

关于句法么半群为幂零的语言结构*

龙冬阳

(中山大学计算机科学系, 广州 510275)

摘 要 本文研究一个语言的句法么半群的幂零性, 给出了一个语言的句法么半群是幂零, 么半群的充要条件, 从而完全决定了这类语言的结构。

关键词 幂零么半群, 语言, 句法同余

分类号 TP301.2, O152.7

众所周知, 一个语言的句法么半群在语言代数学、字的组合、自动机与形式语言理论、码论等中起着非常重要的作用, 通过研究句法么半群, 刻划了许多语言类^[1~4], 特别是对许多正则(或可识)语言的真子类的刻划^[2, 5]引起了语言簇的广泛研究. 幂零半群、幂零么半群是很重要的半群类与么半群类, 它们是研究一般半群、么半群的一个重要概念和工具. 然而对其自身的结构, 研究较少, 鉴于以上原因, 本文讨论一个很有趣的问题: 当一个语言的句法么半群是幂零么半群时, 则语言本身有何特征?

本文是笔者以前工作^[6]的继续, 完整地回答了上述问题. 值得指出的是, [2] 和 [5] 都研究了语言的句法么半群的幂零性, 但其讨论的对象是有限半群或有限么半群的子类所形成的半群簇与么半群簇, 因此, 所得到的语言的句法么半群的刻划都是假定它们都是正则(或可识)语言的前提下进行的. 我们的讨论是任意语言(并不一定是正则)的句法么半群的幂零性. 例如, 从 [2] 中可知, 对于有限半群或有限么半群而言, 幂零元(或谐零)半群与么半群的概念与幂零半群及么半群的概念是重合的; 而对于无限半群或无限么半群而言, 幂零元(或谐零)与幂零是两个不同的概念, 幂零元半群(么半群)并不一定是幂零半群(么半群).

首先给出一些必要的概念与记号, 有关细节可参见文献 [2], [7].

设 A^* 是由字母表 A 生成的自由么半群, A^* 中元素与子集分别称为字母表 A 上的字与语言, 记 $A^+ = A^* \setminus \{1\}$, 表示字母表 A 生成的自由半群, 1 是 A^* 的么元, 称为空字. 对于任一 $w \in A^*$, 记 $|w|$ 为字 w 的长度, 即字 w 中包含字母表 A 中字母的个数. 设 $C \subseteq A^*$, 如果对于 C 中任意 $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m$, 由 $x_1 x_2 \dots x_n = y_1 \dots y_m$ 则必有 $n = m, x_i = y_i, i = 1, \dots, n$, 则称 C 是 A 上的一个码. 如果 $C \subseteq A^*$, $C \cap CA^+ = \Phi (C \cap A^+ C = \Phi)$, 则称 C 为

* 国家自然科学基金资助项目

收稿日期: 1995-03-10 龙冬阳, 男, 38 岁, 副教授

前(后)缀码;如果 C 既是一前缀码, 又是一后缀码, 则称 C 是一双缀码. $C \subseteq A^*$ 称为一个强码, 如果它满足下面 2 条: (i) $x, y_1y_2 \in C \Rightarrow y_1xy_2 \in C^*$; (ii) $x, y_1xy_2 \in C^* \Rightarrow y_1y_2 \in C^*$. 从 [7] 中可知, 每一个强码必是双缀码.

对任意语言 $L \subseteq A^*$, 记

$$\text{res}(L) = \{u \in A^* \mid A^*uA^* \cap L = \emptyset\}$$

称为 L 的剩余, 关于 L 决定 A^* 上一个同余关系 P_L 如下:

$$u \equiv v(P_L) \Leftrightarrow (\forall x, y \in A^*) xuy \in L \leftrightarrow xvy \in L$$

记 $\text{syn}(L) = A^*/P_L$, 称 $\text{syn}(L)$ 为语言 L 的句法么半群. 一个语言 L 称为正则(或可识)的, 当且仅当 L 的句法么半群是有限的. 记 $[w]$ 为包含字 w 的 P_L -类, 即 $[w] = \{v \in A^* \mid v \equiv w(P_L)\}$.

