

AN INTRODUCTION TO  
ABSTRACT INTERPRETATION

P. COUSOT

Patrick.Cousot@ens.fr <http://www.di.ens.fr/~cousot>

Biarritz IFIP-WG 2.3 meeting (1)

23 — 28 mars 2003, Hotel Miramar, Biarritz, France

© P. COUSOT, ALL RIGHTS RESERVED.

An Introduction to Abstract Interpretation, © P. Cousot, 25/3/03— 3:1/58 —◀◀◀▶▶▶ ▶◀!■□?▶ Idx, Toc

2.2.1 MOORE FAMILY-BASED ABSTRACTION

See Sec. 5.1 of [POPL'79].

— Reference —

[POPL'79] P. Cousot & R. Cousot. Systematic design of program analysis frameworks. In *6<sup>th</sup> POPL*, pages 269–282, San Antonio, TX, 1979. ACM Press. 10

An Introduction to Abstract Interpretation, © P. Cousot, 23/3/03— 2:10/102 —◀◀◀▶▶▶ ▶◀!■□?▶ Idx, Toc

2.2 A SHORT INTRODUCTION TO ABSTRACT INTERPRE-  
TATION THEORY (SEE SEC. 5 OF [POPL'79])

— Reference —

[POPL'79] P. Cousot & R. Cousot. Systematic design of program analysis frameworks. In *6<sup>th</sup> POPL*, pages 269–282, San Antonio, TX, 1979. ACM Press. 9, 99

An Introduction to Abstract Interpretation, © P. Cousot, 23/3/03— 2:9/102 —◀◀◀▶▶▶ ▶◀!■□?▶ Idx, Toc

PROPERTIES

- We represent **properties**  $P$  of objects  $s \in \Sigma$  as **sets of objects**  $P \in \wp(\Sigma)$  (which have the property in question);

**Example:** the property “*to be an even natural number*” is  $\{0, 2, 4, 6, \dots\}$

An Introduction to Abstract Interpretation, © P. Cousot, 23/3/03— 2:11/102 —◀◀◀▶▶▶ ▶◀!■□?▶ Idx, Toc

## COMPLETE LATTICE OF PROPERTIES

- The set of properties of objects  $\Sigma$  is a complete boolean lattice:

$$\langle \wp(\Sigma), \subseteq, \emptyset, \Sigma, \cup, \cap, \neg \rangle .$$

## DIRECTION OF APPROXIMATION

- **Approximation from above:** approximate  $P$  by  $\overline{P}$  such that  $P \subseteq \overline{P}$ ;
- **Approximation from below:** approximate  $P$  by  $\underline{P}$  such that  $\underline{P} \subseteq P$  (dual).

## ABSTRACTION

A reasoning/computation such that:

- only some properties can be used;
- the properties that can be used are called “*abstract*”;
- so, the (other **concrete**) properties must be **approximated** by the abstract ones;

## ABSTRACT PROPERTIES

- **Abstract Properties:** a set  $\overline{\mathcal{A}} \subseteq \wp(\Sigma)$  of properties of interest (the only one which can be used to approximate others).



## AVOIDING BACKTRACKING

- We don't want to **exhaustively try all minimal approximations**;
- We want to **use only one of the minimal approximations**;

## BEST ABSTRACTION

- We require that all concrete property  $P \in \wp(\Sigma)$  have a **best abstraction**  $\bar{P} \in \bar{\mathcal{A}}$ :

$$P \subseteq \bar{P} \\ \forall \bar{P}' \in \bar{\mathcal{A}} : (P \subseteq \bar{P}') \implies (\bar{P} \subseteq \bar{P}')$$

- So, by definition of the greatest lower bound/meet  $\sqcap$ :

$$\bar{P} = \sqcap \{ \bar{P}' \in \bar{\mathcal{A}} \mid P \subseteq \bar{P}' \} \in \bar{\mathcal{A}}$$

## WHICH MINIMAL ABSTRACTION TO USE?

- Which **minimal abstraction** to choose?
  - make a **circumstantial choice**<sup>1</sup>;
  - make a definitive **arbitrary choice**<sup>2</sup>;
  - require the existence of a **best choice**<sup>3</sup>.

