Faculty of Technology and Engineering

Computer Engineering / Computer Science and Engineering

Date : 16 / 12 / 2024

Practical List

Academic Year	:	2024-2025	Semester	:	6
Course code	:	CE365 CSE313	Course name	:	Compiler Construction Design of Language Processor

Sr. No.	Aim	Hr	CO
1.	**Practical Definition** String Validation Against Regular Expression **Objective** To implement a program that validates a user-input string against the regular expression a*bb. The program should determine whether the input string is valid or invalid based on the defined pattern. **Language Constraint** The program must be implemented in C language. **Input Requirements** • Accept a character string from the user. • Ensure the input is terminated with a newline character. **Expected output** • If the input string matches the pattern a*bb, the program should output: "Valid String". • If the input string does not match the pattern, the program should output: "Invalid String". **Sample input output** Input: aaabb → Output: Valid string Input: abab → Output: Invalid string **Testcases** ^ | bbbb | aaa | baaabb | aaabbb baaabb | aaaab | abbabb | abb | aaaaabb	2	1

2.	**Practical Definition** String Validation Using Finite Automata **Objective** To implement a program that validates a given string against rules defined in terms of finite automata. **Language Constraint** The program can be implemented in any programming language **Input requirement** • Accept rules in the form of finite automata (e.g., states, transitions, start state, accept states) as input. • Accept a string to be validated against the provided finite automata rules. **Expected output** • If the string adheres to the rules of the finite automata, the program should output: "Valid String". • If the string does not adhere to the rules, the program should output: "Invalid String". **Sample input output** Input: Number of input symbols : 2 Input symbols : a b Enter number of states : 4 Initial state : 1 Number of accepting states : 1 Accepting states : 2 Transition table : 1 to a -> 2 1 to b -> 3 2 to a -> 1 2 to b -> 4 3 to a -> 4 3 to b -> 1 4 to a -> 3 4 to b -> 2 Input string : abbabab Output: Valid string **Testcases** • String over 0 and 1 where every 0 immediately followed by 11 • string over a b c, starts and end with same letter. • String over lower-case alphabet and digits, starts with alphabet only.	2	1

3.	**Practical Definition** Implementation of a Lexical Analyzer for C Language Compiler **Objective** To design and implement a lexical analyser, the first phase of a compiler, for the C programming language. The lexical analyser should perform the following tasks: (1) tokenizing the input string (2) removing comments (3) removing white spaces (4) entering identifiers into the symbol table (5) generating lexical errors. **Language Constraint** The program can be implemented in any programming language **Input requirement** • Accept a C source code file. • The input can contain keywords, identifiers, constants, strings, punctuation, operators, comments, and white spaces. **Expected output** • Tokenized output categorizing tokens into five types: keyword, identifier, constant, punctuation, and operator. • Symbol table with all identified identifiers stored. • Detected lexical errors **Sample input output**	4	1

Input	Output	

int main() {
int a = 5 , 7H;
// assign value
char b = 'x';
/* return
value */
return a + b;
}
	TOKENS
	Keyword: int
	Identifier: main
	Punctuation: (
	Punctuation:)
	Punctuation: {
	Keyword: int
	Identifier: a
	Operator: =
	Constant: 5
	Punctuation : ,
	Punctuation: ;
	Keyword: char
	Identifier: b
	Operator: =
	Constant: 'x'
	Punctuation: ;
	Keyword: return
	Identifier: a
	Operator: +
	Identifier: b
	Punctuation: ;
	Punctuation: }
		LEXICAL ERRORS
		Line 2 : 7H invalid lexeme
		
		SYMBOL TABLE ENTRIES
		1) a
		2) b

Testcases

/* salary calculation*/
void main( )
{
	// user defined data type
	struct student
	{

```
long int bs , da , hra , gs;
//take basic salary as input
scanf("%ld",&bs);
//calculate allowances
da=bs*.40;
hra=bs*.20;
gs=bs+da+hra;
// display salary slip
printf("\n\nbs  : %ld",bs);
printf("\nda  : %ld",da);
printf("\nhra : %ld",hra);
printf("\ngs  : %ld",gs);
}
```

```
int id;
float cgpa;
}
void main( )
{
student s;
s.id = 10;
s.cgpa = 8.7;
}
```

```
//function prototype
void add ( int , int );
void main( )
{
int a , b;
a = 10;
b = 20;
// function call
add ( a , b );
}
void add ( int x , int y )
{
return x + y;
}
```

4. **Practical Definition** | 4 | 1

String validation using Lax tool

Objective - 1

Write a program to identify and extract all numbers from input string and display them one by one in new line.

