The Annals of Statistics
1973 Vol. 1, No. 4, 729–739

HIDING AND COVERING IN A COMPACT METRIC SPACE[1]

BY ROBERT J. MCELIECE AND EDWARD C. POSNER

Jet Propulsion Laboratory, California Institute of Technology

This paper studies the relationship between games of search on a compact metric space X and the absolute epsilon entropy $I(X)$ of X. The main result is that $I(X) = -\log v_L^*$, v_L^* being the lower value of a game on X we call "restricted hide and seek."

1. Introduction. Let X be a set, S a collection of subsets of X with $\bigcup S = X$. The two-person zero-sum game "hide and seek" $G(X, S)$ is played as follows. Player 1 (the "hider") chooses a point $x \in X$, and player 2 (the "seeker") chooses $s \in S$. If $x \in s$ player 1 pays player 2 one unit; otherwise no payoff occurs. Let us denote the value of this game, if it exists, by v. (We assume that X has enough structure so that mixed strategies can be defined.)

Now for each integer N let c_N be the smallest integer such that the Cartesian power X^N can be covered with c_N sets from S^N, and let $c = \lim_{N\to\infty} c_N^{1/N}$. The main theorem of a previous paper of ours [2] was that if X is finite, $v = c^{-1}$. It is the object of this paper to study the relationship between v and c when X is a compact metric space, and S is the set of closed spheres of radius ε.

Our first main result (Theorem 1) is that in this situation, the game G still has a value. For finite X von Neumann's fundamental theorem of finite two-person zero-sum games immediately implies that v exists, and so in [2] this problem did not arise.

Our second main result is that $c = v^{-1}$ is not true in general, but rather that $c = v^{*-1}$, where v^* is the best expected gain the seeker can guarantee himself when he must restrict his sets to a finite subset of S he has chosen in advance. It is always true that $v^* \leqq v$, and for a fixed X, $v^* = v$ except for at most countably many values of ε. In Section 4, however, we give an example of a compact metric space for which $v^* < v$. In Section 5 we prove that $c = v^{*-1}$.

These problems arise in information theory. The logarithm of the limit c is the least average number of bits per sample necessary to describe X modulo S; i.e., to identify an s containing x, when block coding is used, and when there is no a priori probability distribution on X. We shall show at the end of Section 5 that $-\log v$ represents the maximum, over all Borel a priori probability distributions on X, average number of bits per sample necessary to describe X to within an ambiguity of ε, when variable-length coding is used. Thus when

Received December 1971; revised October 1972.

[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.

$v = c^{-1}$ (the usual state of affairs in spite of our counter example) there exist probability distributions on X which render variable-length coding useless.

2. General hide and seek. If the hider chooses his point x according to a probability distribution λ on (a Borel field containing the points of) X, we say he uses strategy λ. Similarly a strategy μ for the seeker is a probability distribution on (a Borel field containing the points of) S. Let $E = \{(x, s) : x \in s\}$, a subset of the product space $X \times S$. The expected value of the payoff, given that the hider plays strategy λ and the seeker plays μ, is $(\lambda \times \mu)(E) = v(\lambda, \mu)$, $\lambda \times \mu$ being the product measure induced by λ and μ on $X \times S$.

If the hider uses a fixed strategy λ, then from his point of view the worst possible expected payoff is $\sup_\mu v(\lambda, \mu)$. Hence he will choose a λ which makes $\sup_\mu v(\lambda, \mu)$ as small as possible. Thus we define the *upper value* of $G(X, S)$ as

$$v_U(X, S) = \inf_\lambda \sup_\mu v(\lambda, \mu) . \tag{2.1}$$

Similarly the seeker will choose a μ which makes $\inf_\mu v(\lambda, \mu)$ as large as possible, and we define the *lower value* of $G(X, S)$ as

$$v_L(X, S) = \sup_\mu \inf_\lambda v(\lambda, \mu) . \tag{2.2}$$

It is an easy exercise to show that $v_L \leqq v_U$. If it happens that $v_L = v_U$ we denote this common value by $v(X, S)$, and say that the game $G(X, S)$ has a value. If the game has a value, then for every $\eta > 0$, there exist strategies λ and μ such that if the hider plays λ, his expected loss is $\leqq v(X, S) + \eta$ no matter how the seeker plays, and if the seeker plays μ his expected gain is $\geqq v(X, S) - \eta$ no matter how the hider plays. If it happens that there exist strategies λ for the hider which guarantee an expected loss no greater than $v(X, S)$, these strategies are called *optimal* strategies. Optimal strategies for the seeker are defined similarly.

