SOME REMARKS ON RESEARCH FOR STRUCTURAL DESIGN OF ASPHALT CONCRETE PAVEMENT SYSTEMS

Karl S. Pister

That construction and operation of successful pavements predates the dream of a "rational" method of design can hardly have escaped the attention of engineers or laymen. The present state of the art is the product of a long history of successes and failures, the former fortunately overshadowing the latter. In fact one may well ask the question: Why does the engineer want a "theory" of pavement design, inasmuch as theories invariably are wrong, have limited applicability, or are too complicated when put to the test of real experience? The answer seems to lie in the observation that, in the present milieu of rapid change, experience quickly becomes obsolete or is often totally lacking. Examples of this are rampant in the pavement field; e.g., witness the increase in traffic volume, the change in construction costs and methods, and the potential problems arising from disappearance of high-quality raw materials with concomitant increased use of new and sometimes marginal substitutes (would you believe crushed glass bottles?). The Via Appia was a first-class Roman road; but, as any tourist can tell you, the service life has long been exceeded. Thus, it appears that this workshop was based on the implicit assumptions that a rational method of pavement design exists, is important to acquire, and is accessible to the minds of engineers. In reviewing the papers prepared for the workshop, I have drawn the conclusion that these assumptions are shared by the speakers, and I expect they are held plausible by most of the participants. However, as we are often painfully aware, sharing a common set of assumptions does not imply any uniqueness for subsequent application and action. This I believe is what needs very careful examination during the workshop sessions. We must cast aside our denominational prejudices and try to examine what indeed we are trying to accomplish from our common point of departure. We will then be in a much more favorable position to discuss the organization of research and development work directed toward sucessful design and management of pavement systems. In this regard the following quotation from Bertrand Russell's "Unpopular Essays" is quite relevant: "So whenever you find yourself getting angry about a difference of opinion, be on your guard; you will probably find, on examination, that your belief is getting beyond what the evidence warrants."

Let me now return to the "implicit assumptions," which in a sense form the basis for my subsequent remarks. It seems to me that the existence of a rational method of design has to be established a posteriori; i.e., it is the task of the engineer to observe, acquire, and organize information and experience obtained from operational physical systems. This point of view, incidentally, is strongly supported in Finn's paper (9). However, these steps are in themselves insufficient, for we must perform the extremely difficult job of "pattern recognition" to visualize the structure of a model that, post facto, seems to "fit" our observations. Such a model, given an adequate mathematical structure, can then be employed to carry out simulation tasks to seek the "best" among alternatives—this being the pragmatic task to which we customarily attribute economic importance, whether in terms of dollars or expenditure of other resources. In turn, this step leads to the development of the final assumption, accessibility of the rational method to the engineer. The kinds of models and methods to be discussed here are from a practical viewpoint accessible only through a digital computer, arising from experience with operating pavement systems.

Finally, before we turn to a more systematic examination of the ideas sketched earlier, it is well to note that, even though we accept the existence of a rational design formula, we do not pretend to believe that, like the commandments of Moses, the formula is unchangeable with time. Man (especialy engineers) by nature is a creature who likes to modify his systems, adapting them to suit his needs. Pavement design and management are problems of adaptive control, a concept that is found to be extremely useful in development of models of system design and behavior. We shall, in fact, try to indicate how insight into design and management of pavement systems may be enhanced by viewing the task as a multistage decision process, utilizing Bellman's dynamic programming techniques as a vehicle. Let us turn now to certain preliminaries and terminology needed to describe the problem under consideration.

A FLEXIBLE PAVEMENT SYSTEM

As a point of departure we adopt the terminology used in HRB Special Report 113 (1).

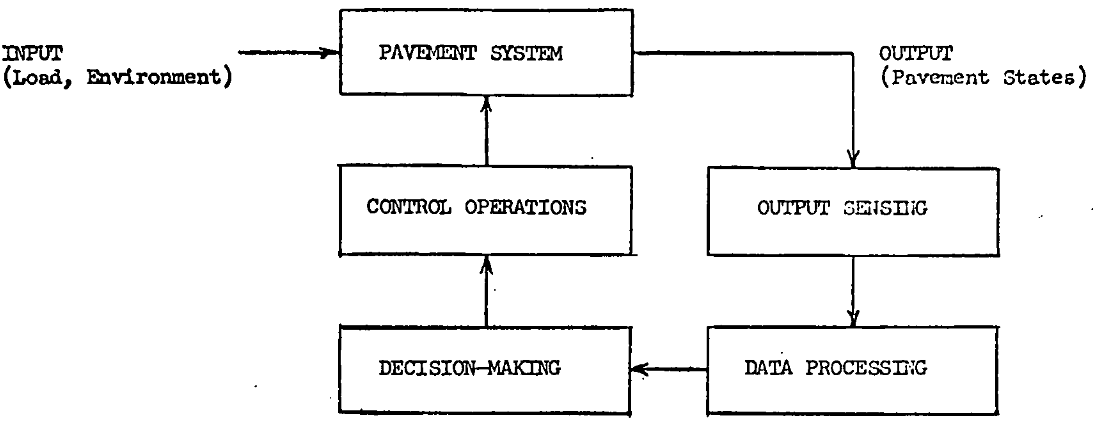
1. A flexible pavement is a pavement structure that maintains intimate contact with and distributes loads to the subgrade and depends on aggregate interlock, particle friction, and cohesion for stability.
2. Pavement structure is the combination of subbase, base course, and surface course placed on a subgrade to support the traffic load and distribute it to the roadbed. To this I would add the qualifying phrase, "under a history of environmental conditions."
3. Serviceability, which embodies the function of a pavement, is the ability of a pavement to serve traffic with safety and comfort and with a minimum of detrimental effects to either vehicle or pavement.
4. The present serviceability index, which is the current (present) measure of the effectiveness of the pavement, is a numerical index of the ability of a pavement in its present condition to serve traffic.
5. Performance is the measure of the accumulated service provided by a pavement, i.e., the adequacy with which a pavement fulfills its purpose. I would add here that performance implicitly includes "service per dollar" or some other type of economic measure.