---

### Reference

[JLC'92] P. Cousot & R. Cousot. Abstract interpretation frameworks. *J. Logic and Comp.*, 2(4):511–547, 1992.

<sup>1</sup> [JLC'92] uses a concretization function.

<sup>2</sup> [JLC'92] uses an abstraction function.

<sup>3</sup> [JLC'92] uses an abstraction/concretization Galois connection (this talk).

## MOORE FAMILY

- So, the hypothesis that any concrete property  $P \in \wp(\Sigma)$  has a **best abstraction**  $\bar{P} \in \bar{\mathcal{A}}$  implies that:

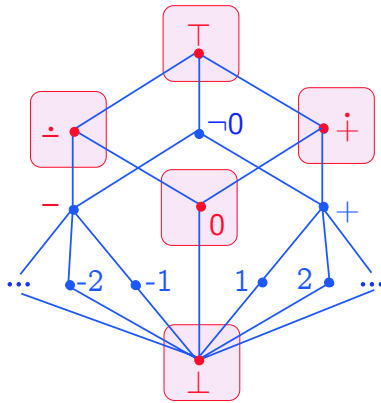
$\bar{\mathcal{A}}$  is a **Moore family**

i.e. it is closed under intersection  $\sqcap$ :

$$\forall S \subseteq \bar{\mathcal{A}} : \sqcap S \in \bar{\mathcal{A}}$$

- In particular  $\sqcap \emptyset = \Sigma \in \bar{\mathcal{A}}$ .

## EXAMPLE OF MOORE FAMILY-BASED ABSTRACTION



## 2.2.2 CLOSURE OPERATOR-BASED ABSTRACTION

See Sec. 5.2 of [POPL'79].

Reference

[POPL'79] P. Cousot & R. Cousot. Systematic design of program analysis frameworks. In 6<sup>th</sup> POPL, pages 269–282, San Antonio, TX, 1979. ACM Press. 26

## THE LATTICE OF ABSTRACTIONS (1)

- The set  $\mathcal{M}(\wp(\wp(\Sigma)))$  of all abstractions i.e. of Moore families on the set  $\wp(\Sigma)$  of concrete properties is the complete **lattice of abstractions**

$$(\mathcal{M}(\wp(\wp(\Sigma))), \supseteq, \wp(\Sigma), \{\Sigma\}, \lambda S. \mathcal{M}(US), \cap)$$

where:

$$\mathcal{M}(\bar{A}) = \{\bigcap S \mid S \subseteq \bar{A}\}$$

is the  $\subseteq$ -least Moore family containing  $\bar{A}$ .

## CLOSURE OPERATOR INDUCED BY AN ABSTRACTION

The map  $\rho_{\bar{A}}$  mapping a concrete property  $P \in \wp(\Sigma)$  to its best abstraction  $\rho_{\bar{A}}(P)$  in  $\bar{A}$  is:

$$\rho_{\bar{A}}(P) = \bigcap \{\bar{P} \in \bar{A} \mid P \subseteq \bar{P}\}.$$

It is a **closure operator**:

- extensive,
- idempotent,
- isotone/monotonic;

such that

$$P \in \bar{A} \iff P = \rho_{\bar{A}}(P)$$

hence

$$\bar{A} = \rho_{\bar{A}}(\wp(\Sigma)).$$

## ABSTRACTION INDUCED BY A CLOSURE OPERATOR

- Any closure operator  $\rho$  on the set of properties  $\wp(\Sigma)$  induces an abstraction:

$$\rho(\wp(\Sigma)).$$

Examples:

- $\lambda P \cdot P$  the most precise abstraction (identity),
- $\lambda P \cdot \Sigma$  the most imprecise abstraction (I don't know).
- Closure operators are isomorphic to the Moore families (i.e. their fixpoints).

