Semigroup Forum Vol. 24 (1982) 25-38

RESEARCH ARTICLE

THE VARIETY GENERATED BY FINITE NILPOTENT MONOIDS

Howard Straubing

Communicated by G. Lallement

1. A finite semigroup S is said to be nilpotent if S has a zero and if $S^n = 0$ for some positive integer n. The family of all finite nilpotent semigroups forms a variety of finite semigroups in the sense of the word used by Eilenberg (see [EIL], particularly Section VIII.2) - that is, it is closed under finite direct products and division. Let us say that a finite monoid M is nilpotent if $M - \{1\}$ is a nilpotent semigroup. The family of all finite nilpotent monoids does not form a variety of finite monoids, since the direct product of two nilpotent monoids is not, in general, a nilpotent monoid. One can, however, consider the smallest variety $\underline{V}$ of finite monoids which contains all the nilpotent monoids. A monoid M belongs to $\underline{V}$ if and only if

$$M \prec M_1 \times \cdots \times M_k$$

for some nilpotent monoids $M_1, \cdots, M_k$. (The symbol $\prec$ means "divides".) The principal result of this article is that $M \in \underline{V}$ if and only if M is aperiodic (that is, all groups contained in M are trivial) and for all $e, s \in M$ such that e is idempotent, $es = se$.

This article is intended as a small but nontrivial contribution to the theory of varieties of finite semigroups as initiated by Eilenberg. The proof of the main

0037-1912/82/0024-0025 $03.00

result uses a number of important ideas from this theory: ultimately equational definitions of varieties, congruences of finite index on a finitely generated free monoid, and syntactic monoids of recognizable languages (the anxious reader is referred to [EIL] and [LAL] for definitions and some idea as to what this is all about!). After the proof I briefly discuss the connections between the present result and some of the older results of this theory.

2. As above, let $\underline{V}$ be the variety of finite monoids generated by the finite nilpotent monoids.

THEOREM: The following four conditions are equivalent.

(a) $M \in \underline{V}$

(b) M is aperiodic, and for all $e, s \in M$ such that $e^2 = e$, $es = se$.

(c) M satisfies the equations

$$x^n = x^{n+1};\ yx^n = x^n y$$

for each sufficiently large positive integer n. (That is, no matter how one substitutes elements of M for the 'variables' x and y, the above equations hold.)

(d) M satisfies the equations

$$x^n = x^{n+1};\ y_0xy_1x \cdots xy_n = x^n y_0 \cdots y_n = y_0 \cdots y_n x^n$$

for each sufficiently large positive integer n.

Proof: (a) $\Longrightarrow$ (b). It suffices to show that the collection $\underline{W}$ of all finite aperiodic monoids M satisfying $es = se$ for all $e,s \in M$ with $e^2 = e$ is a variety which contains all the nilpotent monoids. First of all, every nilpotent monoid M is aperiodic, and the only idempotents in M are 0 and 1, which commute with every element of M. Thus $\underline{W}$ contains all the nilpotent monoids. To show that $\underline{W}$ is a variety, observe that $\underline{W} = \underline{W}' \cap \underline{A}$, shere $\underline{A}$ is the collection of all finite aperiodic monoids and $\underline{W}'$ is the collection of all finite monoids M such that $es = se$ for all $e,s \in M$ with $e^2 = e$. Since $\underline{A}$ is a variety (see [EIL] Section V.3) and since the intersection of two varieties is a variety, it suffices now to show that $\underline{W}'$ is a variety - that is, that $\underline{W}'$ is closed under submonoids, quotients and direct products. It is trivial to verify that $\underline{W}'$ is closed under submonoids and direct products. To see that it is closed under quotients as well, let $M_1 \in \underline{W}'$ and let $\varphi : M_1 \longrightarrow M_2$ be a surjective morphism. If $e,s \in M_2$ with $e^2 = e$, then there exist $e',s' \in M_1$ such that $e'\varphi = e$, $s'\varphi = s$. Furthermore, some power $f = (e')^k$ of e' is idempotent, and $f\varphi = e^k = e$. Since $M_1 \in \underline{W}'$, $fs' = s'f$, and consequently $es = (fs')\varphi = (s'f)\varphi = se$. Thus $M_2 \in \underline{W}'$, so $\underline{W}'$ is closed under quotients.

