INFORMATION AND CONTROL 42, 148–156 (1979)

Star-Free Regular Sets of ω-Sequences

WOLFGANG THOMAS

Mathematisches Institut der Universität, 7800 Freiburg, West Germany

The sets of ω-sequences over a finite alphabet which are definable in an appropriate first-order language are characterized in terms of star-free regular sets of words. This settles a problem of Ladner (1977), *Inform. Contr.* **33**, 281–303.

0. INTRODUCTION AND STATEMENT OF RESULTS

An interesting subclass of the class of regular sets of words was introduced by McNaughton and Papert (1971), namely the class of star-free (regular) sets. Given a finite alphabet Σ, the class of *star-free* sets over Σ consists of all word-sets which can be constructed from the finite subsets of Σ^* by the boolean operations (union, intersection, and complement w.r.t. Σ^*) and the product operation

$$U \cdot V = \{uv \in \Sigma^* \mid u \in U,\ v \in V\}.$$

If also the star-operation

$$U^* = \{u_0 \cdots u_{n-1} \in \Sigma^* \mid n \geqslant 0,\ u_i \in U\}$$

is allowed, one gets the class of *regular* sets over Σ. For technical reasons, we shall denote by SF_Σ (resp. REG_Σ) the star-free (resp. regular) subsets of Σ^+ (i.e. Σ^* without the empty word).

McNaughton and Papert (1971) found several characterizations of SF_Σ; among these an "equivalence" between SF_Σ and the first-order theory of finite linear orders. Connecting this with earlier work of Büchi (1960) concerning regular sets and the (weak) monadic second order theory of finite linear orders, this equivalence was worked out by Ladner (1977) in an appealing way.

To formulate these results we need some terminology. Throughout the paper we consider a fixed finite alphabet of the form $\Sigma = \Sigma_n = \{0, 1\}^n$ where $n \geqslant 1$. Every word $w = a_0 \cdots a_m$ of Σ^+ determines a finite structure $\mathfrak{M}_w = (M, <, P_1, \ldots, P_n)$, where $P_i \subset M$, in the following way: $M = \{0, \ldots, m\}$, $<$ is the usual ordering on M, and we have $k \in P_i$ iff the ith component of a_k is 1. Similarly, a sequence $\alpha \in \Sigma^\omega$ determines a structure $\mathfrak{M}_\alpha = (\omega, <, P_1, \ldots, P_n)$, where ω is the set of natural numbers and $P_i \subset \omega$. The first-order language appropriate for Σ, which we denote by $L_1(\Sigma)$, has as nonlogical constants the binary relation

0019-9958/79/080148–09$02.00/0

symbol $<$ and unary relation symbols $\mathbf{P}_1, \ldots, \mathbf{P}_n$, and is built up from these in the usual way. Similarly, the monadic second order language $L_2(\Sigma)$ is defined. For a sentence φ of $L_1(\Sigma)$ (or of $L_2(\Sigma)$) let Mod φ consist of all words $w \in \Sigma^+$ such that $\mathfrak{M}_w \models \varphi$. $\mathrm{Mod}^\omega\, \varphi$ denotes the set of all $\alpha \in \Sigma^\omega$ such that $\mathfrak{M}_\alpha \models \varphi$. (Here it is understood that the monadic second order variables range over arbitrary subsets of ω.) If $W = \mathrm{Mod}\, \varphi$ for some $\varphi \in L_1(\Sigma)$ (resp. $L_2(\Sigma)$), we say that W is first-order (resp. monadic-second-order) definable. Similarly for a set $A \subset \Sigma^\omega$. The following characterization of the monadic-second-order definable sets of words is a result of Büchi (1960) and Ladner (1977):

THEOREM 0.1. *A set $W \subset \Sigma^+$ is monadic-second-order definable iff it belongs to* REG_Σ.