## THE LATTICE OF ABSTRACTIONS (2)

- The set  $\text{clo}(\wp(\Sigma) \mapsto \wp(\Sigma))$  of all abstractions, i.e. isomorphically, closure operators  $\rho$  on the set  $\wp(\Sigma)$  of concrete properties is the complete lattice of abstractions for pointwise inclusion<sup>4</sup>:

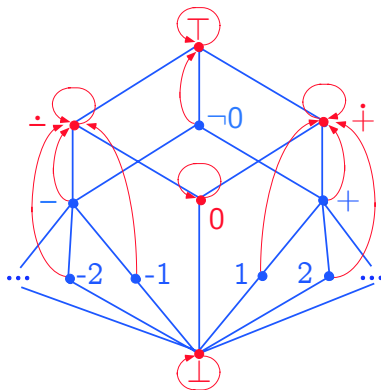
$$\langle \text{clo}(\wp(\Sigma) \mapsto \wp(\Sigma)), \subseteq, \lambda P \cdot P, \lambda P \cdot \Sigma, \lambda S \cdot \text{ide}(\dot{\cup} S), \dot{\cap} \rangle$$

where:

- the lub  $\lambda S \cdot \text{ide}(\dot{\cup} S)$  is the reduced product;
- $\text{ide}(\rho) = \text{lfp}_{\rho}^{\subseteq} \lambda f \cdot f \circ f$  is the  $\subseteq$ -least idempotent operator on  $\wp(\Sigma)$   $\subseteq$ -greater than  $\rho$ .

<sup>4</sup> M. Ward, *The closure operators of a lattice*, Annals Math., 43(1942), 191–196.

## EXAMPLE OF CLOSURE OPERATOR-BASED ABSTRACTION



## 2.2.4 GALOIS CONNECTION-BASED ABSTRACTION

See Sec. 5.3 of [POPL'79]).

Reference

[POPL'79] P. Cousot & R. Cousot. Systematic design of program analysis frameworks. In *6<sup>th</sup> POPL*, pages 269–282, San Antonio, TX, 1979. ACM Press. 38



## GALOIS CONNECTION

- Relaxing the condition that  $\alpha$  is onto:

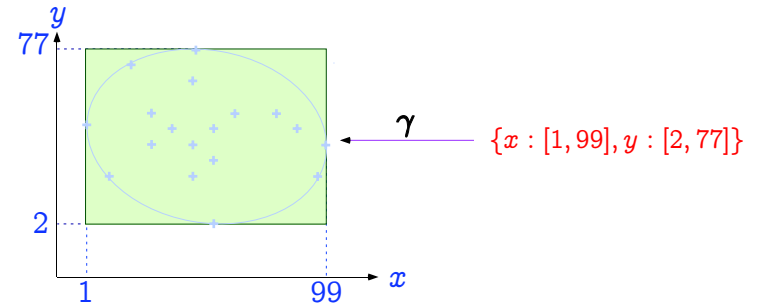
$$\langle \wp(\Sigma), \subseteq \rangle \xleftrightarrow[\alpha]{\gamma} \langle \overline{\mathcal{D}}, \sqsubseteq \rangle$$

that is to say:

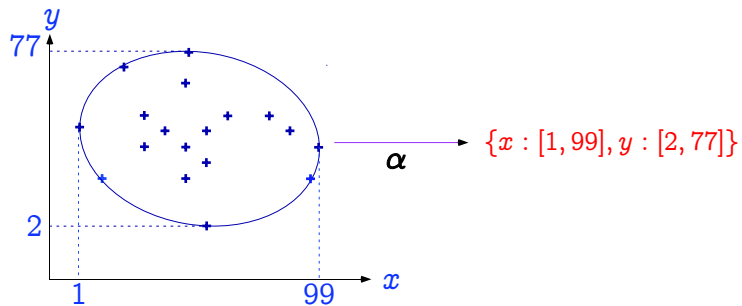
$$\forall P \in \wp(\Sigma), \overline{P} \in \overline{\mathcal{D}} : \alpha(P) \sqsubseteq \overline{P} \Leftrightarrow P \subseteq \gamma(\overline{P});$$

- i.e.  $\rho$  is now  $\gamma \circ \alpha$ ;  
We can now have different representations of the same abstract property.

