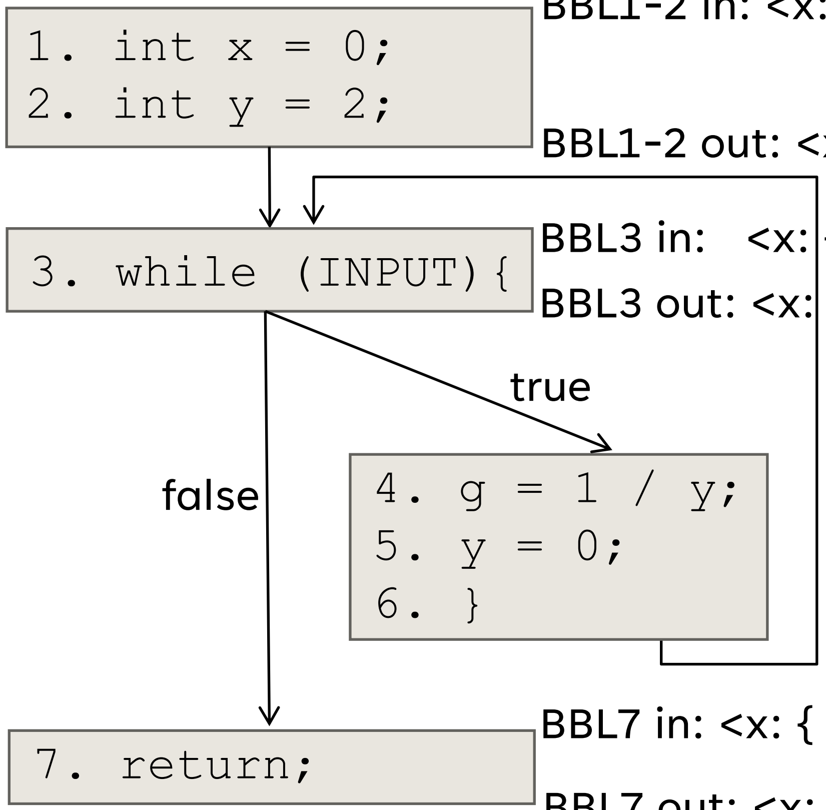
EXERCISE #7

DATAFLOW SATURATION REVIEW

Write your name and answer the following on a piece of paper

- Assume a value-set dataflow analysis starting at BBL 7, then BBL 4-6, then BBL 3, then BBL 1-2. Give the value sets at the top of each block after 1 round of analysis?



BBL1-2 in: <x: { } , y: { }, g: { }> → **<x: { } , y: { }, g: { }>**

BBL1-2 out: <x: { } , y: { }, g: { }> → **<x: { 0 } , y: { 2 }, g: { }>**

BBL3 in: <x: { } , y: { }, g: { }> → **<x: { } , y: { 0 }, g: { }>**

BBL3 out: <x: { } , y: { }, g: { }> → **<x: { } , y: { 0 }, g: { }>**

BBL4-6 in: <x: { } , y: { }, g: { }> → **<x: { } , y: { }, g: { }>**

BBL4-6 out: <x: { } , y: { }, g: { }> → **<x: { } , y: { 0 }, g: { }>**

BBL7 in: <x: { } , y: { }, g: { }> → **<x: { } , y: { }, g: { }>**

BBL7 out: <x: { } , y: { }, g: { }> → **<x: { } , y: { }, g: { }>**

test: 9/8 :

~~9/15~~ : ~~winner~~

9/13 :

ADMINISTRIVIA AND ANNOUNCEMENTS

FORMALIZING DATAFLOW

EECS 677: Software Security Evaluation

Drew Davidson

CLASS PROGRESS

EXPLORING STATIC ANALYSIS

- FINISHED ENOUGH INTUITION THAT WE CAN PERFORM A BASIC ANALYSIS

LAST TIME: SATURATION

REVIEW: STATIC ANALYSIS

Extending our basic dataflow to loops

- No obvious start-point for analysis (circular dependence)

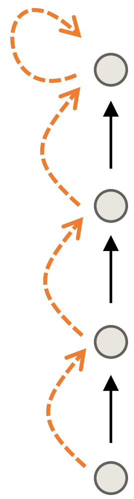
  Chaotic iteration

- No obvious end-point (can't necessarily do with a single pass)

  Run the algorithm until it hits a fixpoint

Reaching Fixpoints Faster

- Intuitively: add some extra over-approximation



Perform an operation until it stops making progress

LECTURE OUTLINE

- Dataflow Frameworks
- Abstract Interpretation

FORMALIZING TERMINATION

DATAFLOW FRAMEWORKS

Our value-set analyses (appeared to) have some nice properties

- Guaranteed termination
- Completeness in values found

A couple of conditions happened to occur:

- A domain *D* of dataflow facts with a particular ordering
  Sets of possible integer values
- An operator to combine distinct dataflow facts
  Union over value-sets
- A dataflow function f_n: $D \rightarrow D$ that defines the effect of BBL_n
  Composition of the individual instruction transfer functions

