

EECS 583 – Class 3

Region Formation, Predicated Execution

University of Michigan

September 8, 2021

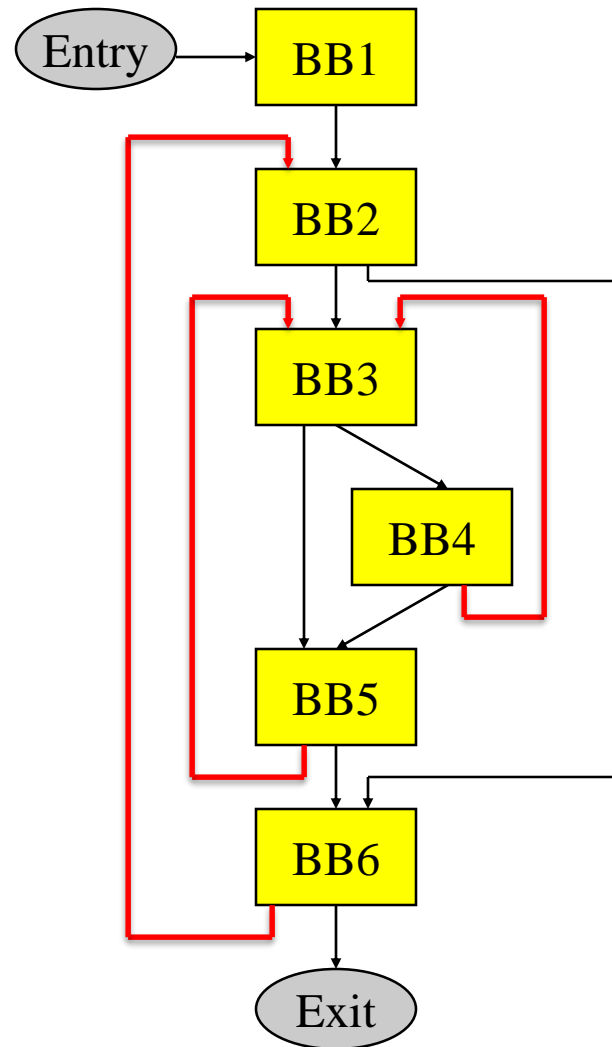
Announcements & Reading Material

- ❖ HW0 due today – Remember nothing to turn in
- ❖ HW1 is out – Due Wed Sep 15
 - » <http://web.eecs.umich.edu/~mahlke/courses/583f21/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.

From Last Time: Natural Loops

- ❖ Cycle suitable for optimization
 - » Discuss optimizations later
- ❖ 2 properties
 - » Single entry point called the header
 - Header dominates all blocks in the loop
 - » Must be one way to iterate the loop (ie at least 1 path back to the header from within the loop) called a backedge
- ❖ Backedge detection
 - » Edge, $x \rightarrow y$ where the target (y) dominates the source (x)

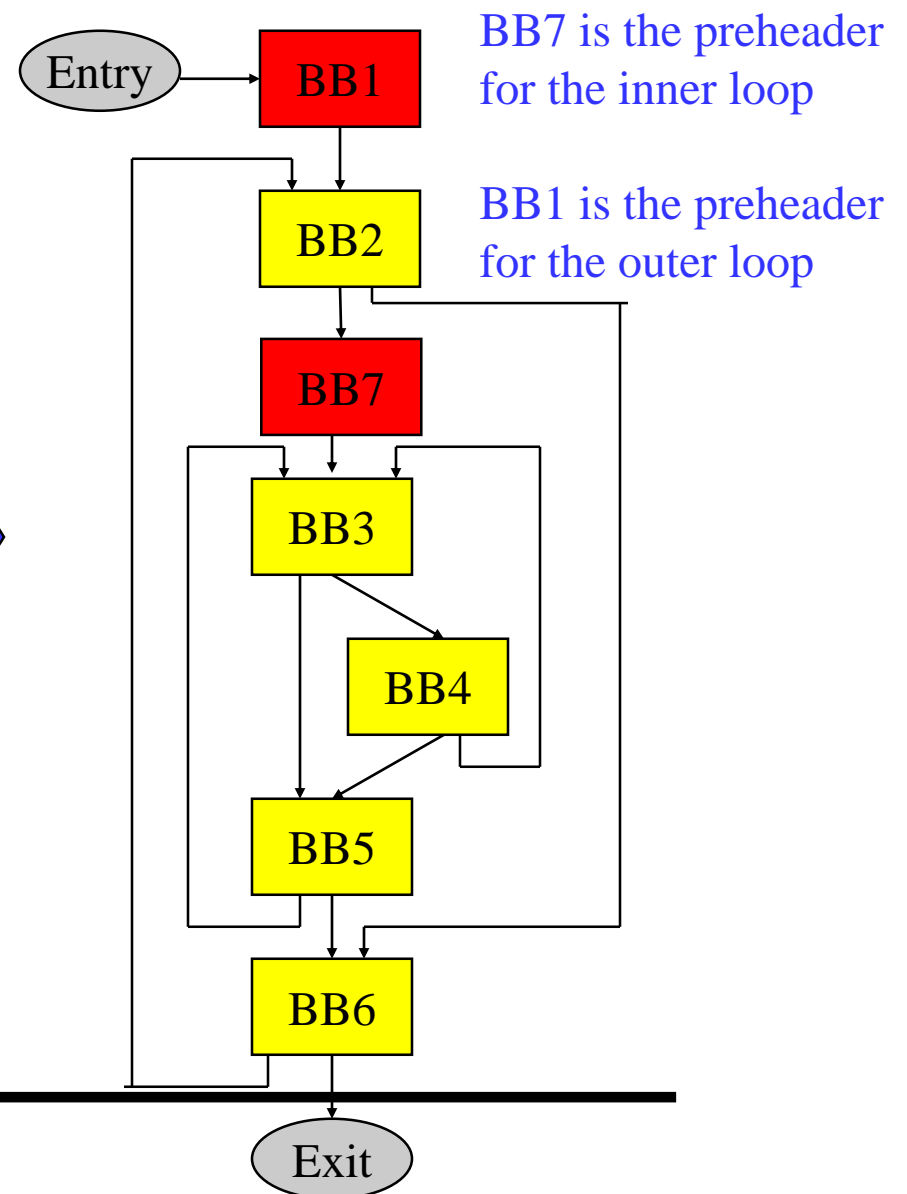
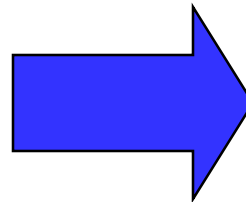
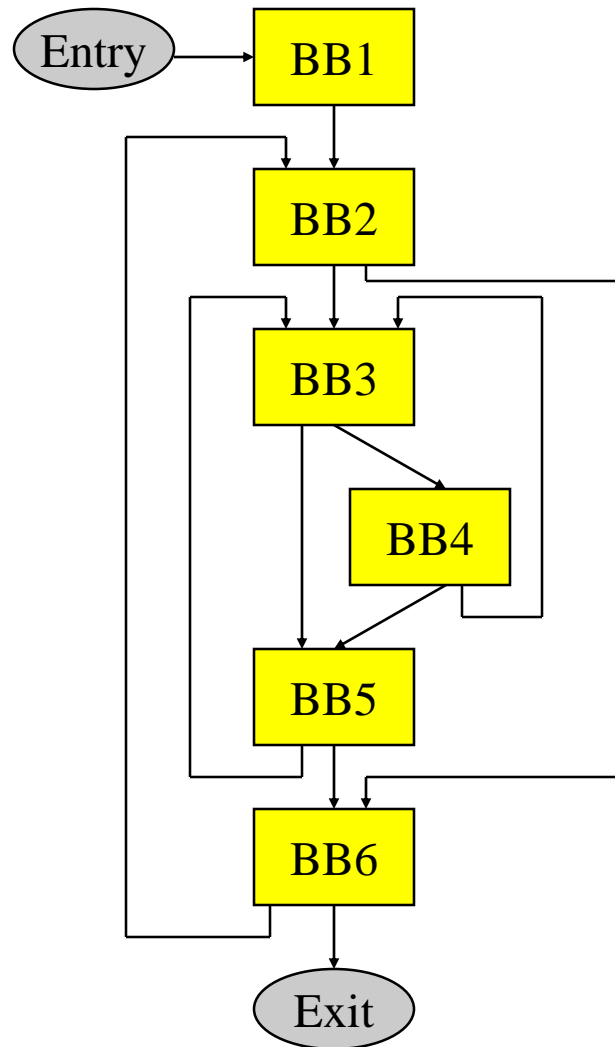
From Last Time: Backedge Example