The corresponding theorem for first-order definability was proved in McNaughton/Papert (1971) and Ladner (1977):

THEOREM 0.2. *A set $W \subset \Sigma^+$ is first-order definable iff it belongs to SF_Σ.*

Büchi (1962) and McNaughton (1966) proved an analogue of 0.1 for ω-sequences. For this, define the class $\mathrm{REG}_\Sigma{}^\omega$ of *ω-regular* sets over Σ to contain all subsets A of Σ^ω for which there are m and regular sets $U_1, \ldots, U_m, V_1, \ldots, V_m$ over Σ with $A = \bigcup_{i=1}^m U_i \cdot V_i^\omega$. (For $V \subset \Sigma^*$, V^ω contains all ω-sequences $v_0 v_1 v_2 \cdots$ where $v_i \in V$.)

Then we have, by Büchi (1962),

THEOREM 0.3. *A set $A \subset \Sigma^\omega$ is monadic-second-order definable iff it belongs to* $\mathrm{REG}_\Sigma{}^\omega$.

As Ladner remarks, the corresponding theorem for first-order definability fails: Take $A = (\{0\} \cdot \{0\} \cup \{1\})^\omega$ (a set of ω-sequences having between any two letters 1 an even number of 0's). Then A is of the form $\bigcup_{i=1}^m U_i \cdot V_i^\omega$ where the U_i, V_i are star-free regular, but not first-order definable.

So Ladner asked for an "algebraic" characterization of the first-order definable sets of ω-sequences. In this paper we obtain such a characterization.

If $W \subset \Sigma^*$, let

$$\lim W = \{\alpha \in \Sigma^\omega \mid \text{infinitely many initial segments of } \alpha \text{ are in } W\}.$$

Choueka (1974) has shown (see also McNaughton (1966, Lemma 2)) that in 0.3 we can replace the condition that A belongs to $\mathrm{REG}_\Sigma{}^\omega$ by the following equivalent one:

$$\begin{array}{l} \text{there are } m \text{ and regular sets } U_1, \ldots, U_m, V_1, \ldots, V_m \\ \text{over } \Sigma \text{ such that } A = \bigcup_{i=1}^m U_i \cdot \lim V_i . \end{array} \qquad (*)$$

Now let the class $SF_\Sigma{}^\omega$ of ω-sequences be defined as follows: A set $A \subset \Sigma^\omega$ is in $SF_\Sigma{}^\omega$ iff it can be written in the form $\bigcup_{i=1}^m U_i \cdot \lim V_i$ where the U_i, V_i are in SF_Σ. We shall show that a set $A \subset \Sigma^\omega$ is first-order definable iff it belongs to $SF_\Sigma{}^\omega$. Moreover, our method of proof also yields the modified form of 0.3 (with (*) instead of the condition that A is ω-regular). Ladner (1977) also suggests a class of "ω-star-free Σ-regular sets" of ω-sequences over Σ, namely the smallest class containing $\varnothing$ and closed under union, complement, and product with a star-free set of words on the left; it is shown there that any such set is indeed first-order definable. Using the characterization mentioned above, we can also show the converse; hence Ladner's class coincides with $SF_\Sigma{}^\omega$. Again a corresponding result will hold for the ω-regular sets.

We assume familiarity with the model-theoretic methods developed by Ehrenfeucht and Fraissé. We use a variant of these techniques due mainly to Läuchli (1966) and Shelah (1975). Apart from the results presented in this paper the proofs have also applications in logic, for instance concerning the decidability of certain subsystems of first-order arithmetic. These matters are treated in Thomas (1979b).

1. Preliminaries

In this section we present without proof some facts from first order logic to be used later on. Let us fix the language $L_1(\Sigma_n)$ with the nonlogical constants $\mathbf{<}, \mathbf{P}_1, \ldots, \mathbf{P}_n$. Thus formulas are always $L_1(\Sigma_n)$-formulas. For the variables $v_0, v_1, v_2, \ldots$ of this language we also write $x, y, z, \ldots$. Structures will always be $\{\mathbf{<}, \mathbf{P}_1, \ldots, \mathbf{P}_n\}$-structures. The letter $\mathfrak{M}$ will be reserved for structures of the form $\mathfrak{M} = (\omega, <, P_1, \ldots, P_n)$.

