INFORMATION AND CONTROL 11, 452–470 (1967)

Automata in General Algebras*

SAMUEL EILENBERG

Columbia University, New York I.B.M. Watson Research Center, Yorktown Heights, New York

JESSE B. WRIGHT

I.B.M. Watson Research Center, Yorktown Heights, New York

1. INTRODUCTION

The study of automata and of context-free languages usually deals with monoids or semigroups. The purpose of this paper is to propose an extension of the domain of study to other algebraic systems. In the process, a better understanding of what takes place in monoids can be achieved.

The basic language is that of the theory of categories. The basic facts are reviewed in sections 2 and 3. The basic ideas of universal algebra are then introduced. The key notions are that of a "theory" and of algebras belonging to a theory. These ideas were laid down by Lawvere. A streamlined version of Lawvere's theory is given in sections 4–10. This part of the paper is regarded as expository and no proofs are given.

Recognizable sets and (deterministic) automata are discussed briefly in sections 11 and 12. In order to consider the analogs of nondeterministic automata, a restriction must be imposed upon the "theories" considered. The normal habitat for this notion is the so-called "linear theories". However, since the main result (Theorem III) is valid only for "free theories" (which are linear), we accept this restriction starting with section 13 and do not introduce linear theories at all. A full treatment of the subject is scheduled to appear in a book by the first of the authors.

Relational algebras (section 13) and relational automata (section 14) supply then the analog of nondeterministic automata, while the notion of a polynomial (section 15) and of an algebraic set (section 16) are the generalizations of that of a grammar and of a context-free language. The main result asserts that, for free theories, recognizable sets and algebraic ones coincide. This result is due to Mezei and Wright. The proofs are given in sections 17 and 18.

It is clear from this introduction that this paper contains nothing that is essentially new, except perhaps for a point of view.

2. CATEGORIES

A category $\mathcal{A}$ consists of

(2.1) a class of elements called *objects* of $\mathcal{A}$ and denoted by A, A_1, A_2, A', etc.;

(2.2) a set $\mathcal{A}(A_1, A_2)$ defined for any pair A_1, A_2 of objects of $\mathcal{A}$. The elements $f \in \mathcal{A}(A_1, A_2)$ are called *morphisms* and are written as $f: A_1 \to A_2$ or

$$A_1 \xrightarrow{f} A_2 ;$$

(2.3) a composition law which to morphisms

$$A_1 \xrightarrow{f} A_2 \xrightarrow{g} A_3$$

assigns a morphism

$$A_1 \xrightarrow{gf} A_3 .$$

The following axioms are postulated.

(2.4) Associativity: Given morphisms

$$A_1 \xrightarrow{f} A_2 \xrightarrow{g} A_3 \xrightarrow{h} A_4 ,$$

we have $h(gf) = (hg)f$.

(2.5) Identity: For every object A there exists a morphism $1_A : A \to A$ such that in

$$A_1 \xrightarrow{f} A \xrightarrow{1_A} A \xrightarrow{g} A_2$$

we have

$$1_Af = f \quad \text{and} \quad g1_A = g.$$

The uniqueness of 1_A follows from (2.4) and (2.5).

A morphism $f: A_1 \to A_2$ is called an isomorphism if there exists a morphism $g: A_2 \to A_1$ such that $gf = 1_{A_1}$, $fg = 1_{A_2}$. The uniqueness of g follows from (2.4) and (2.5) and we write $g = f^{-1}$.

Let $\mathcal{A}$ and $\mathcal{B}$ be categories. A functor $F: \mathcal{A} \to \mathcal{B}$ consists of

(2.6) a function which to each object A of $\mathcal{A}$ assigns an object FA of $\mathcal{B}$;

(2.7) a function which to each morphism $f: A_1 \to A_2$ in $\mathcal{A}$ assigns a morphism $Ff: FA_1 \to FA_2$ in $\mathcal{B}$.

The following axioms are postulated.

(2.8) $F(gf) = (Fg)(Ff)$.

(2.9) $F1_A = 1_{FA}$.

Given functors $F: \mathcal{A} \to \mathcal{B}$ and $G: \mathcal{B} \to \mathcal{C}$, the composite functor $GF: \mathcal{A} \to \mathcal{C}$ is defined in the obvious way.

3. EXAMPLES OF CATEGORIES

The category S of sets has sets as objects and functions as morphisms with composition defined as composition of functions. Two objects in S will play special roles: the empty set $\emptyset$ and the set I consisting of the number 1 alone. For any set A there are unique morphisms

$$\emptyset \to A, \qquad A \to I.$$

We say that $\emptyset$ is an *initial* object and I is a *terminal* object for S.

A set A and an element $a \in A$ determine a unique morphism $I \to A$ with a as value. We shall denote this morphism by the same letter a. Thus morphisms $I \to A$ and elements of A will be identified.

For each integer $n = 0, 1, \cdots$ we denote by $[n]$ the set $\{1, \cdots, n\}$. Thus $[0] = \emptyset$ and $[1] = I$. The sets $[n]$, $n = 0, 1, \cdots$ together with all morphisms between them form a subcategory S_0 of S.

4. THEORIES

A *theory* T is a category such that

(4.1) the objects of T are $[n]$ for $n = 0, 1, \cdots$;

(4.2) S_0 is a subcategory of T; i.e., every morphism in S_0 is also a morphism in T, composition of morphisms in S_0 agrees with that in T, and the identity morphisms $1_{[n]}$ in S_0 are also identity morphisms in T;

(4.3) given morphisms

$$\phi_i: I \to [p] \quad \text{in} \quad T, \quad i = 1, \cdots, n,$$

there exists a unique morphism

$$\phi: [n] \to [p]$$

such that ϕ_i is the composition

$$I \xrightarrow{i} [n] \xrightarrow{\phi} [p]$$

for every $i \in [n]$.