Language Constraint

Lex (Lexical analyser generator)

Input requirement

- Accept a character string, mix of text and numbers, from the user.
- Ensure the input is terminated with a newline character.

Expected output

The program should print out each number found in the input, each on a new line.

Sample input output

Input	Output
a1b22c3	1 22 3

Testcases

power operation -> 12 ** 3 = 1728
You multiply 804569 with 1 then will be :

Objective - 2

Write a program to replace the word "charusat" with “university” in the input text.

Language Constraint

Lex (Lexical analyser generator)

Input requirement

- Accept a character string from the user where the word "charusat" may appear multiple times.
- Ensure the input is terminated with a newline character.

Expected output

The program should print the input text with all occurrences of "charusat" replaced by “university”.

Sample input output

Input	Output
This is charusat.	This is university.

Testcases

Charusat is in Anand district.	I am doing my BTech from CHARSAT.
Charusat , What is charusat?	Every where it is charusat , charusat and only charusat.

Objective – 3
Write a program to count number of characters, word and lines from the input file.

Language Constraint
Lex (Lexical analyser generator)

Input requirement
Read contain from a text file containing multiple word and lines.

Expected output
The program should print total number of characters (including spaces), words (separated by white spaces), lines (end with new line symbol).

Sample input output

Input	Output
The 45 is odd number.	Characters : 22 Words : 5 Line : 1

Testcases

I want to calculate a number. The number of characters, words and lines. All know that \n is ending character of line.
45 + 89 =40

Objective – 4
Write a program which validate the password as per given rules.
- length can be 9 to 15 characters
- includes lower case letter, upper case letter, digit, symbols (*, ; # $ @)
- minimum count for each category must be one

Language Constraint
Lex (Lexical analyser generator)

Input requirement
- Accept a character string from the user which is mix of letters, numbers and symbols.
- Ensure the input is terminated with a newline character.

Expected output
- If the password meets the given rules, the program should print "Valid password".
- If the password does not meet the rules, the program should print "Invalid password".

Sample input output

Input	Output
a@1T	Invalid password

Testcase

aB1@	aaBB11,#cdefg2345	CHARUSAT	
Charusat	CHArusat123	Cspit-2024	
Charusat@2024	Charu$at@20#24	charu*sAT;22	

5.

Practical Definition

Implementation of a Lexical Analyzer for C Language Compiler

Objective

To design and implement a lexical analyser to perform 1st, 2nd, 3rd, and 5th task as per the list given in practical 2.

Language Constraint

Lex (Lexical analyser generator)

Input requirement

- Accept a C source code file.
- The input can contain keywords, identifiers, constants, strings, punctuation, operators, comments, and white spaces.

Expected output

- Tokenized output categorizing tokens into six types: keyword, identifier, constant, string, punctuation, and operator.
- Detection and reporting of lexical errors

Sample input output

Input	Output	

int main() {
int a = 5 , 7H;
// assign value
char b = 'x';
/* return
value */
return a + b;
}
	TOKENS
	Keyword: int
	Identifier: main
	Punctuation: (
	Punctuation:)
	Punctuation: {
	Keyword: int
	Identifier: a
	Operator: =
	Constant: 5
	Punctuation : ,
	Punctuation: ;
	Keyword: char
	Identifier: b
	Operator: =
	Constant: 'x'
	Punctuation: ;
	Keyword: return
	Identifier: a
	Operator: +
	Identifier: b
	Punctuation: ;
	Punctuation: }
		LEXICAL ERRORS
		Line 2 : 7H invalid lexeme
		
