ON THE THEORY OF RANDOM SEARCH

A. RÉNYI

Introduction. The problems of search dealt with in this paper can be described by the following simple model. Let S_n be a *finite* set having $n \geq 2$ distinguishable elements—called points—and suppose that we want to find an unknown point x of the set S_n; the set S_n itself is supposed to be known to us. Let us suppose further that it is not possible to observe x directly, however we may choose some functions $f_1, f_2, \cdots, f_N$ from a given set F of functions defined on S_n, and observe the values $f_1(x), f_2(x), \cdots, f_N(x)$ taken on by these functions at the unknown point x. Of course if F would contain a function f which takes on different values at different points, a single observation of this function would be sufficient. We suppose however that all functions f belonging to the class F are such that the number of different values taken on by f is much smaller than n. (We shall be especially interested in the case when each $f \in F$ takes on only the two values 0 and 1 and n is a large number.) In such a case of course it is necessary to observe the value of a large number of functions f at the point x. Each such observation gives us only partial information on x (namely it specifies a subset A of S_n to which x must belong), but after making a fairly large number of such observations the information obtained accumulates and enables us to determine x. We want to find x by a not too large number of observations. We may e.g. suppose that each observation is connected with a certain cost (or that it requires a definite amount of time) and we want to keep the cost (or duration) of the whole procedure of search relatively low. We shall call a method for the successive choice of the functions $f_1, \cdots, f_N$, which leads in the end to the determination of the unknown x, a *strategy of search*. Obviously one usually tries to choose a strategy with N (the number of functions to be observed) as small as possible. Of two search procedures the one which has a smaller (average) duration is the better one, however there may be other requirements. For instance a simple strategy which can e.g. be easily programmed on a computer is usually preferable to a complicated strategy. If A and B are two strategies such that A requires (in the average) the observation of a somewhat smaller number of functions than B (i.e. A is "better" than B) but the effective carrying out of A

requires much more work and time, then nevertheless B may be preferable to A from a practical point of view.

A strategy will be called a *pure strategy* if it uniquely specifies the choice of the functions $f_1, f_2, \cdots, f_N$ and a *mixed strategy* if the choice of these functions depends to some extent on chance. A pure strategy will be called *predetermined* if the number N and the choice of each of the functions $f_1, \cdots, f_N$ is determined in advance, before beginning the observations; it will be called *sequential* if only the choice of f_1 is decided in advance and the choice of f_k ($k \geq 2$) is made only after observing $f_1(x)$, $f_2(x)$, $\cdots$, $f_{k-1}(x)$ and may depend on these observed values; in the case of a sequential (pure) strategy the number of observations N depends usually also on the value of x.

Problems of search occur in practically every field of human activity. Typical examples are: medical diagnosis, chemical analysis, search for a failure in a complicated mechanism, search for a lost or hidden object, search for a mistake in a long series of computations or in a program for a computer, search for some bibliographical data, search for the root of an equation, the maximum of a function, the parameter of a probability distribution, etc.

The theory of search should be considered—according to the author's point of view—as a chapter of information theory; this chapter of information theory is however not very far developed. Many interesting particular problems of random search have been investigated, but a systematic study of such problems has not yet been given.

In §§1 and 2 of the present paper we deal with certain basic notions which are of importance in search theory (separating systems of functions, different notions of homogeneity of such systems) and with certain combinatorial questions connected with these notions. We deal with these questions in somewhat greater detail than needed in what follows, because these combinatorial questions are interesting in themselves too, and in view of other applications, to be given elsewhere. In §3 we prove a few general theorems concerning the duration of search of random search procedures and shall show that in general these random search methods are almost as good as the best pure strategy, being at the same time usually much simpler. In §4 we give a few examples.

1. **Separating systems of functions.** Clearly in order that there should exist at least one strategy which leads to finding the unknown element x of S_n whatever it may be, it is necessary that the class F of available functions f should be rich enough. Let us introduce the following definition:

(I) *Let us say that a system F of functions defined on the set S_n is separating the elements of S_n (or for the sake of brevity: is a separating system) if it has the following property:*

To any pair of different elements x_1, x_2 ($x_2 \neq x_1$) of S_n there exists in F a function f such that $f(x_2) \neq f(x_1)$.

Clearly if F is not a separating system, i.e. if there are elements x_1 and $x_2 \neq x_1$ in S_n such that for each $f \in F$, $f(x_2) = f(x_1)$, one never can decide whether the unknown element x of S_n which we want to find, is x_1 or x_2 by observing the values of functions $f \in F$ at the point x. On the other hand if F is a family of functions on S_n separating the elements of S_n, then by observing $f(x)$ for all $f \in F$ x will be uniquely determined; thus there exists at least one successful strategy, namely the trivial one consisting in observing the value of $f(x)$ one by one for all functions $f \in F$.

Let F be a system of functions defined on the set S_n which separates the elements of S_n. We shall call F a *minimal separating system* of functions, if no proper subset of F does separate the elements of S_n.

Another way to characterize separating systems of functions is the following. For each $f \in F$ let us call the set $A \subset S_n$ a level-set of f if A contains all elements x of S_n for which $f(x) = a$, where a is any fixed value. Clearly the system F separates the elements of the set S_n if and only if the least algebra of sets containing all level-sets of all functions $f \in F$ is identical with the algebra of all subsets of S_n.

A third equivalent way of characterizing separating systems of functions is the following. To any set F of functions defined on the set S_n there corresponds a matrix M defined as follows: If the elements of the set S_n are $a_1, a_2, \cdots, a_n$ and F consists of the functions $f_1, f_2, \cdots, f_m$, the matrix M contains m rows and n columns and the kth entry of the jth row of M is $f_j(a_k)$. Clearly F is separating the elements of S_n if and only if the column-vectors of the matrix M are all different.