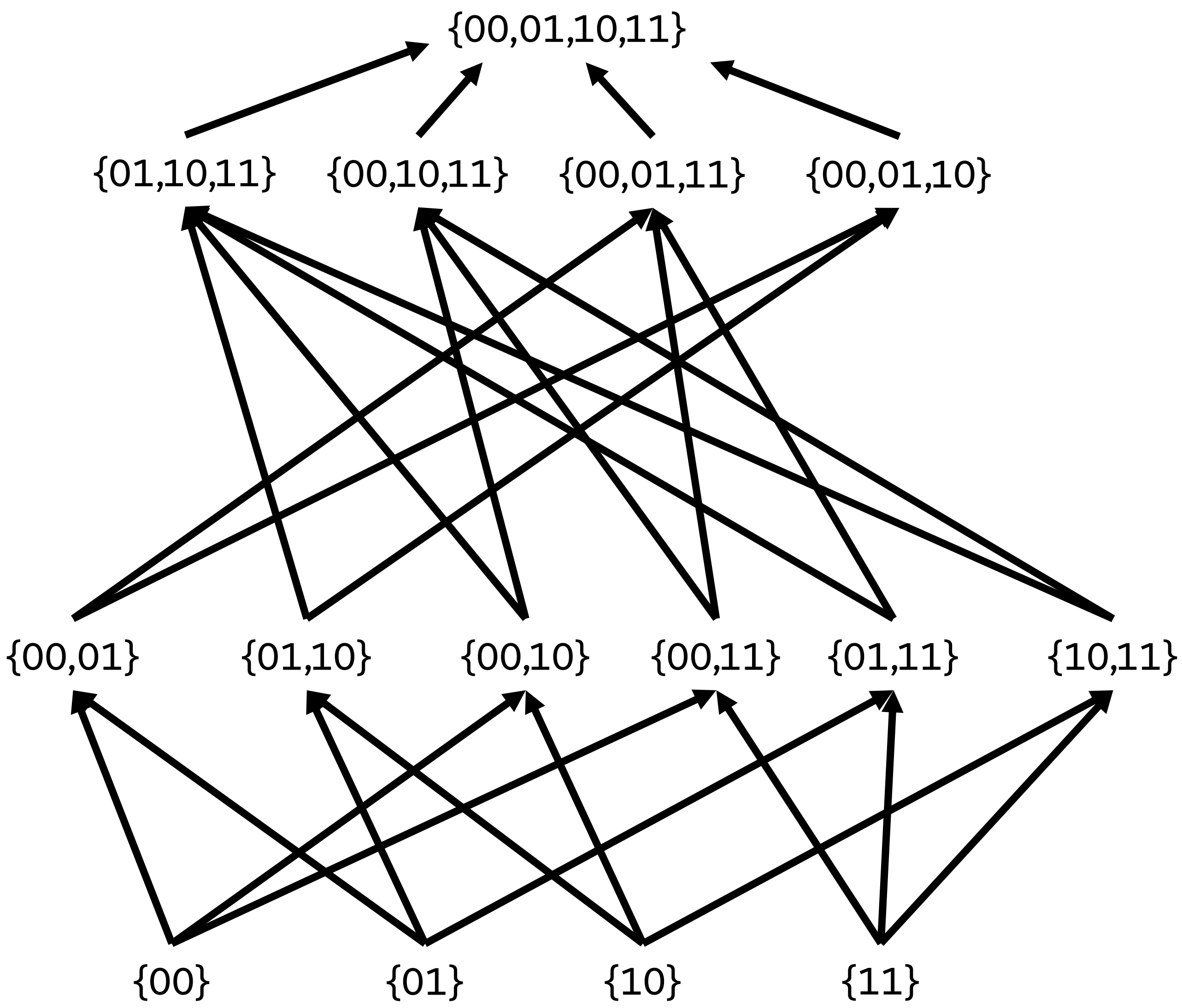
Claims

Bold Claims

FORMALIZING TERMINATION

DATAFLOW FRAMEWORKS

Value-Set "Rank"
(2-bit computer)



DOMAIN NEEDS

DATAFLOW FRAMEWORKS

SOME BASIC DEFINITIONS

A **partially-ordered set** (poset) is a set S and a partial ordering ⊆, such that the ordering ⊆ is:

- Reflexive
- Anti-symmetric
- Transitive

A **lattice** is a poset in which each pair of elements has

- A least upper bound (the *join*)
  - for x and y, the join z is defined such that:
  - x ⊆ z and
  - y ⊆ z and *z is actually an upper bound*
  - for all w such that x ⊆ w and y ⊆ w, w ⊇ z *No upper bound lower than z*
- A greatest lower bound (the *meet*)
  - basically the same deal, but reversed

