Experiment 05 : Intermediate Code Generator

Learning Objective: Student should be able to Apply Intermediate Code Generator using 3Address code.

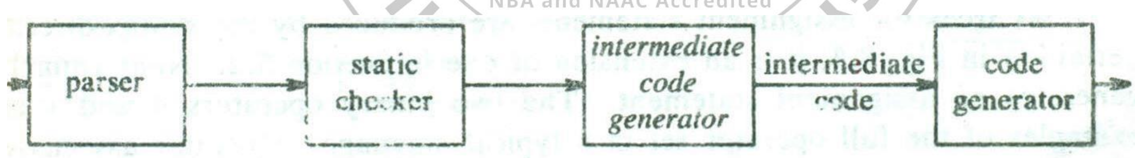
Tools: Jdk1.8,Turbo C/C++, Python, Notepad++

Theory:

Intermediate Code Generation:

In the analysis-synthesis model of a compiler, the front end translates a source program into an intermediate representation from which the back end generates target code. Details of the target language are confined to the backend, as far as possible. Although a source program can be translated directly into the target language, some benefits of using a machine-independent intermediate form are:

1. Retargeting is facilitated; a compiler for a different machine can be created by attaching a back end for the new machine to an existing front end.
2. A machine-independent code optimizer can be applied to the intermediate representation.



Position of intermediate code generator.

(Intermediate Languages) Intermediate Code Representation:

a) Syntax trees or DAG

b) postfix notation

c) Threeaddress code

Types of Three-Address Statements:

Three-address statements are akin to assembly code. Statements can have symbolic labels and there are statements for flow of control. A symbolic label represents the index of a three-address statement in the array holding intermediate code.

Actual indices can be substituted for the labels either by making a separate pass, or by using "back patching,"

Here are the common three-address statements:

I. Assignment statements of the form x= y op Z, where op is a binary arithmetic or logical operation.

2. Assignment instructions of the form x : = op y, where op is a unary operation. Essential unary operations include unary minus, logical negation, shift operators, and conversion operators that, for example, convert a fixed-point number to a floating-point number.

3. Copy statements of the form x : = y where the value of y is assigned to x.

4. The unconditional jump goto L. The three-address statement with label L is the next to be executed.

5. Conditional jumps such as if x relop y goto L. This instruction applies a relational operator «, =. >=, etc.) to x and y, and executes the statement with label L next if x stands in relation relop to y. If not, the three-address statement following if x relop y goto L is executed next, as in the usual sequence.

6. param x and call p, n for procedure calls and return y, where y representing a returned value is optional. Their typical use is as the sequence of three-address statements par am Xl par am X2 paramXn

callp,n

generated as part of a call of the procedure p (Xl, X2, ……….. Xn) The integer n indicating the number of actual parameters in "call p, n" is not redundant because calls can be nested.

7. Indexed assignments of the form X : = y[i] and x[i] : = y. The first of these sets x to the value in the location i memory units beyond location y. The statement x[i] : = y sets the contents of the location i units beyond x to the value of y. In both these instructions, x, y, and irefer to data objects.Address and pointer assignments of the form x : = &y, x : = *y, and *x : = y.

Implementations of Three-Address Statements

A three-address statement' is an abstract form of intermediate code. In a compiler, these statements can be implemented as records with fields for the operator and the operands. Three such representations are quadruples, triples, and indirect triples.

Quadruples:

a) A quadruple is a record structure with four fields: op, arg I, arg 2, and result.

b) The op field contains an internal code for the operator.

c) The three-address statement x : = y op z is represented by placing y in arg1, Z in arg 2, and x in result.

d) Statements with unary operators like x : = -y or x : = y do not use arg 2. Operators like param use neither arg 2 nor result.

e) Conditional and unconditional jumps put the target label in result.

f) The quadruples are for the assignment a : = b* - c + b* - c.
They are obtained from the three-address code in Fig. (a).

g) The contents of fieldsarg I, arg 2, and result are normally pointers to thesymbol-table entries for the names represented by these fields. If so, temporary names must be entered into the symbol table as they are created.

Triples:

a) To avoid entering temporary names into the symbol table,refer to a temporary value by the position of the statement that computes it.

b) Three-address statements can be represented by records with only three fields: op, arg1 and arg2, as in Fig.(b).

c) The fieldsarg1 and arg2, for the arguments of op, are either pointers to the symbol table. Since three fields are used, this intermediate code format is known as triples.

	op	*arg 1*	*arg 2*	*result*
(0)	uminus	c		t_1
(1)	*	b	t_1	t_2
(2)	uminus	c		t_3
(3)	*	b	t_3	t_4
(4)	+	t_2	t_4	t_5
(5)	:=	t_5		a

(a) Quadruples

	op	*arg 1*	*arg 2*
(0)	uminus	c	
(1)	*	b	(0)
(2)	uminus	c	
(3)	*	b	(2)
(4)	+	(1)	(3)
(5)	assign	a	(4)

(b) Triples

Quadruple and triple representations of three-address statements.

A ternary operation like x[i] := y requires two entries in the triple structure, as shown in Fig.(a), while x : = y[i] is naturally represented as two operations in Fig. (b).

	op	*arg 1*	*arg 2*
(0)	[ ]=	x	i
(1)	assign	(0)	y

(a) `x[i] := y`

	op	*arg 1*	*arg 2*
(0)	=[ ]	y	i
(1)	assign	x	(0)

(b) `x := y[i]`

More triple representations.

Indirect Triples:

Another implementation of three-address code that has been considered is that of listing pointers to triples, rather than listing the triples themselves. This implementation is naturally called indirect triples.