We shall prove now the following simple

LEMMA 1. *Let F be a minimal separating system of functions separating the elements of the finite set S_n having n elements. If m denotes the number of elements of F we have $m \leq n - 1$.*

PROOF. The statement of Lemma 1 can be expressed also in the following way: if F is a system of functions separating the elements of the set S_n, and if the number of functions belonging to F is $\geq n$, one can select from F $n-1$ functions such that these selected functions separate the elements of S_n. This statement is evidently true for $n = 2$, because clearly there must be in F a function which is not

constant on S_2, and a single such function clearly separates the two elements of S_2.

Let us suppose that the statement of Lemma 1 holds for some $n \geq 2$; we shall show that in this case it holds also for $n+1$. Let us denote the $n+1$ elements of the set S_{n+1} by $a_1, a_2, \cdots, a_{n+1}$. If the system F separates the elements of the set S_{n+1}, it separates a fortiori the n elements $a_1, \cdots, a_n$; thus by supposition we can select from F $n-1$ functions $f_1, f_2, \cdots, f_{n-1}$ which separate the elements $a_1, a_2, \cdots, a_n$. Now let us consider the vector $(f_1(a_{n+1}), f_2(a_{n+1}), \cdots, f_{n-1}(a_{n+1}))$: if this vector is different from each of the n vectors $(f_1(a_j), f_2(a_j), \cdots, f_{n-1}(a_j))$ $(1 \leq j \leq n)$, then the $n-1$ functions $f_1(x), \cdots, f_{n-1}(x)$ separate all elements of the set S_{n+1}; if not, the vector $(f_1(a_{n+1}), \cdots, f_{n-1}(a_{n+1}))$ is identical with one (and only one) of the n different vectors $(f_1(a_j), \cdots, f_{n-1}(a_j))$; suppose for instance that $f_i(a_{n+1}) = f_i(a_{j_1})$ for $1 \leq i \leq n-1$. Then there is in F at least one function f such that $f(a_{n+1}) \neq f(a_{j_1})$, because otherwise F would not separate a_{n+1} and a_{j_1}. But then the system of functions $f_1, \cdots, f_{n-1}, f$ separates the elements of the set S_{n+1}. Thus Lemma 1 is proved by induction.

The following example shows that the statement of Lemma 1 is best possible. Let S_n consist of the elements $0, 1, \cdots, n-1$; let F consist of the $n-1$ functions $f_i(x)$ $(0 \leq i \leq n-2)$ defined as follows: put

$$f_i(x) = \begin{cases} 1 & \text{if } x = i, \\ 0 & \text{otherwise} \end{cases}$$

for $0 \leq i \leq n-2$. Then F is a minimal separating system for the set S_n.

Let us consider some examples of separating systems. In all the following examples, if not otherwise specified, S_n is the set consisting of the elements $0, 1, \cdots, n-1$.

EXAMPLE 1. F is the set of all 2^n functions on S_n taking on only the values 0 and 1.

EXAMPLE 2. F_k is the set of all $\binom{n}{k}$ functions on S_n which take on the value 1 at k points and the value 0 at the remaining $n-k$ points $(1 \leq k \leq n-1)$.

The set F is not minimal for $2 \leq k \leq n-2$. In his lectures at the University of Budapest, the author has proposed the question: to determine the least number $N(n, k)$ for which there exists a separating system F of functions on S_n such that each $f \in F$ takes on the value 1 at k points and the value 0 at the remaining $n-k$ points. The answer to this question is easy (and has been obtained independently by B. Bollobás, J. Galambos, Gy. Katona, T. Nemetz and D. Szász) if $n \geq k(k+1)/2+1$; in this case one has $N(n, k) = \{2(n-1)/(k+1)\}$ where $\{x\}$ denotes the least integer $\geq x$. On the other hand if n is

even and $k = n/2$ one has clearly $N(n, n/2) = \{\log n\}$. To find an explicit formula for $N(n, k)$ in general seems to be difficult; however Mr. Gy. Katona has obtained rather close lower and upper estimates for $N(n, k)$, which he will publish in a forthcoming paper.

EXAMPLE 3. Suppose $n = 2^r$. Let us represent each integer x ($0 \leq x \leq n-1$) in the binary number system:

$$x = \sum_{k=0}^{r-1} \epsilon_k(x) \cdot 2^k$$

where $\epsilon_k(x) = 0$ or $= 1$, and consider the system $F_{[r]}$ of functions $\epsilon_0(x), \cdots, \epsilon_{r-1}(x)$. Clearly the values $\epsilon_0(x), \cdots, \epsilon_{r-1}(x)$ determine x uniquely; thus $F_{[r]}$ is a (minimal) separating system for the set S_2^r.

EXAMPLE 4. Let the matrix $M_{r,s}$ be defined as follows: $M_{r,s}$ contains r rows and $\binom{r}{s}$ columns, where $0 \leq s \leq r$. Let each column of $M_{r,s}$ contain s ones and $r-s$ zeros, and let all columns of $M_{r,s}$ be different; by other words the columns of $M_{r,s}$ are all possible sequences of r terms consisting of s ones and $r-s$ zeros. Let $f_j(x)$ be equal to the xth element of the jth row of $M_{r,s}$ ($1 \leq j \leq r$; $1 \leq x \leq \binom{r}{s}$). Then the functions $f_1, \cdots, f_r$ separate the elements of the set $\{1, 2, \cdots, \binom{r}{s}\}$. Evidently each row of $M_{r,s}$ contains $\binom{r-1}{s-1}$ ones and $\binom{r-1}{s}$ zeros. Note that if we form from the matrices $M_{r,0}, M_{r,1}, \cdots, M_{r,r}$ a single matrix having r rows and 2^r columns by putting the matrices $M_{r,s}$ side by side, we get the matrix of the minimal separating system of Example 3.

Note that the separating system of functions having the matrix $M_{r,s}$ is not minimal if $r \geq 2$, $s = 0, 1, \cdots, r$, however if $1 \leq s \leq r-1$ and we omit any one of the functions from the separating system having the matrix $M_{r,s}$ we obtain a minimal separating system.

