On the Length of the Smallest Uniform Experiment Which Distinguishes the Terminal States of a Machine*

SEYMOUR GINSBURG

The National Cash Register Company, Hawthorne, California

The problem considered here is to obtain estimates of the length of the smallest experiment (on a machine) which is independent of the unknown initial state and which allows us, by observing the outputs, to distinguish the terminal state.

The estimates obtained depend, of course, on the assumptions placed on the machine. In general, the bounds derived are slightly less than n^2, where n is the number of distinguishable states in the machine. For no really general class of machines is a best bound known.

1. *Preliminary Results*

Many of the basic ideas (for example, the notion of machine) used here are the same as in [3]. A familiarity with [3], while desirable, is not necessary for an understanding of this paper.

By a (deterministic) machine[1] is meant a finite number of *states* $q_1, \cdots, q_n$, a finite number of *inputs* $I_1, \cdots, I_m$, and a finite number of *outputs* $U_1, \cdots, U_p$, with the machine satisfying the following conditions:

(1) The machine is always in, i.e., assumes, exactly one of the states q_i, at a time.

(2) If the machine is in state q_i, then upon application of any input I_j the machine assumes a (new) state q_k (possible q_i). q_k depends only on q_i and I_j.

(3) Associated with each state q_i is an output $U_{v(i)}$ with the property that, if the machine is in state q_i, then upon application of any input I, the output $U_{v(i)}$ occurs.

(4) A new state, also an output, can occur only upon application of an input.

A machine as described above shall be called an (n, m, p) machine. This terminology differs slightly with that found in [3].

Some authors [1, 2] replace condition (3) above with (3′).

(3′) The present output is determined by both the present input and the present state.

Most of the following results hold in both cases.

It is henceforth assumed that whenever a machine is given its inputs, outputs,

* Received August, 1957.

[1] A more rigorous definition of (deterministic) machine is as follows. A machine S is a set $\{I_1, \ldots, I_m, q_1, q_2, \quad q_n, U_1, \ldots, U_p\}$, each I_i being called an input, each q_i a state, and each U_i an output, together with two functions $\delta(I, q)$ and $\lambda(I, q)$, where I is an input, q a state, $\delta(I, q)$ a state, and $\lambda(I, q)$ an output. It shall be assumed that $\lambda(I, q)$ is independent of I, i.e., $\lambda(I, q) = \lambda(q)$. $\delta(I, q)$ is said to be the new state and $\lambda(I, q)$ the output of the machine upon application of input I.

and internal states, i.e., each new state in relation to each input and each present state as well as each output in relation to each present state, are known.

By an *experiment* (of length s) is meant a sequence of inputs $I_1, I_2, \cdots, I_s$. By *experiment which distinguishes the terminal state of the machine* is meant an experiment, possibly depending on the unknown initial state, such that the resulting output allows us to determine the terminal state, i.e., the state of the machine after applying the last input of the experiment. If the experiment E is independent of which unknown state from a set A of admissible states of the machine is the initial state, then E is said to be *uniform* (with respect to A).

Two states q_i and q_j of a machine S are said to be *distinguished* if there exists an experiment E such that the output of q_i under E, i.e., the output with the machine initially in state q_i, is not identical with the output of q_j under E. A machine is said to be *distinguished* if each pair of distinct states can be distinguished.

In [3, Th. 8], Moore showed that for any distinguished machine in an unknown initial state, an experiment, depending on the initial state, can be found which distinguishes the terminal state. We shall be concerned here with constructing uniform experiments and estimating the length of a minimum uniform experiment.

Fundamental to our work are the partitions P_k as defined in [3]. Following Moore, let S be a machine and for each pair of states q_i and q_j of S write $q_i R_k q_j$ if there is no experiment of (at most) length k which distinguishes q_i and q_j. Then R_k is an abstract equivalence relation. Let P_k be the partition associated with this equivalence relation.

Moore [3, p. 145–146] has noted the following two results which are used later.

Lemma 1.1 *For each positive integer* k, P_{k+1} *is a refinement of* P_k, *that is, each class in* P_{k+1} *is a subclass of some class in* P_k.

Lemma 1.2. *Two states* q_i *and* q_j *in the same class of the partition* P_k *are in different classes of* P_{k+1} if and only if *there exists an input which transforms* q_i *and* q_j *into states that are in different classes of* P_k. *If* $C_1 \cup C_2$ *is a subset of a class, say* B_e, *of* P_i, *and* C_1 *and* C_2 *are subsets of distinct classes of* P_{i+1}, *then there exists an input* I *and distinct classes* B_1 *and* B_2 *of* P_i *such that*[2] $I(C_1) \subseteq B_1$ *and* $I(C_2) \subseteq B_2$. *If* B_1 *and* B_2 *are two distinct classes of* P_i, *if* I *is an input, and if* $I(C_1) \subseteq B_1$ *and* $I(C_2) \subseteq B_2$, *then* C_1 *and* C_2 *are subsets of different classes of* P_{i+1}.

When I, $C_1 \cup C_2$, C_1, and C_2 are related as in the second sentence of lemma 1.2, we say that I *splits* $C_1 \cup C_2$ *into classes* C_1 *and* C_2 (of P_{i+1}).

Lemma 1.2 will now be extended for later use.

Lemma 1.3. *If* $C_1 \cup C_2$ *is a subset of a class, say* B_e, *of* P_i, *where* $i \geqq 2$, *and if an input* I *splits* $C_1 \cup C_2$ *into the classes* C_1 *and* C_2 *of* P_{i+1}, *then there exists two distinct classes* B_1 *and* B_2 *of* $P_i - P_{i-1}$ *such that* $I(C_1) \subseteq B_1$, $I(C_2) \subseteq B_2$, *and* $B_1 \cup B_2$ *is a subset of a class of* P_{i-1}.

Proof. By hypothesis, $I(C_1) \subseteq B_1$ and $I(C_2) \subseteq B_2$ for two distinct classes

[2] If I is an input and q is a state of S, then by $I(q)$ is meant the terminal state of q upon application of I. If A is a set of states of S, then by $I(A)$ is meant the set $\{I(q)/q \in A\}$.

B_1 and B_2 in P_i. To see that both B_1 and B_2 are not in P_{i-1}, let us assume that one of them, say B_1, is in P_{i-1}. Two possibilities occur.

