

Dynamic Programming and Principles of Optimality

MOSHE SNIEDOVICH

*Department of Civil Engineering, Princeton University,
Princeton, New Jersey 08540*

Submitted by E. S. Lee

A sequential decision model is developed in the context of which three principles of optimality are defined. Each of the principles is shown to be valid for a wide class of stochastic sequential decision problems. The relationship between the principles and the functional equations of dynamic programming is investigated and it is shown that the validity of each of them guarantees the optimality of the dynamic programming solutions. As no monotonicity assumption is made regarding the reward functions, the results presented in this paper resolve certain questions raised in the literature as to the relation among the principles of optimality and the optimality of the dynamic programming solutions.

1. INTRODUCTION

... It all started in the early 1950s when the principle of optimality and the functional equations of dynamic programming were introduced by Bellman [1, p. 83]. Let us recall Bellman's statement, noting that this statement was made in the context of certain decision processes where the notion of optimality regarding policies was associated with a preassigned criterion function defined over the final state variables. The statement [1, 83] is as follows:

In each process, the functional equation governing the process was obtained by an application of the following intuitive: 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.*

The mathematical transliteration of this simple principle will yield all the functional equations we shall encounter throughout the remainder of the book. A proof by contradiction is immediate. ■

As simple and intuitive as the principle might be, many investigators, including this author, find it difficult to provide a precise statement of its meaning in terms of the elements of the process under consideration. Denardo [2, p. 36], for example, raises five questions with respect to which the principle seems to be ambiguous, and he adds the following statement: "Much of the ambiguity in

the preceding statement of the Principle of Optimality could be removed by a longer and more careful statement along the same lines...."

According to Aris [3, p. 27] "... The rather loose wording of this statement is said to be intentional and indicates that the user should think carefully before applying it...."

It seems, however, as if the wording of the statement follows Bellman's [1, p. 82] statement concerning the spirit of dynamic programming: "... It is extremely important to realize that one can neither axiomatize mathematical formulation nor legislate away ingenuity...."

Many investigators [2, p. 27; 3, p. 27; 4, p. 20] define their own principles of optimality, which they consider to be either a more precise statement of Bellman's principle or even a better alternative.

The purpose of our discussion is *not* to try to clarify Bellman's statement of the principle and certainly not to add another interpretation. The objective of this paper is to discuss a few of the principles currently in use in the literature, to demonstrate the difference among them, and above all to investigate the relationship among them and the functional equations characterizing dynamic programming.

In particular we will investigate the conditions under which the validity of the principles guarantee the optimality of the solutions derived from the functional equations.

It should be noted that although in the literature [5, p. 1198; 6, p. 1179; 7, p. 480; 8, p. 269] authors often justify the optimality of the dynamic programming solutions by invoking the principle of optimality, as indicated by Yakowitz [9, p. 43] and Hinderer [4, pp. 14-15], there seems to be no formal proof to justify such an exposition.

Our discussion will be restricted to discrete sequential decision processes in which the criterion functions are real valued functions. Recent papers by Mitten [10] and Sobel [11] indicate the potential use of dynamic programming in decision processes in which the optimality of policies is established by means of a preference order over the set of all feasible policies.

Sections 2 through 8 are introductory sections; they contain some new results regarding the validity of the principles in processes in which the reward functions are not additive.

The main results concerning the relationship among the principles of optimality and the functional equations are presented in Section 9 and summarized in Section 11.

2. SEQUENTIAL DECISION MODEL

The sequential decision model under consideration consists of the set $\mathbb{N} = \{n = 1, 2, \dots\}$ of *decision stages* and the elements S , D , p_0 , $(p_n, n \in \mathbb{N})$, $(D_n, n \in \mathbb{N})$, and $(r_n, n \in \mathbb{N})$ where:

(1) S , the *state space* is a nonempty countable set of elements s called states.

(2) D , the *decision space*, is a nonempty set of elements d called *decisions*. A *history* h_n is an element of the set \bar{H}_n defined as follows:

$$\bar{H}_1 := S \quad (2.1)$$

and

$$\bar{H}_n := S \times D \times S \times D \times \cdots \times S \quad (2n - 1 \text{ factors}), \quad n > 1. \quad (2.2)$$

We let

$$\bar{H} := S \times D \times S \cdots, \quad (2.3)$$

and we denote its elements by h .

(3) p_0 , the *initial distribution*, is a counting density on S .

(4) p_n , the *transition law* between the n th and $(n + 1)$ st decision stages, is a family of counting densities on S conditioned on $h_n \in \bar{H}_n$, $d_n \in D$. More specifically, $p_n(h_n, d_n, \cdot)$ is a counting density on S , $n \in \mathbb{N}$, $h_n \in \bar{H}_n$, $d_n \in D$.

(5) D_n , the *admissible decision map* associated with the n th decision stage is a map from a certain subset H_n of \bar{H}_n to the set of all subsets of D with the following properties:

$$H_1 := \{s: p_0(s) > 0, s \in S\} \quad (2.4)$$

and

$$H_{n+1} := \{(h_n, d_n, s_{n+1}): h_n \in H_n, d_n \in D_n(h_n), p_n(h_n, d_n, s_{n+1}) > 0\}. \quad (2.5)$$

We denote by H the subset of \bar{H} whose elements $h = (s_1, d_1, s_2, \dots)$ have the property that $(s_1, d_1, \dots, s_n) \in H_n$, $\forall n \in \mathbb{N}$. In other words,

$$H := \{h: h = (s_1, d_1, s_2, d_2, \dots) \in \bar{H}, (s_1, h_1, \dots, s_n) \in H_n, \forall n \in \mathbb{N}\}. \quad (2.6)$$

The set H_n is called the set of *admissible histories* associated with the n th decision stage, whereas $D_n(h_n)$ is called the *set of feasible decisions* associated with the history h_n .

(6) r_n , the *reward function* associated with the n th decision stage is a real valued function on H .

DEFINITION 2.1. A *sequential decision model* is any tuple $W = (S, D, p_0, (p_n: n \in \mathbb{N}), (D_n, n \in \mathbb{N}), (r_n, n \in \mathbb{N}))$, where S is a state space, D is a decision space, p_0 is an initial distribution, $(p_n, n \in \mathbb{N})$ is a sequence of transition laws, $(D_n, n \in \mathbb{N})$ is a sequence of admissible decision maps, and $(r_n, n \in \mathbb{N})$ is a sequence of reward functions. ■

The sequential decision model presented above is a generalization of Hinderer's [4] model in the sense that the reward functions under consideration are *all* defined on H and they are not assumed to be additive.

The procedure used for selecting the decisions is determined by a sequence of maps from H_n to D , $n \in \mathbb{N}$, respectively.

