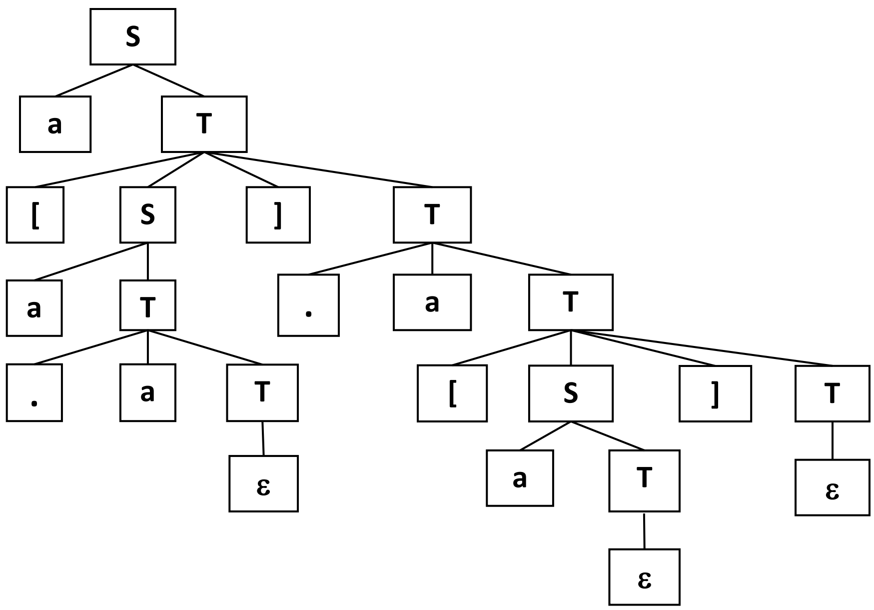
CS 475/575 Exam 2 Fall 2012 Solution

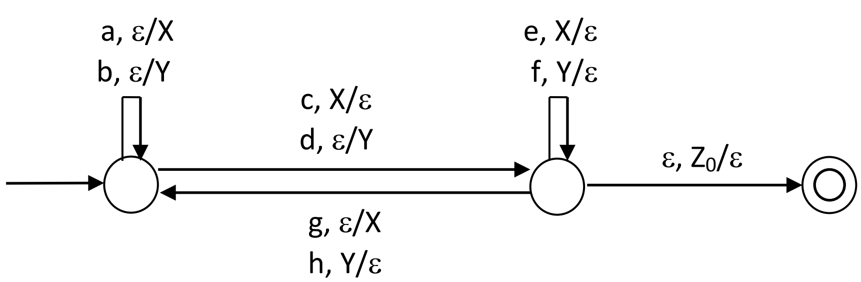
This exam has 120 possible points (includes 20 points extra credit).

1. Draw a parse tree for the string a [a **.** a] **.** a [a] using this context-free grammar: **[10 points]**

$S \rightarrow$ a T
$T \rightarrow$ **[** S **]** T | **.** a T | ε



2. Write all strings with length exactly 4 that are accepted by this pushdown automaton: **[10 points]**



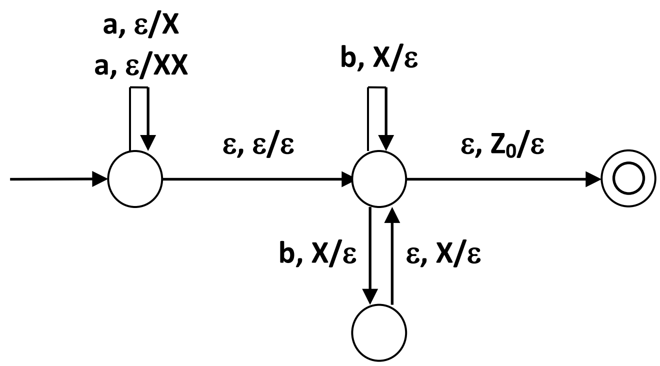
aace	**acgc**	**adfe**	**adhc**	**bacf**
bdff	**dfgc**	**dgcf**	**dhac**	**dhdf**

3. Let language $L_1 = \{a^m b^n \mid m \leq 2n \text{ and } n \leq 2m\}$.

    a. Write a context-free grammar that generates language L_1. Ambiguity is permitted. **[10 points]**

    $S \rightarrow aSb \mid aaSb \mid aSbb \mid \varepsilon$

    b. Draw a pushdown automaton that accepts language L_1 both by final state and by empty stack. Non-determinism is permitted. **[10 points]**

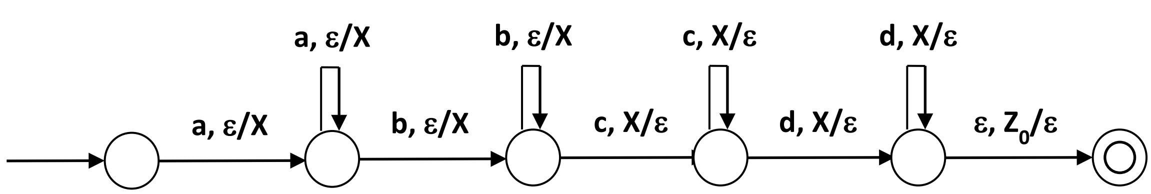


4. Let language $L_2 = \{a^m b^n c^p d^q \mid m\geq1, n\geq1, p\geq1, q\geq1, \text{ and } m+n=p+q\}$. Examples: $a^3b^2c^4d^1$ = aaabbccccd, $a^2b^5c^3d^4$ = aabbbbbcccdddd.

    a. Write an *unambiguous* context-free grammar that generates language L_2. **[10 points]** (Half credit for an ambiguous grammar.)