There is another form of the definitions of v_U and v_L which will be useful in what follows. By the definition of product measure we can write $v(\lambda, \mu)$ as either of the integrals

$$\begin{aligned} v(\lambda, \mu) &= \int_X \mu(\text{star}\,(x))\, d\lambda \\ &= \int_S \lambda(s)\, d\mu , \end{aligned} \tag{2.3}$$

where $\text{star}\,(x) = \{s \in S \mid x \in s\}$. Now if we define the *pure strategy* λ_x for the hider as that strategy which always chooses x; i.e., $\lambda_x(x) = 1$, $\lambda_x(x') = 0$ if $x' \neq x$, we see that $\mu(\text{star}\,(x)) = v(\lambda_x, \mu)$. Similarly if μ_s is a pure strategy for the seeker, $\lambda(s) = v(\lambda, \mu_s)$. Thus from (2.3) we obtain the estimate $v(\lambda, \mu) \leqq \sup_{s \in S} \lambda(s) = \sup_{s \in S} v(\lambda, \mu_s)$. Hence for a fixed λ, $\sup_\mu v(\lambda, \mu) = \sup_{s \in S} v(\lambda, \mu_s)$ and so

$$v_U(X, S) = \inf_\lambda \sup_{s \in S} \lambda(s) . \tag{2.1$'$}$$

Similarly

$$v_L(X, S) = \sup_\mu \inf_{x \in X} \mu(\text{star}\,(x)) . \tag{2.2$'$}$$

Let us remark finally that if the set X is finite, it is a consequence of von Neumann's fundamental theorem of finite two-person, zero-sum games that $G(X, S)$ has a value ([4], Chapter 3).

3. Hide and seek in a compact metric space. For the remainder of the paper X will be a compact metric space and S will be the set of closed spheres of radius ε. (However, most of the results to be proved also hold when S is the set of closed sets of diameter $\leqq \varepsilon$.) The sphere of radius ε around x will be denoted by $s_\varepsilon(x) = \{y \in X: d(y, x) \leqq \varepsilon\}$. This game is denoted by $G(X, \varepsilon)$. In this case strategies for the hider and the seeker will both be Borel probability distributions on X, since the seeker need only specify the center of the sphere he wishes to select. In the product space $X \times X$, the set $E = \{(x, y): d(x, y) \leqq \varepsilon\}$, and for strategies λ and μ, $v(\lambda, \mu) = (\lambda \times \mu)(E)$. Before proceeding we need a result on weak convergence.

Let $B(X)$ be the space of all Borel probability distributions on X, $C(X)$ the space of real-valued continuous functions on X. The topology of weak convergence on $B(X)$ is defined as follows ([5], Chapter II): $\mu_n \rightarrow \mu$ in $B(X)$ if for every $f \in C(X)$ $\int f d_{\mu_n} \rightarrow \int f d_\mu$. $B(X)$ is compact in this topology ([5], page 45) and if F is any closed subset of X and $\mu_n \rightarrow \mu$, then

$$\mu(F) \geqq \limsup_{n\rightarrow\infty} \mu_n(F) \tag{3.1}$$

([5], page 40).

We now consider probability distributions on the product space $X \times X$.

LEMMA 1. *If $\mu_n \rightarrow \mu$ and $\lambda_n \rightarrow \lambda$ then $\mu_n \times \lambda_n \rightarrow \mu \times \lambda$.*

PROOF. This follows immediately from the Stone-Weierstrass Theorem, which guarantees that the functions of the form

$$\sum_{i=1}^{k} f_i(x) g_i(x)\,, \qquad f_i, g_i \in C(X)$$

are dense in $C(X \times X)$ under the sup norm.

LEMMA 2. *If $\lambda_n \rightarrow \lambda$ and $\mu_n \rightarrow \mu$, then* $v(\lambda, \mu) \geqq \limsup_{n\rightarrow\infty} v(\lambda_n, \mu_n)$.

PROOF. From Lemma 1, $\lambda_n \times \mu_n \rightarrow \lambda \times \mu$. Since $v(\lambda, \mu) = (\lambda \times \mu)(E)$ the result follows from property (3.1).

We can now prove the main theorem of this section.

THEOREM 1. *$G(X, \varepsilon)$ has a value $v(\varepsilon)$ which is continuous from above in ε, and the seeker has an optimal strategy. For every $\delta > 0$ the hider has a strategy with finite support which guarantees that he loses no more than $v(\varepsilon) + \delta$. The set of optimal strategies for the seeker is closed in the topology of weak convergence.*

PROOF. Let $\{x_j; j \geqq 1\}$ be a countable dense subset of X, and for each $n \geqq 1$ let $G_n(X, \varepsilon)$ be the game $G(X, \varepsilon)$ with the hider restricted to $x_1, x_2, \cdots, x_n$. Let $B_n(X)$ represent the strategies allowed to the hider in $G_n(X, \varepsilon)$. Define

$$v_n(\varepsilon) = \sup_{\mu \in B(X)} \inf_{\lambda \in B_n(X)} v(\lambda, \mu)\,. \tag{3.2}$$

