underline{P}(X) = \inf \{P(X) : P \in \mathcal{M}\}$ for all $X \in \mathcal{L}(\Omega)$.[7] Conversely, given a coherent lower prevision $\underline{P}$, the strict partial orders $>_{\mathcal{M}(\underline{P})}$ and $>_{\underline{P}}$ are identical, where

$$\mathcal{M}(\underline{P}) = \{P \in \mathcal{P}(\Omega) : (\forall X \in \mathcal{L}(\Omega))(P(X) \geq \underline{P}(X))\}$$

is the set of linear previsions that dominate $\underline{P}$. These strict partial orders have the same maximal elements, and lead to the same notion of optimality.

---

[7] Since $\mathcal{M}$ is compact, this infimum is actually achieved.

But there is in the literature yet another notion of optimality that can be associated with a compact convex set of linear previsions $\mathcal{M}$: a gamble X is considered optimal in a set of gambles $\mathcal{K}$ if it is a maximal element of $\mathcal{K}$ with respect to the strict partial order $>_P$ for *some* $P \in \mathcal{M}$. This notion of optimality is called 'E-admissibility' by Levi [14, Section 4.8]. It does not generally coincide with the ones associated with the strict partial orders $>_\mathcal{M}$ and $>_{\underline{P}}$, unless the set $\mathcal{K}$ is convex [2, Section 3.9]. We are therefore led to consider a third notion of optimality, associated with a lower prevision $\underline{P}$, or a set of linear previsions $\mathcal{M}$.

Definition 10 *Let $x \in \mathcal{X}$, $k \in [0, N]$ and $\mathcal{V} \subseteq \mathcal{U}(x, k)$. A path $(x, k, u_{\cdot}^*) \in \mathcal{V}$ is said to be $\mathcal{M}$-maximal in $\mathcal{V}$ if it is P-maximal in $\mathcal{V}$ for some P in $\mathcal{M}$, i.e., if it is $\geq$-maximal in $\mathcal{V}$ and maximises $P(J(x, k, u_{\cdot}))$ over $\mathcal{V}$ for some $P \in \mathcal{M}$. The set of all $\mathcal{M}$-maximal elements of $\mathcal{V}$ is denoted by $\mathrm{opt}_\mathcal{M}(\mathcal{V})$.*

Interestingly, for any set of paths $\mathcal{V} \subseteq \mathcal{U}(x, k)$:

$$\mathrm{opt}_\mathcal{M}(\mathcal{V}) = \bigcup_{P \in \mathcal{M}} \mathrm{opt}_{>_P}(\mathcal{V}). \tag{3}$$

3 Dynamic Programming

3.1 A General Notion of Optimality

So far, we have discussed three different ways of associating optimal paths with a lower prevision $\underline{P}$, all of which occur in the literature. We now propose to find out whether, for these different types of optimality, we can use the ideas behind the dynamic programming method to solve the corresponding optimal control problems. To do this, we take a closer look at Bellman's analysis as described in Section 1, and we investigate which properties a generic notion of optimality must satisfy for his method to work. Let us therefore assume that there is some property, called *$*$-optimality*, which a path in a given set of paths $\mathcal{P}$ either has or does not have. If a path in $\mathcal{P}$ has this property, we say that it is $*$-optimal in $\mathcal{P}$. We shall denote the set of the $*$-optimal elements of $\mathcal{P}$ by $\mathrm{opt}_*(\mathcal{P})$. By definition, $\mathrm{opt}_*(\mathcal{P}) \subseteq \mathcal{P}$. Further on, we shall apply our findings to the various instances of $*$-optimality described above.

Consider Figure 3, where we want to find the $*$-optimal paths from state a to state e. Suppose that after one time step, we can reach the states b, c and d from state a. The $*$-optimal paths from these states to the final state e are known to be α, γ, and δ and η, respectively. For the dynamic programming approach to work, we need to be able to infer from this a generalised form of the Bellman equation, stating essentially that the $*$-optimal paths from a to e, *a priori* given by $\mathrm{opt}_*(\{\lambda\alpha, \lambda\beta, \mu\gamma, \nu\delta, \nu\epsilon, \nu\eta\})$, are actually also given by $\mathrm{opt}_*(\{\lambda\alpha, \mu\gamma, \nu\delta, \nu\eta\})$, i.e., the $*$-optimal paths in the set of concatenations of λ, μ and ν with the respective $*$-optimal paths α, γ, and δ and η. It is therefore necessary to exclude
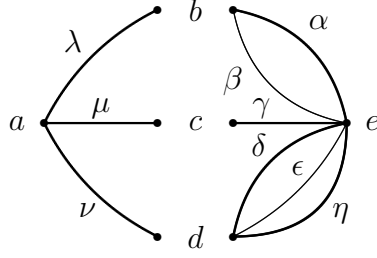
Fig. 3. A More General Type of Dynamic Programming

that the concatenations $\lambda\beta$ and $\nu\epsilon$ with the non-$*$-optimal paths β and ϵ can be $*$-optimal. This amounts to requiring that the operator opt_* should satisfy some appropriate generalisation of Bellman's *principle of optimality* that will allow us to conclude that $\lambda\beta$ and $\nu\epsilon$ cannot be $*$-optimal because then β and ϵ would be $*$-optimal as well. Definition 13 below provides a precise general formulation.