<u>(b) $\Longrightarrow$ (c).</u> Suppose that M is aperiodic and that $es = se$ for all $e, s \in M$ with $e^2 = e$. Since M is aperiodic there exists a positive integer k such that for all x, $x^n = x^{n+1}$ whenever $n \geq k$. It follows, in particular, that $x^n = x^{2n}$ and thus x^n is idempotent when $n \geq k$. The condition on M then implies that for all $x, y \in M$, $x^n y = yx^n$. Thus M satisfies the equations in (c) for every sufficiently large positive integer n.

<u>(c) $\Longrightarrow$ (d).</u> Suppose that M satisfies the equations in (c) for all n greater than or equal to some positive integer k.

I denote by $\leq_{\mathcal{R}}$, $\leq_{\mathcal{L}}$ and $\leq_{\mathcal{J}}$ the $\mathcal{R}$-, $\mathcal{L}$- and $\mathcal{J}$-orderings on M, and by $\equiv_{\mathcal{R}}$, $\equiv_{\mathcal{L}}$, $\equiv_{\mathcal{J}}$ and $\equiv_{\mathcal{H}}$ the relations of $\mathcal{R}$-, $\mathcal{L}$-, $\mathcal{J}$- and $\mathcal{H}$-equivalence. I will show that M is $\mathcal{R}$-trivial - that is, $s \equiv_{\mathcal{R}} t$ implies $s = t$. First, if $s \equiv_{\mathcal{R}} e$ where e is idempotent, then $ex = s$ for some $x \in M$, and thus $s = ex = e^k x = xe^k = xe$ (by the condition (c)) so $s \leq_{\mathcal{L}} e$. In a finite semigroup, two $\mathcal{R}$-equivalent elements which are related in the $\mathcal{L}$-ordering must be $\mathcal{L}$-equivalent as well, hence $s \equiv_{\mathcal{L}} e$, and consequently $s \equiv_{\mathcal{H}} e$. The $\mathcal{H}$-class of the idempotent e is a group contained in M, but the equation $x^k = x^{k+1}$ implies that M is aperiodic. Consequently this group is trivial, so $s = e$. Now suppose $s,t \in M$ and $s \underset{\mathcal{R}}{\equiv} t$. Then $sa = t$ and $tb = s$ for some $a,b \in M$, and thus $s(ab)^k = s$ and $s(ab)^k a = t$. Now $(ab)^k = (ab)^{k+1} \underset{\mathcal{R}}{\leq} (ab)^k a \underset{\mathcal{R}}{\leq} (ab)^k$. Hence $(ab)^k \underset{\mathcal{R}}{\equiv} (ab)^k a$. Since $(ab)^k$ is idempotent, it follows from what was proved above that $(ab)^k = (ab)^k a$, thus $s = s(ab)^k = s(ab)^k a = t$, and M is $\mathcal{R}$-trivial, as claimed. (See [LAL], especially Section 2.3, for a discussion of the facts used in the above argument. The argument itself is taken from [PIN], Proposition 2.1.1.)

Now let r be the length of the longest sequence of elements $s_1,\cdots,s_r$ of M such that $s_{i+1} \underset{\mathcal{R}}{\leq} s_i$ and $s_{i+1} \underset{\mathcal{R}}{\not\equiv} s_i$ for $i = 1,\cdots,r-1$. Let $n \geq \max\{k,r\}$; I will show that M satisfies the equations in (d) for n. If $x,y_0,y_1,\cdots,y_n \in M$, let $t_0 = y_0x$, $t_1 = y_0xy_1x$, $\cdots$, $t_{n-1} = y_0x\cdots y_{n-1}x$, $t_n = y_0x\cdots y_{n-1}xy_n$. Then $t_n \underset{\mathcal{R}}{\leq} t_{n-1} \underset{\mathcal{R}}{\leq} \cdots \underset{\mathcal{R}}{\leq} t_0$, and since $n \geq r$, there must be $i \in \{0,\cdots,n-1\}$ such $t_i \underset{\mathcal{R}}{\equiv} t_{i+1}$. Thus $t_{i+1} = t_iy_{i+1}x \underset{\mathcal{R}}{\equiv} t_iy_{i+1} \underset{\mathcal{R}}{\equiv} t_i$, and since M is $\mathcal{R}$-trivial, it follows that $t_i = t_iy_{i+1} = t_{i+1}$. In particular $t_ix = t_i$, and consequently $t_i x^p = t_i$ for all $p \geq 0$. Thus $t_n = y_0 x \cdots y_i x y_{i+1} \cdots y_n = y_0 x \cdots y_i x^n y_{i+1} x \cdots y_n$. Now by applying the equations in (c), $x^n y_{i+1} x$ can be rewritten $y_{i+1} x^{n+1}$, then $x^{n+1} y_{i+2}$ can be rewritten $y_{i+2} x^{n+2}$, etc. We thus obtain

