DECISION PROCESSES AND FUNCTIONAL EQUATIONS

S. Karlin
Harold N. Shapiro

RM-933

3 September 1952

This is a working paper. It may be expanded, modified, or withdrawn at any time. The views, conclusions, and recommendations expressed herein do not necessarily reflect the official views or policies of the United States Air Force.
SUMMARY

This considers in an abstract manner the general question of existence and uniqueness of solutions of problems in dynamic programming.

DISTRIBUTED PROCESSES AND FUNCTIONAL EQUATIONS

S. Karlin and Harold M. Shapiro

§1. Introduction. Recent investigations connected with bombing models, dynamic programming, and inventory policy, [1], [2], [3], have all led to the consideration of certain functional equations.

In each of these problems there is a notion of strategy, or policy, S and a function $f(S)$ which measures in some sense the value of this strategy. The general question is then that of determining the existence of an $S$ which maximizes $f(S)$. In many cases this question may be translated into the consideration of the existence and uniqueness of the solution to a certain functional equation. In the aforementioned investigations the emphasis has always been placed on the functional equation, and the existence theorems are derived via specially fitted iteration schemes.

This paper assumes a more global point of view and focuses on the structure of the space of strategies. By dealing directly with the underlying model the questions of existence and uniqueness of solutions to the associated functional equations may be settled easily.

We shall present a general formulation of this procedure, and show that it covers the results of the investigations mentioned above.
In passing, it is well to emphasize that none of the work done thus far (including this paper) sheds any light on the important problem of determining the strategies $\mathcal{S}$ which maximize the relevant $\mathcal{E}(\mathcal{S})$.

\section{Decision processes and associated functional equations}

We consider as our underlying structure an abstract system (quite similar to that discussed in [2]) consisting of:

I. A "phase space" $\mathcal{X}$ whose points $\omega$ correspond to the admissible states of the system.

II. A decision space $\mathcal{D}$ whose points $\delta$ correspond to admissible decisions.

III. A payoff function space $\mathcal{L}$ of real valued positive functions $L(\omega, \delta)$ which are functions of the state $\omega \in \mathcal{X}$ and the decision $\delta \in \mathcal{D}$. These will be used to measure the "gain" derived from the decision $\delta$ when presented with the state $\omega$.

(In many problems $\mathcal{L}$ consists of a single function. However, in general, if we wish to allow for changes as might be due to the passage of time alone, $\mathcal{L}$ can consist of many functions.)

IV. A collection $\mathcal{G}$ of "transition operators" $\mathcal{G}$ indexed by $\delta \in \mathcal{D}$ which are mappings of $\mathcal{X}$ into itself.

V. A collection $\mathcal{F}$ of weight functions $F(t)$ defined for $0 \leq t < \infty$ which are such that $F(t)$ is 0 except for a discrete set of points. These $F(t)$ correspond to the admissible distributions of our decisions in time.
Any structure which satisfies I-V we shall call a decision process. For the purposes of this paper we shall consider only decision processes which satisfy the following additional conditions:

(a) \( \mathcal{L} \) is a topological space.
(b) \( \mathcal{D} \) and \( \mathcal{I} \) are compact topological spaces.
(c) \( L(\omega, \xi) \in \mathcal{L} \) is a continuous function over the cartesian product \( \mathcal{L} \times \mathcal{D} \) to the non-negative reals.
(d) For each fixed \( \xi \in \mathcal{L} \), \( \tau_{\xi} \) is a continuous mapping of \( \mathcal{D} \) into \( \mathcal{L} \).

Setting \( \mathcal{D}^0 = \mathcal{D} \), we consider the cartesian product

\[
\mathcal{L}^0 = \left\{ \prod_{t=0}^{\infty} \mathcal{D}_t \right\} \times \mathcal{I}
\]

which we call the strategy space. The points \( \sigma \) of \( \mathcal{L}^0 \) are naturally enough called strategies. It should be noted that a strategy \( \sigma \) is simply a couple \( [\xi(t), F(t)] \) where \( \xi(t) \) is a function from the non-negative reals into \( \mathcal{D} \), and \( F(t) \in \mathcal{I} \). Since \( \mathcal{D} \) and \( \mathcal{I} \) are compact it follows via a theorem of Tychonoff that the strategy space \( \mathcal{L}^0 \) is also compact.