		SYMBOL TABLE ENTRIES
		1) a
		2) b

Testcases

/* salary calculation*/
void main( )
{
long int bs , da , hra , gs;
//take basic salary as input
scanf("%ld",&bs);
//calculate allowances
	// user defined data type
	struct student
	{
	int id;
	float cgpa;
	}
	void main( )

2 | 1

		da=bs*.40;
		hra=bs*.20;
		gs=bs+da+hra;
		// display salary slip
		printf("\n\nbs : %ld",bs);
		printf("\nda : %ld",da);
		printf("\nhra : %ld",hra);
		printf("\ngs : %ld",gs);
		}
			{
			student s;
			s.id = 10;
			s.cgpa = 8.7;
			}		

		//function prototype
		void add (int , int);
		void main( )
		{
		int a , b;
		a = 10;
		b = 20;
		// function call
		add (a , b);
		}
		void add (int x , int y)
		{
		return x + y;
		}			

<table>
<tr>
<td>6.</td>
<td>
Practical definition
String validation using Recursive Descent Parsing (RDP)

Objective
Implement a Recursive Descent Parser (RDP) to validate an input string against the given grammar.
S → (L) | a
L → S L’
L’ → , S L’ | ϵ

Language constraint
The program can be implemented in any programming language

Input Requirement
A string that needs to be validated against the grammar

Expected output
• If the input string is valid according to the grammar, the program should print "Valid string".
• If the input string is invalid according to the grammar, the program should print "Invalid string".

Sample input output
<table>
<tr><th>Input</th><th>Output</th></tr>
<tr><td>(a)</td><td>Valid string</td></tr>
</table>

Testcases
<table>
<tr><td>a</td><td>(a)</td><td>(a,a)</td><td>(a,(a,a),a)</td><td>(a,a),(a,a)</td></tr>
<tr><td>a)</td><td>(a</td><td>a,a</td><td>a,</td><td>(a,a),a</td></tr>
</table>
</td>
<td>2</td>
<td>2</td>
</tr>
</table>

7.	**Program definition** Computing First and Follow Sets for a Context-Free Grammar (CFG) **Objective** Develop a program computes the First and Follow sets for all non-terminal symbols in for the below given grammar. S → A B C | D A → a | ε B → b | ε C → (S) | c D → A C **Language Constraint** The program can be implemented in any programming language **Input requirement** No input **Expected output** First(S) = {a, b, (, c} First(A) = {a, ε} First(B) = {b, ε} First(C) = {(, c} First(D) = {a, (, c} Follow(S) = {), $} Follow(A) = {b, (, c} Follow(B) = {(, c} Follow(C) = {), $} Follow(D) = {), $}	2	2

<table>
<tr>
<td>8.</td>
<td>
Program definition
Predictive Parsing Table Construction and LL(1) Grammar Validation

Objective
Develop a program to construct a predictive parsing table for the given grammar. The program should analyse the generated parsing table to determine whether the grammar is LL(1) or not. If the grammar is LL(1), the program should also validate an input string against the given grammar.

Language Constraint
The program can be implemented in any programming language

Input requirement
• First() and Follow() generated by practical 7
• A string that needs to be validated against the grammar

Expected output
• A predictive parsing table generated for the given grammar.
• A message indicating whether the grammar is LL(1) or not.
• If the input string is valid according to the grammar, the program should print "Valid string".
• If the input string is invalid according to the grammar, the program should print "Invalid string".

Testcases
<table>
<tr><td>abc</td><td>ac</td><td>(abc)</td><td>c</td><td>(ac)</td></tr>
<tr><td>a</td><td>( )</td><td>(ab)</td><td>abcabc</td><td>b</td></tr>
</table>
</td>
<td>3</td>
<td>2</td>
</tr>
</table>

9.	**Program definition** String parsing using YACC **Objective** Develop a YACC program to validate input strings based on the given grammar. The program should parse the string using the grammar rules and determine whether the string is valid or invalid. S → i E t S S' | a S' → e S | ε E → b **Language Constraint** YACC (Syntex Analyser generator) **Input requirement** An input string to validate based on the provided grammar. **Expected output** • If the input string is valid according to the grammar, the program should print "Valid string". • If the input string is invalid according to the grammar, the program should print "Invalid string". **Sample input output** Input: i b t — Output: Invalid string **Testcases** ibtai, ibtaea, a, ibtibta, ibtaeibta ibti, ibtaa, iea, ibtb, ibtibt	3	2

