Introduction to Lattice Theory

Ashok Sreenivas

Tata Research Design and Development Centre

Thanks to Prof. Chris Hankin of Imperial College and Dr. G. Ramalingam of IBM T.J. Watson Research Centre

Approximations and correctness

Consider the problem of identifying the set of initialized variables at a point in the program. Assume that the precise set of initialized variables at a point is $\{a, b, c\}$.

Approximations and correctness

Consider the problem of identifying the set of initialized variables at a point in the program. Assume that the precise set of initialized variables at a point is $\{a, b, c\}$.

The 'guaranteed' version

- Solution $\{a, b\}$ is approximate and correct.
- Solution $\{a, b, c, d\}$ is *incorrect*.
- So, the 'fastest but most useless' analysis for this problem is one that returns $\{\}$ for all points.

Approximations and correctness

Consider the problem of identifying the set of initialized variables at a point in the program. Assume that the precise set of initialized variables at a point is $\{a, b, c\}$.

The 'guaranteed' version

- Solution $\{a, b\}$ is approximate and correct.
- Solution $\{a, b, c, d\}$ is *incorrect*.
- So, the 'fastest but most useless' analysis for this problem is one that returns $\{\}$ for all points.

The 'may be' version

- Solution $\{a, b, c, d\}$ is approximate and correct.
- Solution $\{a, b\}$ is *incorrect*.
- So, the 'fastest but most useless' analysis for this problem returns the universal set.

Sets and relations

A (binary) relation $\mathcal{R}$ between sets S_1 and S_2 is just a subset of $S_1 \times S_2$. Similarly, any subset $\mathcal{R}$ of $S_1 \times S_2$ is a relation between sets S_1 and S_2. That is, for any sets S_1, S_2, $\mathcal{R}$ is a relation between S_1 and S_2 *iff*

$$\mathcal{R} \subseteq S_1 \times S_2$$

If $(s_1, s_2) \in \mathcal{R}$, we also write $s_1 \mathcal{R} s_2$
Let $S_1 = \{a, b, c, d\}$, $S_2 = \{1, 2, 3\}$
$\mathcal{R}_1 = \{(a, 1), (b, 2), (c, 3), (d, 1)\}$?
$\mathcal{R}_2 = \{(a, 1), (b, 1), (c, 1), (c, 2), (c, 3)\}$?
$\mathcal{R}_3 = \{\}$?
$\mathcal{R}_4 = \{(a, a), (b, b), (c, c)\}$?
$\mathcal{R}_5 = \{(a, x), (b, y)\}$?

Some ‘real’ relations

$S_1 = \{$Rajeev, Sanjay, Bhim, Mohandas, Duryodhan$\}$
$S_2 = \{$Pandu, Gandhari, Indira, Feroz, Bhishma$\}$
$\texttt{Child} = \{(\text{Rajeev}, \text{Indira}), (\text{Rajeev}, \text{Feroz}),$
$(\text{Sanjay}, \text{Indira}), (\text{Sanjay}, \text{Feroz}),$
$(\text{Bhim}, \text{Pandu}), (\text{Duryodhan}, \text{Gandhari})\}$

$Z = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$
$<= \{\cdots, (-2, -1), (-2, 0), (-2, 1), (-2, 2), (-1, 0), \cdots\}$
$\texttt{double} = \{\cdots, (-1, -2), (0, 0), (1, 2), (2, 4) \cdots\}$

- Relations can be $N - N$, $1 - N$, $N - 1$ or $1 - 1$.
- Relations can be ‘total’ (all of S_1) or ‘onto’ (all of S_2)
- Functions are just $N - 1$ relations!

Kinds of relations

A relation $\mathcal{R}$ from S to S is

- *reflexive iff* $\forall a \in S.\ (a, a) \in \mathcal{R}$. Example: $\leq$ is reflexive, but $<$ is not.
- *symmetric iff* $(a, b) \in \mathcal{R} \Rightarrow (b, a) \in \mathcal{R}$. Example: $\neq$ and $=$ are symmetric, but $\leq$ is not.
- *anti-symmetric iff* $a \neq b \wedge (a, b) \in \mathcal{R} \Rightarrow (b, a) \notin \mathcal{R}$. Example: $\leq$ is anti-symmetric, while $\neq$, $=$ are not.
- *transitive iff* $(a, b) \in \mathcal{R} \wedge (b, c) \in \mathcal{R} \Rightarrow (a, c) \in \mathcal{R}$. Example: $<$, $\leq$, `Ancestor` are transitive, but $\neq$, `Parent` are not.

