Dynamic Programming for Discrete-Time Systems with Uncertain Gain*

G. DE COOMAN
Ghent University, Belgium

M. C. M. TROFFAES
Ghent University, 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.

Keywords

optimal control, dynamic programming, uncertainty, imprecise 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 effective method for solving optimal control problems for discrete-time systems is the recursive *dynamic programming* method, introduced by Richard Bellman [1].

To explain the ideas behind this method, 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 associated gains by $J_{\alpha\beta}$, $J_{\alpha\gamma}$ and $J_{\alpha\delta}$ respectively. Assume that path $\alpha\gamma$ is optimal: $J_{\alpha\gamma} > J_{\alpha\beta}$ and $J_{\alpha\gamma} > J_{\alpha\delta}$. Then it follows that path γ is the optimal way to go from b to c. To

*This paper presents research results of project G.0139.01 of the Fund for Scientific Research, Flanders (Belgium), and of the Belgian Programme on Interuniversity Poles of Attraction initiated by the Belgian state, Prime Minister's Office for Science, Technology and Culture.

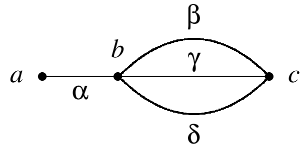

Figure 1: Principle of Optimality

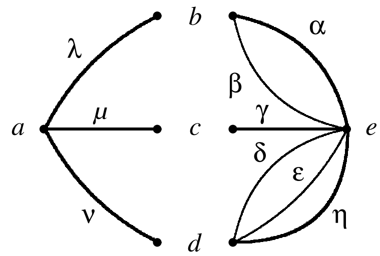

Figure 2: Dynamic Programming

see this, observe that $J_{\alpha\nu} = J_\alpha + J_\nu$ for $\nu \in \{\beta, \gamma, \delta\}$ (gains are assumed to be additive) 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 $\nu\varepsilon$ 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 very 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 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.

It has however been argued by various scholars (see [11, 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 to identify a single probability measure. In those cases, it is more appropriate to model the available information through an *imprecise* probability model, e.g., by a lower prevision, or by a set of probability measures. For applications of this approach, see for instance [4, 10].

Two questions now arise naturally. First, 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. To arrive at this, we study the properties that an abstract notion of optimality should satisfy for the Bellman approach to work.

We recognise that other authors (see for instance [8]) have extended the dynamic programming algorithm to systems with imprecise gain and/or imprecise dynamics. However in doing so, none of them seems to have questioned in what sense their generalised dynamic programming method leads to optimal paths. In this article we approach the problem from the opposite, and in our opinion, more logical side: one should *first* define a notion optimality and investigate whether the dynamic programming argument holds for this notion of optimality, instead of blindly "generalising" Bellman's algorithm. In the remainder of this section, we introduce the basic terminology and notation that will allow us to give a precise formulation of the problems under study. We have omitted proofs of technical results that do not contribute to a better understanding of the main ideas.

1.1 Preliminaries

1.1.1 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 through 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 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.

1.1.2 Paths

A *path* is a triple $(x,k,u_\cdot)$, where $x \in \mathcal{X}$ is a state, $k \in [0,N]$ a time instant, and $u_\cdot\colon [k,N-1] \to \mathcal{U}$ a sequence of controls. A path fixes a unique state trajectory $x_\cdot\colon [k,N] \to \mathcal{X}$, which is defined recursively through $x_k = x$ and $x_{i+1} = f(x_i,u_i,i)$ for every $i \in [k,N-1]$. 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 $\emptyset$ to $\mathcal{U}$ by $u_\emptyset$. If $k = N$, the control $u.$ does nothing: it is equal to $u_\emptyset$.

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 $(x,k,u.)_\ell \oplus (y,\ell,v.)_m$, and 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 $(x,k,u.)_\ell \oplus \mathcal{V}$ for the set

$$\{(x,k,u.)_\ell \oplus (x_\ell,\ell,v.)\colon (x_\ell,\ell,v.) \in \mathcal{V}\}.$$

1.1.3 The Gain Function

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, used to model uncertainty of the gains. 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.