We shall write $\langle \phi_1, \cdots, \phi_n \rangle$ for the morphism ϕ given in (4.3). Thus for any morphism ϕ: $[n] \to [p]$ in T, we have $\phi = \langle \phi 1, \cdots, \phi n \rangle$.

It should be noted that it follows from the above axioms that in T there is only one morphism $O_n : \emptyset \to [n]$ for every n, just as in the case of S_0. However, (contrary to what takes place in S_0) there may be in T morphisms ϕ: $I \to \emptyset$. In fact, these "0-ary operations" play a fundamental role in the sequel.

5. ALGEBRAS

Let T be a theory. A T-algebra A consists of a set A and a rule which to each ϕ: $[n] \to [p]$ in T and each p-tuple $(x_1, \cdots, x_p)$ of elements of A assigns an n-tuple

$$(x_1', \cdots, x_n') = (x_1, \cdots, x_p)\phi$$

of elements of A, subject to the following two axioms:

(5.1) if ϕ is in S_0, then $x_i' = x_{\phi i}$;

(5.2) if ψ: $[k] \to [n]$ in T, then

$$(x_1', \cdots, x_n')\psi = (x_1, \cdots, x_p)(\phi\psi).$$

If we write $x = (x_1, \cdots, x_p)$, then (5.2) may be rewritten as

$$(x\phi)\psi = x(\phi\psi). \tag{5.2$'$}$$

A morphism f: $A \to B$ of T-algebras is a mapping from A to B satisfying

$$f[(x_1, \cdots, x_p)\phi] = (fx_1, \cdots fx_p)\phi, \tag{5.3}$$

or in abbreviated form

$$f(x\phi) = (fx)\phi, \tag{5.3$'$}$$

where $fx = (fx_1, \cdots, fx_p)$.

With composition of morphisms of algebras defined in the ordinary fashion, there results the category $T^\flat$ of T-algebras.

We note that if ϕ: $I \to [p]$ in T, then

$$(x_1, \cdots, x_p)\phi \in A$$

so that ϕ yields a mapping $A^p \to A$ where A^p is the p-fold Cartesian product $A \times \cdots \times A$.

If in the above $p = 0$, i.e., ϕ: $I \to \emptyset$, then ( ) $\phi \in A$ is an element of A determined by ϕ, independent of any "inputs" $a_1, \cdots, a_p$. We denote this element by ϕ_A.

6. FREE ALGEBRAS

Let $A_k = T(I, [k])$ be the set of all morphisms $I \rightarrow [k]$. We convert A_k into a T-algebra as follows: Given $\phi\colon [n] \rightarrow [p]$ in T and given $x_1, \cdots, x_p \in A_k$, we have $x_i : I \rightarrow [k]$ and therefore $\langle x_1, \cdots, x_p\rangle\colon [p] \rightarrow [k]$ in T. Thus, the composition $\gamma = \langle x_1, \cdots, x_p\rangle\phi$ is defined and is a morphism $\gamma\colon [n] \rightarrow [k]$ in T. We define

$$(x_1, \cdots, x_p)\phi = (\gamma 1, \cdots, \gamma n).$$

The verification that A_k is a T-algebra is immediate.

We note that each mapping $i\colon I \rightarrow [k]$ $i = 1, \cdots, n$ in S_0 is an element of A_k and thus in a natural fashion, $[k]$ becomes a subset of A_k. The following fact is fundamental:

(6.1) If A is any T-algebra, then every mapping $f\colon [k] \rightarrow A$ admits a unique extension $\bar{f}\colon A_k \rightarrow A$ to a morphism of T-algebras.
Indeed, we must have for $\phi \in A_k$

$$\bar{f}\phi = (f1, \cdots, fk)\phi \in A.$$

The above shows that the algebras A_k are "free" with $[k]$ as base. In particular, A_0 is the free algebra with an empty base and (6.1) asserts that for any T-algebra A there is a unique morphism

$$\zeta_A : A_0 \rightarrow A.$$

In fact, $\zeta_A\phi = \phi_A$ for $\phi \in A_0$; i.e., for $\phi\colon I \rightarrow \emptyset$. Thus, A_0 is an "initial object" in the category T^b.

7. FREE THEORIES

As is usual in algebra, theories will be defined by "generators and relations" or as "quotient" theories of "free" theories. We start out with this second notion.

Let $\Omega = \{\Omega_n\}$ $n = 0, 1, \cdots$ be a sequence of sets. Consider a theory T such that

$$\Omega_n \subset T(I, [n]).$$

Assume further that with each morphism $\phi\colon [n] \rightarrow [p]$ in T there is associated an integer $d\phi \geqq 0$ (called the *degree* of ϕ) satisfying the following conditions:

(7.1) $d\phi = 0$ if ϕ is in S_0.

(7.2) $d\phi = d(\phi 1) + \cdots + d(\phi n)$.

(7.3) If $\omega \in \Omega_n$, then $d(\phi\omega) = 1 + d\phi$.

(7.4) If $\phi: I \to [p]$ and $d\phi > 0$, then there exists a unique $k \geqq 0$ and a unique factorization

$$I \xrightarrow{\omega} [k] \xrightarrow{\psi} [p]$$

of ϕ with $\omega \in \Omega_k$ and ψ in T.

It is not too difficult to see that the above conditions virtually amount to the construction of a theory T which is unique. We call it the *free theory* with base Ω and denote it by $S_0[\Omega]$.

