THE CONCEPT OF THE AMOUNT OF INFORMATION IN INFORMATION EVALUATION†

By HIROYUKI ITAMI*

1. *Introduction*

The purpose of this paper is to introduce the concept of the amount of information within the framework of information economics and accounting information evaluation as developed in, for example, Demski [4], Feltham [5], and Marschak and Radner [11]. As such, the persent paper also uses a decision theoretic framework.

In the next section, the concept and definition of the amount of information provided by an information system will be formally introduced. It is based on the concept of entropy developed in information theory. Although there have been uses of entropy and their criticism in accounting research (for example, Lev [7], Ronen and Falk [14], Abdel-khalik [1], Lee and Bedford [6], Nakano [13]), their emphasis is either on the use of entropy as a measure of aggregation or on a communication theoretic view of the accounting process. This paper emphasizes the decision theoretic use of the entropy as a measure of the degree of uncertainty of a probability distribution and takes a rather different perspective as will become evident as we proceed.

After describing several properties of the amount of information, the merits of its introduction will be discussed in a practical as well as a theoretical sense. Section 3 also presents some merits of this concept by revealing an appealing relationship between the value of information and the amount of information through an example. A suggestion is made to break down information evaluation process in the information economics framework into the production process of a quantity of information and the consumption and the evaluation process of this quantity of information.

In Section 4 a general theorem linking the value of information and the amount of information in information systems selection is presented and its accounting implications are discussed.

2. *The Amount of Information Provided by an Information System*

Briefly, the amount of information provided by an information system is defined as the average amount of reduction of uncertainty from a prior distribution to a posterior distribu-

* Assistant Professor (*Jokyōju*) of Management Science.

† This paper was presented at the 1975 Annual Meeting of the American Accounting Association at Tucson, Arizona, U.S.A. Partial financial support of the Peat, Marwick & Mitchell Foundation through Stanford Business School while the author was on the faculty of the School is acknowledged.

tion, a reduction which is provided by an information system as a transforming agent of a prior distribution to a posterior distribution. In the following, the degree of uncertainty of a probability distribution is measured by the entropy of the distribution as developed in the area of information theory and the precise definition of the amount of information of an information system will be given.[1]

Let us first introduce several notations for later discussion:

S: the set of states of nature or events

s: a state of nature, an element of S

A: the set of actions available to the decision maker

a: an action, an element of A

$U(s, a)$: utility function of the decision maker

$\varphi(s)$: a prior probability distribution over S representing the decision maker's belief for future uncertain events before he receives any message.

η: an information system

Y: the set of messages from η

y: a message, an element of Y

$P(y/s, \eta)$: a probability distribution of messages given an information system η and a state s.

$\varphi(s/y, \eta)$: a posterior probability distribution of states of nature given a message y from an information system η.

In the context of information economics, the (gross) value of an information system η, denoted by $V(\eta)$, may be defined as:

(1) $$V(\eta)=\sum_{s\varepsilon S}\sum_{y\varepsilon Y}\varphi(s)P(y/s, \eta)\left\{\max_{a\varepsilon A}\sum_{s\varepsilon S}U(s, a)\varphi(s/y, \eta)\right\}-\max_{a\varepsilon A}\sum_{s\varepsilon S}U(s, a)\varphi(s)$$

That is, $V(\eta)$ is the difference between the maximal expected utilities with an information system η and with no information system. Roughly speaking, it is defined as the expected increase of utility by having an information system η.

Likewise, we may define the amount of information provided by η as the expected reduction of the degree of uncertainty with respect to S. But first, we have to define the degree of uncertainty of a probability distribution. The concept of entropy in information theory gives us a neat definition of the degree of uncertainty of a probability distribution.[2] The entropy of a prior distribution, denoted by $H(s)$ is

(2) $$H(s)\equiv-\sum_{s\varepsilon S}\varphi(s)\log\varphi(s)$$

Now after the decision maker has received a message y from η, how much uncertainty is left in his posterior distribution? Using the entropy, it is denoted by $H(s/y, \eta)$,

(3) $$H(s/y, \eta)\equiv-\sum_{s\varepsilon S}\varphi(s/y, \eta)\log\varphi(s/y, \eta)$$

Then, the amount of uncertainty reduced (or sometimes increased) by a message y, denoted by $J(y, \eta)$, is

(4) $$J(y, \eta)\equiv H(s)-H(s/y, \eta).$$

This amount is not necessarily non-negative. In cases of confusing messages, the degree of uncertainty may increase as a result of receiving the message.