Assuming gain additivity, we can also associate a gain with a path $(x,k,u.)$:

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

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

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

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 the time instant 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 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 [6]. We assume that the reader is familiar with lower previsions and coherence (see [11] 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 called the *gain gamble* associated with $(x,k,u.)$.[1] 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 $k \leq \ell \leq m \leq N$, 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 $k \in [0,N]$:

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

For fixed $k \in [0,N-1]$ and $x \in \mathcal{X}_k$, the gain $J(x,k,u.,\omega)$ can also be interpreted as a map from $\mathcal{U}(x,k)$ to $\mathcal{L}(\Omega)$; this map is denoted by $J(x,k)$.

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 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 X* is strictly preferred to *Y (with respect to $\underline{P}$), and write $X >_{\underline{P}} Y$, if*

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

Indeed, if $X \geq Y$ and $X \neq Y$, then the subject should be willing to exchange Y for X, since this transaction can only improve his gain. On the other hand,

[1] To simplify the discussion, we assume this map is bounded.

$\underline{P}(X-Y)>0$ expresses that the subject is willing to pay a 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 lower prevision $\underline{P}$ to express a strict preference between any two *paths* $(x,k,u.)$ and $(x,k,v.)$, based on their gains: if $J(x,k,u.)>_{\underline{P}} J(x,k,v.)$ this means that the uncertain gain $J(x,k,u.)$ is strictly preferred to the uncertain gain $J(x,k,v.)$. We then say that the path $(x,k,u.)$ is strictly preferred to $(x,k,v.)$, and we use the notation $(x,k,u.)>_{\underline{P}}(x,k,v.)$.

$>_{\underline{P}}$ is anti-reflexive and transitive, and therefore 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: any 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 that are not dominated by any other path. Observe that a maximal gamble X in a set $\mathcal{K}$ with respect to $>_{\underline{P}}$ is 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 just the point-wise undominated gambles whose prevision is maximal; they maximise expected gain.

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., $(x,k,u.)\not>_{\underline{P}}(x,k,u^*_\cdot)$ for all $(x,k,u.)\in\mathcal{V}$. We denote the set of the $\underline{P}$-maximal paths in $\mathcal{V}$ by* $\operatorname{opt}_{>_{\underline{P}}}(\mathcal{V})$. *The operator* $\operatorname{opt}_{>_{\underline{P}}}$ *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 just 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 $\operatorname{opt}_{>_{\underline{P}}}(\mathcal{J}(x,k))$, then for all $(x,k,u.)\in\mathcal{U}(x,k)$:

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

$\underline{P}$-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 elements in $\mathcal{J}(x,k)$ (and hence in $\mathcal{U}(x,k)$) is that $\mathcal{J}(x,k)$ should be compact[2] (and of course non-empty). This follows from a general result mentioned in [11, Section 3.9.2]. In fact, Theorem 1 is a stronger result, whose Corollary 1 turns out to be very important in proving that the dynamic programming approach works for $\underline{P}$-maximality (see Section 3.2). Its proof is based on Zorn's lemma.

Theorem 1 *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 $Y>_{\underline{P}}X$.*

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

Corollary 1 *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 another optimality criterion that can be associated with a lower prevision $\underline{P}$. We can use $\underline{P}$ to define another strict order on $\mathcal{L}(\Omega)$:

Definition 3 *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 4 *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 1 *$\underline{P}$-maximinity implies $\underline{P}$-maximality. For a linear prevision P, P-maximinity is equivalent to P-maximality.*

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 1 and Corollary 1 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 [2, 7].

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 leads to a 'robustified' strict partial order $>_{\mathcal{M}}$ on $\mathcal{L}(\Omega)$.

Definition 5 *$X >_{\mathcal{M}} Y$ 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 $\underline{P}(X) = \inf\left\{P(X)\colon P \in \mathcal{M}\right\}$ for all $X \in \mathcal{L}(\Omega)$.[3] Conversely, given a coherent lower prevision $\underline{P}$, the strict partial orders $>_{\mathcal{M}(\underline{P})}$ and $>_{\underline{P}}$ are identical, where

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

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

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 [9, 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 [11, Section 3.9]. We are therefore led to consider a third notion of optimality:

Definition 6 *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 $\mathcal{M}$, or in other words 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 $\mathcal{V} \subseteq \mathcal{U}(x,k)$:

$$\mathrm{opt}_{\mathcal{M}}(\mathcal{V}) = \bigcup_{P \in \mathcal{M}} \mathrm{opt}_{>_P}(\mathcal{V}). \tag{1}$$

