INFORMATION AND CONTROL 24, 144–154 (1974)

Degrees of Finite-State Transformability

GERHARD RAYNA

Department of Mathematics, Lehigh University, Bethlehem, Pennsylvania 18015

The upper semilattice of degrees of transformability by finite-state automata is defined analogously to the upper semilattice of degrees of recursive unsolvability (which arises from transformability by Turing machines). Two infinite sequences from a finite alphabet are considered equivalent if each can be transformed into the other by a finite-state automaton, perhaps after finite initial segments (not necessarily of the same length) are deleted from each. We require the output sequence to be generated at the same rate as the input, with exactly one output character for each input character. If such a transformation is possible in only one direction, an order relation holds between the equivalence classes.

We show that this partially ordered set does indeed form an upper semilattice, exhibit the (unique) minimal class, and prove there is no maximal class. In the course of the proof of the last assertion, the notion of a complete sequence, a sequence in which every block of the alphabet occurs, is introduced and shown to be significant. The richness of the partial ordering is shown by two contrasting examples: We exhibit one section of it in which the partial ordering is dense, and, on the other hand, we exhibit two classes $[x] > [z]$ having no class properly between them.

1. INTRODUCTION

In this note, we study a certain equivalence relation between infinite sequences and a partial ordering of the equivalence classes. Roughly speaking, two sequences are equivalent if each can be transformed into the other, possibly with a fixed delay, by some pair of finite-state automata. If the transformation is possible in one direction but not the other, an order relation holds between their equivalence classes.

A finite-state automaton "reads" an input character at each "pulse" (a discrete succession of times) and immediately "writes" an output character. The input sequence is scanned steadily from left to right, and the output sequence is produced at the same rate.

A transformation carried out by a finite-state automaton is, of course, a computable process. Hence, our partial ordering is compatible with, but much finer than, the partial ordering of degrees of recursive unsolvability [see Kleene and Post (1954)].

In our proofs, we shall use the Mealy formulation of finite state automata: The output character may depend on the input character as well as the state. [See, for example, Mealy (1955) or Gill (1962, p. 7).] If sequence X can be transformed into Y by a Mealy machine, then X can be transformed into Y with unit delay by a Moore machine [see Moore (1956) or Gill (1962, p. 12)] and conversely (with no delay required in the second case). Since our equivalence relation allows finite fixed delays, the results are as applicable to Moore machines.

2. Definitions

In the Mealy formulation, a finite-state automaton is a set T:

Finite input alphabet Σ
Finite output alphabet Σ'
Finite state set S
State mapping function $f: \Sigma \times S \to S$
Output function $\chi: \Sigma \times S \to \Sigma'$.

If an initial state s_0 is distinguished, we call (T, s_0) an initialized automaton. We will frequently use single letters like T to designate initialized automata (instead of automata without distinguished initial state), if it is not necessary to refer explicitly to the name of the distinguished state.

We say that (T, s_0) transforms a sequence $\{x_n\}$, $n = 0, 1, 2,...$ of symbols from Σ into $\{z_n\}$, a sequence of symbols from Σ', if

$$s_{n+1} = f(x_n, s_n) \qquad \text{and} \qquad z_n = \chi(x_n, s_n) \qquad \text{for} \quad n = 0, 1, 2,... .$$

We also say that the input $\{x_n\}$ drives (T, s_0) through the state sequence $\{s_{n+1}\}$. Note that we omit the initial state s_0.

The unqualified word "sequence" will always mean an infinite sequence, indexed 0, 1, 2,... . A finite sequence will be called a block. If its length is N, we may speak of an N-block. We say that an input N-block $\{x_0, x_1, x_2,..., x_{N-1}\}$ drives (T, s_0) through the state block $\{s_1, s_2,..., s_N\}$ and yields the output block $\{z_0, z_1, z_2,..., z_{N-1}\}$. We call s_N the final state.