The quantifier-rank $\mathrm{qr}(\varphi)$ of a formula φ is defined recursively by

$$\mathrm{qr}(\varphi) = 0, \text{ if } \varphi \text{ is atomic (i.e. } x \mathbf{<} y \text{ or } \mathbf{P}_i x \text{ or } x = y),$$

$$\mathrm{qr}(\neg\varphi) = \mathrm{qr}(\varphi),$$

$$\mathrm{qr}(\varphi \vee \psi) = \mathrm{qr}(\varphi \wedge \psi) = \mathrm{qr}(\varphi \to \psi) = \max\{\mathrm{qr}(\varphi), \mathrm{qr}(\psi)\},$$

$$\mathrm{qr}(\exists x\varphi) = \mathrm{qr}(\forall x\varphi) = \mathrm{qr}(\varphi) + 1.$$

Let $\mathfrak{M} = (\omega, <, P_1, \ldots, P_n)$ be given. We shall define for $m \geqslant 0$ and every $k, l \in \omega$ with $k \leqslant l$ the *m-type* $T_m^{\mathfrak{M}}[k, l]$ of the segment $[k, l] := \{i \in \omega \mid k \leqslant i \leqslant l\}$. $T_m^{\mathfrak{M}}[k, l]$ will be a finite object, and from it one will be able to determine effectively for any φ with $\mathrm{qr}(\varphi) \leqslant m$ whether the substructure of $\mathfrak{M}$ with domain $[k, l]$ satisfies φ. Similarly $T_m^{\mathfrak{M}}$ will be defined such that from $T_m^{\mathfrak{M}}$ one can determine whether $\mathfrak{M} \models \varphi$.

DEFINITION 1.1. For $\mathfrak{M} = (\omega, <, P_1, ..., P_n)$, $k, l \in \omega$, $k \leqslant l$, and a sequence $k_0, ..., k_{r-1}$ of elements of $[k, l]$, where $r \in \omega$, let

$$T_0^{\mathfrak{M}}[k, l](k_0, ..., k_{r-1}) = \{\varphi(v_0, ..., v_{r-1}) \mid \varphi \text{ is an atomic formula } v_i = v_j \text{ or } v_i < v_j \text{ or } \mathbf{P}v_i, \text{ where } i, j < r, \text{ such that } \mathfrak{M} \models \varphi(k_0, ..., k_{r-1})\}$$

and

$$T_{m+1}^{\mathfrak{M}}[k, l](k_0, ..., k_{r-1}) = \{T_m^{\mathfrak{M}}[k, l](k_0, ..., k_{r-1}, k_r) \mid k_r \in [k, l]\}.$$

Let $T_m^{\mathfrak{M}}[k, l] = T_m^{\mathfrak{M}}[k, l]\Lambda$, where Λ is the empty sequence. $T_m^{\mathfrak{M}}$ is defined in the same way, taking ω instead of the segment $[k, l]$.

In the following, let $\mathfrak{T}_m$ be the (finite!) set of all formally possible m-types. Let us summarize some properties of $T_m^{\mathfrak{M}}[k, l]$ and $T_m^{\mathfrak{M}}$ (the proofs are standard, by induction on m; for (b_2), (c_2) using Ehrenfeucht-games):

LEMMA 1.2. (a) *For $\mathfrak{M} = (\omega, <, P_1, ..., P_m)$, $m \geqslant 0$, and $k, l \in \omega$, $k \leqslant l$: $T_m^{\mathfrak{M}}[k, l]$ and $T_m^{\mathfrak{M}}$ are finite objects, and for any φ with $\mathrm{qr}(\varphi) \leqslant m$, one can determine effectively from $T_m^{\mathfrak{M}}[k, l]$ (resp. $T_m^{\mathfrak{M}}$) whether φ holds in the substructure of $\mathfrak{M}$ with domain $[k, l]$ (resp. in $\mathfrak{M}$).*