For the initial state \( \omega(0) = \omega \), we define the payoff of the strategy \( \sigma = [\xi(t), F(t)] \) to be an expression of the form

\[
\Phi(\omega, \sigma) = \sum_{t=0}^{\infty} c_t L_t \left( \xi(t), \omega(t) \right)
\]

*J. C. Kendel has independently remarked on the relevance of Tychonoff's Theorem to these matters. Only extreme laziness has prevented him from writing a note completely isomorphic to this one.*
where the \( L_k \) are given in \( \mathcal{X} \), the \( t_1 \) are points where \( F(t) \neq 0 \), \( a_1 = F(t_1) \) and the transition operators are chosen so that

\[
\omega(t) = \tau_{\delta_0(t_1-1)} \omega^*.
\]

We next assume that the right side of (2.1) converges uniformly in \( S \). Then from assumptions (c) and (d) above and the fact that the \( dF(t) \) ascribe weights to only a discrete set of points, it follows that \( \tilde{F}(\omega, S) \) is a continuous function of \( S \), for each \( \omega \in \mathcal{X} \). Since a continuous function over a compact space achieves its maximum we have the existence of

\[
(2.2) \quad \mathcal{M}(\omega) = \max_{S \in \mathcal{X}} \tilde{F}(\omega, S) = \tilde{F}(\omega, S^*).
\]

§3 In this section we proceed to develop a natural functional equation for the case where the decisions at each step take place at the discrete time set \( t = 0, 1, 2, 3, \ldots \) with \( L_n(\delta, \omega) = L(\delta, \omega) \) for all \( n \) and \( a_n = a^n \) for some real number \( 0 < a < 1 \). It is only to be remarked that more general situations can be handled at the expense of elegance, let \( \delta^* \) denote an optimal strategy, then

\[
\mathcal{M}(\omega) = \max_{S \in \mathcal{X}} \tilde{F}(\omega, S) = \max_{S \in \mathcal{X}} \sum_{i=0}^{\infty} a^{i-1} F(\omega(i), S(i))
\]

which reads as

\[
(2.3) \quad \mathcal{M}(\omega) = \max_{\delta} \{ L(\omega, \delta) \cdot a \mathcal{M}(\tau_{\delta_0}(\omega)) \}.
\]
We remark that if (2.3) is presented out of context, then the function described in (2.2) is a solution to (2.3). Since the series converges uniformly in (2.1), we deduce easily that

\[ \lim_{n \to \infty} \max_{\mathcal{F}(n)} c_n \mathcal{H}(T_{\mathcal{F}(n)} \omega) = 0 \]

where \( \mathcal{F}(n) \) denotes successive application to \( \omega \) of the strategy \( \gamma \) \( n \) times. Conversely, if we start with a bounded solution \( M(\omega) \) of (2.3), we have upon expanding

\[
M(\omega) = \max_{\mathcal{F}(n+1)} \left[ L(\omega, \mathcal{S}(0)) + \alpha L(T_{\mathcal{S}(0)} \omega, \mathcal{S}(1)) + \cdots + \alpha^{n+1} L(T_{\mathcal{S}(n)} \omega, \mathcal{S}(n+1)) + \alpha^{n+1} M(T_{\mathcal{S}(n+1)} \omega) \right]
\]

Since all terms are positive, it is evident that

\[ M(\omega) \geq \gamma(\omega) \]

On the other hand

\[ M(\omega) \leq \gamma(\omega) + \lim_{n \to \infty} \max_{\mathcal{F}(n)} M(T_{\mathcal{S}(r)} \omega) \alpha^{n-1} \]

Thus, if we assume that
\[(2.6) \quad \max M(T_\delta(n)\omega)\alpha^{n+1} \to 0\]

we find that
\[M(\omega) = \mathcal{M}(\omega).\]

In all of the applications (2.6) as well as our other assumptions hold trivially, and in these cases the existence and uniqueness of the bounded solution to (2.3) is thereby established.

Finally, we observe that if we require conversely that $\overline{f}(\omega, S)$ be a continuous function of $S$ for each fixed $\omega$, then the convergence of the right side of (2.1) is uniform. This follows by Dini's Theorem, using the fact that $L(\omega, \delta) \geq 0$ and that $\mathcal{D}$ is compact. Furthermore, the continuity of $\overline{f}(\omega, S)$ as a function of $(\omega, S)$ is equivalent here to the uniform convergence of the right side of (2.1) in $\omega$ and $S$ jointly. In this case it follows easily that $\mathcal{M}(\omega)$ is a continuous function of $\omega$.

\section{Applications}

This section is devoted to the consideration of various models from "bombing model" and "dynamic programming" Theory, which are encompassed by the general discussion of the preceding section.

\textbf{Example I.}

In this example the decision process is constituted as follows:
\[ \mathcal{L} = \text{a bounded portion of Euclidean n space.} \]

\[ \mathcal{G} = \text{a set of m objects } \delta_1, \delta_2, \ldots, \delta_m \text{ with the discrete topology assigned.} \]

The functions of \( \mathcal{F} \) are defined so that

\[ J_{\alpha}(\omega, \delta) = I(\omega, \delta) \]

with \( \alpha \) is a fixed real number \( 0 < \alpha < 1 \).

\( \mathcal{F} \) consists of the single function which assigns weight 1 to all the decisions.

