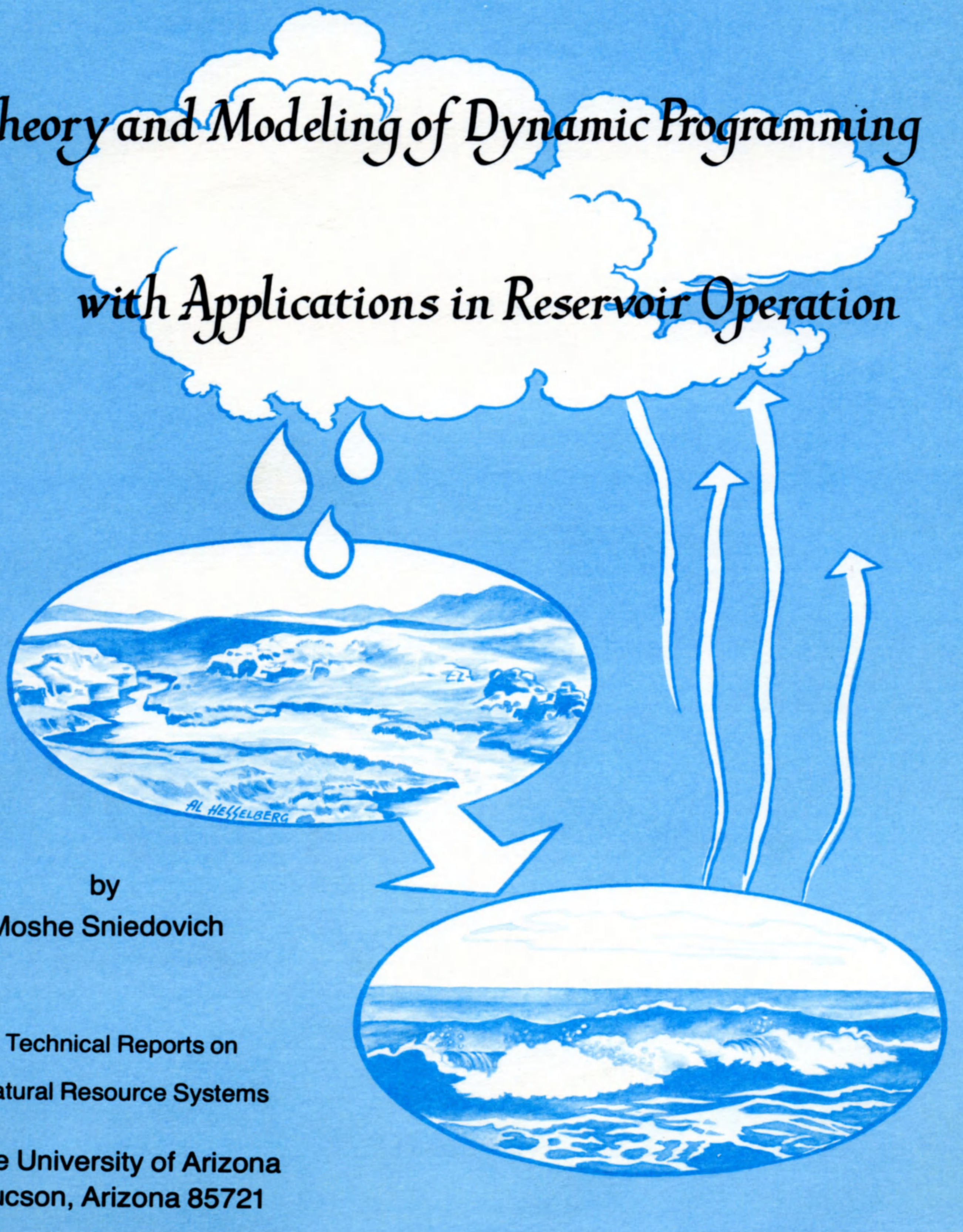
Technical Report No. 27

December, 1976

On the Theory and Modeling of Dynamic Programming with Applications in Reservoir Operation



by

Moshe Sniedovich

Technical Reports on

Natural Resource Systems

The University of Arizona
Tucson, Arizona 85721

ON THE THEORY AND MODELING OF DYNAMIC PROGRAMMING WITH APPLICATIONS IN RESERVOIR OPERATION

by

Moshe Sniedovich

Reports on Natural Resource Systems
No. 27

December 1976

Collaborative effort between the
following departments:

Hydrology and Water Resources
Systems and Industrial Engineering

University of Arizona
Tucson, Arizona 85721

PREFACE

This report constitutes the doctoral dissertation of the same title completed by the author in May, 1976, and accepted by the Faculty of the Department of Hydrology and Water Resources.

The investigation presented in this paper was conducted under the direction of Sidney J. Yakowitz, Professor of Systems and Industrial Engineering.

The research effort was supported in part by funds provided by the National Science Foundation through a grant (GK-35915) on "Space Time Sampling and Equations of Hydrologic Systems" and in part by funds provided by the Office of Water Resources Research through a grant (14-31-0001-5056) on "Practical Use of Decision Theory to Assess Uncertainties about Actions Affecting the Environment."

This report series constitutes an effort to communicate to practitioners and researchers the complete research results, including computer programs and more detailed theoretical developments, that cannot be reproduced in professional journals. These reports are not intended to serve as a substitute for the review and referee process exerted by the scientific and professional community in their journals.

ACKNOWLEDGMENTS

I would like to express my appreciation to Professors L. Duckstein, M. Fogel, D. Davis, and D. Dietrich for their help, advice, and guidance. I would like to extend special thanks to Professor Yakowitz, my dissertation advisor for his effort in guiding me through the avenues of dynamic programming.

I greatly appreciate the support of the Hydrology and Water Resources Department, especially for enabling me to work in the field of operations research.

I would like also to express my thanks to Mrs. Paula Tripp for her typing.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

TABLE OF CONTENTS--Continued

LIST OF TABLES

ABSTRACT

This dissertation contains a discussion concerning the validity of the principle of optimality and the dynamic programming algorithm in the context of discrete time and state multistage decision processes. The multistage decision model developed for the purpose of the investigation is of a general structure, especially as far as the reward function is concerned. The validity of the dynamic programming algorithm as a solution method is investigated and results are obtained for a rather wide class of decision processes. The intimate relationship between the principle and the algorithm is investigated and certain important conclusions are derived.

In addition to the theoretical considerations involved in the implementation of the dynamic programming algorithm, some modeling and computational aspects are also investigated. It is demonstrated that the multistage decision model and the dynamic programming algorithm as defined in this study provide a solid framework for handling a wide class of multistage decision processes.

The flexibility of the dynamic programming algorithm as a solution procedure for nonroutine reservoir control problems is demonstrated by two examples, one of which is a reliability problem.

To the best of the author's knowledge, many of the theoretical derivations presented in this study, especially those concerning the relation between the principle of optimality and the dynamic programming algorithm, are novel.

CHAPTER 1

INTRODUCTION

The dynamic programming algorithm and the principle of optimality as introduced by Bellman in the early 1950's have since been the subject of a continuous research effort, especially as related to stochastic processes. After the pioneering work of Bellman, Howard and others, certain fundamental questions concerning their validity have been raised. It was realized (Karlin, 1955) that any meaningful discussion on these subjects should be conducted in the context of the decision process under consideration. The result was that multistage decision processes started to be classified according to certain properties of the elements of the process such as state spaces, decision sets, reward functions, etc. This study is restricted to processes in which the state space(s) and the set of decision stages are countable, and which are often referred to as discrete processes.

The investigation will concentrate on two subjects: first, the development of a multistage decision model and second, the validity of the principle and the algorithm in the context of the model developed.

In Chapter 2, the multistage decision model is developed and some of its basic properties are analyzed. The concept of sufficient statistics is used to show how the dimensionality of the original (complete) problem may be reduced without affecting the optimality of the rewards.

Chapter 3 includes the formal definitions of the principle of optimality and the dynamic programming algorithm and an extensive study concerning their validity. An important conclusion concerning the validity of the principle of optimality is derived. An attempt is made to clarify the ambiguity concerning the relation between the principle and the algorithm. It is shown that indeed the principle of optimality and the dynamic programming algorithm are intimately related and this relation is specified. The notions of the principle of optimality and the dynamic programming algorithm as introduced in this study are compared with others and their generality is emphasized.

Chapter 4 is devoted to the investigation of a potential method of reducing the computational load often encountered when implementing the algorithm. By means of two simple reservoir control problems, it is demonstrated that analytical considerations may be extremely effective in reducing the computational load.

The study is concluded by investigating the elements of the model from a modeling viewpoint. Two nonroutine reservoir control problems are introduced and it is demonstrated how the model may be used to solve them, using the dynamic programming algorithm. In contrast to certain comments made recently in the hydrologic literature, it is demonstrated how the dynamic programming algorithm can be used to handle probabilistic constraints.

Each of the above chapters concludes with a discussion in which the contribution of this study to the state-of-the-art is specified.

The reference material includes a list of most of the symbols used in Chapter 2 and Chapter 3 (Appendix A). The computer program used to solve the "range" problem introduced in Chapter 5 is presented in Appendix B.

CHAPTER 2

THE MULTISTAGE DECISION MODEL

The multistage decision model introduced in this chapter is developed and formulated so as to provide a convenient mathematical framework for (a) investigating the properties of the optimal decisions, and consequently (b) the construction of a solution procedure. In other words, the model is designed for analytical purposes.

The structure of the model is determined by a sequence of definitions. In this chapter, neither the motivation for choosing the specific definitions nor their physical interpretation are elaborated; this will be done in Chapter 5 where the modeling aspects of the multistage decision process are discussed.

2.1. Mathematical Formulation of the Complete Multistage Decision Model

The model developed in this chapter is a modified version of Hinderer's (1970, pp. 5-47) model. The elements of the model will be first defined followed by a formal definition of the model itself.

Definitions

Definition 2.1. The set $\mathbb{N}$ of *decision stages* is the set of positive integers. More specifically: $\mathbb{N} = \{n: \; n = 1, 2, \ldots\}$.

The set $\mathbb{N}$ identifies the stages in which the decision maker is allowed to make decisions and consequently to implement the corresponding actions. It should be noted that $\mathbb{N}$ consists of countably many elements.

Definition 2.2. The *state space*, Ω_n, associated with the n*th* decision stage is a nonempty countable set containing the elements x_n called *states*. The set $\bar{\Omega} = \bigcup_{n\in\mathbb{N}} \Omega_n$ is called the *universe* and the set $\mathfrak{R} = \{\Omega_n: n\in\mathbb{N}\}$ the *set of state spaces*.ȷ

Definition 2.3. A *trajectory*, $\bar{x}_n$, associated with the n*th* decision stage is a sequence of states. More specifically,

$$\bar{x}_n = (x_1, x_2, \ldots, x_n), \quad x_i\varepsilon\Omega_i, \; i = 1, 2, \ldots, n.$$

The set of all the trajectories associated with the n*th* decision stage will be denoted by $\bar{X}_n$, i.e.:

$$\bar{X}_n = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n, \; n\varepsilon\mathbb{N}$$

and

$$\bar{X}_\infty = \Omega_1 \times \Omega_2 \times \ldots$$

The i*th* coordinate of $\bar{x}_n$ will be denoted by $\bar{x}_n(i)$, i.e., $\bar{x}_n(i) \; \varepsilon\Omega_i$, $i \leq n$.ȷ

Definition 2.4. The *decision set*, $\mathbb{D}$, is a nonempty set containing the elements d called *decisions*.ȷ

The set $\mathbb{D}$ contains all the decisions available to the decision maker. However, not all the elements of $\mathbb{D}$ are available at a given decision stage. Moreover, at any decision stage the decisions available to the decision maker may depend either on the previous and current states and/or the previous decisions already made. In other words, the set of alternative decisions available to the decision maker at the n*th* stage may depend on the history of the process as far as states and decisions are concerned.

Definition 2.5. The *history space*, $\bar{H}_n$, associated with the n*th* decision stage is the set given by:

$$\bar{H}_n = \Omega_1 \times \mathbb{D} \times \Omega_2 \times \ldots \times \mathbb{D} \times \Omega_n, \; n \in \mathbb{N}$$

with

$$\bar{H}_\infty = \Omega_1 \times \mathbb{D} \times \Omega_2 \times \mathbb{D} \times \ldots$$

The elements h_n of $\bar{H}_n$ are called *histories.*

Definition 2.6. The sequence $D = \{D_n\}_{n \in \mathbb{N}}$ of *admissible decision maps* is a sequence of maps from certain sets $H_n \subset \bar{H}_n$ to the set of all nonempty subsets of $\mathbb{D}$ with the property that

$$H_1 = \Omega_1$$

$$H_{n+1} = \{(h,d,x)\colon h \in H_n, \; d \in D_n(h), \; x \in \Omega_{n+1}\}$$

D_n is called the *admissible decision map* associated with the n*th* decision stage and $D_n(h_n)$ the *set of admissible decisions* at (h_n, n), whereas H_n is called the *set of admissible histories* at the n*th* decision stage. Let W_n be defined as:

$$W_n = \{(h_n,d)\colon \; h_n \in H_n, \; d \in D_n(h_n)\}, \; n \in \mathbb{N}$$

H_{n+1} can be written then as

$$H_{n+1} = W_n \times \Omega_{n+1}.$$

The dynamics of the process is assumed to be a statistical one in the sense that given the history $h_n \in H_n$ at the n*th* decision stage, where the decision $d_n \in D_n(h_n)$ is made, the next stage of the system, x_{n+1}, is selected from Ω_{n+1} according to a mass function defined over Ω_{n+1}.

Definiton 2.7. The *law of motion,* F, is a sequence of families of mass functions. More specifically, $F = \{f_n\}_{n \in \mathbb{N}}$, where f_n is a real valued function defined on $H_n \times \mathbb{D} \times \Omega_{n+1}$ with the property that

(1) $0 \leq f_n\,(h_n, d_n, x_{n+1}) \leq 1, \; h_n \in H_n, \; d_n \in D_n(h_n), \; x_{n+1} \in \Omega_{n+1}, \; n \in \mathbb{N}$

and

(2) $\sum_{x_{n+1} \in \Omega_{n+1}} f_n(h_n, d_n, x_{n+1}) = 1, \; h_n \in H_n, \; d_n \in D_n(h_n), \; n \in \mathbb{N}$

f_n is called the *law of motion associated with the nth decision stage* whereas $f_n(h_n, d_n, \cdot)$ is called the *conditional mass function of x_{n+1} given h_n and d_n*.⌋

Notice that the state of the system at the first decision stage is not specified by F.

Definition 2.8. The *initial condition,* P_o, is a real valued function on Ω_1 with the property:

(1) $0 \leq P_o(x_1) \leq 1, \; x_1 \in \Omega_1$

and

(2) $\sum_{x_1 \in \Omega_1} P_o(x_1) = 1$.⌋

The initial state of the system is allowed then to be specified by means of a mass function defined on Ω_1. Obviously if there exists $x^\circ \in \Omega_1$ such that $P_o(x^\circ) = 1$, x° may be considered as the initial state.

In order to establish a preference order over the set H_∞, with each element of H_∞ a reward is associated.

Definition 2.9. The *reward function,* L, is a sequence of real valued functions defined on H_∞. More specifically, $L = \{L_n\}_{n \in \mathbb{N}}$, such that for each $n \in \mathbb{N}$, L_n is a real valued function defined on H_∞.⌋

Now that the elements of the decision model are defined, the formal definition of the complete multistage decision model is introduced.

Definition 2.10. A *complete multistage decision model* (abbreviated CMDM) is any qunituple (Ω, D, F, P_o, L) where: $\Omega = \{\Omega_n : n \in \mathbb{N}\}$ is a set of state spaces, $D = \{D_n\}_{n \in \mathbb{N}}$ is a sequence of admissible decision maps,

$F = \{f_n\}_{n\varepsilon\mathbb{N}}$ is a law of motion, P_o is an initial condition and $L = \{L_n\}_{n\varepsilon\mathbb{N}}$ is a reward function, as defined above.$\lrcorner$

The procedure used by the decision maker while making his decisions at the different decision stages will be defined now:

Definition 2.11. A *strategy*, S, associated with the model (Ω, D, F, P_o, L) is a sequence of maps from H_n to $\mathbb{D}$. More specifically $S = \{S_n\}_{n\varepsilon\mathbb{N}}$, where S_n is a map from H_n to $\mathbb{D}$, $n\varepsilon\mathbb{N}$.$\lrcorner$

When using the strategy $S = \{S_n\}_{n\varepsilon\mathbb{N}}$ and observing the history $h_n \varepsilon H_n$ at the n*th* decision stage, $d_n = S_n(h_n)\varepsilon\mathbb{D}$ is the decision taken. In order for the strategy S to be feasible, it is required that $S_n(h_n)\varepsilon D_n(h_n)$.

Definition 2.12. The strategy S associated with the model (Ω, D, F, P_o, L) is said to be *feasible* if $S_n\,(h_n) = D_n(h_n)$, $\forall n\varepsilon\mathbb{N}$, $h_n\varepsilon H_n$. The set of all the feasible strategies associated with the model will be denoted by SS.$\lrcorner$

The application of the feasible strategy, S, associated with the model (Ω, D, F, P_o, L) generates a process: the process induced by S, which schematically may be described as follows:

The process starts at the first decision stage, n = 1, at some $x_1\varepsilon\Omega_1$ selected from Ω_1 according to the initial condition P_o; then the action $S_1(x_1)$ is taken and the system moves to some state $x_2\varepsilon\Omega_2$ selected according to the conditional mass function $f_1(x_1, S_1(x_1),\cdot)$; then the action $S_2(x_1,S_1(x_1), x_2)$ is taken and the system moves to some point $x_3\varepsilon\Omega_3$ selected according to $f_2(x_1, S_1(x_1), x_2, S_2\,(x_1, S_1(x_1), x_2), \cdot)$, etc.

The first concern of the decision maker is then the construction of a feasible strategy.

<u>Definition 2.13</u>. The *multistage decision problem* associated with the model $(\mathfrak{A}, D, F, P_o, L)$ is the construction of a feasible strategy. The strategy S is said to be a *feasible solution* to the problem if it is feasible.

Obviously, in addition to the feasibility of S the decision maker is also interested in the histories that may be produced by S which affect the rewards associated with the process. In order to select an element of SS that will optimize the reward, an optimality criterion is to be determined.

The discussion will be restricted to situations in which the expected value of the reward is used as a measure of effectiveness. For this purpose, it will be shown that the expected value criterion is meaningful, at least mathematically.

<u>Definition 2.14</u>. Let $(\mathfrak{A}, D, F, P_o, L)$ be a complete model. The product $\Omega = \overset{\infty}{\underset{n=1}{X}} \Omega_n$ is called the *sample space* associated with the model and its elements will be denoted by ω. Three sequences of functions will be associated with the sample space:

(1) $\xi = \{\xi_n\}_{n\varepsilon\mathbb{N}} \qquad \xi_n : \Omega \to \Omega_n \qquad ,n\varepsilon\mathbb{N}$

(2) $\eta = \{\eta_n\}_{n\varepsilon\mathbb{N}} \qquad \eta_n : \Omega \to \overset{n}{\underset{i=1}{X}} \Omega_i \qquad ,n\varepsilon\mathbb{N}$

(3) $\zeta = \{\zeta_n\}_{n\varepsilon\mathbb{N}} \qquad \zeta_n : \Omega \to \overset{\infty}{\underset{i=n}{X}} \Omega_i \qquad ,n\varepsilon\mathbb{N}$

where: $\xi_n(\omega) = x_n$, $\eta_n(\omega) = (x_1, x_2, \ldots, x_n)$, $\zeta_n(\omega) = (x_n, x_{n+1}, \ldots)$, $\omega = (x_1, x_2, \ldots)\varepsilon\Omega$.

ξ_n, η_n, and ζ_n will be referred to as the *present, past, and future state functions* associated with the n*th* decision stage, respectively.

<u>Definition 2.15</u>. Let (Ω, D, F, P_o, L) be a complete model and S a feasible strategy associated with it. Consider the probability space (Ω, ψ, P_s) where:

(1) Ω is the sample space associated with the model;

(2) Ψ is the infinite product σ-algebra determined by the factors consisting of all the subsets of Ω_n, $n \epsilon \mathbb{N}$;

(3) P_s is the unique probability measure on ψ with the property that:

$$P_s\ (x_1, x_2, \ldots, x_n) = P_o(x_1) \cdot f_1(x_1, S_1, (x_1), x_2) \ldots f_{n-1}(h_{n-1}, S_{n-1}(h_{n-1}), x_n)$$

where P_o is the initial condition, f_i is the law of motion associated with the *ith* decision stage, and

$h_{n-1} = (x_1, S_1(x_1), x_2, S_2(x_1, S_1(x_1), x_2), \ldots, x_{n-1})$

(Ω, ψ, P_s) is called the *probability space induced by S* and P_s the *probability measure induced by S.*

The existence and uniqueness of P_s and (Ω, ψ, P_s) are guaranteed by the structure of the model. (For details, see theorem of Kolmogoroff or theorem of Tulcea, cf. Loeve, 1960, p. 137).

<u>Definition 2.16</u>. Let (Ω, D, F, P_o, L) be a complete model, S a feasible strategy and (Ω, ψ, P_s) the probability space induced by S. Let also $h_{n,s}\ (\bar{x}_n)$ be the history associated with $\bar{x}_n$ and S, i.e.:

$h_{n,s}\ (\bar{x}_n) = (x_1, S_1(x_1), x_2, S_2\ (x_1, S_1(x_1), x_2), \ldots, x_n)$ with

$\bar{x}_n = (x_1, x_2, \ldots, x_n)$.

The expected value of L_1 assiciated with the strategy S and denoted by R(S) is called the *total reward associated with S.* More specifically,

$$R(S) = E[\ell_s(\omega)]$$

where $\ell_s(\omega) = L_1\ (h_{\infty,s}\ (\zeta_1(\omega)))$.

The strategy S^* is said to be an *optimal feasible solution* to the problem associated with the model if:

(1) $S^* \varepsilon SS$

and

(2) $R(S^*) \geq R(S), \forall S \varepsilon SS$.

The set of all the optimal feasible strategies will be denoted by SS^* and $R^* = R(S^*)$, $S^* \varepsilon SS^*$ will be called the *optimal feasible total reward* associated with the model.

Suppose that the process starts at n=1 by applying the strategy $S' \varepsilon SS$ and that at the n*th* stage h_n is observed. At this point the strategy $S'' \varepsilon SS$ is applied and the process continues under S'' for all $i \geq n$. The situation the decision maker encounters at (h_n, n) may be considered as a modified problem.

<u>Definition 2.17</u>. Let (Ω, D, F, P_o, L) be a complete model and S a feasible strategy. The conditional expectation of L_n given h_n associated with S denoted by $R_n(h_n,S)$ is called the *reward associated with the strategy S at the modified problem (h_n,n)*. More specifically,

$$R_n(h_n,S) = E[\ell_{n,s}(\omega)|h_n]$$

where $\ell_{n,s}(\omega) = L_n(h_{\infty,s}(\zeta_1(\omega)))$.

The strategy S' is said to be an *optimal feasible solution to the modified problem (h_n,n)* if

(1) $S' \varepsilon SS$ and (2) $R_n(h_n,S') \geq R_n(h_n,S)$, $\forall S \varepsilon SS$,

and for such a strategy $R_n^*(h_n) = R_n(h_n,S')$ is said to be the *optimal feasible reward associated with the modified problem (h_n,n)*.

It should be noted that when considering $R_n(h_n,S)$, it is not required that h_n is actually observed by S.