These definitions taken together constitute what we may call a flexible pavement system: a set of interacting components subject to various inputs (traffic, environment), producing various outputs (as yet unspecified). System performance measures adequacy over the operational lifetime. At this stage there is little to be gained from a more formal definition; however, there is a great deal of conceptual mileage to be gained from this intuitive picture. For example, it makes clear that performance is the real goal of design and operation (through proper management) of the system. Yet, relatively little information concerning pavement behavior, in which performance is the dependent variable, can be found in the literature. This is of course understandable because performance is somewhat ambiguous to define, in spite of its conceptual importance.

Perhaps it is easier to deal with distress, which is really absence of serviceability, and to invent measures of distress, along with a normalizing requirement that "serviceability plus distress equals unity" during the pavement lifetime. No matter what definition prevails, the point here is that one must acquire sets of "systematic and continuous observations of performance (or distress) of full-scale pavements" (9). It is only through such a data acquisition program that any hope of pattern recognition will emerge to guide the formalization of operational rules leading to rational design. For example, without this, mathematical simulation of pavement systems, no matter how fascinating a game in itself, will remain precisely a game with very little payoff to pavement systems. Whether one adopts distress (a structural or mechanistic, designer-oriented concept) or performance (a user-oriented concept), as a practical matter it is expedient to attempt to divide traffic-associated performance from environment-associated performance wherever possible. In this connection frequency of occurrence studies of types of distresses serve to emphasize the behavioral aspects of the pavement system deserving the most detailed study. Such observations necessarily are global in nature; i.e., they constitute integrated or averaged values of pavement response variables, as opposed to local values of the variables. This is in fact an extremely important point that strongly influences the development of mathematical models, as will be seen.

Finally, it may be appropriate here to consider the advantages of dealing with the pavement system problem in two separate, yet highly related, stages: (a) the problem of observing and controlling (managing) an existing pavement system to achieve optimum performance, and (b) the problem of simulating a pavement system by mathematical modeling so that an optimum design configuration can be achieved. These problems cannot really be separated either in planning a research program such as is our task here or in implementing a policy for design and control. As noted previously, a separation of these problems leads to the probability of two separate games being played rather than one; thus caution must be observed. When these problems are examined, it is useful from both conceptual and operation viewpoints to use block diagrams to describe the system under consideration. Figure 1 shows the basic elements of a system whose output is analyzed and evaluated by performance criteria so that control operations (maintenance) can be effected to provide a certain level of serviceability. No attempt is made at this point to inquire in detail into the subsystem components constituting the system, nor to select quantitative measures to describe the system. Obviously, this is a crucial matter for the success of mathematical modeling, and it will be examined more fully later in this paper.

A similar diagram can be constructed for the second problem of mathematical simulation of a pavement system. The emphasis in this problem is on selection of the parameters of the system itself; i.e., for a given range of inputs and desired performance criteria, a policy leading to an optimum selection of model parameters is desired. This is the classical inverse problem of design, a problem whose complexity invariably requires that a certain family of model structures be examined, from which the "best" choice of parameters is selected. (An example is the selection of layer thicknesses and elastic moduli using elastic layered system theory as the mathematical model.) A diagram of the basic phases of mathematical simulation is shown in Figure 2.



INPUT

Traffic-associated
Environment-associated

OUTPUT SENSING

Visual inspection
Transducer measurements
Test vehicles

DECISION-MAKING

Select optimal policy

OUTPUT

Stress
Deformation Performance
Distress

DATA PROCESSING

Evaluate performance quantitatively

CONTROL OPERATIONS

Maintenance
Repair
Replacement

Figure 1. Pavement system control.