The theory $S_0[\Omega]$ has the following two important properties, both of which are easily provable by induction on the degree:

(7.6) Given any theory T', any family of functions

$$\Omega_n \to T'(I, [n]), \qquad n = 0, 1, \cdots$$

admits a unique extension to a morphism

$$S_0[\Omega] \to T'$$

of theories.

(7.7) Given a set A and functions

$$\bar{\omega}: A^n \to A \quad \text{for all} \quad \omega \in \Omega_n, \qquad n = 0, 1, \cdots,$$

there exists a unique $S_0[\Omega]$-algebra structure on A such that

$$(x_1, \cdots, x_n)\omega = \bar{\omega}(x_1, \cdots, x_n).$$

8. CONGRUENCES

Let A be a T-algebra. A congruence Q in A consists of an equivalence relation $\sim$ in A satisfying

$$(a_1, \cdots, a_p)\phi \sim (a_1', \cdots, a_p')\phi$$

for any $\phi: I \to [p]$ in T, provided $a_i \sim a_i'$ for $i = 1, \cdots, p$. It is then clear that the quotient set A/Q (i.e., the set of equivalence classes of A under the equivalence relation) acquires a structure of a T-algebra, uniquely determined by the condition that the natural factorization mapping $A \to A/Q$ be a morphism of T-algebras.

A congrunce Q in a theory T is a family of equivalence relations, one in each set $T([n], [p])$ satisfying the following conditions:

(8.1) If $\phi_1, \phi_2 : [n] \to [p]$ and $\phi_1 \sim \phi_2$, then $\phi_1\psi \sim \phi_2\psi$ for every $\psi: [q] \to [n]$ and $\gamma\phi_1 \sim \gamma\phi_2$ for every $\gamma: [p] \to [q]$.

(8.2) If $\phi_1, \phi_2 : [n] \to [p]$ and $\phi_1 i \sim \phi_2 i$ for every $i = 1, \cdots, n$, then $\phi_1 \sim \phi_2$.

(8.3) If $\phi_1, \phi_2 : I \to [p]$ are in S_0 and if $\phi_1 \sim \phi_2$, then $\phi_1 = \phi_2$.

Condition (8.1) permits us to define a category T/Q in which morphisms are equivalence classes of morphisms in T. Condition (8.3) insures that S_0 is embedded in T/Q and condition (8.2) shows that T/Q is a theory. This is the *quotient theory* of T by the congruence Q. A T/Q-algebra A is simply a T-algebra satisfying

$$(a_1, \cdots, a_p)\phi_1 = (a_1, \cdots, a_p)\phi_2,$$

wherever $\phi_1 \sim \phi_2$. In this way, the category $(T/Q)^\flat$ becomes the subcategory of $T^\flat$ determined by the T-algebras A that are "compatible" with Q.

Conditions (8.1) and (8.2) imply that Q is completely determined by knowing the equivalence relation in $T(I, [p])$ for every p. There result congruences Q_p in the free algebras A_p, $p = 0, 1, \cdots$. One can regard Q as the sequence $\{Q_p\}$ of these congruences and reformulate conditions (8.1)–(8.3) accordingly. The free T/Q-algebras are then simply the quotient algebras A_p/Q_p.

In practice, a congruence Q in T will seldom be given "in toto". It will usually be "generated" by designating for each p certain pairs $\phi_1, \phi_2 : I \to [p]$ and taking the least Q for which these pairs become congruent. It is easy to construct Q so that (8.1) and (8.2) are satisfied. Whether (8.3) is satisfied is impossible to predict. In this connection the following procedure is useful: Use Q (satisfying only (8.1) and (8.2)) to define congruences Q_p in A_p. Then Q satisfies (8.3) if and only if the algebra A_2/Q_2 has at least two points.

9. PRESENTATION OF THEORIES

The two operations described above, namely the formation of a free theory and the passage to the quotient theory by a congruence, form the basic two steps in the formation of theories. The procedure follows closely the method of presenting groups by generators and relations. Instead of discussing it in general, we shall illustrate by examples. The examples chosen are those particularly relevant to theory of automata.

The theory Sg whose algebras will be semigroups may be described as follows: We begin with a free theory T generated by a single morphism $\pi: I \to [2]$. More explicitly, $T = S_0[\Omega]$ where $\Omega_2 = \{\pi\}$ while $\Omega_n = \emptyset$ for $i \neq 2$. The T-algebras are then sets A with a binary multiplication $(a_1, a_2)\pi \in A$ subject to no conditions whatsoever. To introduce the associative law, we must "equate" certain two morphisms $I \to [3]$ in T or equivalently certain two elements of the free T-algebra A_3. These two elements are

$$((1, 2)\pi, 3)\pi, (1, (2, 3)\pi)\pi.$$

This generates a congruence Q in T (*a priori* satisfying only (8.1) and (8.2)). A T-algebra A is compatible with Q if and only if

$$((a_1, a_2)\pi, a_3)\pi = (a_1, (a_2, a_3)\pi)\pi$$

holds for any $a_1, a_2, a_3 \in A$; i.e., if and only if A is a semigroup. The fact that there exist semigroups with more than one element implies that Q satisfies also (8.3). Then Sg is defined as T/Q.

To obtain the "monoid theory" M whose algebras will be monoids, one proceeds as above, but in addition to π: $I \to [2]$, one has an additional generator $\epsilon: I \to \emptyset$. Then in the free algebra A_1, one must "equate" the following three elements

$$(1, \epsilon_1)\pi, 1, (\epsilon_1, 1)\pi,$$

where ϵ_1 is the composition

$$I \xrightarrow{\epsilon} \emptyset \xrightarrow{\sigma} I,$$

σ being the unique morphism.