Note: $\mathcal{R}$ is symmetric and transitive $\not\Rightarrow$ $\mathcal{R}$ is reflexive! Example: the empty relation!!

Equivalence relations

Relation $\equiv$ from S to S is an equivalence relation *iff* it is reflexive, symmetric and transitive. $\equiv$ *partitions* S into disjoint subsets (equivalence classes) where all elements of each sub-set are $\equiv$-related to each other, and no two elements across the subsets are $\equiv$-related. Examples:

- $=$ partitions Z into an infinite number of singleton equivalence classes: $\{\cdots, \{-1\}, \{0\}, \{1\}, \cdots\}$.
- $=_{mod\ n}$ partitions N into n infinitely large equivalence classes: $\{\{0, n, 2n, \cdots\}, \{1, n+1, 2n+1, \cdots\} \cdots \{n-1, 2n-1, 3n-1 \cdots\}\}$.
- `sibling` partitions the entire human population into equivalence classes, where a `sibling` b *iff* a and b have *both* parents in common.

Partial orders

Relation $\sqsubseteq$ from S to S is a *partial order iff* it is reflexive, anti-symmetric and transitive. Note that there may be elements a, b in S such that neither $a \sqsubseteq b$ nor $b \sqsubseteq a$. If either $a \sqsubseteq b$ or $b \sqsubseteq a$ for all a, b, then $\sqsubseteq$ is called a *total* order. Examples:

- $\leq$ is a total order over Z.
- Relation $\subseteq$ is a partial order over any set S of sets. For all sets A, B, C: $A \subseteq A$; $A \subseteq B \wedge A \neq B \Rightarrow B \not\subseteq A$; and $A \subseteq B \wedge B \subseteq C \Rightarrow A \subseteq C$. And of course, there exist sets S_1, S_2 such that neither $S_1 \subseteq S_2$ nor $S_2 \subseteq S_1$.
- Relation $|$ (divides) is a partial order over N. For any natural numbers a, b, c, $a|a$; $a|b \wedge a \neq b \Rightarrow b \not| a$; and $a|b \wedge b|c \Rightarrow a|c$.

Approximations

Partial orders capture the notion of approximations!

Approximations

Partial orders capture the notion of approximations!

- Everything approximates itself (perfectly!).
- a approximates b and $a \neq b \Rightarrow b$ definitely does not approximate a.
- a approximates b and b approximates $c \Rightarrow a$ approximates c. In such cases, b is a ‘more precise’ approximation of c than a.
- There will usually be some a and b such neither of them approximates the other.

Approximations

Partial orders capture the notion of approximations!

- Everything approximates itself (perfectly!).
- a approximates b and $a \neq b \Rightarrow b$ definitely does not approximate a.
- a approximates b and b approximates $c \Rightarrow a$ approximates c. In such cases, b is a 'more precise' approximation of c than a.
- There will usually be some a and b such neither of them approximates the other.

In analysis $a \sqsubseteq b$ is usually defined such that a is more precise than b.

Approximations

Partial orders capture the notion of approximations!

- Everything approximates itself (perfectly!).
- a approximates b and $a \neq b \Rightarrow b$ definitely does not approximate a.
- a approximates b and b approximates $c \Rightarrow a$ approximates c. In such cases, b is a ‘more precise’ approximation of c than a.
- There will usually be some a and b such neither of them approximates the other.

In analysis $a \sqsubseteq b$ is usually defined such that a is more precise than b.
Example: $3 \sqsupset 3.1 \sqsupset 3.14 \sqsupset 3.141 \sqsupset \cdots \sqsupset \pi$.

