Course: CSC707, Automata, Computability and Computational Theory
Homework 6: Context-free languages, context-free grammars, PDA, Pumping lemma.
Submission: Use Wolfware
File Format: LaTeX and PDF, and any images you have

Due Date: 2:00 A.M. (EST), Tuesday, April 13, 2010

---

1. Prove non-context-free using Pumping lemma:
   (a) $L = \{0^i 1^j 2^i 3^j | i, j \geq 1\}$
   (b) $L = \{a^i b^j c^k | 0 \leq i < j < k\}$
   (c) $L = \{a^i b^j | j = i^2\}$
   (d) $\bar{L}$, where $L = \{0^k | k$ is a perfect square $\}$

   (a) **Theorem 1** *$L = \{0^i 1^j 2^i 3^j | i, j \geq 1\}$ is non-context-free.*
   Proof. We can use pumping lemma to prove it.
   1. For $\forall p$, we choose $s = 0^p 1^p 2^p 3^p \in L \bigcap \Sigma^{\geq p}$.
   2. $\forall u, v, x, y, z, s = uvxyz, s.t. |vy| \geq 1$ and $|vxy| \leq p$.
   3. Choose $i = 2$, we need to prove $s = uv^2xy^2z \notin L$.
   4. Consider all the cases of v and y:

   - Case 1: Both v and y contain only a same number:

     $\frac{v}{y} | \frac{0^m}{0^n} | \frac{1^m}{1^n} | \frac{2^m}{2^n} | \frac{3^m}{3^n}$

     When $i = 2$, the number of 0 (or 1,2,3) would be $p + m + n$. By (2), $m + n \geq 1 \Rightarrow p + m + n \neq p$. Thus $s = uv^2xy^2z \notin L$.

   - Case 2: Both v and y contain one number but they are different and $m \geq 1, n \geq 1$:

     $\frac{v}{y} | \frac{0^m}{1^n} | \frac{1^m}{2^n} | \frac{2^m}{3^n}$

     When $i = 2$, the number of 0 and 1 (or 1 and 2, 2 and 3) would be $p + m$ and $p + n$, respectively. Since $p + m \neq p$ and $p + n \neq p$, $s = uv^2xy^2z \notin L$.

   - Case 3: Either v or y contain two numbers and $m \geq 1, n \geq 1, k \geq 1$:

     $\frac{v}{y} | \frac{0^m 1^k}{1^n} | \frac{0^m}{0^n 1^k} | \frac{1^m 2^k}{2^n} | \frac{1^m}{1^n 2^k} | \frac{2^m 1^k}{3^n} | \frac{2^m}{2^n 3^k}$

When $i = 2$, s will have interleaving patterns, such as $0^m 1^k 0^m 1^k$. Thus $s = uv^2xy^2z \notin L$.

5. $L = \{0^i 1^j 2^i 3^j | i, j \geq 1\}$ is non-context-free. ∎

(b) **Theorem 2** $L = \{a^i b^j c^k | 0 \leq i < j < k\}$

Proof. We can use pumping lemma to prove it.

1. For $\forall p$, we choose $s = a^p b^{p+1} c^{p+2} \in L \bigcap \Sigma^{\geq p}$.
2. $\forall u, v, x, y, z, s = uvxyz, s.t. |vy| \geq 1$ and $|vxy| \leq p$.
3. Consider all the cases of v and y:

- Case 1: Both v and y contain only a same number:

  $\frac{v}{y} | \frac{a^m}{a^n} | \frac{b^m}{b^n} | \frac{c^m}{c^n}$

  For $v = a, y = a$, we choose $i = 2$. In this way, the number of a would be $p + m + n$. By (2), $m + n \geq 1 \Rightarrow p + m + n \geq p + 1$. Thus $s = uv^2xy^2z = a^{p+m+n} b^{p+1} c^{p+2} \notin L$.

  For $v = b, y = b$, we choose $i = 2$. In this way, the number of b would be $p + m + n$. By (2), $m + n \geq 1 \Rightarrow p + 1m + n \geq p + 2$. Thus $s = uv^2xy^2z = a^p b^{p+m+n+1} c^{p+2} \notin L$.