(b₁) *For any $\tau \in \mathfrak{T}_m$ there is a bounded formula $\varphi_\tau(x, y)$[1] such that for all $\mathfrak{M}$ as above and $k, l \in \omega$: $\mathfrak{M} \models \varphi_\tau(k, l)$ iff $T_m^{\mathfrak{M}}[k, l] = \tau$. We also write $\varphi_\tau(x, y)$ as "$T_m[x, y] = \tau$".*

(b₂) *For any bounded formula $\psi(x, y)$ with $\mathrm{qr}(\psi) \leqslant m$ there is a set $T_\psi \subset \mathfrak{T}_m$ (effectively obtainable from ψ) such that for all $\mathfrak{M}$ as above and all $k, l \in \omega$ with $k \leqslant l$:*

$$\mathfrak{M} \models \psi(k, l) \quad \text{iff} \quad \mathfrak{M} \models \bigvee_{\tau \in T_\psi} \varphi_\tau(k, l).$$

In fact, we can define T_ψ by: $\tau \in T_\psi$ iff there is a structure $\mathfrak{M}$ and $k, l \in \omega$ with $T_m^{\mathfrak{M}}[k, l] = \tau$ and $\mathfrak{M} \models \psi(k, l)$.

(c₁) *For any $\tau \in \mathfrak{T}_m$ there is a sentence φ_τ such that for all $\mathfrak{M}$ as above: $\mathfrak{M} \models \varphi_\tau$ iff $T_m^{\mathfrak{M}} = \tau$.*

(c₂) *For any sentence ψ with $\mathrm{qr}(\psi) \leqslant m$ there is a set $T_\psi \subset \mathfrak{T}_m$ (effectively obtainable from ψ) such that for all $\mathfrak{M}$ as above:*

$$\mathfrak{M} \models \psi \quad \text{iff} \quad \mathfrak{M} \models \bigvee_{\tau \in T_\psi} \varphi_\tau .$$

We have $\tau \in T_\psi$ iff there is a structure $\mathfrak{M}$ with $T_m^{\mathfrak{M}} = \tau$ and $\mathfrak{M} \models \psi$.

[1] We call a formula $\varphi(x, y)$ *bounded* if in φ each quantifier is relativized to the segment $[x, y]$.

Remark 1.3. If two m-types τ_1 and τ_2 are given and we have $T_m{}^{\mathfrak{M}}[k_0, k_1] = \tau_1$ and $T_m{}^{\mathfrak{M}}[k_1 + 1, k_2] = \tau_2$, then the m-type $T_m{}^{\mathfrak{M}}[k_0, k_2]$ depends only on τ_1 and τ_2 and can be found effectively from τ_1 and τ_2. Thus we can introduce an effectively computable addition of m-types and write e.g. $T_m{}^{\mathfrak{M}}[k_0, k_2] = \tau_1 + \tau_2$. Similarly, for given $\mathfrak{M} = (\omega, <, P_1, \ldots, P_n)$ we may write $T_m{}^{\mathfrak{M}} = \tau_1 + \sum_\omega \tau_2$, if there are numbers $n_0, n_1, n_2, \ldots$ with $n_0 < n_1 < \cdots$ and $T_m{}^{\mathfrak{M}}[0, n_0] = \tau_1$, $T_m{}^{\mathfrak{M}}[n_i + 1, n_{i+1}] = \tau_2$ for $i \geqslant 0$. Also this m-type can be found effectively from τ_1 and τ_2.