3 Dynamic Programming

3.1 A General Notion of Optimality

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

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

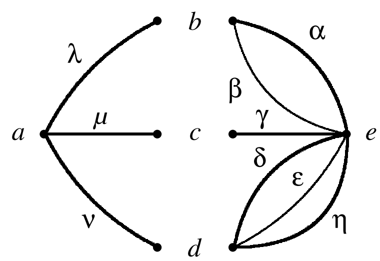

Figure 3: A More General Type of Dynamic Programming

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 $\mathrm{opt}_*(\{\lambda\alpha, \lambda\beta, \mu\gamma, \nu\delta, \nu\varepsilon, \nu\eta\})$, 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 that the concatenations $\lambda\beta$ and $\nu\varepsilon$ with the non-$*$-optimal paths β and ν 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 $\nu\varepsilon$ cannot be $*$-optimal because then β and ε would be $*$-optimal as well. Definition 8 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 $\nu\varepsilon$ from the set of candidate $*$-optimal paths should not have any effect on the actual $*$-optimal paths: we need that

$$\mathrm{opt}_*(\{\lambda\alpha, \lambda\beta, \mu\gamma, \nu\delta, \nu\varepsilon, \nu\eta\}) = \mathrm{opt}_*(\{\lambda\alpha, \mu\gamma, \nu\delta, \nu\eta\}).$$

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 $\{\lambda\alpha, \lambda\beta, \mu\gamma, \nu\delta, \nu\varepsilon, \nu\eta\}$, in the following sense.

Definition 7 *Consider a set $S \neq \emptyset$ and an* optimality operator opt_* *defined on the set $\wp(S)$ of subsets of S such that* $\mathrm{opt}_*(T) \subseteq T$ *for all $T \subseteq S$. Elements of* $\mathrm{opt}_*(T)$ *are called* $*$-optimal in T. opt_* *is called* insensitive to the omission of non-$*$-optimal elements from S *if* $\mathrm{opt}_*(S) = \mathrm{opt}_*(T)$ *for all T such that* $\mathrm{opt}_*(S) \subseteq T \subseteq S$.

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 path.

Proposition 2 *Let S be a non-empty set provided with a family of strict partial orders $>_j$, $j \in J$. Define for $T \subseteq S$, $\operatorname{opt}_{>_j}(T) = \{a \in T : (\forall b \in T)(b \not>_j a)\}$ as the set of maximal elements of T with respect to $>_j$, and let $\operatorname{opt}_J(T) = \bigcup_{j \in J} \operatorname{opt}_{>_j}(T)$. Then $\operatorname{opt}_{>_j}$, $j \in J$ and opt_J are optimality operators. If for some $j \in J$,*

$$(\forall a \in S \setminus \operatorname{opt}_{>_j}(S))(\exists b \in \operatorname{opt}_{>_j}(S))(b >_j a), \tag{2}$$

then $\operatorname{opt}_{>_j}$ is insensitive to omission of non-$>_j$-optimal elements from S. If (2) holds for all $j \in J$, then opt_J is insensitive to omission of non-J-optimal elements from S.

Proof. Consider j in J, and assume that (2) holds for this j. Let $\operatorname{opt}_{>_j}(S) \subseteq T \subseteq S$, then we must prove that $\operatorname{opt}_{>_j}(S) = \operatorname{opt}_{>_j}(T)$. First of all, if $a \in \operatorname{opt}_{>_j}(S)$ then $b \not>_j a$ for all b in S, and *a fortiori* for all b in T, so that $a \in \operatorname{opt}_{>_j}(T)$. Consequently, $\operatorname{opt}_{>_j}(S) \subseteq \operatorname{opt}_{>_j}(T)$. Conversely, let $a \in \operatorname{opt}_{>_j}(T)$ and assume *ex absurdo* that $a \notin \operatorname{opt}_{>_j}(S)$. It then follows from (2) that there is some c in $\operatorname{opt}_{>_j}(S)$ and therefore in T such that $c >_j a$, which contradicts $a \in \operatorname{opt}_{>_j}(T)$.