Let us introduce in the set S_n a probability measure, by supposing that each element of S_n has the same probability $1/n$. In this case each function $f(x)$ defined on S_n can be considered as a random variable. If the set of values of $f(x)$ taken on S_n is the set $\{y_1, \cdots, y_r\}$ and $f(x)$ takes on the value y_j at K_j different points of $S_n (\sum_{j=1}^r K_j = n)$ then the probability distribution of $f(x)$ is given by

(1.1) $$P(f(x) = y_j) = \frac{K_j}{n} = P_j$$

and the entropy of f is

(1.2) $$H(f) = \sum_{j=1}^r P_j \log \frac{1}{P_j}.$$

(Here and in what follows log denotes logarithm with base 2.)

Now we prove the following simple

LEMMA 2. *If F is a system of functions separating the elements of the set S_n then*

$$(1.3) \qquad\qquad \sum_{f \in F} H(f) \geqq \log n.$$

PROOF OF LEMMA 2. Let $f_1, f_2, \cdots, f_m$, be all the functions belonging to the set F, then the vector $\psi(x) = (f_1(x), \cdots, f_m(x))$ takes on different values for different values of $x \in S_n$. Thus $H(\psi) = \log n$. As it is well known that the entropy of the joint distribution of a sequence of random variables can not exceed the sum of the entropies of these random variables, we obtain (1.3) which proves Lemma 2. Note that there is equality in (1.3) if and only if the functions $f_1, \cdots, f_m$, considered as random variables, are independent.

If for a separating system F there is equality in (1.3) we call F an *optimal separating system*. Clearly an optimal separating system is always minimal, but not conversely.

The following remark shows the difference between these two notions: if $F = \{f_1, \cdots, f_N\}$ is a minimal separating system, then omitting one function—say f_K—from F the remaining system $\{f_1, \cdots, f_{K-1}, f_{K+1}, \cdots, f_N\}$ is no more separating, i.e. there exist elements x_1 and $x_2 \neq x_1$ of S_n such that $f_j(x_1) = f_j(x_2)$ for $j \neq K$, $1 \leqq j \leqq N$. By other words, there exists an element x of S_n such that to find x we have to observe all functions f_j belonging to F at the point x, i.e. the process of search is *not always* finished before making N observations. However if F is an optimal system, the process of search is *never* finished before making N observations; as a matter of fact if the random variables $f_1, \cdots, f_N$ are independent (and none of them is constant) then by fixing the values of $N-1$ among these functions, the remaining function takes on all its values under this condition too. It should be added that this property is necessary but not sufficient for F being an optimal system. For instance if the functions of the system F are independent random variables with respect to some other probability measure on S_n different from the uniform measure, then F has the mentioned property, without being an optimal system.

The minimal system of Example 3 is optimal, but the minimal system of Example 4 (obtained by omitting any one of the functions of the system corresponding to the matrix $M_{r,s}$) is not optimal if $r \geqq 3$, $1 \leqq s \leqq r-1$.

An optimal separating system F can clearly be characterized by saying that the partial informations obtained from observing different functions f belonging to F never overlap. Thus an optimal separating system corresponds to a most economic strategy. The ratio $\log n / \sum_{f \in F} H(f)$ may be interpreted as the *efficiency* of the separating systems F.

2. Different notions of homogeneity of a separating system of functions. In our investigations we do not suppose the existence of any algebraical, geometrical or topological structure in the set S_n, and we suppose further that we have no prior knowledge on the unknown element x of S_n, i.e. that before we start the search all possibilities $x = a_K$ ($K = 1, 2, \cdots, n$) have the same prior probability $1/n$. Accordingly it is natural to restrict ourselves to the case in which the separating system F of available functions f is in some sense symmetrical with respect to the elements a_K of S_n. We shall define different sorts of symmetry of systems of functions. First we adopt the following definition:

(II) *A system F of functions f defined on the set $S_n = \{a_1, a_2, \cdots, a_n\}$ is called completely homogeneous, if denoting by π any permutation of the sequence $1, 2, \cdots, n$ for any $f \in F$ the function g defined by*

$$g(a_j) = f(a_{\pi(j)})$$

belongs to F too.

Let us consider some examples. The separating systems of Examples 1 and 2 are clearly completely homogeneous; but the systems of Examples 3 and 4 are not completely homogeneous. A system F of functions which is completely homogeneous and contains at least one function which is not constant is obviously a separating system.

Let M denote the matrix corresponding to a separating system F. Clearly F is completely homogeneous if and only if by permuting in any possible way the columns of M we obtain a matrix M' which can be obtained from M also by a suitable permutation of the rows of M. Thus if F is a completely homogeneous system of functions separating the elements of the set S_n and the number of elements of F is m, then necessarily $m \geq n$, because otherwise there would be more permutations of the columns than permutations of the rows of the corresponding matrix M. As by Lemma 1 for a minimal separating system $m \leq n - 1$, it follows that a minimal separating system can never be completely homogeneous.

Let us introduce now the following definition:

(III) *A system of functions F defined on the set $S_n = \{a_1, a_2, \cdots, a_n\}$*

is weakly homogeneous of order K $(2 \leq K \leq n)$ *if it has the following property*: Let us choose K *different elements* $a_{j_1}, a_{j_2}, \cdots, a_{j_K}$ *of* S_n *and let* R_K *denote the number of functions* $f \in F$ *such that*

$$f(a_{j_1}) = f(a_{j_2}) = \cdots = f(a_{j_K}).$$

The R_K *does not depend on the choice of the elements* $a_{j_1}, a_{j_2}, \cdots, a_{j_K}$. We shall call the number R_K *the parameter of* F.

If a system F of functions is weakly homogeneous of order 2 and $R_2 < R_1$ where R_1 is the number of elements of the set F (that is always true except the trivial case when F consists of constant functions only) then of course F is a separating system.

