

SOME STRUCTURE THEOREMS FOR STATIONARY PROBABILITY MEASURES ON FINITE STATE SEQUENCES¹

BY CHARLES HOBBY AND N. DONALD YLVIKAKER

University of Washington

1. Introduction. Given a stationary probability measure P on doubly infinite sequences over finitely many states, there are measures P_n induced on n consecutive coordinates. These measures P_n are again stationary to the extent that stationarity requirements are satisfied over subsets of n consecutive coordinates. Conversely, given a family of such "stationary" measures $\{P_n\}$ which are consistent, there is an induced stationary measure P on infinite sequences. Generally, given a fixed P_n , there are many stationary measures P which reduce to P_n . We shall consider the restrictions imposed on P by choosing P_n . In the main, our results relate the growth of the number of elements in the support of P_{n+k} to the choice of P_n .

In Section 2, we state some basic properties of these support numbers. Deterministic measures are defined and the structure imposed on P by a deterministic measure P_n is noted. In Section 3, beginning with a fixed P_n , we consider certain extremal measures P which reduce to P_n and show that there is always a deterministic measure P which reduces to P_n . In Section 4, we examine the largest possible growth of support numbers for a fixed P_n . An example is given of a measure P which has the property that its restrictions P_n are concentrated on precisely $n + 1$ vectors.

2. Preliminaries. Let X^n be the n th Cartesian product of the set $X = \{0, 1, \dots, q - 1\}$, let X^I be the space of doubly infinite sequences $\{X(n)\}$ whose coordinates take values in the set X and let \mathcal{G} denote the σ -field of subsets of X^I generated by the cylinder sets. Given a stationary probability measure P on (X^I, \mathcal{G}) we induce a probability measure on X^n by restricting P to sets of the form $\{x(1) = x_1, \dots, x(n) = x_n\}$. This restricted measure will be denoted by P_n and usually abbreviated by writing

$$p_n(x_1, \dots, x_n) = P\{x(1) = x_1, \dots, x(n) = x_n\}.$$

For each choice of n , these measures satisfy

$$\begin{aligned} (*) \quad p_n(x_1, \dots, x_n) &= \sum_{x_0=0}^{q-1} p_{n+1}(x_0, x_1, \dots, x_n) \\ &= \sum_{x_{n+1}=0}^{q-1} p_{n+1}(x_1, \dots, x_n, x_{n+1}) \end{aligned}$$

Received 5 July 1963; revised 20 September 1963.

¹ Prepared with the partial support of the National Science Foundation, Grants G-19915 and G-25211.

for all $\mathbf{x} = (x_1, \dots, x_n) \in X^n$. Conversely, if we are given a family of measures $\{P_n\}$ on X^n which satisfy (*) for each choice of n , it follows from the consistency theorem that a stationary probability measure P is induced on (X^I, \mathcal{G}) and that the restrictions of this P are the measures P_n .

Given a stationary probability measure P on (X^I, \mathcal{G}) , we let N_n be the number of elements of X^n in the support of P_n . That is, N_n is the number of $\mathbf{x} \in X^n$ for which $p_n(\mathbf{x}) > 0$. We clearly have $1 \leq N_n \leq q^n$ and $N_{n+m} \leq N_n N_m$. It is also easy to show that $N_n \leq N_{n+1}$ and if $N_{n+1} = N_n$, then $N_{n+k} = N_n$ for all $k \geq 1$. For the special case $N_n = N_{n+1}$, we make the following

DEFINITION 1. The stationary probability measure P on (X^I, \mathcal{G}) is said to be n -deterministic if $N_{n-1} < N_n = N_{n+1}$, and is said to be non-deterministic if it is not n -deterministic for some n .

An n -deterministic measure P must consequently be concentrated on at most q^n infinite sequences. This structure is summarized in

THEOREM 1. *If P is an n -deterministic stationary probability measure on (X^I, \mathcal{G}) , there is a subset B of X^I containing at most q^n sequences, each of which is periodic of period at most q^n , with $P(B) = 1$.*

The existence of n -deterministic measures P which are supported on sequences of period q^n is shown in another context by Good [2].

