ω-NTB 文法及其秩的若干性质

郭清泉

(山东大学计算机科学系)

ω-NTB GRAMMARS AND SOME PROPERTIES ON THEIR RANK

Guo Qingquan

(*Department of Computer Science, Shandong University*)

ABSTRACT

In this paper we introduce the concept on ω-NTB (ω-Nonterminal Bounded) grammars and their rank and prove that class of ω-NTB languages is identical with the class of ω-ultralinear languages. Moreover, we present some properties on the rank of ω-NTB grammars.

摘　要

本文引入了ω-NTB 文法及其秩的概念，证明了ω-NTB 语言和ω 超线性语言是同一语言类，给出了ω-NTB 文法秩的若干重要性质。

文献[1] 中定义了NTB(非终止符有界) 文法和它的秩。对于ω 语言，我们给出下列定义：

定义1　$\omega-cfgG=(N,\Sigma,P,S,F)$ 称为ω-NTB (ω 非终止符有界) 文法，其中F 为变量的重复集，如果存在正整数k, 使得对于任意的$A\in N, A\overset{*}{\Longrightarrow} w, w\in(N\cup\Sigma)^*$, 则$w$ 中至多有变量的k 次出现。

如果存在一个ω-NTB 文法G, 使得$L(G)=L$, 则称L 为ω-NTB 语言。

定义2　设$G=(N,\Sigma,P,S,F)$ 为ω-NTB 文法，$w\in(N\cup\Sigma)^*$, 则w 在G 中的秩

$$r_G(w)=\max\{r|w\overset{*}{\Longrightarrow}w', w'\in(N\cup\Sigma)^*, \text{且 } w'\text{中有变量的}r\text{次出现}\}$$

定义3　设$G=(N,\Sigma,P,S,F)$ 为ω-NTB 文法，则G 的秩

$$r(G)=\max\{r_G(A)\big|A\in N\}$$

1989 年10月17 日收到，1989 年12 月27 日定稿。

ω-NTB 文法生成的语言L 称为ω-NTB 语言，L 的秩

$$r(L) = \min\{r(G) \mid G为\omega - NTB\ 文法且\ L(G) = L\}$$

定义4 如果ω 超线性文法$G = (N, \Sigma, P, S, F)$ 中的生成式仅为形式

i) 1 型生成式$A \to aBb$,

ii) 2 型生成式$A \to CD$,

iii) 3 型生成式$A \to a$,

其中A, B, C, $D \in N$, $a, b \in \Sigma \cup \{\lambda\}$, 则称$G$ 为ω 超线性范式文法。

文献[3] 中证明，对于任意的ω 超线性文法G, 存在ω 超线性范式文法G', 使得$L(G') = L(G)$。

定理1 设$G = (N, \Sigma, P, S, F)$ 为ω-NTB 文法，则$L(G)$ 为ω 超线性语言。

证明：令$N_i = \{A \in N \mid r_G(A) = i\}$, $i = 1, 2, \cdots, r(G)$, 则诸N_i 两两无交且$N = \bigcup_{i=1}^{r(G)} N_i$。

如果任给$A \in N_i$, 则A 生成式为下列形式之一:

i) $A \to \alpha B\beta$,　　$\alpha\beta \in \Sigma^*$,

由$r_G(B) = r_G(A) = i$, 故$B \in N_i$。

ii) $A \to \prod_{j=1}^{k} \alpha_j B_j \beta_j$, $k \geq 2$, $\alpha_j\beta_j \in \Sigma^*$, $j = 1, 2, \cdots, k$