According to our discussion in Section 2, $v_n(\varepsilon)$ is the largest expected payoff the seeker can guarantee himself if the hider is restricted to $x_1, x_2, \cdots, x_n$. Thus the $v_n(\varepsilon)$ decrease with n and so approach a limit which we call $v(\varepsilon)$. In Section 2 we saw that the inner "inf" in (3.2) could be replaced by an "inf" over the pure strategies; and in this case there are only finitely many pure strategies $\lambda_1, \lambda_2, \cdots, \lambda_n$ so that "inf" can be replaced by "min". Next, for fixed n choose a sequence of seeker's strategies $\{\mu_m\}$ such that $\min_{\lambda \in B_n(x)} v(\lambda, \mu_m) \geqq v_n(\varepsilon) - 1/m$. Since $B(X)$ is compact in the weak topology there will exist a convergent subsequence $\mu_{m_k} \to \mu$ say. Then from $v(\lambda, \mu_{m_k}) \geqq v_n(\varepsilon) - 1/m_k$ for all λ it follows that $v(\lambda, \mu) \geqq \limsup_{k\to\infty} v(\lambda, \mu_{m_k}) \geqq v_n(\varepsilon)$ from Lemma 2. Hence $\min_\lambda v(\lambda, \mu) = v_n(\varepsilon)$ and so we may rewrite (3.2) as

$$v_n(\varepsilon) = \max_{\mu \in B(X)} \min_{\lambda \in B_n(X)} v(\lambda, \mu) . \tag{3.3}$$

Let μ_n achieve the maximum in (3.3). Then $v(\lambda_j, \mu_n) \geqq v_n(\varepsilon)$ for all $j \leqq n$, where λ_j is the pure strategy which always chooses x_j. Let $\mu \in B(X)$ be a limit point of the sequence, such that the sequence $\mu_{n_k} \to \mu$. For a fixed j, if k is sufficiently large, $v(\lambda_j, \mu_{n_k}) \geqq v_{n_k}(\varepsilon)$. Hence another application of Lemma 2 implies $v(\lambda_j, \mu) \geqq v(\varepsilon)$ for all j, and so $v(\lambda, \mu) \geqq v(\varepsilon)$ for any λ whose support is contained in $\{x_j\}$. But such measures are dense in $B(X)$ ([5], page 44). Thus another application of Lemma 2 yields $v(\lambda, \mu) \geqq v(\varepsilon)$ for all $\lambda \in B(X)$. Thus $v_L(X, \varepsilon) \geqq v(\varepsilon)$ by (2.2).

To see that $v_U(X, \varepsilon) \leqq v(\varepsilon)$, observe that $G_n(X, \varepsilon)$ is essentially a finite game. For the set of spheres of radius ε, $s_\varepsilon(x_i)$, $i = 1, 2, \cdots, n$, induces a partition of X into a finite number $m \leqq 2^n$ of Borel sets $E_1, \cdots, E_m$. If $y_i \in E_i$ are fixed representatives of the E_i, then any $\mu \in B(x)$ has the same expected payoff $v(\lambda_j, \mu)$ as the finitely based strategy $\tilde{\mu}$ defined by $\tilde{\mu}(y_i) = \mu(E_i)$, $\mu = 0$ otherwise. Hence, according to von Neumann's theorem, $G_n(X, \varepsilon)$ has a value which must of course be $v_n(\varepsilon)$ by (3.2). By playing an optimal strategy λ_n in G_n, the hider can keep the expected payoff down to $v_n(\varepsilon)$. Since $v_n(\varepsilon) \downarrow v(\varepsilon)$, for every $\delta > 0$, by choosing n large enough, the hider can keep the expected payoff down to $v(\varepsilon) + \delta$. Hence $v_U \leqq v(\varepsilon)$ by (2.1) and so $v(\varepsilon)$ is the value of $G(X, \varepsilon)$.

To prove that the set of optimal strategies for the seeker is closed, suppose μ_n is optimal; i.e., $v(\lambda, \mu_n) \geqq v(\varepsilon)$ for all $\lambda \in B(X)$, and $\mu_n \to \mu$. By Lemma 2, $v(\lambda, \mu) \geqq \limsup_{n\to\infty} v(\lambda, \mu_n) \geqq v(\varepsilon)$ for all λ and so μ is optimal also.

Finally, we must show that $v(\varepsilon)$ is continuous from above in ε. It will be enough to prove that if $\lambda \in B(X)$ and $\delta > 0$ is given, there is an $\eta > 0$ such that $\max_x \lambda(s_{\varepsilon+\eta}(x)) \leqq \max_y \lambda(s_\varepsilon(y)) + \delta$. For then by taking $\inf_\lambda$ on both sides we obtain $v(\varepsilon + \eta) \leqq v(\varepsilon) + \delta$.

Assume by way of contradiction that there exists $\delta > 0$ such that for every $\eta > 0$ there exists $x_\eta \in X$ with $\lambda(s_{\varepsilon+\eta}(x_\eta)) > \max_y \lambda(s_\varepsilon(y)) + \delta$. Choose $x_n \to x$ and $\eta_n \to 0$ such that $\lambda(s_{\varepsilon+\eta_n}(x_n)) > \max_y \lambda(s_\varepsilon(y)) + \delta$. Now for every $\eta > 0$ for sufficiently large n, $s_{\varepsilon+\eta}(x) \supseteqq s_{\varepsilon+\eta_n}(x_n)$, and so $\lambda(s_{\varepsilon+\eta}(x)) \geqq \lambda(s_{\varepsilon+\eta_n}(x_n))$. Thus for every $\eta > 0$, $\lambda(s_{\varepsilon+\eta}(x)) \geqq \max_y \lambda(s_\varepsilon(y)) + \delta$. But $\lambda(s_\varepsilon(x)) = \lim_{\eta\to 0} \lambda(s_{\varepsilon+\eta}(x))$ by countable additivity of λ, which is a contradiction. This completes the proof of Theorem 1.

