Lab Basic Blocks

J. Nelson Amaral

A Basic Block

Only the first instruction of a basic block can be the target of a branch.

Execution can only start here

If any instruction of the basic block is executed, then all of them must be executed.

TARGET: andi \$t2
add \$t
or \$t
bnez \$t

\$t2, \$t0, 0x1 \$t2, \$t2, \$t4 \$t2, \$t3, \$t0 \$t2, ODD

Only the last instruction of the basic block can be a branch or a jump.

Execution can only stop here

A Program Example

```
int odd_series(int p){
int j=0;
for (int i=0; i<p; i++){
    j += i; odd_series:
  if(i%2)
               li $t0, 0 # i <-- 0
  else
                 li $v0, 0 # j <-- 0
    j++;
                blez $a0, DONE # if (p <= 0) goto DONE
           L00P:
return j;
                andi $t2, $t0, 0x1 # $t2 <-- bit0 of i
                  bnez $t2 ODD # if i is odd goto ODD
                  add $v0, $v0, $t0 # j <-- j+i
                       REINIT
           ODD:
                  add $v0, $v0, 1 # j <-- j+1
           REINIT:
                  add $t0, $t0, 1 # i <-- i+1
                  blt $t0, $a0, LOOP # if i<p goto LOOP
           DONE:
                  jr
                       $ra
```

Finding Leaders

Rule 1: The first statement of a procedure is a leader.

```
odd_series:
      li
            $t0, 0
                         # i <-- 0
                       # j <-- 0
       li
            $v0, 0
       blez $a0, DONE # if (p <= 0) goto DONE
L00P:
       andi $t2, $t0, 0x1 # $t2 <-- bit0 of i
       bnez $t2 ODD # if i is odd goto ODD
       add $v0, $v0, $t0 # j <-- j+i
            REINIT
ODD:
       add $v0, $v0, 1 # j <-- j+1
REINIT:
       add $t0, $t0, 1 # i <-- i+1
       blt $t0, $a0, LOOP # if i<p goto LOOP
DONE:
       jr
            $ra
```

Rule 2: Any target of a branch or jump is a leader.

```
odd_series:
       li
            $t0, 0
                          # i <-- 0
                        # j <-- 0
       li
            $v0, 0
       blez $a0, DONE # if (p <= 0) goto DONE
L00P:
       andi $t2, $t0, 0x1 | # $t2 <-- bit0 of i
       bnez $t2 ODD # if i is odd goto ODD
            $v0, $v0, $t0 # j <-- j+i
       add
            REINIT
ODD:
            $v0, $v0, 1 | # j <-- j+1
       add
REINIT:
            $t0, $t0, 1 # i <-- i+1
       add
            $t0, $a0, LOOP # if i<p goto LOOP
       blt
DONE:
            $ra
       jr
```

Rule 3: Any instruction that follows a branch or jump is a leader.

```
odd_series:
             $t0, 0
       li
                           # i <-- 0
                         # j <-- 0
       li
             $v0, 0
       blez $a0, DONE # if (p <= 0) goto DONE
L00P:
       andi $t2, $t0, 0x1 | # $t2 <-- bit0 of i
             $t2 ODD
                           # if i is odd goto ODD
       bnez
             $v0, $v0, $t0 | # j <-- j+i
       add
             REINIT
ODD:
             $v0, $v0, 1 | # j <-- j+1
       add
REINIT:
             $t0, $t0, 1 # i <-- i+1
       add
             $t0, $a0, LOOP # if i<p goto LOOP
       blt
DONE:
             $ra
       jr
```

Instruction	Enco	ding						
beq \$s, \$t, offset	0001	00ss	ssst	tttt	0000	0000	0000	0000
bgez \$s, offset	0000	01ss	sss0	0001	0000	0000	0000	0000
bgezal \$s, offset	0000	01ss	sss1	0001	0000	0000	0000	0000
bgtz \$s, offset	0001	11ss	sss0	0000	0000	0000	0000	0000
blez \$s, offset	0001	10ss	sss0	0000	0000	0000	0000	0000
bltz \$s, offset	0000	01ss	sss0	0000	0000	0000	0000	0000
bltzal \$s, offset	0000	01ss	sss1	0000	0000	0000	0000	0000
bne \$s, \$t, offset	0001	01ss	ssst	tttt	0000	0000	0000	0000
j target	0000	10aa	aaaa	aaaa	aaaa	aaaa	aaaa	aaaa
jal target	0000	11aa	aaaa	aaaa	aaaa	aaaa	aaaa	aaaa
jr \$s	0000	00ss	sss0	0000	0000	0000	0000	1000

Finding the Leaders – Lab #2

Instruction	Encod	ling						
beq \$s, \$t, offset	0001	00ss	ssst	tttt	0000	0000	0000	0000
bgez \$s, offset	0000	01ss	sss0	0001	0000	0000	0000	0000
bgezal \$s, offset	0000	01ss	sss1	0001	0000	0000	0000	0000
bgtz \$s, offset	0001	11ss	sss0	0000	0000	0000	0000	0000
blez \$s, offset	0001	10ss	sss0	0000	0000	0000	0000	0000
bltz \$s, offset	0001	10ss	sss0	0000	0000	0000	0000	0000
bne \$s, \$t, offset	0001	01ss	ssst	tttt	0000	0000	0000	0000
j target	0000	10aa	aaaa	aaaa	aaaa	aaaa	aaaa	aaaa
jal target	0000	11aa	aaaa	aaaa	aaaa	aaaa	aaaa	aaaa
jr \$s	0000	00ss	sss0	0000	0000	0000	0000	1000

Given the Leaders, How do we get the Basic Blocks?

A basic block is the leader followed by all subsequent instructions up to, but not including, the next leader.

```
odd_series:
             $t0, 0
             $v0, 0
       li
                           # if (p <= 0) goto DONE
       blez
             $a0, DONE
L00P:
             $t2, $t0, 0x1 | # $t2 <-- bit0 of i
       andi
                           # if i is odd goto ODD
             $t2 ODD
       bnez
             $v0, $v0, $t0
       add
                           # j <-- j+i
             RFTNTT
ODD:
             $v0, $v0, 1
       add
                           \# j <-- j+1
REINIT:
             add
             $t0, $a0, LOOP # if i<p goto LOOP
       blt
DONE:
       jr
             $ra
```