The quantity $J(y, \eta)$ is an *ex post* concept in the sense that it can be calculated only

[1] This section relies heavily on the work by Lindley (9).

[2] See, for example, Shannon and Weaver (16), Marschak (10), Lev (8).

after a particular message y is known to the decision maker. But, as is evident in the definition of the value of an information system in (1), we are interested mainly in finding the *ex ante* quantity to make any judgment on an information system, not on a particular message. Thus, let us take the expected value of $J(y, \eta)$ and denote it by $I(\eta)$.

(5) $I(\eta) \equiv H(s) - \sum_{s \in S} \sum_{y \in Y} P(y/s, \eta)\varphi(s)H(s/y, \eta)$

Defining

(6) $M(s, \eta) \equiv \sum_{s \in S} \sum_{y \in Y} P(y/s, \eta)\varphi(s)H(s/y, \eta)$

we have

(7) $I(\eta) = H(s) - M(s, \eta)$.

$M(s, \eta)$ is the expected value of the degree of the remaining uncertainty under an information system η and is called equivocation in information theory.

Thus, $I(\eta)$ represents the expected amount of reduction of uncertainty concerning s which is provided by having an information system η. This is our definition of the amount of information provided by an information system. The same quantity appears in information theory and is called the rate of transmission of information along a channel (or information system in our terminology).

$I(\eta)$ has several nice properties as a measure of the amount of information of an information system. The first property to note here is its nonegativity. That is, on the average or *ex ante* basis, an information system always reduces (not increases) the degree of uncertainty concerning s, although the degree of uncertainty may increase or decrease upon receiving any particular message. More formally,[3]

Property 1: $I(\eta) \geq 0$, with equality if and only if $P(y/s, \eta)$ does not depend on s.

The condition for $I(\eta)=0$ essentially says that if and only if messages are generated from η in such a way with no probabilistic connection to the states of nature (or y and s are independent) the amount of information is zero. This is certainly the way a measure of the amount of information should behave.

The second property is the addivity of the amounts of information of two information systems η_1 and η_2. These two information systems are characterized by two probability distributions of messages, $P(y_1/s, \eta_1)$ and $P(y_2/s, \eta_2)$. Now, let us consider an information system whereby one can receive a pair of messages (y_1, y_2) simultaneously. Let us denote this coupled information system by $\eta_1 \cup \eta_2$, its amount of information being $I(\eta_1 \cup \eta_2)$. Now, if we consider two information systems generating messages in tandem, not simultaneously, η_1 being the first of the two, the amount of information of η_1, plus the *additional* amount of information of η_2 after having a message from η_1 would have to be equal to the amount of information of $\eta_1 \cup \eta_2$ to appeal to our intuition of a measurement of quantity. This is exactly the case for our measure of the amount of information.

Property 2: $I(\eta_1 \cup \eta_2) = I(\eta_1) + I(\eta_2/\eta_1)$

Here, $I(\eta_1)$ and $I(\eta_1 \cup \eta_2)$ are defined as usual and $I(\eta_2/\eta_1)$ is defined as

(8) $I(\eta_2/\eta_1) = \sum_{y_1 \in Y} \sum_{s \in S} \varphi(s)P(y_1/s, \eta_1)I(\eta_2/y_1)$

where $I(\eta_2/y_1)$ is the usual definition of I with a prior distribution $\varphi(s/y_1, \eta_1)$ and the message generating probability $P(y_2/s, \eta_2, y_1)$. Thus $I(\eta_2/y_1)$ is the additional amount of information of η_2 after receiving a particular message y_1 from η_1. By averaging it over y_1, we get

[3] For the proof of the following three properties, the reader is referred to Lindley (9).

$I(\eta_2/\eta_1)$. One of the corollaries of this property is that

(9) $I(\eta_1 \cup \eta_2) \geq I(\eta_1)$

because $I(\eta_2/\eta_1) \geq 0$. This implies, quite obviously, that the amount of information of the two information systems coupled together is no less than the amount of information of one of the two information systems alone.

