Journal of the
Operations Research
Society of Japan

VOLUME 9 *April 1967* NUMBER 2

A BAYESIAN VIEWPOINT IN THE EVALUATION OF INFORMATION SYSTEMS*

U. O. GAGLIARDI

Dunlap and Associates, Inc. Darien, Connecticut, U.S.A.

(Received May 25, 1965)

Abstract

In this paper, the insufficiency of the classical statistical viewpoint in evaluating information systems, as exemplified by Shannon's theory of information, is identified. A decision theoretic definition of the effectiveness of an information system is given. The use of the decision theoretic formulation is illustrated in a simple numerical example.

1. Introduction

1.1. Information theory and the value of information systems

Information in general denotes a set of potential messages associated with a given channel or system of information. It can be viewed as something which informs us about the state of a given environment so that the uncertainty associated with such an environment can be expected to be reduced if not completely eliminated. Our desire to reduce the

* This research was completed with the support of the U.S. Air Force under Contract No. AF19(628)-2830.

uncertainty associated with the environment can be based on either purely intellectual motivations (e.g., to satisfy the curiosity of the mind) or the need for making decisions with the hope of achieving some economic goals. In situations where the economic motivation is a dominating factor, the value of an information system has to be assessed in terms of the expected gain in helping us to attain our goals. Since information is the output of an information system, we have to know the nature of our decision task and how the provided information is utilized in the decision process in order to determine the value associated with an information system.

Let Z be a random variable that assumes N values, denoted by $z_1, \cdots, z_N$, respectively; and let $P(z_i)$ be the probability of z_i. Then a statistical parameter called the "entropy" associated with the random variable Z, denoted by $H(Z)$, is defined as

$$H(Z)=-\sum_i P(z_i) \cdot \log P(z_i)$$

When the z's represent the various states of nature, the entropy measures, in some sense, the degree of uncertainty associated with the state of nature. When the z's represent the set of potential messages associated with an information channel, it is used by Shannon (Ref. 1) as a measure of "the amount of information."

Let $X=(x_1, \cdots, x_N)$ denote the set of N possible states of nature, and $Y=(y_1, \cdots, y_M)$ denote the set of M potential messages associated with an information channel. Let $H(X|y_j)=-\sum_i P(x_i|y_j) \cdot \log P(x_i|y_j)$, where $P(x_i|y_j)$ is the conditional probability of x_i given message y_j. The value $H(X|y_j)$ can be interpreted as a measure of the amount of uncertainty (about the state of nature) remaining after receiving the j-th message. Then the equivocation of the channel Y with respect to the source X, denoted by $H(X|Y)$, is defined as

$$H(X|Y)=\sum_j P(y_j) \cdot H(X|y_j)$$

The transmission rate of Y with respect to X, denoted by $R(Y, X)$, is defined as

$$R(Y, X) = H(X) - H(X \mid Y)$$

and the channel capacity, denoted by $C(Y)$, is defined as

$$C(Y) = \operatorname*{Max}_{x} R(Y, X)$$

Since the derivation of the transmission rate and the channel capacity does not take into consideration the user's utility structure, it is clear that it cannot adequately represent the value of an information channel to a particular user faced with a particular decision task-although under certain special assumptions concerning the user's utility structure, the channel capacity can be made to correspond to the value of an information channel to the user (Ref. 2).

Then what does the entropy associated with an information channel represent? According to Marschak (Ref. 3), since the entropy usually increases with the number of distinct potential messages, and the larger the number of distinct potential messages, the larger the number of symbols needed at a minimum to distinguish the messages, then the entropy more appropriately represents the cost of constructing an information instrument. This perhaps explains, at least partially, the fact that the entropy concept was first proposed by the people at Bell Lab. A producer of information instruments cannot hope to take into consideration the various needs of different users of its product; however, he does concern himself with costs associated with producing various information instruments.

1.2. Statistical decision theory

Statistical decision theory is concerned with the derivation of an optimal decision rule (in the face of uncertainty) based on

- The decision maker's utility structure.
- The decision maker's prior probability distribution over the states

of nature.

- An experiment which generates observations.
- A rule of revising the prior distribution upon receiving an observation.
- A definition of what constitutes an optimal decision rule.