DEFINITION 2.2. A *plan* $\delta = (\delta_1, \delta_2, \dots)$ is a sequence of maps $\delta_n: H_n \rightarrow D$, $n \in \mathbb{N}$. The plan $\delta = (\delta_1, \delta_2, \dots)$ is said to be *feasible* if and only if $\delta_n(h_n) \in D_n(h_n)$, $\forall n \in \mathbb{N}$, $h_n \in H_n$. The set of all the feasible plans is denoted by Δ , i.e.,

$$\Delta := \bigcap_{n=1}^{\infty} \Delta_n, \Delta_n := \{\delta_n: \delta_n(h_n) \in D_n(h_n), \forall h_n \in H_n\}. \quad \blacksquare \quad (2.7)$$

The application of the plan $\delta = (\delta_1, \delta_2, \dots) \in \Delta$ generates a process which schematically may be described as follows: The process starts at $n = 1$, where s_1 is selected from S according to the initial distribution p_0 ; then the decision $d_1 = \delta_1(s_1)$ is made and the process moves to the stage $n = 2$ where the state s_2 is selected from S according to the transition law $p_1(s_1, d_1, \cdot)$; then the decision $d_2 = \delta_2(h_2)$, $h_2 = (s_1, d_1, s_2)$ is made and the process moves to the stage $n = 3$ where s_3 is selected from S according to the transition law $p_2(h_2, d_2, \cdot)$, etc. The rewards associated with the realization of $h = (s_1, d_1, s_2, d_2, \dots)$ and the n th decision stage is $r_n(h)$, $n \in \mathbb{N}$. That is, $r_n(h)$ is the reward associated with the history h and the n th decision stage.

The mathematical model describing the above process is the probability space $(\Omega, \mathcal{O}, P_\delta)$ and the random variables x_n , $n \in \mathbb{N}$, where:

(1) $\Omega := S^\infty$ is the sample space, i.e., the set of all sequences $\omega = (s_1, s_2, \dots)$.

(2) \mathcal{O} is the infinite product σ -algebra on Ω determined by the factors consisting of all subsets of S .

(3) x_n is the n th coordinate variable, i.e.,

$$x_n(s_1, s_2, \dots) := s_n, \quad n \in \mathbb{N}. \quad (2.8)$$

(4) P_δ is the unique [12] probability measure on \mathcal{O} satisfying the condition:

$$P_\delta(\gamma_n(\omega) = (s_1, s_2, \dots, s_n)) = p_0(s_1) \cdot p_1(s_1, d_1, s_2) \cdots p_{n-1}(h_{n-1}, d_{n-1}, s_n), \quad (2.9)$$

where $h_1 = s_1$, $d_1 = \delta_1(h_1)$, $h_{m+1} = (h_m, \delta_m(h_m), s_{m+1})$, $d_{m+1} = \delta_{m+1}(h_{m+1})$, and γ_n is the trajectory vector, i.e.,

$$\gamma_n(s_1, s_2, \dots) := (s_1, s_2, \dots, s_n), \quad n \in \mathbb{N}, \quad (2.10)$$

We shall use the symbol $h_{n,\delta}(y_n)$, $y_n = (s_1, s_2, \dots, s_n)$ to denote the history associated with the trajectory y_n and the plan δ , i.e.,

$$h_{n,\delta}(y_n) := (s_1, \delta_1(s_1), s_2, \delta_2(h_2), \dots, s_n), \quad n \in \mathbb{N}, \quad \delta \in \Delta, \quad (2.11)$$

$$h_{m+1} = (h_m, \delta_m(h_m), s_{m+1})$$

with

$$h_\delta(\omega) = (s_1, d_1, s_2, d_2, \dots) \in H, \quad d_1 = \delta_1(s_1), \quad (2.12)$$

$$h_n = (s_1, d_1, \dots, s_n), \quad d_n = \delta_n(h_n).$$

Associated with each element $\delta = (\delta_1, \delta_2, \dots) \in \Delta$ let $Y_n(\delta)$ and $H_n(\delta)$ be the subsets of S^n and H_n , respectively, defined as follows:

$$Y_1(\delta) := H_1(\delta) := H_1, \quad (2.13)$$

$$H_{n+1}(\delta) := \{(h_n, d_n, s_{n+1}) : h_n \in H_n(\delta), d_n = \delta_n(h_n), p_n(h_n, d_n, s_{n+1}) > 0\}, \quad (2.14)$$

and

$$Y_n(\delta) := \{(s_1, s_2, \dots, s_n) : h_{n,\delta}(s_1, s_2, \dots, s_n) \in H_n(\delta)\}. \quad (2.15)$$

The sets $H_n(\delta)$ and $Y_n(\delta)$ will be referred to as the *set of feasible histories and trajectories* associated with the n th decision stage, respectively.

For every sequence $\omega = (s_1, s_2, \dots) \in \Omega$, $n \in \mathbb{N}$, and $\delta \in \Delta$ let

$$l_{n,\delta}(\omega) := r_n(h_\delta(\omega)). \quad (2.16)$$

Thus, $l_{n,\delta}$ is a well-defined real valued random variable with respect to $(\Omega, \mathcal{A}, P_\delta)$ for every $n \in \mathbb{N}$ and $\delta \in \Delta$.

ASSUMPTION 2.1. The expected value $E_\delta[l_{n,\delta}(\omega)]$ of $l_{n,\delta}(\omega)$ with respect to $\delta \in \Delta$ exists for all $n \in \mathbb{N}$ and $\delta \in \Delta$. We denote this expectation by $R_{n,\delta}$. That is

$$R_{n,\delta} := E_\delta[l_{n,\delta}(\omega)] := \int_\Omega l_{n,\delta}(\omega) dP_\delta(\omega). \quad (2.17)$$

Moreover, it is also assumed that the conditional expectation of $l_{n,\delta}(\omega)$ given $h_m \in H_m$, denoted by $R_{n,\delta}(h_m)$, exists for all $m, n \in \mathbb{N}$, $\delta \in \Delta$ and $h_m \in H_m$, and if $(\delta^k : k = 1, 2, \dots)$ and $\delta \in \Delta$ are such that for every $k \in \mathbb{N}$, $\delta_m^k \equiv \delta_m$, $\forall m \leq k$ then $(R_{n,\delta^k}(h_n) : k = 1, 2, \dots)$ converges to $R_{n,\delta}(h_n)$ for all $n \in \mathbb{N}$, $h_n \in H_n$. ■

ASSUMPTION 2.2.

$$(1) \quad R_n := \sup_{\delta \in \Delta} R_{n,\delta} \quad \text{exists } \forall n \in \mathbb{N}, \quad (2.18)$$

$$(2) \quad R_n(h_m) := \sup_{\delta \in \Delta} R_{n,\delta}(h_m) \quad \text{exists for all } m, n \in \mathbb{N}, \quad h_m \in H_m. \quad \blacksquare \quad (2.19)$$

Details concerning sufficient conditions for the validity of these assumptions can be found in [2, 4, 9]. For example, if S and D are finite and the model is truncated [2, p. 30], these assumptions are valid. Models satisfying the two assumptions indicated above will be referred to as *regular* models.

