

AN INTRODUCTION TO ABSTRACT INTERPRETATION

P. COUSOT

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

Biarritz IFIP-WG 2.4 meeting (1)

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

© P. COUSOT, ALL RIGHTS RESERVED.

3. APPLICATION TO STATIC ANALYSIS

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

Reference

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

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 *6th POPL*, pages 269–282, San Antonio, TX, 1979. ACM Press. 10

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 $\{0, 2, 4, 6, \dots\}$

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 $\overline{A} \subsetneq \wp(\Sigma)$ of properties of interest (the only one which can be used to approximate others).

IN ABSENCE OF (UPPER) APPROXIMATION

- What to say when some property has **no** (computable) **abstraction**?
 - loop?
 - block?
 - ask for help?
 - say something!

I DON'T KNOW

- Any property should be approximable from above by **I don't know** (i.e. “true” or Σ).

MINIMAL APPROXIMATIONS

- A concrete property $P \in \wp(\Sigma)$ is **most precisely abstracted** by any minimal upper approximation $\overline{P} \in \overline{\mathcal{A}}$:

$$P \subseteq \overline{P}$$
$$\nexists \overline{P'} \in \overline{\mathcal{A}} : P \subseteq \overline{P'} \subsetneq \overline{P}$$

- So, an abstract property $\overline{P} \in \overline{\mathcal{A}}$ is best approximated by itself.

WHICH MINIMAL APPROXIMATION IS MOST USEFUL?

- Which minimal approximation is **most useful** depends upon the circumstances;
- Example (**rule of signs**):
 - 0 is better approximated as **positive** in “ $3 + 0$ ”;
 - 0 is better approximated as **negative** in “ $-3 + 0$ ”.

AVOIDING BACKTRACKING

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

WHICH MINIMAL ABSTRACTION TO USE?

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

Reference

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

¹ [JLC '92] uses a concretization function.

² [JLC '92] uses an abstraction function.

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

BEST ABSTRACTION

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

$$\begin{aligned} P &\subseteq \overline{P} \\ \forall \overline{P}' \in \overline{\mathcal{A}} : (P &\subseteq \overline{P}') \implies (\overline{P} \subseteq \overline{P}') \end{aligned}$$

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

$$\overline{P} = \bigcap \{ \overline{P}' \in \overline{\mathcal{A}} \mid P \subseteq \overline{P}' \} \in \overline{\mathcal{A}}$$

MOORE FAMILY

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

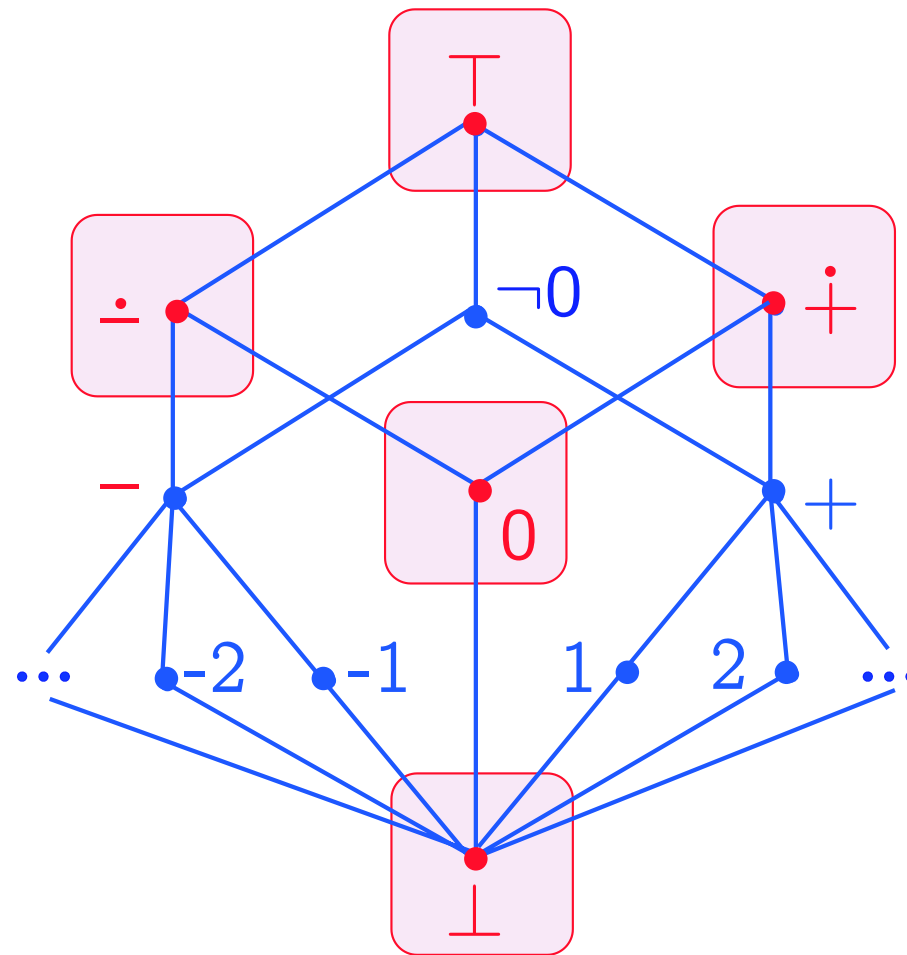
$\overline{\mathcal{A}}$ is a Moore family

i.e. it is closed under intersection \cap :

$$\forall S \subseteq \overline{A} : \bigcap S \in \overline{A}$$

- In particular $\bigcap \emptyset = \Sigma \in \overline{\mathcal{A}}$.

EXAMPLE OF MOORE FAMILY-BASED ABSTRACTION



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

$$\langle \mathcal{M}(\wp(\wp(\Sigma))), \supseteq, \wp(\Sigma), \{\Sigma\}, \lambda S \cdot \mathcal{M}(\cup S), \cap \rangle$$

where:

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

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

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 *6th POPL*, pages 269–282, San Antonio, TX, 1979. ACM Press. 26

CLOSURE OPERATOR INDUCED BY AN ABSTRACTION

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

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

It is a **closure operator**:

- extensive,
- idempotent,
- isotone/monotonic;