The third property of $I(\eta)$ to be discussed here is concerned with the amounts of information from two independent information systems. We say that η_1 and η_2 are independent when

(10) $P(y_1, y_2/s, \eta_1 \cup \eta_2) = P(y_1/s, \eta_1)P(y_2/s, \eta_2)$.

That is, for any given state of nature, s, the joint probability of messages y_1 and y_2 being generated from $\eta_1 \cup \eta_2$ is the product of the individual message generating probabilities of η_1 and η_2. Note that it does not necessarily follow from (10) that y_1 and y_2 are independent. In general

$$\sum_{s \varepsilon S} P(y_1, y_2/s, \eta_1 \cup \eta_2)\varphi(s) \neq (\sum_{s \varepsilon S} P(y_1/s, \eta_1)\varphi(s))(\sum_{s \varepsilon S} P(y_2/s, \eta_2)\varphi(s))$$

Property 3: If η_1 and η_2 are independent information systems,

$$I(\eta_1) + I(\eta_2) \geq I(\eta_1 \cup \eta_2)$$

with equality if and only if y_1 and y_2 are independent.

This property means that when we have two independent information systems, the sum of the amounts of information of individual information systems usually exceeds the amount of information of the coupled information system. Thus it is better to get messages separately (or as separate drawings of messages from Y) rather than get two messages jointly (or as joint drawings of messages from Y). Only when y_1 and y_2 are (completely) independent, the sum of the amounts of information of individual information systems is equal to the amount of information of the coupled information systems.

A corollary of this property, coupled with Property 2, is that if η_1 and η_2 are independent

(11) $I(\eta_2) \geq I(\eta_2/\eta_1)$,

with equality if and only if y_1 and y_2 are independent. The implication of (11) is that the amount of information of an information system, say η_2, decreases if used after having another independent information srstem, say η_1, compared to the case when η_2 is used alone. If we take analogy with the production process, considering information systems as a kind of input into the production of information, (11) indicates diminishing marginal productivity. This point is most apparent if we take the case $\eta_1 = \eta_2$ (that is, getting two messages from the same information system). In this case, (11) clearly indicates diminishing marginal productivity of the same information system input in the production of information where the quantity produced is measured by $I(\eta)$.

Having introduced the concept and the operational definition of the amount of information provided by an information system and shown some of its properties, the next task of this paper is to investigate its relationship to the value of an information system in the information economics context. One of the important points which distinguishes $I(\eta)$ from $V(\eta)$ is that $I(\eta)$ is essentially a value-free quantity whereas $V(\eta)$ inescapably depends on the decision maker's utility function. $I(\eta)$ is also free from the alternative courses of action available to the decision maker, whereas $V(\eta)$ depends on A. One important point which is common to $I(\eta)$ and $V(\eta)$ is that both depend heavily on the prior distribution, $\varphi(s)$, and therefore the set of states of nature, S, which is considered relevant to (a) particular decision situation (s).

The aforementioned differences between $I(\eta)$ and $V(\eta)$ may tempt one to doubt the usefulness of the concept of the amount of information, given the present trends toward greater emphasis on user or decision model orientation in management accounting and information choice.[4] Yet, there seems to be at least two benefits in introducing $I(\eta)$ in the information economics framework, particularly in accounting information evaluation. One of the benefits may be called a rather practical one and its case rests on the very fact that $I(\eta)$ is not so much decision-model oriented as directly as $V(\eta)$. In that sense, $I(\eta)$ is more objective (though still involving a prior, $\varphi(s)$) than $V(\eta)$. When we try to make some judgment on rather institutional (and in many senses social) matters like the selection of accounting alternatives, too much emphasis on the information needs of each individual decision makers will in general lead to impossible selection or no selection, as is most vividly shown in Demski [5]. His result is an almost inevitable consequence of using $V(\eta)$ as a criterion of information choice. Perhaps we should settle for something less subjective as a selection criterion of information systems, although this less ambitious attitude will perhaps make a most ardent individualist among information evaluators gripe very much. If the value of an information system is to be considered too subjective, the next thing which will come to mind as a selection criterion would be the amount of information of an information system. The concept introduced in this section is just that and some reasonable properties that $I(\eta)$ has are also shown.

