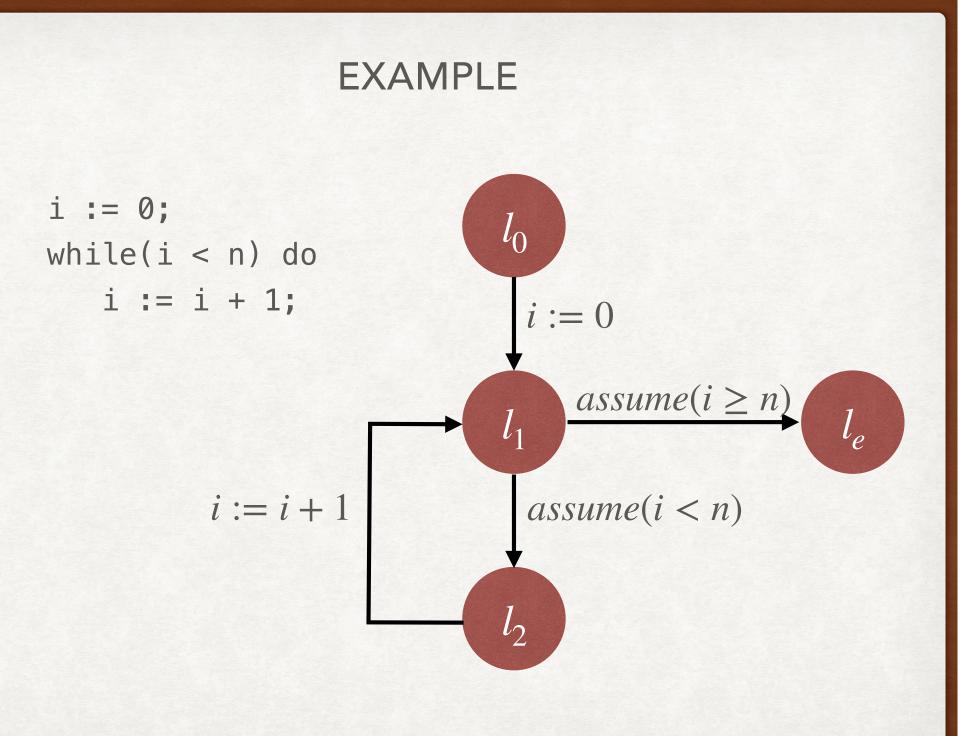
ABSTRACT INTERPRETATION

LABELLED TRANSITION SYSTEM

- We express the program c as a labelled transition system $\Gamma_c \equiv (V,L,l_0,l_e,T)$
 - V is the set of program variables
 - L is the set of program locations
 - l_0 is the start location
 - l_e is the end location
 - $T \subseteq L \times c \times L$ is the set of labelled transitions between locations.



PROGRAMS AS LTS

- There are various ways to construct the LTS of a program
 - We can use control flow graph
 - We can use basic paths as defined by the book (BM Chapter 5). A basic path is a sequence of instructions that begins at the start of the program or a loop head, and ends at a loop head or the end of the program.
- Program State (σ, l) consists of the values of the variables $(\sigma : V \to \mathbb{R})$ and the location.
- An execution is a sequence of program states, $(\sigma_0, l_0), (\sigma_1, l_1), \dots, (\sigma_n, l_n)$, such that for all $i, 0 \le i \le n 1$, $(l_i, c, l_{i+1}) \in T$ and $(\sigma_i, c) \hookrightarrow^* (\sigma_{i+1}, skip)$.
- A program satisfies its specification $\{P\}c\{Q\}$ if $\forall \sigma \in P$, for all executions $(\sigma, l_0), (\sigma_1, l_1), \dots, (\sigma', l_e)$ of $\Gamma_c, \sigma' \in Q$.

INDUCTIVE ASSERTION MAP

• With each location, we associate a set of states which are reachable at that location in any execution.

• $\mu: L \to \Sigma(V)$

 To express that such a map is an inductive assertion map, we will use Strongest Post-condition.

• $\forall (l, c, l') \in T. sp(\mu(l), c) \rightarrow \mu(l')$

• Then, if μ is an inductive assertion map on Γ_c , the Hoare triple $\{P\}c\{Q\}$ is valid if $P \to \mu(l_0)$ and $\mu(l_e) \to Q$.

GENERATING THE INDUCTIVE ASSERTION MAP

• We can express the inductive assertion map as a solution of a system of equations:

