

LINEAR AUTOMATON TRANSFORMATIONS

A. NERODE¹

Let R be a nonempty set, let N consist of all non-negative rational integers, and denote by R^N the set of all functions on N to R . If R is a ring, a map $M: R^N \rightarrow R^N$ is *linear* if $M(r_1f_1 + r_2f_2) = r_1(Mf_1) + r_2(Mf_2)$ for r_1, r_2 in R, f_1, f_2 in R^N . For a finite commutative ring with unit we determine which linear transformations $M: R^N \rightarrow R^N$ can be realized by finite automata.

More precisely, let A, B be finite nonempty sets. A map $M: A^N \rightarrow B^N$ is an *automaton transformation* if there exists a finite set Q , maps $M_Q: A \times Q \rightarrow Q, M_B: A \times Q \rightarrow B$, elements \bar{b} in B, \bar{q} in Q such that corresponding to each f in A^N there exists an h in Q^N satisfying

$$(1) \quad \begin{aligned} h(0) &= \bar{q}, & h(n+1) &= M_Q(f(n), h(n)), & (Mf)(0) &= \bar{b}, \\ & & (Mf)(n+1) &= M_B(f(n), h(n)). \end{aligned}$$

(In automaton language, A is the input alphabet, B is the output alphabet, Q is the set of states, \bar{q} is the initial state, \bar{b} is the initial output, while $M_B(a, q)$ and $M_Q(a, q)$ are respectively the output and state resulting from input a and state q . For the case that A and B coincide with the set consisting of 0 and 1, the concept of automaton transformation is simply a variant of the concept of representable event of Kleene [1].)

Call a matrix $u_{ij}: N \times N \rightarrow R$ *eventually doubly-periodic* if for some positive integers P_1, P_2, p_1, p_2 :

$$(2) \quad u_{ij} = u_{(i+p_1)j} \quad \text{for all } i > P_1 \text{ and all } j,$$

$$(3) \quad u_{ij} = u_{i(j+p_2)} \quad \text{for all } j > P_2 \text{ and all } i.$$

THEOREM 1. *Let R be a finite commutative ring with unit. Then $M: R^N \rightarrow R^N$ is a linear automaton transformation if and only if there exists a matrix $u_{ij}: N \times N \rightarrow R$ such that:*

- (i) for all $j, u_{0j} = 0$;
- (ii) for f in R^N and $n \geq 0, (Mf)(n) = u_{n0}f(0) + u_{(n-1)1}f(1) + \dots + u_{0n}f(n)$;
- (iii) u_{ij} is eventually doubly-periodic.

Define $\tau: R^N \rightarrow R^N$ by $(\tau f)(0) = 0, (\tau f)(n) = f(n-1), n \geq 1$. A map $M: R^N \rightarrow R^N$ is *translation invariant* if for f in $R^N, M\tau f = \tau Mf$. Call a sequence u_0, u_1, \dots *eventually periodic* if there exist positive integers

Presented to the Society January 30, 1958; received by the editors December 26, 1957.

¹ National Science Foundation postdoctoral fellow.

P, p such that $u_{n+p} = u_n$ for $n \geq P$, then p is a *period*.

COROLLARY. *Let R be a finite commutative ring with unit. Then $M: R^N \rightarrow R^N$ is a linear translation invariant automaton transformation if and only if there exists an eventually periodic sequence $u_0 = 0, u_1, u_2, \dots$ of elements of R such that for f in $R^N, (Mf)(n) = u_0f(n) + \dots + u_n f(0)$.*

Consider a linear difference equation

$$(4) \quad \begin{aligned} S_1(n-1)F(n-1) + \dots + S_k(n-k)F(n-k) \\ = G(n) + T_1(n-1)G(n-1) + \dots + T_k(n-k)G(n-k), \end{aligned}$$

where $S_1, \dots, S_k, T_1, \dots, T_k, F, G$ are functions on the set of rational integers (positive and negative) to R which vanish for negative arguments. For fixed $S_1, \dots, S_k, T_1, \dots, T_k$, (4) induces a linear map $M: R^N \rightarrow R^N$ given by the requirement that whenever F, G jointly satisfy (4), and f is a member of R^N such that $f(n) = F(n)$ for $n \geq 0$, then $(Mf)(n) = G(n)$ for $n \geq 0$.