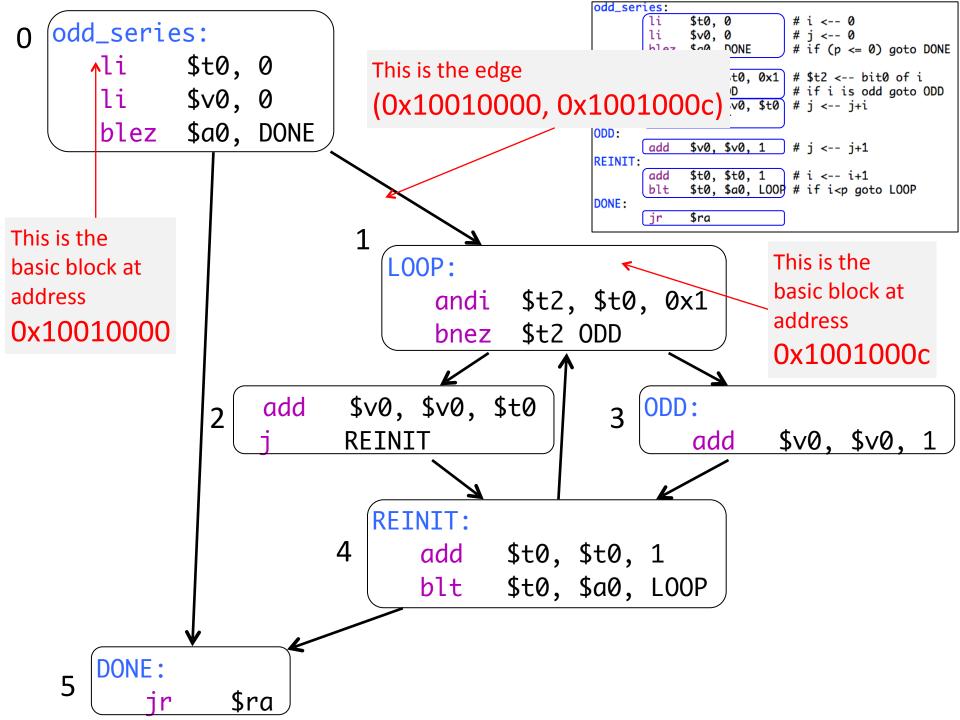
A basic block is the leader followed by all subsequent instructions up to, but not including, the next leader.

```
odd_series:
              $t0, 0
              $v0, 0
                            # if (p <= 0) goto DONE
        blez
              $a0, DONE
L00P:
        andi $t2, $t0, 0x1 | # $t2 <-- bit0 of i
                             # if i is odd goto ODD
              $t2 ODD
        bnez
              $v0, $v0, $t0
        add
                             # j <-- j+i
              RFTNTT
ODD:
              $v0, $v0, 1 | # j <-- j+1
        add
REINIT:
              $t0, $t0, 1 # i <-- i+1
        add
              $t0, $a0, LOOP # if i<p goto LOOP
        blt
DONE:
       jr
              $ra
```

The Actual Code

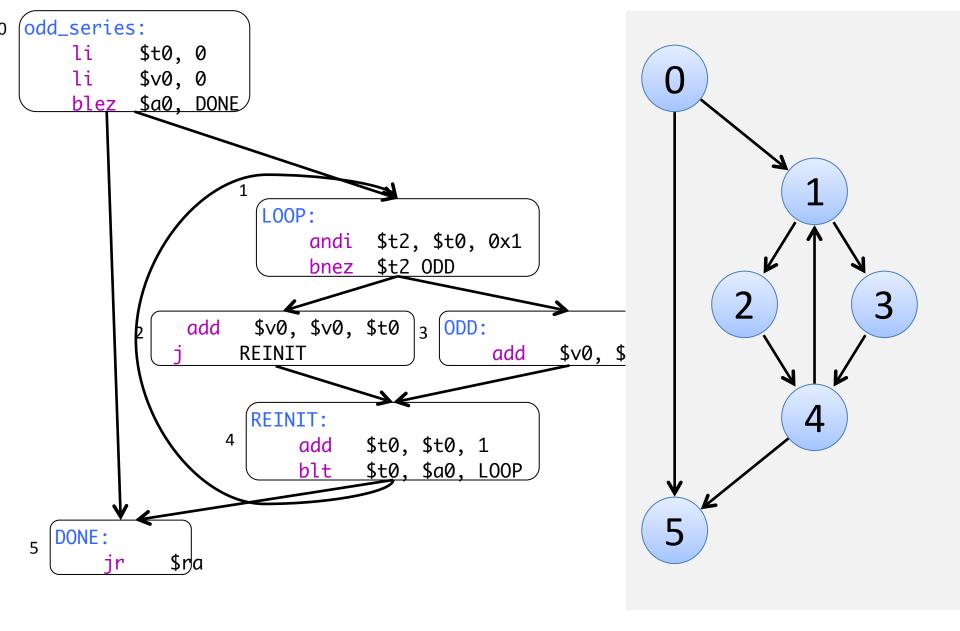
```
100100001
            34080000
                          $8,
                             $0,0
                      ori $2, $0, 0
            34020000
                                                    li $v0, 0 # i
            18800009
                      blez $4 36 [DONE-0x10010008];
                                                  54: blez $a0, DONE #
  1001000cl
            310a0001
                      andi $10,
                                6 block(s) found.
 $t2
  100100101
           15400003
                                Block Leader: 0x10010000, Size: 3
  10010014]
            00481020
                        0×100100 Block Leader: 0x1001000C, Size: 2
            08100029
  100100181
                              Block Leader: 0x10010014, Size: 2
  [1001001c]
            20420001
                              Block Leader: 0x1001001C, Size: 1
  [10010020]
            21080001
  [10010024]
            0104082a
                      slt $1, $8
                                Block Leader: 0x10010020, Size: 3
  10010028]
            1420fff9
                      bne $1, $0
                                Block Leader: 0x1001002C, Size: 1
5 [1001002c] 03e00008
                      jr $31
```