such that

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

hence

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

ABSTRACTION INDUCED BY A CLOSURE OPERATOR

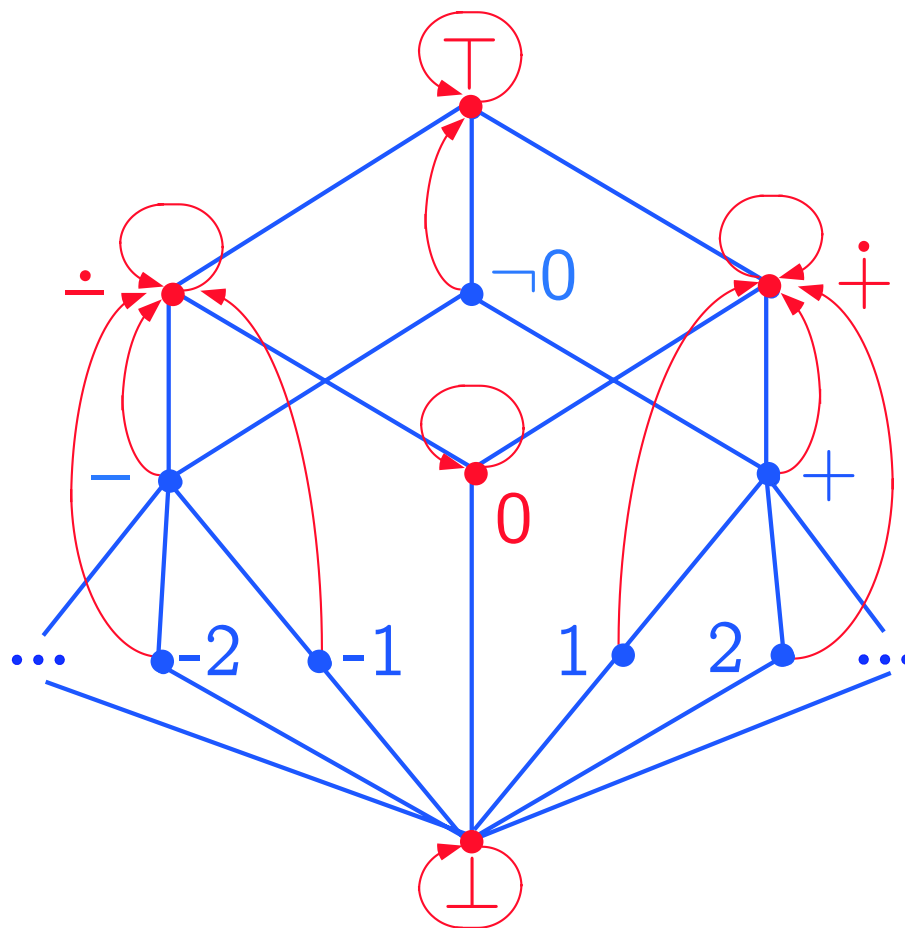
- Any closure operator ρ on the set of properties $\wp(\Sigma)$ induces an abstraction:

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

Examples:

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

EXAMPLE OF CLOSURE OPERATOR-BASED ABSTRACTION



THE LATTICE OF ABSTRACTIONS (2)

- The set $\text{clo}(\wp(\Sigma) \longrightarrow \wp(\Sigma))$ of all abstractions, i.e. isomorphically, closure operators ρ on the set $\wp(\Sigma)$ of concrete properties is the complete **lattice of abstractions** for pointwise inclusion⁴:

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

where:

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

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

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 *6th POPL*, pages 269–282, San Antonio, TX, 1979. ACM Press. 38

CORRESPONDANCE BETWEEN CONCRETE AND ABSTRACT PROPERTIES

- For closure operators ρ , we have:

$$\rho(P) \subseteq \rho(P') \Leftrightarrow P \subseteq P'$$

written:

$$\langle \wp(\Sigma), \subseteq \rangle \xrightleftharpoons[\rho]{1} \langle \rho(\wp(\Sigma)), \subseteq \rangle$$

where 1 is the identity and:

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

means that $\langle \alpha, \gamma \rangle$ is a **Galois connection**:

- $\forall P \in \wp(\Sigma), \overline{P} \in \overline{\mathcal{D}} : \alpha(P) \sqsubseteq \overline{P} \Leftrightarrow P \subseteq \gamma(\overline{P})$;
- α is onto (equivalently $\alpha \circ \gamma = 1$ or γ is one-to-one).

ABSTRACT DOMAIN

- **Abstract Domain:** an isomorphic representation $\overline{\mathcal{D}}$ of the set $\overline{\mathcal{A}} \subseteq \wp(\Sigma) = \rho(\wp(\Sigma))$ of abstract properties (up to some order-isomorphism ι).

GALOIS SURJECTION⁶

- We have the Galois surjection:

$$\langle \wp(\Sigma), \subseteq \rangle \xrightleftharpoons[\iota \circ \rho]{\iota^{-1}} \langle \overline{\mathcal{D}}, \sqsubseteq \rangle$$

- More generally:

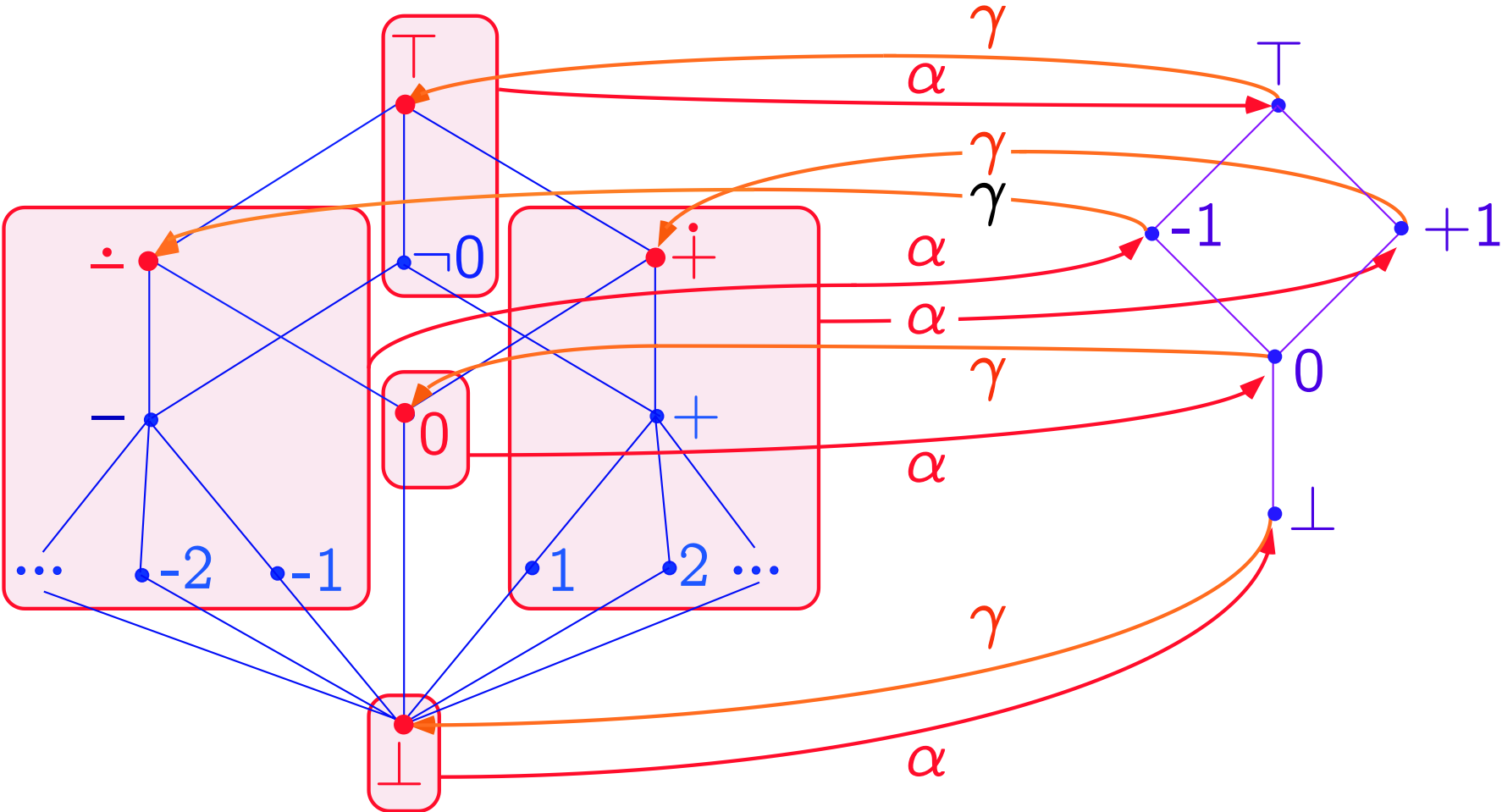
$$\langle \wp(\Sigma), \subseteq \rangle \begin{matrix} \xleftarrow{\gamma} \\ \xrightarrow{\alpha} \end{matrix} \langle \overline{\mathcal{D}}, \sqsubseteq \rangle$$

denoting (again) the fact that:

- $\forall P \in \wp(\Sigma), \overline{P} \in \overline{\mathcal{D}} : \alpha(P) \sqsubseteq \overline{P} \Leftrightarrow P \subseteq \gamma(\overline{P});$
- α is onto (equivalently $\alpha \circ \gamma = 1$ or γ is one-to-one).

⁶ Also called Galois insertion since γ is injective.

EXAMPLE OF GALOIS SURJECTION-BASED ABSTRACTION



GALOIS CONNECTION

- Relaxing the condition that α is onto:

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

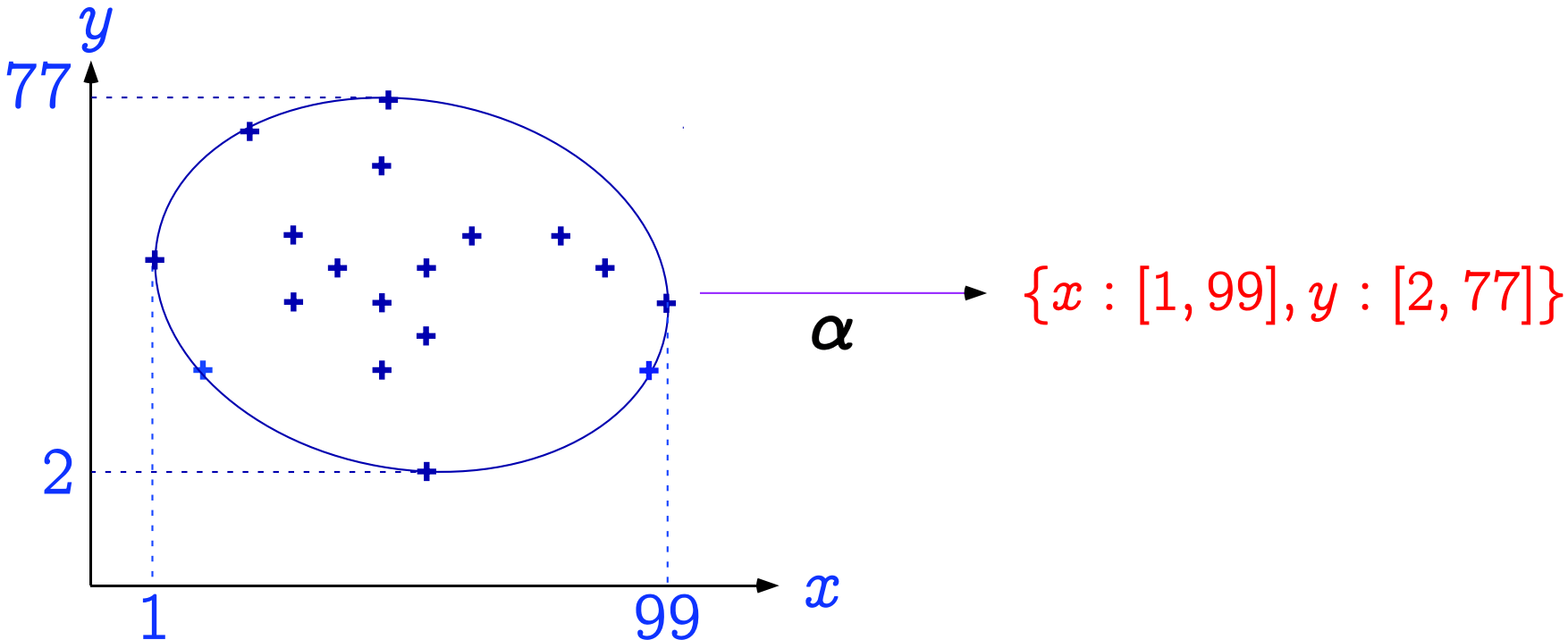
that is to say:

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

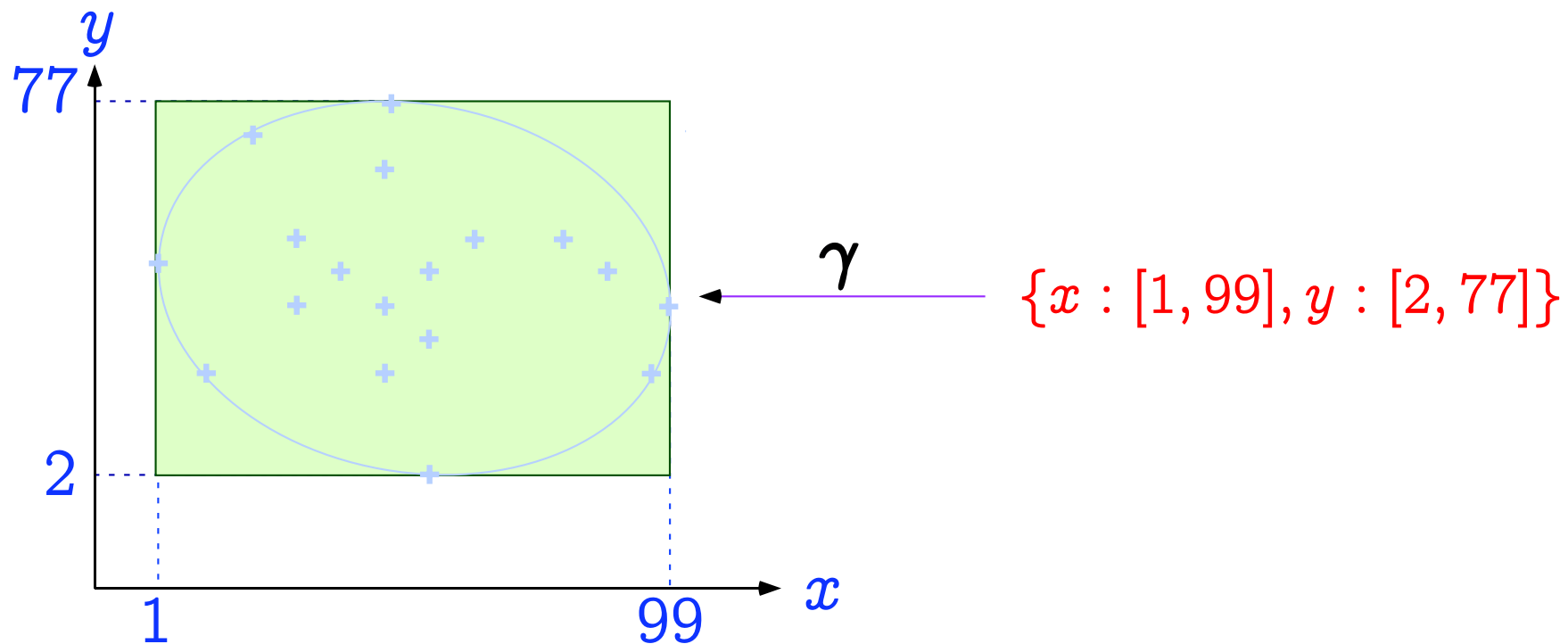
- i.e. ρ is now $\gamma \circ \alpha$;

We can now have different representations of the same abstract property.