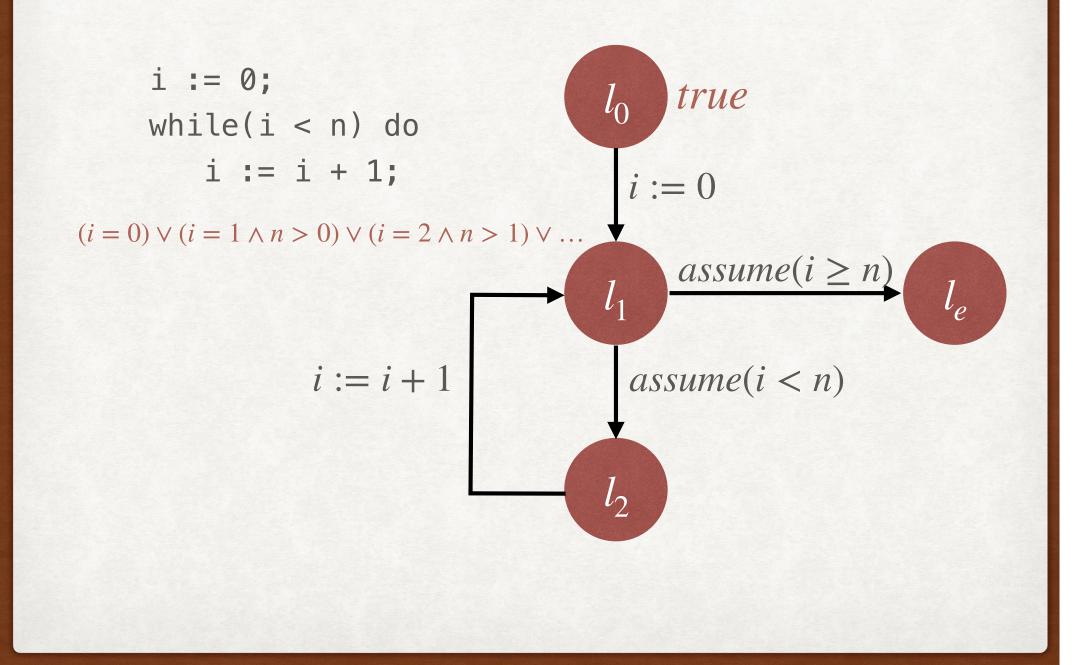
• $X_{l_0} = P$

• For all other locations $l \in L \setminus \{l_0\}, X_l = \bigvee_{(l',c,l) \in T} sp(X_{l'},c)$

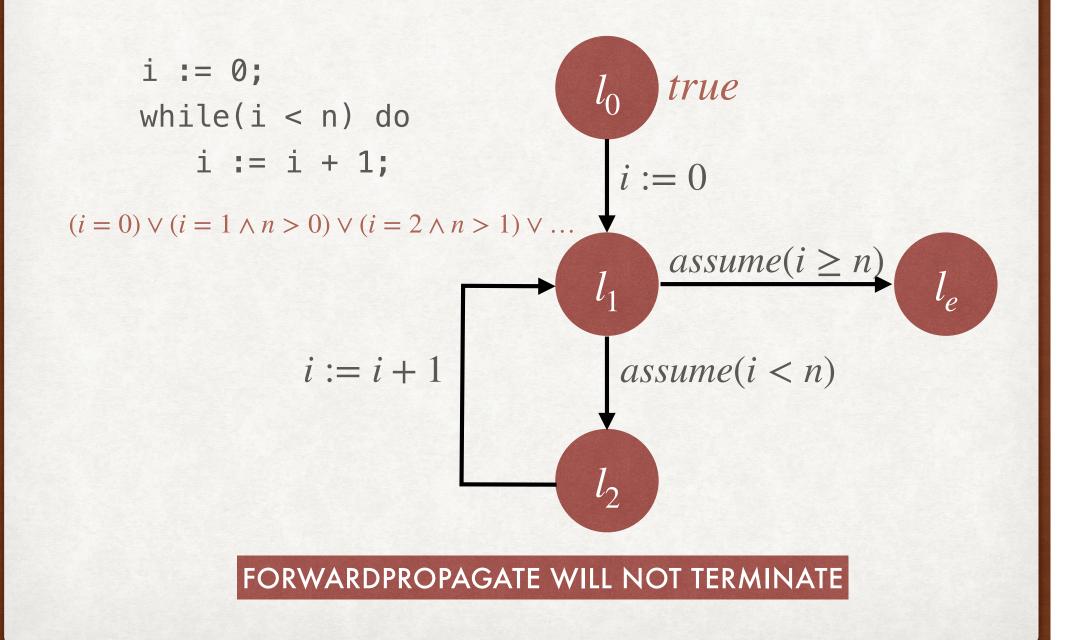
GENERATING THE INDUCTIVE ASSERTION MAP

```
ForwardPropagate(\Gamma_c, P)
S := \{l_0\};
\mu(l_0) := P;
\mu(l) := false, \text{ for } l \in L \setminus \{l_0\};
while S \neq \emptyset do{
     l := Choose S;
     S := S \setminus \{l\};
      foreach (l, c, l') \in T do{
           F := sp(\mu(l), c);
          if \neg(\mathsf{F} \rightarrow \mu(l')) then{
               \mu(l') := \mu(l') \lor F;
               S := S \cup \{l'\};
           }
      }
 }
```

EXAMPLE







ABSTRACT INTERPRETATION: OVERVIEW

- Instead of maintaining an arbitrary set of states at each location, maintain an artificially constrained set of states, coming from an abstract domain D.
 - $\hat{\mu}: L \to D$
- Let *State* $\triangleq V \rightarrow \mathbb{R}$ be the set of all possible concrete states.
 - Abstraction function, $\alpha : \mathbb{P}(State) \rightarrow D$
 - Concretization function, $\gamma : D \rightarrow \mathbb{P}(State)$
- $\hat{\mu}$ over approximates the set of states at every location.
 - For all locations $l, \gamma(\hat{\mu}(l)) \supseteq \mu(l)$
- Use abstract strongest post-condition operator $\hat{sp}: D \times c \rightarrow D$
 - $\gamma(\hat{sp}(d,c)) \supseteq sp(\gamma(d),c)$

