

## UNIT-4

PDI

INTERMEDIATE CODE GENERATOR :ADD R<sub>1</sub>, R<sub>2</sub>ADD R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> → XADD R<sub>1</sub>, M<sub>1</sub>, M<sub>2</sub> → M<sub>2</sub>, M<sub>3</sub> added and store result in R<sub>1</sub>ADD M<sub>1</sub>, M<sub>2</sub>, M<sub>3</sub>ADD M<sub>1</sub>, M<sub>2</sub>

ADD → Binary operator

we write ADD RM } use different opcodes for different type of addition operation  
 ADD M

① ADDR A machine architecture containing

② ADD RM 8 diff op ⇒ 3 bits

③ SUB R 32 diff registers ⇒ 5 bits

④ SUBRM 1024 diff locations ⇒ 10 bits

⑤ MOV

⑥ LOAD

3	5	10	10
---	---	----	----

Maximum instruction length can be 28 bits.

Another Machine

15 diff operations ⇒ 4 bits

64 diff registers ⇒ 6 bits

4K memory size ⇒ 12 bits

4	6	12	12
---	---	----	----

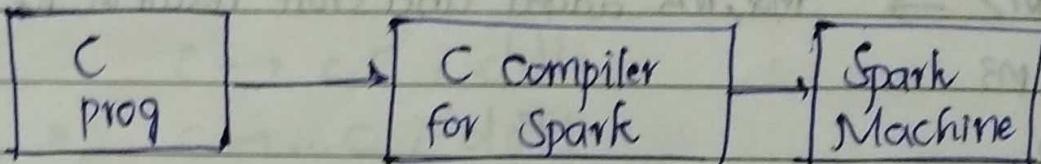
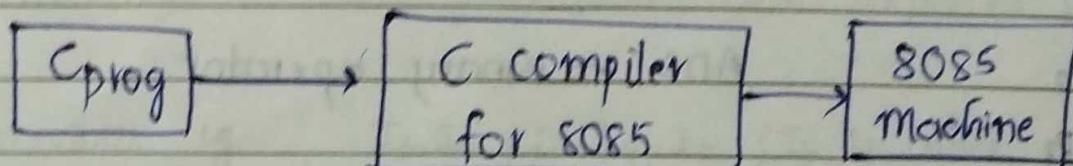
34 bits

Compiler design  
so depends on architectures

## Importance of ICG

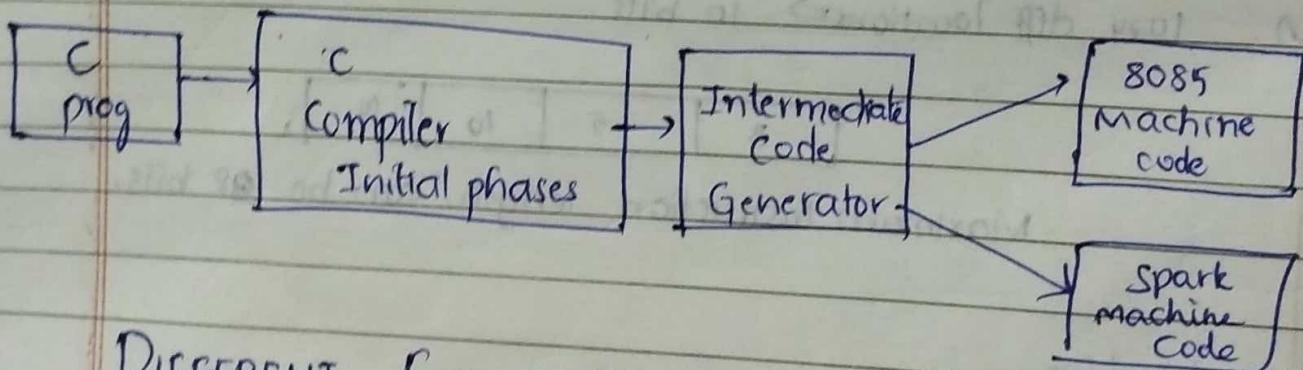
- \* If for any language, a compiler is designed and if it is to be run on different machines, then the entire compiler has to be reconstructed for the other machine.

Eg:



- \* If intermediate code generator is introduced, then, there is no need of re-construction of the compiler with all the phases for a new machine

- \* The code generated in Intermediate code is suitable for any machine and so, the portability of the compiler is increased.



## DIFFERENT FORMS OF INTERMEDIATE CODE GENERATION

1. Abstract Syntax Tree - similar to parse tree
2. Reverse Polish notation (postfix)
3. Three address code



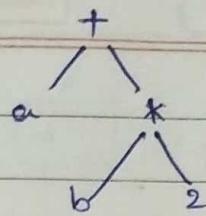
Eg :  $a+b * 2$

Eg : 2 :  $(a+b) * (c+d)$

DATE: / /

PAGE NO.:

parse tree :



Functions : mknode ( op , left , child )

