GOVT. MODEL ENGINEERING COLLEGE

THRIKKAKKARA, ERNAKULAM

.....

NAME: MUHAMMED BILAL M N

BRANCH: COMPUTER SCIENCE AND ENGINEERING

SEMESTER : 7 **ROLL NO : 22CSC69**

Certified that this is the bonafide work done by

.....

Staff- in Charge ***Head of the Department***

Register No..... Date:.....

Year and month..... Thrikkakkara

Internal Examiner **External Examiner**

INDEX

PROGRAM

```
#include<stdio.h>
#include<string.h>
char result[20][20], copy[3], states[20][20];
void add_state(char a[3], int i){
  strcpy(result[i], a);
}
void display(int n){
  int k=0;
  printf("\nEpsilon closure of %s = { ",copy);
  while(k < n){
    printf(" %s,",result[k]);
    k++;
  }
  printf(" } \n");
}
int main(){
  FILE* ipfile;
  ipfile=fopen("input.txt","r");
  char state[3], state1[3],state2[3], input[3];
  int end,i=0,n,k=0;
  printf("\n Enter the no of states: ");
  scanf("%d",&n);
  printf("\n Enter the states:");
  for(k=0;k<3;k++)
    scanf("%s",states[k]);
  for( k=0;k<n;k++){
     i=0;
     strcpy(state,states[k]);
     strcpy(copy,state);
     add_state(state,i++);
     while(1){
       end = fscanf(ipfile,"%s%s%s",state1,input,state2);
       if (end == EOF )
         break;
       if( strcmp(state,state1) == 0 )
         if( strcmp(input,"e") == 0 ) {
            add_state(state2,i++);
            strcpy(state, state2);
         }
    }
    display(i);
    rewind(ipfile);
  }
  fclose(fp);
  return 0;
}
```

Date:

EXPERIMENT - 1

EPSILON(ε) - CLOSURE OF STATES OF NFA

AIM

To write a C program to find ε – closure of all states of any given NFA with ε transition.

ALGORITHM

1. Start.
2. Declare necessary headers.
3. Create an empty 2D array ‘states’ to store the states of the NFA.
4. Read the number of states ‘n’.
5. Enter all the states and store them in ‘states[]’.
6. Open the transition table file ‘input.txt’. The file contains transitions in the format: ‘state1 input_symbol state2’ (where ‘input_symbol’ can be ‘ε’).
7. For each state in ‘states[]’:
    a. Set ‘i = 0’ as an index for ‘result[]’.
    b. Initialize a copy of the current state to ‘copy’.
    c. Create an empty array ‘result[]’ to store the epsilon closure for the current state.
    d. Add the current state itself to its epsilon closure
    e. For each transition in the file (‘state1’, ‘input’, ‘state2’):
        i. If ‘state1’ matches the current state and the ‘input’ is ‘ε’ (epsilon transition):
        ii. Add ‘state2’ to the current state's epsilon closure (‘result[]’).
        iii. Update the current state to ‘state2’ and repeat, adding all states reachable by epsilon transitions.
    f. Call display function to print the set of states in ‘result[]’.
    g. Use ‘rewind()’ to reset the file pointer to the beginning of the file so that it can be read again for the next state.
8. Close the file.
9. Stop.

OUTPUT

Input.txt

```
q0 a q1
q0 e q2
q1 b q0
q1 e q2
q2 a q0
```

```
muhammedbilalmn@Muhammeds-MacBook-Air CD_LAB % gcc pgm1.c
muhammedbilalmn@Muhammeds-MacBook-Air CD_LAB % ./a.out

 Enter the no of states: 3

 Enter the states:q0 q1 q2

Epsilon closure of q0 = {  q0, q2 }

Epsilon closure of q1 = {  q1, q2 }

Epsilon closure of q2 = {  q2 }
muhammedbilalmn@Muhammeds-MacBook-Air CD_LAB %
```

RESULT

Successfully found ε – closure of all states of any given NFA with ε transition.

PROGRAM

```
#include<stdio.h>
#include<stdlib.h>
struct node{
  int st;
  struct node *link;
};
static int set[20], nostate, noalpha, s, notransition, nofinal,
start, finalstate[20],c,r,buffer[20], newstate;
char alphabet[20];
static int e_closure[20][20]={0};
struct node * transition[20][20]={NULL};

int findalpha(char c){
  int i;
  for(i=0;i<noalpha;i++)
    if(alphabet[i]==c)
      return i;
  return(999);
}
void insert_trantbl(int r,char c,int s){
  int j;
  struct node *temp;
  j=findalpha(c);
  if(j==999){
    printf("error\n");
    exit(0);
  }
  temp=(struct node *) malloc(sizeof(struct node));
  temp->st=s;
  temp->link=transition[r][j];
  transition[r][j]=temp;
}
void findclosure(int x,int sta){
  struct node *temp;
  int i;
  if(buffer[x])
     return;
  e_closure[sta][c++]=x;
  buffer[x]=1;
  if(alphabet[noalpha-1]=='e' && transition[x][noalpha-1]!=NULL){
    temp=transition[x][noalpha-1];
    while(temp!=NULL){
    findclosure(temp->st,sta);
    temp=temp->link;
  }
}
}
```

Date:

EXPERIMENT - 2
CONVERT NFA WITH ε-TRANSITIONS TO NFA

AIM

To write a C program to convert NFA with ε transition to NFA without ε transition.

ALGORITHM

main:

1. Input the number of alphabets ‘noalpha’.
2. Store the alphabet in ‘alphabet[]’, ensuring that if epsilon (‘e’) is present, it is the last alphabet.
3. Input the number of states ‘nostate’ and the starting state ‘start’.
4. Input the number of final states ‘nofinal’ and store the states in ‘finalstate[]’.
5. Input the number of transitions ‘notransition’.
6. For each transition (r, c, s), call the ‘insert_trantbl()’ function to store the transition.
7. For each state from 1 to ‘nostate’, reset the buffer and epsilon closure arrays.
8. Call ‘findclosure()’ for each state to compute its epsilon closure.
9. For each state and alphabet, compute the new transitions using ‘unionclosure()’ to merge the epsilon closures.
10. Print the computed transitions.
11. Call ‘findfinalstate()’ to print the final states in the equivalent NFA.

insert_trantbl(int r, char c, int s):

1. Call ‘findalpha(c)’ to find the index of the alphabet ‘c’.
2. If the alphabet does not exist, print an error and exit.
3. Create a new node to represent the transition (with state ‘s’).
4. Insert this node at the head of the linked list for ‘transition[r][index]’ (where ‘index’ is the index of ‘c’ in ‘alphabet[]’).

findalpha(char c):

1. Loop through ‘alphabet[]’ to find the character ‘c’.
2. If found, return its index.
3. If not found, return 999 to indicate an error.

```
void unionclosure(int i){
   int j=0,k;
   while(e_closure[i][j]!=0){
     k=e_closure[i][j];
     set[k]=1;
     j++; newstate++;
   }
}
void findfinalstate(){
  int i,j,k,t;
  for(i=0;i<nofinal;i++)
    for(j=1;j<=nostate;j++)
      for(k=0;e_closure[j][k]!=0;k++)
           if(e_closure[j][k]==finalstate[i])
                print_e_closure(j);
           printf("\n");
}
void print_e_closure(int i){
   int j;
   printf("{");
   for(j=0;e_closure[i][j]!=0;j++)
      printf("q%d,",e_closure[i][j]);
   printf("}\t");
}

void main(){
  int i,j,k,m,t,n;
  struct node* temp;
  printf("Enter the number of alphabets: ");
  scanf("%d",&noalpha);
  getchar();
  printf("NOTE:- [ use letter e as epsilon]\n");
  printf("NOTE:- [e must be last character ,if it is present]\n");
  printf("\nEnter alphabets: ");
  for(i=0;i<noalpha;i++){
    alphabet[i]=getchar();
    getchar();
  }
  printf("Enter the number of states: ");
  scanf("%d",&nostate);
  printf("Enter the start state: ");
  scanf("%d",&start);
  printf("Enter the number of final states: ");
  scanf("%d",&nofinal);
  printf("Enter the final states: ");
  for(i=0;i<nofinal;i++)
    scanf("%d",&finalstate[i]);
  printf("Enter no of transition: ");
  scanf("%d",&notransition);
  printf("NOTE:- [Transition is in the form--> qno alphabet qno
\n",notransition);
```

findclosure(int x, int sta):

1. If the state ‘x’ has already been visited (‘buffer[x] == 1’), return.
2. Add state ‘x’ to the epsilon closure of ‘sta’.
3. Mark ‘buffer[x] = 1’ to indicate that this state has been visited.
4. If the last alphabet is ‘e’ (epsilon) and there is an epsilon transition from state ‘x’:
5. Follow the epsilon transition and recursively call ‘findclosure()’ for the target states of the epsilon transition.

unionclosure(int i):

1. For each state in ‘e_closure[i]’, add it to the ‘set[]’ array.
2. Increment ‘newstate’ to track new states being added.

findfinalstate():

For each final state:

1. For each state ‘j’ (1 to ‘nostate’), check if any state in ‘e_closure[j]’ matches the final states.
2. If there is a match, print the epsilon closure for state ‘j’.

print_e_closure(int i)’:

1. Loop through the states in ‘e_closure[i][]’.
2. Print the states in the form ‘{q1, q2, ...}’.

```
printf("NOTE:- [States number must be greater than zero]\n");
printf("\nEnter transition--\n");
for(i=0;i<notransition;i++){
  scanf("%d %c%d",&r,&c,&s);
  insert_trantbl(r,c,s);
}
printf("\n");
for(i=1;i<=nostate;i++){
  c=0;
  for(j=0;j<20;j++){
    buffer[j]=0;
    e_closure[i][j]=0;
  }
  findclosure(i,i);
}
printf("Equivalent NFA without epsilon\n");
printf("-------------------------------------------\n");
printf("Start state: q1");
printf("\nAlphabets:");
for(i=0;i<noalpha;i++)
  printf("%c ",alphabet[i]);
printf("\nStates:" );
for(i=1;i<=nostate;i++)
   printf("\n\tq%d", i);
printf("\nTransitions:\n");
for(i=1;i<=nostate;i++){
   for(j=0;j<noalpha-1;j++){
      for(m=1;m<=nostate;m++)
         set[m]=0;
      newstate=0;
      for(k=0;e_closure[i][k]!=0;k++){
         t=e_closure[i][k];
         temp=transition[t][j];
         while(temp!=NULL){
           unionclosure(temp->st);
           temp=temp->link;
         }
      }

   if (!newstates) continue;

   printf("\t");
   print_e_closure(i);
   printf("%c\t",alphabet[j] );
   printf("{");
   for(n=1;n<=nostate;n++){
      if(set[n]!=0)
          printf("q%d,",n);
   }
    printf("}\n");
 }
```

```
  }
  printf("Final states:");
  findfinalstate();
}
```

OUTPUT

```
muhammedbilalmn@Muhammeds-MacBook-Air CD_LAB % ./a.out
Enter the number of alphabets: 3
NOTE:- [ use letter 'e' as epsilon]
NOTE:- ['e' must be the last character, if it is present]

Enter alphabets: a b e
Enter the number of states: 3
Enter the start state: 1
Enter the number of final states: 1
Enter the final states: 3
Enter no of transition: 5
NOTE:- [Transition is in the form--> state alphabet state]
NOTE:- [State numbers must be greater than zero]

Enter transitions--
1 e 1
1 e 2
2 e 2
2 a 2
3 b 3

Equivalent NFA without epsilon
--------------------------------
Start state: {q1,q2}

Alphabets: a b

States: (Each new state is a set of states from the original NFA)
        New State (from q1) = {q1,q2}
        New State (from q2) = {q2}
        New State (from q3) = {q3}

Transitions:
        {q1,q2} --a--> {q2}
        {q2} --a--> {q2}
        {q3} --b--> {q3}

Final states: {q3}
muhammedbilalmn@Muhammeds-MacBook-Air CD_LAB %
```

RESULT

Successfully converted epsilon NFA to NFA without epsilon.