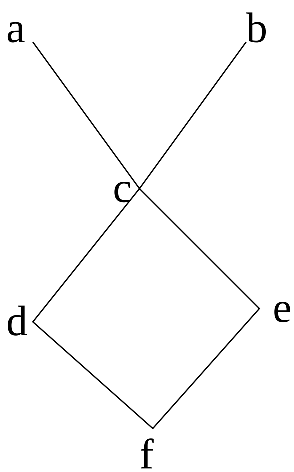
Exercises

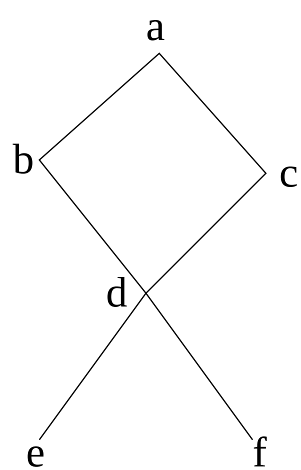
- Which of the following are reflexive, symmetric, anti-symmetric, equivalence relations, partial orders?
  - $\{(a,a),(b,b),(c,c)\}$ over set $S = \{a,b,c\}$
  - $\{(a,a),(b,b),(c,c),(a,b),(b,a)\}$ over set $S = \{a,b,c\}$
  - $\mathcal{R} = \{\}$ over any set S.
  - $\mathcal{R} = S \times S$ over any set S.
  - $f(x) = x^2$
  - $f(x) = -x$
  - $\mathcal{R}$ is defined over S, the set of functions from N to N, as $f\mathcal{R}g$ *iff* $\forall x. f(x)|g(x)$.
- Give examples of other 'real' equivalence classes and partial orders.

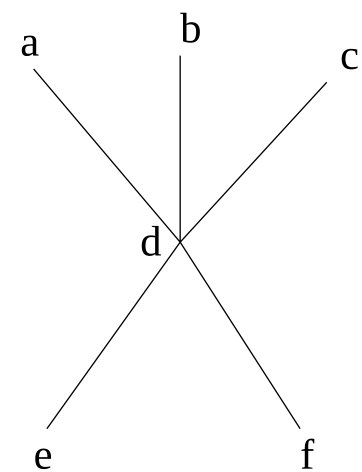
Posets

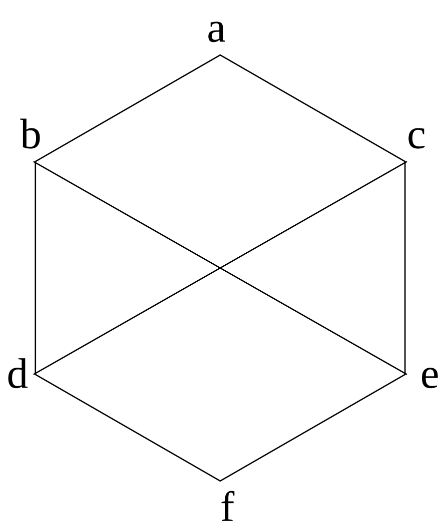
- If partial ordering $\sqsubseteq$ is defined over set S, then $(S, \sqsubseteq)$ is called a *partially-ordered set* or a *poset*.
- $b \sqsupseteq a$ is the same as $a \sqsubseteq b$. Note that if $\sqsubseteq$ is a partial order, then so is $\sqsupseteq$.
- $a \sqsubset b$ is the same as $a \sqsubseteq b$ and $a \neq b$. $\sqsubset$ is *not* a partial order. Similarly $\sqsupset$.
- Element b is a *minimal* element or *lower bound* of poset $(S, \sqsubseteq)$ *iff* $\forall x \in S . x \not\sqsubset b$.
- Similarly t is a *maximal* element or *upper bound* of the poset *iff* $\forall x \in S . x \not\sqsupset t$. Note: Minimal and maximal elements may not be unique for a poset.

