DOCUMENT RESUME

ED 152 235 IR 005 578

AUTHOR	Yovits, M. C.; And Others
TITLE	Development of a Theory of Information Flow and Analysis.
INSTITUTION	Ohio State Univ., Columbus. Computer and Information Science Research Center.
SPONS AGENCY	National Science Foundation, Washington, D.C. Div. of Science Information.
REPORT NO	OSU-CISRC-TR-77-12
PUB DATE	Aug 77
GRANT	DSI-76-21949; GN-41628
NOTE	36p.; Best copy available
EDRS PRICE	MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS	Decision Making; Higher Education; Information Retrieval; *Information Science; *Models

ABSTRACT

This paper describes research being carried out at the Ohio State University leading toward a general theory of information flow and analysis. The objectives of this program sponsored by the National Science Foundation include the following: (1) to develop a theory of information flow and analysis; (2) to identify important parameters and variables in the information process which can be measured and quantified; (3) to develop relationships among the variables which describe their behavior and limitations; (4) to apply this theory to specific practical situations, particularly those involving science information; and (5) to develop both simulation and experimental models for quantification and validation of the theory. (Author/AWP)

(OSU-CISRC-TR-77-12)

DEVELOPMENT OF A THEORY OF INFORMATION FLOW AND ANALYSIS

by

M. C. Yovits, L. L. Rose,
and
J. G. Abilock

Work performed under

Grant Numbers GN 41628 and DSI 76-21949

Division of Science Information

National Science Foundation

Computer and Information Science Research Center
The Ohio State University
Columbus, Ohio 43210

August, 1977

TABLE OF CONTENTS

Development of A Theory of Information Flow and Analysis*

M.C. Yovits, L.L. Rose, J.G. Abilock
The Ohio State University
Columbus, Ohio

Abstract

This paper describes research being carried out at The Ohio State University leading toward a general theory of information flow and analysis. The objectives of the National Science Foundation-sponsored research program† include the following: (1) to develop a theory of information flow and analysis; (2) to identify important parameters and variables in the information process which can be quantified and measured; (3) to develop relationships among the variables which describe their behavior and limitations; (4) to apply this theory to specific practical situations, particularly those involving science information; and (5) to develop both simulation and experimental models for quantification and validation of the theory.

The generalized model of information flow is shown to represent virtually any decision situation. Using this model we derive a number of measures. Two of the more important ones are: $I = m \sum_{i=1}^{m} P(a_i)^2 - 1$ and

$DME = \left[\sum_{i=1}^{m} P(a_i)\ EV^*_i\right] / max\ (EV^*_k)$. With these measures we can quantify the amount of information (I) in a decision state and evaluate decision-maker effectiveness (DME). We can also determine the value of information through its

---

* Portions of this paper were presented at the American Association for the Advancement of Science Meeting in Denver, Colorado, February 25, 1977.

† This material is based upon research supported by the National Science Foundation, Division of Science Information, under Grant Numbers GN 41628 and DSI 76-21949.

effec on *DME*. Rules are derived for use by the decision-maker for assimilating new data in his estimates of values and for use in determining selection probabilities for various courses of action. These rules enable feedback, learning, and alternative selection to be modeled, measured, and evaluated. Research is underway to validate the model realistically and to apply it to practical situations.

A further possible result of this research is the development of a decision calculus which will establish guidelines for decision-making given certain situations. These guidelines should permit quantification of the importance of information in the decision-making process.

Background

In several previous papers (1,2,3), we have discussed some of the properties which should exist if information science is ever to become a "true" science similar to physics or chemistry. It has been pointed out that a number of analytical expressions and concepts should exist which can be used to describe and analyze information flow. A framework, called a "generalized information system" was suggested which permits the development of these concepts and expressions.

The word "information" takes on a variety of meanings depending upon the context in which it is used. Our approach relates information to its effectiveness and thus its use and value. Information is frequently used rather specifically in the sense of the Shannon and Weaver "information theory" (more accurately called "communication theory"). In this sense the context of the message is of no significance; the theory is concerned with the probability of the receipt of any particular message for various conditions of the transmission system. While this may be of interest in information science, it is certainly not the major nor even a large part of information science. Such a treatment does not consider the vital areas of concern, almost all of which are involved with the context, meaning, and effectiveness of the message. For these reasons, the Shannon and Weaver approach is generally regarded as too restrictive to be a basis for the formulation of an information science. At the other extreme, the treatment of information to be synonymous with knowledge appears to be far too broad to lead to meaningful and useful principles or relationships in information science.

In our formulation we treat information to be *data of value in decision-making*. Later in the paper we define information quantitatively and rigorously. While our formulation may somewhat delimit the total range of interest in an intellectual sense, it *does* have virtually universal applicability with regard to any potential applications for information. The authors also feel that any more general definition is not amenable to the quantification and conceptualization necessary to establish meaningful relationships. An implication of this definition then is that information is used *only* for decision-making and that the decision maker has *only* the resource of information available to him. Thus, information and decision-making are very closely bound together in our general model.

Levels of Information

Research in the general area of information theory started in the mid-1940's. The basic theory, as already noted, was established with the fundamental work of Claude Shannon and Warren Weaver *(4,5)*. This theory covers the transmission of messages over a channel, independent of meaning. Their defined measure of information, which relates to the entropy of the interpretation of the sequence of bits transmitted, bears little relation to the layman's concept of information. This is acknowledged by Weaver, who defines three levels of information, the first of which (the technical level) is his concern.

The second level of information is concerned with the meaning of information - its semantic content. Significant research in this area has been performed by Carnap, Bar-Hillel, Winograd, and others *(6,7)*. Their research has attempted to measure the semantic content of simple declarative sentences within their language system. In this context, they do not refer explicitly to the process of communication between individuals. While the level 1 information definition is concerned with the successful transmission of a message from point *a* to point *b*, the level 2 definition of information is concerned with the successful interpretation or understanding of a message once it is received.

The third level of information deals with the effectiveness of information - i.e., how information, once received and understood, is utilized. Hence the level 3 model is imbedded in a decision-making environment, since that is the only framework in which one can observe information utilization. Little research has thus far been accomplished in this area, even though it is probably the area of greatest significance, interest, and application *(8,9,10,11)*. Our research is in this area in an effort to develop a general and useful theory of information *(1,2,3)*.

Effectiveness of Information

Research by several authors has been previously performed in an effort to define an effectiveness or pragmatic information measure. MacKay *(12)* defines the value of any data item to be the base 10 logarithm of the ratio of the performance of the system after and before receipt of the data. Just how one defines "system performance" is left open by MacKay.

Cherry *(13)* argues that information aids the decision-maker by narrowing the range of hypotheses. Information thus reduces decision-maker uncertainty by narrowing his range of viable alternatives. The information measure Cherry introduces is the logarithm of the ratio of the aposteriori and apriori probabilities of selecting an alternative and is based upon Bayes theorem.

Goffman *(14)* also considers a measure for effectiveness of information which satisfies all of the desirable properties of the Shannon measure. This measure quantifies the amount of non-redundant information in a system comprised of many communicators. Goffman's information measure is much like the Shannon entropy measure, but relates more to defining "connectedness" via information flow between the elements of the system.

