CS738: Advanced Compiler Optimizations Foundations of Data Flow Analysis

Amey Karkare

karkare@cse.iitk.ac.in

http://www.cse.iitk.ac.in/~karkare/cs738 Department of CSE, IIT Kanpur



Agenda

- Intraprocedural Data Flow Analysis
 - We looked at 4 classic examples
 - ▶ Today: Mathematical foundations

Categorized along several dimensions

- Categorized along several dimensions
 - the information they are designed to provide

- Categorized along several dimensions
 - the information they are designed to provide
 - the direction of flow

- Categorized along several dimensions
 - the information they are designed to provide
 - the direction of flow
 - confluence operator

- Categorized along several dimensions
 - the information they are designed to provide
 - the direction of flow
 - confluence operator
- Four kinds of dataflow problems, distinguished by

- Categorized along several dimensions
 - the information they are designed to provide
 - the direction of flow
 - confluence operator
- Four kinds of dataflow problems, distinguished by
 - the operator used for confluence or divergence

- Categorized along several dimensions
 - the information they are designed to provide
 - the direction of flow
 - confluence operator
- Four kinds of dataflow problems, distinguished by
 - the operator used for confluence or divergence
 - data flows backward or forward

$\textbf{Confluence} \rightarrow$	U	\cap
Direction \downarrow		
Forward		
Backward		

$\textbf{Confluence} \rightarrow$	U	\bigcap
Direction \downarrow		
Forward	RD	
Backward		

$\textbf{Confluence} \rightarrow$	U	\bigcap
Direction \downarrow		
Forward	R D	Av E
Backward		

$\textbf{Confluence} \rightarrow$	U	\bigcap
Direction \downarrow		
Forward	RD	Av E
Backward	1 \/	

Confluence →	U	\bigcap
Direction \downarrow		
Forward	RD	Av E
Backward	LV	VBE

Why Data Flow Analysis Works?

- Suitable initial values and boundary conditions
- Suitable domain of values
 - Bounded, Finite
- Suitable meet operator
- Suitable flow functions
 - monotonic, closed under composition
- But what is SUITABLE ?

Lattice Theory

Posets

Posets

S: a set

Posets

S: a set

 \leq : a relation

Posets

S: a set

 \leq : a relation

 (S, \leq) is a **poset** if for $x, y, z \in S$

Posets

S: a set

 \leq : a relation

 (S, \leq) is a **poset** if for $x, y, z \in S$

 $ightharpoonup x \le x$ (reflexive)

- Posets
 - S: a set
 - <: a relation
 - (S, \leq) is a **poset** if for $x, y, z \in S$
 - \triangleright $x \le x$ (reflexive)
 - $ightharpoonup x \le y ext{ and } y \le x \Rightarrow x = y ext{ (antisymmetric)}$

- Posets
 - S: a set
 - \leq : a relation
 - (S, \leq) is a **poset** if for $x, y, z \in S$
 - \triangleright $x \le x$ (reflexive)
 - $ightharpoonup x \le y ext{ and } y \le x \Rightarrow x = y ext{ (antisymmetric)}$
 - ▶ $x \le y$ and $y \le z \Rightarrow x \le z$ (transitive)

Linear Ordering

- Linear Ordering
- Poset where every pair of elements is comparable

- Linear Ordering
- Poset where every pair of elements is comparable
- ▶ $x_1 \le x_2 \le ... \le x_k$ is a chain of length k

- Linear Ordering
- Poset where every pair of elements is comparable
- ▶ $x_1 \le x_2 \le ... \le x_k$ is a chain of length k
- We are interested in chains of finite length

Observation

Any finite nonempty subset of a poset has minimal and maximal elements

Observation

- Any finite nonempty subset of a poset has minimal and maximal elements
- Any finite nonempty chain has unique minimum and maximum elements

▶ Set S and meet ∧

- ▶ Set S and meet ∧
- $ightharpoonup x, y, z \in S$

- ▶ Set S and meet ∧
- \triangleright $x, y, z \in S$
 - $ightharpoonup x \land x = x$ (idempotent)

- ▶ Set S and meet ∧
- \triangleright $x, y, z \in S$
 - \triangleright $x \land x = x$ (idempotent)
 - \triangleright $x \land y = y \land x$ (commutative)