(a) Suppose that B_2 is in P_{i-1}. Then by lemma 1.2, I splits $C_1 \cup C_2$ into classes C_1 and C_2 of P_i.

(b) Suppose that B_2 is not in P_{i-1}. Then by lemma 1.1, $B_2 \subseteq B_a$ for some B_a in P_{i-1}. Now $B_a \neq B_1$ since B_1 and B_2 are in P_i, $B_2 \subseteq B_a$, and $B_1 \neq B_2$. Then $I(C_1) \subseteq B_1$ and $I(C_2) \subseteq B_2 \subseteq B_a$, so that I splits $C_1 \cup C_2$ into classes C_1 and C_2 of P_i, again effecting a contradiction.

Both cases lead to a contradiction. We therefore are forced to conclude that B_1 being in P_{i-1} is false and B_1 is in $P_i - P_{i-1}$.

An analogous argument shows that B_2 is in $P_i - P_{i-1}$.

Finally, suppose that $B_1 \cup B_2$ is not a subset of a class of P_{i-1}. Then by lemma 1.1 there exist distinct classes B_c and B_d in P_{i-1} such that $B_1 \subseteq B_c$ and $B_2 \subseteq B_d$. Here again we see that I splits $C_1 \cup C_2$ into classes C_1 and C_2 in P_i, a contradiction.

As a special case we get

Corollary. *For some integer $i \geqq 2$ let P_{i-1} contain exactly two classes, say B_1 and B_2. If an input I splits the subset $C_1 \cup C_2$ of the class B_e of P_i into the classes C_1 and C_2 of P_{i+1}, then either $I(C_1) \subseteq B_1$ and $I(C_2) \subseteq B_2$, or $I(C_1) \subseteq B_2$ and $I(C_2) \subseteq B_1$.*

Another known result which we need is lemma 1.4 below. It is implicit in the proof of theorem 6 of [3]; (b) is stated as theorem 3 of [1].

Lemma 1.4. *Let S be a distinguished (n, m, p) machine.* (a) *If P_j contains at least k classes, with $k \leqq n - 1$, then P_{j+1} contains at least $k + 1$ classes.* (b) *If P_1 contains r classes, then P_j contains at least $j + r - 1$, provided $j + r - 1 \leqq n$. In particular, P_{n-r+1} contains n classes.* (c) *P_1 contains at least two classes.*

Remark. Let S be a distinguished (n, m, p) machine and let P_{n-2} contain exactly $n - 1$ classes. Then (a) of lemma 1.4 yields the fact that for $2 \leqq i \leqq n - 1$, $P_{i+1} - P_i$ contains exactly two elements. Thus the Corollary to lemma 1.3 applies here for each i, $2 \leqq i \leqq n - 1$.

2. *A First Approach*

We now turn to the actual construction of uniform experiments. First though, we prove

Lemma 2.1. *Let S be a distinguished (n, m, p) machine and let A be a set of $k \leqq n$ distinct states of S. Then for each admissible (that is, in A) initial state q_σ, there exists an experiment E (depending on q_σ), of length at most $\frac{(k-1)}{2}(2n - k)$ such that*

(1) if S is initially in state q_σ, then the output (under E) is called U^* and the terminal state q_σ^*;

(2) regardless of the admissible initial state, if the output under E is U^*, then the terminal state is q_σ^*.

Proof. The proof is a variation of theorem 8 of [3].

By lemma 1.4 the partition P_{n-k+1} divides the states of S into at least $n - k + 2$ classes. At least one of the admissible initial states q_i, $q_i \neq q_\sigma$, is in a class different from that in which q_σ occurs. For suppose that all the admissible q_i are in the same class. Since there are only $n - k$ other states in S, there can be at most $n - k + 1$ different classes altogether; this is a contradiction. Hence there exists an experiment E_1, of length at most $n - k + 1$, and an admissible initial state q, $q \neq q_\sigma$, such that the output U_1 under E_1 from q_σ is not the same as the output from q. Let $q_{2,1}$ be the terminal state of q_σ under E_1. Let $q_{2,1}, q_{2,2}, \cdots, q_{2,r(2)}$, be the different terminal states of those admissible initial states which yield U_1 upon application of E_1. Clearly $r(2) \leqq k - 1$. Now repeat this procedure with q_σ replaced by $q_{2,1}$, k by $r(2)$, and A by $\{q_{2,1}, q_{2,2}, \cdots, q_{2,r(2)}\}$. Continue this method by induction in the obvious way, obtaining experiments $E_2, E_3, \cdots$, until the first stage, say the jth, that there is only one terminal state, say $q_{j+1,1}$. Obviously $j \leqq k - 1$. Let $E = E_1E_2 \cdots E_j$. The length of E is at most

$$\sum_2^k (n - i + 1) = \frac{(k-1)}{2}(2n - k).$$

It is readily seen that E satisfies the conclusion of the lemma.

Remarks. (1) Whether or not the bound given in lemma 2.1 is the best possible is unknown.

(2) If the E constructed in lemma 2.1 is such that $j = k - 1$, then E is a uniform experiment which can distinguish the terminal state. For if $k - 1$ stages are required, then for each integer $i > 1$, $r(i) = k - i + 1$, so that at the ith stage precisely one state is ascertained.

Using lemma 2.1 and theorem 2.1 below, we now show the existence of a uniform experiment for a distinguished machine.

Theorem 2.1. *Let S be a distinguished (n, m, p) machine and for some $t \leqq n$ and each $k \leqq t$ let $\alpha(k)$ be a number with the following property: "For* every *set of k states, if S initially is in one of these states, then there exists an experiment (usually depending on the initial state) of length at most $\alpha(k)$ which satisfies* (1) *and* (2) *of Lemma* 2.1." *Then for any set A of t states, say $A = \{q_i/i \leqq t\}$, there exists a uniform experiment E of length at most $\sum_2^t \alpha(k)$ which distinguishes the terminal state of S.*

Proof. By hypothesis there exists an experiment E_1 of length at most $\alpha(t)$ such that

(i) if S is initially in state q_1, then the output is called U_1^* and the terminal state q_1^*; and

(ii) regardless of the admissible initial state, if the output under E_1 is U_1^*, then the terminal state is q_1^*.