Our research considers information in the domain of decision-making and a conceptual decision-maker (DM). It is important to consider information in a decision-making environment inasmuch as it is necessary to measure the effect of the information on the recipient (decision-maker). This can be done only by monitoring the DM's actions and the observables resulting from the decisions made. We have chosen to focus upon all data incoming to the DM in an effort to quantify the informative content of that data. Our information measure thus relates input data to its effect upon the decision-making process as shown below.

It is apparent that information - if considered from the effectiveness point of view - is heavily concerned with the decision process, and therefore it is important to consider information and decision making together. While the fundamental objectives of the current research are to develop a theoretical formulation for a basic theory of information flow and analysis, we believe that this cannot be done without considering in detail the use of the information; that is, for making decisions.

There is, of course, a great body of literature in the decision-making field. Some of the typical references are by such researchers as Morris *(15)*, Raiffa *(16)*, Schlaiffer *(17)*, and Radnor *(18)*. An excellent survey by Bandyopadhyay *(19)* provides further references to research in this area. Unfortunately, we find that the approaches described in the literature are neither general enough nor sufficiently flexible to represent the situation adequately with regard to the use of information. In particular, there is little in the literature pertaining to the non-expert decision maker, whereas information will be of perhaps greatest value in this situation. Furthermore, we believe it is necessary to consider the *entire* decision process which is generally repetitive over periods of time. This process heavily involves feedback information using the results of previous decisions.

The General Model

Our general model of information flow and analysis is illustrated in Figure 1, and we briefly describe its operation. This model can be used to describe most, if not all, information-dependent activities. It provides a way of looking at any information-decision interaction and defining the role and flow of information in the system. The IAD module (Information Acquisition and Dissemination) processes data for the system. Both exogenous (external environmental) and endogenous (internal feedback) data are acquired by the IAD module. Whenever a decision must be made, the DM module (using all of the data available) establishes the possible courses of action and selects the "best" one to execute. The Execution module executes the DM-chosen course of action, according to all pertinent external environmental factors, leading to various outcomes depending on the alternative executed. These outcomes will be some *observable* quantities. They must be observable in physical sense if they are to have any effect. The Transformation module takes all observables of the alternative executed and turns these observations into data. These data are fed back into the IAD module and we have come full circle, following the flow of information in the model.

This model of a generalized information system rests upon three basic hypotheses:

H1. Information is data of value in decision-making;
H2. Information gives rise to observable effects;
H3. Information feedback exists so that the Decision Maker will adjust his model for later similar decisions.

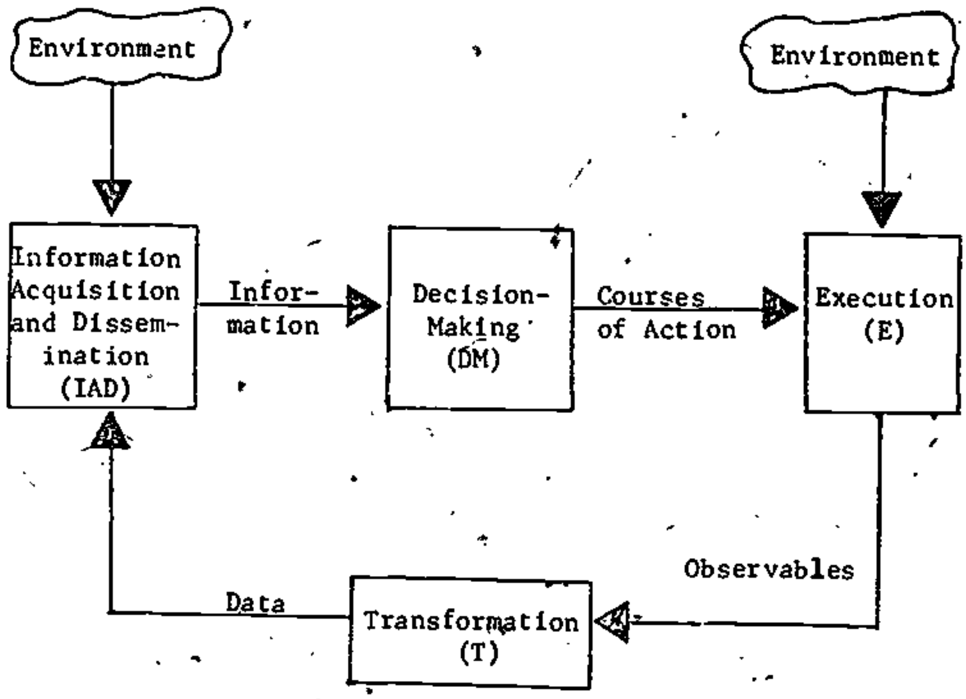

Figure 1: The Generalized Information System Model

The first hypothesis requires that information be used in a decision-making context. If information is received, but never used or applied to a subsequent decision, then its effect does not exist and it cannot be measured. Hypothesis 2 assures that if the decision-maker (DM) does make a decision, then the outcome of that decision can be observed and measured. This precludes decision-making in a vacuum. Observables must exist if the decision-making and the courses of action are to be evaluated. Hypothesis 3 indicates that the DM learns from feedback data resulting from previous decisions. Note that the observed outcomes of repetitive or related decision-making situations provide data upon which future decisions will be made.

Generalized Decision Making

This model is dynamic in nature: even in a stationary decision situation the DM's model may well change with time. A DM learns about his particular decision situation and environment as follows:

1. He makes a decision (chooses a course of action) on the basis of all information available to him;
2. He predicts some probable outcomes;
3. He compares actual resulting observables against his predicted observables (feedback learning);
4. He updates his total model of the situation as a result of this process;
5. He returns to step 1.

For any problem environment, a DM attempts to fulfill two main objectives:

1. To choose the "best" course of action (c.o.a.) according to some criterion given his state of knowledge;
2. To learn the most about the total existing situation from the decision-making process.

It is important to note that a DM has these two main objectives in choosing a course of action. "Classical" decision theory says that a DM always chooses that c.o.a. that maximizes some criteria. This is a considerable oversimplification, and the DM in fact does not automatically choose the c.o.a. with highest expected return. This is true for at least three reasons. First, because the DM may be unsure of his estimates; second, he may wish to learn more about the total situation by executing alternatives other than the one with highest expected return; and third, the situation may be changing with time, or if not the DM must generally assure himself that it is a static situation.

Learning from any c.o.a. executed is in fact a very important objective of decision-making. The DM does this by monitoring the outcomes resulting from the decision. This feedback updates the DM's data base upon which decisions are based. Given the outcome of earlier decisions, then how will the DM update his data base? If the outcome data matches the DM's expectations, then his confidence will rise even though there may be no other change in his estimates. However, if the expected and actual results vary, then the DM will incorporate these new data into his current estimates. He learns from the deviation between his expected values of the outcomes and the actual values. Learning is similar to sampling from a distribution and predicting the underlying theoretical distribution. The better the DM's learning capability, the better his expected performance. The action of learning updates the DM's state of knowledge, with input arriving as either external or feedback data as illustrated in Figure 1. Learning increases the DM's confidence in his perception of the decision situation. Learning more about a given situation further removes uncertainty by giving the DM a better estimate of the vari us possibilities open to him.