- ▶ Set S and meet ∧
- \triangleright $x, y, z \in S$
 - \triangleright $x \land x = x$ (idempotent)
 - \triangleright $x \land y = y \land x$ (commutative)
 - $ightharpoonup x \wedge (y \wedge z) = (x \wedge y) \wedge z$ (associative)

- ▶ Set S and meet ∧
- \triangleright $x, y, z \in S$
 - \triangleright $x \land x = x$ (idempotent)
 - \triangleright $x \land y = y \land x$ (commutative)
 - $ightharpoonup x \wedge (y \wedge z) = (x \wedge y) \wedge z$ (associative)
- Partial order for semilattice

- ▶ Set S and meet ∧
- \triangleright $x, y, z \in S$
 - \triangleright $x \land x = x$ (idempotent)
 - \triangleright $x \land y = y \land x$ (commutative)
 - $ightharpoonup x \wedge (y \wedge z) = (x \wedge y) \wedge z$ (associative)
- Partial order for semilattice
 - $ightharpoonup x \le y$ if and only if $x \land y = x$

Semilattice

- ▶ Set S and meet ∧
- \triangleright $x, y, z \in S$
 - \triangleright $x \land x = x$ (idempotent)
 - \triangleright $x \land y = y \land x$ (commutative)
 - $ightharpoonup x \wedge (y \wedge z) = (x \wedge y) \wedge z$ (associative)
- Partial order for semilattice
 - \triangleright $x \le y$ if and only if $x \land y = x$
 - Reflexive, antisymmetric, transitive

► Top Element (⊤)

- ► Top Element (⊤)
 - $\blacktriangleright \ \forall x \in \mathcal{S}, x \land \top = \top \land x = x$

- ► Top Element (⊤)
 - $\forall x \in S, x \land \top = \top \land x = x$
- ► (Optional) Bottom Element (⊥)

- ► Top Element (⊤)
 - $\forall x \in S, x \land \top = \top \land x = x$
- ► (Optional) Bottom Element (⊥)
 - $\forall x \in \mathcal{S}, x \land \bot = \bot \land x = \bot$

▶ Powerset for a set S, 2^S

- ▶ Powerset for a set S, 2^S
- ► Meet ∧ is ∩

- ▶ Powerset for a set S, 2^S
- ► Meet ∧ is ∩
- ▶ Partial Order is ⊆

- ▶ Powerset for a set S, 2^S
- ► Meet ∧ is ∩
- ▶ Partial Order is ⊆
- ► Top element is *S*

- ▶ Powerset for a set S, 2^S
- ► Meet ∧ is ∩
- ▶ Partial Order is ⊆
- ► Top element is *S*
- ▶ Bottom element is ∅

▶ Powerset for a set S, 2^S

- ► Powerset for a set S, 2^S
- ► Meet ∧ is ∪

- ► Powerset for a set S, 2^S
- ► Meet ∧ is ∪
- ▶ Partial Order is ⊇

- ▶ Powerset for a set S, 2^S
- ► Meet ∧ is ∪
- ▶ Partial Order is ⊇
- ▶ Top element is ∅

- ► Powerset for a set S, 2^S
- ► Meet ∧ is ∪
- ► Partial Order is ⊃
- ► Top element is ∅
- ▶ Bottom element is S

 \triangleright $x, y, z \in S$

- \triangleright $x, y, z \in S$
- glb of x and y is an element g such that

- \triangleright $x, y, z \in S$
- glb of x and y is an element g such that
 - ▶ $g \le x$

- \triangleright $x, y, z \in S$
- glb of x and y is an element g such that
 - ▶ $g \le x$
 - ▶ $g \le y$

- \triangleright $x, y, z \in S$
- glb of x and y is an element g such that
 - ▶ $g \le x$
 - $\triangleright g \leq y$
 - ▶ if $z \le x$ and $z \le y$ then $z \le g$

 $ightharpoonup x, y \in S$

- $ightharpoonup x, y \in S$
- \triangleright (S, \land) is a semilattice

QQ

- \triangleright $x, y \in S$
- \triangleright (S, \land) is a semilattice
- ▶ Prove that $x \land y$ is glb of x and y.

