EECS 583 – Class 3

Region Formation, Predicated Execution

University of Michigan

September 8, 2023

Announcements & Reading Material

- HW0 was due Monday – Remember nothing to turn in
- HW1 is out – Due Monday Sep 18
  - http://web.eecs.umich.edu/~mahlke/courses/583f23/homeworks
- Today's class
  - "Trace Selection for Compiling Large C Applications to Microcode", Chang and Hwu, MICRO-21, 1988.
  - "The Superblock: An Effective Technique for VLIW and Superscalar Compilation", Hwu et al., Journal of Supercomputing, 1993
- Material for Monday
  - "The Program Dependence Graph and Its Use in Optimization", J. Ferrante, K. Ottenstein, and J. Warren, ACM TOPLAS, 1987
    - This is a long paper – the part we care about is the control dependence stuff. The PDG is interesting and you should skim it
    - "On Predicated Execution", Park and Schlansker, HPL Technical Report, 1991.

Homework 1 – Due Mon Sep 18

- Get started ASAP. If you haven't done HW0, you are falling behind!
- Goals: Learn how to profile with LLVM, write stats collection pass
- 583_F23_HW1.tgz
  - hw1pass.cpp: template for your pass
  - 583simple, 583anagram, 583compress: benchmark source code + inputs + expected outputs + run instructions
- Easy to do, but hard to start because of newness
  - Look for Aditya's piazza post for help
    - Skeleton code
    - How to run profiler
    - Simple example with opcode stats
  - Talk to the GSI if you are stuck

Regions

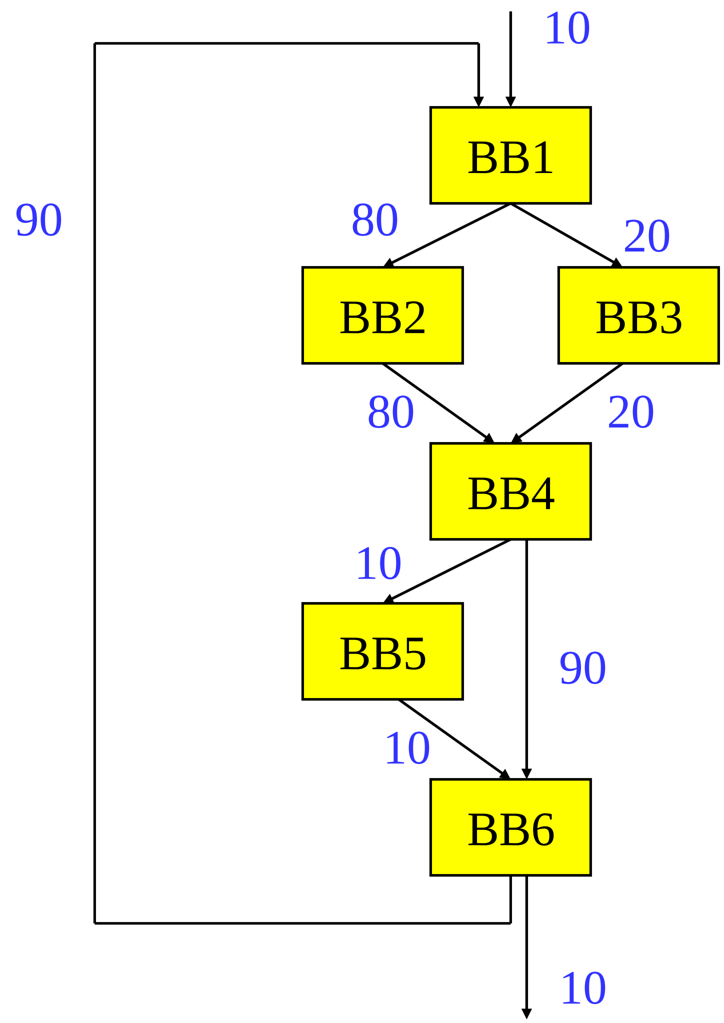
- Region: A collection of operations that are treated as a single unit by the compiler
  - Examples
    - Basic block
    - Procedure
    - Body of a loop
  - Properties
    - Connected subgraph of operations
    - Control flow is the key parameter that defines regions
    - Hierarchically organized
- Problem
  - Basic blocks are too small (3-5 operations)
    - Hard to extract sufficient parallelism
  - Procedure control flow too complex for many compiler xforms
    - Plus only parts of a procedure are important (90/10 rule)

Regions (2)

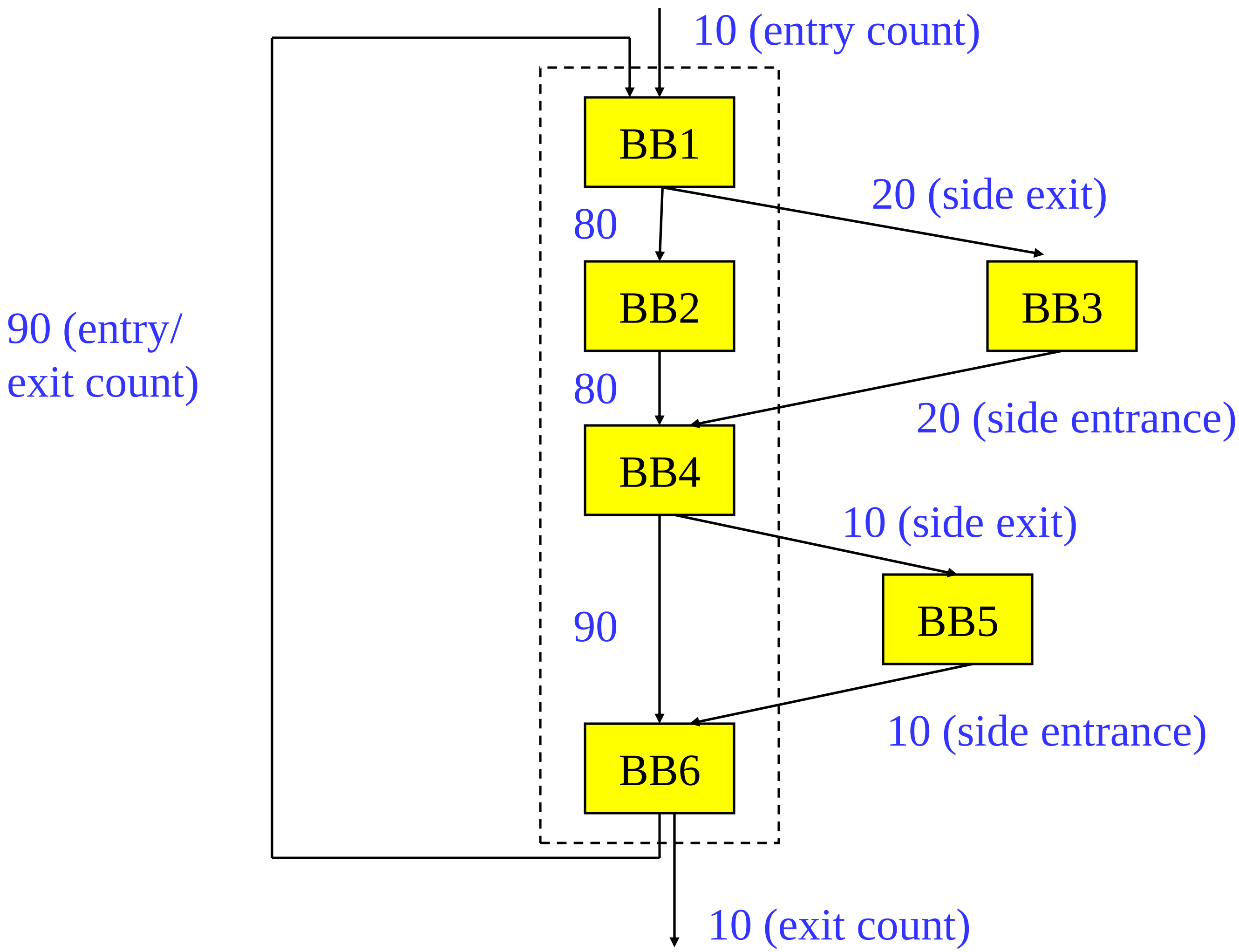
- Want
  - Intermediate sized regions with simple control flow
  - Bigger basic blocks would be ideal !!
  - Separate important code from less important
  - Optimize frequently executed code at the expense of the rest
- Solution
  - Define new region types that consist of multiple BBs
  - Profile information used in the identification
  - Sequential control flow (sorta)
  - Pretend the regions are basic blocks