Decision-Maker Uncertainties

The uncertainties with which a DM must cope at any given time can be classified into three categories: state of nature uncertainty, executional uncertainty, and goal uncertainty. The states of nature encompass the uncontrollable external conditions that will determine the various outcomes. Depending upon the decision to be made, they might include for example weather, economy, competitive environment, governmental regulation. The more knowledgeable the DM is of the probable current environmental conditions (i.e., the prevailing state of nature), the more effectively he can make decisions.

Executional uncertainty appears in two ways. First the DM must identify the c.o.a.'s available to him - his options. Second, he must determine likely outcomes for each c.o.a. under consideration. If any of five outcomes, for example, is possible for a given c.o.a., then what is the probability that a particular one will occur? The DM must determine these probabilities of occurrence of various outcomes for each c.o.a. This represents his best approximation to the actual situation. For any complex system, the relationship between outcomes and courses of action is probabilistic and not deterministic even if the state of nature were known with certainty. In other words, executional uncertainty is an inherent part of the decision-making process.

Lastly, goal uncertainty relates outcomes to goal-achievement. The DM must examine each outcome considered and evaluate it in light of his goals: i.e., he must (if he is to be successful) recognize the value or lack thereof of each possible outcome to the attainment of his particular goals.

Each of the uncertainties discussed has both a structural and relational context. The DM has structural uncertainty about the *number* of relevant states of nature, the *number* of viable alternatives, and the *number* of outcomes that may occur as a result of executing the alternatives. Any structural deficiency will degrade the DM's performance; e.g., not considering a certain c.o.a. or not knowing that a given outcome may result will exascerbate the difference between the DM's expected and actual performance.

Once the structure is identified by the DM, he must then resolve *relational* uncertainties. What is the probability a given state of nature prevails? What is the probability that a particular c.o.a. will result in a specific outcome? What is the value of each outcome relative to goal attainment?

Some Examples

Let us now illustrate typical decision situations and specifically identify the uncertainties facing the DM.

Consider first the case of a family physician, the decision-maker. His decisions relate to proper diagnosis and treatment of the ailments of his patients. An ill patient visits his physician with a set of symptoms. The physician's first step is to identify as best he can the illness giving rise to these symptoms. The illness is the prevailing state of nature. Inasmuch as many diseases may have overlapping symptoms and give rise to the set of those symptoms experienced by the patient, identification of the specific illness clearly is probabilistic.

The next step is to identify a treatment - the course of action. Each treatment will have a number of effects, most of which are known to the physician only probabilistically - *even if he knew with certainty the precise illness* (state of nature). For example, a particular medicine may have unexpected side effects. Instead of curing the patient new symptoms may appear or old ones may be exascerbated. This is executional uncertainty.

The symptoms are given a value structure by the doctor (and by the patient also). Some symptoms are obviously of little interest, and some are vital to patient health. Note that the symptoms are the *only* observables. When the symptoms are measured, they become data and perhaps information. Blood pressure, for example, is a physical manifestation, whereas it becomes data when measured. Further examination of the patient provides information to the physician concerning the actual effects of the treatment in comparison with the expected effects, and if these differ a new course of action may now be chosen and executed thus repeating the entire process. This illustrates the feedback involved. Note that this iterative process is the only way a physician as a DM can learn more about the entire situation and in particular the effectiveness of various treatments.

A second example of some current interest which illustrates our model and corresponding uncertainties involved is that of the economy of the United States. The observables are the rate of inflation, unemployment, growth of the GNP, money supply, etc. The states of nature are inflation, recession, growth rate, energy availability, etc. These are clearly all probabilistic in nature and of course, never known with certainty at any time. In order to achieve some goal, which may be a specific growth rate or a reduced unemployment or a stable economy, certain courses of action are taken by the Government, the DM. More money may be made available to consumers by lowering taxes, a Government-financed job program may be undertaken, interest rates may be changed, etc. The relationships of these courses of action to the observables is probabilistic, and economists can speak only in terms of probabilities of certain actions having certain observable outcomes. This is, of course, the executional uncertainty which *exists even if the state of nature were known with certainty*. The goal structure is, of course, important and changeable. For example, does the President want to control inflation or cut unemployment? Which is more important (has the higher value structure)?

The outcomes are measured periodically and the results (now data) are reported to the President who examines the observable outcomes in comparison with the predicted outcomes. This is feedback. On the basis of his knowledge of the situation (his predicted model) and the data obtained, he readjusts his expectations and may choose a new course of action, and thus he reiterates the process. Again, note that this iterative process is the only way to learn about the entire situation and the effectiveness of various courses of action.

Model Representation

This decision-making process can be analytically modeled in a number of different ways. One procedure that appears to represent this situation well and which can be formally manipulated uses decision matrices as shown in Figure 2. These matrices are both of dimension m by n to correspond to the m alternatives and n outcomes under consideration by the DM. An additional dimension should be added to each matrix (e.g., m by n by r) to express the r possible states of nature. Extending the model to consider other states of nature is straightforward and appears to add little to the discussion herein, so for simplicity we consider only one state of nature in the remainder of this paper. We define the matrices W and V to be the *decision state* of the DM. The decision state relates to a particular decision, is unique to each DM, and will change with time.

The W matrix describes the DM's executional uncertainty. Element w_{ij} represents the DM's *estimate* of the probability that execution of the *i-th* alternative will result in the *j-th* outcome. We denote by W^* the matrix describing the actual probabilities for the executional uncertainty of the decision situation. Each element w^*_{ij} of matrix W^* represents the *actual* probability of occurrence of outcome j if alternative i were to be executed. This is an *aposteriori* probability in the standard sense and represents the fraction of times outcome j would occur if alternative i were executed many times. We repeat that these matrices as indicated are for a single state of nature. There would in reality be a third dimension of dimensionality r for the different states of nature.

The decision-maker develops his estimates of the w_{ij}'s on the basis of experience and whatever data or information he has available. These are his best estimates of the w^*_{ij}'s. The more "expert" the DM is the closer his w_{ij}'s should be to the w^*_{ij}'s.

If $w_{ij} = w^*_{ij}\ \forall\ i,j$, then the DM has correctly assessed the consequence of his possible actions. If $m < m^*$, then the DM is unaware of certain courses of action; if $n < n^*$, then one or more outcomes may occur of which the DM is unaware. If $m > m^*$, then the decision-maker believes that certain

W, Probability Matrix for outcomes for various courses of action

$$
\text{Courses of Action}\quad
\begin{array}{c|cccc}
 & \multicolumn{4}{c}{\text{Outcomes}} \\
 & o_1 & o_2 & \cdots & o_n \\
\hline
a_1 & w_{11} & w_{12} & \cdots & w_{1n} \\
a_2 & w_{21} & w_{22} & \cdots & w_{2n} \\
\cdot & \cdot & \cdot & & \cdot \\
\cdot & \cdot & \cdot & & \cdot \\
\cdot & \cdot & \cdot & & \cdot \\
a_m & w_{m1} & w_{m2} & \cdots & w_{mn}
\end{array}
$$

V, Value Matrix for outcomes for various alternatives

$$
\text{Courses of Action}\quad
\begin{array}{c|cccc}
 & \multicolumn{4}{c}{\text{Outcomes}} \\
 & o_1 & o_2 & \cdots & o_n \\
\hline
a_1 & v_{11} & v_{12} & \cdots & v_{1n} \\
a_2 & v_{21} & v_{22} & \cdots & v_{2n} \\
\cdot & \cdot & \cdot & & \cdot \\
\cdot & \cdot & \cdot & & \cdot \\
\cdot & \cdot & \cdot & & \cdot \\
a_m & v_{m1} & v_{m2} & \cdots & v_{mn}
\end{array}
$$

Figure 2: W and V: The Decision State for One State of Nature

courses of action are viable when in fact they are not. Similarly for $n > n^*$. We see that $w_{ij} \neq w^*_{ij}$ for some i,j suggests a *relational* uncertainty; $n \neq n^*$ or $m \neq m^*$ indicates *structural* uncertainty.

