理论计算机科学基础
第二章习题课

晏荣杰
中科院软件所
计算机科学国家重点实验室

第二章 上下文无关文法

- 主要概念

  1. 上下文无关文法 (*CFG*) : (V, Σ, R, S) 2.1.c, 2.4.b, 2.6.b, 2.14

  推导 (派生)、最左 (最右) 推导、推导树 (语法分析树)、歧义性、乔姆斯基范式

  2. 下推自动机 (PDA) : $(Q, \Sigma, \Gamma, \delta, q_0, F)$ 2.11

  格局、格局的改变、计算的定义、接受的语言

  3. 上下文无关语言 2.2

- 主要结论:

  上下文无关语言类 = PDA 接受的语言类 = CFG 生成的语言类

- CFG 和 PDA 的设计 (不太复杂的例子) 补充作业

第二章 上下文无关文法

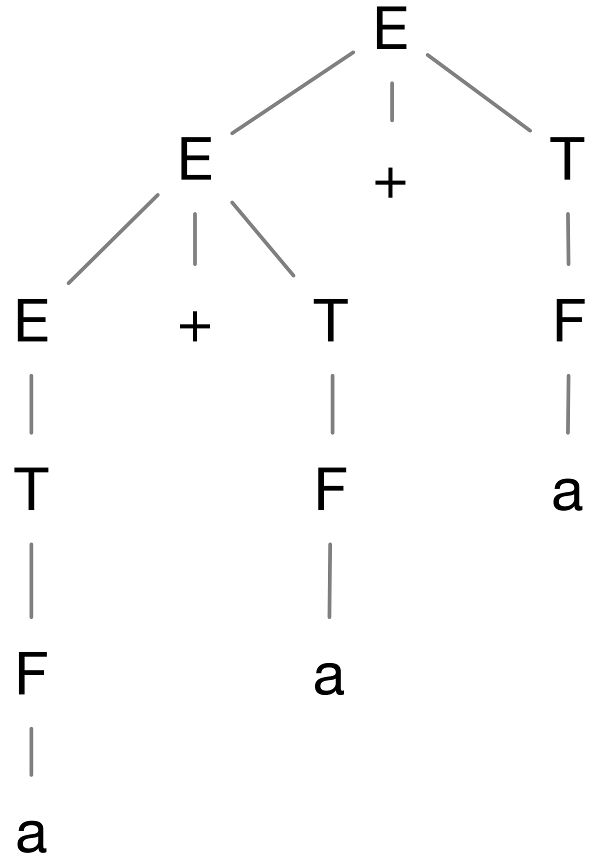
2.1 给出下述字符串的语法分析树与派生：

– C. a+a+a

E->E+T | T

T->TXF | F

F->(E) | a



E=>E+T=>E+F=>E+a=>
E+T+a=>E+F+a=>E+a+a=>
T+a+a=>F+a+a=>a+a+a

第二章 上下文无关文法

2.4设字母表是{0,1}，给出产生下述语言的上下文无关文法

- b. {w | w以相同的符号开始和结束}

- S->0A0 | 1A1
- A->0A| 1A | ε

第二章 上下文无关文法

2.6 给出产生下述语言的上下文无关文法