Important Parts of a Loop

- ❖ Header, LoopBB
- ❖ Backedges, BackedgeBB
- ❖ Exitedges, ExitBB
 - » For each LoopBB, examine each outgoing edge
 - » If the edge is to a BB not in LoopBB, then its an exit
- ❖ Preheader (Preloop)
 - » New block before the header (falls through to header)
 - » Whenever you invoke the loop, preheader executed
 - » Whenever you iterate the loop, preheader NOT executed
 - » All edges entering header
 - Backedges – no change
 - All others, retarget to preheader
- ❖ Postheader (Postloop) - analogous

Find the Preheaders for each Loop

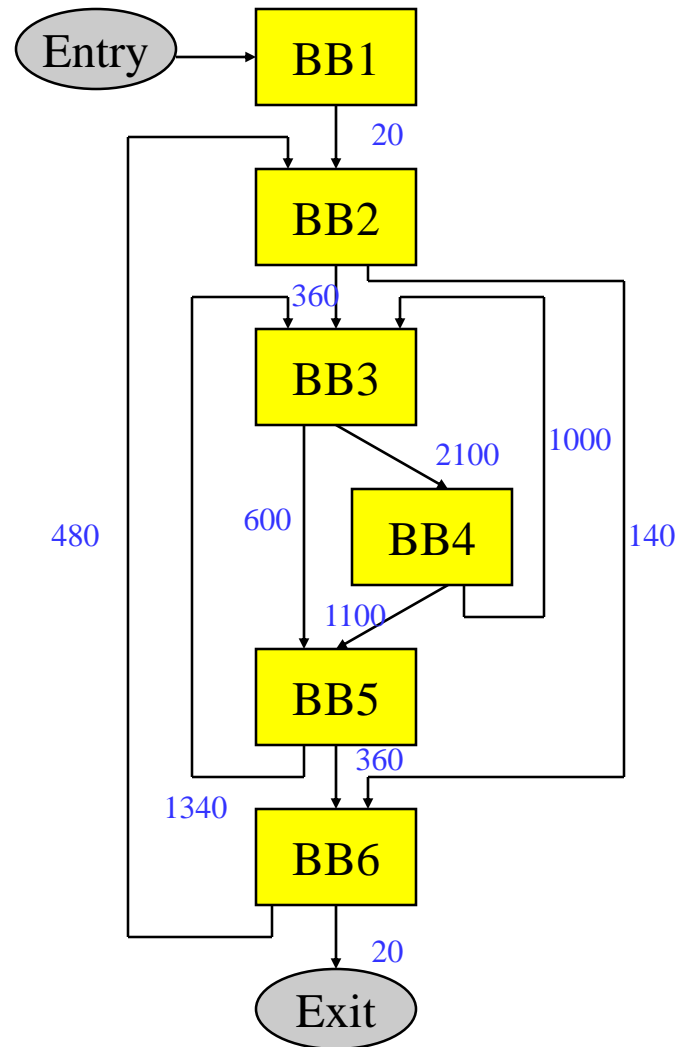


Characteristics of a Loop

- ❖ Nesting (generally within a procedure scope)
 - » Inner loop – Loop with no loops contained within it
 - » Outer loop – Loop contained within no other loops
 - » Nesting depth
 - $\text{depth}(\text{outer loop}) = 1$
 - $\text{depth} = \text{depth}(\text{parent or containing loop}) + 1$
- ❖ Trip count (average trip count)
 - » How many times (on average) does the loop iterate
 - » `for (I=0; I<100; I++)` → trip count = 100
 - » With profile info:
 - $\text{Ave trip count} = \text{weight}(\text{header}) / \text{weight}(\text{preheader})$

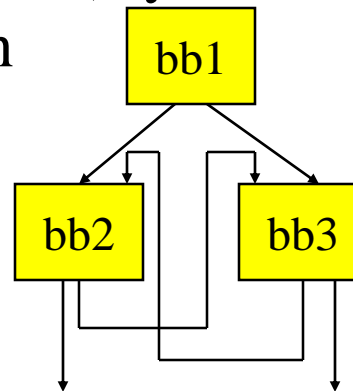
Trip Count Calculation Example

Calculate the trip counts for all the loops in the graph



Reducible Flow Graphs

- ❖ A flow graph is reducible if and only if we can partition the edges into 2 disjoint groups often called forward and back edges with the following properties
 - » The forward edges form an acyclic graph in which every node can be reached from the Entry
 - » The back edges consist only of edges whose destinations dominate their sources
- ❖ More simply – Take a CFG, remove all the backedges ($x \rightarrow y$ where y dominates x), you should have a connected, acyclic graph



Non-reducible!

