1. Introduction

A dynamic programming problem is specified by four objects: $S$, $A$, $q$, $r$, where $S$ is a nonempty Borel set, the set of states of some system, $A$ is a nonempty Borel set, the set of acts available to you, $q$ is the law of motion of the system; it associates (Borel measurably) with each pair $(s, a)$ a probability distribution $q(\cdot | s, a)$ on $S$: when the system is in state $s$ and you choose act $a$, the system moves to a new state selected according to $q(\cdot | s, a)$, and $r$ is a bounded Borel measurable function on $S \times A \times S$, the immediate return: when the system is in state $s$, and you choose act $a$, and the system moves to $s'$, you receive an income $r(s, a, s')$. A plan $\pi$ is a sequence $\pi_1, \pi_2, \ldots$, where $\pi_n$ tells you how to select an act on the $n$-th day, as a function of the previous history $h = (s_1, a_1, \ldots, a_{n-1}, s_n)$ of the system, by associating with each $h$ (Borel measurably) a probability distribution $\pi_n(\cdot | h)$ on (the Borel subsets of) $A$.

Any sequence of Borel measurable functions $f_1, f_2, \ldots$, each mapping $S$ into $A$, defines a plan. When in state $s$ on the $n$-th day, choose act $f_n(s)$. Plans $\pi = \{f_n\}$ of this type may be called Markov plans. A single $f$ defines a still more special kind of plan: whenever in state $s$, choose act $f(s)$. This plan is denoted by $f^{(\infty)}$, and plans $f^{(\infty)}$ are called stationary.

A plan $\pi$ associates with each initial state $s$ a corresponding expected $n$-th period return $r_n(\pi)(s)$ and an expected discounted total return

$$I_\beta(\pi)(s) = \sum_{n=1}^{\infty} \beta^{n-1} r_n(\pi)(s),$$

where $\beta$ is a fixed discount factor, $0 \leq \beta < 1$.

The problem of finding a $\pi$ which maximizes $I_\beta$ was studied in [1]. Three of the principal results obtained were the following.

**RESULT (i).** For any probability distribution $p$ on $S$ and any $\epsilon > 0$, there is a stationary plan $f^{(\infty)}$ which is $(p, \epsilon)$-optimal; that is,

$$p\{I_\beta(f^{(\infty)}) > I_\beta(\pi) - \epsilon\} = 1$$

for all $\pi$.

**RESULT (ii).** Any bounded $u$ which satisfies

$$u(s) \geq \int [r(s, a, \cdot) + \beta u(\cdot)] \, dq(\cdot | s, a)$$

for all $s, a$ is an upper bound on incomes;

This research was supported by the Information Systems Branch of the Office of Naval Research under Contract Nonr-222(53).