We shall need also another property of homogeneity which is stronger than homogeneity of order K; we call this strong homogeneity of order K and define it as follows:

(IV) *Let* F *be a set of functions* f *defined on the set* $S_n = \{a_1, a_2, \cdots, a_n\}$. *Let us suppose that the values of every* $f \in F$ *belong to a set* $Y = \{y_1, \cdots, y_r\}$. *We call the set* F *strongly homogeneous of order* K *if choosing any* K *different points* $a_{j_1}, a_{j_2}, \cdots, a_{j_K}$ *of* S_n *and any sequence of elements* $y_{l_1}, \cdots, y_{l_K}$ *of* Y *(the* y_{l_i} *($i = 1, 2, \cdots, K$) are not necessarily different) the number* R_K *($y_{l_i}, \cdots, y_{l_K}$) of those functions* $f \in F$ *for which*

$$f(a_{j_i}) = y_{l_i} \qquad (i = 1, 2, \cdots, K)$$

does not depend on the choice of the points $a_{j_1}, \cdots, a_{j_K}$ *but it may depend on the values* $y_{l_1}, \cdots, y_{l_K}$.

Evidently a strongly homogeneous set of order K is homogeneous of order K because we have

$$R_K = \sum_{y \in Y} R_K(y, y, \cdots, y).$$

It also is easy to see that if F is strongly homogeneous of order K it is strongly homogeneous of every order $K' < K$. For the weak homogeneity a similar statement does not hold. As a matter of fact every set F of functions f defined on S_n is homogeneous of order n but not necessarily of order $K < n$. It can easily be shown that a completely homogeneous set F is strongly homogeneous of every order K. If each function f in F takes on only the values 0 and 1 and it takes on the value 1 at exactly k points, we can characterize F also as follows: to each function $f \in F$ there corresponds the K-tuple of those numbers j ($1 \leq j \leq n$) for which $f(a_j) = 1$. Thus F can be represented by a sequence of k-tuples of the integers $1, 2, \cdots, n$. Now if F is strongly

homogeneous of order $l < k$ this implies that any l-tuple of the integers $1, 2, \cdots, n$ is contained in the same number r of these k-tuples. Thus these k-tuples form a *balanced incomplete block design*.

EXAMPLE 5. Let F be the class of functions f_K ($K = 1, 2, \cdots, N$) defined on the set $S_n = \{1, 2, \cdots, n\}$ and such that

$$f_K(j) = g_j(K) \quad 1 \leq j \leq n \quad (1 \leq k \leq N)$$

where the $g_j(k)$ ($1 \leq j \leq n$) are functions defined on the space $\Omega = \{1, 2, \cdots, N\}$ such that if we introduce a uniform measure on Ω the g_j's are equivalent (symmetrically dependent) random variables; then F is strongly homogeneous of every order.

EXAMPLE 6. Let n be arbitrary, $r < n$; let us form all possible $\binom{n}{r}$ sequences consisting of r ones and $n - r$ zeros, and let the matrix M of the family F consist of these rows. Then clearly F is strongly homogeneous of every order, because if $\epsilon_1, \cdots, \epsilon_K$ ($K < n$) is any sequence of $l \leq K$ ones and $K - l$ zeros ($K \leq r$, $K - l \leq n - r$) the number of rows of M which have $\epsilon_1, \cdots, \epsilon_K$ in the j_1st, j_2nd, $\cdots, j_K$th column is $\binom{n-k}{r-l}$.

Now let us change the first row of M so that we replace every 1 by 0 and every 0 by 1. Clearly the matrix M^* so obtained is not strongly homogeneous of any order, but it is weakly homogeneous of every order. Evidently if the set of functions F^* has M^* as its matrix, for purposes of search F^* and F are equivalent, because F differs from F^* only in that f_1 has been replaced by $1 - f_1$. This example shows that if F is a weakly homogeneous system of functions there sometimes exists an equivalent strongly homogeneous system; if we call two systems of functions F and F^* equivalent if their elements being f_K and f_K^* ($K = 1, 2, \cdots, N$), for every K and every pair of points x, y of S_n $f_K(x) = f_K(y)$ if and only if $f_K^*(x) = f_K^*(y)$. However not every weakly homogeneous system is equivalent to a strongly homogeneous one, as is shown by the next example.

EXAMPLE 7. The system of functions having the matrix

$$\begin{matrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{matrix}$$

is weakly homogeneous of orders 2 and 3 but it is not equivalent to a strongly homogeneous system. The same holds if $n = 2K$ is arbitrary and the rows of the matrix M of F are all possible sequences of K ones and K zeros starting with 1.

EXAMPLE 8. The system of functions having the matrix

$$
\begin{array}{cccc}
1 & 1 & 1 & 0 \\
1 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 \\
0 & 0 & 1 & 1
\end{array}
$$

is weakly homogeneous of order 3 but not of order 2.

EXAMPLE 9. The set F having the matrix

$$
\begin{array}{ccccccc}
1 & 1 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 1 & 1 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 1 & 1 & 0
\end{array}
$$

is strongly homogeneous of order 2, weakly homogeneous of order 3, but not weakly homogeneous of order 4. In this example $R_1 = 7$, $R_2 = 3$, $R_3 = 1$.

We introduce still one further definition:

(V) *A system of functions F is called homogeneous of order K if for any partition $K = K_1 + K_2 + \cdots + K_s$ of K the number $R_K[K_1, K_2, \cdots, K_s]$ of those $f \in F$ for which*

$$
\begin{aligned}
f(x_{11}) &= f(x_{12}) = \cdots = f(x_{1K_1}) \\
f(x_{21}) &= f(x_{22}) = \cdots = f(x_{2K_2}) \\
&\quad \vdots \\
f(x_{s1}) &= f(x_{s2}) = \cdots = f(x_{sK_s})
\end{aligned}
$$

does not depend on the choice of the points x_{ij} from S_n provided they are all different.

Clearly if F is homogeneous of order K it is also homogeneous of any order $l < K$ further it is weakly homogeneous of any order $l \leq K$. Conversely if F is strongly homogeneous of order K it is also homogeneous of order K. Thus the notion of homogeneity is intermediate between strong and weak homogeneity.

