Concepts, Theory, and Techniques

ORTHOGONAL INFORMATION STRUCTURES—A MODEL TO EVALUATE THE INFORMATION PROVIDED BY A SECOND OPINION

Niv Ahituv
Faculty of Management, Tel Aviv University, Ramat-Aviv, Tel Aviv, Israel 69978
Boaz Ronen
Graduate School of Business Administration, New York University, New York, NY 10006

ABSTRACT

The paper discusses the value of information when a number of independent sources provide information related to a common set of states of nature. The starting point is the information economic model of information structures. The model is augmented to represent independence of informational sources by means of orthogonality of the information structures.

A new mathematical operator, orthogonal product, is defined and its properties are probed. It is shown that this operator maintains some mathematical properties such as closure, association, unity element, null element, and so forth. It is demonstrated how the orthogonal product represents the notion of multisource information.

The paper proves that an orthogonal product is generally more informative than its multipliers, namely, if cost is not considered a constraining factor, then there is a nonnegative value to obtaining a second opinion. An appendix to the paper expands this result to a case of partially dependent signals. The paper concludes with a numerical example and a discussion of the model's applicability for practical problems such as cost estimates.

Subject Areas: Decision Analysis, Information Economics, and Statistical Decision Theory.

INTRODUCTION

The information economic approach to information evaluation is based on an information structure (IS) model developed by Marschak [9], McGuire and Radner [11], Demski [5], and others and later expanded in a number of articles (e.g., [1], [2], and [3]). The model portrays an information system as a stochastic (Markov) matrix of probabilities which transforms states of nature to signals. The decision maker (DM) must select the optimal decision rule under given values of a priori probabilities for states of nature, and given values of payoffs. The information economic model proposes a partial rank ordering of information structures by using Blackwell's Theorem [10].

The IS model basically adheres to a case of one source of information generating signals in a stochastic manner. In reality, however, there are many cases where the DM must consider a number of signals based on the same set of states of nature but generated by "independent" information systems. The IS model does not deal with such cases explicitly. Since the power of any model is in its correspondence to a real decision problem [5], the traditional model needs modification.

In other words, the real state of nature is unknown to the decision maker who must learn about it through signals. However, instead of reacting to a signal provided by a single system, the DM requests signals from a number of sources and reacts only after examining the combined information. The last section of this paper provides an elaborate numerical example of such a case. Here we illustrate pertaining cases through a number of narrative examples.

A manufacturing business of hi-tech electronic products wishes to place an offer in a bid for a new product. The CEO requests a cost estimate for the new product. The estimate is prepared by two independent teams. One team takes a "micro" approach: the product is decomposed into major components which are further divided into items until the entire bill of materials (see [13, ch. 11]) is exploded; then the cost of each elementary component is assessed, and the total is summed up in order to obtain an aggregate estimate. The second team takes a "macro" approach (also known as parametric control): the team attempts to assess overall traits of the new product, for example, weight, volume, and number of electronic cards (slots); based on these few parameters and, of course, on past experience and historical data, the team calculates a rough estimate of the cost. The CEO must consider two distinct signals provided by two independent sources before deciding what he or she believes to be the "real" cost.

A similar example is common in the construction industry. The cost of building a new house usually is estimated in two ways. One way is to try to list all the necessary "ingredients" of the house and sum up their cost; the other way is to calculate the "magic number," normally the floor space, and multiply this number by the cost per square foot. The result will yield a quite good approximation of the cost of the building.

One last example is taken from a totally different area. When people face a crucial decision regarding their own health (i.e., undergoing a major surgery), most of them ask for a second opinion. A second opinion is indeed an additional signal based on the same state of nature but provided by an independent source.

The problem of multisource information has been discussed in a number of articles. Winkler [19], for instance, examined the problem of combining several forecasts of a single variable. Morris [14] [15] treated the same problem in a two-stage Bayesian process. However, the incorporation of the IS model (which is in fact based on decision theory) into the problem of multisource information has not yet been explicitly presented.

This paper applies the IS model to the case where a decision maker has to consider a number of signals provided by "independent" (this term will be defined later more rigorously) information structures. The paper addresses a number of questions. The first one is how independence of ISs can be formulated. The paper coins a new term, ***orthogonal information structures***; a new mathematical operator labeled ***orthogonal product*** is defined and its mathematical properties are analyzed.

A second question deals with the value of the information provided by orthogonal ISs. It is proven that the combined information collected from orthogonal systems generally is more informative than the information produced by each individual system, that is, it is worthwhile to ask for a second opinion (subject, of course, to cost considerations). The Appendix expands this result to a case of partially dependent information structures.

The last part of the paper discusses the applicability of the orthogonal model. This discussion is assisted by a numerical example.

ORTHOGONAL INFORMATION STRUCTURES

We first review briefly the traditional IS model [10] and then incorporate orthogonality into the model.