3. Extensions. We begin here by considering measures P_k on X^k which are restrictions to X^k of stationary probability measures on (X^I, \mathcal{G}) . These restrictions satisfy equations of the form (*) where P_n is the restriction of P_k in turn to sets of the form

$$\{x(1) = x_1, \dots, x(n) = x_n\}, \quad n < k.$$

DEFINITION 2. The restriction P_{n+k} of a stationary probability measure P is said to be an extension of P_n if P_n is the restriction of P_{n+k} to X^n .

It should be noted that restriction always refers to cylinder sets on consecutive integers.

Every solution of the equations (*) gives an extension of the restricted measure P_n . We consider first two special ways of constructing an extension P_{n+1} . (These extensions will, of course, agree if $N_{n-1} = N_n$.)

(a) *A Markov type extension.* For each (x_2, \dots, x_n) in the support of P_{n-1} , we define

$$p_{n+1}(x_1, \dots, x_{n+1}) = \frac{p_n(x_1, \dots, x_n)p_n(x_2, \dots, x_{n+1})}{p_{n-1}(x_2, \dots, x_n)}.$$

This extension gives the largest possible value of N_{n+1} .

(b) *A minimal extension.* For each (x_2, \dots, x_n) in the support of P_{n-1} , we wish to assign probabilities $p_{n+1}(x_1, \dots, x_{n+1})$ so that P_{n+1} is an extension of P_n . Let N_{x_2, \dots, x_n} be the number of choices of x_{n+1} for which $p_n(x_2, \dots, x_{n+1}) > 0$ and let N_{x_2, \dots, x_n}^* be the number of choices of x_1 for which $p_n(x_1, \dots, x_n) > 0$. Nonzero probabilities can be assigned only to the $N_{x_2, \dots, x_n} \cdot N_{x_2, \dots, x_n}^*$

elements of the form (x_1, \dots, x_{n+1}) for which both $p_n(x_1, \dots, x_n) > 0$ and $p_n(x_2, \dots, x_{n+1}) > 0$. We must satisfy the equations

$$\sum_{x_1=0}^{q-1} p_{n+1}(x_1, \dots, x_{n+1}) = p_n(x_2, \dots, x_{n+1})$$

$$\sum_{x_{n+1}=0}^{q-1} p_{n+1}(x_1, \dots, x_{n+1}) = p_n(x_1, \dots, x_n).$$

These equations are precisely the transportation problem equations of linear programming with the interpretation that an amount $p_{n+1}(x_1, \dots, x_{n+1})$ is sent from the origin (x_1, \dots, x_n) to the destination (x_2, \dots, x_{n+1}) , (cf. [3], p. 274). In this context it is known that there are non-negative solutions with at most $N_{x_2, \dots, x_n} + N_{x_2, \dots, x_n}^* - 1$ nonzero choices. Choose then an assignment with the minimum number of positive choices. It follows that

$$N_{n+1} \leq \sum_{\substack{(x_2, \dots, x_n) \\ p_{n-1}(x_2, \dots, x_n) > 0}} (N_{x_2, \dots, x_n} + N_{x_2, \dots, x_n}^* - 1) = 2N_n - N_{n-1}$$

and consequently $N_{n+1} - N_n \leq N_n - N_{n-1}$.

Each of the extensions described above gives a method of extending a measure P_n to a stationary probability measure on (X^I, \mathcal{A}) . The Markov extension gives a measure which is n -step dependent. It is not known whether a repeated minimal extension must terminate the growth of N_{n+k} at some future stage. However, we shall see later (Theorem 3) that every measure P_n does have a deterministic extension.

We examine first the conditions a set S of vectors must satisfy if they are to form the support of some P_n . Consider the mapping ψ_S of S to S defined by

$$\psi_S(x_1, x_2, \dots, x_n) = (x_2, x_3, \dots, x_{n+1}).$$

The mapping ψ_S is at most q -valued. Furthermore, it follows from the recurrence property of stationary measures applied to an extension of P_n to (X^I, \mathcal{A}) that $\mathbf{x} \in \bigcup_{k>0} \{\psi_S^k \mathbf{x}\}$ for all $\mathbf{x} \in S$. In particular, if the sequence $\mathbf{x} = \mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{x}_{k+1} = \mathbf{x}$ consists of distinct \mathbf{x}_i and if $\mathbf{x}_{i+1} \in \{\psi_S \mathbf{x}_i\}$ for $i = 0, \dots, k$ then the subset $S^* = \{\mathbf{x}, \mathbf{x}_1, \dots, \mathbf{x}_k\}$ of S is the support of the measure R_n^* defined by $p_n^*(\mathbf{x}_i) = 1/(k + 1), i = 0, 1, \dots, k$. It is easy to see that the equations (*) hold for P_n^* . We say that a subset S of X^n is *minimal* if S supports some P_n while no proper subset of S supports any P_n .