  For $v = c, y = c$, we choose $i = 0$. In this way, the number of c would be $p - m - n$. By (2), $m + n \geq 1 \Rightarrow p + 2 - (m + n) \leq p + 1$. Thus $s = uv^2xy^2z = a^p b^{p+1} c^{p+2-(m+n)} \notin L$.

- Case 2: Both v and y contain one number but they are different and $m \geq 1, n \geq 1$:

  $\frac{v}{y} | \frac{a^m}{b^n} | \frac{b^m}{c^n}$

  For $v = a, y = b$, we choose $i = 2$. In this way, the number of b would be $p - n$. $p + 1 + n \geq p + 2$. Thus $s = uv^2xy^2z = a^{p+m} b^{p+1_n} c^{p+2} \notin L$.

  For $v = b, y = c$, we choose $i = 0$. In this way, the number of b would be $p + 1 - n$. $p + 1 - m \leq p$. Thus $s = uv^2xy^2z = a^p b^{p+1-m} c^{p+2-n} \notin L$.

- Case 3: Either v or y contain two numbers and $m \geq 1, n \geq 1, k \geq 1$:

  $\frac{v}{y} | \frac{a^m b^k}{b^n} | \frac{a^m}{a^n b^k} | \frac{b^m c^k}{c^n} | \frac{b^m}{b^n c^k}$

  When $i = 2$, s will have interleaving patterns, such as $a^m b^k a^m b^k$. Thus $s = uv^2xy^2z \notin L$.

5. $L = \{0^i 1^j 2^i 3^j | i, j \geq 1\}$ is non-context-free. ■

(c) **Theorem 3** $L = \{a^i b^j | j = i^2\}$

Proof.

We can use pumping lemma to prove it.

1. For $\forall p$, we choose $s = a^p b^{p^2} \in L \bigcap \Sigma^{\geq p}$.

2. $\forall u, v, x, y, z, s = uvxyz, s.t. |vy| \geq 1$ and $|vxy| \leq p$.

3. Consider all the cases of v and y:

- Case 1: Both v and y contain only a same number:

  $\frac{v}{y} | \frac{a^m}{a^n} | \frac{b^m}{b^n}$

  For $v = a, y = a$, we choose $i = 2$. In this way, the number of a would be $p + m + n$. By (2), $m + n \geq 1 \Rightarrow (p + m + n)^2 \neq p^2$. Thus $s = uv^2xy^2z = a^{p+m+n}b^{p^2} \notin L$.

  For $v = b, y = b$, we choose $i = 2$. In this way, the number of b would be $p^2 + m + n$. By (2), $m + n \geq 1 \Rightarrow p^2 + m + n \neq p^2$. Thus $s = uv^2xy^2z = a^p b^{p^2+m+n} \notin L$.

- Case 2: Both v and y contain one number but they are different and $m \geq 1, n \geq 1$:

  $\frac{v}{y} | \frac{a^m}{b^n}$

  For $v = a, y = b$, we choose $i = 2$. In this way, the number of a would be $p + m$ and the number of b would be $p^2 + n$. Since $(p + m)^2 = p^2 + 2mp + m^2 > p^2 + n$. Thus $s = uv^2xy^2z = a^{p+m}b^{p^2+n} \notin L$.

- Case 3: Either v or y contain two numbers and $m \geq 1, n \geq 1, k \geq 1$:

  $\frac{v}{y} | \frac{a^m b^k}{b^n} | \frac{a^m}{a^n b^k}$

  When $i = 2$, s will have interleaving patterns, such as $a^m b^k a^m b^k$ and $a^n b^k a^n b^k$. Thus $s = uv^2xy^2z \notin L$.

■

(d) **Theorem 4** $\bar{L}$*, where* $L = \{0^k | k$ *is a perfect square* $\}$

Proof.

We use pumping lemma to prove it.