Let E be a finite set of events (states of nature), $E=\{e_1, \ldots, e_{nE}\}$. Let $\mathbf{p}$ be a vector of a priori probabilities associated with the events in E, $p^t=(p_1, \ldots, p_{nE})$, where $\Sigma p_i=1$, $p_i \geq 0$, $i=1, \ldots, n_E$. (The superscript t stands for a transpose operator.)

Let Z be a finite set of signals, $Z=\{z_1, \ldots, z_{nZ}\}$. An information structure **Q** is defined as a Markovian (stochastic) matrix of conditional probabilities (dimension $n_E \times n_Z$) in which signals of the set Z will be displayed at the occurrence of an event of E. Thus q_{ij} of **Q** is the probability that for a given event e_i, signal z_j will be displayed.

Let A be a finite set of actions that can be taken by the decision maker, $A=\{a_1, \ldots, a_{nA}\}$. A cardinal payoff function U is defined from $A \times E$ to the real numbers, R^1, associating payoffs to pairs of actions and events, $U{:}A \times E \rightarrow R^1$. The function U can be depicted by an $n_A \times n_E$ matrix, denoted **U**, whose each element a_{ij} reflects the payoff gained when an action a_i is taken and the event turns out to be e_j.

The DM cannot observe the events but only the signals and chooses actions accordingly. The DM's strategy is delineated by an $n_Z \times n_A$ Markov matrix **D**, whose each element d_{ij} determines the probability that the DM takes action a_j on observing signal z_i. Obviously, the DM wishes to optimize **D** to obtain the maximum expected payoff. This is performed by the following algorithm.

Let $\mathbf{p}'$ be a square matrix containing the elements of p in its main diagonal and zeros elsewhere:

$$\mathbf{p}' = \begin{vmatrix} p_1 & 0 & \cdots & 0 \\ 0 & \cdot & & \vdots \\ \vdots & & \cdot & 0 \\ 0 & \cdots & 0 & p_{nE} \end{vmatrix} .$$

The expected payoff gained from **Q**, **U**, and a decision rule **D** is given by tr(**QDUp**′), where "tr" represents the trace operator. Maximization of the above is obtained by solving a linear programming problem for the elements of **D** constrained by the properties of a Markovian matrix (see [3] for an elaborate discussion).

Given two ISs **Q** and **R** operating on the same set of events E, **Q** is defined to be *generally more informative* than **R** if the maximal expected payoff yielded by **R** is not larger than that yielded by **Q** for all payoff matrices **U** and all probability vectors **p**. A partial rank ordering of ISs is provided by Blackwell's Theorem [10] stating that **Q** is generally more informative than **R** if and only if there exists a Markov matrix **M** with appropriate dimensions such that $\mathbf{Q} \times \mathbf{M} = R$; **M** is called the garbling matrix. (Hereafter we will use the terms "informativeness" and "more informative" for the relationship "generally more informative.")

The example below is provided to illustrate the IS model; it follows [1] and [2] closely.

Consider a railroad intersection with a highway in which a two-color traffic light is posted. The pertinent sets are

Events: $E=\{$Train arrives (T), No train (N)$\}$
Probabilities: $p^t=(.1, .9)$
Signals: $Z=\{$Red (R), Green (G)$\}$
Actions: $A=\{$Stop (S), Proceed (P)$\}$

The payoff matrix, **U**, is as follows:

	Event	
Action	T	N
S	0	-10
P	-10^{18}	0

Let **Q** and **R** be two information structures (i.e., traffic lights) as follows:

Q		
	Signal	
Event	R	G
T	1	0
N	0	1

R		
	Signal	
Event	R	G
T	1	0
N	.001	.999

According to Blackwell's Theorem, **Q** is generally more informative than **R** since there exists a Markov matrix **M** such that $\mathbf{QM}=\mathbf{R}$ (in fact, for this example $\mathbf{M}=\mathbf{R}$). Therefore there is no need to calculate the expected payoff in order to compare **Q** and **R**. If we do perform the calculation we indeed will find that the optimal decision rule for both structures is $\mathbf{D}=\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$ and $\mathrm{tr}(\mathbf{QDUp}')=0$ while $\mathrm{tr}(\mathbf{RDUp}')=-.009$. We now introduce the notion of orthogonality. Intuitively, two (or more) ISs are considered to be orthogonal when they observe the same set of states of nature but generate signals independently; in other words, the likelihood of a signal being generated by a certain IS does not depend on the signal produced by the other IS, but only on the conditional probabilities of the IS itself. This now will be formulated more rigorously.

Let **Q** and **R** be two ISs operating on the set of events E and producing the sets of signals $Z=\{z_1, \ldots, z_{nZ}\}$ and $W=\{w_1, \ldots, w_{nW}\}$, respectively.