Note. We have recently learned that much of Theorem 1 (the existence of $v(\varepsilon)$, the seeker's optimal strategy, and the "δ-optimal" hider's strategy) follows from known theorems of game theory. In particular Theorem 2.9.2 and the remarks following it on page 86 of [1] can be combined with our equation (3.1) to give a shorter proof.

We conclude this section with two examples which show the necessity of certain hypotheses in Theorem 1.

EXAMPLE 1. The hider need not have an optimal strategy. X will be a countable subset of the unit circle, under the geodesic metric. Let $X_n = \exp(\pi i/2^{n+1})$. X will consist of the points $\pm x_n$, $\pm ix_n$ for all $n \geqq 0$. Then X is closed and so compact. Let $\varepsilon = \pi/2$; for each $x \in X$ we adopt the abbreviation $s(x) = s_{\pi/2}(x)$. Then if the seeker plays ± 1 each with probability $\frac{1}{2}$ his expected gain against any pure hider's strategy will be $\geqq \frac{1}{2}$ and so $v_L \geqq \frac{1}{2}$. On the other hand, if the hider uses the strategy λ_N defined by $\lambda_N(x) = 1/(2N)$ for $x = \pm x_1, \pm x_2, \cdots, \pm x_N$; $\lambda_N(x) = 0$ otherwise, then $\lambda_N(s) = \frac{1}{2} + 1/(2N)$ if $x = \pm ix_n$ for some $n \leqq N$; $\lambda_N(s(x)) = \frac{1}{2}$ otherwise, and so the hider's expected loss is $\leqq \frac{1}{2} + 1/(2N)$ for any pure seeker's strategy. Thus $v_U \leqq \frac{1}{2} + 1/(2N)$ for any N, and so $G(X, \pi/2)$ has value $\frac{1}{2}$. If, however, the hider had an optimal strategy λ, $\lambda(s(x)) \leqq \frac{1}{2}$ for all $x \in X$, then it would follow from $\lambda(s(x)) + \lambda(s(-x)) = 1 + \lambda(ix) + \lambda(-ix)$ that $\lambda(ix) = \lambda(-ix) = 0$ for all $x \in X$, a contradiction.

EXAMPLE 2. The set of optimal strategies for the hider, if nonempty, need not be closed. Let X be the closed interval $[0, 4]$ under the usual metric, and $\varepsilon = 1$. Then $v(X, \varepsilon) = \frac{1}{2}$, and if λ_n is the strategy $\lambda_n(0) = \lambda_n(2 + 1/n) = \frac{1}{2}$, λ_n is optimal for all $n > 1$. However $\lambda_n \to \lambda$ where $\lambda(0) = \lambda(2) = \frac{1}{2}$, but λ itself is not optimal since if the seeker always picks the sphere centered at 1 his gain against λ is always 1.

EXAMPLE 3. The seeker need not have finitely based nearly optimal strategies such as the hider has; i.e., it is possible that there exists $\delta > 0$ such that if μ is any finitely based strategy (a probability distribution on X which is zero outside a finite subset of X), then $\mu(s_\varepsilon(x)) \leqq v(\varepsilon) - \delta$ for some $x \in X$. This example is best understood in the context of a game we call "restricted hide and seek," introduced in the next section, so we postpone it until then.

4. Restricted hide and seek. In *restricted* hide and seek, the seeker is restricted to finitely based strategies. Of course there is no way the hider can tell if the seeker is cheating and using a strategy with infinite support, so a referee will be needed who knows the support of the seeker's strategy. We denote this game by $G^*(X, \varepsilon)$ and define $v_L^*(X, \varepsilon)$ and $v_U^*(X, \varepsilon)$ as in Section 2. We frequently adopt abbreviated notations such as $v_L^*(X)$, $v_L^*(\varepsilon)$ or simply v_L^* when no confusion can arise. Let $v(\varepsilon-) = \lim_{\varepsilon' \to \varepsilon-} v(\varepsilon')$.

LEMMA 3. $v(\varepsilon-) \leqq v_L^*(\varepsilon) \leqq v_U^*(\varepsilon) = v(\varepsilon)$.

PROOF. That $v_L^* \leqq v_U^*$ follows from the discussion in Section 2. Also, by (2.1′) we know that $v_U^* = \inf_\lambda \sup_{s \in S} v(\lambda, \mu_s)$ since the pure strategy μ_s is certainly finitely based. But this is also $v_U = v$.