2. Characterizing First-Order Definable Sets of ω-Sequences

As a preparation we have to restate Theorem 0.2 in a modified form. Given a structure $\mathfrak{M} = (\omega, <, P_1, \ldots, P_n)$, every segment $[k, l]$ determines a word in Σ^+; thus, referring to a set $W \subset \Sigma^+$, we may say that "the segment $[k, l]$ of $\mathfrak{M}$ belongs to W". Now it should be clear that from 0.2 we have

Theorem 0.2′. *A set $W \subset \Sigma^+$ is in SF_Σ iff there is a bounded formula $\varphi(x, y)$ of $L_1(\Sigma)$ such that for all $\mathfrak{M} = (\omega, <, P_1, \ldots, P_n)$ and all $k, l \in \omega$ with $k \leqslant l$ the segment $[k, l]$ of $\mathfrak{M}$ belongs to W iff $\mathfrak{M} \models \varphi(k, l)$.*

We now can proceed to the proof of the main result:

Theorem 2.1. *A set $A \subset \Sigma^\omega$ is first-order definable iff it belongs to $SF_\Sigma{}^\omega$ (i.e. iff it can be written in the form $\bigcup_{i=1}^m U_i \cdot \lim V_i$ where $U_1, \ldots, U_m, V_1, \ldots, V_m$ are in SF_Σ).*

The direction from right to left is easy, using 0.2′: Given $U_1, \ldots, U_m$, $V_1, \ldots, V_m$ in SF_Σ, we may choose bounded formulas $\varphi_1(x, y), \ldots, \varphi_m(x, y)$, $\psi_1(x, y), \ldots, \psi_m(x, y)$ which define the U_i and V_i as in 0.2′. Then a sentence defining $\bigcup_{i=1}^m U_i \cdot \lim V_i$ is

$$\bigvee_{i=1}^{m} \exists x(\varphi_i(0, x) \wedge \forall y\, \exists z > y\, \psi_i(x + 1, z)). \qquad (*)$$

(Here the term 0 is used to denote the minimal element of a structure, and $x + 1$ to denote the immediate $<$-successor of x. This gives sense since we are are only interested in structures over ω with the usual ordering.)

The other half of the theorem requires rewriting an arbitrary sentence φ of $L_1(\Sigma)$ in the form $(*)$ where the $\varphi_i(x, y)$ and $\psi_i(x, y)$ have to be bounded. For this, we consider, given $\mathfrak{M} = (\omega, <, P_1, \ldots, P_n)$ and $m \geqslant 0$, the following binary relation $\sim$ over ω (short for $\sim_m^{\mathfrak{M}}$):

$$k \sim l \text{ iff there is a } k' > k, l \text{ such that } T_m{}^{\mathfrak{M}}[k, k'] = T_m{}^{\mathfrak{M}}[l, k']. \qquad (+)$$

We say "k and l merge at k'" if (+) holds and call the minimal such k' (if it exists) "the minimal merging point of k, l". $\sim$ is an equivalence relation with only finitely many equivalence classes (for transitivity use the fact that if k and l merge at k' and $k'' > k'$, then k and l merge at k''). An automaton-theoretic version of $\sim$ has been used in McNaughton (1966). It seems that the "logical" version was first considered in Shelah (1975).

For $\sigma, \tau \in \mathfrak{T}_m$ let $\varphi_{(\sigma,\tau)}$ be the sentence

$$\varphi_{(\sigma,\tau)} = \exists x(\text{“}T_m[0, x] = \sigma\text{”} \wedge \forall y\, \exists z > y \\ (\text{“}x + 1 \sim z + 1\text{”} \wedge \text{“}T_m[x + 1, z] = \tau\text{”})),$$

where the parts in quotation marks are formulated according to 1.2b$_1$). The key point in proving 2.1 is the following

LEMMA 2.2. *For any structure* $\mathfrak{M} = (\omega, <, P_1, \ldots, P_n)$ *and* $m \geqslant 0$:

(a) *There is a pair* $(\sigma, \tau) \in \mathfrak{T}_m \times \mathfrak{T}_m$ *such that* $\mathfrak{M} \models \varphi_{(\sigma,\tau)}$.