1. For $\forall p$, we choose $s = 0^{7(x)^2}$, where x is the product of all the primes that are less or equal to p. Since 7 is a prime and x^2 contains two copies of each prime that is less or equal to p, if $p \geq 7$, then

$7(x)^2$ has three 7s as prime factors and cannot be perfect square; if $p < 7$, then $7(x)^2$ has only one 7 as prime factors and cannot be perfect square. Thus, $7(x)^2$ is not a perfect square and $s \in \bar{L} \bigcap \Sigma^{\geq p}$.
2. $\forall u, v, x, y, z, s = uvxyz =, s.t. |vy| \geq 1$ and $|vxy| \leq p$.
3. There is only one case for v and y: $|vy| = m, 1 \leq m \leq p$. In this way, $s = 0^{7(x)^2+(i-1)m}$.
4. Choose $i = (1 + \frac{2x^2}{m})$. By 2, $1 \leq m \leq p$. Thus, m can be expressed as the products of some primes that are less or equal than p. Since x is the product of all the primes that are less or equal to p, x contains all the prime factors of m. Thus, $2x^2$ is the multiple of m and $i = (1+\frac{2x^2}{m})$ is a natural number. In this way, $s = 0^{7(x)^2+(\frac{2x^2}{m})m}$. The number of 0s is $7(x)^2 + (\frac{2x^2}{m})m = 7(x)^2 + 2x^2 = 9x^2 = (3x)^2$, which is a perfect square. Thus, $s \notin \bar{L}$ and $\bar{L}$ is not context-free.
■

2. Design a PDA and provide a context-free grammar (in any form) to accept the following language:
   (a) $L = \{a^n b^{n+m} c^m | n \geq 0, m \geq 1\}$
   (b) The set of all strings over $\{a, b\}$ with exactly twice as many a's as b's.

   Solutions:

   (a) The PDA for L is shown in Figure 1.

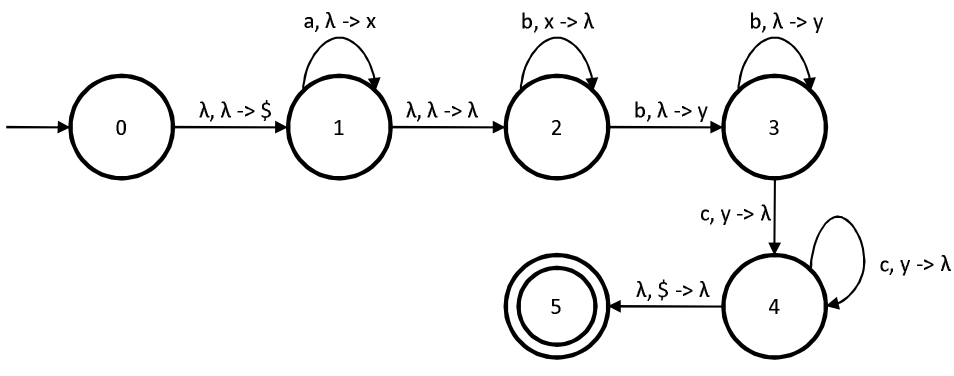

Figure 1: PDA of $L = \{a^n b^{n+m} c^m | n \geq 0, m \geq 1\}$

The CFG for L is:

$S \rightarrow AB$
$A \rightarrow aAb|\lambda$
$B \rightarrow bBc|bc$

(b) The PDA for L is shown in Figure 2.

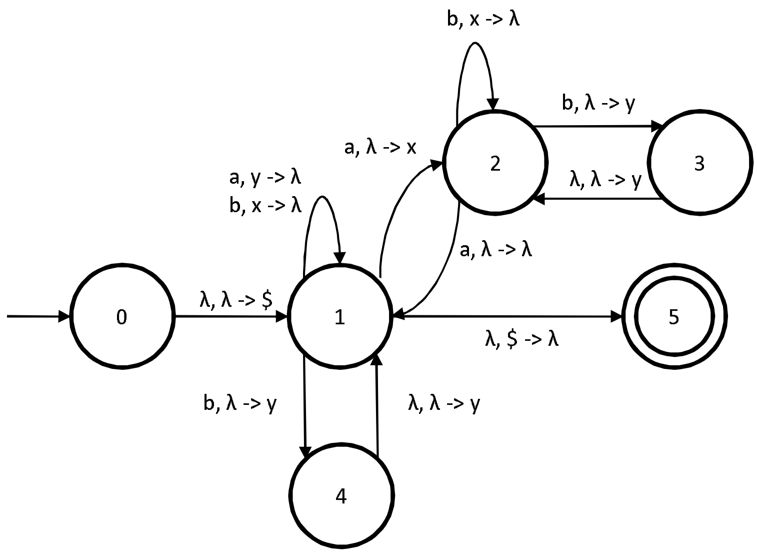

Figure 2: PDA of the set of all strings over $\{a, b\}$ with exactly twice as many a's as b's

