一类ω－正则语言*

苏锦祥

(郑州大学)

A CLASS ω — REGULAR LANGUAGES

Su Jinxiang

(Zhengzhou University)

ABSTRACT

An ω—language is a set consisting of infinite-strings over some alphabet Σ, the ω—language accepted by some ω—finite state automation is called the ω—regular language.

Several sufficient conditions for an ω—language is an ω—regular language are given by author from the point of view of the set in [4]. In this paper, author gives still from the point of view of the set a sufficient condition for an ω—language is an ω—regular language, i.e., if L is an ω—convex language, such that $L = Adh(pref(L)) = Pref(L)Tail(L)$, then the L is an ω—regular language. Thus defined one subclass of the ω—regular languages class.

摘 要

ω—语言是由有穷字母表Σ上的某些无穷串组成的集合。被所谓的ω—有穷自动机接受的ω—语言称为ω—正则语言。在[4]中作者曾从集合的角度给出一ω—语言为ω—正则语言的几个充分条件。在本文作者仍从集合的角度给出一个ω—语言为ω—正则语言的充分条件，即若一ω—凸语言L满足$L = adh(pref(L)) = pref(L)tail(L)$，则L是一ω—正则语言。从而，确定了ω—正则语言类的一个子类。

R. McNaughton 首先提出了被有穷自动机识别的ω—语言的理论。R. Cohen 与 A. Gold 在[1]中综述了ω—有穷自动机的概念，ω—有穷自动机的五种识别方式

* 1989年8月27日收到。

以及在此基础上引进的ω—正则语言的概念。同时也给出了关于ω—正则语言的几个特征，但这些特征实质上或为有穷自动机或为右线性文法。而本文不从识别的角度也不从生成的角度来考查ω—语言的ω—正则性，而是从ω—语言本身的某些集合性质给出ω—语言是ω—正则语言的一个充分条件，从而给出ω—正则语言类的一个子类。

为了引述本文的结果，首先简要地叙述和引进一些有关的概念与记号。

用Σ表示一有穷字母表，由Σ中的字母组成的形如$\delta = a_1a_2\cdots a_n\cdots$的无穷序列称为Σ上的一个ω—字。用$\delta(n)$表示ω—字δ的前n个字母所组成的δ的前缀，即$\delta(n) = a_1\cdots a_n$, $n \geq 1$。用Σ^ω表示Σ上的一切ω—字组成的集合。Σ^ω的任一子集称为Σ上的一个ω—语言。

在[2]中通过字的嵌入定义了Σ^*上的一个半序关系。即，设Σ是有穷字母表，在Σ^*上定义二元关系"≤"，使对于$x, y \in \Sigma^*$, $x \leq y$当且仅当$x = x_1\cdots x_n$, $y = y_1x_1y_2\cdots x_ny_{n+1}$, $x_i \in \Sigma^*$, $i = 1,\cdots,n$; $y_i \in \Sigma^*$, $i = 1,\cdots n+1$。二元关系"≤"显然是集合Σ^*上的一个半序关系。

利用Σ^*上的半序关系"≤"，可定义Σ上的凸语言。设L是Σ上的一个语言，若$x, y \in L$且$x \leq z \leq y$, $z \in \Sigma^*$，则$z \in L$，就称L是Σ上的凸语言[2]。

设Σ是有穷字母表。我们在Σ^ω上定义二元关系"$\leq_\omega$"如下：对$\delta_1, \delta_2 \in \Sigma^\omega$, $\delta_1 \leq_\omega \delta_2$当且仅当$\delta = x_1x_2\cdots x_r\cdots$, $\delta_2 = y_1x_1y_2\cdots x_ny_{n+1}\cdots$, $x_i, y_i \in \Sigma^*$, $i = 1, 2, \cdots$。作者在定义了Σ^ω上的二元关系"$\leq_\omega$"的基础上，引进了ω—凸语言的概念。

定义1. 设L是Σ上的一个ω—语言，若对于$\delta_1, \delta_2 \in L$且$\delta_1 \leq_\omega \delta \leq_\omega \delta_2$, $\delta \in \Sigma^\omega$，则$\delta \in L$，就称L是Σ上的一个ω—凸语言。

定义2. 设$\delta \in \Sigma^\omega$，令

$$pref(\delta) = \{x \in \Sigma^* | \text{存在}\delta' \in \Sigma^\omega, \text{使} x\delta' = \delta\};$$

设$L \subseteq \Sigma^*$，令

$$pref(L) = \{x \in \Sigma^* | \text{存在} y \in \Sigma^*, \text{使} xy \in L\};$$

设$L \subseteq \Sigma^*$，令

$$pref(L) = \{x \in \Sigma^* | \text{存在}\delta' \in \Sigma^\omega, \text{使} x\delta' \in L\},$$

显然，若$L \subseteq \Sigma^\omega$，则$pref(L) = pref(pref(L))$。

定义3. 设$L \subseteq \Sigma^*$，称Σ上的ω—语言

$$\{\delta \in \Sigma^\omega (pref(\delta) \subseteq pref(L)\}$$

为语言L的附着[5]，记作$adh(L)$。

定义4. 一个确定的ω—有穷自动机是一个五元组

$$M = (K, \Sigma, \delta, q_0, F),$$

其中，K—状态的有穷非空集；Σ—输入字母表；$\delta : KX\Sigma \to K$—状态转换函数；q_0—初始状态；$F \subseteq 2^K$—指定状态集族。