2.2. Sufficient Statistic and the Reduced Model

The elements D_n, f_n, and L_n of the complete model assumed to depend on the histories in the sense that $D_n = D_n(h)$, $f_n = f_n(h_n, \cdot, \cdot)$ and $L_n = L_n(h_\infty)$. In many situations this dependence does not require a full knowledge of h but rather may be determined by codensing the information contained by h. The ability to condense the information contained by h and still preserve the basic characteristics of the process may significantly reduce the dimension of the problem associated with the model. This is the motivation for using the concept of sufficient statistic.

Definitions

Definition 2.18. Let (Ω, D, F, P_0, L) be a complete multistage decision model and $T = \{t_n\}_{n\in\mathbb{N}}$ a sequence of maps from H_n to U_n where $\{U_n\}_{n\in\mathbb{N}}$ is a sequence of arbitrary sets. The sequence $T = \{t_n\}_{n\in\mathbb{N}}$ is called *a sufficient statistic of the complete model* if it has the following properties:

(1) t_n is a surjective map $\forall n\in\mathbb{N}$,

(2) $D_n(h) = D_n'(t_n(h))$ $\forall n\in\mathbb{N}$, $h\in H_n$,

(3) $f_n(h,d,x) = f_n'(t_n(h),d,x)$, $\forall n\in\mathbb{N}$, $d\in D_n(h)$, $x\in\Omega_{n+1}$,

(4) $L_n(h_n,d_n,x_{n+1},d_{n+1}, \ldots) = L_n'(t_n(h_n'), d_n', x_{n+1}', d_{n+1}', \ldots)$ for all h_n, $h_n'\in H_n$, x_i, $x_i'\in\Omega_i$, $i \geq n+1$, $d_j\in D_j(h_j)$, $d_j'\in D_j'(t_j(h_j'))$, $j \geq n$ for which

(4.1) $t_n(h_n) = t_n(h_n')$

(4.2) $h_{n+1} = (h_n,d_n,x_{n+1})$, $h_{n+1}' = (h_n',d_n',x_{n+1}')$

$h_i = (h_{i-1}, d_{i-1}, x_i)$, $h_i' = (h_{i-1}',d_{i-1}',x_i')$, $i \geq n+1$

(4.3) $t_i(h_i) = t_i(h_i')$, $i \geq n+1$

for some functions D_n', f'_n and L'_n, $n\varepsilon \mathbb{N}$ such that D_n': $U_n \to \mathbb{D}$, f_n': $U_n \times \mathbb{D} \times \Omega_{n+1} \to \mathbb{R}$ and L_n': $U_n \times \mathbb{D} \times \Omega_{n+1} \times \mathbb{D} \times \Omega_{n+2} \times \ldots \to \mathbb{R}$.

(5) If h_n, $h_n' \varepsilon H_n$, then $t_n(h_n) = t_n(h_n')$ implies

$$t_{n+1}(h_n, d, x) = t_{n+1}(h_n', d, x)$$

for all $d\varepsilon D_n(h_n) = D_n'(t_n[h_n])$, $x\varepsilon\Omega_{n+1}$.

The sequences $D' = \{D_n'\}_{n\varepsilon\mathbb{N}}$, $F' = \{f_n'\}_{n\varepsilon\mathbb{N}}$, and $L' = \{L_n'\}_{n\varepsilon\mathbb{N}}$ are called the *reduced admissible decision maps*, the *reduced law of motion*, and the *reduced reward function*, respectively.

The sufficient statistic T defines certain functions which will be useful when investigating the relationship between D, F, L and D', F' and L'.

Definition 2.19. Let (Ω, D, F, P_o, L) be a complete model and $T = \{t_n\}_{n\varepsilon\mathbb{N}}$ a sufficient statistic associated with it. Let $V = \{V_n\}_{n\varepsilon\mathbb{N}}$ be the sequence of maps defined on $U_n \times \mathbb{D} \times \Omega_{n+1}$ with values in U_{n+1} such that: $V_{n+1}(u_n, d_n, x_{n+1}) = t_{n+1}(h_{n+1})$

for all $h_{n+1} = (h_n, d_n, x_{n+1})$ for which $u_n = t_n(h_n)$, as defined by property (5) of T as described in Definition 2.18. The sequence V is called the *transition function associated with the sufficient statistic T.*

Let $\Gamma = \{\Gamma_n\}_{n\varepsilon\mathbb{N}}$ be the sequence of maps from U_n to the set of all subsets of H_n such that:

$$\Gamma_n(u_n) = \{h_n: \ h_n\varepsilon H_n, \ t_n(h_n) = u_n\} \ , \ n\varepsilon\mathbb{N}, \ u_n\varepsilon U_n.$$

The sequence will be called the *partition function* associated with the sufficient statistic T.

It should be noted that the sequences V and Γ are uniquely determined by the sufficient statistic T, and that there always exists a sufficient statistic, i.e., there always exists the trivial sufficient statistic: $T = \{t_n: t_n(h_n) = h_n, h_n \varepsilon H_n\}_{n\varepsilon\mathbb{N}}$.

Now that the elements D', F', L' and T are defined, the notion of a reduced model is introduced.

Definition 2.20. Let (Ω, D, F, P_o, L) be a complete model, T a sufficient statistic, and D', F', and L' the sequence of reduced admissible decision maps, reduced law of motion and reduced reward function associated with T, respectively. The quintuple (Ω, D', F', P_o, L') is called the *reduced multistage decision model (RMDM) associated with the complete model and T.*

The decision making procedure associated with the RMDM is similar to that used in the CMDM only that in this case $u_n = t_n(h_n)$ is observed rather h_n itself. Notice that since $D'_n(t_n(h_n)) = D_n(h_n)$, the set of alternative decisions available at (u_n, n) is identical with the set available at (h_n, n), $\forall h_n$ for which $t_n(h_n) = u_n$.

Definition 2.21. A *strategy,* G, associated with the reduced model (Ω, D', F', P_o, L') is a sequence of maps from U_n to $\mathbb{D}$. More specifically, $G = \{G_n\}_{n\varepsilon\mathbb{N}}$, such that $G_n: U_n \to \mathbb{D}$, $\forall n\varepsilon\mathbb{N}$. The strategy G is said to be *feasible* if $G_n(u_n)\varepsilon D'_n(u_n)$, $\forall n\varepsilon\mathbb{N}$, $u_n \varepsilon Un$. The set of all the feasible strategies associated with the reduced model is denoted by GG.

Definition 2.22. The *multistage decision problem* associated with the reduced model (Ω, D', F', P_o, L') is the construction of a feasible strategy. The strategy G is said to be a *solution* to the problem if it is feasible.

As in the case of the complete model, each element of GG induces a probability space (Ω, ψ, P_G), where Ω and ψ are as defined in Definition 2.15 and P_G is the unique probability measure as defined by F'.

Definition 2.23. Let $(\Omega, D', F', P_o, L')$ be a reduced model and G a feasible strategy associated with it. The probability space (Ω, ψ, P_G) is called the *probability space induced by G* where Ω is the sample space, ψ is the σ-algebra on Ω as defined in Definition 2.15 and P_G is the unique probability measure on ψ such that:

$$P_G\ (\xi_{n+1}\ (\omega) = x_{n+1} | \eta_n(\omega) = \bar{x}_n) = f'_n(u_n, G_n(u_n), x_{n+1})$$

where: $\omega = (\bar{x}_n, x_{n+1}, x_{n+2}, \ldots)$, $u_n = t_n(h_n)$, and

$$h_n = (\bar{x}_n(1), G_1(t_1(\bar{x}_n(1))), \ldots, \bar{x}_n(n)).$$

P_G is called the *probability measure induced by G*. The expected value of L'_1 associated with the strategy G, denoted by R'(G) is called the *total reward associated with G*. More specifically,

$$R'(G) = E[\ell'_G(\omega)]$$

where: $\ell'_G(\omega) = L_1(t_1(\xi_1(\omega)), G_1(t_1(\xi_1(\omega))), \xi_2(\omega), \ldots)$.

The strategy G* is said to be an *optimal feasible* strategy if:

(1) $G^* \varepsilon GG$

and

(2) $R'(G^*) \geq R'(G)$, $\forall\ G \varepsilon GG$.

The set of all the optimal feasible strategies associated with the model will be denoted by GG*, and $R'^* = R'(G^*)$, $G^* \varepsilon GG^*$ will be called the *optimal feasible total reward* associated with the model. The conditional expectation of L'_n associated with G given u_n, denoted by $R'_n(u_n, G)$ is called the *reward associated with G at the modified problem (u_n, n)*. More specifically:

$$R'_n(u_n, G) = E[\ell'_{n,G}(\omega) \mid u_n]$$

where: $\ell'_{n,G}(\omega) = L'_n\,(t_n(h_{n,G}(\eta_n(\omega))),\ G_n(u_n),\ \xi_{n+1}(\omega),\ . \ . \ .)$

and $h_{n,G}(\bar{x}_n) = (x_1,\ G,\ (t_1(x_1)),\ . \ . \ .\ x_n)$.

The strategy G' is said to be *an optimal feasible solution to the modified problem* (u_n, n) if:

(1) $G' \varepsilon GG$

and

(2) $R'_n(u_n, G') \geq R'_n(u_n, G)$, $\forall\ G \varepsilon GG$,

and for such a strategy $R'^*_n(u_n) = R'_n(u_n, G')$ is said to be the *optimal feasible reward at* (u_n, n).⌋

Suppose that while making the decisions there is a choice between the use of strategies depending on histories vs. strategies depending on the sufficient statistics. Which option will be advantageous as far as the rewards are concerned? In order to show that the two options produce the same results, the concept of images is introduced.

<u>Definition 2.24</u>. Let (Ω, D, F, P_o, L) be a complete model, T a sufficient statistic associated with it, and $(\Omega, D', F', P_o, L')$ the corresponding reduced model. Let I^c: $GG \rightarrow SS$ be the map determined as follows:

$$I^c(G) = \{S_n : S_n(h_n) = G_n(t_n(h_n)),\ h_n \varepsilon H_n\}_{n\varepsilon\mathbb{N}}$$

The strategy $S = I^c(G)$ is called the *complete image* of G.⌋

<u>Lemma 2.1</u>. Let $(\Omega, D', F', P_o, L')$ be the reduced model associated with the complete model (Ω, D, F, P_o, L) and the sufficient statistic T. Then,

(1) $R_n(h_n, I^c(G)) = R'_n(t_n(h_n), G)$, $\forall n \varepsilon N$, $h_n \varepsilon H_n$, $G \varepsilon GG$,

and

(2) $R(I^c(G)) = R'(G), \forall G \varepsilon GG$.

Proof: (1) Let G be any arbitrary element of GG and S the complete image of G, i.e., $S = I^c(G)$. By construction $G \varepsilon GG$ implies that $S \varepsilon SS$ and from the relation between F and F' it is given that $P_S = P_G$. Since from the relation between L and L' it follows that

$$L_n'(u_n, G_n(u_n), x_{n+1}, \ldots) = L_n(h_n, S_n(h_n), x_{n+1}, \ldots)$$

for all $n \varepsilon \mathbb{N}$, $h_n \varepsilon H_n$ for which $t_n(h_n) = u_n$ and $x_i \varepsilon \Omega_1$ $i \geq n$, it also follows that

$$R_n(h_n, S) = R'(t_n(h_n), G), \quad n \varepsilon \mathbb{N}, \ h_n \varepsilon H_n.$$

(2) From the definitions of R and R' it follows that

$$R(S) = \sum_{x_1 \varepsilon \Omega_1} R_1(x_1, S) \cdot P_o(x_1)$$

and

$$R'(G) = \sum_{x_1 \varepsilon \Omega_1} R_1'(t_1(x_1), G) \cdot P_o(x_1) .$$

Using the first part of the Lemma for n=1 it is given that

$$R(S) = R'(G).\rfloor$$

The above lemma implies that as far as the rewards are concerned the complete model is as good as the reduced model.

Definition 2.25. Let (Ω, D, F, P_o, L) be a complete model and T a sufficient statistic associated with it. For each $S \varepsilon SS$ and $n \varepsilon \mathbb{N}$, construct the sequence $\{S^i\}_{i \geq n}$ of strategies as follows:

(1) for i=n set:

$$S_m^n(h_m) = \begin{cases} S_m(h_m), & m \neq n, \quad h_m \varepsilon H_m \\ S_m(h^*(h_m)), & m = n, \ h_m \varepsilon H_m \end{cases}$$

where $h^*(h_n)$ is some arbitrary element of $\Gamma_n(t_n(h_n))$ for which

$R_n(h^*(h_i),S^{i-1}) \geq R_n(h,S^{i-1})$, $\forall h \varepsilon \Gamma_i(t_i(h_i))$ and $R_n(h_i,S^{i-1})$ is the expected value of L_n given h_i associated with the strategy S^{i-1}. Any sequence $\{S^i\}_{i \geq n}$ constructed as described above is said to be *generated by S at n.* The strategy $S' = \lim_{i \to \infty} S^i$ will be denoted by $I_n^r(S)$ and called the *strategy generated by S at n.* For n=1 the strategy $I^r(S) = I_1{}^r(S)$ is called a *reduced image of S.*

Notice that by construction, the uniqueness of $I^r(S)$ is not guaranteed, and depends on the choice of $h^*(h_i)$, if any. However, the above definition guarantees that every reduced image, S', of S has the following property:

$S_n'(h_n) = S_n'(h_n')$, $\forall h_n$, h' for which $t_n(h_n) = t_n(h_n')$ and thus the strategy $G \varepsilon GG$ constructed by setting $G_n(u_n) = S_n'(h_n)$ for any arbitrary $h_n \varepsilon \Gamma_n(u_n)$ is well defined.

Lemma 2.2. Let (Ω, D', F', P_o, L') be the reduced model associated with the complete model (Ω, D', F', P_o, L') and the sufficient statistic T. Let also $\{S^i\}_{i \geq n}$ be a sequence of strategies generated by $S \varepsilon SS$ at n. Then, $R_n(h_n,s^i) \geq R_n(h_n,S)$ $\forall h_n \varepsilon \mathbb{H}$, $i \geq n$.

Proof. The lemma will be proven by induction on i. For i = n the relation $R_n(h_n,S^n) \geq R_n(h_n,S)$, $\forall h_n \varepsilon H_n$ is guaranteed by the structure of S^n, T and R_n. Assume that the inductive hypothesis is true for i = n+1, n+2, . . .m. In particular assume that it is true for i=m, i.e.:

$$R_n(h_n,S^m) \geq R_n(h_n,S), \ \forall h_n \varepsilon Hn.$$

Consider i = m+1 for which

$$R_n(h_n,S^{m+1}) = E\ [R_n(h_n,S^{m+1}\ |x_{n+1}, \ldots, x_{m+1})]$$

and

$$R_n(h_n,S^m) = E\ [R_n(h_n,S^{m}\ |x_{n+1}, \ldots, x_{m+1})]$$

where the expectations are taken with respect to $(x_{n+1}, \ldots, x_{m+1})$. By construction $S_k^{m+1}(h_k) = S_k^m(h_k)$ $k < i$ so that the definition of $\{S^i\}_{i \geq n}$ implies that

$$R_n(h_n, S^{m+1})\big|_{x_{n+1}, \ldots, x_{m+1}} \geq R_n(h_n, S^m)\big|_{x_{n+1}, \ldots x_{m+1}}$$

for all $h_n \varepsilon H_n$, and $(x_{n+1}, \ldots x_{m+1}) \varepsilon \underset{k=n+1}{\overset{m+1}{X}} \Omega_k$.

From the definition of $\{S^i\}_{i \geq n}$ it also follows that:

$$P_{S^m}(x_{n+1}, \ldots, x_{m+1}|h_n) = P_{S^{m+1}}^{m+1}(x_{n+1}, \ldots, x_{m+1}|h_n)$$

for all $h_n \varepsilon H_n$ and $(x_{n+1}, \ldots, x_{m+1}) \varepsilon \underset{k=n+1}{X} \Omega_k$, and thus

$$R_n(h_n, S^{m+1}) \geq R_n(h_n, S), \ \forall h_n \varepsilon H_n$$

and the inductive hypothesis is true for i=m+1, and hence it is true for all $i \geq n$.⌋

The results obtained by Lemma 2.1 and Lemma 2.2 yield then,

<u>Theorem 2.1</u>. Let $(\Omega, D', F', P_o, L')$ be the reduced model associated with the complete model (Ω, D, F, P_o, L) and the sufficient statistic T. Then

(1) $R_n{*}(h_n) = R_n'{*}(t_n(h_n))$, $\forall n \varepsilon \mathbb{N}$, $h_n \varepsilon H_n$

and

(2) $R^* = R'^*$

provided that the above exist.⌋

<u>Proof</u>: (1) Let (h_n, n) be any arbitrary modified problem for which $R_n{*}(h_n)$ exists. In other words,

$$R_n{*}(h_n) = R_n(h_n, S') \quad \text{for some} \quad S' \varepsilon SS.$$

Let G' be the reduced image of S' at n: i.e.,

$$G' = I_n^r(S')$$

From Lemma 2.2 it follows then that

$$R_n'(t_n(h_n), G') \geq R_n^*(h_n)$$

However, from Lemma 2.1 it follows that

$$R_n(h_n, I^c(G')) = R_n'(t_n(h_n), G')$$

which implies then that

$$R_n^*(h_n) = R_n'^*(t_n(h_n)).$$

(2) Let $S^* \varepsilon SS$ and $G^* \varepsilon GG^*$ be any optimal feasible solutions to the problems associated with the complete and reduced model respectively. In other words,

$$R^* = R(S^*), \text{ and } R'^* = R'(G^*)$$

From Lemma 2.1, it is known that

$$R^* \geq R(I^c(G^*)) = R'(G^*) = R^*$$

while from Lemma 2.2 it follows that

$$R'^*(I^r(S^*)) \geq R(S^*) = R^*$$

Thus,

$$R^* = R(S^*) = R(G^*) = R'^*.$$

Theorem 2.1 implies that trying to optimize the rewards the use of the complete and the reduced model will provide the same results. As far as computation is considered, often the reduced model is advantageous as will be demonstrated in Chapter 5.

2.3. Special Types of Multistage Decision Models

Certain properties of Ω, D, F, L, and T are often used as classification criteria. In this section, a number of these criteria are introduced.

Definitions

Definition 2.26. A complete model for which there exists the sufficient statistic:

$$T = \{t_n: \ t_n(h_n) = x_n, \ h_n \varepsilon H_n\}_{n\varepsilon \mathbb{N}}$$

is called a *Markovian* model. If in addition the model is such that: $\Omega_n = \bar{\Omega}, \ \forall n \varepsilon \mathbb{N}, \ D_i = D_j$ and $f_i = f_j, \ \forall \ i,j \varepsilon \mathbb{N}$, the model is said to be *stationary.*⌋

In many situations the actual process is such that the number of decision stages is finite, say N. If the model presented in this chapter is to be used it is necessary then to construct dummy decision stages, reward functions, etc., for decision stages greater than N. The notion of a truncated model as will be introduced now is not restricted to the above situation and also represents a situation where given the problem, that is, the history at n = N, the rest of the decision process is already determined as far as the decisions at decision stages greater than N are concerned.

Definition 2.27. Let (Ω, D, F, P_o, L) be a complete model for which $D = \{D_n\}_{n\varepsilon \mathbb{N}}$ is such that

$$D_n(h_N, d, x_{N+1}, \ldots, x_n) = \delta_n(h_N) \quad , \quad n > N$$

where δ_n is a function defined on H_N with values in $\mathbb{D}$. It is said then that the model is *truncated at N* which is indicated by writing: $(\Omega, D, F, P_o, L)_N$.⌋

As far as the reward function is concerned in many situations it has the following structure.

Definition 2.28. Let (Ω, D, F, P_o, L) be a complete model for which $L = \{L_n\}_{n\varepsilon \mathbb{N}}$ is such that for all $n\varepsilon \mathbb{N}$

$$L_n(h_n,d_n,x_{n+1}, \ldots) = r_n(h_n,d_n,x_{n+1}) + L_{n+1}(h_{n+1},d_{n+1}, \ldots)$$

where r_n is a real valued function defined on $H_n \times \mathbb{D} \times \Omega_{n+1}$. It is said then that L is an *additive* reward function.

Most of the early investigations concerning multistage decision models have been restricted to additive reward function. Hinderer's (1970) model, for example, treats only additive reward functions. At this stage of the analysis, the only properties of L that have been specified are the domain of definition, H_∞, and the range, $\mathbb{R}$.

2.4. Discussion

The multistage decision model presented in this chapter belongs to the class of models often referred to as discrete dynamic programming models or discrete time state models (Blackwell, 1962; Aris, 1964, Maitra, 1968, Miller and Veinoff, 1969, and others).

Following the pioneering work of Bellman (1952, 1953, 1954, 1957) and Howard (1960), the class of multistage decision models has been expanded significantly especially as far as the structure of the reward function is concerned (Mitten, 1964, 1974; Denardo, 1965; Sobel, 1975).

In this section, the basic characteristics of the elements of the model will be discussed including possible modifications for handling processes other than the one for which the model was originally designed for.

The Set of Decision Stages, $\mathbb{N}$

The model is concerned with processes having countably many decision stages. For truncated processes, one can still use the model by constructing dummy decision stages. It is also assumed that the process is a serial one. Nemhauser (1966) has shown how certain non-serial processes may be decomposed into a set of serial processes each of which is treated by methods applicable to serial processes. If there is uncertainty concerning the sequence of decision stages to be realized, it is possible (Denardo, 1965) to embed the decision stages in the state spaces while using dummy variables for the decision stages themselves.

The Set of State Spaces, Ω

The basic characteristic of Ω is that its elements Ω_n are countable sets. For models allowing noncountable state spaces see Blackwell (1965), Sirjaev (1970), and Hinderer (1970). It was purposely determined to explicitly indicate that the state spaces need not be identical. For truncated processes there is a need to construct dummy state spaces for dummy decision stages which for convenience often may consist of one element only.

Admissible Decision Map, D

As will be indicated in Chapter 5, the construction of D is based on two types of constraints: the first has to do with the availability of decisions at the modified problem (h_n,n) while the second involves constraints imposed on Ω_n. In some models, Yakowitz (1969), for example, these two types of constraints are explicitly formulated. At this stage it should be emphasized that in contrast to Askew's (1974) comment, probabilistic constraints can also be handled by the sequence D, as indicated by Bellman and Dreyfus (1962), White (1974), and as will be demonstrated in Chapter 5.

The Law of Motion, F

Deterministic processes as far as the dynamics of the process is concerned may be treated as a degenerate case of the statistical law of motion introduced in the model. It should be noted, however, that even if the law of motion is a deterministic one, an expected value criterion still may be meaningful if some uncertainty is involved in the rewards associated with the process.

Initial Condition, P_o

As in the case of F, a deterministic process may be formulated (as far as the initial condition is concerned) by setting $P_o(x_1) = 1$, for that element $x_1 \varepsilon \Omega_1$ which is the initial state of the process.

The Reward Function, L

While in Hinderer's model the reward function is assumed to be additive, no assumptions are made as to the structure of L_n other than specifying its domain of definition H_∞ and its range $\mathbb{R}$. It should be noted, however, that nonreal valued function may also be considered when modeling multistage decision processes. Mitten (1974) and Sobel (1975), for example, introduce a reward function for which, $L_n(h_\infty) = h_\infty$ for this type of a reward function a modified optimality criterion is needed since the expected value is no longer suitable.

The Sufficient Statistic, T

Hinderer's definition has been modified so as to account for the general form of L used in the model. Since the uniqueness of T is often not guaranteed, its construction may be considered as a modeling problem.