PROGRAM

```
#include<stdio.h>
#include<stdlib.h>
struct node{
 int st;
 struct node* link;
};
struct node1{
  int nst[20];
};
static int nostate, noalpha, s, notransition, nofinal, start, c, r;
static int set[20],finalstate[20], buffer[20];
static int eclosure[20][20] ={0};
int complete=-1;
char alphabet[20];
struct node1 hash[20];
struct node* transition[20][20] = {NULL};
int compare(struct node1 a,struct node1 b){
  for(int i=1;i<=nostate;i++)
    if(a.nst[i]!=b.nst[i])
      return 0;
  return 1;
}
int insertDfaState(struct node1 newstate){
  for(int i=0;i<=complete;i++)
    if(compare(hash[i],newstate))
      return 0;
  complete++;
  hash[complete]=newstate;
  return 1;
}
int findalpha(char c){
  int i;
  for (i=0; i<noalpha; i++)
    if (alphabet[i] == c)
      return i;
  return 999;
}
void insert(int r, char c, int s){
  int j=  findalpha(c);
  if (j==999){
    printf("Error from insert function");
    exit(0);
  }
  struct node* temp;
  temp = (struct node*)malloc(sizeof(struct node));
  temp->st = s;
  temp->link = transition[r][j];
  transition[r][j] = temp;
}
```

Date:

EXPERIMENT - 3
NFA TO DFA CONVERSION

AIM

To write a C program to convert NFA to DFA

ALGORITHM

main()

1. Prompt the user to input the number of alphabets (‘noalpha’).
2. Read and store the alphabets in ‘alphabet[]’. (Ensure ‘e’ is used for epsilon and it should be the last character).
3. Input the number of states (‘nostate’), starting state (‘start’), number of final states (‘nofinal’), and the final state numbers in ‘finalstate[]’.
4. Input the number of transitions (‘notransition’).
5. For each transition (state ‘r’, alphabet ‘c’, and state ‘s’), call the ‘insert()’ function to store the transition.
6. Initialize the hash table ‘hash[]’ where DFA states are stored.
7. Create the initial DFA state from the NFA start state and add it to the DFA state list using ‘insertDfaState()’.
8. Loop through each DFA state, processing transitions for each alphabet.
9. For each new DFA state found, insert it using ‘insertDfaState()’ and print the transitions using ‘printnewstate()’.
10. Print the DFA states and their transitions.
11. Identify and print the final states using ‘findfinalstate()’.

insert(int r, char c, int s):

1. Call ‘findalpha(c)’ to find the index of alphabet ‘c’.
2. If the alphabet is not found, print an error and exit.
3. Create a new node representing the transition (‘s’).
4. Insert the node at the head of the linked list for ‘transition[r][index]’.

```
void printnewstate(struct node1 state){
  printf("{");
  for(int j=1;j<=nostate;j++)
    if(state.nst[j]!=0)
      printf("q%d,",state.nst[j]);
  printf("}\t");
}
void findfinalstate(){
  for(int i=0;i<=complete;i++)
    for(int j=1;j<=nostate;j++)
      for(int k=0;k<nofinal;k++)
        if(hash[i].nst[j]==finalstate[k]){
          printnewstate(hash[i]);
          printf(" ");
          j=nostate;
          break;
      }
}
void main(){
  int i, j, k, m, t, n, l;
  struct node *temp;
  struct node1 newstate = {0}, tmpstate={0};
  printf("NOTE:- use letter e as epsilon\n :- e must be last character if it
is present: \n");
  printf("Enter the number of alphabets: ");
  scanf("%d", &noalpha); //noalpha -> number of alphabets
  printf("Enter the alphabets: ");
  getchar();
  for (i=0; i<noalpha; i++){
    alphabet[i] = getchar(); //character read is stored in alphabet
    getchar(); //newline is read here
  }
  printf("Enter the number of states: ");
  scanf("%d", &nostate); //nostate -> number of states
  printf("Enter the start state: ");
  scanf("%d", &start); //start -> start state number
  printf("Enter the number of final states: ");
  scanf("%d", &nofinal); //nofinal -> number of final states
  printf("Enter the final states: ");
  for(i=0; i<nofinal; i++)
     scanf("%d", &finalstate[i]); //contains the position of final states
  printf("Enter the number of transition: ");
  scanf("%d", &notransition);
  printf("NOTE:- Transition is in the form -> qno alphabet qno\n :- State
number must be greater than zero\nEnter Transition:\n");

  for (i=0; i<notransition; i++){
    scanf("%d %c %d", &r, &c, &s);
    insert(r,c,s);
  }
```

findalpha(char c):

1. Iterate through the ‘alphabet[]’ array to find the character ‘c’.
2. If found, return its index.
3. If not found, return 999 as an error indicator.

insertDfaState(struct node1 newstate):

1. Loop through the existing DFA states in ‘hash[]’.
2. Call ‘compare()’ to check if ‘newstate’ matches an existing DFA state.
3. If ‘newstate’ does not already exist, increment ‘complete’ and add the new state to ‘hash[]’.

compare(struct node1 a, struct node1 b):

1. Loop through each state in ‘a.nst[]’ and ‘b.nst[]’.
2. If any state differs between ‘a’ and ‘b’, return 0 (not equal).
3. If all states match, return 1 (equal).

printnewstate(struct node1 state):

1. Loop through the ‘state.nst[]’ array.
2. For each non-zero state, print the state in the form ‘q<number>‘.

findfinalstate():

1. For each DFA state in ‘hash[]’, loop through its constituent NFA states.
2. For each NFA state in the DFA state, check if it matches any of the NFA final states in ‘finalstate[]’.
3. If a DFA state contains any NFA final state, print that DFA state using ‘printnewstate()’.

```
for(i=0;i<20;i++){
  for(j=0;j<20;j++)
    hash[i].nst[j]=0;
}
i=-1;
printf("\nEquivalent DFA.....\n");
printf(".......................\n");
printf("Transitions of DFA\n");
newstate.nst[start]=start;
insertDfaState(newstate);
while(i!=complete){
  i++;
  newstate=hash[i];
   printnewstate(newstate);
  for(k=0; k<noalpha; k++){
    c=0;
    for(j=1; j<=nostate;j++)
      set[j]=0;
    for(j=1; j<=nostate;j++){
         l = newstate.nst[j];
         if (l){
          temp = transition[l][k];
          while(temp){
            if (!set[temp->st]){
             c++;
             set[temp->st] = temp->st;
            }
            temp = temp->link;
          }
        }
    }
    printf("\n");

    if (c){
      for(m=1; m<=nostate; m++)
        tmpstate.nst[m] = set[m];
      insertDfaState(tmpstate);
      printnewstate(tmpstate);
    }
   else{
      printf("NULL\n");
    }
  }
    Printf("\n");
}
printf("\n\nStates of DFA:\n");
for(i=0;i<=complete;i++)
  printnewstate(hash[i]);
```

```
printf("\n Alphabets: ");
  for(i=0;i<noalpha;i++)
    printf("%c\t",alphabet[i]);
  printf("\n Start State:q%d\n", start);
  printf("Final states: ");
    findfinalstate();
}
```

OUTPUT

```
muhammedbilalmn@Muhammeds-MacBook-Air CD_LAB % gcc pgm3.c
muhammedbilalmn@Muhammeds-MacBook-Air CD_LAB % ./a.out
NOTE:- Use 'e' for epsilon.
NOTE:- 'e' must be the last alphabet character if present.

Enter the number of alphabets: 2
Enter the alphabets: a b
Enter the number of states: 4
Enter the start state: 1
Enter the number of final states: 2
Enter the final states: 3 4
Enter the number of transitions: 8
NOTE:- Transitions are in the form: state alphabet state
Enter Transitions:
1 a 1
1 b 1
1 a 2
2 b 2
2 a 3
3 a 4
3 b 4
4 b 3

--- Equivalent DFA ---
Transitions:
{q1}--a--> {q1,q2}
{q1}--b--> {q1}
{q1,q2}--a--> {q1,q2,q3}
{q1,q2}--b--> {q1,q2}
{q1,q2,q3}--a--> {q1,q2,q3,q4}
{q1,q2,q3}--b--> {q1,q2,q4}
{q1,q2,q3,q4}--a--> {q1,q2,q3,q4}
{q1,q2,q3,q4}--b--> {q1,q2,q3,q4}
{q1,q2,q4}--a--> {q1,q2,q3}
{q1,q2,q4}--b--> {q1,q2,q3}

States of DFA are:
{q1}
{q1,q2}
{q1,q2,q3}
{q1,q2,q3,q4}
{q1,q2,q4}

Start State is: {q1}

Final States are: {q1,q2,q3}{q1,q2,q3,q4}{q1,q2,q4}
muhammedbilalmn@Muhammeds-MacBook-Air CD_LAB %
```

RESULT

Successfully converted NFA to DFA.

PROGRAM

```
#include<stdio.h>
#include<string.h>
#include<stdlib.h>
#define STATES 99
#define SYMBOLS 20
int
N_symbols,N_DFA_states,N_optDFA_sta
tes;
char
DFA_finals[STATES],OptDFA_finals[ST
ATES],StateName[STATES][STATES+1];
int
DFAtab[STATES][SYMBOLS],OptDFAtab[S
TATES][SYMBOLS];
void
load_DFA_table(),print_dfa_table(in
t tab[][SYMBOLS],int nstates,int
nsymbols,const char* finals,int
is_opt),optimize_DFA();
int
init_equiv_class(),get_equiv_class_
index(char state,int n_classes);
void get_next_state(char*
next_state,const char*
current_states,int
symbol),chr_append(char* s,char
ch),sort_string(char*
s),sort_state_names(int n);
int
main(){load_DFA_table();print_dfa_t
able(DFAtab,N_DFA_states,N_symbols,
DFA_finals,0);optimize_DFA();return
0;}
void load_DFA_table(){
N_DFA_states=6;N_symbols=2;strcpy(D
FA_finals,"EF");
DFAtab[0][0]=1;DFAtab[0][1]=2;
DFAtab[1][0]=4;DFAtab[1][1]=5;
DFAtab[2][0]=0;DFAtab[2][1]=0;
DFAtab[3][0]=5;DFAtab[3][1]=4;
DFAtab[4][0]=3;DFAtab[4][1]=5;
DFAtab[5][0]=3;DFAtab[5][1]=4;
}
void optimize_DFA(){
int n_classes=init_equiv_class();
while(1){
printf("\nEQUIV. CLASS CANDIDATE
==> ");
for(int
i=0;i<n_classes;i++)printf("%d:[%s]
",i,StateName[i]);printf("\n");
for(int i=0;i<n_classes;i++){
for(int sym=0;sym<N_symbols;sym++){
char
next_states[STATES+1]="",signature[STATE
S+1]="";
get_next_state(next_states,StateName[i],
sym);
for(int
k=0;next_states[k];k++)chr_append(signat
ure,'0'+get_equiv_class_index(next_state
s[k],n_classes));
printf("%d:[%s]\t--> [%s]
(%s)\n",i,StateName[i],next_states,signa
ture);
}}
int n_new_classes=0,changed=0;char
new_StateName[STATES][STATES+1];
for(int i=0;i<n_classes;i++){
char* p_class=StateName[i];char
subgroups[STATES][STATES+1];int
n_subgroups=0,visited[STATES]={0};
for(int j=0;p_class[j];j++){
if(visited[j])continue;
char sig_j[SYMBOLS+1]="";
for(int sym=0;sym<N_symbols;sym++){int
next_idx=DFAtab[p_class[j]-
'A'][sym];chr_append(sig_j,'0'+get_equiv
_class_index('A'+next_idx,n_classes));}
strcpy(subgroups[n_subgroups],"");chr_ap
pend(subgroups[n_subgroups],p_class[j]);
visited[j]=1;
for(int k=j+1;p_class[k];k++){
if(visited[k])continue;
char sig_k[SYMBOLS+1]="";
for(int sym=0;sym<N_symbols;sym++){int
next_idx=DFAtab[p_class[k]-
'A'][sym];chr_append(sig_k,'0'+get_equiv
_class_index('A'+next_idx,n_classes));}
if(strcmp(sig_j,sig_k)==0){chr_append(su
bgroups[n_subgroups],p_class[k]);visited
[k]=1;}
}
n_subgroups++;
}
for(int
sg=0;sg<n_subgroups;sg++)strcpy(new_Stat
eName[n_new_classes++],subgroups[sg]);
}
if(n_new_classes==n_classes)break;
n_classes=n_new_classes;
for(int
k=0;k<n_classes;k++){sort_string(new_Sta
teName[k]);strcpy(StateName[k],new_State
Name[k]);}
sort_state_names(n_classes);
}
```

Date:

EXPERIMENT - 4
DFA MINIMIZATION

AIM

To write a C program to minimize the given DFA.