There is a *value* for every possible alternative-outcome pair. Therefore a value matrix V, also of size m by n, can be superimposed over the decision matrix W. Each element v_{ij} of matrix V represents the DM's estimate of the value of outcome j as a result of executing alternative i. Note that the value is a function not only of the outcome but also of the course of action chosen. The following examples illustrate why this must be so.

Consider the situation where a document from a collection is chosen on some basis to be relevant or non-relevant (the c.o.a.). If it is chosen to be relevant and it turns out to be relevant (the outcome), then the set of {c.o.a., outcome} has a particular value to the DM. If, on the other hand, the document were chosen to be non-relevant and it turns out to be relevant, then the value to the DM of this set is quite different from the previous situation and is in fact negative.

Consider the simple betting game of flipping a coin (heads or tails). The value matrix (if we win or lose a silver dollar) is as follows:

		Heads	Tails	{outcomes	
Alternative	{Heads	\$1	-\$1		Value Matrix
	{Tails	-\$1	+\$1		

Clearly the outcome values depend upon the alternative selected.

Consider also a three alternative model where the options are to buy, sell, or hold a given stock. The outcomes are that the stock price goes up, down, or remains the same. Presume that the first outcome occurs: i.e., the stock price increases. Considering our three alternatives, we observe the following: had the DM purchased some of this stock he would have realized a profit: had he sold some of this stock he would have realized a loss: had he done nothing he would have realized neither profit nor loss. Hence we see three values for the same outcome: each is alternative-dependent.

Currently for simplicity we assume $V = V^*$, where V^* denotes the actual value matrix: i.e., the DM correctly knows the value of each outcome identified. Removing this assumption adds another dimension to DM learning. Full generality, as we have indicated, considers V to be of size m by n by r for the r possible states of nature. It follows then that the value of an outcome depends also upon the existing state of nature, which may change with time.

We have chosen to use an expected value model for evaluation of the matrices. By this we mean that the expected value of any c.o.a. is computed as the weighted average of the outcome-value pairs. Given the W and V matrices which describe the decision state of the DM, one can compute the DM's expected value of each alternative as follows:

$$EV_i = \sum_{j=1}^{m} w_{ij} \, v_{ij} \tag{1}$$

Other procedures for determining values of c.o.a.'s can be used, but the expected value model is reasonable and tractable and thus convenient to use.

The m EV's computed represent the DM's *assessment* of the expected values of the m alternatives. The *actual expected value* of each alternative we denote by EV^*_i. This value is obtained by using elements w^*_{ij} of matrix W^* for the DM's w_{ij} values in equation (1) above. The EV^*_i then represents the average outcome value which would be expected if alternative i were executed many times. The DM does not know the value of EV^*_i; else he would simply choose the alternative with the highest EV^* and thus maximize his returns. The closer the DM's EV vector is to vector EV^*, the more correct his assessment of the situation and the better decision maker he should be.

Our current model considers only positive EV values: $\forall i,\ EV_i \geq 0$. We would also like to treat negative expected values for the following reason: the DM must be penalized for choosing a bad alternative. Hence his outcome value for this alternative must be negative, as indicated in the relevant/non-relevant document example described earlier. If enough executions of alternative k will result in a loss to the DM, then the resultant EV_k will be negative. Negative terms in our formulations cause certain mathematical problems which must be treated. We plan, in the near future, to expand our model so that negative EV's can be considered and handled by the model.

Making a Decision

The DM makes a decision by selecting a course of action to execute. Just how should he make this decision? Clearly it will be based upon all the data he has stored regarding his current assessment of the decision state: the W and V matrices. We have just shown how an expected value vector, EV, can be derived from the decision state. If the DM is to make a *rational* decision, then it must be based upon these data. Although we do not impose a specific decision strategy upon the DM, we will make one reasonable assumption. This is that the DM will prefer alternative i to alternative j if $EV_i > EV_j$.

The DM must establish a probability for choosing each c.o.a. In classic expected value decision theory, the DM will *always* choose the c.o.a. with the greatest expected value. But selecting the alternative of highest EV is generally the best strategy only if the DM has virtually full confidence in his derived EV values. The DM may have considerable structural or relational uncertainty regarding the situation. Also, he may wish to learn more about the total decision situation. The only way he can do this is to sample c.o.a.'s in addition to the one with maximum EV. Then he can compare predicted and observed outcomes to learn and reevaluate his EV estimates. Furthermore, the situation may be changing with time and the DM simply may not wish to rely completely upon his previously computed EV values.

Thus we desire a method by which a probability can be associated with each alternative, representing the DM's *probability* of *selecting* that alternative. This distribution should have the properties that:

a) $\sum_{i=1}^{m} P(a_i) = 1;$

b) $EV_i > EV_j \Rightarrow P(a_i) \geq P(a_j);$

c) no DM confidence in his knowledge of EV's $\Rightarrow \forall_i\ P(a_i) = \frac{1}{m}$; that is, with no information about the values to be expected, the DM will select a c.o.a. at random;

d) total DM confidence in his knowledge of EV's $\Rightarrow P(a_i) = 1$ for maximum EV_i and zero for all others.

We propose the selection rule

$$P(a_i) = (EV_i)^C \Big/ \sum_{k=1}^{m} (EV_k)^C \quad \ni C \text{ is a non-negative real} \qquad (2)$$

for situations involving positive EV values.

The variable C in equation (2) above is called the confidence factor: the higher the value of C, the higher the DM's confidence in his knowledge of the EV's. If $C = 0$ then $P(a_i) = 1/m\ \forall i$ and we have the random case. The DM has no confidence in his current state of knowledge so he prefers to make a random choice rather than let his EV's determine which c.o.a. to execute. At the other extreme, $C \Rightarrow \infty$ results in the classic decision theory rule: the alternative with highest EV is executed every time.

The following lemma formally demonstrates that this selection rule does indeed satisfy the four essential criteria noted above.

<u>Lemma 1</u>: The selection rule $P(a_i) = (EV_i)^C \Big/ \sum_{k=1}^{m} (EV_k)^C$ for $C,\ EV_i \geq 0$ satisfies the four requisite properties for establishing the DM selection distribution.

<u>Proof</u>:

a) $\sum_{i=1}^{m} P(a_i) = 1.$

Let $P(a_i)$ be defined as in equation (2).