There is one more theory that is of vital interest in the theory of automata. Let M be a fixed monoid. We shall construct a theory whose algebras will be sets on which M operates on the right. To this end we consider generators $m: I \to I$ in a 1–1 correspondence with the elements of M. In the free algebra A_1 we then must equate the pairs

$$(1m)n, 1(mn) \qquad m, n \in M$$

$$1, 1\epsilon$$

where $\epsilon \in M$ is the unit element of M. Usually we adjoin to this theory an additional generator τ: $I \to \emptyset$ without any axioms. There results a theory $\tilde{M} = S_0[M, \tau]$. An $\tilde{M}$-algebra is a set A with a right action $am \in A$ for $a \in A$, $m \in M$ satisfying

$$(am)n = a(mn), \qquad a\epsilon = a$$

and with a selected element $\tau_A \in A$. The initial algebra for this theory is the monoid M itself acting on itself by multiplication and with $\epsilon = \tau_M$.

If M is a free monoid with base $\mu_1, \cdots, \mu_k$, then $\tilde{M}$ is the free theory $S_0[\mu_1, \cdots, \mu_k, \tau]$.

10. OPERATIONS ON THEORIES

An important operation on theories is the construction of the *free product* (also called *direct sum* or *coproduct*) $T = T' \oplus T''$ of two theories T' and T''. This theory is completely determined by the requirement that its algebras are to be sets A equipped with a T'-algebra structure and T''-algebra structure simultaneously, without any further conditions. The existence of such a theory can be established by choosing presentations of T' and T''. In particular, if $T' = S_0[\Omega']$ and $T'' = S_0[\Omega'']$, then $T = S_0[\Omega]$ where Ω is the disjoint union $\Omega' \cup \Omega''$. The free product $T \oplus S_0[\Omega]$ is denoted by $T[\Omega]$; this is the theory obtained from T by adjoining "freely" the operations Ω.

Let T be a theory and A^0 a T-algebra. One can construct a new theory $T[A^0]$ whose algebras will be pairs (A, f) where A is a T-algebra and $f: A^0 \rightarrow A$ is a morphism of T-algebras. A morphism $g: (A, f) \rightarrow (A', f')$ will be a morphism $g: A \rightarrow A'$ of T-algebras satisfying $gf = f'$. The theory $T[A^0]$ may be constructed by first adjoining freely all the elements of A^0 to T as operations $I \rightarrow \emptyset$ and then dividing by a suitable congruence. The theory $T[A^0]$ has the property that the pair $(A^0, 1_{A^0})$ becomes the initial algebra in category $T[A^0]^\flat$. If $A^0 = A_X$ is a free algebra on a base X, then $T[A^0]$ is nothing else than the free extension $T[\Omega]$ with $\Omega_0 = X$, $\Omega_i = \emptyset$ for $i > 0$.

11. RECOGNIZABLE SETS

Let A be a T-algebra and Q a congruence in A. We say that Q is *finite* if A has a finite number of equivalence classes mod Q, or equivalently if A/Q is a finite T-algebra. A subset X of A is said to be *closed* for Q if X is the union of congruence classes mod Q, or equivalently if $a \sim b$ and $a \in X$ imply $b \in X$. A subset X of A which is closed relative to some finite congruence Q is called *recognizable*. The class of recognizable subsets of A is closed with respect to finite Boolean operations.

If M is a monoid, then M may be viewed as a T-algebra for a variety of theories T. If we take for T the "monoid theory", then $T^\flat$ is the category of monoids and $M \in T^\flat$. A congruence Q in A is then an equivalence relation for which $m_1 \sim m_2$ implies $km_1l \sim km_2l$ for all k, $l \in M$. The same notion of congruence in M is obtained if we view M as the initial algebra for the extended theory $T[M]$.

On the other hand, M may also be viewed as the initial algebra for the theory $\tilde{M} = S_0[M, \tau]$ described in section 9. A congruence in M is then an equivalence relation for which $m_1 \sim m_2$ implies $m_1 l \sim m_2 l$ for all $l \in M$.

It is a known fact that both types of congruences in M lead to the same class of recognizable sets.

12. AUTOMATA

Let T be a theory. A *T-automaton* is a pair $\mathbf{A} = (A, t)$ where A is a *finite* T-algebra and t is a subset of A. The T-automata are converted into a category T_a^b by defining a morphism $f \colon \mathbf{A} \to \mathbf{B}$ where $\mathbf{B} = (B, s)$ as a morphism $f \colon A \to B$ of T-algebras such that $f^{-1}s = t$.

The *behavior* $\mathfrak{B}\mathbf{A}$ is defined as a subset of the initial T-algebra A_0 as follows: Let $\zeta_A : A_0 \to A$ be the unique T-morphism. Then $\mathfrak{B}\mathbf{A} = \zeta_A^{-1}t$.

The morphism ζ_A defines a congruence Q in A_0 by defining $a_1 \sim a_2$ whenever $\zeta_A a_1 = \zeta_A a_2$. This congruence is finite since A is finite. Further, $\mathfrak{B}\mathbf{A}$ is closed for Q.

$\mathfrak{B}\mathbf{B}$ is a recognizable subset of A_0. Conversely, let X be any recognizable subset of A_0. Let then Q be a finite congruence in A_0 for which X is closed. Then $A = A_0/Q$ is a finite T-algebra and setting $t = X/Q$ we obtain a T-automaton $\mathbf{A} = (A, t)$. Further, ζ_A is the natural factorization mapping $A_0 \to A_0/Q = A$. Thus, $X = \zeta_A^{-1}t = \mathfrak{B}\mathbf{A}$. This shows that the class of all the behaviors of T-automata coincides with the class of recognizable subsets of A_0.

