

Homework #2

Due Date: Thursday, October 13, 2005, 2:00

Your Name: _____

Your Email: _____

1. Do tokens appear as Terminal or Non-terminal symbols in the grammar?

2. Will lexical errors (e.g., typos) always be caught by the Lexer?

3. Whether or not the Lexer detects an error, what is always returned to the Parser?

4. If OS system calls that read a single character at a time are more expensive than calls that read large chunks of bytes, what might the Lexer do?

5. In order to improve efficiency, one technique involves placing a special value just past the end of an array or character buffer. What is this special value called?

6. When dealing with strings of characters (or symbols), the set of possible characters (or symbols) is called... what? _____

7. What Greek letter is used for this set? _____

8. Is a string always finite (in what we'll be doing)? _____

9. What does ϵ (epsilon) stand for? _____

10. How do we denote the length of string s ? _____

11. What is a language? _____
12. Is a language always finite (in what we'll be doing)? _____
13. Let s be the string "dog". Give all prefixes.

14. Let s be the string "dog". Give all suffixes.

15. Let s be the string "dog". Give all substrings.

16. Let s be the string "dog". Give all subsequences.

17. What is the identity for string concatenation? _____
18. Let s be the string "dog". Give s^0 . _____
19. Let s be the string "dog". Give s^3 . _____
20. Technically, is s^* a string? Why? _____
21. Let L_1 and L_2 be these two languages:
 $L_1 = \{ wx, yz \}$
 $L_2 = \{ ab, cd \}$
Give $L_1 \cup L_2$. _____
23. Give $L_1 L_2$. _____
24. Give $(L_1)^0$. _____
25. Give $(L_1)^1$. _____
26. Give $(L_1)^2$. _____

27. Is L^* a language? _____

28. How do we denote the Kleene closure of L ? _____

29. How do we denote the positive closure of L ? _____

30. If L contains ϵ (epsilon), will the Kleene closure of L contain ϵ ?

31. If L does not contain ϵ (epsilon), will the Kleene closure of L contain ϵ ?

32. If L contains ϵ (epsilon), will the positive closure of L contain ϵ ?

33. If L does not contain ϵ (epsilon), will the positive closure of L contain ϵ ?

34. Let L and D be the following two languages:

$L = \{ a, b, c, \dots, z \}$

$D = \{ 0, 1, \dots, 9 \}$

Give a language (using \cup , concatenation, and Kleene closure) of “The set of strings beginning with a digit, containing one or more letters and/or digits, followed by a digit.” Another way to describe this is “The set of strings of letters and digits, with at least 3 symbols, starting with a digit and ending with a digit.” Here are some examples from the desired language:

567

8x9

5abc6

34h6km9

36. Add full parentheses to the following regular expressions. The first one is an example:

a b c Answer: (a b) c

a b c * _____

a b | c d | e f _____

a | b * _____

a * | b _____

a * b c _____

a b * | c * d _____

37. Give a four different strings in the language described by the regular expression:

$(b | c) d (e^* | f^*) g$

38. Give a regular expression for the set of strings that begin with zero or more a's, followed by 2 or more b's, followed by c or d.

39. If R_1 and R_2 are regular expressions denoting the languages $L(R_1)$ and $L(R_2)$, then what language is denoted by the regular expression $R_1 R_2$?

40. What language is denoted by the regular expression $R_1 | R_2$?

41. What language is denoted by the regular expression R_1^* ?

42. What is a "regular set" or "regular language"?

43. When are two regular expressions considered equal (i.e., “equivalent”)?

44. Simplify these regular expressions:

$a | b | a | a | b$ _____

$ab | ac | ad | ae$ _____

$a \epsilon b \epsilon c d$ _____

$((a b c)^*)^*)^*$ _____

45. Can a sequence of regular definitions be recursive? _____

46. Use shorthand notations for to simplify these regular expressions:

$a b b^* c$ _____

$a (b | \epsilon) c$ _____

$0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9$ _____

47. Consider the set of string of balanced parentheses. Eg

$((()()())()$

Can this language be described by a regular expression? _____