More on the role of sufficient statistics in the modeling of decision process can be found in Sirjaev (1970).

Optimality Criterion

The optimal feasible solutions are defined as those strategies maximizing the total reward, which in itself is the expected value of L_1. If the objective is the minimization of the reward, L_1 is taken as the original objective function multiplied by -1. Alternatively, the definition of the optimal solution may be modified so that an optimal strategy will be such that it minimizes the total reward. Optimality criteria other than the expected value, such as the minimax (Nemhauser, 1966) and the average cost (Derman, 1966) may also be used in the context of the model presented in this chapter by redefining the notion of optimality as far as the strategies and the rewards are concerned.

As far as modeling flexibility is concerned, the model covers a rather wide class of multistage decision processes. Moreover, its use may be extended even more by minor modifications in the structure of its elements.

CHAPTER 3

THE DYNAMIC PROGRAMMING ALGORITHM AND THE PRINCIPLE OF OPTIMALITY

In this chapter, an algorithm for the construction of feasible solution(s) to the multistage decision problem is discussed: the *dynamic programming* (DP) algorithm. It will be shown that the algorithm provides optimal feasible solutions to a certain class of multistage decision problems. Also to be discussed are: the *principle of optimality* (PO) and a class of models for which it holds, and the relation between the principle and the algorithm.

3.1. The Dynamic Programming Algorithm

The DP algorithm traces back to Bellman (1952) where it was used for the construction of optimal feasible strategies for rather simple multistage decision problems. Although the DP algorithm as defined in this chapter is very similar to algorithms defined elsewhere, Yakowitz (1969) for example, it should be noted that it is defined in the context of a multistage decision process which is not necessarily truncated. Since for every CMDM there exists a sufficient statistic and thus a RMDM, the DP algorithm will be formulated for reduced models with the understanding that when used for complete models the trivial sufficient statistic may be used.

Definitions

Definition 3.1. Let $(\Omega, D', F', P_o, L')$ be a RMDM and K an element of $\mathbb{N}$. Consider the following algorithm for constructing the sets GG^n, $n \leq K$ of strategies:

Step 1. For $n = K$ and $u \varepsilon U_K$ construct the set $A^K(u)$ of all strategies $G' \varepsilon GG$ statisfying the condition

$$R'_K(u,G') = \max_{G \varepsilon GG} R'_K(u,G) \tag{3.1}$$

and let $A^K = \{G': \ G' \varepsilon A^K(u),\ u \varepsilon U_K\}$.

Also let $B^K(u)$ be the subset of $D'_K(u)$ such that

$$B^k(u) = \begin{cases} d: \{d = G'_K(u),\ G' \varepsilon A^K(u)\}, & \text{if } A^K(u) \neq \emptyset \\ d: \{d = G'_K(u),\ G' \varepsilon A^K\}, & \text{if } A^K(u) = \emptyset,\ A^K \neq \emptyset \\ D'_K(u) & \text{otherwise} \end{cases}$$

Construct the set GG^K of strategies G^K such that

$$GG^K = \begin{cases} \{G^K: \ G^K \varepsilon A^K,\ G_k{}^K(u)\ \varepsilon B^K(u)\} & \text{if } A^K \neq \emptyset \\ GG & \text{otherwise} \end{cases}$$

Step 2. For $n < K$ and $u \varepsilon U_n$ construct the set $A^n(u)$ of all the strategies $G' \varepsilon GG^{n+1}$ satisfying the condition

$$R'_n(u,\ G') = \max_{G^{n+1} \varepsilon GG^{n+1}} R'_n(u,\ G^{n+1}) \tag{3.2}$$

and let $A^n = \{G': \ G' \varepsilon A^n(u),\ u \varepsilon U_n\}$.

Also, let $B^n(u)$ be the subset of $D'_n(u)$ such that

$$B^n(u) = \begin{cases} d: \{d = G'_n(u),\ G' \varepsilon A^n(u)\}, & \text{if } A^n(u) \neq \emptyset \\ d: \{d = G'_n(u),\ G' \varepsilon A^n\}, & \text{if } A^n(u) = \emptyset,\ A^n \neq \emptyset \\ D'_n(u) & \text{otherwise} \end{cases}$$

Construct the set GG^n of strategies G^n such that

$$GG^n = \begin{cases} \{G^n: \; G^n \varepsilon A^n, \; G_n^{\,n}(u) \varepsilon B^n(u)\}, & \text{if } A^n \neq \emptyset \\ GG^{n+1} & \text{otherwise} \end{cases}$$

Step 3. Construct the set $GG°$ of strategies $G°$ such that

$$GG° = \{G°: \; G° \varepsilon GG^1, \; R'(G°) \geq R'(G^1), \; \forall \; G^1 \varepsilon GG^1\}$$

The above procedure is called the *dynamic programming algorithm* and the sets GG^n, $n \leq K$ the *dynamic programming solutions* for the n*th* decision stage. The decision stage $n = K$ is called the *initial decision stage* associated with the algorithm. The set $GG°$ is called the *set of solution produced by the DP algorithm.* Equation 3.2 associated with the second step of the algorithm is called the *dynamic programming equation.* The dynamic programming equation is said to *hold at (u,n),* $n < K$ if:

$$R_n'(u, \; G^n) = \max_{G \varepsilon GG} R_n(u,G), \; \forall G^n \varepsilon A^n(u)$$

and it is said to *hold* if it holds for every modified problem (u_n,n), $u_n \varepsilon U_n$, $n \leq K$.

Remarks. (1) The structure of the algorithm guarantees that $GG° \neq \emptyset$ and $GG° c GG$. In other words, all the DP solutions are feasible.

(2) The decision stage K where the algorithm starts is not specified. For truncated models, K may be set to N, however, this is not a requirement.

(3) It is still left to be shown under what conditions the elements of $GG°$ are optimal feasible.

(4) Notice that the elements G^K of GG^K are not required to be optimal feasible for all $u_K \varepsilon U_K$ but rather every $G^K \varepsilon GG^K$ is required to be an optimal feasible solution for at least one element $u_K \varepsilon U_K$.

(5) In order to start the algorithm at $n = K$, a method is needed for solving equation 3.1. For truncated models, however, with $K = N$ the

solution of equation 3.1 is often a straightforward procedure. More details concerning the first step of the algorithm may be found in Denardo (1965).

A modified algorithm designed for (but not restricted to) truncated models is now introduced.

Definition 3.2. Let $(\Omega, D', F', P_o, L')$ be a RMDM and K an element of $\mathbb{N}$. Consider the following procedure of constructing the sets GG^n, $n \leq K$ of strategies G^n:

Step 1. For n = K constructs the set GG^K of strategies G^K such that: $$GG^K = \{G^K: G^K \varepsilon GG, R'_K(u, G^K) = R'^{*}_K(u) \text{ for all } u\varepsilon U_K \text{ for which } R'^{*}_K(u) \text{ exists}\}.$$

Step 2. The same as Step 2 in Definition 3.1.

Step 3. The same as Step 3 in Definition 3.1.

The above procedure is called the *modified dynamic programming algorithm.*

Remarks. (1) The modified algorithm does not guarantee that $GG^\circ \neq \emptyset$.

(2) In order to guarantee that $GG^K \neq \emptyset$ and consequently $GG^\circ \neq \emptyset$ it is required that the process is such that GG includes a strategy which is simultaneously optimally feasible for all $u\varepsilon U_K$, for which $R'^{*}_K(u)$ exists.

(3) For truncated models, the modified algorithm is similar to the algorithm defined in Definition 3.1.

The next step is to show that there exists a class of multistage decision problems for which the DP algorithm produces optimal feasible solutions.

Definition 3.3. The reduced multistage decision model (Ω, D', F', P_o, L') is said to be *regular* if:

$$R_n'^*(u_n) = \max_{G\varepsilon GG} R_n'(u_n, G)$$

exists for all $u_n \varepsilon U_n$, $n \varepsilon \mathbb{N}$.⌡

For example, if $D_n'(u_n)$ is finite for all $u_n \varepsilon U_n$, $n\varepsilon\mathbb{N}$ it follows that the model is regular.

Definition 3.4. Let (Ω, D', F', P_o, L') be a RMDM. The reward function $L' = \{L_n'\}_{n\varepsilon\mathbb{N}}$ is said to be *separable under expectation* if there exists a sequence $\{\rho_n\}_{n\varepsilon\mathbb{N}}$ of real valued functions defined on $U_n \times \mathbb{D} \times \mathbb{R}$ such that:

$$R_n'(u_n, G|x_{n+1}) = \rho_n[u_n, G_n(u_n), R_{n+1}' (u_{n+1}, G)]$$

for all $n\varepsilon\mathbb{N}$, $u_n \varepsilon U_n$, $G \varepsilon GG$ and $x_{n+1} \varepsilon \Omega_{n+1}$ where $u_{n+1} = V_n(u_n, G_n(u_n), x_{n+1})$. The reward function is said to be a *type Shoshana reward function* if L' is separable under expectation and $R_n'(u_n, G|x_{n+1})$ is monotone increasing with R_{n+1}', and a *type Moshe reward function* if it is separable under expectation and $R_n'(u_n, G|x_{n+1})$ is strictly monotone increasing with R_{n+1}'. Similarly, the model is said to be a *type Shoshana* and *type Moshe* model if L' is type Shoshana and type Moshe reward function, respectively.⌡

Examples

A number of reward functions are introduced now and their properties are investigated on the basis of the discussion presented above.

Example 3.1. Consider the reward function L' where:

$$L' = \{L_n': L_n'(u_n, d_n, x_{n+1}, d_{n+1}, \ldots) = \sum_{i \geq n} r_i(u_i, d_i)\}_{n\varepsilon\mathbb{N}}$$

and r_i is a real valued function defined on $U_i \times \mathbb{D}$, $i\varepsilon\mathbb{N}$. L_n' can also be written as: $L_n(u_n, d_n, x_{n+1}, \ldots) = r_n(u_n, d_n) + L_{n+1}' (u_{n+1}, d_{n+1}, \ldots)$

Thus, $R'_n(u_n, G|x_{n+1}) = r_n(u_n, G_n(u_n)) + R'_{n+1}(u_{n+1}, G)$ and hence L' is a type Moshe reward function.

Example 3.2. Consider the reward function L' where:

$$L' = \{L'_n: L'_n(u_n, d_n, x_{n+1}, \ldots) = \exp(\sum_{i \geq n} r_n(u_i, d_i)\}_{n \in \mathbb{N}}$$

and r_i is as defined in Example 3.1. L'_n can also be written as:

$$L'_n(u_n, d_n, x_{n+1}, \ldots) = \{\exp(r_n(u_n, d_n))\} \cdot \exp\{\sum_{i \geq n+1} r_i(u_i, d_i)\}$$

$$= \{\exp(r_n(u_n, d_n))\} \cdot L_{n+1}(u_{n+1}, d_{n+1}, x_{n+1}, \ldots)$$

Thus, $R'_n(u_n, G|x_{n+1}) = \{\exp(r_n(u_n, d_n))\} \cdot R'_{n+1}(u_{n+1}, G)$, $d_n = G_n(u_n)$ and hence L' is a type Moshe reward function.

Example 3.3. Consider the reward function L' where:

$$L' = \{L'_n: L_n(u_n, d_n, x_{n+1}, \ldots) = \prod_{i \geq n} d_i\}_{n \in \mathbb{N}} \quad \text{or}$$

$$L'_n(u_n, d_n, x_{n+1}, \ldots) = d_n \cdot L'_{n+1}(u_{n+1}, d_{n+1}, x_{n+2}, \ldots)$$

Thus, $R'_n(u_n, G|x_{n+1}) = d_n \cdot R'_{n+1}(u_{n+1}, G)$, $d_n = G_n(u_n)$.

If $\mathbb{D}$ contains only positive elements then L' is a type Moshe reward function. If $\mathbb{D}$ contains only non-negative elements, then L' is a type Shoshana reward function.

Example 3.4. Consider the reward function L' where:

$$L' = \{L'_n: L'_n(u_n, d_n, x_{n+1}, \ldots) = \max(u_n(1), \max_{i > n}\{x_i\})\}_{n \in \mathbb{N}} \quad \text{where}$$

$u_n(1) = \max_{i \leq n}\{x_i\}$, $n \in \mathbb{N}$. For this case it follows that,

$$L'_n(u_n, d_n, x_{n+1}, \ldots) = L'_{n+1}(u_{n+1}, d_{n+1}, x_{n+2}, \ldots)$$

and thus,

$$R'_n(u_n, G|x_{n+1}) = R'_{n+1}(u_{n+1}, G) \quad , \quad u_{n+1} = V_n(u_n, d_n, x_{n+1}).$$

Hence, L' is a type Moshe reward function.

Theorem 3.1. Let (Ω, D', F', P_o, L') be a regular reduced type Shoshana multistage decision model for which there exists $K\varepsilon\mathbb{N}$ and $G'\varepsilon GG$ such that

$$R'(u, G') = \max_{G\varepsilon GG} R'(u, G), \ \forall u\varepsilon U_K$$

Then,

$$(1) \ R'_n(u, G^n) = \max_{G\varepsilon GG} R'_n(u,G), \ \forall n\leq K, \ u\varepsilon U_n, \ G^n\varepsilon GG^n$$

and

$$(2) \ R'(G^\circ) = \max_{G\varepsilon GG} R'(G), \ \forall G^\circ\varepsilon GG^\circ$$

where GG^n is the set of the solutions produced by the DP algorithm at n and GG° is the set of the DP solutions.

Proof: It should be noted that under the above conditions the DP algorithm and the modified DP algorithm are identical.

(1) The first part of the theorem will be proven by induction on n. For n = K, the inductive hypothesis is true by the conditions specified by the theorem. Assume that the inductive hypothesis is true for n = K-1, K-2, . . .m. In particular, assume that it is true for n = m, i.e.:

$$R'_m(u, G^m) \geq R'_m(u,G), \ \forall u\varepsilon U_m, \ G^m\varepsilon GG^m, \ G\varepsilon GG.$$

Consider n = m-1 for which

$$R'_{m-1}(u,G) = \sum_{x_m\varepsilon\Omega_m} \rho_{m-1} \ (u, \ G_{m-1}(u), \ R_m(u_m,G)) \cdot f'_{m-1}(u,G_{m-1}(u),x_m)$$

From the monotonicity of ρ_{m-1}, the definition of the DP algorithm and the inductive hypothesis at n = m it follows then that:

$$R'_{m-1}(u, G^{m-1}) \geq R'_{m-1}(u,G), \ \forall u\varepsilon U_{m-1}, \ G^{m-1}\varepsilon GG^{m-1}, \ G\varepsilon GG.$$

Thus, the inductive hypothesis is true for n = m-1 and hence it is true for all $n \leq K$.

(2) By definition, $R'(G) = \sum_{x_1 \varepsilon \Omega_1} R_1'(t_1(x_1), G) \cdot P_o(x_1)$

From the first part of the theorem it follows then that

$$R'(G°) = \max_{G \varepsilon GG} R(G), \ \forall G° \varepsilon GG°$$

Notice that from the definition of the DP algorithm

$$R'(G') = R'(G''), \ \forall G', G'' \varepsilon GG°.$$

Remarks. (1) Notice that the conditions specified in Theorem 3.1 do not require that the model would be truncated.

(2) Theorem 3.1 does not provide an answer as to the optimality of $G° \varepsilon GG°$ at modified problems associated with $n > K$.

3.2. The Principle of Optimality

Consider the following situation: an optimal feasible strategy is to be constructed for a given reduced model and suppose that it can be shown that G^* is such a strategy. Suppose now that the process induced by G^* starts and that the modified problem (u,n) is observed. Two basic questions arise:

(1) Is G^* an optimal feasible solution to the modified problem (u,n)? In other words, is it true that

$$R_n'(u, G^*) \geq R_n(u,G), \ \forall G \varepsilon GG.$$

(2) Is G^* an optimal feasible solution to $R'(\cdot|u,n)$? In other words, is it true that

$$R'(G^*|u,n) \geq R'(G|u,n), \ \forall G \varepsilon GG.$$

Theorem 3.1 provides only a partial answer to the first question; that is, if the model is a regular type Shoshana model and $G^* \varepsilon GG°$, then for all $n \leq K$, G^* is optimal feasible for every modified problem (u,n). However, there is no guarantee that for models other than the one specified by Theorem 3.1 this condition holds nor is it guaranteed for optimal strategies which are not produced by the DP algorithm.

The principle of optimality is designed to provide a more complete answer to the questions introduced above. Before presenting the formal definition of the principle, some elements related to it are defined.

Definitions and Theorems

Definition 3.5. The *state observing function* associated with the reduced model $(\Omega, D', F', P_o, L')$ is the sequence $\Theta = \{\Theta_n\}_{n\varepsilon N}$ of maps from GG to the set of all subsets of $\bar{X}_n$ such that

$$\Theta_n(G) = \{\bar{x}_n: \; P_G(\eta_n = \bar{x}_n) > 0\}, \; n\varepsilon N, \; G\varepsilon GG$$

which is called the *set of trajectories observed under G at n.* Similarly the sets

$$H_n(G) = \{h_{n,G}(\bar{x}_n): \bar{x}_n\varepsilon\Theta_n(G)\}$$

and

$$U_n(G) = \{u_n: \; u_n = t_n(h_n), \; h_n\varepsilon H_n(G)\}$$

are called the *set of histories observed under G at n,* and the *set of statistics observed under G at n,* respectively.

It should be noted that the existence of and uniqueness of Θ is guaranteed by the structure of the model and that $H_n(G)$ and $U_n(G)$ are well defined.

Definition 3.6. Let $(\Omega, D', F', P_o, L')$ be a reduced multistage decision model and Θ the state observing function associated with it. Let G* be any optimal feasible strategy associated with the model. The *principle of optimality is said to hold* if with probability one G* is also an optimal feasible solution to every modified problem (u_n, n) for which $u_n \varepsilon U_n(G^*)$.

It will be shown that the principle of optimality holds for type Moshe models. First, however, its validity will be shown for complete type Moshe models.

Theorem 3.2. Let (Ω, D, F, P_o, L) be a complete type Moshe multistage decision model. Then, the principle of optimality holds.

Proof. Let S^* be any arbitrary optimal feasible strategy associated with the model. In contradiction to the statement specified by the theorem assume that there exist $n \varepsilon \natural$, $h_n^\circ \varepsilon H_n(S^*)$ and $S' \varepsilon SS$ such that $R_n(h_n^\circ, S') > R_n(h_n^\circ, S^*)$. Consider the strategy S^{**} defined as follows:

$$S_i^{**}(h_i) = \begin{cases} S_i^*(h_i) & , \quad i < n, \; h_i \varepsilon H_i \\ S_i'(h_i) & , \quad i \geq n \quad h_i \varepsilon H_i(S'|h_n^\circ, n) \\ S_i^*(h_i) & \qquad \text{otherwise} \end{cases}$$

where $H_i(S'|h_n^\circ, n) = \{h_{i,s'}(\bar{x}_i): \; \bar{x}_i \varepsilon \Theta_i(S'|h_n^\circ, n)\}$

is the set of all the histories observed under S' at i given that the modified problem (h_n°, n) is observed. From the structure of S^{**} it follows that $S^{**} \varepsilon SS$ and that

$$R_n(h_n^\circ, S^{**}) > R_n(h_n^\circ, S^*)$$

and

$$R_n(h_n, S^{**}) = R_n(h_n, S^*), \; \forall h_n \varepsilon \{h: h \varepsilon H_n, \; h \neq h_n^\circ\}$$

From the strict monotonicity of a type Moshe reward function and the fact that $h_n^\circ \varepsilon H_n(S^*) = H_n(S^{**})$ it follows then that $R(S^{**}) > R(S^*)$. This, however, contradicts the optimality of S^* and hence there exist no such $n \varepsilon \natural$, $h_n^\circ \varepsilon H_n$, and $S' \varepsilon SS$. It follows then that $R_n(h_n, S^*) \geq R_n(h_n, S)$, $\forall n \varepsilon \natural$, $h_n \varepsilon H_n(S^*)$, $S \varepsilon SS$.

In order to show that the principle of optimality holds for <u>reduced</u> type Moshe models in general, i.e., not necessarily for those associated with a trivial sufficient statistic, the results obtained from Theorem 2.1 and Theorem 3.2 will be combined to yield:

<u>Theorem 3.3</u>. Let (Ω, D, F, P_o, L) be a complete multistage decision model, T a sufficient statistic associated with it and $(\Omega, D', F', P_o, L')$ the corresponding reduced model. Then, if $(\Omega, D', F', P_o, L')$ is a type Moshe model the principle of optimality holds. $\rfloor$

<u>Proof</u>. Let G^* be any arbitrary optimal feasible strategy associated with the reduced model. Assume that there exist $n \varepsilon \mathbb{N}$, $u_n^\circ \varepsilon U_n(G^*)$ and $G' \varepsilon GG$ such that

$$R_n'(u_n^\circ, G') > R_n'(u_n^\circ, G^*)$$

From Theorem 2.1 it follows that $S^* = I^c(G^*)$ is an optimal feasible strategy for the complete model. Thus,

$$R_n(h_n^\circ, I^c(G')) > R_n(h_n^\circ, I^c(G^*))$$

for some $h_n^\circ \varepsilon H_n(I^c(G^*))$. This contradicts, however, Theorem 3.2 and thus there exists no such $n \varepsilon \mathbb{N}$, $u_n^\circ \varepsilon U_n(G^*)$ and $G' \varepsilon GG$ for which

$$R_n'(u_n^\circ, G') > R_n'(u_n^\circ, G^*)$$

and hence the principle of optimality holds. Notice that the structure of $I^c(G^*)$ implies that if $u_n^\circ \varepsilon U_n(G^*)$ then there exists at least one element $h_n^\circ \varepsilon H_n(I^c(G^*))$ such that $t_n(h_n^\circ) = u_n^\circ$. $\rfloor$

In the next section the relation between the principle and the algorithm will be discussed without restricting the investigation to a specific model.

3.3. The Relation Between the Principle of Optimality and the DP Algorithm

Suppose that it can be shown that the principle of optimality holds for a given multistage decision model. Does this information imply that the DP algorithm produces optimal feasible solutions? Similarly suppose that it can be shown that for a certain multistage decision model the DP algorithm produces optimal feasible solutions. Does this imply that the principle of optimality holds?

The DP algorithm and the principle of optimality have been thus far discussed in the context of specific models (type Shoshana and type Moshe). In order to answer the questions presented above for the general case, i.e., not necessarily for type Moshe/Shoshana models, the models under investigation will be such that both GG° and GG^* are not empty.

The questions raised above are extremely important from the theoretical viewpoint and have been raised by many investigators (Yakowitz, 1969; Hinderer, 1970; and others).

Theorems

Theorem 3.4. Let (Ω, D', F', P_o, L) be a reduced multistage decision model for which

(1) $GG^* \neq \emptyset$

and

(2) The principle of optimality holds.

Let K be any arbitrary element of $\mathbb{N}$ and GG° the set of strategies produced by the DP algorithm starting at K. Then,

$GG^{\circ} \subset GG^*$

In other words, the DP algorithm produces optimal feasible strategies.