We shall consider two sequences $\{x_n\}$, $\{y_n\}$ of symbols, each from some finite alphabet (not necessarily the same), to be equivalent iff there exist initialized finite-state automata (T, s_0) and (T', s_0') and nonnegative integers d and d' such that (T, s_0) transforms $\{x_n\}$ into $\{y_{n-d}\}$, i.e., $\{y_n\}$ delayed by fixed delay d, and (T', s_0') transforms $\{y_n\}$ into $\{x_{n-d'}\}$. (For the purpose of this definition, we may choose the negative-subscript terms of the sequences arbitrarily.) This is easily seen to be an equivalence relation. (Recall that the composition of two initialized finite-state automata can be defined, and is another such automaton.) The class of a sequence x is designated $[x]$.

In particular, a sequence is equivalent to itself delayed (with arbitrary initial characters inserted), advanced (with any finite number of characters deleted at the beginning), or with any finite number of characters changed. Each equivalence class is called a "degree of finite-state transformability".

If some initialized automaton transforms the sequence x into the sequence y or into a sequence equivalent to y, we write $[x] \geqslant [y]$. This is easily seen to define a partial ordering on the equivalence classes. Observe that $[x] \geqslant [y]$ iff some initialized automaton transforms x into y delayed by some finite nonnegative delay d.

In the sections which follow, we show some properties of this partial ordering.

3. Elementary Properties

THEOREM. *The class* $[0]$ (*to which the constant sequence* 0 *belongs*) *consists precisely of the ultimately periodic sequences. Furthermore, for any sequence* x, $[x] \geqslant [0]$.

Proof. Let x be ultimately periodic. Then it is a simple matter to describe an initialized automaton which ignores its input and always produces x as its output. Hence $[0] \geqslant [x]$.

Conversely, it is an even simpler matter to describe an initialized automaton which, with arbitrary input, outputs the constant sequence 0; hence $[x] \geqslant [0]$.

This proves that for x ultimately periodic, $[0] = [x]$. The second paragraph also proves the second assertion of the theorem.

THEOREM. *The partially ordered system of degrees of finite-state transformability is an upper semilattice.*

Proof. We must show that for any two classes $[x]$, $[y]$, there exists a class $[z]$ such that $[x] \leqslant [z]$, $[y] \leqslant [z]$, and that if $[w]$ is such that $[x] \leqslant [w]$ and $[y] \leqslant [w]$, then $[z] \leqslant [w]$.

Let $\{x_i\} \in [x]$, $\{y_i\} \in [y]$ be sequences over the alphabets Σ_1, Σ_2, respectively. Then consider the sequence $\{z_i = (x_i, y_i)\}$ over the product alphabet $\Sigma_1 \times \Sigma_2$. Then automata (with just one state) can be defined transforming $\{z_i\}$ into $\{x_i\}$ and into $\{y_i\}$.

On the other hand, assume $\{w_i\}$ is such that automata T_1, T_2 exist transforming it into $\{x_i\}$ delayed and $\{y_i\}$ delayed, respectively. We may assume the delays are equal. Then an automaton can be described, combining T_1 and T_2 "in parallel", which transforms $\{w_i\}$ into $\{(x_i, y_i)\}$ delayed by the same amount.

Note that the same construction appears in Kleene and Post (1954, p. 382), where the sequence $\delta(a)$ is defined as the sequence of ordered pairs $(\alpha(a), \beta(a))$ coded, α, β being sequences of integers, as $2^{\alpha(a)}3^{\beta(b)}$.

4. The Classes Below a Certain Class

We shall analyse completely the structure of the partial ordering of the set of classes below ($\leqslant$) the class of the particular sequence

$$x = \{101001000100001...\}.$$

Consider the 1's in this sequence enumerated, so we may speak of the i-th 1, for $i = 1, 2, 3, \ldots$. The i-th 1 is followed by exactly i 0's.

