# INFINITE SEQUENCES AND FINITE MACHINES

by

David E. Muller

## 1. Introduction

Regular sets of sequences have always been regarded as consisting of just finite sequences from some finite alphabet although the sets themselves may be infinite. This restriction to finite sequences is quite natural if one regards a regular set as the set of input sequences acceptable by some finite machine.

A regular set may also be thought of as a projection of a set of paths in a finite directed linear graph. This linear graph model is the one which is often most convenient to use. For example, if we regard a regular set as the set of possible state sequences of an inputless, nondeterministic machine, then the linear graph model is quite natural. Such machines are considered in the theory of speed-independent circuits. However, state sequences of inputless, nondeterministic machines may either terminate in some equilibrium condition, or else the machine may pass from one state to another without ever reaching equilibrium.

While it is true that an actual machine will eventually stop as a result of human intervention or as the result of some physical breakdown, these actions will remain outside the mathematical description of the machine. Furthermore, we shall see that it is impossible to describe adequately the behavior of some nonstopping machines in terms of finite state sequences alone. Therefore, we shall extend the notion of a regular set to include infinite as well as finite sequences.

## 2. The Linear Graph Model

Let $A$ be a finite alphabet $A(a, b, \ldots)$. We shall write $A^*$ to denote the set of all finite sequences of $A$, including the null sequence $e$. Similarly, we shall denote the set of all infinite sequences by $A^\infty$. The set of all sequences, both finite and infinite will be written $A^* + A^\infty$.

If $x$ is in $A^*$ and $y$ is in $A^* + A^\infty$, then $xy$ will represent the element of $A^* + A^\infty$ formed by concatenating $x$ followed by $y$. If $X \subseteq A^*$ and $Y \subseteq A^* + A^\infty$ then we shall write $XY$ to describe the set of all pairs $xy$ where $x$ is taken from $X$ and $y$ from $Y$. If $X \subseteq A^*$ does not contain the null sequence then we shall permit the expression $X^\infty \subseteq A^\infty$ to mean the set of all infinite sequences formed by concatenating the finite sequences comprising $X$.

A directed linear graph $G = \langle A, R \rangle$ consists of a finite set $A$ of nodes and a set $R \subset A \times A$ of branches. Given two nodes $a$ and $b$, we construct the set $P^*(G,a,b)$ of finite paths passing from the initial node $a$ to the terminal node $b$ by these rules:

A sequence $x = a_1 a_2 \ldots a_n$ in $A^*$ is a path in $P^*(G,a,b)$ iff

1. $(a,a_1)$ is in R

2. $(a_i,a_{i+1})$ is in R for each consecutive pair in the sequence

$3^*$. $a_n = b$

We shall use the convention that the null sequence will be included in the set of paths whenever a = b.

For sets of nodes $\alpha$ and $\beta$, we shall write $P^*(G,\alpha,\beta)$ to denote the union of all sets of paths from a to b, where a,b has the range $\alpha \times \beta$.

To extend the notion of a set of paths to include the infinite case we replace the final node b by a set $\tau$ of nodes. We construct the set $P^\infty(G,a,\tau)$ of infinite paths passing from the initial node a to the terminal set $\tau$ by the following rules:

A sequence $x = a_1 a_2 \ldots$ in $A^\infty$ is a path of $P^\infty(G,a,\tau)$ iff

1. $(a,a_1)$ is in R

2. $(a_i,a_{i+1})$ is in R for each consecutive pair in the sequence

$3^\infty$. the set $\sigma(x)$ of nodes which occur infinitely often in the sequence x is just $\tau$.

Again, we may write $P^\infty(G,\alpha,T)$ if $\alpha$ is a set of nodes and T is a set of subsets of A. The meaning, as before, is the union of all $P^\infty(G,a,\tau)$, where $a,\tau$ has the range $\alpha \times T$.

Contrasting the two definitions of sets of paths, we see that the rules $3^*$ and $3^\infty$ contain the only formal differences. It may strike the reader as strange that rule $3^\infty$ specifies that all paths must contain all nodes of $\tau$ infinitely often. It might seem sufficient to require only that the nodes occurring infinitely often comprise a subset of $\tau$. Both practical and mathematical difficulties would result from such a weaker condition.

Consider the example of the asynchronous, inputless circuit formed from two NOT elements connected as shown below.

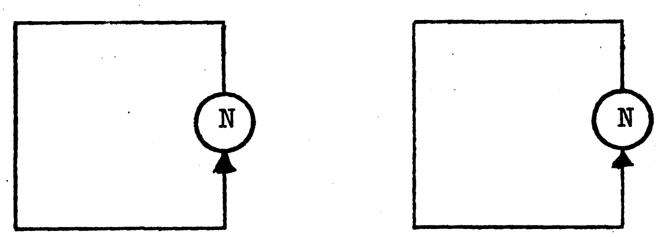


Figure 1

In this circuit A is the set of four possible states (0,0), (0,1), (1,0) and (1,1) describing the conditions of the two NOT elements. Each NOT oscillates between 0 and 1 but they are not synchronized. Thus, the set of state sequences, represented by the set of paths cannot have just the pair, for example, $\{(0,0), (0,1)\}$ as the terminal set $\tau$. Instead, the sets in T must contain states in which both NOT elements are at times both 0 and 1. Such a set is the set of all four states. Another is the pair $\{(0,0), (1,1)\}$.