LEMMA 1. *If the subset S of X^n is minimal, then S supports a unique measure $P_n(S)$. The measure $P_n(S)$ is k -deterministic, $k \leq n - 1$, and it assigns a constant measure over S .*

PROOF. We see from the minimality of S that any sequence $\mathbf{x} = \mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{x}_{k+1} = \mathbf{x}$ with $\mathbf{x}_{i+1} \in \{\psi_S \mathbf{x}_i\}, i = 0, 1, \dots, k$, must contain all of S . If for some $i < j, \mathbf{x}_{j+1} \in \{\psi_S \mathbf{x}_i\}$ then the subsequence $\bar{\mathbf{x}} = \mathbf{x}_0, \dots, \mathbf{x}_i, \mathbf{x}_{j+1}, \dots, \mathbf{x}_k, \mathbf{x}_{k+1} = \mathbf{x}$ has the property that each element is ψ_S of its predecessor. This would give a proper subset of S which could support a measure P_n , contrary to

the minimality of S . Therefore $\mathbf{x}_{j+1} \notin \{\psi_S \mathbf{x}_i\}$ for $i \neq j$, whence any P_n supported on S is k -deterministic, $k \leq n - 1$. It follows easily that such a P_n assigns constant weight to \mathbf{x}_i , $i = 0, 1, \dots, k$.

The unique measure $P_n(S)$ will be called a minimal n -measure. Corresponding to $P_n(S)$ there is a uniquely determined measure $P(S)$ on (X^I, \mathcal{G}) . We shall now show that any measure P_n can be decomposed as a convex combination of minimal n -measures.

THEOREM 2. *If P_n is the restriction of a stationary measure on (X^I, \mathcal{G}) then there are minimal n -measures $P_n(S_\alpha)$, $\alpha = 1, \dots, N$ corresponding to minimal subsets S_α of X^n such that $P_n = \sum_{\alpha=1}^N \lambda_\alpha P_n(S_\alpha)$, where $\lambda_\alpha > 0$ and $\sum_{\alpha=1}^N \lambda_\alpha = 1$.*

PROOF. Let $\mathbf{x} = \mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{x}_{k+1} = \mathbf{x}$ be a sequence drawn from the support S of P_n as before and suppose this sequence has minimal length for all such sequences with $\mathbf{x} \in S$. The set $S_1 = \{\mathbf{x}, \mathbf{x}_1, \dots, \mathbf{x}_k\}$ is clearly minimal. Let p denote the smallest of the numbers $p_n(\mathbf{y})$ for $\mathbf{y} \in S_1$. Define P_n^1 by the equation

$$P_n = (k + 1)pP_n(S_1) + (1 - (k + 1)p)P_n^1.$$

The measure P_n^1 satisfies $(*)$ since both P_n and $P_n(S_1)$ do. The support of P_n^1 does not contain any element \mathbf{y} of S_1 for which $p_n(\mathbf{y}) = p$, hence is a proper subset of the support of P_n . The existence of the required decomposition now follows by induction on the size of the support of P_n .

The fact that every measure P_n has a deterministic extension is an immediate consequence of Theorem 2. We state the result as

THEOREM 3. *If P_n is the restriction of a stationary measure on (X^I, \mathcal{G}) then P_n is the restriction of a measure P of the form $P = \sum_{\alpha=1}^N \lambda_\alpha P(S_\alpha)$, where $\lambda_\alpha > 0$ and $\sum_{\alpha=1}^N \lambda_\alpha = 1$, and where the $P(S_\alpha)$ are the unique extensions of minimal n -measures.*