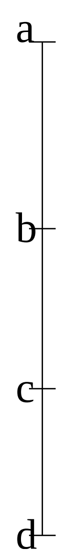
Maximal and minimal elements

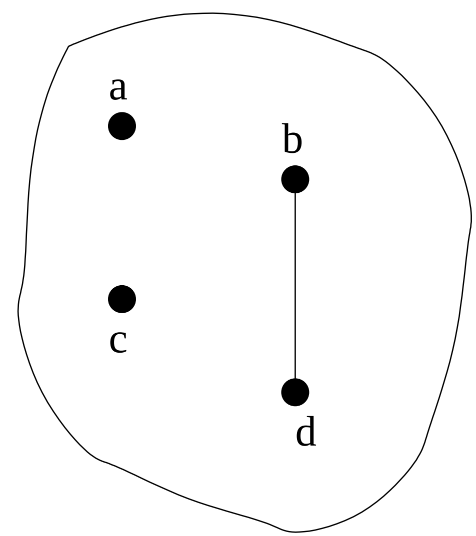










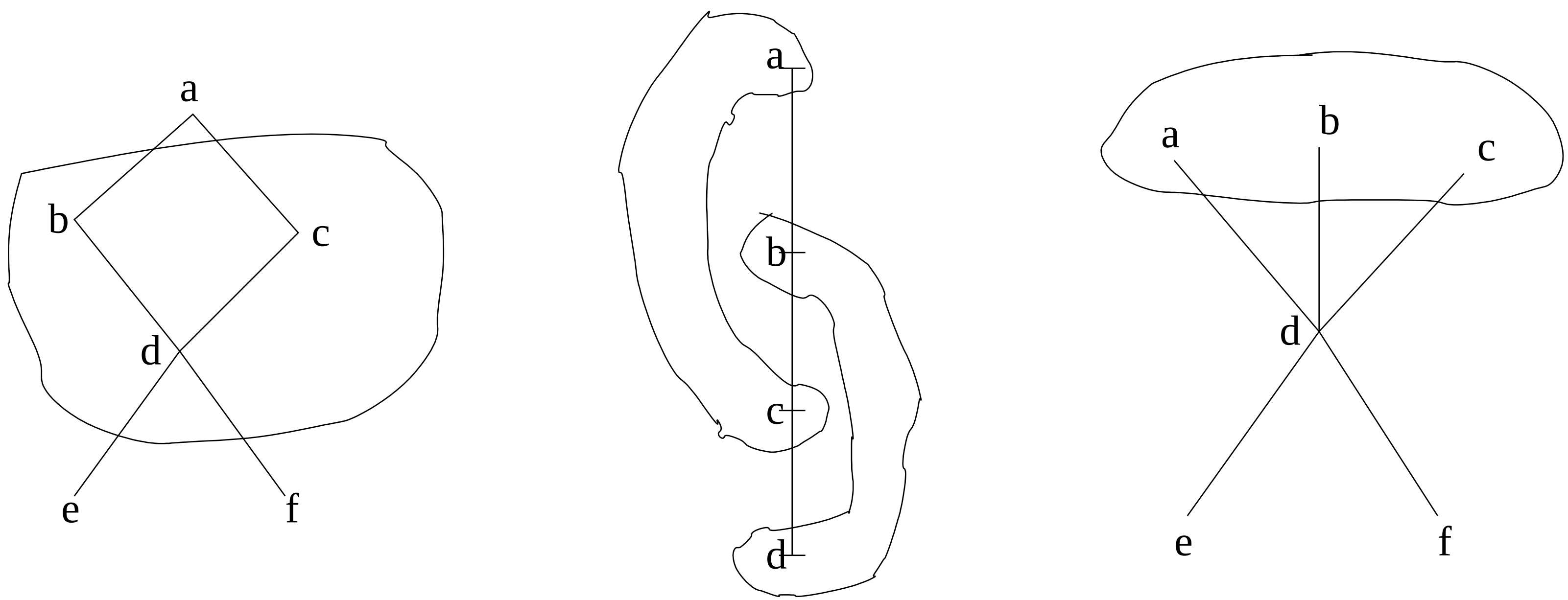


Lower and upper bounds

Consider a poset $(L, \sqsubseteq)$ and a subset Y of L. Element $l \in L$ is an upper bound of Y if $\forall l' \in Y . l' \sqsubseteq l$. Similarly, l is a lower bound if $\forall l' \in Y . l' \sqsupseteq l$. Note: l may not belong to the subset Y. l may not be unique, i.e. Y may have many (or no) lower and upper bounds.

Lower and upper bounds

Consider a poset $(L, \sqsubseteq)$ and a subset Y of L. Element $l \in L$ is an upper bound of Y if $\forall l' \in Y.l' \sqsubseteq l$. Similarly, l is a lower bound if $\forall l' \in Y.l' \sqsupseteq l$. Note: l may not belong to the subset Y. l may not be unique, i.e. Y may have many (or no) lower and upper bounds.