In the following sections, I will further show that there are reasons to believe that $I(\eta)$ are related to $V(\eta)$ in some reasonable ways, if one ever wants to consider $I(\eta)$ only as a proxy for $V(\eta)$ in information evaluation. Indeed, if one considers the practical difficulty of actually computing $V(\eta)$ which at least involves two difficult tasks of finding the utility function and the computation of optimum, one may as well settle for $I(\eta)$ as a criterion of information choice as long as there is some theoretical guarantee that rankings of information systems by $V(\eta)$ and $I(\eta)$ are (almost) identical or very closely related. In this context, the fact that $I(\eta)$ still depends on a prior, $\varphi(s)$, may be considered as a positive factor because it means that $I(\eta)$ at least maintains some user-orientation. Furthermore, as Bayesian probability theory tells us, $\varphi(s)$ will converge to what may be called "objective probability" if repeated observations and probability revisions are allowed. Thus we may consider the dependence of $I(\eta)$ on $\varphi(s)$ as somewhat user-oriented but not so subjective as the dependence of $V(\eta)$ on the utility function.

Another potential benefit of introducing the amount of information within the information economics framework is a theortical one. Usually in most economic analysis when one talks and analyzes value or utility of an object, one has a quantity measure of an object, just like in the theory of consumer's behavior where a utility function usually takes as its arguments variables representing the quantities of each goods consumed. Thus, when one deals with the value or utilities of apples, for example, one does not usually have to deal with each apple individually. Instead, one often measures them in number or pounds, and then starts analysis. Marginal utility of apples is meaningful only when we have some quantity measure of apples.

Basically, one of the reasons for often implicit importance of quantity measures in economic analysis is that reasonably common analysis of an object is made possible by hav-

[4] For example, American Accounting Association [2].

ing a quantity measure without being entangled in minor individual differences of each item of an object. Going back to the example of apples, there are certain economic analysis possible for apples as a whole, regardless whether it is from Vermont, New York, or California. And to be able to do that, a quantity measure seems to be indispensable. Perhaps, similar arguments may be made for economic analysis of information and information systems. The introduction of the amount of information may serve as a vehicle to some common analysis and understanding across various information systems regardless of their individual details and differences. Here again, the major point is that the present framework of information economics may be too individualistic, thus perhaps hindering potential common analysis. I shall come back to this point, a theoretical benefit of introducing the amount of information into the information ecomonics context, later in the next section.

3. *Relationship between the Value of Information and the Amount of Information: An Example*

In this section, the relationship between $V(\eta)$ and $I(\eta)$ is investigated through an example. Since $V(\eta)$ and $I(\eta)$ defined in the previous section both depend on the same elements ($\varphi(s)$, $P(y/s, \eta)$), it is clear that there is some mapping from $I(\eta)$ to $V(\eta)$ for a given decision maker to relate these two concepts. The question is the kinds of properties that this mapping may have. In the following example, it is shown that this mapping is a quite nice and intuitively appealing function. However, under what general conditions these nice properties are generalizable is not yet known and should be one of the major targets of further reaesrch.

Now, let us consider a case of a forecaster who has to supply a forecast of s which can take any numerical value. His prior distribution on s is described by a normal distribution with mean μ and precision (reciprocal of variance) d. The forecaster's action is a forecast, a, for s. His loss function (or negative of utility function) is quadratic, i.e.,

$$U(s, a)=-(a-s)^2$$

For taking his action, a, the forecaster has an information system, η, available to him which can supply n samples for s as its messages, $y_1, y_2, \ldots, y_n$. Suppose that for each y_1 of n sample messages, $P(y_1/s, \eta)$ is also a normal distribution with the same mean s and the same precision h and messages y_1's are independent of each other.

Under these conditions, the posterior distribution of s after receiving messages ($y_1, \ldots, y_n$) from η is also a normal distribution.[5] Its mean is

$$\frac{d\mu+nh\bar{y}}{d+nh}$$

where $\bar{y}=\frac{1}{n}\sum_{i=1}^{n} y_i$, and its precision is $d+nh$. The optimal action is to equate a to the mean of the distribution, thus making the maximal expected utility a negative of the variance of the distribution.[6] Thus we obtain

$$V(\eta)=-\frac{1}{(d+nh)}-\left(-\frac{1}{d}\right)$$

[5] See DeGroot [3], p. 167.
[6] DeGroot [3], p. 228.