ALGORITHM

main()

1. Call the `load_DFA_table()` function to initialize the DFA transition table `DFAtab[]`, the number of DFA states (`N_DFA_states`), and the number of symbols (`N_symbols`).
2. Call `print_dfa_table(DFAtab, N_DFA_states, N_symbols, DFA_finals)` to print the initial DFA transition table.
3. Call `optimize_DFA(DFAtab, N_DFA_states, N_symbols, DFA_finals, StateName, OptDFA)` to perform state minimization of the DFA.
4. The result is stored in `OptDFA[]`, and the number of optimized DFA states is returned as `N_optDFA_states`.
5. Call `get_NEW_finals(NEW_finals, DFA_finals, StateName, N_optDFA_states)` to get the final states of the optimized DFA.
6. Call `print_dfa_table(OptDFA, N_optDFA_states, N_symbols, NEW_finals)` to print the minimized DFA transition table.

load_DFA_table()

1. Initialize the DFA transition table `DFAtab[][]` with specific state transitions.
2. Set the DFA final states `DFA_finals`.
3. Set the number of DFA states (`N_DFA_states`) and the number of symbols (`N_symbols`).

optimize_DFA()

1. Call `init_equiv_class()` to divide DFA states into two equivalence classes: final states and non-final states.
2. Store the equivalence classes in `stnt[][]` and return the number of initial equivalence classes `n`.
3. Repeat until no further splitting of equivalence classes occurs:
    a. Print the current equivalence class candidates using `print_equiv_classes()`.
    b. Call `get_optimized_DFA()` to compute the transitions for the current equivalence classes.
    c. Call `set_new_equiv_class()` to split and refine the equivalence classes if necessary.
4. Return the number of minimized DFA states.

init_equiv_class()

1. Divide DFA states into two groups:
    a. Group 1: Non-final states.
    b. Group 2: Final states.
2. Store these groups as equivalence classes in `statename[][]`.
3. Return the number of initial equivalence classes.

```
N_optDFA_states=n_classes;
char
final_ordered_names[STATES][STATE
S+1];
strcpy(final_ordered_names[0],"A"
);strcpy(final_ordered_names[1],"
BD");
strcpy(final_ordered_names[2],"C"
);strcpy(final_ordered_names[3],"
EF");
for(int
i=0;i<N_optDFA_states;i++)strcpy(
StateName[i],final_ordered_names[
i]);
for(int
i=0;i<N_optDFA_states;i++)
for(int j=0;j<N_symbols;j++){char
next_s[2]={'A'+DFAtab[StateName[i
][0]-
'A'][j],'\0'};OptDFAtab[i][j]=get
_equiv_class_index(next_s[0],N_op
tDFA_states);}
strcpy(OptDFA_finals,"");
for(int
i=0;i<N_optDFA_states;i++)if(strp
brk(StateName[i],DFA_finals))chr_
append(OptDFA_finals,'A'+i);
print_dfa_table(OptDFAtab,N_optDF
A_states,N_symbols,OptDFA_finals,
1);
}
int init_equiv_class(){
strcpy(StateName[0],"");strcpy(St
ateName[1],"");
for(int i=0;i<N_DFA_states;i++){
if(strchr(DFA_finals,'A'+i))chr_a
ppend(StateName[1],'A'+i);
else
chr_append(StateName[0],'A'+i);
}
return 2;
}
int get_equiv_class_index(char
state,int n_classes){
for(int
i=0;i<n_classes;i++)if(strchr(Sta
teName[i],state))return i;
return -1;
}
void get_next_state(char*
next_state,const char*
current_states,int symbol){
for(int
j=0;current_states[j];j++)chr_app
end(next_state,'A'+DFAtab[current
_states[j]-'A'][symbol]);
}
void print_dfa_table(int
tab[][SYMBOLS],int nstates,int
nsymbols,const char* finals,int
is_opt){
printf("\n%s: STATE TRANSITION
TABLE\n",is_opt?"DFA":"DFA");
printf("     | ");for(int
i=0;i<nsymbols;i++)printf("
%d  ",i);
printf("\n-----+");for(int
i=0;i<nsymbols;i++)printf("----
");printf("\n");
for(int
i=0;i<nstates;i++){printf("  %c
| ",'A'+i);for(int
j=0;j<nsymbols;j++)printf("
%c  ",'A'+tab[i][j]);printf("\n")
;}
printf("Final states =
%s\n",finals);
}
void chr_append(char* s,char
ch){int
n=strlen(s);s[n]=ch;s[n+1]='\0';}
void sort_string(char* s){
for(int i=0;s[i];++i)
for(int j=i+1;s[j];++j)
if(s[i]>s[j]){char
t=s[i];s[i]=s[j];s[j]=t;}
}
void sort_state_names(int n){
for(int i=0;i<n-1;i++)
for(int j=i+1;j<n;j++)
if(StateName[i][0]>StateName[j][0
]){char
t[STATES+1];strcpy(t,StateName[i]
);strcpy(StateName[i],StateName[j
]);strcpy(StateName[j],t);}
}
```

get_optimized_DFA()

1. For each pseudo-DFA state in `stnt[][]`:
    a. For each input symbol, compute the next states.
    b. Call `get_next_state()` to compute the next states for the current state.
    c. Call `state_index()` to check whether the next states belong to a single equivalence class or need further splitting.
    d. Update the `newdfa[][]` transition table with the new equivalence class for each state.
2. Return the new number of equivalence classes after the transition computation.

set_new_equiv_class()

1. For each equivalence class in `stnt[][]`:
    a. For each input symbol, check whether the next states belong to a unique equivalence class using the `newdfa[][]` transition table.
    b. If a state needs further splitting, call `split_equiv_class()` to divide the equivalence class into subclasses.
2. Return the new number of equivalence classes.

split_equiv_class()

1. For the selected equivalence class:
    a. Divide the states in the class into multiple subclasses based on their transitions.
2. Update the equivalence class list `stnt[][]` to include the new subclasses.
3. Sort the equivalence classes and return the updated number of classes.

get_NEW_finals()

1. For each equivalence class in `stnt[][]`, check whether all states in the class are final states.
2. If the class contains final states, add the new state to `newfinals[]`.
3. Return the final states for the minimized DFA.

print_dfa_table()

1. Print the header for the DFA transition table (states and input symbols).
2. For each state in the DFA, print the transitions for all input symbols.
3. Print the final states of the DFA.

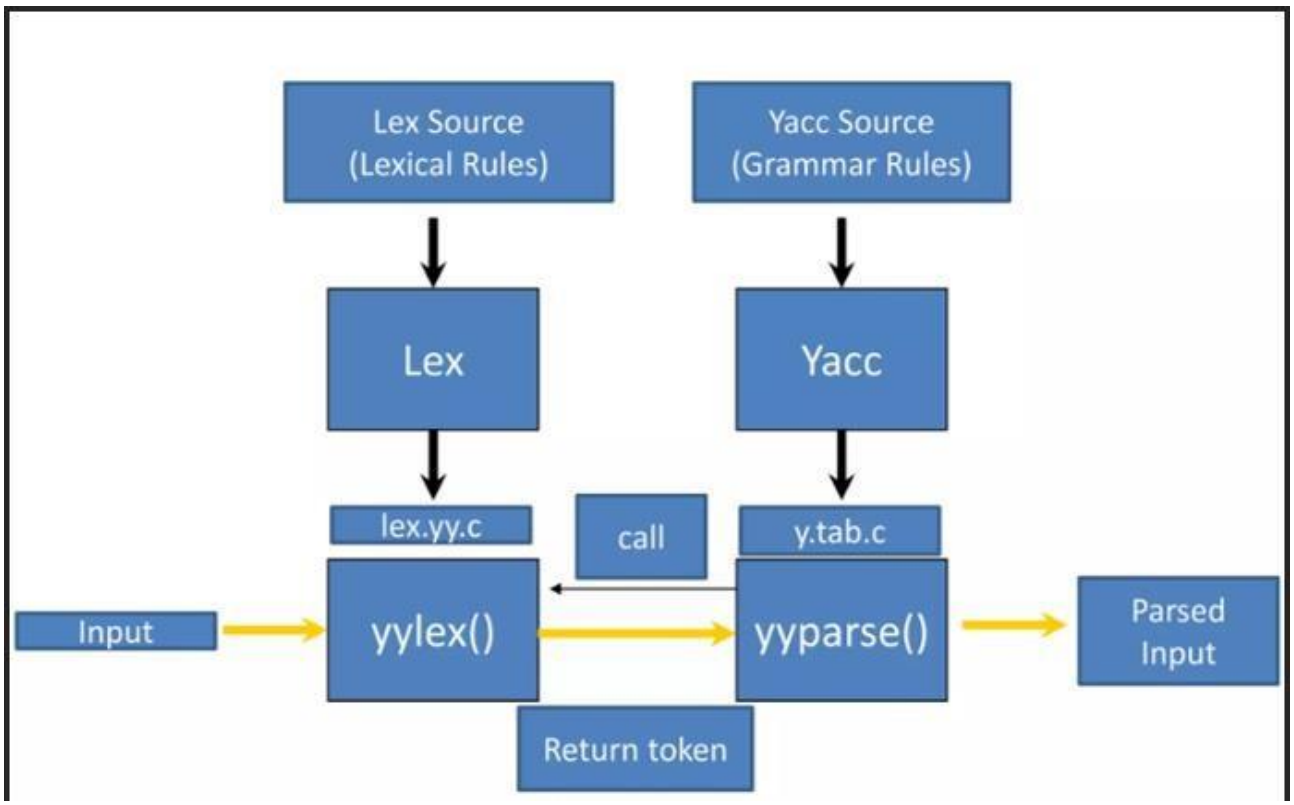
OUTPUT

```
muhammedbilalmn@Muhammeds-MacBook-Air CD_LAB % gcc pgm4.c
muhammedbilalmn@Muhammeds-MacBook-Air CD_LAB % ./a.out

DFA: STATE TRANSITION TABLE
     |  0   1
-----+--------
  A  |  B   C
  B  |  E   F
  C  |  A   A
  D  |  F   E
  E  |  D   F
  F  |  D   E
Final states = EF

EQUIV. CLASS CANDIDATE ==> 0:[ABCD] 1:[EF]
0:[ABCD]        --> [BEAF] (0101)
0:[ABCD]        --> [CFAE] (0101)
1:[EF]  --> [DD] (00)
1:[EF]  --> [FE] (11)

EQUIV. CLASS CANDIDATE ==> 0:[AC] 1:[BD] 2:[EF]
0:[AC]  --> [BA] (10)
0:[AC]  --> [CA] (00)
1:[BD]  --> [EF] (22)
1:[BD]  --> [FE] (22)
2:[EF]  --> [DD] (11)
2:[EF]  --> [FE] (22)

EQUIV. CLASS CANDIDATE ==> 0:[A] 1:[BD] 2:[C] 3:[EF]
0:[A]   --> [B] (1)
0:[A]   --> [C] (2)
1:[BD]  --> [EF] (33)
1:[BD]  --> [FE] (33)
2:[C]   --> [A] (0)
2:[C]   --> [A] (0)
3:[EF]  --> [DD] (11)
3:[EF]  --> [FE] (33)

DFA: STATE TRANSITION TABLE
     |  0   1
-----+--------
  A  |  B   C
  B  |  D   D
  C  |  A   A
  D  |  B   D
Final states = D
muhammedbilalmn@Muhammeds-MacBook-Air CD_LAB %
```

RESULT

Successfully minimized the given DFA.

COMPILATION STEPS



Date:

EXPERIMENT – 5

STUDY OF LEX AND YACC TOOLS

AIM

To understand about Lex and Yacc tools

THEORY

What is Lex?

Lex is a computer program that generates lexical analysers ("scanners" or "lexers"). It is commonly used with the Yacc parser generator and is the standard lexical analyser generator on many Unix and Unix-like systems. Lex reads an input stream specifying the lexical analyser and writes source code which implements the lexical analyser in the C programming language.

Structure of Lex Programs

A Lex program is organized into three main sections: Declarations, Rules, and Auxiliary Functions. Each section serves a distinct purpose in defining the behaviour of the lexical analyser.

- **Declarations**
  The Declarations section comprises two parts:
  - Regular Definitions: Define macros and shorthand notations for regular expressions to simplify rule definitions.
  - Auxiliary Declarations: Contains C code such as header file inclusions, function prototypes, and global variable declarations. This code is enclosed within %{ and %} and is copied verbatim into the generated lex.yy.c file.

- **Rules**
  The Rules section consists of pattern-action pairs, where each pair defines a regular expression pattern to match and the corresponding C code to execute upon a successful match.

- **Auxiliary Functions**
  The Auxiliary Functions section includes additional C code that is not part of the rules. This typically contains the main function and any helper functions required by the lexer. This code is placed after the second %% and is directly copied into the lex.yy.c file.

yylex()

The `yylex()` function, defined by Lex in the `lex.yy.c` file, reads the input stream, matches it against regular expressions, and executes the corresponding actions for each match. It also generates tokens for further processing by parsers or other components, though the programmer must explicitly invoke `yylex()` within the auxiliary functions of the Lex program.

yywrap()

The function `yywrap()` is called by `yylex()` when the end of an input file is reached. If `yywrap()` returns 0, scanning continues; if it returns a non-zero value, `yylex()` terminates and returns 0.

Compilation Steps

Step 1: Create a file with a .l and write your Lex .

Step 2: give the command lex simple_calc.l This command generates a C source file named lex.yy.c.

Step 3: GCC (GNU Compiler Collection) to compile the generated C code. gcc-o simple_calc lex.yy.c -lfl

Step 4: Running the Executable Execute the compiled program using: /simple_calc

What is Yacc?

Yacc (Yet Another Compiler Compiler) generates LALR parsers from formal grammar specifications. It is often used with Lex for building compilers and interpreters, handling the syntactic analysis phase. Yacc reads a grammar file and produces a C source parser that processes token sequences (typically from Lex) to check if they follow the grammar. The modern counterpart of Yacc is Bison.

Structure of YACC Program

```
%{
        C Declarations
%}
        Yacc Declarations
%%
        Grammar Rules
%%
        Additional Code
```

(Comments enclosed in /*…..*/ may appear in any section)

Compilation Steps

Step 1: Create a file with a .y and write your yacc .