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

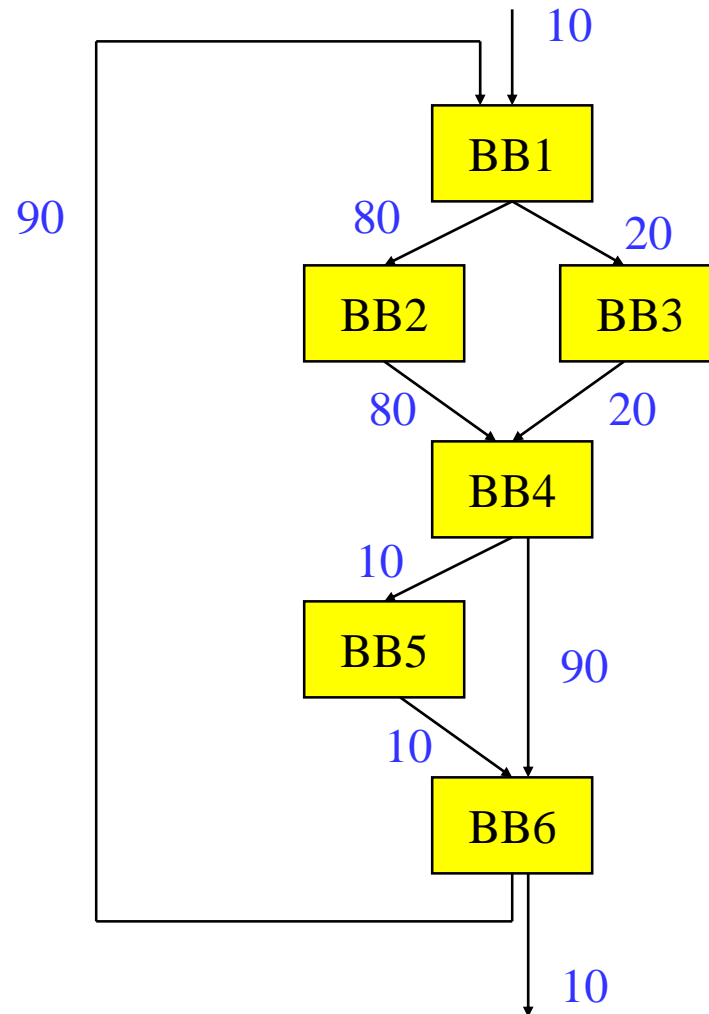
- » Intermediate sized regions with simple control flow
- » Bigger basic blocks would be ideal !!
- » Separate important code from less important
- » Optimize frequently executed code at the expense of the rest

❖ Solution

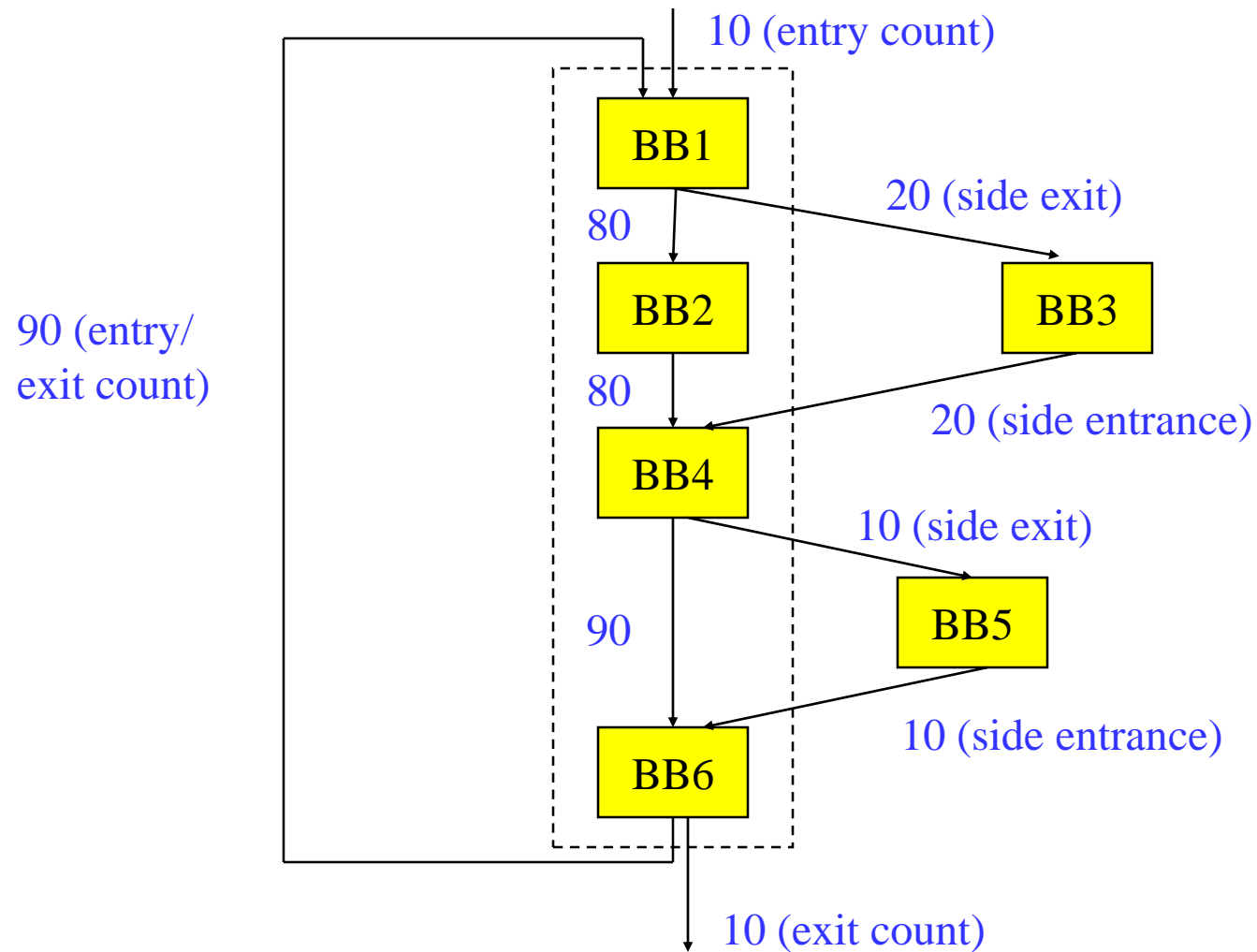
- » Define new region types that consist of multiple BBs
- » Profile information used in the identification
- » Sequential control flow (sorta)
- » Pretend the regions are basic blocks