DEFINITION 2.3. The plan $\delta \in \Delta$ is said to be *optimal* with respect to the *original problem*, or simply *optimal*, if and only if

$$R_{1,\delta} = R := R_1. \quad (2.20)$$

The set of all the optimal plans is denoted by Δ^* and its elements by δ^* . ■

Notice that by definition the notion of optimality as far as plans are concerned is related to r_1 .

DEFINITION 2.4. The plan $\delta \in \Delta$ is said to be *optimal* at the *modified problem* (h_n, n) , $n \in \mathbb{N}$, $h_n \in H_n$, if and only if

$$R_{n,\delta}(h_n) = R_n(h_n). \quad \blacksquare \quad (2.21)$$

Now that the sequential decision model is defined including the notion of optimal plans, three principles of optimality are introduced. The strong principle, the weak principle, and the dynamic programming principle.

DEFINITION 2.5 (The Strong Principle of Optimality). The *strong principle of optimality* is said to hold for the sequential decision model W if and only if $\delta^* \in \Delta^*$ implies that

$$R_{1,\delta^*}(h_n) = R_1(h_n), \quad \forall n \in \mathbb{N}, \quad h_n \in H_n(\delta^*). \quad (2.22)$$

That is, the strong principle is said to be valid if and only if *every* optimal plan $\delta^* \in \Delta^*$ is also optimal, with respect to r_1 , at all the modified problems it generates with positive probability. ■

It should be noted that the strong principle of optimality is concerned *only* with r_1 .

DEFINITION 2.6 (The Weak Principle of Optimality). The *weak principle of optimality* is said to hold for the model W if and only if $\delta^* \in \Delta^*$ implies that

$$R_{n,\delta^*}(h_n) = R_n(h_n), \quad \forall n \in \mathbb{N}, \quad h_n \in H_n(\delta^*). \quad (2.23)$$

That is, the weak principle is said to be valid if and only if every optimal plan $\delta^* \in \Delta^*$ is also simultaneously optimal, with respect to R_n , $n \in \mathbb{N}$, at all the modified problems (h_n, n) it generates with positive probability. ■

Notice that as in the case of the strong principle, (2.23) is restricted to $h_n \in H_n(\delta^*)$, that is, the histories generated by δ^* with positive probability.

DEFINITION 2.7 (The Dynamic Programming Principle). The *dynamic programming principle* is said to hold for the model W if and only if there exists an element $\delta^* \in \Delta^*$ with the property that

$$R_{n,\delta^*}(h_n) = R_n(h_n), \quad \forall n \in \mathbb{N}, \quad h_n \in H_n. \quad (2.24)$$

That is, the dynamic programming principle is said to be valid if and only if there exists an optimal plan which is also optimal with respect to *all* the modified problems. ■

Notice that the dynamic programming principle is related to the *existence* of a certain plan while the other two are related to certain properties possessed by *all* the optimal plans.

No attempt has been made here to interpret Bellman's principle of optimality in the context of the sequential decision model. Instead, the elements of the principle will be compared with those of the strong principle.

3. BELLMAN'S PRINCIPLE

Three observations are of importance while attempting to interpret Bellman's principle in the context of models other than the one introduced by Bellman [1, pp. 81–82].

First, according to Bellman [1, p. 82] "... an *optimal policy* is a policy which maximizes a preassigned function of the final state variables." That is, only *one* objective function is considered. Second, no assumption concerning the structure of the objective function is made, except for some regularity conditions.

Finally, the principle is associated only with states that are generated by the optimal plans.

Thus, not only that the notion of optimality as far as plans are concerned is associated with a single objective function, but this function is assumed neither to be decomposable, nor to possess any monotonicity properties.

As will be shown in the next section, the generality of Bellman's statement can be maintained in the context of the stochastic sequential model presented in this paper. That is, the strong principle of optimality will be shown to hold for *all* regular decision models.

Since each of the modified problems (h_n, n) can be viewed as an original problem with respect to the reward function r_n , Bellman's principle can be applied successively at each of the modified problems.

Thus the sequential decision problem can be viewed as a family of embedded problems to each of which the principle can be applied.

However, in order to prove the validity of the principle at each of the modified problems, it is necessary to impose certain monotonicity properties on the relation between r_n and r_{n+1} .

4. THE STRONG PRINCIPLE OF OPTIMALITY

The strong principle of optimality preserves two important characteristics of Bellman's principle. That is, only one criterion function is under consideration as far as the optimality of plans is considered, i.e., r_1 , and the statement regarding the optimality of plans is related to histories which are observed with positive probability. Moreover, the following also indicates the generality of the principle as far as its validity is concerned.

THEOREM 3.1. *The strong principle of optimality holds for any arbitrary regular decision model.*

Proof. Contrary to the above statement, assume that there exist $\delta^* \in \Delta^*$, $n \in \mathbb{N}$, $h_n^* \in H_n(\delta^*)$, and $\delta \in \Delta$ such that

$$R_{1,\delta}(h_n^*) > R_{1,\delta^*}(h_n^*). \quad (4.1)$$

Construct the plan δ' as follows:

$$\begin{aligned} \delta'_m(h_m) &= \delta_m^*(h_m), & m < n, & \quad h_m \in H_m, \\ &= \delta_m(h_m), & m = n, & \quad h_m = h_n^*, \\ &= \delta^*(h_m), & m = n, & \quad h_m \in \{H_n - \{h_n^*\}\}, \\ &= \delta_m(h_m), & m > n, & \quad h_m = (h_n^*, d_1, \dots, s_m) \in H_m, \\ &= \delta_m^*(h_m), & & \quad \text{otherwise.} \end{aligned} \quad (4.2)$$

Obviously by construction $\delta' \in \Delta$, $R_{1,\delta'}(h_n) > R_{1,\delta^*}(h_n)$, and

$$R_{1,\delta'}(h_n) \geq R_{1,\delta^*}(h_n), \quad \forall h_n \in \{H_n - \{h_n^*\}\}. \quad (4.3)$$

Since h_n^* is observed with positive probability under both δ^* and δ' , and since by construction

$$P_{\delta'}(h_{n,\delta'}(\omega) = h_n) = P_{\delta^*}(h_{n,\delta^*}(\omega) = h_n), \quad \forall h_n \in H_n(\delta^*), \quad (4.4)$$

it follows that $R_{1,\delta'} > R_{1,\delta^*}$. This, however, contradicts the optimality of δ^* . Thus, there exist no such $\delta^* \in \Delta^*$, $n \in \mathbb{N}$, $h_n^* \in H_n(\delta^*)$, and $\delta \in \Delta$ and hence the above statement is true. ■

Remarks. (1) It should be noted that no assumption has been made concerning the structure of the function r_1 other than the one concerning the regularity of the decision model.