Let $q_1^*, q_{2,1}, q_{2,2}, \cdots, q_{2,r(2)}$ be the different terminal states under E_1 for all admissible initial states. Clearly $r(2) \leqq t - 1$. Continuing by induction let us assume that the admissible initial states are the $r(i)$, $i \geqq 2$, states $q_{i,1}, q_{i,2}, \cdots, q_{i,r(i)}$ and that $r(i) \leqq t - i + 1$. By hypothesis there exists an experiment E_i of length at most $\alpha(t - i + 1)$ such that

(iii) if S is initially in state $q_{i,1}$, then the output is called U_i^* and the terminal state q_i^*; and

(iv) regardless of the admissible initial state, if the output under E_i is U_i^*, then the terminal state is q_i^*.

Let q_i^*, $q_{i+2,1}$, $\cdots$, $q_{i+1,r(i+1)}$ be the different terminal states under E_i for all admissible initial states. Clearly $r(i + 1) \leqq t - (i + 1) + 1$. In this way the procedure is continued until the first stage, say the jth, that the terminal states are either q_j^* and $q_{j+1,1}$, or q_j^* alone. Obviously $j \leqq t - 1$.

We now show that the experiment $E = E_1E_2 \cdots E_j$ satisfies the conclusion of the theorem. Summing we see that E is of the length at most $\sum_2^k \alpha(k)$. Suppose that the unknown admissible state of S is q_v. Let U_i be the output from each stage of the experiment, so that the total output is $U = U_1U_2 \cdots U_j$. Let q_r be the terminal state of q_v under E. Starting in the initial state q_1, let w_1 be the terminal state under $E_2E_3 \cdots E_j$. For each $2 \leqq i \leqq j - 1$, let w_i be the terminal state of q_i^* under $E_{i+1} \cdots E_j$. In view of (ii) and (iv) it is obvious that either

(a) there is a smallest integer, call it i, such that $U_i = U_i^*$, in which case $q_r = w_i$ if $i \leqq j - 1$ and $q_r = q_j^*$ if $i = j$; or

(b) for no integer i is $U_i = U_i^*$, in which case $q_r = q_{j+1,1}$. Q.E.D.

Lemma 2.1 states that one possible value for $\alpha(k)$ is $\frac{(k-1)}{2}(2n - k)$. Then

$$\begin{aligned}
\sum_{k=2}^{t} \alpha(k) &= \sum_{k=2}^{t} \frac{(k-1)}{2}(2n-k) \\
&= \frac{1}{2}\sum_{2}^{t} [k(2n+1) - 2n - k^2] \\
&= \frac{1}{2}\left[(2n+1)\sum_{2}^{t} k - 2n(t-1) - \sum_{2}^{t} k^2\right] \\
&= \frac{1}{2}\left[(2n+1)\left(\frac{t-1}{2}\right)(t+2) - 2n(t-1) - \left\{\frac{t}{6}(t+1)(2t+1) - 1\right\}\right] \\
&= \frac{1}{2}\left[n(t^2 - t) + \frac{2t - t^3}{6}\right].
\end{aligned}$$

Hence we have

THEOREM 2.2. *Let S be a distinguished (n, m, p) machine. Then for any set A of $k \leqq n$ states, there exists a uniform experiment E of length at most*

$$\tfrac{1}{2}\left[n(k^2 - k) + \frac{2k - k^3}{6}\right]$$

which distinguishes the terminal state of S.

Letting $k = n$, we get

THEOREM 2.3. *Let S be a distinguished (n, m, p) machine. Then there exists a uniform experiment E of length at most $\frac{n(2n-1)(n-1)}{6}$ which distinguishes the terminal state of S.*

3. *Knowledge of* P_1

In sections 3 and 4 we lower the bounds given in theorems 2.1 and 2.2. Our estimates will be based on knowledge of the number of classes in P_1. We shall see that whereas theorem 2.2 yields an estimate of approximately $\frac{1}{2}nk^2$, the results of this and the next section yield an estimate of less than nk, n being the number of distinguishable states in the machine and k the number of admissible initial states.

LEMMA 3.1. *Let S be any distinguished (n, m, p) machine and let P_1 contain at least r classes. Then for any set $A = \{q_i/i \leqq k\}$ of $k \leqq n$ states*

(a) *if $n - r + 1 \leqq 2$, then there exists a uniform experiment of length at most $(k - 1)(n - r + 1)$ which distinguishes the terminal state of S;*

(b) *if $n - r + 1 \geqq 3$, then there exists a uniform experiment of length at most $(k - 1)(n - r + 1) + 2 - k$ which distinguishes the terminal state of S.*

PROOF. The proof of (a) is found in lemma 4.1. We therefore shall consider only the proof of (b) here.

For $k = 1$ and $k = 2$, (b) is obviously true. Suppose that (b) is true for all $i \leqq k - 1 < n - r + 2$. Then $n - k - r + 3 \geqq 1$, so that $P_{n-k-r+3}$ exists. Now the partition $P_{n-k-r+3}$ divides the states of S into at least $n - k - r + 3 + r - 1 = n - k + 2$ disjoint non-empty classes. By the argument as in lemma 2.1, we see that all the q_i in A cannot be in the same class. Thus there exists an experiment E_1 of length at most $n - k - r + 3$ which divides the q_i, $1 \leqq i \leqq k$, into (at least two non-empty) classes $A_1, A_2, \cdots, A_s$. Let B_1 consist of the terminal states, under E_1, of those states initially in A_1, and B_2 the terminal states of those states initially in $\bigcup_{i \geq 2} A_i$.

