CS243 Final Examination

Winter 2019

March 22, 2019

This is an open-book, open-notes, open-laptop, *closed*-network exam. Please do not post anything on Piazza until the solutions are put up on the class website.

You have 3 hours to work on this exam. The examination has 7 problems worth 160 points. Please budget your time accordingly. Write your answers in the space provided on the exam. If you use additional scratch paper, please turn that in as well.

Your Name: _____ SUNet ID: _____

The following is a statement of the Stanford University Honor Code:

a. The Honor Code is an undertaking of the students, individually and collectively:

   (i) that they will not give or receive aid in examinations; that they will not give or receive unpermitted aid in class work, in the preparation of reports, or in any other work that is to be used by the instructor as the basis of grading;

   (ii) that they will do their share and take an active part in seeing to it that others as well as themselves uphold the spirit and letter of the Honor Code.

b. The faculty on its part manifests its confidence in the honor of its students by refraining from proctoring examinations and from taking unusual and unreasonable precautions to prevent the forms of dishonesty mentioned above. The faculty will also avoid, as far as practicable, academic procedures that create temptations to violate the Honor Code.

c. While the faculty alone has the right and obligation to set academic requirements, the students and faculty will work together to establish optimal conditions for honorable academic work.

Signature: _____

Problem	#1	#2	#3	#4	#5	#6	#7	**Total**
Score								
Max	20	20	20	20	25	25	30	160

Problem 1. True or False? Briefly justify your answer in 1 to 5 lines [20 Points]

a. When moving an instruction in a downward code motion, if the source block dominates the destination block, but the destination block does not post-dominate the source block, an extra operation may be executed.
True. If the destination block were contained in a loop that the instruction would not previously have been executed within, extra operations may be executed.

b. In partial garbage collection, none of the garbage in the stable set is collected, but all the garbage in the target set is collected.

False - not all garbage in target set is collected. Garbage in the target set that is pointed to by garbage in the stable set is not collected.

c. All primitive operations (Apply, Restrict, Exists) on Binary Decision Diagrams can be performed in time polynomial on the size (number of nodes) of the BDD.

True. They are exponential in the **number of variables** of the BDD, but polynomial in the number of nodes.

d. There is no disadvantage to allocating registers before global instruction scheduling.

False. Can introduce dependencies that prevent the global instruction scheduling from moving operations

e. For the expression $(x_1$ `OR` $x_2)$ `AND` $(\neg x_1$ `OR` $\neg x_2)$, it is not possible to create a BDD with just 4 nodes (including the 0 and 1 nodes).

True - regardless of the variable order, the nodes representing the second variable will have opposing edges depending on the value of the first variable. So we cannot merge any nodes.

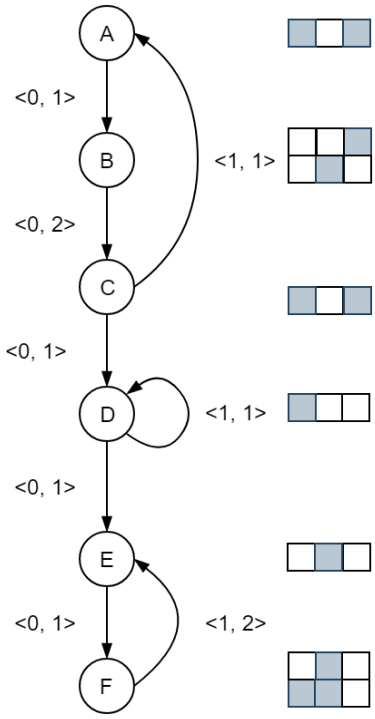
Problem 2. Partial Redundancy Elimination [20 Points]

Show the result of applying partial redundancy elimination to the following program. You do not need to show the intermediate steps.

After applying PRE, we arrive at the same graph. A common error was to place the temporaries where they are not actually anticipated. Solutions that have equivalent code but do not perform the meet of the postponable set received full credit.

Problem 3. Software Pipelining [20 Points]

In the following data dependence graph of a loop, each node is labeled with its resource reservation table, and each dependence edge is labeled with a tuple: the iteration difference and the delay.



a. What is the lower bound imposed on the initiation interval by resource constraints?

4

b. What is the lower bound imposed on the initiation interval by precedence constraints?

4 from the first cycle

c. Find the optimal software pipelined schedule for this loop. What is the minimum initiation interval? Show the modulo reservation table and the code for a single iteration of the pipelined loop. (Show `nop`s if necessary).
The minimum initiation interval is 4.

Modulo reservation table:

A	F	A
F	F	B
D	B	
C	E	C

Single iteration: A, B, nop, C, nop, nop, D, E, F, nop

d. Can this schedule be generated using the heuristic algorithm described in class? Why or why not?