Definition 1: Signals z_j and w_k are orthogonal if and only if $\Pr((z_j/e_i)/(w_k/e_i)) = \Pr(z_j/e_i)=q_{ij}$ and $\Pr((w_k/e_i)/(z_j/e_i))=\Pr(w_k/e_i)=r_{ik}$ for all i (i.e., the probability that z_j is triggered by an occurrence of e_i does not depend on whether w_k has been displayed or not, and vice versa).

Definition 2: Information structures **Q** and **R** are orthogonal when all their signals are orthogonal one to the other.

It is obvious that the relationship of orthogonality is symmetric (by definition) and transitive. The next section shows how to compose an integrated IS out of two orthogonal ones.

ORTHOGONAL PRODUCT

This section shows how to combine the information provided by distinct orthogonal ISs. This is done by defining a mathematical operator, orthogonal product (or orthogonal multiplication), and inquiring into its traits.

Let **Q** and **R** be two ISs defined in the same way as in the previous section.

Definition 3: **S** is the orthogonal product of **Q** and **R** (denoted $\mathbf{S}=\mathbf{Q}@\mathbf{R}$) if **S** is a matrix of n_E rows and $n_w \times n_z$ columns whose elements are as follows: for every i, $i=1, \ldots, n_E$

$$
\begin{aligned}
s_{i1} &= q_{i1} \times r_{i1} \\
s_{i2} &= q_{i1} \times r_{i2} \\
&\vdots \\
s_{inW} &= q_{i1} \times r_{inW} \\
s_{inW+1} &= q_{i2} \times r_{i1} \\
&\vdots
\end{aligned}
$$

The orthogonal product **S** maps a set of n_E events into a set of $n_W \times n_Z$ signals. The following numerical example clarifies the notion of orthogonal product.

Example: Let **Q** and **R** be the following 2×2 orthogonal IS:

$$
\begin{array}{c|cc|}
 & z_1 & z_2 \\
e_1 & .9 & .1 \\
e_2 & .2 & .8
\end{array} = \mathbf{Q}
\qquad
\begin{array}{c|cc|}
 & w_1 & w_2 \\
e_1 & .8 & .2 \\
e_2 & .4 & .6
\end{array} = \mathbf{R}
$$

Let $\mathbf{S} = \mathbf{Q}@\mathbf{R}$. Then **S** is computed as

$$
\begin{array}{c|cccc|}
 & z_1\&w_1 & z_1\&w_2 & z_2\&w_1 & z_2\&w_2 \\
e_1 & .72 & .18 & .08 & .02 \\
e_2 & .08 & .12 & .32 & .48
\end{array} = \mathbf{S}
$$

S observes the original set E and produces four signals which indicate what can be displayed to the DM: z_1 and w_1, z_1 and w_2, and so forth. It now is the task of the DM to devise the optimal decision rule for each individual pair of signals. But first let us discuss some mathematical properties of the orthogonal product.

Property 1: **S** is a Markovian matrix (this is the property of closure, i.e., the set of IS is closed under the operator of orthogonal multiplication; see [20]).

Proof: Obviously, each element of **S** is nonnegative; it is sufficient to show that the sum of all the elements in a row of **S** equals 1.

$$\sum_m s_{im} = \sum_j \sum_k q_{ij} \times r_{ik} = \sum_j q_{ij} \times \sum_k r_{ik} = 1 \times 1 = 1.$$

Property 2: The orthogonal product is an associative operation, that is, $\mathbf{Q}@(\mathbf{R}@\mathbf{L}) = (\mathbf{Q}@\mathbf{R})@\mathbf{L}$. This property is directly consequential to the definition of the operator.

Property 3: The unity element of the orthogonal product is the vector $\mathbf{1} = \begin{vmatrix} 1 \\ \vdots \\ 1 \end{vmatrix}$ since for any **Q** it yields $\mathbf{1}@\mathbf{Q} = \mathbf{Q}$.

Note that the unity element may have an informational interpretation as well. An IS of that kind does not provide any information; it always produces the same signal. Thus, combining it with another IS should not supply any additional knowledge.

Optimizing the decision rule for an orthogonal product matrix is similar to the process carried out for a "simple" IS, namely solving an LP problem. This is obvious since the product is indeed an IS. Let us continue with the previous example to demonstrate the optimization process.