Suppose that B_1 contains v states and B_2 at most $k - v$ states. If both B_1 and B_2 contain just one state, then these states are known and we are finished. If just one of them, say B_1, contains only one element, say q_a, then the terminal state of q_a under any experiment will be known. By our induction hypothesis, there exists a uniform experiment E_2 of length at most $[(k - 1) - 1](n - r + 1) + 2 - (k - 1)$ which distinguishes the terminal state of B_2. The length of $E = E_1E_2$ is at most

$$\begin{aligned}(n - k - r + 3) + (k - 2)(n - r + 1) + 3 - k \\ &= (k - 1)(n - r + 1) + 5 - 2k \\ &\leqq (k - 1)(n - r + 1) + 2 - k \qquad \text{since } k \geqq 3.\end{aligned}$$

Clearly E satisfies the conclusion of the lemma. An analogous result holds if B_2 contains just one and B_1 more than one element. Suppose that both B_1 and B_2 contain at least two elements. By our induction hypothesis there exists a uniform experiment E_2 of length at most $(v - 1)(n - r + 1) + 2 - v$ which distinguishes the terminal state of each initial state in B_1. Let B_3 be the terminal states, under E_2, of the states initially in B_2. By our induction hypothesis there exists a uniform experiment E_3 of length at most $(k - v - 1)(n - r + 1) + 2 - (k - v)$ which distinguishes the terminal state of each initial state in B_3. Let $E = E_1E_2E_3$. The length of E is at most

$$\begin{aligned}(n - k - r + 3) &+ [(v - 1)(n - r + 1) + 2 - v] \\ &+ [(k - v - 1)(n - r + 1) + 2 - (k - v)] \\ &\leqq (n - k - r + 3) + (v - 1)(n - r + 1) + (k - v - 1)\, n - r + 1), \\ &\qquad \text{since } v \geqq 2 \text{ and } k - v \geqq 2, \\ &= (k - 1)(n - r + 1) + 2 - k.\end{aligned}$$

By induction (b) is true for all $k \leqq n - r + 2$.

Now suppose that (b) is true for $j \leqq k - 1$, where $k \geqq n - r + 3$. It is readily seen that all q_i in A cannot be in the same class of P_1. Repeat the procedure given above, replacing $n - k - r + 3$ by 1. The experiment $E = E_1E_2E_3$ obtained is of length at most

$$\begin{aligned}1 + [(v - 1)(n - r + 1) + 2 - v] &+ [(k - v - 1)(n - r + 1) + 2 - (k - v)] \\ &= 5 - k + (k - 2)(n - r + 1) \\ &\leqq 2 - k + (k - 1)(n - r + 1) \qquad \text{since } n - r + 1 \geqq 3.\end{aligned}$$

In this way the induction is continued to $k = n$. Q.E.D.

Using lemma 3.1 we now obtain a sequence of bounds on the length of a minimal uniform experiment for the case when $k \leqq n - r + 2$.

THEOREM 3.1. *Let S be any distinguished (n, m, p) machine and let P_1 contain at least r classes with $n - r + 1 \geqq 3$. Then for any set $A = \{q_i/i \leqq k\}$ of $k \leqq n$ states, where $k \leqq n - r + 2$, and for each positive integer u, there exists a uniform experiment E_u of length at most $f_u(k, n) = (k - 1)(n - r + 1) + 2^{u+1} - 2 - uk$ which distinguishes the terminal state of S.*

PROOF. For $u = 1$ the conclusion is given by lemma 3.1. Now assume the theorem is true for all $u \leqq w$. By the usual argument, there exists an experiment E_1^* of length at most $n - k - r + 3$ which partitions the admissible initial states into (at least two non-empty) classes $A_1, A_2, \cdots, A_s$. Let B_1 and B_2 be as in lemma 3.1. By our induction hypothesis there exists a uniform experiment E_2^* of length at most $(v - 1)(n - r + 1) + 2^{w+1} - 2 - wv$ which distinguishes the terminal states of B_1. Let B_3 be the terminal states of B_2 under E_2^*. Another application of our induction hypothesis yields a uniform experiment E_3^* of length at most $(k - v - 1)(n - r + 1) + 2^{w+1} - 2 - w(k - v)$ which distinguishes the terminal states of B_3. Then $E_{w+1} = E_1^*E_2^*E_3^*$ is a uniform experiment which distinguishes the terminal states of A. The length of E_{w+1} is at most

$$\begin{aligned}(n - k - r + 3) &+ (v - 1)(n - r + 1) + 2^{w+1} - 2wv \\ &+ (k - v - 1)(n - r + 1) + 2^{w+1} - 2 - w(k - v) \\ &= (k - 2)(n - r + 1) + 2 \cdot 2^{w+1} - 4 - wk + n - k - r + 3 \\ &= (k - 1)(n - r + 1) + 2^{w+2} - 2 - (w + 1)k,\end{aligned}$$

which shows that the theorem is true for $w + 1$. Q.E.D.

REMARK. Let k, r, and n be fixed. To find the appropriate u which yields a minimum value among all $f_u(k, n)$, we set $\frac{df}{du} = 0$. Solving, we find that

$$u_{\min} = \frac{\log k - \log \log 4}{\log 2}.$$

Letting $[u_{\min}]$ be the greatest integer $\leqq u_{\min}$, the desired minimum of the $f_u(k, n)$ occurs at either $u = [u_{\min}]$ or $u = [u_{\min}] + 1$.

Below we list the appropriate formula for $2 \leqq k \leqq 9$.

k	bound	k	bound
2	$(n - r) + 1$	6	$5(n - r) - 1$
3	$2(n - r) + 1$	7	$6(n - r) - 2$
4	$3(n - r) + 1$	8	$7(n - r) - 3$
5	$4(n - r)$	9	$8(n - r) - 5$

If $r = 2$, then $k \leqq n - r + 2 = n$.

THEOREM 3.2. *Let S be a distinguished (n, m, p) machine and let $A = \{q_i/i \leqq k\}$ be any set of $k \leqq n$ states. Then for each positive integer u, there exists a uniform experiment of length at most $g_u(k, n) = (k - 1)(n - 1) + 2^{u+1} - 2 - uk$ which distinguishes the terminal state of S. In particular, when $k = n$ the experiment is of length $n^2 + n(u + 2) + 2^{u+1} - 1$.*

PROOF. Letting $r = 2$, the theorem follows from theorem 3.1 if $n - r + 1 \geqq 3$, i.e., $n \geqq 4$. For $n = 1, 2$, and 3, the theorem is easily verified by case analysis and the fact that for $k = n = 3$ a desired uniform experiment of length 3 can be found.