We can define symmetric concepts

- ▶ We can define symmetric concepts
 - ► ≥ order

- ▶ We can define symmetric concepts
 - ► ≥ order
 - ▶ Join operation (\(\forall\))

- ▶ We can define symmetric concepts
 - ► ≥ order
 - ▶ Join operation (\/)
 - Least upper bound (lub)

 $ightharpoonup (S, \land, \lor)$ is a lattice

 (S, \land, \lor) is a lattice iff for each **non-empty finite** subset Y of S

► (S, \(\lambda\), \(\forall\) is a lattice
iff for each non-empty finite subset Y of S
both \(\lambda\) Y and \(\forall\) Y are in S.

- ► (S, \(\lambda\), \(\forall\) is a lattice
 iff for each non-empty finite subset Y of S
 both \(\lambda\) Y and \(\forall\) Y are in S.
- \triangleright (S, \land, \lor) is a complete lattice

- ► (S, \(\lambda\), \(\forall\) is a lattice
 iff for each non-empty finite subset Y of S
 both \(\lambda\) Y and \(\forall\) Y are in S.
- \triangleright (S, \land, \lor) is a complete lattice iff for each subset Y of S

- ► (S, \(\lambda\), \(\forall\) is a lattice
 iff for each non-empty finite subset Y of S
 both \(\lambda\) Y and \(\forall\) Y are in S.
- ► (S, \land, \lor) is a complete lattice iff for each subset Y of S both $\land Y$ and $\lor Y$ are in S.

▶ Complete lattice (S, \land, \lor)

- ▶ Complete lattice (S, \land, \lor)
 - For every pair of elements x and y, both $x \land y$ and $x \lor y$ should be in S

- ightharpoonup Complete lattice (S, \land, \lor)
 - For every pair of elements x and y, both $x \land y$ and $x \lor y$ should be in S
 - Example : Powerset lattice

Lattice

- ▶ Complete lattice (S, \land, \lor)
 - For every pair of elements x and y, both $x \wedge y$ and $x \vee y$ should be in S
 - Example : Powerset lattice
- ► We will talk about **meet** semi-lattices only

Lattice

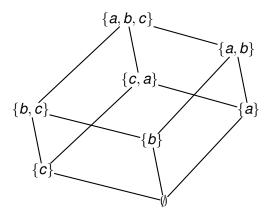
- ▶ Complete lattice (S, \land, \lor)
 - For every pair of elements x and y, both $x \wedge y$ and $x \vee y$ should be in S
 - ► Example : Powerset lattice
- We will talk about meet semi-lattices only
 - except for some proofs

Graphical view of posets

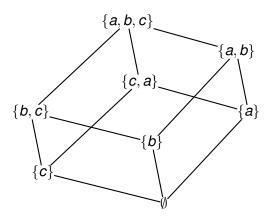
- Graphical view of posets
- ► Elements = the nodes in the graph

- Graphical view of posets
- ► Elements = the nodes in the graph
- ▶ If *x* < *y* then *x* is depicted lower than *y* in the diagram

- Graphical view of posets
- Elements = the nodes in the graph
- ▶ If x < y then x is depicted lower than y in the diagram</p>
- An edge between x and y (x lower than y) implies x < y and no other element z exists s.t. x < z < y (i.e. transitivity is excluded)



Lattice Diagram for $(\{a,b,c\},\cap)$



Lattice Diagram for $(\{a, b, c\}, \cap)$

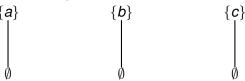
 $x \land y$ = the highest z for which there are paths downward from both x and y.

What if there is a large number of elements?

Combine simple lattices to build a complex one

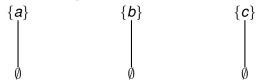
What if there is a large number of elements?

- Combine simple lattices to build a complex one
- Superset lattices for singletons



What if there is a large number of elements?