Control Flow Graphs

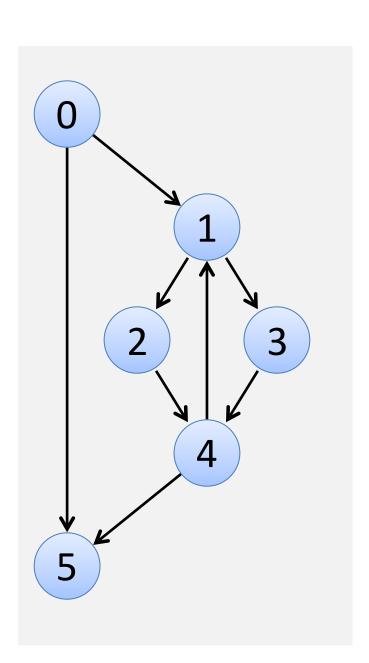


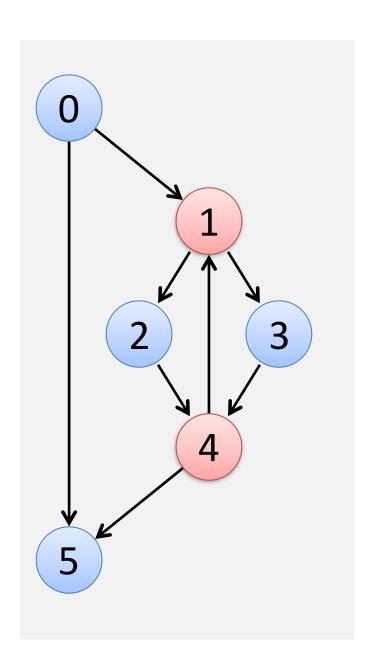
The Actual Code

```
100100001
           34080000
                    ori $8, $0, 0
                     ori $2, $0, 0
                                                 li $v0, 0 # j
           34020000
                     blez $4 36 [DONE-0x10010008]; 54: blez $a0, DONE #
  100100081
            18800009
  if (p
                                           1001000c] 310a0001
                     andi $10, $8, 1
 $t2
                                   Edges:
                     bne $10, $0, 12
  [10010010] 15400003
                                   0x10010000 --> 0x1001000C
                     add $2, $2, $8
  10010014]
           00481020
  10010018]
           08100029
                      0×10010020
                                [REOx10010000 --> 0x1001002C
5 [1001001c] 20420001
                     addi $2, $2,
                                   0x1001000C \longrightarrow 0x10010014
                     addi $8, $8, 1
  [10010020]
           21080001
                                   0x1001000C --> 0x1001001C
 [10010024] 0104082a
                     slt $1, $8, $4
                                    0x10010014 --> 0x10010020
  [10010028] 1420fff9
                     bne $1, $0, -28
                                    0x1001001C --> 0x10010020
5 [1001002c] 03e00008
                     jr $31
                                   0x10010020 --> 0x1001000C
                                    0x10010020 --> 0x1001002C
```

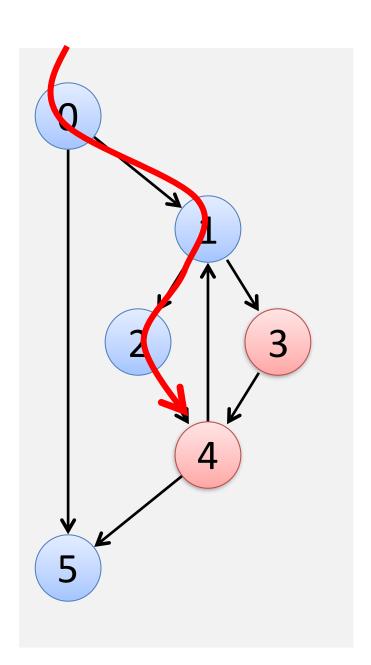


Dominators



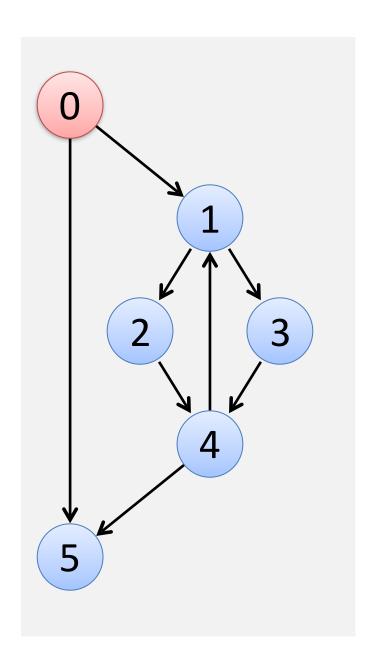


1 <u>dominates</u> 4 because you cannot execute 4 unless you have already executed 1



1 <u>dominates</u> 4 because you cannot execute 4 unless you have already executed 1

3 <u>does not dominate</u> 4 because there is a path $(0\rightarrow 1\rightarrow 2\rightarrow 4)$ that reaches 4 without executing 3.

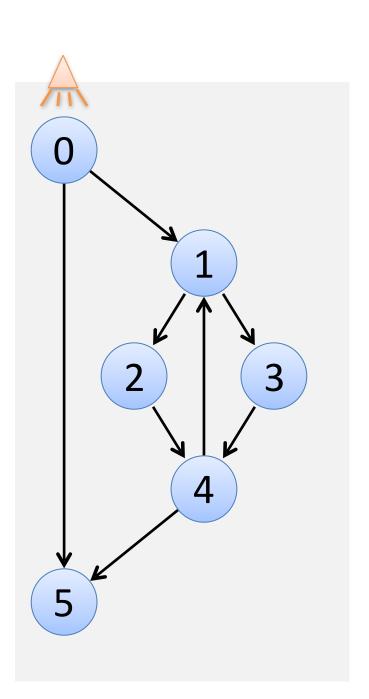


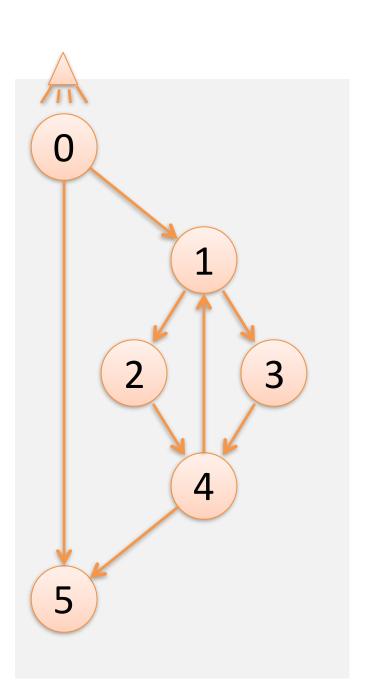
0 dominates all the basic blocks.

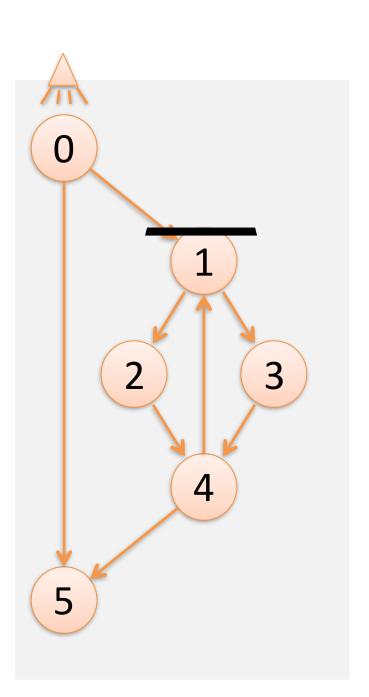
1 <u>dominates</u> 4 because you cannot execute 4 unless you have already executed 1

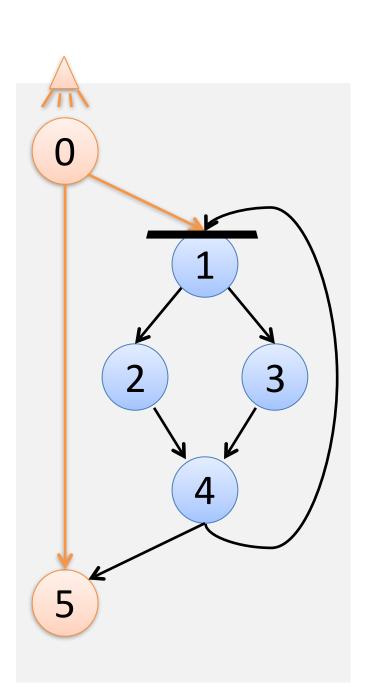
3 <u>does not dominate</u> 4 because there is a path $(0\rightarrow 1\rightarrow 2\rightarrow 4)$ that reaches 4 without executing 3.