Suppose **Q**, **R**, and **S** are as follows:

$$\mathbf{Q}=\begin{vmatrix} .8 & .2 \\ .5 & .5 \end{vmatrix} \quad \mathbf{R}=\begin{vmatrix} .5 & .5 \\ .8 & .2 \end{vmatrix} \quad \mathbf{S}=\mathbf{Q}@\mathbf{R}=\begin{vmatrix} .4 & .4 & .1 & .1 \\ .4 & .1 & .4 & .1 \end{vmatrix}.$$

Suppose the a priori probabilities are $p^t=(.6, .4)$, the set of actions is $A=\{a_1, a_2\}$, and the payoff matrix **U** is the following:

$$\begin{array}{c|cc|} & s_1 & s_2 \\ a_1 & 20 & -15 \\ a_2 & -30 & 40 \end{array} = \mathbf{U}.$$

Given **Q** alone, the optimal decision rule is $D_Q{}^*=\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$ and the expected payoff would be $T_Q=11$. For **R**, the optimal decision rule is $D_R{}^*=\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$ and the expected payoff would be $T_R=8.6$. Given both ISs, the optimal decision rule for **S** is

$$\begin{array}{c|cc|} & a_1 & a_2 \\ z_1\&w_1 & 1 & 0 \\ z_1\&w_2 & 0 & 1 \\ z_2\&w_1 & 1 & 0 \\ z_2\&w_2 & 1 & 0 \end{array} = D_S{}^*$$

and the expected payoff is $T_S=11.8$.

Note that the marginal value of the information provided by the "second opinion" (i.e., **S**) is .8 relative to **Q**, and 3.2 relative to **R**. However, before we elaborate on the value of orthogonal information, let us discuss two more properties of the orthogonal product.

Property 4: The "maximum-entropy" matrix is the "null element" of the orthogonal product in terms of informativeness (i.e., contribution to the expected payoff). A maximum-entropy matrix (see [16]) is a stochastic matrix whose elements are all equal. For instance, suppose the matrix's dimensions are $n_E \times n_Z$; then every element t_{ij} equals $1/n_Z$ for all i and j. For example, assume $\mathbf{T}=\begin{vmatrix} .5 & .5 \\ .5 & .5 \end{vmatrix}$ and $\mathbf{S}=\mathbf{Q}@\mathbf{T}$, where **Q** is taken from the previous example. **S** is in fact a flattening of **Q**:

$$\mathbf{S}=\begin{vmatrix} .45 & .45 & .05 & .05 \\ .1 & .1 & .4 & .4 \end{vmatrix}$$

The optimal decision rule $\mathbf{D}=\begin{vmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{vmatrix}$ will yield the same expected payoff as attained by **Q**.

Property 5: The matrices $\mathbf{S}=\mathbf{Q}@\mathbf{R}$ and $\mathbf{S}_1=\mathbf{R}@\mathbf{Q}$ are informatively equivalent.

Proof: It is obvious that **S** and $\mathbf{S}_1$ consist of the same columns arranged in a different order. It is proven in [9] that permutations of IS columns do not affect the informativeness of the IS.

Intuitively, the interpretation of Property 5 is clear. Knowing that the columns of the orthogonal matrix have been permuted, the DM has only to permute the rows of the decision matrix accordingly. This property implies that it really does not matter which opinion is considered first and which is considered second as long as two opinions are indeed asked for.

The next section will now inquire into the informativeness of the orthogonal product vis-á-vis its multipliers.

INFORMATIVENESS OF THE ORTHOGONAL PRODUCT

Is it always better to acquire a second opinion? A positive answer to that question is not quite intuitive. According to Blackwell's Theorem, a garbled IS cannot perform better than the original matrix. An orthogonal product is after all an IS produced from two original stochastic matrices, so perhaps it will not perform better than its "parent" matrices. However, unlike a garbled IS, the orthogonal product is not generated through common algebraic multiplication but rather by a different operation. We will show now that an orthogonal product is more informative than its multipliers.

Theorem 1: Let **Q** and **R** be two ISs operating on a common set of events. Let **S** be the orthogonal product of **Q** and **R**. Then **S** is generally more informative than **Q**, and **S** is generally more informative than **R**.

Proof: The proof will be handled in a constructive fashion, namely we will build garbling matrices that transform **S** to **Q** or to **R**. This will constitute a sufficient condition for applying Blackwell's Theorem to prove the above assertion.

$$\text{Let } \mathbf{Q}=\begin{vmatrix} q_{11} & \cdots & q_{1nZ} \\ \vdots & & \\ q_{nE1} & \cdots & q_{nEnZ} \end{vmatrix} \qquad \mathbf{R}=\begin{vmatrix} r_{11} & \cdots & r_{1nW} \\ \vdots & & \\ r_{nE1} & \cdots & r_{nEnW} \end{vmatrix}$$

$$\mathbf{S}=\mathbf{Q}@\mathbf{R}=\begin{vmatrix} q_{11}r_{11} & \cdots & q_{11}r_{1nW} & \cdots & q_{1nZ}r_{1nW} \\ \vdots & & & & \\ q_{nE1}r_{nE1} & \cdots & q_{nE1}r_{nEnW} & \cdots & q_{nEnZ}r_{nEnW} \end{vmatrix}$$