A simple consequence of theorem 3.2 is

THEOREM 3.3. *Let S be a distinguished (n, m, p) machine and let $A = \{q_i/i \leqq k\}$ be any set of $k \leqq n$ states. Let the number of different terminal states of the admissible initial states under an experiment E of length α be $k - j$. Then for each positive integer u there exists a uniform experiment E_u of length at most $\alpha + g_u(k - j, n)$ which distinguishes the terminal state of S. In particular, if there exists an input which changes two distinct admissible initial states to the same state, then E_u is of length at most $1 + g_u(k - 1, n)$.*

4. *Knowledge of the Last P_i*

In this section our estimates will be based on knowledge of the last P_i, i.e., the first P_i such that $P_i = P_{i+1}$.

LEMMA 4.1. *Let S be a distinguished (n, m, p) machine and let P_{n-s} contain n classes. Then for any set $A = \{q_i/i \leqq k\}$ of $k \leqq n$ states*

(a) *for $1 \leqq k \leqq s + 1$ there exists a uniform experiment E of length at most $(k - 1)(n - s)$ which distinguishes the terminal state of S;*

(b) *for $s + 1 < k \leqq n$ there exists a uniform experiment E of length at most $(k - 1)(n - s) + (s + 1 - k)$ which distinguishes the terminal state of S.*

Furthermore, both $(k - 1)(n - s)$ and $(k - 1)(n - s) + (s + 1 - k)$ may serve as bounds for all $k \leqq n$, i.e., for either (a) *or* (b).

PROOF. The last statement is obvious since we then replace the bounds in (a) and (b) by larger bounds.

(a) For $k = 1$ no experiment is needed. Hence (a) is trivially true. For $k = 2$, the conclusion is true since, in view of the hypothesis, any two states can be distinguished by an experiment of length at most $n - s$. Suppose that (a) is true for all $k \leqq j < s + 1$. Let $k = j + 1$. Since $n - k + 1 \geqq n - (s + 1) + 1 = n - s$, by the hypothesis of the theorem there exists an experiment E_1 of length at most $n - s$ which divides the admissible q_i into (at least two non-empty) classes $A_1, A_2, \cdots, A_r$. The rest of the proof of (a) parallels the argument given in lemma 3.1.

(b) Let $k > s + 1$. Our argument is by induction. Proceeding as in lemma 3.1 we obtain an experiment E_1 of length at most $n - k + 1$ and sets B_1 and B_2. Let B_1 contain v elements. By induction there exists a uniform experiment E_2 of length at most $(v - 1)(n - s)$ which distinguishes the terminal states of B_1. Letting B_3 be as in lemma 3.1, there exists a uniform experiment E_3 of length at most $(k - v - 1)(n - s)$ which distinguishes the terminal states of B_3. Then $E = E_1E_2E_3$ is a uniform experiment of length at most

$$(n - k + 1) + (v - 1)(n - s) + (k - v - 1)(n - s)$$
$$= (k - 1)(n - s) + (s + 1 - k)$$

which distinguishes the terminal state of S. Q.E.D.

Using lemma 4.1 and an argument similar to that given in theorem 3.1, the following result (whose proof we omit) may readily be shown.

THEOREM 4.1. *Let S be a distinguished (n, m, p) machine and let P_{n-s} contain n classes. Then for any set $A = \{q_i/i \leqq k\}$ of $k \leqq n$ states and each positive integer u,*

(a) *for $1 \leqq k \leqq s + 1$ there exists a uniform experiment E_u of length at most $(k - 1)(n - s)$ which distinguishes the terminal state of S;*

(b) *for $s + 1 < k \leqq n$ there exists a uniform experiment E_u of length at most $h_u(k, n) = (2^u - 1)(s + 1) - uk + (k - 1)(n - s)$ which distinguishes the terminal state of S.*

Furthermore, both $(k - 1)(n - s)$ and $h_u(k, n)$ may serve as bounds for all $k \leqq n$, i.e., for either (a) *or* (b).

REMARKS. Let k, n, and s be fixed. To find the appropriate u which yields a minimum value among all $h_u(k, n)$, we set $\frac{dh}{du} = 0$. Solving, we find that the minimum occurs when $u = \max\{1, a\}$, or $u = \max\{1, a + 1\}$, where

$$a = \left[\frac{\log k - \log\log 2 - \log(s + 1)}{\log 2}\right].$$

(2) Let S be a distinguished (n, m, p) machine and suppose that P_1 contains exactly r classes. Then for $n - r + 1 \geqq 3$ and $k \leqq n - r + 2$ theorem 3.1 gives an estimate of $(k - 1)(n - r + 1) + 2^{u+1} - 2 - uk$. Letting $n - s = n - r + 1$, so that $s = r - 1$, theorem 4.1 gives an estimate of $(2^u - 1)r - uk + (k - 1) \times (n - r + 1)$. For $r \geqq 3$, the first estimate is smaller.

5. *Composition of Machines*

In this section we consider machines which are constructed from "simpler" machines. The estimates obtained for these new machines are simple consequences of previous results.

We now note the following result, whose proof is a trivial variation of the proof of theorem 7 of [3].

LEMMA 5.1. *Let S be a distinguished (n, m, p) machine which is a (direct) sum*[3] $\sum_{i=1}^{d} T_i$, *each T_i being an (n_i, m, p) machine. If n_a and n_b, where $n_a \geqq n_b$, are the two largest n_i (possibly equal) in the sequence of integers $\{n_i\}$, then $P_{n_a+n_b-1}$ consists of n classes.*

PROOF. It is sufficient to show that any two states can be distinguished by an experiment of length at most $n_a + n_b - 1$. To this end let q_x and q_y be any two states in S. Suppose that both q_x and q_y are in the same machine T_i, say T_z. Since T_z has at most n_a states, q_x and q_y can be distinguished by an experiment of length at most $n_a - 1 \leqq n_a + n_b - 1$. Suppose that q_x is in T_σ and q_y in T_τ, with $T_\sigma \neq T_\tau$. Then the machine $T(\sigma, \tau)$, which is defined as the (direct) sum of T_σ and T_τ, is a distinguished $(n_a + n_b, m, p)$ machine, due to the maximality properties of n_a and n_b. Hence q_x and q_y, considered in $T(\sigma, \tau)$, can be distinguished by an experiment E of length at most $n_a + n_b - 1$. Then q_x and q_y, considered in S, are distinguished by E.