- b．语言$\{a^n b^n \mid n \geq 0\}$的补集

解：把补集中的字符串分三种情况：

1. 某个b在a前。

   $A \rightarrow XbaX$

   $X \rightarrow \varepsilon \mid aX \mid bX$

2. a全在b前，但a多b少。

   $B \rightarrow a \mid aB \mid aBb$

3. a全在b前，但a少b多。

   $C \rightarrow b \mid Cb \mid aCb$

最后合起来： $S \rightarrow A \mid B \mid C$

解： 首先生成任意字符串

$T \rightarrow aT \mid bT \mid \varepsilon$

然后，在外围限制a与b出现的次序

$S \rightarrow aSb \mid bT \mid Ta$

因此，结果为：

$S \rightarrow aSb \mid bT \mid Ta$

$T \rightarrow aT \mid bT \mid \varepsilon$

第二章 上下文无关文法

2.14 用定理2.6中给出的过程，把下述CFG转换成等价的乔姆斯基范式文法。

- $A \to BAB|B|\varepsilon$
- $B \to 00|\varepsilon$

解：添加新起始变元S_0,

$S_0 \to A$

$A \to BAB|B|\varepsilon$

$B \to 00|\varepsilon$

去掉$B \to \varepsilon$

$S_0 \to A$

$A \to BAB|AB|BA|A|B|\varepsilon$

$B \to 00$

去掉$A \to \varepsilon$,

$S_0 \to A|\ \varepsilon$

$A \to BAB|AB|BA|A|B|BB$

$B \to 00$

去掉$A \to A$, $A \to B$

$S_0 \to A\ |\ \varepsilon$

$A \to BAB|AB|BA|00|BB$

$B \to 00$

去掉$S_0 \to A$,

$S_0 \to BAB|AB|BA|00|BB\ |\ \varepsilon$

$A \to BAB|AB|BA|00|BB$

$B \to 00$

添加新变元

$S_0 \to VB|AB|BA|UU|BB\ |\ \varepsilon$

$A \to VB|AB|BA|UU|BB$

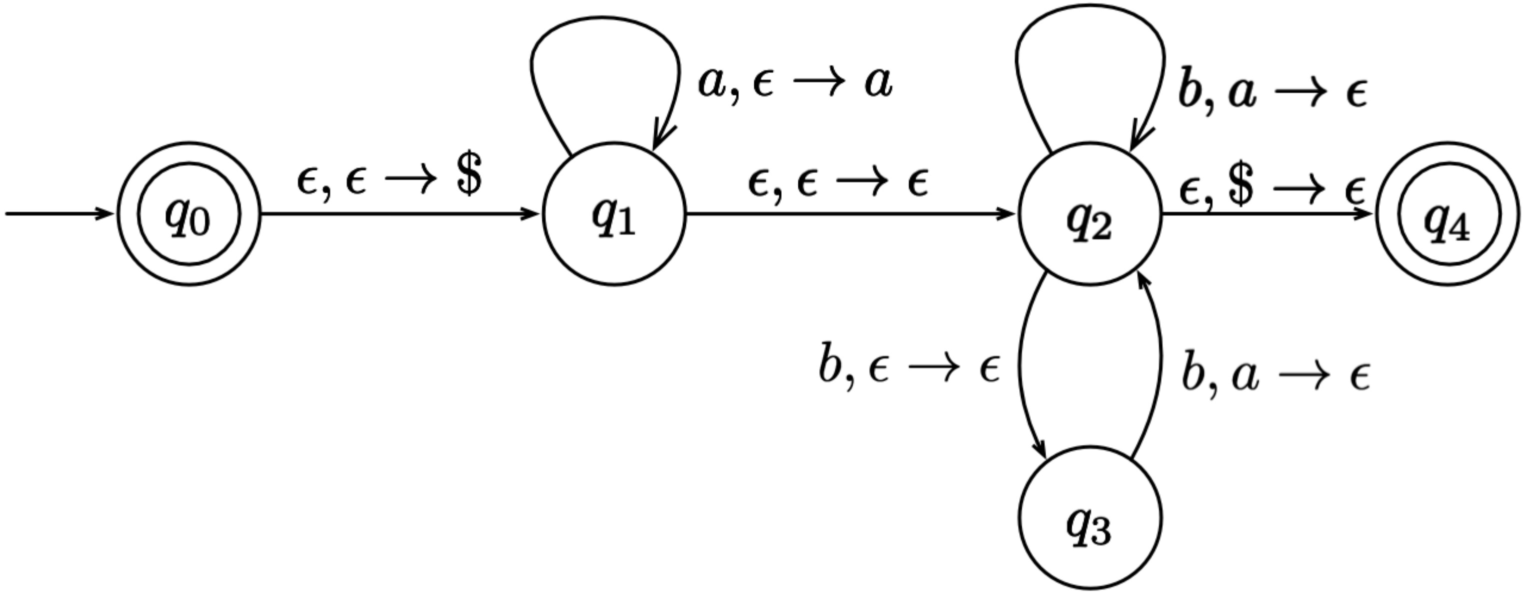
$B \to UU$

$V \to BA$

$U \to 0$

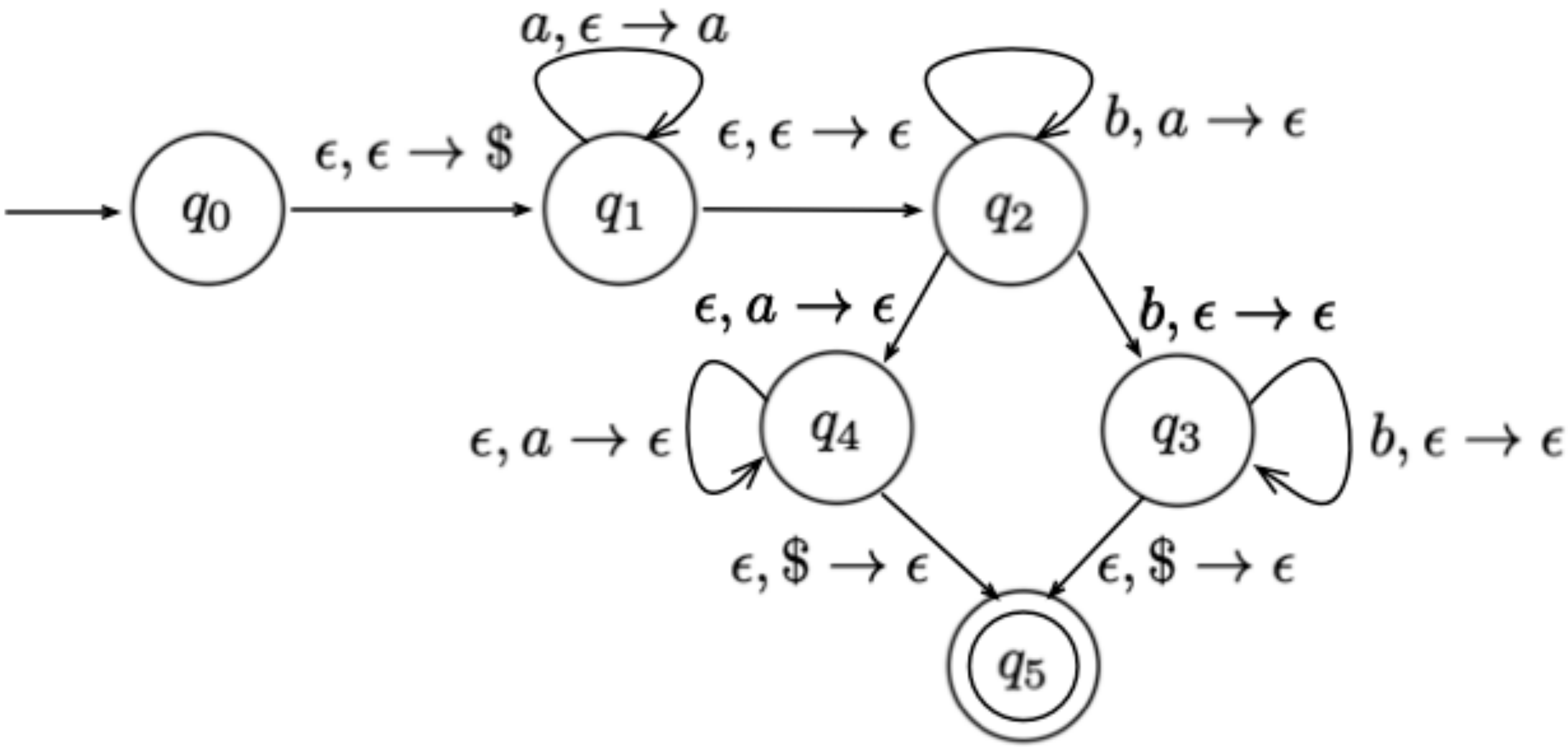
第二章 上下文无关文法

- 1.给出一台识别语言$\{a^m b^n \mid m,n \geq 0, 且 m \leq n \leq 2m\}$的PDA。 对于输入字符串 w = *aabbb*, 给出 PDA 接受 w 的一条计算路径。