(2) The term "strong" has been introduced so as to emphasize that the principle holds for *all* regular models. ■

To the best of the author's knowledge this form of the principle has never been introduced in the literature, either as an interpretation of Bellman's principle or otherwise.

As simple as the strong principle may be, it is the author's view that it provides the foundation not only to the other two principles but to the dynamic programming approach as a whole. Although it will be pure speculation to suggest that the strong principle is indeed the extension of Bellman's principle to stochastic sequential processes, the author does believe that the strong principle maintains few of the basic qualities of Bellman's principle. This subject will be discussed in Section 10.

5. THE WEAK PRINCIPLE OF OPTIMALITY

In many sequential decision processes there is an intimate relationship among the elements r_n , $n \in \mathbb{N}$. Consider, for example, the additive reward function r_1 , where

$$r_1(s_1, d_1, s_2, d_2, \dots) = \sum_{n=1}^{\infty} q_n(h_n, d_n), \quad h_n = (s_1, d_1, \dots, s_n). \quad (5.1)$$

If r_1 is the original criterion function under consideration one can generate a sequence $(r_n, n > 2)$ as follows:

$$r_n(s_1, d_1, s_2, d_2, \dots) := \sum_{m=n}^{\infty} q_m(h_n, d_m), \quad n > 1, \quad (5.2)$$

so that

$$r_n(s_1, d_1, s_2, \dots) = q_n(h_n, d_n) + r_{n+1}(s_1, d_1, s_2, \dots) \quad (5.3)$$

or

$$r_n(h) = q_n(h_n, d_n) + r_{n+1}(h). \quad (5.4)$$

For this case it can be easily verified that

$$R_{n,\delta}(h_{n+1}) = q_n(h_n, d_n) + R_{n+1,\delta}(h_{n+1}). \quad (5.5)$$

In certain situations, then, there is a close relationship between r_n and r_{n+1} , and consequently between $R_{n,\delta}(h_{n+1})$ and $R_{n+1,\delta}(h_{n+1})$.

DEFINITION 5.1. Let W be a regular sequential decision model for which there exists a sequence $\rho = (\rho_1, \rho_2, \dots)$ of real valued functions on $H_n \times D \times \mathbb{R}$, $\mathbb{R} := (-\infty, +\infty)$, respectively, such that

$$R_{n,\delta}(h_{n+1}) = \rho_n(h_n, d_n, R_{n+1,\delta}(h_{n+1})) \quad (5.6)$$

for all $n \in \mathbb{N}$, $\delta \in \Delta$, $h_n \in H_n$, and $d_n \in D$.

The model is said then to be *decomposed* by ρ .

If in addition $\rho_n(h_n, d_n, \cdot)$ is a monotone/strictly monotone increasing function the model is said to be monotone/strictly monotone with respect to ρ . ■

It is obvious that any regular model having $(r_n, n \in \mathbb{N})$ defined by (5.2) as its sequence of reward function is strictly monotone. More details about monotone models can be found in [2, 13].

THEOREM 5.1. *The weak principle of optimality holds for all regular strictly monotone models.*

Proof. Contrary to the above statement assume that there exist $n \in \mathbb{N}$, $\delta^* \in \Delta^*$, $h_n^* \in H_n(\delta^*)$, and $\delta \in \Delta$ such that

$$R_{n,\delta}(h_n^*) > R_{n,\delta^*}(h_n^*). \quad (5.7)$$

Let δ' be the plan defined by (4.2). Using the strict monotonicity of ρ_m , $m < n$, it can easily be verified by backward induction on m that for every $m \leq n$ there exists at least one element $h_m^* \in H_m(\delta^*)$ with the property that

$$(1) \quad R_{m,\delta'}(h_m^*) > R_{m,\delta^*}(h_m^*) \quad (5.8)$$

and

$$(2) \quad R_{m,\delta'}(h_m^*) \geq R_{m,\delta^*}(h_m), \quad \forall h_m \in \{H_m - \{h_m^*\}\}. \quad (5.9)$$

This implies that $R_{1,\delta'} > R_{1,\delta^*}$, which contradicts the optimality of δ^* . Thus, there exist no such $n \in \mathbb{N}$, $\delta^* \in \Delta^*$, $h_n^* \in H_n(\delta^*)$, and $\delta^* \in \Delta$, and hence the above statement is true. ■

To the best of the author's knowledge, Theorem 5.1 provides the most general conditions for the validity of the weak principle of optimality. Theorem 5.1 may be viewed as a generalization of Hinderer's [4] results to nonadditive processes.

It should be noted that the weak principle of optimality is not equivalent to the principle used by Denardo [2, p. 37].

In order to demonstrate that the strict monotonicity property of ρ_n , $n \in \mathbb{N}$ is indeed essential for the validity of the weak principle consider the following example.

EXAMPLE 5.1. Consider the decision model W for which $S = \{0, 1\}$; $D = \{0, 1\}$; $p_0(0) = 1$, $p_n(h_n, 1, 1) = 1$, $\forall n \in \mathbb{N}$, $h_n \in H_n$; $p_n(h_n, 0, 0) = 1$, $\forall n \in \mathbb{N}$, $h_n \in H_n$; $D_n(h_n) = D$, $\forall n \in \mathbb{N}$, $h_n \in H_n$; and

$$r_n(s_1, d_1, s_2, \dots) = \max_{m \geq n} \{s_m\}. \quad (5.10)$$

Notice that

$$(1) \quad r_n(s_1, d_1, s_2, \dots) = \max\{s_n, r_{n+1}(s_1, d_1, s_2, \dots)\}, \quad (5.11)$$

and since the process is deterministic also

$$(2) \quad R_{n,\delta}(h_{n+1}) = \max\{s_n, R_{n+1,\delta}(h_{n+1})\}. \quad (5.12)$$

It follows then that the model is monotone with respect to $\rho = (\rho_1, \rho_2, \dots)$, where

$$\rho_n(h_n, d_n, a) = \max\{s_n, a\}. \quad (5.13)$$

Consider the plan $\delta^* = (\delta_1^*, \delta_2^*, \dots)$, where

$$\begin{aligned} \delta_n^*(h_n) &= 1, & n &= 1, & h_n &\in H_n, \\ &= 0, & n &> 1, & h_n &\in H_n. \end{aligned}$$

Obviously $R_{\delta^*} = 1$, which is an optimal value, and hence δ^* is optimal. Consider the modified problem $(h_3, 3)$, where $h_3 = (0, 1, 1, 0, 0)$ for which

$$R_{3,\delta^*}(h_3) = 0.$$

It is obvious then that δ^* is not optimal at the modified problem, $(h_3, 3)$, since $R_3(h_3) = 1$. Thus, the weak principle of optimality does not hold for W . ■

6. THE DYNAMIC PROGRAMMING PRINCIPLE

We will show now that the dynamic programming principle holds for all regular monotone models. As this principle is concerned with the *existence* of a certain plan it is not actually needed to specify it. However, the proof of the validity of the principle will also indicate the specific structure of this plan, and consequently a potential mechanism for its construction.