LUBs and GLBs

- $l \in L$ is a least upper bound (LUB) of a subset Y *iff* l is an upper bound of Y and $l \sqsubseteq l'$ for all other upper bounds l' of Y.
- Similarly, greatest lower bound (GLB) of a subset Y is a lower bound that is $\sqsupseteq$ all other lower bounds.
- LUB of a set Y is denoted as $\bigsqcup Y$ and is also called the *join* operator.
- GLB of a set Y is denoted as $\bigsqcap Y$ and is also called the *meet* operator.
- A subset Y may be such that $\bigsqcup Y$ or $\bigsqcap Y$ do not exist. But if they exist, they are *unique*.
- $\bigsqcup\{y_1, y_2\}$ is also written $y_1 \sqcup y_2$. Similarly $\bigsqcap\{y_1, y_2\} = y_1 \sqcap y_2$.

Lattices

- A poset $(L, \sqsubseteq)$ is a complete lattice $(L, \sqsubseteq, \bigsqcup, \bigsqcap, \bot, \top)$ *iff* all subsets Y of L have greatest lower bounds as well as least upper bounds.
- $\bot = \bigsqcup \phi = \bigsqcap L$ is the *least element* of L.
- $\top = \bigsqcap \phi = \bigsqcup L$ is the *greatest element* of L.
- $x \sqcup \top = \top$, $x \sqcup \bot = x$
- $x \sqcap \bot = \bot$, $x \sqcap \top = x$

More on lattices

- For any set S, $(2^S, \subseteq, \bigcup, \bigcap, \phi, S)$ is a complete lattice.
- Is $(Z, \leq)$ a complete lattice? And what about $(N, \leq)$?

More on lattices

- For any set S, $(2^S, \subseteq, \bigcup, \bigcap, \phi, S)$ is a complete lattice.
- Is $(Z, \leq)$ a complete lattice? And what about $(N, \leq)$?

The following statements are equivalent:
$x \sqcup y = y$; $x \sqcap y = x$; $x \sqsubseteq y$

More on lattices

- For any set S, $(2^S, \subseteq, \bigcup, \bigcap, \phi, S)$ is a complete lattice.
- Is $(Z, \leq)$ a complete lattice? And what about $(N, \leq)$?

The following statements are equivalent:
$x \sqcup y = y$; $x \sqcap y = x$; $x \sqsubseteq y$

Properties of $\sqcup, \sqcap$:

Idempotence $x \sqcup x = x \sqcap x = x$

Commutativity $x \sqcup y = y \sqcup x$ and $x \sqcap y = y \sqcap x$

Associativity $x \sqcup (y \sqcup z) = (x \sqcup y) \sqcup z$. Similarly for $\sqcap$.

Absorption $x \sqcup (x \sqcap y) = x$

Exercises

Prove the following:

- If $\sqsubseteq$ is a partial order, then so is $\sqsupseteq$.
- When they exist, $\bigsqcup Y$ and $\bigsqcap Y$ are unique for any subset Y of a poset $(L, \sqsubseteq)$.
- $(x \sqcup y = y) \Leftrightarrow (x \sqsubseteq y)$
- The idempotence, commutativity, associativity and absorption properties of $\sqcup$.

Chains

Any totally ordered subset S of poset $(L, \sqsubseteq)$ is called a *chain*.

That is, $\forall l_1, l_2 \in S.\ (l_1 \sqsubseteq l_2) \vee (l_2 \sqsubseteq l_1)$.

Note: An empty subset of L is also a chain!!

A *sequence* of elements $l_1, l_2, \cdots$ is an *ascending chain iff* $i < j \Rightarrow l_i \sqsubseteq l_j$. Similarly, a sequence is a *descending chain iff* $i < j \Rightarrow l_i \sqsupseteq l_j$.

Example ...

Chains $\cdots$

The *height* of a poset $(L, \sqsubseteq)$ is h if the largest chain in the lattice contains $h+1$ elements.

Poset $(L, \sqsubseteq)$ has a finite height *iff* all chains are finite, i.e. all ascending and descending chains are of the form $l_1, l_2, \cdots l_k, l_{k+1}, l_{k+2}, \cdots$ where $l_j = l_k \; \forall j \geq k$.

Obviously finite posets have finite heights!

Examples of infinite posets with finite and infinite heights?