Region Type 1 - Trace

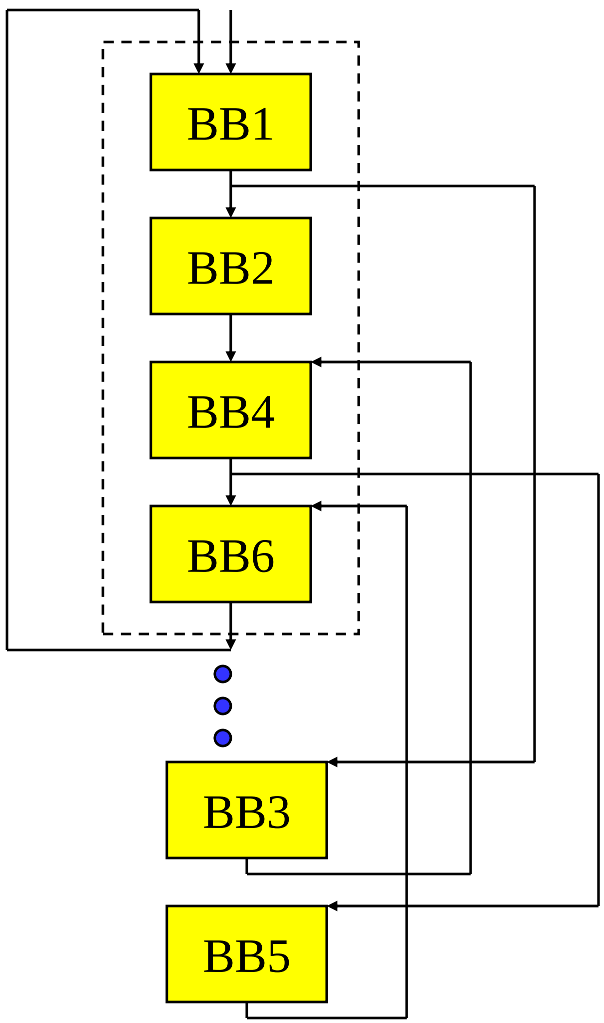
- <u>Trace</u> - Linear collection of basic blocks that tend to execute in sequence
  - "Likely control flow path"
  - Acyclic (outer backedge ok)
- <u>Side entrance</u> – branch into the middle of a trace
- <u>Side exit</u> – branch out of the middle of a trace
- Compilation strategy
  - Compile assuming path occurs 100% of the time
  - Patch up side entrances and exits afterwards
- Motivated by scheduling (i.e., trace scheduling)



Linearizing a Trace

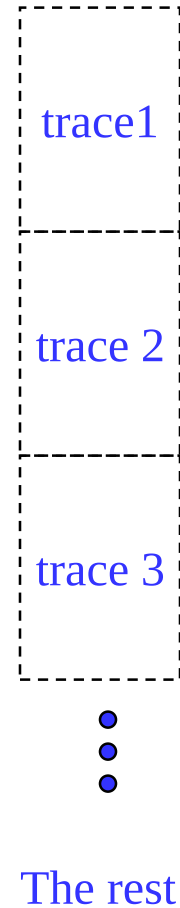


Intelligent Trace Layout for Icache Performance



Trace view

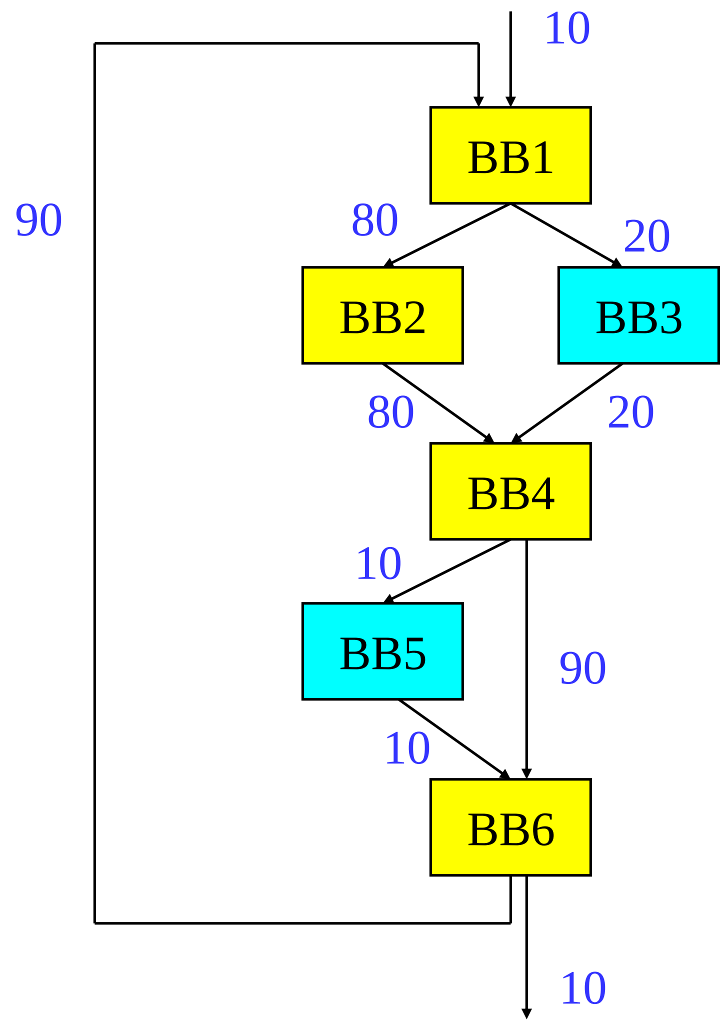
Intraprocedural code placement
Procedure positioning
Procedure splitting



