



Please PRINT using keyboard letters :

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. Draw a *deterministic* finite αὐτόματον for each of the following **flex** expressions. Use as few states as possible. If you use Thompson's construction, etc., for this question, do it on scratch paper. Show only the final answer. **[5✓]**

(i)  $ab^+|c$ (ii)  $a^*bc^*$ (iii)  $ab?c^*$ (iv)  $d^+.d^*$ (v)  $x(x|d)^*$ 

2. Write a scanner using **flex**, which recognizes the following language. Assume a global variable **lineno**, which is used to count lines. Assume, but do not write, a function **setyylval**, which takes a token code as an argument and properly assigns to **yylval**. **[5✓]**

- (a) A sequence of one or more decimal digits is referred to by the token code **INT**.
- (b) The four arithmetic operators **+**, **-**, **\***, and **/**, for which their own character codes are used.
- (c) Spaces, tabs, and newlines are ignored, but newlines are counted.
- (d) A comment is two slashes (**//**) followed by all characters up to but not including the next newline.
- (e) Identifiers (**IDENT**) consist of letters (upper or lower case) and digits, but may not begin with a digit.
- (f) Any invalid characters should cause the function **badchar** to be called. Nothing is returned in this case.

3. Draw abstract syntax trees for each of the following C expressions. [2✓]

(a)  $a*b*c+d$

(b)  $(a+b)*(c+d)$

4. Given the grammar shown here, write a hand-coded parser for **expr**. Assume the existence of the functions **match**, **maketree**, and **term**. Remember that the notation  $\{ \}$  indicates zero or more occurrences of what is inside the braces. [2✓]

$\text{expr} \rightarrow \text{term} \{ + \text{term} \}$

5. Given the regular expression  $ab|c^*$

- (a) Using Thompson's construction, draw the NFA, and fill in each state with a small integer. [2✓]
- (b) For each state, list its  $\epsilon$ -closure. [2✓]
- (c) Convert the NFA into a DFA using the subset algorithm. Do not minimize. [2✓]

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		$\times 1 =$	$= a$
number of wrong answers		$\times \frac{1}{2} =$	$= b$
number of missing answers		$\times 0 =$	$0$
column total $c = \max(a - b, 0)$	11		$= c$

- What **flex** expression matches one or more **x**s followed by one **y**?  
(A) **x\*y**  
(B) **x+y**  
(C) **x?y**  
(D) **x|y**
- How many tokens are there in the following line of Java code?  
`out.printf ("Hello, world!\n");`  
(A) 6  
(B) 7  
(C) 8  
(D) 9
- Which one of the following patterns can accept the fewest number of letters?  
(A) **a\*b\***  
(B) **a\*b+**  
(C) **a+b\***  
(D) **a+b+**
- According to Thompson's construction, which of the following will generate the largest number of states?  
(A) **(a)**  
(B) **a\***  
(C) **ab**  
(D) **a|b**
- What kind of language does a parser generally use?  
(A) regular  
(B) context free  
(C) context sensitive  
(D) unrestricted
- In C, what is printed by:  

```
int main () {
    printf ("%s\n", "3333" + 2);
    return 0;
}
```

  
(A) 33  
(B) 3333  
(C) Nothing: this program won't compile.  
(D) Segmentation Fault (core dumped)
- Which one of the following **flex** actions is obviously wrong?  
(A) `return "=";`  
(B) `return '=';`  
(C) `return *yytext;`  
(D) `return EQ;`
- What is the worst case number of states in an NFA constructed via Thompson's construction for a regex  $r$  when used to scan a string  $s$ ? (The notation  $|r|$  means: the length of  $r$ .)  
(A)  $O(2^{|r|})$   
(B)  $O(|r|)$   
(C)  $O(|r|^2)$   
(D)  $O(|r| \times |s|)$
- Putting reserved words into a **flex** grammar instead of embedding them in the string table will have what effect?  
(A) Make the DFA run more quickly..  
(B) Make the DFA run more slowly.  
(C) Make the DFA take up less space in memory.  
(D) Make the DFA take up more space in memory.
- Which of the following context-free grammar rules shows that **+** is left associative?  
(A)  $E + T \rightarrow E$   
(B)  $E \rightarrow E + E$   
(C)  $E \rightarrow E + T$   
(D)  $E \rightarrow T + E$
- Who wrote the first C compiler?  
(A) Alfred Aho, Peter Weinberger, and Brian Kernighan.  
(B) Brian Kernighan and Dennis Ritchie.  
(C) Grace Hopper and John Backus.  
(D) Ken Thompson and Dennis Ritchie.