Computer aided verification

lecture 13

Abstract interpretation II

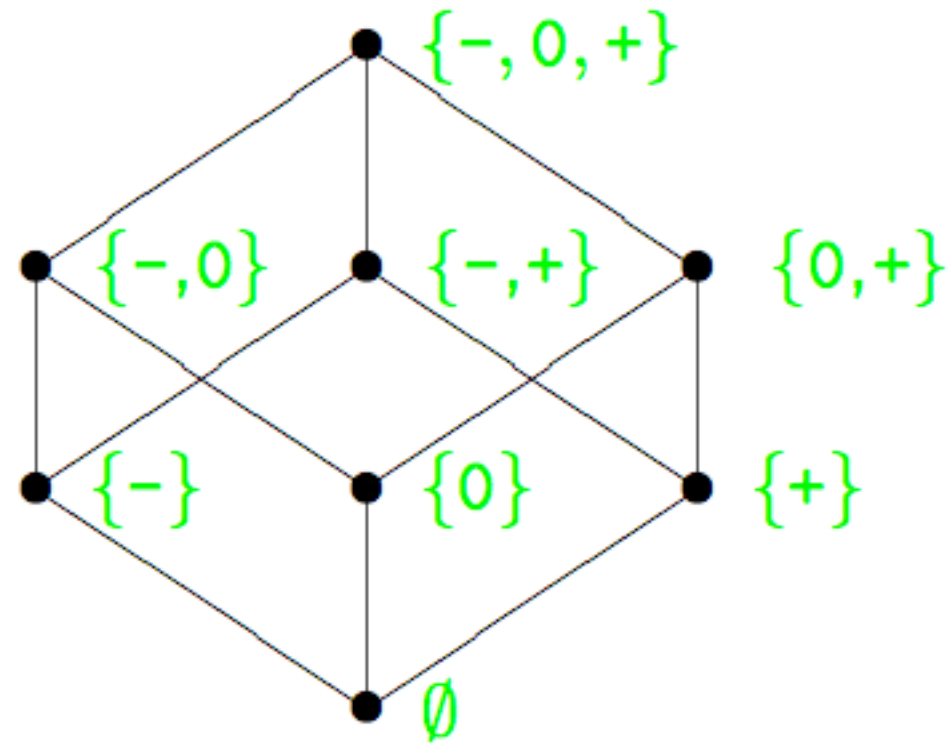
Literature

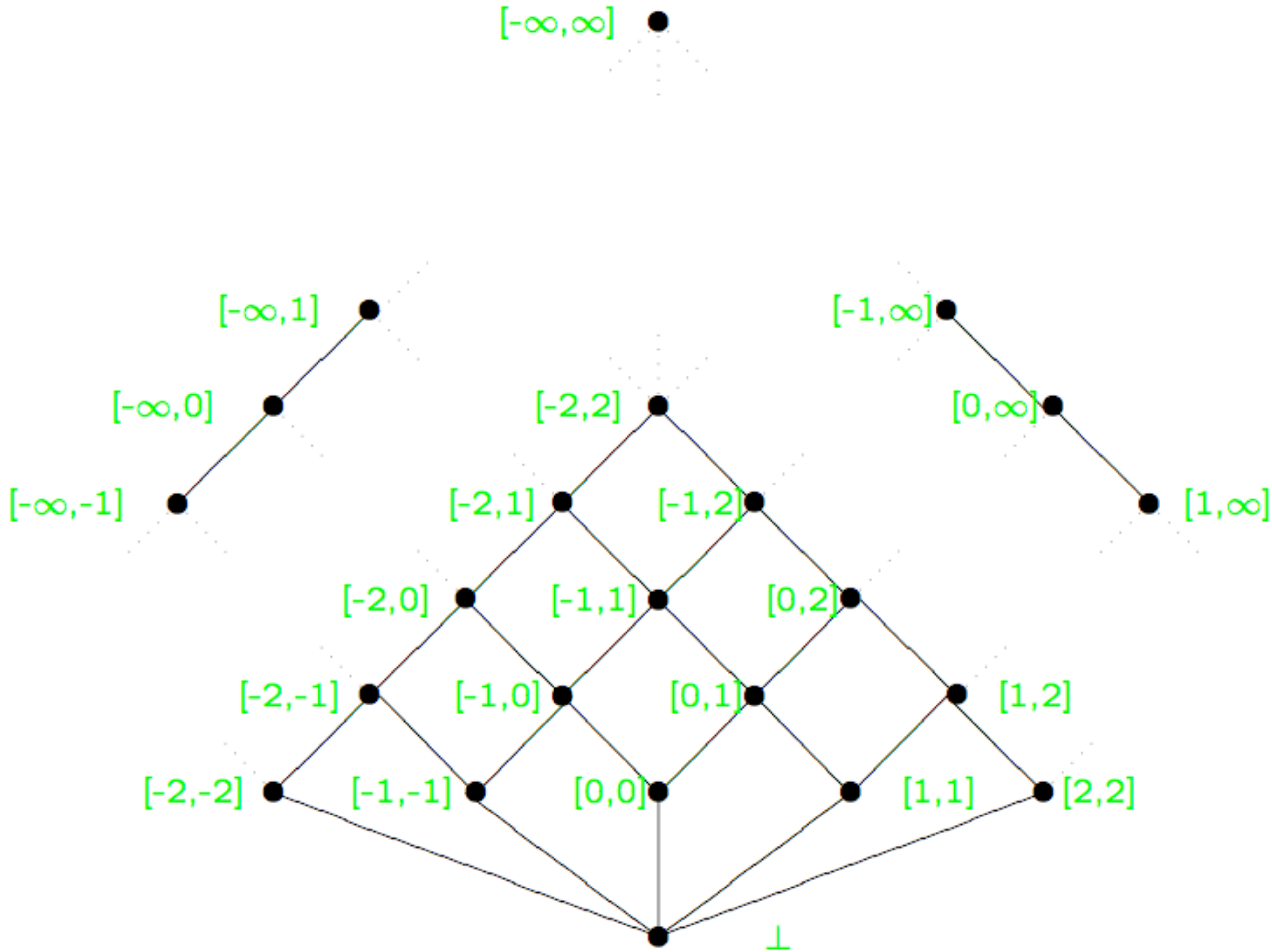
- F. Nielson, H.R. Nielson, C. Hankin, Principles of Program Analysis, Springer, 2005.
- http://www.imm.dtu.dk/~riis/PPA/slides4.pdf
- N. D. Jones, F. Nielson, Interpretation: a Semantics-Based Tool for Program Analysis. Handbook of Logic in Computer Science, tom 4, str. 527-636, 1995.
- V. D'Silva, D. Kroening, G. Weissenbacher, A Survey of Automated Techniques for Formal Software Verification. IEEE Trans. on CAD of Integrated Circuits and Systems 27 (7):1165-1178, 2008.

Abstract domains

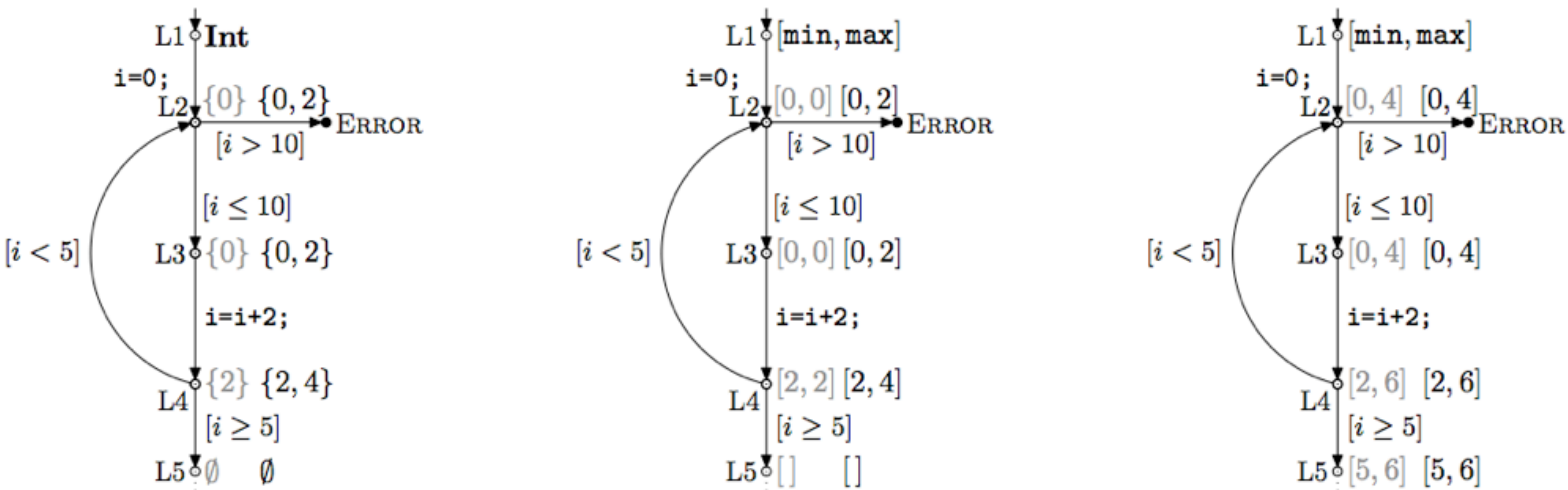
Non-relational domains

- signs $\mathcal{P}(-, 0, +)$
- intervals $[n, m]$
- parity
- congruence modulo k




[-∞,∞]
[-∞,1]
[-1,∞]
[-∞,0]
[-2,2]
[0,∞]
[-∞,-1]
[-2,1]
[-1,2]
[1,∞]
[-2,0]
[-1,1]
[0,2]
[-2,-1]
[-1,0]
[0,1]
[1,2]
[-2,-2]
[-1,-1]
[0,0]
[1,1]
[2,2]
⊥

```
int i = 0;
do {
  assert(i <= 10);
  i = i+2;
} while (i < 5);
```



Relational domains

- DBM (difference bounds matrices) $x - y \leq c$
- octagon $\pm x \pm y \leq c$
- octahedra $\pm x_1 \ldots \pm x_n \leq c$
- polyhedra $a_1 x_1 + \ \ldots \ + a_n x_n \leq c$
- ellipsoid $ax^2 + bxy + cy^2 \leq n$
- linear congruence $ax + by = c \bmod k$

Expressive power

precision ↓

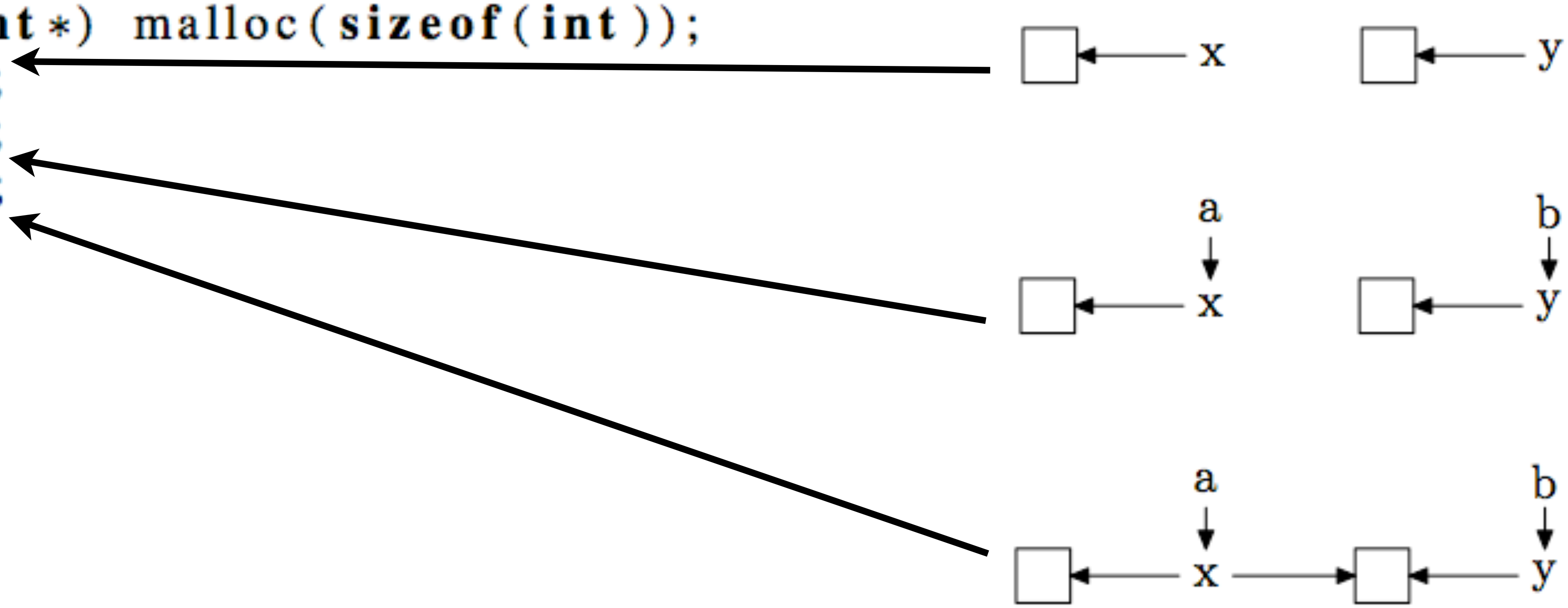
- signs $0 \leq x$
- intervals $c \leq x \leq d$
- DBM (difference bounds matrices) $x - y \leq c$
- octagon $\pm x \pm y \leq c$
- octahedra $\pm x_1 \ldots \pm x_n \leq c$
- polyhedra $a_1 x_1 + \ \ldots \ + a_n x_n \leq c$

Pointer analyses domains

- points-to graphs

Example: alias analysis

```
int **a, **b, *x, *y;
x = (int*) malloc(sizeof(int));
y = (int*) malloc(sizeof(int));
a = &x;
b = &y;
*a = y;
```



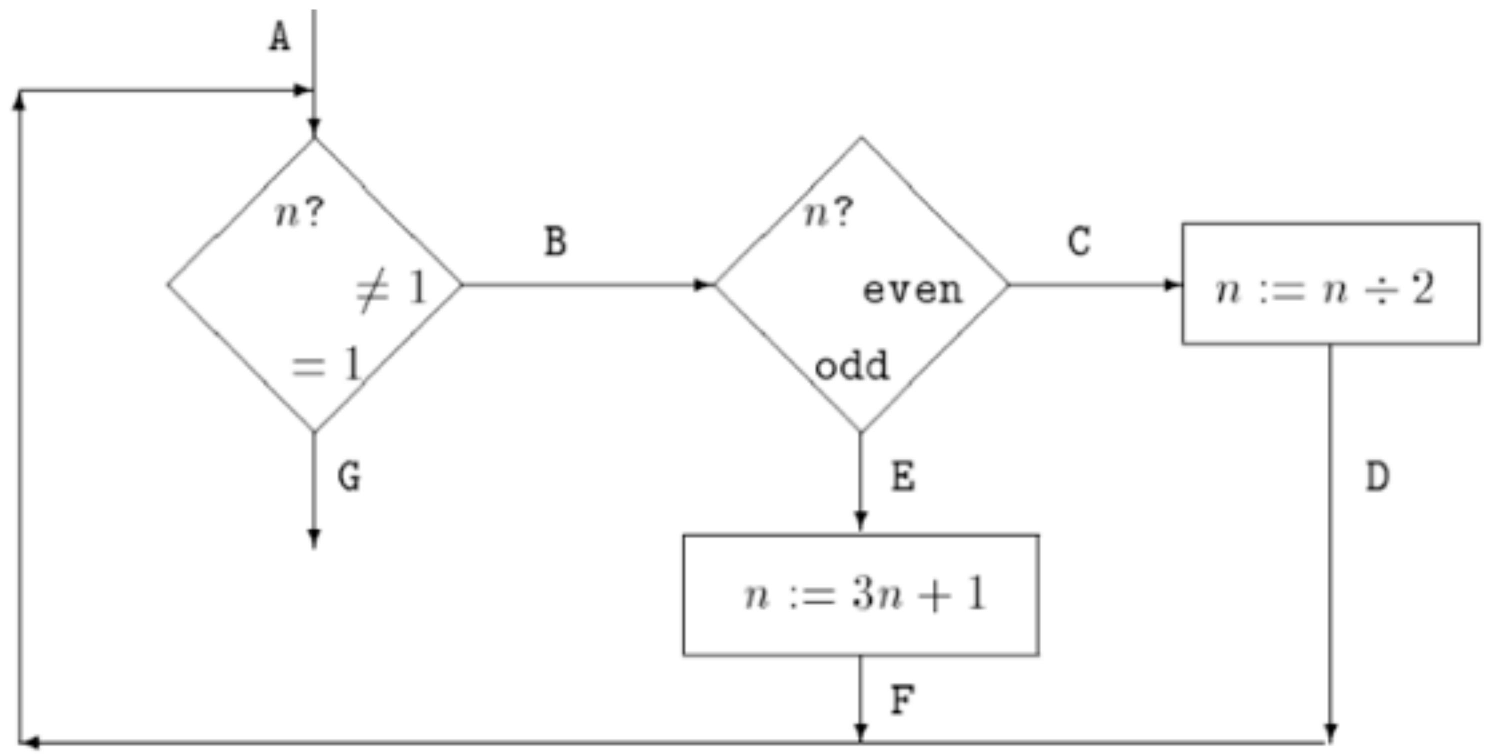
a and b do not point to the same location

x and y may point to the same location

Composition of analyses

Abstract semantics

A: **while** $n \neq 1$ **do**
  B: **if** n even
    then (C: $n := n \div 2$; D:)
    else (E: $n := 3 * n + 1$; F:)
  fi
od
G:

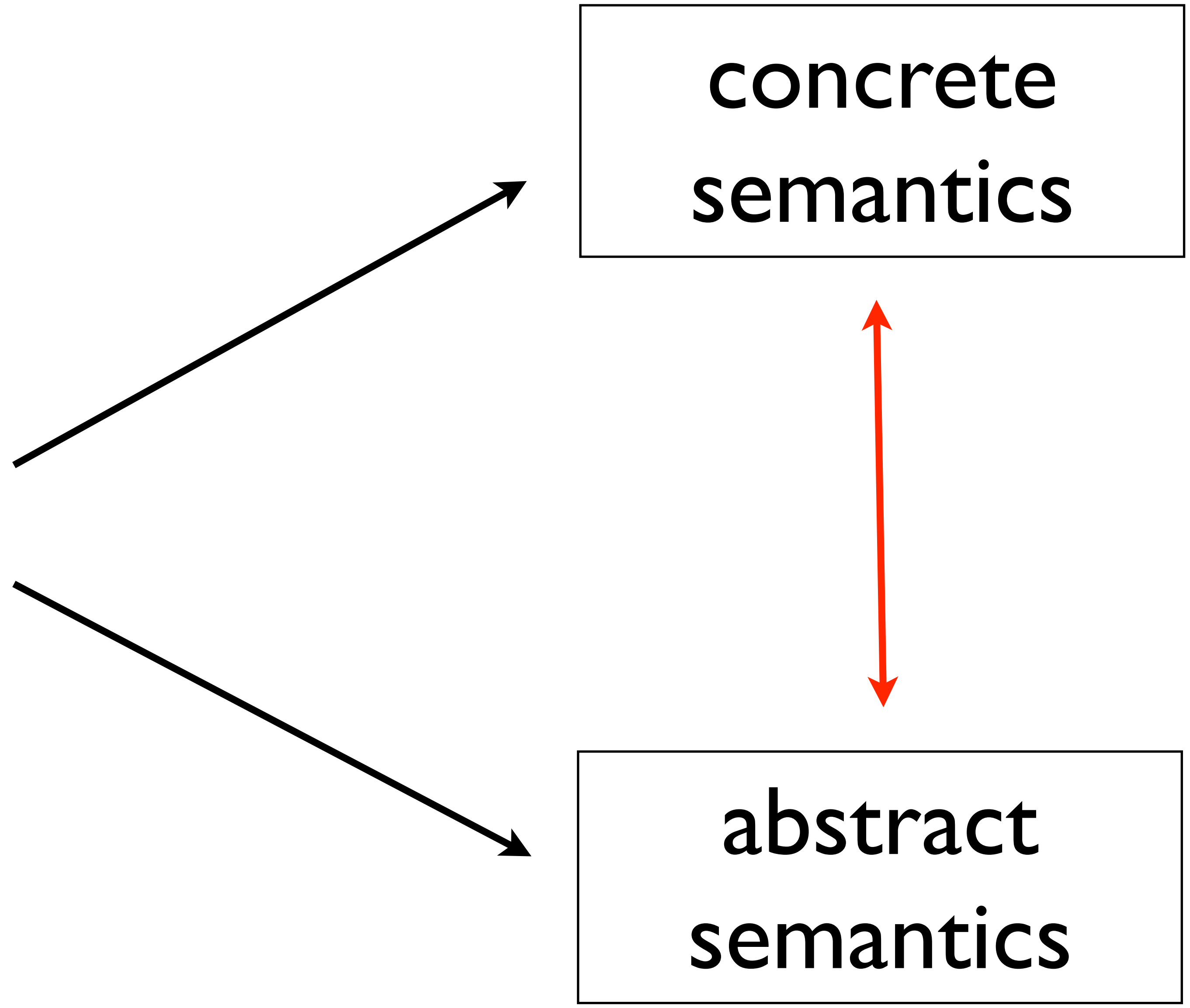


$$S = \{\text{A, B, C, D, E, F, G}\}$$

$$\text{State} = S \times \text{Store}$$

$$\text{Store} = \text{Var} \to \text{Val}$$

```
A: while n ≠ 1 do
    B: if n even
        then (C: n := n ÷ 2; D: )
        else (E: n := 3 * n + 1; F: )
      fi
    od
G:
```



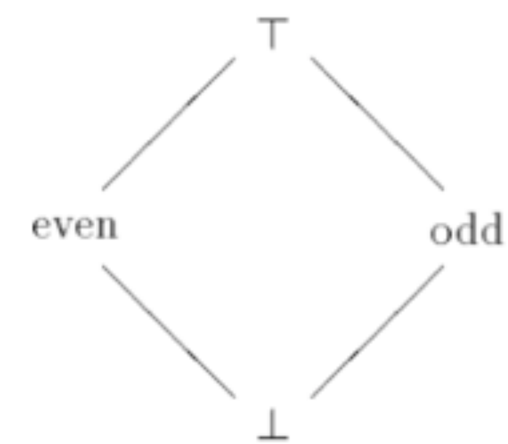
Domains

concrete semantics

abstract semantics

$$V \;=\; \mathrm{Store} \;=\; \mathrm{Var} \to \mathbb{Z}$$

$$L \;=\; \mathrm{Var} \to \{\bot, \mathrm{even}, \mathrm{odd}, \top\}$$



Abstract semantics

$n := n \div 2;$

$$\bot \mapsto \bot$$
$$\text{odd}, \text{even}, \top \mapsto \top$$

$n := 3 * n + 1;$

$$\bot \mapsto \bot$$
$$\text{odd} \mapsto \text{even}$$
$$\text{even} \mapsto \text{odd}$$
$$\top \mapsto \top$$

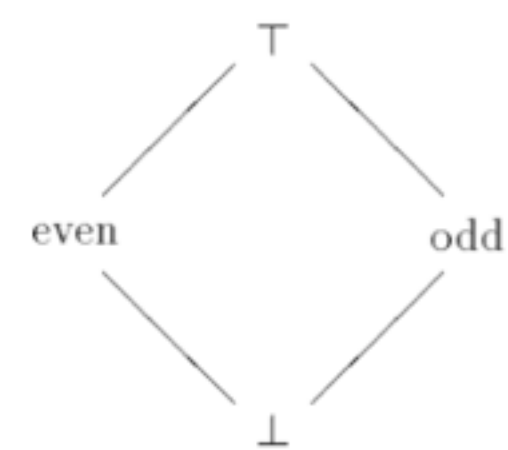
concrete semantics

abstract semantics

$$V \;=\; \text{Store} \;=\; \text{Var} \to \mathbb{Z}$$

do these two semantics agree?

$$L \;=\; \text{Var} \to \{\bot, \text{even}, \text{odd}, \top\}$$



Representation function

concrete semantics

$$V \;=\; \text{Store} \;=\; \text{Var} \to \mathbb{Z}$$

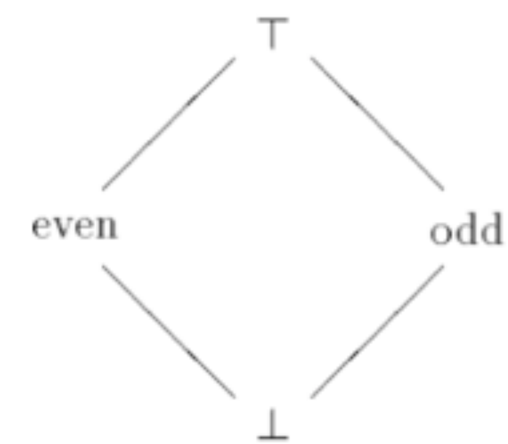
$\beta : V \to L$

monotonic

$$\beta(v) = \begin{cases} \text{even} & \text{if v even} \\ \text{odd} & \text{if v odd} \end{cases}$$

abstract semantics

$$L \;=\; \text{Var} \to \{\bot, \text{even}, \text{odd}, \top\}$$



Representation function

concrete semantics

the best approximation

$$V \;=\; \text{Store} \;=\; \text{Var} \to \mathbb{Z}$$

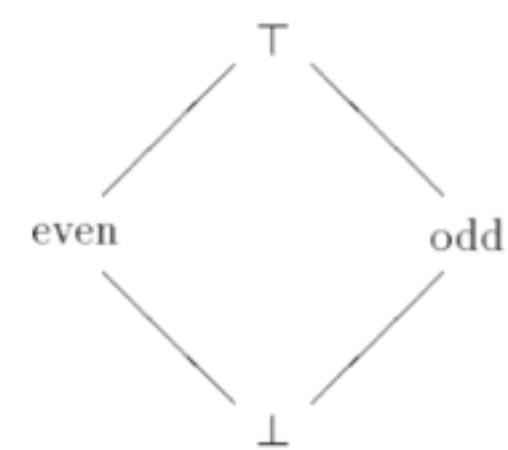
$\beta : V \to L$

monotonic

$$\beta(v) = \begin{cases} \text{even} & \text{if v even} \\ \text{odd} & \text{if v odd} \end{cases}$$

abstract semantics

$$L \;=\; \text{Var} \to \{\bot, \text{even}, \text{odd}, \top\}$$



Representation function

concrete semantics

$$V \;=\; \text{Store} \;=\; \text{Var} \to \mathbb{Z}$$

$\beta : V \to L$

monotonic

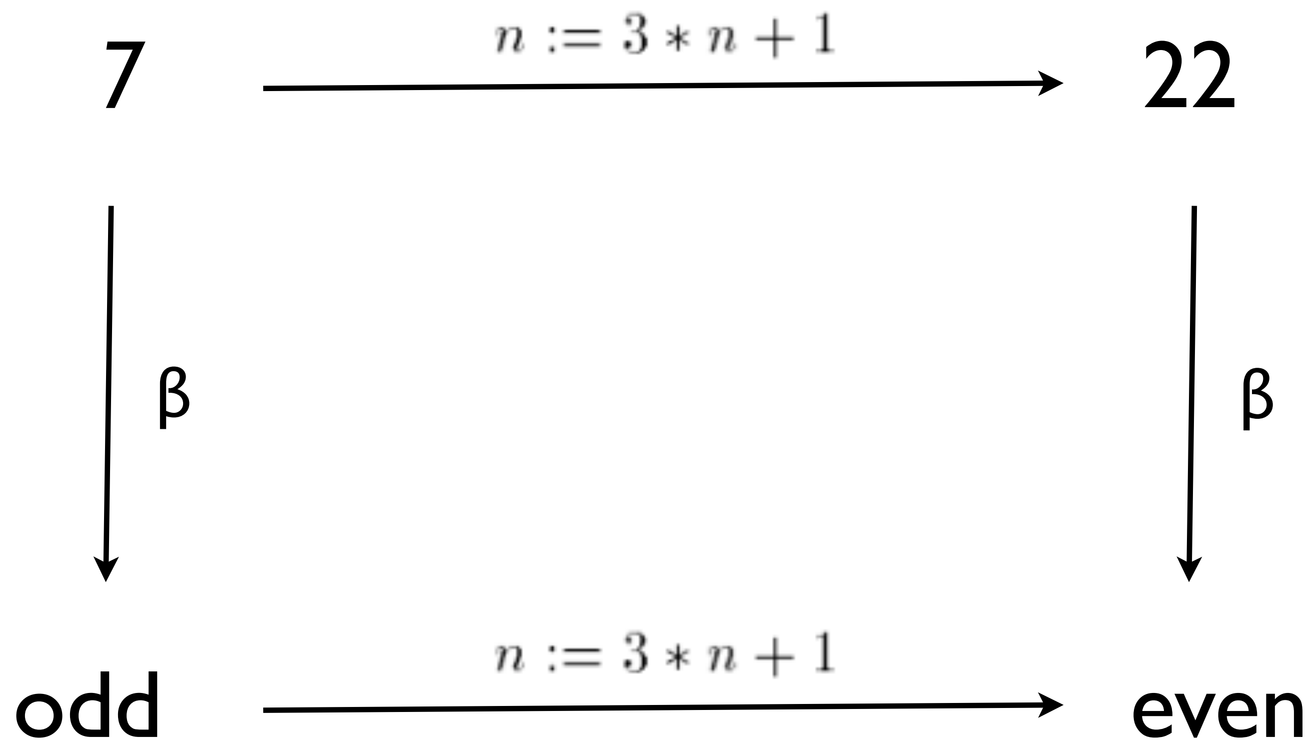
$$\beta(v) = \begin{cases} \{\text{even}\} & \text{if v even} \\ \{\text{odd}\} & \text{if v odd} \end{cases}$$