Proof: It will be shown by induction on $n \leq K$ that for every $G^* \varepsilon GG^*$ there exists an element $G^n \varepsilon GG^n$ such that:

$$G^n_m(u_m) = G^*_m(u_m), \ \forall m \geq n, \ u_m \varepsilon U_m(G^*)$$

For $n = K$ the principle of optimality implies that

$$R'_K(u_K, G^*) \geq R'_K(u_K, G), \ \forall u_K \varepsilon U_K(G^*), \ G \varepsilon GG$$

Since $U_K(G^*) \neq \emptyset$ it follows then from the structure of the algorithm that there exists $G^K \varepsilon GG^K$ such that

$$G^K_m(u_m) = G^*_m(u_m), \ \forall m \geq K, \ u_m \varepsilon U_m(G^*),$$

and hence the inductive hypothesis is true for $n = K$. Assume that the inductive hypothesis is true for $n = K-1, K-2, \ldots, i$. In particular assume that it is true for $n = i$, i.e., for each $G^* \varepsilon GG^*$ there is an element $G^i \varepsilon GG^i$ such that

$$G^i_m(u_m) = G^*_m(u_m), \ m \geq i, \ u_m \varepsilon U_m(G^*)$$

Consider $n = i-1$, for which the principle of optimality implies that:

$$R'_{i-1}(u_{i-1}, G^*) \geq R'_{i-1}(u_{i-1}, G), \ \forall u_{i-1} \varepsilon U_{i-1}(G^*), \ G \varepsilon GG$$

From the structure of the algorithm then it follows that there exists $G^{i-1} \varepsilon GG^{i-1}$ such that

$$G^{i-1}_m(u_m) = G^*_m(u_m) \quad , \quad m \geq i-1, \ u_m \varepsilon U_m(G^*)$$

Notice that by the inductive hypothesis at $n = i$ it is guaranteed that such strategy exists.

Thus the inductive hypothesis is true for $n = i-1$ and hence it is true for all $n \leq K$. In particular it is true for $n = 1$, i.e., for each $G^* \varepsilon GG^*$ there is a strategy $G^1 \varepsilon GG^1$ such that

$$G^1_n(u_n) = G^*_n(u_n), \ \forall n \varepsilon \mathbb{N}, \ u_n \varepsilon U_n(G^*)$$

which implies that

$$R'(G^1) = R'(G^*) = R'^*$$

and thus $G^1 \varepsilon GG^\circ$. Since from the definition of the DP algorithm it follows that

$$R'(G') = R'(G''), \quad \forall G', G'' \varepsilon GG^\circ$$

$R'(G\ ) = R'(G^*)$ implies then that

$$R'(G^\circ) = R'^*, \ \forall G^\circ \varepsilon GG^\circ$$

and thus $GG^\circ c GG^*$.

An interesting question concerning therelation between the DP algorithm and the principle of optimality is the following one: suppose that for a given reduced model GG° is shown to be a subset of GG^*, in other words, it can be shown that all the strategies produced by the DP algorithm are optimal feasible. Does this imply that the principle of optimality holds?

The answer to the above question is provided by the following theorem.

Theorem 3.5. Let $(\Omega, D', F', P_o, L')$ be a RMDM for which the DP algorithm produces optimal feasible strategies, i.e.: $GG^\circ c GG^*$ for some $K \varepsilon \mathbb{N}$. Then, the principle of optimality does not necessarily hold.

Proof: The theorem will be proven by constructing a counter-example. Consider the complete multistage decision model (Ω, D, F, P_o, L) whose elements are as follows:

$$\Omega = \{\Omega_n: \ \Omega_n = \{x: \ x=0,1\}, \ n=1,2,3,4, \ \Omega_n = \{1\} \ n \geq 5)\}$$

$$D = \{D_n: \ D_n(h) = (d: \ d = d', d''), \ n = 1,2,3, \ D_n(h) = (d'), \ n \geq 4\}$$

$$F = \{f_n: \ f_n(h,d',1) = 1, \ f_n(h,d'',0) = 1\}_{n \varepsilon \mathbb{N}}$$

$P_o(1) = 1$

$L = \{L_n: \ L_n(h_\infty) = \pi_{i \geq n} x_i\}_{n \in \mathbb{N}}.$

It can be easily verified that $T = \{t_n(h) = x_n, \ h_n \varepsilon H_n\}_{n \in \mathbb{N}}$ is a sufficient statistic and thus $(\Omega, D', F', P_o, L')$ is a RMDM, where

$D' = \{D'_n: \ D'_n(x) = (d: \ d = d', d''), \ n = 1,2,3, \ D'_n(x) = (d'), \ n \geq 4\}$

$F' = \{f'_n: \ f'_n(x, d', 1) = 1, \ \ f_n(x, d'', 0) = 1\}_{n \in \mathbb{N}}$

$P_o(1) = 1$

$L' = \{L'_n: \ L'_n(x_n, d_n, x_{n+1}, \ . \ . \ .) = \pi_{i \geq n} x_i\}_{n \in \mathbb{N}}$

By inspection, it can be verified that

$$R'_n(x_n, \ G|x_{n+1}) = x_n \ \cdot \ R'_{n+1}(x_{n+1}, \ G)$$

and since $x_n \geq 0$, L' is a type Shoshana reward function.

It can be easily verified that

$$R^*_n(x) = \begin{cases} 1, & n \geq 5 \\ 0, & n \leq 3 \end{cases}$$

Since L' is a type Shoshana reward function and the model is both regular and truncated, Theorem 3.1 can be used to conclude that all the strategies produced by the DP algorithm are optimal feasible. Consider the strategy G^* having the following form:

$$G_1(x_1) = \begin{cases} d'' \ , & x_1 = 1 \\ d' \ , & x_1 = 0 \end{cases}$$

$$G_2(x_2) = \begin{cases} d' \ , & x_2 = 1 \\ d' \ , & x_2 = 0 \end{cases}$$

$$G^*_3(x_3) = \begin{cases} d' \ , & x_3 = 1 \\ d' \ , & x_3 = 0 \end{cases}$$

$$G^*_n(x_n) = d' \ , \quad n \geq 4$$

It can be verified that G* is feasible and that the only history observed under it is

$$H_\infty(G^*) = \{1, d'', 0, d', 1, d', 1, d', \ldots 1, d', 1, \ldots\}$$

Thus, $R(G^*) = 0$ and hence $G^* \varepsilon GG^*$. However, at $n = 3$ and $x_3 = 1$ the strategy G* yields:

$$R_3'(1, G^*) = 1 > R_3^*(1) = 0$$

Thus, the optimal feasible strategy G' is not optimal at a modified problem observe by it with positive probability and hence the principle of optimality does not hold.

Suppose that for a certain RMDM there is an optimal feasible strategy which is also optimal feasible at all the modified problems observed by it with positive probability. Does this imply that the principle of optimality holds?

Theorem 3.6. Let $(\Omega, D', F', P_o, L')$ be a RMDM and $GG^* \neq \emptyset$ its set of optimal feasible strategies. Then the fact that $R_n(u_n, G^*) = R_n^*(u_n)$, $\forall n \varepsilon \mathbb{N}$, $u_n \varepsilon U_n(G^*)$ for some $G^* \varepsilon GG^*$ does not guarantee that the principle of optimality holds.

Proof. The counterexample introduced in Theorem 3.5 indicates that at least one element of GG* is simultaneously optimal at all the modified problems it produces. However, as shown in Theorem 3.5, the principle of optimality does not hold.

Remarks. An important conclusion derived from Theorem 3.5 is that the optimality of the DP solutions does not guarantee that the principle of optimality holds. In other words, if the modified problem (u_n, n) is observed with positive probability by some $G^\circ \varepsilon GG^\circ$ there is no guarantee that there exists an element in GG° which is optimal feasible at this point. Moreover, suppose that the modified problem (u_n, n) is observed with u_n such that $u_n \notin U_n(GG^*)$ where $U_n(GG^*)$ is the subset of U_n whose elements are observed with positive probability by at least one element of GG^*. Is there some $G^* \varepsilon GG^*$ which is optimal feasible at (u_n, n)? A partial answer to this question will be provided in the following section.

3.4. The Optimality Equations and Hinderer's Comment

As was indicated earlier, Hinderer considers additive reward function so that:

$$R_n(h_n, S) = r_n(h_n, G_n(h_n)) + \sum_{x_{n+1} \varepsilon \Omega_{n+1}} R_{n+1}(h_{n+1}, S) \cdot f_n(h_n, G_n(h_n), x_{n+1})$$

where: $h_{n+1} = (h_n,\ G_n(h_n),\ x_{n+1})$.

Let $R_n^*(h_n)$ be the optimal feasible reward associated with the modified problem (h_n, n). It can be shown (Bellman, 1957; Dynkin, 1965, and others) that for regular models any optimal feasible strategy satisfies what Hinderer (1970, p. 21) calls the *systems of optimality equations:*

$$R_n^*(h_n) = \max_{d \varepsilon D_n(h_n)} [r_n(h_n, d) + \sum_{x_{n+1} \varepsilon \Omega_{n+1}} R_{n+1}^*(h_{n+1}) \cdot f_n(h_n, d, x_{n+1})]$$

for all $n \varepsilon \mathbb{N}$, $h_n \varepsilon H_n$.

While discussing the relation between the principle of optimality and the systems of optimality equations (OE), Hinderer states the following:

> . . . the importance of the principle does not rest so much on the fact that it furnishes a necessary condition for the optimality of a policy but in the fact that it is often regarded as a convenient tool for deriving the *optimality equation* (OE) . . . which

> on its part is the starting point for many investigations in dynamic programming. However, to the best of our knowledge there has never been given a rigorous proof of the OE in the general case by means of the principle, though the proofs of the OE and the principle show some similarities . . . we shall give a proof of the OE by means of the principle under rather restrictive assumptions . . . we also remark that sometimes in the literature the principle and the OE are regarded as the same statement, though these are definitely two different things . . .(Hinderer, 1970, p. 14).

Translating Hinderer's comment to the context of the multistage decision model developed in this study requires first the definition of the term "optimality equations."

Definitions and Theorems

<u>Definition 3.7</u>. Let (Ω, D, F, P_o, L) be a CMDM for which L' is separable under expectation. We say that the *system of optimality equations*:

$$C_n(h_n) = \max_{d \varepsilon D_n(h_n)} \sum_{x_{n+1} \varepsilon \Omega_{n+1}} \rho_n(h_n, d, R^*_{n+1}(h_{n+1})) \cdot f_n(h_n, d, x_{n+1})$$

holds if

$$R^*_n(h_n) = C_n(h_n) \quad \forall n \varepsilon \mathbb{N}, \; h_n \varepsilon H_n$$

where $R_n{}^*(h_n)$ is the optimal feasible reward associated with the modified problem (h_n, n).

It will be shown that the system of optimality equations holds for a regular type Shoshana model by showing first that for a regular type Shoshana model there is an optimal feasible strategy which is also optimal for all modified problems.

<u>Theorem 3.7</u>. Let (Ω, D, F, P_o, L) be a regular type Shoshana model and SS its set of feasible strategies. Then there is an element $S' \varepsilon SS$ such that:

$$R_n(h_n, S') = R^*_n(h_n) \quad , \quad \forall n\in\mathbb{N},\ h_n\in H_n$$

where $R^*_i(h_i)$ is the optimal feasible reward associated with the modified problem (h_i,i).

Proof. Let $SS^{h_n,n}$ be the set of all the optimal feasible strategies $S^{h_n,n}$ associated with the modified problem (h_n,n). Consider any arbitrary modified problem (h,n) and any arbitrary element $S^{h,n}$ of $SS^{h,n}$. Construct the sequence $\{S^i\}_{i\geq n}$ of strategies as follows:

For $i = n$ set

$$S^n_j(h_j) = S^{h,n}_j(h_j) \quad , \quad \forall j\in\mathbb{N},\ h_j\in H_j.$$

For $i > n$ set

$$S^i_j(h_j) = \begin{cases} S^{i-1}_j(h_j) & j < i \quad , \quad h_j\in H_i \\ S^{h_i,i}_j(h_j) & j \geq i \quad , \quad h_j = (h_i, d_i, \ldots) \end{cases}$$

where $S^{h_i,i}$ is an arbitrary element of $SS^{h_i,i}$. By induction on $i \geq n$ it will be shown that:

$$R_n(h,S^i) \geq R_n(h,S^{h,n}).$$

For $i = n$ the inductive hypothesis is true by the structure of S^n. Assume that the inductive hypothesis is true for $i = n+1, n+2, \ldots m$. In particular assume that it is true for $i = m$, i.e.:

$$R_n(h,S^m) \geq R_n(h,S^{h,n})$$

Consider $i = m+1$ for which the structure of S^{m+1} implies that

$$R_{m+1}(h_{m+1}, S^{m+1}) \geq R_{m+1}(h_{m+1}, S^m), \ \forall h_{m+1}\in H_{m+1}$$

From the monotomicity of L it follows that

$$R_n(h,S^{m+1}) \geq R_n(h,S^m)$$

and thus

$$R_n(h, S^{m+1}) \geq R_n(h, S^{h,n})$$

and the inductive hypothesis is true for $i = m+1$ and hence it is true for all $i \geq n$. Let $S^{*h,n} = \lim_{i\to\infty} S^i$ for which

$$R_n(h,\ S^{*h,n}) \geq R_n(h,\ S^{h,n}) = R_n^*(h).$$

Notice that $S^{*h,n}$ is feasible. Now construct the strategy S' as follows:

$$S'_m(h_m) = \begin{cases} S_m(h_m) & m < n \quad , \quad h_m \varepsilon H_m \\ S_m^{*h_n,n}(h_m) & m \geq n \quad , \quad h_m = (h_n,\ d_n,\ \ldots) \end{cases}$$

where S is any arbitrary element of SS. From the inductive hypothesis it follows that

$$R_n(h_n,\ S') \geq R_n(h_n,\ S^{h_n,n}) \quad , \quad \forall n \varepsilon \mathbb{N},\ h_n \varepsilon H_n$$

and hence the theorem is true.

It will be shown now that the system of optimality equations holds for any regular type Shoshana model.

<u>Theorem 3.8</u>. Let (Ω, D, F, P_o, L) be a regular type Shoshana complete multistage decision model. Then, the system of optimality equation holds.

<u>Proof</u>. By definition,

$$C_n(h_n) = \max_{d \varepsilon D_n(h_n)} \sum_{x_{n+1} \varepsilon \Omega_{n+1}} \rho_n\ (h_n, d, R_{n+1}^*(h_{n+1}))\ f_n(h_n, d, x_{n+1})$$

with $h_{n+1} = (h_n,\ d,\ x_{n+1})$

Since any optimal feasible strategy at (h_n, n), say S^{*n}, is feasible, it follows that:

$$R_n^*(h_n) \leq C_n(h_n).$$

Suppose that there exists $d^* \varepsilon D_n(h_n)$ such that

$$\sum_{x_{n+1} \varepsilon \Omega_{n+1}} \rho_n(h_n, d^*, R_{n+1}^*(h_{n+1}))\ f_n(h_n, d^*, x_{n+1}) > R_n^*(h_n)$$

This contradicts Theorem 3.7, since it implies that there is no optimal feasible strategy for both h_n and all the elements $h_m \varepsilon H_m$. Thus,

$$C_n(h_n) \leq R_n^*(h_n)$$

which yields that

$$R_n^*(h_n) = C_n(h_n), \ \forall n \varepsilon \mathbb{N}, \ h_n \varepsilon H_n$$

and hence the theorem is true.

<u>Remarks</u>. (1) It was shown that the system of optimality equations holds for regular type Shoshana models (Theorem 3.8). It was also shown (Theorem 3.5) that the principle of optimality does not necessarily hold for type Shoshana models, and thus the principle of optimality is <u>not a necessary condition</u> for the system of optimality equations to hold.

(2) Let $H_n(SS^*)$ be the set of all histories observed with positive probability under at least one optimal feasible strategy. Then one can use the principle of optimality to show that:

$$R_n^*(h) = C_n(h_n), \ \forall n \varepsilon \mathbb{N}, \ h_n \varepsilon H_n(SS^*).$$

This, however, does not provide an answer as to histories not included in $H_n(SS^*)$.

(3) The last two theorems are concerned with complete models. Using the analysis presented in Chapter 2, it can be shown that the last two theorems are <u>valid also for reduced models.</u>

It will be shown now how the validity of the optimality equations can be proven by <u>means of the principle of optimality</u> by imposing certain conditions on the structure of the model.

<u>Lemma 3.2</u>. Let (Ω, D, F, P_o, L) be a regular type Moshe complete multistage decision model with F such that:

$$f_n(h_n, d, x_{n+1}) > 0 \quad , \quad \forall n \varepsilon \mathbb{N}, \ h_n \varepsilon H_n, \ d \varepsilon D_n(h_n), \ x_{n+1} \varepsilon \Omega_{n+1}.$$

Then, the system of optimality equations holds.

Proof. Notice that since any type Moshe model is also a type Shoshana model, Theorem 3.8 can be used to show that the above lemma is true. However, the objective is to prove the lemma by means of the principle of optimality. It is known from Theorem 3.2 that the principle of optimality holds for a type Moshe model. Also, it is known that:

$$R_n^*(h_n) \leq C_n(h_n) \qquad \forall n \in \mathbb{N},\ h_n \in H_n$$

Since for the model under consideration $H_n(SS^*) = H_n$ the definition of the principle of optimality implies that

$$R_n^*(h_n) \geq C_n(h_n) \qquad \forall n \in \mathbb{N},\ h_n \in H_n$$

because otherwise there will be a contradiction to the validity of the principle. Thus,

$$R_n^*(h_n) = C_n(h_n), \ \forall n \in \mathbb{N},\ h_n \in H_n .$$

The assumptions made in Lemma 3.2 are not as restrictive as those made by Hinderer. Moreover, Lemma 3.2 deals with a type Moshe model while Hinderer restricts his proof only to models with additive reward functions.

3.5. Bellman's Multistage Decision Model

When modeling a certain multistage decision process, L_1 is usually uniquely determined by the process. The other elements of L, i.e., $\{L_n\}_{n>1}$, are constructed in such a way that the resulting L may be handled by the available solution methods such as DP. Thus, in many cases L is determined subjectively, so to speak. Suppose that a certain decision process is investigated and L_1 is the function of interest. One can always set:

$$L_n(h_\infty) = L_1(h_\infty), \ \forall n \in \mathbb{N},\ h_\infty \in H_\infty$$

or for the reduced model

$$L'_n(u_n,d_n,x_{n+1}, \ldots) = L'_1(t_1(x_1),d_2,x_2, \ldots x_n,d_n,x_{n+1}, \ldots)$$

for all $n\varepsilon\mathbb{N}$, $u_n = t_n(h_n)$, $h_n = (x_1,d_1,x_2, \ldots x_n)$.

The above structure of $L' = \{L'_n\}_{n\varepsilon\mathbb{N}}$ implies that

$$R'_n(u_n,G) = \sum_{x_{n+1}\varepsilon\Omega_{n+1}} R_{n+1}(u_{n+1},G) \cdot f'_n(u_n,G_n(u_n),x_{n+1})$$

with $u_{n+1} = V_n(u_n, G_n(u_n), x_{n+1})$.

Obviously L' is a type Moshe reward function so that the principle of optimality holds and the DP algorithm produces optimal feasible solutions. Since no assumption concerning L'_1 is made, it follows then that the principle of optimality holds for all those multistages for which L' is as described above, no matter what the structure of L'_1 is.

Definition, Theorem, and Example

Definition 3.8. Let (Ω, D', F;, P_o, L') be a RMDM for which L' is such that

$$L' = \{L'_n: L'_n(t_n(h_n),d_n,x_{n+1}, \ldots) = L'_1(t_1(x_1),d_1, \ldots,x_n, d_n, \ldots)$$

with $h_n = (x_1,d_1,x_2, \ldots, x_n)\varepsilon H_n\}_{n\varepsilon\mathbb{N}}$.

Then L' is said to be a *type Bellman reward function* and the model a *type Bellman model*.$\lrcorner$

Theorem 3.9. Let (Ω, D', F, P_o, L') be a type Bellman model for which $GG^* \neq \emptyset$. Then the principle of optimality holds and $GG^\circ \subset GG^*$.

Proof. From the definition of a type Bellman model it follows that: $R'_n(u_n,G|x_{n+1}) = R_{n+1}(u_{n+1},G)$, $u_{n+1} = V_n(u_n,G_n(u_n),x_{n+1})$ so that L' is a type Moshe model and thus from Theorem 3.3 and 3.4 it follows that the principle of optimality holds and that $GG^\circ \subset GG^*$.$\lrcorner$

The last theorem is rather interesting, because it implies that the principle of optimality holds for all multistage decision models having a type Bellman reward function. Since when formulating the model often only L_1' is specified, it implies that every multistage decision model may be formulated as a type Bellman model. This implies that the principle of optimality holds essentially for all multistage decision models in the sense that every multistage decision model may be formulated also as a type Bellman model. However, as far as the practical implications of the above discussion are concerned, it should be noted that for such models the amount of computation involved in the implementation of the DP algorithm is very close to the amount needed for total enumeration. Thus, the validity of Theorem 3.9 is significant as far as theory and modeling are concerned but does not improve the situation as far as solutions procedure are concerned. For a type Bellman model, it can be written then that:

$$R'(G|u_{n,n}) = R_n'(u_n,G) \quad \forall n\varepsilon\mathbb{N},\ u_n\varepsilon U_n,\ G\varepsilon GG$$

where $R'(G|u_n,n)$ is the total reward given that the modified problem (u_n,n) is observed, and from n on the strategy G is used. Thus, u_n should include all the information needed for evaluating R' which in most cases results in a rather large set U_n as far as the number of elements in U_n is concerned. This implies that many dynamic programming equations have to be solved and thus a relative heavy computational load is expected.

The advantage then of not using a type Bellman model has to do with the dimension of U_n.

Example 3.5. Consider the complete multistage decision model (Ω, D, F, P_o, L) for which L_1 has the following structu :

$$L_1(x_1,d_1,x_2, \ldots) = \sum_{i\geq 1} r_i(x_i,d_i)$$

where r_i is a real valued function defined on Ω_i X $\mathbb{D}$. Obviously, if the objective is the maximization of R, and if a type Bellman reward function is used as L, any sufficient statistic to be introduced should have the property that one of the coordinates of u_n should indicate the quantity $\sum_{i=1}^{n-1} r_i(x_i,d_i)$. This implies that at least this coordinate of u_n may take many values, especially for large n, so that at some n = K, where the DP algorithm starts (suppose that the model is truncated at K) many dynamic programming equations have to be solved.

If instead L is such that

$$L_n(x_1,d_1, \ldots) = \sum_{i\geq n} r_i(x_i,d_i)$$

then when constructing a sufficient statistic, none of the coordinates of u_n is required to indicate the quantity $\sum_{i=1}^{n-1} r_i(x_i,d_i)$ and thus the dimension of U_n is reduced as compared with the type Bellman model. The computational and modeling aspects of the DP algorithm will be discussed in Chapter 4 and Chapter 5, respectively.

3.6. General Discussion

One of the basic difficulties involved in comparing different formulations of the DP algorithm and the principle of optimality has to do with the different contexts (models) in which they are defined. The comparison becomes even more difficult due to the lack of formal definitions of the algorithm and the principle in certain works. It was chosen to relate this work to works done by Bellman (1954, 1957), Denardo (1965) and Hinderer (1970) with the understanding that the differences in the models used in each case prevents a complete and detailed comparison.

