$Id: cmps104a-2009q4-exam1.mm,v 9.64 2009-11-03 17:22:19-08 - - $

page 1 page 2 page 3 Total / 30

Please print clearly :

Name :

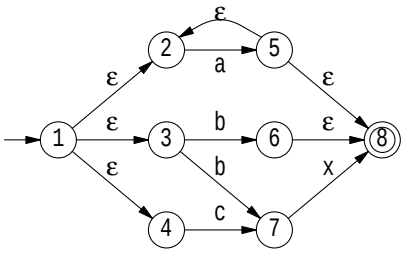
Login : @ucsc.edu

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. Given the following NFA :

   (a) Compute the ε-closure of each state and fill in the table. **[2✔]**

   (b) Then use the Subset Construction to convert the NFA to a DFA. Do your scratch work elsewhere and show only the final DFA. Use the Subset Construction ***exactly*** and do ***not*** minimize. Do not show trap (garbage) states. **[3✔]**



state *s*	ε-closure(*s*)
1	
2	
3	
4	
5	
6	
7	
8	

2. Using as few states as possible, draw ***deterministic*** finite αὐτόματα for each of the following `flex` regexes. Use any method you like and show only the final answer. **[5✔]**

   (i) `a+|b+`

   (ii) `a+b|c`

   (iii) `ab+cd`

   (iv) `a*b*c`

   (v) `a?b+c*`

3. Write `flex` regular expressions for each of the following. **[3✔]**

   (i) A string literal which starts and ends with a quote (") and may have zero or more characters between them. A character may be any character except for a newline, quote, or backslash (\). Alternately, it may contain one of the escape sequences `\"`, `\n`, or `\\`.

   (ii) A floating point literal which consists of a decimal point with one or more digits in front of it and one or more digits after it. An exponent is optional, and consists of an upper- or lower-case letter E, followed optionally by a plus or minus sign, followed by one or more digits.

   (iii) A Perl identifier, which starts with a dollar sign (`$`), an at sign (`@`), or a percent sign, (`%`), followed by any sequence of upper- and lower-case letters, digits, and underscores.

4. Consider a grammar similar to ETF, but with different precedence and associativity.

   (a) Define an ***unambiguous*** grammar with the start symbol A. An ***A*** is a sequence of one or more ***B***s each separated from the next by an asterisk (`*`), binary operator, which is right associative. A ***B*** is a sequence of one or more ***C***s connected by the binary operator slash (`/`), which is left associative, and has a higher precedence than `*`. A ***C*** is either an identifier or an `A` inside of parentheses. **[3✔]**

   (b) For each of two expressions presented here, draw the abstract syntax tree and the parse tree. Use the grammar specified above, not ETF and not any grammar obvious to a C programmer. **[4✔]**

expression : `a/b/c*d`	`(a*b)/e`
AST :	
parse tree :	

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. If an NFA is constructed from a regex r and is used to scan a string s, then its memory requirement M and running time T are :
   (A) $M = O(2^{|r|})$; $T = O(|s|)$
   (B) $M = O(|r|)$; $T = O(|r| \times |s|)$
   (C) $M = O(|r| \times |s|)$; $T = O(|r|)$
   (D) $M = O(|s|)$; $T = O(2^{|r|})$

2. Which **flex** pattern will match the following string ?
   aaaaaabaababab
   (A) **a*b***
   (B) **a+b+**
   (C) **(a+b)***
   (D) **(ab)+**

3. The variable that is used to send semantic information from **yylex** to **yyparse** is :
   (A) **yydebug**
   (B) **yylval**
   (C) **yytext**
   (D) **yywrap**

4. For a grammar $G = \langle V_N, V_T, P, S \rangle$, P is a set of productions of the form $A \to \beta$, where :
   (A) $A \in S^*$; $\beta \in V_T^*$
   (B) $A \in V_N$; $\beta \in (V_N \cup V_T)^*$
   (C) $A \in V_T$; $\beta \in (V_N \cap V_T)^+$
   (D) $A \in V_N^+$; $\beta \in V_T^*$

5. In case of ambiguity, and assuming that the rule at the top of the stack and the lookahead symbol both have precedence and associativity, then the parser should shift if the rule has a [x] precedence than the incoming symbol, or if both precedences are the same and they are [y] associative.
   (A) [x] = higher ; [y] = left
   (B) [x] = higher ; [y] = right
   (C) [x] = lower ; [y] = left
   (D) [x] = lower ; [y] = right

6. If a minimized DFA is constructed from each of the following **flex** regexes, which will have the fewest states ?
   (A) **x**
   (B) **x***
   (C) **x+**
   (D) **x?**

7. The **flex** expression **ab|c*** is equivalent to :
   (A) **(((ab)|c)*)**
   (B) **((a(b|c))*)**
   (C) **((ab)|(c*))**
   (D) **(a((b|c)*))**

8. Which of the following **flex** regexes will match an ANSI C identifier ?
   (A) **[A-Za-z]_[A-Za-z_0-9]***
   (B) **[A-Za-z_0-9]*[A-Za-z_]**
   (C) **[A-Za-z_][A-Za-z_0-9]***
   (D) **[_A-z][A-z0-9_]+**

9. Which is the weakest set of grammars in the Chomsky hierarchy that is capable of determining whether or not parentheses are matched ?
   (A) unrestricted
   (B) context-sensitive
   (C) context-free
   (D) regular

10. What kind of analysis does **bison** perform on a grammar ?
    (A) LR(0)
    (B) SLR(1)
    (C) LALR(1)
    (D) LR(1)

11. The "___ book" is a name commonly used to refer to the textbook used in this course.
    (A) Camel
    (B) Daemon
    (C) Dragon
    (D) Raccoon

The Antikythera 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. http://en.wikipedia.org/wiki/Antikythera_mechanism

Total : 31 ✔