$$t_n = y_0 x_i \cdots y_i x y_{i+1} \cdots y_n = y_0 x \cdots y_i y_{i+1} \cdots y_n x^{n'} \quad \text{for}$$

some $n' \geq n$. Further,

$$x y_i y_{i+1} \cdots y_n x^{n'} = x^{n'+1} y_i y_{i+1} \cdots y_n,$$

$$x y_{i-1} x^{n'+1} = x^{n'+2} y_{i-1}, \quad \text{etc., and finally}$$

we obtain

$$t_n = x^{n''} y_0 \cdots y_n \ .$$

Since $n'' \geq n \geq k$, the equation $x^k = x^{k+1}$ implies $x^{n''} = x^n$. Thus $t_n = x^n y_0 \cdots y_n$. A similar argument shows $t_n = y_0 \cdots y_n x^n$. Thus M satisfies the equations in (d) for all $n \geq \max\{k,r\}$.

(d) $\Longrightarrow$ (a). The set $\underline{V}'$ of all finite monoids which satisfy the equations of (d) for sufficiently large n is a variety of finite monoids ([EIL], Section V.2) and in this part of the proof it must be shown that $\underline{V}' \subseteq \underline{V}$. Since every variety is generated by the syntactic monoids it contains ([EIL], VII.1.8) it is sufficient to show that if A is a finite alphabet and $L \subseteq A^*$ is a recognizable language such that $M(L) \in \underline{V}'$, then $M(L) \in \underline{V}$. To this end let L be such a language and let $\varphi : A^* \longrightarrow M(L)$ be the syntactic morphism of L.

For each $a \in A$, $u \in A^*$, let $|u|_a$ denote the number of occurrences of the letter a in the word u. Let $u\beta_m = \{a \in A \mid |u|_a \geq m\}$ (where m is chosen so that $M(L)$ satisfies the equations in (d) for all $n \geq m$), and let $\overline{u}$ denote the word obtained from u upon erasing all occurrences of the letters of $u\beta_m$. (For example, if $u = abacba$ and $m = 3$, then $u\beta_m = \{a\}$ and $\overline{u} = bcb$.) Let $w_1, w_2 \in A^*$. I define

$$w_1 \underset{m}{\approx} w_2$$

if and only if

$$w_1\beta_m = w_2\beta_m \text{ and } \overline{w_1} = \overline{w_2} .$$

It is evident that $\underset{m}{\approx}$ is an equivalence relation, and it is easy to verify that for all $w_1, w_2 \in A^*$ and $a \in A$, $w_1 \underset{m}{\approx} w_2$ implies $w_1a \underset{m}{\approx} w_2a$ and $aw_1 \underset{m}{\approx} aw_2$. Thus $\underset{m}{\approx}$ is a congruence on A^*. Furthermore, the index of $\underset{m}{\approx}$ is finite, as the sets $\{u\beta_m \mid u \in A^*\}$ and $\{\overline{u} \mid u \in A^*\}$ are both finite (the first is the set of subsets of A, the second the set of words w such that $|w|_a \leq m$ for all $a \in A$).

I claim that if $w_1 \underset{m}{\approx} w_2$ then $w_1\varphi = w_2\varphi$. Indeed, suppose that we order the alphabet $A = \{a_1, \cdots, a_k\}$. Now if $w \in A^*$ either $w\beta_m \neq \emptyset$ or $\overline{w} = w$. In the former case, let b_1 be the <u>smallest</u> element of $w\beta_m$ with respect to the order on A. Then

$$w = u_0 b_1 u_1 b_1 \cdots b_1 u_n$$

where $u_i \in (A - \{b_1\})^*$ for $i = 0, \cdots, n$ and where $n \geq m$. Since $M(L)$ satisfies the equations (d) we obtain

$$w\varphi = (b_1\varphi)^n (u_0 \cdots u_n)\varphi = (b_1\varphi)^m (u_0 \cdots u_n)\varphi .$$

Now if $w\beta_m = \{b_1\}$ then $u_0 \cdots u_n = \overline{w}$. If, on the other hand, there are elements of $w\beta_m$ different from b_1 then we take the smallest element b_2 of $w\beta_m - \{b_1\}$ and we repeat the above procedure with $u_0 \cdots u_n$. Finally we arrive at

$$w\varphi = (b_1\varphi)^m \cdots (b_\ell\varphi)^m (\overline{w}\varphi)$$

where $w\beta_m = \{b_1, \cdots, b_\ell\}$ and $b_1 < \cdots < b_\ell$. It follows at once that if $w_1 \underset{m}{\approx} w_2$ then $w_1\varphi = w_2\varphi$.