For example, let us use an array statement to list pointers to triples in the desired order.

	statement
(0)	(14)
(1)	(15)
(2)	(16)
(3)	(17)
(4)	(18)
(5)	(19)

	op	arg 1	arg 2
(14)	uminus	c	
(15)	*	b	(14)
(16)	uminus	c	
(17)	*	b	(16)
(18)	+	(15)	(17)
(19)	assign	a	(18)

Indirect triples representation of three-address statements

Input:

a : = b * - c + b* - c.

Output:

	op	arg 1	arg 2	result
(0)	uminus	c		t_1
(1)	*	b	t_1	t_2
(2)	uminus	c		t_3
(3)	*	b	t_3	t_4
(4)	+	t_2	t_4	t_5
(5)	:=	t_5		a

(a) Quadruples

	op	arg 1	arg 2
(0)	uminus	c	
(1)	*	b	(0)
(2)	uminus	c	
(3)	*	b	(2)
(4)	+	(1)	(3)
(5)	assign	a	(4)

(b) Triples

Application: Intermediate code can be easily produced to the target code.

Design:

Result and Discussion:

Learning Outcomes: The student should have the ability to

LO1 **Define** the role of Intermediate Code Generator in Compiler design.
LO2: ***Describe*** the various ways to implement Intermediate Code Generator.
LO3: ***Specify*** the formats of 3 Address Code.
LO4: Illustrate the working of Intermediate Code Generator using 3-Address code
.

Course Outcomes: Upon completion of the course students will be able to Evaluate the synthesis phase to produce object code optimized in terms of high execution speed and less memory usage.

Conclusion:

For Faculty Use

Correction Parameters	Formative Assessment [40%]	Timely completion of Practical [40%]	Attendance / Learning Attitude [20%]	
Marks Obtained				

CODE

```
from tabulate import tabulate
def threeaddr(s):   l=s.split("
")
  l=l[2:]

  op=['+','-','*','/','^']
arg1=[]   arg2=[]
res=[]   oper=[]
n=(len(l))//2   if(l[n]
not in op):     while(l[n]
not in op):       n=n-1
p1=l[:n]   p2=l[n+1:]
  ind=1
  '''
  oper.append('OPERATOR:')
arg1.append('ARGUMENT 1:')
arg2.append('ARGUMENT2:')
  res.append('RESULT:')
  '''
  if(len(l)==3):
    oper.append(l[n])
arg1.append(l[0])
arg2.append(l[2])
res.append("t"+str(ind))
oper.append("=")
arg1.append(s[0])
arg2.append("t"+str(ind))
res.append("t"+str(ind+1))
ans=[]
z1=zip(oper,arg1,arg2,res)     for
a1,a2,a3,a4 in list(z1):
      aq=[]
aq.append(a1)
aq.append(a2)
aq.append(a3)
aq.append(a4)
      ans.append(aq)
    print("QUADRAPLE TABLE:")
    print(tabulate(ans,                    headers=["OPERATORS","ARG                    1","ARG
2","RESULT"],tablefmt='orgtbl'))
else:
    m=0     for i in p1:       if(i[0]
in op and len(i)>1):
```

```
oper.append("unary"+i[0])
arg1.append(i[1])
arg2.append("nill")
res.append("t"+str(ind))
ind=ind+1        if(i in op and
len(i)==1):

           oper.append(i)
arg1.append(p1[m-1])
#print(p1.index(i)+1)
arg2.append(p1[m+1])
res.append("t"+str(ind))
my="t"+str(ind)
ind=ind+1        m=m+1     j=0
for i in p2:         if(i[0] in op
and len(i)>1):

          oper.append("unary"+i[0])
arg1.append(i[1])
arg2.append("nill")
res.append("t"+str(ind))
ind=ind+1         if(i in op and
len(i)==1):

           oper.append(i)
arg1.append(p2[j-1])
arg2.append(p2[j+1])
res.append("t"+str(ind))
you="t"+str(ind)             ind=ind+1
       j=j+1
oper.append(l[n])
arg1.append(my)
arg2.append(you)
res.append("t"+str(ind))
oper.append("=")
arg1.append(s[0])
arg2.append("t"+str(ind))
res.append("t"+str(ind+1))
z=zip(oper,arg1,arg2,res)
ans=[]
    for a1,a2,a3,a4 in list(z):
aq=[]       aq.append(a1)
aq.append(a2)
```

```
aq.append(a3)
aq.append(a4)
      ans.append(aq)
    print("QUADRAPLE TABLE:")
    print(tabulate(ans,                    headers=["OPERATORS","ARG                    1","ARG
2","RESULT"],tablefmt='orgtbl'))

    # print(a1,a2,a3,a4)
s=input("Enter code:") threeaddr(s)
```

OUTPUT:

```
Enter code:a = b * -c + b * -c
QUADRAPLE TABLE:
| OPERATORS   | ARG 1   | ARG 2   | RESULT   |
|-------------+---------+---------+----------|
| *           | b       | -c      | t1       |
| unary-      | c       | nill    | t2       |
| *           | b       | -c      | t3       |
| unary-      | c       | nill    | t4       |
| +           | t1      | t3      | t5       |
| =           | a       | t5      | t6       |
```

```
Enter code:a = b + c
QUADRAPLE TABLE:
| OPERATORS   | ARG 1   | ARG 2   | RESULT   |
|-------------+---------+---------+----------|
| +           | b       | c       | t1       |
| =           | a       | t1      | t2       |
```