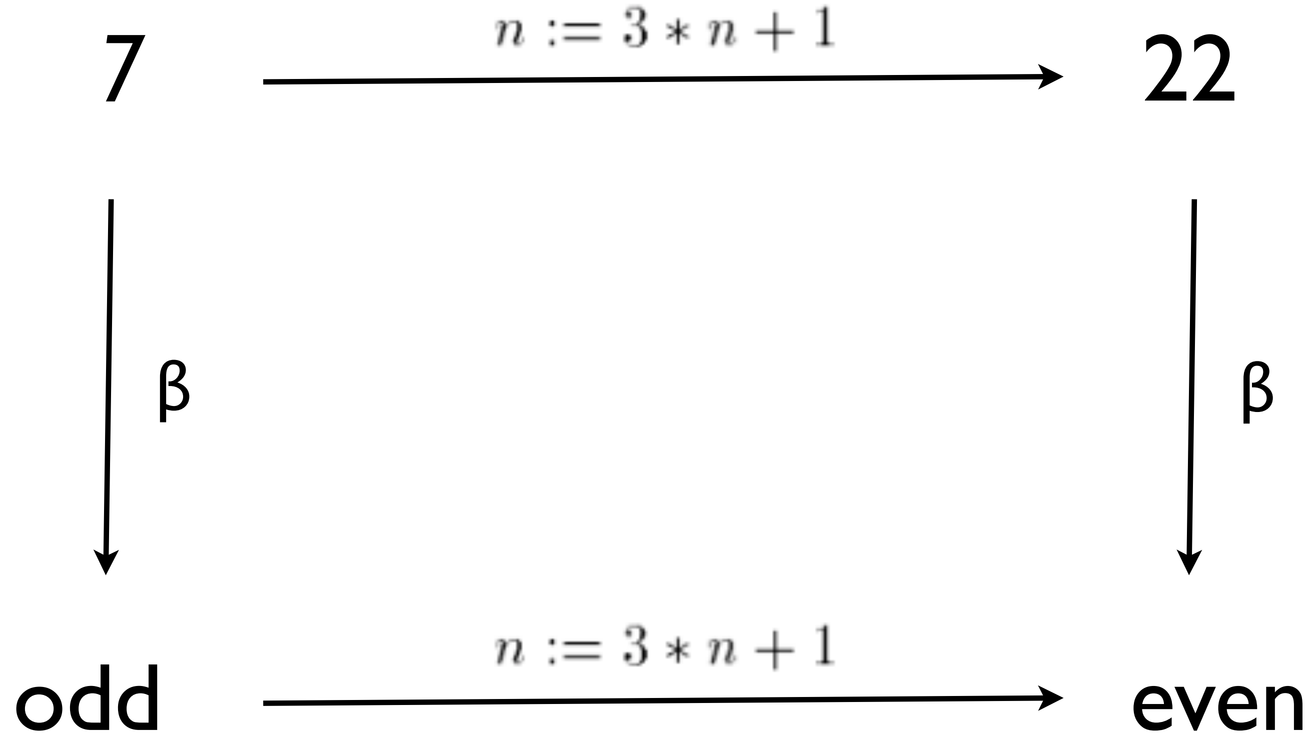
abstract semantics

$$L \;=\; \text{Var} \to \{\bot, \text{even}, \text{odd}, \top\}$$

$$\cong$$

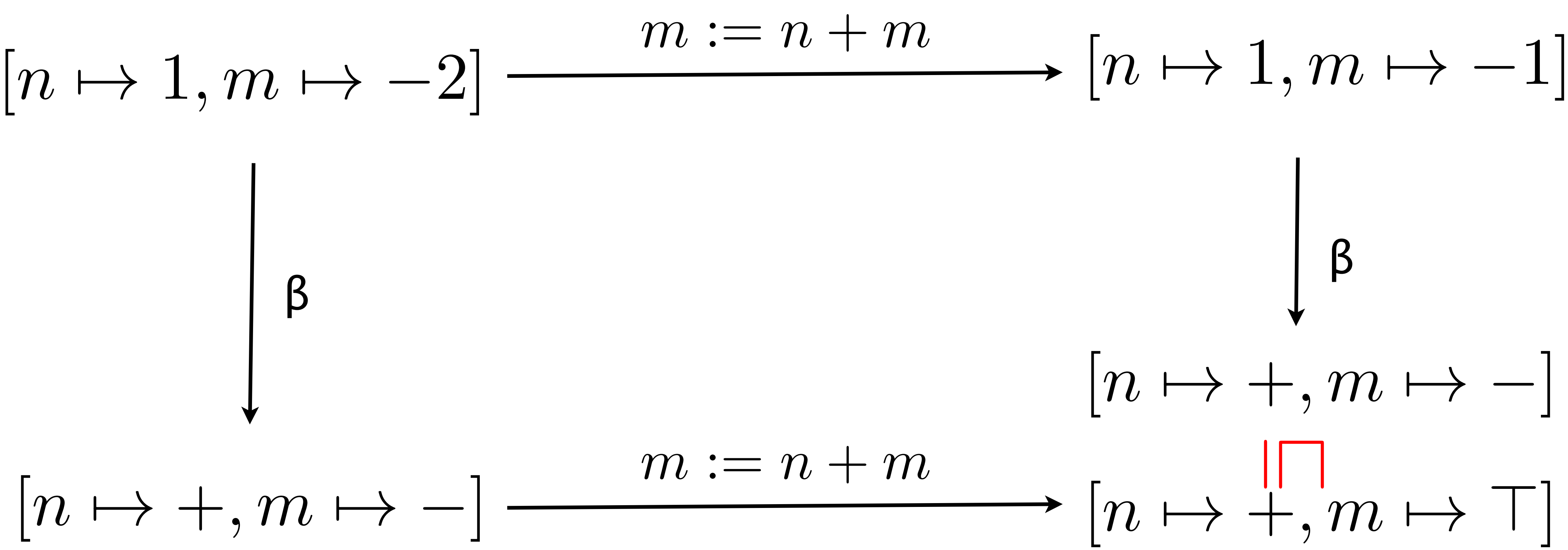
$$\mathcal{P}(\text{even, odd})$$

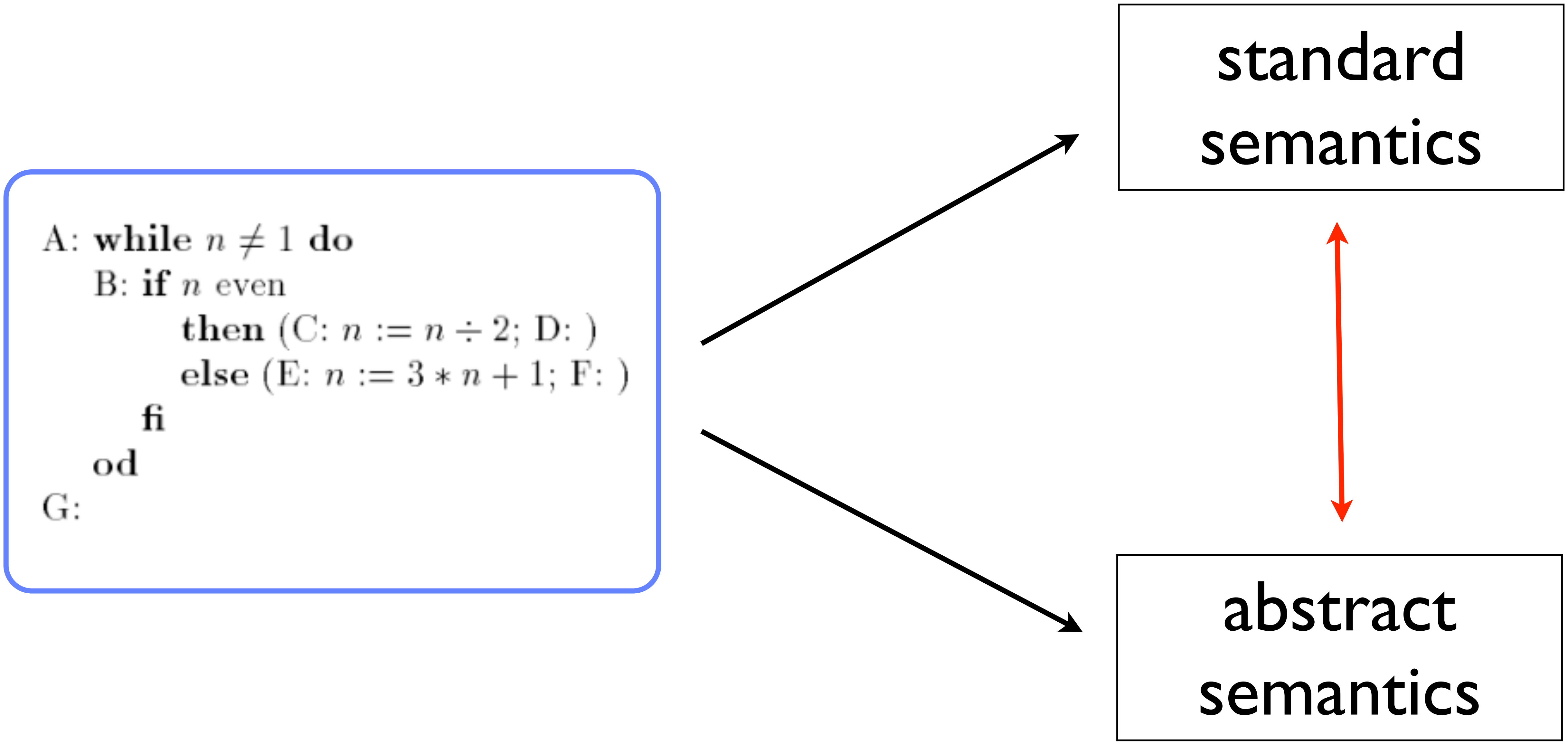

7
n := 3 * n + 1
22
β
β
odd
n := 3 * n + 1
even

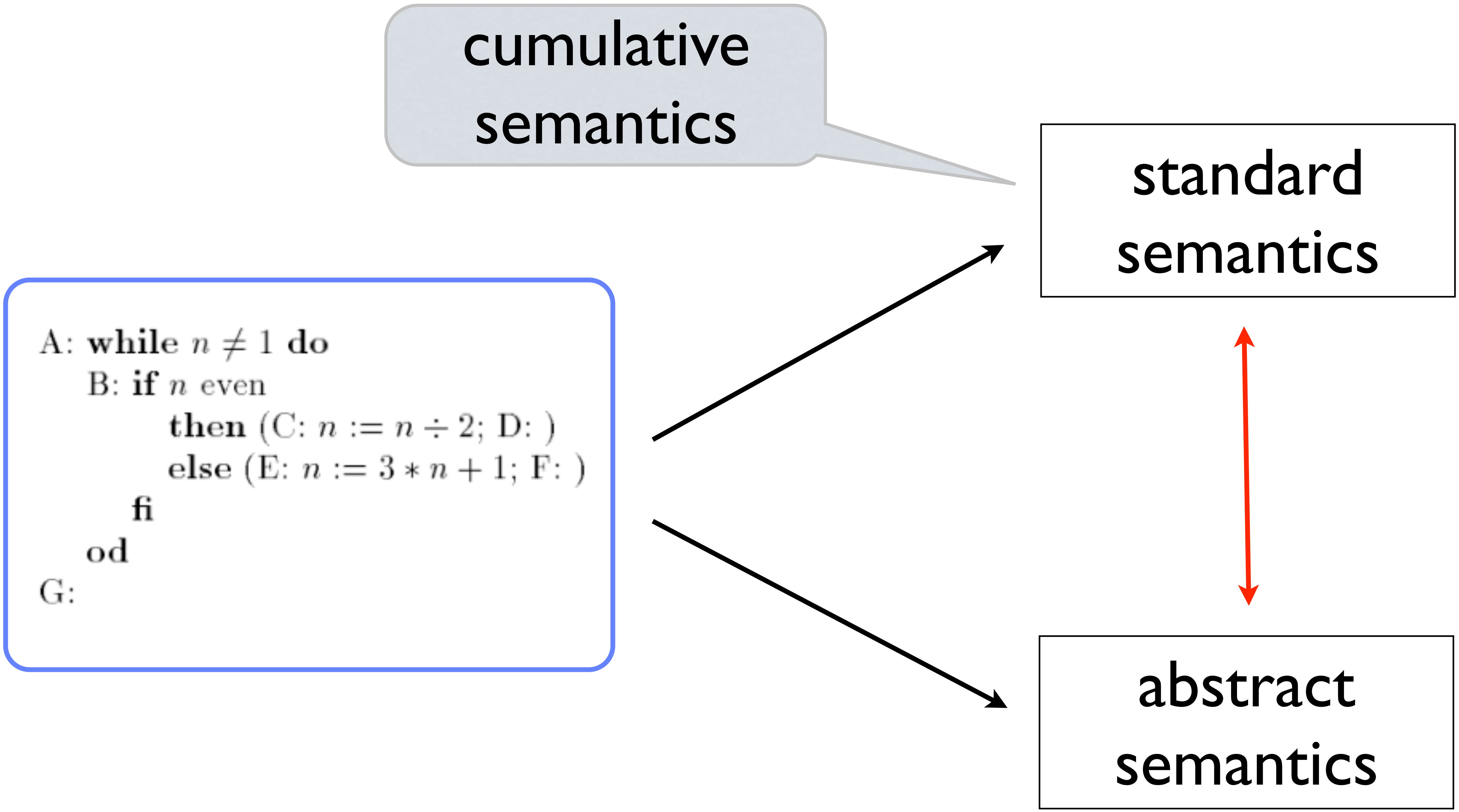


β is not always a homomorphism!

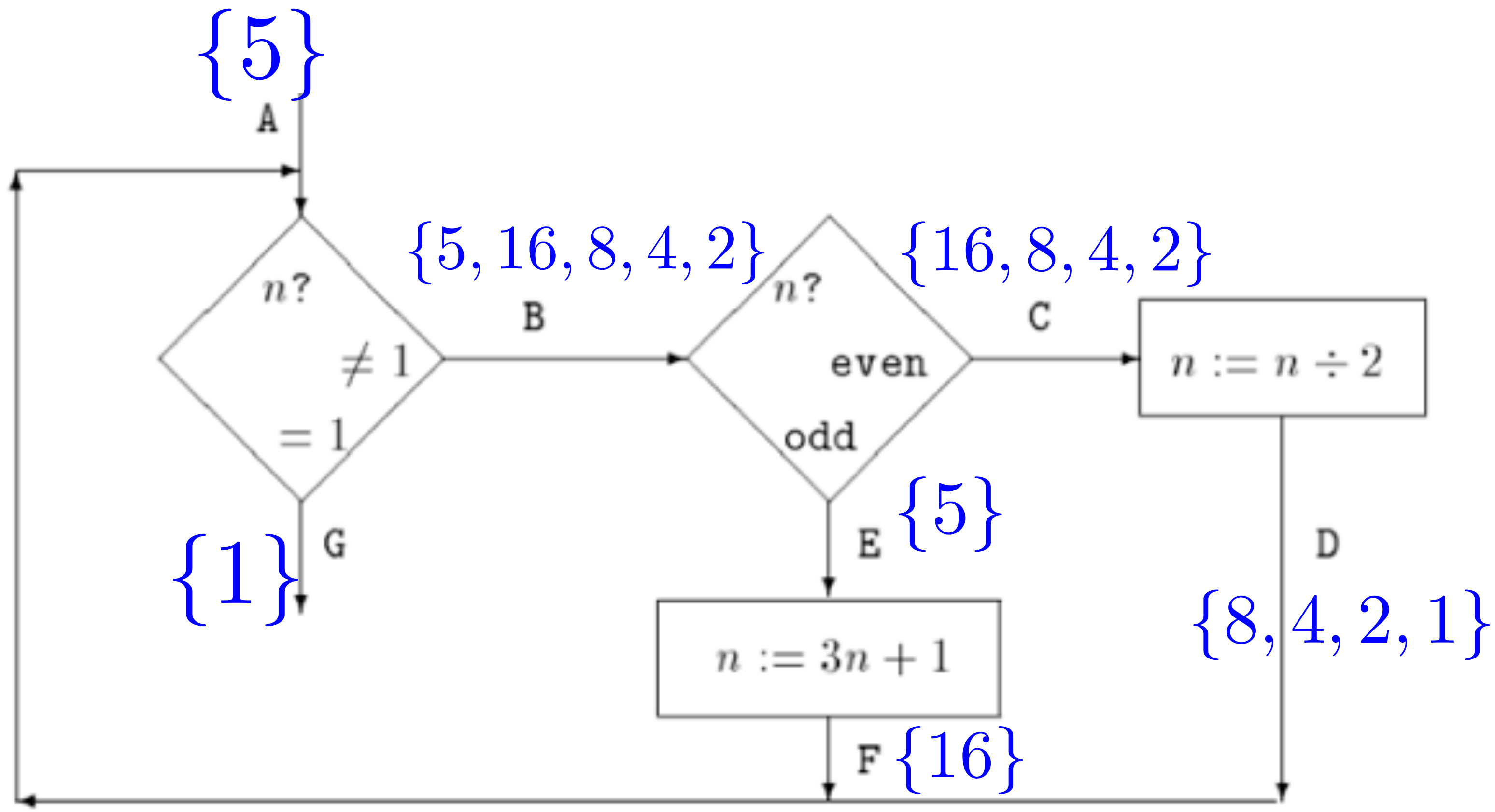
β is not always a homomorphism!




A: while $n \neq 1$ do
B: if n even
then (C: $n := n \div 2$; D:)
else (E: $n := 3 * n + 1$; F:)
fi
od
G:
standard semantics
abstract semantics


cumulative semantics
standard semantics
abstract semantics
A: while $n \neq 1$ do
B: if n even
then (C: $n := n \div 2$; D:)
else (E: $n := 3 * n + 1$; F:)
fi
od
G:

A: **while** $n \neq 1$ **do**
  B: **if** n even
    then (C: $n := n \div 2$; D:)
    else (E: $n := 3 * n + 1$; F:)
  fi
od
G:



Abstraction function

concrete semantics

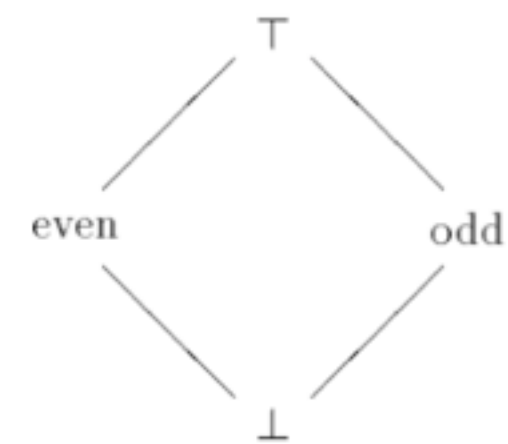
$$\mathcal{P}(V)$$

abstraction

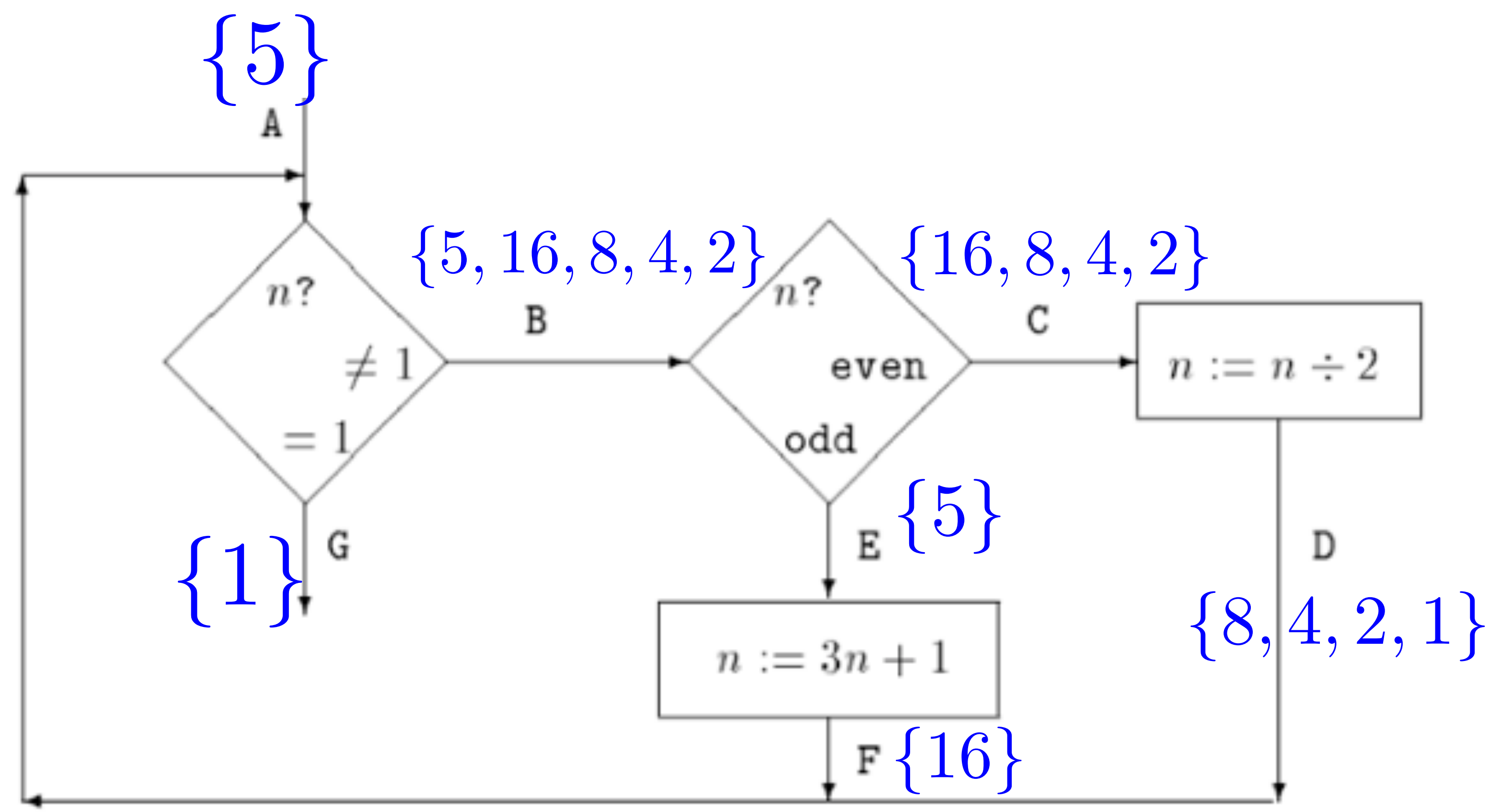
$$\alpha : \mathcal{P}(V) \to L \qquad \alpha(X) = \sqcup\{\beta(v) \mid v \in X\}$$