GENERATING THE INDUCTIVE ASSERTION MAP

```
ForwardPropagate(\Gamma_c, P)
S := \{l_0\};
\mu(l_0) := P;
\mu(l) := false, \text{ for } l \in L \setminus \{l_0\};
while S \neq \emptyset do{
     l := Choose S;
     S := S \setminus \{l\};
      foreach (l, c, l') \in T do{
           F := sp(\mu(l), c);
          if \neg(\mathsf{F} \rightarrow \mu(l')) then{
               \mu(l') := \mu(l') \lor F;
               S := S \cup \{l'\};
           }
      }
 }
```

ABSTRACT FORWARD PROPAGATE

```
AbstractForwardPropagate(\Gamma_c, P)
 S := \{l_0\};
 \hat{\mu}(l_0) := \alpha(P);
 \hat{\mu}(l) := \bot, for l \in L \setminus \{l_0\};
 while S \neq \emptyset do{
      l := Choose S;
      S := S \setminus \{l\};
       foreach (l, c, l') \in T do{
            F := \hat{sp}(\hat{\mu}(l), c);
            if \neg(\mathsf{F} \leq \hat{\mu}(l')) then{
                 \hat{\mu}(l') := \hat{\mu}(l') \sqcup F;
                 S := S \cup \{l'\};
            }
       }
  }
```

ABSTRACT FORWARD PROPAGATE

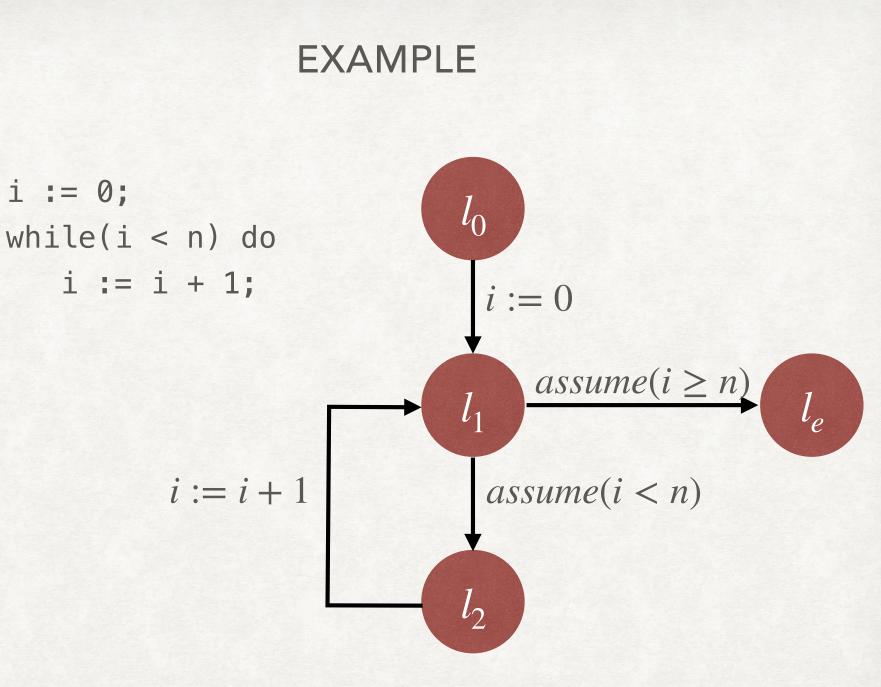
```
AbstractForwardPropagate(\Gamma_c, P)
 S := \{l_0\};
\hat{\mu}(l_0) := \alpha(P);
 \hat{\mu}(l) := \bot, for l \in L \setminus \{l_0\};
 while S \neq \emptyset do{
      l := Choose S;
      S := S \setminus \{l\};
       foreach (l, c, l') \in T do{
            F := \hat{sp}(\hat{\mu}(l), c);
            if \neg(\mathsf{F} \leq \hat{\mu}(l')) then{
                 \hat{\mu}(l') := \hat{\mu}(l') \sqcup F;
                 S := S \cup \{l'\};
            }
       }
```

}

Abstract Domain D is a lattice (D, \leq, \sqcup)