In deriving such an optimal decision rule, one introduces a measure of performance over the set of all possible decision rules which is the expected value of some suitable function of the utilities. A natural application of this theory in evaluating information systems is to model the information system as a scattering process (an experiment) in which any particular state of nature can give rise to several possible messages (observations), and the value associated with an information system is obtained by computing the expected gain which results, using the optimal decision rule, from employing the system.

It then follows that an appropriate measure of effectiveness of a given information system to a particular user is the net gain in the expected utility, resulting from employing the information system as an aid in decision-making processes over and above the expected utility which results when no information system is employed. If the effectiveness of an information system is defined in this manner, it will be logically determined by, and thus consistent with, the utilities chosen and the statistical hypotheses concerning the prior uncertainty about the environment.

Since effectiveness-cost equilibrium will determine the design recommendations, we ultimately have in this theory a means to make design recommendations which are consistent with the aforementioned elements. On the other hand, if one would adopt, following a common practice, measures of performance of components as the basis for performing trade-off analyses, one may recommend a design which does not agree with user preferences. Such a design may, in fact, imply a utility structure which the user would not agree with, if described explicitly to him. In the worst case, the design so obtained may imply mutually contradictory statements of preference so that there is no utility structure which is consistent with the system design.

2. The Bayesian Effectiveness of Information Systems

2.1. The bayesian value of an information system

The usual model of decision making under uncertainty assumes that there are certain states of nature that are relevant to our decision, certain acts that are open to us for choice, and a utility index associated with each act-state pair.

Let X^i denote the i-th state of nature, $i=1, \cdots, N$; A^k denote the k-th act open to us, $k=1, \cdots, L$; and u_{ki} be the utility index assigned to the act-state pair $(A^k, X^i), u_{ki}=U(A^k, X^i)$.

An information structure can be most conveniently characterized as follows:

$$\begin{array}{c} \\ X^1 \\ \cdot \\ \cdot \\ \cdot \\ X^N \end{array} \begin{array}{c} \begin{array}{ccccc} Y^1 & \cdot & \cdot & \cdot & Y^M \end{array} \\ \begin{pmatrix} q_{11} & \cdot & \cdot & \cdot & q_{1M} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ q_{N1} & \cdot & \cdot & \cdot & q_{NM} \end{pmatrix} \end{array}$$

where $Y^j, j=1, \cdots, M$, is the j-th message transmitted to us by the information system, and $q_{ij}=P(Y^j|X^i)$ is the conditional probability of the k-th message given the fact that the true state of the nature is X^i.

A rule which assigns an act to each of the possible messages is called a decision rule. We shall denote it by $A=\alpha(Y)$.

The Bayesian decision rule implies the following assumptions: (1) There is a certain prior probability associated with each state of nature; we shall denote it by $P(X^i), i=1, \cdots, N$. (2) For each message observed, a posterior probability distribution over the states of nature can be derived by using Bayes theorem. Let $P(X^i|Y^j)$ denote the posterior probability of X^i given the fact that Y^j has been observed. Then,

$$P(X^i|Y^j)=P(X^i)\cdot P(Y^j|X^i)/\sum_r P(X^r)\cdot P(Y^j|X^r).$$

(3) Let $V(A^k|Y^j)=\sum_i P(X^i|Y^j)u_{ki}$ be the expected value of A^k given the fact that Y^j has been observed. Then the Bayesian decision rule says that for each message Y^j one should select the act $A=\hat{\alpha}(Y^j)$ such that

$$V[\hat{\alpha}(Y^j)|Y^j]=\max_k V(A^k|Y^j).$$

Let $P(Y^j)=\sum_r P(X^r)\cdot P(Y^j|X^r)$ be the probability of observing the j-th message given the prior probability distribution over X and the information system n. Then the Bayesian value of n is

$$\hat{V}(n)=\sum_{j=1}^{M} P(Y^j)\sum_{i=1}^{N} P(X^i|Y^j)\,u[\hat{\alpha}(Y^j), X^i]. \qquad (2.\ 1\text{-}1)$$

2.2. The bayesian effectiveness

As defined in Section 2.1, the Bayesian decision rule selects the act $A=\hat{\alpha}(Y^j)$ such that

$$V[\hat{\alpha}(Y^j)|Y^j]=\max_k V(A^k|Y^j)$$

i.e., such that

$$\sum_{i=1}^{N} P(X^i|Y^j)\,u[\hat{\alpha}(Y^j), X^i]=\max_k \left[\sum_{i=1}^{N} P(X^i|Y^j)\,u(A^k, X^i)\right]$$