3. Extended Regular Expressions

If A and A' are two finite sets and $h: A \to A'$ is a mapping from A onto A', then we may construct the corresponding projection h from $A^*$ to $(A')^*$. Similarly, we may form the projection h from $A^\infty$ to $(A')^\infty$. In both cases the rule to be followed is that the sequence $a_1 a_2 \ldots$ in $A^* + A^\infty$ shall map to $a_1', a_2', \ldots$ in $(A')^* + (A')^\infty$ where $a_1'$ is the image of $a_1$, $a_2'$ the image of $a_2$, etc., under h.

If $\Gamma \subseteq A^*$ is the projection h of a set of paths of the form $P^*(G,\alpha,\beta)$ for some graph G and sets $\alpha,\beta$, then we shall say that $\Gamma$ is finitely regular. This property is usually called regularity, but we wish to distinguish several forms of regularity. Thus, if $\Gamma \subseteq A^\infty$ is the projection h of a set of paths of the form $P^\infty(G,\alpha,T)$, then we shall say that $\Gamma$ is infinitely regular. A union of a finitely and infinitely regular sets we shall call a generally regular set.

It is a trivial result to graph theory that any generally regular set $\Gamma \subseteq A^* + A^\infty$ is a projection h of a set of paths of the form

$$P^*(G,\alpha,\beta) + P^\infty(G,\alpha,T) = P(G,\alpha,\beta + T) .$$

Theorem 1: Any generally regular set $\Gamma$ may be expressed using a finite number of symbols from $A + \{\cdot, \cup, *, \infty\}$, subject to the rules already described.

Proof: This theorem is well known for the case in which $\Gamma$ is finitely regular, and the operations then required are just concatenation union $\cup$, and repetition $*$. Thus, we need only prove the theorem for the case in which $\Gamma$ is infinitely regular. To do this we shall consider the case in which $\Gamma$ is a projection h of a set of paths $P^\infty(G,a,\tau)$ where a is a single node and $\tau$ a single set of nodes.

Let $\{p_1, p_2, \ldots, p_s\} = \tau$ and let $\Delta_1$ be the projection h of $P^*(G,a,p_1)$. This set is finitely regular and hence the known result mentioned earlier allows us to write $\Delta_1$ as a regular expression.

Let $P_1$ be that subset of $P^*(G,p_1,p_1)$ consisting of paths lying just in $\tau$ in which each node of $\tau$ appears at least once. This set $P_1$ is not a set of paths in the sense defined above, but is a projection of a set of paths and hence finitely regular by the following argument.

From the graph G and set $\tau$ we construct a new graph which we shall call $G' = \langle A', R' \rangle$ having a set of paths whose projection is $P_1$. Each node of $A' \subseteq 2^{\tau} \times \tau$ consists of a pair $(\tau_i, p_j)$, where $\tau_i$ is a subset of $\tau$. The sets $A'$ and $R'$ are constructed recursively by first placing $p_1' = \{\Phi, p_1\}$ in $A'$, where $\Phi$ is the empty set. We place the pair $(\tau_i, p_j)$, $(\tau_i \cup \{p_k\}, p_k)$ in $R'$ whenever $(p_j, p_k)$ is in R, $p_k$ is in $\tau$ and $(\tau_i, p_j)$ has previously been placed in $A'$. We then place $(\tau_i \cup \{p_k\}, p_k)$ in $A'$. This process is continued until no new branches can be added to $R'$.

A projection $h'$ is now formed which maps $(\tau_i, p_j)$ in $A'$ to $p_j$ for each $(\tau_i, p\ )$ in $A'$. We now show that $P_1$ is the projection $h'$ of $P^*(G', p_1', (\tau, p_1))$. To each path in $P_1$ starting at $p_1$ in $\tau$, there corresponds a path of $G'$ starting at $p_1'$ in $A'$. At each node $p_j$ in the path in $P_1$ there is a corresponding node $(\tau_i, p_j)$ of the path in $G'$, where the set $\tau_i$ represents the set of all nodes including $p_j$ which have been encountered in the path of $P_1$ up to $p_j$. Thus, the set of paths $P_1$ corresponds exactly to those paths of $G'$ which start at $p_1'$ node of $A'$ and terminate at $(\tau, p_1)$.

Thus, $P_1$ is a finitely regular set and so is its projection $\Pi_1$ under h since it, too, may be regarded as a projection of $P^*(G', p_1', (\tau, p_1))$. By the known result mentioned earlier in this proof we may write $\Pi_1$ as a finite regular expression.

We conclude the proof by showing that $\Gamma = \Delta_1 \Pi_1^{\infty}$. Since $\Pi_1$ does not contain the null sequence e, we may meaningfully write the expression $\Pi_1^{\infty}$.

Remembering that the sequences of $\Gamma$ are the projections h of paths of $P^{\infty}(G, a, \tau)$, we consider a path x of $P^{\infty}(G, a, \tau)$ and show that its projection is in $\Delta_1 \Pi_1^{\infty}$. We may write x in the form $yp_1z$, where the infinite path z contains only nodes of $\tau$ and all of those occur infinitely many times. We note that $yp_1$ is in $P^*(G, a, p_1)$ and hence its projection is in $\Delta_1$. Next, we consider z. A finite initial segment $z_1$ of z can be found which contains all nodes of $\tau$ since all these nodes occur in z. In fact, we may continue the segment so that it terminates in $p_1$. The same remark may be made about the remaining portion of z. Hence, z may be written $z_1, z_2, \ldots$, where each $z_i$ contains all nodes of $\tau$ and terminates in $p_1$. Each $z_i$ is in $P_1$ and its projection is in $\Pi_1$. The sequence z is in $P_1^{\infty}$, and its projection is in $\Pi_1^{\infty}$. Hence, the projection of x is in $\Delta_1 \Pi_1^{\infty}$.