Then $\sum_{i=1}^{m} P(a_i) = \sum_{i=1}^{m} \left[(EV_i)^C \Big/ \sum_{k=1}^{m} (EV_k)^C \right].$

Hence $\sum_{i=1}^{m} P(a_i) = \sum_{i=1}^{m} (EV_i)^C \Big/ \sum_{k=1}^{m} (EV_k)^C = 1.$

b) $EV_i > EV_j \Rightarrow P(a_i) \geq P(a_j)$.

Let $EV_i > EV_j$.

Then $(EV_i)^C \geq (EV_j)^C \quad \ni C \geq 0$.

Hence $(EV_i)^C / \sum_{k=1}^{m} (EV_k)^C \geq (EV_j)^C / \sum_{k=1}^{m} (EV_k)^C$;

i.e., $P(a_i) \geq P(a_j)$.

c) No DM confidence in his knowledge of

EV's $\Rightarrow \forall_i \; P(a_i) = \frac{1}{m}$.

Let $C = 0$.

Then we have $\forall_i \; P(a_i) = (EV_i)^0 / \sum_{k=1}^{m} (EV_k)^0$.

Hence $\forall_i \; P(a_i) = 1 / \sum_{k=1}^{m} 1 = \frac{1}{m}$.

d) Total DM confidence in his knowledge of

EV's $\Rightarrow \exists\, i \ni P(a_i) = 1$ and all other $P(a_k) = 0$.

Assume $\exists\, i \ni \forall_k \; EV_i > EV_k$ for $i \neq k$.

Then consider the term $P(a_i)$ as $C \to \infty$.

$$\lim_{C \to \infty} P(a_i) = \lim_{C \to \infty} \left[(EV_i)^C / \sum_{k=1}^{m} (EV_k)^C \right]$$

$$= \lim_{C \to \infty} \left\{ 1 / \left[\sum_{k=1}^{m} (EV_k)^C / (EV_i)^C \right] \right\}$$

$$= \lim_{C \to \infty} \left\{ 1 / \sum_{k=1}^{m} \left[(EV_k)^C / (EV_i)^C \right] \right\}.$$

But $\forall k \neq i \; \lim_{C \to \infty} (EV_k)^C / (EV_i)^C = \lim_{C \to \infty} (EV_k / EV_i)^C = 0$,

since $EV_i > EV_k$.

Lastly, for $k = i$ $\lim_{C \to \infty} (EV_k)^C / (EV_i)^C$

$$= \lim_{C \to \infty} \left(\frac{EV_k}{EV_i}\right)^C = \lim_{C \to \infty} \left(\frac{1}{1}\right)^C = 1.$$

Therefore $\lim_{C \to \infty} P(a_i) = 1$ and by a) above,

$$\forall k \neq i \quad \lim_{C \to \infty} P(a_k) = 0.$$

QED

Decision-Maker Learning

In our approach, we use equations (1) and (2) in order to calculate Expected Values and Probabilities of alternatives available to the DM. The following sequence of events then occurs. This procedure is a general one for all expected value models and is not dependent upon specific formulations of the probabilities. In actual practice, the DM may determine probabilities mainly by judgment rather than from calculation.

1. DM estimates or predicts EV_i for $i = 1, 2, \ldots, m$ on the basis of all of his past experiences using equation (1);
2. DM establishes $P(a_i)$ for $i = 1, 2, \ldots, m$ from his estimates of the EV_i and his current level of confidence in the data. In our approach he uses equation (2) to derive these estimates;
3. DM now executes an alternative, k, in accordance with his estimates of the probabilities developed in step 2;
4. Given the prevailing state of nature, some outcome occurs as a result of executing alternative k. Remember that the DM is only probabilistically knowledgeable about which outcome will occur.
5. The resultant outcome is now fed back to the DM in accordance with Figure 1, permitting the DM to upgrade his assessment of the situation.
6. On the basis of this feedback information the DM updates his estimate of EV_k and his assessment of the probabilities of the outcomes for that alternative. Note that the DM's assessment of the decision situation and all the

values thereof are a function of his previous experience and the data obtained from the decisions he has made.

7. Go back to step 2, update the P's using the new EV_k and continue.

Step 6 in the above characterization of DM activity, using feedback data to update the DM's state of knowledge (or decision state), is what we call "relational learning". This learning can be considered to be effected by making row changes to the W matrix. Or, in a more macroscopic sense, we simply alter the current EV assessment corresponding to the alternative executed. In either case (macro or micro) the end result is an updated DM estimate of EV_k. No other EV's are altered since only one c.o.a. has been executed; the feedback data clearly relates only to the alternative executed. Of course from equation (2) we note that all his probabilities are altered due to a change in one EV and/or a change in his confidence.

Learning can be modeled by adjusting the DM's expected value of the alternative just executed. The value of the observed outcome is averaged into the DM's EV approximation in some way. A number of learning rules could be used to update the EV's. We have, for convenience, chosen the following learning rule inasmuch as it does seem to be descriptive of the actual process. In this situation c.o.a. k has been executed at time t resulting in the actual value $V_k(t)$ occurring. The DM then updates his old estimate of the expected value for c.o.a. k by

$$EV_k(t+1) = [1-\lambda_k(t)]\, EV_k(t) + \lambda_k(t)\, V_k(t), \quad 0 \leq \lambda \leq 1. \tag{3}$$

That is, the expected value of alternative k at time $t+1$ is simply a weighted average of the expected value at time t and the actual outcome value obtained at time t, $V_k(t)$.

The learning parameter λ should be a decreasing function of time and confidence. For instance, a possible definition of $\lambda_k(t)$ might be:

$$\lambda_k(t) = 1/[(C_k+1) \times \textit{number of trials for c.o.a. } k]. \tag{4}$$

Variable C_k is the DM's confidence factor in his estimation of EV_k and goes from 0 to infinity.

Note that when the DM has little confidence in his estimates, or when he has little data on which to base an estimate, his learning will be large (almost one) and when his confidence is high the learning will be small (almost zero).

Two important aspects of this learning algorithm should be noted. First, it should be clear that learning causes a chain-reaction: altering a row of the *W* matrix alters the respective *EV* which in turn alters the DM's probabilities for selecting an alternative (by our selection rule).

Second, the DM's confidence plays an important role in both learning and selection. Although the parameter, confidence, is used in both equation (2) for selection procedure and equation (4) for learning, the two may differ inasmuch as they describe different phenomena. Furthermore, in equation (2) the variable C is an *overall* confidence in the entire situation, whereas in equation (4) the variable C_k applies to the confidence that a DM has in his knowledge of the outcome of a particular c.o.a. Further research on this matter, both conceptual and experimental, is underway.

The Basic Information Measure

Given our model of information flow and analysis, we can now develop definitions of important terms which can then be quantified and measured and which are of course consistent with the basic theory. We first define our fundamental measure of information from the standpoint of effectiveness. All levels of information have one important point in common: information reduces some uncertainty. Uncertainty in the effectiveness sense should relate to the DM's choice of a c.o.a. How certain is the DM about which alternative should be chosen?

The uncertainty in choosing a course of action clearly relates directly to the DM's probabilities of executing each course of action. For example, if all of the probabilities are the same, then the DM is totally uncertain which alternative should be selected. At the other extreme, if the DM is completely certain as to his course of action then for some i, $P(a_i) = 1$ and all of the other probabilities are zero.