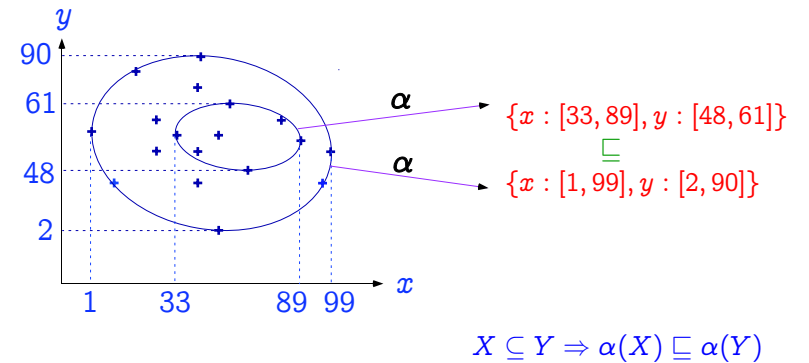
## CONCRETIZATION $\gamma$



## ABSTRACTION $\alpha$



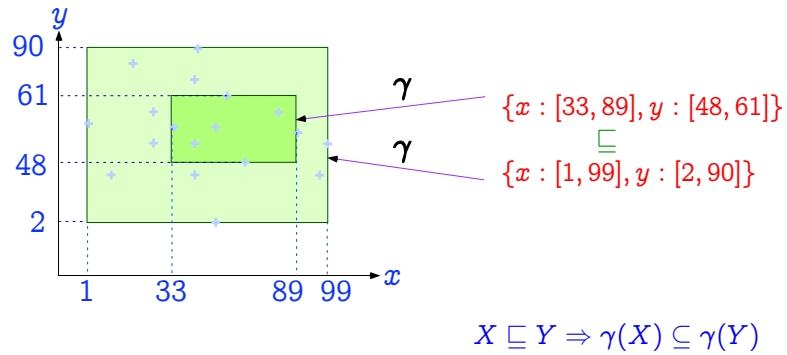
## THE ABSTRACTION $\alpha$ IS MONOTONE



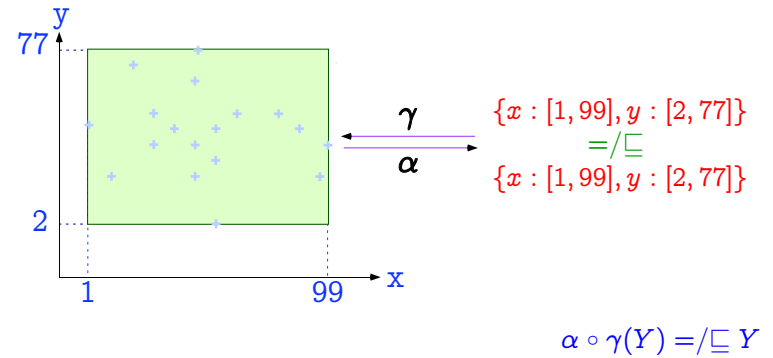
$$X \subseteq Y \Rightarrow \alpha(X) \sqsubseteq \alpha(Y)$$



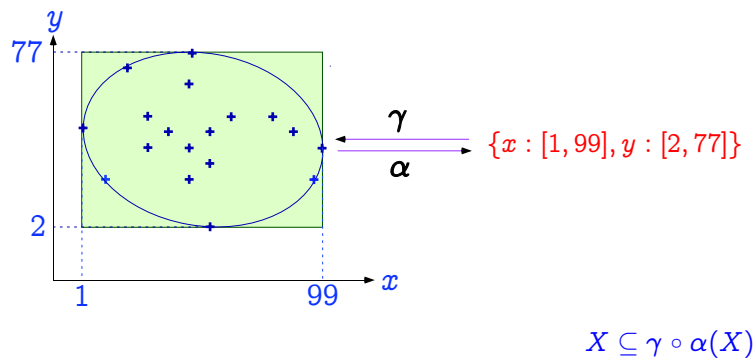
### THE CONCRETIZATION $\gamma$ IS MONOTONE