(b) *If* $\mathfrak{M} \models \varphi_{(\sigma,\tau)}$, *then* $T_m^{\mathfrak{M}} = \sigma + \Sigma^\omega\, \tau$.

Proof. (a) Since there are only finitely many equivalence classes of $\sim$, there is an infinite such class, say I. Let $n_0 + 1 \in I$ and $\sigma := T_m^{\mathfrak{M}}[0, n_0]$. Consider for any $\tau' \in \mathfrak{T}_m$ the set $I_{\tau'} = \{k \in I \mid T_m^{\mathfrak{M}}[n_0 + 1, k] = \tau'\}$. Some set $I_{\tau'}$ must be infinite. Let τ be such a τ'. Then $\mathfrak{M} \models \varphi_{(\sigma,\tau)}$.

(b) Assuming $\mathfrak{M} \models \varphi_{(\sigma,\tau)}$, we can choose $n_0 \in \omega$ such that $T_m^{\mathfrak{M}}[0, n_0] = \sigma$ and $I_\tau = \{n \mid n_0 + 1 \sim n + 1$ and $T_m^{\mathfrak{M}}[n_0 + 1, n] = \tau\}$ is infinite. We now define the required sequence $n_0, n_1, n_2, \ldots$ (where n_0 is chosen as above) in the following way: If $n_0, \ldots, n_i$ are defined (with $n_0 + 1 \sim n_j + 1$ for $j \leqslant i$) let n_{i+1} be the smallest $n > n_i$ such that $n_0 + 1 \sim n + 1$ and for all $j \leqslant i$ $T_m^{\mathfrak{M}}[n_0 + 1, n] = T_m^{\mathfrak{M}}[n_j + 1, n] = \tau$. By infinity of I_τ and construction of $n_0, \ldots, n_i$ such an n_{i+1} exists. We have $T_m^{\mathfrak{M}}[n_i + 1, n_{i+1}] = \tau$ for $i \geqslant 0$; hence $T_m^{\mathfrak{M}} = \sigma + \sum_\omega \tau$.

By Lemma 2.2, the m-type of model $\mathfrak{M} = (\omega, <, P_1, \ldots, P_n)$ is determined by some formula $\varphi_{(\sigma,\tau)}$. Combining this with 1.2c$_2$), we obtain

LEMMA 2.3. *For any sentence* φ *with* $\mathrm{qr}(\varphi) \leqslant m$ *there is a finite disjunction, namely*

$$\delta_\varphi = \bigvee\{\varphi_{(\sigma,\tau)} \mid (\sigma, \tau) \in \mathfrak{T}_m \times \mathfrak{T}_m,\ \textit{there is a model}\ \mathfrak{M} = (\omega, <, P_1, \ldots, P_n) \\ \textit{such that}\ \mathfrak{M} \models \varphi_{(\sigma,\tau)}\ \textit{and}\ \varphi_{(\sigma,\tau)} \models \varphi\},$$

such that for all $\mathfrak{M} = (\omega, <, P_1, \ldots, P_n)$:

$$\mathfrak{M} \models \varphi \quad \textit{iff} \quad \mathfrak{M} \models \delta_\varphi.$$

Remark 2.4. We can find δ_φ from φ in an effective way: Call $(\sigma, \tau) \in \mathfrak{T}_m \times \mathfrak{T}_m$ *suitable* if σ, τ are finitely satisfiable and $\tau + \tau = \tau$. This condition can be checked effectively (use 1.3!). From 2.2 it follows that there is a model $\mathfrak{M} = (\omega, <, P_1, \ldots, P_n)$ satisfying $\varphi_{(\sigma,\tau)}$ iff (σ, τ) is suitable. Since for φ with $\mathrm{qr}(\varphi) \leqslant m$ and suitable (σ, τ) we have either $\varphi_{(\sigma,\tau)} \models \varphi$ or $\varphi_{(\sigma,\tau)} \models \neg\varphi$ ($\varphi_{(\sigma,\tau)}$ determines an m-type!), we see that given φ the set of suitable (σ, τ) with $\varphi_{(\sigma,\tau)} \models \varphi$ is found effectively. These (σ, τ) yield the desired disjunction.