A basic block dominates itself.

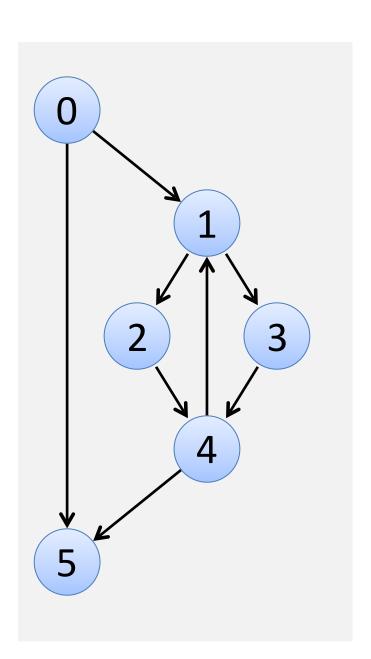




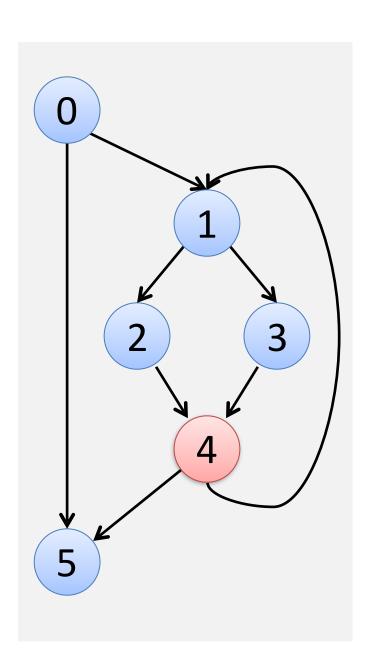




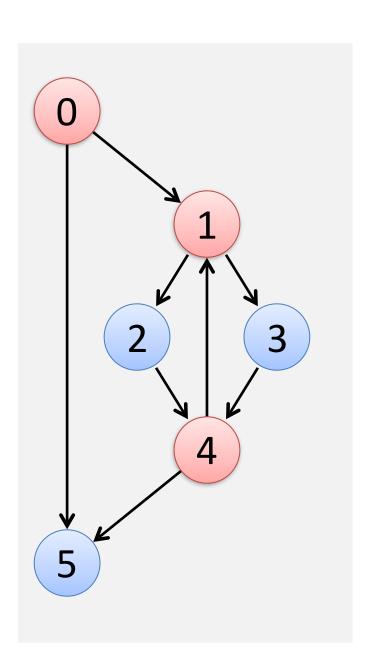
1 dominates nodes 1, 2, 3 and 4.



Which nodes dominate node 4?



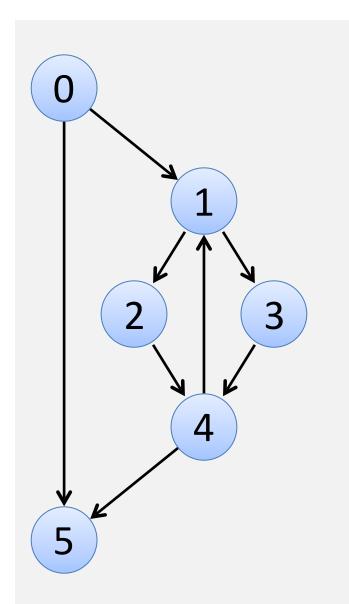
Which nodes dominate node 4?



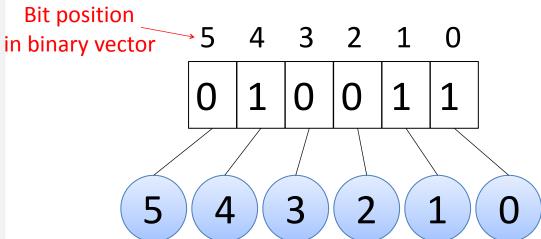
Which nodes dominate node 4?

$$Dom(4) = \{0, 1, 4\}$$

This is called the *Dominator Set* of 4

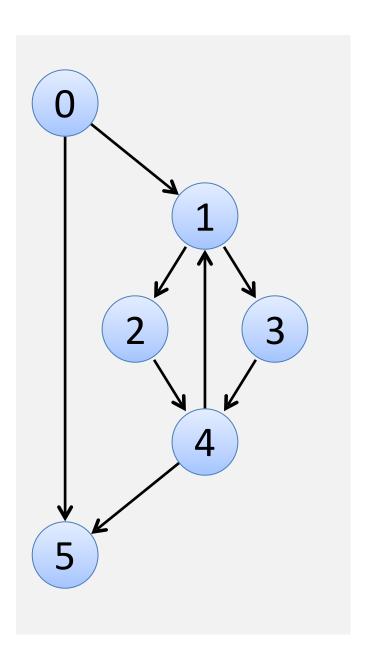


We can use a bit in a binary vector to represent each node in the CFG



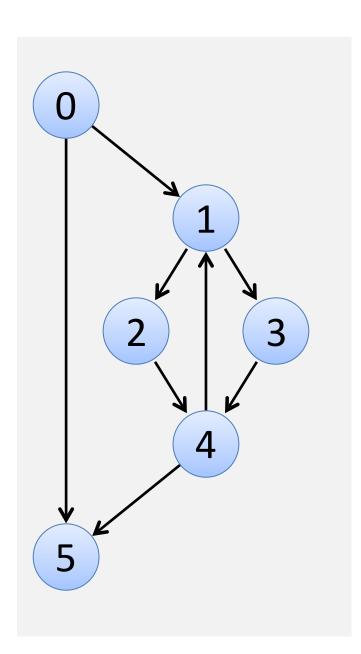
We can now represent the dominator set of 4 as a binary vector

$$Dom(4) = \{0, 1, 4\}$$



Dominator Bit Vectors:

How to Compute Dominator Sets?

