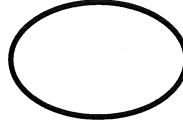
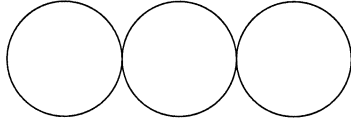


page 1

page 2

page 3

Total / 32



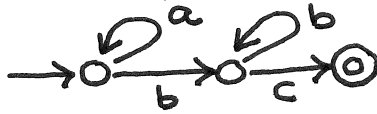
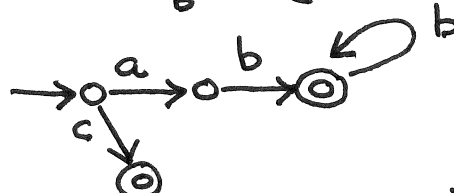
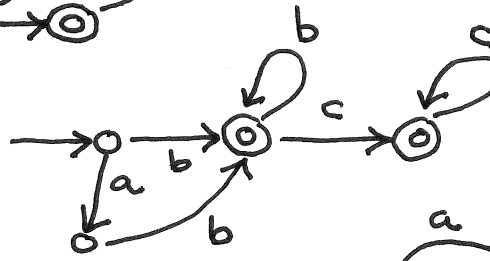
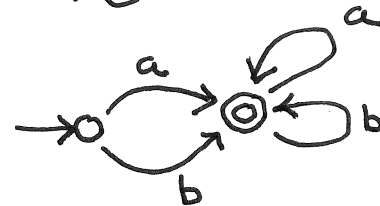
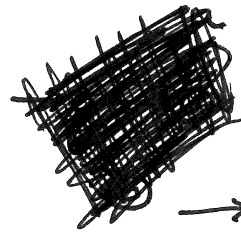
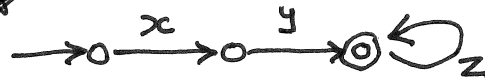
Please print clearly:

Name: SOLUTION

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. Draw *deterministic* finite αὐτόματα for each of the following flex regular expressions. Use as few states as possible [5✓]

(i) a^*b+c (ii) $ab+|c$ (iii) $a?b+c^*$ (iv) $(a+|b)^+$ (v) xyz^* 

2. Given the nondeterministic finite αὐτόματων shown here :

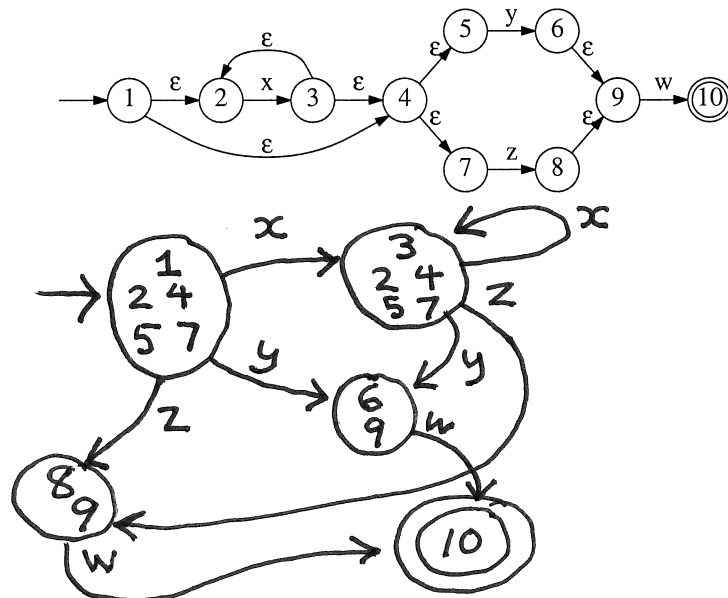
(a) Write the regular expression that was used by Thompson's construction to create this NFA. [1✓]

$$x^*(y|z)w$$

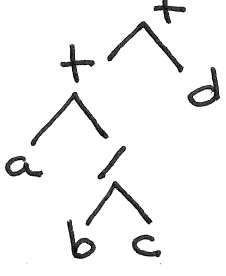
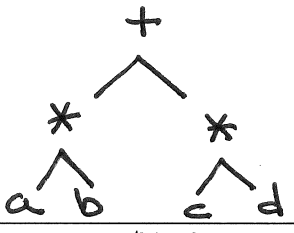
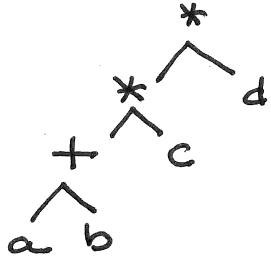
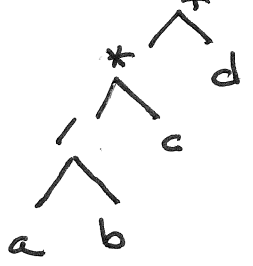
(b) Fill in the table of ϵ -closures for each state. [2✓]

(c) Use the subset algorithm to construct the equivalent *deterministic* finite αὐτόματων. Inside each state of the DFA, write the numbers of the NFA states to which it corresponds. Do not minimize. Draw the DFA underneath the NFA. [2✓]

state s	ϵ -closure (s)
1	1 2 4 5 7
2	2
3	3 2 4 5 7
4	4 5 7
5	5
6	6 9
7	7
8	8 9
9	9
10	10



3. Given the ETF grammar listed on the left, draw the abstract syntax trees (ASTs), not the parse trees, for each of the expressions. [4✓]

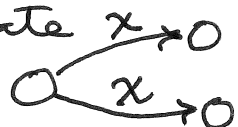
$E \rightarrow E + T$ $E \rightarrow T$ $T \rightarrow T * F$ $T \rightarrow F$ $F \rightarrow (E)$ $F \rightarrow i$	a+b/c*d	a*b+c*d
		
	(a+b)*c*d	a/b*c+d
		

4. State the two things that are permitted for a nondeterministic finite αὐτόματον (NFA) that are prohibited in a deterministic finite αὐτόματον (DFA). Draw two *small* diagrams of NFAs that illustrate each of these things. [1✓]

ϵ -transitions



more than one out trans with same label in any given state



5. Write **flex** regular expressions for each of the following: [5✓]

- a. A hexadecimal integer constant in C, C++, or Java. (They are the same in all three languages.)