- Combine simple lattices to build a complex one
- Superset lattices for singletons



Combine to form superset lattice for multi-element sets

 \blacktriangleright (S, \land) is product lattice of (S_1, \land_1) and (S_2, \land_2) when

• (S, \bigwedge) is product lattice of (S_1, \bigwedge_1) and (S_2, \bigwedge_2) when $S = S_1 \times S_2$ (domain)

• (S, \bigwedge) is product lattice of (S_1, \bigwedge_1) and (S_2, \bigwedge_2) when $S = S_1 \times S_2$ (domain) For (a_1, a_2) and $(b_1, b_2) \in S$

► (S, \bigwedge) is product lattice of (S_1, \bigwedge_1) and (S_2, \bigwedge_2) when $S = S_1 \times S_2$ (domain) For (a_1, a_2) and $(b_1, b_2) \in S$ $(a_1, a_2) \bigwedge (b_1, b_2) = (a_1 \bigwedge_1 b_1, a_2 \bigwedge_2 b_2)$

▶ (S, \bigwedge) is product lattice of (S_1, \bigwedge_1) and (S_2, \bigwedge_2) when $S = S_1 \times S_2$ (domain) For (a_1, a_2) and $(b_1, b_2) \in S$ $(a_1, a_2) \bigwedge (b_1, b_2) = (a_1 \bigwedge_1 b_1, a_2 \bigwedge_2 b_2)$ $(a_1, a_2) \leq (b_1, b_2)$ iff $a_1 \leq_1 b_1$ and $a_2 \leq_2 b_2$

▶ (S, \bigwedge) is product lattice of (S_1, \bigwedge_1) and (S_2, \bigwedge_2) when $S = S_1 \times S_2$ (domain) For (a_1, a_2) and $(b_1, b_2) \in S$ $(a_1, a_2) \bigwedge (b_1, b_2) = (a_1 \bigwedge_1 b_1, a_2 \bigwedge_2 b_2)$ $(a_1, a_2) \leq (b_1, b_2)$ iff $a_1 \leq_1 b_1$ and $a_2 \leq_2 b_2$ \leq relation follows from \bigwedge

▶ (S, \bigwedge) is product lattice of (S_1, \bigwedge_1) and (S_2, \bigwedge_2) when $S = S_1 \times S_2$ (domain) For (a_1, a_2) and $(b_1, b_2) \in S$ $(a_1, a_2) \bigwedge (b_1, b_2) = (a_1 \bigwedge_1 b_1, a_2 \bigwedge_2 b_2)$ $(a_1, a_2) \leq (b_1, b_2)$ iff $a_1 \leq_1 b_1$ and $a_2 \leq_2 b_2$ \leq relation follows from \bigwedge

Product of lattices is associative

▶ (S, \bigwedge) is product lattice of (S_1, \bigwedge_1) and (S_2, \bigwedge_2) when $S = S_1 \times S_2$ (domain) For (a_1, a_2) and $(b_1, b_2) \in S$ $(a_1, a_2) \bigwedge (b_1, b_2) = (a_1 \bigwedge_1 b_1, a_2 \bigwedge_2 b_2)$ $(a_1, a_2) \leq (b_1, b_2)$ iff $a_1 \leq_1 b_1$ and $a_2 \leq_2 b_2$ \leq relation follows from \bigwedge

- Product of lattices is associative
- ▶ Can be generalized to product of N > 2 lattices

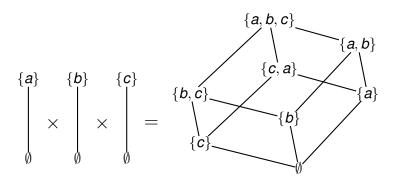
▶ (S, \bigwedge) is product lattice of (S_1, \bigwedge_1) and (S_2, \bigwedge_2) when $S = S_1 \times S_2$ (domain) For (a_1, a_2) and $(b_1, b_2) \in S$ $(a_1, a_2) \bigwedge (b_1, b_2) = (a_1 \bigwedge_1 b_1, a_2 \bigwedge_2 b_2)$ $(a_1, a_2) \leq (b_1, b_2)$ iff $a_1 \leq_1 b_1$ and $a_2 \leq_2 b_2$ \leq relation follows from \bigwedge

- Product of lattices is associative
- ▶ Can be generalized to product of N > 2 lattices
- \triangleright $(S_1, \bigwedge_1), (S_2, \bigwedge_2), \dots$ are called component lattices