10.	**Program definition** Evaluating Arithmetic Expression with Bottom-Up Approach Using SDD **Objective** Develop a program to evaluate arithmetic expressions containing operators using a bottom-up parsing approach and below given Syntax-Directed Definitions (SDD) for semantic evaluation. The program will compute the result of the expression by building a parse tree using and will incorporate semantic rules to evaluate sub-expressions during parsing.	2	4

L → E n	Print (E.val)
E → E + T	E.Val = E.val + T.val
E → E - T	E.Val = E.val - T.val
E → T	E.val = T.val
T → T * F	T.val = T.val * F.val
T → T / F	T.val = T.val / F.val
T → F	T.val = F.val
F → G ^ F	F.val = G.val ^ F.val
F → G	F.val = G.val
G → (E)	G.val = E.val
G → digit	G.val = digit.lexval

Language Constraint
An input string

Input requirement
An arithmetic expression in the form of a string that can contain

- Operands: Integers (e.g., 3, 5, 10) or decimals (e.g., 2.5, 0.75)
- Operators: +, -, *, /, ^ for addition, subtraction, multiplication, division, and exponentiation
- Parentheses: (and) for grouping sub-expressions

Expected output

- The evaluated result of the arithmetic expression.
- If the input expression is invalid, the program should display “Invalid expression”.

Sample input output

Input	Output
(3 + 5) * 2 ^ 3	64

Testcases

(3 + 5) * 2	3 + 5 * 2	3 + 5 * 2 ^ 2	3 + (5 * 2)	3 + 5 ^ 2 * 2
3 * (5 + 2)	(3 + 5) ^ 2	3 ^ 2 ^ 3	3 ^ 2 + 5 * 2	3 + ^ 5
(3 + 5 * 2	(3 + 5 * 2 ^ 2 - 8) / 4 ^ 2 + 6			

11. | 2 | 4

Program definition
Generate Intermediate Code Using Quadruple Table

Objective
Develop a program that break down the input string according to the grammar and produce a sequence of quadruples representing the intermediate code for the expression.
E → E + T | E – T | T
T → T * F | T / F | F
F → (E) | digit

Language Constraint
The program can be implemented in any programming language

Input requirement
An arithmetic expression in the form of a string that can contain
- Operands: Integers (e.g., 3, 5, 10) or decimals (e.g., 2.5, 0.75)
- Operators: +, -, *, / for addition, subtraction, multiplication, division, and exponentiation
- Parentheses: (and) for grouping sub-expressions

Expected output
A quadruple table representing the intermediate code for the given expression

Sample input output

Input	Output			
9 + 42 * 8	Operator	Operand 1	Operand 2	Result
	*	42	8	t1
	+	9	t1	t2

Testcases

5 + 6 – 3	7 – (8 * 2)	(9 – 3) + (5 * 4 / 2)
(3 + 5 * 2 - 8) / 4 – 2 + 6		86 / 2 / 3

<table>
<tr>
<td>12.</td>
<td>
Program definition
Code Optimization Using Constant Folding

Objective
Develop a program that identifies constant expressions at compile-time and replaces them with their evaluated results to enhance execution efficiency.

Language Constraint
The program can be implemented in any programming language

Input requirement
An arithmetic expression in the form of a string that can contain
• Operands: Integers (e.g., 3, 5, 10) or decimals (e.g., 2.5, 0.75) or variables (e.g. a , abc , cgpa)
• Operators: +, -, *, / for addition, subtraction, multiplication, division, and exponentiation

Expected output
Display the optimized expression after applying constant folding

Sample input output
<table>
<tr><td>Input</td><td>Output</td></tr>
<tr><td>5 + x - 3 * 2</td><td>5 + x - 6</td></tr>
</table>

Testcases
<table>
<tr><td>2 + 3 * 4 - 1</td><td>x + (3 * 5) - 2</td><td>(22 / 7) * r * r</td></tr>
</table>
</td>
<td>2</td>
<td>5</td>
</tr>
</table>