Conversely, we need only to note that all sequences in $\Delta_1 \Pi_1^{\infty}$ are clearly in $\Gamma$.

The most general $\Gamma$ which is infinitely regular is a union of projections of sets of paths and hence would be expressed as a union of expressions of the form $\Delta_1 \Pi_1^{\infty}$. However, the number of these terms would always be finite and the proof is complete.

## 4. The Sequential Machine Model

In the case of finitely regular sets, one may show that such sets are exactly characterized as sets acceptable by some finite sequential machine or automaton. Although there seems to be no physical meaning which we can attach to the notion of acceptance or rejections of an infinite tape or input sequence, it is desirable to extend the mathematical concepts to include the infinite case.

In the finite case, we define a machine $\alpha_1$ by the 5-tuple $\langle Q,A,f,d,P\rangle$, where Q is the finite set of states, A the finite input alphabet, f is the transition function mapping Q x A to Q, d is the initial state which is a particular element of Q, and P is a subset of Q called the class of terminal states.

Such a machine $\alpha_1$ defines a set $\Gamma(\alpha_1) \subseteq A^*$ of acceptable input sequences by the following rules. Extend f to a mapping from Q x $A^*$ to Q by the recursive definition: For all q in Q, x in $A^*$, and a in A, we let $f(q,e) = q$, $f(q,xa) = f(f(q,x),a)$. Then we place x in $\Gamma(\alpha_1)$ iff $f(d,x)$ is in P.

As explained less precisely earlier, a set $\Gamma$ is finitely regular if and only if it can be represented in the form $\Gamma(\alpha_1)$ for some machine $\alpha_1$.

In the infinite case we define a machine $\alpha_2$ as a 5-tuple $\langle Q,A,f,d,\psi\rangle$, where the first four quantities have the same meaning as in the finite case. However, $\psi$, which replaces P, is a set of nonempty subsets of Q and is called the class of terminal sets.

We now define recursively the state sequence $F(q,x)$ which follows any state q and is included by any input sequence x. For all q in Q let $F(q,e) = e$, and for all x in $A^*$ and a in A, let $F(q,xa) = F(q,x)f(q,xa)$. Hence, if x and y are both in $A^*$, we may write $F(q,xy) = F(q,x)F(f(q,x),y)$. Induction may be used to extend the definition so that if x is in $A^\infty$, the function $F(q,x)$ will correspondingly lie in $Q^\infty$.

We define the set $\Gamma(\alpha_2) \subseteq A^\infty$ of input sequences acceptable to $\alpha_2$ to contain all sequences x such that the set $\sigma(F(q,x))$ of states occurring infinitely often in $F(q,x)$ is one of the sets in $\psi$.

As in the finite case we shall prove that $\Gamma$ is infinitely regular if and only if there is a machine $\alpha_2$ such that $\Gamma = \Gamma(\alpha_2)$. This theorem has a rather length proof, and until this result has been established we shall refer to sets of the form $\Gamma(\alpha_2)$ as machine regular.

It is possible to combine the two types of machines described above to give a general type $\alpha_3$ defined as a 5-tuple $\langle Q,A,f,d,P + \psi\rangle$, where $P + \psi$ is a set which is a formal sum of two sets of the types described above. We then define $\Gamma(\alpha_3)$ as the union $\Gamma(\alpha_1) \cup \Gamma(\alpha_2)$, where

$\alpha_1 = \langle Q,A,f,d,P\rangle$ and $\alpha_2 = \langle Q,A,f,d,\psi\rangle$. The type of machine regularity formed in this way will be seen to be equivalent to general regularity.

Theorem 2: A projection h of a set $\Gamma(\alpha_2) \subseteq A^\infty$ is acceptable to some machine $\alpha_2'$.

Proof: The relatively simple proof of this theorem in the finite case cannot be extended to the infinite case. We take a projection h gives as a mapping from A onto A' and a machine $\alpha_2 = \langle Q,A,f,d,\psi\rangle$ and we seek to construct another machine $\alpha_2' = \langle Q',A',f',d',\psi'\rangle$ such that h: $\Gamma(\alpha_2) \rightarrow \Gamma(\alpha_2')$.

First, form $2^Q - \psi = \{S_1,S_2,\ldots,S_r\}$, the class of subsets of Q not in $\psi$. For each such set $S_i = \{q_1,q_2,\ldots,q_{t_i}\}$ we set up an arbitrary ordering of its elements as indicated. This order is regarded as cyclic with $q_1$ following $q_{t_i}$.

Second, for each set $S_i$, we use a special procedure to construct a set $\Sigma_i$ consisting of pairs $(D_i^j,L_i^j)$, where each $D_i^j$ is a subset of Q x $S_i$ and $L_i^j$ is a mapping from $D_i^j$ into $\{1,2,\ldots,t_i+1\}$.

The set Q' of states of $\alpha_2'$ will be taken as $Q' = \Sigma_1 \times \Sigma_2 \times \ldots \times \Sigma$ For each $\Sigma_i$, we shall further define a function $f_i'$: $\Sigma_i \times A' \rightarrow \Sigma_i$ and shall take $f' = f_1' \times f_2' \times \ldots \times f_r'$.

