Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Letter to the Editor—Information Theory and Search Theory as Special Cases of Decision Theory

Donald F. Mela,

INFORMATION THEORY AND SEARCH THEORY AS SPECIAL CASES OF DECISION THEORY

Donald F. Mela

Operations Evaluation Group, Washington, D C

(Received August 7, 1961)

IT IS fairly widely realized at this time that a procedure that yields the highest probability of detection in a search problem will not necessarily be the one that maximizes the expected information gained, as given by the standard definition of information theory (Reference 1 is the only document in which the writer has seen this fact in print) It is the purpose of this note to indicate with some simple examples that the connection between the information-theory approach and the search-theory approach is tenuous and that search problems can better be regarded as an application of the more general theory of statistical decisions

As a first example of a case where maximizing information does not lead to the correct search procedure, consider the following Suppose that an object is equally likely to be in one of two areas, and that the probability of detection for one look in an area is ½ Suppose that two looks can be made If we look twice in one area the probability of seeing the object is ¾ when the object is in the area and 0 when the object of search is not in the area, so that the probability of detection is ½ ¾+½ 0=⅜ On the other hand if we take one look in each area, the probability is ½ ½+½ ½=½ that the object will be detected

Next consider the information gains from the two possible procedures Suppose we look twice in area A If the object is not found, then the a posteriori probability that the object is in area A is ⅕ and the a posteriori probability that the object is in area B is ⅘ If the object is found, the probabilities are 1 and 0 Then the expected gain in information is

$$\Delta(2,0)=I(\tfrac{1}{2},\tfrac{1}{2})-[\tfrac{3}{8}\,I(1,0)+\tfrac{5}{8}\,I(\tfrac{1}{5},\tfrac{4}{5})]=0\ 380,$$

where

$$I(p_1, p_2, \quad , p_n)\equiv-\textstyle\sum_{i=1}^{i=n} p_i \log p_i$$

A similar calculation for the case of one look in each area yields an information increase of

$$\Delta(1,1)=I(\tfrac{1}{2},\tfrac{1}{2})-[\tfrac{1}{2}\,I(1,0)+\tfrac{1}{2}\,I(\tfrac{1}{2},\tfrac{1}{2})]=0\ 347$$

Thus the procedure of two looks in one area, which yields more information, results in the lower detection probability

On the other hand it is easy to change this example to one where the procedure that yields more information is preferred Suppose that an action is to be based upon the results of the two looks We might wish to commit forces to one area on the basis of the looks If the object is discovered by the two looks, then of course the forces can be committed correctly If the object is not found after two looks, then the posterior probability distribution is (1/2,1/2) if each area has received one look, and it is (1/5,4/5) if area *A* has received two looks Then the probability of committing the forces correctly is calculated easily and turns out to be 3/4 if each area is looked at once, and 7/8 if one area receives both looks Thus in the present case, the policy that leads to more information gain is also the policy that leads to the better decision

A slightly more complicated version of this example provides a situation in which the best course is neither to maximize detection probability nor to maximize information gain Suppose that the object lies with probability 1/3 in one of three areas, and that three looks are available, each look producing a detection probability of 1/2 Again suppose that the problem is to commit a force correctly on the basis of the three looks There are three procedures available

(3, 0, 0) look three times in one area
(2, 1, 0) look twice in one area, once in another
(1, 1, 1) look once in each area

The following table summarizes the calculations similar to those for the preceding example

Procedure	Detection probability	Expected information gain	Probability of correct commitment
(3, 0, 0)	7/24	0 478	5/8
(2, 1, 0)	5/12	0 541	3/4
(1, 1, 1)	1/2	0 549	2/3

In this case, the procedure that maximizes detection probability also produces the maximum expected information increase, but is not the procedure yielding the highest probability of correct commitment of forces

From the foregoing examples it does not seem likely that there is any intimate connection between search theory and information theory This becomes clearer when the search problems are considered as a part of the more general theory of statistical decisions In a typical decision problem, one of a set of information-gathering processes may be employed, which transforms our initial or a priori knowledge about a situation into a final or a posteriori knowledge The knowledge is represented generally by a priori and a posteriori probability distributions The choice of the information-gathering process (search or experiment) depends upon the value (utility) and the probabilities of the various possible a posteriori distributions

For example, referring to the first example above, one procedure is to look twice in area *A* This results in the a posteriori target distribution (1,0) with probability 3/4 and the a posteriori distribution (1/5,4/5) with probability 1/4 If the objective is detection with certainty, then the first distribution has value 1 and the second has value 0 If the objective is information increase, then the a posteriori distributions are valued using the function $-\sum p_i \log p_i$ Finally if the objective is to commit a force to the correct area on the basis of an a posteriori distribution (p_1,p_2), the distribution (p_1,p_2) has the value $\max(p_1,p_2)$, since the best procedure is to commit the force to the area with the higher a posteriori probability of containing the target

Work on search theory, such as that in references 2 and 3, has dealt with the problem of maximizing detection probability However, in many cases some action other than detection is involved, and the information measure may or may not lead to the proper optimization

In conclusion, there appears to be no reason why the information theory measure should be preferred Each problem will by its nature require its special valuation of the possible results of a search Thus 'search theory' should be considered in connection with the general theory of statistical decisions rather than with information theory

REFERENCES

1 Paul B Coggins, "Application of Information Theory to Search," Memorandum for Director of Research, OEG (LO)1975–52, 22 Oct 1952 (Unpublished)

2 B O Koopman, "The Theory of Search," *Opns Res* **4**, 324–346, 503–531, 1956, *Opns Res* **5**, 613–626 (1957)

3 Jacques de Guenin, "Optimum Distribution of Effort An Extension of the Koopman Theory," *Opns Res* **9**, 1–7 (1961)

COMMENTS ON A PAPER BY THOMAS HEALY

John H Roseboom

Booz, Allen Applied Research, Inc , Glenview, Illinois

(Received August 10, 1961)

IN THE May-June issue of Operations Research, Thomas Healy presents findings which he uses to demonstrate that a *prescribed* subdivision of activities will change probabilities in a PERT network analysis *

As a member of the original PERT team, I would like to call attention to and take issue with some aspects of his paper

On a broad level, I do not think that it is generally wise to put unrealistic emphasis on the precise probability numbers yielded in a PERT computation Insights and direction should stem from an observation of the approximate level of the computed probabilities

However, there are other things that should be said about Healy's paper These have a bearing at two levels At a formal level, Healy makes initial assumptions that guarantee his results

The assumption is made that in the process of subdividing, the sum of the optimistic and pessimistic times of the subdivided activities are equal to the op-

* "Activity Subdivision and PERT Probability Statements," **9**, 341–347 (1961)