3.7. The DP Algorithm

As indicated by Bellman (1957, p. 85), the DP algorithm for stochastic processes with countably many decision stages and state elements has the following form (using our notation):

$$R^*|_{h_n,n} = \max_{d \varepsilon D_n(h_n)} \sum_{x_{n+1} \varepsilon \Omega_{n+1}} R^*|_{h_{n+1},n+1} \cdot f_n(h_n,d,x_{n+1}).$$

where $R^*|_{h_n,n}$ is the optimal value of the total reward given that (h_n,n) is observed and an optimal feasible strategy as far as $R|_{h_{n+1},n+1}$ is concerned is used. Thus, as defined by Bellman, the DP algorithm is used for type Bellman models. When applying the algorithm, Bellman had demonstrated that for certain reward functions, similar results may be obtained by using the relation:

$$R_n^*(h_n) = \max_{d \varepsilon D_n(h_n)} \sum_{x_{n+1} \varepsilon \Omega_{n+1}} \rho_n(h_n,d,R_{n+1}^*(h_{n+1}))\ f_n(h_n,d,x_{n+1})$$

with $h_{n+1} = (h_n,d,x_{n+1})$.

Most of the examples used by Bellman in his early publications were such that L was additive.

Mitten (1964) and later Denardo (1965) have shown that the DP algorithm may also be used for reward functions having (a) certain monotomocity (type Shoshana) property, and (b) certain convergence properties.

The model introduced in Chapter 2 does not require any convergence properties but instead for type Shoshana models it requires that there will be a simultaneous optimal feasible solution at all modified problems associated with the K-th decision stage, where the DP algorithm starts. Moreover, for type Moshe models it was shown that the DP algorithm may be used even if the above condition is not satisfied, and that the only requirement for this case is that for each $u \in U_K$ there will be at least one optimal feasible solution.

Hinderer (1970) does not discuss the DP algorighm in the framework of a solution procedure but rather uses the system of optimality equations to describe the relation between the optimal solutions of successive modified problems. As indicated earlier, Hinderer's model is restricted to additive reward functions only.

3.8. The Principle of Optimality

It is extremely important to read the definition of the principle of optimality (not necessarily the version introduced here) in the context of the model used to describe the decision process under consideration. Much of the criticism surrounding Bellman's "version," for example, (Denardo, 1965, p. 36; Hinderer, 1970, p. 14; and others) could have been partially avoided by interpreting it in the context it was originally introduced. It is not suggested here that Bellman's definition is absolutely clear in the context it was introduced, but rather that certain amount of the ambiguity often related to it may have been avoided.

As introduced by Bellman (1957, p. 83) for a deterministic type Bellman model, the definition is as follows: "PRINCIPLE OF OPTIMALITY. An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision . . ." (Bellman (1957, p. 83). The formal proof provided by Bellman to support the above statement is as follows: "A proof by induction is immediate" (Bellman, 1957, p. 83).

While the above statements are true for type Bellman models (Theorem 3.9) it can be shown (Theorem 3.5) that there are models for which the above statements do not hold. As indicated by Yakowitz (1969, p. 43), "The principle of optimality must be proved to be consistent with the criterion already established."

It seems as if the basic cause for the ambiguity surrounding Bellman's principle has to do with the notion of optimality used when defining the optimal solution. Since Bellman defines his optimal policy as: "An optimal policy is a policy which maximizes a preassigned function on the final state variables . . ." (Bellman, 1957, p. 82), the above statement is meaningful then only in the context of optimality criteria related to the final state variables and thus, the principle of optimality as defined by Bellman should be read in the context of type Bellman models.

Denardo (1965, p. 37) introduced another version of the principle of optimality which in the context of the model developed here may be described by Theorem 3.7. In other words, Denardo's principle states that for certain decision models there exists an optimal feasible solution which is also optimal feasible at all modified problems.

Hinderer (1970, p. 9) is using the notion of $\bar{p}$-optimal to indicate that it is not required from the optimal strategy to be optimal feasible at every $x_1 \varepsilon \Omega_1$. This is exactly the notion of an optimal strategy used here (see Definition 2.18 and 2.26). As far as Hinderer's principle of optimality is concerned (Hinderer, 1970, p. 19 [Theorem 3.8]), again it is equivalent to the notion introduced here which may be also considered as a modification of Yakowitz's (1969) version.

To the best of the author's knowledge, Theorem 3.3 provides the most general conditions for the validity of the principle of optimality in the context of discrete models.

3.9. The Relation Between the DP, PO, and OE

As was indicated by Hinderer (see Section 4 of this chapter), there appears to be certain ambiguity concerning the relation between the principle of optimality, the dynamic programming algorithm, and the system of optimality equations. Similarly, Yakowitz (1969, p. 43), when discussing the relation between the DP algorithm and the principle of optimality states, in the context of his adaptive control process (ACP):

> A rather puzzling situation is that in the engineering literature, DP is used freely, and for the most part correctly, to obtain solutions to statistical problems. When authors justify their procedures, it is usually by appealing to the principle of optimality . . . which is often copied verbatim. Such an exposition is odd in two respects: First, the principle of optimality should not be stated axiomatically, since the ACP problem already has sufficient structure to define a solution. The principle of optimality must be proved to be consistent with the criterion already established. This author is unaware of such a published proof and finds it difficult to supply. Second, the principle tells us that any solution to a problem must have the property that it is also a solution to all modified problems which occur. This does not imply that a strategy constructed by the DP to have his property is necessarily a solution. That a strategy so constructed is a solution is the statement of the DP theorem for ACP's, which was relatively easy to prove. In our analysis, we have fully justified the use of DP without reference to the principle of optimality . . . (Yakowitz, 1969, p. 43).

Theorem 3.4 as introduced in this chapter may be used to "explain the engineers'" appeal to the principle of optimality. In other words, if one can show that for a given model the principle of optimality holds (as defined in this chapter), then the strategies produced by the DP algorithm are optimal feasible.

It should be noted that while for type Shoshana models it was shown that the DP solutions are optimal feasible <u>without appealing to the principle of optimality</u>, for the <u>general case it was necessary to appeal to the principle</u>.

To the best of the author's knowledge, Theorem 3.4 is the <u>only</u> proof available in the literature to the claim that the existence of an optimal solution and the validity of the principle of optimality imply that the DP solutions are optimal feasible.

On the other hand, Theorem 3.5 implies that the principle of optimality is not a necessary condition for the DP algorithm to provide optimal feasible solutions, and thus it is concluded that at most, the principle of optimality is a sufficient condition (together with the condition $GG^* \neq \emptyset$) for the DP solutions to be optimal feasible.

It is, therefore, recommended that statements like: "The mathematical formulation of the principle of optimality is called dynamic programming . . ." and "Dynamic programming is a method of decomposition based upon Bellman's principle of optimality . . ." (Beveridge and Schechter, 1970, p. 679), be carefully examined before introduced into textbooks.

CHAPTER 4

THE ROLE OF ANALYTICAL CONSIDERATIONS IN THE IMPLEMENTATION OF THE DP ALGORITHM -- AN EXAMPLE

One of the limitations of the DP algorithm as a solution procedure for multistage decision problems has to do with the amount of computation and storage requirements involved in solving the DP equations. Consider, for example, a Markovian model truncated at N for which $\Omega_i = \Omega_j = \bar{\Omega}$, $i, j \leq N$ and $D_n(x) = \phi \ \forall\ n \leq N$, $x \varepsilon \bar{\Omega}$. Let $M(\phi)$ and $M(\bar{\Omega})$ be the number of elements in ϕ and $\bar{\Omega}$ respectively and assume that they are finite. One can solve the multistage decision problem associated with the above model by total enumeration. In this case, the set GG may be constructed by construting $N_T = M(\phi)^{N \cdot M(\bar{\Omega})}$ strategies and conducting $N_T - 1$ comparisons. For $N = M(\phi) = M(\bar{\Omega}) = 100$, which are often encountered in large scale systems, this amounts to $N_T = 10^{20,000}$ which is a rather heavy computational load. If instead, the DP algorithm is used, starting at N, $N \cdot M(\bar{\Omega}) = 10^4$ dynamic programming equations have to be solved, each of which requires $M(\phi) = 100$ iterations, which amounts to $N_{DP} = 10^6$ computations of rewards, and $N \cdot M(\bar{\Omega}) \cdot M(\phi - 1) = 99 \cdot 10^4 \approx 10^6$ comparisons. Although the above problem can be handled rather easily by the present generation of computers, using the DP algorithm, the computation cost may be high especially in cases where sensitivity analysis is needed. In any event, it is desired to develop procedures for solving the DP equations, which when solving a given equation will not require a complete search over the elements of $\mathfrak{p}$. In other words, procedures other than "crude dynamic programming" are desired.

In this chapter, the potential role of analytical considerations in reducing the computational load associated with the implementation of the DP algorithm is demonstrated. As will be demonstrated later, in certain models there is a close relationship between the optimal strategies associated with the modified problems at a given decision stage. More specifically, suppose for example that the modified problem (u_n,n) is to be solved and that $G^{u_n,n}$ is found to be an optimal feasible strategy. In certain situations the optimality of $G^{u_n,n}$ at (u_n,n) implies that the optimal strategy at (u_n',n) for some $u' \varepsilon U_n$, may be determined by searching only on a "small" subset of $D_n'(u_n')$, determined by $G^{u_n,n}$.

It will be demonstrated how certain characteristics of the elements of the model may significantly reduce the amount of computation involved in the implementation of the DP algorithm. It should be realized however, that the investigation presented here should be considered a demonstration of the potential role of analytical consideration in the implementation of the DP algorithm and not a method as such. Without underestimating the computation load that the present generation of computers can handle and more so with respect to the future generations, the role of analytical methods should not be ignored.

The two examples to be introduced in the discussion are simple cases of the class of models to be referred to as "mass balance type of models."

4.1. Mass Balance Type of Models

Many natural and management processes are based on the principle of mass conservation, as far as the dynamics of the process is concerned. The mass balance equation can be written schematically as follows:

Change in storage = Input - Output

Often the input and/or the output are decision variables involved in a multistage decision process. For example, many reservoir control processes are characterized by a mass balance equation and so are many inventory situations. Let x_n be the storage level at time n, O_n the output during (n, n+1) and q_n the input during (n, n+1). Thus,

$$x_{n+1} = x_n + q_n - O_n$$

The investigation will be restricted to situations where the output O_n is determined (either deterministically or statistically) by a decision variable d_n and q_n is a realization of a random variable $\tilde{q}_n$ whose distribution function is known. As often done when implementing the DP algorithm, x_n, d_n, q_n, and O_n are assumed to be integers. In terms of the decision variable, d_n, the mass balance equation can be written as follows:

$$x_{n+1} = \begin{cases} M_n & x_n + q_n - d_n \geq M_{n+1} \\ x_n + q_n - d_n & m_{n+1} < x_n + q_n - d_n < M_{n+1} \\ m_n & m_{n+1} \geq x_n + q_n - d_n \end{cases}$$

where M_{n+1} and m_{n+1} are the maximum and minimum storage levels allowed at time n+1, respectively. The relation between d_n and O_n can be written then as follows:

$$O_n(d_n, q_n, x_n) = \begin{cases} M_n - m_n & x_n + q_n - d_n \geq M_{n+1} \\ d_n & M_{n+1} > x_n + q_n - d_n > m_{n+1} \\ x_n + q_n - m_n & x_n + q_n - d_n \leq m_{n+1} \end{cases}$$

It is given then that $m_n \leq x_n \leq M_n$, $\forall n \varepsilon \mathbb{N}$ where m_n and M_n are positive integers.

Let P_n be the probability mass function of $\tilde{q}_n$ and assume that $P_n(q_n) = 0$, $\forall q_n > MQ_n$, for some positive integer MQ_n. P_n can be used then to construct the conditional probability mass function of ξ_{n+1} given ξ_n and d_n where ξ_i is the random variable whose realization is denoted by x_i, $i \varepsilon \mathbb{N}$, i.e.:

$$P_r(x_{n+1} | x_n, d_n) = \begin{cases} P_n(q_n \varepsilon B_n) & x_{n+1} = m_{n+1} \\ P_n(q_n = x_{n+1} - x_n + d_n) & m_{n+1} < x_{n+1} < M_{n+1} \\ P_n(q_n \varepsilon A_n) & x_{n+1} = M_{n+1} \end{cases}$$

where: $B_n = \{q_n : q_n \leq x_n - d_n - m_{n+1}\}$ and $A_n = \{q_n : q_n \geq M_{n+1} - x_n + d_n\}$

For the purpose of the investigation, all RMDM having the above form for f'_n will be defined as mass balance type of models.

Definition 4.1

Let $(\mathfrak{R}, D', F', P_o, L')$ be a Markovian multistage decision model with the following structure:

(1) $\mathfrak{R} = \{\Omega_n : \Omega_n = \{m_n, m_{n+1}, \ldots, M_n\}, n \varepsilon \mathbb{N}\}$

(2) $D'_n(x_n) = \{d: d = md_n(x_n), md_n(x_n)+1, \ldots, MD_n(x_n)\}_{n \varepsilon \mathbb{N}}$

(3)

$$f'_n(x_n, d, x_{n+1}) = \begin{cases} P_n(q_n \varepsilon B_n(x_n, d)) & , \quad x_{n+1} = M_{n+1} \\ P_n(q_n = x_{n+1} - x_n + d_n) & , \quad m_{n+1} \quad x_{n+1} \quad M_{n+1} \\ P_n(q_n \varepsilon A_n(x_n, d)) & , \quad x_{n+1} = m_{n+1} \end{cases}$$

where $\tilde{q}_n$ is a random variable with a probability mass function P_n, $n\in\mathbb{N}$,

$A_n(x_n,d) = \{q: \; x_n + q_n - d_n \le m_{n+1}\}$ and

$B_n(x_n,d) = \{q: \; x_n + q_n - d_n \ge M_{n+1}\}$.

(4) $L'_n(x_n,d_n,x_{n+1},\ldots) = \sum_{i\ge n} r_i(x_i,d_i), \quad n\in\mathbb{N}.$

where: m_n, M_n, md_n, MD_n are all integers and r_i is a real valued function defined on $\mathbb{R}^2$, $\forall i\in\mathbb{N}$. Then, $(\Omega, D', F', P_o, L')$ is said to be a *standard mass balance type of model.*

In the context of reservoir control processes the elements of a mass balance type of model may be interpreted as follows:

n = time of release

x_n = storage level in the reservoir at time n.

d_n = target release for the period (n, n+1)

q_n = inflow to the reservoir during (n, n+1)

$r_n(x_n,d_n)$ = the expected value of the reward when the target release is d_n and the storage level is x_n.

m_n = minimum storage level allowed at time n.

M_n = maximum storage level allowed at time n.

$md_n(x_n)$ = minimum target release allowed at time n if the storage level is x_n.

$MD_n(x_n)$ = maximum target release allowed at time n if the storage level is x_n.

If the time-horizon under consideration is finite, say N, $r_n(x_n,d_n) \equiv 0$ $\forall n>N$, and dummy state and decision elements are constructed for $n > N$.

Notice that since $r_n(x_n,d_n)$ was defined as the expected value of the reward associated with x_n and d_n, it may include a penalty often imposed whenever shortage or overflow is realized. The objective is to construct a release strategy that will maximize the expected value of the total benefit (reward). As far as the initial condition is concerned, a special form for P_o is not required.

Two types of reward functions will be considered; one corresponds to $\{r_n\}_{n \leq N}$ for which every r_n is concave and the other to the case where r_n are all convex.

Example 4.1

Let $(\Omega, D', F', P_o, L')_N$ be a truncated mass balance type of model for which

(1) $m_n = o$, $md_n(x_n) = o$, $MD_n(x_n) = \max\{x_n, MD_n\}$, $n = 1,2, \ldots N$, $x_n \varepsilon \Omega_n$.

(2) $q_n \varepsilon Q_n = \{0, 1, \ldots, MQ_n\}$, $n \varepsilon \mathbb{N}$.

(3) $r_n(x_n, d_n) = r_n'(d_n)$, $n = 1,2, \ldots N$, $\quad x_n \varepsilon \Omega_n$

where r_n' is a concave monotone increasing function, and $r_n'(d_n) \equiv 0 \quad n > N$. It will be shown that there exists an optimal feasible strategy $G^* \varepsilon GG$ such that:

$$G_n^*(x+1) \varepsilon \{G_n^*(x), G_n^*(x_n)+1\}, \forall n \varepsilon \mathbb{N},$$

<u>Proof</u>: The following notation will be used:

(1) $y_n = x_n - d_n$, (Notice that $M_n \geq y_n \geq 0$).

(2) $R_n^*(x) = \max\limits_{G \varepsilon GG} R_n'(x, G)$

(3) $\ell_n(y_n) = \sum\limits_{q_n \varepsilon Q_n} R_{n+1}^*(y_n + q_n) \cdot P_n(q_n)$

(4) For simplicity $r_i(d_i)$ will be used for $r_i'(d_i)$.

Using the inductive hypothesis:

(t) (t.1) ℓ_n is monotone increasing function of y_n, and $2\ell_n(y) \geq \ell_n(y-1) + \ell_n(y+1)$

(t.2) $G_n^*(x+1) \varepsilon \{G_n^*(x), G_n^*(x)+1\}$, $\forall x \varepsilon \Omega_{n+1}$, for some $G^* \varepsilon GG^*$

(t.3) R_n^* is monotone increasing function of x_n and $2 R_n^*(x) \geq R_n^*(x-1) + R_n^*(x+1)$.

it will be shown that at least one of solutions obtained by the DP algorithm, starting at K = N statisfies the relation indicated above and since the model is a type shoshana model, from theorem 3.1 it follows that this solution is optimal feasible. Let start the DP algorithm at K = N, in which by the definition of the algorithm all the elements of GG^K are optimal feasible for all the modified problems (x_K, K) with $R_K^*(x) = r_K(\min \{x, MD_n\})$ and $G_K^K (x_K) = \min \{x_K, MD_n\}$. By inspection, using the structure of r_K it follows that the inductive hypothesis is true for n = K. Assume that the inductive hypothesis is true for n = K-1, K-2, . . .m. In particular assume that it is true for n = m, i.e.,

(1) ℓ_n is monotone increasing and $2\ell_n(y) \geq \ell_n(y-1) + \ell_n(y+1)$.

(2) $G_m^m (x) \leq G_m^m(x+1) \leq G_m^m(x) + 1$, $\forall x \varepsilon \Omega_n$ for some $G_m^m \varepsilon GG^m$

(3) R_m^* is monotone increasing and $2R_n^*(x) \geq R_n^*(x-1) + R^*(x+1)$

Consider now n = m-1 = i. By definition, ℓ_i can be written as follows:

$$\ell_i(y) = \sum_{q \varepsilon Q_i} R_m^* (y+q) \cdot P_i(q) \quad , \quad 0 \leq y_i \leq M_i$$

Consider any element y from the set {1, 2, . . ., M_i-1}, for which

$$\ell_i(y+1) = \sum_{q \varepsilon Q_i} R_m^* (y+1+q) \cdot P_i(q)$$

$$\ell_i(y) = \sum_{q \varepsilon Q_i} R_m^* (y+q) \cdot P_i(q)$$

$$\ell_i(y-1) = \sum_{q \in Q_i} R^*_m(y-1+q) \cdot P_1(q)$$

Since R^*_m is monotone increasing (under the inductive hypothesis at n=m) it follows that ℓ_i is also monotone increasing.

Let $\Delta(y) = 2\ell_i(y) - \ell_i(y+1) - \ell_i(y-1)$, which can also be written as $\Delta(y) = \sum_{q \in Q_i} [2R^*_m(y+q) - [R^*_m(y+1+q) + R^*_m(y-1+q)]] \cdot P_i(q)$.

From the inductive hypothesis at n=m it follows then that:

$$2\, R^*_m(y+q) \geq R^*_m(y+1+q) + R^*_m(y-1+q)$$

so that $\Delta(y) = 2\ell_i(y) - \ell_i(y+1) - \ell_i(y-1) \geq 0$ and thus (t.1) is true for $i = m-1$. Let x be any element of the set $\{1, 2, \ldots, M_i-1\}$ and let $d^i = G^i_i(x)$ for some arbitrary $G^i \varepsilon GG^i$. Thus,

$$r_i(d^i) + \ell_i(x-d^i) \geq r_i(d) + \ell_i(x-d),\ \forall d \varepsilon D_n'(x)$$

Notice that since $D_n'(x)$ is finite and L' is a type Moshe function, $R^*_i(x) = r_i(d^i) + \ell_i(x-d^i) \geq r_i(d) + \ell_i(x-d)$, $\forall d \varepsilon D_n'(x)$ and $R^*_i(x)$ exists. In particular, $r_i(d^i) + \ell_i(x-d^i) \geq r_i(d) + \ell_i(x-d)$, $\forall d \leq d^i$. Since from (t.1) at i it is given that $\ell_i(x+1-d^i) - \ell_i(x-d^i) \geq \ell_i(x+1-d) - \ell_i(x-d)$, $\forall d \leq d^i$ it follows then that, $r_i(d^i) + \ell_i(x \pm 1-d^i) \geq r_i(d) + \ell_i(x \pm 1-d)$, $\forall d < d^i$ which implies that $G^i_i(x+1) \geq d^i = G_i{}^i(x)$. Notice that if d^i is feasible for (x,i) it is also feasible for (x+1,i). Suppose now that there exists $\delta \geq 2$ such that $(d^*+\delta) \varepsilon D_i(x+1)$ for which

$$r_i(d^*+\delta) + \ell_i(x+1 - [d^i+\delta]) > r_i(d^i+1) + \ell_1(x+1 - [d^i+1])$$

Thus,

$$r_i(d^i+\delta) + \ell_i(x+1 - [d^i+\delta]) > r_i(d^i+1) + \ell_i(x-d^i)\ .$$

From the monotomicity and concavity of r_i it follows then that

$$r_i(d^i+ \delta-1) + \ell_i(x - [d^i+ \delta-1]) > r_i(d^i) + \ell_i(x-d^i)$$

This, however, contradicts the optimality of d^i for (x,i). Thus,

$G_i^i(x+1) \leq d^i+1$ and hence $G_i^i(x+1) \varepsilon \{G_i^i(x), G_i^i(x) + 1\}$ so that (t.2) holds for n=i. In order to show that (t.3) holds, the relation between $R_i^*(x-1)$, $R_i^*(x)$ and $R_i^*(x-1)$ will be investigated for any $x\varepsilon\{1, 2, \ldots, M_i-1\}$. Let G^i be any element of GG^i for which $G_i{}^i(x) \leq G_i{}^i(x+1) \leq G_i{}^i(x) + 1$, $\forall x \varepsilon \Omega_i$. (It has already been shown that such a strategy exists.) Let $G_i{}^i(x-1) = d^-$, $G^i(x) = d^i$ and $G_i{}^i(x+1) = d^+$ for some arbitrary $x\varepsilon\{1, 2, \ldots M_i-1\}$. The possible combinations of d^-, d^i, and d^+ are the following:

(1) $d^- = d^i = d^+$

(2) $d^- = d^i-1,\ d^+ = d^i$

(3) $d^- = d^i,\ d^+ = d^i$

(4) $d^- = d^i-1,\ d^+ = d^i + 1$

<u>Case 1</u>. $d^- = d^i,\ d^+ = d^i$

For this case,

$$R_i^*(x+1) = r_i(d^i) + \ell_i(x+1-d^i)$$

$$R_i^*(x) = r_i(d^i) + \ell_i(x-d^i)$$

$$R_i^*(x-1) = r_i(d^i) + \ell_i(x-1-d^i)$$

From (t.1) it follows then that R_i^* is monotone increasing and

$$2R_i^*(x) \geq R_i^*(x+1) + R_i^*(x-1)$$

<u>Case 2</u>. $d^- = d^i-1,\ d^+ = d^i$

For this case,

$$R_i^*(x+1) = r_i(d^i) + \ell_i(x+1-d^i)$$

$$R_i^*(x) = r_i(d^i) + \ell_i(x-d^i)$$

$$R_i^*(x-1) = r_i(d^i-1) + \ell_i(x-d^i)$$

and

$\Delta(x\_1, x) = R_i^*(x+1) - R_i^*(x) = \ell_i(x+1-d^i) - \ell_i(x-d^i)$

$\Delta(x, x-1) = R_i^*(x) - R_i^*(x-1) = r_i(d^i) - r_i(d^i-1)$

so that

$$\Delta\Delta = \Delta(x,x-1) - \Delta(x+1,x)$$

$$= r_i(d^i) + \ell_i(x-d^i) - [r_i(d^i-1) + \ell_i(x-(d^i-1)]$$

Since d^i is optimal at (x,i),

$$r_i(d^i) + \ell_i(x-d^i) \geq r_i(d^i-1) + \ell_i(x-(d^i-1)$$

and thus $\Delta\Delta \geq 0$ which implies that

$$2R_i^*(x) \geq R_i^*(x+1) + R_i^*(x-1).$$

<u>Case 3</u>. $d^- = d^i$, $d^+ = d^i+1$.