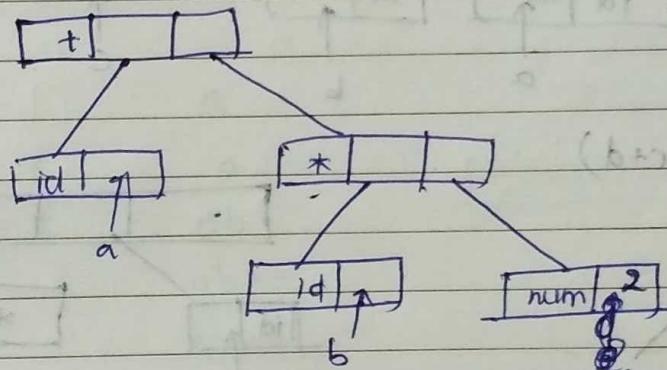
$\downarrow$   
operator  
 $\downarrow$   
subtree

mkleaf ( id, entry )

or mkleaf ( num , value )  
const

$a+b*2$

$\rightarrow$  binary op



2) postfix expression :

ch Stack

a

+

b

\*

2

Postfix exp.

a

a

ab

ab\*

ab\*2\*

ab\*2\*+

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

ch stack

Postfix

(

a

+

b

)

\*

(

c

+

d

)

Teacher's Signature

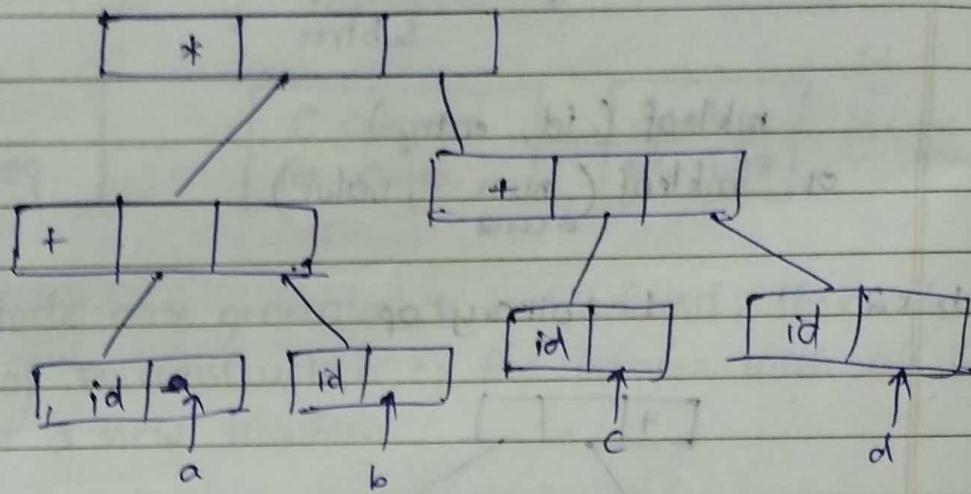
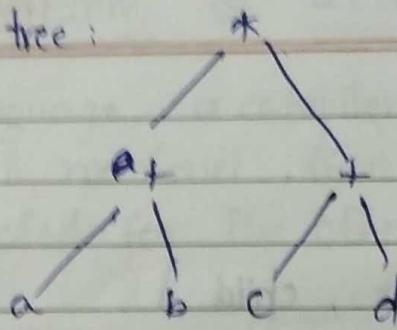
ab+cd\*+

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

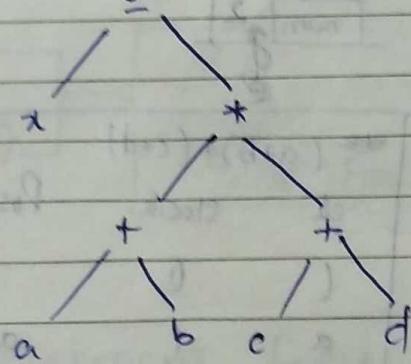
DATE: / /

PAGE NO.:

Parse tree :