THEOREM 2. *Let R be a finite commutative ring with unit. Then $M: R^N \rightarrow R^N$ is a linear automaton transformation if and only if induced by a linear difference equation (4) with $S_1, \dots, S_k, T_1, \dots, T_k$ eventually periodic for $n \geq 0$.*

COROLLARY. *Let R be a finite commutative ring with unit. Then $M: R^N \rightarrow R^N$ is a translation invariant linear automaton transformation if and only if induced by a linear difference equation (4) with $S_1, \dots, S_k, T_1, \dots, T_k$ constant for $n \geq 0$.*

We will need three lemmas to prove Theorems 1 and 2.

LEMMA 1. *Let R be a finite commutative ring with unit. Endow R with the discrete, R^N with the product topology. Then $L: R^N \rightarrow R$ is linear and continuous if and only if there exists a finite sequence W_0, \dots, W_m of elements of R such that for f in $R^N, Lf = W_0f(0) + \dots + W_mf(m)$.*

PROOF. It is an easy consequence of the compactness of R^N and the continuity of L that there exists an m such that $Lf_1 = Lf_2$ whenever f_1, f_2 are in R^N and agree for $n \leq m$. If we put $\delta_k(n) = 1$ or 0 as $n = k$ or not, then we may take $W_k = L\delta_k$ for $k \leq m$.

Call $M: A^N \rightarrow B^N$ *causal* if: for f_1, f_2 in $A^N, (Mf_1)(0) = (Mf_2)(0)$; for f_1, f_2 in A^N and $k > 0$, if $f_1(n) = f_2(n)$ for $n < k$, then $(Mf_1)(k) = (Mf_2)(k)$. Denote by $\sigma(A)$ the set of finite sequences (x_0, \dots, x_j) consisting of elements from a finite set A . Call two such sequences

$(x_0, \dots, x_j), (y_0, \dots, y_k)$ *state-equivalent* (relative to M) if for any f in A^N , $(Mf_1)(n+j+1) = (Mf_2)(n+k+1)$ for all $n \geq 0$, where f_1, f_2 are chosen satisfying: $f_1(n) = x_n$ for $0 \leq n < j$, $f_1(n) = f(n-j)$ for $n \geq j$, $f_2(n) = y_n$ for $0 \leq n < k$, $f_2(n) = f(n-k)$ for $n \geq k$. (Note that the state-equivalence of two sequences does not depend on the last member of either.) Define an *intrinsic state* for M to be an equivalence class under state-equivalence.

LEMMA 2. *Let A, B be finite nonempty sets. Then $M: A^N \rightarrow B^N$ is an automaton transformation if and only if M is causal and M possesses only a finite number of intrinsic states. Further, the least number of states required in order to induce M as in (1) is the number of intrinsic states.*

PROOF. Suppose that M is an automaton transformation. Then M is certainly causal due to (1). We show that M possesses no more intrinsic states than the number of elements of Q . If $X = (x_0, \dots, x_j)$ is in $\sigma(A)$, define q_X to be the $h(j)$ determined from (1) by letting $f(n) = x_n$ for all $n < j$. Then X, Y in $\sigma(A)$ are state-equivalent whenever $q_X = q_Y$.

Conversely, if M is causal and possesses only a finite set Q of intrinsic states, define $\bar{b}, \bar{q}, M_B, M_Q$ as follows.

- (i) Let $\bar{b} = (Mf)(0)$ for any f in A^N .
- (ii) Let \bar{q} be the intrinsic state of any finite sequence of length 1.
- (iii) Let $M_Q(a, q_1) = q_2$ if for some X in q_1, Y in q_2 , we have $X = (x_0, \dots, x_j), Y = (y_0, \dots, y_{j+1}), x_n = y_n$ for all $n < j, y_j = a$. Let $M_B(a, q_1) = (Mf)(j+1)$ if f is a member of A^N such that $f(n) = y_n$ for $n \leq j$.