Thus we choose the *variance* of the P's as a basis for our measure of information. Note that the variance is zero when all of the P's are the same and a maximum when one P is unity and the others are zero.

The mean square variance of the probabilities, $\sigma^2(P)$, is by definition:

$$\sigma^2(P) = \sum_{i=1}^{m} [P(a_i) - \mu(P)]^2/m, \qquad (5)$$

and the mean, $\mu(P)$, is:

$$\mu(P) = \left[\sum_{i=1}^{m} P(a_i)\right]/m = \frac{1}{m}, \qquad (6)$$

where m is the number of viable courses of action.

Any basic measure of information should possess a number of fundamental properties. First, it should be defined in terms of some fundamental unit of measure. It should be at a maximum when the variance of the elements is maximal, and it should be at a minimum when the variance is minimal. These bounds should be well defined. The measure should be indifferent to the order of consideration of the elements. It must be stable; i.e., the consideration of an additional c.o.a. of low probability should not have a significant effect upon the measure. Finally, the measure should be sequentially additive. Whether applied once to the entire set of elements or respectively to a set of mutually exclusive and exhaustive subsets, the measure should yield the same result. A measure of information which possesses the aforementioned properties would provide a sound basis for our theory of effective information.

We have indicated that any effectiveness measure of information must relate to the variance of the DM's probabilities of selection of the alternatives. However, it is not the variance directly but the normalized variance which is important. The simplest way to consider the variance of a population in a normalized form is to divide it by the square of the mean, that is to consider the r.m.s. deviation in terms of units of the mean. It is clear that we must consider a normalized variance since a given variance about the mean will be much less significant when the mean is large than when the mean is small.

Hence we define our fundamental measure of information to be:

$$I = \sigma^2(P)/\mu^2(P). \qquad (7)$$

Since $\mu^2(P) = 1/m^2$ (see equation (6)), we observe that this measure relates the variance to m, the number of courses of action.

Referring to equations (5), (6), and (7), we see that

$$I = \sigma^2(P)/\mu^2(P) = \left[\sum_{i=1}^{m} \left(P(a_i) - \frac{1}{m}\right)^2/m\right] / \left(\frac{1}{m}\right)^2$$

$$= m \sum_{i=1}^{m} P(a_i)^2 - 1 .$$

This quantity possesses the desired properties for an information measure as we now demonstrate. Thus, we define this quantity to be *the amount of information in the decision state*, viz

$$I = m \sum_{i=1}^{m} P(a_i)^2 - 1 . \tag{8}$$

Properties of Information Measure

The information measure as defined possesses most of the desired properties of a fundamental measure of information effectiveness.

This quantity has a minimum of zero when all the $P(a_i)$'s are equal to $\frac{1}{m}$ (pure chance). This is complete uncertainty. The quantity has a maximum of $m - 1$ in the case of complete certainty where one of the $P(a_i)$'s is one and the others are zero.

When there are only two possible courses of action, the quantity I will assume values from zero to one. It will be equal to one under condition of certainty, i.e., when the probability of choosing one course of action is one and the other probability is zero. Accordingly, we will define the unit of information in terms of a deterministic two-choice situation. This unit we call a *binary choice unit*, or b.c.u.

When there are m possible courses of action, then the maximum amount of information from equation (8) is seen to be $m - 1$ b.c.u.'s. This is in agreement with a well-known combinatoric principle that a minimum of $m - 1$ deterministic choices from pairs of alternatives is required when there are m alternatives to consider. More explicitly, if $m - 1$ choices are required and the maximum amount of information in each choice is one, then the maximum amount of information is $m - 1$. Analogously the minimum amount of information is zero.

Clearly the measure is indifferent to the order in which the elements are considered. Since addition is associative, the P's can be summed in any order yielding a unique result. Hence the alternatives can be considered in any order.

The consideration of additional courses of action of low probability have an insignificant effect upon the measure. This is seen from equation (8) inasmuch as the sum of the P^2 will be essentially unchanged with the addition of such a c.o.a. However, the m will change by one and thus the relative information change is proportional to $\frac{1}{m}$. Thus the addition of a new c.o.a. does not *significantly* alter I for reasonable size m.

An additional question to be considered is the additivity of our measure. The measure itself clearly does not possess simple linear additivity since it is based on a nonlinear distribution, e.g.. the variance. This must be the case inasmuch as the probability of choosing any course of action must depend on *all* the other courses of action. However, the measure does have an important additivity property involving expected values.

The probability of choosing any course of action is a function of the decision-maker's expectation of all the values as we have seen in equation (2). The values (or the expectation of the values) on the other hand are independent of each other. We show below that the values of subsets of the courses of action are indeed additive and can be considered separately. Thus in going from the values to the probabilities to information we show an important additivity property. Hence the decision-maker's estimate of performance (defined in equation (12) below) for disjoint subsets A, B of the courses of action has the following additivity property:

The DM's estimate of performance with sets (A, B) equals his estimate of performance for set (A) multiplied by the probability of selecting set (A) plus his estimate of performance for set (B) multiplied by the probability of selecting set (B). Thus in order to add information in disjoint sets we must first consider the fundamental values.

The proof of this statement follows:

Consider set S of alternatives $a_1, a_2, \ldots, a_m$ and its associated EV_i and $P_S(a_i)$ for $i = 1, 2, \ldots, m$. Then by definition from equation (12) below

$$DP_S = \sum_{i=1}^{m} P_S(a_i)\, EV_i \, .$$

Now consider a partitioning of S into disjoint subsets A and B whose union is S. As order of evaluation for DP_S is unimportant, rename the alternatives so that $a_1, a_2, \ldots, a_k \in A$ and $a_{k+1}, a_{k+2}, \ldots, a_m \in B$. Now consider computing DP_A: note that $\sum_{i=1}^{k} P_S(a_i) \neq 1$, since $k < m$. If only these alternatives are to be considered then they must be factored uniformly so that they sum to 1. Hence we set $P_A(a_i) = P_S(a_i) / \sum_{j=1}^{k} P_S(a_j)$, which uniformly increases the probabilities for alternatives in $A \ni \sum_{i=1}^{k} P_A(a_i) = 1$.

Thus $DP_A = \sum_{i=1}^{k} P_A(a_i)\, EV_i = \sum_{i=1}^{k} \left[P_S(a_i)\, EV_i \Big/ \sum_{j=1}^{k} P_S(a_j) \right]$.

