CS150 Homework 5

Solution keys, spring, 2025

Problem 1. (10 points)

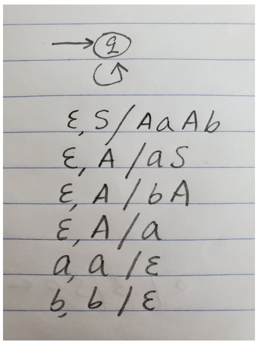
Convert the grammar

$$S \to AaAb$$
$$A \to aS \mid bA \mid a$$

to a PDA that accepts the same language by empty stack.

PDA $(\{q\}, \{a, b\}, \{S, A, a, b\}, \delta, q, S)$ where

$$\delta(q, a, a) = \{(q, \epsilon)\}$$
$$\delta(q, b, b) = \{(q, \epsilon)\}$$
$$\delta(q, \epsilon, S) = \{(q, AaAb)\}$$
$$\delta(q, \epsilon, A) = \{(q, aS), (q, bA), (q, a)\}$$



Note that you should show the PDA as a diagram although our solution keys list a sequence of transitions.

Problem 2. (Exercise 6.3.4, 10 points)

Convert the PDA of Exercise 6.1.1 (treated as an empty-stack PDA) to a context-free grammar.

> **Exercise 6.1.1:** The PDA $P = (\{q, p\}, \{0, 1\}, \{Z_0, X\}, \delta, q, Z_0)$ has the following transition function:
>
> 1. $\delta(q, 0, Z_0) = \{(q, XZ_0)\}$.
> 2. $\delta(q, 0, X) = \{(q, XX)\}$.
> 3. $\delta(q, 1, X) = \{(q, X)\}$.
> 4. $\delta(q, \epsilon, X) = \{(p, \epsilon)\}$.
> 5. $\delta(p, \epsilon, X) = \{(p, \epsilon)\}$.
> 6. $\delta(p, 1, X) = \{(p, XX)\}$.
> 7. $\delta(p, 1, Z_0) = \{(p, \epsilon)\}$.
>
> Starting from the initial ID (q, w, Z_0).

$$S \to [qZq] \mid [qZp]$$

The following four productions come from rule (1).

$$\begin{aligned}
[qZq] &\to 0[qXq][qZq] \\
[qZq] &\to 0[qXp][pZq] \\
[qZp] &\to 0[qXq][qZp] \\
[qZp] &\to 0[qXp][pZp]
\end{aligned}$$

The following four productions come from rule (2).

$$\begin{aligned}
[qXq] &\to 0[qXq][qXq] \\
[qXq] &\to 0[qXp][pXq] \\
[qXp] &\to 0[qXq][qXp] \\
[qXp] &\to 0[qXp][pXp]
\end{aligned}$$

The following two productions come from rule (3).

$$[qXq] \to 1[qXq]$$
$$[qXp] \to 1[qXp]$$

The following production comes from rule (4).

$$[qXp] \to \epsilon$$

The following production comes from rule (5).

$$[pXp] \to \epsilon$$

The following four productions come from rule (6).

$$[pXp] \to 1[pXq][qXp]$$
$$[pXp] \to 1[pXp][pXp]$$
$$[pXq] \to 1[pXq][qXq]$$
$$[pXq] \to 1[pXp][pXq]$$

The following production comes from rule (7).

$$[pZp] \to 1$$

Problem 3. (Exercise 7.1.3, 20 points)

Begin with the grammar:

$$
\begin{aligned}
S &\to 0A0 \mid 1B1 \mid BB \\
A &\to C \\
B &\to S \mid A \\
C &\to S \mid \epsilon
\end{aligned}
$$

(a) Eliminate ϵ-productions.

(b) Eliminate any unit productions in the resulting grammar.

(c) Eliminate any useless symbols in the resulting grammar.

(d) Put the resulting grammar into Chomsky Normal Form.

(a) $V_N = \{C, A, B, S\}$

$$
\begin{aligned}
S &\to 0A0 \mid 00 \mid 1B1 \mid 11 \mid BB \mid B \\
A &\to C \\
B &\to S \mid A \\
C &\to S
\end{aligned}
$$

(b) Because $C \to S$, $A \to S$ ($A \to C$, $C \to S$) and $B \to S$ ($B \to A$, $A \to S$), we can replace A, B, C with S.

$$S \to 0S0 \mid 00 \mid 1S1 \mid 11 \mid SS$$

Or we can have four identical sets of productions for *S, A, B, C*!
Note the above simplification is not part of the standard transformation algorithm.