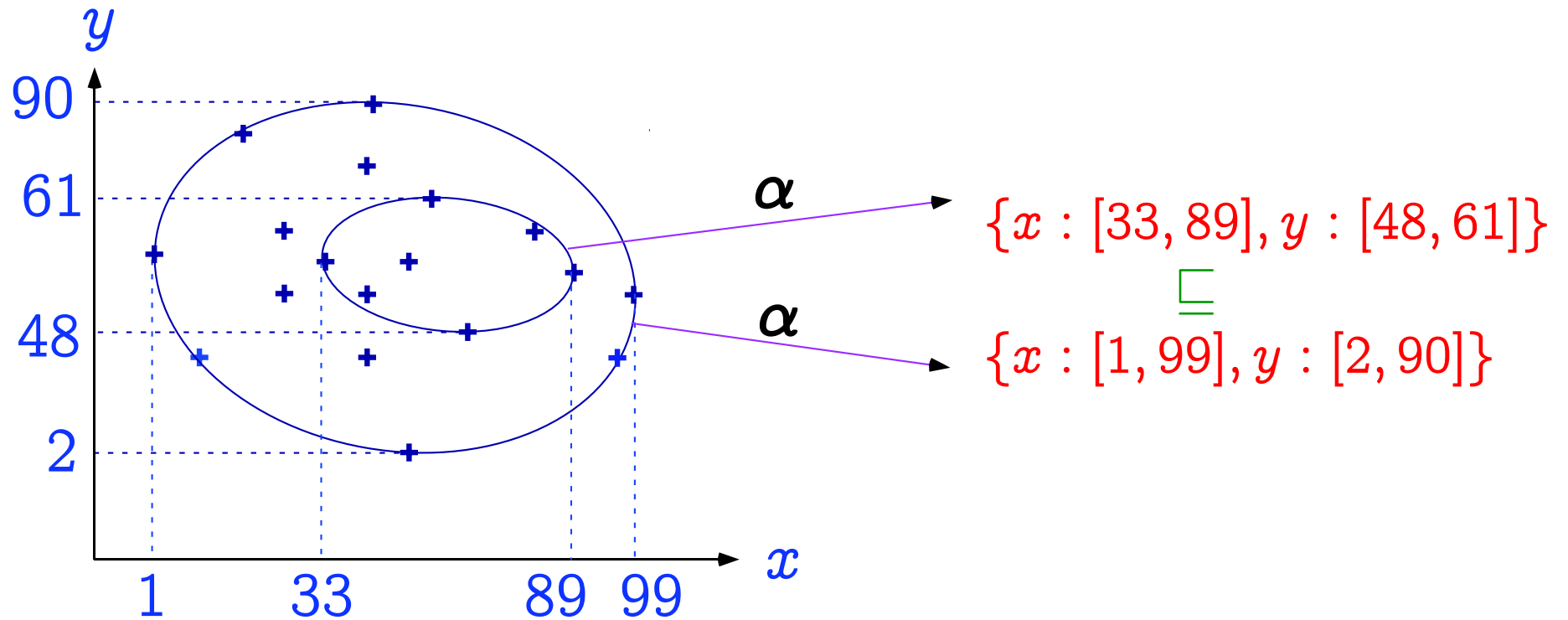
ABSTRACTION α



CONCRETIZATION γ

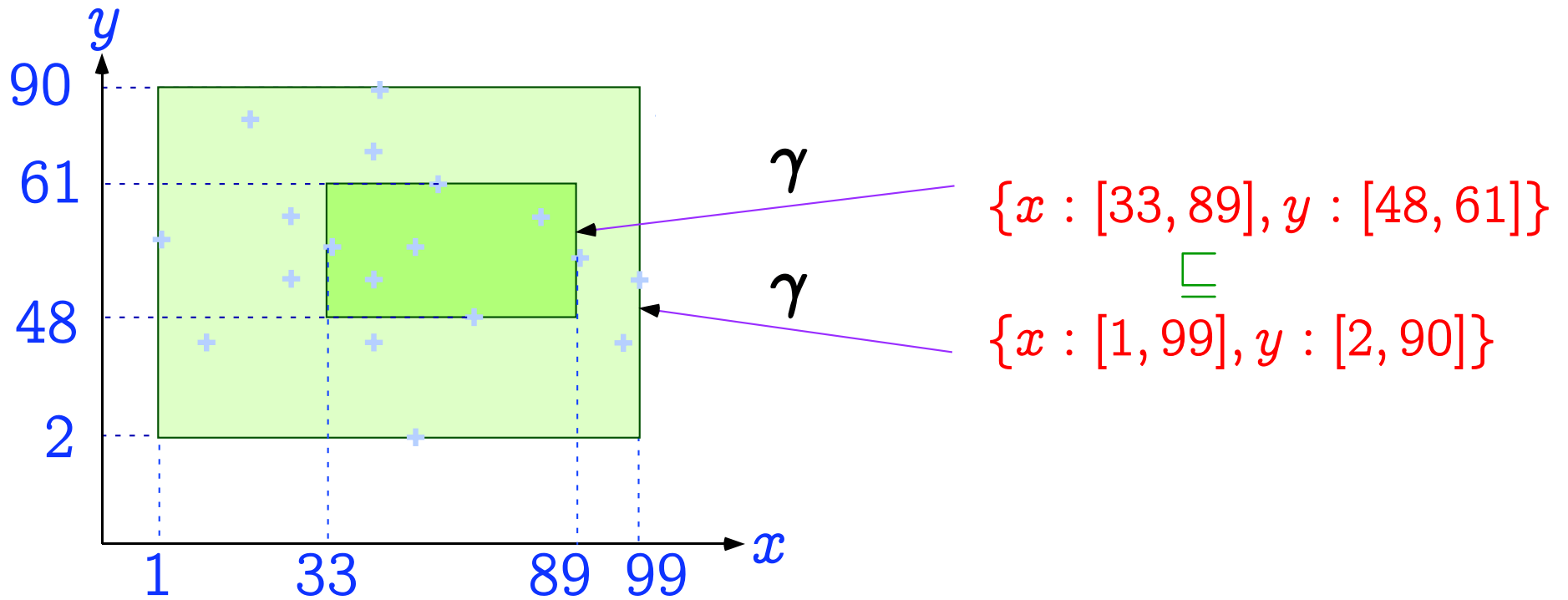


THE ABSTRACTION α IS MONOTONE



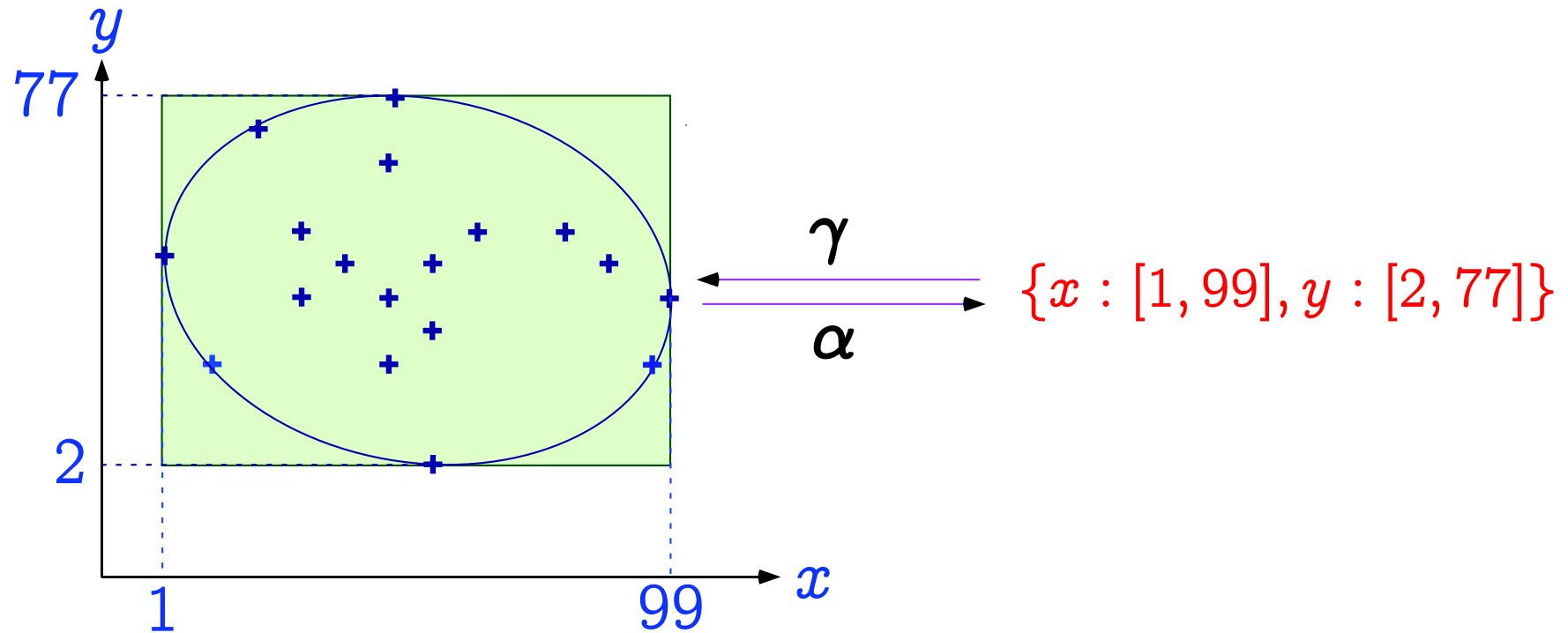
$$X \subseteq Y \Rightarrow \alpha(X) \sqsubseteq \alpha(Y)$$

THE CONCRETIZATION γ IS MONOTONE



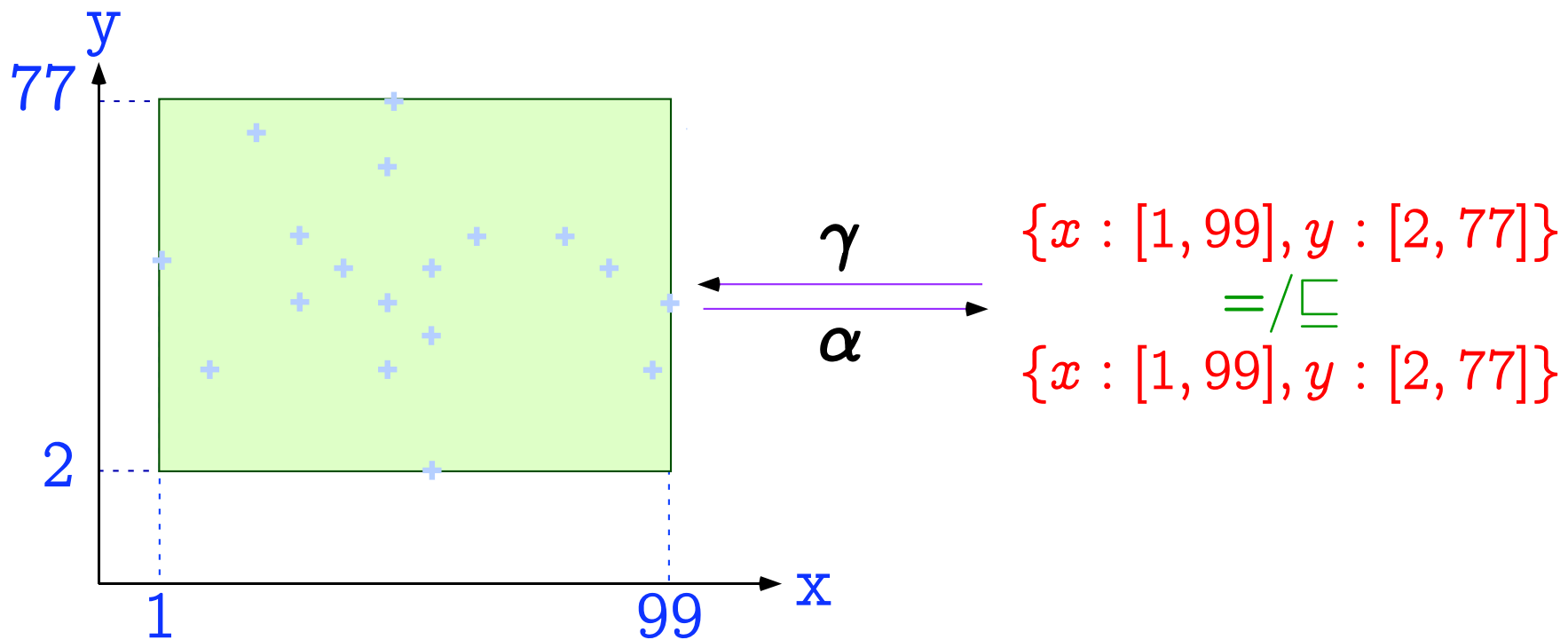
$$X \sqsubseteq Y \Rightarrow \gamma(X) \subseteq \gamma(Y)$$

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



$$X \subseteq \gamma \circ \alpha(X)$$

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



$$\alpha \circ \gamma(Y) =_{/\sqsubseteq} Y$$

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 *6th POPL*, pages 269–282, San Antonio, TX, 1979. ACM Press. 51