由$r_G(\prod_{j=1}^{k} \alpha_j B_j \beta_j) = \sum_{j=1}^{k} r_G(B_j) = r_G(A) = i$, 故对于任意的$j$, $j = 1, 2, \cdots, k$, $r_G(B_j) < i$, 从而$\prod_{j=1}^{k} \alpha_j B_j \beta_j \in (\Sigma \cup N_1 \cup \cdots \cup N_{i-1})^*$。

iii) $A \to \alpha$, $\alpha \in \Sigma^*$,

于是，$\{N_1, N_2, \cdots, N_{r(G)}\}$ 是G 的一个超线性分解，从而G 为ω 超线性文法，$L(G)$ 为ω 超线性语言。

定理2 设$G = (N, \Sigma, P, S, F)$ 是ω 超线性范式文法，则$L(G)$ 是ω-NTB 语言，且$r(L(G)) \leq 2^{n-1}$, 这里n 为G 的超线性分解子集数。

证明：设$\{N_1, N_2, \cdots, N_n\}$ 是G 的一个超线性分解，对于任意的$A \in N_i$, 考察A 在G 中的派生

$$A \stackrel{*}{\Longrightarrow} w$$

若派生中使用1 型和3 型生成式，w 中变量出现数不会增加；若派生中使用一次2 型生成式，w 中变量出现数增加1, 同时，出现的变量将属于$(N_1 \cup N_2 \cup \cdots \cup N_{i-1})$。于是$r_G(w) \leq 2^{i-1}$, 即$r_G(A) \leq 2^{i-1}$。当$A \in N$ 时得到$r(G) \leq 2^{n-1}$。从而$L(G)$ 为ω-NTB 语言。

对于任意的ω 超线性文法，可以化归为等价的ω 超线性范式文法(见[3]), 于是我们有下面的重要推论:

推论1 ω-NTB 语言和ω 超线性语言是同一语言类。

推论2 对于任意的ω-NTB 文法$G = (N, \Sigma, P, S, F)$, 可以直接构造有穷转向的ω-pdaM, 使得$T(M) = L(G)$, 且M 的转向次数不超过$2^{r(G)-1}$。

证明：由定理1 可知，G 为ω 超线性文法。不妨假设G 为ω 超线性范式，$\{N_1, N_2, \cdots, N_{r(G)}\}$ 是G 的一个超线性分解，则满足条件的有穷转向的ω-pda M 可以如下构造:

$$M = (Q, \Sigma, \Gamma, \delta, [q, 0], S, F')$$

其中$Q = \{[q,Z] \mid Z \in \{0\} \cup N\}$, $\Gamma = N \cup \Sigma$, $F' = \{[q,Z] \mid Z \in F\}$, 以及$\delta$: 对于任意的$Z \in \{0\} \cup N$,

i)1 型动作: $([q,B], bB) \in \delta([q,Z], a, A)$, iff$(A \to aBb) \in P$,

ii)2 型动作: $([q,0], CB) \in \delta([q,Z], \lambda, A)$, iff$(A \to aBC) \in P$,

iii)3 型动作: $([q,0], \lambda) \in \delta([q,Z], a, A)$, iff$(A \to a) \in P$,

iv)4 型动作: $([q,0], \lambda) \in \delta([q,Z], a, a)$.

为了考察有穷转向的ω-pda M 的转向次数和生成$T(M)$ 的ω-NTB 文法的秩之间的关系，先给出下面的引理:

引理 对于任意$2k-1$ 次转向的$pdaM$, 可以直接构造NTB 文法G, 使得$L(G) = N(M)$, 且$r(G) \le 2k-1$。

证明: 设$M = (Q, \Sigma, \Gamma, \delta, q_0, Z_0)$, 由文献[1], 如下构造的文法$G$ 满足$L(G) = N(M)$:

$$G = (N, \Sigma, P, S)$$

其中$N = \{S\} \cup \left(\bigcup_{1 \le i < j \le 2k} Q_i \times \Gamma \times Q_j \times (\Gamma \cup \{\lambda\})\right)$, 这里$Q_j = \{q \in Q \mid h(q) = i\}$, $h(q)$ 表示到达状态q 所需要的转向次数加1, 以及P:

i) $S \to [q_0, Z_0, q', \lambda], h(q_0) = 1$,

ii) $[q, Z, q', \lambda] \to x$, 若 $(q', \lambda) \in \delta(q, x, Z), h(q)$ 为奇数且 $h(q') - h(q) = 1$,