THEOREM 6.1. *The dynamic programming principle holds for all regular monotone models.*

Proof. We will first prove by induction on n that for each $n \in \mathbb{N}$ there exists a plan $\delta^n \in \mathcal{A}^*$ with the property that

$$R_{i,\delta^n}(h_i) = R_i(h_i), \quad \forall i \leq n, \quad h_i \in H_i. \quad (6.1)$$

For each $n \in \mathbb{N}$, $h_n \in H_n$, let δ^{h_n} be any arbitrary element of Δ for which

$$R_{n, \delta^{h_n}}(h_n) = R_n(h_n). \quad (6.2)$$

Since the model is regular, δ^{h_n} exists for all $n \in \mathbb{N}$ and $h_n \in H_n$. Associated with the set $\{\delta^{h_n}: n \in \mathbb{N}, h_n \in H_n\}$ define:

$$\begin{aligned} \pi(\delta, n) &:= \{\delta': \delta'_m(h_m) = \delta_m(h_m), \quad m < n, \quad h_m \in H_m, \\ &= \delta_m^{h_n}(h_m), \quad m = n, \quad h_m = h_n \in H_n, \\ &= \delta_m^{h_n}(h_m), \quad m > n, \quad h_m = (h_n, \dots, s_m) \in H_m\}. \end{aligned} \quad (6.3)$$

By definition, $\pi(\delta, n)$ has the following properties:

$$(1) \quad \pi(\delta, n) \in \Delta, \quad \forall n \in \mathbb{N}, \quad \delta \in \Delta, \quad (6.4)$$

$$(2) \quad R_{n, \pi(\delta, n)}(h_n) = R_n(h_n), \quad \forall n \in \mathbb{N}, \quad h_n \in H_n, \quad \delta \in \Delta, \quad (6.5)$$

$$(3) \quad R_{1, \pi(\delta, 1)}(h_1) = R_{1, \pi(\delta', 1)}(h_1), \quad \forall h_1 \in H_1, \quad \delta, \delta' \in \Delta. \quad (6.6)$$

Define:

$$\delta^1 := \pi(\delta, 1), \quad \delta \in \Delta \quad (6.7)$$

and

$$\delta^n := \pi(\delta^{n-1}, n), \quad n > 1. \quad (6.8)$$

By definition, $\delta^1 \in \Delta$ and

$$R_{1, \delta^1}(h_1) = R_1(h_1), \quad \forall h_1 \in H_1. \quad (6.9)$$

Thus, $R_{\delta^1} = R$ and hence $\delta^1 \in \Delta^*$. The inductive hypothesis is true then for $n = 1$. Assume that the inductive hypothesis is true for $n = 2, 3, \dots, m$. In particular, assume that for $n = m$, $\delta^m \in \Delta^*$ and

$$R_{i, \delta^m}(h_i) = R_i(h_i), \quad \forall i \leq m, \quad h_i \in H_i. \quad (6.10)$$

Consider $n = m + 1$ for which by construction

$$R_{m+1, \delta^{m+1}}(h_{m+1}) = R_{m+1}(h_{m+1}), \quad \forall h_{m+1} \in H_{m+1}. \quad (6.11)$$

Since the model is monotone it follows then that

$$R_{i, \delta^{m+1}}(h_i) = R_i(h_i), \quad \forall i \leq m + 1, \quad h_i \in H_i. \quad (6.12)$$

Since δ^{m+1} is feasible, it follows that the inductive hypothesis is true for $n = m + 1$, and hence it is true for all $n \in \mathbb{N}$.

Let $(\delta^n, n \in \mathbb{N})$ be the sequence of plan generated by π and any arbitrary element $\delta \in \Delta$, i.e. $\delta' = \pi(\delta, 1)$, $\delta \in \Delta$. From the definition of π the sequence $(\delta^n, n \in \mathbb{N})$ is guaranteed to be unique. Construct the plan $\delta^* = (\delta_1^*, \delta_2^*, \dots)$ as follows:

$$\delta_n^*(h_n) = \delta_n^{h_n}(h_n), \quad n \in \mathbb{N}, \quad h_n \in H_n. \quad (6.13)$$

Then by construction,

$$\delta_m^n = \delta_m^*, \quad \forall m \leq n \quad (6.14)$$

Thus, $\delta^* \in \Delta^*$ and

$$R_{n,\delta^*}(h_n) = R_n(h_n), \quad \forall n \in \mathbb{N}, \quad h_n \in H_n. \quad \blacksquare \quad (6.15)$$

As will be indicated in the sequel, the validity of the dynamic programming principle plays an important role in many dynamic programming investigations.

As a matter of fact, Mitten [10] incorporates the validity of this principle in the definition of an optimal plan. That is, a plan is said to be optimal if and only if it is optimal at *all* the modified problems.

It should be noted that the validity of the dynamic programming principle does not guarantee the validity of the weak principle. The following example demonstrates this point.

EXAMPLE 6.1. Consider the model W for which $S = \{0, 1\}$; $D = \{0, 1\}$; $D_n(h_n) = D$, $\forall n \in \mathbb{N}$, $h_n \in H_n$; $p_0(0) = 1$; $p_n(h_n, d_n, d_n) = 1$, $n \in \mathbb{N}$, $h_n \in H_n$; and $r_n(h) = \prod_{m=n}^{\infty} s_m$, $n \in \mathbb{N}$; Since all the elements of S are nonnegative,

$$R_{n,\delta}(h_n) = s_n \cdot R_{n+1,\delta}(h_{n+1}), \quad \forall n \in \mathbb{N}, \quad \delta \in \Delta, \quad h_n \in H_n,$$

so that the model is monotone. Obviously, $R = 0$, and $\Delta^* = \Delta$.

The plan $\delta^* = (\delta_1^*, \delta_2^*, \dots)$ for which $\delta_n^*(h_n) = 1$, $\forall n \in \mathbb{N}$, $h_n \in H_n$, is obviously optimal and also optimal with respect to all the modified problems. However, any other plan in Δ^* is *not* simultaneously optimal at all the modified problems. For example, the plan $\delta = (\delta_1, \delta_2, \dots)$, $\delta_n(h_n) = 0$, $\forall n \in \mathbb{N}$, $h_n \in H_n$ is optimal although it is *not* optimal with respect to any modified problem. \blacksquare

Thus, the validity of the dynamic programming principle does not guarantee the validity of the weak principle.