### THE $\alpha \circ \gamma$ COMPOSITION IS REDUCTIVE



### THE $\gamma \circ \alpha$ COMPOSITION IS EXTENSIVE



### COMPOSITION OF GALOIS CONNECTIONS

The composition of Galois connections:

$$\langle L, \leq \rangle \xleftrightarrow[\alpha_1]{\gamma_1} \langle M, \sqsubseteq \rangle$$

and:

$$\langle M, \sqsubseteq \rangle \xleftrightarrow[\alpha_2]{\gamma_2} \langle N, \preceq \rangle$$

is a Galois connection:

$$\langle L, \leq \rangle \xleftrightarrow[\alpha_2 \circ \alpha_1]{\gamma_1 \circ \gamma_2} \langle N, \preceq \rangle$$

## 2.2.5 FUNCTION ABSTRACTION

See Sec. 7.2 of [POPL'79].

Reference

[POPL'79] P. Cousot & R. Cousot. Systematic design of program analysis frameworks. In 6<sup>th</sup> POPL, pages 269–282, San Antonio, TX, 1979. ACM Press. 51

An Introduction to Abstract Interpretation, © P. Cousot, 23/3/03— 2:51/102 — ◀ ◁ ▷ ▶ ▶ ◀ ◁ ▷ ▶ ◀ ◁ ▷ ▶ Idx, Toc

## 2.2.6 FIXPOINT ABSTRACTION

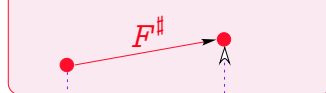
See Sec. 7.1 of [POPL'79].

Reference

[POPL'79] P. Cousot & R. Cousot. Systematic design of program analysis frameworks. In 6<sup>th</sup> POPL, pages 269–282, San Antonio, TX, 1979. ACM Press. 53

An Introduction to Abstract Interpretation, © P. Cousot, 23/3/03— 2:53/102 — ◀ ◁ ▷ ▶ ▶ ◀ ◁ ▷ ▶ ◀ ◁ ▷ ▶ Idx, Toc

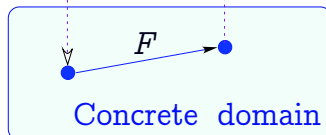
### Abstract domain



### FUNCTION ABSTRACTION

$$F^\# = \alpha \circ F \circ \gamma$$

i.e.  $F^\# = \rho \circ F$

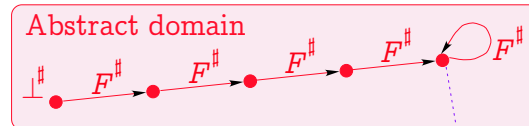


$$\langle P, \sqsubseteq \rangle \xrightarrow[\alpha]{\gamma} \langle Q, \sqsubseteq \rangle \Rightarrow$$

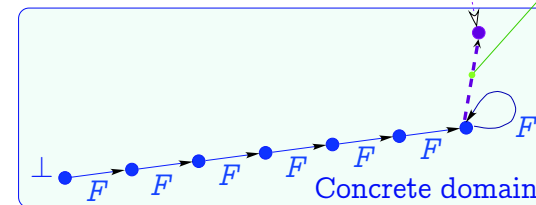
$$\langle P \xrightarrow{\text{mon}} P, \dot{\sqsubseteq} \rangle \xrightarrow[\lambda F \circ \alpha \circ F \circ \gamma]{\lambda F^\# \circ \gamma \circ F^\# \circ \alpha} \langle Q \xrightarrow{\text{mon}} Q, \dot{\sqsubseteq} \rangle$$

An Introduction to Abstract Interpretation, © P. Cousot, 23/3/03— 2:52/102 — ◀ ◁ ▷ ▶ ▶ ◀ ◁ ▷ ▶ ◀ ◁ ▷ ▶ Idx, Toc

### APPROXIMATE FIXPOINT ABSTRACTION



Approximation relation  $\sqsubseteq$



$$\alpha(\text{lfp } F) \sqsubseteq \text{lfp } F^\#$$

An Introduction to Abstract Interpretation, © P. Cousot, 23/3/03— 2:54/102 — ◀ ◁ ▷ ▶ ▶ ◀ ◁ ▷ ▶ ◀ ◁ ▷ ▶ Idx, Toc