13. RELATIONAL ALGEBRAS

In order to generalize the notion of a nondeterministic automaton, we restrict ourselves to the case that the theory T is free: $T = S_0[\Omega]$. In view of (7.7), a T-algebra A is then described by functions $(x_1, \cdots, x_n)\omega \in A$ for $x_1, \cdots, x_n \in A$ and $\omega \in \Omega_n$.

We define a *relational* T-algebra A to consist of a set A together with functions which to $x_1, \cdots, x_n \in A$ and $\omega \in \Omega_n$ assign a *subset* $(x_1, \cdots, x_n)\omega$ of A. If $X_1, \cdots, X_n$ are subsets of A, then we set

$$(X_1, \cdots, X_n)\omega = \bigcup(x_1, \cdots, x_n)\omega, \tag{13.1}$$

the union extended over all n-tuples $(x_1, \cdots, x_n)$ in A such that $x_i \in X_i$, $i = 1, \cdots, n$. In this way, the set $\hat{A}$ of all the subsets of A becomes a T-algebra.

Conversely, assume that on the set $\hat{A}$ we have a T-algebra structure satisfying the "distributive" law (13.1) where A is regarded as a subset of $\hat{A}$. Then restricting each $\omega \in \Omega_n$ to A gives the relational T-algebra A.

A morphism $f: A \to B$ of relational T-algebras is defined as a morphism

$$f: \hat{A} \to \hat{B} \quad \text{in} \quad T^\flat$$

satisfying the distributivity condition

$$fX = \bigcup fx, \qquad x \in X. \tag{13.2}$$

A function $\hat{A} \to \hat{B}$ satisfying (13.2) is the same thing as a relation from A into B; it is described by the subset R of $A \times B$ defined as follows:

$$R = \{(a, b) \mid b \in fa\}.$$

The relational T-algebras form a category denoted by $T^\natural$. The category of T-algebras is a subcategory of $T^\natural$. Further, the passage from A to $\hat{A}$ yields a functor $\Lambda : T^\natural \to T^\flat$.

An important fact to note is that the initial algebra A_0 for the category $T^\flat$ remains an initial algebra also within the larger category $T^\natural$. Indeed, if $A \in T^\natural$, then $\hat{A} \in T^\flat$ and we have a unique $\zeta_{\hat{A}} : A_0 \to \hat{A}$. This defines $\zeta_A : A_0 \to A$ in $T^\natural$, which is unique since $\zeta_{\hat{A}}$ is.

14. RELATIONAL AUTOMATA

We define a relational automaton $\mathbf{A} = (A, t)$ exactly as above, except that $A \in T^\natural$. The behavior is defined as

$$\mathcal{B}\mathbf{A} = \zeta_A^{-1} t = \{x \mid x \in A_0, \zeta_A x \cap t \neq \emptyset\}.$$

It is now clear that if we define the automaton

$$\mathbf{A} = (\hat{A}, t'), \qquad t' = \{X \mid X \subset A, X \cap t \neq \emptyset\},$$

then $\mathcal{B}\mathbf{A} = \mathcal{B}\mathbf{A}$. We thus have the generalization of the known fact that nondeterministic automata recognize the same sets as deterministic automata.

15. POLYNOMIALS

Let T be a free theory. A *polynomial*

$$P: [n] \to [p]$$

is an n-tuple $P = (P_1, \cdots, P_n)$ where $P_1, \cdots, P_n$ are finite subsets of

$T(1, [p])$. The elements of P_i are called the *constituents* of P, and if all these constituents have degree 1, then we say that P has degree 1.

Let A be a relational T-algebra, and let $X = (X_1, \cdots, X_p)$ be a p-vector of subsets of A. We define

$$XP_i = \bigcup_{\phi \in P_i} (X_1, \cdots, X_p)\phi$$

$$XP = (XP_1, \cdots, XP_n).$$

Thus, XP is an n-vector of subsets of X. Therefore, P defines a function

$$P_A : \hat{A}^p \rightarrow \hat{A}^n.$$

In the p-fold product $\hat{A}^p$ of $\hat{A}$ we define inclusion and union coordinate by coordinate. We then have the following important property of P_A :

(15.1) If in $\hat{A}^p$ we have

$$X^0 \subset X^1 \subset \cdots \subset X^k \subset \cdots,$$

then

$$(\bigcup_k X^k)P_A = \bigcup_k (X^k P_A).$$

For the proof it suffices to consider the case $n = 1$ and $P = P_1 = \phi : I \rightarrow [p]$ is a monomial (i.e., P has a single constituent). In this case the desired relation is proved in a straightforward manner by induction on the degree of ϕ.

Property (15.1) implies that P_A is monotone; i.e., that $XP_A \subset YP_A$ whenever $X \subset Y$ in $\hat{A}^p$.