Procedure view

Issues With Selecting Traces

- Acyclic
  - Cannot go past a backedge
- Trace length
  - Longer = better ?
  - Not always !
- On-trace / off-trace transitions
  - Maximize on-trace
  - Minimize off-trace
  - Compile assuming on-trace is 100% (ie single BB)
  - Penalty for off-trace
- Tradeoff (heuristic)
  - Length
  - Likelihood remain within the trace



Trace Selection Algorithm

```
i = 0;
mark all BBs unvisited
while (there are unvisited nodes) do
    seed = unvisited BB with largest execution freq
    trace[i] += seed
    mark seed visited
    current = seed
    /* Grow trace forward */
    while (1) do
        next = best_successor_of(current)
        if (next == 0) then break
        trace[i] += next
        mark next visited
        current = next
    endwhile
    /* Grow trace backward analogously */
    i++
endwhile
```

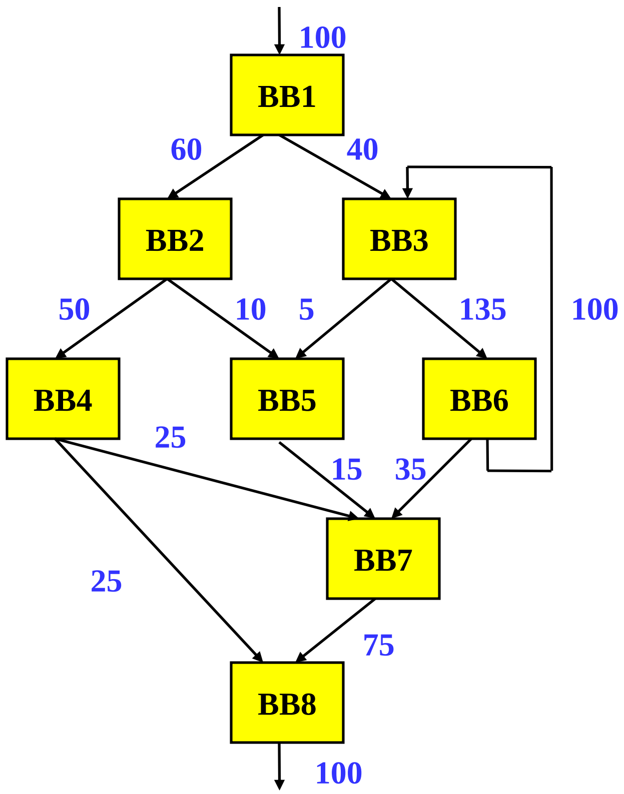
Best Successor/Predecessor

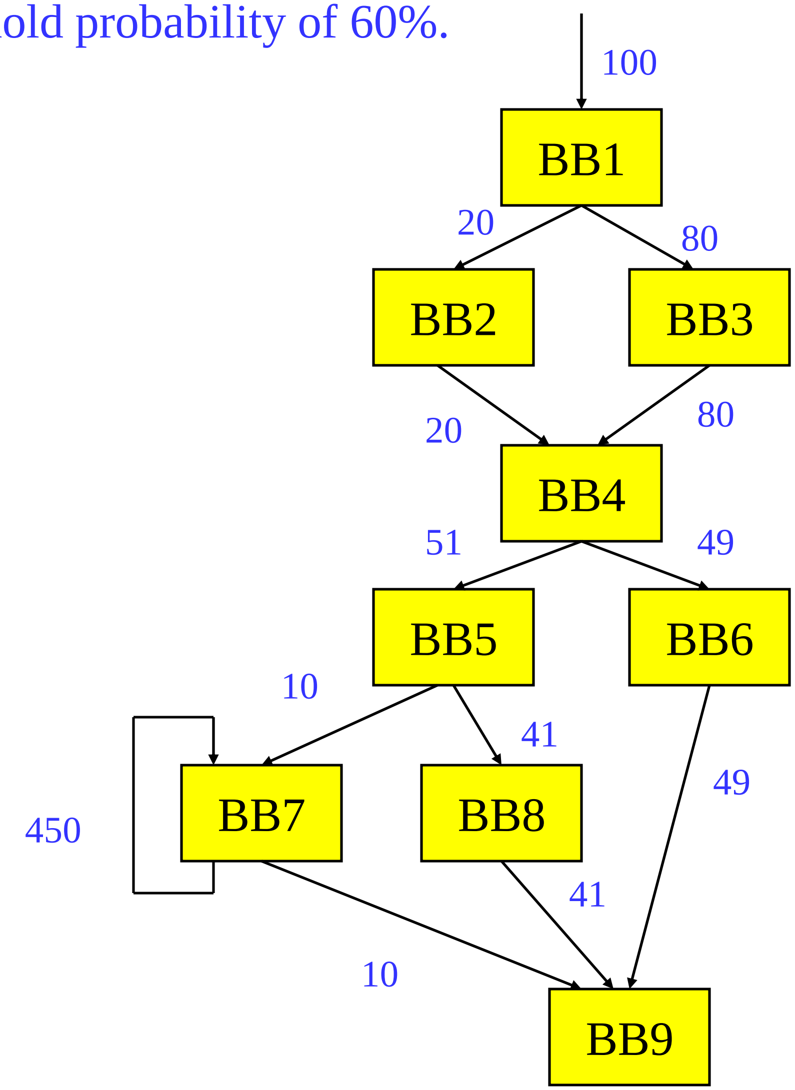
- Node weight vs edge weight
  - edge more accurate
- THRESHOLD
  - controls off-trace probability
  - 60-70% found best
- Notes on this algorithm
  - BB only allowed in 1 trace
  - Cumulative probability ignored
  - Min weight for seed to be chose (ie executed 100 times)

```
best_successor_of(BB)
   e = control flow edge with highest
        probability leaving BB
   if (e is a backedge) then
      return 0
   endif
   if (probability(e) <= THRESHOLD) then
      return 0
   endif
   d = destination of e
   if (d is visited) then
      return 0
   endif
   return d
end procedure
```

Example Problems

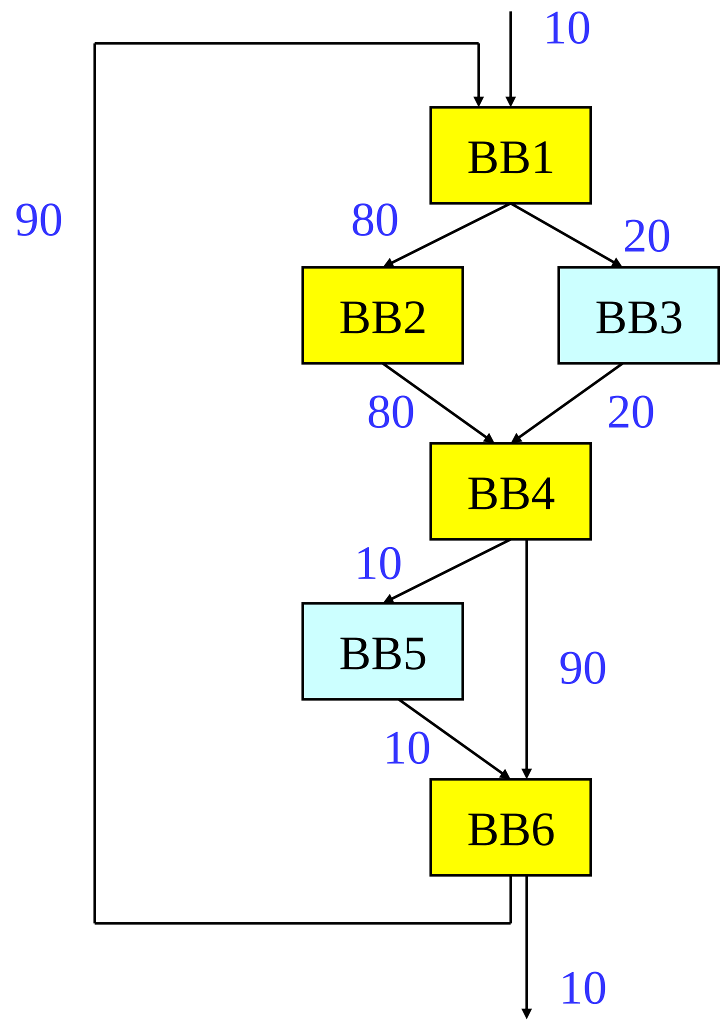
Find the traces. Assume a threshold probability of 60%.