设 M 是一个带有零元 0 和么元 $1 (\neq 0)$ 的么半群, 如果关于任意 $x (\neq 1) \in M$, 必存在正整数 n_x 使得 $x^{n_x} = 0$, 则称 M 是幂零元(或谐零)的; 如果存在 n 使得 $(M \setminus \{1\})^n = \{0\}$, 则称 M 是幂零的. 幂零么半群必是幂零元(或谐零)的么半群, 反之一般不真.

下面给出一个语言 L 其句法么半群 $\text{syn}(L)$ 是幂零么半群时的结构. 因为 \emptyset 和 A^* 的句法么半群仅含一个元素, 所以, 我们的讨论总是假定语言 $L \neq \emptyset, A^*$.

定理 设 A 是一有限字母表, $L \subseteq A^*$ 是任意一个语言, 则 L 的句法么半群 $\text{Syn}(L)$ 是幂零么半群, 当且仅当存在 $C \subseteq A$ 及正整数 m, n_1, n_2, \dots, n_m 和 $a_{11}, a_{12}, \dots, a_{1n_1} \in (A \setminus C) \cup \{1\} (1 \leq i \leq m)$ 使得

$$L = \bigcup_{i=1}^m (C^* a_{i1} C^* a_{i2} C^* \dots C^* a_{in_i} C^*)$$

或者
$$L = A^* \setminus \left(\bigcup_{i=1}^m (C^* a_{i1} C^* a_{i2} C^* \dots C^* a_{in_i} C^*) \right).$$

为了证明定理, 先给出两个引理

引理 1^[6] 设 $L \subseteq A^*$. 则语言 L 的句法么半群 $\text{syn}(L)$ 是幂零么半群且空字 1 所在 P_L -类 $[1] = \{1\}$, 当且仅当 L 是有限语言或者 $A^* \setminus L$ 是有限语言.

引理 2^[6] 设 $L \subseteq A^*$, 空字 1 所在 P_L -类 $[1] = C^*$, 其中 C 为 A 上的某一个强码, 如果语言 L 的句法么半群 $\text{syn}(L)$ 是幂零元的么半群, 那么 $C \subseteq A$.

引理 3^[7] 设 $L \subseteq A^*$. 则存在 A 上强码 C , 使得 $[1] = C^*$, 其中 $[1]$ 表示空字 1 所在的 P_L -类.

定理的证明 对于任意语言 $L \subseteq A^*$, 因空字 1 所在的 P_L -类中的字的个数 $|[1]| \geq 1$. 如果 $|[1]| = 1$, 即 $[1] = \{1\}$, 由引理 1 知, $\text{syn}(L)$ 是幂零么半群当且仅当 L 是有限语言或者 $A^* \setminus L$ 是有限语言. 因此, 取 $C = \emptyset \subseteq A, C^* = \{1\}$, 易知结论显然成立. 下面不妨设 $|[1]| > 1$.

充分性 假设 $L = \bigcup_{i=1}^m (C^* a_{i1} C^* a_{i2} C^* \dots C^* a_{in_i} C^*) = C^* a_{11} C^* a_{12} C^* \dots C^* a_{1n_1} C^* \cup C^* a_{21} C^* a_{22} C^* \dots C^* a_{2n_2} C^* \cup \dots \cup C^* a_{m1} C^* a_{m2} C^* \dots C^* a_{mn_m} C^*$. 据 P_L 之定义, 易知对于所有 $u, v \in C^*, u \equiv v(P_L)$, 特别 $1 \equiv u(P_L)$. 这说明 $C^* \subseteq [1]$, 同在空字 1 所在的 P_L -类中. 如果 $C = A$, 则 $L = A^*, \text{syn}(L) = \{[1]\}$, 结论是平凡的. 当 $C \subseteq A$ 且 $C \neq \emptyset$ 时, 由 L 的构造, $\text{res}(L) = \{w \in A^* \mid A^*wA^* \cap L = \emptyset\} \neq \emptyset$. 事实上, 取一个充分长的字 u , 它以