We now consider a polynomial

$$P : [n] \rightarrow [n].$$

Then the transformation

$$P_A : \hat{A}^n \rightarrow \hat{A}^n$$

may be iterated, yielding

$$P_A^k : \hat{A}^n \rightarrow \hat{A}^n,$$

for which (15.1) also holds. In particular, if $\emptyset \in \hat{A}^n$ is the n-tuple $(\emptyset, \cdots, \emptyset)$, we have

$$\emptyset \subset \emptyset P_A \subset \emptyset P_A^2 \subset \cdots \subset \emptyset P_A^k \subset \cdots$$

and we define

$$\bar{P}_A = \bigcup_k \emptyset P_A{}^k.$$

The following fact should be regarded as well known (as well as easily provable):

(15.2) $\bar{P}_A$ is the least solution of the equation

$$XP_A = X,$$

for $X \in A^n$, as well as the least solution of the inequality

$$XP_A \subset X.$$

In the special case where $A = A_0$ is the initial algebra in $T^\natural$, we shall write $\bar{P}$ instead of $\bar{P}_{A_0}$. This special case is all-important because of

$$\text{If } \zeta_A : A_0 \rightarrow A, \qquad \text{then } \bar{P}_A = \zeta_A \bar{P}. \tag{15.3}$$

Here, ζ_A denotes the mapping $\zeta_A : \hat{A}_0{}^n \rightarrow \hat{A}^n$ defined by the mapping $\zeta_A : \hat{A}_0 \rightarrow \hat{A}$ given by the relation $\zeta_A : A_0 \rightarrow A$. The fact stated in (15.3) follows readily from the commutative diagram

$$\begin{tikzcd}
\hat{A}_0{}^n \arrow[r, "P_{A_0}"] \arrow[d, "\zeta_A"'] & \hat{A}_0{}^n \arrow[d, "\zeta_A"] \\
\hat{A}^n \arrow[r, "P_A"'] & \hat{A}^n
\end{tikzcd}$$

and the facts that

$$\zeta_A \emptyset = \emptyset, \qquad \zeta_A \bigcup_k X^k = \bigcup_k \zeta_A X^k.$$

16. ALGEBRAIC SETS

A subset X of the initial algebra A_0 for a free theory $T = S_0[\Omega]$ is called *algebraic* if there exists an integer n and a polynomial $P: [n] \rightarrow [n]$ such that $X = \bar{P}_1$; i.e., X is the first coordinate of the least solution of the equation $YP_{A_0} = Y$ for $Y \in \hat{A}_0{}^n$.

The following properties of algebraic sets should be regarded as known:

(16.1) Each element x of A_0 is an algebraic set.

(16.2) The empty set is algebraic.

(16.3) If X_1 and X_2 are algebraic sets, then so is $X_1 \cup X_2$.

(16.4) If $\phi: I \rightarrow [p]$ in T and $X_1, \cdots, X_p$ are algebraic sets, then so is $(X_1, \cdots, X_p)\phi$.

We are now in a position to state the main results of this paper.

THEOREM 1. *For each algebraic set X in the initial algebra A_0 over a free theory $T = S_0[\Omega]$, there exists an integer $n > 0$ and a polynomial $P: [n] \to [n]$ of degree* 1 *such that $X = \bar{P}_1$.*

The proof will be given in section 17.

Now assume that the theory $T = S_0[\Omega]$ is free on a finite base; i.e., that each of the sets Ω_k is finite and that $\Omega_k = \emptyset$ for all but a finite number of integers $k \geqq 0$.

Let A be a relational T-algebra with $[n]$ as underlying set. We associate with A a polynomial,

$$A^P: [n] \to [n]$$

of degree 1 as follows: A morphism $\phi: I \to [n]$ of degree 1 is a composition

$$I \xrightarrow{\omega} [p] \xrightarrow{x} [n],$$

where $\omega \in \Omega_p$ and $x = (x1, \cdots, xp)$, a p-tuple of elements in $[n]$; i.e., in A. We define $x\omega \in P_i$ if and only if $i \in (x1, \cdots, xp)\omega$ according to the relational T-algebra structure A. This clearly gives a bijection between the relational T-algebra structures A on $[n]$ and polynomials $P: [n] \to [n]$ of degree 1.

THEOREM 2. *If the relational T-algebra A on $[n]$ and the polynomial $P: [n] \to [n]$ of degree* 1 *are related as above, then*

$$\bar{P}_i = \zeta_A^{-1} i.$$

We recall here that $\zeta_A : A_0 \to A = [n]$ is a relation so that

$$\zeta_A^{-1} i = \{y \mid y \in A_0, i \in \zeta_A y\}.$$

The proof will be given in section 18.

From the two theorems asserted above, we can now prove:

THEOREM 3. *If $T = S_0[\Omega]$ is a free theory on a finite base, then in the initial T-algebra A_0 the recognizable sets and the algebraic sets coincide.*

Proof. Let $X \subset A_0$ be recognizable. Then $X = \mathfrak{B}\mathbf{A}$ where $\mathbf{A} = (A, t)$ is an automaton. Since the T-algebra A is finite, we may, without loss, assume that the underlying set of A is $[n]$ for some $n > 0$. Let $P: [n] \to [n]$ be the associated polynomial of degree 1. Then by Theorem 2

$$X = \mathfrak{B}\mathbf{A} = \zeta_A^{-1} t = \bigcup_{i \in t} \zeta_A^{-1} i = \bigcup_{i \in t} \bar{P}_i.$$

Since each $\bar{P}_i$ is algebraic, it follows from (16.3) that X is algebraic.

Conversely, let X be an algebraic set in A_0. Then by Theorem 1 we have $X = \bar{P}_1$ for some polynomial $P\colon [n] \to [n]$ of degree 1. Let A be the relational T-algebra structure associated with P and let $A = (A, t)$ be the relational automaton with $t = \{1\}$. Then by Theorem 2

$$\mathcal{B}\mathbf{A} = \zeta_A^{-1}1 = \bar{P}_1 = X$$

so that X is recognizable.

17. PROOF OF THEOREM 1

We shall establish two auxiliary propositions.

PROPOSITION 1. *Given a polynomial $P\colon [n] \to [n]$, there exists a polynomial $Q\colon [n] \to [n]$ such that*

(i) *The constituents of Q are precisely the constituents of P of degree >0.*

(ii) $\bar{Q} = \bar{P}$.