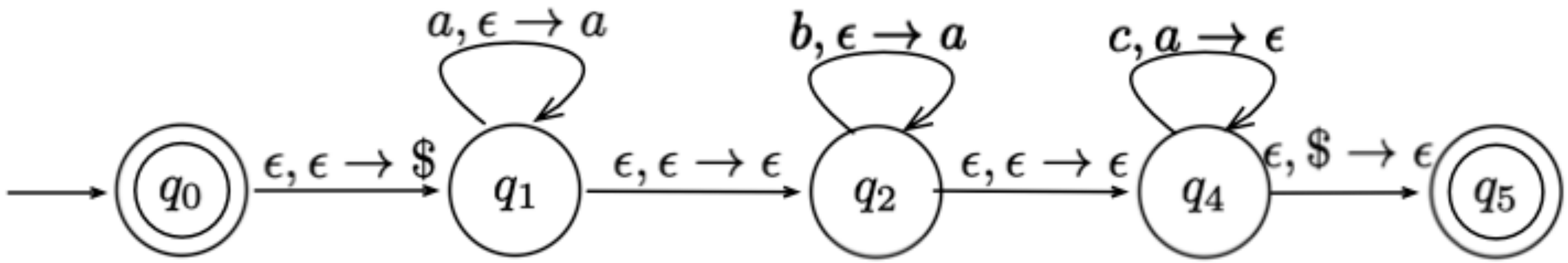
$<q_0,aabbb,\varepsilon> \vdash <q_1,aabbb,\$> \vdash <q_1, abbb,a\$> \vdash <q_1, bbb,aa\$> \vdash <q_2, bbb,aa\$> \vdash <q_2, bb,a\$> \vdash <q_3, b,a\$> \vdash <q_2,\varepsilon, a\$> \vdash <q_4,\varepsilon,\varepsilon>$

第二章 上下文无关文法

- 2. 给出一台识别语言$\{a^m b^n \mid m,n\geq 0$, 且$m\neq n\}$的PDA。对于 输入字符串 *w* = *aabbb*, 给出 PDA 接受 *w* 的一条计算路径。
- 