Now we prove the following simple but rather surprising

LEMMA 3a. *A set F of functions f taking on only the values 0 and 1*

which is weakly homogeneous of order 2 is also weakly homogeneous of order 3.

PROOF.[1] Let A_i $(1 \leq i \leq n)$, denote the set of those $f \in F$ for which $f(a_i) = 1$. Let us introduce a uniform measure on the set F. Then the A_i can be considered as random events, and the condition that F is weakly homogeneous of order 2 means that

$$P(A_i A_j) + P(A_i \overline{A}_j)$$

has the same value for any choice of the different indices i and j. (Here the product of two events denotes the joint occurrence of these events and $\overline{A}$ denotes the event that A does not occur.) Let us denote this value by C_2, i.e. we put

$$C_2 = P(A_i A_j) + P(A_i \overline{A}_j) \qquad (1 \leq i < j \leq n).$$

To prove Lemma 2 we have to show that the value of $P(A_i A_j A_k)$ $+ P(\overline{A}_i \overline{A}_j \overline{A}_k) = C_3$ does not depend on the choice of the (different) indices i, j, k. As however

$$P(A_i A_j A_k) = 1 - P(A_i) - P(A_j) - P(A_k) + P(A_j) + P(A_k)$$
$$+ P(A_k) - P(A_i A_j A_k)$$

and similarly

$$P(A_i A_j A_K) = 1 - P(A_i) - P(A_j) - P(A_k) + P(A_i A_j) + P(A_i A_k)$$
$$+ P(A_j A_k) - P(A_i A_j A_k)$$

it follows

$$(2.1) \qquad 2C_3 = 2(P(A_i A_j A_k) + P(A_i A_j \overline{A}_k)) = 3C_2 - 1.$$

This proves Lemma 3a.

Note that if R_1 denotes the number of elements of F, R_2 the number of functions $f \in F$ for which $f(a_i) = f(a_j)$ for any pair (i, j) with $1 \leq i < j \leq n$ and R_3 the number of functions $f \in F$ for which $f(a_i)$ $= f(a_j) = f(a_k)$ for any triple (i, j, k) $(1 \leq i < j < k \leq n)$, then according to what has been proved above we have

$$(2.2) \qquad\qquad\qquad 3R_2 - R_1 = 2R_3.$$

It follows from (2.2) that

$$(2.3) \qquad\qquad\qquad \frac{1}{3} \leq \frac{R_2}{R_1} \leq 1.$$

[1] My first proof of Lemma 3a was slightly more complicated. Later, essentially the same simpler proof was found independently by Mrs. Vera T. Sós and Mr. T. Nemetz; the proof given here is a third version of their proof.

It is easy to prove that a weakly homogeneous system of order k is not necessarily of order $k+1$ if $k \geq 3$, e.g. for $k=3$ this can be seen from Example 9; thus the statement of Lemma 3a cannot be generalized by replacing 2 and 3 by k and $k+1$. However the following generalization of Lemma 3a is valid.

LEMMA 3b. *If the system F of functions with values 0 and 1 is weakly homogeneous of all orders $l \leq 2k$ then it is also weakly homogeneous of order $2k+1$ $(k=1, 2, \cdots)$.*

PROOF. Using the notations of the proof of Lemma 3a according to our suppositions

$$P(A_{i_1}A_{i_2} \cdots A_{i_l}) + P(A_{i_1}A_{i_2} \cdots A_{i_l}) = C_l$$

does not depend on the choice of the different indices $i_1, i_2, \cdots, i_l$ for $l \leq 2k$. Now similarly as we have proved (2.1) one can show that

$$2(P(A_{i_1}A_{i_2} \cdots A_{i_{2k+1}}) + P(A_{i_1}A_{i_2} \cdots A_{i_{2k+1}}))$$

(2.4)
$$= -(2k-1) + \sum_{l=2}^{2k} (-1)^l \binom{2k+1}{l} C_l.$$

Thus Lemma 2b follows. Putting $c_l = R_l < R_1$ we obtain as a generalization of (2.2)

$$(2.5) \qquad 2R_{2k+1} = -(2k-1)R_1 + \sum_{l=2}^{2k} (-1)^l \binom{2k+1}{l} R_l.$$

EXAMPLE 9. If the events A_i $(1 \leq i \leq n)$ are symmetrically dependent, then the system F is clearly weakly homogeneous of every order. The same holds a fortiori if the events A_i are independent and have the same probability p. In this case

$$(2.6) \qquad R_l = R_1(p^l + (1-p)^l) \qquad (l = 1, 2, \cdots)$$

and (2.5) is trivially satisfied.

Note that in case (2.6) holds we have $R_2 \geq R_1/2$.

It can be shown that weak homogeneity of every order l with $2 \leq l \leq 2k-1$ of a system of functions taking on the values 0 and 1 does not necessarily imply its weak homogeneity of order $2k$. This is shown e.g. by the following example, due to T. Nemetz (oral communication).

EXAMPLE 10. Let the function f defined on $S_{2k+1} = \{0, 1, 2, \cdots, 2k\}$ belong to the family F if and only if the number of integers $1 \leq i \leq 2k$ for which $f(i) = 0$ is even. Then clearly F is homogeneous of every order $l \leq 2k-1$ with $R_1 = 2^{2k}$, $R_l = 2^{2k-l+1}$ if $l \leq 2k-1$. However F is

not homogeneous of order $2k$ because the number of functions of $f \in F$ for which $f(1) = f(2) = \cdots = f(2k)$ is 4 while the number of functions $f \in F$ for which $f(0) = f(1) = \cdots = f(2k-1)$ is 2.