It remains to show that $v(\varepsilon-) \leqq v_L^*(\varepsilon)$. Fix η, $0 < \eta < \varepsilon$ and let E_i, $i = 1, 2, \cdots, K$ be a disjoint family of measurable subsets of X whose union is X and $E_i \subseteqq s_\eta(x_i)$ for suitable $x_i \in E_i$. Let μ be an optimal seeker's strategy for $G(X, \varepsilon - \eta)$, and let $\bar{\mu}$ be the finitely based strategy $\bar{\mu}(x_i) = \mu(E_i)$, $i = 1, 2, \cdots, K$, $\bar{\mu}(x) = 0$ otherwise. Also, for any $x \in X$ let $\{E_j^x\}$ be a minimal cover of $s_{\varepsilon-\eta}(x)$ by the sets E_i. Then $s_{\varepsilon-\eta}(x) \subseteqq \bigcup E_j^x \subseteqq s_\varepsilon(x)$ and so $\mu(s_{\varepsilon-\eta}(x)) \leqq \mu(\bigcup E_j^x) = \bar{\mu}(\bigcup E_j^x) \leqq \bar{\mu}(s_\varepsilon(x))$. Taking the inf of this last equation over $x \in X$ we obtain $v(\varepsilon - \eta) \leqq \inf_x \bar{\mu}(s_\varepsilon(x)) \leqq v_L^*(\varepsilon)$. Since this is true for all $\eta > 0$, $v(\varepsilon-) \leqq v_L^*(\varepsilon)$ is asserted.

LEMMA 4. $v_L^*(\varepsilon) = v(\varepsilon)$ *with at most countably many exceptions.*

PROOF. $v(\varepsilon)$ is monotone non-decreasing in ε and so is continuous with at most countably many exceptions. Since $v(\varepsilon-) = v(\varepsilon)$ if v is continuous at ε, Lemma 4 follows from Lemma 3.

If X has only two points x, y and $d(x, y) = 1$, then $v(X, 1-) = \frac{1}{2}$ but $v(X, 1) = v_L^*(X, 1) = 1$. It is much more difficult to give an example which shows that v_L^* may be strictly less than v_L. We now give such an example.

EXAMPLE 3. There exists a compact metric space X such that $v(X, 1-) < v_L^*(X, 1) < v(X, 1)$.

Let C be a circle of circumference 4, d the geodesic metric on C, and let H_C be the space of closed subsets of C under the Hausdorff metric d':

$$d'(E, F) = \max\left(\max_{e \in E} \min_{f \in F} d(e, f), \max_{f \in F} \min_{e \in E} d(e, f)\right).$$

H_C is a compact metric space under d' [3]. The set Z of all closed subsets of C of Lebesgue measure 2 is a closed, hence compact, subspace of H_C and is therefore separable. Let $\{B_i, i \geqq 1\}$ be a countable dense subset of Z. No finite subset $\{b_k\}$ of C has the property that every B_i contains a b_k, for $\{b_k\}$ can be covered by an open set of arbitrarily small Lebesgue measure and so there exists a set $B \in Z$ and $d_0 > 0$ such that $d(B, b_k) \geqq d_0$ for all k. Thus a B_i such that $d'(B, B_1) < d_0$ cannot contain a b_k.

The space X of this example will have C as a subspace, the metric restricted to C being the geodesic metric. It also contains points a; a_i $i \geqq 1$ where

$$\begin{aligned} d(a, c) &= 1 \qquad \text{for} \quad c \in C \\ d(a, a_i) &= 2^{-i} \\ d(a_i, a_j) &= |2^{-i} - 2^{-j}|\,. \\ d(a_i, c) &= 1 + \min(d(B_i, c), 2^{-i-1}) \qquad \text{for all} \quad c \in C\,. \end{aligned}$$

In addition X contains three points c_1', c_2', c_3' which are to be thought of as outside the circle C and equally spaced in angle. The point on C closest to c_i' is labeled c_i. The metric is extended as follows:

$$\begin{aligned} d(c_i', a) &= d(c_i', a_j) = \tfrac{15}{8} && \text{for all } i, j, \\ d(c_i', c_i) &= \tfrac{7}{8}, \\ d(c_i', c) &= \tfrac{7}{8} + d(c_i, c) && \text{if } d(c_i, c) \leqq \tfrac{1}{8} \\ &= 1 && \text{if } \tfrac{1}{8} \leqq d(c_i, c) \leqq 1 \\ &= d(c_i, c) && \text{if } d(c_i, c) \geqq 1 \quad \text{for } c \in C, \\ d(c_i', c_j') &= 2, \quad i \neq j. \end{aligned}$$

We assert that (X, d) as defined above is indeed a compact metric space, but omit the tedious verification that d satisfies the triangle inequality. Compactness is best verified by checking sequential compactness, which is equivalent to compactness for a metric space. The space X is sketched in Figure 1, topologically embedded in three-space.

