COMP 272 Theory of Computation
April 30, 1998, **Quiz 2 Solution**

Problem 1.

Construct a context-free grammar for the language represented by the following regular expression:

$$a(a \cup bb)^*ba$$

The grammar is $G = (V, \Sigma, R, S)$ where

$V = \{a, b, S, T, X\}$
$\Sigma = \{a, b\}$
$R = \{\ S \to aTba,$
$\quad T \to XT,$
$\quad T \to e,$
$\quad X \to a,$
$\quad X \to bb\ \}.$

Problem 2.

Construct a pushdown automaton that accepts the language generated by the grammar $G = (V, \Sigma, R, S)$, where $V = \{a, b, S, A\}$, $\Sigma = \{a, b\}$, and

$$R = \{S \to aAA, \quad A \to aS, \quad A \to bS, \quad A \to a\}$$

The pushdown automaton is $M = (K, \Sigma, \Gamma, \Delta, s, F)$ where:

$K = \{s, f\}$
$\Sigma = \{a, b\}$
$\Gamma = \{a, b, S, A\}$
$F = \{f\}$
$\Delta = \{\ ((s, e, e), (f, S)),$
$\quad ((f, e, S), (f, aAA)),$
$\quad ((f, e, A), (f, aS)),$
$\quad ((f, e, A), (f, bS)),$
$\quad ((f, e, A), (f, a)),$
$\quad ((f, a, a), (f, e)),$
$\quad ((f, b, b), (f, e))\ \}.$

Problem 3.

For each language below, determine whether or not it is context-free.

(a) $\{a^m b^n c^p\ :\ m = n \text{ OR } n = p \text{ OR } m = p\}$
The language is context-free. Write the corresponding grammar for verification.

(b) $\{a^m b^n c^p\ :\ m \neq n \text{ OR } n \neq p \text{ OR } m \neq p\}$
The language is context-free. Write the corresponding grammar for verification.

(c) $\{a^m b^n c^p\ :\ m = n \text{ AND } n = p \text{ AND } m = p\}$
The language is NOT context-free. Basically, it is just $\{a^n b^n c^n : n \geq 0\}$.

(d) $\{w \in \{a,b,c\}^* : w$ does not contain equal number of occurrences of a, b, and $c\}$

Let m be the number of a's in the string
n be the number of b's in the string
p be the number of c's in the string

Then, the language L is $\{w \in \{a,b,c\}^* :$ NOT $(m = n$ AND $n = p)\}$,
which is just $\{w \in \{a,b,c\}^* : (m \neq n$ OR $n \neq p)\}$,
which in turn is the union of L_1 and L_2, where
$L_1 = \{w \in \{a,b,c\}^* : m \neq n\}$ and $L_2 = \{w \in \{a,b,c\}^* : n \neq p\}$.
Since L_1 and L_2 are accepted by some pushdown automata M_1 and M_2 respectively (verify this), so $L = L_1 \cup L_2$ is context-free.

Problem 4.

Consider the context-free grammar $G = (V, \Sigma, R, S)$, where $V = \{a, b, S\}$, $\Sigma = \{a, b\}$, and

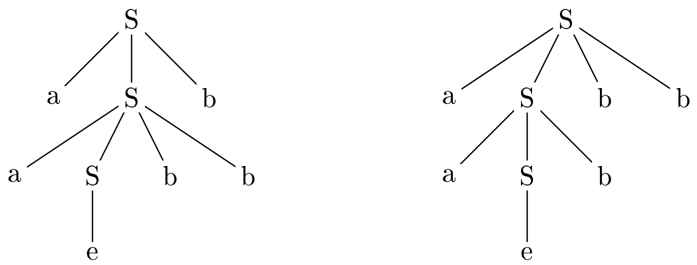
$$R = \{S \to e, \quad S \to aSb, \quad S \to aSbb\}$$

(a) Describe informally (that is, in plain English) the language generated by this context-free grammar.

The language consists of strings in the form $a^n b^m$ such that $n \leq m \leq 2n$.

(b) Show that this context-free grammar is ambiguous.

Consider the string $aabbb$. Two different parse trees can be drawn for this string:



(c) Construct an equivalent unambiguous context-free grammar.

The idea is to introduce another non-terminal such that one grammar rule is to be exhaustively applied first before the other.
The modified grammar is $G' = (V', \Sigma, R', S)$ where:

$V' = \{a, b, S, T\}$
$R' = \{\ S \to aSb,$
$S \to T,$
$T \to aTbb,$
$T \to e\ \}.$

Problem 5.

Construct a pushdown automaton for the following language:

$\{w \in \{a,b,c\}^* :$ (Number of a's in w) $=$ (Number of b's in w) $+$ (Number of c's in w)$\}$

The pushdown automaton is $M = (K, \Sigma, \Gamma, \Delta, s, F)$ where:

$$
\begin{aligned}
K &= \{s, q, f\} \\
\Sigma &= \{a, b, c\} \\
\Gamma &= \{a, b, c, m, n, t\} \\
F &= \{f\} \\
\Delta &= \{ \ ((s, e, e), (q, t)), \\
&\quad ((q, a, t), (q, mt)), \\
&\quad ((q, a, m), (q, mm)), \\
&\quad ((q, b, m), (q, e)), \\
&\quad ((q, c, m), (q, e)), \\
&\quad ((q, b, t), (q, nt)), \\
&\quad ((q, c, t), (q, nt)), \\
&\quad ((q, a, n), (q, e)), \\
&\quad ((q, b, n), (q, nn)), \\
&\quad ((q, c, n), (q, nn)), \\
&\quad ((q, e, t), (f, e)) \ \}.
\end{aligned}
$$

Problem 6.

Consider the context-free grammar $G = (V, \Sigma, R, S)$, where $V = \{a, b, S\}$, $\Sigma = \{a, b\}$, and

$$R = \{S \to a, \quad S \to Sa, \quad S \to bSS, \quad S \to SbS\}$$

Prove using mathematical induction that every string in the language generated by this grammar has more a's than b's.

We prove this by induction on the length of the string.

If the string is of length 1, the only possible derivation is $S \Rightarrow a$. It is obvious that the number of a's, which is 1, is larger than the number of b's, which is 0. Hence the proposition is true for the basis step.

Assume that the proposition is true for any string in the language L with length less than or equal to $n \geq 1$.

Consider any string w where $|w| = n + 1$. There are several derivation possibilities: [1]

1. $S \Rightarrow \underline{S}a \Rightarrow^* w$
2. $S \Rightarrow b\underline{SS} \Rightarrow^* w$
3. $S \Rightarrow \underline{S}b\underline{S} \Rightarrow^* w$

Observe that any string derived from the underlined non-terminal has length $\leq n$ (otherwise $|w| > n + 1$, leading to contradiction). It is not difficult to check that there are more a's than b's in each case (Take the first case as an example. By the induction hypothesis, the string derived from $\underline{S}$ has more a's than b's. The extra a produced from the first derivation will not alter this property. Other cases are to be verified similarly).

By mathematical induction, any string in the language has more a's than b's.

[1] Non-terminals that are underlined are for illustration purpose only. They are identical to non-terminals without underlining otherwise.