Letting $n - s = n_a + n_b - 1$, so that $s = n + 1 - (n_a + n_b)$, in theorem 4.1 we get

THEOREM 5.1. *Under the hypothesis of lemma 5.1, for any set $A = \{q_i/i \leqq k\}$ of $k \leqq n$ states, and each positive integer u*

(a) *for $1 \leqq k \leqq n + 2 - (n_a + n_b)$ there exists a uniform experiment E_u of length $(k - 1)(n_a + n_b - 1)$ which distinguishes the terminal state of S;*

(b) *for $s + 1 < k \leqq n$ there exists a uniform experiment E_u of length $h_u(k, n) = (2^u - 1)[n + 2 - (n_a + n_b)] - uk + (k - 1)(n_a + n_b - 1)$ which distinguishes the terminal state of S.*

Furthermore, both $(k - 1)(n_a + n_b - 1)$ and $h_u(k, n)$ may serve as bounds for all $k \leqq n$ in either (a) *or* (b).

REMARKS. (1) From remark 1 following theorem 4.1, the minimum $h_u(k, n)$ occurs when $u = \max\{1, a\}$ or $u = \max\{1, a + 1\}$, where

$$a = \left[\frac{\log k - \log\log 2 - \log\{n + 2 - (n_a + n_b)\}}{\log 2}\right].$$

(2) The terminal state is in the same submachine as the initial state. Hence the experiment in theorem 5.1 distinguishes the submachine containing the initial state.

[3] Let $W = \{S_i/i \leqq s\}$ be a family of (n_i, m, p) machines. By a relettering if necessary we may assume that all the machines have the same inputs and outputs. Label the states of each S_i as $q_{r(i)+1}, \cdots, q_{r(i)+n_i}$, where $r(1) = 0$ and $r(i) = \sum_{k<i} n_k$ for $i \geqq 2$. Then the (direct) sum $\sum_{S_i \in W} S_i$ is the machine whose states consist of all q_i, $1 \leqq i \leqq r(s) + n_s$, the input affecting the states in each S_i considered a submachine of S as if S_i were by itself, i e. independent of the other machines.

Theorem 5.1 (a) and remark (2) above yield

COROLLARY 1. *Suppose that W is a set consisting of d(n, m, p) machines* T_i *and that each state in any* T_i *can be distinguished from any other state of any* T_j. *Then there exists a uniform experiment E of length at most* $(nd - 1)(2n - 1)$ *which, when applied to an unknown initial state of an unknown machine* T_j, *distinguishes both the terminal state of that machine and the machine itself.*

In theorem 9 of [3], Moore has defined a certain class of distinguished machines which he calls $R_{n,m,p}$. As intimated there the machines in $R_{n,m,p}$ have the property described in the first sentence of corollary 1. Moore has shown that the number of machines in $R_{n,m,p}$ is no more than $n^{nm}p^n/n!$. [It is not difficult to lower that bound, but this is another matter.] Hence we have

COROLLARY 2. *There exists a uniform experiment E of length at most*

$$\left(\frac{n^{nm+1}p^n}{n!} - 1\right)(2n - 1) < \frac{2n^{nm+2}p^n}{n!}$$

which, when applied to an unknown state of an unknown machine in $R_{n,m,p}$, *distinguishes both the terminal state of that machine and the machine itself.*

Another way of combining several machines to form a new machine is by means of the "product".

DEFINITION. For $1 \leqq i \leqq t$ let S_i be an (n_i, m_i, p_i) machine, the typical state, input, and output being q_j^i, I_k^i, and U_r^i respectively. Then the *product* $S = \prod_{i=1}^t S_i$ or $S = S_1 \times S_2 \times S_3 \times \cdots \times S_t$ is the $(\Pi n_i, \Pi m_i, \Pi p_i)$ machine defined as follows. The states of S consist of all t-tuples $(q_i^1, q_j^2, \cdots, q_k^t)$; the inputs consist of all t-tuples $(I_i^1, I_j^2, \cdots, I_k^t)$; and the outputs consist of all t-tuples $(U_i^1, U_j^2, \cdots, U_k^t)$. The input $I = (I^1, I^2, \cdots, I^t)$ changes the state $(q^1, q^2, \cdots, q^t)$ to the state $(I^1(q^1), I^2(q^2), \cdots, I^t(q^t))$; and the output from state $(q^1, \cdots, q^\sigma, \cdots, q^t)$ is $(U^1, \cdots, U^\sigma, \cdots, U^t)$, where U^σ is the output from state q^σ.

We next note the following result:

"Let i be fixed and let $E_i = \{I_{i,1}, \cdots, I_{i,\xi}, \cdots, I_{i,\nu}\}$ be an experiment of S_i which distinguishes the two states q_a^i and q_b^i of S_i. Let q_a and q_b be any two states of $S = \Pi S_i$ whose ith coordinates are q_a^i and q_b^i respectively. Then any experiment $I_1, \cdots, I_\xi, \cdots, I_\nu$, having the property that the ith coordinate of each I_ξ is $I_{i,\xi}$, distinguishes q_a and q_b in S."

From this observation we immediately infer the ensuing two facts.

(1) If each S_i is distinguished, then so is $S = \Pi S_i$.

(2) Let A consist of any $k \leqq n = \Pi n_i$ states of $S = \Pi S_i$, and for each i let T_i consist of the ith coordinates of the states in A, i.e.,

$$T_i = \{q_j^i / \text{for some } q_\xi \text{ in } A, \text{ the } i\text{th coordinate of } q_\xi \text{ is } q_j^i\}.$$

For each i let E_i be a uniform experiment of length γ_i which distinguishes the terminal state of T_i. Then there exists a uniform experiment E of length $\gamma = \max_i \{\gamma_i\}$ which distinguishes the terminal state of S.

In conjunction with theorems 3.1 and 4.1, (2) above yields bounds on the length of a uniform experiment for k states in ΠS_i. We leave the details to the reader.

In passing, we note that product is distributive over sum, i.e., $A \times (B \times C) = (A \times B) + (A \times C)$.

6. *Permutation Machines*

In this, the final section, we deviate from the topic of lengths of uniform experiments. Here we introduce a special machine and investigate the associated P_i.

DEFINITION: The machine S is called a *permutation* machine if each input merely permutes the states of the machine.