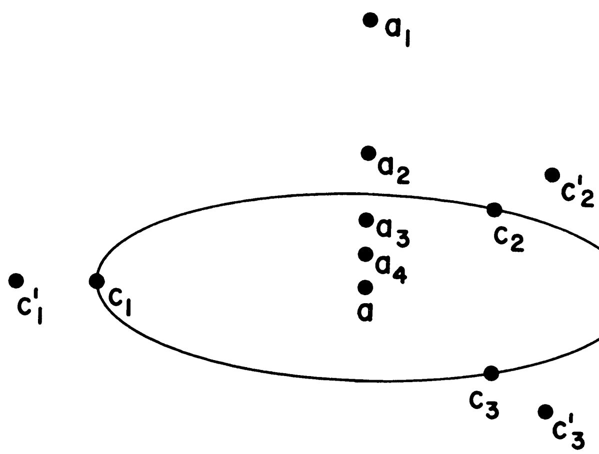

FIG. 1. The space X of counterexample 3.

We now produce Example 3 by showing that

$$v(X; 1-) = \tfrac{1}{3}, \qquad v_L^*(X, 1) = \tfrac{2}{5}, \qquad v(X, 1) \geqq \tfrac{1}{2}.$$

The last inequality is easily obtained: if the seeker plays $\frac{1}{4}$ Lebesgue measure on C his expected gain is at least $\frac{1}{2}$ against any pure hider's strategy. (Notice also that if the hider plays c_1', c_2', c_3' with probability $\frac{1}{7}$ each and C with $\frac{1}{4}$ Lebesgue measure with probability $\frac{4}{7}$ he will keep his expected loss $\leqq \frac{4}{7}$. Thus, $\frac{1}{2} \leqq v \leqq \frac{4}{7}$. The exact value of the game apparently depends in detail upon the choice of the sets B_i.)

To compute $v(1-)$, observe that the hider can hold his losses to at most $\frac{1}{3}$ by playing a and two antipodal points on C with probability $\frac{1}{3}$ each. But the seeker can guarantee himself an expected gain of $\frac{1}{3}$ by playing a with probability $\frac{1}{3}$ and $\frac{1}{4}$ Lebesgue measure on C with probability $\frac{1}{3}$.

Finally we compute v_L^*. The key property of the metric d is that for any finite subset $\{b_k\}$ of C there exists a j such that $d(a_j, b_k) > 1$ for all k. Hence for any strategy μ of the seeker

$$\min_{x \in X} \mu(s_1(x)) \leqq \mu(\{a\} \cup \bigcup_{i=1}^{\infty} \{a_i\}) = \alpha \quad \text{say.}$$

Since $s_1(a) \supseteqq s_1(a_j)$ for $j \geqq 1$ there is no point in the seeker choosing a strategy with $\mu(a_j) > 0$ for any j. Thus we assume $\mu(a) = \alpha$, $\mu(a_j) = 0$ for $j \geqq 1$. Furthermore, we assume $\mu(c_i') = 0$ since $s_1(c_i') \subseteqq s_1(c_i)$ $i = 1, 2, 3$. Now

$$\sum_{i=1}^{3} \mu(s_1(c_i')) \leqq 2(1 - \alpha)$$

since each point of C is contained in at most two of the sets $s_1(c_i')$. Thus for some i, $\mu(s_1(c_i')) \leqq \frac{2}{3}(1 - \alpha)$ and so

$$\min_{x \in X} \mu(s_1(x)) \leqq \min\left(\tfrac{2}{3}(1 - \alpha), \alpha\right);$$

$$\max_{\mu} \min_{x \in X} \mu(s_1(x)) \leqq \max_{0 \leqq \alpha \leqq 1} \min\left(\tfrac{2}{3}(1 - \alpha), \alpha\right) = \tfrac{2}{5}.$$

But with $\mu(a) = \frac{2}{5}$, $\mu(\bar{c}_i) = \frac{1}{5}$, $i = 1, 2, 3$ where the $\bar{c}_i$ are points on C halfway between the c_i, the seeker will win at least $\frac{2}{5}$ and so $v_L^* = \frac{2}{5}$ is asserted.

5. Absolute epsilon entropy. Let us use the term "ε-set" to describe a subset of a compact metric space which is contained in some sphere of radius ε. The *epsilon entropy* $H_\varepsilon(X)$ is then defined to be $\log_2 N$, where X can be covered with N ε-sets, but no fewer. $H_\varepsilon(X)$ can be interpreted information theoretically as the minimum average number of bits per sample needed to describe X to within an error of at most ε.

If (X_i, d_i) $i = 1, 2, \cdots, n$ are compact metric spaces we shall make the Cartesian product $X_1 \times X_2 \times \cdots \times X_n$ into a compact metric space by defining $d((x_1, \cdots, x_n), (x_1', \cdots, x_n')) = \max_i d(x_i, x_i')$. With this definition products of ε-sets are ε-sets and projections of ε-sets onto the coordinate spaces X_i are ε-sets; hence it is a suitable definition for dealing with uniform approximation. If $X_i = X$ for all i we shall write X^n instead of $X_1 \times \cdots \times X_n$.