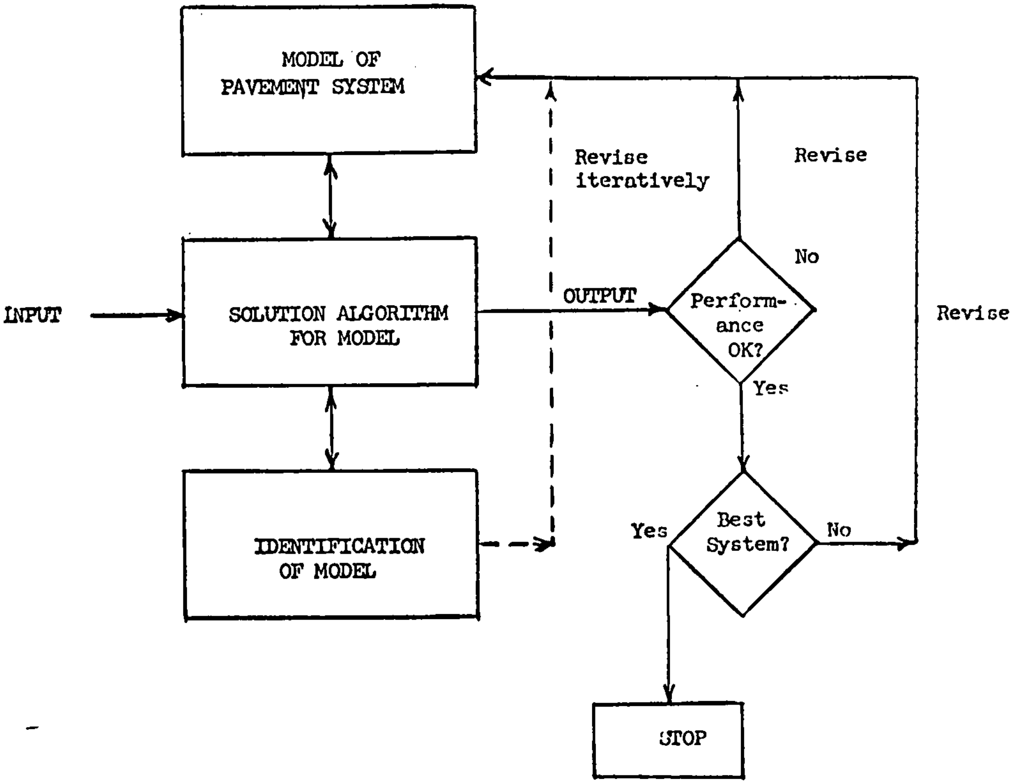


Figure 2. Simulation of pavement system.

A treatment of the design and management of pavement systems as a control problem can be found in the paper by Hudson (10), as well as in an earlier report by Hudson et al. (4). These papers call attention to the need to view the problem in the context just described. In addition, however, they suggest a structure by which quantitative results can be developed. Because of the extreme complexity of the system, the method is necessarily rather primitive (even if somewhat involved). Nevertheless, it should be considered an important step forward. What I wish to emphasize in referring to this work here is the need to develop a general systems model not only describing the broad problem of design and management but also looking critically into the "black box" so that each subsystem is understood and described in the most comprehensive way possible. This workshop provides the opportunity of bringing together people who look at various "black boxes," and it is incumbent on us to bind ideas together into a cohesive view of the real problem. Short of this we will return to playing our own games of solitaire.

The systems viewpoint described in this section is drawn from ideas presented in a report by Hudson et al. (2), which in part grew out of an earlier version of pavement systems analysis (3). Since that report, substantial progress in applications has been made, notably by Hudson and his colleagues (4) and in the very interesting works of Lemer and Moavenzadeh (5) and Moavenzadeh (11). What seems most appropriate to the writer at this time is to exploit to the fullest the structures and methodologies currently existing in the field of "systems." I use this term advisedly, because there is often a certain hesitation or snickering among engineers at mention of the word. Let me make clear my intentions. First of all, the notion of a system has been helpful to organize and place into proper interrelationship the myriad factors influencing behavior of a pavement structure. Furthermore, as more is learned about actual pavement performance through systematic field observation, systems engineering provides the fabric on which an interaction matrix can be constructed, i.e., the possibility of assigning weighting factors to system subcomponents so that more enlightened research and development can be undertaken in areas with highest payoff. This phase of the application of systems theory can be qualitative or semiquantitative (e.g., the insights gained by examining block diagrams, categorizing the type of distress, and the like) and still be of considerable value in guiding design practice and in orienting research and development.

The second aspect of systems, one to which I wish to devote attention later in the paper, is that commonly associated with system control processes and dynamic programming. The notion of pavement design as a feedback control process is given in the paper by Hudson (10), while Nair (12) points to the probable desirability of examining applications of dynamic programming. With this encouragement I believe it worthwhile to examine briefly the mathematical theory of control to see what light may be shed on our design problem. The formulation of a workable model will require the utmost support from each of the areas represented in this workshop. However, I hope to avoid the mistake of adding a new game to the plethora already at our disposal for entertainment of highway engineers and researchers. Let me first review in more detail the modeling of pavement system simulation (the second problem), which in turn leads logically to the first problem of management through control.

MATHEMATICAL MODELING OF THE DESIGN PROCESS FOR A PAVEMENT SYSTEM

The essential ingredients of design by mathematical simulation are as follows:

1. A description of the configuration and the input-output relations of the system along with a "parameterized" structure defining these quantitatively;
2. A statement of the operating conditions (input);
3. An algorithm for predicting the evolution of the system, i.e., its output or performance;
4. A criterion function by which performance can be judged; and
5. Modification of the system to seek an optimum performance.

The task of the designer is that of searching to select values of parameters of the system (within logical constraints), which in turn lead to optimum performance as judged by the criterion function. However appealing this view of design may be conceptually, it implicitly contains the seeds of its own destruction. Except for the most trivial cases it cannot actually be made operational at this time because of our lack of understanding of realistic input-output relations for the system (13, 14, 15), the difficulty in finding a prediction algorithm (12), and the problem of defining a criterion function (10, 11, 16). These are the subsystem black-box problems to which I previously referred. They deserve our careful attention. In the meantime we have to be content with the best of a bad situation, but we should be careful that our modeling is done by appropriate principles.

Historically the principles used to develop models of pavement behavior have come from continuum mechanics, particularly mechanics of solids, and there does not appear to be any serious challenger at this moment. Let us review briefly how reality appears to a solid mechanician. First, certain state variables must be introduced for the system, in our case, the stress matrix $\underset{\sim}{\sigma}$, the strain matrix $\underset{\sim}{\varepsilon}$, temperature T, and possibly moisture content M. Because stress and strain each require six components for their description, we have 14 local state variables, i.e., 14 scalar functions of time at each and every particle (point) in the pavement system. These variables must satisfy certain basic principles of mechanics, such as balance of momentum, conservation of mass, balance of energy, and the entropy production inequality. In addition, when a process is defined for a particular kind of material, a constitutive equation must be identified. The kinds of processes of interest to us here are primarily mechanical—e.g., deformations—although temperature and moisture content may also change. Constitutive equations tell us how these state variables are related during such processes. Different kinds of materials have different kinds of constitutive equations for the same process. The task of "material characterization" is that of determining the nature of the constitutive equations for materials of interest in pavement design. Unfortunately, the processes for which these equations must be found are not known a priori. Thus, the problem of characterization must be attacked iteratively; i.e., assumed equations are used to predict the output (process) of the system so that these processes can in turn be used for characterization experiments, from which data assumed versus actual behavior can be adjusted iteratively, by changing the constitutive model, to a desired degree of accuracy.