But, perhaps surprisingly for someone familiar with the traditional form of dynamic programming, opt_* should satisfy an *additional* property: the omission of the non-$*$-optimal paths $\lambda\beta$ and $\nu\epsilon$ from the set of candidate $*$-optimal paths should not have any effect on the actual $*$-optimal paths: we need that

$$\text{opt}_*\left(\{\lambda\alpha, \lambda\beta, \mu\gamma, \nu\delta, \nu\epsilon, \nu\eta\}\right) = \text{opt}_*\left(\{\lambda\alpha, \mu\gamma, \nu\delta, \nu\eta\}\right).$$

This is obviously true for the simple type of optimality that we have looked at in Section 1, but it need not be true for the more abstract types that we want to consider here. Equality will be guaranteed if opt_* is insensitive to the omission of non-$*$-optimal elements from $\{\lambda\alpha, \lambda\beta, \mu\gamma, \nu\delta, \nu\epsilon, \nu\eta\}$, in the following sense.

Definition 11 *Consider a set $S \neq \emptyset$ and an* optimality operator opt_* *defined on the set $\wp(S)$ of subsets of S such that $\text{opt}_*(T) \subseteq T$ for all $T \subseteq S$. Elements of $\text{opt}_*(T)$ are called $*$-optimal in T. The optimality operator opt_* is called* insensitive to the omission of non-$*$-optimal elements from S *if $\text{opt}_*(S) = \text{opt}_*(T)$ for all T such that $\text{opt}_*(S) \subseteq T \subseteq S$.*

The following proposition gives an interesting sufficient condition for this insensitivity in case optimality is associated with a (family of) strict partial order(s): it suffices that every non-optimal path is strictly dominated by an optimal one.

Proposition 12 *Let S be a non-empty set provided with a family of strict partial orders $>_j$, $j \in J$. Define for $T \subseteq S$, $\text{opt}_{>_j}(T) = \{a \in T : (\forall b \in T)(b \not>_j a)\}$ as the set of maximal elements of T with respect to $>_j$, and let $\text{opt}_J(T) = \bigcup_{j \in J} \text{opt}_{>_j}(T)$. Then $\text{opt}_{>_j}$, $j \in J$ and opt_J are optimality operators. If for some $j \in J$,*

$$(\forall a \in S \setminus \text{opt}_{>_j}(S))(\exists b \in \text{opt}_{>_j}(S))(b >_j a), \qquad (4)$$

then $\text{opt}_{>_j}$ is insensitive to the omission of non-$>_j$-optimal elements from S. If Condition (4) holds for all $j \in J$, then opt_J is insensitive to the omission of non-J-optimal elements from S.

13

PROOF. Consider j in J, and assume that Condition (4) holds for this j. Let $\mathrm{opt}_{>_j}(S) \subseteq T \subseteq S$, then we must prove that $\mathrm{opt}_{>_j}(S) = \mathrm{opt}_{>_j}(T)$. First of all, if $a \in \mathrm{opt}_{>_j}(S)$ then $b \not>_j a$ for all b in S, and *a fortiori* for all b in T, whence $a \in \mathrm{opt}_{>_j}(T)$. Consequently, $\mathrm{opt}_{>_j}(S) \subseteq \mathrm{opt}_{>_j}(T)$. Conversely, let $a \in \mathrm{opt}_{>_j}(T)$ and assume *ex absurdo* that $a \notin \mathrm{opt}_{>_j}(S)$. It then follows from (4) that there is some c in $\mathrm{opt}_{>_j}(S)$ and therefore in T such that $c >_j a$, which contradicts $a \in \mathrm{opt}_{>_j}(T)$.

Next, assume that (4) holds for all $j \in J$. Let $\mathrm{opt}_J(S) \subseteq T \subseteq S$, then we must prove that $\mathrm{opt}_J(S) = \mathrm{opt}_J(T)$. Consider any $j \in J$, then $\mathrm{opt}_{>_j}(S) \subseteq \mathrm{opt}_J(S) \subseteq T \subseteq S$, so we may infer from the first part of the proof that $\mathrm{opt}_{>_j}(S) = \mathrm{opt}_{>_j}(T)$. By taking the union over all $j \in J$, we find that indeed $\mathrm{opt}_J(S) = \mathrm{opt}_J(T)$.

