Computer Science Comprehensive Exam—Fall 2014
Compiler Construction

Instructions: Do *not* put your name on the exam, please answer all the questions directly on the exam itself. You may write on the back of the pages. You may need scratch paper to work out the answers before writing them on the exam. Answer **all** the questions. Explain answers as fully as possible, give examples or define terms, if appropriate.

1. For each statement below, circle either "true" or "false" as appropriate.

   (a) true / false The set of languages recognized by SLR grammars is a proper subset of the languages recognized by context-free grammars.

   (b) true / false The set of languages recognized by LR(1) grammars is a proper subset of the languages recognized by LL(1) grammars.

   (c) true / false The set of languages recognized by LALR(1) grammars is a proper subset of the languages recognized by LR(1) grammars.

   (d) true / false A shift/shift conflict is possible in an LR parsing table.

   (e) true / false There exists a regular expression r for which the formal language denoted by r is exactly the set of syntactically correct Java programs.

2. Give a (simple) grammar for which the following two LR(1) items appear in the same state of the LR(1) parsing automaton. Or, explain why it cannot happen. A, P, Q, and R are non-terminals; and x is a terminal.

$$A \to P \bullet Q\,R, x \qquad A \to P\,Q \bullet R, x$$

3. What is an activation record? What is a procedure closure? What is the difference?

4. Consider the regular expression $a((b|a^*c)d)^*|d^*a$.

   (a) A regular expression can be understood better as an abstract syntax tree. Give the asbstract syntax tree for the regular expression above.

   (b) Convert the reuglar expression to an NFA. (Do not show your work; write the answer in the space below.

5. Consider the following grammar:

$$
\begin{aligned}
S &\rightarrow u\,B\,D\,z \\
B &\rightarrow B\,v \\
B &\rightarrow w \\
D &\rightarrow E\,F \\
E &\rightarrow y \\
E &\rightarrow \epsilon \\
F &\rightarrow x \\
F &\rightarrow \epsilon
\end{aligned}
$$

(a) Compute nullable, FIRST and FOLLOW for all nonterminals.

	nullable	FIRST	FOLLOW
S			
B			
D			
E			
F			

(b) Compute the FIRST of the right-hand side of all productions.

	α	FIRST(α)
1	$S \rightarrow u\,B\,D\,z$	
2	$B \rightarrow B\,v$	
3	$B \rightarrow w$	
4	$D \rightarrow E\,F$	
5	$E \rightarrow y$	
6	$E \rightarrow \epsilon$	
7	$F \rightarrow x$	
8	$F \rightarrow \epsilon$	

(c) Fill in the LL(1) parse table for the grammar.

	w	u	v	x	y	z
S						
B						
D						
E						
F						

(d) Is the grammar LL(1)? Explain.

(e) Is the grammar LL(2)? Explain.

6. Consider the following grammar (uppercase letters are nonterminals):

$$
\begin{array}{ll}
0 & S' \to S\,\$ \\
1 & S \to a\,A\,d \\
2 & S \to b\,B\,d \\
3 & S \to a\,B\,c \\
4 & S \to b\,A\,c \\
5 & A \to c \\
6 & B \to c
\end{array}
$$

Answer all the following questions on the remainder of the page and on the back. Circle the word "yes" or "no" as appropriate, and don't forget to explain.

(a) Give a diagram of the LR(1) states and transitions.

(b) Give the entire LR(1) parsing tables.

(c) Is the grammar LR(1)? yes / no Explain.

(d) Give the entire LALR(1) parsing tables.

(e) Is the grammar LALR(1)? yes / no Explain.