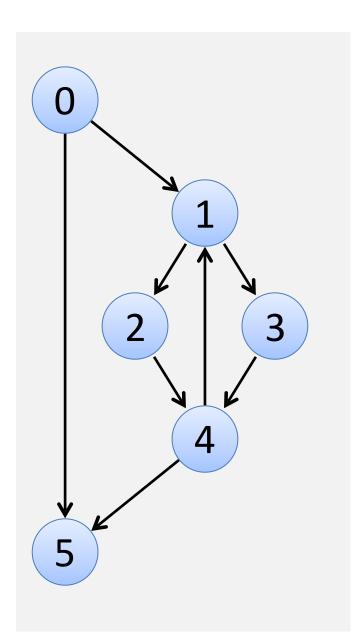


Initialization:

0 is the only dominator of 0

All other nodes are dominated by every node.

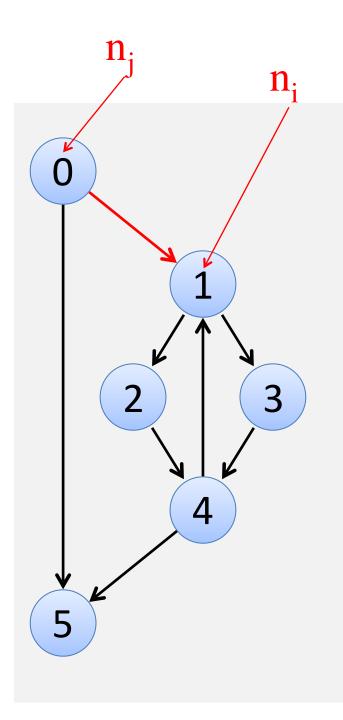
Node	Dominators							
	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	1	1	1	1	1	1		
2	1	1	1	1	1	1		
3	1	1	1	1	1	1		
4	1	1	1	1	1	1		
5	1	1	1	1	1	1		



For each edge (n_j, n_i) in CFG:

 $Dom(n_i) = \{n_i\} \cup (Dom(n_i) \cap Dom(n_i))$

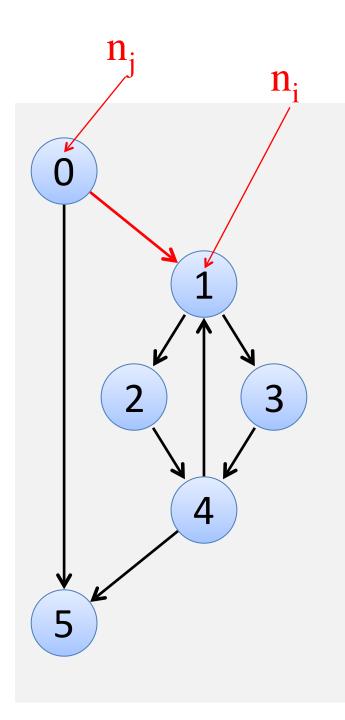
Node	Dominators							
	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	1	1	1	1	1	1		
2	1	1	1	1	1	1		
3	1	1	1	1	1	1		
4	1	1	1	1	1	1		
5	1	1	1	1	1	1		



For each edge (n_i, n_i) in CFG:

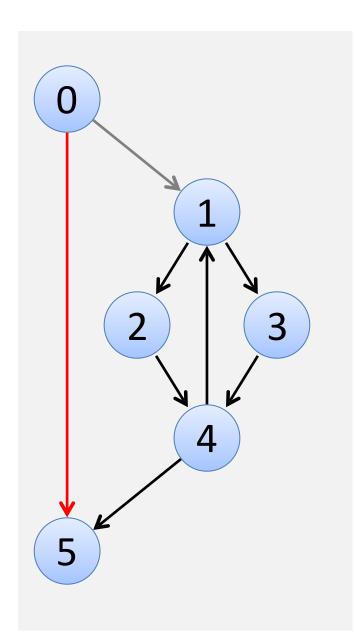
$$Dom(n_i) = \{n_i\} \cup (Dom(n_i) \cap Dom(n_j))$$
$$Dom(1) = \{1\} \cup (\{0,1,2,3,4,5\} \cap \{0\})$$
$$Dom(1) = \{1\} \cup \{0\} = \{0,1\}$$

Node	Dominators								
	5	4	3	2	1	0			
0	0	0	0	0	0	1			
1	1	1	1	1	1	1			
2	1	1	1	1	1	1			
3	1	1	1	1	1	1			
4	1	1	1	1	1	1			
5	1	1	1	1	1	1			

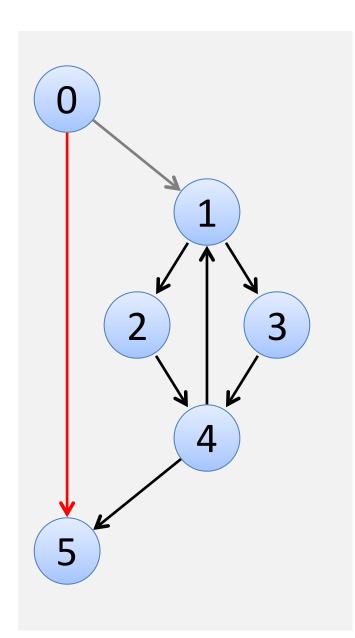


$$Dom(n_i) = \{n_i\} \cup (Dom(n_i) \cap Dom(n_j))$$
$$Dom(1) = \{1\} \cup (\{0,1,2,3,4,5\} \cap \{0\})$$
$$Dom(1) = \{1\} \cup \{0\} = \{0,1\}$$

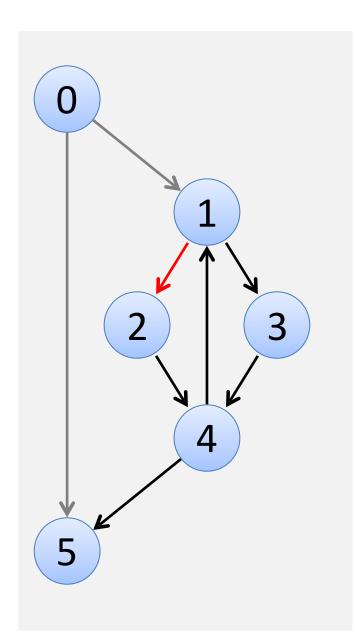
Node	Dominators							
Noue	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	0	0	0	0	1	1		
2	1	1	1	1	1	1		
3	1	1	1	1	1	1		
4	1	1	1	1	1	1		
5	1	1	1	1	1	1		



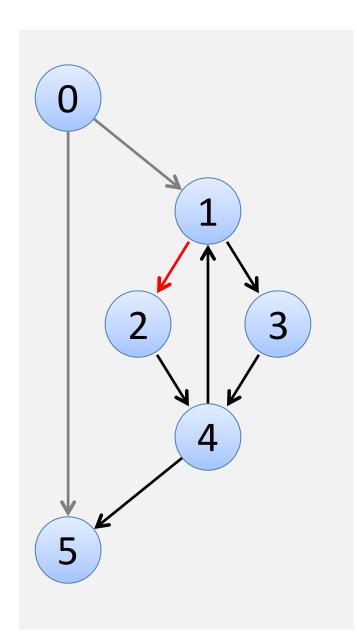
Node	Dominators							
Node	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	0	0	0	0	1	1		
2	1	1	1	1	1	1		
3	1	1	1	1	1	1		
4	1	1	1	1	1	1		
5	1	1	1	1	1	1		