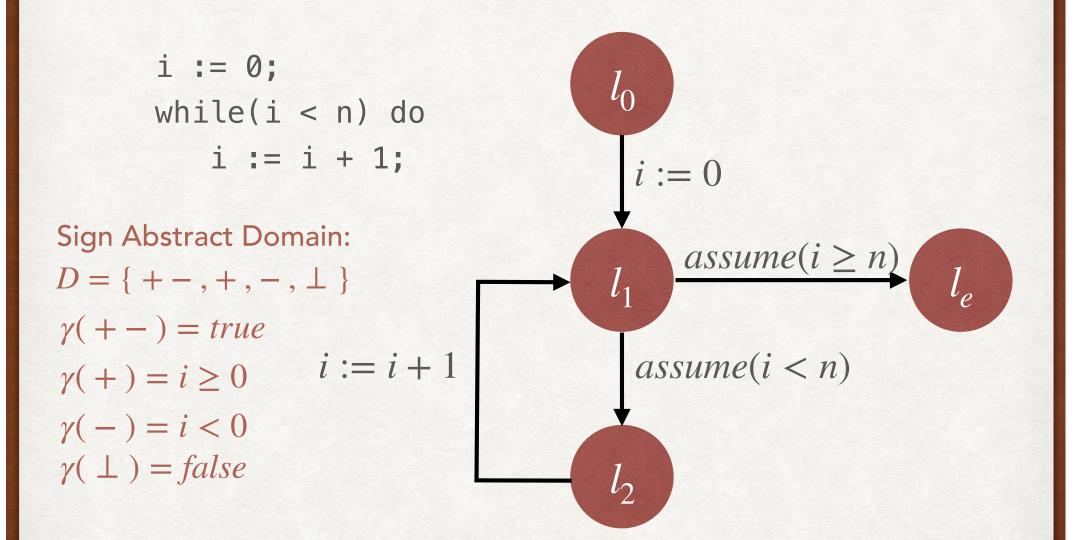
ABSTRACT INTERPRETATION: OVERVIEW

- At the end, we will check whether $\hat{\mu}(l_e) \leq \alpha(Q)$.
 - Equivalently, $\gamma(\hat{\mu}(l_e)) \subseteq Q$

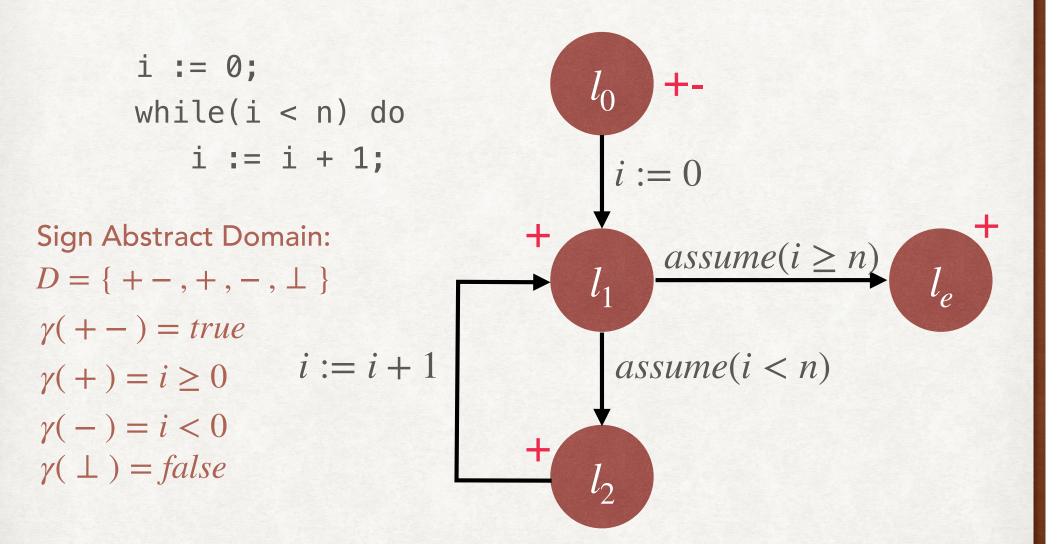


Suppose we want to prove the post-condition : $i \ge 0$

EXAMPLE



EXAMPLE



ABSTRACT INTERPRETATION: OVERVIEW

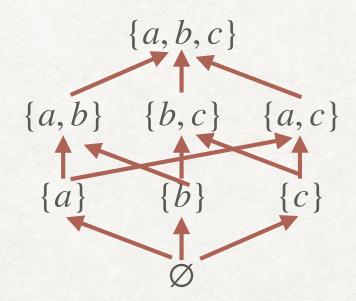
- Desirable properties of Abstract Interpretation
 - Soundness: $\hat{\mu}$ over approximates the set of states at every location.
 - Guaranteed termination of AbstractForwardPropagate
- We will use concepts from lattice theory to characterise the conditions required for these properties.

PARTIAL ORDER

- Given a set D, a binary relation $\leq \subseteq D \times D$ is a partial order on D if
 - \leq is reflexive: $\forall d \in D . d \leq d$
 - \leq is anti-symmetric: $\forall d, d' \in D . d \leq d' \land d' \leq d \rightarrow d = d'$
 - \leq is transitive: $\forall d_1, d_2, d_3 \in D, d_1 \leq d_2 \land d_2 \leq d_3 \rightarrow d_1 \leq d_3$
- Examples
 - \leq on \mathbb{N} is a partial order.
 - Given a set $S \subseteq On \mathbb{P}(S)$ is a partial order.