PROPOSITION 2. *Given a polynomial $P\colon [n] \to [n]$, there exists a polynomial $Q\colon [m] \to [m]$, $n \leqq m$ such that*

(i) *All constituents of Q have degree $\leqq 1$.*

(ii) *Q_i has the same constituents of degree 0 as P_i, $i = 1, \cdots, n$.*

(iii) *Q_i, $n < i < m$ has no constituents of degree 0.*

(iv) *$\bar{Q}_i = \bar{P}_i$ for $i = 1, \cdots, n$.*

It is now clear how Theorem 1 follows from these two propositions. Given an algebraic set X in A_0, choose a polynomial $P\colon [n] \to [n]$ such that $X = \bar{P}_1$. Then apply both propositions consecutively and in either order. There results a polynomial $Q\colon [m] \to [m]$, $n \leqq m$ of degree 1 such that $\bar{Q}_1 = X$.

In the proofs that follow, it will be convenient to use the symbol $+$ for $\cup$ and Σ for $\bigcup$.

Proof of Proposition 1. We represent the polynomial P in the form $P = R + M$ where $R\colon [n] \to [n]$ consists of all the constituents of P of degree >0, while M consists of the constituents of P of degree 0. The morphisms $j\colon 1 \to [n]$ of degree 0 are $j = 1, \cdots, n$. Thus M may be represented as the $n \times n$ matrix $\{M_{ij}\}$ whose coordinates M_{ij} are 1 or 0 depending on whether $j\colon I \to [n]$ is a constituent of M_i (i.e., a constituent of P_i). We now form the matrix

$$N = E + M + M^2 + \cdots + M^k + \cdots$$

as follows: We regard $\mathcal{B} = \{0, 1\}$ as a semi-ring with the operation table

$$0 + 0 = 0, \qquad 0 + 1 = 1 + 0 = 1 + 1 = 1$$
$$11 = 1, \qquad 01 = 10 = 00 = 0$$

and regard M as a matrix with coefficients in $\mathfrak{B}$. Matrices are multiplied and added in the usual fashion. E denotes the matrix with 1 on the diagonal and zero everywhere else. The sequence of matrices

$$M^{(k)} = E + M + M^2 + \cdots + M^k$$

$k = 0, 1, \cdots$ is ascending and, therefore, for some k we have $M^{(k)} = M^{(l)}$ for all $l \geqq k$. The matrix N is then defined as $M^{(k)}$, for k sufficiently large. We now define the polynomial Q: $[n] \rightarrow [n]$ as $Q = RN$; i.e.,

$$Q_i = \sum_k R_k N_{ki}.$$

Since $N_{ii} = 1$ we have $R_i \subset Q_i$. Condition (i) of Proposition 1 is then clearly satisfied and we now prove that $\bar{Q} = \bar{P}$.

Assume that $X = (X_1, \cdots, X_n)$ is a vector of subsets of A_0 such that $XP \subset X$. Then since $P = R + M$, we must have $XR \subset X$ and $XM \subset X$. Therefore, $XM^k \subset X$ for all k and thus $XN \subset X$. It follows that

$$XQ = (XR)N \subset XN \subset X.$$

This proves that $\bar{Q} \subset \bar{P}$. To prove the converse, assume that $XQ = X$. Since $R \subset Q$, we have $XR \subset X$. Since $NM \subset N$, we have

$$QM = RNM \subset RM = Q.$$

Therefore,

$$XP = XR + XM = XR + XQM \subset X + QX = X.$$

This shows that $\bar{P} \subset \bar{Q}$. Thus $\bar{P} = \bar{Q}$.

Proof of Proposition 2. Assume that $k > 1$ is the highest degree of the constituents of P and let ϕ: $I \rightarrow [n]$ be a constituent of P of degree k. We may write

$$P = R + \phi M,$$

where R is a polynomial not containing ϕ as a constituent while $M = (M_1, \cdots, M_n)$ is a vector with components 0, 1 defined by $M_i = 1$ or $M_i = 0$ depending on whether or not ϕ is a constituent of P_i. The morphism ϕ has a factorization

$$I \xrightarrow{\omega} [p] \xrightarrow{\psi} [n]$$

with $\omega \in \Omega_p$ and $d\psi = k - 1$. Since $d\psi = \sum d(\psi i)$ $i = 1, \cdots, p$ we must have $d(\psi i) > 0$ for some i. Without loss of generality, assume that $d(\psi p) > 0$.

Finding $\bar{P}$ is equivalent to finding the minimal solution of the system of equations

$$X = XP \tag{17.1}$$

in subsets $X = (X_1, \cdots, X_n)$ of A_0. The equation (17.1) may be rewritten as

$$X = XR + (X\psi_1, \cdots, X\psi_p)\omega M, \tag{17.2}$$

where $\psi_i = \psi i$.

We now consider the system of $n + 1$ equations with $n + 1$ unknowns as follows:

$$\begin{aligned} X &= XR + (X\psi_1, \cdots, X\psi_{p-1}, X_{n+1})\omega M \\ X_{n+1} &= X\psi_p, \end{aligned} \tag{17.3}$$

where $X = (X_1, \cdots, X_n)$. It is clear that if $X = (X_1, \cdots, X_n)$ is the minimal solution of (17.2), then $(X_1, \cdots, X_n, X_{n+1})$ with $X_{n+1} = X\psi_p$ is the minimal solution of (17.3).