THEOREM 6.1. *Let S be a distinguished (n, m, p) permutation machine such that P_{n-2} has exactly $n - 1$ classes. Then each partition P_k has one of the following forms:*

(a) *P_1 consists of two classes B_1 and B_2 of power t and $n - t$ respectively, t and $n - t$ each being relatively prime with n.*

(b) *P_{n-1} consists of n classes, each consisting of just one state.*

(c) *P_k consists of the $k + 1$ classes*

$$N_1^k, N_2^k, \cdots N_{\alpha(k)}, R_1^k, \cdots, R_{\gamma(k)}^k$$

each N_i^k being of power[4] *v_k and each R_i^k being of power x_k, with $x_k < v_k$. Then P_{k+1} is obtained from P_k by splitting one of the N_i^k into two classes C_k and D_k of power x_k and $v_k - x_k$ respectively, of P_{k+1}.*

(d) *P_k consists of the $k + 1$ classes*

$$N_1^k, \cdots, N_{\alpha(k)}^k, Q_1^k, \cdots, Q_{\beta(k)}^k, R_1^k, \cdots, R_{\gamma(k)}^k,$$

where each N_i^k is of power v_k, each Q_i^k is of power $v_k - x_k$, and each R_i^k is of power x_k, with $x_k < v_k$. Then P_{k+1} is obtained from P_k by splitting one of the N_i^k into two classes C_k and D_k, of power x_k and $v_k - x_k$ respectively, of P_{k+1}.

PROOF. Since P_{n-2} contains exactly $n - 1$ classes, by lemma 1.4 for each $i \leqq n - 1$, P_i consists of exactly $i + 1$ classes.

(a) Suppose that t is not relatively prime to n. Then $t = zx$ and $n = zy$, where x, y, and z are positive integers and $z > 1$. Thus $n - t = z(y - x)$, where $y - x$ is a positive integer. In other words, the power of each class in P_1 is an integral multiple of z. Using induction let us assume that for each $i \leqq k < n - 1$, P_i consists of $i + 1$ classes, the power of each class being an integral multiple of z. We now extend this statement to P_{k+1}.

To this end let the classes in P_k be $C_1, C_2, \cdots C_{k+1}$. By lemma 1.2 and the fact that P_{k+1} has exactly $k + 2$ classes, there exists an input, say I, which decomposes just one of the C_i of P_k, say C_a, into two classes D_1 and D_2 of P_{k+1}. Let C_b and C_c be classes of P_k such that $I(D_2) \subseteq C_b$ and $I(D_1) \subseteq C_c$. Suppose that $I(D_2) = C_b$. Let zw_1 and zw_2 be the powers of C_a and C_b respectively, w_1 and w_2 being positive integers. Since I is a permutation, the power of D_2 is zw_2. Then the power of D_1 is $zw_1 - zw_2 = z(w_1 - w_2)$. Thus the induction is continued to $k + 1$. Now suppose that $C_b - I(D_2)$ is non-empty. Since I is a

[4] By the power of a set is meant the number of elements in the set.

permutation, $I^{-1}[C_b - I(D_2)]$ is non-empty.[5] Let C_d be a class of P_k such that $I^{-1}[C_b - I(D_2)] \subseteq C_d$. Since I decomposes C_a into D_1 and D_2, $C_b \neq C_c$. Since $I(D_1) \subseteq C_c \neq C_b$, it is clear that $C_a \neq C_d$. As I cannot decompose both C_a and C_d, it therefore follows that $I^{-1}[C_b - I(D_2)] = C_d$. By our induction hypothesis, the power of C_a, C_b, and C_d is zw_1, zw_2, and zw_3 respectively, each w_i being a positive integer. From the fact that I is a permutation, thus a one-to-one onto function, the power of $C_b - I(D_2)$ is zw_3. Hence the power of $I(D_2)$, and thus D_2, is $zw_2 - zw_3 = z(w_2 - w_3)$. Since $D_1 = C_a - D_2$, the power of D_1 is $zw_1 - (zw_2 - zw_3) = z(w_1 - w_2 + w_3)$. Thus the power of each class of P_{k+1} is an integral multiple of z. By induction this statement becomes true for all P_k, in particular for P_{n-1}. Since $z > 1$, the classes of P_{n-1} do not consist of just one element which is a contradiction. We conclude that t and, thus, $n - t$ are relatively prime to n.

(b) Statement (b) is known [3].

(c) Let P_k consist of the classes N_i^k and R_i^k as given in (c) of the statement of the theorem. Two possibilities exist.

(i) Suppose that no input splits any of the N_i^k into two classes of P_{k+1}. In view of (b) and the fact that $P_{s+1} - P_s$ consists of just two elements for each $s \leqq n - 2$, there exists a smallest integer, call it j, $j > k$, such that one of the N_i^k, say N_ξ^k, in P_j is split by an input I into two classes F_1 and F_2 in P_{j+1}. Let H_1 and H_2 be the two distinct classes in P_j such that $I(F_1) \subseteq H_1$ and $I(F_2) \subseteq H_2$. Since P_j is obtained by decomposing classes of P_k, there exist two classes, say H_3 and H_4, not necessarily distinct, in P_k such that $H_1 \subseteq H_3$ and $H_2 \subseteq H_4$. Then $I(F_1) \subseteq H_3$ and $I(F_2) \subseteq H_4$. Since N_ξ^k is not split into two classes of P_{k+1}, it follows that $H_3 = H_4$. Then $I(F_1) \cup I(F_2) = I(F_1 \cup F_2) = I(N_\xi^k) \subseteq H_4$. Since N_ξ^k is of power v_k, $v_k > x_k$, and $v_k > v_k - x_k$, H_4 must be of power v_k. Thus H_4 must be one of the N_i^k, say N_λ^k. Since H_1 and H_2 are subsets of N_λ^k, then N_λ^k must have split prior to P_j; this is a contradiction. Hence this case cannot occur.

(ii) There exists an input I which splits one of the N_i^k, say N_δ^k, into two classes, C_k and D_k, of P_{k+1}. In what follows, we show that either C_k or D_k is of power x_k. A relabelling then makes C_k of power x_k, and thus D_k of power $v_k - x_k$.