Region Type 1 - Trace

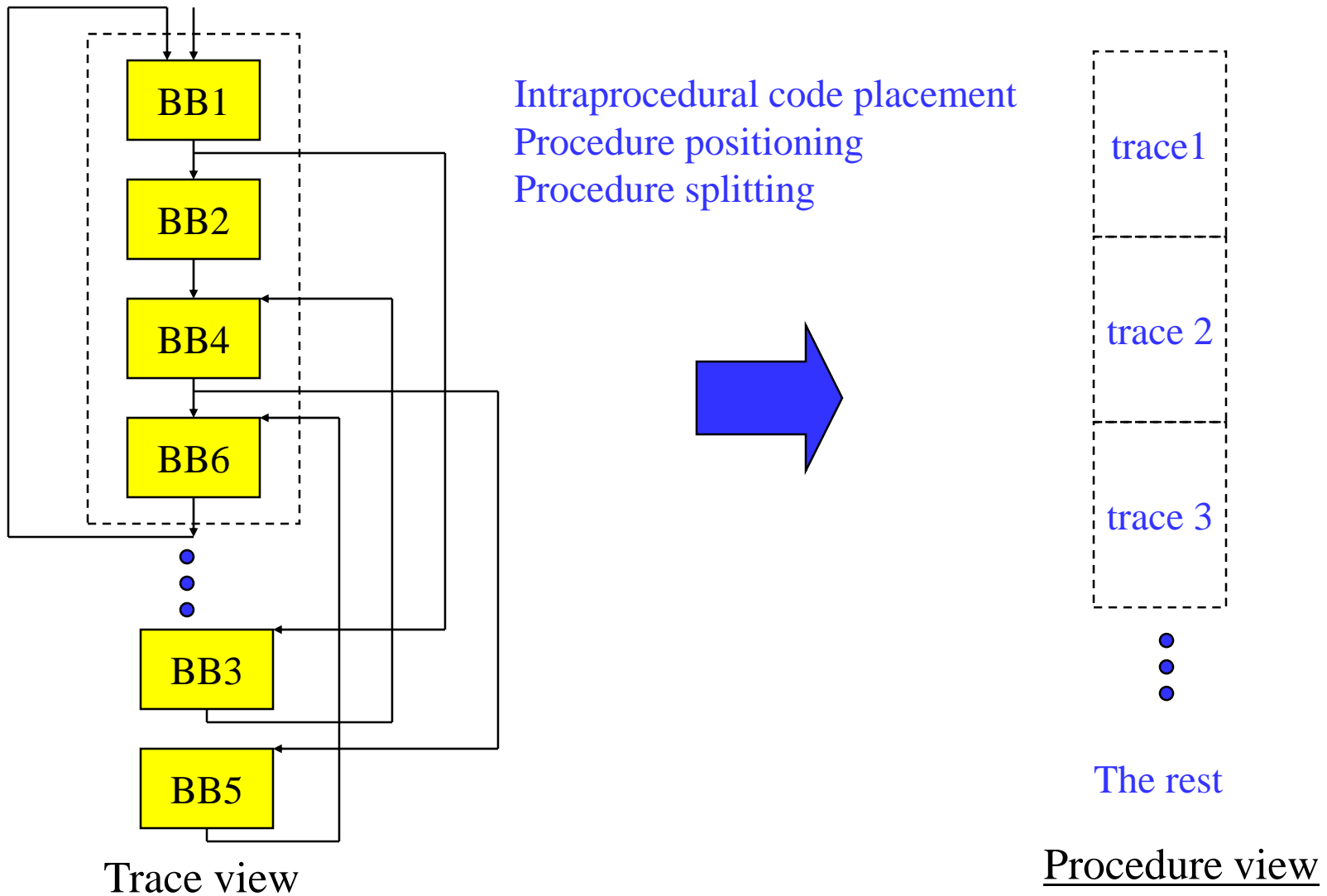
- ❖ Trace - Linear collection of basic blocks that tend to execute in sequence
 - » “Likely control flow path”
 - » Acyclic (outer backedge ok)
- ❖ Side entrance – branch into the middle of a trace
- ❖ Side exit – branch out of the middle of a trace
- ❖ Compilation strategy
 - » Compile assuming path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



Linearizing a Trace

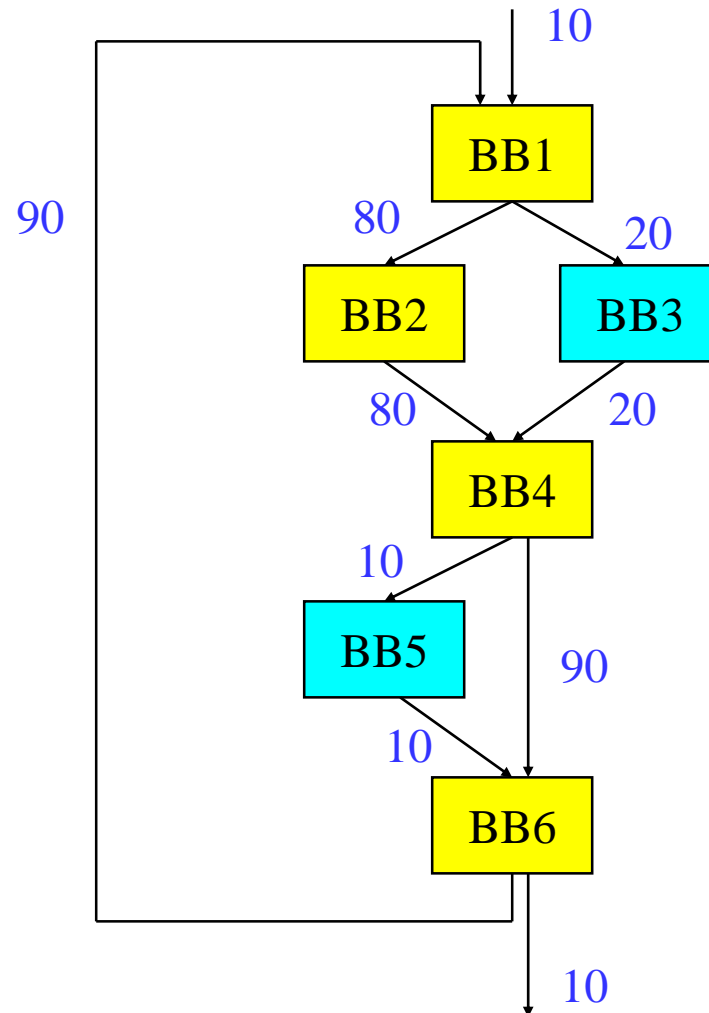


Intelligent Trace Layout for Icache Performance



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
```

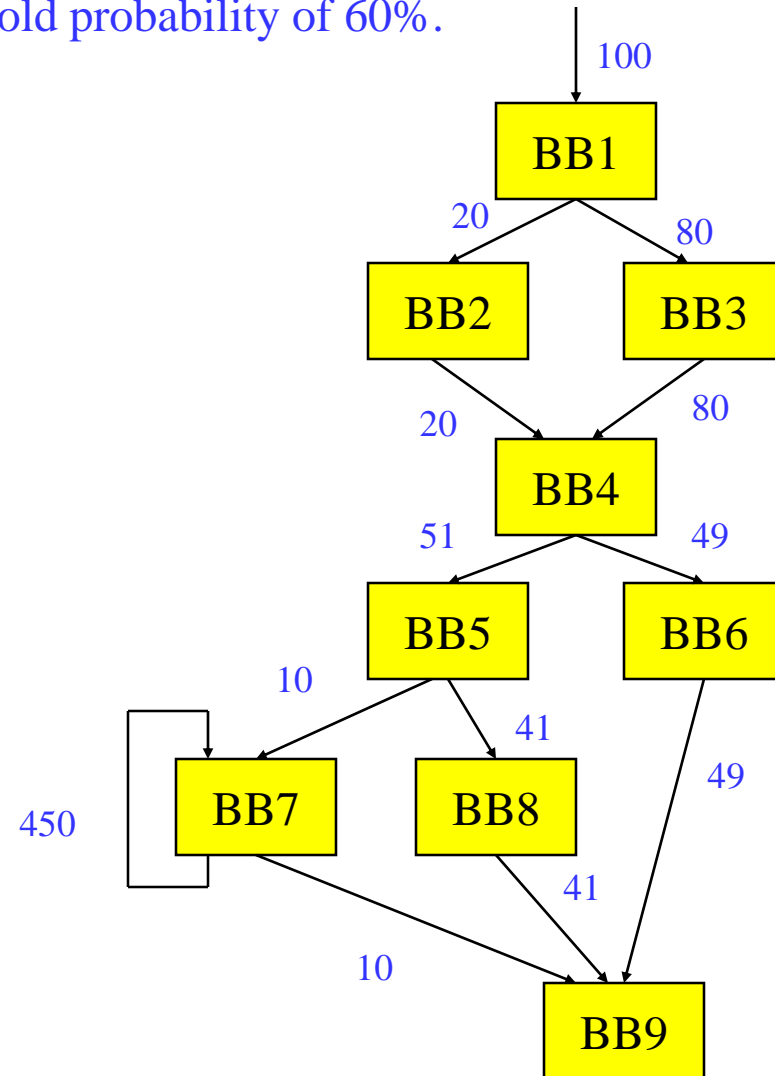
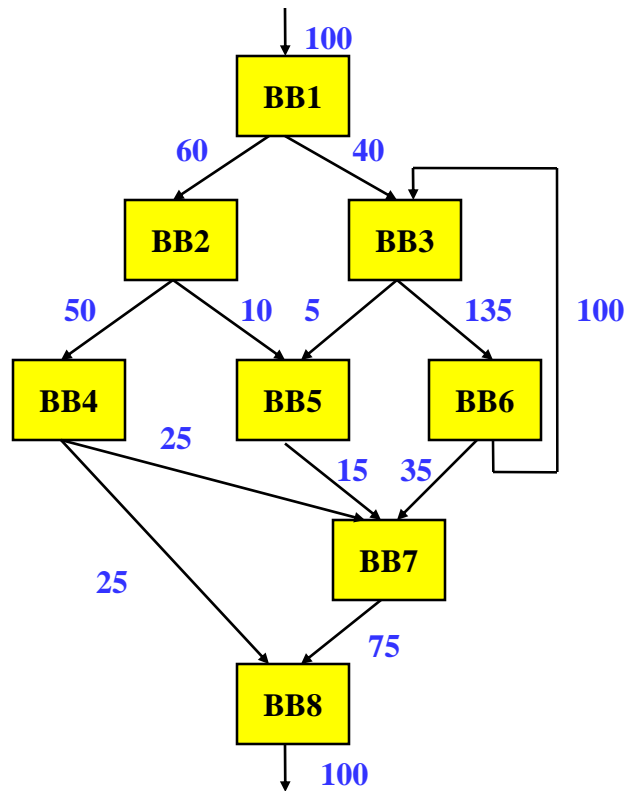
Best Successor/Predecessor

- ❖ 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
```