Step 2: give the command yacc -d simple_calc.y This command generates a C source file named y.tab.c.

Step 3: Use GCC (GNU Compiler Collection) to compile the generated C code. gcc-o simple_calc lex.yy.c -lfl

Step 4: Running the Executable Execute the compiled program using: /simple_calc

RESULT

Familiarised and understood the working of YACC and LEX tools.

PROGRAM

```
#include <stdio.h>
#include <string.h>
#include <ctype.h>

char line[100];

int is_operator(char c) {
    switch (c) {
        case '+':
        case '-':
        case '*':
        case '/':
        case '=':
            printf("%c - Operator\n", c);
            return 1;
    }
    return 0;
}

int is_delimiter(char c) {
    switch (c) {
        case '{':
        case '}':
        case '(':
        case ')':
        case '[':
        case ']':
        case ',':
         case ';':
            printf("%c - Delimiter\n", c);
            return 1;
    }
    return 0;
}

int is_keyword(char buffer[]) {
    char keywords[32][10] = {"auto", "break", "case", "char", "const",
"continue", "default","do", "double", "else", "enum", "extern", "float", "for",
"goto","if", "int", "long", "register", "return", "short", "signed","sizeof",
"static", "struct", "switch", "typedef", "union","unsigned", "void",
"volatile", "while"};
      int i;
       for (i = 0; i < 32; ++i) {
         if (strcmp(keywords[i], buffer) == 0) {
             return 1;
         }
    }
    return 0;
}
```

Date:

EXPERIMENT – 6

IMPLEMENTATION OF A LEXICAL ANALYSER

AIM

Design and implement a lexical analyzer for given language using C and the lexical analyzer should ignore redundant spaces, tabs and newlines.

ALGORITHM

1. Start
2. Create a character array line[100] to hold each line of input.
3. Define functions to identify operators, delimiters, keywords, integers, and floating-point numbers.
4. Open the input file input.txt for reading.
5. While there are lines to read in the file:
    - 5.1 Read a line into the line array.
    - 5.2 Initialize flags for comment detection (flag1 for single-line comments, flag2 for multi-line comments) to false.
    - 5.3 For each character in the line, if the character is '/' and the next character is '/', set flag1 to true and skip processing the line.
    - 5.4 For each character in the line, if the character is '/' and the next character is '*', skip lines until you find the closing '*/' and set flag2 to true and skip processing the line.
    - 5.5 Initialize an empty token string and an index counter set to 0.
    - 5.6 For each character in the line:
        - 5.6.1 If token is a keyword, print it as a keyword.
        - 5.6.2 Else if token is an integer or floating-point number, print it as a number.
        - 5.6.3 Else, print it as an identifier.
        - 5.6.4 Clear the token string and reset the index counter.
        - 5.6.5 If the character is whitespace (space, tab, or newline), continue to the next character.
        - 5.6.6 If the character is not an operator, delimiter, or whitespace, add it to the token string and increment the index.
6. Close the input file after processing all lines.
7. Stop

```
int is_integer(char buffer[]) {
    for (int i = 0; buffer[i] != '\0'; i++) {
        if (!isdigit(buffer[i])) {
            return 0;
        }
    }
    return 1;
}

int is_float(char buffer[]) {
    int dot_count = 0;
    for (int i = 0; buffer[i] != '\0'; i++) {
        if (buffer[i] == '.') {
            dot_count++;
            if (dot_count > 1) return 0;
        } else if (!isdigit(buffer[i])) {
            return 0;
        }
    }
    return dot_count == 1;
}

void main() {
    char c;
    FILE *f = fopen("input.txt", "r");
    while (fgets(line, sizeof(line), f)) {
        int flag1 = 0;
        // Check for single line comment '//'
        for (int i = 0; i < strlen(line); i++) {
            if (line[i] == '/' && line[i + 1] == '/') {
                flag1 = 1;
                break;
            }
        }
        if (flag1) continue;
        // Multi-line comment '/**/'
        int flag2 = 0;
        for (int i = 0; i < strlen(line); i++) {
            if (line[i] == '/' && line[i + 1] == '*') {

// Skip all lines until '*/' has occurred

                while (fgets(line, sizeof(line), f)) {
                    for (int j = 0; j < strlen(line); j++) {
                        if (line[j] == '*' && line[j + 1] == '/') {
                            flag2 = 1;
                        }
                    }
                    if (flag2) break;
                }
            }
        }
        if (flag2) continue;
```

```
        printf("\n%s\n", line);
        char token[100];
        int index = 0;
        strcpy(token, "");

        for (int i = 0; i < strlen(line); i++) {
            // Check if character is an operator, delimiter, space, tab, or
newline
            if (is_operator(line[i]) || is_delimiter(line[i]) || line[i] == ' '
|| line[i] == '\t' || line[i] == '\n') {
                // Check if the token is an identifier, keyword, integer, or
float
                if (strcmp(token, "") != 0) {
                      if (is_keyword(token))
                        printf("%s - Keyword\n", token);
                    else if (is_integer(token) || is_float(token))
                        printf("%s - Number\n", token);
                    else
                        printf("%s - Identifier\n", token);
                    strcpy(token, "");
                    index = 0;
                }
            } else {
                token[index++] = line[i];
                token[index] = '\0';
            }
        }
    }
    fclose(f);
}
```

Input.txt

```
int a = b+10;
return a;
```

OUTPUT

```
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % gcc pgm5.c
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % ./a.out

  --- Analyzing Line: int a = b + 10;
  int - Keyword
  a - Identifier
  = - Operator
  b - Identifier
  + - Operator
  10 - Integer
  ; - Delimiter

  --- Analyzing Line: return a;return - Keyword
  a - Identifier
  ; - Delimiter
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 %
```

RESULT

Successfully run the program and obtained desired output.

PROGRAM

```
%{
#include<stdio.h>
#include<string.h>
int i = 0;
int flag=0;
char name[50]="";
%}

%%
[a-zA-Z]* {
      for(i=0;i<yyleng;i++){
            if(yytext[i]=='a' && yytext[i+1]=='t' && yytext[i+2]=='u' &&
yytext[i+3]=='l'){
                  flag=1;
            }
      }
      strcpy(name,yytext);
}
[\n] {
 if(flag==1){printf("%s contains 4 letters of name as substring\n",name);
 flag=0;
 strcpy(name,"");
 printf("Enter the string    ");
}else{
printf("%s does not contain 4 letters of name as  substring\n",name);
strcpy(name,"");
printf("Enter the string    ");
}
}
%%
int main()
{ printf("Enter the string    ");
      // The function that starts the analysis
      yylex();
      return 0;
}
```

OUTPUT

```
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % lex 7.l
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % gcc lex.yy.c
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % ./a.out
Enter the string bilal
bilal contains 4 letters of name as substring
Enter the string bilaljohn
bilaljohn contains 4 letters of name as substring
Enter the string muhammed
muhammed does not contain 4 letters of name as substring
Enter the string ^Z
zsh: suspended  ./a.out
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 %
```

Date:

EXPERIMENT - 7

LEX PROGRAM TO EXCLUDE NAME PREFIX

AIM

To write a lex program to recognize all strings which does not contain first four characters of your name as a substring.

ALGORITHM

1. The program begins execution in the main function, which prompts the user to enter a string and calls yylex() to start scanning input
2. Lex Rule: The rule [a-zA-Z]* matches any sequence of alphabetic characters (including an empty string).
3. Substring Check: The program loops through the current token yytext to see if it contains the first four letters of your name as a substring.
4. If the first four letters of your name is found within the input string, a flag is set to indicate its presence. The matched string is then stored in the name variable for later use.
5. After processing the string, the program checks the flag and displays messages based on whether the first four letters of your name were found or not.

RESULT

Successfully run the program and obtained desired output.

PROGRAM

Valid.l

```
%{
    #include "y.tab.h"
%}

%%
[a-zA-Z_][a-zA-Z_0-9]* {return
letter;}
[0-9] {return digit;}
. {return yytext[0];}
\n {return 0;}
%%

int yywrap(){
return 1;
}
```

Valid.y

```
%{
    #include<stdio.h>
    int valid=1;
    int yyerror();

%}

%token digit letter

%%
start : letter s
s :     letter s
      | digit s
      |
      ;
%%
int yyerror(){
    printf("\nIts not a
identifier!\n");
   valid=0;
   return 0;}

int main(){
     printf("\nEnter a name to
tested for identifier ");
    yyparse();
    if(valid){
    printf("\nIt is an
identifier!\n");
    }
}
```

OUTPUT

```
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % lex Valid.l
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % yacc -d Valid.y
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % gcc y.tab.c lex.yy.c
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % ./a.out

Enter a name to tested for identifier
bilal

It is an identifier!
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % ./a.out

Enter a name to tested for identifier
1bilal

Its not an identifier!
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 %
```

Date:

EXPERIMENT - 8

YACC PROGRAM TO RECOGNIZE VALID VARIABLES

AIM

To write a YACC program to recognize a valid variable which starts with a letter followed by any number of letters or digits.

ALGORITHM

1. The program begins in main() where the user is prompted to input a string.
2. yyparse() is called to start the parsing process.
3. The lexical analyzer identifies tokens based on patterns: letter for identifiers and digit for numbers.
4. The lexer returns letter for [a-zA-Z_][a-zA-Z_0-9]* and digit for [0-9] to the YACC parser.
5. The parser uses the CFG rules:
   - start → letter s
   - s → letter s | digit s | ε
6. The parser checks if the input follows these rules, defining valid identifiers starting with a letter.
7. If the input matches the CFG, the parsing continues successfully.
8. If the input violates the CFG, yyerror() is called, setting valid = 0 and printing an error message.
9. After parsing, main() checks the valid flag.
10. If valid, it prints "It is an identifier"; otherwise, it prints an error message.

RESULT

Successfully run the program and obtained desired output.

PROGRAM

Calc.y

```
%{
#include <stdio.h>
#include <stdlib.h>
// Declare external functions
void yyerror(char *s);
int yylex(void);
%}
%token NUMBER
%%
start:  expression              { printf(" Result is : %d\n", $1); }
        ;
expression:
      expression '+' term          { $$ = $1 + $3; }
    | expression '-' term          { $$ = $1 - $3; }
    | term                         { $$ = $1; }
    ;
term:
      term '*' factor              { $$ = $1 * $3; }
    | term '/' factor              { $$ = $1 / $3; }
    | factor                       { $$ = $1; }
    ;
factor:
      '(' expression ')'           { $$ = $2; }     // Value inside
parentheses
    | NUMBER                       { $$ = $1; }     // The value of
the number
    ;
%%
int main() {
     printf("Enter an arithmetic expression: ");
    yyparse();   // Parse the input
    return 0;
}
void yyerror(char *s) {
    fprintf(stderr, "Error: %s\n", s);
}
```

Date:

EXPERIMENT - 9

CALCULATOR WITH LEX AND YACC

AIM

To implement a calculator using lex and yacc.

ALGORITHM

1. The program begins in main(), prompting the user to input an arithmetic expression.
2. yyparse() is called to start parsing the input.
3. The lexical analyzer matches patterns for numbers ([0-9]+) and arithmetic operators (+, -, *, /) and returns tokens to the parser.
4. The parser uses the following CFG:
    - start → expression
    - expression → expression '+' term | expression '-' term | term
    - term → term '*' factor | term '/' factor | factor
    - factor → '(' expression ')' | NUMBER
5. The parser processes tokens based on the CFG rules, performing arithmetic operations as it matches expressions.
6. Intermediate results are calculated and stored using the semantic actions associated with each rule.
7. If an invalid token or character is encountered, the lexer prints an error message.
8. After successfully parsing, the result of the expression is printed.
9. If any syntax errors occur, yyerror() is invoked, printing an error message.The program terminates after displaying the result or an error message.

Calc.l

```
%{
#include "y.tab.h" // Include the header file generated by Yacc
%}

%%
[0-9]+     { yylval = atoi(yytext); return NUMBER; }
"+"        { return '+'; }
"-"        { return '-'; }
"*"        { return '*'; }

"/"        { return '/'; }
"("        { return '('; }
")"        { return ')'; }
[ \t]+     { /* Ignore spaces and tabs */ }
.          { printf("Unexpected character: %s\n", yytext); }
%%

int yywrap() {
    return 1;
}
```

OUTPUT

```
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % lex Calc.l
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % yacc -d Calc.y
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % gcc y.tab.c lex.yy.c
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % ./a.out
Enter an arithmetic expression: 4+3
 Result is : 7
```

```
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % lex Calc.l
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % yacc -d Calc.y
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % gcc y.tab.c lex.yy.c
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % ./a.out
Enter an arithmetic expression: (6*(4+2))
 Result is : 36
```

RESULT

Successfully run the program and obtained desired output.