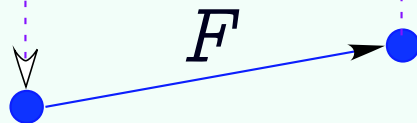
Abstract domain



FUNCTION ABSTRACTION

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

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



Concrete domain

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

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

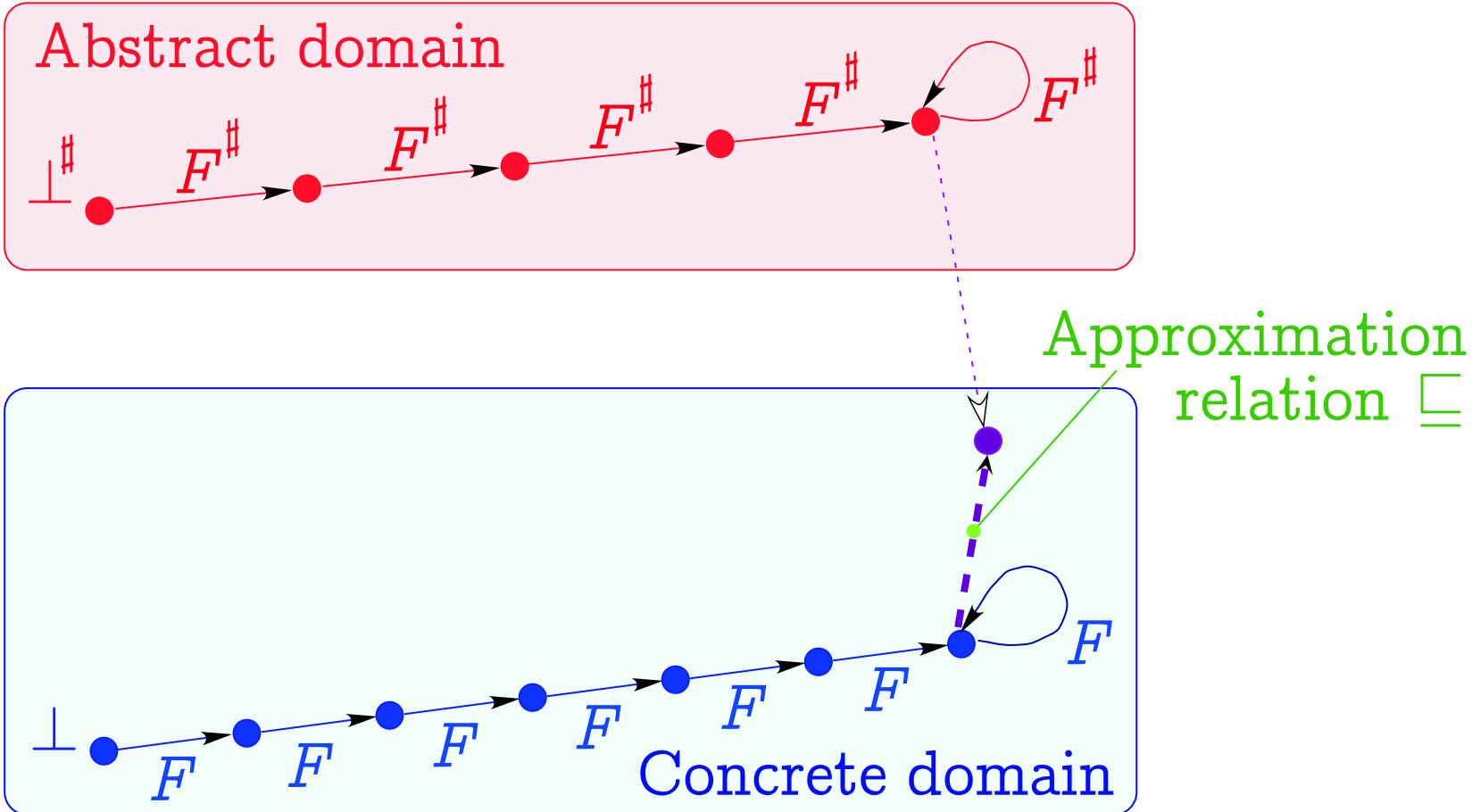
2.2.6 FIXPOINT ABSTRACTION

See Sec. 7.1 of [POPL '79].

Reference

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

APPROXIMATE FIXPOINT ABSTRACTION



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

APPROXIMATE/EXACT FIXPOINT ABSTRACTION

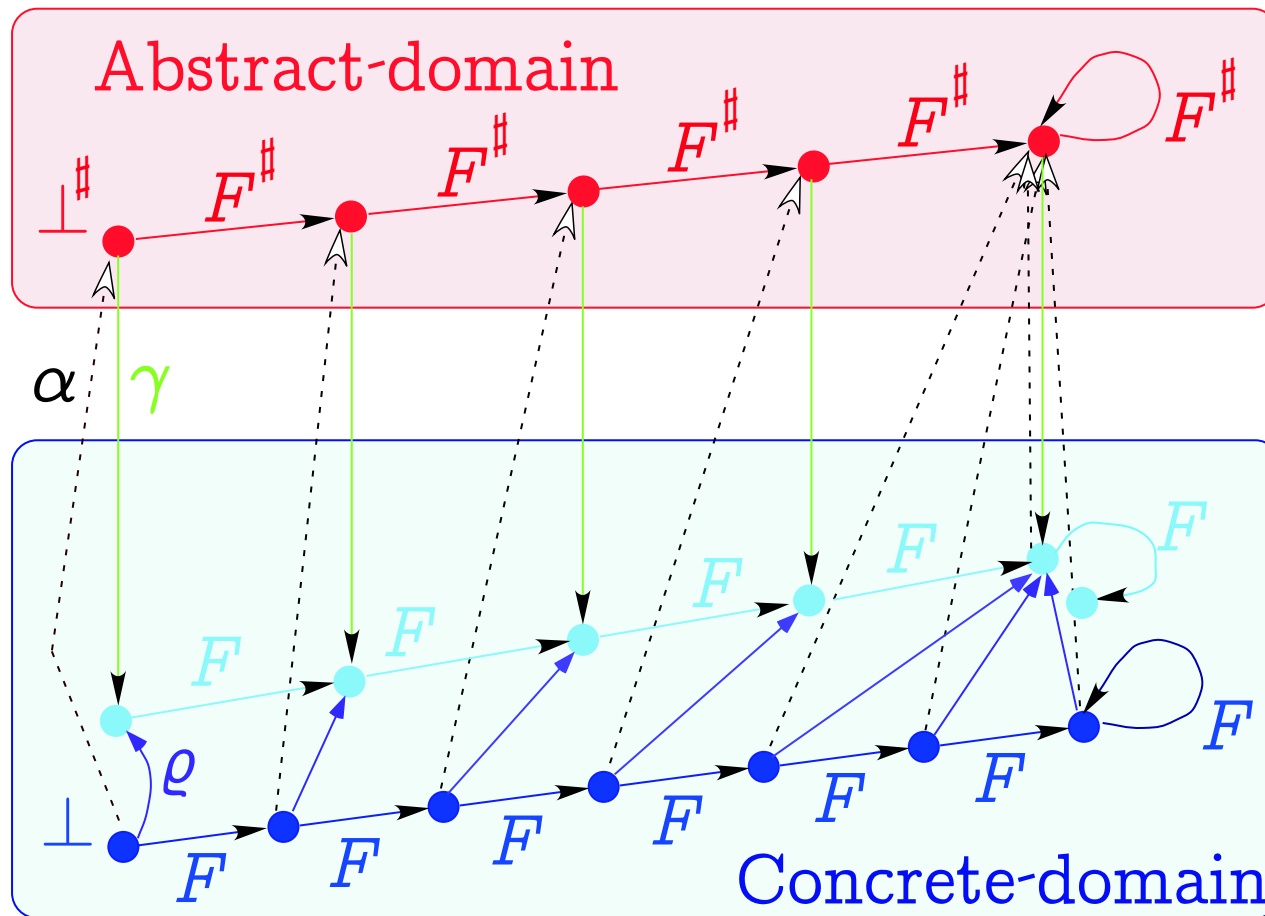
Exact Abstraction:

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

Approximate Abstraction:

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

EXACT FIXPOINT ABSTRACTION



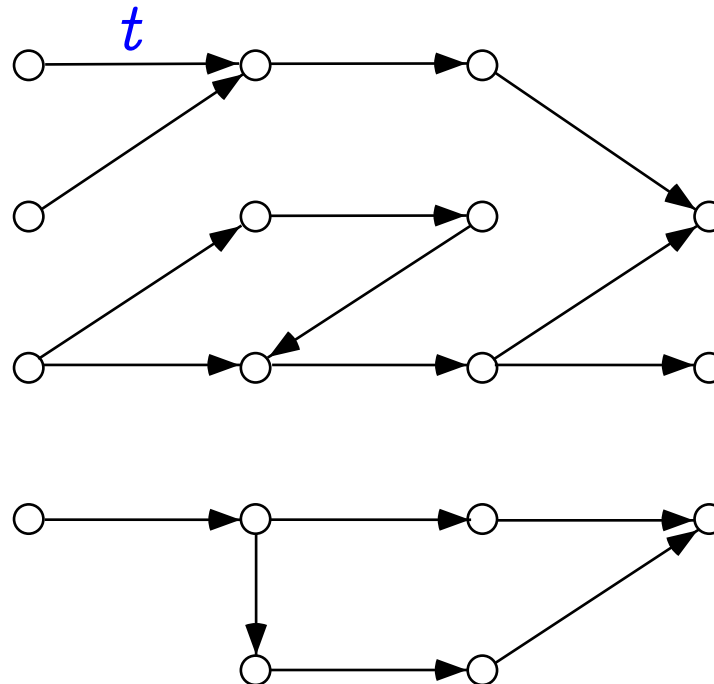
$$\alpha \circ F = F^\# \circ \alpha \Rightarrow \alpha(\text{lfp } F) = \text{lfp } F^\#$$

2.3 APPLICATION TO REACHABILITY

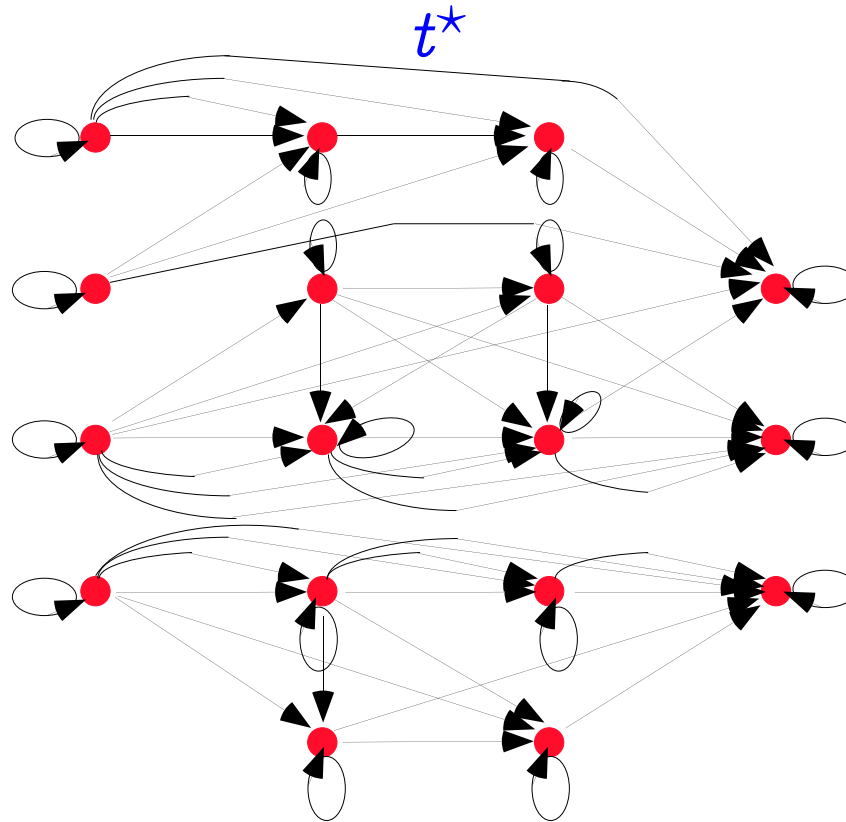
TRANSITION SYSTEMS

- $\langle S, t \rangle$ 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

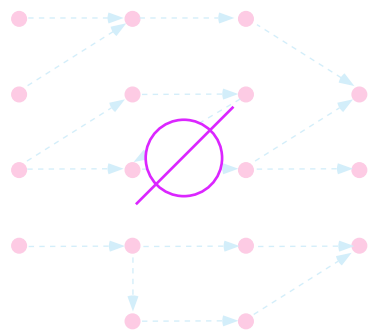
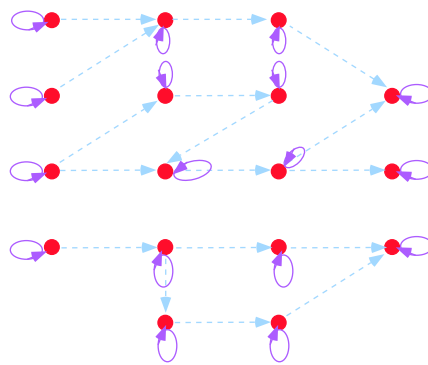
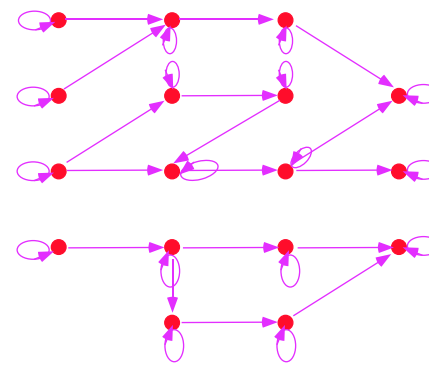
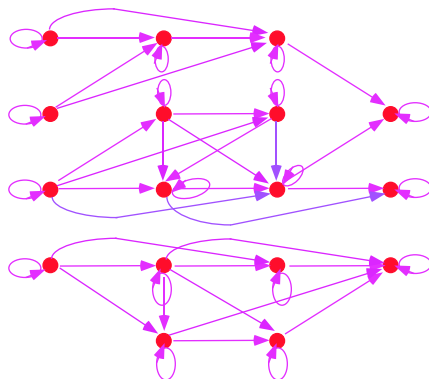
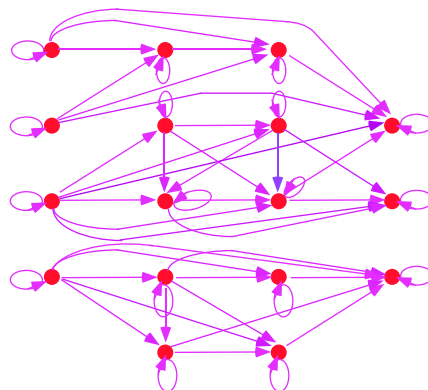
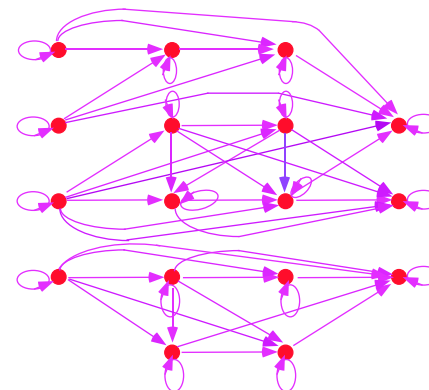