The right-hand side of (17.3) yields a polynomial $Q: [n + 1] \to [n + 1]$ whose constituents are: (1°) compositions

$$I \xrightarrow{\gamma} [n] \xrightarrow{f} [n + 1],$$

where γ is a constituent of P different from ϕ and f is an inclusion; (2°) morphisms $\tau: I \to [n + 1]$ satisfying $0 < d\tau < k$. By the above, $\bar{Q}_i = \bar{P}_i$ for $i = 1, \cdots, n$.

An iteration of the above procedure yields the conclusion of Proposition 2.

18. PROOF OF THEOREM 2

Let

$$Q_i = \{\phi \mid \phi: I \to \emptyset, \quad i \in \zeta_A\phi\}.$$

Then $Q_i \subset A_0$ and we wish to show that $Q = \bar{P}$ where $Q = (Q_1, \cdots, Q_n)$.

We first show that $\bar{P} \subset Q$. For this, it suffices to show that $QP \subset Q$; i.e., that

$$(Q_1, \cdots, Q_n)\phi \subset Q_i \quad \text{whenever} \quad \phi \in P_i.$$

Let $\phi \in P_i$. Then ϕ is the composition

$$I \xrightarrow{\omega} [p] \xrightarrow{x} [n]$$

with $\omega \in \Omega_p$, x is a mapping in S_0 and

$$i \in (x1, \cdots, xp)\omega.$$

Thus we must show that

$$(Q_{x1}, \cdots, Q_{xp})\omega \subset Q_i. \tag{18.1}$$

Let then

$$\psi_j \in Q_{xj} \qquad j = 1, \cdots, p$$

$$\psi = (\psi_1, \cdots, \psi_p) : [p] \rightarrow \emptyset.$$

Then $xj \in \zeta_A\psi_j$ for $j = 1, \cdots, p$, and

$$i \in (x1, \cdots, xp)\omega \subset (\zeta_A\psi_1, \cdots, \zeta_A\psi_p)\omega = \zeta_A(\psi\omega).$$

Consequently, $\psi\omega \in Q_i$, so that (18.1) holds.

To show the opposite inclusion, we must prove that

(18.2) if $\phi: I \rightarrow \emptyset$ and $i \in \zeta_A\phi$, then $\phi \in \bar{P}_i$.

This will be done by induction with respect to the degree of ϕ. First let $d\phi = 1$. Then $\phi \in \Omega_0$ and $i \in \phi_A$. Then the composition

$$I \xrightarrow{\phi} \emptyset \rightarrow [n]$$

is in P_i so that $\phi = \phi_{A_0} \in \emptyset P_i$. Thus $\phi \in \bar{P}_i$ as required.

Now assume $d\phi = k > 1$. Let

$$I \xrightarrow{\omega} [p] \xrightarrow{\psi} [n]$$

be the factorization of ϕ with $\omega \in \Omega_p$ and $d\psi = k - 1 > 0$. Then $p > 0$. We have

$$i \in \zeta_A\phi = \zeta_A\psi\omega = \zeta_A(\psi 1, \cdots, \psi p)\omega = (\zeta_A\psi_1, \cdots, \zeta_A\psi p)\omega.$$

Let then

$$x: [p] \rightarrow [n] = A \quad \text{in} \quad S_0$$

be such that

$$xj \in \zeta_A\psi j \quad j = 1, \cdots, p \tag{18.3}$$

$$i \in (x_1, \cdots, xp)\omega. \tag{18.4}$$

It follows from (18.4) that the composition

$$I \xrightarrow{\omega} [p] \xrightarrow{x} [n]$$

is in P_i. Condition (18.3), in view of $d\psi j < k$, implies that $\psi j \in \bar{P}_{xj}$. Thus $\phi = (\psi 1, \cdots, \psi p)\omega \in (\bar{P}_{x1}, \cdots, \bar{P}_{xp})\omega = (\bar{P}_1, \cdots, \bar{P}_n)x\omega \subset (\bar{P}_1, \cdots, \bar{P}_n)P_i \subset \bar{P}_i$. Consequently, $\phi \in \bar{P}_i$ as required.

Acknowledgment

The authors wish to express their appreciation to Dr. C. C. Elgot for his careful reading of the manuscript and many suggested modifications of the text.

Received: September 27, 1967

REFERENCES

Arden, D. N. (1961). Delayed-logic and finite-state machines. Course notes 6.531. Department of Electrical Engineering, M.I.T.

Buchi, J. R. and Wright, J. B. (1960). "Mathematical Theory of Automata." course notes, Communications Sciences 403, University of Michigan.

Chomsky, N. and Schützenberger, M. P. (1963). The Algebraic Theory of Context-Free Languages. *In* "Computer Programming and Formal Systems" (edited by P. Braffort and D. Hirschberg), North-Holland, Amsterdam, p. 118.

Elgot, C. C. and Mezei, J. (1965). On relations defined by generalized finite automata. *IBM J. Res. Develop.* Vol. **9.**

Ginsburg, S. and Rice, H. G. (1962). Two families of languages related to ALGOL. *J. Assoc. Comp. Mach.* **9,** p. 350.

Lawvere, F. W. (1963). Functorial semantics of algebraic theories. *Proc. Natl. Acad. Sci. U.S.A.* **50,** 869–872.

Mezei, J. and Wright, J. B. (1965). Generalized ALGOL-like languages. IBM Research Report RC 1528. Also to appear in *Inform. Control* as Algebraic automata and context-free sets.

Rabin, M. O. and Scott, D. (1959). Finite automata and their decision problems. *IBM J. Res. Develop.* Vol. **3.**

Thatcher, J. W. and Wright, J. B. (1966). Generalized finite automata theory with an Application to a Decision Problem of second-order logic. IBM Research Report RC 1713.