Let $\mathbf{M}_1$ and $\mathbf{M}_2$ be two Markov matrices each having $n_Z \times n_W$ rows and n_Z or n_W columns, respectively, constructed as follows:

$$\mathbf{M}_1 = \left|\begin{array}{cccccc}
1 & 0 & \cdot & \cdot & \cdot & 0 \\
1 & 0 & \cdot & \cdot & \cdot & 0 \\
\vdots & & & & & \\
1 & 0 & \cdot & \cdot & \cdot & 0 \\
0 & 1 & 0 & \cdot & \cdot & \cdot\ 0 \\
0 & 1 & 0 & \cdot & \cdot & \cdot\ 0 \\
\vdots & & & & & \\
0 & 1 & 0 & \cdot & \cdot & \cdot\ 0 \\
\vdots & & & & & \\
0 & \cdot & \cdot & \cdot & 0 & 1 \\
0 & \cdot & \cdot & \cdot & 0 & 1 \\
\vdots & & & & & \\
0 & \cdot & \cdot & \cdot & 0 & 1
\end{array}\right|
\begin{array}{l}
\\ \text{1st } n_W \text{ rows} \\ \\ \\ \\ \\ \text{2nd } n_W \text{ rows} \\ \\ \\ \\ \\ \text{last } n_W \text{ rows} \\ \\
\end{array}
\qquad
\mathbf{M}_2 = \left|\begin{array}{cccccc}
1 & 0 & \cdot & \cdot & \cdot & 0 \\
0 & 1 & 0 & \cdot & \cdot & \cdot\ 0 \\
\vdots & & & & & \\
0 & \cdot & \cdot & \cdot & 0 & 1 \\
1 & 0 & \cdot & \cdot & \cdot & 0 \\
0 & 1 & 0 & \cdot & \cdot & \cdot\ 0 \\
\vdots & & & & & \\
0 & \cdot & \cdot & \cdot & 0 & 1 \\
\vdots & & & & & \\
1 & 0 & \cdot & \cdot & \cdot & 0 \\
0 & 1 & 0 & \cdot & \cdot & \cdot\ 0 \\
\vdots & & & & & \\
0 & \cdot & \cdot & \cdot & 0 & 1
\end{array}\right|
\begin{array}{l}
\\ \text{1st } n_Z \text{ rows} \\ \\ \\ \\ \\ \text{2nd } n_Z \text{ rows} \\ \\ \\ \\ \\ \text{last } n_Z \text{ rows.} \\ \\
\end{array}$$

It is easy to see that $\mathbf{S} \times \mathbf{M}_1 = \mathbf{Q}$ and $\mathbf{S} \times \mathbf{M}_2 = \mathbf{R}$. Hence, $\mathbf{S}$ is generally more informative than both $\mathbf{Q}$ and $\mathbf{R}$.

Since sometimes two ISs can be equivalent in terms of their informativeness (e.g., $\left|\begin{array}{cc} .9 & .1 \\ .2 & .8 \end{array}\right|$ is equivalent to $\left|\begin{array}{cc} .1 & .9 \\ .8 & .2 \end{array}\right|$), it is important to note that this is not the case here. The relationship "generally more informative" is a one-way relationship between the orthogonal product and its multipliers (except for some "irregular" cases presented in the next section). In order to substantiate this proposition it is sufficient to provide a numerical example.

Examine the sample matrices $\mathbf{Q}$ and $\mathbf{S}$ of the previous section. It is easy to see that the matrix $\mathbf{M}_3$ that solves the set of linear equations $\mathbf{Q} \times \mathbf{M}_3 = \mathbf{S}$ is not Markovian, hence $\mathbf{Q}$ is not generally more informative than $\mathbf{S}$.

We now will examine some immediate results of the above theorem.

THE VALUE OF A SECOND OPINION

Some immediate conclusions can be derived from Theorem 1. First, it is clear that the orthogonal product of $n+1$ orthogonal ISs is generally more informative than the product of any n matrices out of them. This may imply that the acquisition of an additional orthogonal opinion is always worthwhile. However, one has to consider the cost of obtaining the additional information vis-á-vis the marginal expected payoff.

Suppose the cost is not a constraining factor. How far should one seek for additional opinions? A clear stopping rule is when one manages to obtain a "complete and perfect" IS, that is, the unity matrix. Such a matrix provides the maximal expected payoff so there is no need for further inquiry. This can also be displayed in terms of an orthogonal product: Let

$$\mathbf{Q} = \left|\begin{array}{cc} .9 & .1 \\ .2 & .8 \end{array}\right| \qquad \mathbf{I} = \left|\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right|.$$

Their orthogonal product is $S_1 = \begin{vmatrix} .9 & 0 & .1 & 0 \\ 0 & .2 & 0 & .8 \end{vmatrix}$. Suppose the optimal decision rule for **I** was $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$. Obviously, a decision rule in the form of $\begin{vmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{vmatrix}$ will yield the same expected payoff.