For the purpose of this analysis, call a sequence $y = \{y_n\}$ *special* if it is like x except that the 1's whose ordinals are in some periodic pattern are replaced by 0's. x itself is also considered special. Formally, y is special if there are integers s and

$$0 \leqslant a_1 < a_2 < \cdots < a_s < M$$

defining y as follows:

$$y_n = 1 \quad \text{if} \quad x_n = 1 \quad \text{and} \quad x_n \text{ is the } i\text{-th } 1,$$
$$\text{where } i \equiv a_1 \text{ or } a_2 \text{ or } \ldots \text{ or } a_s \bmod M;$$

$$y_n = 0 \quad \text{otherwise.}$$

Theorem. *If y is a special sequence and x is the sequence just described, then $[y] \leqslant [x]$. Conversely, every class below $[x]$ contains a special sequence and, indeed, exactly one.*

Proof. Given a special sequence y defined by s, $a_1, \ldots, a_s$, M, it is easy to define an automaton which counts input 1's mod M and outputs just those 1's whose ordinals are in the desired congruence classes. Hence $[y] \leqslant [x]$.

Conversely, let (T, s_0) be an initialized automaton transforming x into y. Starting T in any state s_i, the ultimately periodic input sequence $\{100000 \ldots\}$ drives T through some ultimately periodic state sequence S_i. Let N_i be the length of the initial nonperiodic part of this state sequence and M_i the period of the periodic part. Choose N greater than all N_i and M a common multiple of all M_i. Hence, both the state sequence S_i and the corresponding output sequence Y_i can be considered to consist of an initial part of length $N - 1$ and a periodic part of period M, these parameters independent of the initial state s_i.

The input sequence x consists of a succession of $(k + 1)$-blocks B_k consisting of 1 followed by k 0's, $k = 1, 2, \ldots$. Let T_k be the corresponding block of the state sequence of (T, s_0) and C_k the corresponding block of the output sequence y. T_k and C_k must be initial segments of one of the above described ultimately periodic sequences of states and output characters, respectively, viz., of S_i and Y_i, if s_i is the state of the automaton at the beginning of B_k.

In the argument that follows, assume throughout that $k > N + M$, so that at least one complete copy of the periodic part is included in all the initial segments of ultimately periodic sequences which are mentioned.

If T_k and $T_{k'}$ are both initial segments of the state sequence S_i and $k \equiv k'$ mod M, then the final states in T_k and $T_{k'}$ are the same, because of periodicity. Hence, the successor blocks T_{k+1} and $T_{k'+1}$ are also initial segments of some common S_j and, of course, $(k + 1) \equiv (k' + 1)$ mod M also. It follows that the sequence of S_i's of which the T_k's are initial segments is ultimately periodic, and so then is the sequence of Y_i's of which the output blocks C_k are initial segments.

Now, suppose that C_k is an initial segment of Y_i and that the block consisting of C_k followed by C_{k+1} also happens to be an initial segment of Y_i. In this case, we call C_{k+1} *invisible*; in the opposite case, *visible*. Whether C_{k+1} is visible or not evidently depends only on Y_i and k mod M, since k mod M determines the point in the periodic pattern of Y_i where the end of C_k comes and also determines the sequence Y_j of which C_{k+1}, by itself, is initial segment. Since the sequence of Y_i's is ultimately periodic, so is the succession of pairs $(Y_i, k \bmod M)$. Hence, the visibility/invisibility pattern is ultimately periodic.

If there are no visible blocks, the output sequence is ultimately periodic, and nothing remains to be proved. If there are visible blocks, for every such C_{k+1} let d_k be the index of the first place in C_{k+1} at which C_k followed by C_{k+1} differs from the corresponding place in Y_i. Clearly, d_k depends only on the pair $(Y_i, k \bmod M)$, so only finitely many different values occur. Let d be their maximum. Then a finite automaton can easily be specified which transforms the output sequence into the special sequence, delayed by d, in which only the 1's corresponding to the beginnings of visible blocks are retained, and another which transforms that special sequence into the original output sequence with no delay. Hence, the output sequence is equivalent to a special sequence.