PARTIAL ORDER - EXAMPLES $S = \{a, b, c\}$

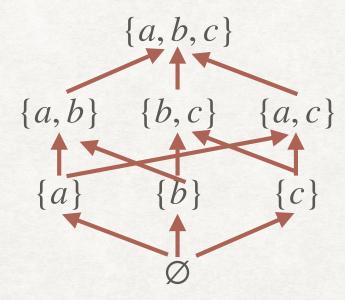
 $\mathbb{P}(S) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$



Partially Ordered Set: $(\mathbb{P}(S), \subseteq)$

PARTIAL ORDER - EXAMPLES

 $S = \{a, b, c\}$ $\mathbb{P}(S) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$



Hasse diagram:

- Doesn't show reflexive edges (self-loops)
- Doesn't show transitive edges

Partially Ordered Set: $(\mathbb{P}(S), \subseteq)$

PARTIAL ORDER - MORE EXAMPLES

- Which of the following are partially ordered sets (posets)?
 - $(\mathbb{N} \times \mathbb{N}, \{(a, b), (c, d) \mid a \le c\})$
 - $(\mathbb{N} \times \mathbb{N}, \{(a, b), (c, d) \mid a \le c \land b \le d\})$
 - $(\mathbb{N} \times \mathbb{N}, \{(a, b), (c, d) \mid a \le c \lor b \le d\})$

LEAST UPPER BOUND

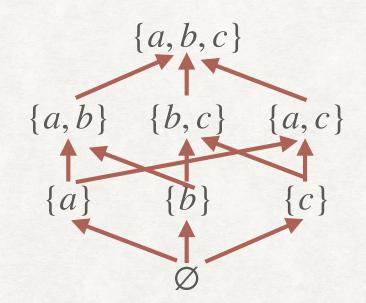
- Given a poset (D, ≤) and X ⊆ D, u ∈ D is called an upper bound on X if ∀x ∈ X.x ≤ u.
 - $u \in D$ is called the least upper bound (lub) of X, if u is an upper bound of X, and for every other upper bound u', $u \leq u'$.
 - We use the notation ⊔ X to denote the least upper bound of X.
 Also called the join of X.
 - Homework: Prove that the least upper bound, if it exists, is unique.

GREATEST LOWER BOUND

- Given a poset (D, ≤) and X ⊆ D, l ∈ D is called a lower bound on X if ∀x ∈ X. l ≤ x.
 - $l \in D$ is called the greatest lower bound (glb) of X, if l is a lower bound of X, and for every other lower bound l', $l' \leq l$.
 - We use the notation ⊓ X to denote the greatest lower bound of X. Also called the meet of X.
 - Homework: Prove that the greatest lower bound, if it exists, is unique.

LUB - EXAMPLE

 $S = \{a, b, c\}$ $\mathbb{P}(S) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$



- Consider $X = \{\{a\}, \{b\}\}$
- {a,b}, {a,b,c} are both upper bounds of X
- {*a*, *b*} is the least upper bound.

LATTICE

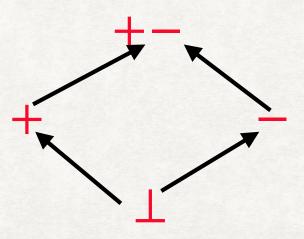
- A lattice is a poset (D, ≤) such that ∀x, y ∈ D, x ⊔ y and x ⊓ y exist.
 - A join semi-lattice is a poset (D, ≤) such that ∀x, y ∈ D, x ⊔ y exists.
 - A meet semi-lattice is a poset (D, ≤) such that ∀x, y ∈ D, x ⊓ y exists.
- A complete lattice is a lattice such that $\forall X \subseteq D$, $\sqcup X$ and $\sqcap X$ exists.
- Example: $(\mathbb{P}(S), \subseteq)$ is a complete lattice.

LATTICE - MORE EXAMPLES

- What is the simplest example of a poset that is not a lattice?
 - $(\{a,b\},\{(a,a),(b,b)\})$
- What is an example of a lattice which is not a complete lattice?
 - (\mathbb{N}, \leq)

LATTICE - MORE EXAMPLES

- What is the simplest example of a poset that is not a lattice?
 - $(\{a,b\},\{(a,a),(b,b)\})$
- What is an example of a lattice which is not a complete lattice?
 - (\mathbb{N}, \leq)
- Sign Lattice:



SOME PROPERTIES OF LATTICES

- (D, \leq) is a lattice, $x, y, z \in D$
 - If $x \leq y$, then $x \sqcup y = y$ and $x \sqcap y = x$.
 - $x \sqcup x = x$ and $x \sqcap x = x$
 - $(x \sqcup y) \sqcup z = x \sqcup (y \sqcup z) = \sqcup \{x, y, z\}$
 - If D is finite, then D is also a complete lattice.