It has been noted in a previous section that the "maximum-entropy" matrix cannot improve the expected payoff. Nevertheless, the above theorem makes it obvious that this matrix does not worsen the level of informativeness when it orthogonally multiplies a "regular" IS. It maintains, in fact, the same level of informativeness; this of course is mathematically possible since the rank ordering imposed by Blackwell's Theorem is not a strict relationship.

The maximum-entropy matrix also can be used to demonstrate the notion of "little improvement." For instance, the matrix $\begin{vmatrix} .5 & .5 \\ .5 & .5 \end{vmatrix}$ does not add to the informativeness of any existing IS; however, if this matrix is slightly modified to look like $\begin{vmatrix} .5+\epsilon_1 & .5-\epsilon_1 \\ .5-\epsilon_2 & .5+\epsilon_2 \end{vmatrix}$ where ϵ_1 and ϵ_2 are small numbers, its orthogonal multiplication with any other IS represents a "little bit" of added knowledge. In ordinary words, the maximum-entropy matrix represents a situation of maximum uncertainty, and any deviation from it likely will constitute an improvement.

This concludes the discussion on the value of an orthogonal IS. In the Appendix we present some properties of partially orthogonal structures. The next section discusses an application of orthogonal systems for cost estimation.

APPLYING THE MODEL

The purpose of this concluding section is to raise some ideas on an application of the orthogonality model. This will be done in the context of cost estimation problems.

Cost estimate of large (and usually unique and nonrepetitive) projects is a severe problem in areas such as construction, public utility companies, and ship building, aircraft, software development, electronics, and high-tech industries. Deviations might be three or four times larger than initial estimates (see [7] for examples of hardware, software, power plants, and aircraft developments; [12] for examples of subway construction and military systems developments). There are numerous reasons for the deviations. The prominent ones are unforeseen exogenous factors (e.g., environmental, political, legal), mismanagement, deliberate deviation in order to get a contract, and wrong estimation techniques.

We will show now how the concept of orthogonal IS can be applied to improve the estimation process. This will be done using a numerical example.

Let E be a set of events representing the "real" cost of a project (ex post); for example, $E = \{\$100{,}000;\ \$200{,}000;\ \$300{,}000\}$, where the figures indicate possible costs of a project. Note that the fineness of event classification is assumed to be subject to judgment and may be revised. Assume the events are arranged such that the associated cost figures are sorted in an ascending order. Let $p^t=(1/3, 1/3, 1/3)$ be the a priori probabilities assigned to the events of E. The probabilities may be subjective or based on past experience.

The decision problem is to estimate the cost. It is assumed that the DM would like to tell what will be the real cost. Hence the decision rule is a matrix whose rows correspond to signals provided by an information system (which will be discussed later); the columns correspond to estimates of the costs (events), and the elements indicate the probability that the DM estimates a certain cost value under a given signal. For example:

	Estimate			
	$100K	$200K	$300K	
signal 100,000	d_{11}	d_{12}	d_{13}	
signal 200,000	d_{21}	d_{22}	d_{23}	$=\mathbf{D}$
signal 300,000	d_{31}	d_{32}	d_{33}	

The payoff matrix **U** displays a cardinal profit function which relates estimates to occurrences of real events (ex post). It can be assumed that as the deviation increases so does the penalty the company pays. Therefore, the elements (payoff values) of the main diagonal of the matrix will be more in favor of the DM while "remote" elements will reduce the profit (or increase the loss) monotonically. The following example delineates a matrix reflecting losses due to wrong estimates; note that underestimates and overestimates do not necessarily incur similar losses:

$$\mathbf{U}=\begin{vmatrix} 0 & -150{,}000 & -180{,}000 \\ -20{,}000 & 0 & -120{,}000 \\ -70{,}000 & -50{,}000 & 0 \end{vmatrix}$$

The objective of the DM is to determine the most appropriate decision rule, namely to estimate the cost as accurately as possible when a certain signal emerges from the IS. Note that the DM does not necessarily need to follow the signal; if the DM does not trust the IS, he or she may place an estimate not concurring with the signal emerging from the system (see [3] for a case like that).

In order to perform a reasonable estimate, the DM employs some teams that ought to provide him or her with sufficient data. In the case of cost estimation, a very common approach is to decompose the project into components to obtain the bill of material for the project. This method is labelled "bottom-up" or "work breakdown structure" (WBS); it was formulated by the U.S. Army in MIL STD 881A [18] (also see [4]). Once the elementary components have been identified and their cost has been determined, the figures are aggregated upward to obtain the total cost, to which one must add labor and other direct costs as well.

This method is considered relatively accurate; however it consumes much time and labor. Its accuracy deteriorates in R&D projects or in projects where humanpower is a major factor (e.g., software development).