Two points need further amplification here: What guidelines are there for selecting constitutive models, and how can the system output be predicted? These questions are discussed in more detail by Westmann (13) and Nair (12), so I will include only a brief statement. Constitutive equations are expected to satisfy a fundamental principle of determinism; i.e., the past determines the future. In solid mechanics this means that the (local) stress state at the present time may depend on all the past states up to the present (history) of strain, temperature, and moisture content. Symbolically, this can be written

$$\underset{\sim}{\sigma}(\underset{\sim}{x}, t) = \underset{s=-\infty}{\overset{s=t}{\underset{\sim}{F}_1}} \left[\underset{\sim}{\varepsilon}(\underset{\sim}{x}, s), T(\underset{\sim}{x}, s), M(\underset{\sim}{x}, s); \underset{\sim}{x}, t\right] \tag{1}$$

Without going into many detailed points that may be raised (see any modern book on solid mechanics), we can say that we have here a statement that the stress matrix at a particle, $\underset{\sim}{x}$, and the current time, t, depends on all past strain states, temperatures, and moisture contents at the particle. The constitutive rule (functional), $\underset{\sim}{F}_1$, may depend on the particle, $\underset{\sim}{x}$, and on time, t; this is clearly so in a layered system and in cases where asphalt properties degrade with time. Generally, it is assumed that, while stress depends on M and T (moisture content and temperature), they can be determined separately from diffusion equations unaffected by input fluctuations of stress. The literature is replete with work reported to have completed the task of finding $\underset{\sim}{F}_1$; however, the facts do not support this contention. One cannot verify most of the work simply because the actual process (sequences of states) in a pavement system is unknown. What one can measure is only a set of selected output variables such as surface deflection under wheel loads, an output known to be notoriously insensitive to constitutive model parameters.

I do not wish to deprecate serious attempts to understand and model constitutive behavior. These are urgently needed to provide the comprehensive subsystem support to which I referred earlier. I wish only to caution against the overenthusiastic approach often used by those seeking support for games and to emphasize that it makes no sense to use additively a set of measurements taken in part with a micrometer and in part with a yardstick. Our resources will be better used if we try to measure the entire problem with the yardstick first and then try to determine the size of the components more accurately, to speak analogically.

Troubles do not vanish with the constitutive problem. We must next confront the task of developing an algorithm for predicting the evolution of the states of the system. In mechanics this is accomplished through the device of an initial-boundary value problem. The term initial suggests that the evolution of the pavement state variables will depend on a starting point in time, whereas boundary suggests that the geometrical confines of the system are acted on, in our case by traffic loads and environment. The solution of such a problem depends on mathematical analysis—primarily numerical analysis performed on a digital computer. The output of the computer corresponds to the output of the pavement system in the sense that states of the system are determined from the input as functions of time and location in the system. Unfortunately this information by itself is inadequate; one must append a criterion function by which output can be judged good or bad. This raises very serious questions that are addressed in the papers by Moavenzadeh (11) and McCullough (16): What is a suitable criterion function? How does one define "failure" locally? How does failure (or distress) propagate in space and time? When does an accumulation of local failures constitute global failure (distress) in the system? These questions deserve a great deal more attention than they have received, and it will only be through serious cooperative efforts of theoreticians and field engineers that any hope of solution will emerge. In symbolic terms one can pose the question: How can the local distress be calculated as a function of time? A possible form is

$$D(\underset{\sim}{x}, t) = \underset{s=-\infty}{\overset{s=t}{F_2}} [\underset{\sim}{\sigma}(\underset{\sim}{x}, s), \underset{\sim}{e}(\underset{\sim}{x}, s), T(\underset{\sim}{x}, s), M(\underset{\sim}{x}, s)] \tag{2}$$

The scalar-valued functional F_2 at each particle $\underset{\sim}{x}$ assigns a value, D, at time, t, to the set of stress, strain, temperature, and moisture content histories. That number is called here the distress. The structure of the functional F_2 is unclear; some structures appear in the paper by Moavenzadeh (11). The second important question relates to the the propagation of distress from particle to particle, leading to global and often observable damage in the pavement. This notion is embodied in the definition of a second functional defined now over all particles in the system at time t:

$$D_s(t) = F_3[D(\underset{\sim}{x}, t)] \quad \text{over all } \underset{\sim}{x} \epsilon S \tag{3}$$

This formalism calls attention to the idea that $D_S(t)$, the distress in ths system at time t, depends on the accumulation of histories of distress at all particles in the system; i.e., it is some kind of spatial influence function. $D(x, t)$ is related to the notion of (local) distress index defined by Hudson et al. (2), whereas Eq. 3 is a measure of present distress, or perhaps of a volume density of distress, in the system as a whole. In this sense it is complementary to the definition of present serviceability index referred to earlier, and one could write

$$D_s(t) + PSI(t) = 1 \tag{4}$$

if a suitable normalization is performed. System performance, PF, an integrated (over time) concept suggests the definition

$$PF(t) = \int_0^t PSI(s)\, ds = \int_0^t [1 - D_s(s)]\, ds \tag{5}$$

This view, even if it could be implemented through appropriate mathematical structures, still has the disadvantage that it is mechanistically oriented; i.e., the user is excluded from influencing the evaluation of performance. More sophisticated qualitative incorporation of the user is mentioned in other papers (4, 5). These aspects of performance need much more attention than they have received. At the moment it would appear that very limited progress has been made in quantifying the concepts described by Eqs. 2 and 3. One could mention several examples:

1. For linear elastic and viscoelastic layered system models, calculation of maximum surface deflections under simple wheel load patterns is an example of a "functional defined over all particles." If this deflection is limited by an inequality, a crude distress model corresponding to Eq. 3 is obtained. Surface curvature can be treated similarly.
2. For certain types of linear viscoelastic layered system models, permanent surface deflections can be calculated. These would represent a model similar to Eq. 3, in which history is incorporated.

When one moves beyond linear elastic and linear viscoelastic models, a substantial amount of empiricism is introduced. Although there is no evil in empiricism, it should not be listed under the rubric of mechanics. It is a useful procedure that one must employ to obtain a solution of a real system problem in the face of complexity. This leads to the next consideration, that of introducing an element of "external disturbance" into the design process. What has been done to date seems to fall into the pattern of using rational, yet inadequate, models of pavement behavior, observing that these simulations do not correspond to real system behavior and that no rational criteria for distress (or performance) exist and then making the best of a bad situation, namely, allowing the engineer to use his judgment to assign criteria required to achieve as near an optimum as possible for the system design. In other words, the engineer is a short-circuit of the rational design process. Our attempts should be directed toward using the engineer in this role but supplying him with the best possible data on which to base his judgments, thereby minimizing the possibility of irrational short-circuits. The engineer is a finite, fallible control system. In spite of improvements in modeling subsystems and in developing more sophisticated models of the pavement system, predicted performance will seldom match actual performance of a pavement system. In other words, the real system performance leaves something to be desired. This is precisely what constitutes the notion of control in a decision process. Because we do not like the manner in which the system is evolving, we intervene to change its performance. Such control may take the form of patching, seal coats, overlays, or, in an extreme case, complete replacement. We shall now try to sketch more abstractly the mathematical structure of this type of system and indicate a possible direction of research in this area.

PAVEMENT MANAGEMENT AS A MULTISTAGE DECISION PROCESS

We have examined the manner in which mathematical simulation can be carried out in the areas shown in Figure 2. In this section we turn our attention to the problem of observing and controlling an existing pavement system to attain optimum performance along the lines shown in Figure 1. Much of what has been said already pertains to this problem; however, it is necessary to adopt a more modest view of measure of performance in order to expect numerical results, a fact already noted in connection with work reported by Hudson et al. (4). The complexity of the system with which we are dealing suggests the desirability of beginning with a qualitative discussion of the problem and proceeding to incorporate more factors into the model in order to approach a more realistic simulation of the pavement system.

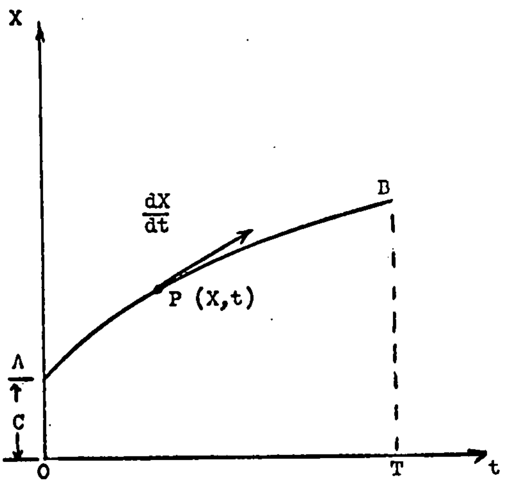
The basic idea in treating the control of a dynamic system as a multistage decision process (6) can be most easily grasped geometrically. Suppose that $X(t)$ denotes the position vector of a particle moving along a space curve. The trajectory of the particle is to be determined in such a manner that the "cost" of moving the particle over its trajectory is minimized. Intuitively, such a process can be thought of as a guidance (control) process in which continuous "steering" directions are required. Thus, a multistage decision process is defined to consist of the following operations:

1. Observing the system state $X(t)$ at time t;
2. Processing this information and making a decision utilizing a control rule; and
3. Modifying the evolution of the system by exerting the control. In the example chosen, at a point $P(X, t)$ along the particle trajectory we wish to determine dX/dt, i.e., the "steering direction" as a function of position (state) and time, such that the cost of the trip is minimized. In symbolic form we can pose this problem by seeking a function $G(X, t)$ such that

$$\frac{dX}{dt} = G(X, t), \text{ with the initial condition } X(0) = C \qquad (6)$$

along with

$$K[X(t)] = \text{minimum}, \ 0 \leq t \leq T \qquad (7)$$



Equation of Evolution

$$\frac{dX}{dt} = G(X,t)$$

Initial Condition

$$X(0) = C$$

Criterion Function

$$K\left[X(t)\right] = \text{Min.}\ ,\quad 0 \le t \le T$$

Figure 3. Trajectory controlled for minimum cost.