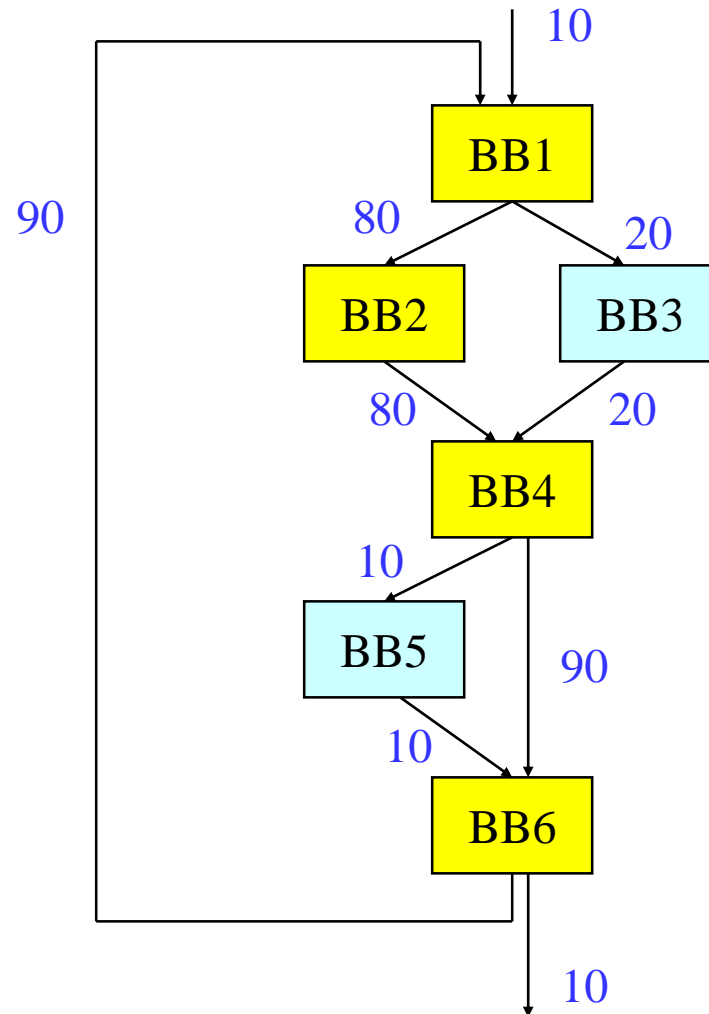
Example Problems

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



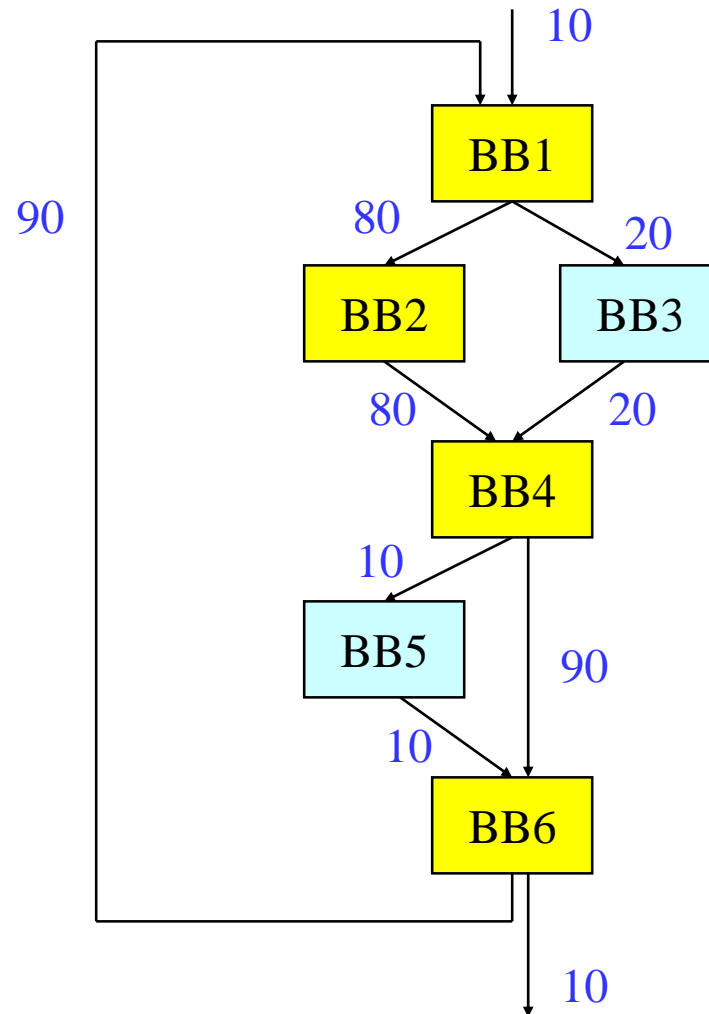
Traces are Nice, But ...