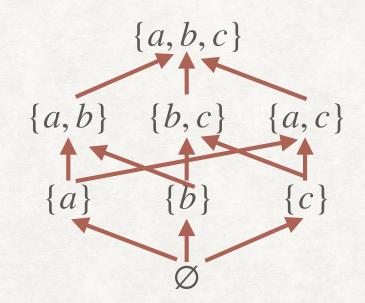
MINIMUM AND MAXIMUM

- Given a poset (D, ≤), x ∈ D is called the minimum element if ∀y ∈ D.x ≤ y.
 - Also called the bottom element. Denoted by \perp .
- Given a poset (D, ≤), x ∈ D is called the maximum element if ∀y ∈ D. y ≤ x.
 - Also called the top element. Denoted by T.
- Complete lattices are guaranteed to have top and bottom elements.
 - $\sqcup D = \top, \sqcap D = \bot$
 - $\sqcup \emptyset = \bot, \sqcap \emptyset = \top$

MONOTONIC FUNCTIONS

- Given two posets (D_1,\leq_1) and (D_2,\leq_2) , function $f:D_1\to D_2$ is called monotonic (or order-preserving) if
 - $\forall x, y \in D_1 . x \leq_1 y \to f(x) \leq_2 f(y)$
- In the special case when $D_1 = D_2 = D$, $f: D \rightarrow D$ is monotonic if
 - $\forall x, y \in D \, . \, x \le y \to f(x) \le f(y)$

MONOTONIC FUNCTIONS - EXAMPLE $S = \{a, b, c\}$ $\mathbb{P}(S) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$



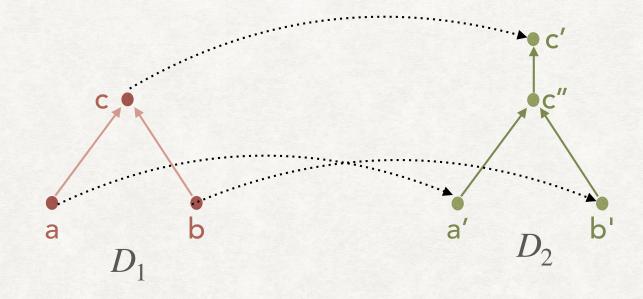
- Consider $f : \mathbb{P}(S) \to \mathbb{P}(S)$, $f(X) = X \cup \{a\}$.
 - *f* is monotonic.
- What about $f(X) = X \cap \{a\}$?
- Example of a non-monotonic function on $\mathbb{P}(S)$?

JOIN PRESERVING

• Given posets (D_1, \leq_1) and (D_2, \leq_2) , a monotonic function $f: D_1 \to D_2$, and $S \subseteq D_1$, if $\sqcup_1 S$ and $\sqcup_2 f(S)$ exist, then $\sqcup_2 f(S) \leq_2 f(\sqcup_1 S)$.

JOIN PRESERVING

• Given posets (D_1, \leq_1) and (D_2, \leq_2) , a monotonic function $f: D_1 \to D_2$, and $S \subseteq D_1$, if $\sqcup_1 S$ and $\sqcup_2 f(S)$ exist, then $\sqcup_2 f(S) \leq_2 f(\sqcup_1 S)$.



JOIN PRESERVING

• Given posets (D_1, \leq_1) and (D_2, \leq_2) , a monotonic function $f: D_1 \rightarrow D_2$, and $S \subseteq D_1$, if $\sqcup_1 S$ and $\sqcup_2 f(S)$ exist, then $\sqcup_2 f(S) \leq_2 f(\sqcup_1 S)$.

Proof: Let $u = \bigsqcup_1 S$.

Then $\forall x \in S \, . \, x \leq_1 u$. This implies that $\forall x \in S \, . \, f(x) \leq_2 f(u)$.

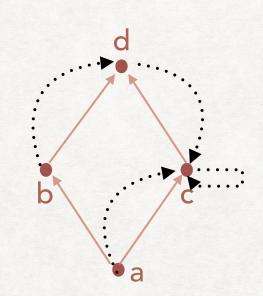
Thus f(u) is an upper bound of f(S).

Hence, $\sqcup_2 f(S) \leq_2 f(u)$.

FIXPOINTS

- A fixpoint of a function f: D → D is an element x ∈ D such that f(x) = x.
- A pre-fixpoint of a function f : D → D is an element x ∈ D such that x ≤ f(x).
- A post-fixpoint of a function f: D → D is an element x ∈ D such that f(x) ≤ x.

FIXPOINTS - EXAMPLE

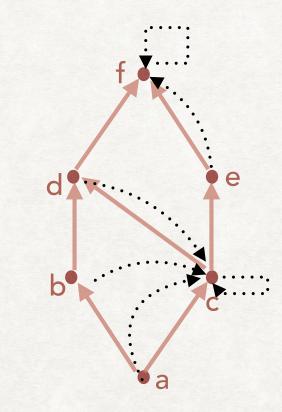


- Fixpoint : c
- Pre-fixpoints : a,b,c
- Post-fixpoint : c,d

KNASTER-TARSKI FIXPOINT THEOREM

- Let (D, ≤) be a complete lattice, and f: D → D be a monotonic function on (D, ≤). Then:
 - f has at least one fixpoint.
 - f has a least fixpoint (lfp), which is the same as the glb of the set of post-fixpoints of f, and a greatest fixpoint (gfp) which is the same as the lub of the set of pre-fixpoints of f.
 - The set of fixpoints of f itself forms a complete lattice under \leq .