## APPROXIMATE/EXACT FIXPOINT ABSTRACTION

Exact Abstraction:

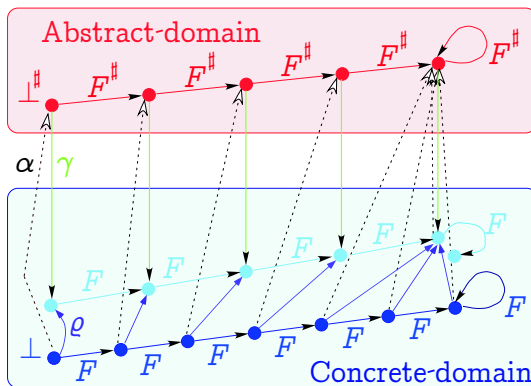
$$\alpha(\text{lfp } F) = \text{lfp } F^\sharp$$

Approximate Abstraction:

$$\alpha(\text{lfp } F) \sqsubseteq^\sharp \text{lfp } F^\sharp$$

## 2.3 APPLICATION TO REACHABILITY

## EXACT FIXPOINT ABSTRACTION

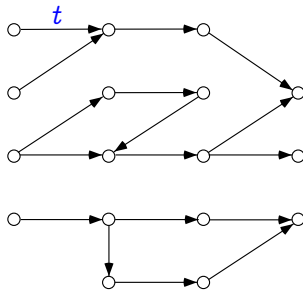


$$\alpha \circ F = F^\sharp \circ \alpha \Rightarrow \alpha(\text{lfp } F) = \text{lfp } F^\sharp$$

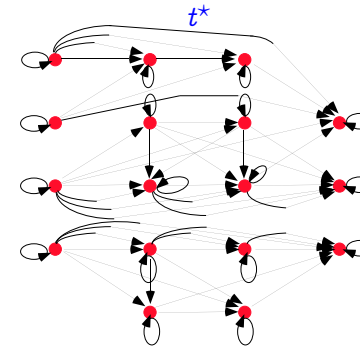
## TRANSITION SYSTEMS

- $\langle S, t \rangle$  where:
  - $S$  is a set of states/vertices/...
  - $t \in \wp(S \times S)$  is a transition relation/set of arcs/...

### EXAMPLE OF TRANSITION SYSTEM



### EXAMPLE OF TRANSITIVE REFLEXIVE CLOSURE



### REFLEXIVE TRANSITIVE CLOSURE

- Composition:
  - $t \circ t' \stackrel{\text{def}}{=} \{ \langle s, s'' \rangle \mid \exists s' : \langle s, s' \rangle \in t \wedge \langle s', s'' \rangle \in t' \}$
- Powers:
  - $t^0 \stackrel{\text{def}}{=} \{ \langle s, s \rangle \mid s \in S \}$
  - $t^{n+1} \stackrel{\text{def}}{=} t^n \circ t \quad n \geq 0$
- Reflexive transitive closure:
  - $t^* = \bigcup_{n \geq 0} t^n$

### REFLEXIVE TRANSITIVE CLOSURE IN FIXPOINT FORM

$$t^* = \text{lfp}^{\subseteq} \lambda X. t^0 \cup X \circ t$$

Proof

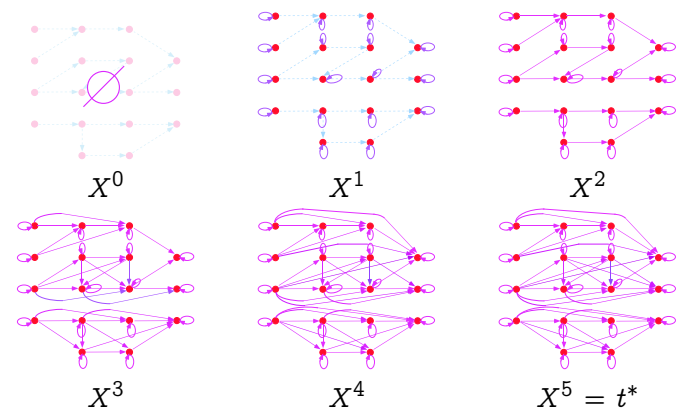
$$\begin{aligned} X^0 &= \emptyset \\ X^1 &= t^0 \cup X^0 \circ t = t^0 \\ X^2 &= t^0 \cup X^1 \circ t = t^0 \cup t^0 \circ t = t^0 \cup t^1 \\ &\dots \dots \\ X^n &= \bigcup_{0 \leq i < n} t^i \quad (\text{induction hypothesis}) \end{aligned}$$