According to (2.3) if F is weakly homogeneous of order 2 then $R_2/R_1 \geq 1/3$. This inequality is best possible without restriction on n as is seen from Example 7. However a better lower bound can be given, which depends on n. We prove

LEMMA 4. *Let F be a system of functions with values 0 and 1 on the set S_n which is weakly homogeneous of order 2. Let R_1 denote the number of elements of F and R_2 the number of functions $f \in F$ for which $f(a_i) = f(a_j)$ $(i \neq j)$. Then we have*

$$(2.7) \qquad \frac{R_2}{R_1} \geq \frac{(n-2)}{2(n-1)}.$$

PROOF. Let $f_1, f_2, \cdots, f_{R_1}$ denote the elements of F and $a_1, a_2, \cdots, a_n$ the elements of S_n. Let us put

$$X = \sum_{j=1}^{R_1} \sum_{k=1}^{n} f(a_k).$$

By Cauchy's inequality we have

$$(2.8) \qquad \begin{aligned} X^2 &= \left(\sum_{j=1}^{R_1} \sum_{k=1}^{n} f_j(a_k) \right)^2 \leq R_1 \sum_{j=1}^{R_1} \left(\sum_{k=1}^{n} f_j(a_k) \right)^2 \\ &= R_1 X + R_1 \sum_{k=1}^{n} \sum_{l=1}^{n} \sum_{j=1}^{R_1} f_j(a_k) f_j(a_l), \qquad k \neq l. \end{aligned}$$

Now by supposition

$$(2.9) \qquad \sum_{j=1}^{R_1} (1 - f_j(a_k) - f_j(a_l) + 2f_j(a_k)f_j(a_l)) = R_2$$

for all pairs k, l with $k \neq l$. Thus it follows that

$$(2.10) \qquad n(n-1)R_2 = R_1 n(n-1) - 2(n-1)X + 2 \sum_{K=1}^{n} \sum_{l=1}^{n} f_j(a_k)f_j(a_l),$$

$$k \neq l.$$

Combining (2.8) and (2.10) we get

$$(2.11) \qquad X^2 \leq R_1 X + \tfrac{1}{2} R_1(n(n-1)(R_2 - R_1) + 2(n-1)X).$$

Thus it follows

$$(2.12) \quad \frac{R_2}{R_1} \geq 1 - \frac{2(n(X/R_1) - (X/R_1)^2)}{n(n-1)} \geq 1 - \frac{n}{2(n-1)} = \frac{n-2}{2(n-1)}$$

which proves Lemma 4.

3. **On the duration of the random search.** Let F be a set of R_1 functions on the set $S_n = \{a_1, a_2, \cdots, a_n\}$ which separates the elements of S_n. Let x be an unknown element of S_n and let us suppose that we search for x in the following way: We choose first a function f_1 from F at random so that each element of F has the same probability $1/R_1$ to be chosen. We observe the value of $f_1(x)$, and after this choose again a function f_2 from F so that the choice of f_2 is independent from the choice of f_1 and each element f of F (including f_1) has the same probability $1/R_1$ to be chosen as f_2. We observe $f_2(x)$ and choose an element f_3 of F independently from the choices of f_1 and f_2, so that each element f of F has the same probability $1/R_1$ to be taken for f_3. (Thus the possibility that f_3 is equal to f_1 or f_2 is not excluded.) We observe $f_3(x)$, etc. This process is repeated until $f_N(x)$ is selected. Let $P_1(n, N, F, x)$ denote the probability that the sequence $f_1(x), \cdots, f_N(x)$ determines x uniquely. Let $P_2(n, N, F)$ denote the probability that the sequence $f_1(x), \cdots, f_N(x)$ determines x whatever it may be. It has to be emphasized that $P_2(n, N, F)$ and $P_1(n, N, F, x)$ are usually different (even if $P_1(n, N, F, x)$ does not depend on x) and in general we can say only that $P_2(n, N, F) \leq \min_x P_1(n, N, F, x)$. This is shown for instance by the following:

EXAMPLE 11. Let S_3 be the set $\{1, 2, 3\}$ and let F consist of the three functions f_1, f_2 and f_3 where

$$f_1(1) = 0,\ f_1(2) = 1,\ f_1(3) = 1,$$
$$f_2(1) = 1,\ f_2(2) = 0,\ f_2(3) = 1,$$
$$f_3(1) = 1,\ f_3(2) = 1,\ f_3(3) = 0.$$

Evidently F separates the elements of S_3 because all columns of the corresponding matrix

$$\begin{matrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{matrix}$$

are different. Further F is weakly homogeneous of every order with $R_2 = 1$, $R_3 = 0$ and even strongly homogeneous of every order (with $R_2(1, 1) = R_2(0, 1) = R_2(1, 0) = 1, R_2(0, 0) = 0$ and $R_3(0, 1, 1) = R_3(1, 0, 1) = R_3(1, 1, 0) = 1$). Now let us compute $P_1(3, 1, F, x)$. If $x = i$, x is

uniquely determined by the first chosen function if and only if this function is $f_i(x)$ $(i=1, 2, 3)$ and otherwise not; thus for each fixed value of x, $P_1(3, 1, F, x) = 1/3$. On the other hand $P_2(3, 1, F) = 0$ because no function $f \in F$ determines x uniquely whatever x may be.

The difference between $P_1(n, N, F, x)$ and $P_2(n, N, F)$ can be made clear also as follows. Suppose the N functions which we have chosen by our random procedure are $f_1, \cdots, f_N$. Consider the matrix M having N rows and n columns, in which the kth element of the jth row is $f_j(a_k)$.

Then a fixed unknown element x of S_n is determined by the observations $f_1(x), \cdots, f_N(x)$ if and only if *the xth column* of this matrix M (containing the entries $f_j(x)$ $(j = 1, 2, \cdots, N)$) *is different from every other column* of M, while the observations $f_1(x), \cdots, f_N(x)$ determine x whatever it may be if and only *if any two columns* of M *are different*, and this second condition imposes a much more severe restriction on M than the first and thus the probability of M having the second property is usually definitely less than the probability of M having the first property.

Thus we have to solve two different problems: to evaluate $\min_x P_1(n, N, F, x)$ and $P_2(n, N, F)$. First we give lower estimates for these quantities.