Given a particular pair $(D_i^j,L_i^j)$ of $\Sigma_i$ and a' of A', let $(D_i^k,L_i^k)$ be their image under $f_i'$. To form $(D_i^k,L_i^k)$, we begin by contracting $L_i^j$ to form $(L_i^j)_c$. Arrange the images of the elements of $D_i^j$ under $L_i^j$ in increasing order. Renumber these images consecutively starting with 1. These renumbered images produce the mapping $(L_i^j)_c$ in which the "gaps" have been removed from $L_i^j$.

Pick an element $(p_u,q_v)$ of $D_i^j$. If there is an element a in A such that h(a) = a' and such that $f(p_u,a) = p_w$ for some $p_w$ in Q, then if $p_w \neq q_v$, place $(p_w,q_v)$ in $D_i^k$. Make the image of $(p_w,q_v)$ under $L_i^k$ the same as the image of $(p_u,q_v)$ under $(L_i^j)_c$, unless there is another pair $(p_t,q_v)$ in $D_i^j$ having a smaller image under $(L_i^j)_c$ which also gives rise to $(p_w,q_v)$ in $D_i^k$.

If, after constructing $f(p_u,a) = p_w$, it is found that $p_w = q_v$, then place $(p_w,q_{v+1})$ in $D_i^k$ and make the image of $(p_w,q_{v+1})$ under $L_i^k$ just one greater then the greatest image produced by $(L_i^j)_c$.

The above process is carried out for all a in A such that h(a) = a' and all pairs $(p_u,q_v)$ of $D_i^j$, thus forming $D_i^k$ and $L_i^k$.

The initial state d' is an r-tuple whose ith component $(D_i^j,L_i^j)$ is chosen so that $D_i^j$ is a set consisting of the single pair $(d,q_1)$ The mapping $L_i^j$ which maps $D_i^j$ into $\{1, 2, \ldots, t_i + 1\}$ is chosen so that the image of the single element of $D_i^j$ is 1. This choice is made for i = 1, 2, ..., r, but in the pair $(d,q_1)$ the first component is the same for all i, while the second component varies with i and in each case is the number one state of $S_i$.

We finally choose $\psi'$ to consist of all sets $P'$ which are subsets of $Q'$, such that for $i = 1, 2, \ldots, r$ at least one of the two following properties holds.

(1) Either there is some state in $P'$ whose i'th coordinate $(D_i^j, L_i^j)$ is such that $D_i^j$ is not a subset of $S_i \times S_i$,

or

(2) For all states in $P'$ the i'th coordinate $(D_i^j, L_i^j)$ is such that some pair $(p_u, q_v)$ in $D_i^j$ maps to 1 under $L_i^j$.

Thus, we have completed the specification of $\alpha_2'$. By our construction the machine is finite.

We now prove that $h\colon\ \Gamma(\alpha_2) \to \Gamma(\alpha_2')$ by showing that if $x'$ is any sequence not in $\Gamma(\alpha_2')$ then all $x$ such that $h(x) = x'$ must lie outside $\Gamma(\alpha_2)$. Pick $x' = a_1'a_2'\ldots$ to be in $(A')^\infty$ but not in $\Gamma(\alpha_2')$. Let $F'(d',x') = q_1'q_2'\ldots$ be the corresponding sequence of states. For any $q_j'$ we may consider the i'th coordinate $(D_i^j, L_i^j)$. The set $D_i^j$ contains pairs $(p_u, q_v)$ such that $p_u$ in $Q$ is the j'th state reached by $\alpha_2$ if some sequence $a_1a_2\ldots a_j$ is applied whose image under h is $a_1'a_2'\ldots a_j'$. In other words $f(d, a_1a_2\ldots a_j) = p_u$ for some such sequence by the construction of $f'$.

We now scrutinize the definition of $\psi'$. Since $x'$ is not in $\Gamma(\alpha_2')$ it must have a terminal set $\sigma(F'(d',x'))$ of states which occur infinitely often in the corresponding state sequence, and $\sigma(F'(d',x'))$ is not in $\psi'$. Hence, there is some i in the range $1, 2, \ldots, r$ which is characteristic of $\sigma(F'(d',x'))$ such that:

(1') In the i'th coordinate $(D_i^j, L_i^j)$ of each state in $\sigma(F'(d',x'))$, the set $D_i^j$ is a subset of $S_i \times S_i$,

and

(2') There is some state in $\sigma(F'(d',x'))$ whose i'th coordinate $(D_i^j, L_i^j)$ has the property that no pair $(p_u, q_v)$ in $D_i^j$ maps to 1 under $L_i^j$.

The first of these properties ensures that all sequences x such that $h(x) = x'$ have the property that $\sigma(F(d,x))$ is a subset of $S_i$. This follows from the fact that to each sequence $F(d,x)$ there corresponds a sequence of pairs $(p_u, q_v)$ lying in $D_i^j$ of the corresponding state in $F'(d',x')$ such that $p_u$ is the state of $F(d,x)$.

The second property ensures that there is no state $q_v$ of $S_i$ such that $q_v$ fails to appear in $\sigma(F(d,x))$ for some x such that $h(x) \doteq x'$. For were this the case, then either $q_v$ or else some other state $q_w$ of $S_i$ also not in $\sigma(F(d,x))$ would appear in pairs $(p_u, q_v)$ of $D_i^j$ which map to 1 under $L_i^j$. In the sequence $F'(d',x')$, such pairs would continue to map to 1 because we could never have $p_u$ become equal to $q_v$. Thus, we see that $\sigma(F(d,x)) = S_i$.