The *absolute epsilon entropy* $I_\varepsilon(X)$ is defined by

$$I_\varepsilon(X) = \lim_{n \to \infty} \frac{1}{n} H_\varepsilon(X^n).$$

That the limit exists is a consequence of the simple property $H_\varepsilon(X^{n+m}) \leqq H_\varepsilon(X^n) + H_\varepsilon(X^m)$. $I_\varepsilon(X)$ can be interpreted as the minimum average number of bits per sample needed to describe X to within ε when an unlimited number of samples can be stored prior to transmission.

Theorem 2, the main result of this paper, identifies $I_\varepsilon(X)$ in terms of the game "restricted hide and seek".

THEOREM 2. $I_\varepsilon(X) = -\log v_L^*(X; \varepsilon)$.

Before giving the proof of Theorem 2 we need two lemmas.

LEMMA 5. $H_\varepsilon(X) \geqq -\log v_L^*(X; \varepsilon)$.

PROOF. Let U be a covering of X with N ε-sets, which we assume are in fact spheres of radius ε, $H_\varepsilon(X) = \log N$. Let V be the refinement of U to a partition of X, $V = \{v_i : 1 \leqq i \leqq M\}$, where $M \leqq 2^N$, and choose a point x_i in each v_i. Then the probability distributions on V and those on X with support $\{x_i\}$ are in obvious 1-1 correspondence.

If P is any probability distribution on V, then for some $u \in U$, $P(u) \geqq 1/N$ since $\sum P(u) \geqq 1$, U being a cover of X. Hence $\min_P \max_{u \in U} P(u) \geqq 1/N$; i.e.,

$$N \geqq 1/v(V; U) ,$$

where $v(V; U)$ is the value of the game $G(V; U)$. But

$$\begin{aligned} v(V; U) &= \max_{P \in B(U)} \min_{v \in V} P(\text{star}\,(v)) \\ &= \max_{P \in B(U)} \min_{x \in X} P(\text{star}\,(x)) , \end{aligned}$$

where star $(x) = \{u \in U : x \in u\}$. If we define $\mu_P \in B(X)$ by $\mu_P(y_i) = P(s_\varepsilon(y_i))$ if y_i is a center for one of the elements of U, $\mu_P(y) = 0$ otherwise, then μ_P is finitely based and $P(\text{star}\,(x)) = \mu_P(s_\varepsilon(x))$. Therefore

$$v(V; U) = \max_{P \in B(U)} \min_{x \in X} \mu_P(s_\varepsilon(x)) .$$

But since $v_L{}^*$ is the sup of $\min_{x \in X} \mu(s_\varepsilon(x))$ over all finitely based $\mu \in B(X)$, it thus follows that

$$v_L{}^*(X; \varepsilon) \geqq v(U; V) \geqq 1/N$$

from which Lemma 5 follows.

LEMMA 6. $v_L{}^*(X \times Y, \varepsilon) = v_L{}^*(X, \varepsilon)\, v_L{}^*(Y, \varepsilon)$.

PROOF. Let μ_x and μ_y be finitely based strategies for the seeker in X and Y such that $\inf_{x \in X} \mu_x(s_\varepsilon(x)) \geqq v_L{}^*(X) - \eta$ and $\inf_{y \in Y} \mu_y(s_\varepsilon(y)) \geqq v_L{}^*(Y) - \eta$. Then since $s_\varepsilon(x) \times s_\varepsilon(y) = s_\varepsilon(x, y)$ in $X \times Y$, it follows that $\inf_{(x,y) \in X \times Y} \mu_x \times \mu_y(s_\varepsilon(x, y)) \geqq v_L{}^*(X) v_L{}^*(Y) - 2\eta + \eta^2$. Since $\mu_x \times \mu_y$ is finitely based and η is arbitrary, this shows $v_L{}^*(X \times Y) \geqq v_L{}^*(X) v_L{}^*(Y)$.

Now let μ be any finitely based strategy on $X \times Y$. Then the support of μ is contained in a finite set of the form $\{x_1, \cdots, x_m\} \times \{y_1, \cdots, y_m\}$. Define marginal strategies μ_X and μ_Y by $\mu_X(S) = \mu(S \times Y)$ if $S \subseteqq X$ and $\mu_Y(T) = \mu(X \times T)$ if $T \subseteqq Y$. Since μ_X is finitely based there exists a point $x_0 \in X$ such that $\mu_X(s_\varepsilon(x_0)) \leqq v_L{}^*(X)$; similarly there is a $y_0 \in Y$ such that $\mu_Y(s_\varepsilon(y_0)) \leqq v_L{}^*(Y)$. Then

$$\mu(s_\varepsilon(x_0, y_0)) = \mu_X(s_\varepsilon(x_0)) \mu_Y(s_\varepsilon(y_0)) \leqq v_L{}^*(X) v_L{}^*(Y) ,$$

and so $v_L{}^*(X \times Y) \leqq v_L(X) v_L(Y)$. This completes the proof of Lemma 6.

We now complete the proof of Theorem 2.