Let us introduce the following notation:

$$U=\begin{pmatrix} u_{11} & \cdot & \cdot & \cdot & u_{1N} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ u_{L1} & \cdot & \cdot & \cdot & u_{LN} \end{pmatrix}$$

where $u_{ki}=U(A^k, X^i)$

$$Q=\begin{pmatrix} q_{11} & \cdot & \cdot & \cdot & q_{1M} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ q_{N1} & \cdot & \cdot & \cdot & q_{NM} \end{pmatrix}$$

where $q_{ij}=P(Y^j|X^i)$

$$P=\begin{pmatrix} p_{11} & \cdot & \cdot & \cdot & p_{1M} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ p_{N1} & \cdot & \cdot & \cdot & p_{NM} \end{pmatrix}$$

where $p_{ij}=P(X^i|Y^j)$

$$=\frac{P(X^i)\cdot P(Y^j|X^i)}{\sum_{r=1}^{N} P(X^r)\cdot P(Y^j|X^r)}$$

$$=\frac{P(X^i)\cdot P(Y^j|X^i)}{P(Y^j)}$$

The j-th column of P, denoted by $[P]_j$,

$$\begin{pmatrix} p_{ij} \\ \cdot \\ \cdot \\ \cdot \\ p_{Nj} \end{pmatrix}=\begin{pmatrix} P(X^i|Y^j) \\ \cdot \\ \cdot \\ \cdot \\ P(X^N|Y^j) \end{pmatrix}$$

is the conditional probability distribution over the states of nature if Y^j has been observed. It is then clear that the j-th column of UP, denoted by $[UP]_j$, is the set of expected utilities associated with various acts conditional on the occurrence of Y^j.

We shall now define the operator*. If $[B]$ represents a column vector

$$\begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_M \end{pmatrix}, \text{ then } [B]^* = \max_i \{b_i\}.$$

Let B be a matrix

$$\begin{pmatrix} b_{11} & \cdot & \cdot & \cdot & b_{1N} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ b_{M1} & \cdot & \cdot & \cdot & b_{MN} \end{pmatrix}$$

Then $B^* = ([B]_1^* \cdots [B]_N^*)$, where $[B]_j$ denotes the j-th column of B.

With the aid of the operator*, we can define the Bayesian decision rule as $\hat{\alpha}(Y^j) = A^{\hat{k}}$ such that $V(A^{\hat{k}}|Y^j) = [UP]_j^*$. Then the Bayesian value of an information system n is given by

$$\hat{V}(\kappa) = \sum_{j=1}^{M} P(Y^j)\, [UP]_j^*$$

where $P(Y^j) = \sum_{r=1}^{N} P(X^r) \cdot P(Y^j|X^r)$.

Let $\hat{E}(\kappa)$ be the Bayesian effectiveness associated with an information system κ. Then

$$\hat{E}(\kappa) = \hat{V}(\kappa) - \hat{V}(\kappa^0)$$

where κ^0 denotes the null information system.

Consider the k-th component of $[UP]_j$. It is

$$\sum_{i=1}^{N} U(A^k, X^i)\, P(X^i|Y^j) = \frac{1}{P(Y^j)} \sum_{i=1}^{N} U(A^k, X^i)\, P(X^i)\, P(Y^j|X^i).$$

Let $\bar{U}=UD$

where $D=\begin{pmatrix} P(X^1) & \cdot & \cdot & \cdot & 0 \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & P(X^N) \end{pmatrix}$ is a diagonal matrix whose diagonal elements ore $P(X^1), \cdots, P(X^N)$. Then the k-th component of $[UP]_j$ is simply $\frac{1}{P(Y^j)}[\bar{U}Q]_{kj}$, where $[\bar{U}Q]_{kj}$ is the kj-th element of $\bar{U}Q$ and

$$[UP]_j = \frac{1}{P(Y^j)}[\bar{U}Q]_j.$$

Since $\left(\frac{1}{P(Y^j)}[\bar{U}Q]_j\right)^* = \frac{1}{P(Y^j)}[\bar{U}Q]_j^*$, it follows that

$$\begin{aligned} V(\kappa) &= \sum_{j=1}^{M} P(Y^j)\,[UP]_j^* \\ &= \sum_{j=1}^{M} P(Y^j)\left(\frac{1}{P(Y^j)}[\bar{U}Q]_j\right)^* \\ &= \sum_{j=1}^{M} [UQ]_j^* \\ &= (\bar{U}Q)^*\,\xi \\ &= [UDQ]^*\,\xi \end{aligned}$$

where ξ is a column vector with M components whose values are all equal to 1.