(c) There is no useless symbol that we can remove.

$$S \to 0S0 \mid 00 \mid 1S1 \mid 11 \mid SS$$

(d) Replace 0 with Q_0 and 1 with Q_1, we get

$$\begin{aligned}
S &\to Q_0SQ_0 \mid Q_0Q_0 \mid Q_1SQ_1 \mid Q_1Q_1 \mid SS \\
Q_0 &\to 0 \\
Q_1 &\to 1
\end{aligned}$$

Then we introduce new variables to split long productions.

$$\begin{aligned}
S &\to Q_0X_1 \mid Q_0Q_0 \mid Q_1X_2 \mid Q_1Q_1 \mid SS \\
Q_0 &\to 0 \\
Q_1 &\to 1 \\
X_1 &\to SQ_0 \\
X_2 &\to SQ_1
\end{aligned}$$

Problem 4. (Exercise 7.2.1 (b)(c), 20 points)

Use the CFL pumping lemma to show each of these languages not to be context-free:

(b) $\{a^n b^n c^i \mid i \leq n\}$.

(c) $\{0^p \mid p \text{ is a prime}\}$. *Hint:* Adapt the same ideas used in Example 4.3, which showed this language not to be regular.

(b) Let p be the pumping const and consider the string $z = a^p b^p c^p$. We may write $z = uvwxy$, where v and x may be "pumped," and $|vwx| \leq p$, $|vx| > 0$. Because $|vwx| \leq p$, vwx cannot contain three different types of terminals (a, b, and c) and cannot contain both a and c as well.

1. If vwx contains a or b, then uv^0wx^0y will decrease the number of a's or b's, resulting in more number of c's than a's or b's, which is not in the language.
2. If vwx contains b or c, then uv^iwx^iy $(i > 1)$ will increase the number of b's and c's, without increasing the number of a's, which is not in the language.

By contradiction, we conclude this language is not context-free.

(c) Let n be the pumping-lemma constant and consider string $z = 0^p$ and $p \geq n + 2$. We can write $z = uvwxy$, and $|vwx| \leq n$ and $|vx| > 0$. Let $|vx| = m$, consider $uv^{p-m}wx^{p-m}y$, then the length of it is $(p - m) \times m + (p - m) = (p - m) \times (m + 1)$ which is obviously not a prime. This contradicts the assumption, so the language is not context-free.

Problem 5. (Exercise 7.3.2, 20 points)

Consider the following two languages:

$$L_1 = \{a^n b^{2n} c^m \mid n, m \geq 0\}$$
$$L_2 = \{a^n b^m c^{2m} \mid n, m \geq 0\}$$

(a) Show that each of these languages is context-free by giving grammars for each.

(b) Is $L_1 \cap L_2$ a CFL? Justify your answer.

(a) L_1:

$$S \to XC$$
$$X \to aXbb \mid \epsilon$$
$$C \to cC \mid \epsilon$$

L_2:

$$S \to AY$$
$$Y \to bYcc \mid \epsilon$$
$$A \to aA \mid \epsilon$$

(b) Their intersection becomes $L_3 = \{a^n b^{2n} c^{4n} \mid n \geq 0\}$ and is NOT context-free.

Let n be the pumping-lemma constant and consider string $z = a^n b^{2n} c^{4n}$. We can write $z = uvwxy$, where $|vwx| \leq n$ and $|vx| > 0$. Because $|vwx| \leq n$, vwx cannot contain 3 different symbols (a, b, and c) or include both a and c as well.

1. If vwx contains a and b, then uv^iwx^iy ($i > 1$) will increase number of a's and b's without increasing number of c's and then it cannot be in the language.
2. If vwx contains b and c, then uv^iwx^iy ($i > 1$) will increase number of b's and c's without increasing number of a's and then it cannot be in the language.
3. If vwx contains only a or b or c, then uv^iwx^iy ($i > 1$) will increase only the number of that symbol without increasing the other two, and then it cannot be in the language.

All the above lead to contradiction, so the language is not context-free.

Problem 6. (Exercise 7.4.3 (b)(c), optional, 0 points)

Using the grammar G of Example 7.34, use the CYK algorithm to determine whether each of the following strings is in $L(G)$.

$$\begin{aligned} S &\to AB \mid BC \\ A &\to BA \mid a \\ B &\to CC \mid b \\ C &\to AB \mid a \end{aligned}$$

(b) *baaab*.

(c) *aabab*.

(b)

$\{S, C\}$				
$\{S, A, C\}$	$\{S, C\}$			
$-$	$\{S, A, C\}$	$\{B\}$		
$\{S, A\}$	$\{B\}$	$\{B\}$	$\{S, C\}$	
$\{B\}$	$\{A, C\}$	$\{A, C\}$	$\{A, C\}$	$\{B\}$
b	a	a	a	b

$baaab \in L(G)$.

(c)

$\{S, C\}$				
$\{S, A, C\}$	$\{B\}$			
$\{B\}$	$\{B\}$	$\{S, C\}$		
$\{B\}$	$\{S, C\}$	$\{S, A\}$	$\{S, C\}$	
$\{A, C\}$	$\{A, C\}$	$\{B\}$	$\{A, C\}$	$\{B\}$
a	a	b	a	b

$aabab \in L(G)$.