No. The algorithm would greedily schedule D, and as a result would fail in scheduling the SCC consisting of nodes E and F.

Problem 4. Pointer Analysis [20 Points]

Perform a context-insensitive, flow-insensitive pointer analysis on the code below. The call graph should be constructed on-the-fly as the algorithm discovers the points-to relationships. List the `hP` tuples that are produced by the analysis.

```
public class A {
    public A x;
    public void foo() {
        this.x = new A(); // h1
    }
}

public class B extends A {
    public void foo() {
        this.x = new B(); // h2
    }
}

public class Main {
    public static void main(String [] args) {
        A a = new B(); // h3
        a.x = new A(); // h4
        a.x.foo();
        a.foo();
    }
}
```

pts(a, h3)

htype(h3, B)
htype(h4, A)
htype(h2, B)

hP(h3, x, h4)
hP(h3, x, h2)

hP(h4, x, h1)
hP(h2, x, h2)

hP(h4, x, h2)
hP(h2, x, h1)

Problem 5. Smart Dead Code Elimination [25 Points]

The data flow analysis algorithms discussed in class ignore the value of conditions in the control flow graph. This problem explores how we can improve dead code elimination by analyzing the value of conditions.

Consider a programming language with the following operations:

- An assignment of the form "`x = true;`" or "`x = false;`" where `x` is a variable name.
- A boolean operation of the form "`x = not y;`" or "`x = y && z;`," or "`x = y || z;`" with obvious semantics.
- An assertion of the form "`assert(x);`" or "`assert(not x);`" asserts that the value of variable `x` is true or false, respectively. The program terminates if the assertion fails; thus the assertion can be assumed to be true if the execution proceeds to the next statement.
- A condition of the form "`if (x)`". Each condition has two control edges: one is labeled `true`, taken when `x` is true, and the other labeled `false`, taken when `x` is false.
- An unconditional "`branch`" statement which has a single control flow edge leading to the successor basic block.

Your task is to design a compiler that eliminates dead code, including the removal of unnecessary conditional statements, by statically determining if certain edges can never be taken. Assume that all variables are live at the end of the program.

For this problem, you may change the control flow graph, add instructions, and apply multiple passes of data flow analysis. If you are using a known data flow analysis, you need to mention only its name; if you are defining a new data flow analysis, answer the questions below. Do NOT use SMT in your solution.

a. What is the direction of the analysis?

b. What is the domain and boundary condition?

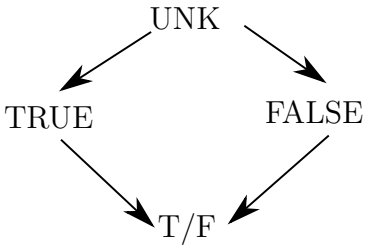
c. Describe the meet operator.

d. Describe the transfer function.

e. Is your data flow framework monotone?

f. Is your data flow framework distributive?

Boolean propagation: forward dataflow analysis with (per-variable) domain {True, False, UNK, T\F }.
Semi-lattice:



Boundary condition: Initialize all variables to T\F.
Transfer function:

- `assert(x)`, `x = true`: set `x` to T
- `assert(not x)`, `x = false`: set `x` to F
- `x = not y`: set `x` to:

y	x
UNK	UNK
T	F
F	T
T\F	T\F

- `x = y && z`: set `x` to:

	UNK	T	F	T\F
UNK	UNK	UNK	UNK	UNK
T	UNK	T	F	T\F
F	UNK	F	F	F
T\F	UNK	T\F	F	T\F

- `x = y || z`: set `x` to:

	UNK	T	F	T\F
UNK	UNK	UNK	UNK	UNK
T	UNK	T	T	T
F	UNK	T	F	T\F
T\F	UNK	T	T\F	T\F

The framework is monotone, but not distributive.

CFG preprocessing: for each condition of the form "`if (x)`", add a basic block along each of the two edges leaving the conditional statement. `assert(x);` on the control edge corresponding to the `true` branch, and a statement `assert(not x);` on the control edge corresponding to the `false` branch.

Passes (repeat until the CFG is no longer modified):

- Boolean Propagation

- CFG modification: remove all edges going out of `assert(x);` blocks where `IN[b][x] == False`, and remove all edges going out of `assert(not x);` blocks where `IN[b][x] == True`.

  For every condition `if (x)`, if `IN[b][x] == True`, then remove this block, adding an edge from every predecessor of b to the first block of the true branch of b. If `IN[b][x] == False`, then remove this block, adding an edge from every predecessor of b to the first block of the false branch of b.

- Remove all basic blocks that are no longer reachable from entry.