Nodo	Dominators							
Node	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	0	0	0	0	1	1		
2	1	1	1	1	1	1		
3	1	1	1	1	1	1		
4	1	1	1	1	1	1		
5	1	0	0	0	0	1		

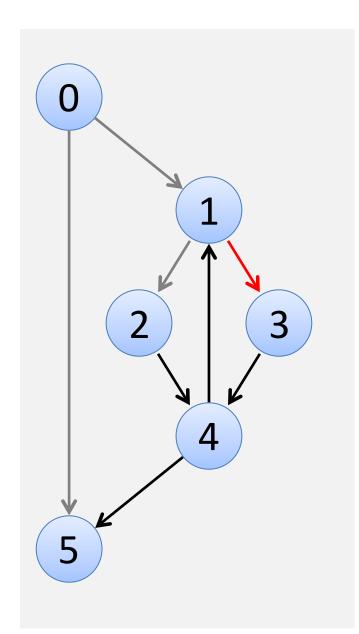


Node	Dominators							
Noue	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	0	0	0	0	1	1		
2	1	1	1	1	1	1		
3	1	1	1	1	1	1		
4	1	1	1	1	1	1		
5	1	0	0	0	0	1		

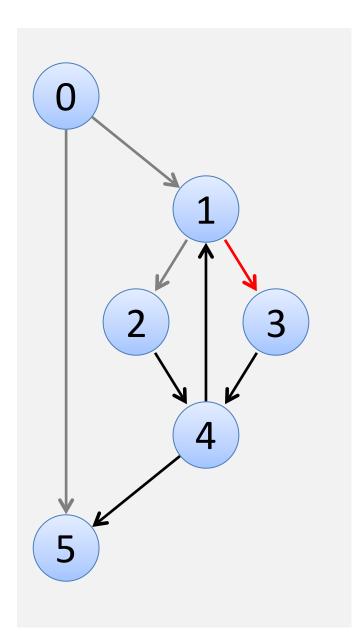


 $\mathrm{Dom}(n_i) = \{n_i\} \cup (\mathrm{Dom}(n_i) \cap \mathrm{Dom}(n_i))$

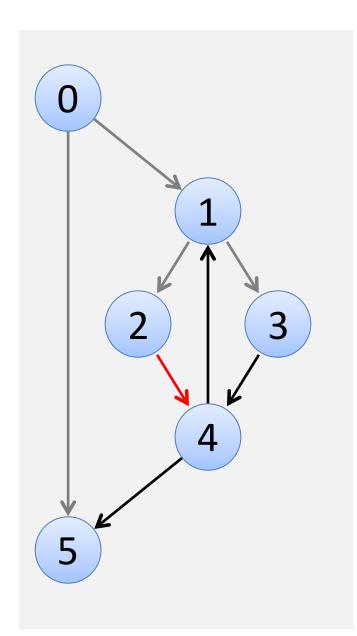
Node	Dominators							
Noue	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	0	0	0	0	1	1		
2	0	0	0	1	1	1		
3	1	1	1	1	1	1		
4	1	1	1	1	1	1		
5	1	0	0	0	0	1		



Node	Dominators							
Noue	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	0	0	0	0	1	1		
2	0	0	0	1	1	1		
3	1	1	1	1	1	1		
4	1	1	1	1	1	1		
5	1	0	0	0	0	1		

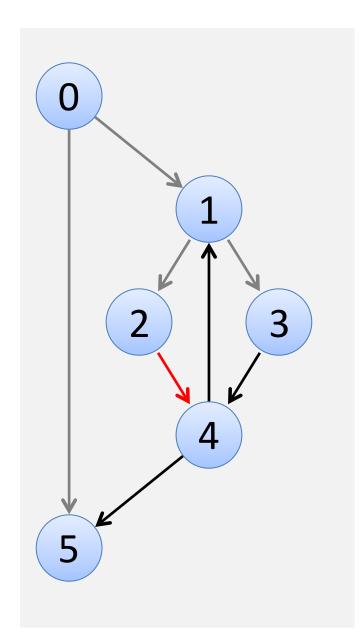


Nodo	Dominators							
Node	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	0	0	0	0	1	1		
2	0	0	0	1	1	1		
3	0	0	1	0	1	1		
4	1	1	1	1	1	1		
5	1	0	0	0	0	1		



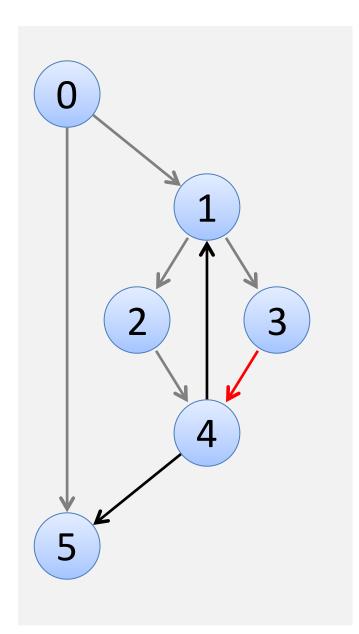
 $\mathrm{Dom}(n_i) = \{n_i\} \cup (\mathrm{Dom}(n_i) \cap \mathrm{Dom}(n_i))$

Nodo	Dominators							
Node	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	0	0	0	0	1	1		
2	0	0	0	1	1	1		
3	0	0	1	0	1	1		
4	1	1	1	1	1	1		
5	1	0	0	0	0	1		

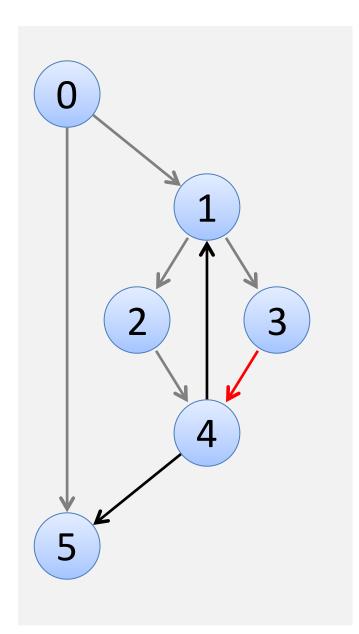


 $\mathrm{Dom}(n_i) = \{n_i\} \cup (\mathrm{Dom}(n_i) \cap \mathrm{Dom}(n_i))$

Nodo	Dominators								
Node	5	4	3	2	1	0			
0	0	0	0	0	0	1			
1	0	0	0	0	1	1			
2	0	0	0	1	1	1			
3	0	0	1	0	1	1			
4	0	1	0	1	1	1			
5	1	0	0	0	0	1			