$a_{11}a_{12}\cdots a_{1n_1}a_{21}a_{22}\cdots a_{2n_2}\cdots a_{m1}\cdots a_{mn_m}$ 为其子字, 这 u 满足 $A^*uA^* \cap L = \emptyset$. 于是 $\text{res}(L)$ 中的字是同余的, 且它们所在的 P_L -类就是 $\text{syn}(L)$ 的零元 0 . 即 $\text{syn}(L)$ 是带有零元 0 的么半群. 对于任意 $w \in A^* \setminus C^*$, 可以证明 $1 \not\equiv w(P_L)$, 因为取 $x_i = a_{i1}a_{i2}\cdots a_{in_i}$, $n_i = \max\{n_1, n_2, \dots, n_m\}$ 易知 $1x_i = x_i \in L$, 但 $wx_i \notin L$, 因此 $1 \not\equiv w(P_L)$, 这说明 $[1] = C^*$, 对于上面所取的 n_i , 若 $w \in A^*(A \setminus C)A^*(A \setminus C)A^*\cdots(A \setminus C)A^*$, 即 w 至少含有 $n_i + 1$ 个 $A \setminus C$ 中的字母, 易推得 $[w] = 0$, 即 $w \in \text{res}(L)$. 因此, 对于任意 $n_i + 1$ 个字 $w_1, w_2, \dots, w_{n_i+1} \in A^* \setminus C^*$, $w_1w_2\cdots w_{n_i+1} \in \text{res}(L)$, $[w_1w_2\cdots w_{n_i+1}] = 0$. 又 $[w] \neq [1]$ 当且仅当 $w \in A^* \setminus C^*$, 故任意 $[w_1], \dots, [w_{n_i+1}] \in \text{syn}(L) \setminus \{[1]\}$, 都有 $[w_1][w_2]\cdots[w_{n_i+1}] = [w_1w_2\cdots w_{n_i+1}] = 0$, 即 $\text{syn}(L)$ 是幂零么半群.

若 $L = A^* \setminus (\bigcup_{i=1}^m (C^*a_{i1}C^*a_{i2}C^*\cdots C^*a_{in_i}C^*))$. 因为 $\text{syn}(L) = \text{syn}(A^* \setminus L)$ 及前面的结论可知, $\text{syn}(L)$ 是幂零么半群.

必要性 假设 $\text{syn}(L)$ 是幂零么半群, 而且 $|[1]| > 1$. 但这时并没有假设 L 是正则语言, 即 $\text{syn}(L)$ 不一定有限.

据引理 2 和引理 3 知: 存在 $C \subseteq A$ 使得, $[1] = C^*$. 因 $\text{syn}(L)$ 是幂零的, 所以存在正整数 k , 使得 $(\text{syn}(L) \setminus \{[1]\})^k = \{0\}$, 故对于 $A^* \setminus C^*$ 中任意 k 个字 x_1, x_2, \dots, x_k 均有 $[x_1][x_2]\cdots[x_k] = [x_1x_2\cdots x_k] = 0$.

下面再分两种情况讨论:

(i) $0 \not\subseteq L$, 即 0 元所在的 P_L -类的字均不在 L 中. 因此, 对任意 $x_1, x_2, \dots, x_k \in A^* \setminus C^*$, $x_1x_2\cdots x_k \notin L$. 考查 $A^* = C^* \cup C^*(A \setminus C)C^* \cup C^*(A \setminus C)C^*(A \setminus C)C^* \cup C^*(A \setminus C)C^*(A \setminus C)C^*(A \setminus C)C^* \cup \dots$, 即 A^* 可表为 $C^*(A \setminus C)C^*(A \setminus C)C^*\cdots C^*(A \setminus C)C^*$ 这种形式的语言的无交并. 由 $x_1x_2\cdots x_k \notin L$, 那么对于任意一个长度大于 k 且该字中至少含有 $A \setminus C$ 中 k 个字母的字 x , 必有 $x \notin L$. 故 L 中的字最多含有 $A \setminus C$ 中的 $k-1$ 个字母, 因此 $L \subseteq C^* \cup C^*(A \setminus C)C^* \cup C^*(A \setminus C)C^*(A \setminus C)C^* \cup \dots \cup C^*(A \setminus C)C^*\cdots C^*(A \setminus C)C^*$. 假如存在 $u \in C^*$ 使得 $u \in L$, 由 $[1] = C^*$, 故 $C^* \subseteq L$. 假如存在 $u \in C^*(A \setminus C)C^*\cdots C^*(A \setminus C)C^*$, 即 u 含有 $A \setminus C$ 中 i 个字母 ($1 \leq i \leq k-1$) 使得 $u \in L$, 由 $u = u_1b_1u_2b_2\cdots u_ib_ib_{i+1}$, 其中 $u_1, \dots, u_{i+1} \in C^*$, $b_1, \dots, b_i \in A \setminus C$. 又 $u_i \equiv u_j(P_L)$, 所以 $u = u_1b_1u_2b_2\cdots u_ib_ib_{i+1} \equiv v_1b_1v_2b_2\cdots v_ib_iv_{i+1}(P_L)$, 其中 v_1, v_2, \dots, v_{i+1} 是 C^* 中任意 $i+1$ 个字. 再由 $u \in L$ 则 $v_1b_1v_2b_2\cdots v_ib_iv_{i+1} \in L$, 故 $C^*(A \setminus C)C^*(A \setminus C)\cdots(A \setminus C)C^* \subseteq L$; 这说明 L 包含在有限个形如 $C^*(A \setminus C)C^*\cdots C^*(A \setminus C)C^*$ 的语言的无交并中. 综合上面的分析, 存在正整数 m, n_1, n_2, \dots, n_m 及 $a_{i1}, a_{i2}, \dots, a_{in_i} \in (A \setminus C) \cup \{1\}$, $1 \leq i \leq m \leq k-1, 1 \leq n_i \leq k-1$ 使得