$$\begin{aligned}
X^{n+1} &= t^0 \cup X^n \circ t \\
&= t^0 \cup \left( \bigcup_{0 \leq i < n} t^i \right) \circ t \\
&= t^0 \cup \bigcup_{0 \leq i < n} (t^i \circ t) \\
&= t^0 \cup \bigcup_{1 \leq i+1 < n+1} t^{i+1} \\
&= t^0 \cup \left( \bigcup_{1 \leq j < n+1} t^j \right) \circ t \\
&= \bigcup_{0 \leq i < n+1} t^i \\
&\dots \quad \dots
\end{aligned}$$

$$\begin{aligned}
X^{\omega+1} &= t^0 \cup X^\omega \circ t \\
&= t^0 \cup \left( \bigcup_{n \geq 0} t^n \right) \circ t \\
&= t^0 \cup \bigcup_{n \geq 0} (t^n \circ t) \\
&= t^0 \cup \bigcup_{n \geq 0} t^{n+1} \\
&= t^0 \cup \bigcup_{k \geq 1} t^k \\
&= \bigcup_{n \geq 0} t^n \\
&= t^*
\end{aligned}$$

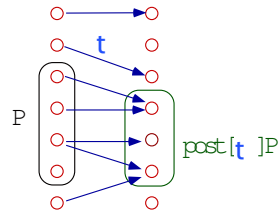
$$\begin{aligned}
X^\omega &= \bigcup_{n \geq 0} X^n \\
&= \bigcup_{n \geq 0} \bigcup_{0 \leq i < n} t^i \\
&= \bigcup_{n \geq 0} t^n \\
&= t^*
\end{aligned}$$

### ITERATES



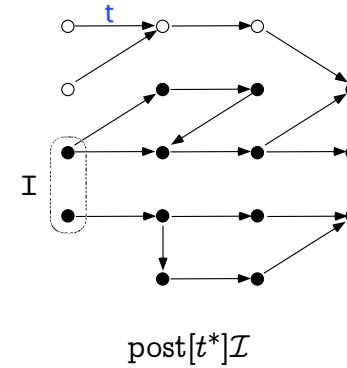
## POST-IMAGE

$$\text{post}[t]I = \{s' \mid \exists s \in I : \langle s, s' \rangle \in t\}$$



We have  $\text{post}[\bigcup_{i \in \Delta} t^i]I = \bigcup_{i \in \Delta} \text{post}[t^i]I$  so  $\alpha = \lambda t \cdot \text{post}[t]I$  is the lower adjoint of a Galois connection.

## REACHABLE STATES



## POSTIMAGE GALOIS CONNECTION

Given  $I \in \wp(S)$ ,

$$\langle \wp(S \times S), \subseteq \rangle \xrightleftharpoons[\lambda t \cdot \text{post}[t]I]{\gamma} \langle \wp(S), \subseteq \rangle$$

$$\begin{aligned} & \text{post}[t]I \subseteq R \\ \Leftrightarrow & \{s' \mid \exists s \in I : \langle s, s' \rangle \in t\} \subseteq R \\ \Leftrightarrow & \forall s' \in S : (\exists s \in I : \langle s, s' \rangle \in t) \Rightarrow (s' \in R) \\ \Leftrightarrow & \forall s', s \in S : (s \in I \wedge \langle s, s' \rangle \in t) \Rightarrow (s' \in R) \\ \Leftrightarrow & \forall s', s \in S : \langle s, s' \rangle \in t \Rightarrow ((s \in I) \Rightarrow (s' \in R)) \\ \Leftrightarrow & t \subseteq \{\langle s, s' \rangle \mid (s \in I) \Rightarrow (s' \in R)\} \stackrel{\text{def}}{=} \gamma(R) \end{aligned}$$