We note that the representation of P in Theorem 3 implies that P is $(n + k)$ -deterministic if k is the smallest value for which the $P_{n+k}(S_\alpha)$ are supported on mutually disjoint subsets of X^{n+k} . This must occur for some $k \leq \max_\alpha \tilde{S}_\alpha - 1$ where \tilde{S}_α is the number of vectors in S_α . Since each S_α corresponds to an at most $(n - 1)$ -deterministic measure, $\tilde{S}_\alpha \leq q^{n-1}$, and hence $k \leq q^{n-1} - 1$. That is, P_n has an extension P which is $(n + k)$ -deterministic for some $k \leq q^{n-1} - 1$. There are, in fact, minimal n -sets of q^{n-1} vectors, one corresponding to each $(n - 1)$ -deterministic measure of the type constructed by Good. Indeed, they are quite numerous (cf. [1]).

4. The growth of N_n . In this section we shall examine the effect of a fixed measure P_n on the growth of the sequence $\{N_{n+k}\}$. Two extreme possibilities will be considered.

The Markov type extension of the fixed measure P_n has the property that it maximizes N_{n+k} over all extensions P_{n+k} of P_n . If P_n is such that $N_n = q^n$, a Markov extension gives $N_{n+k} = q^{n+k}$ for all k . A more interesting problem arises when $N_n < q^n$ for here there are impossible sequences, each of which leads separately to many impossible sequences at higher stages.

Let $\mathbf{x}_1, \dots, \mathbf{x}_{N_n}$ be the elements of X^n in the support of P_n . Let A be the $N_n \times N_n$ matrix of elements a_{ij} where

$$a_{ij} = 1 \text{ if there is an element of } X^{n+1} \text{ beginning with } \mathbf{x}_i \text{ and ending with } \mathbf{x}_j, \\ = 0 \text{ otherwise.}$$

The i th row sum of A is the quantity $N_{\mathbf{x}_i}$ in the first Markov extension while the j th column sum is $N_{\mathbf{x}_j}^*$. The behaviour of A is similar to that of a 1-step transition matrix for the Markov extension. In particular, the i, j th element in A^k is the number of elements of X^{n+k-1} beginning with \mathbf{x}_i and ending with \mathbf{x}_j and we have $N_{n+k} = \mathbf{1}'A^k\mathbf{1}$ where $\mathbf{1}' = (1, \dots, 1)$. This observation enables us to find the limiting value of $(\log N_n)/n$ for the Markov extension in terms of A . This limiting behaviour is suggested in [4], p. 8 but we give a detailed proof here.

THEOREM 4. *For the Markov extension, $\{N_{n+k}\}$ satisfies a recurrence relation which is the minimal equation of A , and $\lim_m (\log N_m)/m = \log \lambda$ where λ is the largest real eigenvalue of A .*

PROOF. A satisfies its minimal equation and the equation remains valid if we multiply by A^{k-1} . Multiplying this equation on the left by $\mathbf{1}'$ and on the right by $\mathbf{1}$, we obtain the recurrence relation satisfied by the N_{n+k} 's. To obtain the second conclusion, suppose first that A is indecomposable. By the Frobenius theorem, if λ is the spectral radius of A , then λ is a simple eigenvalue of A and has an associated positive eigenvector u . There exist positive numbers c and d for which $cu \leq \mathbf{1} \leq du$. Consequently

$$c^2 u' A^k u = c^2 \lambda^k \leq \mathbf{1}' A^k \mathbf{1} \leq d^2 u' A^k u = d^2 \lambda^k,$$

and therefore

$$\frac{\log c^2}{n+k} + \frac{k \log \lambda}{n+k} \leq \frac{\log N_{n+k}}{n+k} \leq \frac{\log d^2}{n+k} + \frac{k \log \lambda}{n+k}.$$

The conclusion is then immediate when A is indecomposable. If A is decomposable, $\mathbf{x}_1, \dots, \mathbf{x}_{N_n}$ can be separated into equivalence classes of communicating states. The matrix A will contain an indecomposable submatrix corresponding to each of these equivalence classes. A suitable rearrangement of \mathbf{x}_i will transform A into a block diagonal form where each block on the diagonal corresponds to one of the equivalence classes and hence is indecomposable. If these diagonal blocks are, say, B_1, B_2, \dots, B_r , then

$$N_{n+k} = \sum_{i=1}^r \mathbf{1}' B_i^k \mathbf{1},$$

where $\mathbf{1}'$ represents a vector of 1's of appropriate length. We may suppose that λ , the largest positive eigenvalue of A , is an eigenvalue of B_1 . Then

$$N_{n+k} = \mathbf{1}' B_1^k \mathbf{1} \cdot (1 + \sum_{i=2}^r \mathbf{1}' B_i^k \mathbf{1} / \mathbf{1}' B_1^k \mathbf{1}).$$

Since the second factor is bounded, $(\log N_{n+k})/(n+k) \rightarrow (\log \mathbf{1}' B_1^k \mathbf{1})/(n+k) \rightarrow \log \lambda$. This completes the proof.