Product Lattice: Example

$$\begin{cases}
a \\
b \\
x \\
y
\end{cases}
\times
\begin{cases}
b \\
x \\
y
\end{cases}$$

Product Lattice: Example



Height of a Semilattice

▶ Length of a chain $x_1 \le x_2 \le ... \le x_k$ is k

Height of a Semilattice

- ▶ Length of a chain $x_1 \le x_2 \le ... \le x_k$ is k
- Let $K = \max$ over lengths of all the chains in a semilattice

Height of a Semilattice

- ▶ Length of a chain $x_1 \le x_2 \le ... \le x_k$ is k
- Let $K = \max$ over lengths of all the chains in a semilattice
- ▶ Height of the semilattice = K 1

 \triangleright (D, S, \land, F)

- $ightharpoonup (D, S, \bigwedge, F)$
- D: direction Forward or Backward

- \triangleright (D, S, \land, F)
- ▶ D: direction Forward or Backward
- \triangleright (S, \land): Semilattice Domain and meet

- \triangleright (D, S, \land, F)
- ▶ D: direction Forward or Backward
- \triangleright (S, \land): Semilattice Domain and meet
- F: family of transfer functions of type S → S (see next slide)

▶ F: family of functions $S \rightarrow S$. Must Include

- ▶ F: family of functions $S \rightarrow S$. Must Include
 - functions suitable for the boundary conditions (constant transfer functions for *Entry* and *Exit* nodes)

- ▶ F: family of functions $S \rightarrow S$. Must Include
 - functions suitable for the boundary conditions (constant transfer functions for *Entry* and *Exit* nodes)
 - ▶ Identity function *I*:

$$I(x) = x \quad \forall x \in S$$

- ▶ F: family of functions $S \rightarrow S$. Must Include
 - functions suitable for the boundary conditions (constant transfer functions for *Entry* and *Exit* nodes)
 - ▶ Identity function *I*:

$$I(x) = x \quad \forall x \in S$$

Closed under composition:

$$f,g\in F,\quad f\circ g\quad \Rightarrow\quad h\in F$$

Monotonic Functions

 \triangleright (S, \leq): a poset

Monotonic Functions

- \triangleright (S, \leq): a poset
- ▶ $f: S \rightarrow S$ is monotonic iff

$$\forall x, y \in S \quad x \leq y \Rightarrow f(x) \leq f(y)$$

Monotonic Functions

- \triangleright (S, \leq): a poset
- ▶ $f: S \rightarrow S$ is monotonic iff

$$\forall x, y \in S \quad x \leq y \Rightarrow f(x) \leq f(y)$$

Composition preserves monotonicity

Monotonic Functions

- \triangleright (S, \leq): a poset
- ▶ $f: S \rightarrow S$ is monotonic iff

$$\forall x, y \in S \quad x \leq y \Rightarrow f(x) \leq f(y)$$

- Composition preserves monotonicity
 - ▶ If f and g are monotonic, $h = f \circ g$, then h is also monotonic

Monotone Frameworks

▶ (D, S, \land, F) is monotone if the family F consists of monotonic functions only

$$f \in F$$
, $\forall x, y \in S$ $x \le y \Rightarrow f(x) \le f(y)$

Monotone Frameworks

 \triangleright (D, S, \land, F) is monotone if the family F consists of monotonic functions only

$$f \in F$$
, $\forall x, y \in S$ $x \leq y \Rightarrow f(x) \leq f(y)$

Equivalently

$$f \in F$$
, $\forall x, y \in S$ $f(x \land y) \leq f(x) \land f(y)$

Monotone Frameworks

▶ (D, S, \land, F) is monotone if the family F consists of monotonic functions only

$$f \in F$$
, $\forall x, y \in S$ $x \leq y \Rightarrow f(x) \leq f(y)$

Equivalently

$$f \in F$$
, $\forall x, y \in S$ $f(x \land y) \leq f(x) \land f(y)$

Proof? : QQ in class



Let f be a monotonic function on a complete lattice (S, \land, \lor) . Define

- Let f be a monotonic function on a complete lattice (S, \land, \lor) . Define
 - ▶ $red(f) = \{v \mid v \in S, f(v) \le v\}$, pre fix-points