An alternative approach is called parametric costing [6]. This method is based on identifying some crude parameters that constitute a significant statistical correlation with the cost of a project in a certain industry. For instance, a well-known method was developed by Large, Campbell, and Cates [8] for estimating the cost of manufacturing a new aircraft. Their assessment for the cost of building 100 combat airplanes of the same model is given by a simple formula:

$$C = 4.2 \times W^{.73} \times S^{.74}$$

where C is the cost per 100 units, W is the aircraft weight, and S is its maximum speed. Another example is an adaptive system named PRICE [17]. The system user can calibrate it to fit the organization's particular circumstances. PRICE can handle estimates in a number of areas such as hardware, electronics, and software.

The problem with information systems for parametric costing is that they are not costless. They provide "quick and dirty" information—that is, faster but probably less accurate—at a certain additional cost. The DM can use them as decision support systems either to obtain fast responses or to crosscheck the signal provided by the regular IS. Still the question remains: How much should one pay for a "second-opinion" system?

The orthogonal IS model cannot advise us how much to pay, but it can tell the worth of a second opinion by figuring out the marginal expected payoff emanating from the orthogonal product of the two systems. The DM must judge whether the cost is worthwhile. Let us turn now to the numerical example to demonstrate this.

Suppose the ISs for the bottom-up and the parametric approaches are the following matrices **Q** and **R**, respectively:

$$\mathbf{Q} = \begin{vmatrix} .9 & .1 & 0 \\ .05 & .9 & .05 \\ 0 & .1 & .9 \end{vmatrix} \qquad \mathbf{R} = \begin{vmatrix} .7 & .2 & .1 \\ .1 & .8 & .1 \\ .1 & .2 & .7 \end{vmatrix}.$$

Based on the IS model, the optimal decision rule for both **Q** and **R** is the following matrix:

$$\mathbf{D^*} = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}.$$

The expected payoffs for **Q** and **R** are -8,000 and -22,000, respectively. In terms of gross expected payoff **Q** is preferred; however one must consider the cost of carrying out a bottom-up analysis and the time it might consume. Nevertheless, let us see if taking **Q** or **R** as a second opinion yields some significant marginal payoff.

Let S be the orthogonal product, **S**=**Q@R**:

$$\mathbf{S} = \begin{vmatrix} .63 & .18 & .09 & .07 & .02 & .01 & 0 & 0 & 0 \\ .005 & .04 & .005 & .09 & .72 & .09 & .005 & .04 & .005 \\ 0 & 0 & 0 & .01 & .02 & .07 & .09 & .18 & .63 \end{vmatrix}.$$

The optimal decision rule for **S** is

$$\mathbf{D}_S{}^* = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{vmatrix}$$

The expected payoff is –6,066.66. The marginal gross payoff is 1,933.34 relative to **Q** and 15,933.34 relative to **R**. These values should be compared to the cost and time factors associated with obtaining the additional information. Note that it makes intuitive sense that when **R** is available, the additional value of a second opinion provided by **Q** is greater than the other way around, since initially **R** appears to be less "exact" than **Q**.

The last important question is how to calibrate the model for practical use. Initially it should be based on past experience and judgment. However, once the model is programmed and installed on a computer (including the LP routine to solve the optimal decision rule), it can be used not only to assess the value of a second opinion but also to analyze the sensitivity of the solution to various assumptions regarding the model's components, that is, the ISs, the payoffs, and the a priori probabilities. In fact the programmed model can serve as a decision support system for obtaining a solution as well as for testing the initial assumptions. [Received: June 17, 1986. Accepted: December 16, 1986.]

REFERENCES

[1] Ahituv, N. A comparison of information structures for a "rigid decision rule" case. *Decision Sciences*, 1981, *12*, 399-416.

[2] Ahituv, N. Describing the information system life cycle as an adjustment process between information and decisions. *International Journal of Policy Analysis and Information Systems*, 1982, *6*(2), 133-145.

[3] Ahituv, N., & Wand, Y. Comparative estimation of information under two business objectives. *Decision Sciences*, 1984, *15*, 31-51.

[4] Buffa, E. S. *Modern production/operations management* (7th ed.). New York: Wiley, 1983.

[5] Demski, J. S. *Information analysis*. Reading, MA: Addison-Wesley, 1972.

[6] Dumas, A. J. Parametric costing and institutionalized inefficiency. *AIIE Transactions*, 1979, *11*(2), 147-154.

[7] Kharbanda, O. P., & Stallworthy, E. A. *How to learn from project disasters*. Aldershot, England: Grower, 1983.

[8] Large, J. P., Campbell, H. G., & Cates, D. *Parametric equations of estimating aircraft costs* (R-1963-PA&E). Santa Monica, CA: Rand Corporation, 1975.

[9] Marschak, J. Economics of information systems. *Journal of the American Statistical Association*, 1971, *66*, 192-219.

[10] McGuire, C. B. Comparison of information structures. In C. B. McGuire & R. Radner (Eds.), *Decision and organization*. Amsterdam: North Holland, 1972.