Product lattices

- Given two posets $(L_1, \sqsubseteq_1)$ and $(L_2, \sqsubseteq_2)$, $(L, \sqsubseteq)$ is also a partial order where

$$L = \{(l_1, l_2) | l_1 \in L_1 \wedge l_2 \in L_2\}$$

and

$$(l_{11}, l_{12}) \sqsubseteq (l_{21}, l_{22}) \text{ iff } l_{11} \sqsubseteq_1 l_{21} \wedge l_{12} \sqsubseteq_2 l_{22}$$

- Prove the above!

Product lattices $\cdots$

- If each L_i is a complete lattice, then so is $(L, \sqsubseteq, \bigsqcup, \bigsqcap, \bot, \top)$ as follows:

  $$\bigsqcup Y = (\bigsqcup_1 \{l_1 | \exists l_2 : (l_1, l_2) \in Y\}, \bigsqcup_2 \{l_2 | \exists l_1 : (l_1, l_2) \in Y\}$$

  Similarly for $\bigsqcap$
  $\bot = (\bot_1, \bot_2)$
  $\top = (\top_1, \top_2)$
- L often referred to as $L_1 \times L_2$, the cartesian product of L_1 and L_2.
- Cartesian products can be extended to any number of posets or lattices, i.e. $L_1 \times L_2 \times L_3 \times \cdots \times L_k$

Functions

Consider posets $(L_1, \sqsubseteq_1)$, $(L_2, \sqsubseteq_2)$, and a function $f : L_1 \rightarrow L_2$.

- f is a *monotonic* (or *monotone*) function *iff* $\forall x, y . x \sqsubseteq y \ \Rightarrow f(x) \sqsubseteq f(y)$
- f is a completely *additive* (or *distributive*) function if $\forall Y \subseteq L_1 . f(\bigsqcup_1 Y) = \bigsqcup_2 \{f(l') | l' \in Y\}$ whenever $\bigsqcup_1 Y$ exists.
- Similarly, it is completely *multiplicative* if $\forall Y \subseteq L_1 . f(\bigsqcap_1 Y) = \bigsqcap_2 \{f(l') | l' \in Y\}$ whenever $\bigsqcap_1 Y$ exists.
- A function is *strict* if $f(\bot_1) = \bot_2$. f is completely additive $\Rightarrow f$ is strict.

Functions, fixed points $\cdots$

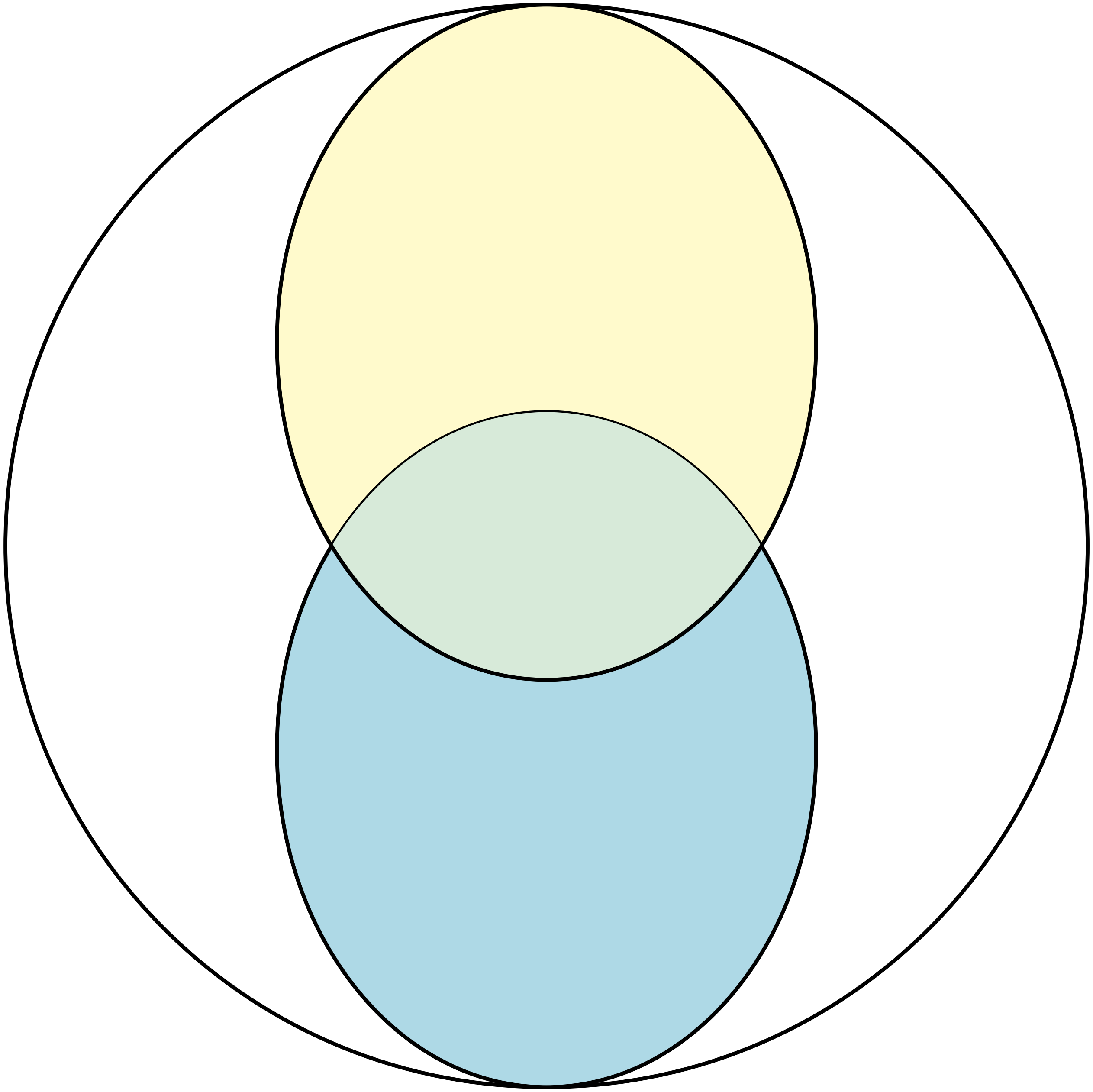
Consider a monotone function $f : L \rightarrow L$ on a complete lattice $(L, \sqsubseteq, \bigsqcup, \bigsqcap, \perp, \top)$.