abstract semantics

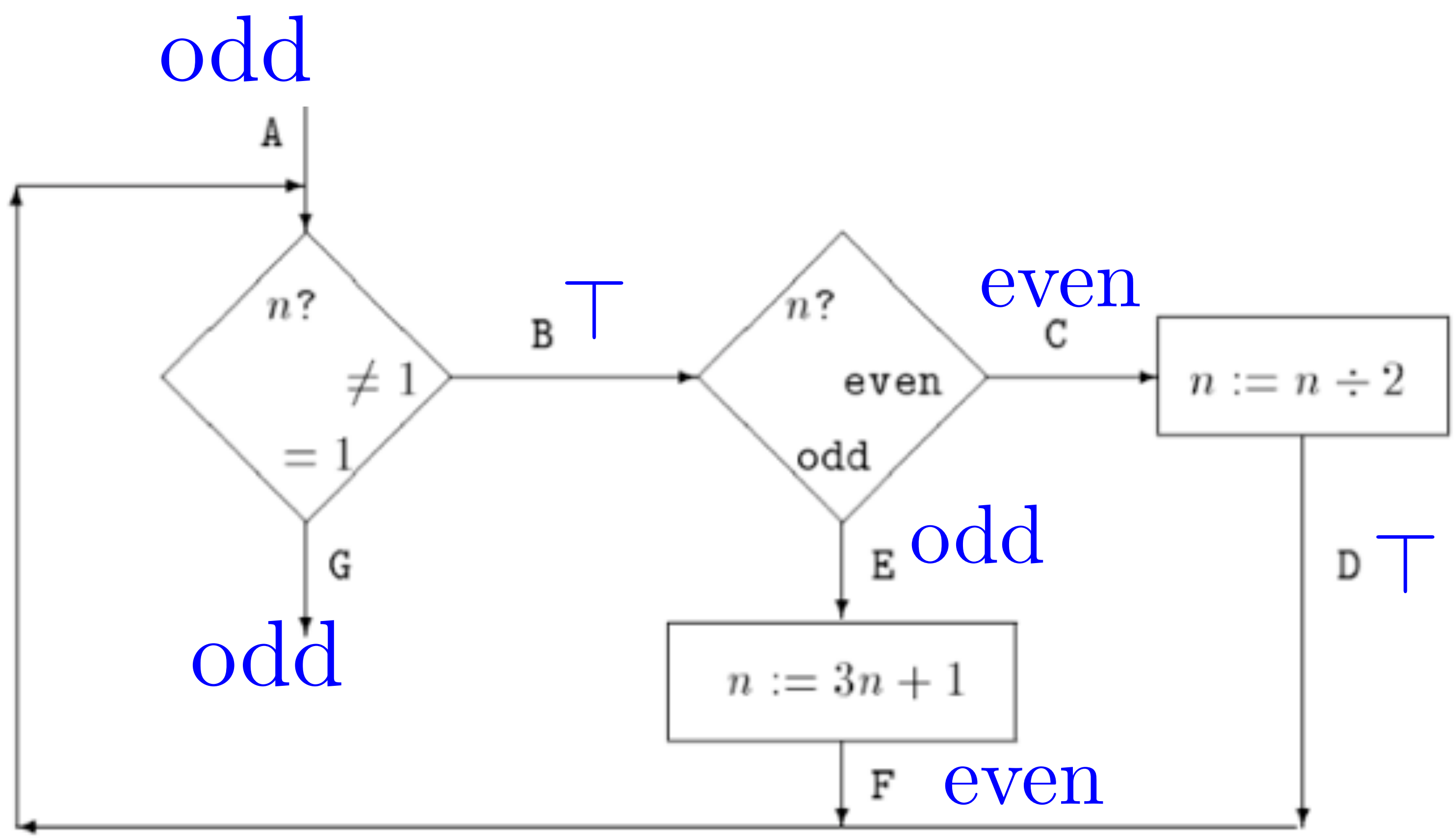
$$L \;=\; \mathrm{Var} \to \{\bot, \mathrm{even}, \mathrm{odd}, \top\}$$

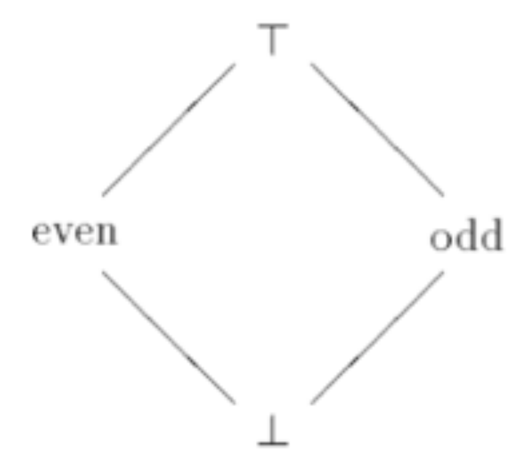


A: **while** $n \neq 1$ **do**
  B: **if** n even
    then (C: $n := n \div 2$; D:)
    else (E: $n := 3 * n + 1$; F:)
  fi
od
G:



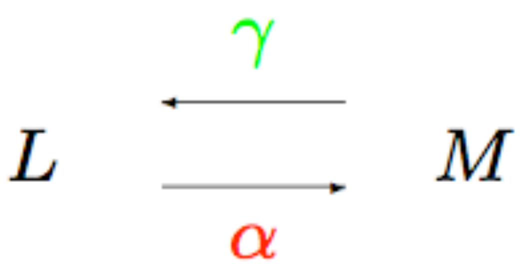
A: **while** $n \neq 1$ **do**
B: **if** n even
then (C: $n := n \div 2$; D:)
else (E: $n := 3 * n + 1$; F:)
fi
od
G:





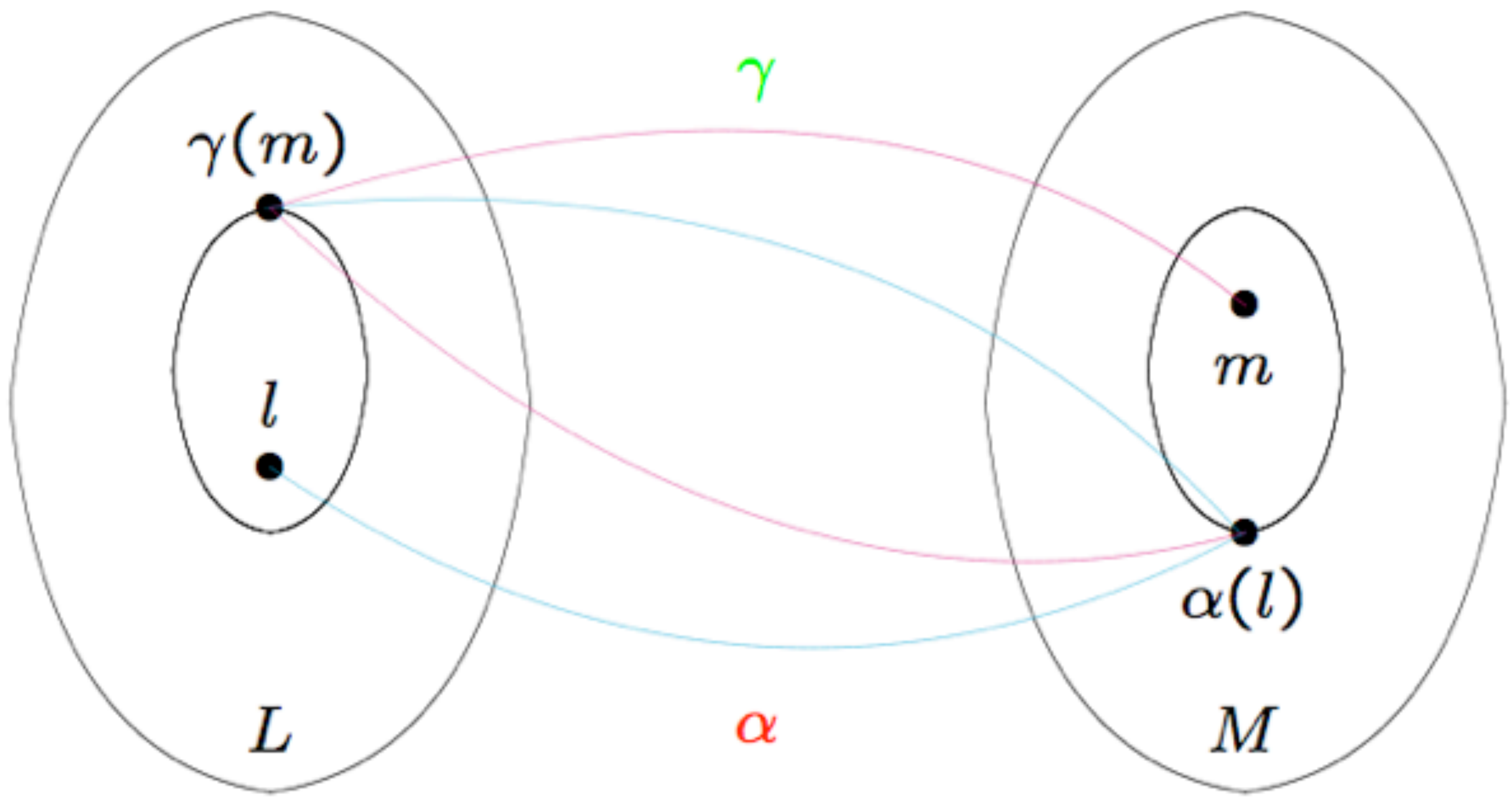
Galois connection

Galois connection



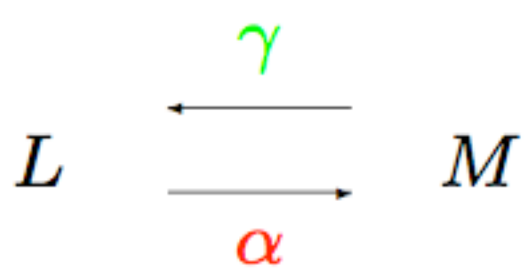
α - abstraction function

γ - concretization function



$l \sqsubseteq \gamma(\alpha(l))$ $\alpha(\gamma(m)) \sqsubseteq m$

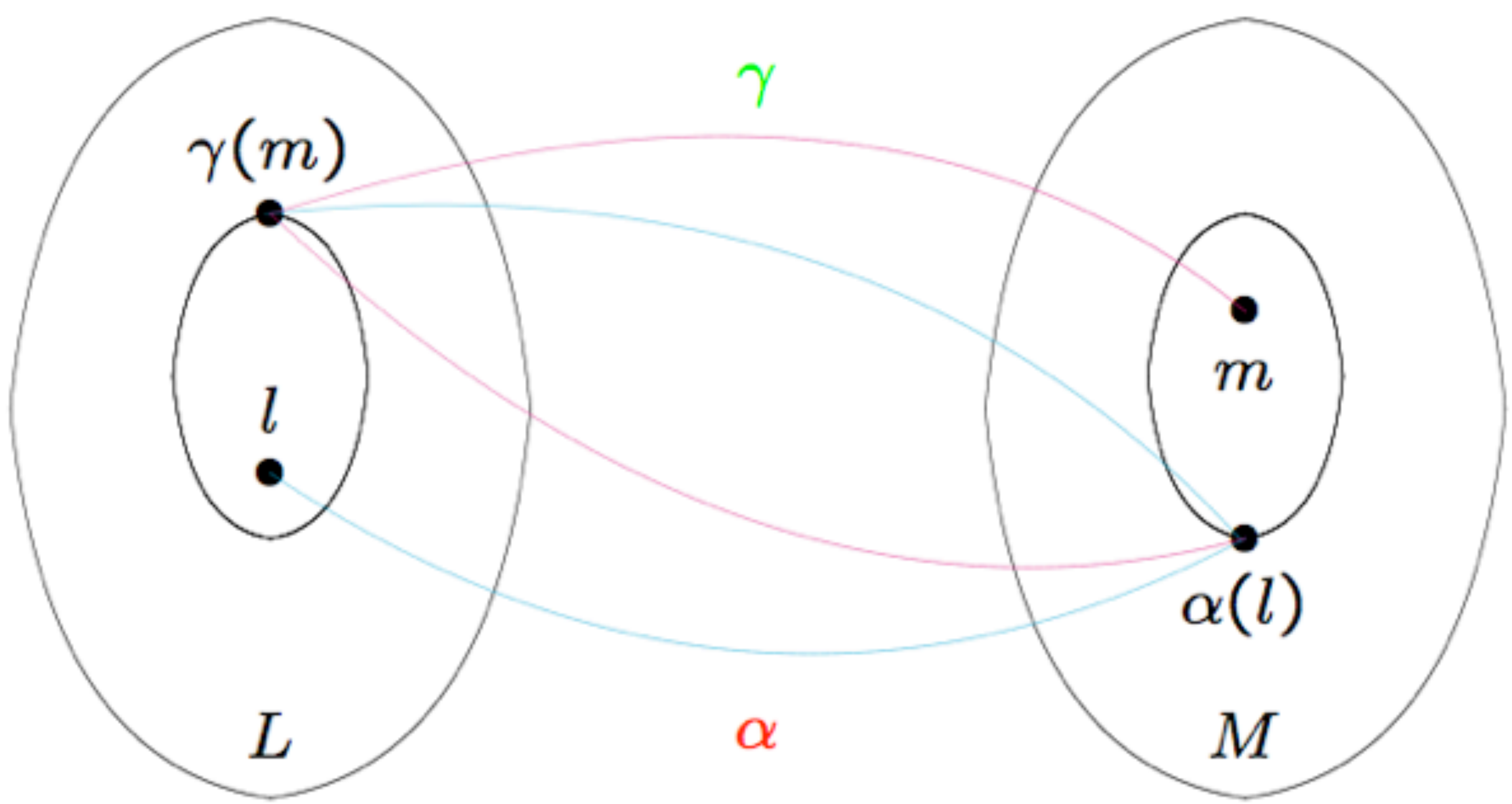
Galois connection



α - abstraction function

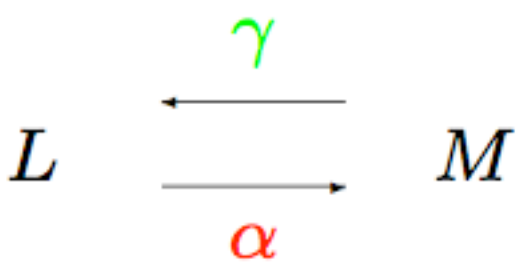
γ - concretization function

monotonic



$$l \sqsubseteq \gamma(\alpha(l)) \qquad \alpha(\gamma(m)) \sqsubseteq m$$

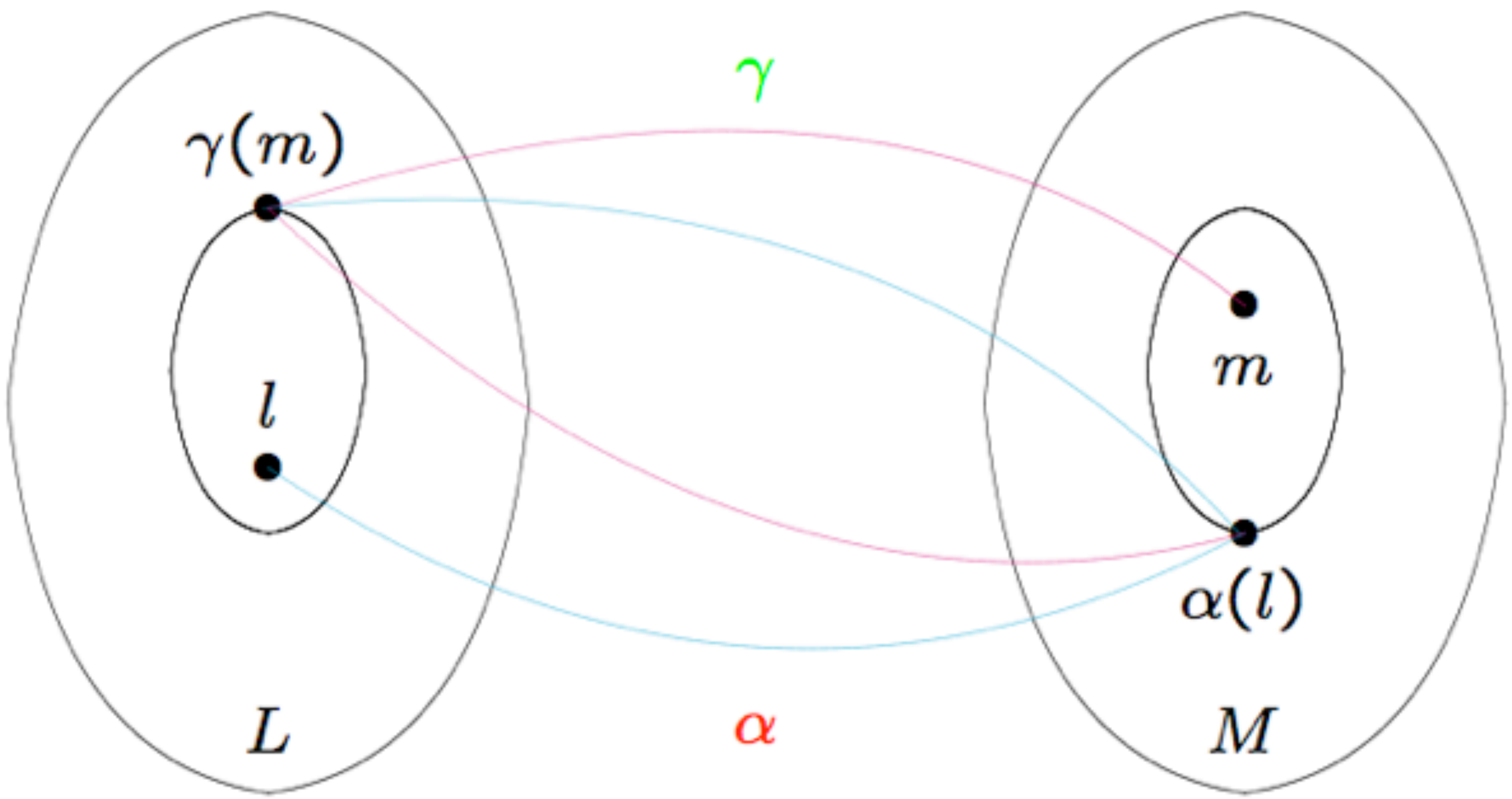
Galois connection



α - abstraction function

γ - concretization function

monotonic



$$l \sqsubseteq \gamma(m) \iff \alpha(l) \sqsubseteq m$$

Concrete and abstract domain

set of represented values

$$\mathcal{P}(V) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} L$$

the most exact abstraction

Example

$$\mathcal{P}(\mathbb{Z}) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \text{Intervals}$$