Conversely, any set of sequences x of $A^\infty$ which all have the property that $\sigma(F(d,x)) = S_i$ must map into sequences x' of $(A')^\infty$ such that $\sigma(F'(d',x'))$ has the properties (1') and (2') listed above. The property (1') is clear from the previous argument. To see that (2') also holds, we note that in each sequence F(d,x) each state $q_v$ of $S_i$ occurs infinitely often. Hence, if at any point in the sequence $F'(d',x')$ in the i'th component $(D_i^j, L_i^j)$ the pair $(p_u, q_v)$ maps to 1 under $L_i^j$, then at some later point j' in the sequence F(d,x) we will obtain another state $p_u = q_v$ and the pair $(p_u, q_{v+1})$ will be formed with a mapping under $L_i^{j'}$ to some number greater than 1. Therefore, the set of pairs in $D_i^j$ mapping to 1 will eventually become empty and will do so an infinite number of times in the sequence F'(d',x'). Thus, (2') holds.

We have shown that each sequence x in $A^\infty$ which is not in $\Gamma(\alpha_2)$ maps to some x' which is not in $\Gamma(\alpha_2')$, and each x' in $(A')^\infty$ which is not in $\Gamma(\alpha_2')$ has only inverse images x of $A^\infty$ which are not in $\Gamma(\alpha_2)$. Thus, h: $\Gamma(\alpha_2) \to \Gamma(\alpha_2')$ and the theorem is proved.

A slightly more powerful result can be proved if we define a nondeterministic machine $\alpha_{N2} = \langle Q, A, f_N, d, \psi \rangle$ as a 5-tuple in which the function $f_N$ is a mapping from Q x A to the set $2^Q$ of subsets of Q. For any $x = a_1 a_2 \ldots$ in $A^\infty$ we shall say that $y = q_1 q_2 \ldots$ in $Q^\infty$ is in $F_N(d,x)$ if $q_1$ is in $f_N(d, a_1)$, and for each i = 2, 3, ..., $q_i$ is in $f_N(q_{i-1}, a_i)$. We take the set $\Gamma(\alpha_{N2})$ of sequences acceptable to $\alpha_{N2}$ to consist of all x in $A^\infty$ such that there is a sequence y in $F_N(d,x)$ for which $\sigma(y)$ is in $\psi$.

We now show that corresponding to $\alpha_{N2}$, as given above, there is an ordinary (deterministic) machine $\alpha_2$ such that $\Gamma(\alpha_2) = \Gamma(\alpha_{N2})$.

Begin by constructing $\alpha_2' = \langle Q', A', f', d', \psi' \rangle$, in which $Q' = Q \cup \{m\}$, where m is special state called the dead state, which is not in Q, A' = Q x A. We construct f' so that $q_i, a_j'$ maps to $q_k$ whenever $a_j' = (q_k, a_j)$ for some $a_j$ in A and provided $q_k$ is in $f_N(q_i, a_j)$. If $a_j'$ is of the form $(q_k, a_j)$ and $q_k$ is not in $f_N(q_i, a_j)$, then map $q_i, a_j'$ to m. Further, let $d' = d$ and $\psi' = \psi$.

We see that A' may be projected to A, and that the corresponding machine $\alpha_2$ obtained by theorem 2 is such that $\Gamma(\alpha_2) = \Gamma(\alpha_{N2})$.

The same result may be obtained by a slight variation in the proof of theorem 2.

## 5. Equivalence of the Three Models

Theorem 3: Every set $\Gamma \subseteq A^* + A^\infty$ of sequences is acceptable to some machine $\alpha_3 = \langle Q, A, f, d, P + \psi \rangle$ if and only if it is a projection h of some set of paths $P(G, \alpha, \beta + T)$ of a linear graph G.

Proof: Let us first assume that $\Gamma$ is acceptable to the machine $\alpha_3 = \langle Q, A, f, d, P + \psi \rangle$. Form G as a graph whose set of nodes is Q x A and whose set R of branches is constructed by the rule that $(q_1, a_1)$, $(q_2, a_2)$

is a pair of nodes in R iff $f(q_1,a_2) = q_2$. Let $\alpha$ be the set of all nodes $(d,a)$ where a is any element of A. Let $\beta$ consist of all pairs $(q,a)$ such that q is in P and let T consist of all sets $\mu$ of nodes such that the projection of $\mu$ to Q is an element of $\psi$. Any path of $P(G,\alpha,\beta + T)$ is a sequence of the form $(F(d,x),x)$ where $F(d,x)$ terminates in P if x is finite and $\sigma(F(d,x))$ is in $\psi$ if x is infinite. The projection of $P(G,\alpha,\beta + T)$ to the right-hand member of the pair, namely x, is just $\Gamma$.

Secondly, let us assume that $\Gamma$ is a projection h of some set $P(G,\alpha,\beta + T)$ of paths. We wish to construct a machine to accept $\Gamma$. However, in this construction we shall make use of theorem 2 and first construct a machine to accept a set $\Gamma'$ such that $\Gamma$ is a projection of $\Gamma'$. In fact, this projection will turn out to be just the projection h.

If G is a linear graph with set N of nodes and R of branches, then form $\alpha_3' = \langle Q',A',f',d',P' + \psi'\rangle$ where $Q' = d' + k + N$, in which d' and k are called initial and dead states respectively and are separate from the nodes N. Take A' to be equal to N and construct f' so that for all inputs n in N we let

1. $f'(q,n) = n$ if $(q,n)$ is in R.
2. $f'(q,n) = k$ if q is in N but $(q,n)$ is not in R.
3. $f'(d,n) = n$ if n is in $\alpha$.
4. $f'(d,n) = k$ if n is not in $\alpha$.
5. $f'(k,n) = k$

If we further take $P' = \beta$ and $\psi' = T$, then the set of input sequences acceptable to $\alpha_3$ is just $P(G,\alpha,\beta + T)$ and the theorem is proved.