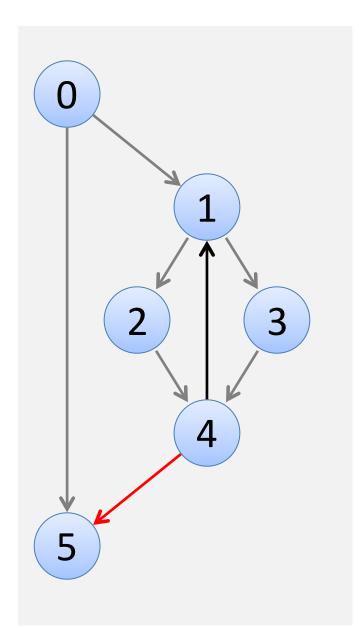


Node	Dominators							
Noue	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	0	0	0	0	1	1		
2	0	0	0	1	1	1		
3	0	0	1	0	1	1		
4	0	1	0	1	1	1		
5	1	0	0	0	0	1		

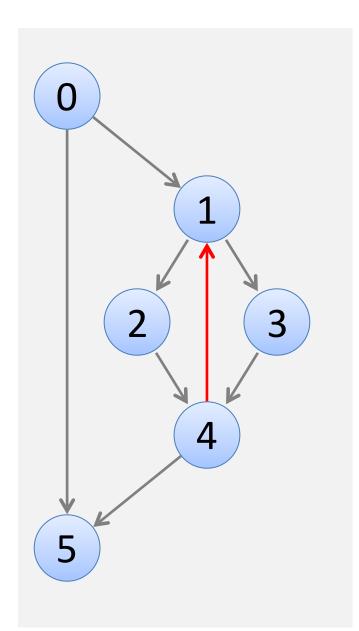


 $\mathrm{Dom}(n_i) = \{n_i\} \cup (\mathrm{Dom}(n_i) \cap \mathrm{Dom}(n_i))$

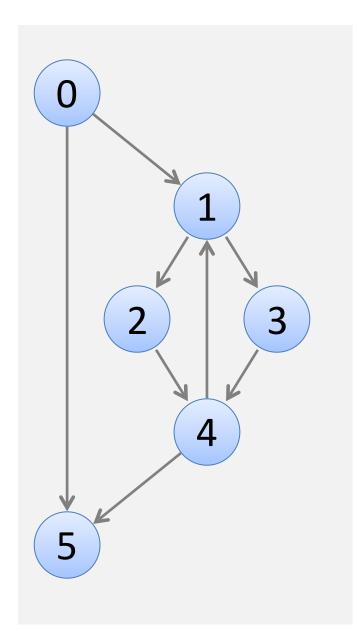
Node	Dominators							
Noue	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	0	0	0	0	1	1		
2	0	0	0	1	1	1		
3	0	0	1	0	1	1		
4	0	1	0	0	1	1		
5	1	0	0	0	0	1		



Node	Dominators							
Node	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	0	0	0	0	1	1		
2	0	0	0	1	1	1		
3	0	0	1	0	1	1		
4	0	1	0	0	1	1		
5	1	0	0	0	0	1		



Node	Dominators							
	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	0	0	0	0	1	1		
2	0	0	0	1	1	1		
3	0	0	1	0	1	1		
4	0	1	0	0	1	1		
5	1	0	0	0	0	1		



Node	Dominators							
	5	4	3	2	1	0		
0	0	0	0	0	0	1		
1	0	0	0	0	1	1		
2	0	0	0	1	1	1		
3	0	0	1	0	1	1		
4	0	1	0	0	1	1		
5	1	0	0	0	0	1		

Bit Vectors

Use as many words as necessary to store the bit vector in memory.

Examples:

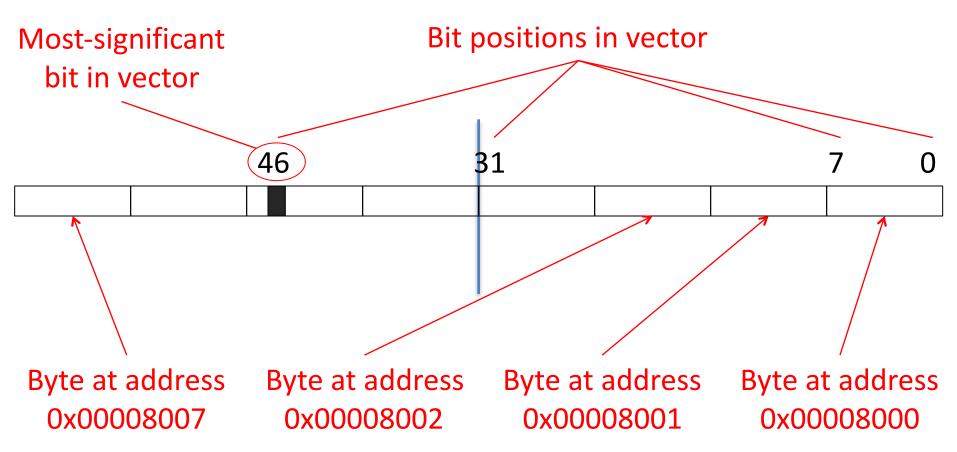
A 17-bit vector occupies one word.

A 42-bit vector occupies two words.

A 65-bit vector occupies three words.

Bit Ordering

Consider a 47-bit vector stored at address 0x00008000.



The Assignment (input)

The MIPS code is guaranteed to contain a single procedure.

This is the sentinel indicating the end of the procedure.

\$a0 This is the binary code for the odd series example in this presentation. \$a0 contains a memory address At that address is the binary representation of the first

instruction

Memory

0xffffffff 0x03e00008

0x1420fff9

0x0104082a

0x21080001

0x20420001

0x08100029

0x00481020

0x15400003

0x310a0001

0x18800009

0x34020000

0x34080000

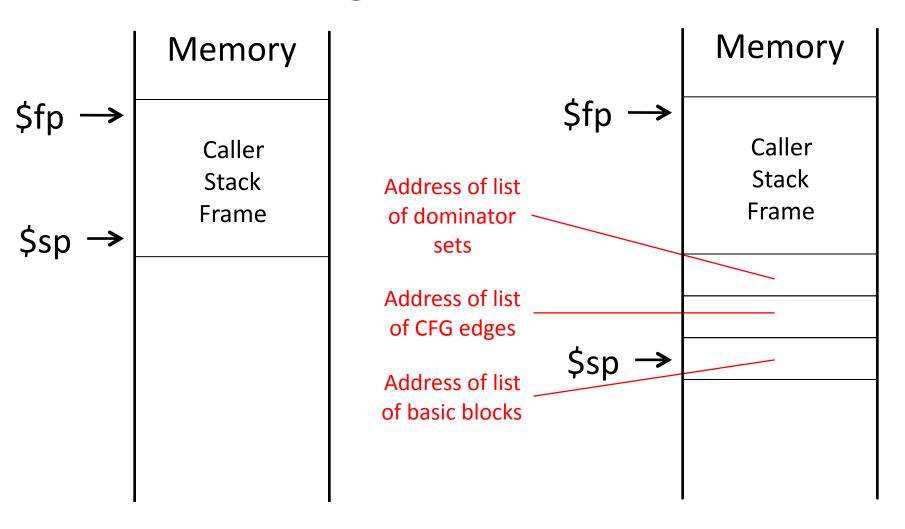
The Assignment (output)

\$v0: number of basic blocks in the procedure.