$$L = \bigcup_{i=1}^m (C^*a_{i1}C^*a_{i2}C^*\cdots C^*a_{in_i}C^*)$$

(ii) $0 \subseteq L$, 即 0 元所在的 P_L -类中每一个字均在 L 中, 因此 $0 \subseteq A^* \setminus L$. 据 $\text{syn}(L) = \text{syn}(A^* \setminus L)$ 以及(i)的讨论, 则存在正整数 m, n_1, \dots, n_m 及 $a_{i1}, a_{i2}, \dots, a_{in_i} \in (A \setminus C) \cup \{1\}$ 使得

$$A^* \setminus L = \bigcup_{i=1}^m (C^*a_{i1}C^*a_{i2}C^*\cdots C^*a_{in_i}C^*) \tag{7}$$

故 $L = A^* \setminus (\bigcup_{i=1}^m (C^*a_{i1}C^*a_{i2}C^*\cdots C^*a_{in_i}C^*))$

据(i)与(ii), 这就证明了定理的必要性.

从定理的必要性证明即可得:

推论 1 设 A 是一有限字母表, $L \subseteq A^*$ 是任意语言, 若 $\text{syn}(L)$ 是幂零么半群, 则 $\text{syn}(L)$ 必是有限的幂零么半群, 特别地, L 必是正则的语言.

据语言 $L \subseteq A^+$ 的句法半群 $\text{syn}(L) = A^+ / P_L$ 之定义^[2], 即刻可得

推论 2^[2] 设 A 是有限字母表, $L \subseteq A^+$ 是任意语言. 那么 L 的句法半群 $\text{syn}(L) = A^+ / P_L$ 是幂零半群当且仅当 L 有限或者 $A^* \setminus L$ 有限.

注 1 [2] 中定理 3.3 的证明中是先已假定了 $L \subseteq A^+$ 是正则(或可识)语言, 因为大前提是研究有限半群的子类形成的半群簇对应的语言簇, 而这里的语言簇均指可识语言的子类, 我们的证明完全不同于[2]中定理 3.3, 且结论更强.

注 2 给定有限字母表 A , 则 A 上语言 L 的句法么半群(半群)一旦是幂零的就一定有限, 因此不存在无限的幂零的句法么半群(半群). 当然存在无限的幂零么半群(半群), 它们不能成为 A 上的任何语言的句法么半群(半群).

据定理的证明, 易得

推论 3 设 A 是有限字母表, k 是一正整数, $L \subseteq A^*$ 是任一语言. 那么 L 的句法么半群 $\text{syn}(L)$ 满足 $(\text{syn}(L) \setminus \{[1]\})^k = \{0\}$, 当且仅当存在 $C \subseteq A$, 及正整数 m, n, \dots, n_m 和 $a_{i1}, a_{i2}, \dots, a_{in_i} \in (A \setminus C) \cup \{1\} (1 \leq i \leq m)$ 使得

$$L = \bigcup_{i=1}^m (C^* a_{i1} C^* a_{i2} C^* \dots C^* a_{in_i} C^*)$$

或者

$$L = A^* \setminus \left(\bigcup_{i=1}^m (C^* a_{i1} C^* a_{i2} C^* \dots C^* a_{in_i} C^*) \right)$$

其中, $1 \leq i \leq m < k, 1 \leq n_i < k$.