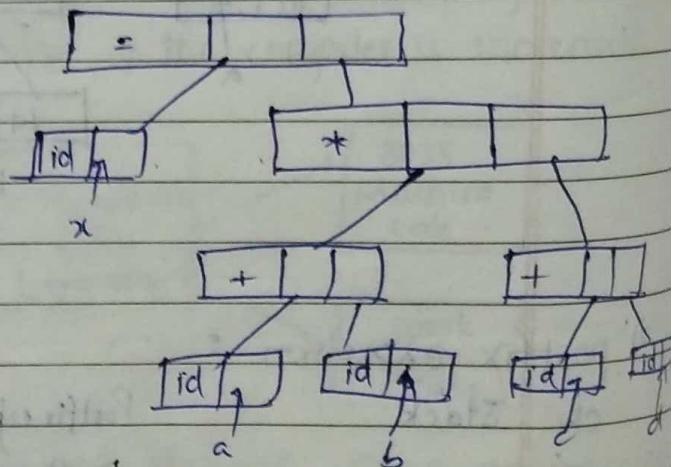


$x = (a+b)*(c+d)$



③  $* = ((-a)+b)+((-c)+d)$

\* Unary has highest priority



④  $a = (b+c)*(d/e)$

⑤  ~~$d = (a \& b) \& c \& ; (a|b) \& c$~~

⑥  $(a|b) \& (!c | !d)$

⑦  $((a+b) * (c+d)) - ((a+b) * d)$

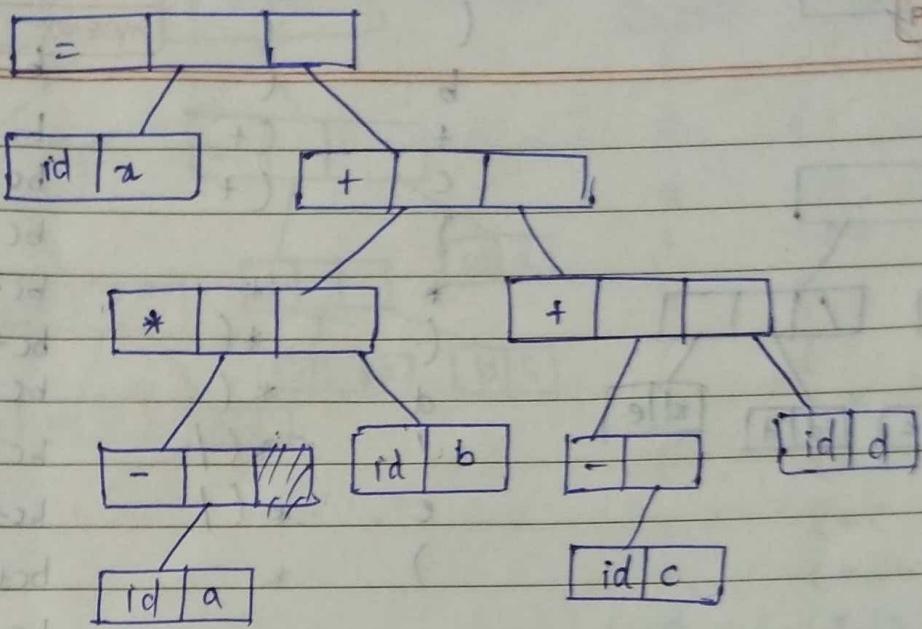
⑧  $x = (a * (b+c+(-a)))$

⑨  $x = (a * b + c + d) * b - d$

3)

$$x = (-a * b) + (-c * d)$$

DATE: \_\_\_\_\_  
PAGE NO.: \_\_\_\_\_

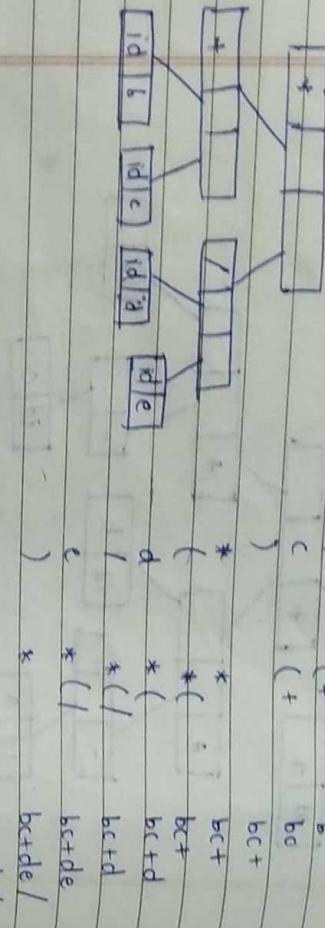
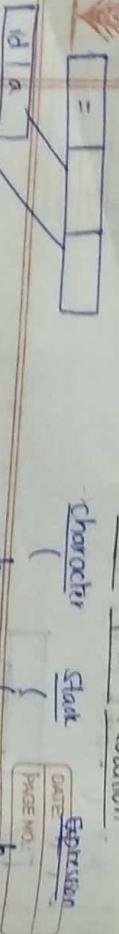


### Reverse Polish Notation:

character	stack	expression
(	(	
-	(-	
a	(-1)	a
*	(-* )	a
b	( -* )	ab
)	&	ab*-
+	+	ab*-
*	+ (	ab*-
-	+ (-	ab*-
c	+ (-	ab*-c
*	+ (-*	ab*-c
d	+ (-*	ab*-cd
)	+	ab*-cd*-
		ab*-cd*-+

$$d = (b+c)*(d/e)$$

### Reverse Polish Notation



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

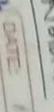
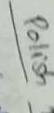
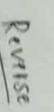
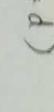
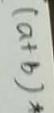
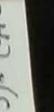
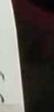
### Reverse Polish Notation

Character	Stack	Expression
=		
+		
a	a	
b	a b	
c	a b c	
*	a b c *	
d	a b c * d	
e	a b c * d / e	

$$b+c+d/e$$

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

Reverse Polish Notation	Character Stack	DATE / EXPRESSION
-		
*		
+		
a	a	
b	a b	
c	a b c	
d	a b c d	
*	a b c d *	
+	a b c d +	
a	a b c d + a	
b	a b c d + a b	
c	a b c d + a b c	
d	a b c d + a b c d	



10/05/22  
Thursday

## THREE ADDRESS CODES

DATE: / /  
PAGE NO.: / /

\* Three address code can be represented in 3 different forms

→ 1. Quadruples  
→ 2. Triples  
→ 3. Indirect triples.

\* for any of these 3 representations, the 3 address code have to be generated initially

\* In the 3 address code, each statement contains atmost 3 address locations.