It follows that each congruence class of the syntactic congruence of L is a union of $\underset{m}{\approx}$-classes. Since L is itself a union of classes of the syntactic congruence, we obtain $L = \bigcup_{i=1}^{p} L_i$, where each L_i is a

$_m$-class. (The union is finite because the syntactic congruence and $\underset{m}{\approx}$ are both of finite index.) Thus

$$M(L) \prec M(L_1) \times \cdots \times M(L_p)$$

(see [EIL], VII.2.2). It therefore suffices to show that the syntactic monoid of each $\underset{m}{\approx}$-class belongs to $\underline{V}$. To this end, let $w \in A^*$, and let K_w denote the $\approx_m$-class containing w. I define a morphism $\gamma : A^* \longrightarrow A^*$ by

$$a\gamma = \begin{cases} 1 & \text{if } a \in w\beta_m \\ a & \text{if } a \notin w\beta_m \end{cases}$$

for all $a \in A$. Then $w \underset{m}{\approx} w'$ if and only if $w\beta_m = w'\beta_m$ and $w'\gamma = \overline{w}$. Thus $K_w = K_1 \cap K_2$, where

$$K_1 = \overline{w}\gamma^{-1}$$

$$K_2 = \{w' \in A^* \mid w'\beta_m = w\beta_m\} .$$

Since $M(K_1 \cap K_2) \prec M(K_1) \times M(K_2)$ ([EIL], VII.2.2) it is enough to show that $M(K_1) \in \underline{V}$ and $M(K_2) \in \underline{V}$. The syntactic monoid of the language consisting of the single word $\overline{w}$ is a nilpotent monoid (this follows directly from the fact that the syntactic <u>semigroup</u> of any finite subset of A^+ is a nilpotent semigroup - see [EIL], Section VIII.2), and

$$M(K_1) = M(\overline{w}\gamma^{-1}) \prec M(\underline{w})$$

([EIL], VII.2.2), so $M(K_1) \in \underline{V}$. To show that

$M(K_2) \in \underline{V}$, observe that $w'\beta_m = w\beta_m$ iff $|w'|_a \geq m$ for all $a \in w\beta_m$ and $|w'|_a < m$ for all $a \in A - w\beta_m$. Thus

$$K_2 = \bigcap_{a \in w\beta_m} \{w' \in A^* \mid |w'|_a \geq m\} - \bigcup_{a \notin w\beta_m} \{w' \in A^* \mid |w'|_a \geq m\}.$$

Now it is easy to prove that the syntactic monoid of the language $\{w' \in A^* \mid |w'|_a \geq m\}$ is the cyclic nilpotent monoid $\{1, a, a^2, \cdots, a^m = 0\}$. From the above expression for K_2 and Proposition VII.2.2 of [EIL], it follows that $M(K_2)$ divides the direct product of $|A|$ copies of this monoid. Thus $M(K_2) \in \underline{V}$. This completes the proof.

3. The collection N of all finite nilpotent semigroups forms a variety of finite semigroups, and Eilenberg ([EIL], Section VIII.2) describes the corresponding variety of languages. That is, he describes for each finite alphabet A the family $A^+\mathcal{N}$ of all recognizable languages in A^+ (the free semigroup generated by A) whose syntactic semigroups are nilpotent. In this instance, $A^+\mathcal{N}$ is the family of all finite and cofinite subsets of A^+. Equivalently, $A^+\mathcal{N}$ is the boolean closure of the family of all finite subsets of A^+. Thus $\underline{N}$ is the smallest variety of finite semigroups such that the corresponding variety of languages contains all the finite languages.