We are now ready for a precise formulation of the dynamic programming approach for solving optimal control problems associated with general types of optimality. We assume that we have some type of optimality, called $*$-optimality, that allows us to associate with the set of admissible paths $\mathcal{U}(x, k)$ starting at time k in initial state x, an optimality operator opt_* defined on the set $\wp(\mathcal{U}(x, k))$ of subsets of $\mathcal{U}(x, k)$. For each such subset $\mathcal{V}$, $\mathrm{opt}_*(\mathcal{V})$ is then the set of admissible paths that are $*$-optimal in $\mathcal{V}$. The principle of optimality states that the optimality operators associated with the various $\mathcal{U}(x, k)$ should be related in a special way.

Definition 13 (Principle of Optimality) $*$-*optimality satisfies the* principle of optimality *if it holds for all* $k \in [0, N]$, $x \in \mathcal{X}_k$, $\ell \in [k, N]$ *and* $(x, k, u.)$ *in* $\mathcal{U}(x, k)$ *that if* $(x, k, u.)$ *is* $*$-*optimal in* $\mathcal{U}(x, k)$, *then* $(x_\ell, \ell, u.)$ *is* $*$-*optimal in* $\mathcal{U}(x_\ell, \ell)$.

This may also be expressed as:

$$\mathrm{opt}_*\left(\mathcal{U}(x, k)\right) \subseteq \bigcup_{(x,k,u.)_\ell \in \mathcal{U}(x,k)_\ell} (x, k, u.)_\ell \oplus \mathrm{opt}_*\left(\mathcal{U}(x_\ell, \ell)\right).$$

The Bellman equation now states that applying the optimality operator to the right hand side suffices to achieve equality. (Usually this is stated with $\ell = k + 1$.)

Theorem 14 (Bellman Equation) *Let* $k \in [0, N]$ *and* $x \in \mathcal{X}_k$. *Assume that* $*$-*optimality satisfies the principle of optimality, and that the optimality operator* opt_* *for* $\mathcal{U}(x, k)$ *is insensitive to the omission of non-*$*$-*optimal elements from* $\mathcal{U}(x, k)$. *Then for all* $\ell \in [k, N]$:

$$\mathrm{opt}_*\left(\mathcal{U}(x, k)\right) = \mathrm{opt}_*\left(\bigcup_{(x,k,u)_\ell \in \mathcal{U}(x,k)_\ell} (x, k, u)_\ell \oplus \mathrm{opt}_*\left(\mathcal{U}(x_\ell, \ell)\right)\right),$$

that is, a path is $*$-*optimal if and only if it is a* $*$-*optimal concatenation of an admissible path* $(x, k, u.)_\ell$ *and a* $*$-*optimal path of* $\mathcal{U}(x_\ell, \ell)$.

PROOF. Fix k in $[0, N]$, $\ell \in [k, N]$ and $x \in \mathcal{X}_k$. Define

$$\mathcal{V}_1 = \bigcup_{(x,k,u)_\ell \in \mathcal{U}(x,k)_\ell} (x, k, u)_\ell \oplus \text{opt}_* \left(\mathcal{U}(x_\ell, \ell) \right), \text{ and,}$$

$$\mathcal{V}_2 = \bigcup_{(x,k,u)_\ell \in \mathcal{U}(x,k)_\ell} (x, k, u)_\ell \oplus \left(\mathcal{U}(x_\ell, \ell) \setminus \text{opt}_* \left(\mathcal{U}(x_\ell, \ell) \right) \right).$$

Obviously, $\mathcal{U}(x, k) = \mathcal{V}_1 \cup \mathcal{V}_2$ and $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$. We have to prove that $\text{opt}_* \left(\mathcal{U}(x, k) \right) = \text{opt}_* \left(\mathcal{V}_1 \right)$. By the principle of optimality, no path in $\mathcal{V}_2$ is $*$-optimal in $\mathcal{U}(x, k)$, so $\mathcal{V}_2 \cap \text{opt}_* \left(\mathcal{U}(x, k) \right) = \emptyset$. This implies that $\text{opt}_* \left(\mathcal{U}(x, k) \right) \subseteq \mathcal{V}_1 \subseteq \mathcal{U}(x, k)$, and since opt_* is assumed to be insensitive to the omission of non-$*$-optimal elements from $\mathcal{U}(x, k)$, it follows that $\text{opt}_* \left(\mathcal{U}(x, k) \right) = \text{opt}_* \left(\mathcal{V}_1 \right)$.

Let us now apply these general results to the specific types of optimality introduced in the previous section. For all three optimality operators $\text{opt}_{>_P}$, $\text{opt}_{\mathcal{M}}$ and $\text{opt}_{\sqsupseteq_P}$, we shall check whether we can use a Bellman equation to solve the corresponding optimal control problem.

3.2 $\underline{P}$-Maximality

We first consider the optimality operator $\text{opt}_{>_P}$ that selects from a set of gambles (or paths) S those gambles (or paths) that are the maximal elements of S with respect to the strict partial order $>_P$. The following lemma roughly states that the preference amongst paths with respect to $>_P$ is preserved under concatenation and truncation. It yields a sufficient condition for the principle of optimality with respect to $\underline{P}$-maximality to hold. Moreover, the lemma, and the principle of optimality, do not necessarily hold for preference with respect to $\underline{P}$-maximinity.