注 3 [2] 和[5] 中给出了许多可识语言的句法么半群(半群)的幂零性与语言本身特征的讨论. 然而对于非正则语言的句法么半群(半群)的幂零性与其自身的特征的联系是否也象正则语言类一样, 这是有待进一步研究的问题, 在此不再赘述.

参 考 文 献

- 1 Brzozowski J A, Simon I. Characterizations of locally testable events. *Discrete Math*, 1973, 4: 243 ~ 271
- 2 Pin J E Varieties of Formal Languages. North Oxford Academic Published Ltd. 1986
- 3 Simon I. Piecewise testable events. *Lecture Notes in Computer Science*, 1975, 33: 214 ~ 222
- 4 Thierrin G. Hypercodes, right convex languages and their syntactic monoids. *Proc Amer Math Soc*, 1981, 83: 255 ~ 258
- 5 Straubing H. The variety generated by finite nilpotent monoids. *Semigroup Forum* 1982, 24: 25 ~ 38
- 6 Long D Y (龙冬阳). On Nilpotency of the Syntactic Monoid of a Languages, World Scientific, 1994, 279 ~ 293
- 7 Shyr H J. Free Monoids and Languages. Soochow Univ. Taiwan, 1979

A Language Whose Syntactic Monoid Is Nilpotent

Long Dongyang*

Abstract In the paper we mainly study the nilpotency of the syntactic monoid of a language. A necessary and sufficient condition for the syntactic monoid of a language to be nilpotent is given and consequently the structure of languages whose syntactic monoid are nilpotent is completely determined.

Keywords nilpotent monoids, languages, syntactic congruence

· 简 讯 ·

钟盛标教授捐赠我校图书馆研究论文一批

钟盛标先生曾于 1946~1949 年任国立中山大学物理系教授兼系主任。最近,钟先生从加拿大将其研究论文 24 篇捐赠我校图书馆收藏,其中包括 1932 年他与严济慈先生合作的“臭氧在 3050 至 3400 埃间之吸收系数”等 10 篇有关光谱学的论文,7 篇晶体物理论文,以及原子核物理、磷光学、带光谱和摄影术等领域的学术论文。这些论文均发表在《法国科学院周刊》、《法国物理与镭期刊》、《自然周刊》(德国)、《自然周刊》(英国)、《化学物理期刊》(美国)、《美国光学期刊》、《美国矿物学者期刊》、《中国物理学报》、《中国科学记录》等权威杂志上,具有较高的学术价值。

钟先生对物理学的最重要贡献之一在于光谱学方面。早在 30 年代初期,他就在北平研究院随严济慈先生研究光谱学,内容涉及臭氧、铷和铯的吸收光谱及电场对光谱的影响。后来,他又研究了硒和碲及其氧化物、水杨酸与其碱性盐水溶液及其铯盐蒸气、银蒸气、货币金属等在多种条件下的光谱等。他的光谱学研究成果,常被有关分子光谱和分子结构的书籍引用。抗战期间开始的晶体物理学研究是钟先生对物理学的另一重要贡献。在这方面,他主要研究了水晶的腐蚀图样及其应用,并发明了确定水晶轴向的几种新方法,为抗战困难条件下通讯器材的研制提供了方便。除实验研究之外,钟先生执教多年,桃李满园,为海内外培育了许多优秀的教育和研究人材。

钟先生是广东梅县人,幼年随父至新加坡。后回国就读于北京大学物理系,1934 年由中华文化基金会资助进入法国巴黎大学深造,获法国国家科学博士学位,之后在巴黎居里实验室研究原子核物理。1938 年回国,历任北平研究院物理学研究所研究员,中山大学物理系教授兼系主任,台湾大学物理系教授兼理学院院长等。1956 年前往新加坡,任南洋大学物理系主任兼理学院院长。1973 年退休后赴加拿大安渡晚年。

钟先生向我校图书馆捐赠学术论文将有助于我校的学术交流,并为中国物理学发展史提供珍贵的史料。谨此致以谢意。

(周显元)

* Department of Computer Science, Zhongshan University, Guangzhou 510275