KNASTER-TARSKI FIXPOINT THEOREM



- Complete Lattice
- Monotonic Function
- Pre-fixpoints: a,c,e,f
- Post-fixpoints: c,d,f
- Fixpoints: c,f

- $Pre = \{x \mid x \le f(x)\}$
 - We will show that ⊔ *Pre* is a fixpoint.
 - Notice that *Pre* cannot be empty. Why?

Proof: Let $u = \sqcup Pre$.

Consider $x \in Pre$. Then, $x \leq u$. Hence, $f(x) \leq f(u)$. Since $x \leq f(x)$, we have $x \leq f(u)$. Thus, f(u) is an upper bound of *Pre*. Since *u* is the least upper bound of *Pre*, we have $u \leq f(u)$.

 $u \le f(u) \Rightarrow f(u) \le f(f(u))$. Hence, f(u) is a pre-fixpoint. Therefore, $f(u) \le u$.

This proves that u = f(u).

- $Pre = \{x \mid x \le f(x)\}$
 - \Box *Pre* is the greatest fixpoint.

Proof: Consider another fixpoint *g*.

Then, g is also a pre-fixpoint. Hence, $g \leq \sqcup Pre$.

- $Post = \{x | f(x) \le x\}$
 - $\sqcap Post$ is a fixpoint of f.
 - \square *Post* is the least fixpoint.

HOMEWORK

• $P = \{x | f(x) = x\}$

• We will show that (P, \leq) is a complete lattice.

Proof Sketch: (P, \leq) is a partial order.

Let $X \subseteq P$. Let *u* be the $\sqcup X$ in *D*. Consider $U = \{a \in D \mid u \leq a\}$

Then (U, \leq) is a complete lattice. [Homework]

Further, $f(U) \subseteq U$. [Homework]

Hence, f is a monotonic function on complete lattice (U, \leq) . By previous part of Knaster-Tarski Theorem, the least fixpoint of f in U exists.

Let v be the least fixpoint of f in U. Then v is the least upper bound of X in P. [Homework]

Similarly, we can show that $\sqcap X$ also exists in P. [Homework]

CHAINS

- Given a poset (D, \leq) , $C \subseteq D$ is called a chain if $\forall x, y \in C . x \leq y \lor y \leq x$.
- A poset (D, ≤) satisfies the ascending chain condition, if for all sequences x₁ ≤ x₂ ≤ ..., ∃k. ∀n ≥ k.x_n = x_k.
 - We say that the sequence stabilizes to x_k .
- A poset (D, ≤) satisfies the descending chain condition, if for all sequences x₁ ≥ x₂ ≥ ..., ∃k. ∀n ≥ k.x_n = x_k.
 - A poset that satisfies the descending chain condition is also called well-ordered.
- Poset (D, ≤) is said to have finite height if it satisfies both the ascending and descending chain conditions.
 - Does (\mathbb{N}, \leq) have finite height?

COMPUTING LFP

- Consider a complete lattice (D, ≤) and a monotonic function f: D → D.
- Consider the sequence \perp , $f(\perp), f^2(\perp), f^3(\perp), \ldots$
 - If it stabilizes, it will converge to a fixpoint of f.
 - Further, this fixpoint will be the least fixpoint of f.
- Hence, if (D, \leq) satisfies the ascending chain condition, we can compute lfp(f) by finding the stable value of $\perp, f(\perp), f^2(\perp), f^3(\perp), \ldots$
- Homework: If $a \in Pre$, and the sequence $a, f(a), f^2(a), \ldots$ stabilizes, it will converge to the least fixpoint greater than a (denoted by $lfp_a(f)$).

CONTINUOUS FUNCTIONS

- Given two posets (D_1, \leq_1) and (D_2, \leq_2) , function $f: D_1 \to D_2$ is called continuous if for all chains $C \subseteq D_1$ such that $\sqcup_1 C$ exists, then $\sqcup_2 f(C)$ also exists, and $\sqcup_2 f(C) = f(\sqcup_1 C)$.
- Are all continuous functions monotonic?
 - Consider $x \leq_1 y$. Now, $f(x \sqcup_1 y) = f(x) \sqcup_2 f(y)$. Hence, $f(y) = f(x) \sqcup_2 f(y)$. This implies that $f(x) \leq f(y)$.
- Are all monotonic functions continuous?

CONTINUOUS FUNCTIONS

- Given a complete lattice (D, \leq) and continuous function $f: D \to D, lfp(f) = \bigsqcup_{i \geq 0} f^i(\perp)$ • $f(\bigsqcup_{i \geq 0} f^i(\perp)) = \bigsqcup_{i \geq 1} f^i(\perp) = \bigsqcup_{i \geq 0} f^i(\perp)$
 - Consider a fixpoint g. Then g is an upper bound of the chain \perp , $f(\perp), f^2(\perp), \ldots$

• Hence, $\bigsqcup_{i \ge 0} f^i(\perp) \le g$

DISTRIBUTIVE AND INFINITELY DISTRIBUTIVE FUNCTIONS

- Given two posets (D_1, \leq_1) and (D_2, \leq_2) , function $f: D_1 \to D_2$ is called distributive if for $x, y \in D_1$ such that $x \sqcup_1 y$ exists, then $f(x) \sqcup_2 f(y)$ also exists, and $f(x \sqcup_1 y) = f(x) \sqcup_2 f(y)$.
- Given two posets (D_1, \leq_1) and (D_2, \leq_2) , function $f: D_1 \to D_2$ is called infinitely distributive if for all $X \subseteq D_1$ such that $\sqcup_1 X$ exists, then $\sqcup_2 f(X)$ also exists, and $\sqcup_2 f(X) = f(\sqcup_1 X)$.