iii) $[q, Z, q', Y] \to x[q_1, Z_1, q', Y]$, 若 $(q_1, Z_1) \in \delta(q, x, Z)$,

iv) $[q, Z, q', Z'] \to [q, Z, q'', Z'']x$, 若 $(q', Z') \in \delta(q'', x, Z'')$,

v) $[q, Z, q', Z',] \to x[q_1, Z_1, q', \lambda]$, 若 $(q_1, Z', Z_1) \in \delta(q, x, Z), h(q_1)$为奇数, $h(q_1) < h(q')$,

vi) $[q, Z, q', \lambda] \to [q, Z, q'', Z'']x$, 若$(q', \lambda) \in \delta(q'', x, Z''), h(q')$为偶数, $h(q) < h(q'')$,

vii)$[q, Z, q', Y] \to [q, Z, q_1, Z_1][q_1, Z_1, q', Y], h(q) < h(q_1) < h(q')$。

由G 的构造可知G 为NTB 文法。注意到在派生中使用生成式i)~vi) 不会增加变量的出现数，使用一次生成式vii), 变量的出现数增加1, 又对于任意的$q \in Q$, 有$1 \le h(q) \le 2k$, 且对于任意变量$[q, Z, q', Z'] \in N$, 应有$h(q) < h(q')$, 从而$r(G) \le 2k-1$。

定理3 对于任意n 次转向的ω-pda M, 存在ω-NTB 文法G, 使得$L(G) = N(M)$, 且$r(G) \le n+1$。

证明: 由于$N(M) = \bigcup_{i=1}^{l} U_i V_i^{\omega}$, 其中$V_i (i = 1, 2, \cdots, l)$ 是右线性语言，$U_i (i = 1, 2, \cdots, l)$ 是转向次数不大于n 的有穷转向的pda 接受的语言。

首先设生成U_i 的NTB 文法$G'_i = (N'_i, \Sigma, P'_i, S'_i)$, $i = 1, 2, \cdots, l$, 以及生成V_i^{ω} 的ω 右线性文法$G''_i = (N''_i, \Sigma, P''_i, S''_i, F_i)$, $i = 1, 2, \cdots, l$; 又令$G_i = (N_i, \Sigma, P_i, S_i, F_i)$, $i = 1, 2, \cdots, l$, 其中$N_i = N'_i \cup N''_i \cup \{S_i\}$, 以及$P_i = P'_i \cup P''_i \cup \{S_i \to S_i, S''_i\}$, 最后令$G = (N, \Sigma, P, S, F)$, 其中$N = \bigcup_{i=1}^{l} N_i \cup \{S\}$, $P = \bigcup_{i=1}^{l} P_i \cup \{S \to S_i \mid 1 \le i \le l\}$, 以及重复集的集合$F = \{F_i \mid 1 \le i \le l\}$。容易证明$L(G) = N(M)$。

这样，G 为ω-NTB 文法，且由G 的构造和引理可知，$r(G'_i) \le n$, $i = 1, 2, \cdots, l$, 从而$r(G_i) \le n+1$, 以及$r(G) = \max\{r(G_i) \mid 1 \le i \le l\} \le n+1$。

为了研究ω-NTB 文法的秩的其他性质，再引入下面的定义:

定义5 f' 转换器是一个六元组$S_{f'}=(Q,\Sigma,\Delta,\delta,S_0,F')$，其中$Q,\Sigma$ 和Δ 分别是有穷非空的状态，输入和输出集合；S_0 是开始状态；$F'\subseteq 2^Q$ 是指定状态重复集的集合；δ 是$Q\times\Sigma^*\times\Delta^*\times Q$ 的有穷子集。

如果$\delta\subseteq Q\times(\Sigma\cup\{\lambda\})\times(\Delta\cup\{\lambda\})\times Q$，则称$S_{f'}$ 为1– 限制型f' 转换器。