Next, assume that (2) holds for all $j \in J$. Let $\operatorname{opt}_J(S) \subseteq T \subseteq S$, then we must prove that $\operatorname{opt}_J(S) = \operatorname{opt}_J(T)$. Consider any $j \in J$, then $\operatorname{opt}_{>_j}(S) \subseteq \operatorname{opt}_J(S) \subseteq T \subseteq S$, so we may infer from the first part of the proof that $\operatorname{opt}_{>_j}(S) = \operatorname{opt}_{>_j}(T)$. By taking the union over all $j \in J$, we find that indeed $\operatorname{opt}_J(S) = \operatorname{opt}_J(T)$. □

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 $\mathcal{V}$, $\operatorname{opt}_*(\mathcal{V})$ 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 8 (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 $(x,k,u_\cdot)$ in $\mathcal{U}(x,k)$ that if $(x,k,u_\cdot)$ is $*$-optimal in $\mathcal{U}(x,k)$, then $(x_\ell,\ell,u_\cdot)$ is $*$-optimal in $\mathcal{U}(x_\ell,\ell)$.*

This may also be expressed as:

$$\operatorname{opt}_*(\mathcal{U}(x,k)) \subseteq \bigcup_{(x,k,u_\cdot)_\ell \in \mathcal{U}(x,k)_\ell} (x,k,u_\cdot)_\ell \oplus \operatorname{opt}_*(\mathcal{U}(x_\ell,\ell)).$$

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 2 (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 $\ell \in [k,N]$:*

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

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

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

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

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

3.2 $\underline{P}$-Maximality

Let us now apply these general results to the specific types of optimality introduced before. We first consider the optimality operator $\mathrm{opt}_{>_{\underline{P}}}$ 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 $>_{\underline{P}}$. The following lemma roughly states that the preference amongst paths with respect to $>_{\underline{P}}$ 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 1 *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 3 (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 $\ell \in [k,N]$.*

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 1 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 1 and Proposition 2, we see that if $\mathcal{J}(x,k)$ is compact, then the optimality operator $\mathrm{opt}_{>_{\underline{P}}}$ associated with $\mathcal{U}(x,k)$ is insensitive to the omission of non-$>_{\underline{P}}$-optimal elements. Together with Proposition 3 and Theorem 2, this allows us to infer a Bellman equation for $\underline{P}$-maximality.

Corollary 2 *Let $k \in [0,N]$ and $x \in \mathcal{X}_k$. If $\mathcal{J}(x,k)$ is compact, then for all $\ell \in [k,N]$*

$$\mathrm{opt}_{>_{\underline{P}}}(\mathcal{U}(x,k)) = \mathrm{opt}_{>_{\underline{P}}} \bigcup_{(x,k,u)_\ell \in \mathcal{U}(x,k)_\ell} (x,k,u)_\ell \oplus \mathrm{opt}_{>_{\underline{P}}}(\mathcal{U}(x_\ell,\ell)), \tag{3}$$

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 2 results in a procedure to calculate all $\underline{P}$-maximal paths. Indeed, $\mathrm{opt}_{>_{\underline{P}}}(\mathcal{U}(x,N)) = \{u_\emptyset\}$ for every $x \in \mathcal{X}_N$, and $\mathrm{opt}_{>_{\underline{P}}}(\mathcal{U}(x,k))$ can be calculated recursively through Eq. (3). It also provides a method for constructing a $\underline{P}$-maximal feedback: for every $x \in \mathcal{X}_k$, choose any $(x,k,u^*_\cdot(x,k)) \in \mathrm{opt}_{>_{\underline{P}}}(\mathcal{U}(x,k))$. Then $\phi(x,k) = u^*_k(x,k)$ realises a $\underline{P}$-maximal feedback.

3.3 $\mathcal{M}$-Maximality

We now turn to the optimality operator $\mathrm{opt}_{\mathcal{M}}$, satisfying (1). By Proposition 2 and (1), it follows 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 3, $\mathrm{opt}_{\mathcal{M}}$ 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 2 and the discussion at the end of Section 2.2

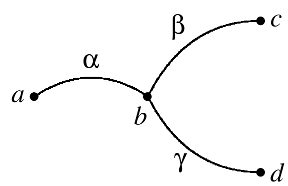

Figure 4: A Counterexample

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 dynamic programming may not work here—not even with a vacuous uncertainty model. 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 gain, lower expected gains are not additive.