- ❖ 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



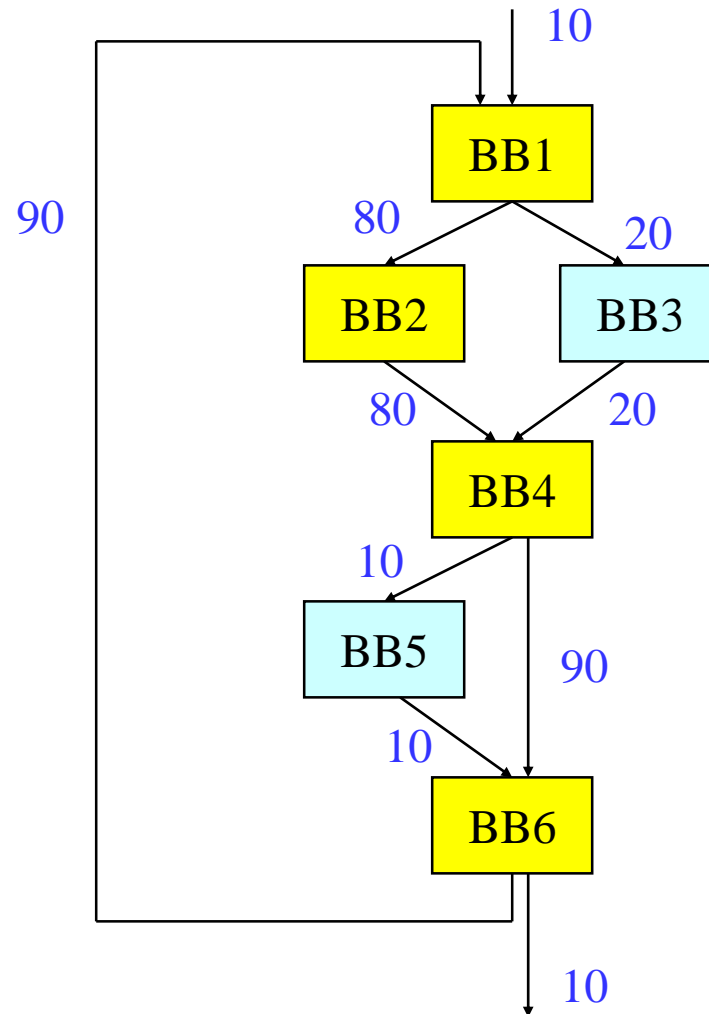
Region Type 2 - Superblock

- ❖ 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

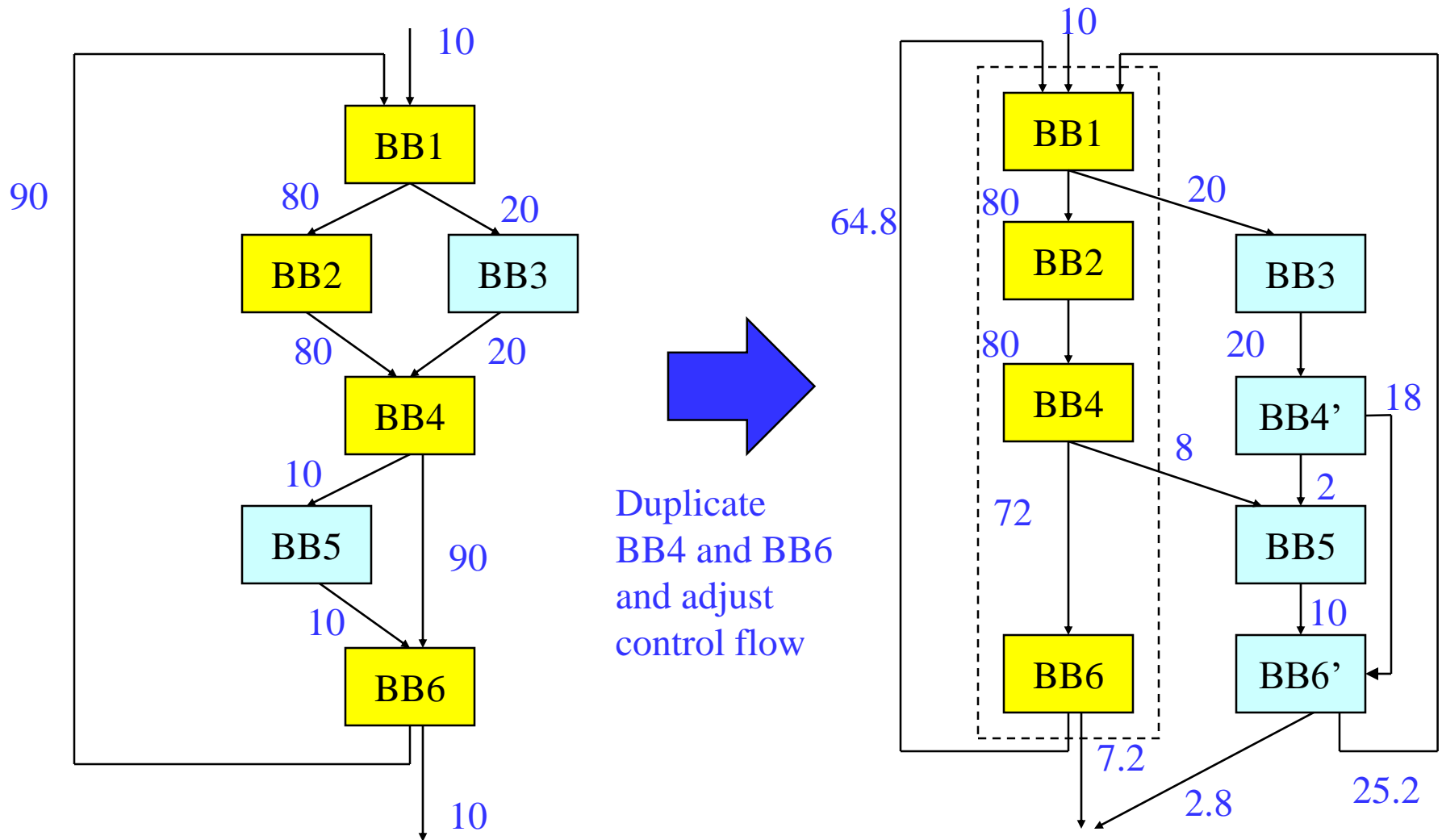


Tail Duplication

- ❖ 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

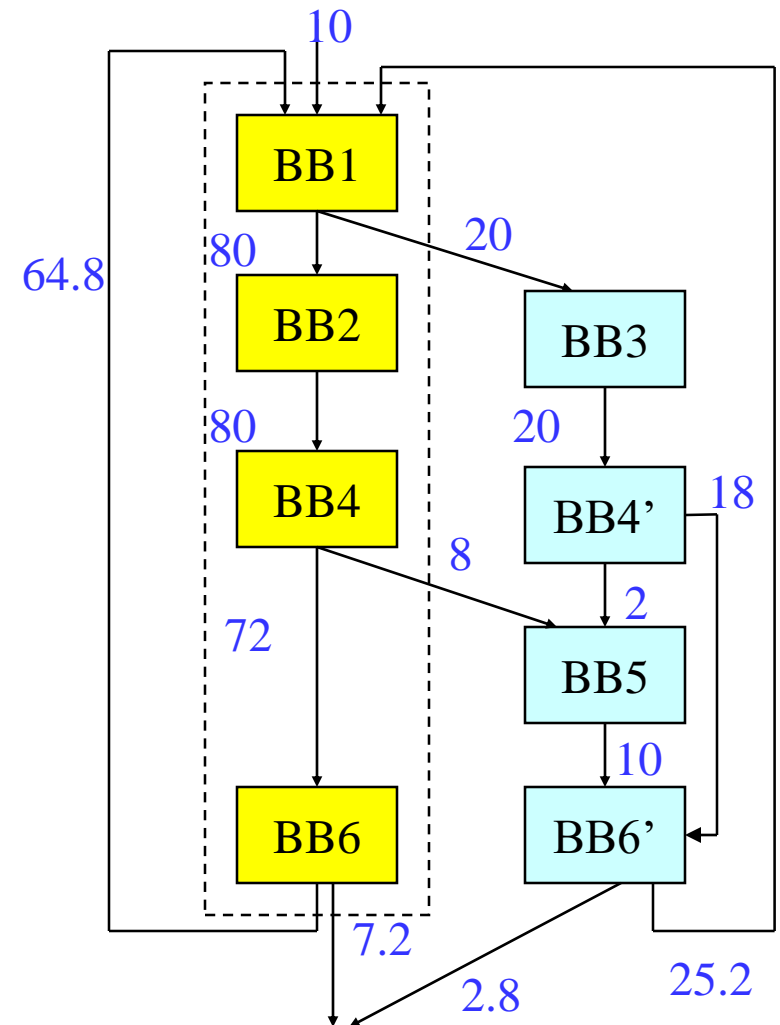


Superblock Formation



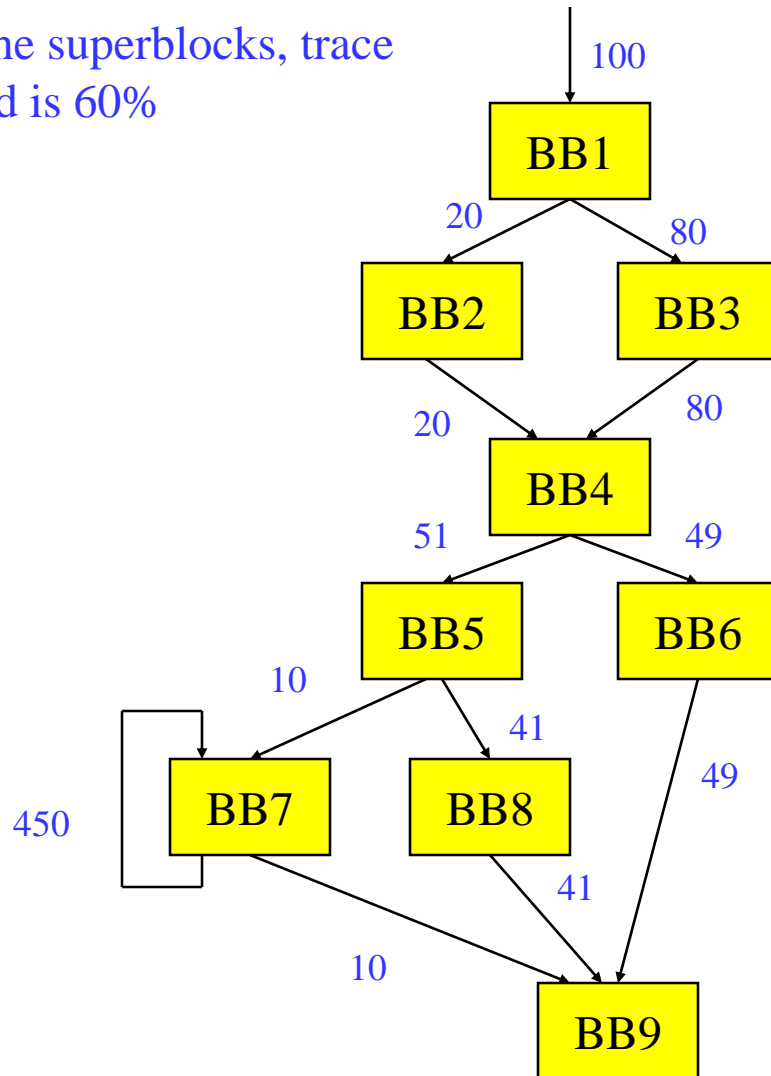
Issues with Superblocks

- ❖ 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



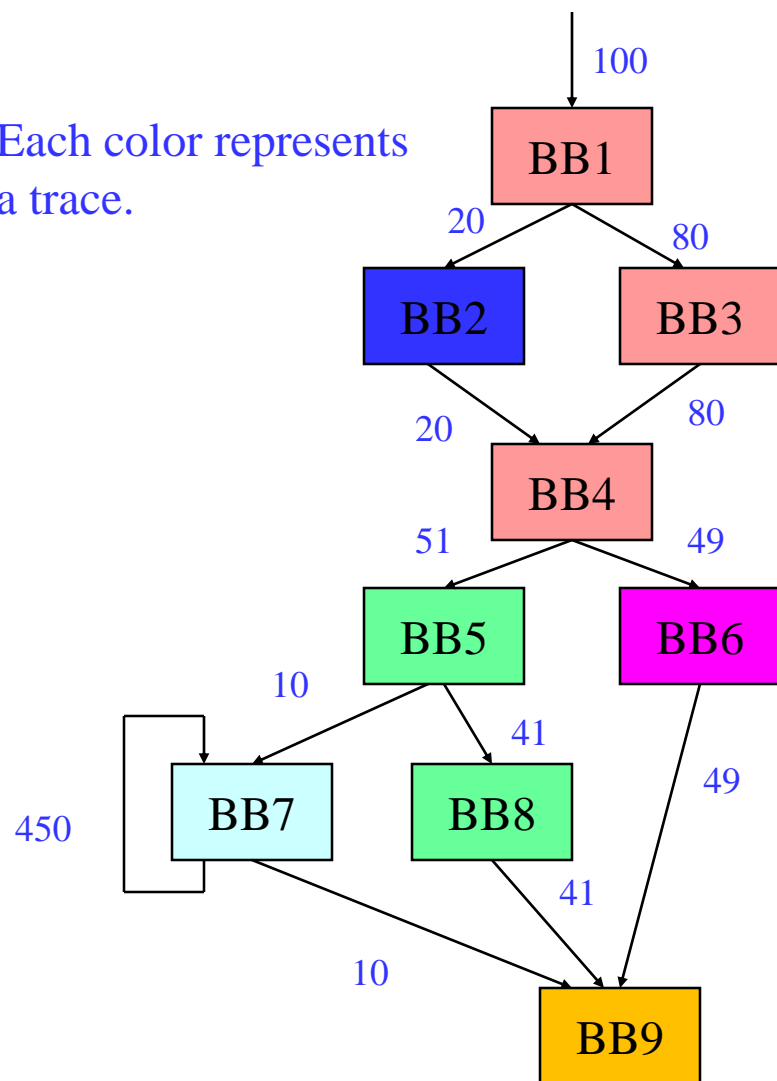
Class Problem