One assertion of the theorem remains to be verified: that no class contains two different special sequences. If y and z are special and different, then in y there are infinitely many occurrences of 1's which are missing in z (or vice versa). The number of consecutive 0's on either side of these 1's is unbounded. Now, no automaton with, say, k states can transform an input block of $k + 1$ 0's into a block consisting of k 0's followed by a 1; hence, no automaton can fill in all the missing 1's of z to yield y. An obvious refinement of this argument shows that y with constant delay can not be obtained either. Hence, $[z] \not\geqslant [y]$, so $[z] \neq [y]$.

5. There Are No Maximal Classes

In this section, we shall prove that there exist no sequences y such that $[z] > [y]$ is never the case. This result could be deduced from the known fact that there are no maximal elements in the partial ordering of degrees of recursive enumerability, but the methods in the following elementary proof are of interest in their own right.

Throughout this discussion, we are assuming a fixed alphabet. The question of finding z with $[z] > [y]$ is trivial, if z is allowed to be written in a larger alphabet than y. For example, if $y = \{y_i\}$, z could be $\{(y_i, w_i)\}$, where $w = \{w_i\}$ is one of the uncountably many sequences of 0's and 1's for which $[y] \not> [w]$. (Here z is written in an alphabet double the size of that of y.)

Obviously, we must assume the alphabet has at least two characters.

We call a sequence of symbols from some alphabet *complete* if for every k, every block of length k of symbols from that alphabet occurs in the sequence. It follows that every block occurs infinitely often: for example, the (nk)-block consisting of n consecutive copies of the k-block will occur for arbitrary n.

Lemma. *If y is not complete, its class is not maximal.*

Proof. Let B be a k-block which does not appear in y and A one which occurs infinitely often among the "integrally positioned k-blocks", i.e., among

k-blocks starting at index 0, k, $2k$, $3k$,..., in y. We shall construct a sequence z such that $[z] \geqslant [y]$ but $[z] \nleqslant [y]$.

The sequence z will be formed from y by replacing some of the integrally positioned copies of A by copies of B. An automaton can be described easily which will map z into y delayed by k: Starting at the completion of every integrally positioned input k-block, it must reproduce it unchanged if it was not B and output A if it was B. This shows that $[z] \geqslant [y]$.

We now describe how to choose the copies of A to replace so as to assure $[z] \nleqslant [y]$. We must construct z in such fashion that for no initialized automaton T (over the specified alphabet) and no delay d does the input y yield z delayed by d as output. But the pairs (T, d) are enumerable. Construct z stepwise as follows. For $i = 1, 2, 3,...$, let (T, d) be the i-th such pair. Replace the i-th integrally positioned copy of A by B if necessary to assure that the part of z through the end of that k-block, delayed by d, differs from the initial segment of the output of initialized automaton T with input sequence y.

Having proved that the only possible maximal classes are those containing only complete sequences, we shall now show that no such class is maximal. Therefore, there are no maximal classes.

THEOREM. *If a class contains a complete sequence, it is not maximal.*

Proof. (It is convenient to assume, in this proof, that the alphabet contains at least three symbols, say 0, 1, 2. If it does not, replace the word "symbol" in the proof by "pair of successive symbols". Since the alphabet has at least 2 symbols, it has at least 4, so at least 3, pairs of symbols.)

It is easy to define an initialized automaton to replace the character following each maximal string of consecutive copies of the symbol 2 by the one which followed the previous such string; at the first substitution, an arbitrarily chosen character, say 0, is inserted. [If the alphabet has K symbols, we can define an initialized automaton with $2K$ states corresponding to pairs (B, X), where B is "TRUE" or "FALSE" as the last character was 2 or not 2, respectively, and X is the character which followed the previous maximal string of 2's. The initial state is (FALSE, 0).] For any input sequence x, let Dx be the output of this initialized automaton. We shall show that if x is complete, then $[Dx] \ngeqslant [x]$, so $[Dx] < [x]$, since, of course, $[Dx] \leqslant [x]$.

Let y be any sequence. Then $[y] = [Dx]$ for a suitably chosen sequence x, viz. the sequence formed by replacing the character following each maximal string of consecutive 2's by the one which follows the next such string (or, say, 0 if there are no more 2's in the sequence). (Indeed, the sequences y and Dx agree except possibly for the character following the first such string.)