7. THE FUNCTIONAL EQUATIONS

Let W be a regular model and δ any arbitrary feasible plan. If the model is decomposed by $\rho = (\rho_1, \rho_2, \dots)$ then by definition, for all $n \in \mathbb{N}$, $h_n \in H_n$,

$$R_{n,\delta}(h_n) = \sum_{s_{n+1} \in S} \rho_n(h_n, \delta_n(h_n), R_{n+1,\delta}(h_n, \delta_n(h_n), s_{n+1})) \cdot p_n(h_n, \delta_n(h_n), s_{n+1}). \quad (7.1)$$

This relationship between $R_{n,\delta}(h_n)$ and $R_{n+1,\delta}(h_{n+1})$ is the basis for the consideration of the optimality equations and the dynamic programming algorithm as solution procedures for sequential decision problems.

DEFINITION 7.1. Let W be a regular decision model decomposed by $\rho = (\rho_1, \rho_2, \dots)$. The system of optimality equations is said to hold if and only if

$$R_n(h_n) = \sup_{d_n \in D_n(h_n)} \sum_{s_{n+1} \in S} \rho_n(h_n, d_n, R_{n+1}(h_n, d_n, s_{n+1})) \cdot p_n(h_n, d_n, s_{n+1}), \quad \forall n \in \mathbb{N}, \quad h_n \in H_n. \quad (7.2)$$

These equations will also be referred to as the *functional equations*. ■

For convenience, define

$$G_n(h_n) := \sup_{d_n \in D_n(h_n)} \sum_{s_{n+1} \in S} \rho_n(h_n, d_n, R_{n+1}(h_n, d_n, s_{n+1})) \cdot p_n(h_n, d_n, s_{n+1}), \quad n \in \mathbb{N}, \quad h_n \in H_n. \quad (7.3)$$

THEOREM 7. Let W be a regular decision model decomposed by $\rho = (\rho_1, \rho_2, \dots)$. Then if the model is monotone with respect to ρ the system of optimality equations holds.

Proof. Since the model is monotone with respect to ρ , Theorem 6.1 implies that there exists a plan $\delta^* \in \Delta^*$ with the property that

$$R_{n,\delta^*}(h_n) = R_n(h_n), \quad \forall n \in \mathbb{N}, \quad h_n \in H_n. \quad (7.4)$$

For this plan Eq. (6.1) implies that for any $n \in \mathbb{N}$, $h_n \in H_n$,

$$R_{n,\delta^*}(h_n) = \sum_{s_{n+1} \in S} \rho_n(h_n, \delta_n^*(h_n), R_{n+1}(h_n, \delta_n^*(h_n), s_{n+1})) \cdot p_n(h_n, \delta_n^*(h_n), s_{n+1}), \quad \forall n \in \mathbb{N}, \quad h_n \in H_n. \quad (7.5)$$

The mononicity of $\rho_n(h_n, \delta_n^*(h_n), \cdot)$ and the fact that $\delta_n^*(h_n) \in D_n(h_n)$ imply then that

$$R_{n,\delta^*}(h_n) \leq G_n(h_n), \quad \forall n \in \mathbb{N}, \quad h_n \in H_n. \quad (7.6)$$

However, if for some $n \in \mathbb{N}$, and $h'_n \in H_n$, $R_{n,\delta^*}(h'_n)$ is strictly less than $G_n(h'_n)$, we have a contradiction to the optimality of δ^* at (h'_n, n) . That is, let d'_n be any element of $D_n(h'_n)$ for which

$$R_{n,\delta^*}(h'_n) < \sum_{s_{n+1} \in S} \rho_n(h'_n, d'_n, R_{n+1}(h'_n, d'_n, s_{n+1})) \cdot p_n(h'_n, d'_n, s_{n+1}). \quad (7.7)$$

Construct the plan $\delta = (\delta_1, \delta_2, \dots)$ as follows:

$$\begin{aligned} \delta_m(h_m) &= d'_n, & m &= n, \quad h_m = h'_n, \\ &= \delta_m^*(h_m), & &\text{otherwise.} \end{aligned} \quad (7.8)$$

Obviously, $R_{m,\delta}(h'_n) > R_{n,\delta^*}(h'_n)$, and thus we have a contradiction to the optimality of δ^* at (h'_n, n) . Thus

$$R_{n,\delta^*}(h_n) = G_n(h_n), \quad \forall n \in N, \quad h_n \in H_n, \quad (7.9)$$

and from (7.4) it follows that

$$R_n(h_n) = G_n(h_n), \quad \forall n \in \mathbb{N}, \quad h_n \in H_n. \quad \blacksquare \quad (7.10)$$

It should be noted that in the proof we explicitly make use of the monotonicity property of ρ_n .

8. THE DYNAMIC PROGRAMMING ALGORITHM

For regular monotone models the optimality equations establish an intimate relationship between $R_n(h_n)$, ρ_n , and $R_{n+1}(h_{n+1})$. However, in order to use these equations for the determination of $R_n(h_n)$ and R , it is necessary to determine first the values $R_n(h_n)$, $\forall h_n \in H_n$ for some $N \in \mathbb{N}$. Obviously, if the model is truncated at some $N' \in \mathbb{N}$ the recursive procedure can be started at $N = N'$. Moreover, as indicated by Denardo [2, p. 30], it is often possible to start the recursive procedure even if the model is not naturally truncated.

Before we formally introduce the dynamic programming algorithm, let us observe that if W is a regular model, Δ its set of feasible plans, and n any arbitrary element of \mathbb{N} , then there exists an element $\delta \in \Delta$ with the property that

$$R_{n,\delta}(h_n) = R_n(h_n), \quad \forall h_n \in H_n. \quad (8.1)$$

More specifically, for any $n \in \mathbb{N}$, $h_n \in H_n$, let δ^{h_n} be any element of Δ for which $R_{n,\delta^{h_n}}(h_n) = R_n(h_n)$. Then any plan $\delta = (\delta_1, \delta_2, \dots)$ having the property

$$\begin{aligned} \delta_m(h_m) &= \delta_m^{h_n}(h_m), & m = n, & \quad h_m = h_n \in H_m, \\ &= \delta_m^{h_n}(h_m), & m > n, & \quad h_m = (h_n, \dots, s_m) \in H_m, \end{aligned} \quad (8.2)$$

satisfies Eq. (8.1).

We will use the existence of such a plan to show that the dynamic programming algorithm is well defined.

DEFINITION 8.1 (The Dynamic Programming Algorithm). Let W be a regular model decomposed by $\rho = (\rho_1, \rho_2, \dots)$. Consider the following procedure for determining the sequences F_n , Δ^n , $n \leq N$, and the value F .

Step 1. Determine some $N \in \mathbb{N}$.