We appeal to the construction of a minimal extension of P_n to X^{n+1} to construct an example of a nondeterministic measure P on (X^I, \mathcal{G}) for which $N_n < N_{n+1}$ and $N_{n+1} - N_n$ is bounded. As noted previously, it is possible to extend P_n to P_{n+1} with $N_{n+1} - N_n \leq N_n - N_{n-1}$. If this procedure is repeated, $N_{n+k} - N_{n+k-1} \leq N_n - N_{n-1}$ for all k . We shall construct an extension in which $N_{n+k} - N_{n+k-1} = N_n - N_{n-1}$.

Since our object is to show the existence of nondeterministic yet small extensions, we consider the case of states 0 and 1 alone. A similar construction can be given for q states.

In constructing P_{n+1} from P_n we start with $\mathbf{x} = (x_1, \dots, x_{n-1})$ in the support of P_{n-1} . The numbers $N_{\mathbf{x}}$ and $N_{\mathbf{x}}^*$ then denote respectively the number of $p_n(\mathbf{x}, 0)$, $p_n(\mathbf{x}, 1)$ and $p_n(0, \mathbf{x})$, $p_n(1, \mathbf{x})$ which are positive, where we use the notation $(\mathbf{x}, 0) = (x_1, \dots, x_{n-1}, 0)$, etc. Now if $N_{\mathbf{x}} + N_{\mathbf{x}}^* = 2$, exactly one of the probabilities $p_{n+1}(0, \mathbf{x}, 0)$, $p_{n+1}(0, \mathbf{x}, 1)$, $p_{n+1}(1, \mathbf{x}, 0)$, and $p_{n+1}(1, \mathbf{x}, 1)$ is nonzero. If $N_{\mathbf{x}} + N_{\mathbf{x}}^* = 3$, the probabilities $p_{n+1}(i, \mathbf{x}, j)$, for $i, j = 0, 1$, are again determined with exactly two of them being nonzero. In the case $N_{\mathbf{x}} + N_{\mathbf{x}}^* = 4$, we want to find conditions under which precisely three of these probabilities may be made nonzero. We have

$$p_n(0, \mathbf{x}) + p_n(1, \mathbf{x}) = p_n(\mathbf{x}, 0) + p_n(\mathbf{x}, 1)$$

and we must satisfy the three equations

$$p_{n+1}(0, \mathbf{x}, 0) + p_{n+1}(0, \mathbf{x}, 1) = p_n(0, \mathbf{x})$$

$$p_{n+1}(0, \mathbf{x}, 0) + p_{n+1}(1, \mathbf{x}, 0) = p_n(\mathbf{x}, 0)$$

$$p_{n+1}(0, \mathbf{x}, 1) + p_{n+1}(1, \mathbf{x}, 1) = p_n(\mathbf{x}, 1).$$

If $p_n(0, \mathbf{x}) = p_n(\mathbf{x}, 0) = p_n(\mathbf{x}, 1)$ then $p_{n+1}(0, \mathbf{x}, 0) = p_{n+1}(1, \mathbf{x}, 1)$ and $p_{n+1}(0, \mathbf{x}, 1) = p_{n+1}(1, \mathbf{x}, 0)$ so that it is not possible in this case to have three nonzero choices. In all other cases it may be checked that such a choice is possible.

There is consequently the possibility of an extension with $N_{n+k} - N_{n+k-1} = N_n - N_{n-1}$ for all $k > 0$ if whenever $N_{\mathbf{x}} + N_{\mathbf{x}}^* = 4$, $p(0, \mathbf{x})$, $p(1, \mathbf{x})$, $p(\mathbf{x}, 0)$ and $p(\mathbf{x}, 1)$ are not all equal.