Furthermore, if y is complete, so, of course, is x. Hence, given any complete sequence y, we have

$$[y] = [Dx] \leqslant [x].$$

We shall now demonstrate that $[Dx] < [x]$. We need the following lemma, whose proof will appear below.

LEMMA. *For any finite-state automaton T and any nonnegative integer d, there exists a block B of length greater than d such that no matter in which of its states T is started, the output of T with DB as input differs from B delayed by d in at least one of the places where they are both defined.*

Using this lemma, it is easy to complete the proof. Let x be any complete sequence. Then it contains each such block B infinitely often. Let (T, s_0) be any initialized automaton and d any delay. Then the output of (T, s_0) given Dx as input differs from x delayed by d infinitely often, in particular, at least once in each occurrence of the block B corresponding, by the lemma, to that pair (T, d). Hence $[Dx] \not\geqslant [x]$.

Proof of the Lemma

We are given an automaton T and a positive integer d. Enumerate the states of T: $s_1, \ldots, s_n$.

Construct B as a succession of n copies of the blocks $X = 2000 \ldots 0$ and $Y = 2100 \ldots 0$, each of length $d + 2$, selected as follows:

Whether B begins with X or Y, DB will begin with X (because of the arbitrary choice of 0 that was made for what to substitute for the character after the first string of 2's). Assume T is started in state s_1, and consider the last output character for input block X. If it is 0, let B begin with a copy of Y; otherwise, a copy of X.

To complete the inductive construction, assume B has been defined through its first k subblocks X and Y by consideration of starting states s_1 through s_k. DB begins with X followed by the part of B already defined; so the first $(k + 1)(d + 2)$ characters are known. Now, assume T is started in state s_{k+1} with this known block of DB as input. Consider the last character of the output. If it is 0, let the next part of B be Y; otherwise, X. Continue in this way until the last state, s_n, is used [and so B has length $n(d + 2)$].

If B is so defined and T started in state s_k with DB as input, the output differs from B delayed by d in at least the $k(d + 2)$ position.

6. An Atomic Step

In Section 4, a subset of the partially ordered set was analysed, and it was observed that between any two classes $[x] > [z]$ in it, there was a class properly between: $[x] > [y] > [z]$. In this section, we show that this is not true in general. Specifically, we construct a sequence x (over the alphabet 0, 1) whose class is only atomically greater than the lowest class, i.e., $[x] > [0]$ but there exists no $[y]$ such that $[x] > [y] > [0]$.

We shall define a sequence of blocks I_i, with each I_i a proper initial segment of the next, I_{i+1}. The sequence x will be defined as the limit of this sequence of blocks, i.e., as the unique sequence of which each I_i is an initial segment.

The blocks I_i, together with certain blocks A_i, B_i, will be defined by induction for $i = 1, 2, 3, \ldots$. The details of the construction will be given at the end of this section. At this point, we state certain properties these blocks will have.

(1) A_1 begins with 0, B_1 with 1; they are the same length.

(2) A_{i+1} and B_{i+1} are each successions of copies of the smaller blocks A_i and/or B_i, beginning with A_i and B_i, respectively. They are of equal length, greater than i. Neither is a part of any periodic sequence of period i.

(3) I_{i+1} consists of I_i followed by copies of A_i and/or B_i.

(4) Enumerate the initialized automata (T, s_0). If the i-th machine (T, s_0) is brought to final state s by the input I_i, then, if started in state s, T is returned to state s as final state by both input block A_i and input block B_i.

The sequence $x = \lim I_i$ is not ultimately periodic. For suppose it had an ultimate period i. The sequence x can be considered to be the initial block I_{i+1} followed by a succession of copies of A_{i+1} and B_{i+1}. But, by Property 2, neither A_{i+1} nor B_{i+1} is a part of any periodic sequence of period i.

