

CMSC 430 (Spring'06)

Practice Problems 2

Use the following 3-address code and and Java stack code instructions for answering code generation questions.

3-addr Instruction	Effect
load R1 x	R1 ← x
store x R1	x ← R1
add R1 R2 R3	R1 ← R2 + R3
sub R1 R2 R3	R1 ← R2 - R3
mult R1 R2 R3	R1 ← R2 * R3
neg R1 R2	R1 ← -(R2)

Java Stack Code	Effect
nop	none
ldc_int c	push constant c onto stack
iload index(x)	push local variable X onto stack
istore index(x)	pop stack, store in local variable X
iadd	pop 2 elems off stack, add, push
isub	pop 2 elems off stack, subtract, push
imult	pop 2 elems off stack, multiply, push
ineg	pop stack, negate, push
goto L	jump to handle L
ifeq L	pop stack, jump to handle L if zero
if_icmpeq L	pop 2 elems, jump to L if equal
if_icmpgt L	pop 2 elems, jump to L if 1st greater
dup	duplicate top of stack
pop	pop top of stack
swap	swap top two positions of stack

1. Code generation.

You are generating code for a Java stack machine. You are given the following grammar attributes and helper functions:

Attribute	Holds
AstNode.code	list of instructions
Function	Effect
genInst(X)	create new instruction X returns handle to instruction
append(...)	concatenates lists of instructions

- What grammar actions needed to generate code for a C-style IF statement in the following production?

$$\text{stmt} \rightarrow \text{IF} (\text{exp}) \text{stmtList} ;$$

$$\{ \text{stmt.code} = ??; \}$$
- What grammar actions needed to generate code for a C-style FOR loop in the following production?

$$\text{stmt} \rightarrow \text{FOR} (\text{stmt} ; \text{exp} ; \text{stmt}) \text{stmt} ;$$

$$\{ \text{stmt.code} = ??; \}$$
- Write grammar actions needed to generate control code for an AND expression in the following production, using numerical value representation of booleans. Use *short circuiting*.

$$\text{exp} \rightarrow \text{exp}_1 \text{ AND } \text{exp}_2$$

$$\{ \text{exp.code} = ??; \}$$

- Write grammar actions needed to generate control code for an NOR expression in the following production, using numerical value representation of booleans. Use *short circuiting*.

$$\text{exp} \rightarrow \text{exp}_1 \text{ NOR } \text{exp}_2$$

$$\{ \text{exp.code} = ??; \}$$
- Write grammar actions needed to generate control code for an \geq (GEQ) expression in the following production, using numerical value representation of booleans.

$$\text{exp} \rightarrow \text{exp}_1 \text{ GEQ } \text{exp}_2$$

$$\{ \text{exp.code} = ??; \}$$

2. Complex code generation.

- Name two issues and solutions to generating code for function calls in C.
- Name two issues and solutions to generating code for array references in C.
- What code must the compiler generate for the code

$$\text{int } i, a[100] ;$$

$$\dots$$

$$a[i + 5] = 4 ;$$
- What code must the compiler generate for the code

$$\text{int } \text{foo} (\text{int } x) ;$$

$$\dots$$

$$x = \text{foo}(i + 2) ;$$
 - Assuming $\text{foo}()$ is call-by-value?
 - Assuming $\text{foo}()$ is call-by-reference?

3. Optimizations

- How can compiler transformations improve a program?
- What does the compiler need to consider when applying optimizations?
- What are the different scopes of compiler optimizations? What are the tradeoffs when considering what scope of optimizations to use?

4. Local optimizations

Consider the following code.

- (1) a := 1
- (2) b := f + a
- (3) c := a
- (4) d := f + a
- (5) e := f + c
- (6) f := b
- (7) g := f + a

- Build a DAG for the code.

5. Control flow analysis

For the following problems, consider this code:

```

<S1>      a := 1
<S2>      b := 2
<S3>      L1:  c := a + b
<S4>      d := c - a
<S5>      if (...) goto L3
<S6>      L2:  d := b * d
<S7>      if (...) goto L3
<S8>      d := a + b
<S9>      e := e + 1
<S10>     goto L2
<S11>     L3:  b := a + b
<S12>     e := c - a
<S13>     if (...) goto L1
<S14>     a := b * d
<S15>     b := a - d

```

- What are the basic blocks?
- What is the control flow graph?
- Depth-first order selects nodes in the order they are visited (start by visiting the root node) and then recursively visiting every child of each node (if the child has not been visited before). Note that the order in which children are visited is random.

What are all the possible results of depth-first traversal on the control flow graph?

- Using depth-first order, is it possible to visit a child before its parent? For which depth-first ordering(s) of the control flow graph does this occur?
- Postorder selects nodes (starting from root) *after* visiting every child of that node (if the child has not been visited before). Note that the order in which children are visited is random.

What are all the possible results of Postorder traversal for the control flow graph?

- Reverse Postorder simply reverses the node ordering found by a Postorder traversal of the graph. What are the possible Reverse Postorder traversals of the control flow graph?
- Using Reverse Postorder, is it possible to visit a child before its parent? Why or why not?

6. Reaching definitions

Reaching definitions for a point in the program p is defined as the set of definitions of a variable for which there is some path from the definition to p with no other definition of that variable. Calculate reaching definitions for the code in the control-flow graph problem.