Step 2. For $n = N$ determine the following:

$$(1) \quad \Delta^N := \{\delta: \delta \in \Delta, R_{n,\delta}(h_N) = R_N(h_N), \forall h_N \in H_N\}, \quad (8.3)$$

$$(2) \quad F_N(h_N) := R_N(h_N), \quad h_N \in H_N. \quad (8.4)$$

Step 3. For $n < N$ determine the following:

$$(1) \quad F_n(h_n) := \sup_{d_n \in D_n(h_n)} F_n(h_n, d_n), \quad (8.5)$$

where

$$F_n(h_n, d_n) := \sum_{s_{n+1} \in S} \rho_n(h_n, d_n, F_{n+1}(h_n, d_n, s_{n+1})) \cdot p_n(h_n, d_n, s_{n+1}),$$

$$h_n \in H_n, \quad d_n \in D_n(h_n), \quad (8.6)$$

and

$$(2) \quad \Delta^n = \{\delta: \delta \in \Delta^{n-1}, F_n(h_n, \delta_n(h_n)) = F_n(h_n), \forall h_n \in H_n\}. \quad (8.7)$$

Step 4. Determine F and Δ^0 as follows.

$$F := \sum_{h_1 \in H_1} F_1(h_1) p_0(h_1) \quad (8.8)$$

and

$$\Delta^0 := \Delta^1.$$

The recursive procedure defined by these four steps will be referred to as the *dynamic programming algorithm* and the elements $\delta^0 \in \Delta^0$ as the *dynamic programming plans* or *solutions*. ■

COROLLARY 8.1. *Let W be any regular monotone model and N any arbitrary element of N . Then,*

$$(1) \quad F_n(h_n) = R_n(h_n), \quad \forall n \leq N, \quad h_n \in H_n, \quad (8.9)$$

$$(2) \quad F = R, \quad (8.10)$$

and

$$(3) \quad \Delta^0 \subset \Delta^*. \quad (8.11)$$

Proof. The proof is directly from Theorem 7.1. That is, since the model is regular and monotone, the optimality equations hold and there exists an optimal plan which is also optimal with respect to all the modified problems. By construction, the dynamic programming algorithm recovers any such plan. ■

It should be noted that in Corollary 8.1 it is assumed that the model is *monotone*.

In the next section we will prove the optimality of the dynamic programming plans under less restrictive conditions regarding the model.

The dynamic programming algorithm as defined in this section is used for theoretical purposes. From the computation viewpoint, sufficient statistics [4, p. 36] can be used to reduce the domain of definition of F_n and thus reduce the number of iterations involved in the implementation of the algorithm.

9. SOME RESULTS

In the previous sections results obtained by Hinderer [3, Sect. 3] have been generalized to nonadditive models.

We will show now that the monotonicity assumption in Corollary 8.1 can be replaced by a less restrictive condition.

THEOREM 9.1. *Let W be any regular model decomposed by some $\rho = (\rho_1, \rho_2, \dots)$ for which there exists a plan $\delta^* \in \Delta$ with the following properties:*

$$(1) \quad \delta^* \in \Delta^* \quad (9.1)$$

and

$$(2) \quad R_{n, \delta^*}(h_n) = R_n(h_n), \quad \forall n \in \mathbb{N}, \quad h_n \in H_n(\delta^*). \quad (9.2)$$

Then, for any $N \in \mathbb{N}$

$$(1) \quad F_n(h_n) = R_n(h_n), \quad \forall n \leq N, \quad h_n \in H_n(\delta^*), \quad (9.3)$$

$$(2) \quad F = R, \quad (9.4)$$

and

$$(3) \quad \Delta^0 \subset \Delta^*. \quad (9.5)$$

Proof. (1) We will prove the first part of the theorem by induction on n , showing that for each $n \leq N$,

(1) there exists $\delta' \in \Delta^n$ such that

$$\delta'_m(h_m) = \delta_m^*(h_m), \quad \forall m \in \mathbb{N}, \quad h_m \in H_m(\delta^*), \quad (9.6)$$

and

$$(2) \quad F_n(h_n) = R_n(h_n), \quad \forall h_n \in H_n(\delta^*). \quad (9.7)$$

The validity of the inductive hypothesis at $n = N$ is easily verified by Eq. (9.2), and the definition of the second step of the algorithm.

Assume that the inductive hypothesis is true for $n = N - 1, N - 2, \dots, m$. In particular for $n = m$ assume that

(1) there exists $\delta' \in \Delta^m$ such that

$$\delta'_i(h_i) = \delta_i^*(h_i), \quad \forall i \in \mathbb{N}, \quad h_i \in H_i(\delta^*), \quad (9.8)$$

and

$$(2) \quad F_m(h_m) = R_m(h_m), \quad \forall h_m \in H_m(\delta^*). \quad (9.9)$$

Consider $n = m - 1$ and any arbitrary element $h_{m-1} \in H_{m-1}(\delta^*)$. From the inductive hypothesis at $n = m$ and the fact that $\delta_{m-1}^*(h_{m-1}) \in D_{m-1}(h_{m-1})$, it follows that $F_{m-1}(h_{m-1}) \geq R_{m-1, \delta^*}(h_{m-1})$. However, since by construction

$$F_{m-1}(h_{m-1}) = R_{m-1, \delta}(h_{m-1}), \quad \forall \delta \in \Delta^{m-1}, \quad (9.10)$$

and $\Delta^{m-1} \subset \Delta$, if $F_{m-1}(h_{m-1})$ is strictly greater than $R_{m-1, \delta^*}(h_{m-1})$, we have a contradiction to the optimality of δ^* at $h_{m-1} \in H_{m-1}(\delta^*)$. Thus, $F_{m-1}(h_{m-1}) = R_{m-1, \delta^*}(h_{m-1}) = R_{m-1}(h_{m-1})$, $\forall h_{m-1} \in H_{m-1}(\delta^*)$, which implies that there exists a plan $\delta' \in \Delta^{m-1}$ with the property that $\delta'_{m-1}(h_{m-1}) = \delta_{m-1}^*(h_{m-1})$, $\forall h_{m-1} \in H_{m-1}(\delta^*)$.

The inductive hypothesis is true then for $n = m$, and hence it is true for all $n \leq N$.

(2) To show that $F = R$ notice that from the first part of the theorem

$$F_1(h_1) = R_1(h_1), \quad \forall h_1 \in H_1(\delta^*). \quad (9.11)$$

However, since $H_1(\delta^*) = H_1$, it follows that $F = R$.

(3) By construction for all $n \leq N$, and $h_n \in H_n(\delta^*)$,

$$R_{n, \delta'}(h_n) = R_{n, \delta}(h_n) = R_n(h_n), \quad \forall \delta, \delta' \in \Delta^n. \quad (9.12)$$

In particular for $n = 1$,

$$R_{1, \delta^1}(h_1) = R_{n, \delta^*}(h_1), \quad \forall h_1 \in H_1, \quad \delta^1 \in \Delta^1. \quad (9.13)$$

Since δ^* is optimal and by definition $\Delta^0 = \Delta^1$, it follows that

$$R_{\delta^0} = R_{\delta^*} = R, \quad \forall \delta^0 \in \Delta^0 \quad (9.14)$$

and hence $\Delta^0 \subset \Delta^*$. ■

It is extremely important to note that the model considered in Theorem 9.1 is *not assumed to be monotone*.