$0[Xx][0-9a-fA-F]^+$

- b. A string constant in the FORTRAN language, which starts with a single quote (') and ends with a single quote, and which may contain any characters in between, except for a newline. But if a single quote appears in the string, it must be doubled, for example, 'Don't', doubles the single quote.

$'([^\backslash n]|'')^*'$

- c. A C++ decimal integer constant which begins with a decimal digit other than zero, and is followed by any number of decimal digits, including none. Between any pair of digits may occur a single quote (').

$[1-9]('?[0-9])^+$

- d. A C, C++, or Java comment of the double slash (//) kind, which includes all following characters up to but not including the next newline character.

$//.*$ or $//[^\backslash n]^*$

- e. An identifier in some language. It consists of any number of upper- or lower-case letters, digits, and underscores, but may not begin with a digit.

$[A-Za-z_][A-Za-z_0-9]^*$

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. [12✓]

number of correct answers		$\times 1 =$	$= a$
number of wrong answers		$\times \frac{1}{2} =$	$= b$
number of missing answers		$\times 0 =$	0
column total	12		$= c$
$c = \max(a - b, 0)$			

1. Which of the following semantic actions in a **flex** grammar is obviously wrong?

(A) { return '+'; }
 (B) { return *yytext; }
 (C) { return PLUS; }
 (D) { return "+"; }

2. If N is the set of languages that can be recognized by an NFA, and D is the set of languages that can be recognized by a DFA, then :

(A) $N \subset D$
 (B) $N = D$
 (C) $N \supset D$
 (D) None of the above.

3. What is prohibited in a deterministic finite αὐτόματον?

(A) cycles in the graph
 (B) epsilon transitions
 (C) multiple final states
 (D) reserved words

4. The parser generated by **bison** is a :

(A) finite-state machine
 (B) linear-bounded automaton
 (C) push-down automaton
 (D) Turing machine

5. The scanner generated by **flex** is a :

(A) finite-state machine
 (B) linear-bounded automaton
 (C) push-down automaton
 (D) Turing machine

6. For a context free grammar $G = \langle V_N, V_T, P, S \rangle$: P is a set containing productions of the form $A \rightarrow \beta$, where :

(A) $A \in V_N$ and $\beta \in (V_N \cup V_T)^*$
 (B) $A \in V_N$ and $\beta \in (V_N \cup V_T)^+$
 (C) $A \in V_N$ and $\beta \in (V_N \cup V_T)^*$
 (D) $A \in V_T$ and $\beta \in (V_N \cup V_T)^*$

7. If an **NFA** constructed from a regular expression r , whose length is $|r|$, is used to scan a string s , whose length is $|s|$, then its memory requirements M and running time to scan the string T are :

(A) $M = O(2^{|r|})$ and $T = O(|r|)$
 (B) $M = O(2^{|r|})$ and $T = O(|r| \times |s|)$
 (C) $M = O(|r|)$ and $T = O(|r| \times |s|)$
 (D) $M = O(|r|)$ and $T = O(|s|)$

8. What **flex** regex describes an identifier in C?

(A) $[a-zA-Z_0-9]^*[a-zA-Z_]$
 (B) $[a-zA-Z_0-9]^+$
 (C) $[a-zA-Z_]^+$
 (D) $[a-zA-Z_][a-zA-Z_0-9]^*$

9. What is the same as the **flex** pattern . (dot)?

(A) $[/n]$
 (B) $[\backslash n]$
 (C) $[\wedge/n]$
 (D) $[\wedge\backslash n]$

10. The subset algorithm :

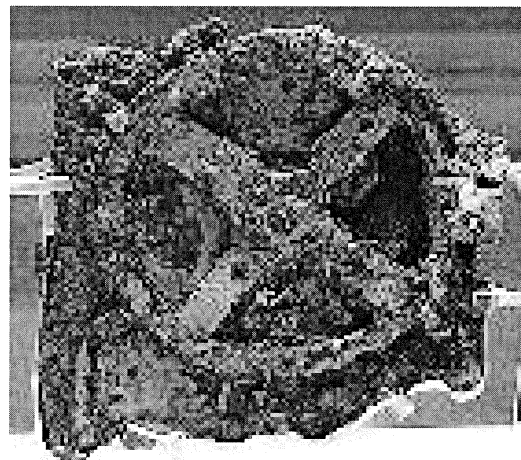
(A) converts a regular expression into an NFA
 (B) converts an NFA into a DFA
 (C) converts an NFA into a regular expression
 (D) minimizes a DFA

11. In a C compiler, removing comments from the source code is done by which phase?

(A) code generator
 (B) parser
 (C) preprocessor
 (D) scanner

12. Is half of two plus two equal to two or three?

(A) two
 (B) three
 (C) yes
 (D) no



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