Lemma 15 *Let $k \in [0, N]$ and $\ell \in [k, N]$. Consider the paths $(x, k, u.)_\ell$ in $\mathcal{U}(x, k)_\ell$ and $(x_\ell, \ell, v.)$, $(x_\ell, \ell, w.)$ in $\mathcal{U}(x_\ell, \ell)$. Then $(x_\ell, \ell, v.) >_{\underline{P}} (x_\ell, \ell, w.)$ if and only if $(x, k, u.)_\ell \oplus (x_\ell, \ell, v.) >_{\underline{P}} (x, k, u.)_\ell \oplus (x_\ell, \ell, w.)$.*

PROOF. Let X, Y and Z be gambles on Ω. The statement is proven if we can show that $Y >_{\underline{P}} Z$ implies $X + Y >_{\underline{P}} X + Z$. Assume that $Y >_{\underline{P}} Z$. If $\underline{P}(Y - Z) > 0$, then $\underline{P}((X + Y) - (X + Z)) = \underline{P}(Y - Z) > 0$. If $Y \geq Z$, then $X + Y \geq X + Z$, and finally, if $Y \neq Z$, then $X + Y \neq X + Z$. It follows that $X + Y >_{\underline{P}} X + Z$.

Proposition 16 (Principle of Optimality) *Let $k \in [0, N]$, $x \in \mathcal{X}_k$ and $(x, k, u_\cdot^*) \in \mathcal{U}(x, k)$. If $(x, k, u_\cdot^*)$ is $\underline{P}$-maximal in $\mathcal{U}(x, k)$ then $(x_\ell, \ell, u_\cdot^*)$ is $\underline{P}$-maximal in $\mathcal{U}(x_\ell, \ell)$ for all $\ell \in [k, N]$.*

15

PROOF. If $(x_\ell, \ell, u^*_\cdot)$ is not $\underline{P}$-maximal, there is a path $(x_\ell, \ell, u_\cdot)$ such that $(x_\ell, \ell, u_\cdot) >_{\underline{P}} (x_\ell, \ell, u^*_\cdot)$. By Lemma 15 we find that

$$(x, k, u^*_\cdot)_\ell \oplus (x_\ell, \ell, u_\cdot) >_{\underline{P}} (x, k, u^*_\cdot)_\ell \oplus (x_\ell, \ell, u^*_\cdot) = (x, k, u^*_\cdot).$$

This means that $(x, k, u^*_\cdot)_\ell \oplus (x_\ell, \ell, u_\cdot)$ is preferred to $(x, k, u^*_\cdot)$, and therefore $(x, k, u^*_\cdot)$ cannot be $\underline{P}$-maximal, a contradiction.

As a direct consequence of Corollary 5 and Proposition 12, we see that if $\mathcal{J}(x, k)$ is compact, then the optimality operator $\mathrm{opt}_{>_P}$ associated with $\mathcal{U}(x, k)$ is insensitive to the omission of non-$>_{\underline{P}}$-optimal elements. Together with Proposition 16 and Theorem 14, this allows us to infer a Bellman equation for $\underline{P}$-maximality.

Corollary 17 *Let $k \in [0, N]$ and $x \in \mathcal{X}_k$. If $\mathcal{J}(x, k)$ is compact, then for all $\ell \in [k, N]$*

$$\mathrm{opt}_{>_P}(\mathcal{U}(x, k)) = \mathrm{opt}_{>_P}\left(\bigcup_{(x,k,u)_\ell \in \mathcal{U}(x,k)_\ell} (x, k, u)_\ell \oplus \mathrm{opt}_{>_P}(\mathcal{U}(x_\ell, \ell))\right), \qquad (5)$$

that is, a path is $\underline{P}$-maximal if and only if it is a $\underline{P}$-maximal concatenation of an admissible path $(x, k, u_\cdot)_\ell$ and a $\underline{P}$-maximal path of $\mathcal{U}(x_\ell, \ell)$.

Corollary 17 results in a procedure to calculate all $\underline{P}$-maximal paths. Indeed, $\mathrm{opt}_{>_P}(\mathcal{U}(x, N)) = \{u_\emptyset\}$ for every $x \in \mathcal{X}_N$, and $\mathrm{opt}_{>_P}(\mathcal{U}(x, k))$ can be calculated recursively through Eq. (5). It also provides a method for constructing a $\underline{P}$-maximal feedback: for every $x \in \mathcal{X}_k$, choose any $(x, k, u^*_\cdot(x, k)) \in \mathrm{opt}_{>_P}(\mathcal{U}(x, k))$. Then $\phi(x, k) = u^*_k(x, k)$ realises a $\underline{P}$-maximal feedback.

