$Id: cmps104a-2012q4-exam2.mm,v 1.73 2012-11-07 13:04:02-08 - - $

page 1 page 2 page 3 Total / 31

Please print clearly :

Name :

Login : @ucsc.edu

No books ; No calculator ; No computer ; No email ; No internet ; No notes ; No phone. Neatness counts ! Do your scratch work elsewhere and enter only your final answer into the spaces provided.

1. For each of the two expressions in the table at the right, draw parse trees, ***not*** abstract syntax trees for the E-T-F grammar. **[4✔]**

```
E → E + T
E → T
T → T * F
T → F
F → ( E )
F → i
```

`a * ( b + c )`	`a * b + c`

2. In the style of the LALR(1) handout discussed in class :
   (a) Construct the characteristic finite state machine (CFSM), sets of items and transition diagram, showing shifts, reductions, and acceptance. **[5✔]**
   (b) Is the grammar LR(0) ? If not, which state(s) have LR(0) conflicts ? **[1✔]**

```
0. S → $ A $
1. A → f ( A )
2. A → E
3. E → A + A
4. E → x
```

3. Write a grammar in `bison` that recognizes the following language. Show declarations from part 1 and the grammar from part 2 of a `bison` grammar file. Keep the number of rules small by using an ambiguous grammar, with suitable precedence and associativity declarations. Do not show any semantic actions. **[5✔]**
   (a) A program is a sequence of zero or more statements.
   (b) A statement is an assignment or a query.
   (c) An assignment is an identifier, an equal sign, an expression, and ends with a semi-colon.
   (d) A query is an identifier followed by a question mark.
   (e) An expression is a mixture of identifiers, numbers, left and right parentheses, and the operators `+`, `-`, `*`, and `/`, with the same precedence and associativity of C.
   (f) Make other assumptions sensible to a C programmer.

4. Write a lexical grammar using `flex` for the language described above. **[5✔]**
   (a) Identifiers are sequences of letters, digits, and underscores, but must begin with a letter, and may not end with an underscore.
   (b) A number is a sequence of one or more digits, with an optional decimal point which may appear before the first digit, after the last digit, or in between digits.
   (c) A number may have an optional exponent which is an upper- or lower-case letter E, an optional sign, followed by one or more digits.
   (d) Spaces, tabs, and newlines are ignored.
   (e) Comments begin with a hash (`#`) and are followed by all characters up to but not including a trailing newline character.

Multiple choice. To the ***left*** of each question, write the letter that indicates your answer. Write ***Z*** if you don't want to risk a wrong answer. Wrong answers are worth negative points. **[11✔]**

number of correct answers		× 1 =	$= a$
number of wrong answers		× ½ =	$= b$
number of missing answers		× 0 =	0
column total $c = \max(a - b, 0)$	11		$= c$

1. The following grammar:
   $A \to (\ A\ )$
   $A \to x$
   (A) is LR(0) but not SLR(1).
   (B) is SLR(1) but not LR(0).
   (C) is both LR(0) and SLR(1).
   (D) is neither LR(0) nor SLR(1).

2. Which of the following items in a state will cause a shift?
   (A) $E \to \bullet\ E + T$
   (B) $E \to E \bullet + T$
   (C) $E \to E + \bullet\ T$
   (D) $E \to E + T \bullet$

3. Given the following two rules in a grammar, and assuming a language like C, when the lookahead symbol is an **else**, we should mark the rule and the **else** [x] associative, so that we perform a [y] action.
   $S \to$ if $(\ E\ )\ S$ **else** S
   $S \to$ if $(\ E\ )\ S$
   (A) [x] = left, [y] = reduce
   (B) [x] = left, [y] = shift
   (C) [x] = right, [y] = reduce
   (D) [x] = right, [y] = shift

4. Which of the following grammars will recognize an arbitrary number of xs followed by one y?
   (A) $A \to A\ x \mid y$
   (B) $A \to A\ y \mid x$
   (C) $A \to x\ A \mid y$
   (D) $A \to y\ A \mid x$

5. Putting reserved words into the lexical grammar instead of into a separate reserved word table will have what effect?
   (A) Make the DFA run more slowly.
   (B) Put too many ε-transitions into the DFA.
   (C) Make the DFA larger.
   (D) Generate shift/reduce error messages.

6. Which grammar is ambiguous?
   (A) $A \to A\ A \mid x$
   (B) $A \to A\ x \mid x$
   (C) $A \to x\ A \mid x$
   (D) $A \to x\ x \mid x$

7. If we have a grammar $G = \langle V_N, V_T, P, S \rangle$, and the notation $\mid x \mid$ means "length of x" or "size of x": when reducing by the rule $(A \to \alpha) \in P$, how many items are popped from the stack?
   (A) $\mid P \mid$
   (B) $\mid V_N \mid$
   (C) $\mid V_T \mid$
   (D) $\mid \alpha \mid$

8. The semantic information is communicated from `yylex()` to `yyparse()` via what variable?
   (A) `yyleng`
   (B) `yylval`
   (C) `yytext`
   (D) `yywrap`

9. How many tokens are there in the following Java code?
```
/* Say hello */
out.printf ("%s%n", "Hello, World!");
```
   (A) 7
   (B) 9
   (C) 11
   (D) 13

10. If D is the set of languages recognizable by a DFA and N is the set of languages recognizable by an NFA, then:
    (A) $D \subset A$
    (B) $D = A$
    (C) $D \supset A$
    (D) None of the above.

11. The ____ mechanism, built ca. 150–100 BCE, is the oldest known complex scientific calculator, and is sometimes called the first known analog computer, with operational instructions written in Greek.
    (A) Ἀντιγόνη
    (B) Ἀντίκλεια
    (C) Ἀντικύθηρα
    (D) Ἀντίνοος

(answer 11 was in this space on test 1)