Now we conclude the proof of 2.1. In view of 2.3, it is only to be verified that any of the formulas $\varphi_{(\sigma,\tau)}$ can be written in the form $\exists x(\varphi(\sigma, x) \wedge \forall y \exists z > y\ \psi(x+1, z)$ where φ and ψ are bounded. In the formula $\varphi_{(\sigma,\tau)}$

$$\exists x(\text{“}T_m[0, x] = \sigma\text{”} \wedge \forall y \exists z > y(\text{“}x + 1 \sim z + 1\text{”} \wedge \text{“}T_m[x + 1, z] = \tau\text{”}))$$

the only critical point is the unbounded quantifier hidden in “$x + 1 \sim z + 1$”. But clearly $\varphi_{(\sigma,\tau)}$ is equivalent with

$$\exists x(\text{“}T_m[0, x] = \sigma\text{”} \wedge \forall y \exists u > y$$

“u is the minimal merging point of $x + 1$ and an element $z + 1$ between $x + 1$ and u such that $T_m[x + 1, z] = \tau$”)

The part in quotation marks can be expressed by a formula $\psi(x + 1, u)$ which is bounded. Hence 2.1 is proved.

We mention without proof that a completely analogous argument leads to the characterization of monadic-second-order definable sets of ω-sequences as those of the form $\bigcup_{i=1}^m U_i \cdot \lim V_i$ where the U_i and V_i are regular. For the modifications which are necessary to obtain the analogue of 2.2 and of 2.4, see §2 of Shelah (1975) and Thomas (1979a).

3. Equivalence with Ladner's Definition

Let $LSF_\Sigma{}^\omega$ be the class (introduced in Ladner (1977)) defined as the closure of the empty set of ω-sequences over Σ under the operations of union, complement and concatenation with a star-free set of words on the left.

Theorem 3.1. $SF_\Sigma{}^\omega = LSF_\Sigma{}^\omega$.

Proof. It is established by induction over $LSF_\Sigma{}^\omega$ that any $A \in LSF_\Sigma{}^\omega$ is first-order definable (see Ladner (1977, Theorem 5.6)). Hence, by 2.1, $LSF_\Sigma{}^\omega \subset SF_\Sigma{}^\omega$. For the converse we have to show that for any $W \in SF_\Sigma$ the set $\lim W$ belongs to $LSF_\Sigma{}^\omega$. (Then it is immediate that an arbitrary set of the form $\bigcup_{i=1}^m U_i \cdot \lim V_i$, where the U_i, V_i are in SF_Σ, belongs to $LSF_\Sigma{}^\omega$.) Since

LSF_{Σ^ω} is closed under complement, let us consider instead of $\lim W$ the set $\Sigma^\omega - \lim W$. We shall find star-free sets $U_1, \ldots, U_k$, $V_1, \ldots, V_k$ such that

$$\Sigma^\omega - \lim W = \bigcup_{i=1}^{k} (U_i \cdot (V_i \cdot \Sigma^\omega)^c), \tag{*}$$

where ^{c} denotes complement. Since $\Sigma^\omega = \varnothing^c$, this will show that $\Sigma^\omega - \lim W$ and thus $\lim W$ belongs to LSF_{Σ^ω}. In order to find the representation (*), it will suffice to express the condition $\alpha \in \Sigma^\omega - \lim W$ in the form

$$\mathfrak{M}_\alpha \models \bigvee_{i=1}^{k} \exists x(\varphi_i(0, x) \wedge \forall y > x \neg \psi_i(x+1, y)) \tag{+}$$

where the φ_i and ψ_i are bounded formulas of $L_1(\Sigma)$. This is achieved as follows:

$$\begin{aligned} &\alpha \in \Sigma^\omega - \lim W \\ &\text{iff} \quad \exists m\, \forall n > m\ \alpha(0) \cdots \alpha(n) \notin W \\ &\text{iff} \quad \mathfrak{M}_\alpha \models \exists x\, \forall y > x\ \psi(0, y) \end{aligned}$$