A similar term is derived for set B; i.e.,

$$DP_B = \sum_{i=k+1}^{m} P_B(a_i)\, EV_i = \sum_{i=k+1}^{m} \left[P_S(a_i)\, EV_i \Big/ \sum_{j=k+1}^{m} P_S(a_j) \right].$$

If a DM now wishes to combine these two evaluations to consider all m alternatives collectively, then it is only natural to consider the sets A and B as two alternative sets with associated probabilities of selection $P(A)$ and $P(B)$ and estimated values DP_A and DP_B. Thus,

$DP_{A \cup B} = P(A) \times DP_A + P(B) \times DP_B$. But

$P(A) = \sum_{i=1}^{k} P_S(a_i)$ and $P(B) = \sum_{i=k+1}^{m} P_S(a_i)$. So substituting $P(A)$, $P(B)$, DP_A and DP_B into our equality, we have

$$DP_{AUB} = \sum_{i=1}^{k} P_S(a_i) \left[\sum_{i=1}^{k} P_S(a_i)\, EV_i \Big/ \sum_{j=1}^{k} P_S(a_j) \right] +$$

$$\sum_{i=k+1}^{m} P_S(a_i) \left[\sum_{i=k+1}^{m} P_S(a_i)\, EV_i \Big/ \sum_{j=k+1}^{m} P_S(a_j) \right]$$

$$= \sum_{i=1}^{m} P_S(a_i)\, EV_i = DP_S \,.$$

QED

Amount of Information in a Data Set

A measure of the amount of information in a data set or message can be arrived at by computing the difference in the amount of information in the decision state after and before receipt of the data. That is, the amount of information is determined by considering the impact these new data have on the decision-maker's decision state. In symbolic terms, *$I(D)$, the amount of information in data set D,* is

$$I(D) = I_{t+1} - I_t \,, \tag{9}$$

where I_{t+1} and I_t are the amounts of information in the decision state after and before receipt of the data set.

It should be noted that the amount of information in a data set may be either positive or negative. In general, positive information sharpens or refines the decision-maker's understanding of the situation in that it either reduces the number of structural components in the model or reduces the dispersion in one or more of the various probability distributions in the model. Negative information, on the other hand, either increases the number of structural components (e.g., the addition to the model of a previously unknown alternative or outcome) or increases the dispersion in the various distributions. Negative information, despite a possible connotation of the term, does represent information that is of significance to the decision-maker.

This measure will in general vary with time; it will in general differ with different DM's and with different situations. This variation is consistent with the fact that data cannot be evaluated out of context. For instance, a given document which is relevant to a scientist working on a problem in nuclear physics is probably non-relevant to an economist. Knowing next month's projected output is of value to a DM worried about finding buyers or projecting future output, but after that month has passed the DM will use the actual output data, not the old projections. Lastly, each DM is different and thus may react differently to the same data; hence a different $I(D)$. This is a measure personal to the DM that is time and situation dependent.

Related Measures

We have now defined a quantitative measure of the *amount* of information in terms of the fundamental unit, *b.c.u.* From an effectiveness standpoint, we are further concerned with the *value* of the information. In order to establish a measure of the value of information, we first must define a term which describes the capability of a decision maker to achieve his goals. Such a term we will call decision-maker effectiveness or *DME*. We can then define the *value* of information in terms of the corresponding change in the *DME*. This will then be a measure of the effectiveness with which the DM uses the information available to him.

We will define *DME* in terms of the average expected performance of the DM at any given time. The DM executes the c.o.a.'s in accordance with his probability distributio. The *average* outcome of alternative i is EV_i^*, the actual expected value. Thus, average DM performance, *AP*, can be expressed as:

$$AP = \sum_{i=1}^{m} P(a_i)\ EV_i^* \ . \tag{10}$$

If the DM makes a series of decisions with these probabilities, then his performance, *on the average*, will be that projected by *AP*. Actual DM performance at time t is a specific value $V_k(t)$ which is obtained from the alternative executed, a_k. Note that $V_k(t)$ will in fact be one of v_{kj}'s associated with alternative k as seen from observation of Figure 2. The values obtained by execution of c.o.a., a_k, will vary according to the probabilities defined by the starred matrix and will vary about the mean EV_k^* with a variance determined by the values and their probabilities. Observation of each row of the value matrix indicates the extreme values.

Because of the fluctuations possible in *actual* DM performance, we define the decision-maker effectiveness in terms of the *average* DM performance as compared to the maximum possible DM performance. The maximum possible DM performance occurs when $max\ EV_k^*$ is chosen with certainty. Thus

$$DME = AP/max(EV_k^*),$$

$$DME = \sum_{i=1}^{m} P(a_i)\ (EV_i^*)/max(EV_k^*). \qquad (11)$$

DME is a dimensionless unit which goes from zero to one and tells us how well the DM is able to meet his goals. Note that *AP* is clearly additive in the same way as we showed above for *DP*, the DM's estimate of his performance.

A rational DM should increase his *DME* with time on the average, as $max\ (EV_k^*)$ is constant and average performance $(\sum_i P(a_i)\ EV_i^*)$ should increase as feedback data enhances the DM's probabilities of choice, $P(a_i)$. We see that if $P(a_k) = 1$ for the maximum valued EV_k^*, then $DME = 1$, the maximum.

Of course, the DM does not know the EV^*'s with certainty. He can only approximate them probabilistically. Therefore, the DM can only estimate his *DME* and, in fact, his best approximation to his performance is given by the term we define to be the DM expected performance, *DP*. That is

$$DP = \sum_i P(a_i)\ EV_i\ , \qquad (12)$$

where EV_i is the DM's estimate of EV_i^*. As the DM becomes more expert, his *DP* approaches *AP*.

The DM does not know the $max\ EV^*$ and therefore he can only estimate his *DME* from equation (12). This would take the form of *DP* divided by his estimate of the maximum (EV_i). Every decision maker has a perception of his own effectiveness for a given situation. We are interested empirically in relating the DM's estimate of his effectiveness to that calculated from our equations above for *DME* and *DP*. We are planning both simulations and experiments to determine the similarity of these two terms.

Value of Information

Decision-maker effectiveness, as defined in equation (11), provides the basis for our measure of the *value* of information, $Q(D)$. As our measures are concerned with the effectiveness level of information, it is natural to define the value of information to be related to its effect upon the performance of the DM. Did the new data increase or decrease the DM's performance, and by how much?

Because of the variations possible in actual DM performance at a given time, we believe it is only meaningful to consider the performance on an *average* basis. This, of course, is accompanied with a mean square deviation, σ^2 about the average. Hence $DME_{t+1} - DME_t$ would show the change in DM effectiveness (see equation 11) due to information received. Thus we define $Q(D)$, *the value of the information in data set D*, as:

$$Q(D) = DME_{t+1} - DME_t \,. \tag{13}$$

The measure DME is well defined and ranges from 0 to 1; hence our value measure $Q(D)$ ranges from -1 to +1. Data that results in decreased DM performance are of negative information value; if they do not effect performance, then they are of no value, and $Q(D) = 0$.

The value measure $Q(D)$ is described in terms of a basic unit: DME. It has prescribed upper and lower bounds that are consistant with the intended meaning of $Q(D)$. The measure is associative and additive if the AP is considered directly. Thus, this measure, $Q(D)$, possesses all of the properties we require of a fundamental information measure.

Hence we have developed two fundamental metrics which measure *information amount*, $I(D)$, and *information value*, $Q(D)$. While the quantity $I(D)$ shows the change in the DM's plans for dealing with the situation, the quantity $Q(D)$ defines the value accrued from the DM's change in probabilities. These two measures are thus distinct, yet they are interdependent. However, the relationship between them is quite complex. Information amount deals with the DM's actions; information value relates to DM performance. Further research will determine under what assumptions general relationships can be guaranteed.