Example 1 *Consider the dynamical system depicted in Figure 4. 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* $J(\alpha) = \langle 2,0\rangle$*,* $J(\beta) = \langle 0,-1\rangle$ *and* $J(\gamma) = \langle -2,0\rangle$ *(there is zero gain associated with the final state). Then* $\alpha\beta \not\sqsupset_{\underline{P}} \alpha\gamma$*: indeed,* $\langle 2,-1\rangle$ *does not dominate* $\langle 0,0\rangle$ *point-wise, and* $\inf\langle 2,-1\rangle \not> \inf\langle 0,0\rangle$ *or equivalently* $\langle 0,0\rangle$ *maximises the worst expected gain. Hence, we find that* $\alpha\gamma$ *is* $\underline{P}$*-maximin. But* $\beta \sqsupset_{\underline{P}} \gamma$*: indeed,* $\inf\langle 0,-1\rangle > \inf\langle 0,-2\rangle$ *which means that* γ *is not* $\underline{P}$*-maximin. Thus the "principle of* $\underline{P}$*-maximin optimality" does not hold here.*

3.5 Yet Another Type of Optimality

We end this discussion with another type of optimality associated with a strict partial order, introduced by Harmanec in [8, Definition 3.4]. In our setting (precisely known system dynamics), its definition basically reduces to

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

It can be shown easily that if $\mathcal{J}(x,k)$ is compact, the optimality operator induced by $>^{\star}_{\underline{P}}$ for $\mathcal{U}(x,k)$ is insensitive to the omission of non-$>^{\star}_{\underline{P}}$-optimal paths from $\mathcal{U}(x,k)$. But, as the following counterexample shows, we cannot guarantee that the principle of optimality holds for $>^{\star}_{\underline{P}}$-optimality, and therefore the dynamic programming approach may not work here—not even with a vacuous uncertainty model. Again, this is because the partial order $\sqsupset_{\underline{P}}$ is not compatible with gain additivity. It also indicates that the solution of the Bellman-type equation advocated in [8] will not necessarily lead to optimal paths, in the sense we described above.

Example 2 *Consider the dynamical system depicted in Figure 4. 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, $\inf\langle 2,0\rangle \not> \sup\langle 1,-1\rangle$. 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. Thus the "principle of $>^\star_{\underline{P}}$-maximal optimality" does not hold for this example.

4 Conclusion

The main conclusion of our work is that the method of dynamic programming can be extended to systems with imprecise gain. 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: for this notion the principle of optimality does not hold in general.

Throughout the paper we 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 issues 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 $x\colon \Omega \to \mathcal{X}$. A discussion of these matters has been omitted from the present paper due to limitations of space.

Acknowledgements

This paper presents research results of project G.0139.01 of the Fund for Scientific Research, Flanders (Belgium), and of the Belgian Programme on Interuniversity Poles of Attraction initiated by the Belgian state, Prime Minister's Office for Science, Technology & Culture. The scientific responsibility rests with the authors.

References

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

[2] BERGER, J. O. The robust Bayesian viewpoint. In *Robustness of Bayesian Analyses*, J. B. Kadane, Ed. Elsevier Science, Amsterdam, 1984.

[3] BERNARDO, J. M., H., D. J., V., L. D., AND SMITH, A. F. M., Eds. *Bayesian Statistics*. University Press, Valencia, 1980.

[4] CHEVÉ, M., AND CONGAR, R. Optimal pollution control under imprecise environmental risk and irreversibility. *Risk Decision and Policy 5* (2000), 151–164.

[5] DE COOMAN, G., COZMAN, F. G., MORAL, S., AND WALLEY, P., Eds. *ISIPTA '99 – Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications* (Ghent, 1999), Imprecise Probabilities Project.

[6] DE FINETTI, B. *Theory of Probability: a Critical Introductory Treatment*. Wiley, London, 1975.

[7] GIRON, F. J., AND RIOS, S. Quasi-Bayesian behaviour: A more realistic approach to decision making? In Bernardo et al. [3], pp. 17–38.

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

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

[10] UTKIN, L. V., AND GUROV, S. V. Imprecise reliability models for the general lifetime distribution classes. In de Cooman et al. [5], pp. 333–342.

[11] WALLEY, P. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

Gert de Cooman is a member of the SYSTeMS research group of Ghent University.
Address: Technologiepark – Zwijnaarde 914, B-9052, Zwijnaarde, Belgium.
E-mail: gert.decooman@ugent.be

Matthias Troffaes is a member of the SYSTeMS research group of Ghent University.
Address: Technologiepark – Zwijnaarde 914, B-9052, Zwijnaarde, Belgium.
E-mail: matthias.troffaes@ugent.be