- What is the dataflow equation for REACH?
- In what direction is REACH calculated? I.e., does information flow forwards or backwards in the CFG?
- Calculate GEN, KILL for each basic block.
- What is a good initial value for REACH for each basic block?

- Solve the data-flow equations in reverse Postorder. Show your work.

7. Live variables

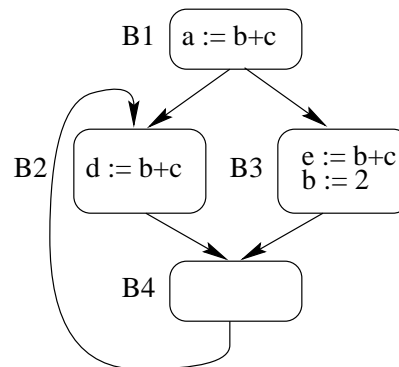
Live variables for a point in the program p is defined as the set of variables x for which there is some path from p to a use of x with no definition to x on the path. Calculate live variables for the code in the control-flow graph problem.

- We define $LIVE(b)$ for a basic block b to be the set of live variables at the end of b . What is the dataflow equation for LIVE?
- In what direction is LIVE calculated? I.e., does information flow forwards or backwards in the CFG?
- Show GEN, KILL for each basic block.
- What is a good initial value for LIVE for each basic block?
- Solve the data-flow equations in rPostorder. Show your work.

8. Available expressions

Available expressions is a data-flow analysis problem whose solution is used to guide global common subexpression. It calculates AVAIL, the expressions available at the beginning of each basic block.

Consider the following code. Assume that $b+c$ is the only expression of interest:



- What is the data-flow equation for AVAIL?
- Give GEN and KILL (needed by AVAIL) for each basic block.
- What is a good initial value for AVAIL for each basic block?
- Calculate AVAIL. Show all steps, including values for AVAIL and the order basic blocks are analyzed.

9. Data-flow lattices

Prove the following properties of lattices:

- Show that $a \leq b$ and $b \leq c$ implies $a \leq c$
- Show that $a \leq (b \wedge c)$ implies $a \leq b$

10. Data-flow frameworks

- (a) When estimating each of the following sets, tell whether too-large or too-small estimates are conservative. Explain your answer in terms of the intended use of information.
- Available expressions
 - Reaching definitions
 - Live variables
- (b) What properties are necessary to ensure an iterative data-flow analysis framework terminates?
- (c) What properties are necessary to ensure an iterative data-flow analysis framework terminates with the meet-over-all-paths solution?

11. Instruction scheduling

Consider scheduling the code below using list scheduling. All instructions must complete before executing the *jmp* instruction. Assume the following instruction latencies:

- 2-cycle latency for load
- 1-cycle latency otherwise

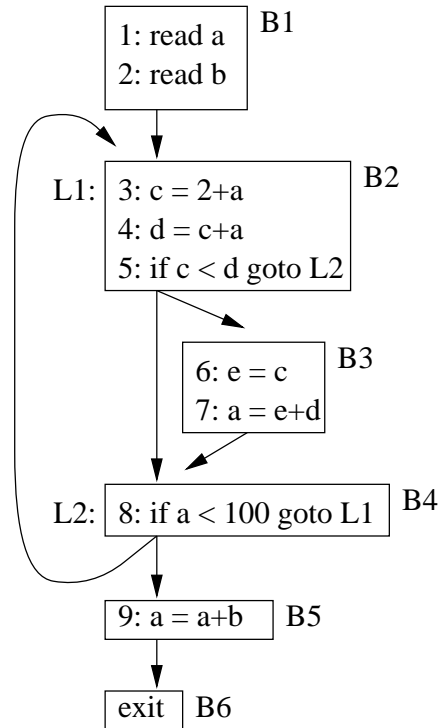
<op> <dst, s1, s2>

```
1  load  r1, a
2  add   r2, r1, #4
3  store x, r2
4  load  r3, b
5  mult  r4, r3, r2
6  load  r1, c
7  add   r5, r1, r3
8  store y, r5
9  load  r6, d
10 mult  r7, r5, #1
11 store z, r7
12 jmp
```

- (a) Build the precedence graph for the instructions. Mark dependences as flow, anti, or output. You can ignore transitive dependences.
- (b) Calculate the critical path for the instructions.
- (c) Schedule the instructions for a single-issue processor, using forward list scheduling. Showing candidates instructions at each cycle. Prioritize candidates using 1) critical path, 2) latency of instruction, 3) number of children.
- (d) Schedule the instructions as above, for a two-issue VLIW processor.
- (e) How could you change register assignments to improve instruction schedules in the code?

12. Register allocation

Consider the flow graph below. Each statement is labeled by its statement number and each basic block is labeled in the upper right hand corner.



Use statement numbers or basic block numbers to indicate live ranges as appropriate.

- (a) What are the live ranges for a global top-down allocator?
- (b) Draw the interference graph for the live ranges.
- (c) Use the graph-simplification method to find a coloring for this graph.
- (d) Can you color this graph with fewer colors?
- (e) If spilling is needed, which live range would be spilled first? Why?
- (f) Draw the spill code needed if the value for *c* is spilled.