By Lemma 5, $H_\varepsilon(X^N) \geqq -\log v_L{}^*(X^N, \varepsilon)$. By Lemma 6 $v_L{}^*(X^N, \varepsilon) = (v_L{}^*(X, \varepsilon))^N$ and so $H_\varepsilon(X^N) \geqq -N \log v_L{}^*(X, \varepsilon)$. Hence from the definition of $I_\varepsilon(X)$ we obtain

$$I_\varepsilon(X) \geqq -\log v_L{}^*(X, \varepsilon) .$$

To prove the opposite inequality let $\eta > 0$ be arbitrary and let μ be a probability distribution on X with a finite base $\{x_1, x_2, \cdots, x_n\}$, such that

$$\min_{x \in X} \mu(s_\varepsilon(x)) \geqq v_L^* - \eta . \tag{5.1}$$

Let $s_i = s_\varepsilon(x_i)$ and let $U = \{s_1, s_2, \cdots, s_n\}$. Let $V = \{v_1, v_2, \cdots, v_m\}$ be the partition of X induced by U. The value of the finite game $G(V, U)$ is given by

$$v(V, U) = \max_{P \in B(U)} \min_{v \in V} P(\text{star}\,(v)) . \tag{5.2}$$

If c_N denotes the smallest integer such that V^N can be covered with c_N subsets from U^N, the main result of [2] implies that

$$\lim_{N\to\infty} \frac{1}{N} \log c_N = -\log v(V, U) . \tag{5.3}$$

Now since coverings of V^N with elements of U^N induce in an obvious way coverings of X^N with spheres of radius ε, (5.3) implies that

$$I_\varepsilon(X) \leqq -\log v(V, U) . \tag{5.4}$$

Next define $Q \in B(U)$ by $Q(s_i) = \mu(x_i)$. Then clearly $Q(\text{star}\,(v_i)) = \mu(s_\varepsilon(x))$ if $x \in v_i$, and so from (5.1) and (5.2) we see that

$$v(V, U) \geqq \min_{v \in V} Q(\text{star}\,(v)) = \min_{x \in X} \mu(s_\varepsilon(x)) \geqq v_L^* - \eta .$$

Thus from (5.4) we obtain $I_\varepsilon(X) \leqq -\log (v_L^*(X, \varepsilon) - \eta)$. But η was arbitrary and so

$$I_\varepsilon(X) \leqq -\log v_L^*(X, \varepsilon) ,$$

and this completes the proof of Theorem 2.

We conclude the paper with two corollaries to Theorem 2. Let p be a Borel probability measure on X, and let $H_{\varepsilon;p}(X)$ be the infimum, over all partitions $X = \bigcup_i A_i$, $A_i \cap A_j = \emptyset$ if $i \neq j$, each A_i being a Borel ε-set of X, of the Shannon entropy $-\sum_i p(A_i) \log p(A_i)$. $H_{\varepsilon;p}$ is called the $\varepsilon; p$ entropy of X [6]. Also define the absolute $\varepsilon; p$ entropy of X by

$$I_{\varepsilon;p}(X) = \lim_{n\to\infty} \frac{1}{n} H_{\varepsilon;p^n}(X^n) ,$$

p^n being the product measure induced on X^n by p. $I_{\varepsilon;p}(X)$ then represents the minimum average number of bits per sample necessary to describe X with an error not exceeding ε, with p as the a priori probability distribution on X, when arbitrarily long variable-length codes are used. Combining Theorem 2 with Theorem 2 of [2], which had $-\log v(X, \varepsilon) = \sup_p I_{\varepsilon;p}(X)$, we conclude:

COROLLARY 1. $I_\varepsilon(X) = \sup_p I_{\varepsilon,p}(X)$ *whenever* $v_L^*(X, \varepsilon) = v(X, \varepsilon)$; *in particular equality holds for all but at most countably many* ε.

Hence most of the time "nature" can choose p on X which is so "bad" that prior knowledge of p could not be used to increase the transmission rate.

Our final result is a simple consequence of Theorem 2 and Lemma 6, and tells us that one cannot save anything by encoding two sources simultaneously:

COROLLARY 2. $I_\varepsilon(X \times Y) = I_\varepsilon(X) + I_\varepsilon(Y)$.

Acknowledgmcnt. The authors wish to thank Professors Thomas Ferguson and Shmuel Zamir for many helpful suggestions.

REFERENCES

[1] FERGUSON, T. S. (1966). *Mathematical Statistics.* Academic Press, New York.

[2] MCELIECE, R. J. and POSNER, E. C. (1971). Hide and seek, data storage, and entropy. *Ann. Math. Statist.* **42** 1706–1716.

[3] MICHAEL, E. (1951). Topologies on spaces of subsets. *Trans. Amer. Math. Soc.* **71** 151–182.

[4] VON NEUMANN, J. and MORGENSTERN, O. (1947). *Theory of Games and Economic Behavior,* (2nd ed.). Princeton Univ. Press.

[5] PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces.* Academic Press, New York.

[6] POSNER, E. C. and RODEMICH, E. R. (1971). Epsilon entropy and data compression. *Ann. Math. Statist.* **42** 2079–2125.

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
4800 OAK GROVE DRIVE
PASEDENA, CALIFORNIA 91103