$\alpha(X) =$ the smallest interval containing *X*

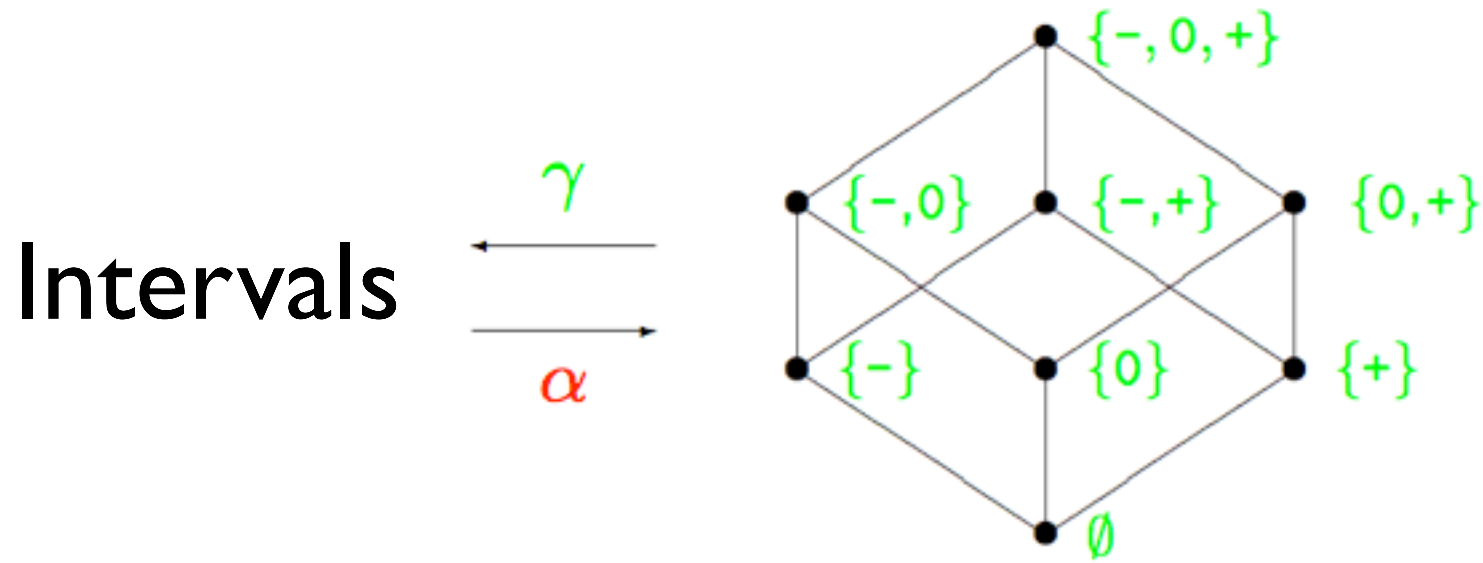
$$\gamma(I) = I$$

Two abstract domains

$$L \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} M$$

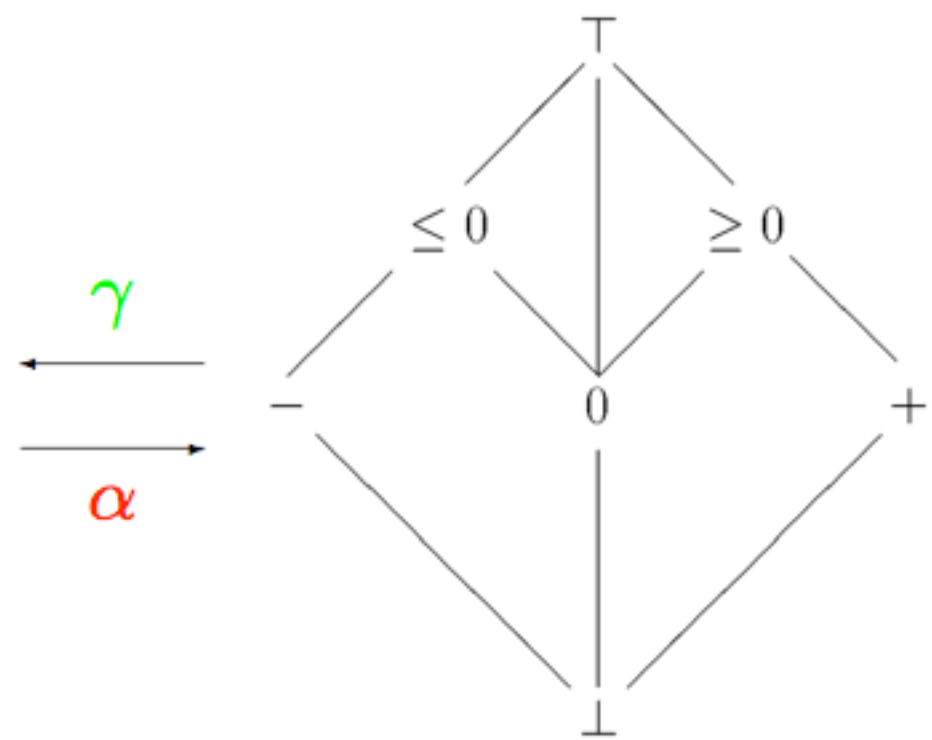
M is more abstract (less exact) than *L*

Example



Example

Intervals



no {-,+}

Representation function β induces a connection

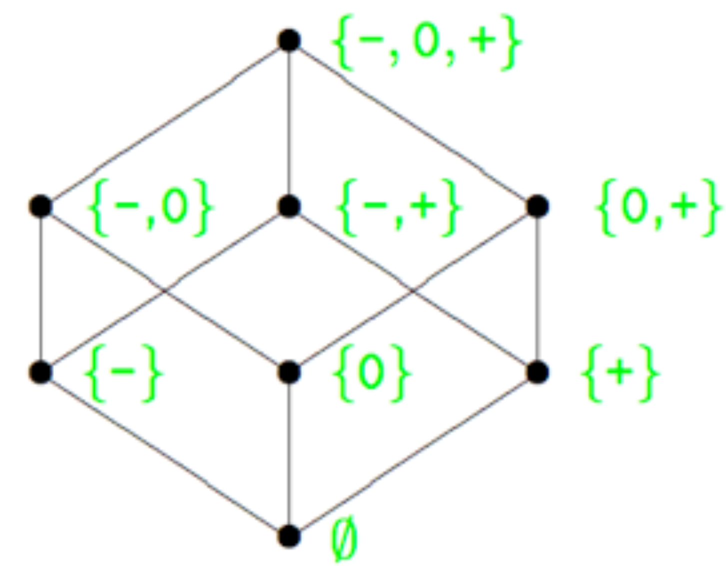
$$\mathcal{P}(V) \;\underset{\alpha}{\overset{\gamma}{\leftrightarrows}}\; L$$

$$\begin{aligned} \alpha(X) &= \sqcup\{\beta(v) \mid v \in X\} \\ \gamma(l) &= \{v \in V \mid \beta(v) \sqsubseteq l\} \end{aligned}$$

Example

$$\beta : \mathbb{Z} \to \{-, 0, +\}$$

$$\mathcal{P}(V) \underset{\alpha}{\overset{\gamma}{\rightleftarrows}} \mathcal{P}(\{-, 0, +\})$$



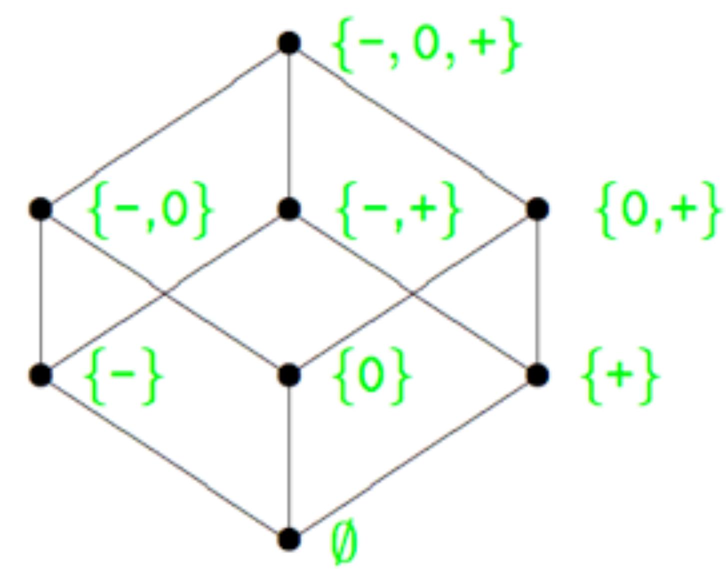
$$\alpha(X) = \{\beta(z) \mid z \in X\}$$

$$\gamma(S) = \{z \in \mathbb{Z} \mid \beta(z) \in S\}$$

Example

$$\beta : \mathbb{Z} \to \{-, 0, +\} \quad \subseteq \mathcal{P}(\{-, 0, +\})$$

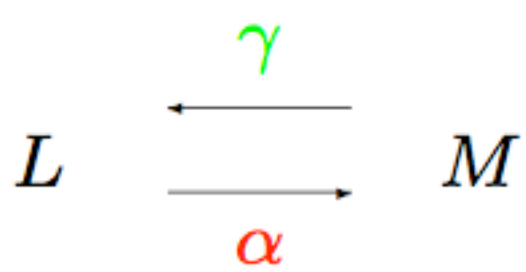
$$\mathcal{P}(V) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \mathcal{P}(\{-, 0, +\})$$



$$\alpha(X) = \{\beta(z) \mid z \in X\}$$

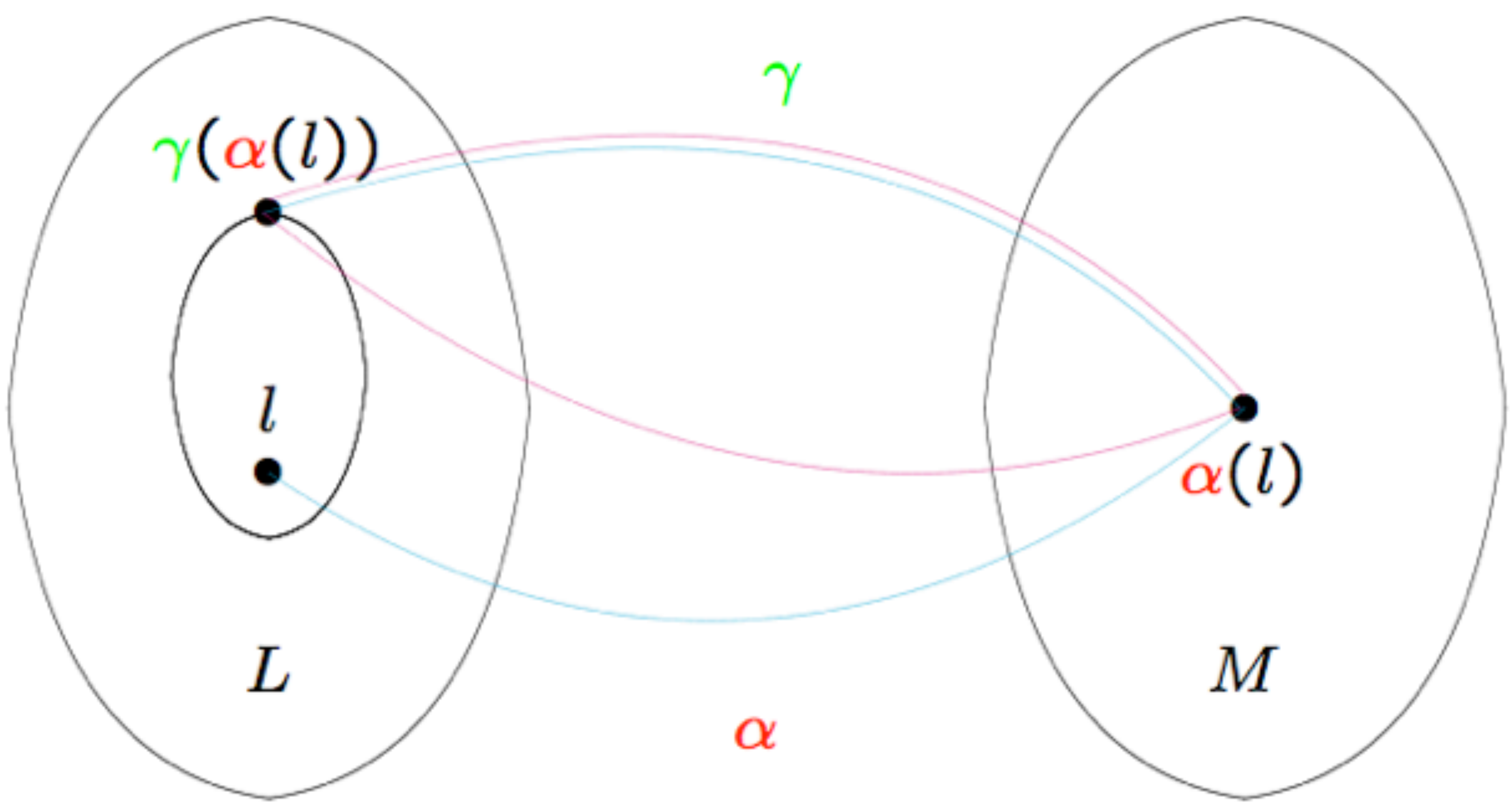
$$\gamma(S) = \{z \in \mathbb{Z} \mid \beta(z) \in S\}$$