For this case:

$$R_i^*(x+1) = r_i(d^i+1) + \ell_i(x-d^i)$$

$$R_i^*(x) = r_i(d^i) + \ell_i(x-d^i)$$

$$R_i(x-1) = r_i(d^i) + \ell_i(x-1-d^i)$$

so that

$$\Delta(x+1,x) = r_i(d^i-1) - r_i(d^i)$$

$$\Delta(x,x-1) = \ell_i(x-d^i) - \ell_i(x-1-d^i)$$

and

$$\Delta\Delta = r_i(d^i) + \ell_i(x-d^i) - [r_i(d^i+1) + \ell_i(x-1-d^i)]$$

Since $d^- = d^i$ it follows that $d^i + 1 \leq x$ so that $d^i + 1 \varepsilon D_n'(x)$ and thus $\Delta\Delta \geq 0$ which implies that:

$$2R_i^*(x) \geq R_i^*(x+1) + R_i^*(x-1).$$

<u>Case 4</u>. $d^- = d^i-1$, $d^+ = d^i+1$.

For this case:

$$R_i^*(x+1) = r_i(d^i+1) + \ell_i(x-d^i)$$

$$R_i^*(x) = r_i(d^i) + \ell_i(x-d^i)$$

$$R_i^*(x-1) = r_i(d^i-1) + \ell_i(x-d^i)$$

Since r_i is monotone increasing concave function it follows then that $2R_i^*(x) \geq R_i^*(x+1) + R_i^*(x-1)$ and R_i^* is monotone increasing.

It is still necessary to show that R_i^* is monotone increasing for Cases 2 and 3. By definition:

$$R_i^*(x+1) \geq r_i(d^i) + \ell_i(x+1-d^i) \quad \text{and}$$
$$R_i^*(x) = r_i(d^i) + \ell_i(x-d^i)$$

Since ℓ_i is monotone increasing so is R_i^*. Thus, (t.3) is true and hence t is true for all $n \leq K$.

Since GG° contains only optimal feasible solutions and every $G^1 \varepsilon GG^1$ is optimal feasible at all $x_1 \varepsilon \Omega_1$ (L is type Moshe model) it follows then that $GG° = GG^1$ and hence there exists $G° \varepsilon GG^*$ with the above properties.

Remarks. (1) It should be noted that not all the elements of GG° have the above property. However, if r_i is strictly monotone increasing that since L' is a type Moshe function it can be shown that all the elements of GG° are with the above property.

(2) From the DP algorithm viewpoint, the above results indicate that while solving the DP equation for (x,n) the search may be restricted to two possible values for $G_n^n(x)$, i.e.: $G_n^n(x) \varepsilon \{G_n^n(x-1),\ G_n^n(x-1) + 1\}$. Since for $x_n = 0$, $D_n'(0) = \{0\}$, $G_n^n(0)$ can be set to zero, n = 1, 2, . . ., N and then the DP equation for $x_n = 1, 2, \ldots, M_n$ can be solved in a successive manner using the fact that $G_n^n(x+1) \varepsilon \{G_n^n(x),\ G_n^n(x) + 1\}$.

Example 4.2

Consider the model introduced in Example 4.1 with the following modifications:

(1) $MD_N = M_N$

and

(2) r_n' is monotone increasing convex function, $\forall n = 1, 2, \ldots$ N and $r_n' \equiv 0 \quad n > N$.

It will be shown that there exists an optimal feasible strategy G* such that

$$G_n^*(x) \varepsilon \{0, \min[x, MD_n]\}, \; n = 1, 2, \ldots N.$$

Proof. Using the notation introduced in Example 4.1, it will be shown that the following inductive hypothesis is true for $n \leq N$.

(t)
(t.1) ℓ_n is monotone increasing and $2\ell_n(x) \leq \ell_n(x+1) + \ell_n(x-1)$
(t.2) $G_n^n(x) \; \varepsilon\{0, \min(x, MD_n)\}$, $\forall x \varepsilon \Omega_n$, for some $G^n \varepsilon GG^n$,
(t.3) R_n^* is monotone increasing and $2R_n^*(x) \leq R_n^*(x+1) + R_n^*(x-1)$.

Let start the DP algorithm at n = N = K. Obviously,

$$R_K^*(x_n) = r_K(x_K),$$

and there exists $G^K \; GG^K$ such that $G_K^K(x_K) = x_K$. It is also followed that $\ell_K(y) \equiv 0$, $\forall y \varepsilon \Omega_{K+1}$. Thus, (t) is true for n = N = K. Assume that (t) is true for n = K-1, K-2, . . ., m. In particular, assume that it is true for n = m, i.e.:

(1) ℓ_m is monotone increasing and $2\ell_n(x) \leq \ell_n(x+1) + \ell_n(x-1)$

(2) $G_m^m(x) \varepsilon \{0, \min(x, MD_m)\}$ for some $G^m \varepsilon GG^m$ $\forall x \varepsilon \Omega_m$.

(3) R_m^* is monotone increasing and $2R_m^*(x) \leq R_m^*(x+1) + R_m^*(x-1)$.

Consider n = m-1 = i.

To show that (t_1) is true the same procedure as was used in Example 4.1 may be used but this time from the property of R^*_m it follows that ℓ_i is increasing and $2\ell_i(x) \leq \ell_i(x+1) + \ell_i(x-1)$. Let G^i be any arbitrary element of GG^i, and $G^i_i(x) = d^i$ for an arbitrary element of the set $\{1, 2, \ldots, M_i-1\}$. Thus,

$$r_i(d^i) + \ell_i(x-d^i) \geq r_i(d) + \ell_i(x-d), \ \forall d \varepsilon\ D'_i(x).$$

Notice that since L' is a type Moshe function it follows that:

$$R^*_i(x) = r_i(d^i) + \ell_i(x-d^i).$$

In particular,

$$r_i(d^i) + \ell_i(x-d^i) \geq r_i(d) + \ell_i(x-d), \ \forall d \leq d^i.$$

Since r_i is monotone increasing convex function, it follows that:

$$r_i(d^i+1) + \ell_i(x-d^i) \geq r_i(d+1) + \ell_i(x-d)$$

From (t.1) at i it is implied that

$$r_i(d^i+1) + \ell_i(x - (d^i+1)) \geq r_i(d+1) + \ell_i(x-(d+1)), \ \forall d \leq d_i.$$

In particular for $d = d^i - 1$:

$$r_i(d^i+1) + \ell_i(x-(d^i+1)) \geq r_i(d^i) + \ell_i(x-d^i)$$

Thus if d^i is optimal and $d^i < \min\{x, MD_i\}$, $d^i + 1$ is also optimal. Moreover, if $(d^i+\delta)\varepsilon D'_i(x)$, then,

$$r_i(d^i+\delta) + \ell_i(x-(d^i+\delta)) \geq r_i(d+\delta) + \ell_i(x-(d+\delta)), \ \forall d \leq d^i$$

so that if d^i is optimal so is $\min\{x, MD_i\}$. Thus there exists a strategy $G^i \varepsilon GG^i$ such that

$$G^i_i(x) \ \varepsilon\ \{0, \min(x, MD_i)\}.$$

and (t.2) is true for n = i.

To show that (t.3) is true for n = i, the relation between $R^*_i(x-1)$, $R^*_i(x)$ and $R^*_i(x+1)$ is to be investigated for some arbitrary element $x\varepsilon(1, 2, \ldots, M_i-1)$. It is known that there exists a strategy

$G^i \varepsilon GG^i$ such that $d^i = G_i^i(x) \varepsilon \{0, \min(x, MD_i)\}$ for some $x \varepsilon \{1, 2, \ldots, M_i - 1\}$. Let $C(x,d) = r_i(d) + \ell_i(x-d)$ and assume that $d^i = 0$. Thus

(1) $R_i^*(x) = C(x, 0)$

(2) $R_i^*(x+1) \geq C(x+1,0)$

(3) $R_i^* (x-1) \geq C(x-1,0)$

It follows then that:

$$\Delta(x+1,x) = R_i^*(x+1) - R_i^*(x) \geq c(x+1,0) - c(x,0)$$

$$\Delta(x,x-1) = R_i^*(x) - R_i^*(x-1) \leq c(x,0) - c(x-1,0).$$

and thus,

$$\Delta\Delta = \Delta(x+1,x) - \Delta(x,x-1) \geq c(x+1,0) + c(x-1,0) - 2c(x,0)$$

or $\quad \Delta\Delta \geq \ell_i(x+1) + \ell_i(x-1) - 2\ell_i(x)$

and from (t.1) it follows then that $\Delta\Delta \geq 0$ which implies that $2R_i^*(x) \leq R_i^*(x+1) + R_i^*(x-1)$. Suppose that $d^i = \min\{x, MD_i\} = MIN$. If $x > MD_n$ then $MIN = MD_i$ and thus $d^i = MD_i$ so that

$$R_i^*(x) = c(x,MD_i)$$

$$R_i^*(x+1) \geq c(x+1,MD_i)$$

$$R_i^*(x-1) \geq c(x-1, MD_i)$$

and,

$\Delta\Delta \geq \ell_i(x+1 - MD_i) + \ell_i(x-1 - MD_i) - 2\ell_i(x-MD_i)$ and from (t.1) it follows then that $\Delta\Delta \geq 0$ so that again $2R_i^*(x) \leq R_i^*(x+1) + R_i^*(x-1)$. If $x < MD_i$ then $d^i = x$ and $R_i^*(x) = c(x,x)$. It is also known that for this case,

$$R_i^*(x+1) \geq c(x+1, x+1) = r_i(x+1) + \ell_i(0)$$

$$R_i^*(x-1) \geq c(x-1, x-1) = r_i(x-1) + \ell_i(0)$$

so that for this case,

$$\Delta\Delta \geq r_i(x+1) + r_i(x-1) - 2r_i(x)$$

and since r_i is convex, $\Delta\Delta \geq 0$ which implies that $2R_i^*(x) \leq R_i^*(x+1) + R_i^*(x-1$. Suppose that $x = MD_i$ and thus $d^i = x = MD_i$, for which:

$$R_i^*(x) = c(x,x)$$

$$R_i^*(x+1) \geq c(x+1, x)$$

$$R_i^*(x-1) \geq c(x-1, x-1)$$

so that

$$\Delta\Delta \geq r_i(x+1) + r_i(x-1) - 2r_i(x) + \ell_i(1) - \ell_i(0).$$

From the convexity of r_i and (t.1), it follows then that $\Delta\Delta \geq 0$ which implies that $2R_i^*(x) \leq R_i^*(x+1) + R_i^*(x-1)$. To show that R_i^* is monotone increasing let x be any arbitrary element of $\{1, 2, \ldots, M_i-1\}$, and $d^i = G_i^i(x)$ for some strategy $G^i \varepsilon GG^i$ with the properties mentioned above. Thus,

$$R_i^*(x+1) \geq c(x+1, d^i) = r_i(d^i) + \ell_i(x+1 - d^i) \geq r_i(d^i) + \ell_i(x-d^i)$$

and hence $R_i^*(x+1) \geq R_i^*(x)$. Thus (t.3) is true then for n=1 and hence the inductive hypothesis is true for all $n \leq N = K$. Since L' is a type Moshe function it follows that $GG° = GG^1$ and that $GG° \subset GG^*$. Thus at least one strategy $G^* \varepsilon GG^*$ is such that

$$G_n^*(x) \varepsilon \{0, \min(x, MD_n)\}, \; n = 1, 2, 3, \ldots, N, \; x \varepsilon \Omega_n.$$

It will be shown that in addition to the above property of G^*, it also has the following characteristics:

(1) If for some $x > MD_n$ $G_n^*(x) = 0$ then $G_n^*(x') = 0 \quad \forall x' > x$.

and (2) If for some $x > MD_n$ $G_n^*(x) = MD_n$ then $G_n^*(x') = MD_n$ $\forall \; MD_n \leq x' \leq x$

Proof: (1) Since $G_n^*(x) = 0$ is optimal for (x,n) it follows that,

$$r_n(0) + \ell_n(x) \geq r_n(MD_n) + \ell_n(x-MD_n)$$

Using the property specified by (t.1),

$$r_n(0) + \ell_n(x+\delta) \geq r_n(MD_n) + \ell_n(x+\delta - MD_n) \quad , \quad \delta \geq 0$$

and hence $G_n^*(x^1 = x + \delta) = 0$

(2) Since $G_n^*(x) = MD_n$ is optimal at (x, n) it follows that

$$r_n(MD_n) + \ell_n(x-MD_n) \geq r_n(0) + \ell_n(x)$$

Using the property specified by (t.1) it follows that:

$$r_n(MD_n) + \ell_n(x-\delta-MD) \geq r_n(0) + \ell_n(x-\delta), \ \forall\delta \leq x-MD_n$$

Thus, $G_n^*(x' = x-\delta) = MD_n$ is optimal for all $MD_n \leq x' \leq x$.

The above discussion concerning the properties of G* indicates that additional reduction in the computation may be achieved by using the following procedure:

(1) For $x_n = MD_n$ check the relation between $c_1 = c(x_n,0)$ and $c_2 = c(x_n,x_n)$. If $c_2 \geq c_1$ set $G_n^*(x) = 0$, $\forall x > MD_n$. If $c_2 > c_1$ set $G_n^*(x)$ to either 0 or MD_n by solving the DP equations.

(2) For $x_n = M_n$ check the relation between $c_1 = c(x_n,0)$ and $c_2 = c(x_n, MD_n)$. If $c_2 \geq c_1$ set $G_n^*(x) = MD_n$, $\forall x \leq M_n$.

(3) Go to (1) and repeat the procedure for x' = x+1.

(4) Go to (2) and repeat the procedure for x' = x-1.

In other words, computational savings may be achieved by solving the DP equations in an <u>alternating</u> manner (as far as x_n is concerned) in the range (MD_n, M_n).

Example 4.3

Let $r_n'(d_n) = c_n \cdot d_n$ where c_n is a positive constant, i.e., r_n is a linear reward function. Since r_n satisfies the assumptions made in both Example 4.1 and Example 4.2, it follows that:

$$G_n^*(x) = \begin{cases} 0 & x \leq x_n^* \\ \min\,\{x, MD_n\} & x > x_n^* \end{cases}$$

for some $G^* \varepsilon GG^*$, where x_n^* is some critical value of the storage level. In this case, the solution for G^* involves the construction of the set $\{x_n^*:\ n = 1, 2, \ldots, N\}$, which can be done by using the DP algorithm.

As was indicated above, the examples considered in this chapter were introduced to demonstrate the role of analytical considerations in the implementation of the DP algorithm. An interesting question related to the above examples is the following one: how will the structure of G^* considered above be affected by permitting d_n to be greater than x_n and imposing some penalty for cases where $x_n + q_n < d_n$?

4.2. Discussion

The investigation presented in this chapter should be considered as an example rather than a method. The only objective considered when formulating the above decision model and demonstrating some solution procedures was to demonstrate that analytical considerations may be a basis for computational procedure for overcoming the dimensionality curse. More specifically, results obtained by Bellman (1957, pp. 19-25) and Nemhauser (1966, pp. 53-55) for deterministic process with continuous reward functions have been extended. Convex and concave reward functions are often used in the design and operation of water resources systems (Dorfman, 1962) so that the results obtained in this chapter may be applicable to practical problems in reservoir control.

It seems as if a combination of numerical analysis procedures (Larson, 1968; Heidari, 1970) and analytical ones like the one presented in this chapter may be used to overcome difficulties concerning the dimensionality of the DP problems.

It should be emphasized that the examples presented in this chapter on the context of reservoir control are of a general form and may also be used in the context of allocation and inventory problems.

CHAPTER V

THE MODELING OF A MULTISTAGE DECISION PROCESS

One of the advantages of the DP algorithm as a solution procedure is that it can handle a rather wide class of multistage decision problems. However, before starting the first step of the algorithm, the problem under consideration should be formulated as a multistage decision problem. Moreover, in order to guarantee that the DP algorithm indeed provides optimal feasible solutions, the formulated problem should have certain properties as far as the structure of its element is concerned. As an example, it was shown that certain type Shoshana models can be handled by the DP algorithm (ignoring for a moment the computational aspects). Thus, if the problem under consideration can be formulated by a type shoshana model, it is guaranteed that all the solutions obtained by the DP algorithm are optimal feasible. There are indications, however, that in practice the modeling of a multistage decision process is not a trivial matter.

In this chapter, a modeling framework to be used while formulating the problem under consideration as a multistage decision problem is introduced. In order to emphasize the importance of the modeling stage, consider the following illustrative example.

5.1. Example

Consider the following problem:

$$\max c = \prod_{i=1}^{N} y_i, \text{ subject to: } y_i \varepsilon Y_i, \ i = 1, 2, \ldots, N.$$

where:

$Y_1, Y_2, \ldots Y_N$ are subsets of the set of integers.

The following may be considered as a potential model for handling the problem.

Attempt #1

Let $(\Omega, D, F, P_o, L)_N$ be a CMDM where:

$\Omega = \{\Omega_n: \Omega_n = (1), n\varepsilon\mathbb{N}\}$

$D = \{D_n: D_n(h_n) = Y_n\}_{n\varepsilon\mathbb{N}} \qquad Y_i = \{1\} \quad i > N$

$F = \{f_n: f_n(h_n, d, 1) = 1, h_n\varepsilon H_n, d_n\varepsilon D_n(h_n)\}_{n\varepsilon\mathbb{N}}$

$P_o(1) = 1$

$L = \{L_n: L_n(h_\infty) = \prod_{i\geq n} d_i\}_{n\varepsilon\mathbb{N}}$

If all the elements of $\bigcup_{i=1}^{N} Y_i$ are non-negative, it can be easily verified that the model is a type Shoshana model and that if the DP algorithm starts at K = N optimal feasible solutions are obtained. However, if the above condition is not satisfied, the model is not a type Shoshana model and thus, there is no guarantee that the DP solutions are optimal feasible.

In order to formulate the problem as a problem associated with a type Shoshana model for cases where $\bigcup_{i=1}^{N} Y_i$ contains negative elements consider the following.

Attempt #2

Consider the complete model $(\Omega, D, F, P_o, L)_N$ where Ω, D, F, and P_o are as defined above and

$$L = \{L_n: L_n(h_\infty) = \prod_{i=1}^{\infty} d_i = \prod_{i=1}^{N} d_i\}_{n\varepsilon\mathbb{N}}$$

Using the sufficient statistic,

$$T = \{t_n: \quad t_n(h_n) = \prod_{i=1}^{n-1} d_i\}_{n\varepsilon\mathbb{N}}$$

it can be shown that the reduced model associated with (Ω, D, F, P_o, L) and T is a type Shoshana model and that the DP solutions are optimal feasible. For this model,

$$U_n = \{u_n: \quad u_n = \prod_{i=1}^{n-1} d_i\} \quad , \qquad n = 1, 2, \ldots, N$$

and thus depending on Y_i, i=1, 2, . . ., N,U_n may include a relatively large number of elements. Thus, although the above (reduced) model may be used as a framework for solving the problem, the DP algorithm may require solutions to a large number of DP equations.

In order to reduce the dimensionality of the problem, consider the following.

Attempt #3

Consider the complete model $(\Omega, D, F, P_o, L)_N$ where Ω, D, F, and P_o are as defined above, and

$$L = \{L_n: \quad L_n(h_\infty) = \text{SIGN}\left(\prod_{i=1}^{n-1} d_i\right) \cdot \prod_{i \geq n} d_i\}_{n\varepsilon\mathbb{N}}$$

where

$$\text{SIGN}(t) = \begin{cases} -1 & t < 0 \\ 0 & t = 0 \\ 1 & t > 0 \end{cases} \quad , \qquad t\varepsilon\mathbb{R}.$$

Consider the following sufficient statistic:

$$T = \{t_n: \quad t_n(h_n) = \text{SIGN}\left(\prod_{i=1}^{n-1} d_i\right)\}_{n\varepsilon\mathbb{N}} \; .$$

It can be easily verified that the reduced model associated with the complete model and T is a type shoshana model and that the DP algorithm provides optimal feasible solutions. Notice that in this case $U_n = \{-1, 0, 1\}$, $\forall n \in N$ and in most cases (unless the problem is extremely simple) contains less elements than the one introduced in the previous attempt.

Thus, the reduced model defined in the third attempt seems to be more efficient than the previous ones. Notice, however, that the third model has the disatrange that at the modified problem (u_n, n) the original reward function trying to maximize is not treated explicitly. Thus the choice between the models may be determined by the information desired when solving the modified problems -- taking into consideration the computational implicatins of such a choice.

From the modeling viewpoint it is important to realize that often more than one model is available to mathematically desirable the process under consideration. When making the decision concerning the model to be used, it is important to investigate the implications of such a decision as far as computation and other aspects of the situation are concerned.

In addition, often the problem under consideration is not presented in an explicit mathematical form so that there is also a need (from the modeling viewpoint) to present the problem under consideration in an explicit mathematical form.

The modeling framework developed in this chapter is designed for what may be called the preparation stage in which the process under consideration is mathematically formulated.

5.2. Modeling Framework

The modeling of a multistage decision process is often far from being a routine procedure. It starts with the identification of the objects related to the process, followed by the investigation concerning the relation between them which often includes feedbacks to the first step, and then ends with the formulation of the model. Once the model is mathematically formulated, potential solution procedures are considered.

In practice, however, there is a tendency to reach the solution procedure, the DP equations as an example, as soon as possible so that often the first two steps of the modeling procedure are either over-simplified or totally ignored. This type of "short-cuts" in the modeling procedure often limits the use of the DP algorithm as a solution procedure as will be indicated later.

The elements of the multistage decision model will be investigated now from a modeling viewpoint.

Decision Stages

The set of decision stages often consists of either time and/or space elements. When identifying the decision stages, it is essential to also identify the direction of the process as far as its evolution is concerned. For example, there is a need to identify loops, branches, (if any) and determine the direction of the process as far as the decision stages are concerned. If the process is non-serial, it should be converted into a set of serial processes linked together. If the process is truncated, the last decision stage should be carefully defined. Finally, the decision stages are ordered, usually by indexing.

State Spaces

Once the set of decision stages is defined with each of its elements, a state space is to be defined. The state spaces are not necessarily identical, although in most cases they consist of elements of the same type. When constructing the state space for a given decision stage, say n, the following considerations should be made:

(1) The n*th* state space should include all the elements needed to describe the situation of the system at the n*th* stage, as related to the dynamics of the process under consideration.

(2) If certain constraints are imposed on the system at time n, they should be specified by the elements of the state space.

(3) The state space should include all the elements needed in order to determine the set of decisions available at that stage.