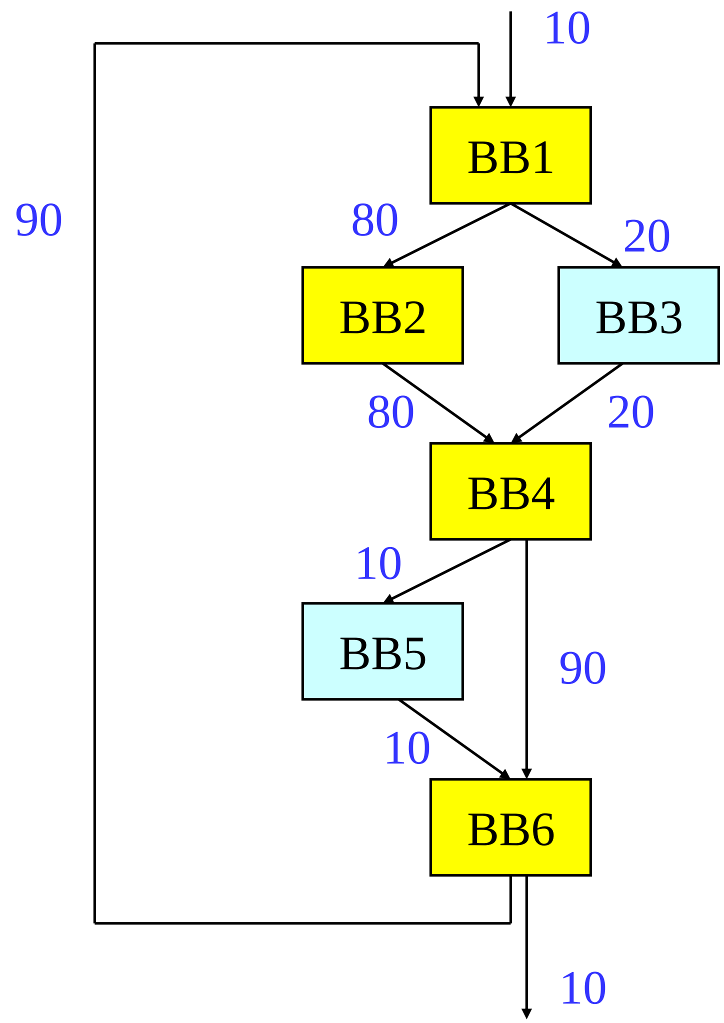
Traces are Nice, But …

- Treat trace as a big BB
  - Transform trace ignoring side entrance/exits
  - Insert fixup code
    - aka bookkeeping
  - Side entrance fixup is more painful
  - Sometimes not possible so transform not allowed
- Solution
  - Eliminate side entrances
  - The superblock is born



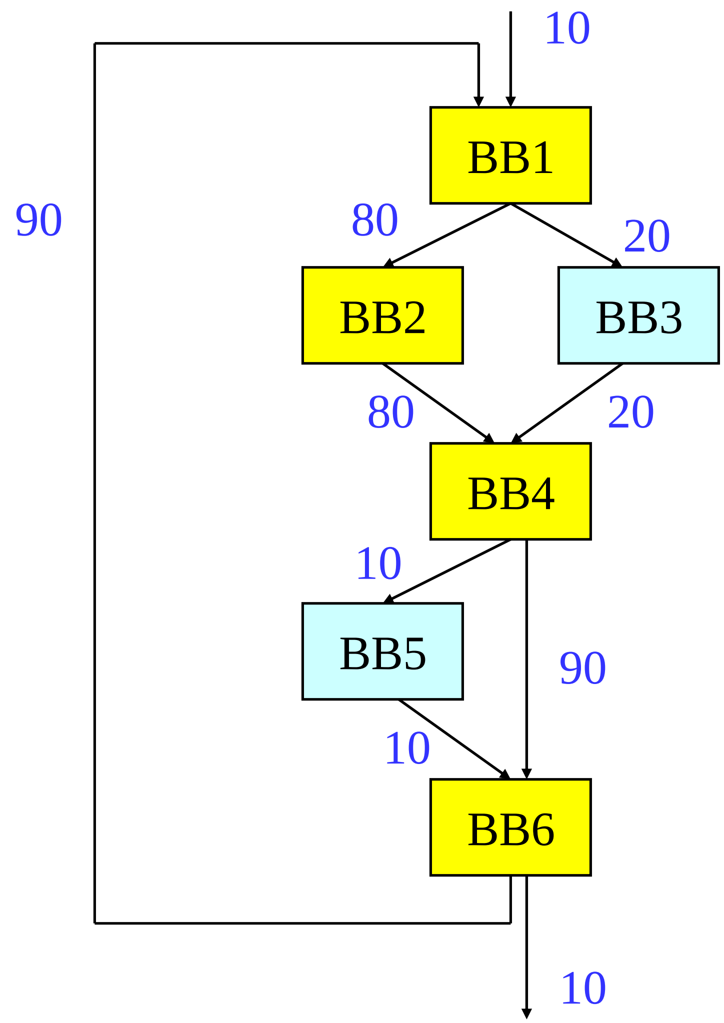
Region Type 2 - Superblock

- Superblock - Linear collection of basic blocks that tend to execute in sequence *in which control flow may only enter at the first BB*
  - "Likely control flow path"
  - Acyclic (outer backedge ok)
  - Trace with no side entrances
  - Side exits still exist
- Superblock formation
  - 1. Trace selection
  - 2. Eliminate side entrances