(where ψ defines the star-free set $\Sigma^+ - W$)

$$\text{iff} \quad \mathfrak{M}_\alpha \models \exists x\, \forall y > x \bigvee_{\tau \in T_\psi} \text{“}T_m[0, y] = \tau\text{”}$$

(where T_ψ is chosen as in 1.2b$_2$))

$$\begin{aligned} \text{iff} \quad \mathfrak{M}_\alpha \models \bigvee_{\tau_1 \in T_1} &\exists x(\text{“}T_m[0, x] = \tau_1\text{”} \wedge \forall y > x \\ &\bigvee_{\tau_2 \in T(\tau_1)} T_m[x+1, y] = \tau_2\text{”}) \end{aligned}$$

(where $\tau_1 \in T_1$ iff there is τ_2 with $\tau_1 + \tau_2 \in T_\psi$, and $\tau_2 \in T(\tau_1)$ iff $\tau_1 + \tau_2 \in T_\psi$). The last sentence can be written as required in (+); hence 3.1 is proved.

As an analysis of the proof of 3.1 together with 1.2 and 1.3 shows, the transition from “LSF_{Σ^ω}-expressions” to corresponding “SF_{Σ^ω}-expressions” or to corresponding first-order formulas and vice versa is effective.

The proof of 3.1 carries over (with “regular” instead of “star-free”) to yield the result that $\mathrm{REG}_{\Sigma^\omega}$ coincides with the closure of the empty set of ω-sequences over Σ under union, complement, and concatenation with a regular set on the left.

RECEIVED: November 3, 1978

Note added in proof. After submission of this paper, Prof. Ladner informed the author that he had independently obtained the result that every first-order definable set $A \subset \Sigma^\omega$ belongs to LSF_{Σ^ω}.

References

Büchi, J. R. (1960), Weak second-order arithmetic and finite automata, *Z. Math. Logik Grundlagen Math.* **6**, 66–92.

Büchi, J. R. (1962), On a decision method in restricted second order arithmetic, *in* "Methodology and Philosophy of Science, Proceedings, 1960 Stanford Intern. Congr." (E. Nagel *et al.*, Eds.), pp. 1–11, Stanford Univ. Press, Stanford, Calif., 1962.

Choueka, Y. (1974), Theories of automata on ω-tapes: A simplified approach, *J. Comput. System Sci.* **8**, 117–141.

Ladner, R. E. (1977), Application of model-theoretic games to discrete linear orders and finite automata, *Inform. Contr.* **33**, 281–303.

Läuchli, H. (1966), A decision procedure for the weak second order theory of linear order, *in* "Contributions to Mathematical Logic, Proc. Logic Coll. Hannover 1966" (H. A. Schmidt *et al.*, Eds.), pp. 189–197, North-Holland, Amsterdam, 1968.

McNaughton, R. (1966), Testing and generating infinite sequences by a finite automaton, *Inform. Contr.* **9**, 521–530.

McNaughton, R., and Papert, S. (1971), "Counter-Free Automata," MIT Press, Cambridge, Mass.

Shelah, S. (1975), The monadic theory of order, *Ann. of Math.* **102**, 379–419.

Thomas, W. (1979a), On the bounded monadic theory of well-ordered structures, *J. Symbolic Logic*, to appear.

Thomas, W. (1979b), Relations of finite valency over $(\omega, <)$, in preparation.