Let P_0 be the P matrix associated with the null information system κ_0. Since $[UP_0]_j$ is the weighted average—with the weights $\{P(X^i)\}$—of the columns of U and is independent of j, we shall denote it by $[U_0]$. Then

$$V(\kappa_0)=\sum_{j=1}^{M} P(Y^j)\,[U_0]^*=[U_0]^*$$

and
$$E(\kappa)=(\bar{U}Q)^*\,\xi-[U_0]^*=[UDQ]^*\,\xi-(U\pi)^*$$

where π is the vector whose i-th component is the prior probability of of $X^i, P(X^i)$.

The above formulation assumes discrete sets of states, messages, and acts. For more general formulations, consult References 4 and 5.

3. Illustrative Example

In this section, the evaluation approach developed above will be used in connection with a simple information system to illustrate its application.

3.1. Alternate system concepts

Let us assume that we are interested in evaluating the relative effectiveness of three alternative information system concepts for use in an air defense task against threats which have the following characteristics;

Table 1.

State Vector Number	State Vector Elements	INTERPRETATION		
	$\{X_1, X_2, X_3\}$	Altitude	Speed	Threat
0	0, 0, 0	No Reading	Slow Speed	No Threat
1	0, 0, 1	No Reading	High	No Threat
2	0, 1, 0	Low	Slow	Bomber
3	0, 1, 1	Low	Fast	Light Plane
4	1, 0, 0	Medium	Slow	Bomber
5	1, 0, 1	Medium	Fast	Lighl Plane
6	1, 1, 0	High	Slow	No Threat
7	1, 1, 1	High	Fast	Missile

The alternative information system concepts are characterized by the

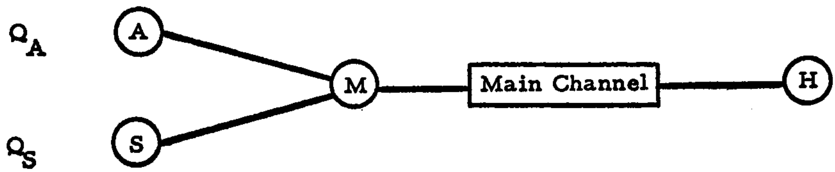


Fig. 1.

following system assumptions:

System 1. The general system configuration is shown in Figure 1, and the system assumptions are:

Altitude sensor —noisy with error rate to neighboring unit ρ_a.
Speed sensor —noisy with error rate ρ_s.
Sensor channels—perfect, no noise, $r=0$ ($Q_{SC}=1$)
Main channel —binary, symmetric, noisy with orror rate $r=.1$. High order terms are neglected in calculating QC.

System 2. The same system as System 1 except for the main channel. The main channel is noisy with an error rate $r=.1$. Single errors are eliminated by the adoption of a single error correcting code.

System 3. The same system as System 1 except for the main channel. The main channel is assumed to be perfect, $r=0(Q_C=1)$.

3.2. Decision Theoretic Formulation

The air defense task described above can be given a decision theoretic formulation in the following manner. Let us assume that the action set contains only three actions:

A_0—Do nothing.
A_1—Engage the enemy with a fighter.
A_2—Engage the enemy with a missile.

A utility structure which is a function of the targets and weapon system parameters can be defined as follows:

$$
U=\left\{\begin{array}{cccc}
X^0 & X^1 & X^2 & X^3 \\
0 & 0 & T_B & 0 \\
-C_{MF} & -C_{MF} & \begin{array}{c} P_{EB}V_B-C_{MF} \\ (1-P_{FB})\,T_B \end{array} & \begin{array}{c} P_{FF}V_F \\ -C_{MF} \end{array} \\
-C_M & -C_M & T_B-C_M & -C_M
\end{array}\right.
$$

$$
\left.\begin{array}{cccc}
X^4 & X^5 & X^6 & X^7 \\
T_B & 0 & 0 & T_M \\
\begin{array}{c} P_{FB}V_B-C_{MF} \\ (1-P_{FB})\,T_B \end{array} & \begin{array}{c} P_{FF}V_F \\ -C_{MF} \end{array} & -C_{MF} & T_M-C_{MF} \\
\begin{array}{c} \bar{P}_{MB}V_B-C_M \\ (1-\bar{P}_{MB})\,T_B \end{array} & \begin{array}{c} \bar{P}_{MF}V_F \\ -C_M \end{array} & -C_M & \begin{array}{c} V_M-C_M \\ (1-P_{MM})\,T_M \end{array}
\end{array}\right\}
$$

where the meaning and assumed value for each quantity in the above matrix is as follows:

- Cost of one fighter mission $C_{MF} = .1$
- Cost of one defensive missile $C_M = 1.0$
- Threat posed by an enemy bomber which penetrates the defenses $T_B = -20.0$
- Threat posed by an enemy missile which penetrates the defenses $T_M = -20.0$
- Value of destroying enemy bomber $V_B = 5.0$
- Value of destroying enemy missile $V_M = 3.0$
- Value of destroying enemy light plane $V_F = 2.0$
- Probability of kill for fighter against bomber engagement (any altitude) $P_{FB} = .75$
- Probability of kill of fighter against light plane engagement (any altitude) $P_{FF} = .5$
- Probability of kill for missile-bomber engagement for medium altitude bombers $\bar{P}_{MB} = .7$

- Probability of kill of missile-light plan engagement for medium altitude light planes $\bar{P}_{MF}=$ .7
- Probability of kill for missile-missile engagement $P_{MM}=$ .95

The substitution of these values in utility matrix yields the following matrix, which is used in this numerical example.

$$U=\{u_{ij}\}=\begin{pmatrix} 0 & 0 & -20.00 & 0 & -20.00 & 0 & 0 & -20.0 \\ -.1 & -.1 & -1.35 & .9 & -1.35 & .9 & -.1 & -20.1 \\ -1.0 & -1.0 & -21.00 & -1.0 & -3.5 & .4 & -1.0 & 1.0 \end{pmatrix}$$

For this example we will assume that the prior distribution is uniform. Then the final element required to complete the decision theoretic formulation is to compute the Q matrix which models the specific information system to be used. The computation of the Q model for the information system shown in Figure 1 can be performed in two steps.

The first step is to construct the Q model for the sensor subsystem. This is accomplished by combining the Q models for the speed and altitude sensors by means of a special commutative operation derived in Reference 4, according to the following formula:

$$Q_{\text{Sensor}}=Q_S * Q_A$$

where

$$Q_S=X_s^i \quad \begin{array}{c|cc|} & \multicolumn{2}{c}{Y_s^j} \\ & 0 & 1 \\ \hline 0 & 1-\rho_s & \rho_s \\ 1 & \rho_s & 1-\rho_s \end{array}$$

$$
Q_A = X_a^i \quad
\begin{array}{c|cccc|}
 & 00 & 01 & 10 & 11 \\
\hline
00 & (1-\rho_a) & \rho_a & 0 & 0 \\
01 & 1/2\rho_a & (1-\rho_a) & 1/2\rho_a & 0 \\
10 & 0 & 1/2\rho_a & (1-\rho_a) & 1/2\rho_a \\
11 & 0 & 0 & \rho_a & (1-\rho_a)
\end{array}
\qquad Y_a^j
$$

and ρ_s and ρ_a are respectively the error probabilities for the speed and altitude sensors. The altitude sensor model reflects the continuous nature of the initial measurement process (errors are assumed to occur only between adjacent levels). The Q_{Sensor} matrix can be shown to be:

$$
Q_{\text{Sensor}} = \left|\begin{array}{c:c} (1-\rho_s)Q_A & \rho_s Q_A \\ \hdashline \rho_s Q_A & (1-\rho_s)Q_A \end{array}\right|
$$

The second step in the derivation is to compute the system Q which is accomplished by using the equation:

$$
Q = Q_{\text{Sensor}}\, Q_{\text{Channel}}
$$

where the operation indicated is regular matrix product. The Q_{Sensor} matrix is identical for all three systems and is a function of the parameters ρ_s and ρ_a.