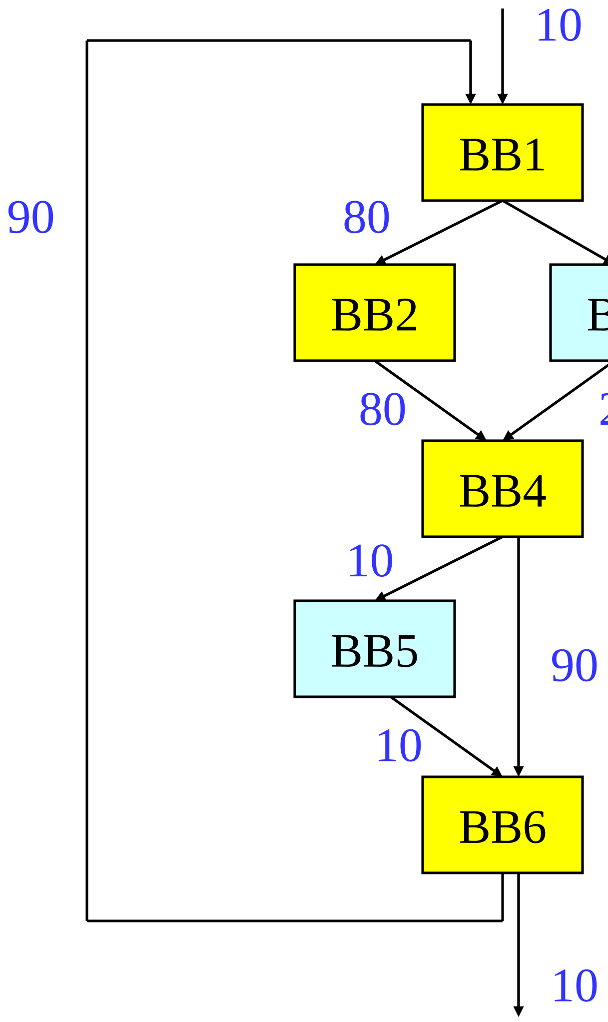


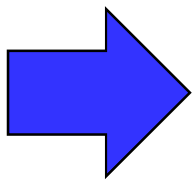
Tail Duplication

- To eliminate all side entrances replicate the "tail" portion of the trace
  - Identify first side entrance
  - Replicate all BB from the target to the bottom
  - Redirect all side entrances to the duplicated BBs
  - Copy each BB only once
  - Max code expansion = 2x-1 where x is the number of BB in the trace
  - Adjust profile information