A **complete lattice** is a lattice in which all subsets have a meet and join

Example 1: S: English words, ⊆ substring

Poset: ✓ Lattice: ✗

Example 2: S: English words, ⊆ shorter or equal in length

Poset: ✗ Lattice: ✗

Example 3: S: integers, ⊆ as lte

Poset: ✓ Lattice: ✓

Example 4: S: integers, ⊆ as lt

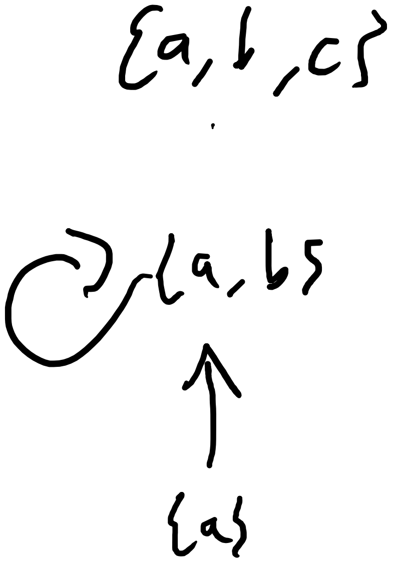
Poset: ✗ Lattice: ✗

Example 5: S: set of all sets of letters, ⊆ is subset

Poset: ✓ Lattice: ✓

FUNCTION NEEDS

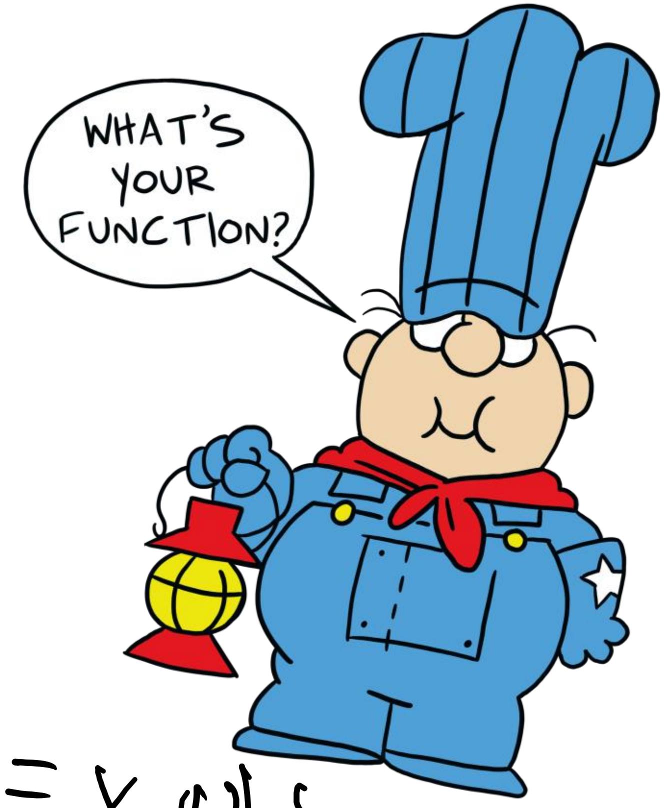
DATAFLOW FRAMEWORKS



Some Basic Definitions

A function f is a **monotonic function** if
$x \subseteq y$ implies $f(x) \subseteq f(y)$

An element z is a **fixpoint** of f iff $z = f(z)$



$f(x) = x \cup \{b\}$

WHY DOES THIS MATTER?

DATAFLOW FRAMEWORKS

Every finite lattice is complete

Practical Upshot

If L is a complete lattice and f is monotonic, then f has a greatest fixpoint and a least fixpoint

If L has no infinite ascending chains, the least fixpoint can be computed by iterative application of f

↑ analysis will terminate

WHY DOES THIS MATTER?

DATAFLOW FRAMEWORKS

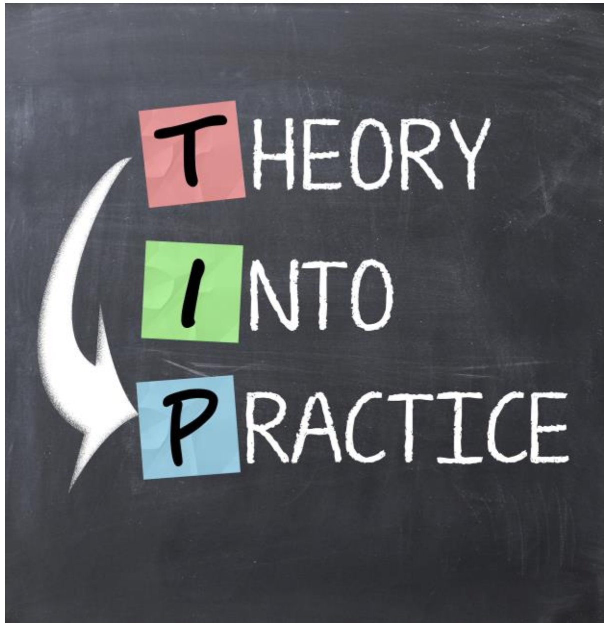
Making the Theory Work for Us

A complete lattice with no infinite chains can be solved via iteration

Catches

Sometimes our domain DOESN'T have these properties

Sometimes the iteration is too lengthy



LECTURE OUTLINE

- Dataflow Frameworks
- Abstract Interpretation

ANALYSIS PRECISION

ABSTRACT INTERPRETATION

Precision / Efficiency Tradeoff

With a complete lattice we can, in theory, eventually terminate

That's not a very strong guarantee!

The shallower the lattice, the faster the fixpoint

Choose to approximate the lattice