LEMMA 3. *If $S_1, \dots, S_k, T_1, \dots, T_k$ are eventually periodic for $n \geq 0$, then (4) induces a linear automaton transformation.*

PROOF. We wish to apply Lemma 2; it suffices to show that M has only a finite number of intrinsic states, since any M induced by Equation (4) is causal. Let p_i, p'_i be periods for $S_i, T_i, i = 1, \dots, k$. Then for n_1 sufficiently large, the intrinsic state of a finite sequence (x_0, \dots, x_{n+1}) is determined for $n \geq n_1$ by $F(n-1), \dots, F(n-k), G(n-1), \dots, G(n-k), n \bmod p_1, \dots, n \bmod p_k, n \bmod p'_1, \dots, n \bmod p'_k$. Thus for $n \geq n_1$, finite sequences fall into at most $z^{2k} p_1 \dots p_k p'_1 \dots p'_k$ distinct intrinsic states, where z is the number of elements of R . Thus M has altogether only a finite number of intrinsic states.

We now prove Theorems 1 and 2. If M is a linear automaton transformation, then for each $n \geq 0$, the map $L_n: R^N \rightarrow R$ given by

$L_n f = (Mf)(n)$ is linear and continuous. Thus Lemma 1 applies and there exists a matrix $W_{nk}: N \times N \rightarrow R$ such that for each $n \geq 0$ we can find an $m \geq 0$ satisfying $(Mf)(n) = W_{n0}f(0) + \dots + W_{nm}f(m)$, for all f in R^N . Causality implies $W_{nk} = 0$ for $k \geq n$. Setting $u_{ij} = W_{(i+j)j}$ we need only verify (2) and (3) to satisfy Theorem 1.

(5) Suppose that $M: A^N \rightarrow B^N$ is an automaton transformation, and that f is a member of A^N such that $f(0), f(1), f(2), \dots$ is eventually periodic. Then $(Mf)(0), (Mf)(1), (Mf)(2), \dots$ is eventually periodic. Moreover, if q_n is the intrinsic state of $(f(0), \dots, f(n))$, then q_0, q_1, q_2, \dots is eventually periodic.

We employ (5) to prove (2) and (3). Since the k th column of u_{ij} consists of the entries $0, (M\delta_k)(k+1), (M\delta_k)(k+2), (M\delta_k)(k+3), \dots$ it follows that this column is completely determined by the intrinsic state of a k -term sequence consisting of $k-1$ zero entries followed by a one. Since this sequence has the same intrinsic state as a k -term sequence consisting of zeros, (5) applies to show that this intrinsic state is an eventually periodic function of k , and hence proves (3).

With this done, (2) is easy since it now suffices to show that the k th column is itself eventually periodic. But (5) applied to $M\delta_k$ yields this.

Conversely, suppose that M is defined by a matrix u_{ij} satisfying (i), (ii), (iii) of Theorem 1. Define functions U_i by $U_i(j) = u_{ij}$ for $j \geq 0, U_i(j) = 0$ for $j < 0$. Then the following linear difference equation induces M when recast in form (4). (In the notation of (2), put $k = p_1 + P_1$.)

$$\begin{aligned} &U_1(n-1)F(n-1) + \dots + U_k(n-k)F(n-k) \\ &\quad - U_1(n-p_1-1)F(n-p_1-1) - \dots - U_{P_1}(n-k)F(n-k) \\ &= G(n) - G(n-p_1). \end{aligned}$$

By (3), U_1, \dots, U_k are eventually periodic for $n \geq 0$; hence by Lemma 3, M is an automaton transformation. This proves both Theorem 1 and Theorem 2.

REFERENCE

1. S. C. Kleene, *Representation of events in nerve nets and finite automata*, Automata Studies, Princeton University Press, 1956, pp. 3-41.

INSTITUTE FOR ADVANCED STUDY