第二章 上下文无关文法

- 3.给出一台识别语言$\{a^i b^j c^k \mid i,j,k\geq 0,$且$k=i+j\}$的PDA。



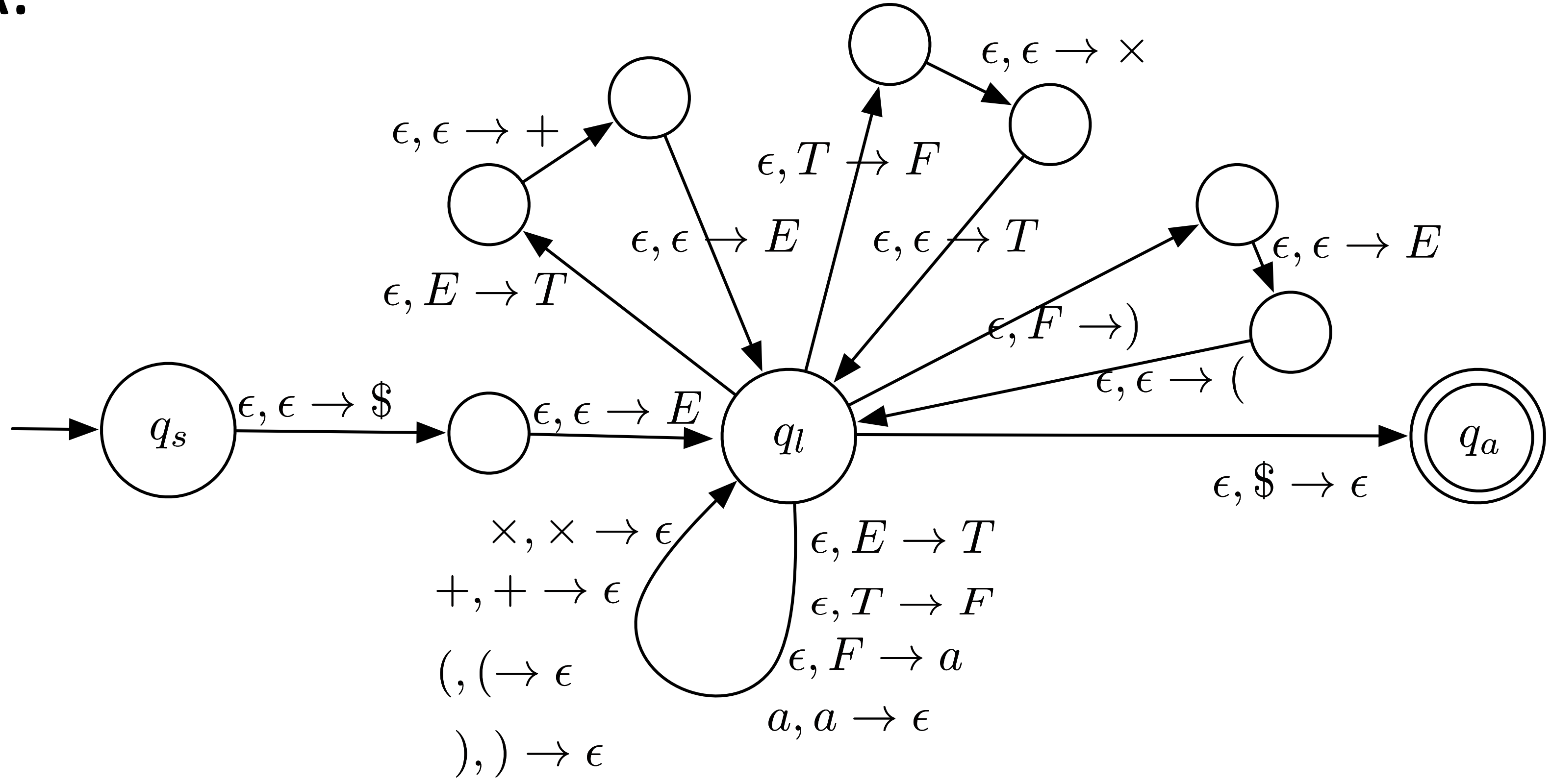
第二章 上下文无关文法

- 2.11 用定理2.12中给出的过程，把练习2.1中的CFG G4转化成等价的PDA.

  E->E+T | T

  T->TxF | F

  F->(E) | a



例. 设 D 是语言 $\{ww \mid w \in \{0,1\}^*\}$, 用泵引理证明 D 不是 CFL.

(反证法) 假定 D 是 CFL，则它有一个泵长度 p。
选择ω 为字符串 $0^p1^p0^p1^p$, 因为 $\omega \in D$ 且 $|\omega| \geq p$, 由泵引理可知
ω 可以分成 5 段 $\omega = uvxyz$, 使得下列 3 条都成立：
(1) $|vxy| \leq p$,
(2) $|vy| \geq 1$,
(3) 对于任意 $n \geq 0$, 有 $uv^nxy^nz \in D$
情况 1： 若 vxy 中全是 0 且位于 ω 的前 p 位.
令 $k = |vy|$, 由 (2) 知 $k \geq 1$。取 $n = 2$, 则由 (3) 可知 $uv^nxy^nz =$
$0^{p+k}1^p0^p1^p \in D$. 但由于 $k \geq 1$, 故 $0^{p+k}1^p0^p1^p$ 的前半段和后半段
的 0 的个数不等，这与 D 的定义矛盾。
情况 2： 若 vxy 中有 0 也有 1，且位于 ω 的前 $2p$ 位.
不妨设 vy 中有 1(否则证明同情况 1)。取 $n = 0$, 则由 (3) 可知
$uxz \in D$，但 uxz 的前半段最多有 $p-1$ 个 1，而其后半段至少有
p 个 1. 与 D 的定义矛盾。

第二章 上下文无关文法

- 2.2 a. 利用语言$A=\{a^m b^n c^n \mid m,n\geq 0\}$和$B=\{a^n b^n c^m \mid m,n\geq 0\}$以及例2.20，证明上下文无关语言在交的运算下不封闭。

证明：a.先说明A,B均为上下文无关文法

对A构造CFG C_1

$S\rightarrow aS|T|\varepsilon$

$T\rightarrow bTc|\varepsilon$

对B,构造CFG C_2

$S\rightarrow Sc|R|\varepsilon$

$R\rightarrow aRb|\varepsilon$

由此知 A,B均为上下文无关语言。

但是由例2.20, $A\cap B=\{a^n b^n c^n|n\geq 0\}$不是上下文无关语言，所以上下文无关语言在交的运算下不封闭。

第二章 上下文无关文法

- 2.2 a. 利用语言A={$a^m b^n c^n$ | m,n≥0}和A={$a^n b^n c^m$ | m,n≥0}以及例2.20，证明上下文无关语言在交的运算下不封闭。
- b. 利用(a)和DeMorgan律(定理0.10)，证明上下文无关语言在补运算下不封闭。

证明：用反证法。假设CFL在补运算下封闭，则对于(a)中上下文无关语言A,B，他们的补$\bar{A}$、$\bar{B}$也为CFL，且CFL对并运算封闭，所以 A、B的补语言的并$\bar{A} \cup \bar{B}$也是CFL，所以 A、B的补语言的并的补$\overline{\bar{A} \cup \bar{B}}$也是CFL。

由DeMorgan定律， A、B的补语言的并的补为$\overline{\bar{A} \cup \bar{B}}$ ＝A∩B，由此A∩B是CFL，这与(a)的结论矛盾，所以CFL对补运算不封闭。