Now there exist two classes H_1 and H_2 of P_k such that $I(C_k) \subseteq H_1$ and $I(D_k) \subseteq H_2$. Suppose that H_1 and H_2 are both of power v_k. Since $P_{k+1} - P_k$ has just two elements, N_δ^k is the only class in P_k which is split into two classes of P_{k+1}. Combining this with the fact that $x_k < v_k$ we see that for $i \neq \delta$, $I(N_i^k) = N_{\sigma(i)}^k$ for some $\sigma(i)$, with $N_{\sigma(i)}^k$ different for different i. Clearly no $N_{\sigma(i)}^k$ can be either H_1 and H_2. Hence there are only $\alpha(k) - 2$ classes $N_{\sigma(i)}^k$ and $\alpha(k) - 1$ classes N_i^k, which is a contradiction. Therefore at least one of the H_i, say H_1, is of power x_k, i.e., is one of the R_i^k.

If $x_k = 1$, then H_1 is of power 1, thus $I(C_k)$ and C_k are both of power 1, so that we are through. Suppose that $x_k > 1$. Assume now that $H_1 - I(C_k)$ is non-empty. We shall show that this assumption leads to a contradiction so

[5] By I^{-1} is meant the inverse function of I.

that $H_1 - I(C_k)$ is empty, i.e., $I(C_k) = H_1$ whence $I(C_k)$, thus C_k, is of power x_k. Since $H_1 - I(C_k)$ is non-empty, it is of power less than x_k. Hence $F = I^{-1}[H_1 - I(C_k)]$ is of power less than x_k. Let H_3 be the class in P_k containing F and let $G = H_3 - F$. Since H_3 has at least x_k elements, G is non-empty. Since $I(F) = H_1 - I(C_k)$ whereas no part of N_δ^k maps into $H_1 - I(C_k)$, it follows that $H_3 \neq N_\delta^k$. As $I^{-1}[H_1] = C_k \cup F$, G is not mapped into H_1 by I. Consequently I splits H_3 as well as N_δ^k. This is a contradiction. Hence $I(C_k) = H_1$.

(d) Let P_k consist of the classes N_i^k, Q_i^k, and R_i^k, as given in (d) of the theorem. We proceed by induction. In going from P_{k-1} to P_k our induction yields a set N_δ^{k-1} in P_{k-1} which is split into Q_σ^k and R_τ^k in P_k.

(iii) Suppose that no input splits any of the N_i^k into two classes of P_{k+1}. An argument parallel to that given in (i) above yields a contradiction, so that (iii) does not occur.

(iv) Suppose that there exists an input I which splits one of the N_i^k, say N_δ^k, into two sets C_k and D_k of P_{k+1}. By the corollary to lemma 1.3 and a relabelling if necessary, $I(C_k) \subseteq Q_\sigma^k$ and $I(D_k) \subseteq R_\tau^k$. Since $C_k \cup D_k$ and $Q_\sigma^k \cup R_\tau^k$ both are of power v, and since I is a permutation, $I(C_k) = Q_\sigma^k$ and $I(D_k) = R_\tau^k$. Thus C_k is of power x_k and D_k of power $v_k - x_k$. Hence the induction is extended and the theorem is completely proved.

Remarks. (1) The partititions P_i of the machine in theorem 6.1 occur sequentially in the following manner. P_1 consists of two sets, one of power t and one of power $n - t$, each relatively prime to n. By a relabelling if necessary we may assume that $t < n - t$. P_2 is obtained by decomposing the class with $n - t$ elements into two classes of powers t and $n - 2t$ respectively. This process is continued until classes of powers t and $n - ut = r < t$ are obtained. The classes of t elements then are decomposed (one at a time) into classes of powers r and $t - r = s$ respectively. When this is completed the classes with max (r, s), say s, are decomposed (one at a time) into classes with r and $s - r$ elements respectively. The procedure in the previous sentence is then repeated, with s replaced by $s - r$. This process is continued until at P_{n-1} each class contains just one element.

(2) Theorem 6.1 is no longer true if the condition on P_{n-2} is removed. For example, let S be as follows:

Present State	New State: Input 0	New State: Input 1	Present State	Output
q_1	q_2	q_2	q_1	1
q_2	q_3	q_3	q_2	1
q_3	q_4	q_4	q_3	0
q_4	q_5	q_5	q_4	0
q_5	q_6	q_6	q_5	0
q_6	q_1	q_1	q_6	0

$P_1 = \{(q_1, q_2), (q_3, q_4, q_5, q_6)\}$; $P_2 = \{(q_1), (q_2), (q_6), (q_3, q_4, q_5)\}$;
$P_3 = \{(q_1), (q_2), (q_5), (q_6), (q_3, q_4)\}$; $P_4 = \{(q_1), (q_2), (q_3), (q_4), (q_5), (q_6)\}$.

As a corollary to theorem 6.1 we have

THEOREM 6.2. *Let S be a distinguished (n, m, p) permutation machine such that P_{n-2} has exactly $n - 1$ classes. Furthermore, suppose that there exists one and only one state, say q_1, whose output is U. Then for each positive integer $k < n$, the partition P_k consists of k classes of exactly one element and one class of $n - k$ elements.*

The proof is obvious since the condition about U means that one of the classes in P_1 contains precisely one element.

REMARK: Theorem 6.2 is no longer true if the hypothesis on S being a permutation machine is removed. For example, let S be the following machine:

Present State	New State		Present State	Output
	Input 0	Input 1		
q_1	q_1	q_2	q_1	1
q_2	q_5	q_3	q_2	0
q_3	q_4	q_4	q_3	0
q_4	q_3	q_1	q_4	0
q_5	q_2	q_1	q_5	0

Then

$$P_1 = \{(q_1), (q_2, q_3, q_4, q_5)\}; \qquad P_2 = \{(q_1), (q_2, q_3), (q_4, q_5)\};$$
$$P_3 = \{(q_1), (q_2), (q_3), (q_4, q_5)\}; \qquad P_4 = \{(q_1), (q_2), (q_3), (q_4), (q_5)\}.$$

BIBLIOGRAPHY

1. WILLIAM CADDEN, Sequential Circuit Theory. Ph.D. Thesis, Princeton, 1956.
2. GEORGE MEALY, A Method for Synthesizing Sequential Circuits. *The Bell Tech. J.*, Sept. 1955, pp. 1045–1079.
3. EDWARD MOORE, *Gedanken—Experiments on Sequential Machines.* Annals of Mathematics Studies, No. 34, Automata Studies, pp. 129–153.