Summary of Definitions

In this paper we have developed a number of important definitions characterizing information and decision-making. This section reviews those definitions so that they can be considered together in a meaningful manner. We list each term, describe its meaning, and repeat its definition as given previously.

a) Expected Value of a c.o.a., a_i: the expected value of an alternative is the weighted average of the outcome values.

$$EV_i = \sum_{j=1}^{m} w_{ij} v_{ij} . \quad (1)$$

b) Selection Rule for c.o.a.'s: the probability of selecting alternative i is proportional to its relative expected value raised to the DM's confidence.

$$P(a_i) = (EV_i)^C / \sum_{k=1}^{m} (EV_k)^C . \quad (2)$$

c) Learning Rule for expected value of c.o.a. (a_k): the new expected value of alternative k, given the outcome value of a prior execution, is a weighted combination of the old expected value and the observed outcome value.

$$EV_k(t+1) = [1-\lambda_k(t)] EV_k(t) + \lambda_k(t) V_k . \quad (3)$$

d) Information in the Decision State: the amount of information in the decision state of the DM in b.c.u.'s is

$$I = m \sum_{i=1}^{m} P(a_i)^2 - 1 . \quad (8)$$

e) Information Amount: the amount of information in a data set D is defined as the change in the information of the DM's decision state after and before receipt of the data.

$$I(D) = I_{t+1} - I_t . \quad (9)$$

f) DM Actual Performance: actual DM performance at time t is the actual value of outcome j, which is obtained from the alternative executed, a_k:

$$DM\ actual\ performance\ (t) = V_k(t) = v_{kj}\ .$$

g) DM Average Performance: the average performance of a DM is the weighted average of the actual expected outcome values.

$$AP = \sum_{i=1}^{m} P(a_i)\ EV_i^*\ . \tag{10}$$

h) Decision-maker Effectiveness: the effectiveness of a DM is the ratio of his average performance to the maximum expected outcome value.

$$DME = \sum_{i=1}^{m} P(a_i)\ (EV_i^*)/max(EV_k^*)\ . \tag{11}$$

i) DM Expected Performance: the DM's estimate of his average performance is the weighted average of his expected outcome values.

$$DP = \sum_{i=1}^{m} P(a_i)\ EV_i\ . \tag{12}$$

j) Information Value: the value of the information in a data set D to a given DM is defined to be his change in effectiveness.

$$Q(D) = DME_{t+1} - DME_t\ . \tag{13}$$

Conclusions and Further Research

The fundamental goal of this research is to develop a comprehensive, usable theory of information at the effectiveness level. Thus far we have defined a detailed model for information flow and analysis. We have suggested that most decision-making situations can be modeled within this formulation. We have developed an underlying mathematical framework to define the decision state of the decision maker: matrices W and K. From this framework we have algorithmically defined procedures to model, we believe realistically, DM data assimilation, DM selection, and DM execution of alternatives. The entire model implies iteration many times so that the DM can better estimate the decision state and various associated parameters. With this framework we believe that we can accurately describe the use of information in an effectiveness sense and the role of information in the total decision process.

We have defined herein a number of important information-related measures. Two of these of particular importance are the amount of information *I(D)* and value of information *Q(D)*, which quantify the information in a data set *D*. Other measures that can be derived from our model include the amount of information in the decision state *(I)*, decision-maker effectiveness *(DME)*, decision-maker expected performance *(DP)* and DM average performance *(AP)*. These measures provide us with a quantitative bssis for analyzing and defining information flow and for identifying limitations of this flow.

Continuing research involves establishing relationships among these quantities and the significance of each to the information flow process. We are seeking generalized information relationships in an effort to establish fundamental guidelines for information flow, analysis, storage, and processing. In addition, we feel that generalized rules for making decisions under various conditions - a decision calculus - will emerge from this model as well.

We are planning to apply this theoretical development to practical situations and indicate how the quantities can be defined, measured, and used in a practical way. In particular, we are developing examples using a bibliographical retrieval system, a production control situation, and a general economic model.

Of course, of primary concern is the question of validation in an empirical sense of our theory and measures. We have developed a simulation model to verify some of our basic measures and procedures and to determine their value. Two versions of the simulation model have been implemented in FORTRAN: a batch model and an interactive model. The batch model is running on an IBM 370/168 and is being used for running long repetitive decision situations with many trials. Aggregate data from test cases are being studied to determine if the results are consistent with the theory and to find further relationships between the basic quantities defined in our information flow model. The interactive model is running on a PDP-10 and enables direct input to be made by a human decision maker. We plan to conduct actual experiments using a document collection and real decision makers in order to establish the various paramaters and their validity.

References

1. Yovits, M.C., and Ernst, R.L., "Generalized Information Systems: Consequences for Information Transfer," *People and Information*, ed. by H.P. Pepinsky, Pergamon Press, New York, 1969.

2. Whittemore, B.J., and Yovits, M.C., "A Generalized Conceptual Development for the Analysis and Flow of Information," *JASIS*, May-June 1973, Vol. 24, No. 3, 221-231.

3. Yovits, M.C., and Abilock, J.G., "A Semiotic Framework for Information Science Leading to the Development of a Quantitative Measure of Information," *Information Utilities; Proceedings of the 37th ASIS Annual Meeting*, Vol. 11, Atlanta, 1974.

4. Shannon, C.E., and Weaver, W., *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana, Ill., 1949.

5. Weiner, N., *The Human Use of Human Beings*, Houghton Mifflin Co., Boston, 1950.

6. Carnap, R., and Bar-Hillel, Y., "An Outline of a Theory of Semantic Information," Technical Report #247, Research Laboratory of Electronics, MIT, 1952.

7. Winograd, T., "Understanding Natural Language," *Cognitive Psychology*, 3, 1972.

8. Zunde, P., "Information Measurement and Value," *Third Annual Symposium of the American Society for Cybernetics*, Gaithersburg, Md., October 1969.

9. MacKay, D.M., *Information, Mechanism, and Meaning*, MIT Press, Cambridge, Mass., 1969.

10. Ackoff, R.L., "Towards a Behavioral Theory of Communication," *Management Science*, 4 (1958), 218-234.

11. Marshak, J., "Problems in Information Economics," *Management Controls*, ed. by C.P. Bonini, et al, McGraw-Hill, New York, 1964, 38-74.

12. MacKay, D.M., "Quantal Aspects of Scientific Information," *Philosophical Magazine*, Vol. 41, No. 314, March, 1950, 289-311.

13. Cherry, C., *On Human Communication*, MIT Press, Cambridge, Mass., 1966.

14. Goffman, W., "On the Dynamics of Communication," *AAAS Conference Proceedings*, Denver, Colorado, February 1977.

15. Morris, W.T., *The Analysis of Management Decisions*, Richard D. Irwin, Inc., Homewood, Ill., 1964.

16. Raiffa, H., *Decision Analysis, Introductory Lecture on Choices Under Uncertainty*, Addison-Wesley, Reading, Mass., 1971.

17. Schlaiffer, R., *Analysis of Decisions under Uncertainty*, McGraw-Hill, New York, 1969.

18 Radnor, R., "Normative Theory and Individual Decision - An Introduction," *Decision and Organization*, ed. by C.B. McGuire and R. Radnor, North Holland Publishing Co., Amsterdam, 1972.

19. Bandyopadhyay, R., "Information for Organizational Decisionmaking - A Literature Review," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-7, No. 1, January 1977.