- A *fixed point* of f is some $l \in L$ such that $f(l) = l$.
- What, if any, are the fixed point(s) of the following functions over Z?
  - $f(x) = x + 3$
  - $f(x) = x^2$
  - $f(x) = x$
  - $f(x) = x!$ (factorial, over N)
- $Fix(f) = \{l \in L | f(l) = l\}$, the set of fixed points of $f : L \rightarrow L$.

Reductive and extensive regions

- $f : L \rightarrow L$ is *reductive* at $l \in L$ if $f(l) \sqsubseteq l$.
  $Red(f) = \{l \in L | f(l) \sqsubseteq l\}$
- $f : L \rightarrow L$ is *extensive* at $l \in L$ if $f(l) \sqsupseteq l$.
  $Ext(f) = \{l \in L | f(l) \sqsupseteq l\}$
- The function f itself is reductive (extensive) if $Red(f) = L$ ($Ext(f) = L$).
- Example: $f(x) = x + 3$ over $(Z, \leq)$ is extensive, while $f(x) = x - 3$ is reductive.

Reductive, extensive regions



Computation of fixed points

For a monotone function f over a complete lattice L:

- Least fixed point $lfp(f) = \bigsqcap Fix(f)$
- Greatest fixed point $gfp(f) = \bigsqcup Fix(f)$

Tarski's theorem: For a complete lattice $(L, \sqsubseteq, \sqcup, \sqcap, \bot, \top)$ and monotone $f : L \rightarrow L$
$lfp(x) = \bigsqcap Red(f) \in Fix(f)$
$gfp(x) = \bigsqcup Ext(f) \in Fix(f)$

- $lfp(f)$ is a fixed point of f that is $\sqsubseteq$ all other fixed points. Similarly, $gfp(f)$.
- If all chains in L are finite, $lfp(f)$ can be computed as the limit of the chain $f^n(\bot)$, i.e. $\bot, f(\bot), f^2(\bot) \cdots$. In other words, $lfp(f) = f^k(\bot)$ such that $f^k(\bot) = f^{k+1}(\bot)$

Analysis and lattices

- Lattices can be used to model approximations
- $a \sqsubseteq b$ means a is *more* precise than b in the semantics and some analysis literature.
- But in data flow analysis literature, it means a is *less* precise than b. This means $\cdots$
- $\sqsubseteq$ and $\sqsupseteq$ are interchanged;
- $\perp$ and $\top$ are interchanged;
- $\sqcup$ and $\sqcap$ are interchanged;
- lfp and gfp are interchanged.
- Basically, the lattice is hung ‘upside down’.