The space of transition operators \( \mathcal{G} \) need not be specified explicitly other than that they have the properties required by the discussion of section (1).

In this example the relevant functional equation has the form

\[ (2.1) \quad \mathcal{H}(\omega) = \max_{1 \leq k \leq n} \left[ I(\omega, \delta_k) + \alpha \mathcal{H}(T_{\delta_k} \omega) \right] \]

Existence and uniqueness of bounded solutions of \((2.1)\) are an immediate consequence of our earlier discussion.

**Example II.**

\( \mathcal{L} = \text{the interval of the real line } 0 \leq \omega \leq x_0 \)

\( \mathcal{J} = \mathcal{L} \)
Let $h(x)$ and $g(x)$ denote monotonically increasing continuous functions on $-\infty < x \leq x_0$ with $h(x) = g(x) = 0$ for $x \leq 0$. Also, assume for some $c$, $0 < c < 1$, we have a continuous function $b(x)$, $b(x) \leq cx$; and

\begin{equation}
\sum_{n} g(c^n x_0) < \infty, \quad \sum_{n} h(c^n x_0) < \infty.
\end{equation}

The family of operators $T_\delta$ is defined as follows: For $\delta \in \mathcal{S}$, (since $\mathcal{S}$ is a real number) we define

$$T_\delta \omega = b(\delta) \omega$$

As for the functions of $\mathcal{L}$, we have

$$L_t(\delta, \omega) = I_0(\delta, \omega) = g(\delta) + h(\omega - \delta).$$

The set $\mathcal{J}$ is as in Example 1.

The functional equation in this case has the form

\begin{equation}
\mathcal{M}(\omega) = \max_{0 \leq \delta \leq \omega} \left\{ g(\delta) + h(\omega - \delta) \cdot M(b(\delta) \omega) \right\}.
\end{equation}

The value of $\delta$ is taken to be 1 here. The conditions (2.2) together with $0 < c < 1$ insure the necessary uniform convergence, and the existence and uniqueness of a solution to (2.3) follow immediately from the discussion of $\mathcal{S}_1$. 
Example III. (The optimal inventory policy)

This section is devoted to a non-deterministic case.

Let $x, 0 \leq x < \infty$, represent the commodities available at the initial period and suppose one orders up to $y$. The demand $z$ is a random variable with distribution $F(z)$ which is independent of the time interval. Let $V_i(x, y)$ represent the total expected loss during the $i$-th time interval with $x$ available initially and an ordering given up to $y$. The space $\mathcal{L}$ coincides with the half infinite line $0 \leq x < \infty$. We now impose a restriction on the choice of ordering $y$ in that we require $0 \leq y \leq M$ where $M$ is a fixed bound. This assumption is made in all the papers treated to date \[2\], \[3\]. It can easily be justified as realistic. We also suppose that $V_i(x, y) = \infty$ for $y < x$ and is uniformly bounded for $y \geq x$. Define $L_0(x, y) = V(x, y)$.

The ordering policy $0 \leq y_i \leq M$ is actually a random variable depending on the amount of goods $x$ available at the $i$-th state. (The quantity of goods available is also considered as a random variable, and it is these random variables which constitute the space $\mathcal{L}$.)

The space of all bounded random functions defined on the sample space describing the goods available at a given state constitutes a compact set $C$ in the weak * topology over the Banach space of all integrable functions defined on the sample space. This is the well known theorem of Alaoglu. The transformation $T_y x$ is a mapping of the random variable state at the commencement of the $i$-th period into the state of the remainder of the goods after ordering $y_i$ and the $i$-th
period has been completed. We define for \( n \geq 1 \) \( L_n(x, y) = \mathcal{E}(x, y) = \) expected loss with respect to \( x \) when the initial state is \( x \) and the ordering policy is \( y \). The value of \( \alpha \) is \( 0 < \alpha < 1 \). The total loss for the initial state \( x \) following an optimal policy is

\[
\mathcal{L}(x) = \inf_{\delta} \left[ L_0(x, y_0) + \alpha \mathcal{E}(T_{y_0} x, y_1) + \alpha^2 \mathcal{E}(T_{y_1} x, y_2) + \ldots \right]
\]

It is now assumed that \( \mathcal{E}(x, y) \) is a continuous function on \( C \). The convergence is uniform and the general discussion implies the existence and uniqueness of the solution to the functional equation (2.3) which can be written in this case as

\[
\mathcal{L}(x) = \inf_{y \geq x} \left\{ V(x, y) + \alpha \mathcal{L}(C) \left[ 1 - F(y) \right] + \alpha \int_{-\infty}^{y} \mathcal{L}(y-t) dF(t) \right\}.
\]

**Bibliography**


(3) " " Arrow, Harris, Mochel. Cowles Commission Papers, New Series, No. 44.

hjt