The problem, as shown in Figure 3, can also be posed in a slightly different manner to emphasize the aspect of control, leading to an algorithm of dynamic programming. Figure 3 shows that the direction dX/dt represents the control variable, which we define as

$$Y(t) \equiv \frac{dX}{dt} \tag{8}$$

Figure 3 also shows that the cost K will depend on the initial state C and the "life" of the trajectory, T. Therefore, we replace Eq. 7 with an equivalent statement

$$F(C, T) = \min_{Y(t)} K[X(t)] \tag{9}$$

In words, we seek a set of "controls," $Y(t)$, which minimizes the cost of a trajectory, parameterized by the initial state and life (duration) of the process. [As pointed out by Bellman (6), this is equivalent to defining a geodesic in the trajectory space in terms of its tangents.] Dynamic programming provides the computational algorithm for determining the set of controls for the process. This set constitutes an optimal policy for guidance of the process, which must satisfy the requirement that the criterion function K attain a minimum value.

In applications to pavement systems two things are apparent: (a) the state and control vectors, as well as the criterion function, are extremely complex; and (b) observations and decisions are made at a finite number of times, i.e., the process is discrete. This leads us to consider a model of discrete deterministic multistage decision processes, in which we now consider a sequence of states $X_1, X_2, \ldots, X_N$ and a sequence of control vectors (decisions) $Y_1, Y_2, \ldots, Y_N$. We define the evolution of the N-stage process by the equation

$$X_n = G(X_{n-1}, Y_{n-1}),\ n = 2, 3, \ldots, N \tag{10}$$

The meaning of Eq. 10 is that, at the initial state X_1 of the process, a decision Y_1 is made. This results in a new state X_2 given by Eq. 10:

$$X_2 = G(X_1, Y_1) \tag{11}$$

and so on, each new state depending on the immediately previous state and decision. Equation 10 corresponds to Eq. 6 in the continuous case. The set of decisions must be chosen such that the criterion function K is minimized, i.e., if

$$K = K(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_N) \tag{12}$$

the purpose of the decision process is to choose the Y_n so as to minimize Eq. 12. At this point the structure of the transformation function G is unspecified, except that it must lead to a unique state. (Although X_n in Eq. 10 depends only on the previous state, it is possible to extend the structure to incorporate hereditary effects, at the expense of computational complexity.) It may depend on the age of the process. The criterion function is presumed to possess a so-called Markovian property so that after k decision, the effect of the remaining (N - k) decisions on the value of K depends only on the system state at time k and the subsequent decision. An additive cost function K satisfies this requirement, i.e.,

$$K = f(X_1, Y_1) + f(X_2, Y_2) + \ldots + f(X_N, Y_N) \tag{13}$$

Principle of Optimality

We now consider how to establish the optimum set of decisions (optimal policy) for the problem posed by using dynamic programming. Bellman (7) states: "An optimal policy has the property that whatever the initial state and the initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision." The proof of this intuitive concept is virtually self-evident. Figure 3 shows that, if AB constitutes an optimal path, having arrived at P, PB must also constitute an optimal path. By using this principle we can deduce a recurrence equation for constructing an optimal policy, given the transformation function G, criterion function K, and the initial state of the system. Given (X_1, Y_1) the system is transformed to state X_2 according to Eq. 10; i.e., $X_2 = G(X_1, Y_1)$ and the N-stage process is reduced to an (N - 1)- stage process. Analogous to Eq. 9, Eq. 14 has the minimizing condition embedded in the initial state X_1 and "duration of process," N:

$$F_N(X_1) = \overset{\min}{Y_n} \ (K) = \overset{\min}{Y_n} \ [f(X_1+Y_1) + \ldots + f(X_N, Y_N)] \tag{14}$$

From the principle of optimality, and the Markovian structure of K, the cost of the last (N - 1) stages after making the first decision Y_1 will be

$$F_{N-1}(X_2) = F_{N-1}\,[G(X_1, Y_1)] \tag{15}$$

Thus, it follows that

$$F_N(X_1) = f(X_1, Y_1) + F_{N-1}\,[G(X_1, Y_1)] \tag{16}$$

From Eq. 14 this choice of Y_1 must be such that the right side of Eq. 16 is minimized. Thus,

$$F_N(X_1) = \overset{\min}{Y_1} \ \{f(X_1, Y_1) + F_{N-1}\,[G(X_1, Y_1)]\} \tag{17}$$

Allowing N to range over the values 2, 3, ... produces a recurrence relation connecting members of the sequence $[F_N(X_1)]$, thus specifying an optimal policy. We note the important result: The problem of selecting N decision vectors Y_n in an N-dimensional policy space is reduced to sequential selection of N vectors in a one-dimensional space. The computational significance of this result is obvious. One may well ask why dynamic programming should be selected over a straightforward search procedure that explores all possible policies and selects the policy leading to minimum cost. The answer is that the principle of optimality limits the choice of policies to those in the neighborhood of the policy for a minimum of the criterion function. Policies of no importance are thereby eliminated, along with attendant savings in computational time, a factor of prime importance in multidimensional state vector problems.

We turn now to a simple example chosen to illustrate application of the dynamic programming algorithm (Eq. 17) and associated concepts.

Example: An Elementary Model of Management as a Multistage Decision Process

Let us suppose (contrary to the consensus of speakers at this workshop) that performance serviceability index can be measured and is in fact the sole performance state variable X. Furthermore, it is supposed that the state variable is observed over the life of the pavement at some specified number of times. If no control over the system is exercised, a history of traffic and environmental inputs will cause a monotonic decrease in PSI, which is symbolically shown by Eq. 10 with $Y_n \equiv 0$,

$$X_n = G(X_{n-1}),\ n = 2, 3, \ldots, N \tag{18}$$

and is also shown in Figure 4, labeled 0 to denote zero cost of control. The initial state of the system X_1, normalized to unity, and the structure of the transformation G between states clearly depend on the initial design of the system. Furthermore, the transformation G also depends on the load and environmental inputs carried between state observations. Such a function clearly has to be born of field and road test experience. In order to "manage" our model system let us now introduce the notion of control via Eq. 10, where the set of decisions, Y_n, constitutes alternative maintenance, repair, or replacement operations. We seek an optimal policy for selecting these decisions in the face of certain restrictions, which are in part arbitrary but essential. Here, for simplicity, we choose as our criterion function minimum cost (Eq. 13). In Eq. 13, K represents the accumulated cost of performing the Y_N control operations. (The added effect of "cost of money" can also be included here.) As an added constraint to the

