The Annals of Mathematical Statistics
1970, Vol. 41 No. 1, 78–86

DISCRETE-TIME MARKOVIAN DECISION PROCESSES WITH INCOMPLETE STATE OBSERVATION

BY YOSHIKAZU SAWARAGI AND TSUNEO YOSHIKAWA

Kyoto University

1. Introduction. Discrete-time Markovian decision processes (MDP's) with complete state observation and an infinite planning horizon, have been investigated by many authors (for example [4], [5], [9], [10]).

MDP's with incomplete state observation have also been investigated by several authors [1], [2], and [7]. Dynkin [7] has treated a very general discrete-time problem which includes MDP's with and without complete state observation as special cases. However, the relation between [7] and [5], [10] is not clear. Åström [2] and Aoki [1] have treated the case of a finite planning horizon (control interval).

In this paper it is shown that MDP's with incomplete state observation, countable possible states, uncountable available actions and an infinite planning horizon, can be transformed to MDP's with complete state observations and uncountable possible states. The latter MDP's are those which have been treated in [5] and [10]. The states of the latter MDP's are the probability distributions on the set of the states of the former. Similar transformations have been pointed out by several authors [1]–[3]. However, the above transformation should be formulated explicitly.

2. Probabilistic definitions and notation. In this section we develop the probabilistic notation to be used throughout the paper. We follow [5] and [10] as closely as possible.

A *Borel set* X is a Borel subset of a complete separable metric space. For any Borel set $X, \mathscr{B}(X)$ denotes the σ-field of Borel subsets of X. *Measurable* means measurable with respect to $\mathscr{B}(X)$. A *probability* on a non-empty Borel set X is a probability measure defined on $\mathscr{B}(X)$, and the set of all probabilities on X is denoted by $P(X)$. If X and Y are non-empty Borel sets, a *conditional probability* on Y given X is a function $q(\cdot \mid \cdot)$ such that for each $x \in X$, $q(\cdot \mid x)$ is a probability on Y and for each $E \in \mathscr{B}(Y)$, $q(E \mid \cdot)$ is a Baire function on X. The set of all conditional probabilities on Y given X is denoted by $Q(Y \mid X)$. $p \in P(X)$ and $q \in Q(Y \mid X)$ are also denoted by $p([x])$ and $q([y] \mid x)$ respectively, using the coordinate variables x, y of X, Y, in order to indicate explicitly the spaces where these probabilities are defined. We denote the Cartesian product of X and Y by XY. Every probability $p \in P(XY)$ has a factorization $p = p'q$, where $p' \in P(X)$ is the marginal distribution of the first coordinate variable under p, and $q \in Q(Y \mid X)$ is a version of the conditional distribution of the second coordinate variable given the first. $F(X)$ denotes either the set of all bounded Baire functions on X or the set of all non-positive, extended real-valued Baire functions on X. Unless otherwise noted, statements made about elements of $F(X)$ are valid for either definition.

Received December 2, 1968.