Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Approximate Reasoning
39 (2005) 257–278
INTERNATIONAL JOURNAL OF APPROXIMATE REASONING
www.elsevier.com/locate/ijar

Dynamic programming for deterministic discrete-time systems with uncertain gain

Gert de Cooman *, Matthias C.M. Troffaes

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

Received 1 February 2004; accepted 1 October 2004
Available online 26 November 2004

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; 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

* Corresponding author. Tel.: +329 264 56 53; fax: +329 264 58 40.
E-mail address: gert.decooman@rug.ac.be (G. de Cooman).

0888-613X/$ - see front matter © 2004 Published by Elsevier Inc.
doi:10.1016/j.ijar.2004.10.004

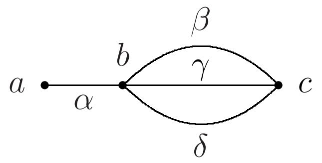

Fig. 1. Principle of optimality.

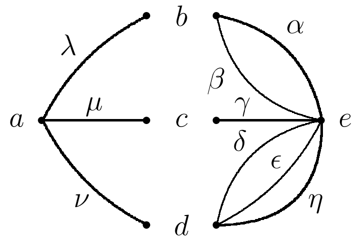

Fig. 2. Dynamic programming.

solving optimal control problems for discrete-time systems is the recursive *dynamic programming* technique, introduced by Bellman [1].

To explain the ideas behind it, we refer to Figs. 1 and 2. In Fig. 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 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 Fig. 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