Therefore,

(12) $V(\eta)=\dfrac{nh}{d(d+nh)}$

On the other hand, the entropy of a normal distribution with mean μ and precision d is[7]

(13) $H(s)=\dfrac{1}{2}\log\dfrac{2\pi e}{d}$

Likewise,

$$H(s/y,\ \eta)=\frac{1}{2}\log\frac{2\pi e}{d+nh}$$

Since $H(s/y,\ \eta)$ does not depend on y,

(14) $M(s,\ \eta)=\dfrac{1}{2}\log\dfrac{2\pi e}{d+nh}$

From (13), (14) and (7) we get

(15) $I(\eta)=\dfrac{1}{2}\log\dfrac{d+nh}{d}$

Now, from (12) and (15), we can derive the functional relationship between $V(\eta)$ and $I(\eta)$ as follows.[8] From (15)

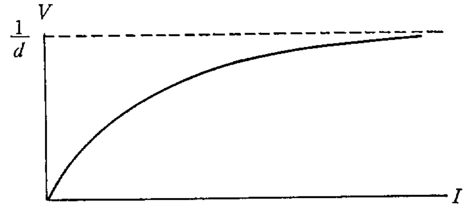
$$nh=d(e^{2I}-1)$$

Substituting this into (12), we get

(16) $V=\dfrac{1}{d}(1-e^{-2I})$

That is, at least in this example, the value of information is a unique and monotone increasing function of the amount of information, as is perhaps desirable. It is also a concave function, meaning that the law of diminishing marginal value is at work between the value and amount of information. Graphic representation of this functional relationship is given in Figure 1.

FIG. 1



The fact that the value of information depends only on the amount of information of an information system[9] seems to have great significance. First of all, it implies that an information system has value solely because of its ability to reduce the degree of uncertainty by probability revision. This seems to be a generalizable fact as far as the value of informa-

[7] Lev [8], p. 56.

[8] For notational convenience, $V(\eta)$ and $I(\eta)$ are written sometimes simply as V and I in the following.

[9] As far as the characteristics of an information system is concerned. Of course, V depends on d and the form of the loss function.

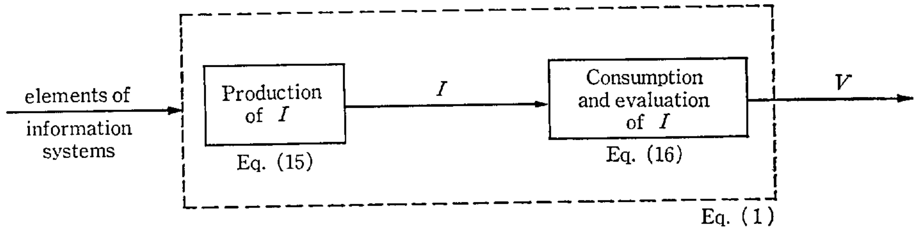
tion in information economics context is concerned. Secondly, a ranking of information systems by V are identical to the one by I because V is a monotone increasing function of I. This further implies that selection information systems may be made entirely on the basis of the amount of information provided by information systems, although the cost of information systems have to be considered separately. Thus, $I(\eta)$ may be used as a surrogate criterion for the value of information. In the above example, alternative sets of information systems available to the forecaster may be different combinations of the sample size (n) and precision (h). A more concrete example may be a choice among different market reserach strategies with differing sample sizes and reliability of each sample, or a choice among different cost estimates. At any rate, different n and h have different effects on the value of information entirely through their effects on the amount of information. Choice may be made without referring to the loss or utility function of the forecaster. As mentioned in the pervious section, non-reliance of an information system choice on the utility function of the receiver of the messages is certainly appealing, if at least in a practical sense.