PROGRAM

Bnf.y

```
%{
#include<string.h>
 #include<stdio.h>
struct quad { char op[5];
 char arg1[10];
 char arg2[10];
char result[10];
}QUAD[30];
struct stack { int items[100];
 int top; }stk;
int
Index=0,tIndex=0,StNo,Ind,tInd;
extern int LineNo;
%}
%union
{
char var[10];
}
%token <var> NUM VAR RELOP
%token MAIN IF ELSE WHILE TYPE
%type <var> EXPR ASSIGNMENT
CONDITION IFST ELSEST WHILELOOP
%left '-' '+'
%left '*' '/'
%%
PROGRAM : MAIN BLOCK
;
BLOCK: '{' CODE '}'
;
CODE: BLOCK
| STATEMENT CODE
| STATEMENT
;
STATEMENT: DESCT ';'
| ASSIGNMENT ';'
| CONDST
| WHILEST
;
DESCT: TYPE VARLIST
;
VARLIST: VAR ',' VARLIST
| VAR
;
ASSIGNMENT: VAR '=' EXPR{
strcpy(QUAD[Index].op,"=");
 strcpy(QUAD[Index].arg1,$3);
  strcpy(QUAD[Index].arg2,"");

strcpy(QUAD[Index].result,$1);

strcpy($$,QUAD[Index++].result)
;
}
;
EXPR: EXPR '+' EXPR
{AddQuadruple("+",$1,$3,$$);}
| EXPR '-' EXPR
{AddQuadruple("-",$1,$3,$$);}
| EXPR '*' EXPR
{AddQuadruple("*",$1,$3,$$);}
| EXPR '/' EXPR
{AddQuadruple("/",$1,$3,$$);}
| '-' EXPR
{AddQuadruple("UMIN",$2,"",$$);
}
| '(' EXPR ')' {strcpy($$,$2);}
| VAR
| NUM
;
CONDST: IFST{
Ind=pop();
sprintf(QUAD[Ind].result,"%d",I
ndex);
 Ind=pop();
sprintf(QUAD[Ind].result,"%d",I
ndex);
}
| IFST ELSEST
;
IFST: IF '(' CONDITION ')' {
strcpy(QUAD[Index].op,"==");
strcpy(QUAD[Index].arg1,$3);

strcpy(QUAD[Index].arg2,"FALSE"
);
  strcpy(QUAD[Index].result,"-
1");
   push(Index);
Index++;
}
BLOCK {
strcpy(QUAD[Index].op,"GOTO");
 strcpy(QUAD[Index].arg1,"");
strcpy(QUAD[Index].arg2,"");
```

Date:

EXPERIMENT - 10

BNF TO YACC CONVERSION

AIM

To write a program to convert the BNF rules into YACC form and write code to generate abstract syntax tree

ALGORITHM

1. The program begins in the main() function, where input is taken from a file (if provided), and yyparse() is invoked to start parsing.
2. The parser processes the source code, using a set of grammar rules and a CFG, which defines program structures like PROGRAM, BLOCK, CODE, STATEMENT, EXPR, CONDITION, and control structures like IF, ELSE, and WHILE.
3. The lexical analyzer scans the input code and returns tokens such as NUM, VAR, RELOP, and control keywords like IF, WHILE, and TYPE to the YACC parser.
4. During parsing, the program builds quadruples (3-address code) for expressions and conditions using the function AddQuadruple().
5. The quadruples are stored in the global array QUAD[], and temporary variables are generated using tIndex for intermediate expressions.
6. For assignments and expressions (e.g., addition, subtraction), the AddQuadruple() function creates the necessary quadruple, updating QUAD[] with the operator, operands, and result.
7. Control structures like IF, ELSE, and WHILE involve conditional jumps. These are handled by pushing and popping indices of quadruples onto the stack and updating their result fields to indicate where control should jump.
8. After parsing, the program prints all generated quadruples in a tabular format showing the operator, arguments, and result for each operation.
9. The stack functions push() and pop() manage control flow for conditional and loop statements by handling jump addresses.
10. The program concludes by printing any errors encountered, using yyerror() to indicate the line number where an error occurred.

```
strcpy(QUAD[Index].result,"-
1");
push(Index);
Index++;
};
ELSEST: ELSE{
tInd=pop();
 Ind=pop();
  push(tInd);
sprintf(QUAD[Ind].result,"%d",I
ndex);
}
BLOCK{
Ind=pop();
sprintf(QUAD[Ind].result,"%d",I
ndex);
};
CONDITION: VAR RELOP VAR
{AddQuadruple($2,$1,$3,$$);
StNo=Index-1;
}
| VAR
| NUM
;
WHILEST: WHILELOOP{
Ind=pop();
sprintf(QUAD[Ind].result,"%d",S
tNo);
 Ind=pop();
sprintf(QUAD[Ind].result,"%d",I
ndex);
}
;
WHILELOOP: WHILE '(' CONDITION
')' {
strcpy(QUAD[Index].op,"==");
strcpy(QUAD[Index].arg1,$3);
strcpy(QUAD[Index].arg2,"FALSE"
);
strcpy(QUAD[Index].result,"-
1");
push(Index);
Index++;
}
BLOCK {
strcpy(QUAD[Index].op,"GOTO");
strcpy(QUAD[Index].arg1,"");
strcpy(QUAD[Index].arg2,"");
strcpy(QUAD[Index].result,"-
1");
push(Index);
Index++;
}
;
%%
extern FILE *yyin;
int main(int argc,char *argv[])
{
FILE *fp;
int i;
    if(argc>1) {
        fp=fopen(argv[1],"r");
            if(!fp) {
                printf("\n File
not found");
                exit(0);
                }
        yyin=fp;
        }
yyparse();
printf("\n\n\t\t --------------
--------------","\n\t\t Pos
Operator Arg1 Arg2 Result",
"\n\t\t ------------------------------ ");
for(i=0;i<Index;i++)
{
printf("\n\t\t %d\t %s\t %s\t
%s\t%s",i,QUAD[i].op,QUAD[i].ar
g1,QUAD[i].arg2,QUAD[i].result)
;
}
printf("\n\t\t ----------------
---------");
printf("\n\n");
return 0; }

void push(int data){
    stk.top++;
        if(stk.top==100)
        {
            printf("\n Stack
overflow\n");
            exit(0); }

stk.items[stk.top]=data;
        }
int pop() {
    int data;
        if(stk.top==-1){
```

```
            printf("\n Stack
underflow\n");
            exit(0);
            }
    data=stk.items[stk.top--];
    return data;
 }
void AddQuadruple(char
op[5],char arg1[10],char
arg2[10],char result[10]){
    strcpy(QUAD[Index].op,op);

strcpy(QUAD[Index].arg1,arg1);

strcpy(QUAD[Index].arg2,arg2);

sprintf(QUAD[Index].result,"t%d
",tIndex++);

strcpy(result,QUAD[Index++].res
ult);
    }

yyerror() {
printf("\n Error on line
no:%d",LineNo);
}
```

Bnf.l

```
%{
#include"y.tab.h"
#include <stdio.h>
#include<string.h>
int LineNo=1;
%}
identifier [a-zA-Z][_a-zA-Z0-
9]*
number [0-9]+|([0-9]*\.[0-9]+)
%%
main\(\) return MAIN;
if return IF;
else return ELSE;
while return WHILE;
int|char|float return TYPE;
{identifier}
{strcpy(yylval.var,yytext);
return VAR;}
{number}
{strcpy(yylval.var,yytext);
return NUM;}
\< |
\> |
\>= |
\<= |
== {strcpy(yylval.var,yytext);
return RELOP;}
[ \t] ;
\n LineNo++;
. return yytext[0];
%%
int yywrap() {
    return 1;
}
```

Input file(test.c)

```
main() { int a,b,c;
 if(a<b) { a=a+b; }
 while(a<b) { a=a+b; }
 if(a<=b) { c=a-b; }
 else { c=a+b;}
}
```

OUTPUT

```
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % lex bnf.l
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % yacc -d bnf.y
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % gcc y.tab.c lex.yy.c
muhammedbilalmn@Muhammeds-MacBook-Air cycle2 % ./a.out test.c

                ----------------------------------------
                Pos     Operator        Arg1    Arg2    Result
                ----------------------------------------
                0       <               a       b       t0
                1       ==              t0      FALSE   5
                2       +               a       b       t1
                3       =               t1              a
                4       GOTO                            5
                5       <               a       b       t2
                6       ==              t2      FALSE   10
                7       +               a       b       t3
                8       =               t3              a
                9       GOTO                            5
                10      <=              a       b       t4
                11      ==              t4      FALSE   15
                12      -               a       b       t5
                13      =               t5              c
                14      GOTO                            17
                15      +               a       b       t6
                16      =               t6              c
                ----------------------------------------

muhammedbilalmn@Muhammeds-MacBook-Air cycle2 %
```

RESULT

Successfully run the program and obtained desired output.

PROGRAM

For.y

```
%{
      #include <stdio.h>
      int valid = 1;
%}

%token FOR PARANTHESIS OPERAND OPERATOR COMMA SEMICOLON NEWLINE
CURLYBRACE

%%
start: FOR PARANTHESIS A A B PARANTHESIS CURLYBRACE start CURLYBRACE
NEWLINE| ;
A: OPERAND OPERATOR OPERAND SEMICOLON | OPERAND OPERATOR OPERAND
COMMA A | SEMICOLON;
B: OPERAND OPERATOR | OPERAND OPERATOR COMMA B |;
%%

int yyerror()
{
      valid = 0;
      printf("Invalid.\n");
      return 1;
}

void main()
{
      printf("Enter string:\n");
      yyparse();

      if (valid)
            printf("Valid.\n");
}
```

For.l

```
%{
#include "y.tab.h"
%}

%%
for                         return FOR;
[\(\)]                            return PARANTHESIS;
[a-zA-Z0-9]*                      return OPERAND;
"="|"<"|">"|">="|"<="|"=="|"++"|"--"  return OPERATOR;
\;                          return SEMICOLON;
,                           return COMMA;
[\{\}]                   return CURLYBRACE;
```

Date:

EXPERIMENT – 11

FOR STATEMENT SYNTAX VALIDATOR

AIM

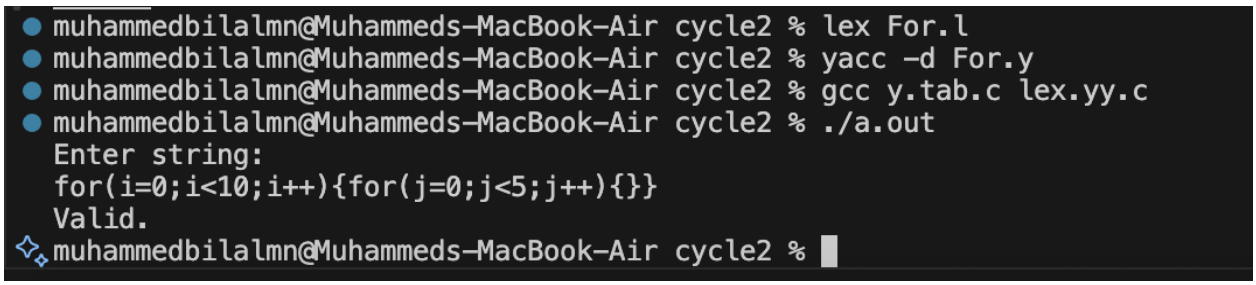
To write a yacc program to check syntax of for statement

ALGORITHM

1. The program execution begins in the main() function, where the user is prompted to enter a string.
2. The yyparse() function is called to initiate the parsing process, which triggers both the lexical analyzer and the YACC parser.
3. The lexical analyzer scans the input string, recognizing tokens (FOR, PARANTHESIS, OPERAND, OPERATOR, COMMA, SEMICOLON, NEWLINE) based on regular expressions and returns them to the YACC parser.
4. The YACC parser attempts to match the tokens against the predefined context-free grammar (CFG):
    - start → FOR PARANTHESIS A A B PARANTHESIS NEWLINE
    - A → OPERAND OPERATOR OPERAND SEMICOLON | OPERAND OPERATOR OPERAND COMMA A | SEMICOLON
    - B → OPERAND OPERATOR | OPERAND OPERATOR COMMA B
5. If the input conforms to the CFG, the parsing continues smoothly.
6. If any part of the input violates the grammar rules, the yyerror() function is invoked, which sets valid = 0 and prints "Invalid."
7. Once the parsing completes, control returns to the main() function.
8. The valid flag is checked. If it remains 1, indicating no errors, "Valid" is printed.
9. The yywrap() function is called at the end of input processing to signal completion.
10. The program successfully terminates after printing the validation result.

```
\n                              return NEWLINE;
.                               ;
%%

int yywrap()
{
      return 1;
}
```

OUTPUT



RESULT

Successfully run the program and obtained desired output.