\$v1: number of edges in the CFG of the procedure.

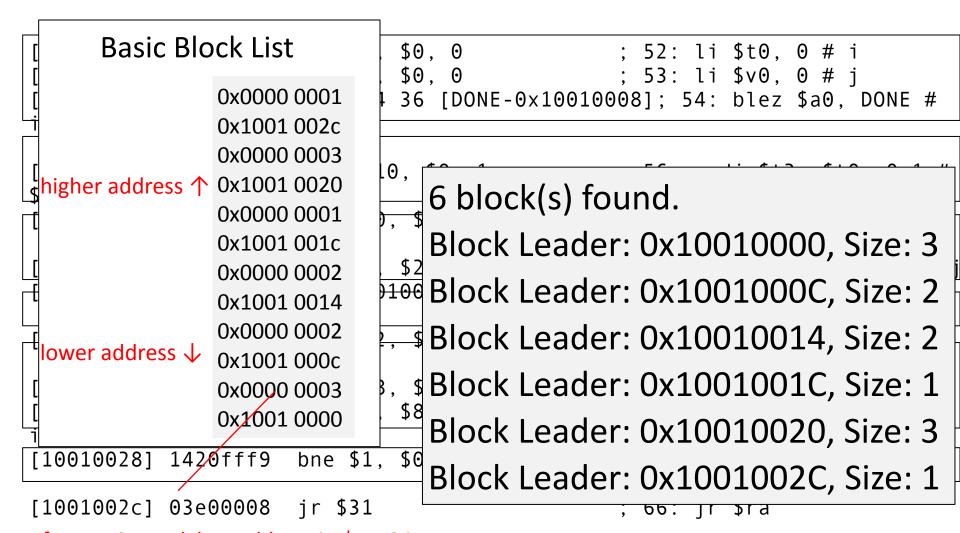
Three additional memory addresses returned into the stack.

Returning Addresses in Stack

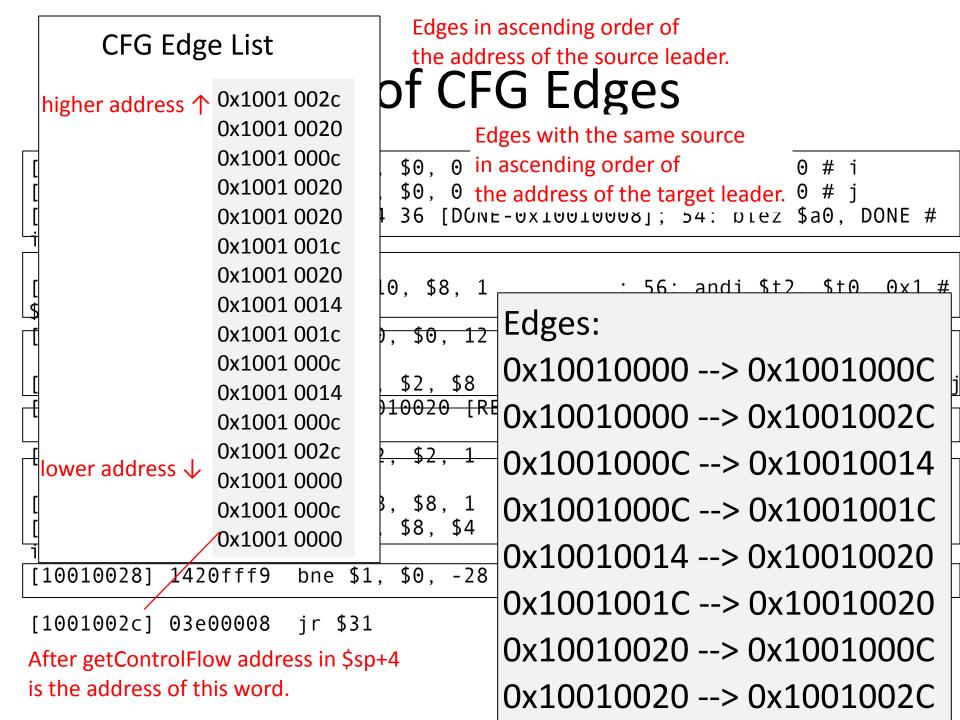


Basic blocks in ascending order of the address of their leaders.

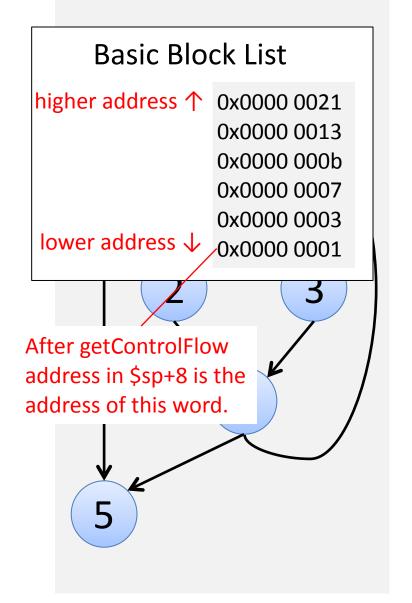
List of Basic Blocks



After getControlFlow address in \$sp+0 is the address of this word.



List of Dominator Sets



Dominator Bit Vectors:

If the CFG has more than 32 basic blocks, then each binary vector will occupy more than one word.

Dominator sets in ascending order of the address of the corresponding basic-block's leader.

Testing

- Test Cases
 - A few test cases are provided under Resources
- Student-Generated Test Cases
 - TestCaseGeneration
- Printing the Output of your solution
 - MIPS code provided for printing

University of Alberta Code of Student Behavior

http://www.governance.ualberta.ca/en/CodesofConductandResidenceCommunityStandards/CodeofStudentBehaviour.aspx

30.3.2(2) Cheating

30.3.2(2) a No Student shall in the course of an examination or other similar activity, obtain or attempt to obtain information from another Student or other unauthorized source, give or attempt to give information to another Student, or use, attempt to use or possess for the purposes of use any unauthorized material.