The functional relationship in (16) or its graphical representation in Figure 1 may be considered a relationship that corresponds to a utility function of usual goods or commodity in economic theory. Both of them represent how much utility or value a decision maker attaches to certain amount of goods or information. As indicated above, the function in (16) exhibits properties similar to a usual utility function in economic theory. That is, $\frac{\partial V}{\partial I}>0$ and $\frac{\partial^2 V}{\partial I^2}<0$. Furthermore, there is a saturation level, $\frac{1}{d}$, in Figure 1, which is also found in some utility functions in the theory of consumers' behavior. The fact that $V=0$ when $I=0$ also reinforces the similarity. Thus, by introducing the concept of the amount of information, possibility is suggested of treating an information system in the same way as ordinary goods and using the same analytical frameworks or approaches of various branches of economic theory.

Continuing the analogy with usual economic analysis, the equation (15) may be considered as a kind of production function representing the process of producing I amount (or bits) of information from the two production inputs, the sample size (n) and the sample precision (h). In contrast the equation (16) represents the utility function of the end-user or consumer of the goods called information, in the amount of I. As can be easily checked, the information production function (15) has properties similar to the production functions commonly used in economic analysis (for example, Cobb-Douglas production function). As the amounts of inputs vary (n and h), (15) exhibits $\frac{\partial I}{\partial n}>0$, $\frac{\partial I}{\partial h}>0$ and $\frac{\partial^2 I}{\partial n^2}<0$, $\frac{\partial^2 I}{\partial h^2}<0$. Namely, the production process of the amount of information from two inputs has a property of decreasing marginal productivity. Also, when both $n=0$ and $h=0$, $I=0$, implying that with no input there is no output. One important point to note in this connection is that the information production function (15) still depends on the prior distribution, through the prior precision d in this example. Common production functions in economic analysis do not depend on who the end user of the produced goods is. It is perhaps on this point that economic analysis of information differs significantly from economic analysis of ordinary goods.

Thus, usual information systems evaluation in a decision theoretic framework like in (1) may be said to be a comprehensive framework of evaluation which goes straight from the primitive inputs of information systems (e.g., n and h) to the value of information systems to the end user. Instead, we may break down the evaluation process behind (1) into two processes: the production process and the consumption (and evaluation) process of the produced information. Diagramatically, it may look like the following.

FIG. 2



Clearly, the usual approach of information economics is to treat the dotted box as a single process without further breakdown.

Yet, the breakdown suggested in Figure 2 seems to be particularly relevant to accountants whose job is to produce and supply information, not to consume one. Perhaps we need to understand more how we are producing information and how much information can be produced with a certain configuration of information systems elements and so on. That is, the understandings of the left box in solid line. These understandings seem to be desirable before we embark on a more ambitious task of how the produced information satisfy the information needs of the decision maker and thus how information systems are evaluated by the decision maker, although there is no arguing about the fact that this is the ultimate objective of information analysis. It is also true that the suggested breakdown of total information analysis may lead to suboptimal analysis. But, just as there usually is division of labor in an organization between the decision maker and the accountant (i.e., the consumer of information and the producer of information), division of labor in information analysis may be more efficient in the final analysis. Price for the breakdown (or division of analysis) may be worth paying if we consider the seeming difficulties of carrying out total information analysis (dotted box) in both practical and theoretical domains.

4. *Relationship Between the Value of Information and the Amount of Information: A General Theorem*

In the previous section, the value of information V has been shown to be a function of the amount of information I through an example. A point is emphasized, among other things, that there is a possibility that we can use the amount of information I as a surrogate of V in selecting different information alternatives. In this section a general theorem supporting, at least partially, this possibility will be shown.

The theorem is the following.[10]

[10] For proof, the reader is referred to Lindley [9], Theorem 9, or Marschak and Miyasawa [11], section 12.

Theorem: Consider two information systems η_1 and η_2. If $V(\eta_1) \geq V(\eta_2)$ for any utility function U and prior distribution φ, then $I(\eta_1) \geq I(\eta_2)$ for any prior distribution.

When $V(\eta_1) \geq V(\eta_2)$ regardless of the utility function and prior distribution, η_1 is sometimes called "more informative" than η_2. The above theorem says that for η_1 to be more informative than η_2, it is necessary that η_1 provides greater amount of information than η_2 regardless of the prior distribution. Although a greater amount of information is not a sufficient condition for greater informativeness, the above theorem gives us one minimal test in terms of the amount of information that an information system has to pass to be more informative than others.