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

TAC      Quadruple Representation (4 fields)

t<sub>1</sub> = a+b      s<sub>no</sub> op arg<sub>1</sub> arg<sub>2</sub> result

t<sub>2</sub> = c+d      o + a b t<sub>1</sub>

t<sub>3</sub> = t<sub>1</sub>\*t<sub>2</sub>      1 \* t<sub>1</sub> t<sub>2</sub> t<sub>3</sub>

t<sub>4</sub> = t<sub>3</sub>      2 = t<sub>3</sub> x

$$3) \quad x = (-a*b) + (-c*d)$$

TAC      Quadruple      Varony -

t<sub>1</sub> = -a      s.no op arg<sub>1</sub> arg<sub>2</sub> result

t<sub>2</sub> = -c      o u- a t<sub>1</sub>

t<sub>3</sub> = t<sub>1</sub>\*b      1 u- c t<sub>2</sub>

t<sub>4</sub> = t<sub>2</sub>\*d      2 \* t<sub>1</sub> b t<sub>3</sub>

t<sub>5</sub> = t<sub>3</sub>+t<sub>4</sub>      3 \* t<sub>2</sub> d t<sub>4</sub>

x = t<sub>5</sub>      4 + t<sub>3</sub> t<sub>4</sub> x

$$4) \quad x = (b+c)*(d+e)$$

TAC:      s<sub>no</sub> op arg<sub>1</sub> arg<sub>2</sub> result

t<sub>1</sub> = b+c      o + b c t<sub>1</sub>

t<sub>2</sub> = d+e      1 + d e t<sub>2</sub>

t<sub>3</sub> = t<sub>1</sub>\*t<sub>2</sub>      2 \* t<sub>1</sub> t<sub>2</sub> t<sub>3</sub>

x = t<sub>3</sub>      3 = t<sub>3</sub> x

5)  $d = (a+b)*c$       Quadruple  
TAC:      s<sub>no</sub> op arg<sub>1</sub> arg<sub>2</sub>  
t<sub>1</sub> = a+b      o + a b t<sub>1</sub>  
t<sub>2</sub> = t<sub>1</sub>\*c      1 \* t<sub>1</sub> c t<sub>2</sub>  
d = t<sub>2</sub>      2 = t<sub>2</sub> d

DATE: / /  
PAGE NO.: / /

6)  $(a+b)*c - (a+b)*d$       Quadruple  
TAC:      s<sub>no</sub> op arg<sub>1</sub> arg<sub>2</sub> result  
t<sub>1</sub> = a+b      o + a b t<sub>1</sub>  
t<sub>2</sub> = c+d      1 + c d t<sub>2</sub>  
t<sub>3</sub> = t<sub>1</sub>\*t<sub>2</sub>      2 \* t<sub>1</sub> t<sub>2</sub> t<sub>3</sub>  
t<sub>4</sub> = t<sub>1</sub>\*d      3 \* t<sub>1</sub> d t<sub>4</sub>  
t<sub>5</sub> = t<sub>3</sub>-t<sub>4</sub>      4 - t<sub>3</sub> t<sub>4</sub> t<sub>5</sub>

TAC:      s<sub>no</sub> op arg<sub>1</sub> arg<sub>2</sub> result  
t<sub>1</sub> = a+b      o + a b t<sub>1</sub>  
t<sub>2</sub> = c+d      1 + c d t<sub>2</sub>  
t<sub>3</sub> = t<sub>1</sub>\*t<sub>2</sub>      2 \* t<sub>1</sub> t<sub>2</sub> t<sub>3</sub>  
t<sub>4</sub> = t<sub>1</sub>\*d      3 \* t<sub>1</sub> d t<sub>4</sub>  
t<sub>5</sub> = t<sub>3</sub>-t<sub>4</sub>      4 - t<sub>3</sub> t<sub>4</sub> t<sub>5</sub>

7)  $x = (a*(b+c)) - (a*(b+c))$

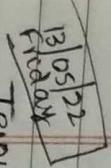
TAC:      s<sub>no</sub> op arg<sub>1</sub> arg<sub>2</sub> result  
t<sub>1</sub> = -a      o - a t<sub>1</sub>  
t<sub>2</sub> = b+c      1 + b c t<sub>2</sub>  
t<sub>3</sub> = t<sub>1</sub>\*t<sub>2</sub>      2 \* t<sub>1</sub> t<sub>2</sub> t<sub>3</sub>  
t<sub>4</sub> = a\*t<sub>3</sub>      3 \* a t<sub>3</sub> t<sub>4</sub>  
x = t<sub>4</sub>      4 = t<sub>4</sub> x

TAC:      s<sub>no</sub> op arg<sub>1</sub> arg<sub>2</sub> result  
t<sub>1</sub> = -a      o - a t<sub>1</sub>  
t<sub>2</sub> = b+c      1 + b c t<sub>2</sub>  
t<sub>3</sub> = t<sub>1</sub>\*t<sub>2</sub>      2 \* t<sub>1</sub> t<sub>2</sub> t<sub>3</sub>  
t<sub>4</sub> = a\*t<sub>3</sub>      3 \* a t<sub>3</sub> t<sub>4</sub>  
x = t<sub>4</sub>      4 = t<sub>4</sub> x

$$x = (a * b + c * d) * b - d$$

Quadruple

DATE: / /  
PAGE NO./RESULT:



### TRIPLES

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

### INDIRECT TRIPLES

s.no	op	arg1	arg2	s.no	op	arg1	arg2
(1)	*	a	b	1	*	c	d
(2)	+	c	d	2	+	t2	t2
(3)	*	b	t3	3	*	t4	t4
(4)	-	t4	d	4	=	t5	t5
(5)	=	t5	x	x			

$$x = ((a+b)*(c+d)) - ((a+b)*(d))$$

### INDIRECT TRIPLES

s.no	op	arg1	arg2	s.no	op	arg1	arg2
(1)	+	a	b	(1)	+	a	b
(2)	*	c	d	(2)	!	d	
(3)	*	(1)	(2)	(3)	!	(11)	(12)
(4)	*	(2)	(1)	(4)	!	(10)	(13)
(5)	=	x	(2)	(5)	=	x	(14)

$$x = ((a+b)*(c+d)) - ((a+b)*d)$$

### TRIPLES

s.no	op	arg1	arg2	s.no	op	arg1	arg2
(1)	+	a	b	(1)	+	a	b
(2)	*	c	d	(2)	+	c	d
(3)	*	(1)	(2)	(3)	*	(10)	(11)
(4)	*	(2)	(1)	(4)	*	(10)	d
(5)	=	x	(14)	(5)	=	x	(14)

### INDIRECT TRIPLES

s.no	op	arg1	arg2	s.no	op	arg1	arg2
(1)	-	a		(1)	-	a	
(2)	+	b	c	(2)	+	b	c
(3)	*	d	e	(3)	*	d	e
(4)	=	x	(2)	(4)	=	x	(3)

$$x = (b+c)*(d/e)$$

### TRIPLES

s.no	op	arg1	arg2	s.no	op	arg1	arg2
(1)	+	b	c	(1)	+	b	c
(2)	*	d	e	(2)	*	d	e
(3)	=	x	(2)	(3)	=	x	(2)

### INDIRECT TRIPLES

s.no	op	arg1	arg2	s.no	op	arg1	arg2
(1)	+	b	c	(1)	+	b	c
(2)	*	d	e	(2)	*	d	e
(3)	=	x	(2)	(3)	=	x	(2)

INDIRECT TRIPLES  
Indirect table  
PAGE NO.:

Teacher's Signature:

9)

$$x = (a * b + c + d) * b - d$$

TRIPLES

~~16/05/22  
Monday~~

END	op	arg1	arg2	s-no	op	arg1	arg2
(1)	*	a	b	(6)	*	a	b
(1)	+	c	d	(1)	+	c	d
(2)	+	(2)	(1)	(2)	+	(16)	(11)
(3)	*	(6)	(2)	(3)	*	(6)	(12)
(4)	-	(3)	d	(4)	-	(B)	d
(5)	=	x	(4)	(5)	=	x	(14)

UNIT - 3

DATE:	/ /
PAGE NO.:	

16/05/22  
Monday

Syntax DIRECTED

- The value of an attribute at any node of parse tree is defined by a semantic rule associated with the production used as the node.
- There are 2 different attributes.
  - i) Synthesized attributes
  - ii) Inherited attributes
- The value of a synthesized attribute at any node is computed from the values of the attribute at the child nodes of that node in the parse tree

Attributed Grammar with synthesized attributes

Attributed Grammar with synthesized attributes	SDD / Attributed Grammar
$E \rightarrow E_1 + T$	$E \cdot \text{val} = E_1 \cdot \text{val} + T \cdot \text{val}$
$T \rightarrow T_1 * F$	$E \cdot \text{val} = T_1 \cdot \text{val}$
$F \rightarrow (E)$	$T_1 \cdot \text{val} = T_1 \cdot \text{val} * F \cdot \text{val}$
$F \rightarrow \text{digit}$	$F \cdot \text{val} = E \cdot \text{val}$
	$F \cdot \text{val} = \text{digit} \cdot \text{lexvalue}$

NOTE: A parse tree representing the values of its attributes or function using SDD is known as Annotated parse tree.

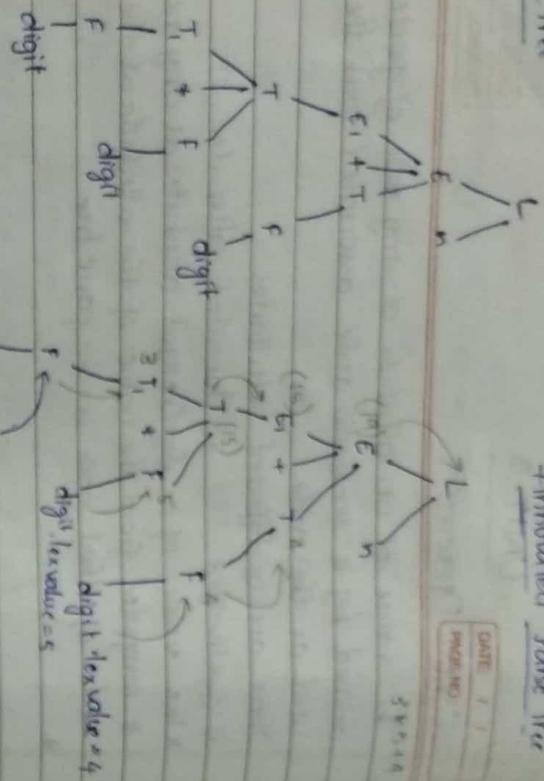
- Construct annotated parse tree for the string  $3 * 5 + 4$  using the above given grammar.