PROGRAM

```
#include<stdio.h>
#include <string.h>
void main() {
  char stack[20], ip[20], opt[10][10][1], ter[10];
  int i, j, k, n, top = 0, col, row;

  for (i = 0; i < 10; i++) {
    stack[i] = NULL;
    ip[i] = NULL;
    for (j = 0; j < 10; j++) {
      opt[i][j][1] = NULL;
    }
  }
  printf("Enter the no.of terminals :\n");
  scanf("%d", &n);
  printf("\nEnter the terminals :\n");
  scanf("%s", &ter);
  printf("\nEnter the table values :\n");
  for (i = 0; i <n; i++) {
    for (j = 0; j < n; j++) {
      printf("Enter the value for %c %c:", ter[i], ter[j]);
      scanf("%s", opt[i][j]);
    }
  }
  printf("\n**** OPERATOR PRECEDENCE TABLE ****\n");
  for (i = 0; i < n; i++) {
printf("\t%c",ter[i]);
  } printf("\n");
  for (i = 0; i < n; i++) {
    printf("\n%c", ter[i]);
    for (j = 0; j < n; j++) {
      printf("\t%c", opt[i][j][0]);
    }
  }
  stack[top] = '$';
  printf("\nEnter the input string:");
  scanf("%s", ip);
  i = 0;
  printf("\nSTACK\t\t\tINPUT STRING\t\t\tACTION\n");
  printf("\n%s\t\t\t%s\t\t\t", stack, ip);
  while (i <= strlen(ip)) {
    for (k = 0; k < n; k++) {
      if (stack[top] == ter[k])
        col = k;
```

Date:

EXPERIMENT - 12

OPERATOR PRECEDENCE PARSER

AIM

To develop an operator precedence parser for a given language.

ALGORITHM:

1. Initialize arrays stack[20], ip[20], opt[10][10][1], and ter[10]. Set variables i, j, k, n, top = 0, col, row.

2. Set all elements of stack and ip to NULL. Set opt[i][j][1] to NULL in a nested loop.

3. Prompt the user to enter the number of terminals n.

4. Read the terminal symbols into the ter array.

5. In a nested loop, input the precedence table values (<, =, >) for each terminal pair.

6. Store the values in the array opt[i][j].

7. Print the operator precedence table for all terminal pairs.

8. Set stack[0] = '$' (bottom of the stack).

9. Print the initial stack and prompt the user to input the string ip.

10. Start a loop while i <= strlen(ip).

11. Find the column col and row row in the precedence table using stack[top] and ip[i].

12. If stack[top] == '$' and ip[i] == '$', print "String is accepted". Exit the loop.

13. If opt[col][row][0] == '<' or '=', push opt[col][row][0] and ip[i] onto the stack. Print "Shift ip[i]". Increment i.

14. If opt[col][row][0] == '>', pop symbols from the stack until stack[top] == '<'. Pop one more symbol. Print "Reduce".

15. If no valid precedence entry, print "String is not accepted". Exit the loop.

16. Print the current stack and remaining input after every shift or reduce.

17. Repeat steps 10-16 until the input string is either accepted or rejected.

18. Exit after accepting or rejecting the string.

```
       if (ip[i] == ter[k])
         row = k;
     }
     if ((stack[top] == '$') && (ip[i] == '$')) {
       printf("String is accepted\n");
       break;
     } else if ((opt[col][row][0] == '<') || (opt[col][row][0] ==
'=')) {
       stack[++top] = opt[col][row][0];
       stack[++top] = ip[i];
       printf("Shift %c", ip[i]);
       i++;
     } else {    if (opt[col][row][0] == '>') {
         while (stack[top] != '<') {
           --top;
         }
         top = top - 1;
         printf("Reduce");
       } else {
         printf("\nString is not accepted");
         break;
       }
     }
     printf("\n");
     for (k = 0; k <= top; k++) {
       printf("%c", stack[k]);
     }
     printf("\t\t\t");
     for (k = i; k < strlen(ip); k++) {
       printf("%c", ip[k]);
     }
     printf("\t\t\t");
   }
   getchar();
}
```

OUTPUT

```
muhammedbilalmn@Muhammeds-MacBook-Air cycle3 % gcc 12.c
Enter the no.of terminals:
4

Enter the terminals:
i+*$

Enter the table values:
Enter the value for i i:=
Enter the value for i +:>
Enter the value for i *:>
Enter the value for i $:>
Enter the value for + i:<
Enter the value for + +:>
Enter the value for + *:<
Enter the value for + $:>
Enter the value for * i:<
Enter the value for * +:>
Enter the value for * *:>
Enter the value for * $:>
Enter the value for $ i:<
Enter the value for $ +:<
Enter the value for $ *:<
Enter the value for $ $:>

**** OPERATOR PRECEDENCE TABLE ****
        i       +       *       $

i       =       >       >       >
+       <       >       <       >
*       <       >       >       >
$       <       <       <       >
Enter the input string: i+i*i

STACK                   INPUT STRING                    ACTION

$                       i+i*i$                  Shift i
$i                      +i*i$                   Reduce
$                       +i*i$                   Shift +
$+                      i*i$                    Shift i
$+i                     *i$                     Reduce
$+                      *i$                     Shift *
$+*                     i$                      Shift i
$+*i                    $                       Reduce
$+*                     $                       Reduce
$+                      $                       Reduce
$                       $                       String is accepted
muhammedbilalmn@Muhammeds-MacBook-Air cycle3 %
```

RESULT

Successfully simulated the working of operator precedence parser.

PROGRAM

```
#include <stdio.h>
#include <string.h>
int n;
char prods[50][50];
char firsts[26][50];
int is_first_done[26], is_follow_done[26];
char follows[26][50];
int isTerminal(char c)
{
    if (c < 65 || c > 90)
        return 1;
    return 0;
}
void first(char nonterm)
{
    int index = 0;
    char curr_firsts[50];
    for (int i = 0; i < n; i++)
    {
        if (prods[i][0] == nonterm)
        {
            int curr_prod_index = 2;
            int flag = 0;
            while (prods[i][curr_prod_index] != '\0' && flag == 0)
            {
                flag = 1;
                if (isTerminal(prods[i][curr_prod_index]))
                {
                      curr_firsts[index] = prods[i][2];
                    index++;
                    break;
                }
                if (!is_first_done[prods[i][curr_prod_index] - 65])
                    first(prods[i][curr_prod_index]);

                int in = 0;
                while (firsts[prods[i][curr_prod_index] - 65][in] !=
'\0')
                {
                    curr_firsts[index] =
firsts[prods[i][curr_prod_index] - 65][in];
                     if (firsts[prods[i][curr_prod_index] - 65][in]
== 'e')
                    {
                        curr_prod_index++;
                        flag = 0;
                    }
                    index++;
                    in++;
```

Date:

EXPERIMENT - 13
SIMULATION OF FIRST AND FOLLOW OF GRAMMAR

AIM

To simulate first and follow of any given grammar.

ALGORITHM:

1. Input the number of productions and the grammar rules.
2. Initialize arrays for storing First and Follow sets, and flags to track if the computation for each non-terminal is done.
3. For each non-terminal, find the First set:
    - 3.1 If the first symbol in the production is a terminal, add it to the First set.
    - 3.2 If the first symbol is a non-terminal, recursively compute its First set.
    - 3.3 If the First set of the non-terminal contains 'ε', continue checking the next symbol in the production.
4. Mark the First set computation for the non-terminal as done.
5. For the start symbol, add '$' to the Follow set.
6. For each production, find the Follow set:
    - 6.1 If a non-terminal is followed by a terminal, add that terminal to the Follow set.
    - 6.2 If a non-terminal is followed by another non-terminal, add the First set of the latter (excluding 'ε') to the Follow set.
    - 6.3 If a non-terminal appears at the end of a production or is followed by a non-terminal whose First set contains 'ε', add the Follow set of the left-hand side of the production to the current non-terminal's Follow set.
7. Mark the Follow set computation for the non-terminal as done.
8. Print the First and Follow sets for each non-terminal.

```
                }
            }
        }
    }
    curr_firsts[index] = '\0';
    index++;
    strcpy(firsts[nonterm - 65], curr_firsts);
    is_first_done[nonterm - 65] = 1;
}
void follow(char nonterm)
{
    int index = 0;
    char curr_follows[50];
    if (nonterm == prods[0][0])
    {
        curr_follows[index] = '$';
        index++;
    }
    for (int j = 0; j < n; j++)
    {
        int k = 2;
        int include_lhs_flag;
        while (prods[j][k] != '\0')
        {
            include_lhs_flag = 0;
            if (prods[j][k] == nonterm) {
                if (prods[j][k + 1] != '\0') {
                    if (isTerminal(prods[j][k + 1])) {
                        curr_follows[index] = prods[j][k + 1];
                        index++;
                        break;
                    }
                    int in = 0;
                    while (firsts[prods[j][k + 1] - 65][in] != '\0')
                    {
                        if (firsts[prods[j][k + 1] - 65][in] ==
'e'){
                           include_lhs_flag = 1;
                            in++
                           continue;
               }

                                    int temp_flag = 0;
                    for (int z = 0; z < index; z++)
                            if (firsts[prods[j][k + 1] - 65][in] ==
curr_follows[z]) {
                                temp_flag = 1;
                                in++;
                                break;
                            }
```

```
                        if (temp_flag)
                            continue;
                        curr_follows[index] = firsts[prods[j][k + 1]
- 65][in];
                        index++;
                        in++;
                    }
                }
                if (prods[j][k + 1] == '\0' || include_lhs_flag ==
1)   {
                    if (prods[j][0] != nonterm) {
                        if (!is_follow_done[prods[j][0] - 65])
                            follow(prods[j][0]);
                        int x = 0;
                        while (follows[prods[j][0] - 65][x] !=
'\0') {
                    int temp_flag = 0;
                            for (int z = 0; z < index; z++)
                                if (follows[prods[j][0] - 65][x] ==
curr_follows[z]) {
                                    temp_flag = 1;
                                    x++;
                                    break;
                                }
                            if (temp_flag)
                                 continue;
                            curr_follows[index] =
follows[prods[j][0] - 65][x];
                            index++;
                            x++;
                        }
                    }
                }
     }
             k++;
             }
    }
    curr_follows[index] = '\0';
    index++;
   strcpy(follows[nonterm - 65], curr_follows);
    is_follow_done[nonterm - 65] = 1;
}
```

```
int main(){
    printf("Enter the number of productions\n");
    scanf("%d", &n);
    printf("Enter productions: \n");

 for (int i = 0; i < n; i++)
        scanf("%s", prods[i]);
    for (int i = 0; i < 26; i++)
        is_first_done[i] = 0;
    for (int i = 0; i < n; i++)
        if (is_first_done[prods[i][0] - 65] == 0)
            first(prods[i][0]);
           for (int i = 0; i < n; i++)
        if (is_follow_done[prods[i][0] - 65] == 0)
            follow(prods[i][0]);
    printf("Firsts:\n");
    for (int i = 0; i < 26; i++)
        if (is_first_done[i])
            printf("%c : %s\n", i + 65, firsts[i]);
    printf("Follows:\n");
    for (int i = 0; i < 26; i++)
        if (is_follow_done[i])
            printf("%c : %s\n", i + 65, follows[i]);
}
```

OUTPUT

```
muhammedbilalmn@Muhammeds-MacBook-Air cycle3 % gcc 13.c
muhammedbilalmn@Muhammeds-MacBook-Air cycle3 % ./a.out
Enter the number of productions: 8
Enter the productions (e.g., E=E+T, use 'e' for epsilon):
E=TF
F=aTF
F=e
T=GH
H=mGH
H=e
G=lEr
G=i

--- FIRST Sets ---
FIRST(E) = { li }
FIRST(F) = { ae }
FIRST(T) = { li }
FIRST(H) = { me }
FIRST(G) = { li }

--- FOLLOW Sets ---
FOLLOW(E) = { $r }
FOLLOW(F) = { $r }
FOLLOW(T) = { a$r }
FOLLOW(H) = { a$r }
FOLLOW(G) = { ma$r }
muhammedbilalmn@Muhammeds-MacBook-Air cycle3 %
```

RESULT

Successfully simulated first and follow of any given grammar.

PROGRAM

```
#include<stdio.h>
#include<string.h>
#include<stdlib.h>
#include<ctype.h>
char ip_sym[15],ip_ptr=0,op[50],tmp[50];
void e_prime();
void e();
void t_prime();
void t();
void f();
void advance();
int n=0;
void e(){
       strcpy(op,"TE'");
       printf("E=%-25s",op);
       printf("E->TE'\n");
       t();
       e_prime();
}
void e_prime(){
      int i,n=0,l;
      for(i=0;i<=strlen(op);i++)
          if(op[i]!='e')
       tmp[n++]=op[i];
      strcpy(op,tmp);
      l=strlen(op);
      for(n=0;n < l && op[n]!='E';n++);
      if(ip_sym[ip_ptr]=='+') {
                  i=n+2;
       do {
             op[i+2]=op[i];
             i++;
               }while(i<=l);
        op[n++]='+';
        op[n++]='T';
        op[n++]='E';
        op[n++]=39;
        printf("E=%-25s",op);
        printf("E'->+TE'\n");
        advance();
        t();
        e_prime();
 }
else {

     op[n]='e';
       for(i=n+1;i<=strlen(op);i++)
       op[i]=op[i+1];
       printf("E=%-25s",op);
       printf("E'->e");}
```

Date:

EXPERIMENT - 14

RECURSIVE DESCENT PARSER

AIM

To construct a recursive descent parser for an expression.