We can now answer the question: What is the smallest variety of finite monoids such that the corresponding variety of languages contains all the finite languages? This is the variety $\underline{V}$ studied above, and our theorem gives an effective method for determining if a given finite monoid M belongs to $\underline{V}$: M must be aperiodic and satisfy the condition $es = se$ for all $e, s \in M$ such that e is idempotent. The family $A^*\mathcal{V}$ of recognizable languages in A^* whose syntactic monoids are in $\underline{V}$ contains more than just the finite and co-finite sets. The last part of the proof of the theorem yields the following result: A recognizable language $L \subseteq A^*$ belongs to $A^*\mathcal{V}$ if and only if L is a union of $\approx_n$-classes for some n. A description which is perhaps more informative is the following:

> $A^*\mathcal{V}$ is the boolean closure C of the family of sets of the form $B^*a_1B^*a_2\cdots B^*a_kB^*$ where $a_1,\cdots,a_k \in A$ and $B = A - \{a_1,\cdots,a_k\}$.

(The letters $a_1,\cdots,a_k$ are not assumed to be distinct.) To see this, recall that in the last part of the proof it was shown that each congruence class was in the boolean closure of the family of sets of the form

$u\gamma^{-1}$ (where $u \in A^*$ and $\gamma : A^* \longrightarrow A^*$ the morphism which erases letters not appearing in u)

and $\{w \in A^* \mid |w|_a \geq n\}$.

Now $u\gamma^{-1} = B^*a_1B^*\cdots B^*a_kB^*$ where $u = a_1\cdots a_k$ and $B = A - \{a_1,\cdots,a_k\}$. Furthermore,

$$\{w \mid |w|_a \geq n\} = A^* - \bigcup_{0 \leq p < n} \{w \in A^* \mid |w|_a = p\}$$

and $\{w \in A^* \mid |w|_a = p\} = [(A-\{a\})^*a]^p(A-\{a\})^*$. It follows that $\{w \in A^* \mid |w|_a \geq n\} \in C$. Consequently each congruence class belongs to C_1 and thus $A^*\mathcal{V} \subseteq C$. The opposite inclusion follows from the facts that $A^*\mathcal{V}$ is closed under boolean operations and that $M(B^*a_1B^*\cdots B^*a_kB^*)$ is a nilpotent monoid.

Finally, let us point out the connection between the present theorem and a result of I. Simon concerning the variety of $\mathcal{J}$-trivial monoids ([SIM] and Section VIII.9 of [EIL]). It was shown in the course of the proof that every member M of $\underline{V}$ is $\mathcal{R}$-trivial, however it is clear that the same argument could have been used to show that M is $\mathcal{L}$-trivial as well. Thus every member of $\underline{V}$ is in fact $\mathcal{J}$-trivial, so $\underline{V} \subseteq \underline{J}$, where $\underline{J}$ is the variety of all finite $\mathcal{J}$-trivial monoids. Let A^* be the free monoid on a finite alphabet A and let k be a positive integer. A congruence $\sim_k$ on A^* is defined as follows: $w_1 \sim_k w_2$ if and only if w_1 and w_2 have the same subwords of length $\leq k$. That is, each factorization of w_1 of the form

$$w_1 = u_0 a_1 u_1 \cdots a_p u_p$$

where $u_i \in A^*$, $a_i \in A$ and $p \leq k$ implies the existence of a factorization

$$w_2 = v_0 a_1 v_1 \cdots a_p v_p$$

of w_2, and vice-versa. (The word $u = a_1 \cdots a_p$ is said to be a subword of w_1 and w_2.) Simon's theorem says that a recognizable language $L \subseteq A^*$ has its syntactic monoid in $\underline{J}$ if and only if L is a union of $\sim_k$-classes for some k. The inclusion $\underline{V} \subseteq \underline{J}$ is reflected in the relation between the congruences $\approx$ and $\sim$. Indeed, the reader can prove that for every $n > 0$, the congruence $\approx_n$ is refined by the congruence $\sim_k$, where $k = (n-1)|A| + 1$.

4. Many thanks to Jean-Eric Pin, who persuaded me to write this article, and who supplied a crucial step in the proof.

REFERENCES

[EIL] Samuel Eilenberg, Automata, Languages and Machines, vol. B, Academic Press, New York, 1976.

[LAL] Gerard Lallement, Semigroups and Combinatorial Applications, Wiley-Interscience, New York, 1979.

[PIN] Jean-Eric Pin, Variétés de Langages et Variétés de Semigroupes, These d'Etat, Universite de Paris VI, 1981.

[SIM] Imre Simon, Piecewise Testable Events, Proc. 2nd GI Conf., Lecture Notes in Computer Science, vol. 33, Springer-Verlag, 1975, 214-222.

Reed College
Portland, Oregon USA

Received May 20, 1981 and October 20, 1981 in final form.