设有无穷序列$p_i\in Q$，$i\geq 0$；$u_i\in\Sigma^*$，$v_i\in\Delta^*$，$i\geq 1$，使得$\sigma_1=\prod_{i=1}^{\infty}u_i$，$\sigma_2=\prod_{i=1}^{\infty}v_i$，及$(p_i,u_{i+1},v_{i+1},p_{i+1})\in\delta,i\geq 0$，则有推导$r$

$$\sigma_2:(p_0,\sigma_1)=(p_0,\prod_{i=1}^{\infty}u_i)\vdash(p_1,\prod_{i=2}^{\infty}u_i)\vdash\cdots$$

$S_{f'}(\sigma_1)=\{\sigma_2\big|$ 存在推导r 为$\sigma_2:(p_0,\sigma_1)=(p_0,\prod_{i=1}^{\infty}u_i)\vdash(p_1,\prod_{i=2}^{\infty}u_i)\vdash\cdots$，且INS (r) $\in F'\}$，这里INS (r) 表示推导r 的状态重复集。

$$S_{f'}(L)=\bigcup_{\sigma_1\in L}S_{f'}(\sigma_1)$$

定理4 如果L 为ω–NTB 语言，$S_{f'}$ 是f' 转换器，则$S_{f'}(L)$ 是ω–NTB 语言，且其秩$r(S_{f'}(L))\leq 2^{r(L)-1}+1$。

证明：不妨设$S_{f'}=(Q,\Sigma,\Sigma,\delta,S_0,F')$ 是1– 限制型f' 转换器，且对于任意的$p\in Q$，有$(p,\lambda,\lambda,p)\in\delta$；又设接受$L$ 的有穷转向的$\omega-pdaM=(Q',\Sigma,\Gamma,\delta',q_0,Z_0,F_2')$，$T(M)=L$，另外，对于任意的$(q,Z)\in Q'\times\Gamma$，有$(q,Z)\in\delta'(q,\lambda,Z)$。令

$$M'=(Q'',\Sigma,\Gamma,\delta'',q_0',Z_0,F'')$$

其中$Q''=Q\times Q'$，$q_0'=[S_0,q_0]$，$F''=\{F_1\times F_2\big|F_1\in F_1'\cdot F_2\in F_2'\}$，以及$\delta'':([p',q'],\gamma)\in\delta''([p,q],a,Z),iff$ 存在$b\in\Sigma\cup\{\lambda\}$，使$(p,b,a,p')\in\delta$，且$(q',\gamma)\in\delta'(q,b,Z)$。

根据M' 的构造可知，$T(M')=S_{f'}(T(M))=S_{f'}(L)$。由推论2，$M$ 的转向次数不大于$2^{r(L)-1}$，而M' 的转向次数不大于M 的转向次数，于是由定理3 得到$r(S_{f'}(L))\leq 2^{r(L)-1}+1$。

定理5 ω–NTB 语言对与ω 右线性语言的交，gsm 映射和正则替换封闭，且它们的秩均不大于$2^{r(L)-1}+1$，这里，$r(L)$ 是ω–NTB 语言L 的秩。

证明：1) 任给ω–NTB 语言L 和ω 右线性语言R，设$\omega-fsaA=(Q,\Sigma,\delta,S_0,F')$，$T(A)=R$。令

$$S_{f'}=(Q,\Sigma,\Sigma,\delta',S_0,F')$$

其中$\delta'=\{(q,a,a,q')\big|\delta(q,a)=q',a\in\Sigma\}$。

我们有$L\cup R=S_{f'}(L)$，以而$L\cap R$ 是ω–NTB 语言，由定理4 知$r(L\cap R)\leq 2^{r(L)-1}+1$。

2) 设gsm 映射$S=(Q,\Sigma,\Delta,\delta,\lambda,S_0)$。令

$$S_{f'}=(Q,\Sigma\cup\Delta,\Sigma\cup\Delta,\delta',S_0,2^Q)$$

其中$\delta':(p,a,b,q)\in\delta'$，iff $\delta(p,a)=q$ 且$\lambda(p,a)=b$。

（下转第19 页）