[11] McGuire, C. B., & Radner, R. (Eds.). *Decision and organization*. Amsterdam: North Holland, 1972.

[12] McNicols, G. R. Independent parametric costing: What? Why? How? In *Proceedings of the 26th Annual Conference of the American Institute of Industrial Engineers*. Norcross, GA, 1975.

[13] Monks, J. G. *Operations management—Theory and problems* (2nd ed.). New York: McGraw-Hill, 1982.

[14] Morris, P. A. Combining expert judgments: A Bayesian approach. *Management Science*, 1977, *23*, 679-693.
[15] Morris, P. A. An axiomatic approach to expert resolution. *Management Science*, 1983, *29*, 24-32.
[16] Shannon, C. H., & Weaver, W. *The mathematical theory of communication.* Urbana: University of Illinois Press, 1949.
[17] Sipper, D., & Zeidmann, A. *Cost estimate of research and development of hi-tech electronic systems.* Unpublished research report (in Hebrew), Tel Aviv University, Faculty of Engineering, July 1983.
[18] U.S. Department of Defense. *Military standard-work breakdown structures for defense material items* (MIL STD 881A). Washington, DC: Government Printing Office, 1975.
[19] Winkler, R. L. Combining probability distribution from dependent information sources. *Management Science*, 1981, *27*, 479-488.
[20] Wylie, C. R., & Barrett, L. C. *Advanced engineering mathematics* (5th ed.). Auckland, New Zealand: McGraw-Hill International, 1982.

APPENDIX

Partly Independent Information Structures

Although our main interest lies in independent opinions (i.e., orthogonal IS), it is important to note that the orthogonal product operator can be applied to cases where the signals are not fully independent; consequently, a theorem similar to Theorem 1 can be proven. This is discussed briefly here.

Definition 4: Let **Q** and **R** be two ISs operating on the same set of events, E, and producing the sets of signals Z and W, respectively. **S** is called the general product of **Q** and **R** (denoted **S**=**Q@R**) if **S** is a matrix of n_E rows and $n_W \times n_Z$ columns whose elements are as follows: for every i, $i=1, \ldots, n_E$

$$
\begin{aligned}
S_{i1} &= \Pr(Z_1 \cap W_1/e_i) \\
S_{i2} &= \Pr(Z_1 \cap W_2/e_i) \\
&\vdots \\
S_{inW} &= \Pr(Z_1 \cap W_{nW}/e_i) \\
S_{inW+1} &= \Pr(Z_2 \cap W_1/e_i) \\
&\vdots
\end{aligned}
$$

The general product **S** maps a set of n_E events into a set of $n_W \times n_Z$ signals. However, unlike the orthogonal product where the signals of the two different ISs were assumed to be independent, here that assumption is not made. In the "worst" case, the structures are fully dependent and the product structure should not provide any additional valuable information. In the "best" case the structures indeed are independent and the discussion presented in the main body of the paper pertains. In any event the general product matrix cannot be less informative than its "parent" structures. This is given in the following theorem.

Theorem 2: Let **Q** and **R** be two ISs operating on a common set of events. Let **S** be the general product of **Q** and **R**. Then **S** is generally more informative than **Q**, and **S** is generally more informative than **R**.

The proof follows the same constructive fashion and the same arguments presented in the proof of Theorem 1, so we shall not repeat it here.

In conclusion, the orthogonal product is, in fact, a special case of the general product. We could have started by proving Theorem 2 and then Theorem 1 would have become a corollary. Nevertheless, since we believe the orthogonal case is more interesting than the general one, we developed the paper the way we did.

Niv Ahituv is an Associate Professor and Chairperson of the Computers and Information Systems Program at the Faculty of Management, Tel Aviv University, Israel. Formerly, he lectured at the Claremont Graduate School, the University of Calgary, and the University of British Columbia and managed

the data processing department at the Bank of Israel. Dr. Ahituv holds a degree of B.Sc. in mathematics and an M.B.A., M.Sc., and Ph.D. in information systems. His articles have appeared in *Computers and Operations Research, The Computer Journal, Information and Management, MIS Quarterly, Interfaces, Decision Sciences, Datamation, Communications of the ACM,* and others. He has co-authored a book, *Principles of Information Systems for Management.* Dr. Ahituv's main areas of interest are economics of computers, information economics, and information systems management and development.

Boaz Ronen is a lecturer at the Faculty of Management, Tel Aviv University, and a visiting Assistant Professor at the Information Systems Area, Graduate School of Business Administration, New York University. His research and teaching areas are in MIS and production/operations management. Dr. Ronen worked for ten years in the high-tech electronics industry as a project manager. He holds degrees of B.Sc. in electronic engineering, M.Sc. in management sciences, and Ph.D. in MIS. His articles have been published in such journals as *Operations Research* and *Journal of Operations Research Society.*