The CFG for L is:

$S \rightarrow aSaSb|aSbSa|bSaSa|\lambda$

3. Give a context-free grammar in Chomsky Normal Form that generates the following language:

   (a) The set of all strings over $\{a, b\}$ with exactly twice as many a's as b's.

   (b) $L = \{w \in (a + b + c)^* | n_a(w) + n_b(w) \neq n_c(w)\}$, where $n_a(w)$ is the number of a's in w.

   (c) $L = \{a^i b^j c^k | i \neq j \text{ or } j \neq k\}$

Solutions:

(a) The CFG for L is:

$S \rightarrow aSaSb|aSbSa|bSaSa|\lambda$

The context-free grammar in Chomsky Normal Form for L is:
$S_0 \rightarrow S|\lambda$

$A \to a$
$B \to b$
$S \to AA_1|AA_3|AA_5|AA_7|BB_1|BB_2|AA_4|A_4A|BA_8$
$A_1 \to AA_2$
$A_2 \to SB$
$A_3 \to SA_4$
$A_4 \to AB$
$A_5 \to BA_6$
$A_6 \to SA$
$A_7 \to SA_8$
$A_8 \to BA$
$B_1 \to BA_6$
$B_2 \to SA_8$

(b) The CFG for L is:

$S \to S_1|S_2|S_3$
$A \to aA|a$
$B \to bB|b$
$C \to cC|c$
$M_{ac} \to aM_{ac}c|cM_{ac}a|ac|ca$
$M_{bc} \to bM_{bc}c|cM_{bc}b|bc|cb$
$S_1 \to AS_1M_{ac}S_1M_{bc}|AS_1M_{bc}S_1M_{ac}|$
$M_{ac}S_1AS_1M_{bc}|M_{ac}S_1M_{bc}S_1A|M_{bc}S_1AS_1M_{ac}|M_{bc}S_1M_{ac}S_1A$
$S_2 \to BS_2M_{ac}S_2M_{bc}|BS_2M_{bc}S_2M_{ac}|$
$M_{ac}S_2BS_2M_{bc}|M_{ac}S_2M_{bc}S_2B|M_{bc}S_2BS_2M_{ac}|M_{bc}S_2M_{ac}S_2B$
$S_3 \to CS_3M_{ac}S_3M_{bc}|CS_3M_{bc}S_3M_{ac}|$
$M_{ac}S_3CS_3M_{bc}|M_{ac}S_3M_{bc}S_3C|M_{bc}S_3CS_3M_{ac}|M_{bc}S_3M_{ac}S_3C$

The context-free grammar in Chomsky Normal Form for L is:

$S_0 \to S$
$A \to a$
$B \to b$
$C \to c$
$A_a \to AA_a|a$
$B_b \to BB_b|b$
$C_c \to CC_c|c$
$M_{ac} \to AM_{ac2}|CM_{ac3}|AC|CA$
$M_{ac2} = M_{ac}C$
$M_{ac3} = M_{ac}A$
$M_{bc} \to BM_{bc2}|CM_{bc3}|BC|CB$