Theorem 4: Let $\Gamma$ and $\Delta$ be two sets of sequences acceptable to two general-type machines. Then the following combinations are each acceptable to some machine.

(a) $\Gamma \subseteq A^\infty$, $\Delta \subseteq B^\infty$ then $\Gamma \times \Delta \subseteq (A \times B)^\infty$.

(b) $\Gamma \subseteq A^*$, $\Delta \subseteq A^* + A^\infty$ then $\Gamma\Delta \subseteq A^* + A^\infty$.

(c) $\Gamma \subseteq A^* - \{e\}$, then $\Gamma^\infty \subseteq A^\infty$.

(d) $\Gamma \subseteq A^* + A^\infty$, $\Delta \subseteq A^* + A^\infty$, then all Boolean combinations of these sets.

Proof: To prove all these sets acceptable we form the sets of paths $P(G^\Gamma,\alpha^\Gamma,\beta^\Gamma + T^\Gamma)$ and $P(G^\Delta,\alpha^\Delta,\beta^\Delta + T^\Delta)$ whose projections $h^\Gamma$ and $h^\Delta$ yield $\Gamma$ and $\Delta$ respectively, as described in theorem 3.

To prove (a) we form $\Gamma \times \Delta$ as the projection $h^\Gamma \times h^\Delta$ of $P(G^\Gamma \times G^\Delta, \alpha^\Gamma \times \alpha^\Delta, \beta^\Gamma \times \beta^\Delta + T^{\Gamma \times \Delta})$, where $T^{\Gamma \times \Delta}$ is defined as consisting of all sets of the form $\mu \times \nu$ where $\mu$ is in $T^\Gamma$ and $\nu$ is in $T^\Delta$.

To prove (b) we note that since $\Gamma \subseteq A^*$ there is no set $T^\Gamma$. We form a new graph $G^{\Gamma\Delta}$ from $G^\Gamma = \langle N^\Gamma, R^\Gamma \rangle$ and $G^\Delta = \langle N^\Delta, R^\Delta \rangle$. Let $N^{\Gamma\Delta} = N^\Gamma + N^\Delta$ and let $R^{\Gamma\Delta} = R^\Gamma + R^\Delta$ + all pairs $(n_1, n_2)$ such that $n_1$ is in $\beta^\Gamma$ and $n_2$ is such that there is a pair $(n, n_2)$ of $R^\Delta$ for which $n$ is in $\alpha^\Delta$. Then project $N^\Gamma$ using $h^\Gamma$ and $N^\Delta$ using $h^\Delta$ to obtain $\Gamma\Delta$ from $P(G^{\Gamma\Delta}, \alpha^\Gamma, \beta^\Delta + T^\Delta)$.

To prove (c) assume that $G^1 = \langle N^1, R^1 \rangle$ and $G^2 = \langle N^2, R^2 \rangle$ are two isomorphic graphs and that both $P^*(G^1, \alpha^1, \beta^1)$ and $P^*(G^2, \alpha^2, \beta^2)$ project to $\Gamma$ under $h^1$ and $h^2$ respectively. We note that $\alpha^1$ and $\beta^1$ are disjoint as are $\alpha^2$ and $\beta^2$.

Form $G^3 = \langle N^1 + N^2, R^3 \rangle$, where $R^3 = R^1 + R^2$ + all pairs $(n_2, n_1)$ such that $n_2$ is in $\beta_2$ and $n_1$ is such that $(n, n_1)$ is in $R'$ for some $n$ in $\alpha_1$, and also all pairs $(n_1', n_2')$ such that $n_1'$ is in $\beta_1$ and $n_2'$ is such that $(n, n_2')$ is in $R^2$ for some $n$ in $\alpha_2$.

Let T consist of all subsets of $N^1 + N^2$ which contain at least one node from each of the two sets $N^1$ and $N^2$.

To prove (d) consider first the form of complementation defined by $A^* + A^\infty - \Gamma$. If $\Gamma$ is acceptable to $\alpha_3 = \langle Q, A, f, d, P + \psi \rangle$, then replace $P + \psi$ by $(Q - P) + 2^Q - \psi - \Phi)$, where $\Phi$ is the empty set.

If $\Gamma \subset A^\infty$ and we have no set P and we may form $A^\infty - \Gamma$ if we replace $\psi$ by $(2^Q - \psi - \Phi)$.

Consider next the operation $\Gamma \cup \Delta$ and return to the linear graph model. Construct $G^{\Gamma \cup \Delta} = \langle N^\Gamma + N^\Delta, R^\Gamma + R^\Delta \rangle$. Form the projection $h = h^\Gamma + h^\Delta$ of $P(G^{\Gamma \cup \Delta}, \alpha^\Gamma + \alpha^\Delta, \beta^\Gamma + \beta^\Delta + \psi^\Gamma + \psi^\Delta)$.

Other Boolean operations, such as intersection may be shown to preserve regularity by the results given here. Each may be formed as a combination of the operations of union and complementation.