3.3 *M-Maximality*

We now turn to the optimality operator $\mathrm{opt}_{\mathcal{M}}$, defined through (3). If we recall Proposition 12, we see that $\mathrm{opt}_{\mathcal{M}}$ is insensitive to the omission of non-$\mathcal{M}$-maximal elements of $\mathcal{U}(x, k)$ whenever $\mathcal{J}(x, k)$ is compact. By Proposition 16, $\mathrm{opt}_{\mathcal{M}}$ satisfies the principle of optimality (indeed, if a path is $\mathcal{M}$-maximal, then it must be P-maximal for some $P \in \mathcal{M}$, and by the proposition any truncation of it is also P-maximal, hence also $\mathcal{M}$-maximal). This means that the Bellman equation also holds for $\mathcal{M}$-maximality under similar conditions as for $\underline{P}$-maximality. As already mentioned in Section 2.3, both types of optimality coincide if $\mathcal{J}(x, k)$ is convex.

3.4 *$\underline{P}$-Maximinity*

Finally, we come to the type of optimality associated with the strict partial order $\sqsupset_{\underline{P}}$. It follows from Proposition 12 and the discussion at the end of Section 2.2
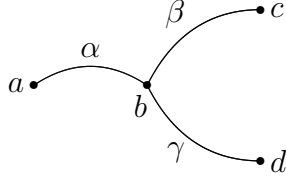
Fig. 4. A Counterexample

that if $\mathcal{J}(x, k)$ is compact, the optimality operator $\mathrm{opt}_{\sqsupset_{\underline{P}}}$ for $\mathcal{U}(x, k)$ is insensitive to the omission of non-$\sqsupset_{\underline{P}}$-optimal paths from $\mathcal{U}(x, k)$. But, as the following counterexample shows, we cannot guarantee that the principle of optimality holds for $\sqsupset_{\underline{P}}$-optimality, and therefore the dynamic programming approach may not work here. Essentially, this is because the partial order $\sqsupset_{\underline{P}}$ is not a vector ordering on $\mathcal{L}(\Omega)$—it is not compatible with gain additivity: contrary to expected gains, lower expected gains are not additive.

Example 18 *Consider the dynamical system depicted in Figure 4. Let $\Omega = \{\sharp, \flat\}$ and denote the gamble $\sharp \mapsto x$, $\flat \mapsto y$ by $\langle x, y \rangle$. Let $\underline{P}$ be the vacuous lower prevision on Ω, defined by $\underline{P}(\langle x, y \rangle) = \min\{x, y\}$. Assume that $J(\alpha) = \langle 2, 0 \rangle$, $J(\beta) = \langle 0, -1 \rangle$ and $J(\gamma) = \langle -2, 0 \rangle$ (there is zero gain associated with the final state). Then $\alpha\beta \not\sqsupseteq_{\underline{P}} \alpha\gamma$: indeed, $\langle 2, -1 \rangle$ does not dominate $\langle 0, 0 \rangle$ point-wise, and $\underline{P}(\langle 2, -1 \rangle) = \min\{2, -1\} \not> \min\{0, 0\} = \underline{P}(\langle 0, 0 \rangle)$ or equivalently $\langle 0, 0 \rangle$ maximises the smallest expected gain. Hence, we find that $\alpha\gamma$ is $\underline{P}$-maximin. But $\beta \sqsupset_{\underline{P}} \gamma$: indeed, $\underline{P}(\langle 0, -1 \rangle) = \min\{0, -1\} > \min\{-2, 0\} = \underline{P}(\langle -2, 0 \rangle)$, which means that γ is not $\underline{P}$-maximin. The "principle of $\underline{P}$-maximin optimality" does not hold here.*

The following theorem gives a sufficient condition for $\underline{P}$-maximality to satisfy the principle of optimality. It seems that this condition is implicitly assumed to hold in most of the literature studying maximin-strategies by dynamic programming.

The idea that underlies this theorem is simple: the principle of optimality will hold if there is additivity of lower expected gains. In order to formulate the theorem in a way that is sufficiently general, we need to introduce a new concept. Assume that the set Ω is a Cartesian product of non-empty sets $\Omega_0, \Omega_1, \ldots, \Omega_N$. Let $\underline{P}$ be a lower prevision defined on $\mathcal{L}(\Omega)$. Then we call $\underline{P}$ *externally additive* relative to $\Omega_0, \Omega_1, \ldots, \Omega_N$ if for all X_k in $\mathcal{L}(\Omega_k)$, where $k = 0, \ldots, N$, it holds that

$$\underline{P}\left(\sum_{k=0}^{N} X_k\right) = \sum_{k=0}^{N} \underline{P}(X_k),$$

where we have identified gambles on the Ω_k with the corresponding gambles on Ω that only depend on the ω_k (their so-called cylindrical extensions).