Create the superblocks, trace
threshold is 60%

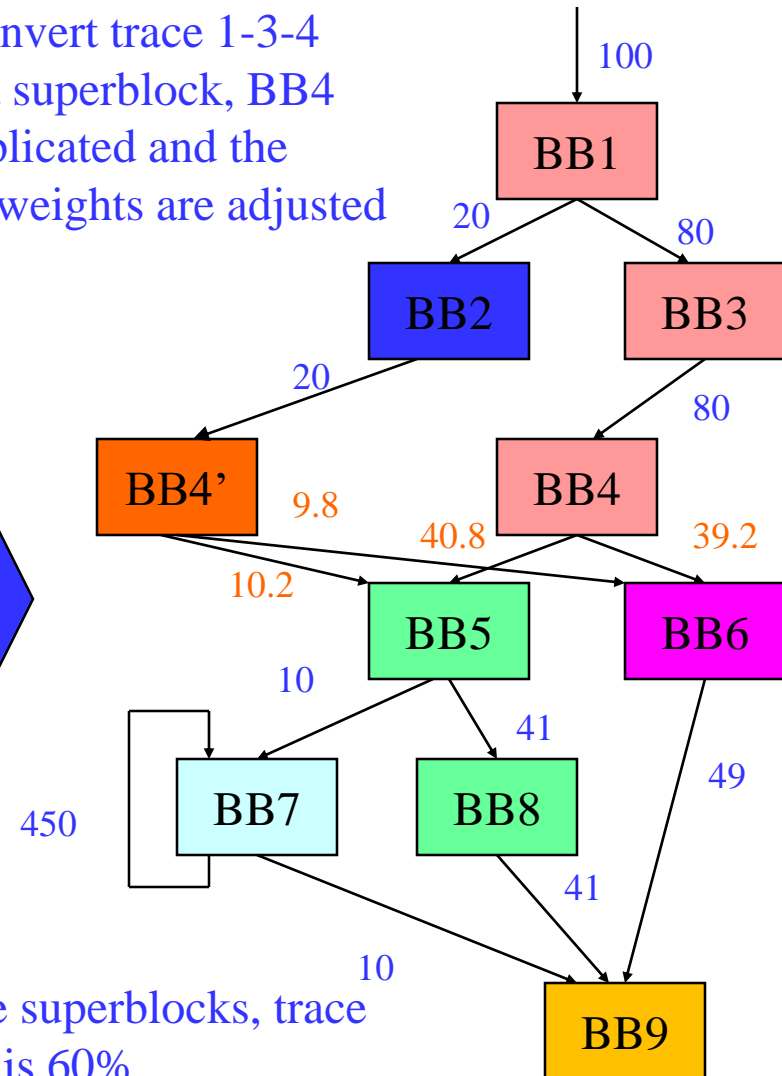
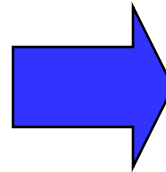


Class Problem Solution – Superblock Formation

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%

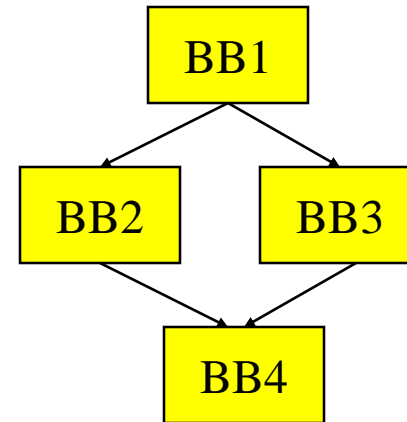
An Alternative to Branches: Predicated Execution

- ❖ 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
```

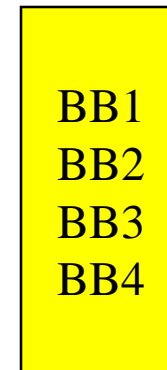
```
BB1  add a, b, c
BB1  bgt a, 0, L1
BB3  div e, f, g
BB3  jump L2
BB2  L1: add e, f, g
BB4  L2: sub h, i, j
```