## FIXPOINT ABSTRACTION, ONCE AGAIN

Let  $F \in L \xrightarrow{m} L$  and  $\overline{F} \in \overline{L} \xrightarrow{m} \overline{L}$  be respective monotone maps on the cpos  $\langle L, \perp, \sqsubseteq \rangle$  and  $\langle \overline{L}, \overline{\perp}, \overline{\sqsubseteq} \rangle$  and  $\langle L, \sqsubseteq \rangle \xrightarrow[\alpha]{\gamma} \langle \overline{L}, \overline{\sqsubseteq} \rangle$  such that  $\alpha \circ F \circ \gamma \sqsubseteq \overline{F}$ . Then<sup>10</sup>:

- $\forall \delta \in \mathbb{O} : \alpha(F^\delta) \sqsubseteq \overline{F}^\delta$  (iterates from the infimum);
- The iteration order of  $\overline{F}$  is  $\leq$  to that of  $F$ ;
- $\alpha(\text{lfp}^\square F) \sqsubseteq \text{lfp}^\square \overline{F}$ ;

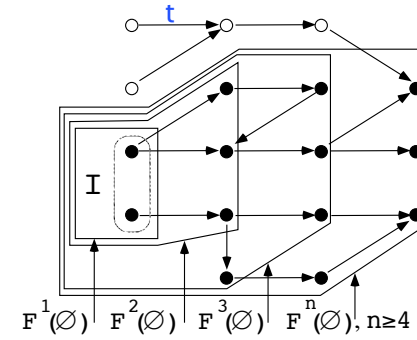
**Soundness:**  $\text{lfp}^\square \overline{F} \sqsubseteq \overline{P} \Rightarrow \text{lfp}^\square F \sqsubseteq \gamma(\overline{P})$ .

<sup>10</sup> P. Cousot & R. Cousot. *Systematic design of program analysis frameworks*. ACM POPL'79, pp. 269–282, 1979. Numerous variants!



$$\begin{aligned}
& \text{post}[X \circ t]I \\
&= \{s' \mid \exists s \in I : \langle s, s' \rangle \in (X \circ t)\} \\
&= \{s' \mid \exists s \in I : \langle s, s' \rangle \in \{\langle s, s'' \rangle \mid \exists s' : \langle s, s'' \rangle \in X \wedge \langle s', s'' \rangle \in t\}\} \\
&= \{s' \mid \exists s \in I : \exists s'' \in S : \langle s, s'' \rangle \in X \wedge \langle s', s'' \rangle \in t\} \\
&= \{s' \mid \exists s'' \in S : (\exists s \in I : \langle s, s'' \rangle \in X) \wedge \langle s', s'' \rangle \in t\} \\
&= \{s' \mid \exists s'' \in S : s'' \in \{s'' \mid \exists s \in I : \langle s, s'' \rangle \in X\} \wedge \langle s', s'' \rangle \in t\} \\
&= \{s' \mid \exists s'' \in S : s'' \in \text{post}[X]I \wedge \langle s', s'' \rangle \in t\} \\
&= \text{post}[t](\text{post}[X]I) \\
&= \text{post}[t](\alpha(X))
\end{aligned}$$

### EXAMPLE OF ITERATION



$$\begin{aligned}
& \alpha \circ (\lambda X \cdot t^0 \cup X \circ t) \\
&= \dots \\
&= \lambda X \cdot \text{post}[t^0]I \cup \text{post}[X \circ t]I \\
&= \lambda X \cdot I \cup \text{post}[t](\alpha(X)) \\
&= \lambda X \cdot \overline{F}(\alpha(X))
\end{aligned}$$

by defining:

$$\overline{F} = \lambda X \cdot I \cup \text{post}[t](X)$$

proving:

$$\text{post}[t^*](I) = \text{lfp}^{\subseteq} \lambda X \cdot I \cup \text{post}[t](X) \quad (2)$$