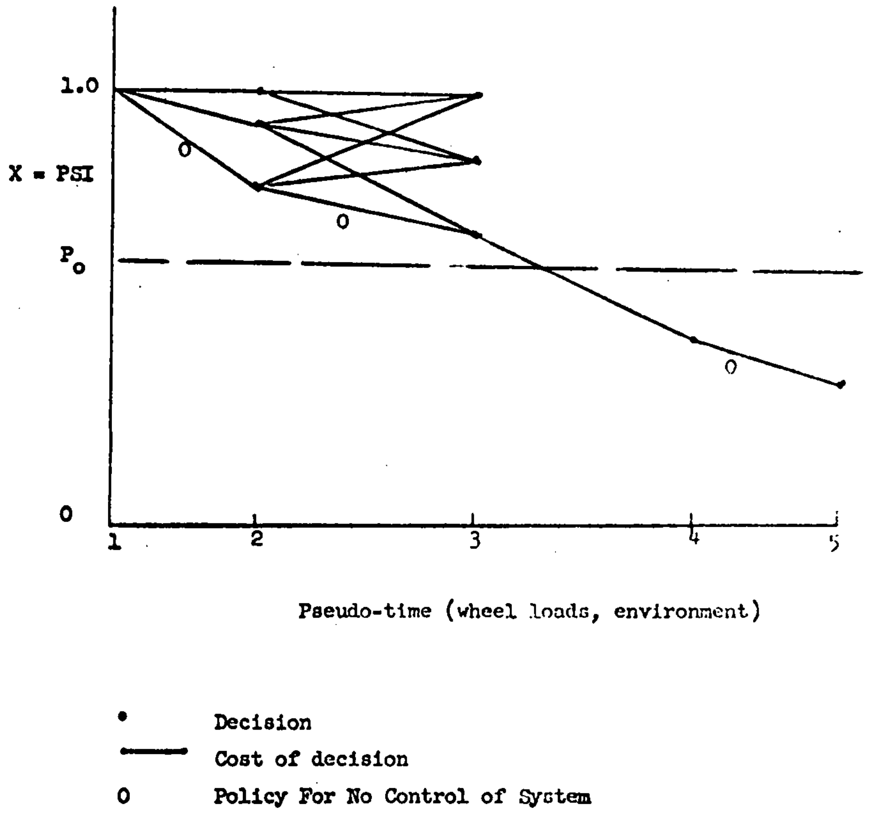


Figure 4. Routing graph of system performance model.

problem, we stipulate that the performance of the system be such that the mean value of the PSI exceeds some minimum value P_O, i.e.,

$$\frac{1}{T} \sum_{n=1}^{N} X_n(\Delta t_n) \geq P_o \tag{19}$$

where again T denotes system life and Δt_n denotes an interval between states, both measured in some pseudo-time. We must now select a sequence of decisions $\{Y_n\}$; i.e., find an optimal policy such that Eq. 13 is minimized subject to the constraint (Eq. 19). This is a straightforward problem in dynamic programming using the algorithm shown in Eq. 17 modified by a Lagrange multiplier to handle the constraint (7). One forms the modified function obtained by combining the criterion function K in Eq. 14 with the constraint condition in Eq. 19, using an undetermined Lagrange multiplier, λ.

$$F_N(X_1) = \min_{Y_n} \left\{ \left[f(X_1, Y_1) + \ldots + f\left(X_N, Y_N\right) \right] - \lambda \left[\frac{1}{T} \sum_{n=1}^{N} X_n\left(\Delta t_n\right) - P_o \right] \right\} \tag{20}$$

when a value is chosen for λ, the dynamic programming algorithm (Eq. 17) is used to obtain a policy for the decision sequence Y_n. After this the inequality (Eq. 19) is checked. The process is then repeated by selecting values of λ and repeating the same process until a policy is found for which Eq. 19 is best satisfied; this constitutes the optimal policy. The parameter λ is an important index of price of the control process; in this case it shows the trade-off in cost per unit of performance required to maintain a certain average value of PSI during the pavement life. Other types of constraints can be treated in a similar fashion. In such problems, certain concepts associated with graph theory can be helpful (8). One can view the choice of decisions as a routing problem. At each state, a set of points denotes new states produced by decisions (Fig. 4). The paths from state to state can be associated with costs of control, and one must select the path of optimal control, bearing in mind that the minimum performance criterion has to be satisfied for the set of decisions.

The model considered is clearly an elementary one, but it can be considerably embellished. When the mechanics of management of systems are better understood, the initial design (inverse) problem might be incorporated as part of the decision process. In this instance the optimal policy is to be found over a set of parameters describing control variables as well as design parameters of the system itself. An example of this type can be found in the report by Hudson et al. (4).

Uncertainty

Thus far we have made the tacit assumption that all aspects of the systems with which we are working are deterministic. This applies equally to input, system model, and control. In other words we are certain of the input, which in turn leads to a certain output, and a control applied to the system produces a certain change of state. Use of the term certain is equivalent to assigning a probability of unity in each of these instances. It is a euphemism to assert that pavement system analysis and design is an uncertain problem. Aspects of this overall problem are discussed by Sherman (15) and Moavenzadeh (11), and a suggested treatment of the overall systems problem is mentioned by Lemer and Moavenzadeh (5).