Theorem 19 *Suppose that $\Omega = \Omega_0 \times \Omega_1 \times \cdots \times \Omega_N$ and assume that the gain gambles $g(x_k, u_k, k)$ are a function of ω_k only, and similarily, that $h(x_N)$ is a function of ω_N only. Let the coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega)$ be externally additive relative to the sets $\Omega_0, \Omega_1, \ldots, \Omega_N$, Then the principle of optimality holds for $\underline{P}$-maximinity.*

17

PROOF. Under the conditions of the theorem, it holds that

$$\underline{P}(J(x,k,u.)) = \underline{P}\left(\sum_{k=0}^{N-1} g(x_k,u_k,k) + h(x_N)\right) = \sum_{k=0}^{N-1} \underline{P}(g(x_k,u_k,k)) + \underline{P}(h(x_N)).$$

This tells us that from the perspective of optimal control, the system is equivalent to a classical optimal control system with precisely known gains $g'(x_i,u_i,i) := \underline{P}(g(x_i,u_i,i))$ and $h'_N(x_N) := \underline{P}(h(x_N))$. Hence, the principle of optimality holds.

The conditions of this theorem will of course not be satisfied in general, but they will hold in a number of important special cases. Sometimes the structure of the problem may impose some type of independence for the lower previsions that model the gain uncertainty at different time points. This usually occurs when the system lends itself to a game-theoretic interpretation, see for instance [5,15]. We mention in passing that external additivity will typically be satisfied if the lower prevision $\underline{P}$ on $\mathcal{L}(\Omega)$ is some type of *independent product* of marginal coherent lower previsions $\underline{P}_k$ defined on the sets $\mathcal{L}(\Omega_k)$. Special cases that lead to external additivity are for instance the *forward irrelevant product* [16], the *independent natural extension* [2,17,18], the *Kuznetsov extension* [19,20] and the *strong independent product* [17], also called the *type-I product* [2,18]. Finally, for a number of very simple imprecise probability models, such as the ones used in Section 4, external additivity is implicitly satisfied.

3.5 Yet Another Type of Optimality

We end this discussion with another type of optimality associated with a strict partial order, sometimes called *interval dominance*, and suggested in a dynamic programming algorithm by Harmanec in [9, Definition 3.4]. In our setting (precisely known system dynamics), its definition basically reduces to

$$X >_{\underline{P}}^{\star} Y \qquad \text{if} \qquad \underline{P}(X) > \overline{P}(Y) \text{ or } (X \geq Y \text{ and } X \neq Y).$$

It can be shown easily that if $\mathcal{J}(x,k)$ is compact, the optimality operator induced by $>_{\underline{P}}^{\star}$ for $\mathcal{U}(x,k)$ is insensitive to the omission of non-$>_{\underline{P}}^{\star}$-optimal paths from $\mathcal{U}(x,\overline{k})$. But, as the following counterexample shows, we cannot guarantee that the principle of optimality holds for $>_{\underline{P}}^{\star}$-optimality, and therefore the dynamic programming approach may not work here. Again, this is because the partial order $\sqsupset_{\underline{P}}$ is not compatible with gain additivity. It also indicates that by solving the Bellman-type equation advocated in [9], we will not necessarily get paths that are optimal in the sense described above.

Example 20 *Consider again the dynamical system depicted in Figure 4. As before, let $\Omega = \{\sharp,\flat\}$, let $\underline{P}$ be the vacuous lower prevision on Ω, and denote the gamble $\sharp \mapsto x$, $\flat \mapsto y$ by $\langle x,y \rangle$. Assume that $J(\alpha) = \langle 2,0 \rangle$, $J(\beta) = \langle 0,0 \rangle$ and $J(\gamma) = \langle -1,-1 \rangle$ (there is zero gain associated with the final state). Then*

$\alpha\beta \not>^\star_{\underline{P}} \alpha\gamma$: *indeed,* $\langle 2, 0 \rangle$ *does not dominate* $\langle 1, -1 \rangle$ *point-wise, and,* $\underline{P}(\langle 2, 0 \rangle) = \min\{2, 0\} \not> \max\{1, -1\} = \overline{P}(\langle 1, -1 \rangle)$. *Hence, we find that* $\alpha\gamma$ *is* $>^\star_{\underline{P}}$*-maximal. But* $\beta >^\star_{\underline{P}} \gamma$: *indeed,* $\langle 0, 0 \rangle$ *dominates* $\langle -1, -1 \rangle$ *point-wise, which means that* γ *is not* $>^\star_{\underline{P}}$*-maximal. The "principle of* $>^\star_{\underline{P}}$*-maximal optimality" does not hold for this example.*

4 A numerical example

Suppose we have a total amount of money x at our disposal, which we can invest into two companies, denoted by 0 and 1. We denote our investment in company 0 by u_0, and in company 1 by u_1. Observe that x, u_0 and u_1 are non-negative real numbers, and $u_0 + u_1 \leq x$. The total gain is