In what follows we shall repeatedly use the following well-known inequalities of elementary probability theory. Let $A_1, \cdots, A_m$ be arbitrary events in a probability space; let the product of events denote the event consisting in the joint occurrence of the factors and let $\overline{A}$ denote the event contrary to the event A. Put $s_k = \sum P(A_{i_1} A_{i_2} \cdots A_{i_k})$ $(1 \leq k \leq m)$ where the summation is extended over all k-tuples $(i_1, i_2, \cdots, i_k)$ which can be selected from the set $\{1, 2, \cdots, m\}$, and put $s_0 = 1$. Then for $l = 0, 1, 2, \cdots$ the inequalities

(3.1)
$$\sum_{k=0}^{2l+1} (-1)^k s_k \leq P(A_1 \cdots A_m) \leq \sum_{k=0}^{2l} (-1)^k s_k$$

hold, and one has

(3.2)
$$P(A_1 \cdots A_m) = \sum_{k=0}^{m} (-1)^k s_k.$$

THEOREM 1. *If the system F of functions defined on the set $S_n = \{1, 2, \cdots, n\}$ is weakly homogeneous of order 2, we have for all $x \in S_n$*

(3.3)
$$P_1(n, N, F, x) \geq 1 - (n-1) \left(\frac{R_2}{R_1} \right)^N$$

and

$$(3.4) \qquad P_2(n, N, F) \geqq 1 - \binom{n}{2}\left(\frac{R_2}{R_1}\right)^N.$$

PROOF. Let B_{ij} $(i \neq j)$ denote the event that $f(i) = f(j)$ for all the N selected functions. Then we have

$$(3.5) \qquad P_1(n, N, F, x) = P\left(\prod_{1 \leq i \leq n;\, i \neq x} \overline{B}_{ix}\right)$$

and

$$(3.6) \qquad P_2(n, N, F) = P\left(\prod_{1 \leq i < j \leq n} \overline{B}_{ij}\right).$$

In view of $P(B_{ij}) = (R_2/R_1)^N$ we obtain from (3.1) immediately (3.3) and (3.4).

COROLLARY. *If F_n is for each n a separating system for the set S_n which is weakly homogeneous of order 2, having $R_1(n)$ elements and the parameter $R_2(n)$, and if*

$$(3.7) \qquad \lim_{n \to \infty} \left(N_n \log \frac{R_1(n)}{R_2(n)} - \log n\right) = +\infty$$

then

$$(3.8) \qquad \lim_{n \to \infty} \min_{x \in S_n} P_1(n, N_n, F_n, x) = 1.$$

If instead of (3.5) we suppose that

$$(3.9) \qquad \lim_{n \to \infty} \left(N_n \log \frac{R_1(n)}{R_2(n)} - 2 \log n\right) = +\infty$$

we have

$$(3.10) \qquad \lim_{n \to \infty} P_2(n, N_n, F_n) = 1.$$

Note that in order to be able to show that (3.10) holds, we had to choose N_n asymptotically twice as large as was needed to ensure the validity of (3.8).

To judge the "quality" of the random search procedures in question, one has to take into account that if only functions with values 0 and 1 are used, then $H(f) \leqq 1$ and thus by Lemma 2 the duration of the best systematic search can not be less than $\log n$.

Thus Theorem 1 asserts that if we use a random search by means

of a family F of R_1 functions which is homogeneous of order 2, with parameters R_2 and if n is large, the duration of the search will be with probability near to 1 only about $(\log(R_1/R_2))^{-1}$ times, respectively about $2(\log(R_1/R_2)^{-1}$ times longer than with the best systematic method, according to whether we want to have a procedure which leads to the solution in a single given case, or one which is universally applicable. This shows that random search processes are under mild restrictions really almost as good as the best systematic method. The factor $(\log(R_1/R_2))^{-1}$ is of course usually larger than 1, but in case R_2/R_1 is equal to $\frac{1}{2}$ (or near to $\frac{1}{2}$) it is equal (near to) 1, and thus in such cases the random search procedure (for a single search) is with probability near to 1 if n is large asymptotically as good as the best systematic one.

Similarly we can get also an upper estimate for $P_1(n, N, F, x)$.

THEOREM 2a. *If the system F of functions with values 0 and 1 defined on the set is weakly homogeneous of order 2 (and thus by Lemma 3a also of order 3) we have for all $x \in S_n$*

$$(3.11) \quad P_1(n, N, F, x) \leq 1 - (n-1)\left(\frac{R_2}{R_1}\right)^N + \binom{n-1}{2}\left(\frac{R_3}{R_1}\right)^N .$$

PROOF. (3.11) follows immediately from (3.5) and (3.1).

REMARK. By (2.2) we have $R_3/R_1 \leq \frac{1}{2}(3(R_2/R_1)-1)$. Thus it follows that

$$\frac{R_3 R_1}{R_2^2} \leq \frac{3(R_2/R_1) - 1}{2(R_2/R_1)^2} \leq \frac{9}{8}.$$

Further, that in case

$$\frac{R_2}{R_1} \leq \frac{1}{2}$$

we have

$$\frac{R_3 R_1}{R_2^2} \leq 1.$$

Thus we get the following:

COROLLARY. *If for each n F_n is a system of functions with values 0 and 1 on the set S_n which is weakly homogeneous of order 2 and if $R_2(n)/R_1(n) \leq 1/2$ further*

$$\lim_{n_j \to \infty} \left(N_{n_j} \log \frac{R_1(n_j)}{R_2(n_j)} - \log n_j\right) = c > 0$$

then if x_{n_j} is any element of S_{n_j}

$$1 - e^{-c} \leqq \liminf_{n_j \to \infty} P_1(n_j, N_{n_j}, F_{n_j}, x_{n_j}) \leqq \limsup_{n \to \infty} P_1(n_j, N_{n_j}, F_{n_j}, x_{n_j})$$

$$\leqq 1 - e^{-c} + \frac{e^{-2c}}{2}.$$

A similar upper estimate can be given for $P_2(n, N, F)$ if F is supposed to be homogeneous of order 4.