Perhaps, the greatest significance of the theorem for accounting information evaluation emerges in connection with the search for an optimal accounting information system for an unidentified user. Simply stated, an information system is said to be optimal for an unidentified user if it is more informative (in the above sense) than any other alternative information system.[11] As already mentionad in Section 2, accounting information systems are often institutions whether they are for external reporting purposes or for internal uses by management, and have in many cases to be designed without any clear-cut single user as the ultimate judge of the value of information systems. Although the requirements of an optimal accounting information system for an unidentified user are rather strong, the above theorem indicates that the amount of information can be used as a check for the necessary condition. Alternatively, the theorem assures us that an information system has to have the greatest amount of information among alternative information systems to be ever optimal for an unidentified user. Here is another reason why it is perhaps sensible to use the amount of information as a surrogate of the value of information in information systems choice when a utility function is not so clearly given to an information system evaluator.

Thus, this theorem demonstrates one of the ways how the concept of the amount of information is useful and viable for information analysis and evaluation.

5. *Conclusion*

In this paper, some cases have been presented for introducing the concept of the amount of information into the information economics framework of information evaluation. Accounting information systems as (social) institutions seem to be in need of a quantity measure free of value judgment, if only, at present stage, to further their theoretical analysis. Since the purpose of this paper has been largely introductory, any specific result directly linked to accounting information systems is not presented. Further research efforts have to follow in order to indicate in any definite sense whether the concept of the amount of information is really viable and useful.[12] Yet, at least, it seems to indicate one direction of future research in accounting information evaluation.

---

[11] See Marshall [13]. He has shown that the optimality condition in his sense is equivalent to the statistical sufficiency condition.

[12] Results of further mathematical analyses on the relationship between the value and the amount of information will appear in the author's forthcoming paper, "On the Relationship between the Value and the Amount of Information" (to appear in the next issue of this journal).

REFERENCES

1. Abdel-khalik, A. Rashad, "The Entropy Law, Accounting Data, and Relevance to Decision-Making," *The Accounting Review*, April 1974, pp. 271–283.
2. American Accounting Association, *A Statement of Basic Accounting Theory*, Sarasota, Florida: American Accounting Association, 1966.
3. DeGroot, Morris H., *Optimal Statistical Decisions*, New York: McGraw-Hill, 1970.
4. Demski, Joel S., *Information Analysis*, Reading, Massachusetts: Addison-Wesley, 1972.
5. Demski, Joel S., "The General Impossibility of Normative Accounting Standards," *The Accounting Review*, October 1973, pp. 718–723.
6. Feltham, Gerald A., *Information Evaluation*, Sarasota, Florida: American Accounting Association, 1972.
7. Lee, Lucy C., and Norton M. Bedford, "An Information Theory Analysis of the Accounting Process," *The Accounting Review*, April 1969, pp. 256–275.
8. Lev, Baruch, *Accounting and Information Theory*, Sarasota, Florida: American Accounting Association, 1969.
9. Lindley, D.V., "On a Measure of the Information Provided by an Experiment," *Annals of Mathematical Statistics*, Vol. 27, 1956, pp. 986–1005.
10. Marschak, Jacob, "Economics of Information Systems," *Journal of American Statistical Association*, March 1971, pp. 192–219.
11. Marschak, Jacob, and K. Miyasawa, "Economic Comparability of Information Systems," *International Economic Review*, June 1968, pp. 137–174.
12. Marschak, Jacob, and Roy Radner, *Economic Theory of Teams*, New Haven, Connecticut: Yale University Press, 1972.
13. Marshall, Ronald M., "Determining an Optimal Accounting Information System for an Unidentified User," *Journal of Accounting Research*, Autumn 1972, pp. 286–307.
14. Nakano, Isao, "Noise and Redundancy in Accounting Communications," *The Accounting Review*, October 1972, pp. 693–708.
15. Ronen, Joshua, and Gideon Falk, "Accounting Aggregation and the Entropy Measure: An Experimental Approach," *The Accounting Review*, October 1973, pp. 696–717.
16. Shannon, Claude E., and W. Weaver, *The Mathematical Theory of Communication*, Urbana, Illinois: The University of Illinois Press, 1964.