$$T(x, u_0, u_1) = \omega_0 u_0 + \omega_1 u_1 + \omega_2(x - u_0 - u_1),$$

where $\omega_0 > 0$, $\omega_1 > 0$ are gain factors (for companies 0 and 1), and $\omega_2 > 0$ is the devaluation factor (of the money we have not invested). We wish to maximise the gain, but, we are uncertain about ω_0, ω_1 and ω_2.[8] We know that $\omega_0 = 1 + g_0 + \epsilon$ and $\omega_1 = 1 + g_1 + \epsilon$. g_0 and g_1 model the productivities of the companies, and ϵ models economical variations that affect each company in the same way, such as the global economical state. We do not make any assumption about the dependence between g_0, g_1, ϵ and ω_2. We only know that $g_0 \in [0.0, 0.3]$, $g_1 \in [0.1, 0.2]$, $\epsilon \in [-0.1, 0.2]$ and $\omega_2 \in [0.85, 0.95]$. This leads to the following lower prevision on $\mathcal{L}(\Omega_0 \times \Omega_1 \times \Omega_2)$:

$$\underline{P}(X) = \inf \Big\{ X(1 + g_0 + \epsilon, 1 + g_1 + \epsilon, \omega_2) :$$
$$g_0 \in [0.0, 0.3], g_1 \in [0.1, 0.2], \epsilon \in [-0.1, 0.2], \omega_2 \in [0.85, 0.95] \Big\}.$$

We refer to [18] for a detailed discussion about why this lower prevision really captures the available information. We now wish to find all u_0 and u_1 such that the gain $J(x, u_0, u_1)$ is $\underline{P}$-maximal. Observe that this is a two-dimensional optimisation problem.

We formulate this problem in terms of a dynamical system. If we define $x_0 = x$ and, recursively $x_{k+1} = x_k - u_k$, the total gain is precisely equal to $J(x, u_., 0)$, with $g(x_k, u_k, k, \omega) = \omega_k u_k$ and $h(x_2, \omega) = \omega_2 x_2$. Each state x_k represents the money we can invest in companies $\ell \geq k$, and should therefore be non-negative. There is gain additivity, and the set of admissible gain gambles is compact. Corollary 17 applies: we can solve this problem using dynamic programming.

For $k = 1$, we find that the control $u_1 = x_1$ is optimal from state x_1 at time 1. Indeed, first observe that all controls are maximal with respect to the point-wise

---

[8] To ensure that gain gambles are bounded, we can assume that the ω_i belong to some (sufficiently large) bounded closed real intervals.

order. In that case, optimality of u_1 is equivalent to $\overline{P}(J(x_1, u_1, 1) - J(x_1, v_1, 1)) \geq 0$ for all v_1. This holds iff

$$\sup\Big\{(1 + g_1 + \epsilon - \omega_2)(u_1 - v_1):$$
$$g_1 \in [0.1, 0.2], \epsilon \in [-0.1, 0.2], \omega_2 \in [0.85, 0.95]\Big\} \geq 0,$$

and thus, iff $u_1 \geq v_1$ for all v_1. Hence, optimal paths maximise u_1. The highest u_1 we can choose such that x_2 is still non-negative is $u_1 = x_1$.

For $k = 0$, the dynamic programming argument says that we only have to consider concatenations of $(x_0, u_0, 0)_1$ with optimal paths from state $x_1 = x_0 - u_0$, of which there is only one, $(x_1, x_1, 1)$, as we showed. Again all controls are maximal with respect to the point-wise order. But

$$\overline{P}\Big(J((x_0, u_0, 0)_1 \oplus (x_0 - u_0, x_0 - u_0, 1)) - J((x_0, v_0, 0)_1 \oplus (x_0 - v_0, x_0 - v_0, 1))\Big) \geq 0$$

also holds for any u_0 and any v_0. Indeed, the inequality is equivalent to

$$\sup\Big\{(g_0 - g_1)(u_0 - v_0): g_0 \in [0.0, 0.3], g_1 \in [0.1, 0.2], \epsilon \in [-0.1, 0.2]\Big\} \geq 0$$

which obviously holds for any choice of u_0 and v_0. Thus, all paths $(x_0, u_0, 0)_1 \oplus (x_0 - u_0, x_0 - u_0, 1)$ are optimal.

In conclusion, the information implies that we should invest all money x, but we cannot infer how we should divide x over the two companies.

By our dynamic programming approach we have have managed to solve this two-dimensional optimisation problem by reducing it to two one-dimensional ones, which are each very easy to solve. In the more general case of uncertain investment with n companies, we initially have a n-dimensional optimisation problem, and dynamic programming reduces this to n very simple one-dimensional optimisation problems.

5 Conclusion

The main conclusion of our work is that the method of dynamic programming can in principle be extended to deterministic systems with an uncertain gain, where the uncertainty about the gain is modelled by a coherent lower prevision, or by a set of linear previsions (probability measures).

But our general study of what conditions a generalised notion of optimality should satisfy for the Bellman approach to work is of some interest in itself too. In particular, besides an obvious extension of the well-known principle of optimality, another condition emerges that relates to the nature of the optimality operators *per se*: the optimality of a path should be invariant under the omission of non-optimal paths from the set of paths under consideration. If optimality is induced by a

strict partial ordering of paths, then this second condition is satisfied whenever the existence of dominating optimal paths for non-optimal ones is guaranteed.