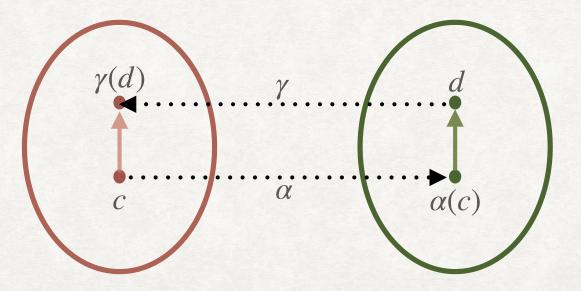
GALOIS CONNECTION

- Given posets (C, \leq_1) and (D, \leq_2) , a pair of functions (α, γ) , $\alpha : C \to D$ and $\gamma : D \to C$ is called a Galois connection if
 - $\forall c \in C . \forall d \in D . \alpha(c) \leq_2 d \Leftrightarrow c \leq_1 \gamma(d)$

• Also written as: $(C, \leq_1) \stackrel{\alpha}{\rightleftharpoons} (D, \leq_2)$

GALOIS CONNECTION

- Given posets (C, \leq_1) and (D, \leq_2) , a pair of functions (α, γ) , $\alpha : C \to D$ and $\gamma : D \to C$ is called a Galois connection if
 - $\forall c \in C . \forall d \in D . \alpha(c) \leq_2 d \Leftrightarrow c \leq_1 \gamma(d)$
- Also written as: $(C, \leq_1) \stackrel{\alpha}{\rightleftharpoons} (D, \leq_2)$



PROPERTIES OF GALOIS CONNECTION

- $x \leq_1 \gamma(\alpha(x))$
 - Proof: $\alpha(x) \leq_2 \alpha(x)$. Applying the definition of Galois connection, $x \leq_1 \gamma(\alpha(x))$.
- $\alpha(\gamma(y)) \leq_2 y$
 - Proof: $\gamma(y) \leq_1 \gamma(y)$. Applying the definition of Galois connection, $\alpha(\gamma(y)) \leq_2 y$.

PROPERTIES OF GALOIS CONNECTION

• α is monotonic.

- **Proof:** Consider $c_1, c_2 \in C$ such that $c_1 \leq_1 c_2$.
- We know that $c_2 \leq \gamma(\alpha(c_2))$. By transitivity, $c_1 \leq \gamma(\alpha(c_2))$. Hence, by definition of Galois connection, $\alpha(c_1) \leq_2 \alpha(c_2)$.
- γ is monotonic.
 - Proof: Homework.

ONTO GALOIS CONNECTION

- Recall: *State* $\triangleq V \rightarrow \mathbb{R}$. The concrete domain *C* will be $(\mathbb{P}(State), \subseteq)$.
- The abstract domain D will be a collection of artificially constrained set of states. We can represent this as $D \subseteq C$.
- The abstraction function α will map $c \in C$ to the smallest set $d \in D$ such that $c \subseteq d$.
- The concretization function γ will simply be $\gamma(d) = d$.
 - Prove that this is a Galois Connection.

ONTO GALOIS CONNECTION - EXAMPLE

- Assume that $V = \{v\}$.
 - Hence, $State = \mathbb{R}$, The concrete domain C is $(\mathbb{P}(\mathbb{R}), \subseteq)$
- Sign Abstract Domain: $D = \{+-, +, -, \bot\}$.
 - $+- \triangleq \mathbb{R}$
 - $+ \triangleq \{n \in \mathbb{R} \mid n \ge 0\}$
 - $\triangleq \{n \in \mathbb{R} \mid n < 0\}$
 - ⊥ ≜ Ø
- Clearly $D \subseteq C$.

ONTO GALOIS CONNECTION - EXAMPLE

• Define the Galois Connection: $(\mathbb{P}(\mathbb{R}), \subseteq) \stackrel{\alpha}{\rightleftharpoons} (D, \subseteq)$

- $\alpha(c) = + \text{ if } min(c) \ge 0$
- $\alpha(c) = \text{ if } max(c) < 0$
- $\alpha(\emptyset) = \bot$
- Otherwise, $\alpha(c) = + -$.
- $\gamma(d) = d$.
- Example: $\alpha(\{3,5\}) = +, \alpha(\{3,6,-1,0\}) = + -$

ONTO GALOIS CONNECTION

- The abstraction function α will map $c \in C$ to the smallest set $d \in D$ such that $c \subseteq d$.
- The concretization function γ will simply be $\gamma(d) = d$.
- Notice that $\alpha(\gamma(d)) = d$.
 - Also called Onto Galois Connection.
 - From now onwards, we will assume that Galois Connections are Onto.