The Q_{Channel} matrix for the first system is given by

$$
Q_{\text{Channel1}} = \left|\begin{array}{cccccccc}
1-3r & r & r & 0 & r & 0 & 0 & 0 \\
r & 1-3r & 0 & r & 0 & r & 0 & 0 \\
r & 0 & 1-3r & r & 0 & 0 & r & 0 \\
0 & r & r & 1-3r & 0 & 0 & 0 & r \\
r & 0 & 0 & 0 & 1-3r & r & r & 0 \\
0 & r & 0 & 0 & r & 1-3r & 0 & r \\
0 & 0 & r & 0 & r & 0 & 1-3r & r \\
0 & 0 & 0 & r & 0 & r & r & 1-3r
\end{array}\right|
$$

where r represents the channel error probability and takes the value .10.

The Q_{Channel} for the second system is given by

$$Q_{\text{Channel}_2}=\frac{1}{1-3r(1-r)^2}\begin{vmatrix}(1-r)^3 & 0 & 0 & r^2(1-r) & 0 & r^2(1-r) & r^2(1-r) & r^3\\ 0 & (1-r)^3 & r^2(1-r) & 0 & r^2(1-r) & 0 & r^3 & r^2(1-r)\\ 0 & r^2(1-r) & (1-r)^3 & 0 & r^2(1-r) & r^3 & 0 & r^2(1-r)\\ r^2(1-r) & 0 & 0 & (1-r)^3 & r^3 & r^2(1-r) & r^2(1-r) & 0\\ 0 & r^2(1-r) & r^2(1-r) & r^3 & (1-r)^3 & 0 & 0 & r^2(1-r)\\ r^2(1-r) & 0 & r^3 & r^2(1-r) & 0 & (1-r)^3 & r^2(1-r) & 0\\ r^2(1-r) & r^3 & 0 & r^2(1-r) & 0 & r^2(1-r) & (1-r)^3 & 0\\ r^3 & r^2(1-r) & r^2(1-r) & 0 & r^2(1-r) & 0 & 0 & (1-r)^3\end{vmatrix}$$

where r is the channel error probability and the null entries reflects the single error correcting capability of the channel and $\frac{1}{1-3r(1-r)^2}$ is a suitable normalizing factor.

The Q_{Channel} for System 3 is simply an 8×8 identity matrix.

3.3. Alternative system evaluation

The use of the decision theoretic approach to evaluate the alternative system concepts results in the effectiveness values shown in Figures 2, 3 and 4 presented at the end of this section. The effectiveness values are graphed as functions of ρ_a and ρ_s. From an examination of these cross sections of the effectiveness function, it is apparent that a 10 per cent error rate in the channel is much more detrimental to system performance than a 10 per cent error rate in both sensors. In fact, the effectiveness of a system with perfect sensors and a non-error correcting channel with a 10 per cent error rate is 1.97 (origin of Figure 2) while the effectiveness of a system with a perfect channel and sensors with 10 per cent error rates is 2.36. This indicates that the accuracies of various components of the same system are not required to be identical in an optimal design and that the only rational way to establish the appropriate accuracies is by determining an optimum cost effectiveness design point. The comparison of Figures 2 and 3 permits the system designer to evaluate the effect of an error correcting scheme on system performance and thus to estimate whether it is worth the additional cost.

Finally, let us remark that the simplicity of the above example results from a desire to illustrate the use of the decision theoretic formulation with a minimum of effort. In actual practice, the user may have to construct much more complex models to analyze a realistic information system. Considerably more complex models than the Q derived in this example can be handled by utilizing the resources of modern algebraic languages such as ALGOL, FORTRAN, or NPL without excessive cost. The availability of these languages makes it feasible to use this technique for the evaluation of realistically complex systems. To demonstrate this point the author, under the Air Force contract which supported this research, is applying the technique to the study of the acquisition and threat evaluation subsystems of the Ballistic Missile Early Warning System (BMEWS).

REFERENCES

[1] Shannon, C. "A Mathematical Theory of Communication." *Bell System Technical Journal*, 1948.

[2] Kelly, J. "A New Interpretation of Information Rate." *Bell System Technical Journal*, 1956.

[3] Marschak, J. "Remarks on the Economics of Information." Contribution to Scientific Research in Management, 1960.

[4] Gagliardi, U.O., L. Holt, and C. Ying, "Mathematical Programming Techniques for Information System Design." Technical Documentary Report No. ESD-TDR-64-530, Dunlap and Associates, Inc., 1964.