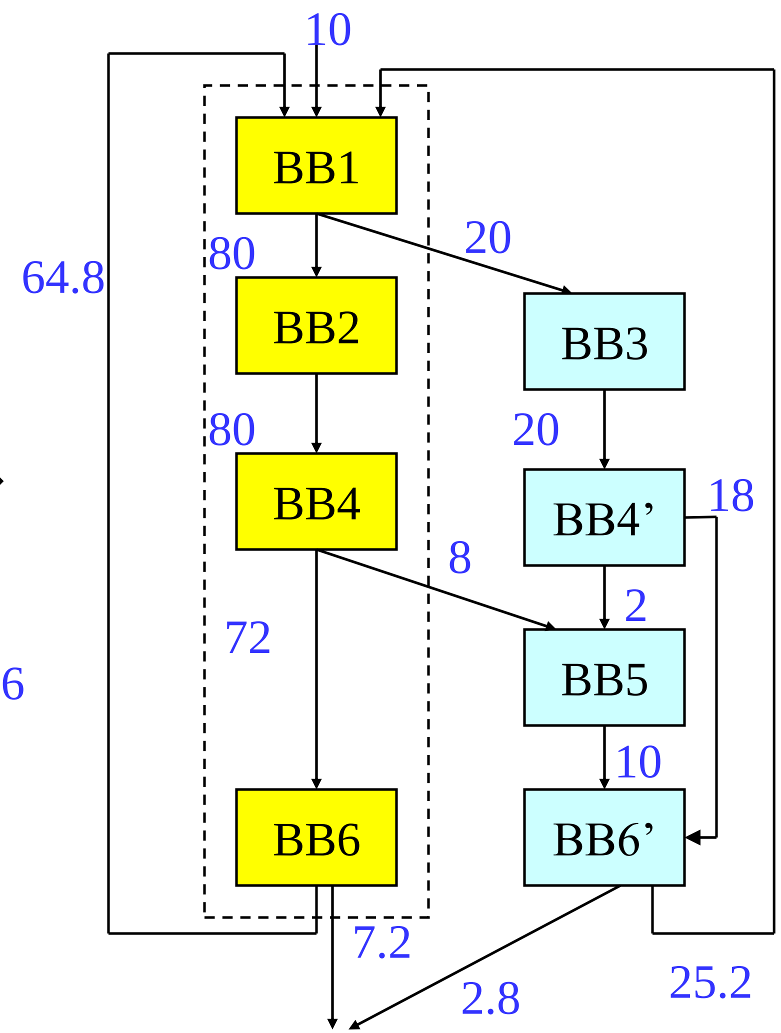


Superblock Formation



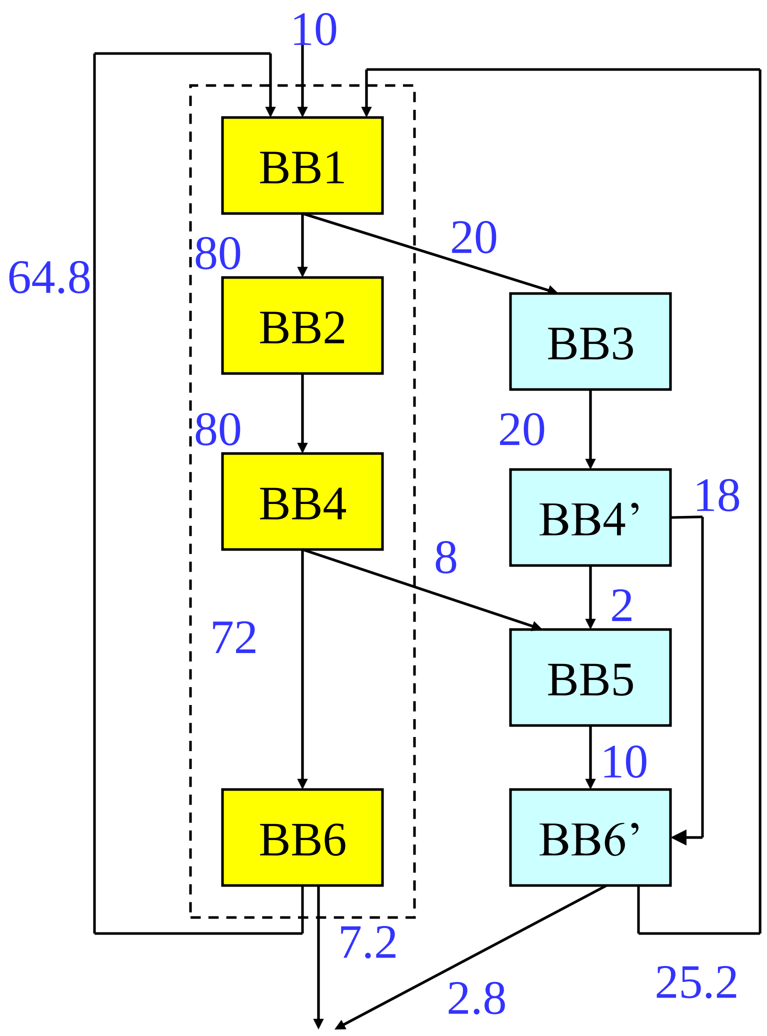


Duplicate
BB4 and BB6
and adjust
control flow



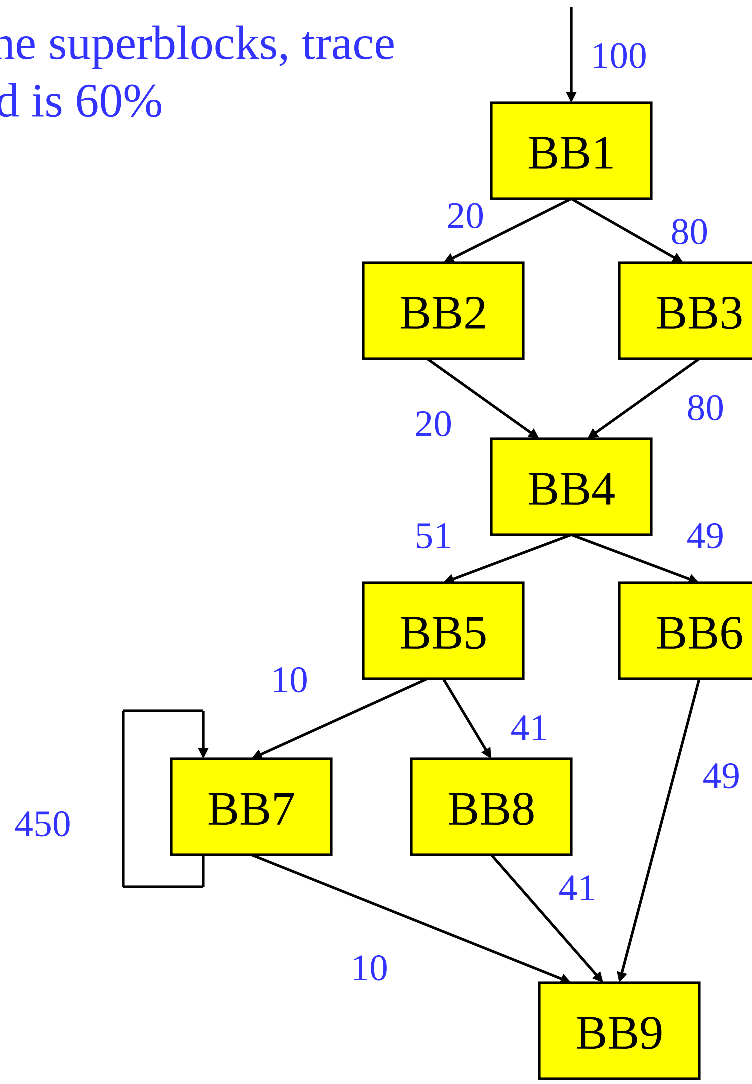
Issues with Superblocks

- Central tradeoff
  - Side entrance elimination
    - Compiler complexity
    - Compiler effectiveness
  - Code size increase
- Apply intelligently
  - Most frequently executed BBs are converted to SBs
  - Set upper limit on code expansion
  - 1.0 – 1.10x are typical code expansion ratios from SB formation



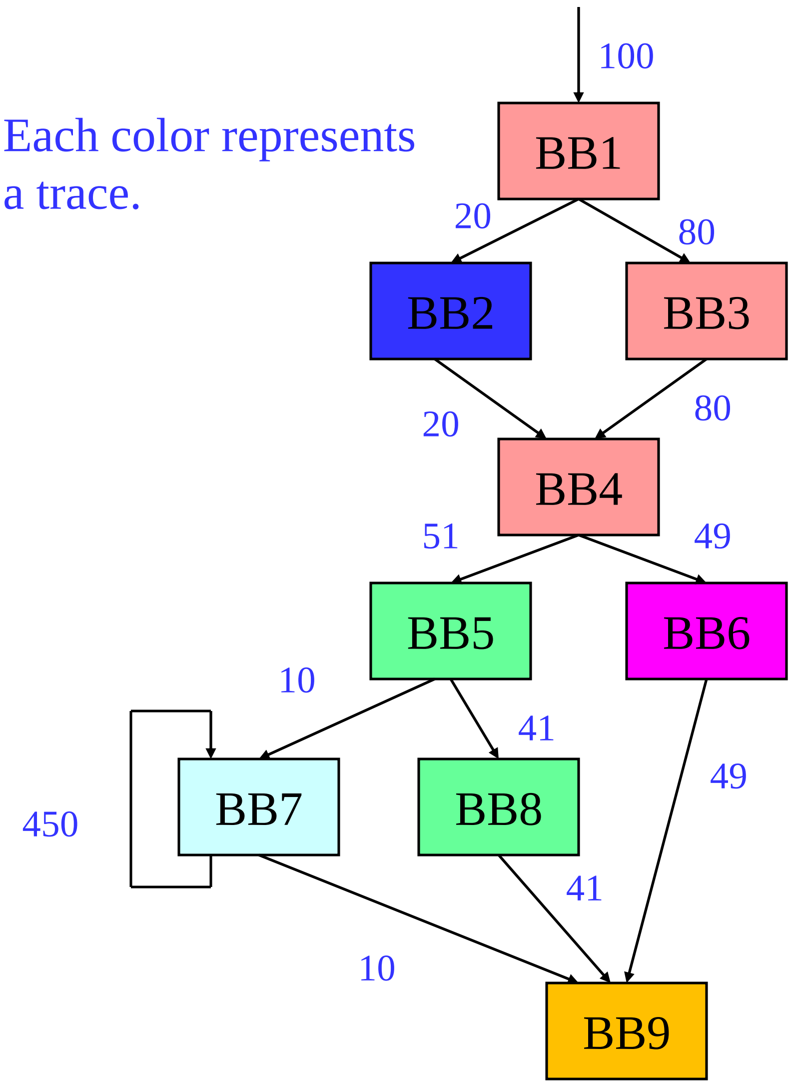
Class Problem

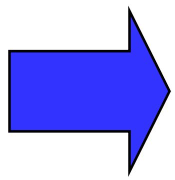
Create the superblocks, trace threshold is 60%