Problem 6. Affine Transforms/Parallelization [25 Points]
Consider the following code:

```
for (i=1; i<n; i++) {
    A[i,i] = B[i];
    for (k=i+1; k<n; k++) {
        A[i,k] = A[i-1,k] / A[i,i];
    }
}
```

a. Is there any communication-free parallelism that can be obtained using affine transforms? If yes, show the result. You need not optimize the code generated; simply wrap all the code in a conditional statement.

There is no communication-free parallelism that can be obtained using affine transforms.

b. Find the outermost fully permutable loop nest in this program. You need not optimize the code generated; simply wrap all the code in a conditional statement.

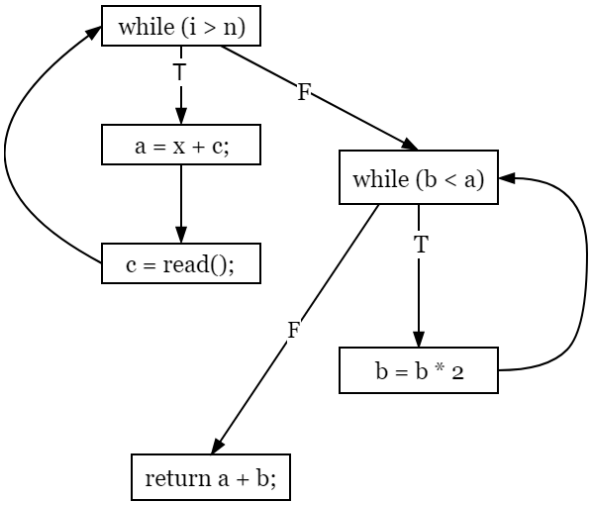
Solution:

```
for (i=1; i<n; i++) {
    for (k=i; k<n; k++) {
        if (k == i) A[i,i] = B[i];
        if (k > i) A[i,k] = A[i-1,k] / A[i,i];
    }
}
```

Problem 7. While Loops [30 Points]

Your task is to create 2 tools that warn if a `while` loop may not terminate. You may assume that the `while`⟨`cond`⟩ loop has only one exit, which is taken when the ⟨`cond`⟩ evaluates to `false`.

a. Your first tool, T1, will warn about a possible infinite loop if none of the variables used in ⟨`cond`⟩ is updated in the loop. Consider the following example; each edge is annotated by the value of the condition evaluated in the source block.



For this example, a warning should be issued only for `while(i > n)`, as neither `i` nor `n` is updated within the loop.

A warning should be issued on each `while`⟨`cond`⟩ that may not terminate. **You may use any concepts learned in this course, other than SMT.** If you are using a known analysis, you must be specific about what that analysis is. If you are defining a new dataflow algorithm, answer the following questions:

(i) What is the direction of the analysis?

(ii) What is the domain and boundary condition?

(iii) Describe the meet operator and transfer function.

(iv) Is your data flow framework monotone?

(v) Is your data flow framework distributive?

Find the natural loops in the graph. For each loop, given that a `while(cond)` serves as the header, say `x` is the variable being tested in the condition, check all nodes in the loop for an update of `x`. If no such instruction exists, generate a warning on the `while(cond)` statement.

b. Your task is to build a different tool T2 using SMT. Consider this following program:

```
while(i < N) {
   i--;
   prev = data[i];
   i++;
   diff[i] = data[i] - prev;
}
```

We see that the value of `i` is updated within a loop, but its value remains unchanged at the end of each iteration. Unroll the loop twice, and write the SMT formula that checks if the condition tested at the end of two iterations may differ from the condition evaluated on entry of the loop.
P:
$\phi_1 = (i_0 < N);$
$i_1 = i_0 - 1$
$prev_0 = data_0[i_1];$
$i_2 = i_1 + 1;$
$diff_0[i_2] = data_0[i_2] - prev_0;$
$\phi_2 = (i_2 < N);$
$i_3 = i_2 - 1;$
$prev_0 = data_0[i_3];$
$i_4 = i_3 + 1;$
$diff_0[i_4] = data_0[i_4] - prev_0;$
$i_5 = \phi_2 \ ? \ i_4 \ : \ i_2;$
$i_6 = \phi_1 \ ? \ i_5 \ : \ i_0;$
$\phi_3 = (i_6 < N);$
SMT formula: P $\wedge(\phi_1 \wedge \neg\phi_3)$

c. Now consider applying the same approach to an arbitrary loop, what can you conclude if there is a solution to your SMT formula? The loop can possibly terminate, as we have found a satisfying assignment, though we cannot conclusively say it will always terminate.

d. What can you conclude if there is no solution to your SMT formula? As we only unroll twice, we can only conclude that the loop will not terminate within two iterations.