The Annals of Mathematical Statistics
1971, Vol. 42, No. 5, 1706–1716

HIDE AND SEEK, DATA STORAGE, AND ENTROPY[1]

BY ROBERT J. MCELIECE AND EDWARD C. POSNER
California Institute of Technology

1. Introduction. In this paper we will study the relationship between games of search and the optimum storage of information. In Section 2 we shall treat the case of finite sets, and in Section 3 a generalization to compact metric spaces. The result is a synthesis of the epsilon entropy theory of approximation (Lorentz (1966)), with the theory of data transmission and compression (Posner and Rodemich (1971)).

Let X be a set with $n = |X|$ elements, and let $\mathscr{S} = \{S_1, S, \cdots, S_m\}$ be a finite collection of subsets of X, with $\cup S_j = X$. Regard the $x \in X$ as "data points," and the S_j as "subsets of allowed uncertainty," such that when a data point is selected, one is not interested in exactly which x it is, but rather in knowing an S_j (there may be more than one) in which it lies. The class $\mathscr{S}$ is regarded as chosen by the experimenter.

Under these assumptions, if m' is the fewest number of the S_j's which are needed to cover X, then at least $\lceil \log_2 m' \rceil$ bits are needed in order to identify an S_j in which an arbitrary x lies. The method of specification to achieve $\lceil \log_2 m' \rceil$ is, of course, to merely specify the index j of the set S_j which contains the given data point, and one specifies j by using a $\lceil \log_2 m' \rceil$-tuple of zeros and ones. This problem is considered in Balinsky (1968) pages 214–221.

However, if N data points can be stored before it is attempted to specify a sequence of N S_j's, a saving may be possible. Let X^N be the cartesian Nth power of X, and let $\mathscr{S}^N$ be the class of subsets of X^N of the form $S_{i_1} \times \cdots \times S_{i_N}$. Here $\lceil \log_2 M \rceil$ bits are needed to specify the sequence of sets corresponding to an unknown sequence of N data points, where M is the fewest number of sets from $\mathscr{S}^N$ needed to cover X^N. Thus $1/N\lceil \log_2 M \rceil$ can be interpreted as the number of bits per sample necessary to specify an S_j when a "block code" of (constant) length N is used.

Thus we are led to several definitions, which generalize those in Posner and Rodemich (1971); here is also found more information-theory, as well as a list of prior references for some of these concepts. The (one shot) *$\mathscr{S}$-entropy* $H_{\mathscr{S}}(X)$ is defined as

$$H_{\mathscr{S}}(X) = \min_{\mathscr{T} \subset \mathscr{S}} \log |\mathscr{T}|$$

where the minimization is taken over those subsets of $\mathscr{S}$ which cover X. (We

Received May 19, 1970.

[1] This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. NAS 7–100, sponsored by the National Aeronautics and Space Administration. We are indebted to the referee for pointing out the reference to Integer Linear Programming.