    $S \rightarrow aTd$
    $T \rightarrow aTd \mid aUc \mid bVd \mid bWc$
    $U \rightarrow aUc \mid bWc$
    $V \rightarrow bVd \mid bWc$
    $W \rightarrow bWc \mid \varepsilon$

    b. Draw a *deterministic* pushdown automaton that accepts language L_2 both by final state and by empty stack. **[10 points]** (Half credit for a non-deterministic machine.)

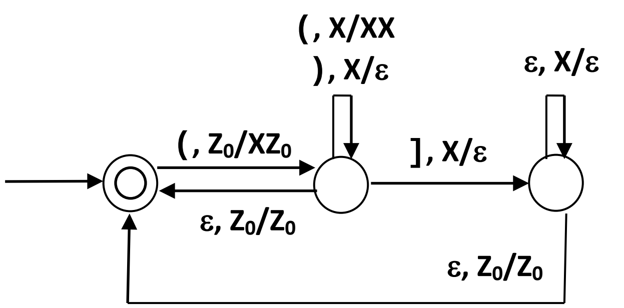


5. Let language L_3 contain strings of balanced left and right parentheses '(' and ')' with the additional provision that a right bracket ']' may be used to match all (one or more) preceding unmatched left parentheses. Examples: `(()())`, `((((()]`, `(((]()`, `()(()(]((())(()(((]`.

   a. Write an *unambiguous* context-free grammar that generates language L_3. **[10 points]** (Half credit for an ambiguous grammar.)

   S → (M)S | (U]S | ε
   M → (M)M | ε // M ≡ every '(' must be matched with ')'
   U → (U | MU | ε // U ≡ some '(' might be unmatched

   b. Draw a *deterministic* pushdown automaton that accepts language L_3 by final state, with only the bottom-of-stack symbol Z_0 remaining on the stack if the string is accepted. **[10 points]** (Half credit for a non-deterministic machine.)



6. Eliminate all useless symbols, ε-productions, and unit productions from this context-free grammar. Your grammar should be equivalent to the original grammar. Bonus if you convert the grammar to Chomsky normal form. **[10 points + 4 points]**

S → UX | TU | YT | RaR
T → b | cT | Sd
U → e | V
V → W | f
W → g | h
X → ε | TZ | XY
Y → XX | ZS | YY
Z → ZQ | QZ
Q → i | ε
R → jk | ε

S → e | f | g | h | TU
| b | cT | Sd
| RaR | Ra | aR | a
T → b | cT | Sd
U → e | f | g | h
R → jk

S → e | f | g | h | TU
| b | CT | SD
| RP | RA | AR | a
T → b | CT | SD
U → e | f | g | h
R → JK
P → AR
A → a
C → c
D → d
J → j
K → k

7. Let $L_4 = \{ a^m b^n c^q \mid q = \max(m,n) \}$. Examples: $a^2b^4c^4$ = aabbbbcccc, $a^4b^2c^4$ = aaaabbcccc, $a^3b^3c^3$ = aaabbbccc. Use the pumping theorem to show that L_4 is not context-free. **[16 points]**

   First complete this statement of the pumping theorem for context-free languages:
   For every context-free language L, there exists some constant p such that
   for every string s with $\mathbf{s \in L}$ and $\mathbf{|s| \geq p}$,
   it is possible to write **s = uvwxy** such that $\mathbf{|vwx| \leq p}$, $\mathbf{|vx| \geq 1}$,
   and for every integer $i \geq 0$, $\mathbf{u v^i w x^i y \in L}$.

   Next apply the pumping theorem to show that $L_4 = \{ a^m b^n c^q \mid q = \max(m,n) \}$ is not context-free.
   Choose string s = $\mathbf{a^p b^p c^p}$.
   Determine the possible cases and show a contradiction in each case:

   If either v or x contains two distinct symbols then $u v^2 w x^2 y \notin a^* b^* c^*$, hence $u v^2 w x^2 y \notin L_4$.
   So each of v and x must consist entirely of one symbol (either a's or b's or c's).

   If neither v nor x contains c's then $u v^2 w x^2 y$ has $q = p < \max(m,n)$, hence $u v^2 w x^2 y \notin L_4$.

   Finally, if either v or x does contains c's then vx cannot also contain both a's and b's, so $u v^0 w x^0 y$ has $q < p = \max(m,n)$, hence $u v^0 w x^0 y \notin L_4$.

8. Trace the CYK dynamic programming algorithm for input string "abcbab" using this Chomsky normal form grammar. Complete the table below. **[10 points]**

   $S \to XV \mid WU$
   $T \to WX \mid VT$
   $U \to VW \mid XS$
   $V \to a \mid UX \mid XW$
   $W \to b \mid TW \mid WV$
   $X \to c \mid SV \mid VX$

	1	2	3	4	5	6
1	**V**	**U**	**T,V**	**U,W**	**U,W**	**S**
2	–	**W**	**T**	**W**	**W**	**S**
3	–	–	**X**	**V**	**V**	**U**
4	–	–	–	**W**	**W**	**S**
5	–	–	–	–	**V**	**U**
6	–	–	–	–	–	**W**