Theorem 4 allows us to state that all "regular expressions," i.e., those involving the operations $\cdot$, $\cup$, $*$, and $\infty$, when used in the manner described, and possibly also involving projection and x are regular in the sense that they represent projections of sets of paths of linear graphs and are acceptable to finite state machines.

Since single symbols of the alphabet A are regular, and regular expressions are just combinations of these symbols, all finite regular expressions represent projections of sets of paths are acceptable to some machine.

Taking the three theorems (1, 3, 4) together we see that the three definitions for a regular set are equivalent in the infinite as well as in the finite case. Summarizing, we have shown the following equivalences with the help of known results.

Theorem 1: Linear graph regularity $\rightarrow$ Reg. expression regularity.

Theorem 3: Sequential machine regularity $\rightarrow$ Linear graph regularity.

Theorem 4: Reg. expression regularity $\rightarrow$ Seq. machine regularity.

We have also shown that some other operations besides those usually included in regular expressions preserve regularity.

## 6. Regularity Preserving Transformations

Let $\Gamma \subseteq A^* + A^\infty$ be any set of sequences in A and x be some sequence in $A^*$. Then we introduce the notation $x \backslash \Gamma$ to represent the subset of $A^* + A^\infty$ consisting of all sequences y such that xy is in $\Gamma$. Similarly, if z is some sequence in $A^* + A^\infty$, then we shall write $\Gamma/z$ to represent the subset of $A^*$ consisting of all sequences y such that yz is in $\Gamma$. We shall also write $\Delta \backslash \Gamma$ to represent the union of all $x \backslash \Gamma$ as x ranges over $\Delta$, and $\Gamma/\Delta$ to represent the union of all $\Gamma/z$ as z ranges over $\Delta$.

Theorem 5: If $\Gamma \subseteq A^* + A^\infty$ is generally regular, then

(a) the number of distinct sets $x \backslash \Gamma$ is finite as x ranges over $A^*$, and

(b) each such set $x \backslash \Gamma$ is generally regular, and

(c) the number of distinct sets $\Gamma/z$ is finite as z ranges over $A^* + A^\infty$, and

(d) each such set $\Gamma/z$ is finitely regular, and

(e) $\Delta \backslash \Gamma$ and $\Gamma/\Delta$ are regular even if $\Delta$ is not regular.

Proof: If $\Gamma \subseteq A^* + A^\infty$ is generally regular there is some machine $\alpha_3 = \langle Q, A, f, d, P + \psi \rangle$ which accepts $\Gamma$. We see that the machine $\alpha_{x3} = \langle Q, A, f, f(d,x), P + \psi \rangle$ accepts $x \backslash \Gamma$, so $x \backslash \Gamma$ is also generally regular. The number of distinct states of the form f(d,x) is finite and hence the number of distinct $x \backslash \Gamma$ is also finite.

Let P' be the set of states q of Q such that z is acceptable to $\langle Q, A, f, q, P + \psi \rangle$. We may then form $\alpha_{1z} = \langle Q, A, f, d, P' \rangle$ which accepts $\Gamma/z$. If P' is empty take $\Gamma/z$ as also empty. Thus $\Gamma/z$ is regular and since the number of possible subsets P' of Q is finite, the number of $\Gamma/z$ is also finite.

Finally, we note that $\Delta \backslash \Gamma$ and $\Gamma/\Delta$ are finite unions of regular sets and hence are always regular. Also, we see that the number of regular sets of either type for fixed $\Gamma$ is not greater than $2^n$, where n is the order of Q.

Remark: If $\Gamma \subseteq A^*$ and the number of distinct sets $x\backslash\Gamma$ is finite as x ranges over $A^*$, then $\Gamma$ is finitely regular. However, if $\Gamma \subseteq A^\infty$ and the number of distinct sets $x\backslash\Gamma$ is finite as x ranges over $A^*$, then $\Gamma$ may or may not be infinitely regular.

Proof: The proof of the first part of the remark is well known and involves constructing a machine $\alpha_1$ which accepts $\Gamma$. We introduce one state into Q for each distinct set $x\backslash\Gamma$ as x ranges over $A^*$. We form f as the function mapping $(x\backslash\Gamma)$, a to $a\backslash(x\backslash\Gamma) = xa\backslash\Gamma$. We let $d = e\backslash\Gamma = \Gamma$, and we form P as consisting of all states $x\backslash\Gamma$ containing e. The proof is quite direct and the result is often stated in terms of right invariant equivalence classes. The machine $\alpha_1$ constructed in this fashion is called the reduced, or distinguished machine which accepts $\Gamma$ and can be shown to have a minimum number of states among all such machines that do.

It is interesting to note that in the infinite case there is no such powerful theorem. It also seems that there is no unique minimum state machine which accepts an infinitely regular $\Gamma$.

We exhibit the following example of $\Gamma \subseteq A^\infty$ which is not regular and for which the number of distinct sets $x\backslash\Gamma$ is finite as x ranges over $A^*$

Let $A = \{0,1\}$. Let Z represent the set of all infinite sequences in $A^\infty$ obtained by removing any finite segment from the left-hand end of the binary expansion of some irrational number. For example, we may take Z as consisting of all sequences of the form $0^i\ 1\ 0^j\ 1\ 0^{j+1}\ 1\ 0^{j+2}\ \ldots$, where i and j range over integers such that $0 \leq i < j$.

Form $\Gamma = A^*Z$. In other words, $\Gamma$ consists of all sequences which begin with an arbitrary finite sequence of 0's and 1's and terminate in a sequence from Z. For all x in $A^*$ we see that $x\backslash\Gamma = A^*Z$. Hence, there is just one set of the form $x\backslash\Gamma$ as x ranges over $A^*$.