ALGORITHM:

1. Start.
2. Declare necessary headers.
3. Define global variables:
    a) Define an array input[] to store the input string.
    b) Pointer to track the current position in the input string.
    c) Temporary arrays like production[] to store and manipulate intermediate parsing results.
4. Define parse_expression():
    a) Apply the rule E->T E'.
    b) Print the current production .
    c) Call parse_term() to handle T.
    d) Call parse_expression_prime() to process E'.
5. Define parse_expression_prime():
    a) Apply the rule E'->+TE'| ε.
    b) If the current input symbol is +, modify the production to E'->+TE', print it, and recursively call parse_term() and parse_expression_prime().
    c) If no + is found, apply epsilon (ε) and terminate.
6. Define parse_term():
    a) Apply the rule T->FT'.
    b) Print the current production .
    c) Call parse_factor() to handle F.
    d) Call parse_term_prime() to process T'.

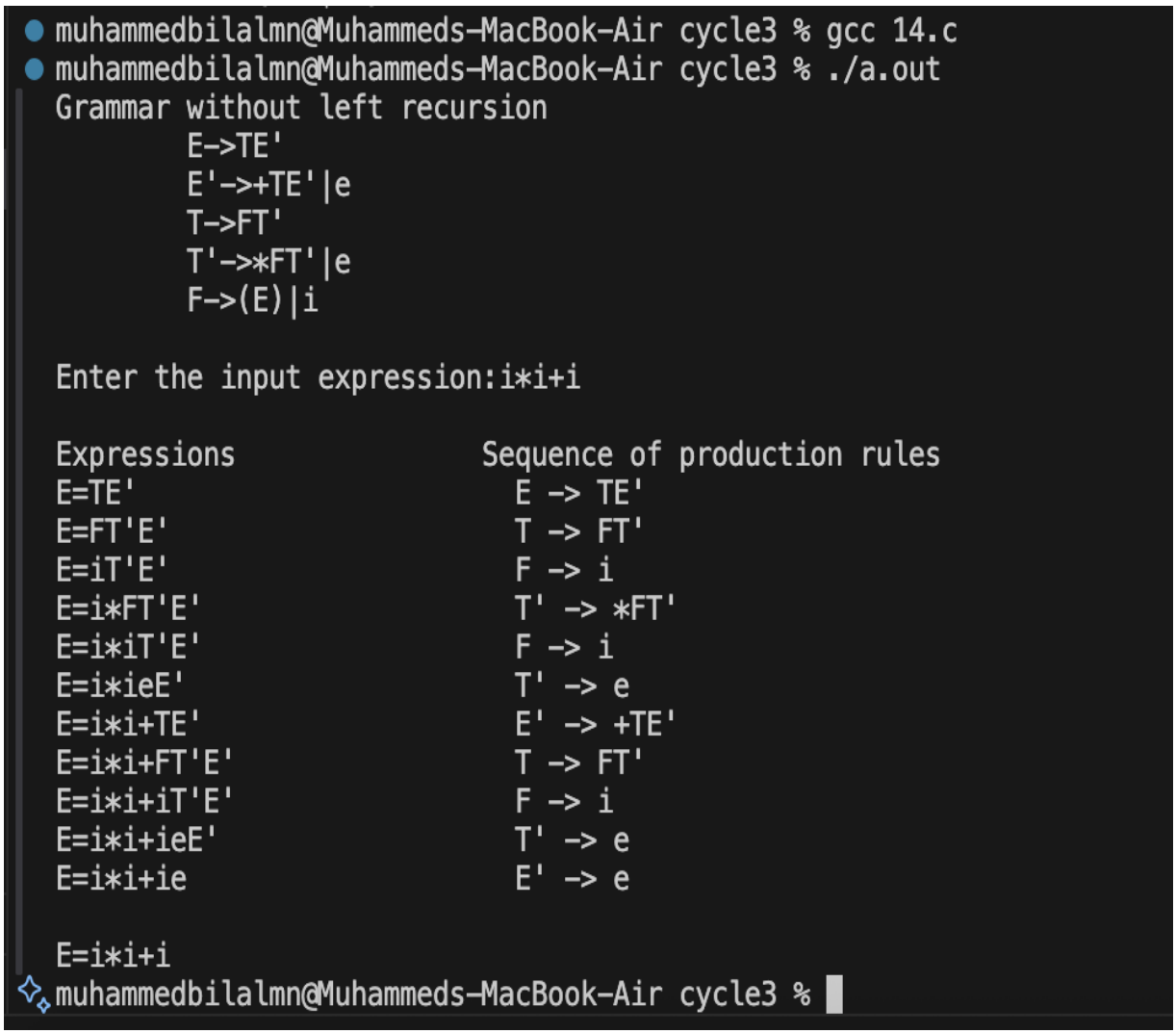
```
}
void t(){
 int i,n=0,l;
 for(i=0;i<=strlen(op);i++)
  if(op[i]!='e')
   tmp[n++]=op[i];
 strcpy(op,tmp);
 l=strlen(op);
for(n=0;n < l && op[n]!='T';n++);
 i=n+1;
 do
 {
  op[i+2]=op[i];
  i++;
 }while(i < l);
 op[n++]='F';
 op[n++]='T';
 op[n++]=39;
 printf("E=%-25s",op);
 printf("T->FT'\n");
 f();
 t_prime();
}

void t_prime()
{
int i,n=0,l;
for(i=0;i<=strlen(op);i++)
    if(op[i]!='e')
 tmp[n++]=op[i];
strcpy(op,tmp);
l=strlen(op);
for(n=0;n < l && op[n]!='T';n++);
if(ip_sym[ip_ptr]=='*')
 {
   i=n+2;
 do
 {
 op[i+2]=op[i];
 i++;
 }while(i < l);
  op[n++]='*';
  op[n++]='F';
  op[n++]='T';
  op[n++]=39;
  printf("E=%-25s",op);
  printf("T'->*FT'\n");
  advance();
  f();
  t_prime();
 }
```

7. Define parse_term_prime():
    a) Apply the rule T'->*FT'|ε.
    b) If the current input symbol is *, modify the production to T'->*FT', print it, and recursively call parse_factor() and parse_term_prime().
    c) If no * is found, apply epsilon (ε) and terminate.
8. Define parse_factor():
    a) Apply the rule F->(E)|id.
    b) If the current input symbol is id, accept it, update the production, and print it.
    c) If the current input is (, recursively call parse_expression() to handle the contents inside parentheses, followed by checking for the closing).
9. Define advance_input():
    a) Move the input pointer to the next character.
10. In main function():
    a) Read input string from the user.
    b) Call parse_expression() to initiate parsing.
    c) Print syntax errors if the input doesn't match the expected grammar rules.
    d) Output the successful parsing sequence if no errors are found.
11. Stop.

```
 else
 {
   op[n]='e';
  for(i=n+1;i<=strlen(op);i++)
 op[i]=op[i+1];
 printf("E=%-25s",op);
 printf("T'->e\n");
 }
}

void f()
{
int i,n=0,l;
for(i=0;i<=strlen(op);i++)
    if(op[i]!='e')
 tmp[n++]=op[i];
 strcpy(op,tmp);
l=strlen(op);
for(n=0;n < l && op[n]!='F';n++);
 if((ip_sym[ip_ptr]=='i')||(ip_sym[ip_ptr]=='I'))
 {
 op[n]='i';
 printf("E=%-25s",op);
 printf("F->i\n");
 advance();
 }
 else
 {
  if(ip_sym[ip_ptr]=='('){
   advance();
   e();
   if(ip_sym[ip_ptr]==')')
   {
    advance();
     i=n+2;
 do {
     op[i+2]=op[i];
 i++;
 }while(i<=l);
  op[n++]='('; op[n++]='E';
  op[n++]=')';
  printf("E=%-25s",op);
  printf("F->(E)\n");
   }
  }
  else{
   printf("\n\t syntax error");
     exit(1);
  }
 }
}
```

```
void advance(){
 ip_ptr++;
}
void main(){
 int i;
  printf("\nGrammar without left recursion");
 printf("\n\t\t E->TE' \n\t\t E'->+TE'|e \n\t\t T->FT' ");
 printf("\n\t\t T'->*FT'|e \n\t\t F->(E)|i");
 printf("\n Enter the input expression:");
 scanf("%s",ip_sym);
 printf("Expressions");
 printf("\t Sequence of production rules\n");
  e();
 for(i=0;i < strlen(ip_sym);i++){
  if(ip_sym[i]!='+'&&ip_sym[i]!='*'&&ip_sym[i]!='('&& ip_sym[i]!=')'
&&ip_sym[i]!='i'&&ip_sym[i]!='I'){
   printf("\nSyntax error");
   break;
  }
  for(i=0;i<=strlen(op);i++)
       if(op[i]!='e')
           tmp[n++]=op[i];
       strcpy(op,tmp);
       printf("\nE=%-25s",op);
 }
}
```

OUTPUT


muhammedbilalmn@Muhammeds-MacBook-Air cycle3 % gcc 14.c
muhammedbilalmn@Muhammeds-MacBook-Air cycle3 % ./a.out
Grammar without left recursion
E->TE'
E'->+TE'|e
T->FT'
T'->*FT'|e
F->(E)|i
Enter the input expression:i*i+i
Expressions
Sequence of production rules
E=TE'
E -> TE'
E=FT'E'
T -> FT'
E=iT'E'
F -> i
E=i*FT'E'
T' -> *FT'
E=i*iT'E'
F -> i
E=i*ieE'
T' -> e
E=i*i+TE'
E' -> +TE'
E=i*i+FT'E'
T -> FT'
E=i*i+iT'E'
F -> i
E=i*i+ieE'
T' -> e
E=i*i+ie
E' -> e
E=i*i+i
muhammedbilalmn@Muhammeds-MacBook-Air cycle3 %

RESULT

Successfully constructed a recursive descent parser for an expression.

PROGRAM

```
#include <stdio.h>
#include <string.h>
#define MAX_INPUT 100
#define MAX_STACK 100
void check();
char input[MAX_INPUT], stack[MAX_STACK], action[20];
int input_len, stack_top = -1;
int main() {
    printf("GRAMMAR is:\nE -> E+E | E*E | (E) | id\n");
    printf("Enter input string: ");
    scanf("%s", input);
    input_len = strlen(input);
    printf("\nStack\t\tInput\t\tAction\n");
    for (int i = 0; i < input_len; i++) {
        if (input[i] == 'i' && input[i+1] == 'd') {
            stack[++stack_top] = 'i';
            stack[++stack_top] = 'd';
            stack[stack_top + 1] = '\0';
            printf("$%s\t\t%s$\t\tSHIFT->id\n", stack, input + i +
2);
            check();
            i++;  // Skip the 'd' in the next iteration
        } else {
            stack[++stack_top] = input[i];
            stack[stack_top + 1] = '\0';
            printf("$%s\t\t%s$\t\tSHIFT->%c\n", stack, input + i +
1, input[i]);
            check();
        }
    }
      if (stack_top == 0 && stack[0] == 'E') {
        printf("\nInput string is VALID.\n");
    } else {
        printf("\nInput string is INVALID.\n");
    }
    return 0;
}
void check() {
    int i, j, handle_size;
    char *handle;
    while (1) {
        if (stack_top >=1 && stack[stack_top-1] == 'i' &&
stack[stack_top] == 'd') {
            handle = "id";
            handle_size = 2;
        }
```

Date:

EXPERIMENT - 15

SHIFT REDUCE PARSER

AIM

To simulate the working of a shift reduce parser.

ALGORITHM:

1. Initialization

   1.1. Initialize ‘input[]’, ‘stack[]’, ‘action[]’, and ‘stack_top = -1’.

   1.2. Set ‘input_len’ as the length of the input string.

2. Display Grammar

   2.1. Display the grammar:

   - ‘E -> E + E’, ‘E -> E * E’, ‘E -> (E)’, ‘E -> id’.

3. Input the String

   3.1. Prompt the user for input.

   3.2. Read the input into ‘input[]’.

4. Shift Operation

   4.1. For each character in ‘input[]’:

   4.2. If ‘id’ is found:

   - Push ‘id’ to the stack.
   - Print stack, remaining input, and action ‘SHIFT->id’.
   - Call ‘check()’ and skip ‘d’ by incrementing ‘i’ by 1.