Galois embedding



α - abstraction function

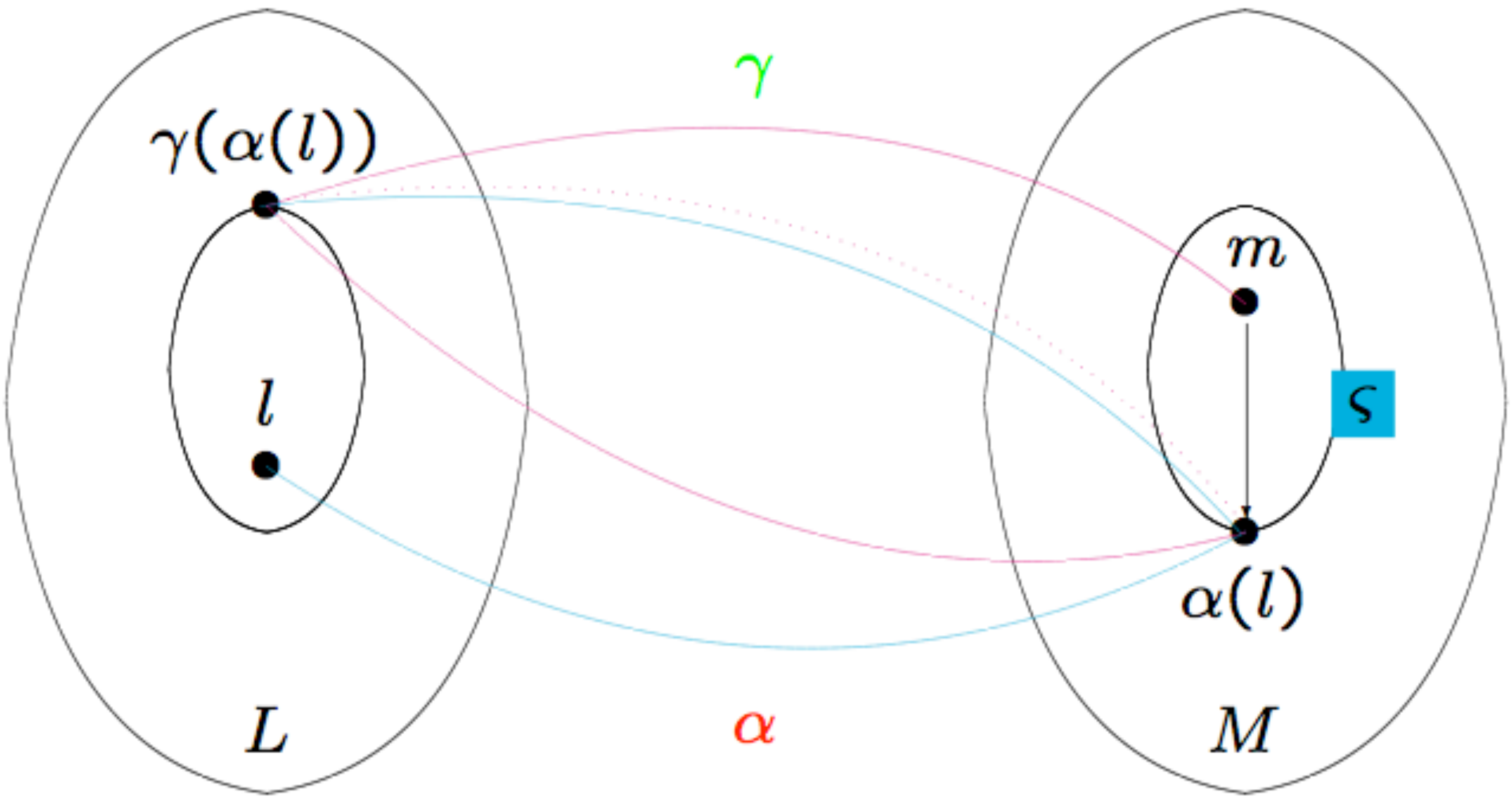
γ - concretization function



$l \sqsubseteq \gamma(\alpha(l))$ $\alpha(\gamma(m)) = m$

Reduction

elimination of unnecessary abstract values



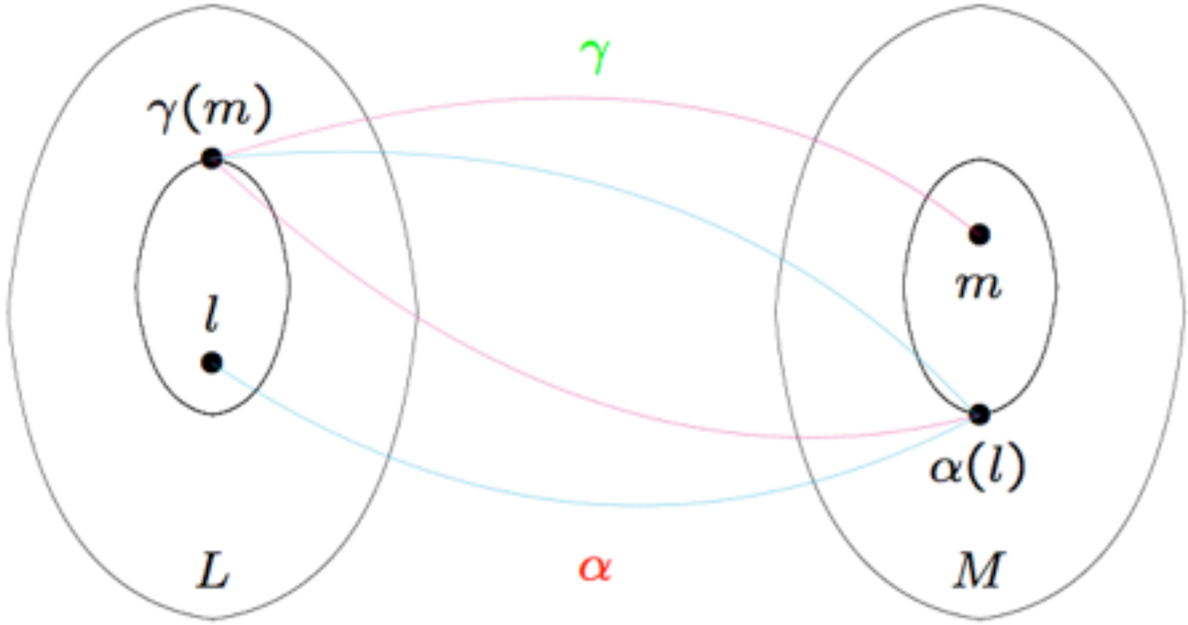
$$\zeta(m) = \sqcap\{m' \mid \gamma(m') = \gamma(m)\}$$

Right notion

- good mathematical properties
  - left adjoint preserves least upper bounds
  - right adjoint preserves greatest upper bounds
  - uniqueness
  - a monotonic function that preserves upper/ lower bounds induces a connection

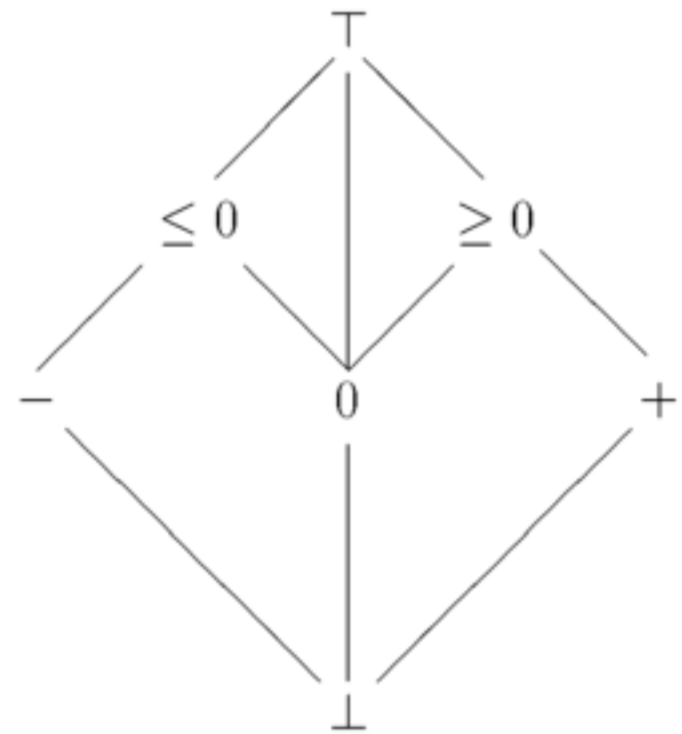
Right notion

- equivalent definitions:
  - closures
  - Moore families



- connections compose
  - which open a way to build more expressible analyses from simpler ones
- connection induces the most exact abstract semantics

$$\mathcal{P}(\mathbb{Z}) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \text{Intervals} \underset{\alpha}{\overset{\gamma}{\leftrightarrows}}$$



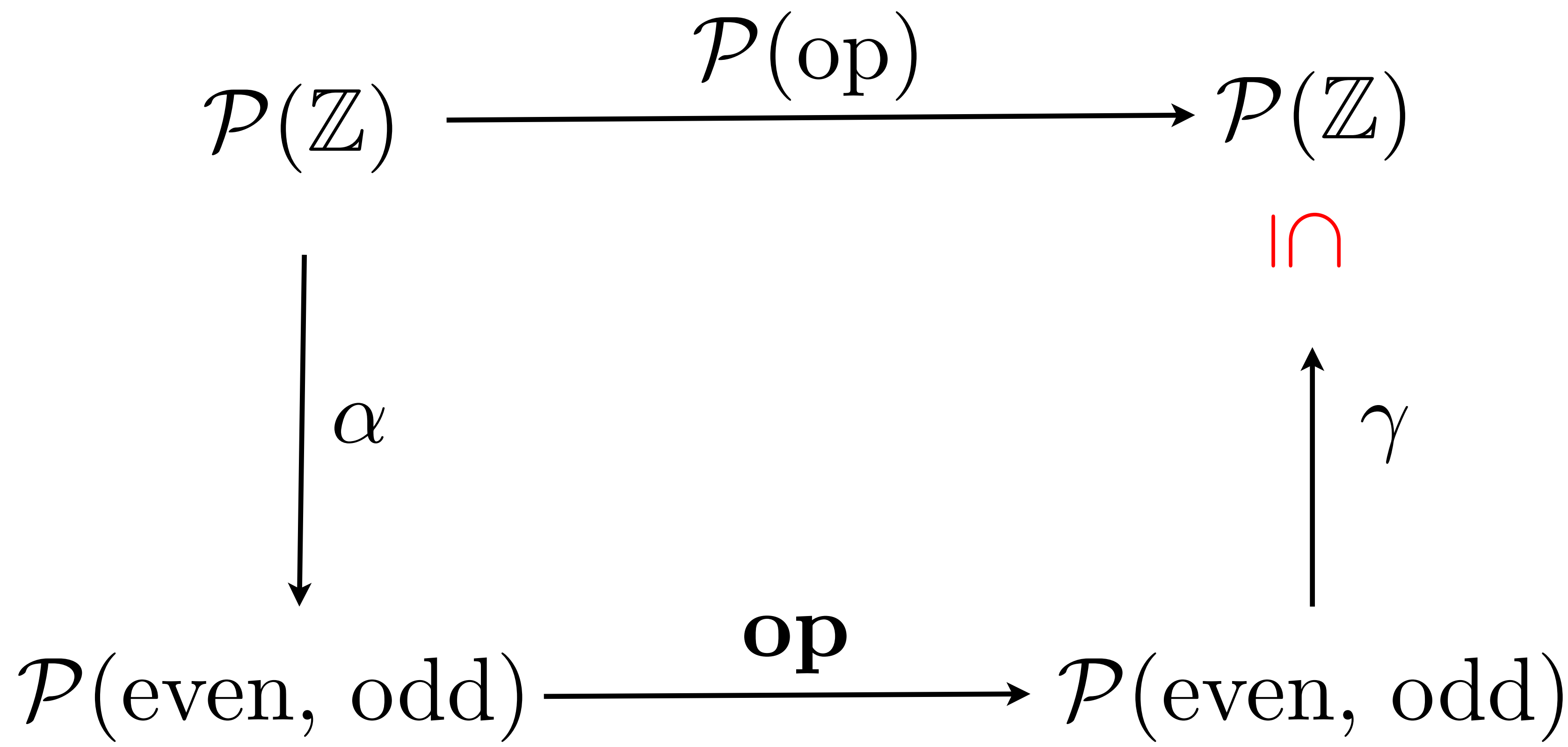
$$\alpha(S) = \begin{cases} \bot & \text{if } S = \{\} \text{ else} \\ + & \text{if } S \subseteq \{1,\ 2,\ 3, ...\} \text{ else} \\ \geq 0 & \text{if } S \subseteq \{0,\ 1,\ 2,\ 3, \ldots\} \text{ else} \\ - & \text{if } S \subseteq \{\text{-}1, \text{-}\ 2,\ \text{-}3, \ldots\} \text{ else} \\ \leq 0 & \text{if } S \subseteq \{0,\ \text{-}1,\ \text{-}2,\ \text{-}3, \ldots\} \text{ else} \\ \top & \end{cases}$$

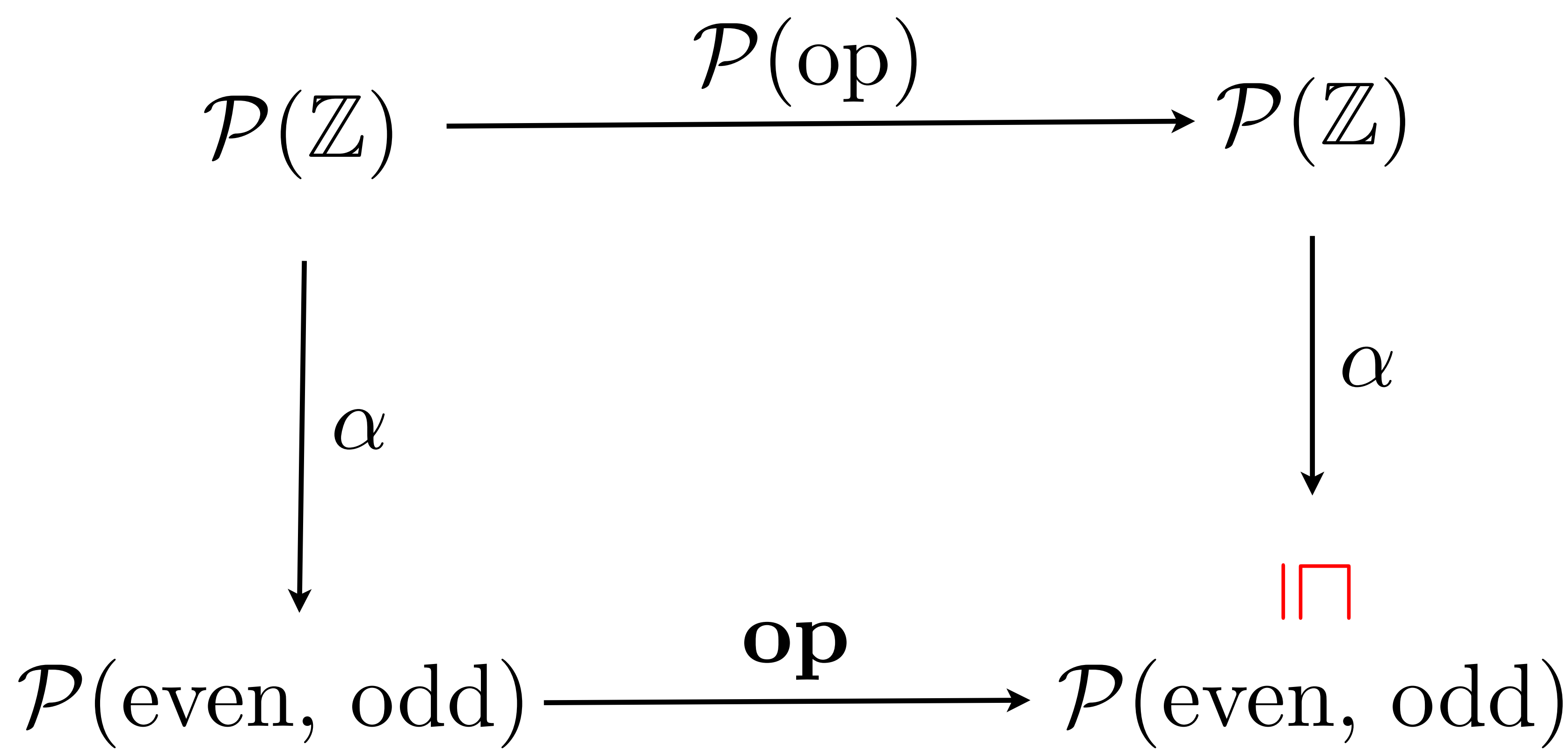
$$\begin{aligned} \gamma(0) &= \{0\} \\ \gamma(+) &= \{1, 2, 3, ...\} \\ \gamma(-) &= \{-1, -2, -3, \ldots\} \\ \gamma(\bot) &= \{\} \\ \gamma(\geq 0) &= \{0, 1, 2, 3, ...\} \\ \gamma(\leq 0) &= \{0, -1, -2, -3, ...\} \\ \gamma(\top) &= \{\ldots, -2, -1, 0, 1, 2, 3, \ldots\} \end{aligned}$$

Safe approximation

$$\mathrm{op} : \mathbb{Z} \to \mathbb{Z}$$

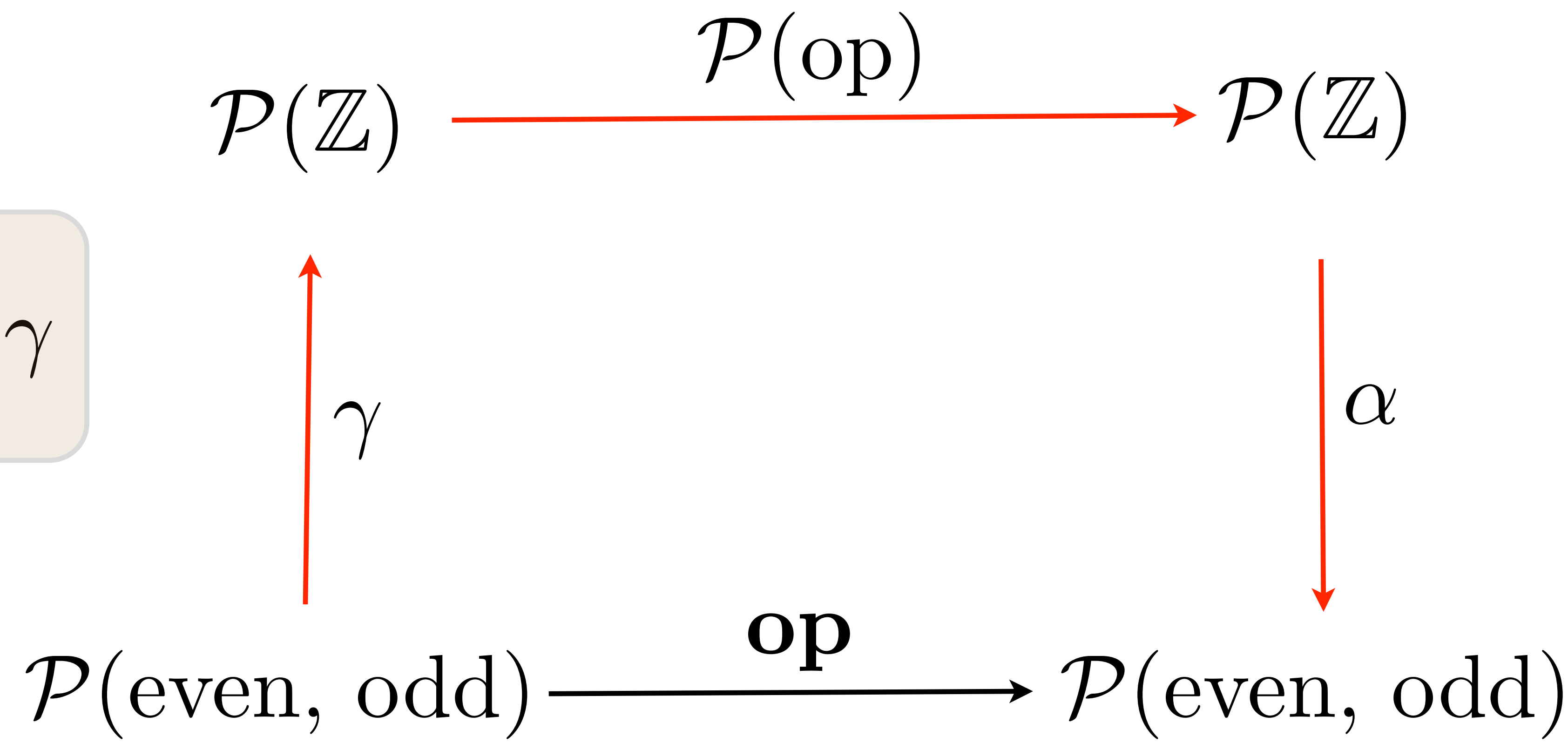
$$\mathcal{P}(\mathrm{op}) : \mathcal{P}(\mathbb{Z}) \to \mathcal{P}(\mathbb{Z})$$