# Parse Tree

## Annotated Parse Tree

DATE: / /  
 PAGE NO.:

3454A



String: 1)  $(3+4)* (5+6)^n$

2)  $1 * 2 * 3 * (4 + 5)^n$

digit.lex value = 3

String: 1)  $(3+4)* (5+6)^n$

2)  $1 * 2 * 3 * (4 + 5)^n$

digit.lex value = 3

## Inherited Attributes

eg:  $D \rightarrow TL$   
 $L \rightarrow int$   
 $T \rightarrow real^1$   
 $T \rightarrow L_1, id$   
 $L \rightarrow id$

$L \rightarrow L_1, id$

## Attributed Grammar with inherited attributes

$L.in = T.type$

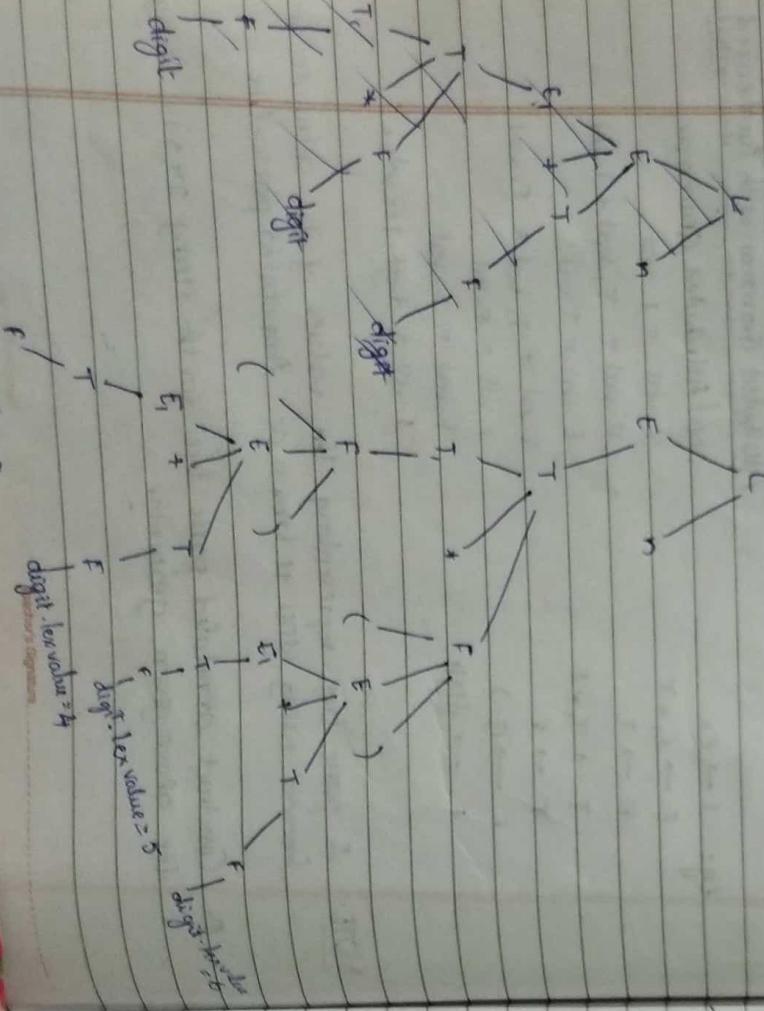
$T.type = int$

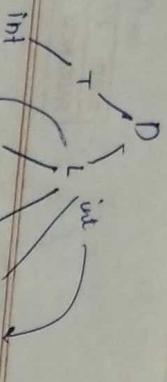
$T.type = real$

$L.in = L.in.addType(id.entry, L.in)$

$addType(id.entry, L.in)$

int a, a2, a3





20/5/22  
Friday

int

$L_1$

int

$L_1$

int

$L_1$

int

$L_1$

int

$L_1$

int

20/5/22  
Friday

### Dependency Graph

NOTE: The interdependencies among the inherited and synthesized attributes can be represented in a parse tree with the help of the directed graph known as Dependency Graph.

e.g.  $E \rightarrow E + T$       Q) Construct SDD for the given CFG or  
 $E \rightarrow E - T$       also construct the syntax tree for the  
 $E \rightarrow T$       expression  $a - 4 + c$

$T \rightarrow (E)$

$T \rightarrow id$

$T \rightarrow num$

Syntax trees a-4+c

$p_1 = \text{mkleaf}(\text{id}, \text{entry } a)$        $E_i.\text{ptr} = \text{mknode}(+, E_i.\text{ptr}, T_i.\text{ptr})$