[5] Raiffa, H. and R. Schlaifer, *Applied Statistical Decision Theory*. Harvard Business School, Division of Research, Boston, 1961.

SYSTEM 1

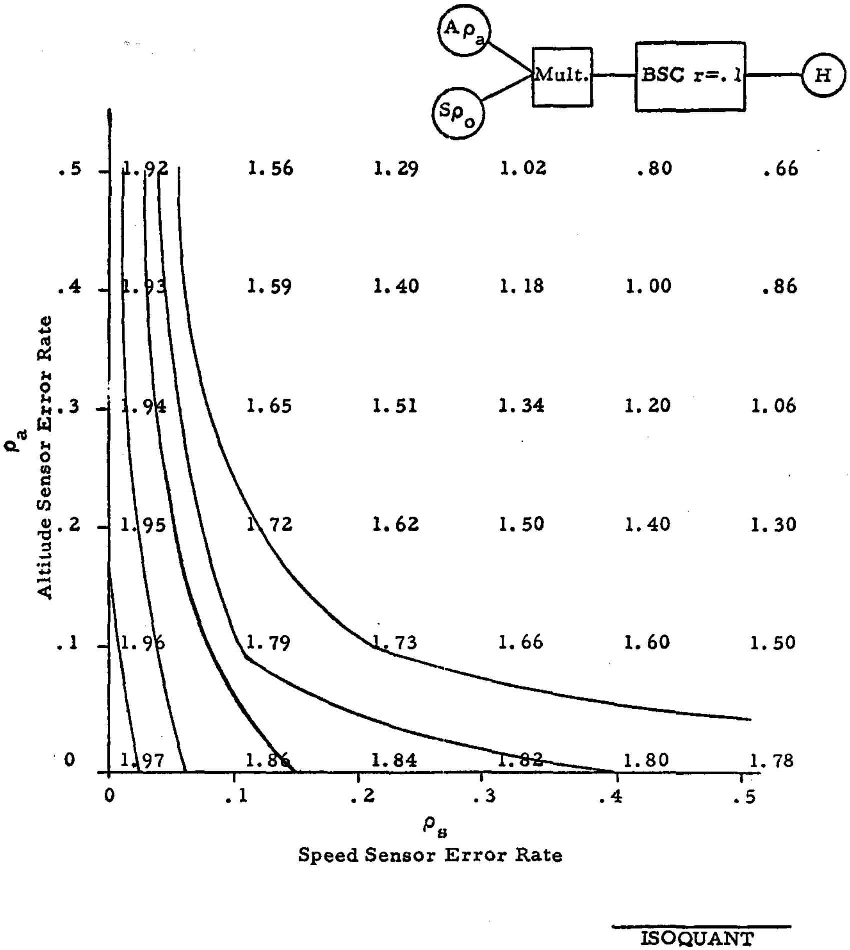


Fig. 2.

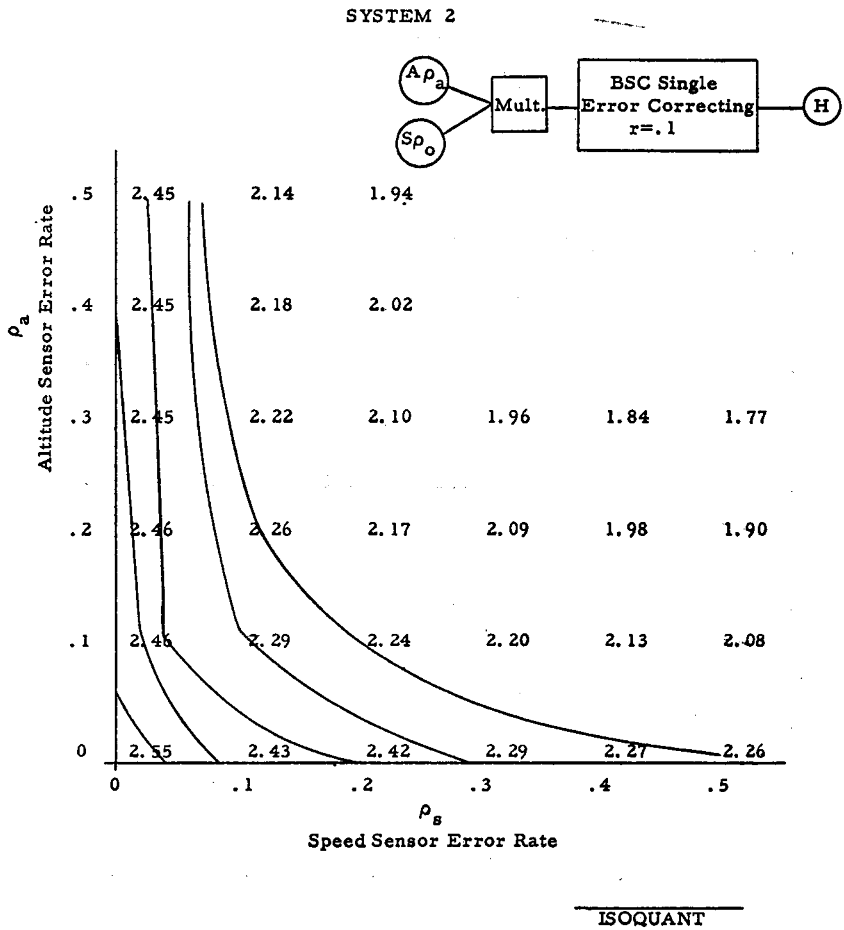
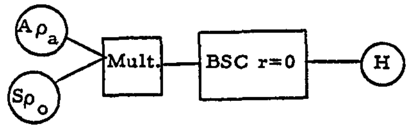