(4) The state space should include all the elements associated with the n*th* decision stage that may affect the reward associated with this stage.

Although as a routine it is preferable to include more elements than needed rather than to exclude some, it is recommended to verify that no redundant elements are included in the state spaces.

Decision Sets

The decision set associated with the n*th* decision stage given a certain realization of the process up to this stage should include all the decision elements that are feasible under these conditions. The feasibility of a decision is chedked according to two different criteria. First, the availability of the decision is checked, i.e., it is to be determined whether the decision is indeed available to the decision maker at that point of the process. Then it should be checked whether the decision satisfies the constraints imposed on the system. From the modeling viewpoint, it is recommended to construct the set $D_n(h_n)$ by intersecting two sets: the set of decisions available to the decision maker at (h_n,n) and the set of decisions satisfying the constraints imposed on the processes. It should be noted that when checking whether a certain decision satisfies the constraints often, the law of motion governing the process is to be examined.

The Law of Motion

When constructing the law of motion of the process, it is recommended to determine first whether the law of motion under consideration is deterministic or else statistical. More specifically, if the elements of the state spaces are multidimensional variables, it is recommended to identify those coordinates of the state element that are governed by a statistical law of motion and those governed by a deterministic law. Once the law of motion is defined, it is recommended to reexamine the state spaces and the decision sets in order to verify (1) that they are complete and satisfy the constraints, and (2) that they do not include redundant elements.

Initial Condition

Although the initial condition is introduced in the discussion as a function describing (statistically) the initial conditions of the process, it can also be used for the purpose of sensitivity analysis. Thus, even for deterministic processes a "statistical form" of the initial condition may be considered when the effects of the initial condition of the process are to be investigated.

Reward Function

Two basic characteristics of the reward function should be first specified; the domain of definitions of L_1 and its range. More specifically, in many situations L_1 is defined on a subset of H_∞ and its range is a subset of $\mathbb{R}$. Once the structure of L_1 is determined, the possibility of decomposing it into a sequence $\{L_i\}_{i>1}$ of real valued functions so that $L = \{L_n\}_{n\in\mathbb{N}}$ will have certain desired properties, for example, additivity, separability under expectation, etc. should be investigated. Notice that often L_1 is uniquely determined by the process under consideration while the decomposition of L_1 is not necessarily unique. It is important then to examine all the potential decompotions of L_1. Once L is constructed, it is recommended to reexamine the structure of the state spaces and the decision sets in order to make sure that they are complete as far as the domain of definition of L_1 is concerned.

Sufficient Statistic

The construction of (non-trivial) sufficient statistics, if any, is motivated primarily by computational considerations. The non-uniqueness of the sufficient statistic suggests the notion of "minimal sufficient statistic." Generally speaking, the efficiency of a sufficient statistic may be measured, so to speak, by the number of dynamic programming equations one has to solve when implementing the DP algorithm (if the algorithm can be used for the particular problem), as compared with the number of equations needed for the complete model. Thus when making the decision as to the sufficient statistic to be used, the number of elements in U_n, $n \in \mathbb{N}$, may be used as a decision criterion.

The discussion presented above should not by any means be considered as a set of instructions to be followed whenever the modeling of a multistage decision process is considered. Rather, it should serve as a guide when constructing the elements of the model. The points made in the discussion will be illustrated by the modeling of two reservoir control problems.

5.3. Reservoir Control Models

The models to be introduced in the following sections should be considered as illustrative ones. No elaboration on the physical justification for choosing certain reward functions will be made, and no justification for the use of the expected value approach as an optimality criterion will be given. The only objective is the demonstration of the modeling flexibilities provided by the model developed in Chapter 2. The first example demonstrates the flexibility of the model and the DP algorithm as far as the handling of probabilistic constraints is concerned, and is based on a comment (Sniedovich and Davis, 1976) related to a paper by Askew (1974). The second example demonstrates the flexibility of the model and the DP algorithm as a modeling and solution procedure, as far as the structure of the reward function is concerned; the problem associated with the minimization of the expected value of the range of fluctuation of the storage level in a reservoir will be considered.

5.4. A Reliability Problem in Reservoir Control

The operation of a reserovir of capacity MC is to be determined for the next N years. The maximum target release associated with the n*th* year is given by MR_n, $n \leq N$. The inflow to the reservoir is described by the sequence $\{\tilde{q}_n\}_{n=1}^{N}$ of independent random variables whose probability mass functions $\{p_n\}_{n=1}^{N}$ are known. Let $Q_n = \{q_n: \ q_n = 0, 1, 2, \ldots MQ_n\}$ be the set of values $\tilde{q}_n$ takes with positive probability. Suppose that at the n*th* year, $n < N$, the storage level x_n is observed and the target release d_n is selected. The decision maker may face the following situations:

(1) $d_n > x_n$. For this case the following process takes place: First, the quantity x_n is released, followed by some input q_n determined by p_n. Then the quantity $\min\,[d_n - x_n, q_n]$ is released. If $q_n < d_n - x_n$ the <u>shortage</u> $\Delta_s = d_n - x_n - q_n$ is experienced and the storage level at the beginning of the next year is zero. If $x_n + q_n - d_n > MC$, the <u>overflow</u> $\Delta_o = x_n + q_n - d_n - MC$ is experienced and the storage level at the beginning of the next year is MC. If $0 < x_n + q_n - d_n < MC$ there is neither a shortage nor an overflow and the storage level at the beginning of the next year is $x_n + q_n - d_n$.

(2) $x_n \geq d_n$. For this case the following process takes place: First, the quantity d_n is released, followed by some input q_n determined by p_n. If $x_n + q_n - d_n > MC$ the overflow $\Delta_o = x_n + q_n - d_n - MC$ is experienced and the storage level at the beginning of the next year is MC. If $x_n + q_n - d_n < MC$ no overflow is experienced and the new storage level is $x_n + q_n - d_n$.

In order to eliminate non-feasible situations, it is assumed that given the storage level x_n the only target releases d_n to be considered are those satisfying the condition: $d_n \leq x_n + MQ_n$. The reward associated with each year is a function of (a) the target release, and (b) the shortage/overflow experienced during that year.

The objective is to construct a release strategy for the N year period which will maximize the expected value of the sum of the yearly rewards, subject to the "safety factor" σ defined as the minimal probability of no shortage allowed during the N years period. In other words, the probability of at least one shortage during the period [1,N] should be less than $1 - \sigma$.

The above situation will be formulated as a multistage decision process using the model developed in Chapter 2 and it will be shown that the DP algorithm may be used for the construction of the set of optimal feasible strategies.

The elements of the multistage decision model representing the above process are to be constructed.

Decision Stages

Since the period of interest consists of finitely many years, it is obvious that the model is truncated, which will be indicated by indexing the set of decision stages, by N, that is $\mathcal{N}_N = \{n: \; n = 1, 2, \ldots\}$.

State Spaces

As far as the motion of the process is concerned, that is, the changes in the storage levels in the reservoir, the n*th* state space should include elements describing the storage level in the reservoir.

However, since the reward associated with the *nth* year depends on the magnitude of the shortage/overflow if any, it should also include elements describing these events. As a first attempt, consider the following state spaces:

$$\Omega_1^1 = \{x_1: \ x_1 = 0, 1, 2, \ldots, MC\}$$

$$\Omega_n^1 = \{x_n: \ x_n \varepsilon(-MR_{n-1}, -MR_{n-1}+1, \ldots, 0, 1, \ldots, MC, MC+1, \ldots MC+MQ_{n-1})\},$$
$$n=2,\ldots,N+1$$

Thus, $x_n < 0$ indicates that a shortage of x_n occurred and $x_n > MC$ indicates that an overflow of $x_n - MC$ occurred. Notice that the discritization of the stage spaces is often done subjectively. For n > N the stage spaces may be constructed arbitrarily, for example, $\Omega_i^1 = \{0\}, \forall i > N+1$.

Decision Sets

It will be assumed (for simplicity) that the range $[0, MR_n]$ is discretized such that the elements d_n are expressed in the same units as used in the description of the state spaces. Furthermore, let MR be the maximum yearly release capacity over the period: [1,N], i.e.:

$MR = \max\limits_{n=1,2,\ldots N} \{MR_n\}$. Thus,

$$\mathcal{D} = \{d: \ d= 0, 1, 2, \ldots, MR\}.$$

Notice, however, that not all the elements of $\mathcal{D}$ are available at a given year. Let $D_n(h_n)$ be the set of admissible decisions associated with (h_n, n). It is known that every element d_n of $D_n(h_n)$ is such that:

$$d_n \ \varepsilon\{0, 1, 2, \ldots, MR_n\}.$$

However, in order to satisfy the safety factor σ all the elements of $D_n(h_n)$ should guarantee that σ is not violated. Given the history h_n and the decision $d_n \varepsilon \{0, 1, 2, \ldots, MR_n\}$ the probability of failure during the n*th* year is computed as follows:

$$P_r(\text{failure during the } n\text{th year} | h_n, d_n) = \sum_{q_n \varepsilon A_n} p_n(q_n),$$

$$A_n = \{q_n: \; y_n(x_n) + q_n - d_n\}$$

where:

$$y_n(x_n) = \begin{cases} 0 & x_n \leq 0 \\ x_n & 0 < x_n < MC \\ MC & x_n \geq MC \end{cases}$$

Since the safety factor σ should be satisfied during the entire period [1,N], the state spaces should include additional information so that when making the decision at (h_n, n) only feasible decisions will be considered. For this purpose, let e_n be the probability of no failure during the period [1, n-1]. This probability is uniquely determined by the strategy used during the period [1, n-1], the sequence $\{p_i\}_{i=1}^{n-1}$ and the initial storage level or else the distribution of the initial storage level. Since $\mathbb{D}$ consists of finitely many elements, there are finitely many strategies and hence at each decision stage say n, they are finitely many feasible values for e_n. Let E_n be the set of the values e_n may take. Consider the following stage spaces:

$$\Omega_1 = \Omega_1^1 \times E_1 \quad , \quad E_1 = \{1\}$$

$$\Omega_n = \Omega_n^1 \times E_n \quad , \quad n-2, \ldots, N$$

$$\Omega_n = \Omega_n^1 \times E_N \quad , \quad n > N.$$

As will be shown later, all the elements of E_n are in the range $[\sigma, 1]$, $n \varepsilon \mathbb{N}$.

Suppose now that the history $h_n = [(x_1, e_1), d_1, (x_2, e_2), d_2, \ldots, (x_n, e_n)]$ is observed, the decision d_n is made and the value of e_{n+1} is to be computed. By definition:

$$e_{n+1}\Big|_{h_n,d_n,q_n} = \begin{cases} 0 & q < d_n - y_n(x_n) \\ e_n & q_n \geq d_n - y_n(x_n) \end{cases}$$

or

$$e_{n+1}\Big|_{h_n,d_n} = e_n \cdot P_r\,(q_n \geq d_n - y_n(x_n))$$

$$= e_n \cdot \sum_{q_n \geq d_n - y_n(x_n)} p_n(q_n)$$

In order to satisfy the safety factor σ all the elements d_n of $D_n(h_n)$ should be such that:

$$e_{n+1}\Big|_{h_n,d_n} = e_n \cdot \sum_{q_n \geq d_n - y_n(x_n)} p_n(q_n) \geq \sigma$$

It follows then that

$$D_n(h_n) = \{d_n : d_n \varepsilon \{0,1,\ldots MR_n\},\ e_n \cdot \sum_{q_n \geq d_n - y_n(x_n)} p_n(q_n) \geq \sigma,\ e_n \varepsilon E_n\},\ n=1,2,\ldots N$$

For $n > N$ set: $D_n(h_n) = \{0\}$.

The Law of Motion

Using the above definitions of $y_n(x_n)$ and e_n, it follows that:

$$f_n(h_n, d_n, (x_{n+1}, e_{n+1})) = \begin{cases} \cdot \displaystyle\sum_{q_n \geq MC - y_n(x_n) + d_n} p_n(q_n), & x_{n+1} = MC \\ \cdot\ p_n(q_n) & 0 < x_n + 1 < MC \qquad e_{n+1} = e_n \cdot \displaystyle\sum_{q_n \geq d_n - y_n(x_n)} p_n(q_n) \\ \cdot \displaystyle\sum_{q_n \leq d_n - y_n(x_n)} p_n(q_n) & x_{n+1} = 0 \\ \cdot\ 0 & \text{otherwise} \end{cases}$$

Notice that $\sigma \le e_{n+1} \le e_n$ so that $e_1 = 1$ implies that $\sigma \le e_n \le 1$. For $n > N$ any arbitrary mass function f_n may be used.

Initial Condition

If the initial storage level is known, say $x°$, let $P_o(x°,1) = 1$. If the initial storage level is described by a probability mass function, P_o', over the range $\{0, 1, 2, \ldots MC\}$ set

$$P_o(x_1,1) = P_o'(x_1), \quad x_1 \varepsilon \{0, 1, 2, \ldots MC\}.$$

Notice that by definition $E_1 = \{1\}$.

Reward Function

Let $r_n(d_n, x_{n+1})$ be the reward associated with the nth year given that the target release is d_n and the modified storage level of n+1 is x_{n+1}. Thus,

$$r_n(d_n,x'_{n+1}) = r_n(d_n,x''_{n+1}) \quad , \quad \forall \qquad 0 \le x''_n, x'_{n+1} \le MC$$

The objective function may be written then as follows:

$$L_1(h_\infty) = \sum_{i=1}^{N} r_i(d_i,x_{i+1}) \qquad , \qquad r_i \equiv 0 \quad , \quad i > N.$$

The reward function $L = \{L_n\}_{n\varepsilon\mathbb{N}}$ can be written then as:

$$L_n(h_\infty) = \sum_{i\ge n}^{N} r_i(d_i,x_{i+1}) \qquad , \qquad n \le N$$

$$L_n(h_\infty) \equiv 0 \quad , \quad n > N.$$

so that

$$L_n(h_\infty) = r_n(d_i, x_{i+1}) + L_{n+1}(h_\infty) \qquad , \qquad n\varepsilon\mathbb{N}$$

and thus

$$R_n(h_n, S|[x_{n+1},e_{n+1}]) = r_n(S_n(h_n), x_{n+1}) + R_{n+1}(h_{n+1},S)$$

so that $L = \{L_n\}_{n\varepsilon\mathbb{N}}$ is a type Moshe reward function.

Sufficient Statistic

The structure of $\Omega = \{\Omega_n : n \varepsilon \mathbb{N}\}$, $D = \{D_n\}_{n\varepsilon\mathbb{N}}$, $F = \{f_n\}_{n\varepsilon\mathbb{N}}$ and $L = \{L_n\}_{n\varepsilon\mathbb{N}}$ indicate that the process is Markovian and thus we may consider the sufficient statistic:

$$T = \{t_n : t_n(h_n) = (x_n, e_n), h_n \varepsilon H_n\}_{n\varepsilon\mathbb{N}}.$$

Solution Procedure

In order to construct the set of optimal feasible solutions, the DP algorithm starting at $n = K = N$ may be used. Notice that the close form relation between e_n and e_{n+1} make it possible to view e_n as a continuous variable, so to speak, when solving the dynamic programming equations. In other words, from the computational viewpoint, the elements of E_n are not required to be specified although this can be done. As indicated above, the objective in this section is not to construct solution algorithms but rather to <u>formulate</u> the problem under consideration using the model developed in Chapter 2.

Often the safety factor is used in the context of a sensitivity analysis. Its value may be changed so as to investigate its effect on the total reward. Then, the safety factor that together with the corresponding optimal total reward constitute the most desirable combination may be chosen as the optimal solution. For details, see Askew (1974).

5.5. The "Range" Problem in Reservoir Control

Consider the reservoir described in the previous example. It is desired to construct a release strategy such that the expected value of the <u>range of fluctuation</u> around the critical level x° is minimized over the N year period, given that the initial storage level in the reservoir is $x_1 = x°$.

A multistage decision model for the above process will be formulated and it will be shown that the DP algorithm may be used as a solution procedure. The elements of the model may be constructed as follows:

Decision Stages

The decision set is denoted by $\mathbb{N}_N$ to indicate that the process is truncated at N. Each $n \varepsilon \mathbb{N}_N$ corresponds to a certain year, or more precisely, each $n \leq N$ corresponds to a certain year in the period [1, N].

State Spaces

Since the magnitude of the shortage/overflow are of no interest and since no constraints relating to the state spaces are to be considered, define the state spaces as follows:

$$\Omega = \{\Omega_n : \ \Omega_n = (0, 1, 2, \ldots, MC)\} \ , \ n \varepsilon \mathbb{N}$$

Decision Sets

The only constraints related to the feasible release is expressed by: $D_n(h_n) = \{d: \ d \varepsilon (0, 1, 2, \ldots, MR_n)\}$, $n \leq N$, $h_n \varepsilon H_n$. For $n > N$ set: $D_n(h_n) = \{0\}$, $n > N$.

The Law of Motion

The law of motion of the process is determined by $\{P_n\}_{n=1}^N$.

More specifically,

$$f_n(h_n, d_n, x_{n+1}) = \begin{cases} \sum\limits_{q_n \geq MC - x_n + d_n} p_n(q_n) & x_{n+1} = MC \\ p_n(q_n = x_{n+1} - x_n + d_n) & 0 < x_{n+1} < MC \\ \sum\limits_{q_n \leq d_n - x_n} p_n(q_n) & x_{n+1} = 0 \end{cases}$$

for $n \leq N$. For $n > N$ the conditional mass functions on Ω_n are:

$p_n(0) = 1$, $n > N$.

Initial Condition

Since the process starts with the initial storage level x°, it follows that $P_o(x^\circ) = 1$.

Reward Function

The objective is to minimize the expected value of the function:

$$L_1(x_1, d_1, x_2, \ldots, x_N, \ldots) = \max_{N \geq i \geq 1} \{x_i\} - x^\circ + x^\circ - \min_{N \geq i \geq 1} \{x_i\}$$

$$= \max_{N \geq i \geq 1} \{x_i\} - \min_{N \geq i \geq 1} \{x_i\}$$

Notice that since $x_1 = x^\circ$, it is guaranteed that $\max_{N \geq i \geq 1} \{x_i\} \geq x^\circ$ and $\min_{N \geq i \geq 1} \{x_i\} \leq x^\circ$ so that L_1 as defined above indeed represents the actual objective function. The reward function $L = \{L_n\}_{n \in \mathbb{N}_n}$ may be defined then as:

$$L_n(x_1, d_1, \ldots) = L_n(x_1, d_1, \ldots), \ \forall\, n \in \mathbb{N}_N.$$

Sufficient Statistic

Since the value of the original objective function depends on the values x_i, $i \leq N$ take, the information contained by h_i may be condensed so that at the *ith* decision stage it will not be required to consider the entire history h_i. Consider then the sufficient statistic:

$$T = \{t_n\colon \ t_n(h_n) = (\max_{n \geq i \geq 1} \{x_i\}, \ \min_{n \geq i \geq 1} \{x_i\} \ , x_n)\} \ _{n \in \mathbb{N}_N}$$

In other words,

$$t_n\colon \ H_n \to A_x^3 = U_n \qquad , \quad A_x^3 = \Omega_i \times \Omega_i \times \Omega_i, \ i \in \mathbb{N}_N.$$

with $t_1(x^\circ) = (x^\circ, x^\circ, x^\circ)$.

Notice that $u_n = (u_n(1),\ u_n(2),\ u_n(3))$, contains the element $u_n(3) = x_n$ which is required in order to determine the law of motion at the n*th* stage. Thus,

$$u_n(1) = \max_{n \geq i \geq 1} (x_i), \quad u_n(2) = \min_{n \geq i \geq 1} (x_i)\ , \quad u_n(3) = x_n \quad .$$

The reduced reward function is then:

$$L' = \{L_n:\ L_n(u_n, d_n, x_{n+1}, \ldots) = \max \{u_n(1),\ \max_{n \geq i \geq n_i} (x_i)\} - \min \{u_n(2),\ \min_{n \geq i > n} (x_1)\}_{n \in N_n}$$

It can be easily verified that:

$$L'_n(u_n,\ d_n,\ \ldots) = L'_{n+1}(u_{n+1},\ d_{n+1},\ \ldots)$$

so that

$$R'_n(u_n,\ G,\ x_{n+1}) = R_{n+1}\ (u_{n+1},\ G)$$

where

$$u_{n+1} = (\max \{u_n(1),\ x_{n+1}\},\ \min \{u_n(2),\ x_{n+1}\},\ x_{n+1}).$$

It follows then that the model under consideration is a truncated Moshe type model and thus the DP algorithm (starting at n = N = K) may be used and will provide optimal feasible strategies.

Solution Procedure

Notice that the complete model may be used when complementing the DP algorithm. However, the reduced model is much more efficient as far as computation is concerned since it involves less modified problems.

Computation Example

Consider the following values for the elements introduced in the above problem:

MC = maximum storage capacity of the reservoir = 10 units,

MR_n = maximum release capacity at the n*th* year = 3 units, $\forall n \leq N$,

N = number of years of operation = 15

$x°$ = critical storage level = 7 units

p_n = the probability mass function of the inflow, $\tilde{q}_n$, n=1, 2, . . .N.

$$p_n(0) = 0.20,\ p_n(1) = 0.30,\ p_n(2) = 0.30,\ p_n(3) = 0.20,\ n=1,2,\ldots N.$$

The multistage decision problem associated with the above values was solved by the DP algorithm (starting at K = N) using the computer program presented in Appendix A. The optimal value of the reward function was found to be R* = 2.92. Portion of the optimal feasible strategy is presented in Table 1.

5.6. Discussion

When discussing the modeling aspects involved in the implementation of the DP algorithm, Bellman (1957, p. 82) indicates:

> We have purposely left the description a little vague, since it is the spirit of the approach to these processes that is significant rather than the letter of some rigid formulation. It is extremely important to realize that one can neither axiomatize mathematical formulation or legislate away ingenuity. In some problems, the state variables and the transformations are forced upon us; in others, there is a choice in these matters and the analytic solution stands or falls upon this choice. In still others, the state variables and sometimes the transformations must be artificially constructed. Experience alone, combined with often laborious trial and error will yield suitable formulations of involved processes . . . (Bellman, 1957, p. 82).

Instead of trying to develop a modeling procedure for a DP problem such as the one proposed by Aris (1964, p. 29) it was suggested that an improvement in the modeling phase of the problem may be best achieved by understanding the role of each of the elements of the multistage decision model. Thus, instead of developing a modeling procedure consisting of

Table 1. Portion of the Optimal Feasible Strategy $G_n^*(u)$ Associated with the Range Problem.