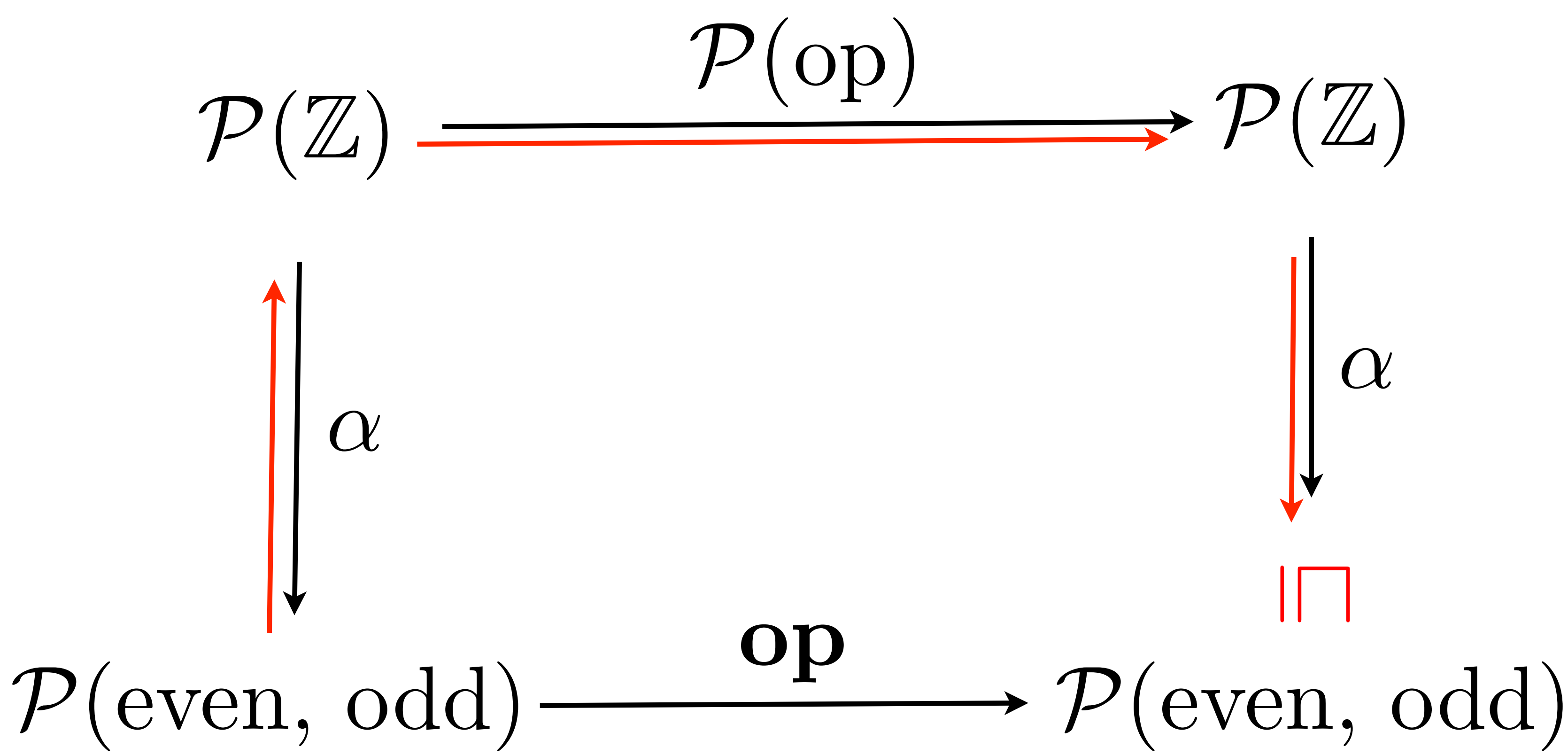




The most exact approximation

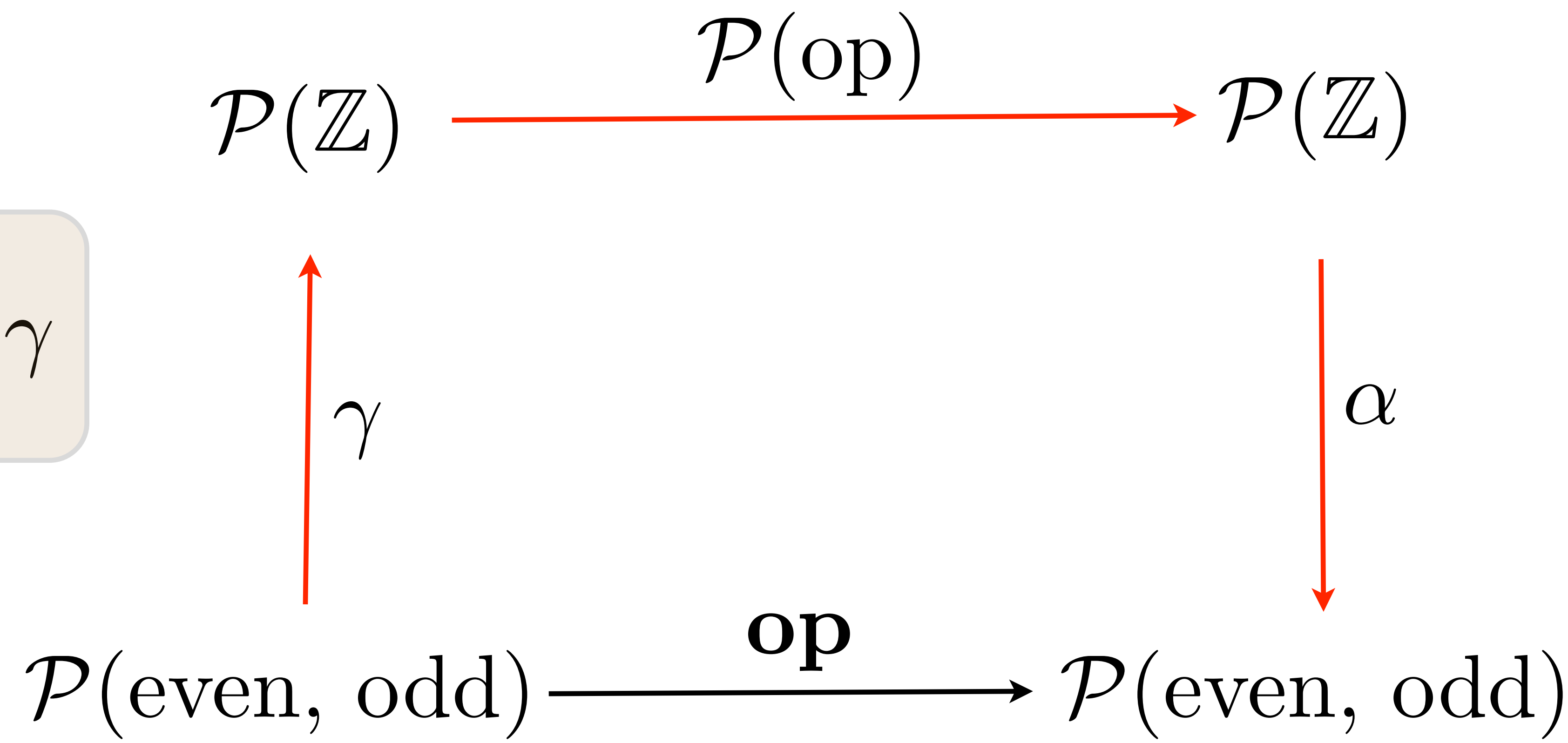
$$\mathbf{op} \;:=\; \alpha \circ \mathcal{P}(\mathrm{op}) \circ \gamma$$



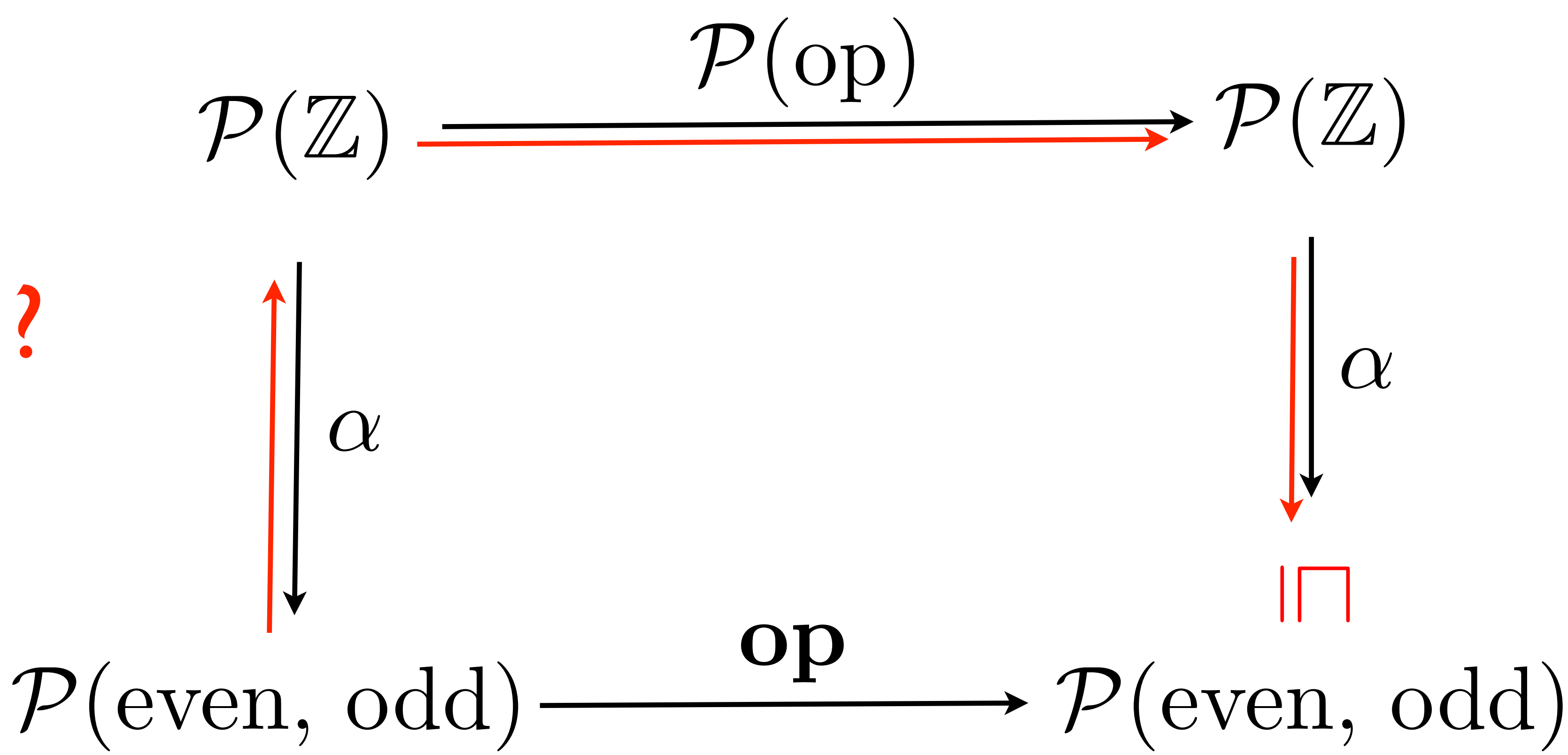


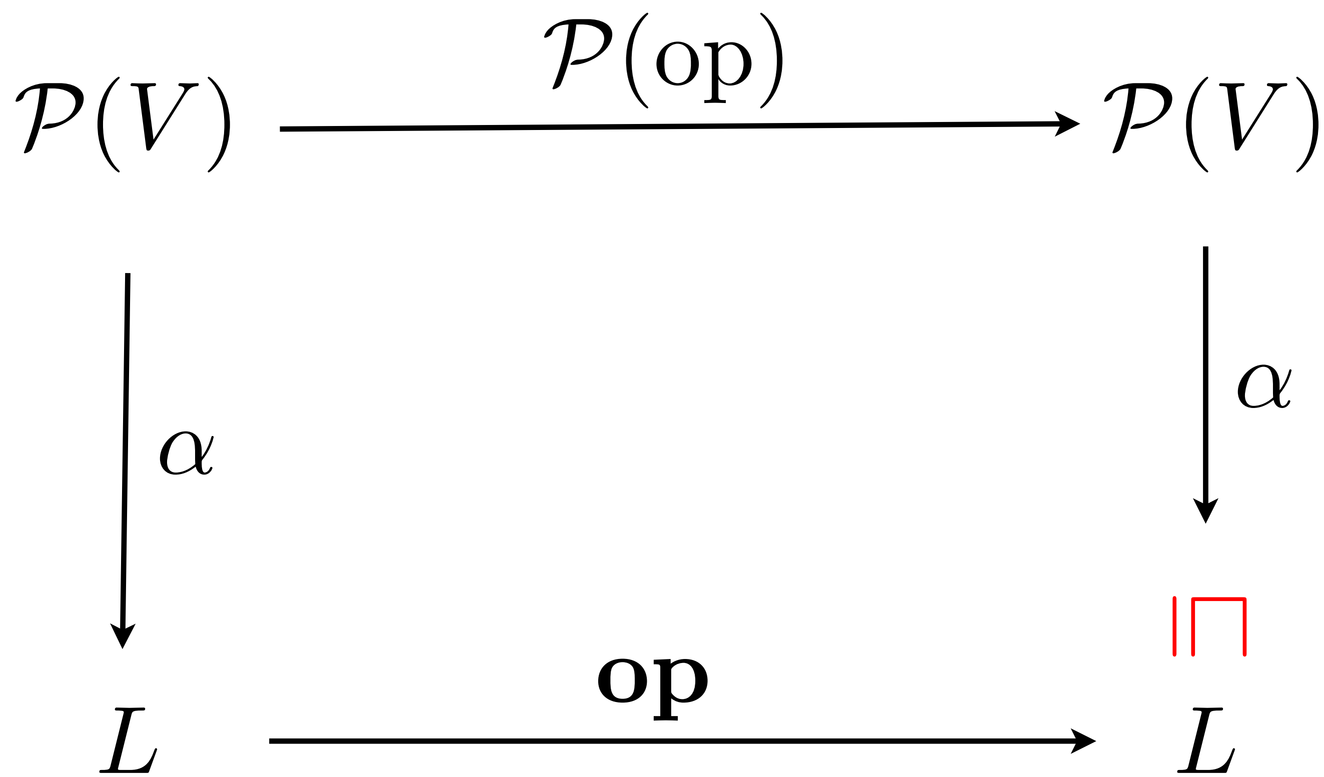
The most exact approximation

$$\mathbf{op} \;:=\; \alpha \circ \mathcal{P}(\mathrm{op}) \circ \gamma$$

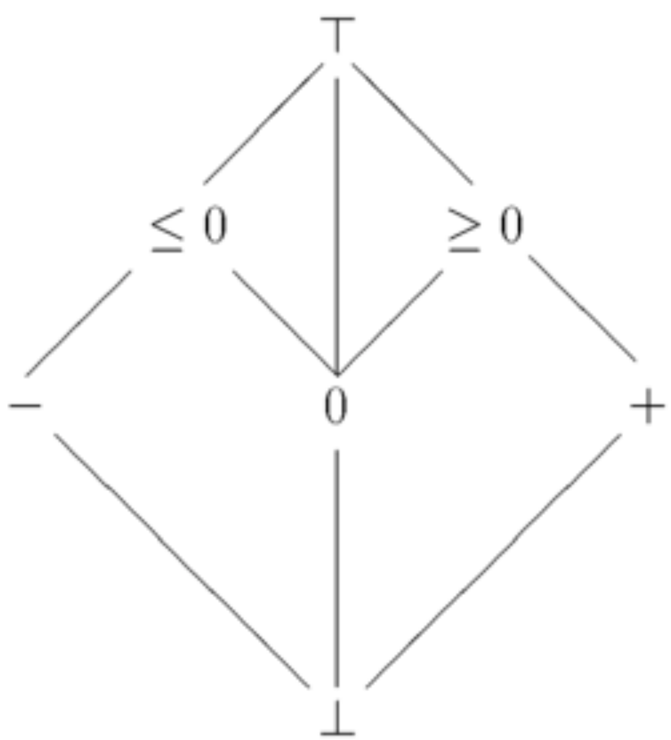


how to compute **op** ?