The following are derived directly from Theorem 9.1.

COROLLARY 9.2. *Let W be any regular model decomposed by some $\rho = (\rho_1, \rho_2, \dots)$ for which the weak principle of optimality holds. Then for any $N \in \mathbb{N}$*

$$(1) \quad F_n(h_n) = R_n(h_n), \quad \forall n \leq N, \quad h_n \in H_n^* := \{h_n^* : h_n^* \in H_n(\delta^*), \delta^* \in \Delta^*\},$$

$$(2) \quad R_{n, \delta^0}(h_n) = R_n(h_n), \quad \forall n \leq N, \quad h_n \in H_n(\delta^0), \quad \delta^0 \in \Delta^0,$$

$$(3) \quad F = R,$$

and

$$(4) \quad \Delta^0 \subset \Delta^*.$$

Proof. Notice that since the model is regular there exists at least one optimal plan and since the weak principle of optimality holds, Theorem 9.1 can be applied for each $\delta^* \in \Delta^*$. ■

COROLLARY 9.3. *Let W be any regular model decomposed by some $\rho = (\rho_1, \rho_2, \dots)$ for which the dynamic programming principle holds. Then, for any $N \in \mathbb{N}$*

- (1) $F_n(h_n) = R_n(h_n), n \leq N, h_n \in H_n,$
- (2) $R_{n, \delta^0}(h_n) = R_n(h_n), \forall n \leq N, h_n \in H_n, \delta^0 \in \Delta^0,$
- (3) $F = R,$

and

- (4) $\Delta^0 \subset \Delta^*.$

Proof. Since the dynamic programming principle holds, it follows that there exists an optimal plan which is also optimal with respect to *all* the modified problems. Then, we can extend Theorem 9.1 to all $n \leq N, h_n \in H_n$. ■

Remarks. (1) Notice that the validity of any of the principles guarantees the optimality of the dynamic programming plans.

(2) It should be noted that the validity of the principles is *sufficient* for the optimality of the dynamic programming solution. ■

In the next section the role of the strong principle of optimality is discussed.

10. THE ROLE OF THE STRONG PRINCIPLE OF OPTIMALITY

If the objective of a dynamic programming investigation is to determine R and Δ^* , then one can set $r_n = r_1, \forall n > 1$. This implies the existence of the trivial decomposition defined by

$$\rho_n(h_n, d_n, R_{n+1}(h_n, d_n, s_n)) = R_{n+1}(h_n, d_n, s_n). \quad (10.1)$$

Obviously the model under consideration is strictly monotone, and thus one can consider the dynamic programming algorithm as a potential solution procedure.

However, in many situations there exist more efficient decompositions as far as computation is considered.

Thus, since the strong principle holds for *all* regular models, and since one can always start with the trivial decomposition, it follows that in *principle* the dynamic programming algorithm is a *potential* solution procedure for all regular problems.

The difficulties associated with the implementation of the algorithm are related to its first step, possible decompositions and the existence of efficient sufficient statistics.

One can consider then the strong principle of optimality as more of a "philosophical" statement providing the motivation and the approach for the use of the dynamic programming algorithm.

It also should be noted that the strong principle of optimality has certain similarities to Bellman's principle. More specifically, both are concerned with a single criterion function, both disregard the monotonicity of the model, and both are related to modified problems generated (with positive probabilities) by optimal plans.

As was emphasized in the previous sections, the author in no way "suggests" that the strong principle is a precise statement of Bellman's principle. On the contrary, the author finds Bellman's principle to be too powerful to be exactly defined. Any precise definition of the principle involves the construction of a specific decision model which in turn restricts the generality of the principle.

11. CONCLUSIONS

Each of the principles discussed in this paper plays an important role in many dynamic programming investigations. To the best of the author's knowledge, Theorem 5.1 provides the most general conditions for the validity of the weak principle of optimality.

The results presented in Section 9 indicate that the validity of each of the principles is sufficient for the dynamic programming solution to be optimal and thus resolve certain questions raised in the literature as to the relationship between the principle and the dynamic programming algorithm.

Moreover, Theorem 9.1 indicates that the existence of a certain optimal plan is sufficient for the optimality of the dynamic programming solutions.

As no monotonicity assumption was made in Section 9, it seems as if the *validity* of the principles and not the monotonicity properties of the reward function is the basic characteristic of dynamic programming problems.

The question as to the exact interpretation of Bellman's principle in the context of the sequential decision model considered here has not been resolved. It seems as if only Professor R. Bellman can provide a definite answer to this controversial question.

ACKNOWLEDGMENT

The author expresses his gratitude to Professor Sidney J. Yakowitz, Department of Systems and Industrial Engineering, University of Arizona, Tucson, Arizona, for his stimulating and constructive guidance.

REFERENCES

1. R. BELLMAN, "Dynamic Programming," Princeton Univ. Press, Princeton, N. J., 1957.
2. E. V. DENARDO, "Sequential Decision Processes," Ph.D. Dissertation, Northwestern University, 1965.
3. R. ARIS, "Discrete Dynamic Programming," Blaisdell, New York, 1964.
4. K. HINDERER, "Foundation of Non-stationary Dynamic Programming with Discrete Time Parameter," Springer-Verlag, New York, 1970.
5. W. S. BUTCHER, Y. Y. HAIMES, AND W. A. HALL, Dynamic programming for the optimal sequencing of water supply projects, *Water Resour. Res.* 5 (1969), 1196-1204.
6. T. L. MORIN, Pathology of a dynamic programming, sequencing algorithm, *Water Resour. Res.* 9 (1973), 1178-1185.
7. T. L. MORIN AND A. M. O. ESOGBUE, Some efficient dynamic programming algorithms for the optimal sequencing and scheduling of water supply projects, *Water Resour. Res.* 7 (1971), 479-484.
8. S. ARUNKUMAR, Characterization of optimal operating policies for finite dams, *J. Math. Anal. Appl.* 49 (1975), 267-274.
9. S. J. YAKOWITZ, "Mathematics of Adaptive Control Processes," American Elsevier, New York, 1969.
10. L. G. MITTEN, Preference order dynamic programming, *Management Sci.* 21 (1974), 43-46.
11. J. M. SOBEL, Ordinal dynamic programming, *Management Sci.* 21 (1975), 967-975.
12. M. LOEVE, "Probability Theory," 2nd ed. Van Nostrand, Princeton, N. J., 1960.
13. L. G. MITTEN, Composition principles for synthesis of optimal multistage processes, *Operations Res.* 12 (1964), 610-619.