- Let f be a monotonic function on a complete lattice (S, \land, \lor) . Define
 - ▶ $red(f) = \{v \mid v \in S, f(v) \le v\}$, pre fix-points
 - ightharpoonup ext $(f) = \{v \mid v \in S, f(v) \ge v\}$, post fix-points

- Let f be a monotonic function on a complete lattice (S, \land, \lor) . Define
 - ▶ $red(f) = \{v \mid v \in S, f(v) \le v\}$, pre fix-points
 - ightharpoonup ext $(f) = \{v \mid v \in S, f(v) \ge v\}$, post fix-points
 - $fix(f) = \{v \mid v \in S, f(v) = v\}$, fix-points

- Let f be a monotonic function on a complete lattice (S, \land, \lor) . Define
 - ▶ $red(f) = \{v \mid v \in S, f(v) \le v\}$, pre fix-points
 - ▶ $ext(f) = \{v \mid v \in S, f(v) \ge v\}$, post fix-points
 - $fix(f) = \{v \mid v \in S, f(v) = v\}$, fix-points

Then,

▶ $\bigwedge \operatorname{red}(f) \in \operatorname{fix}(f)$. Further, $\bigwedge \operatorname{red}(f) = \bigwedge \operatorname{fix}(f)$

- Let f be a monotonic function on a complete lattice (S, \land, \lor) . Define
 - ▶ $red(f) = \{v \mid v \in S, f(v) \le v\}$, pre fix-points
 - ▶ $ext(f) = \{v \mid v \in S, f(v) \ge v\}$, post fix-points
 - $fix(f) = \{v \mid v \in S, f(v) = v\}$, fix-points

- $ightharpoonup \land \operatorname{red}(f) \in \operatorname{fix}(f)$. Further, $\land \operatorname{red}(f) = \land \operatorname{fix}(f)$
- ▶ $\bigvee \text{ext}(f) \in \text{fix}(f)$. Further, $\bigvee \text{ext}(f) = \bigvee \text{fix}(f)$

- Let f be a monotonic function on a complete lattice (S, \land, \lor) . Define
 - ▶ $red(f) = \{v \mid v \in S, f(v) \le v\}$, pre fix-points
 - ▶ $ext(f) = \{v \mid v \in S, f(v) \ge v\}$, post fix-points
 - $fix(f) = \{v \mid v \in S, f(v) = v\}$, fix-points

- ▶ $\bigwedge \operatorname{red}(f) \in \operatorname{fix}(f)$. Further, $\bigwedge \operatorname{red}(f) = \bigwedge \operatorname{fix}(f)$
- $ightharpoonup \bigvee \operatorname{ext}(f) \in \operatorname{fix}(f)$. Further, $\bigvee \operatorname{ext}(f) = \bigvee \operatorname{fix}(f)$
- fix(f) is a complete lattice

▶ $f: S \rightarrow S$ is a **monotonic** function

- ▶ $f: S \rightarrow S$ is a **monotonic** function
- $ightharpoonup (S, \wedge)$ is a **finite height** semilattice

- ▶ $f: S \rightarrow S$ is a **monotonic** function
- \triangleright (S, \land) is a **finite height** semilattice
- ▶ \top is the top element of (S, \land)

- ▶ $f: S \rightarrow S$ is a **monotonic** function
- \triangleright (S, \land) is a **finite height** semilattice
- ightharpoonup T is the top element of (S, \land)
- ► Notation: $f^0(x) = x, f^{i+1}(x) = f(f^i(x)), \forall i \ge 0$

- ▶ $f: S \rightarrow S$ is a **monotonic** function
- \triangleright (S, \land) is a **finite height** semilattice
- ightharpoonup T is the top element of (S, \land)
- ► Notation: $f^0(x) = x, f^{i+1}(x) = f(f^i(x)), \forall i \ge 0$
- The greatest fixed point of f is

$$f^k(\top)$$
, where $f^{k+1}(\top) = f^k(\top)$

// monotonic function f on a meet semilattice

```
// monotonic function f on a meet semilattice x := \top;
```

```
// monotonic function f on a meet semilattice x := T; while (x \neq f(x)) x := f(x);
```

```
// monotonic function f on a meet semilattice x := T; while (x \neq f(x)) \times := f(x); return x;
```