From theorem 1 we note that if $\Gamma$ is regular it can be expressed as a union of terms of the form $\Gamma_1(\Gamma_2)^\infty$, where $\Gamma_1$ and $\Gamma_2$ are finitely regular. Thus, if x and y are any elements of $\Gamma_1$ and $\Gamma_2$ respectively then $xy^\infty$ is in $\Gamma$, but by construction $xy^\infty$ cannot be in $A^*Z$.

An example of the failure of uniqueness of the minimum state machine is now given. Let $A = \{0,1\}$, and let $\Gamma = A^*0^\infty$. Then construct $\alpha_2 = \langle Q, A, f, d, \psi \rangle$, where $Q = \{q_1, q_2\}$, $d = q_1$, $\psi = \{\{q_2\}\}$. We may choose f as either defined by the table

| $q \backslash a$ | 0 | 1 |
|---|---|---|
| $q_1$ | $q_2$ | $q_1$ |
| $q_2$ | $q_2$ | $q_1$ |

or the table

| $q \backslash a$ | 0 | 1 |
|---|---|---|
| $q_1$ | $q_2$ | $q_2$ |
| $q_2$ | $q_2$ | $q_1$ |

We are also free to choose either of the two states as the initial state.

Theorem 6: General regularity is preserved under the input to output transformations produced by Ginsburg's generalized sequential machines.

Proof: Ginsburg's generalized sequential machine M is a 5-tuple $(Q,A,B,f,g)$ where Q, A, and B are finite sets of states, inputs, and outputs respectively. The transition function f is a mapping from $Q \times A$ to Q as in the machines described previously. The output function g maps $Q \times A$ to $B^*$, rather than B as in the more conventional case.

If we pick a particular state d of Q as our initial state, then consider the effect of an input sequence x in $A^*$. We may define $F(d,x)$ as before and also define the extended output function $L(d,x)$ recursively by the rules $L(d,e) = e$, $L(d,xa) = L(d,x)g(f(d,x),a)$ for x in $A^*$ and a in A. Thus, for fixed d, the function $L(d,x)$ produces a mapping $L_d$ from $A^*$ to $B^*$. By induction, we extend this to a mapping from $A^\infty$ to $B^* + B^\infty$. There is some difficulty in that the output function g may map to e, the null sequence. However, we have not defined what is meant by $e^\infty$, an infinite repetition of the null sequence. To complete the induction we introduce the special rule that we take $e^\infty = e$. If $\Delta \subseteq A^* - \{e\}$, then we define specially $(\Delta \cup \{e\})^\infty = \Delta^\infty \cup \Delta^*$ for any set $\Delta$.

Let $\Gamma_A \subseteq A^\infty$ be infinitely regular, and let $\Gamma_B \subseteq B^* + B^\infty$ be its image under $L_d$. We wish to prove that $\Gamma_B$ is generally regular. A similar result is already known when $\Gamma_A$ is finitely regular and hence we do not need expressly to take $\Gamma_A$ as generally regular.

Let $G = \langle N,R\rangle$ be a linear graph such that $\Gamma_A$ is a projection h of a set of paths $P^\infty(G,\alpha,T)$. Form a new linear graph $G' = \langle N',R'\rangle$, where $N' = N \times Q$. Form R' by placing $(n_1,q_1)$, $(n_2,q_2)$ in R' whenever $(n_1,n_2)$ is in R and $f(q_1,h(n_1)) = q_2$. Further construct $\alpha'$ so as to consist of all pairs $(n,d)$ such that n is in $\alpha$, and T' of all sets which project to an element of T.

We see that the set $P^\infty(G',\alpha',T')$ is regular and may be projected to the set of input-state sequences x, $F(d,x)$ obtained as x ranges over $\Gamma_A$. We can therefore represent $P^\infty(G',\alpha',T')$ as a finite regular expression involving elements of N', and the operations of theorem 1.

In the regular expression for $P^{\infty}(G',\alpha',T')$ each element $(n,q)$ of $N'$ may be replaced by its finite output sequence $g(q,h(n))$. This is again a regular expression when the convention $e^{\infty} = e$ is used and represents $\Gamma_B$ which is thus regular.

BIBLIOGRAPHY

1. Elgot, C., "Decision Problems of Finite Automata Design and Related Arithmetics," Transactions of the American Mathematical Society, 98, 21-51, (1961)

2. Ginsburg, S., "Some Remarks on Abstract Machines," Transactions of the American Mathematical Society, 96, 400-444, (1960)

3. Ginsburg, S., An Introduction to Mathematical Machine Theory, Addison Wesley Publish Company, (1962)

4. Kleene, S. C., "Representation of Events in Nerve Nets and Finite Automata," Automata Studies, pp. 3-41, Princeton University Press, (1956)

5. Lee, C. Y., "Automata and Finite Automata," Bell System Technical Journal, 39, 1267-1295, (1960)

6. McNaughton, R. F., and Yamada, H., "Regular Expressions and State Graphs for Automata," I.R.E. Transactions on Electronic Computers, EC-9, 39-47, (1960).

7. Muller, D. E., and Bartky, M. S., "A Theory of Asynchronous Circuits," Proceedings of an International Symposium on the Theory of Switching, pp. 204-243, Harvard University Press, (1959)

8. Rabin, M. O., and Scott, D., "Finite Automata and Their Decision Problems," I.B.M. Journal of Research and Development, 3, 114-125, (1959)