$$\begin{aligned}
X^{n+1} &= t^0 \cup X^n \circ t \\
&= t^0 \cup \left(\bigcup_{0 \leq i < n} t^i \right) \circ t \\
&= t^0 \cup \bigcup_{0 \leq i < n} (t^i \circ t) \\
&= t^0 \cup \bigcup_{1 \leq i+1 < n+1} (t^{i+1}) \\
&= t^0 \cup \left(\bigcup_{1 \leq j < n+1} t^j \right) \circ t \\
&= \bigcup_{0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

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

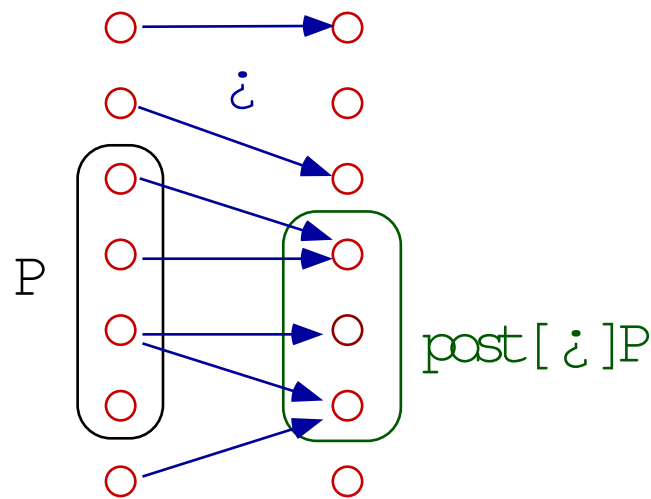
$$\begin{aligned}
X^{\omega+1} &= t^0 \cup X^{\omega} \circ t \\
&= t^0 \cup \left(\bigcup_{n \geq 0} t^n \right) \circ t \\
&= t^0 \cup \bigcup_{n \geq 0} (t^n \circ t) \\
&= t^0 \cup \bigcup_{n \geq 0} t^{n+1} \\
&= t^0 \cup \bigcup_{k \geq 1} t^k \\
&= \bigcup_{n \geq 0} t^n \\
&= t^*
\end{aligned}$$

ITERATES

 x^0  x^1 
$$X^2$$
 x^3  x^4 
$$X^5 = t^*$$

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.

POSTIMAGE GALOIS CONNECTION

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

$$\langle \wp(S \times S), \subseteq \rangle \begin{array}{c} \xleftarrow{\gamma} \\ \xrightarrow{\lambda t \cdot \text{post}[t]I} \end{array} \langle \wp(S), \subseteq \rangle$$

$$\text{post}[t]I \subseteq R$$

$$\Leftrightarrow \{s' \mid \exists s \in I : \langle s, s' \rangle \in t\} \subseteq R$$

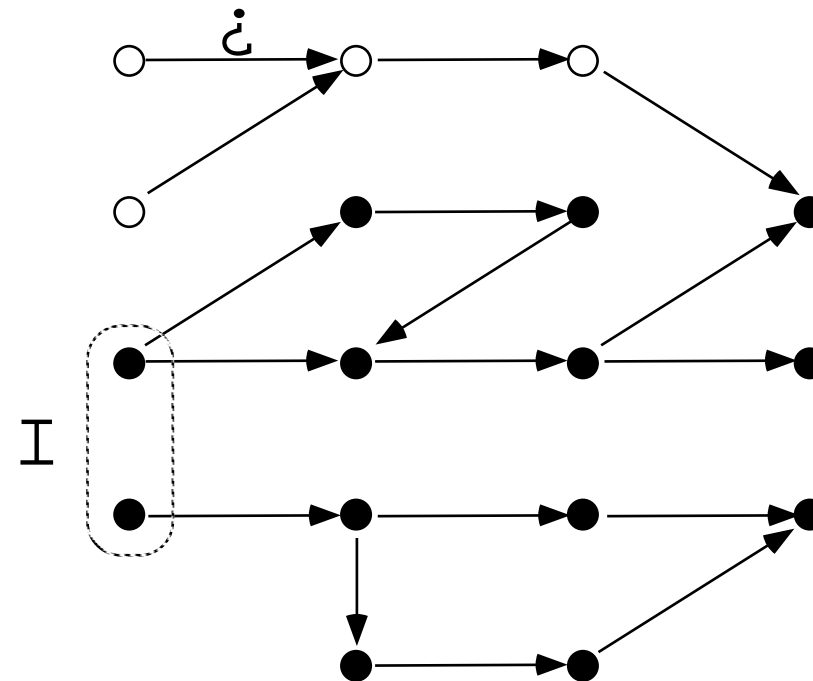
$$\Leftrightarrow \forall s' \in S : (\exists s \in I : \langle s, s' \rangle \in t) \Rightarrow (s' \in R)$$

$$\Leftrightarrow \forall s', s \in S : (s \in I \wedge \langle s, s' \rangle \in t) \Rightarrow (s' \in R)$$

$$\Leftrightarrow \forall s', s \in S : \langle s, s' \rangle \in t \Rightarrow ((s \in I) \Rightarrow (s' \in R))$$

$$\Leftrightarrow t \subseteq \{\langle s, s' \rangle \mid (s \in I) \Rightarrow (s' \in R)\} \stackrel{\text{def}}{=} \gamma(R)$$

REACHABLE STATES


$$\text{post}[t^*]\mathcal{I}$$

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 $\langle L, \sqsubseteq \rangle \xrightleftharpoons[\alpha]{\gamma} \langle \overline{L}, \overline{\sqsubseteq} \rangle$ such that $\alpha \circ F \circ \gamma \sqsubseteq^* \overline{F}$. Then¹⁰:

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

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

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

FIXPOINT ABSTRACTION (CONTINUED)

Moreover, the *commutation condition* $\overline{F} \circ \alpha = \alpha \circ F$ implies¹¹:

- $\overline{F} = \alpha \circ F \circ \gamma$, and
- $\alpha(\text{lfp}^{\sqsubseteq} F) = \text{lfp}^{\sqsubseteq} \overline{F}$;

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

¹¹ P. Cousot & R. Cousot. *Systematic design of program analysis frameworks*. ACM POPL'79, pp. 269–282, 1979. Numerous variants!

$$\text{post}[t^0]I$$

$$= \{s' \mid \exists s \in I : \langle s, s' \rangle \in t^0\}$$

$$= \{s' \mid \exists s \in I : \langle s, s' \rangle \in \{\langle s, s \rangle \mid s \in S\}\}$$

$$= \{s' \mid \exists s \in I\}$$

$$= I$$

$\text{post}[X \circ t]I$

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

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

by defining:

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

proving:

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

EXAMPLE OF ITERATION