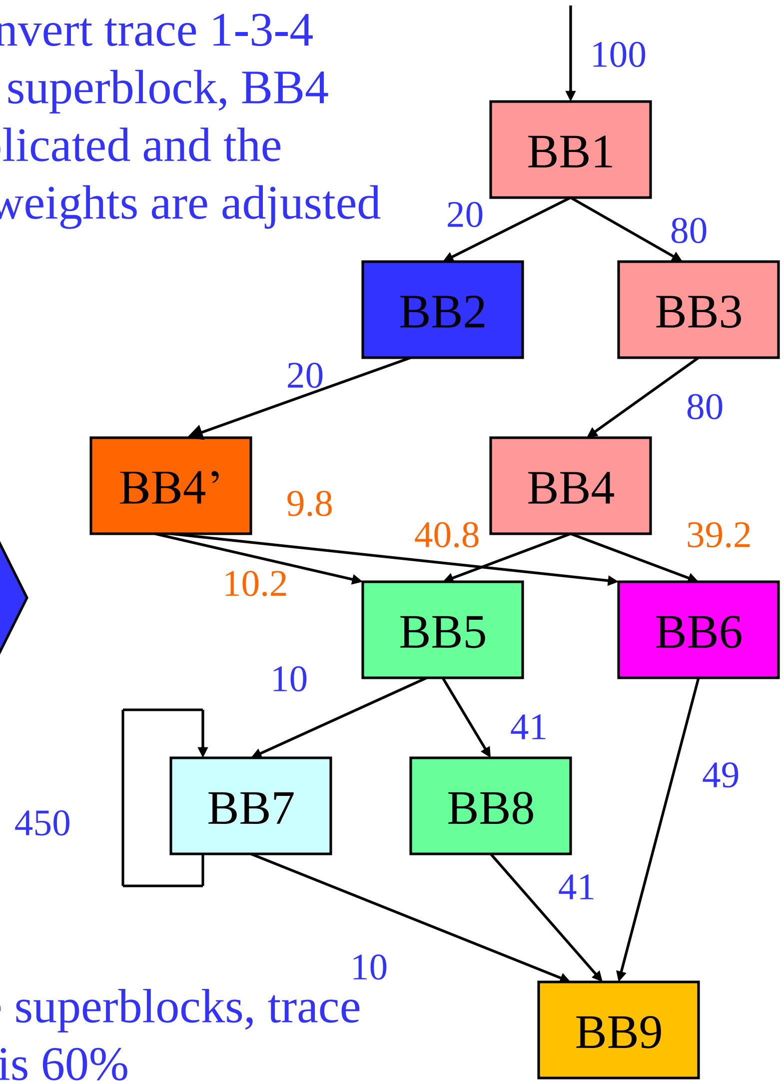
Class Problem Solution – Superblock Formation

Each color represents a trace.





To convert trace 1-3-4 into a superblock, BB4 is duplicated and the edge weights are adjusted



Create the superblocks, trace threshold is 60%

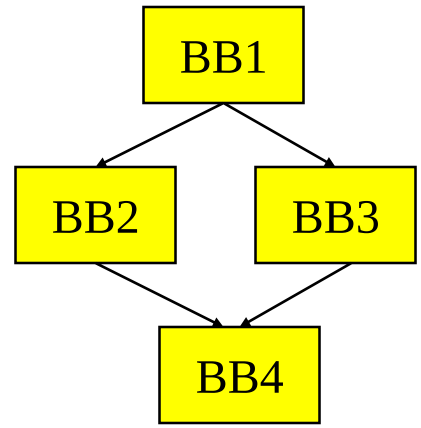
An Alternative to Branches: Predicated Execution

- Hardware mechanism that allows operations to be conditionally executed
- Add an additional boolean source operand (predicate)
  - » ADD r1, r2, r3 if p1
    - if (p1 is True), r1 = r2 + r3
    - else if (p1 is False), do nothing (Add treated like a NOP)
    - p1 referred to as the guarding predicate
    - Predicated on True means always executed
    - Omitted predicated also means always executed
- Provides compiler with an alternative to using branches to selectively execute operations
  - » If statements in the source
  - » Realize with branches in the assembly code
  - » Could also realize with conditional instructions
  - » Or use a combination of both

Predicated Execution Example

```
a = b + c
if (a > 0)
    e = f + g
else
    e = f / g
h = i - j
```

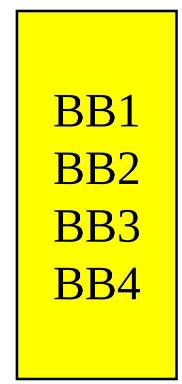
```
BB1    add a, b, c
BB1    bgt a, 0, L1
BB3    div e, f, g
BB3    jump L2
BB2    L1: add e, f, g
BB4    L2: sub h, i, j
```



Traditional branching code

p2 → BB2
p3 → BB3

```
BB1    add a, b, c if T
BB1    p2 = a > 0 if T
BB1    p3 = a <= 0 if T
BB3    div e, f, g if p3
BB2    add e, f, g if p2
BB4    sub h, i, j if T
```

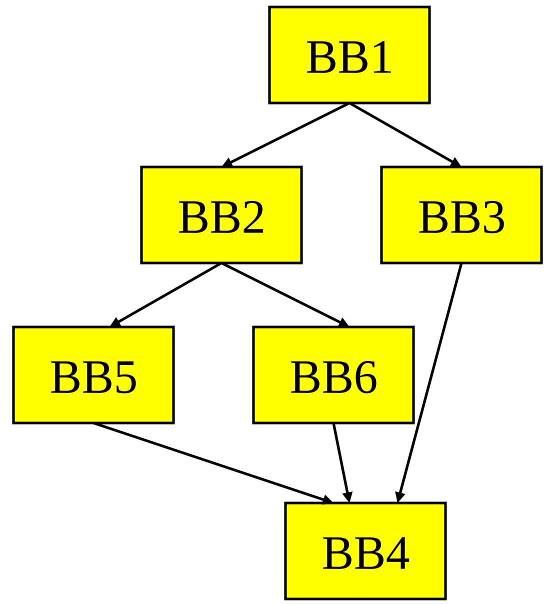


Predicated code

What About Nested If-then-else's?

```
a = b + c
if (a > 0)
    if (a > 25)
        e = f + g
    else
        e = f * g
else
    e = f / g
h = i - j
```

```
BB1    add a, b, c
BB1    bgt a, 0, L1
BB3    div e, f, g
BB3    jump L2
BB2    L1: bgt a, 25, L3
BB6    mpy e, f, g
BB6    jump L2
BB5    L3: add e, f, g
BB4    L2: sub h, i, j
```



Traditional branching code

Nested If-then-else's – No Problem

```
a = b + c
if (a > 0)
    if (a > 25)
        e = f + g
    else
        e = f * g
else
    e = f / g
h = i - j
```

BB1	add a, b, c if T
BB1	p2 = a > 0 if T
BB1	p3 = a <= 0 if T
BB3	div e, f, g if p3
BB3	p5 = a > 25 if p2
BB3	p6 = a <= 25 if p2
BB6	mpy e, f, g if p6
BB5	add e, f, g if p5
BB4	sub h, i, j if T

BB1
BB2
BB3
BB4
BB5
BB6

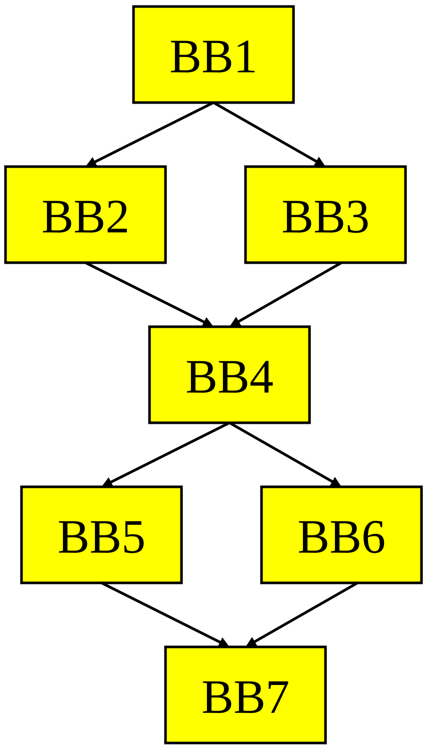
Predicated code

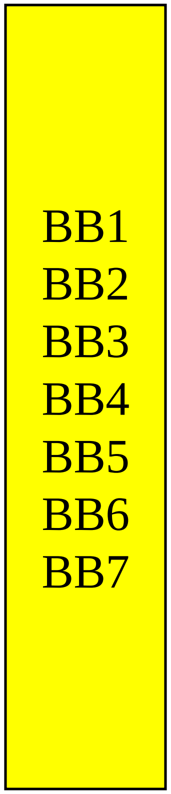
What do we assume to make this work ??

if p2 is False, both p5 and p6 are False

So, predicate setting instruction should set result to False if guarding predicate is false!!!