$p_2 = \text{mkleaf}(\text{num}, 4)$

$E_i.\text{ptr} = T_i.\text{ptr}$

$p_3 = \text{mknode}(-, p_1, p_2)$

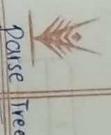
$T_i.\text{ptr} = E_i.\text{ptr}$

$p_4 = \text{mkleaf}(\text{id}, \text{entry } c)$

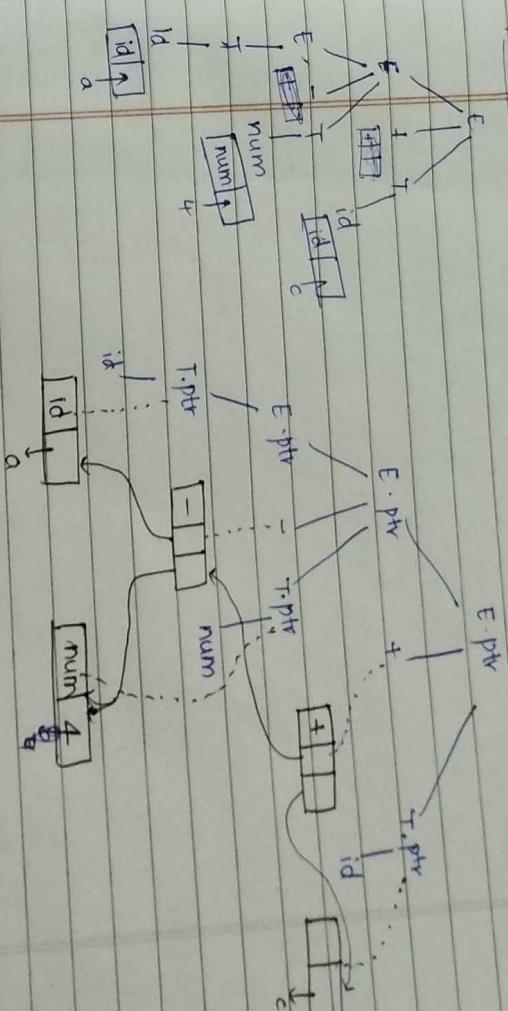
$T_i.\text{ptr} = \text{mkleaf}(\text{id}, \text{id}. \text{entry})$

$p_5 = \text{mknode}(+, p_3, p_4)$

$T_i.\text{ptr} = \text{mkleaf}(\text{num}, \text{value})$



SDD for CFG to  
construct  
parse tree





Q. Represent activation records for the recursive function factorial (3)

### Access to Non-local Names:

~~WLOG B3~~  
Temporary

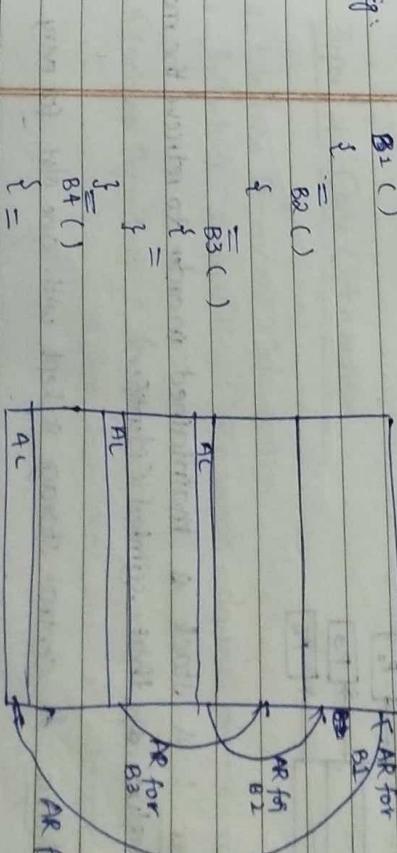
- \* Return value is the value returned by the procedure
- \* It consists of the values of the actual parameters
- \* It points to the activation record of the calling procedure
- \* Access link: It refers to the non-local data in other activation records.
- \* Saved Machine status: Before the procedure is called, some control information will be stored in the registers
- \* Local Variables: It holds all the local variables of a procedure.
- \* Temporaries: All the temporary variables generated by the compiler in order to execute the procedure are stored in the field.

1. Static Scope (or Lexical Scope):  
Two types of representations:  
1. Access Link, 2. Display

2. Dynamic Scope:  
1. Deep Access 2. Shallow access.

1. Access Link:  
In access link, the activation record is represented using the access links.

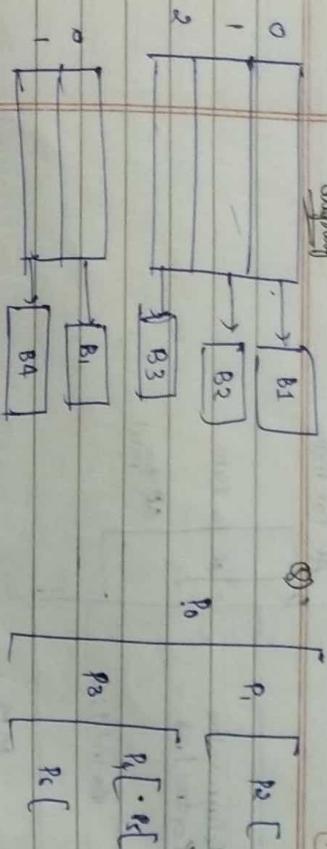
Access Link Representation of static scope



Eg:  
main () {  
 float sum(x,y)  
 {  
 =  
 =  
 a = sum(b,c);  
 int mul(p,q)  
 b = mul(c,f);  
 =  
 }  
}

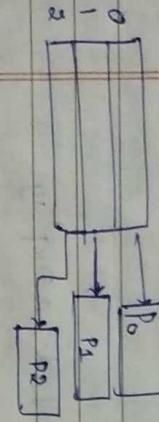
2. In this representation, an auxiliary array is used

DATE: / /  
PAGE NO.:



Represent the access link and display representations for the given procedures

Access Link



16 Oct 22  
Thursday

## CODE OPTIMIZATION

UNIT - V

\* Intermediate code is the input for this phase.