In concluding I wish only to call attention to the need to examine the modeling problems of design and management of pavement systems in the light or, perhaps better, the darkness of uncertainty. The root of the problem is the notion of determinism-cause and effect, combined with the perversity of nature and man. For example, the input variables, traffic and environment, are clearly nondeterministic (stochastic) in the sense that one must attach a probability distribution to these inputs. Similarly, the pavement system itself, by virtue of its constituent materials and methods of construction, possesses a stochastic character; even a deterministic input to such a system will produce a stochastic output. Furthermore, application of a control likewise leads to an uncertainty in outcome.

How does all this uncertainty affect our efforts to develop a rational basis for design and management of pavement systems? Briefly, we can consider the previous apparatus used in this section only with reinterpretation of the primitive elements. We can define a discrete stochastic multistage decision process by asserting that a decision Y_n determines a set of possible outcomes (states) instead of a unique outcome. The state vector X_n is now a stochastic vector in the sense that its components are probability distributions. Furthermore, the transformation leading to state changes, i.e., G( ) in Eq. 10, is a stochastic transformation. In addition the criterion function, depending now on stochastic variables, is itself a stochastic quantity, as are the decision vectors, which depend on the system states. The condition of "minimum of the criterion function" can be replaced by minimum of the expected value of the stochastic criterion function, which leads to the notion of an optimal policy for a discrete stochastic process: Select a sequence of decision vectors $[Y_n(X_n)]$ such that the expected value of the criterion function is minimized. For the special case of a Markov process this problem has been studied in some detail (7). Whether a Markov model is adequate for the pavement system problem is another matter. This is an area needing considerable exploration.

SUMMARY

In this brief review of the status of research primarily associated with development of a more rational basis for design and management of flexible pavements, I have tried to emphasize two basic ideas: (a) the need to develop an overall structure for the entire pavement system, an assembly of many black boxes; and (b) the need to explore in some detail the contents of the various boxes to develop mathematical models of each subsystem, leading eventually to a model of the system in its entirety. A number of suggested directions in each of these categories have been discussed, and a more detailed treatment is given in other papers. I have separated the problem of design from the problem of management only for purposes of clarifying their treatment. One can and must eventually regard the two as one problem when more reliable models of system behavior become available, recognizing that this activity is a pattern recognition problem of special complexity.

What seems to me to be incumbent on those attending this workshop, as well as managers of research and development funds in general, is the development of a systems model for management of state and federal programs directed toward the problem under consideration. One needs to examine carefully the matter of "sensitivity" of various black boxes with regard to system performance. Although it may be commendable to study one box alone in the name of science, it is surely poor engineering practice to channel a lot of support to a subsystem with a low sensitivity factor vis-à-vis the total system performance. Such decisions regarding support should be made in the light of information and reason: They are difficult and agonizing, but that is what managers are expected to live with.

REFERENCES

1. Standard Nomenclature and Definitions for Pavement Components and Deficiencies. HRB Spec. Rept. 113, 1970.
2. Hudson, W. R., Finn, F. N., et al. Systems Approach to Pavement Design, Systems Formulation, Performance Definitions and Materials Characterization. Materials Research and Development, Inc., Oakland, Calif., Final Rept., March 1968.
3. Pister, K. S. A Systems Analysis Plan for Research in Structural Design of Pavements. Materials Research and Development, Inc., Oakland, Calif., TM 67-5, Jan. 1967.
4. Hudson, W. R., et al. A Systems Approach Applied to Pavement Design and Research. Center for Highway Research, Austin, Texas, Research Rept. 123-1, March 1970.

5. Lemer, A. C., and Moavenzadeh, F. The Analysis of Highway Pavement Systems. Dept. of Civil Eng., Material Research Lab., M.I.T., Cambridge, Professional Paper P69-12, Sept. 1969.
6. Bellman, R. Introduction to Mathematical Theory of Control Processes. Vol. 1, Academic Press, 1967.
7. Bellman, R. Adaptive Control Processes: A Guided Tour. Princeton Univ. Press, 1961.
8. Kaufmann, A. Graphs, Dynamic Programming and Finite Games. Academic Press, 1967.
9. Finn, F. N. Observations of Distress in Full-Scale Pavements. Presented at the Workshop on Structural Design of Asphalt Concrete Pavement Systems and included in this Special Report.
10. Hudson, W. R. Serviceability Performance and Design Considerations. Presented at the Workshop on Structural Design of Asphalt Concrete Pavement Systems and included in this Special Report.
11. Moavenzadeh, F. Damage and Distress in Highway Pavements. Presented at the Workshop on Structural Design of Asphalt Concrete Pavement Systems and included in this Special Report.
12. Nair, K. Solutions and Solution Techniques for Boundary Value Problems. Presented at the Workshop on Structural Design of Asphalt Concrete Pavement Systems and included in this Special Report.
13. Westmann, R. A. Fundamentals of Material Characterization. Presented at the Workshop on Structural Design of Asphalt Concrete Pavement Systems and included in this Special Report.
14. Deacon, J. A. Materials Characterization—Experimental Behavior. Presented at the Workshop on Structural Design of Asphalt Concrete Pavement Systems and included in this Special Report.
15. Sherman, G. B. In Situ Materials Variability. Presented at the Workshop on Structural Design of Asphalt Concrete Pavement Systems and included in this Special Report.
16. McCullough, B. F. Distress Mechanisms—General. Presented at the Workshop on Structural Design of Asphalt Concrete Pavement Systems and included in this Special Report.