$M_{bc2} = M_{bc}C$
$M_{bc3} = M_{bc}B$
$S \to AS_{11}|AS_{14}|M_{ac}S_{17}|M_{ac}S_{19}|M_{bc}S_{112}|M_{bc}S_{115}|$
$BS_{21}|BS_{24}|M_{ac}S_{27}|M_{ac}S_{29}|M_{bc}S_{212}|M_{bc}S_{215}$
$CS_{31}|CS_{34}|M_{ac}S_{37}|M_{ac}S_{39}|M_{bc}S_{312}|M_{bc}S_{315}$
$S_1 \to AS_{11}|AS_{14}|M_{ac}S_{17}|M_{ac}S_{19}|M_{bc}S_{112}|M_{bc}S_{115}$
$S_{11} \to S_1S_{12}$
$S_{12} \to M_{ac}S_{13}$
$S_{13} \to S_1M_{bc}$
$S_{14} \to S_1S_{15}$
$S_{15} \to M_{bc}S_{16}$
$S_{16} \to S_1M_{bc}$
$S_{17} \to S_1S_{18}$
$S_{18} \to AS_{13}$
$S_{19} \to S_1S_{110}$
$S_{110} \to M_{bc}S_{111}$
$S_{111} \to S_1A$
$S_{112} \to S_1S_{113}$
$S_{113} \to AS_{114}$
$S_{114} \to S_1M_{ac}$
$S_{115} \to S_1S_{116}$
$S_{116} \to M_{ac}S_{111}$
$S_2 \to BS_{21}|BS_{24}|M_{ac}S_{27}|M_{ac}S_{29}|M_{bc}S_{212}|M_{bc}S_{215}$
$S_{21} \to S_2S_{22}$
$S_{22} \to M_{ac}S_{23}$
$S_{23} \to S_2M_{bc}$
$S_{24} \to S_2S_{25}$
$S_{25} \to M_{bc}S_{26}$
$S_{26} \to S_2M_{bc}$
$S_{27} \to S_2S_{28}$
$S_{28} \to BS_{23}$
$S_{29} \to S_2S_{210}$
$S_{210} \to M_{bc}S_{211}$
$S_{211} \to S_2B$
$S_{212} \to S_2S_{213}$
$S_{213} \to BS_{214}$
$S_{214} \to S_2M_{ac}$
$S_{215} \to S_2S_{216}$
$S_{216} \to M_{ac}S_{211}$
$S_3 \to CS_{31}|CS_{34}|M_{ac}S_{37}|M_{ac}S_{39}|M_{bc}S_{312}|M_{bc}S_{315}$

$S_{31} \to S_3S_{32}$
$S_{32} \to M_{ac}S_{33}$
$S_{33} \to S_3M_{bc}$
$S_{34} \to S_3S_{35}$
$S_{35} \to M_{bc}S_{36}$
$S_{36} \to S_3M_{bc}$
$S_{37} \to S_3S_{38}$
$S_{38} \to CS_{33}$
$S_{39} \to S_3S_{2310}$
$S_{310} \to M_{bc}S_{311}$
$S_{311} \to S_2C$
$S_{312} \to S_3S_{313}$
$S_{313} \to CS_{314}$
$S_{314} \to S_3M_{ac}$
$S_{315} \to S_3S_{316}$
$S_{316} \to M_{ac}S_{311}$

(c) The CFG for L is:

$S \to A_{i>j}C|B_{i<j}C|AC_{j>k}|AD_{j<k}$
$A \to aA|\lambda$
$B \to bB|\lambda$
$C \to cC|\lambda$
$A_{i>j} \to aA_{i>j}b|aA$
$B_{i<j} \to aB_{i<j}b|bB$
$C_{j>k} \to bC_{j>k}c|bB$
$D_{j<k} \to bD_{j<k}c|cC$

The context-free grammar in Chomsky Normal Form for L is:
$S_0 \to S$
$S \to A_{i>j}C_c|B_{i<j}C_c|A_aC_{j>k}|A_aD_{j<k}$
$A \to a$
$B \to b$
$C \to c$
$A_a \to AA_a$
$B_b \to BB_b$
$C_c \to CC_c$
$A_{i>j} \to AA_1|AA_a|A$
$A_1 \to A_{i>j}B$
$B_{i<j} \to AB_1|BB_b|B$
$B_1 \to B_{i<j}B$

$C_{j>k} \to BC_1|BB_b|B$
$C_1 \to C_{j>k}C$
$D_{j<k} \to BD_1|CC_c|C$
$D_1 \to D_{j<k}C$