设$\delta = a_1 a_2 \cdots a_n \cdots, a_i \in \Sigma, i = 1, 2, \cdots$，若状态的无穷序列$r: q_0, q_1, \cdots, q_n, \cdots$，使$q_{i+1} = \delta(q_i, a_{i+1}), i = 0, 1, \cdots$，则称无穷序列r是确定的$\omega$—有穷自动机M在$\omega$—字$\delta$上的运行轨迹，显然，如果状态序列$r: q_0, q_1, \cdots, q_n, \cdots$，是确定的$\omega$—有穷自动机$M$在$\omega$—字$\delta$上的运行轨迹，则它确定一个映射$f_r: N \to k$，其中，N表示全体自然数的集合，使$f_r(i) = q_{i-1}, i = 1, 2, \cdots$。兹令$Inf_r = \{q \in K | card(f_r^{-1}(q) \geq \omega\}$，即$Inf_r$表示在状态序列r中出现无穷多次的状态集合。

称一个Σ上的ω—字δ被确定的有穷自动机

$$M = (K, \Sigma, \delta, q_0, F)$$

接受，当且仅当M在δ上的运行轨迹r，使$Inf_r \in F$。被确定的ω—有穷自动机M接受的一切ω—字的集合记作$T(M)$，即

$$T(M) = \{\delta \in \Sigma^\omega | M\text{在}\delta\text{上的运引轨迹}r, \text{使}Inf_r \in F\}.$$

若对Σ上的ω—语言L，存在一个确定的ω—有穷自动机M，使$T(M) = L$，则说ω—语言L是一个ω—正则语言。

定理. 设L是Σ上的ω—凸语言，若$L = adh(pref(L)) = pref(L)tail(L)$，则L是$\Sigma$上的$\omega$—正则语言。

证明. 首先证明；若Σ上的ω—语言L满足$L = adh(pref(L))$且$pref(L)$是Σ上的正则集，则L是Σ上的一个ω—正则语言。

兹设$pref(L)$是Σ上的正则集，则存在一个确定的有穷自动机

$$M = (K, \Sigma, \delta, q_0, F),$$

使$T(M) = pref(L)$。令构造一个确定的ω—有穷自动机

$$M' = (K', \Sigma', \delta', q_0', F'),$$

其中$K' = K, \Sigma' = \Sigma, \delta' = \delta, q_0' = q_0, F' = 2^F - \{\phi\}$。

下面证明$T(M') = L$。设ω—字$\delta \in T(M')$，即M'在δ上的运行轨迹$r: q_0, q_1, \cdots, q_j, \cdots$所确定的映射$f_r$，使$Inf_r \in F'$。于是在状态序列$q_0, q_1, \cdots, q_j, \cdots$中出现无穷多次的状态均为有穷自动机M的终止状态。设在序列$q_0, q_1, \cdots, q_j, \cdots$中属于$Inf_r$的状态所组成的子列为$q_{i_1}, q_{i_2}, \cdots, q_{i_j} \cdots$。显然，对任意的正整数n，必存在正整数m与k，使$\delta(q_0, \delta(n+m)) = q_{i_k}$。但是，因为$q_{i_k} \in F$，所以知$\delta(n+m) \in pref(L)$，即存在$\delta' \in L$。使$\delta'(n+m) = \delta(n+m)$。当然$\delta(n) = \delta'(n)$，从而推得$\delta(n) \in pref(L)$。上述结果表明$pref(\delta) \subseteq pref(L)$，所以$\delta \in adh(pref(L))$。又因为$L = adh(pref(L))$，所以$\delta \in L$，即证明了$T(M') \subseteq L$。

反之，设$\delta \in L$，由于对任意的自然数n，$\delta(n) \in pref(L)$，则$\delta(q_0, \delta(n)) \in F$。于是，$M'$在$\delta$上的运行轨迹$r: q_0, q_1, \cdots, q_n, \cdots$，使$Inf_r \in F'$，所以$\delta \in T(M')$。这就证明了$L \subset T(M')$。从而，证明了$T(M') = L$。即证明了：若一个$\Sigma$上的$\omega$—语言L，

使得 $L = adh(pref(L))$ 且 $pref(L)$ 是 Σ 上的正则集，则 L 必是一个 Σ 上的 ω— 正则语言。

下面证明定理的结论。设 $x_1, x_2 \in pref(L)$ 且 $x_1 \leq y \leq x_2$, $y \in \Sigma^*$。因为 $x_1 \in pref(L)$, 所以必存在一个 $\delta' \in \Sigma^\omega$, 使 $x_1\delta' \in L$。显然，$\delta' \in tail(L)$。于是，由定理假设 $L = pref(L)tail(L)$ 知 $x_2\delta' \in L$。但是，$x_1\delta' \leq_\omega y\delta \leq_\omega x_2\delta'$, 从而，由 L 是 Σ 上的 ω—凸语言知 $y\delta' \in L$, 所以推得 $y \in pref(L)$。这就证明了 $pref(L)$ 是 Σ 上的一个凸语言。然而，在[3]中已证明了凸语言是正则集，因此，$pref(L)$ 是 Σ 上的一个正则集。综上，我们证明了满足定理条件的 ω— 语言是 Σ 上的 ω— 正则语言。

参考文献

[1] R. S. Cohen and A. Y. Gold, Theory of ω—Languages, I: Characterizations of ω—Context—free Languages, J. Comput. System Sci. 15 (1977).

[2] G. Thierrin, Convex Languages, Automata, Languages and Programming, North—HOLLAND PUBLISHING COMPANY, (1972).

[3] 苏锦祥 关于ω—语言的正则性，计算机学报，Vol. 7, No. 5 (1984)。

[4] 苏锦祥 ω—正则语言族的几个子类，计算机研究与发展，Vol. 22, No. 4 (1985)。

[5] I. Litovsky and E. Timmerman, On Generators of Rational ω—Power Languages, Theoret. Comput. Sci. 53 (1987).