Analysis and fixed points

- Analysis equations of the form:
  $I_o(n) = f_n(I_i(n))$
  $I_i(n) = {f'}_n(I_o(n_1), I_o(n_2) \cdots I_o(n_k))$

Analysis and fixed points

- Analysis equations of the form:
  $I_o(n) = f_n(I_i(n))$
  $I_i(n) = {f'}_n(I_o(n_1), I_o(n_2) \cdots I_o(n_k))$
- If there are k nodes, there are $2k$ pieces of information, namely $I_i(1), I_o(1), I_i(2), I_o(2) \cdots I_i(k), I_o(k)$ defined mutually recursively by the above equations.

Analysis and fixed points

- Analysis equations of the form:
  $I_o(n) = f_n(I_i(n))$
  $I_i(n) = {f'}_n(I_o(n_1), I_o(n_2) \cdots I_o(n_k))$
- If there are k nodes, there are $2k$ pieces of information, namely $I_i(1), I_o(1), I_i(2), I_o(2) \cdots I_i(k), I_o(k)$ defined mutually recursively by the above equations.
- Let I define a tuple of the above $2k$ pieces of information. That is $I = \langle I_i(1), I_o(1), \cdots I_i(k), I_o(k) \rangle$.

Analysis and fixed points

- Analysis equations of the form:
  $I_o(n) = f_n(I_i(n))$
  $I_i(n) = {f'}_n(I_o(n_1), I_o(n_2) \cdots I_o(n_k))$
- If there are k nodes, there are $2k$ pieces of information, namely $I_i(1), I_o(1), I_i(2), I_o(2) \cdots I_i(k), I_o(k)$ defined mutually recursively by the above equations.
- Let I define a tuple of the above $2k$ pieces of information. That is $I = \langle I_i(1), I_o(1), \cdots I_i(k), I_o(k) \rangle$.
- So, the above equations can be together written as:
  $I = F(I)$ where $F(I) = \langle {f'}_1(I), f_1(I), \cdots {f'}_k(I), f_k(I) \rangle$

Analysis and fixed points

- Analysis equations of the form:
  $I_o(n) = f_n(I_i(n))$
  $I_i(n) = f'_n(I_o(n_1), I_o(n_2) \cdots I_o(n_k))$
- If there are k nodes, there are $2k$ pieces of information, namely $I_i(1), I_o(1), I_i(2), I_o(2) \cdots I_i(k), I_o(k)$ defined mutually recursively by the above equations.
- Let I define a tuple of the above $2k$ pieces of information. That is $I = \langle I_i(1), I_o(1), \cdots I_i(k), I_o(k) \rangle$.
- So, the above equations can be together written as: $I = F(I)$ where $F(I) = \langle f'_1(I), f_1(I), \cdots f'_k(I), f_k(I) \rangle$
- So, the answer we seek, namely I is nothing but a fixed point of F! And the most precise answer is the least fixed point of F!!

Home work!

1. Let L be a complete lattice $(L, \sqsubseteq_L, \bigsqcup_L, \bigsqcap_L, \bot_L, \top_L)$. Consider the space of total functions over L, say F. That is, F consists of *all* total functions from L to L. Prove that $(F, \sqsubseteq_F, \bigsqcup_F, \bigsqcap_F, \bot_F, \top_F)$ is also a complete lattice, where:
   $f \sqsubseteq_F g$ *iff* $\forall x \in L. f(x) \sqsubseteq_L g(x)$
   $\forall Y \subseteq F. \bigsqcup_F Y = \lambda x. \bigsqcup_L \{f(x) | f \in Y\}$
   $\forall Y \subseteq F. \bigsqcap_F Y = \lambda x. \bigsqcap_L \{f(x) | f \in Y\}$
   $\bot_F = \lambda x. \bot_L$
   $\top_F = \lambda x. \top_L$

2. Prove that the limit of the chain $\bot, f(\bot), f^2(\bot) \cdots$ for monotone f over a complete lattice with only finite chains is indeed $lfp(f)$.

$f = \lambda x.e$ is the same as $f(x) = e$.