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
```

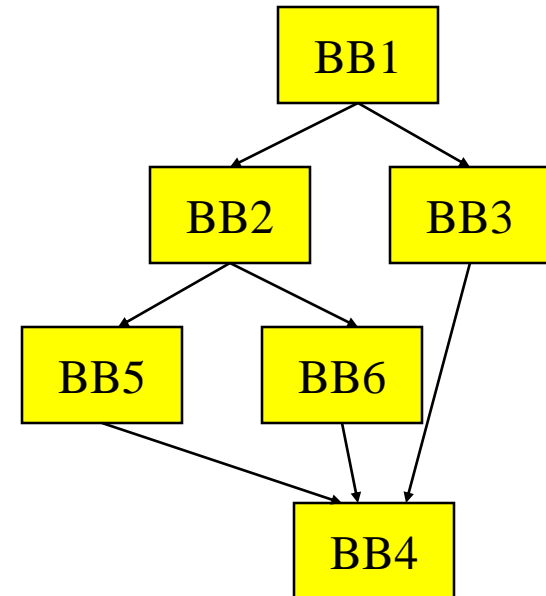


Predicated code

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

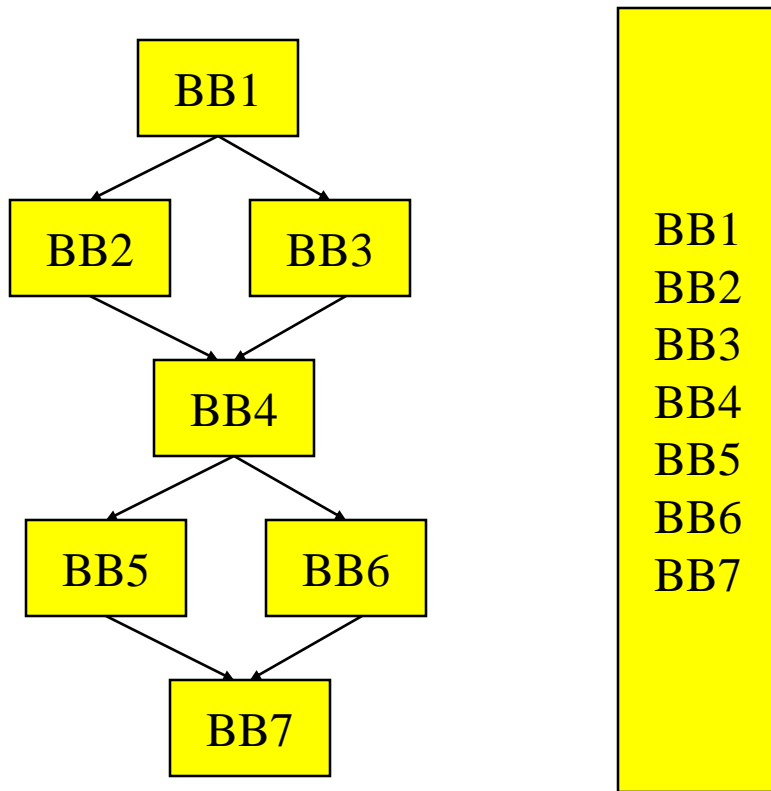
Predicated code

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

Homework 1

- ❖ Get started ASAP. If you haven't done HW0, you are falling behind!
- ❖ Easy to do, but hard to start because of newness
 - » Look for piazza post for help
 - Skeleton code
 - How to run profiler
 - Simple example with opcode stats
 - » Talk to the GSIs if you are stuck