THEOREM 2b. *If the system F is homogeneous of order 4 on the set S_n, then we have*

$$P_2(n, N, F) \leqq 1 - \binom{n}{2}\left(\frac{R_2}{R_1}\right)^N + \frac{3}{2}\binom{n}{3}\left(\frac{R_3}{R_1}\right)^N$$

$$+ 3\binom{n}{4}\left(\frac{R_4(2, 2)}{R_1}\right)^N.$$

If we suppose that F is weakly homogeneous of every order, by using (3.2) we obtain an exact formula for $P_1(n, N, F, x)$, the value of which does not depend in this case on x.

THEOREM 3. *If the system F is weakly homogeneous of every order, then $P_1(n, N, F, x) = P_1(n, N, F)$ does not depend on x and we have*

$$P_1(n, N, F) = \sum_{K=0}^{n-1} (-1)^K \binom{n-1}{K}\left(\frac{R_{K+1}}{R_1}\right)^N.$$

As regards $P_2(n, N, F)$, to obtain a corresponding exact formula, we have to suppose that F is homogeneous of every order. The formula is however rather complicated, and therefore we do not write it out in detail here.

Let us add one final remark. A useful generalization of these results is obtained if we suppose that the values $f_1(x), \cdots, f_N(x)$ are observed not directly, but through a noisy channel. This problem was considered in a special case in [5].

4. Examples.

EXAMPLE 12. Let $F(n)$ be the set of all 2^n functions with values 0 and 1 on S_n. Clearly $F(n)$ is homogeneous of every order with $R_K = 2^{n-K+1}$ $(K = 1, 2, \cdots, n)$. Thus we get from Theorem 3

$$P_1(n, N, F(n)) = \sum_{K=0}^{n-1} (-1)^K \binom{n-1}{K}\frac{1}{2^{NK}} = \left(1 - \frac{1}{2^N}\right)^{n-1}.$$

This formula can of course be also directly verified (see [1] and [2]). It follows that if $N_{n_j} = \log_2 n_j + c + \epsilon_j$ where $\epsilon_j \to 0$ ($\log_2$ denotes logarithm with base 2) we have

$$\lim_{n_j \to \infty} P_1(n_j, N_{n_j}, F(n_j)) = \exp\left(-\frac{1}{2^c}\right).$$

In this case it is easy (see [1] and [2]) to determine $P_2(n, N, F(n))$ explicitly:

$$P_2(n, N, F(n)) = \prod_{l=1}^{n-1}\left(1 - \frac{l}{2^N}\right).$$

Thus if

$$N_{n_j} = 2\log_2 n_j + c + \epsilon_j \quad \text{where} \quad \epsilon_j \to 0$$

we have

$$\lim_{n_j \to \infty} P_2(n_j, N_{n_j}, F(n_j)) = \exp\left(-\frac{1}{2^{c+1}}\right).$$

Note that the best systematic strategy requires the observation of $\{\log_2 n\}$ functions $f \in F(n)$. Thus if we are satisfied to find the unknown with probability 0.99, the number of necessary observations is only by 7 greater for the random strategy as for the best systematic strategy, independently of the value of n.

If for instance S_n is the set of all persons now living, then $\{\log_2 n\}$ = 32. Thus with our random strategy we can find a person among all those living with ≤ 39 random dichotomies (instead of 32 dichotomies corresponding to the best systematic strategy).

EXAMPLE 13 (see [3]). Suppose that F consists of all functions taking on the values $0, 1, \cdots, s-1$ on S_n. Then we obtain from Theorem 1 that $P_1(n, N, F)$ is near to 1 if $N - \log n / \log s$ is large, while $P_2(n, N, F)$ is near to 1 if $N - 2\log n / \log s$ is large. Note that now $H(f) = \log_2 s$ for every $f \in F$, thus the best systematic strategy needs the observation of at least $\log n / \log s$ function $f \in F$.

EXAMPLE 14. Suppose that $F_n(P)$ consists of all functions which take on the value r for $l_r = p_r n$ different elements of S_n ($r = 1, 2, \cdots, S$) where $P = \{p_1, \cdots, p_S\}$ is a probability distribution. Then $F_n(P)$ is weakly homogeneous of order 2 with $R_2/R_1 \sim \sum_{r=1}^{S} p_r^2$. Thus by Theorem 1 $P_1(n, N, F_n(P))$ will be near to 1 if $N - \log_2 n / H_2(P)$ is large, where

$$H_2(P) = \log_2\left(\sum_{r=1}^{S} p_r^2\right)^{-1},$$

$H_2(P)$ is called (see [4]) *the entropy of order* 2 of the probability distribution P.

The best systematic strategy needs the observation of about $\log_2 n/H_1(P)$ function $f \in F$ where

$$H_1(P) = \sum_{r=1}^{s} p_r \, \log_2 \frac{1}{p_r} \, .$$

Thus the efficiency of the random search is in this case $\sim H_2(P)/H_1(P)$.

It follows from the result of our paper [3] that for this example if

$$\lim_{n_j \to \infty} (N_{n_j} H_2(P) - 2 \log_2 n_j) = c$$

we have

$$\lim_{n_j \to \infty} P_2(n_j, N_{n_j}, F_{n_j}) = \exp\left(-\frac{1}{2^c}\right).$$

REFERENCES

1. A. Rényi, *On random generating elements of a finite Boolean algebra*, Acta Sci. Math. (Szeged) 22 (1961), 75–81.

2. ———, *Statistical laws of accumulation of information*, Bull. Inst. Internat. Statist. 39 (1962), 311–316.

3. ———, *Az információ-akkumuláció statisztikus törvényszerüségéröl*, Magyar Tud. Akad. Mat. Fiz. Oszt. Közl. 12 (1962), 15–33.

4. ———, *On measures of entropy and information*, Proceedings of the 4th Berkeley Symposium, Vol. I, pp. 547–561, Univ. of California Press, Berkeley, Calif., 1961.

5. ———, *On a problem of information theory*, Publications of the Math. Inst. Hungarian Academy of Sciences 6 (1961), 505–516.

BUDAPEST, HUNGARY