Benefits/Costs of Predicated Execution





Benefits:
- No branches, no mispredicts
- Can freely reorder independent operations in the predicated block
- Overlap BB2 with BB5 and BB6

Costs (execute all paths)
-worst case schedule length
-worst case resources required

HPL-PD Compare-to-Predicate Operations (CMPPs)

- How do we compute predicates
  - » Compare registers/literals like a branch would do
  - » Efficiency, code size, nested conditionals, etc
- 2 targets for computing taken/fall-through conditions with 1 operation

p1, p2 = CMPP.cond.D1a.D2a (r1, r2) if p3

p1 = first destination predicate
p2 = second destination predicate
cond = compare condition (ie EQ, LT, GE, …)
D1a = action specifier for first destination
D2a = action specifier for second destination
(r1,r2) = data inputs to be compared (ie r1 < r2)
p3 = guarding predicate

CMPP Action Specifiers

Guarding predicate	Compare Result	UN	UC	ON	OC	AN	AC
0	0	0	0	-	-	-	-
0	1	0	0	-	-	-	-
1	0	0	1	-	1	0	-
1	1	1	0	1	-	-	0

UN/UC = Unconditional normal/complement

This is what we used in the earlier examples

guard = 0, both outputs are 0

guard = 1, UN = Compare result, UC = opposite

ON/OC = OR-type normal/complement

AN/AC = AND-type normal/complement

OR-type, AND-type Predicates

```
p1 = 0
p1 = cmpp_ON (r1 < r2) if T
p1 = cmpp_OC (r3 < r4) if T
p1 = cmpp_ON (r5 < r6) if T
```

```
p1 = (r1 < r2) | (!(r3 < r4)) |
      (r5 < r6)
```

Wired-OR into p1

Generating predicated code
for some source code requires
OR-type predicates

```
p1 = 1
p1 = cmpp_AN (r1 < r2) if T
p1 = cmpp_AC (r3 < r4) if T
p1 = cmpp_AN (r5 < r6) if T
```

```
p1 = (r1 < r2) & (!(r3 < r4)) &
      (r5 < r6)
```

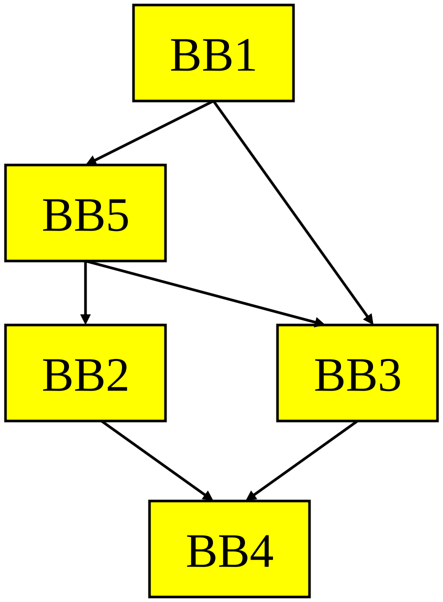
Wired-AND into p1

Talk about these later – used
for control height reduction

Use of OR-type Predicates

```
a = b + c
if (a > 0 && b > 0)
    e = f + g
else
    e = f / g
h = i - j
```

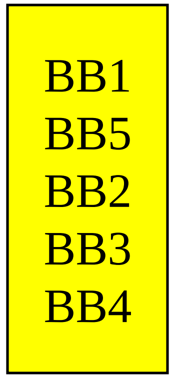
```
BB1    add a, b, c
BB1    ble a, 0, L1
BB5    ble b, 0, L1
BB2    add e, f, g
BB2    jump L2
BB3    L1: div e, f, g
BB4    L2: sub h, i, j
```



Traditional branching code

p2 → BB2
p3 → BB3
p5 → BB5

```
BB1    add a, b, c if T
BB1    p3, p5 = cmpp.ON.UC a <= 0 if T
BB5    p3, p2 = cmpp.ON.UC b <= 0 if p5
BB3    div e, f, g if p3
BB2    add e, f, g if p2
BB4    sub h, i, j if T
```



Predicated code

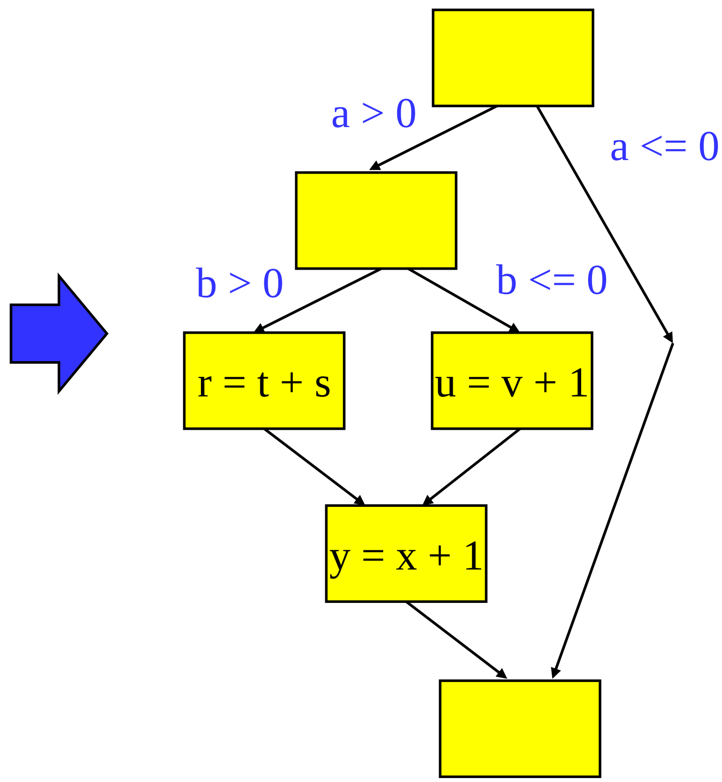
Homework Problem – Answer on next slide but don't cheat!

```
if (a > 0) {
    if (b > 0)
        r = t + s
    else
        u = v + 1
    y = x + 1
}
```

a. Draw the CFG
b. Predicate the code removing all branches

Homework Problem Answer

```
if (a > 0) {
    if (b > 0)
        r = t + s
    else
        u = v + 1
    y = x + 1
}
```



p1 = cmpp.UN(a > 0) if T
p2, p3 = cmpp.UNUC(b > 0) if p1
r = t + s if p2
u = v + 1 if p3
y = x + 1 if p1

a. Draw the CFG
b. Predicate the code removing all branches