Sufficient Statistic			n														
u(1)	u(2)	u(3)	1	2	3	4	5	6.	7	8	9	10	11	12	13	14	15
7	4	5	–	–	1	1	1	1	1	1	1	1	1	1	1	1	1
7	4	6		–	2	2	2	2	2	2	2	2	2	2	2	2	2
7	4	7		–	3	3	3	3	3	3	3	3	3	3	3	3	3
7	5	5		0	0	1	0	0	1	0	1	0	0	0	0	0	0
7	5	6		–	1	2	1	1	2	1	2	1	1	1	1	1	1
7	5	7		–	2	3	2	2	3	2	3	2	2	2	2	2	2
7	6	6		1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	6	7	–	–	2	2	2	2	2	2	2	2	2	2	2	2	2
7	7	7	1	1	1	1	1	1	1	2	1	1	1	2	1	1	1
8	5	5	–	–	0	0	0	0	0	0	0	0	0	0	0	0	0
8	5	6		–	1	1	1	1	1	1	1	1	1	1	1	1	1
8	5	7		–	2	2	2	2	2	2	2	2	2	2	2	2	2
8	5	8		–	3	3	3	3	3	3	3	3	3	3	3	3	3
8	6	6		–	0	1	0	0	1	0	1	0	0	0	0	0	0
8	6	7		–	1	2	1	1	2	1	2	1	1	1	1	1	1
8	6	8		–	2	3	2	2	3	2	3	2	2	2	2	2	2
8	7	7		–	1	1	1	1	1	1	1	1	1	1	1	1	1
8	7	8		2	2	2	2	2	2	2	2	2	2	2	2	2	2
9	6	6		–	0	0	0	0	0	0	0	0	0	0	0	0	0
9	6	7		–	1	1	1	1	1	1	1	1	1	1	1	1	1
9	6	8		–	2	2	2	2	2	2	2	2	2	2	2	2	2
9	6	9		–	3	3	3	3	3	3	3	3	3	3	3	3	3
9	7	7		–	0	1	0	0	1	0	1	0	0	0	0	0	0
9	7	8		–	1	2	1	1	2	1	2	1	1	1	1	1	1
9	7	9		2	2	3	2	2	3	2	3	2	2	2	2	2	2
10	0	0		–	0	0	0	0	0	0	0	0	0	0	0	0	0
10	5	5		–	0	0	0	0	0	0	0	0	0	0	0	0	0
10	7	7		–	0	0	0	0	0	0	0	0	0	0	0	0	0
10	7	10	–	0	0	0	0	0	0	0	0	0	0	0	0	0	0

"steps" to be followed, it was chosen to investigate the structure of the elements of the model and the relations between them as far as modeling is concerned.

There is evidence (Askew, 1974) that while the DP is often used in water resources management, and in most cases correctly, certain basic modeling problems still prevent a full usage of DP.

The reliability problem presented in this chapter was treated by Askew (1974, p. 1100), "these constraints limit the magnitude of parameters that are functions of system variables computed over the entire life of the system; therefore, they cannot be introduced as normal constraints . . ." As indicated above, the validity of Askew's comment depends on the choice of the system variables. In other words, the ability to take care of certain constraints depends on the state spaces used in the model. It should be emphasized that the state spaces may include "artificial" variables that sometimes have "nothing" to do with the physical process under consideration. Bellman and Dreyfus (1962) for example, indicate the possibility of handling probabilistic constraints by the use of DP. However, the modeling aspects involved in such processes have not been treated in detail.

The range problem introduced in this chapter demonstrates the flexibility of the DP algorithm as a solution procedure. Most of the stochastic decision processes treated by the DP in the literature are characterized by additive reward functions. It was demonstrated that more complex reward functions may also be treated by the DP algorithm.

The two non-routine problems were introduced in the discussion mainly for the purpose of demonstrating some of the modeling aspects involved in transforming a multistage decision process to a model having the format of the one developed in Chapter 2.

APPENDIX A

LIST OF SYMBOLS

All the symbols used in the discussion are defined when they first appear in the text. The list presented below includes most of the symbols used in Chapter 2 and Chapter 3. The symbols not included in the list are those used in specific examples and are defined in the context of the discussion.

SYMBOL		PAGE OF FIRST APPEARANCE IN THE TEXT
$C_n(h_n)$	The solution of the optimality equation associated with (h_n,n)	43
d	A decision, an element of ϕ	2
d_n	A decision associated with the nth decision stage	2
D	The set of admissible decision maps associated with a complete model	6
D_n	The admissible decision map associated with the complete model and the nth decision stage	6
$D_n(h_n)$	The set of admissible decisions associated with the history h_n at the nth decision stage	6
D'	The sequence of reduced admissible decision maps associated with the reduced model	14
D'_n	The reduced admissible decision map associated with the nth stage	13
$D'_n(u_n)$	The set of admissible decisions associated with (U_n,n)	13
ϕ	A decision set	2

SYMBOL		PAGE OF FIRST APPEARANCE IN THE TEXT
$E[\cdot]$	The expected value of $[\cdot]$	10
f_n	The conditional mass function associated with the nth decision stage	6
f'_n	The conditional mass function associated with the reduced model at the nth decision stage	13
$f_n(h_n,d_n,\cdot)$	The conditional mass function on Ω_{n+1} given h_n,d_n.	6
$f'_n(u_n,d_n,\cdot)$	The conditional mass function defined on Ω_{n+1} given u_n,d_n.	13
F	The law of motion associated with the complete model	6
F'	A reduced law of motion	14
G	A strategy associated with the reduced model	14
G*	An optimal feasible strategy associated with the reduced model	15
G_n	The decision map associated with the strategy G at the nth decision stage	14
GG	The set of feasible strategies associated with the reduced model	14
GG*	The set of optimal feasible strategies associated with the reduced model	15
GG°	The set of strategies produced by the DP algorithm	28
GG^n	The set of strategies associated with the DP algorithm at the nth decision stage	27
h_n	A history associated with the nth decision stage, an element of H_n	6
$h_{n,s}(\bar{x}_n)$	The history determined by the strategy s at the nth decision stage given $\bar{x}_n$	10
H_n	The set of admissible histories associated with the nth decision stage	6

SYMBOL		PAGE OF FIRST APPEARANCE IN THE TEXT
$H_n(G)$	The set of histories in H_n observed with positive probability under G	6
$\bar{H}_n$	A history space associated with the n*th* decision stage	5
i	A positive integer, an element of $\mathbb{N}$	5
I^c	The map from GG to SS as determined by T	16
$I^c(G)$	The complete image of G	16
I^r	A map from SS to GG	18
$I^r(S)$	The reduced image of S	18
I^r_n	A map from SS to SS associated with the n*th* decision stage	18
j	A positive integer, an element of $\mathbb{N}$	12
k	A positive integer, an element of $\mathbb{N}$	27
K	The decision stage where the DP algorithm starts	27
ℓ'_G	The random variable defined on Ω as determined by G	16
ℓ_S	The random variable defined on Ω as determined by S	10
L	The reward function associated with the complete model	7
L'	The reduced reward function	14
L_n	The reward function associated with the n*th* decision stage	7
L'_n	The reward function associated with the reduced model at the n*th* decision stage	13
n	A decision stage, an element of $\mathbb{N}$	4
$\mathbb{N}$	The set of positive integers, also the set of decision stages	4

SYMBOL		PAGE OF FIRST APPEARANCE IN THE TEXT
N	A positive integer, an element of $\mathbb{N}$	21
P_G	The probability measure induced by G	15
P_o	The initial condition of the process, a mass function on Ω_1	7
P_s	The probability measure induced by S	10
$P_r(\cdot)$	The probability of the event $(\cdot)$	59
r_n	A real valued function associated with the *nth* decision stage	22
$\mathbb{R}$	The set of real numbers	7
R(S)	The total reward associated with the strategy S	10
R'(G)	The total reward associated with the strategy G	15
R*	The total optimal feasible total reward	11
R'*	The optimal feasible total reward associated with the reduced model	15
$R_n(h_n,S)$	The reward associated with the strategy S at (h_n,n)	11
$R'_n(u_n,G)$	The reward associated with the strategy G at the *nth* decision stage given u_n	15
$R^*_n(h_n)$	The optimal feasible reward associated with (h_n,n)	11
$R'^*_n(u_n)$	The optimal feasible reward associated with (u_n,n)	16
S	A strategy associated with the complete model	8
S_n	The map associated with the strategy S at the *nth* decision stage	8
$S_n(h_n)$	The decision determined by S at (h_n,n)	8

SYMBOL		PAGE OF FIRST APPEARANCE IN THE TEXT
S^*	An optimal feasible strategy	11
SS	The set of feasible strategies associated with the complete model	8
SS^*	The set of optimal feasible strategies associated with the complete model	11
$\{S^i\}_{i \geq n}$	The sequence of strategies generated by S at n	17
t_n	The n*th* element of T; the sufficient statistic associated with the n*th* decision stage	12
T	A sufficient statistic	12
u_n	An element of U_n	12
U_n	The range of t_n	12
$U_n(G)$	The set of statistic observed with positive probability at n under G	34
v_n	An element of U_{n+1} as determined by V_n	13
V	The transition function associated with the sufficient statistic	13
V_n	The transition function associated with the sufficient statistic at the n*th* decision stage	13
$V_n(u_n,d_n, x_{n+1})$	The value of t_{n+1} as defined by u_n, d_n, and x_{n+1}	13
x_n	A state associated with the n*th* decision stage, an element of Ω_n	5
$\bar{x}_n$	A trajectory, an element of $\bar{X}_n$	5
$\bar{x}_n(i)$	The i*th* coordinate of $\bar{x}_n$	5
$\bar{X}_n$	The set of all the trajectories associated with the n*th* decision stage	5
ξ	A sequence of random variables on Ω	9

SYMBOL		PAGE OF FIRST APPEARANCE IN THE TEXT
ξ_n	The present state function, the *nth* element of ξ	9
η	A sequence of random variables on Ω	9
η_n	The past state function, the *nth* element of η	9
Θ	The state observing function	34
Θ_n	The state observing function associated with the *nth* decision stage	34
$\Theta_n(G)$	The set of trajectories in $\bar{X}_n$ observed with positive probability under the strategy G	34
ζ	A sequence of random variables on Ω	9
ζ_n	The future state function, the *nth* element of ζ	9
ρ_n	A real valued function associated with the complete model at the *nth* decision stage	30
ρ'_n	A real valued function associated with the reduced model at the *nth* decision stage	30
Ψ	A σ-algebra on Ω	9
ω	An element of Ω	9
Ω	The sample space associated with the multi-stage decision model	9
Ω_n	The state space associated with the *nth* decision stage	5
$\bar{\Omega}$	The universe: the union of all the state spaces	5
$\mathfrak{R}$	The set of all the state spaces	5
$(\mathfrak{R}, D, F, P_o, L)$	A complete multistage decision model	7
$(\mathfrak{R}, D', F', P_o, L')$	A reduced multistage decision model	14

SYMBOL		PAGE OF FIRST APPEARANCE IN THE TEXT
⌋	End of proof, definition, comment, etc.	4

Abbreviations

CMDM	Complete Multistage Decision Model
DP	Dynamic Programming
OE	Optimality Equation
RMDM	Reduced Multistage Decision Model

APPENDIX B

COMPUTER PROGRAM

On the following pages, the program DYNO is listed. The program is designed for the range problem presented in Chapter 5.

```
      PROGRAM DYNO(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)
***********************************************************
*                                                         *
*                                                         *
*  THE PROGRAM DYNO,BY MEANS OF THE DYNAMIC PROGRAMING    *
*  ALGORITHM,CONSTRUCTS THE OPTIMAL FEASIBLE RELEASE AND  *
*  COMPUTES THE OPTIMAL FEASIBLE REWARDS ASSOCIATED WITH  *
*  IT FOR THE RESERVOIR CONTROL PROBLEM PRESENTED IN THE  *
*  DISCUSSION IN SECTION 5.2.2---FOR USE WITH CDC 6400    *
*  COMPUTER.                                              *
*                                                         *
*  PROGRAMMER---MOSHE SNIEDOVICH,DEPARTMENT OF HYDROLOGY  *
*               AND WATER RESOURCES,THE UNIVERSITY OF     *
*               ARIZONA,OCTOBER,1975.                     *
*                                                         *
*                                                         *
*  DATA CARDS FOLLOW THE FOLLOWING FORMATS                *
*    CARD 1                                               *
*       COL  1-5   NM   NUMBER OF YEARS,FORMAT I5         *
*            6-10  IXM  MAXIMUM STORAGE OF THE RESERVOIR  *
*                       FORMAT I5                         *
*           11-15  MR   MAXIMUM RELEASE,FORMAT I5         *
*           16-20  IXO  CRITICAL STORAGE LEVEL,FORMAT I5  *
*           21-25  IQM  MAXIMUM INFLOW,FORMAT I5          *
*    CARD 2                                               *
*       COL  1-80  AP   THE PROBABILITY MASS FUNCTION OF  *
*                       THE INFLOW IQ,IQ=1,2,...IQM.      *
*                       FORMAT IQM(F5.2).                 *
*                                                         *
*    NOTE----IXM,MR,IXO AND IQM ARE TAKEN TO BE GREATER   *
*            THAN THE ACTUAL QUANTITIES BY ONE UNIT.      *
***********************************************************
      COMMON NN,IXM,MR,IXO,IQM,AP(4),
     1ANGE(4,8,11),RANGE(4,8,11)
C    READING DATA
      READ(5,1)NN,IXM,MR,IXO,IQM
      READ(5,2)(AP(I),I=1,IQM)
C    PRINTING THE DATA
      WRITE(6,11)NN,IXM,MR,IXO,IOM
      WRITE(6,2)(AP(I),I=1,IQM)
C    LAST DECISIN STAGE N=NN
      DO10IMAX=IXO,IXM
      DO20IMIN=1,IXO
C    COMPUTING THE REWARD FOR THE LAST STAGE
      DO30IX=IMIN,IMAX
 30   RANGE(IMAX,IMIN,IX)=FLOAT(IMAX-IMIN)
 20   CONTINUE
 10   CONTINUE
C    DECISIN STAGES NN-1 TO 2
      NM=NN-2
      DO100M=1,NM
```

```
      N=NN-M+1
C     HEADING THE OUTPUT TABLE
      WRITE(6,3)N
      DO110IMAX=IXO,IXM
      IA=IMAX
      DO120IMIN=1,IXO
      IB=IMIN
      DO130IX=IMIN,IMAX
      IR=MR
      IF(IX.LT.MR)IR=IX
      AMIN=99999.9
      IRO=0
C     ITERATING OVER ALL FEASIBLE DECISIONS
      DO140ID=1,IR
      CALL EXPECT(IMAX,IMIN,IX,ID,EXP)
      SUM=EXP
      IF(AMIN.LE.SUM)GO TO 140
      AMIN=SUM
      IRO=ID
  140 CONTINUE
C     STORING THE OPTIMAL DECISION
      ANGE(IMAX,IMIN,IX)=AMIN
      IAA=IMAX-1
      IBB=IMIN-1
      ICC=IX-1
      IDD=IRO-1
C     PRINTING THE OPTIMAL DECISION AND THE OPTIMAL REWARD
      WRITE(6,4)IAA,IBB,ICC,IDD,AMIN
  130 CONTINUE
  120 CONTINUE
  110 CONTINUE
      WRITE(6,5)
C     RESTORING THE OPTIMAL REWARD
      DO99I=IXO,IXM
      DO98J=1,IXO
      DO97K=J,I
   97 RANGE(I,J,K)=ANGE(I,J,K)
   98 CONTINUE
   99 CONTINUE
  100 CONTINUE
C     DECISION STAGE N=2
      N=2
C     HEADING THE OUTPUT TABLE
      WRITE(6,3)N
C     DETERMINING THE FEASIBLE SITUATIONS
      IO=IXO-MR
      DO200IX=IO,IXM
      IR=MR
      IA=IXO
      IB=IXO
      IF(IX.LT.IXO)IB=IX
```

```
      IMIN=IB
      IF(IX.GT.IXO)IA=IX
      IMAX=IA
      IF(IX.LT.MR)IR=IX
      AMIN=999999.9
      IRO=0
C    ITERATING OVER ALL FEASIBLE DECISIONS
      DO220ID=1,IR
      CALL EXPECT(IMAX,IMIN,IX,ID,EXP)
      SUM=EXP
      IF(AMIN.LE.SUM)GO TO 220
      AMIN=SUM
      IRO=ID
  220 CONTINUE
C    STORING THE OPTIMAL DECISION
      ANGE(IMAX,IMIN,IX)=AMIN
      IAA=IMAX-1
      IBB=IMIN-1
      ICC=IX-1
      IDD=IRO-1
C    PRINTING THE OPTIMAL DECISION AND THE OPTIMAL REWARD
      WRITE(6,4)IAA,IBB,ICC,IDD,AMIN
  200 CONTINUE
      WRITE(6,5)
C    RESTORING THE OPTIMAL REWARD
      DO999I=IXO,IXM
      DO988J=1,IXO
      DO977K=J,I
  977 RANGE(I,J,K)=ANGE(I,J,K)
  988 CONTINUE
  999 CONTINUE
C    DECISION STAGE N=1
C    DETERMINING THE FEASIBLE SITUATIONS
      IMIN=IXO
      IMAX=IXO
      IX=IXO
      AMIN=999999.9
      IRO=0
C    ITERATING OVER ALL FEASIBLE DECISIONS
      DO310ID=1,IR
      CALL EXPECT(IMAX,IMIN,IX,ID,EXP)
      SUM=EXP
      IF(AMIN.LE.SUM)GO TO 310
      AMIN=SUM
      IRO=ID
  310 CONTINUE
      TOTAL=AMIN
      IXX=IXO-1
      IDD=IRO-1
C    PRINTING THE OPTIMAL DECISION AND THE OPTIMAL REWARD
      WRITE(6,6)IXX,IDD,TOTAL
```

```
C    FORMAT STATMENTS
C     INPUTS FORMAT
 1     FORMAT(5I5)
 2     FORMAT(4F5.1)
C     HEADING OF THE OUTPUT TABLE
 3     FORMAT(1H1,47X,*OPTIMAL RELEASE G AND REWARD R FOR N*
      1,*=*,I2,//35X,60(*-*),
      2/35X,*I*,T53,*U*,T70,*I*,T96,*I*,
      3/35X,*I*,T70,*I*,T77,*G(U)*,T87,*R(U,G)*,T96,*I*,
      4/35X,*I*,T42,*U(1)*,T52,*U(2)*,T62,*U(3)*,T70,*I*,T96
      4,*I*,/T36,60(*-*))
C     THE ROWS OF THE OUTPUT TABLE
 4     FORMAT(T36,*I*,T43,I2,T53,I2,T63,I2,T70,*I*,
      1T78,I2,T88,F5.2,T96,*I*)
C     THE OUTPUT FOR N=1
 6     FORMAT(1H1,30X,*THE OPTIMAL RELEASE FOR *,I2,
      1* IS *,I2,
      2* AND THE OPTIMAL TOTAL REWARD IS *,F5.2,
      3/25X,80(*-*))
C     THE LAST LINE OF THE OUTPUT TABLE
 5     FORMAT(T36,60(*-*))
C     THE DATA PRINTOUT
 11    FORMAT(1H1,10X,*INPUT DATA*,//5I5)
       STOP
       END
```

```
      SUBROUTINE EXPECT(IMAX,IMIN,IX,ID,EXP)
C   SUBROUTINE EXPECT COMPUTES THE EXPECTED VALUE,EXP,OF
C   THE REWARD ASSOCIATED WITH THE MODIFIED PROBLEM
C   (IMAX,IMIN,IX) AT TIME N AND THE DECISION ID,ASSUMING
C   THAT AN OPTIMAL FEASIBLE STRATEGY IS USED FOR TIMES
C   GREATER THAN N.
      COMMON NN,IXM,MR,IXO,IQM,AP(4),
     1ANGE(4,8,11),RANGE(4,8,11)
      IW=IX-ID
      EXP=0.
      DO10IQ=1,IQM
      IA=IMAX
      IB=IMIN
      IY=IW+IQ
      IZ=IY
      IF(IY.GT.IXM)IZ=IXM
      IF(IZ.LT.IMIN)IB=IZ
      IF(IZ.GT.IMAX)IA=IZ
 10   EXP=EXP+AP(IQ)*RANGE(IA,IB,IZ)
      RETURN
      END
```

LIST OF REFERENCES

Aris, R., Discrete Dynamic Programming, Blaisdell, New York, 1964.

Askew, Arthur J., Chance Constrained Dynamic Programming and the Optimization of Water Resource Systems, Water Resources Research, Vol. 10, No. 6, December 1974, pp. 1099-1106.

Bellman, R., On the Theory of Dynamic Programming, Proceedings, National Academy of Sciences, Vol. 38, 1952, pp. 716-719.

Bellman, R., An Introduction to the Theory of Dynamic Programming, Rand Report R-245, The Rand Corporation, Santa Monica, California, June 1953.

Bellman, R., The Theory of Dynamic Programming, Bulletin of American Mathematical Society, Vol. 60, 1954, pp. 503-516.

Bellman, R., Dynamic Programming, Princeton University Press, Princeton, N. J., 1957.

Bellman, R., and S. Dreyfus, Applied Dynamic Programming, Princeton University Press, Princeton, N. J., 1962.

Beveridge, Gordon S. G. and Robert S. Schechter, Optimization: Theory and Practice, McGraw-Hill Book Company, New York, 1970.

Blackwell, D., Discrete Dynamic Programming Annuals of Mathematical Statistics, Vol. 33, 1962, pp. 719-726.

Blackwell, D., Discounted Dynamic Programming, Annuals of Mathamatical Statistics, Vol. 36, 1965, pp. 226-235.

Denardo, E. V., Sequential Decision Processes, Ph. D. Dissertation, Industrial Engineering and Management Science, Northwestern University, Evanston, Ill., 1965.

Derman, C., Denumerable State Markovian Decision Processes -- Average Cost Criterion, Annuals of Mathematical Statistics, Vol. 37, 1966, pp. 1545-1554.

Dorfman, R., Design of Water Resources Systems, (Arthur Mass, ed.) Harvard University Press, Cambridge, Mass., 1962, pp. 88-158.

Dynkin, E. B., Controlled Random Sequences, Theory of Probability and Its Applications, Vol. 10, 1965, pp. 1-14.

Heidari, M., A Differential Dynamic Programming Approach to Water Resources Systems, Ph. D. Dissertation, Department of Civil Engineering, University of Illinois, Urbana, Ill., 1970.

Hinderer, K., Foundation of Non-Stationary Dynamic Programming with Discrete Time Parameter, Springer-Verlag, New York, 1970.

Howard, R. A., Dynamic Programming and Markov Processes, Technology Press and Wiley, New York, 1960.

Karlin, S., The Structure of Dynamic Programming Models, Naval Res. Logist, Quart. 2, 1955, pp. 285-294.

Larson, Robert E., State Increment Dynamic Programming, American Elsevier Publishing Company, Inc., New York, 1968.

Loeve, M., Probability Theory, Van Nostrand, Princeton, N. J., 2nd Edition, 1960.

Maitra, A., Discounted Dynamic Programming on Compact Metric Spaces, Sankhya 30A, 1968, pp. 211-216.

Miller, B. L. and A. F. Veinott, Discrete Dynamic Programming with a Small Interest Rate, Annuals of Mathematical Statistics, Vol. 40, 1969, pp. 366-370.

Mitten, L. G., Composition Principles for Synthesis of Optimal Multistage Processes, Operations Research, Vol. 12, No. 4, 1964, pp. 610-619.

Mitten, L. G., Preference Order Dynamic Programming, Management Science, Vol. 21, No. 1, 1974, pp. 43-46.

Nemhauser, G. L., Introduction to Dynamic Programming, Wiley, New York, 1966.

Sirjaev, A. N., Some New Results in the Theory of Controlled Random Processes, Transactions of the 4th Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, Prague 1965 (Russian), English translation in Selected Translations in Mathematical Statistics and Probability 8, 1970, pp. 49-130.

Sniedovich, M. and D. R. Davis, Comment on "Chance Constrained Dynamic Programming and the Optimization of Water Resource Systems," to appear in Water Resources Research, 1976.

Sobel, J. M., Ordinal Dynamic Programming, Management Science, Vol. 21, No. 9, 1975, pp. 967-975.

White, D. J., Dynamic Programming and Probabilistic Constraints, Operation Research, Vol. 22, No. 3, 1974, pp. 654-664.

Yakowitz, S. J., Mathematics of Adaptive Control Processes, American Elsevier Publishing Company, Inc., New York, 1969.