\* Optimized Intermediate code is the output.

\* The main tasks of code optimizer are

1. The semantic equivalence of the program should not be changed.
2. It should produce efficient object code.

\* The main challenges for Code Optimizer

1. The optimising algorithm must be efficient enough in such a way that the target code should not become time and space inefficient
2. The algorithm should preserve its semantic equivalence

### Types of Code Optimization

1. Machine Dependent Optimization
2. Machine Independent Optimization - depends on programming language

1. Deep Access : A stack is maintained in order to retrieve the most recent value of that symbol (Identifier)

2) Shallow Access: A central storage is kept with one slot for any variable

B1( )  
int a = 15  
p1(a);

B2( )  
= B1  
int a = 20;  
p2(a);

B3( )  
= B2  
int a = 10;  
p3(a);

Shallow Access  
10 value of a in B3  
20 value of a in B2

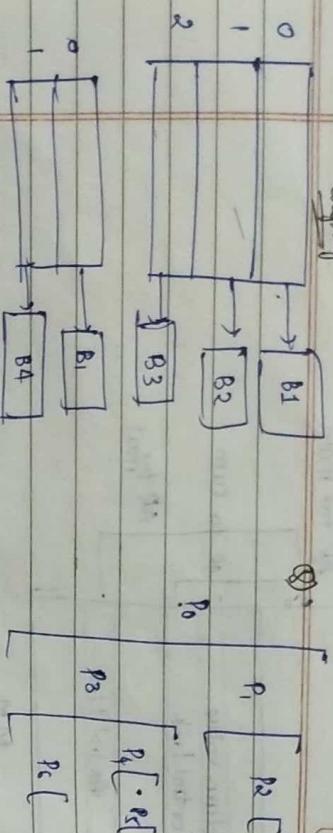
Deep Access  
15 value of a in B1  
20 value of a in B2

DATE: / /  
PAGE NO.:

## Display :

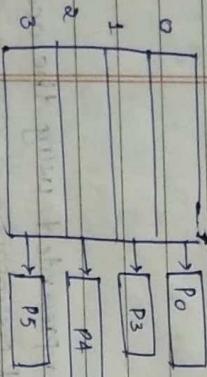
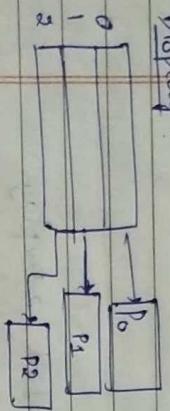
In this representation, an auxiliary array is used

DATE: / /  
PAGE NO.:



Represent the access link and display representations for the given procedures

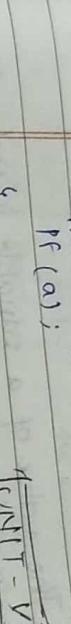
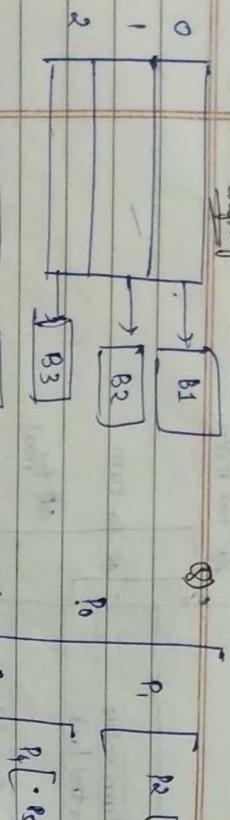
### Access Link



- 1) Deep Access : A stack is maintained in order to retrieve the most recent value of that symbol (identifier)
- 2) Shallow Access : A central storage is kept with one slot for every variable

## Deep Access

DATE: / /  
PAGE NO.:



## CODE OPTIMIZATION

Intermediate code is the input for this phase.

\* Optimised Intermediate code is the output.

\* The main tasks of code optimizer are

- 1. The semantic equivalence of the program should not be changed.
- 2. It should produce efficient object code.
- 3. The main challenges for Code Optimizer must be efficient enough in such a way that the target code should not become time and space inefficient.
- 4. The algorithm should preserve its semantic equivalence

## Types of Code Optimization

- 1. Machine Dependent Optimization
- 2. Machine Independent Optimization - depends on programming language

- 1. Machine Dependent Optimization : It depends on the architecture of the machine i.e. the number of registers storage
- 2) Shallow Access : A central storage is kept with one slot for every variable

DATE: / /  
PAGE NO.:

DATE: / /  
PAGE NO.:





2. Flow of Control Optimization: Unnecessary jumps can be eliminated.

Eg: go to l1;  
~~goto l2~~

l1 :

goto l2  
=

l2 :  
=

Optimized code

goto l2;  
l1 :

go to l2  
=

l2 :  
=

3. Algebraic Simplification

$x = x + 0$  or  $x = x \times 1$  or  $x = x / 1$  can be reduced to by using algebraic identities.

4. Strength Reduction:

\* can be replaced with '+' and '\*' can be replaced by:

5. Machine Idioms: increment or decrement instructions can be used instead of executing the binary operations.

Eg:  $i = i + 1$  can be replaced by ~~INC i~~ INC i