Another important observation is that, in contradistinction to $\underline{P}$-maximality and $\mathcal{M}$-maximality, the dynamic programming method cannot be used to solve optimisation problems corresponding to $\underline{P}$-maximinity in general: for this notion the principle of optimality is not guaranteed to not hold, in particular when the external additivity property is not satisfied.

It is possible to refine our results by considering an additional equivalence relation on paths expressing some notion of *indifference*—relating, for instance, paths with the same expected gain. This allows us to partition a set $\mathrm{opt}_*(\mathcal{V})$ of optimal elements into equivalence classes of mutually indifferent paths. Any two paths in $\mathrm{opt}_*(\mathcal{V})$ that belong to different equivalence classes are necessarily *incomparable*: the available information, modelled through $\underline{P}$, does not allow us to choose between these two paths. A discussion of such matters presents no great conceptual difficulties, but has been omitted from the present paper due to limitations of space.

Throughout the paper we have assumed the system dynamics to be deterministic, that is, independent of ω. This greatly simplifies the discussion, still encompasses a large number of interesting applications, and does not suffer from the computational problems often encountered when dealing with non-deterministic dynamical systems—simply because in general the number of possible (random) paths tends to grow exponentially with the size of the state space $\mathcal{X}$. However, we should note that dropping this assumption still leads to a Bellman-type equation, connecting operators of optimality associated with *random* states $x \colon \Omega \to \mathcal{X}$. We intend to present our results about and views on this issue elsewhere.

References

[1] R. Bellman, Dynamic Programming, Princeton University Press, Princeton, 1957.

[2] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, London, 1991.

[3] M. Chevé, R. Congar, Optimal pollution control under imprecise environmental risk and irreversibility, Risk Decision and Policy 5 (2000) 151–164.

[4] L. V. Utkin, S. V. Gurov, Imprecise reliability for some new lifetime distribution classes, Journal of Statistical Planning and Inference 105 (1) (2002) 215–232.

[5] J. K. Satia, J. Roy E. Lave, Markovian decision processes with uncertain transition probabilities, Operations Research 21 (3) (1973) 728–740.

[6] C. C. White, H. K. Eldeib, Markov decision processes with imprecise transition probabilities, Operations Research 42 (4) (1994) 739–749.

[7] J. E. Smith, Generalized Chebychev inequalities: Theory and applications in decision analysis, Operations Research 43 (5) (1995) 807–825.

[8] R. Givan, S. Leach, T. Dean, Bounded-parameter Markov decision processes, Artificial Intelligence 122 (2000) 71–109.

[9] D. Harmanec, Generalizing Markov decision processes to imprecise probabilities, Journal of Statistical Planning and Inference 105 (1) (2002) 199–213.

[10] B. De Finetti, Theory of Probability: A Critical Introductory Treatment, Wiley, New York, 1974–5, two volumes.

[11] G. de Cooman, M. C. M. Troffaes, Lower previsions for unbounded random variables, technical report, in progress (2004).

[12] J. O. Berger, The robust Bayesian viewpoint, in: J. B. Kadane (Ed.), Robustness of Bayesian Analyses, Elsevier Science, Amsterdam, 1984.

[13] F. J. Giron, S. Rios, Quasi-Bayesian behaviour: A more realistic approach to decision making?, in: J. M. Bernardo, J. H. DeGroot, D. V. Lindley, A. F. M. Smith (Eds.), Bayesian Statistics, University Press, Valencia, 1980, pp. 17–38.

[14] I. Levi, The Enterprise of Knowledge. An Essay on Knowledge, Credal Probability, and Chance, MIT Press, Cambridge, 1983.

[15] A. Nilim, L. E. Ghaoui, Robustness in Markov decision problems with uncertain transition matrices, in: S. Thrun, L. Saul, B. Schölkopf (Eds.), Advances in Neural Information Processing Systems 16, MIT Press, Cambridge, MA, 2004.

[16] G. de Cooman, E. Miranda, A weak law of large numbers for coherent lower previsions, in: Proceedings of the Tenth International Conference IPMU 2004 (Information Processing and Management of Uncertainty in Knowledge-Based Systems, 4–9 July 2004, Perugia, Italy), Vol. 1, Editrice Università La Sapienza, Rome, Italy, 2004, pp. 451–458.

[17] I. Couso, S. Moral, P. Walley, Examples of independence for imprecise probabilities, Risk Decision and Policy 5 (2000) 165–181.

[18] G. de Cooman, M. C. M. Troffaes, Coherent lower previsions in systems modelling: products and aggregation rules, Reliability Engineering and System Safety 85 (2004) 113–134.

[19] V. P. Kuznetsov, Interval Statistical Models, Radio i Svyaz Publ., Moscow, 1991, in Russian.

[20] F. G. Cozman, Computing lower expectations with Kuznetsov's independence condition, in: J.-M. Bernard, T. Seidenfeld, M. Zaffalon (Eds.), ISIPTA '03 – Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications, Carleton Scientific, 2003, pp. 117–186.