SYSTEM 2
$A\rho_a$
$S\rho_o$
Mult.
BSC Single
Error Correcting
r=.1
H
ρ_a
Altitude Sensor Error Rate
.5
.4
.3
.2
.1
0
2.45
2.14
1.94
2.45
2.18
2.02
2.45
2.22
2.10
1.96
1.84
1.77
2.46
2.26
2.17
2.09
1.98
1.90
2.46
2.29
2.24
2.20
2.13
2.08
2.55
2.43
2.42
2.29
2.27
2.26
0
.1
.2
.3
.4
.5
ρ_s
Speed Sensor Error Rate
ISOQUANT

Fig. 3.

SYSTEM 3



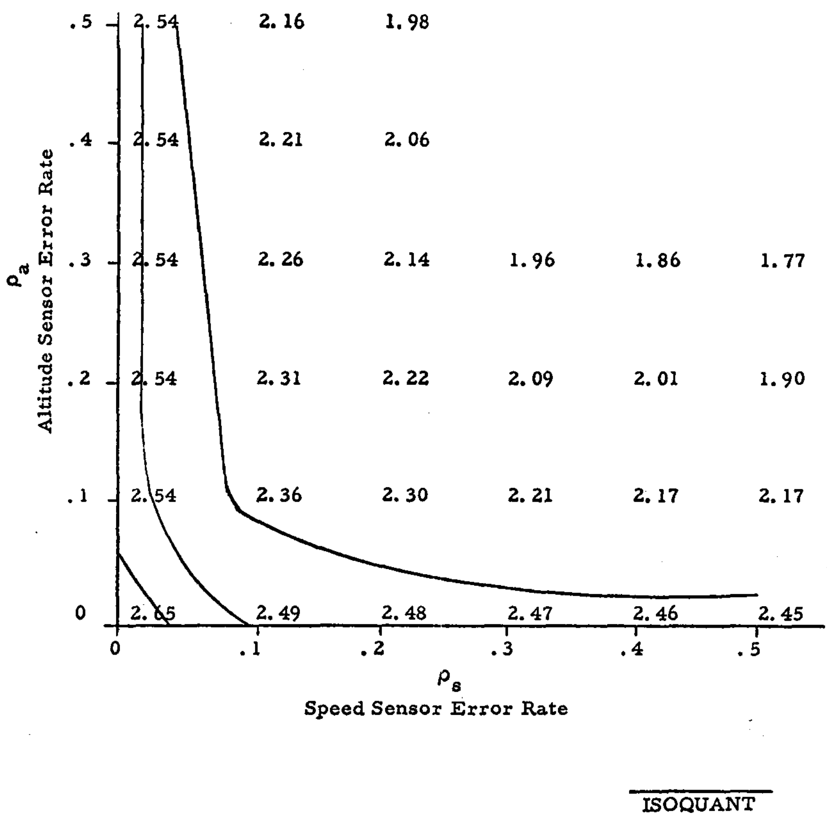


Fig. 4.