Consider now a measure P_1 on X^1 defined by $p_1(0) = b$, $p_1(1) = 1 - b$, b irrational and $\frac{1}{2} < b < 1$.

THEOREM 5. P_1 has an extension P on (X^I, \mathcal{G}) for which $N_n = n + 1$.

PROOF. We proceed by induction to show that P_{k-1} has an extension P_k with $N_k = k + 1$, $k = 2, 3, \dots$. Let P_2 be defined by $p_2(0, 0) = 2b - 1$, $p_2(0, 1) = p_2(1, 0) = 1 - b$. P_2 is an extension of P_1 and $N_2 = 3$. Suppose then that P_{k-1} has an extension P_k with $N_k = k + 1$, $k \leq n$. Let \mathbf{x} and \mathbf{x}^* be the (unique) elements in the support of P_{n-1} for which

$$p_n(\mathbf{x}, 0)p_n(\mathbf{x}, 1) > 0 \text{ and } p_n(0, \mathbf{x}^*) p_n(1, \mathbf{x}^*) > 0.$$

If $\mathbf{x} \neq \mathbf{x}^*$, $N_{\mathbf{x}'} + N_{\mathbf{x}^*} < 4$ for all \mathbf{x}' in the support of P_{n-1} , thus there is an extension to P_{n+1} with $N_{n+1} = n + 2$. If $\mathbf{x} = \mathbf{x}^*$ there is such an extension unless $p_n(0, \mathbf{x}) = p_n(1, \mathbf{x}) = p_n(\mathbf{x}, 0) = p_n(\mathbf{x}, 1)$, so suppose this to be the case. We know from Theorem 2 that P_n has a decomposition as $P_n = \sum_{\alpha} \lambda_{\alpha} P_n(S_{\alpha})$ for certain minimal subsets S_{α} . Note that each minimal subset S_{α} must contain one (but not both) of $(\mathbf{x}, 0)$, $(\mathbf{x}, 1)$. For if S_{β} , say, contains neither $(\mathbf{x}, 0)$ nor $(\mathbf{x}, 1)$ both of them will appear in $\bigcup_{\alpha \neq \beta} S_{\alpha}$. But the residual measure on this union is supported by at most n elements, hence must be k -deterministic, $k \leq n - 1$. This is a contradiction since such a measure cannot contain both $(\mathbf{x}, 0)$ and $(\mathbf{x}, 1)$ in its support. The same contradiction is obtained if one of $(\mathbf{x}, 0)$, $(\mathbf{x}, 1)$ is in both S_{β} and $\bigcup_{\alpha \neq \beta} S_{\alpha}$. Therefore the decomposition of P_n is obtained from two minimal subsets, say S_1, S_2 , where $(\mathbf{x}, 0) \in S_1$ and $(\mathbf{x}, 1) \in S_2$. Furthermore, if $p = p_n(\mathbf{x}, 0) = p_n(\mathbf{x}, 1)$, then $p_n(\mathbf{y}) \geq p$ for every \mathbf{y} in the support of P_n . Suppose S_1 contains $k + 1$ elements. Then

$$P_n = p(k + 1)P_n(S_1) + (1 - p(k + 1))P_n(S_2).$$

In particular, $p = p_n(\mathbf{x}, 1) = [1 - p(k + 1)] \cdot 1/(n - k)$ since $(\mathbf{x}, 1) \notin S_1$. Therefore $p = 1/(n + 1)$. Since P_n is supported by $n + 1$ elements, this gives $p_n(\mathbf{y}) = 1/(n + 1)$ for every \mathbf{y} in the support of P_n . Adding the probabilities of all n -vectors with first coordinate zero must give b , so $b = r/(n + 1)$ for some integer r , a contradiction.

REFERENCES

- [1] DE BRUIJN, N. G. (1946). A combinatorial problem. *Indag. Math.* **8** 461-467.
- [2] GOOD, I. J. (1946). Normal recurring decimals. *J. London Math. Soc.* **21** 169-172.
- [3] HADLEY, G. (1962). *Linear Programming*. Addison-Wesley, Reading, Mass.
- [4] SHANNON, CLAUDE E. and WEAVER, WARREN (1949). *The Mathematical Theory of Communication*. Univ. of Illinois Press.