EXAM

Please read all instructions, including these, carefully

• There are 8 questions on the exam, with multiple parts. You have 3 hours to work on the exam.
• The exam is open book, open notes.
• Please write your final answers in the space provided on the exam. You may use the backs of the exam pages as scratch paper, or use additional pages (available at the front of the room).
• Each problem has a straightforward solution. Solutions will be graded on correctness and clarity. Partial solutions will be given partial credit.

NAME: ______________________________________________________

<table>
<thead>
<tr>
<th>Problem</th>
<th>Max points</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
1. Let $G$ be the following grammar:

\[
\begin{align*}
S & \rightarrow ABC \\
A & \rightarrow Aa \mid aB \\
B & \rightarrow Bb \mid \varepsilon \\
C & \rightarrow Cc \mid \varepsilon
\end{align*}
\]

Answer true or false to the following questions:

a. $G$ can be used to specify an LL(1) parser
b. $a$ is in $\text{FIRST}(S)$
c. $c$ is in $\text{FIRST}(Aa)$
d. $c$ is in $\text{FOLLOW}(A)$

2. Modify the following language to make it suitable for parsing with an $\text{LL}(1)$ parser. Note that the terminal symbols in this grammar are $l$, $($, $*$, $a$, and $b$.

\[
\begin{align*}
E & \rightarrow E \mid T \\
E & \rightarrow T \\
T & \rightarrow TP \\
T & \rightarrow P \\
P & \rightarrow P* \\
P & \rightarrow F \\
F & \rightarrow (E) \\
F & \rightarrow a \\
F & \rightarrow b
\end{align*}
\]
3. Consider the following simple context free grammar for the same language:

\[
S \rightarrow Aa \mid bAc \mid Bc \mid bBa \\
A \rightarrow d \\
B \rightarrow d
\]

a. Is this grammar LR(1)? Why or why not?

b. Is it LALR(1)? Why or why not?
4. Consider the $LL(1)$ grammar, where non-terminals are given in brackets ($<>$):

- $<\text{ROBOT}> \rightarrow \text{WALK} <\text{DIR}> <\text{ROBOT}>
- <\text{ROBOT}> \rightarrow \text{PICK} <\text{OBJ}> <\text{ROBOT}>
- <\text{ROBOT}> \rightarrow \text{DROP} <\text{OBJ}> <\text{ROBOT}>
- <\text{ROBOT}> \rightarrow \text{STOP}
- <\text{DIR}> \rightarrow \text{NORTH}
- <\text{DIR}> \rightarrow \text{EAST}
- <\text{DIR}> \rightarrow \text{SOUTH}
- <\text{DIR}> \rightarrow \text{WEST}
- <\text{OBJ}> \rightarrow \text{BOX}
- <\text{OBJ}> \rightarrow \text{BALL}$

Draw the parse tree of the sentence WALK EAST PICK BOX DROP BALL STOP.
5. Consider the following grammar:

\[
S \rightarrow Bxxzx
\]
\[
\mid Cxxxxz
\]
\[
\mid xBxxxx
\]
\[
B \rightarrow w
\]
\[
C \rightarrow w
\]

For some values of \(i, j, k\), the above grammar is LR(\(i\)) and SLR(\(j\)) but not SLR(\(k\)).

a. What is the smallest correct value for \(i\)?

b. What is the smallest correct value for \(j\)?

c. What is the largest correct value for \(k\)?
6. A context-free grammar over 0 and 1 has three non-terminals $S$, $A$, and $B$. We know the following about the FIRST and FOLLOW sets of these non-terminals:

$$\text{FIRST} (S) = \{0, 1\}$$
$$\text{FOLLOW} (S) = \{0, \$\}$$
$$\text{FIRST} (A) = \{\epsilon\}$$
$$\text{FOLLOW} (A) = \{0, 1\}$$
$$\text{FIRST} (B) = \{\epsilon\}$$
$$\text{FOLLOW} (B) = \{0, 1\}$$

a. Construct a context-free grammar with these FIRST and FOLLOW sets.
b. Construct the $LL(1)$ parsing table for your grammar. Is your grammar $LL(1)$?

c. Show the parsing steps for the string 10 with your $LL(1)$ parsing table, and trace a left-most derivation of this string.
7. Consider the following context-free grammar:

\[
\begin{align*}
E' & \rightarrow E \\
E & \rightarrow E \ a \ T \ | \ T \\
T & \rightarrow T \ F \ | \ F \\
F & \rightarrow F \ b \ | \ c
\end{align*}
\]

a. Construct the \( LR(0) \) items and the DFA to recognize the viable prefixes for this grammar. Notice that the grammar is already augmented, and \( E' \) is the new start symbol.

b. Using the DFA that you constructed in part (a), build the \( SLR(1) \) parsing table for this grammar.
c. Show the parsing steps for the string $cbcac$.

d. Can you add a single additional production to the grammar to make it not $SLR(1)$? Give a brief explanation of your answer.
8. Consider the following context-free grammar that generates regular expressions.

a. Define a syntax-directed transition that records the maximum number of *nested* kleene star operators of a regular expression \( R \) in its attribute \( R.depth \). For example, the regular expression \((a)^* \mid ((b)\mid a)^*\) has depth 2. The semantic actions for the three base cases are given.

\[
R \rightarrow a \\
\{ R.depth = 0; \}
\]

\[
R \rightarrow b \\
\{ R.depth = 0; \}
\]

\[
R \rightarrow \varepsilon \\
\{ R.depth = 0; \}
\]

\[
R \rightarrow R_1 \ R_2 \\
\]

\[
R \rightarrow R_1 \mid R_2 \\
\]

\[
R \rightarrow (R_i) \\
\]

\[
R \rightarrow (R_i)^* \\
\]

b. Is \( R.depth \) inherited or synthesized? Explain your answer.
c. Show the parse tree for the string \((a^*b^*) \mid (b^*a^*)b^*\) and decorate the tree with the appropriate attributes and values.