   4.3. For other characters:

   - Push the character to the stack.
   - Print stack, remaining input, and action ‘SHIFT-><character>‘.
   - Call ‘check()’.

```
else if (stack_top >= 2 && stack[stack_top-2] == 'E' &&
stack[stack_top-1] == '+' && stack[stack_top] == 'E') {
            handle = "E+E";
            handle_size = 3;
        } else if (stack_top >= 2 && stack[stack_top-2] == 'E' &&
stack[stack_top-1] == '*' && stack[stack_top] == 'E') {
            handle = "E*E";
            handle_size = 3;
        } else if (stack_top >= 2 && stack[stack_top-2] == '(' &&
stack[stack_top-1] == 'E' && stack[stack_top] == ')') {
            handle = "(E)";
            handle_size = 3;
        } else {
            return;  // No reduction possible
        }
        // Perform reduction
        stack_top -= handle_size - 1;
        stack[stack_top] = 'E';
        stack[stack_top + 1] = '\0';
        printf("$%s\t\t%s$\t\tREDUCE->%s\n", stack, input +
strlen(input), handle);
    }
}
```

OUTPUT

```
muhammedbilalmn@Muhammeds-MacBook-Air cycle3 % gcc 15.c
muhammedbilalmn@Muhammeds-MacBook-Air cycle3 % ./a.out
GRAMMAR is:
E -> E+E | E*E | (E) | id
Enter input string: id+(id*id)

Stack          Input          Action
-------------------------------------------------
$              id+(id*id)$
$id            +(id*id)$              SHIFT -> id
$E             +(id*id)$              REDUCE E -> id
$E+            (id*id)$               SHIFT -> +
$E+(           id*id)$        SHIFT -> (
$E+(id         *id)$          SHIFT -> id
$E+(E          *id)$          REDUCE E -> id
$E+(E*         id)$           SHIFT -> *
$E+(E*id               )$             SHIFT -> id
$E+(E*E        )$             REDUCE E -> id
$E+(E          )$             REDUCE E -> E*E
$E+(E)         $              SHIFT -> )
$E+E           $              REDUCE E -> (E)
$E             $              REDUCE E -> E+E
-------------------------------------------------

Input string is VALID.
muhammedbilalmn@Muhammeds-MacBook-Air cycle3 %
```

5. Reduction (check() Function)

5.1. Check for handles on the stack:

- If ‘id’ → reduce to ‘E’.

- If ‘E + E’, ‘E * E’, ‘(E)’ → reduce to ‘E’.

5.2. If a handle is found:

- Perform reduction and print stack, input, and action ‘REDUCE-><handle>‘.

5.3. Repeat until no further reductions are possible.

6. Final Validation

6.1. If stack contains only ‘E’, print "VALID".

6.2. Otherwise, print "INVALID".

RESULT

Successfully simulated the working of a shift reduce parser.

PROGRAM

```
#include <stdio.h>
#include <string.h>
void gen_code_for_operator(char *inp, char operator, char * reg){
int i = 0, j = 0;
    char temp[100];
    while (inp[i] != '\0'){
        if (inp[i] == operator){
            printf("%c\t%c\t%c\t%c\n", operator, * reg, inp[i - 1],
inp[i + 1]);
            temp[j - 1] = *reg;
            i += 2;
            (*reg)--;
            continue;
        }
        temp[j] = inp[i];
        i++;
        j++;
    }
    temp[++j] = '\0';
    strcpy(inp, temp);
}
void gen_code(char *inp){
    // Operator precedence - /, *, +, -, =
    char reg = 'Z';
    gen_code_for_operator(inp, '/', &reg);
    gen_code_for_operator(inp, '*', &reg);
    gen_code_for_operator(inp, '+', &reg);
    gen_code_for_operator(inp, '-', &reg);
    gen_code_for_operator(inp, '=', &reg);
}
int main(){
    char inp[100];
    printf("Enter expression:\n\n");
    scanf("%s", inp);
    printf("Op \tDestn\tArg1\tArg2\n");
    gen_code(inp);
}
```

Date:

EXPERIMENT - 16

INTERMEDIATE CODE GENERATOR

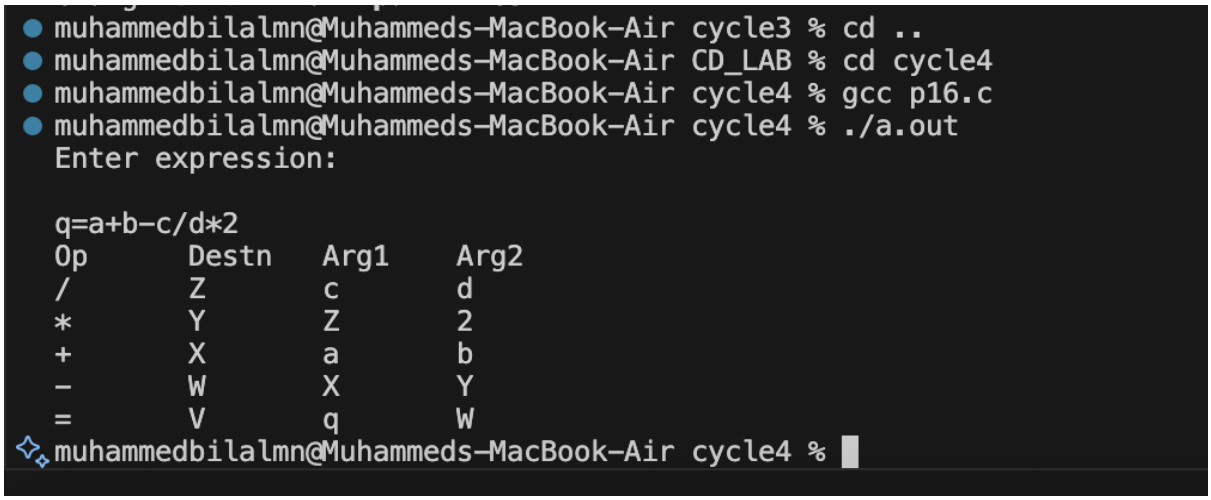
AIM

To implement an intermediate code generator for simple expressions.

ALGORITHM

1. Start
2. Declare necessary headers.
3. Declare a variable reg initialized to 'Z' to store intermediate results.
4. Read the input string inp(arithmetic expression)
5. Define the order of operator precedence: /, *, +, -, =
6. For each operator from highest precedence to lowest
    a) Traverse the input string inp character by character
    b) If an operator matches the current operator:
        i) Print the operator and the operands (characters before and after the operator) as three-address code.
        ii) Assign the result to the current register (reg).
        iii) Update the expression by replacing the operands and operator with the register.
        iv) Decrement the register to use the next available register for further operations.
    c) Continue the traversal until the end of the string.
7. Repeat step 6 for the next operator in the precedence order, continuing until all operators are processed
8. Stop

OUTPUT


muhammedbilalmn@Muhammeds-MacBook-Air cycle3 % cd ..
muhammedbilalmn@Muhammeds-MacBook-Air CD_LAB % cd cycle4
muhammedbilalmn@Muhammeds-MacBook-Air cycle4 % gcc p16.c
muhammedbilalmn@Muhammeds-MacBook-Air cycle4 % ./a.out
Enter expression:
q=a+b-c/d*2
Op Destn Arg1 Arg2
/ Z c d
* Y Z 2
+ X a b
- W X Y
= V q W
muhammedbilalmn@Muhammeds-MacBook-Air cycle4 %

RESULT

Successfully implement an intermediate code generator for simple expressions.

PROGRAM

```
#include <stdio.h>
#include <string.h>
#include <ctype.h>

struct tac {
    char op[5];
    char arg1[10];
    char arg2[10];
    char result[10];
};

struct tac code[] = {
    {"*", "b", "c", "t1"},
    {"+", "t1", "d", "a"},
    {"if", "a", "b", "L1"},
    {"=", "10", "", "c"},
    {"goto", "L2", "", ""},
    {"label", "L1", "", ""},
    {"=", "20", "", "c"},
    {"label", "L2", "", ""}
};

int is_number(const char* str) {
    if (str == NULL || *str == '\0') return 0;
    while (*str) {
        if (!isdigit(*str)) return 0;
        str++;
    }
    return 1;
}

int main() {
    int num_instructions = sizeof(code) / sizeof(code[0]);
    printf("_start:\n");
    for (int i = 0; i < num_instructions; i++) {
        struct tac current = code[i];
        printf("; TAC: %s = %s %s %s\n", current.result, current.arg1,
current.op, current.arg2);
        if (strcmp(current.op, "+") == 0 || strcmp(current.op, "-") == 0
||
            strcmp(current.op, "*") == 0 || strcmp(current.op, "/") == 0)
{
            if (is_number(current.arg1)) printf("    MOV AX, %s\n",
current.arg1);
            else printf("    MOV AX, [%s]\n", current.arg1);
            if (is_number(current.arg2)) printf("    MOV BX, %s\n",
current.arg2);
            else printf("    MOV BX, [%s]\n", current.arg2);
            if (strcmp(current.op, "+") == 0) printf("    ADD AX, BX\n");
            if (strcmp(current.op, "-") == 0) printf("    SUB AX, BX\n");
            if (strcmp(current.op, "*") == 0) printf("    MUL BX\n");
            if (strcmp(current.op, "/") == 0) {
                printf("    XOR DX, DX\n");
                printf("    DIV BX\n");
```

Date:

EXPERIMENT - 17

BACKEND OF COMPILER

AIM

Implement the back end of the compiler which takes the three address code and produces the 8086 assembly language instructions that can be assembled and run using an 8086 assembler. The target assembly instructions can be simple move, add, sub, jump etc.

ALGORITHM

1. Start
2. Declare necessary headers and a structure to represent Three-Address Code (TAC) with fields for operator, two arguments, and a result.
3. Define the input TAC program as a pre-filled array of the declared structures.
4. Print the initial boilerplate for the 8086 assembly program, including the starting label (e.g., _start:).
5. Traverse the array of TAC instructions from the first instruction to the last.
6. For each TAC instruction, check its operator and generate corresponding assembly code:
    a. If the operator is arithmetic (+, -, *, /):
        i. Generate an instruction to move the first argument (arg1) from memory into a register (e.g., AX).
        ii. Generate an instruction to move the second argument (arg2) from memory into another register (e.g., BX).
        iii. Generate the corresponding 8086 arithmetic instruction (ADD, SUB, MUL, DIV).
        iv. Generate an instruction to move the result from the register (e.g., AX) back to the result variable's memory location.
    b. If the operator is assignment (=):
        i. Generate an instruction to move the argument (arg1) into a register (e.g., AX).
        ii. Generate an instruction to move the value from the register to the result variable's memory location.
    c. If the operator is a conditional jump (if):
        i. Generate instructions to load both arguments (arg1, arg2) into two registers.
        ii. Generate a CMP instruction to compare the registers.
        iii. Generate the appropriate conditional jump instruction (e.g., JL for less than) to the target label, which is stored in the result field.
    d. If the operator is an unconditional jump (goto):

```
        }
            printf("    MOV [%s], AX\n\n", current.result);
        } else if (strcmp(current.op, "=") == 0) {
            if (is_number(current.arg1)) printf("    MOV AX, %s\n",
current.arg1);
            else printf("    MOV AX, [%s]\n", current.arg1);
            printf("    MOV [%s], AX\n\n", current.result);
        } else if (strcmp(current.op, "if") == 0) {
            if (is_number(current.arg1)) printf("    MOV AX, %s\n",
current.arg1);
            else printf("    MOV AX, [%s]\n", current.arg1);
            if (is_number(current.arg2)) printf("    MOV BX, %s\n",
current.arg2);
            else printf("    MOV BX, [%s]\n", current.arg2);
            printf("    CMP AX, BX\n");
            printf("    JL %s\n\n", current.result);
        } else if (strcmp(current.op, "goto") == 0) {
            printf("    JMP %s\n\n", current.arg1);
        } else if (strcmp(current.op, "label") == 0) {
            printf("%s:\n\n", current.arg1);
        }
    }
    printf("\n; Program exit\n");
    printf("    MOV AX, 0x4c00\n");
    printf("    INT 0x21\n");
    return 0;
}
```

OUTPUT

```
muhammedbilalmn@Muhammeds-MacBook-Air cycle4 % gcc p17.c
muhammedbilalmn@Muhammeds-MacBook-Air cycle4 % ./a.out
_start:

; TAC: t1 = b * c
    MOV AX, [b]
    MOV BX, [c]
    MUL BX
    MOV [t1], AX

; TAC: a = t1 + d
    MOV AX, [t1]
    MOV BX, [d]
    ADD AX, BX
    MOV [a], AX

; TAC: L1 = a if b
    MOV AX, [a]
    MOV BX, [b]
    CMP AX, BX
    JL L1

; TAC: c = 10 =
    MOV AX, 10
    MOV [c], AX

; TAC:  = L2 goto
    JMP L2

; TAC:  = L1 labelL1
; TAC: c = 20 =
    MOV AX, 20
    MOV [c], AX

; TAC:  = L2 labelL2

; Program exit
    MOV AX, 0x4c00
    INT 0x21
muhammedbilalmn@Muhammeds-MacBook-Air cycle4 %
```

i.Generate a JMP instruction to the target label stored in the arg1 field.

e. If the operator is a label definition (label): i) Print the label name from the arg1 field, followed by a colon.

7. Continue the traversal until all TAC instructions in the array have been processed.
8. Print the final boilerplate assembly code required for a clean program exit (e.g., the DOS exit interrupt call).
9. Stop

RESULT

Successfully implemented the backend of the compiler.