Suppose we have a sequence y such that $[x] \geqslant [y]$. By replacing y by another member of the equivalence class $[y]$, if necessary, we may assume that some initialized automaton (T, s_0) yields y as output for x as input. Let i be the index of (T, s_0) in our enumeration of initialized automata.

The sequence x can be regarded as consisting of the initial block I_i followed by a succession of copies of A_i and B_i which are blocks of the same length, say N. Whenever one of these blocks A_i or B_i begins, T is in the same state, say s. Hence, the output sequence y consists, once the initial segment I_i of the input is passed, of a succession of copies of only two blocks, each of length N, say C (the response to input block A_i when T is started in state s) and D (similarly, the response to B_i).

We have, evidently, $[x] \geqslant [y] \geqslant [0]$. We are to show that one of the weak inequalities here is actually an equality.

There are two possibilities. If C and D are identical, then y is ultimately periodic, so $[y] = [0]$. If C and D are not identical, then an initialized automaton can be described to replace C by A_i and D by B_i, with a delay of no more than N, so x and y are equivalent: $[x] = [y]$.

This completes the proof, except for the actual construction of the blocks used in the argument. We define I_{i+1}, A_{i+1}, and B_{i+1} by induction. To start the induction, define I_0 to be the empty block (length zero), A_0 to be the character 0, and B_0 to be 1.

Let the $(i + 1)$st initialized automaton be (T, s_0). Let the symbol $(A \mid B)^*$ designate the set of finite nonempty blocks which are successions of copies of A_i and/or B_i. For any state s of T, let $\langle s \rangle$ be the set of states which are final states of (T, s) for input blocks in $(A \mid B)^*$.

Let s_1 be the final state of (T, s_0) for input block I_i. Choose a state $s_2 \in \langle s_1 \rangle$ such that the cardinality of $\langle s_2 \rangle$ is minimal. It follows that $s_2 \in \langle s_2 \rangle$. Let I_{i+1} be I_i followed by a block in $(A \mid B)^*$ such that the final state of (T, s_0), for input I_{i+1}, is s_2.

By induction, A_i and B_i are of length at least i. They differ in the first character. Hence, if A_iA_i is part of a periodic sequence of period i, A_iB_i is not. Let A_iX_i be one of these which is not. Let s_A be the final state of (T, s_2) for input block A_iX_i. Clearly, $s_A \in \langle s_2 \rangle \subseteq \langle s_1 \rangle$, so, by the minimality of $\langle s_2 \rangle$, $\langle s_A \rangle = \langle s_2 \rangle$. Therefore, $s_2 \in \langle s_A \rangle$. Let A'_{i+1} be an element of $(A \mid B)^*$, beginning with A_iX_i, for which the final state of (T, s_2) is s_2.

Since A_iX_i is not part of any periodic sequence of period i, neither is A'_{i+1}. Its length is greater than $2i$, so greater than $i + 1$.

Similarly, let B'_{i+1} be an element of $(A \mid B)^*$ beginning with B_i and not part of any periodic sequence of period i, for which the final state of (T, s_2) is s_2.

If A'_{i+1} has length k and B'_{i+1} length l, let $m = lcm(k, l)$, and let A_{i+1} be m/k copies of A'_{i+1} and B_{i+1} m/l copies of B'_{i+1}. Then A_{i+1}, B_{i+1} have the same structure and "s_2-to-s_2" property the original primed blocks had and, furthermore, are of the same length.

Since A_1 begins with 0 and B_1 with 1, so, by induction, do A_{i+1} and B_{i+1}.

This completes the verification of the required properties.

RECEIVED: August 2, 1972

References

Gill, A. (1962), "Introduction to the Theory of Finite-State Machines," McGraw–Hill, New York.

Kleene, S. C., and Post, E. L. (1954), The upper semi-lattice of degrees of recursive unsolvability, *Ann. of Math.* **59**, 379.

Mealy, G. H. (1955), A method for synthesizing sequential circuits, *Bell System Tech. J.* **34**, 1045.

Moore, E. F. (1956), Gedanken-experiments on sequential machines, "Automata Studies," *Ann. of Math. Studies* **34**, 129–153, Princeton.