$\mathcal{P}(V)$
$\mathcal{P}(\mathrm{op})$
$\mathcal{P}(V)$
α
α
$\sqsubseteq$
L
$\mathbf{op}$
L

Example

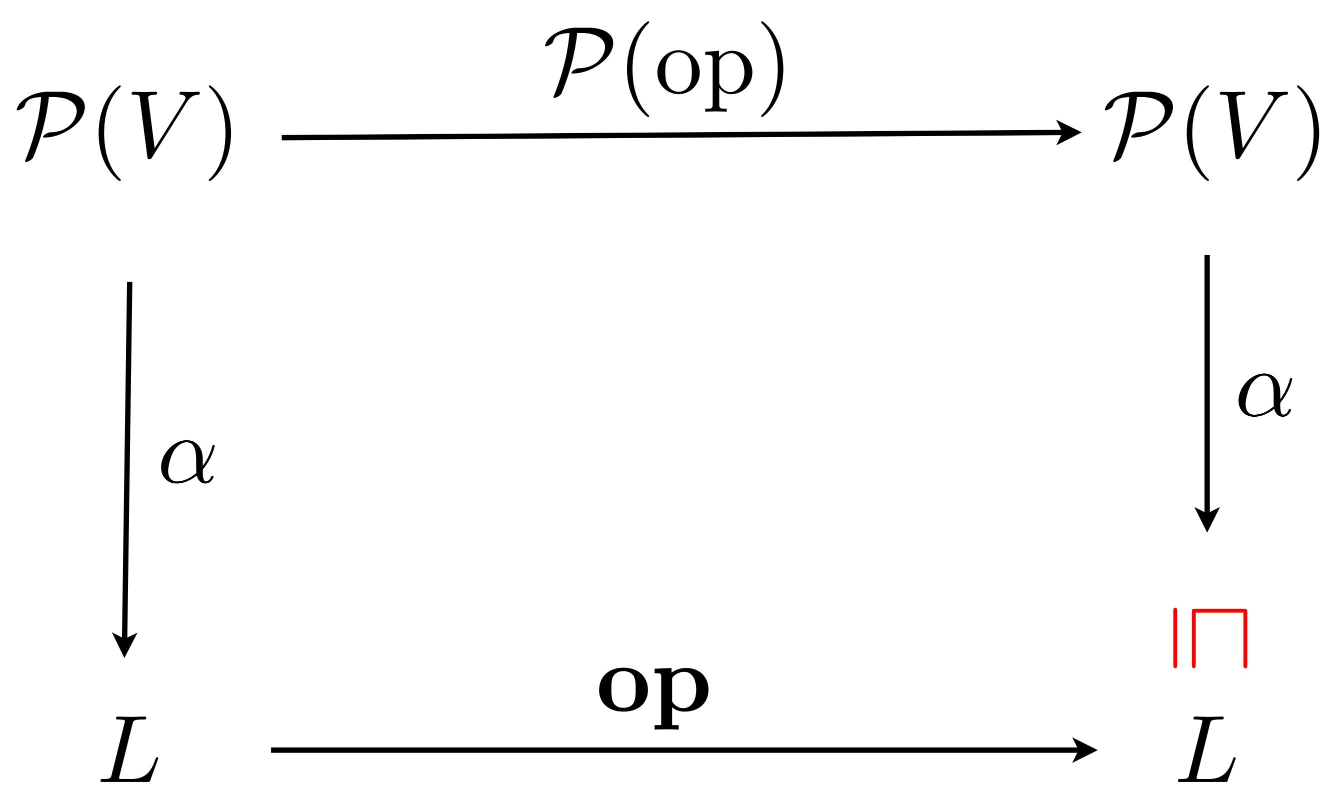


$$
\begin{aligned}
\gamma(0) &= \{0\}\\
\gamma(+) &= \{1, 2, 3, ...\}\\
\gamma(-) &= \{-1, -2, -3, \ldots\}\\
\gamma(\bot) &= \{\}\\
\gamma(\geq 0) &= \{0, 1, 2, 3, ...\}\\
\gamma(\leq 0) &= \{0, -1, -2, -3, ...\}\\
\gamma(\top) &= \{\ldots, -2, -1, 0, 1, 2, 3, \ldots\}
\end{aligned}
$$

$$
\alpha(S) = \begin{cases}
\bot & \text{if } S = \{\} \text{ else}\\
+ & \text{if } S \subseteq \{1,\ 2,\ 3,...\} \text{ else}\\
\geq 0 & \text{if } S \subseteq \{0,\ 1,\ 2,\ 3,\ldots\} \text{ else}\\
- & \text{if } S \subseteq \{\text{-}1, \text{-}\ 2,\ \text{-}3,\ldots\} \text{ else}\\
\leq 0 & \text{if } S \subseteq \{0,\ \text{-}1,\ \text{-}2,\ \text{-}3,\ldots\} \text{ else}\\
\top &
\end{cases}
$$

$$+\ :\ \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$$

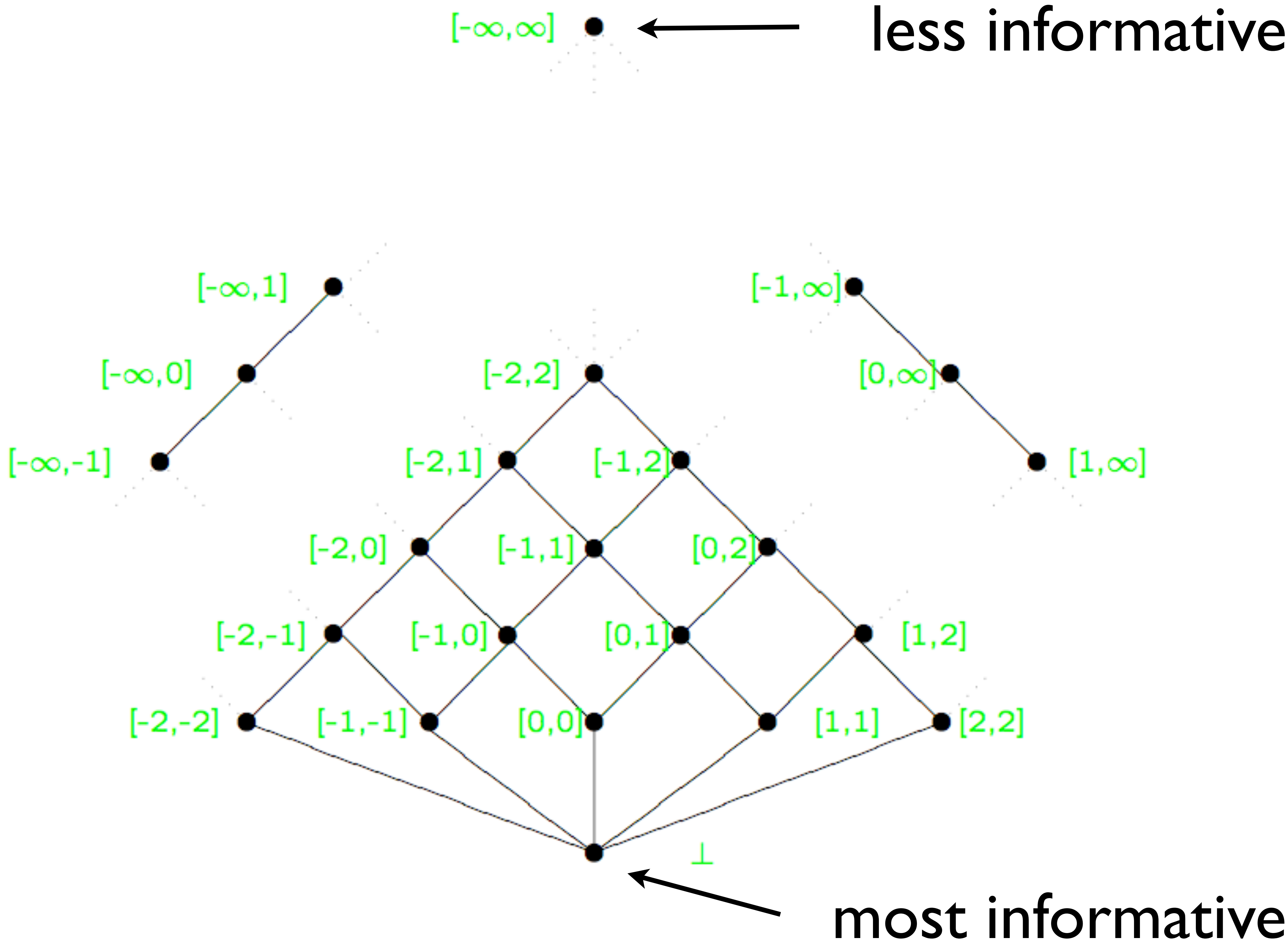
$+'$	$\bot$	$-$	0	$+$	≥ 0	≤ 0	$\top$
$\bot$	$\bot$	$\bot$	$\bot$	$\bot$	$\bot$	$\bot$	$\bot$
$-$	$\bot$	$-$	$-$	$\top$	$\top$	$-$	$\top$
0	$\bot$	$-$	0	$+$	≥ 0	≤ 0	$\top$
$+$	$\bot$	$\top$	$+$	$+$	$+$	$\top$	$\top$
≥ 0	$\bot$	$\top$	≥ 0	$+$	≥ 0	$\top$	$\top$
≤ 0	$\bot$	$-$	≤ 0	$\top$	$\top$	≤ 0	$\top$
$\bot$	$\bot$	$\top$	$\top$	$\top$	$\top$	$\top$	$\top$



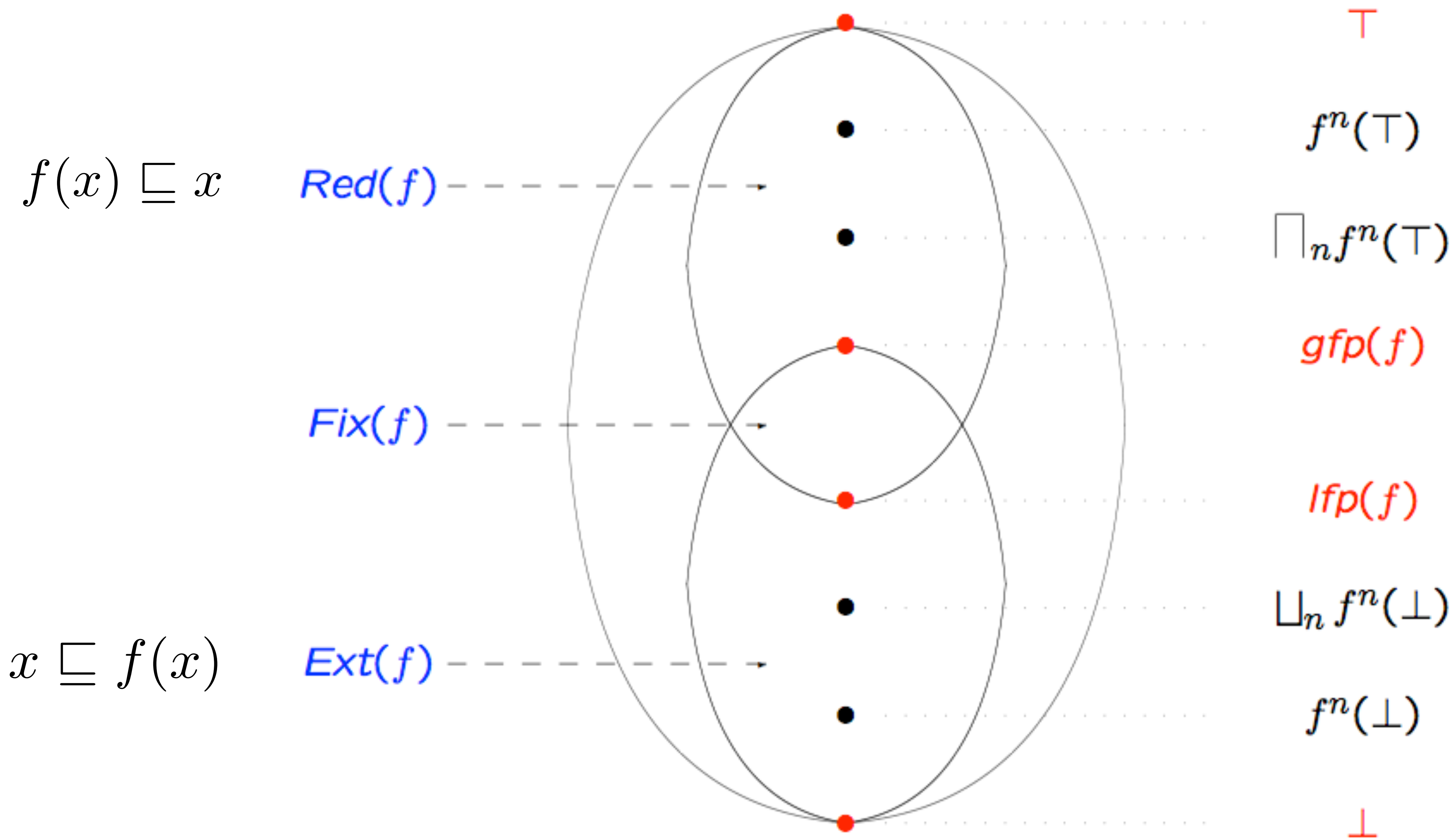
Widening/narrowing

Complete lattice L^S

$$f : L^S \to L^S$$

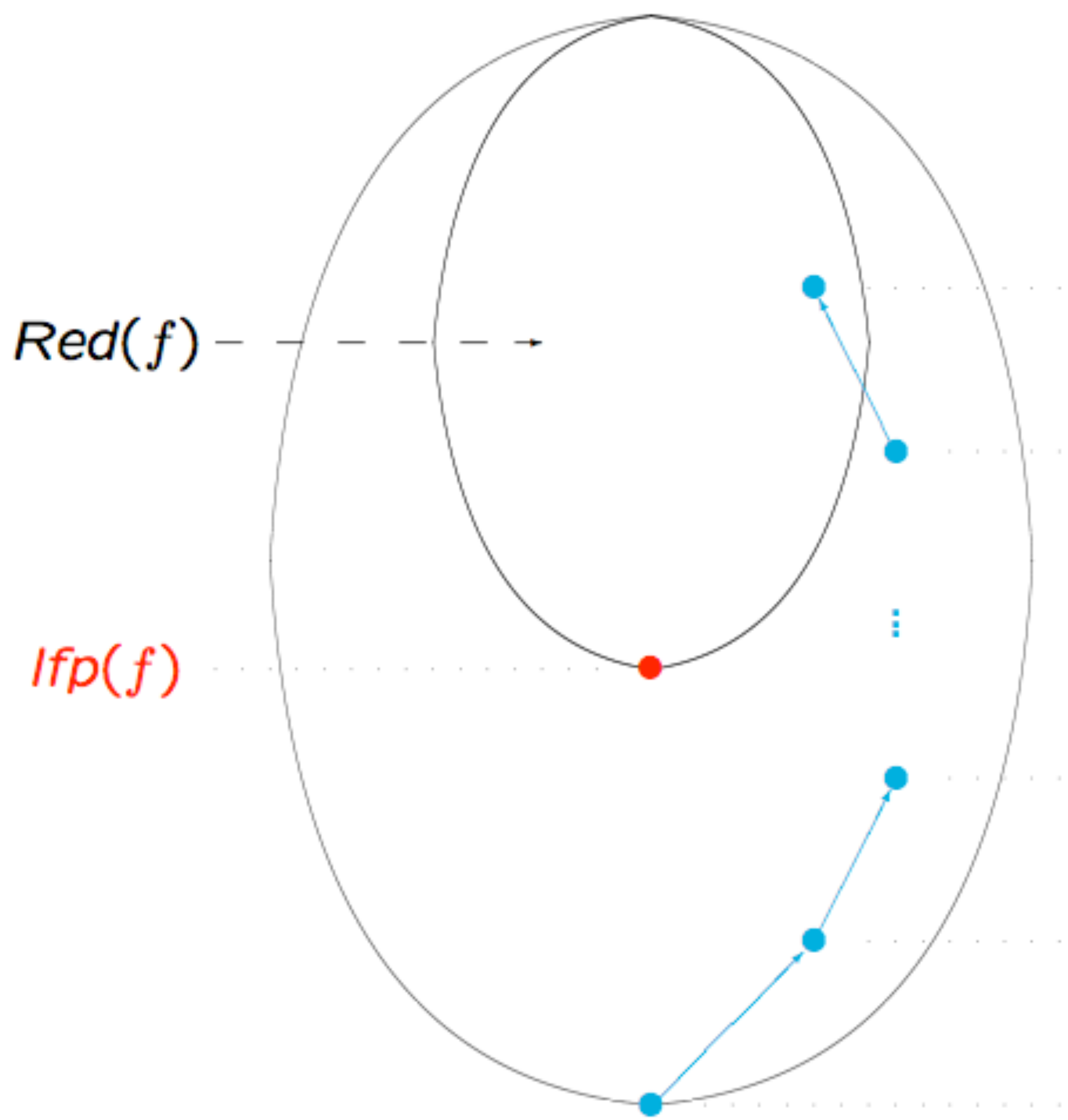

[-∞,∞]
less informative
[-∞,1]
[-1,∞]
[-∞,0]
[-2,2]
[0,∞]
[-∞,-1]
[-2,1]
[-1,2]
[1,∞]
[-2,0]
[-1,1]
[0,2]
[-2,-1]
[-1,0]
[0,1]
[1,2]
[-2,-2]
[-1,-1]
[0,0]
[1,1]
[2,2]
⊥
most informative

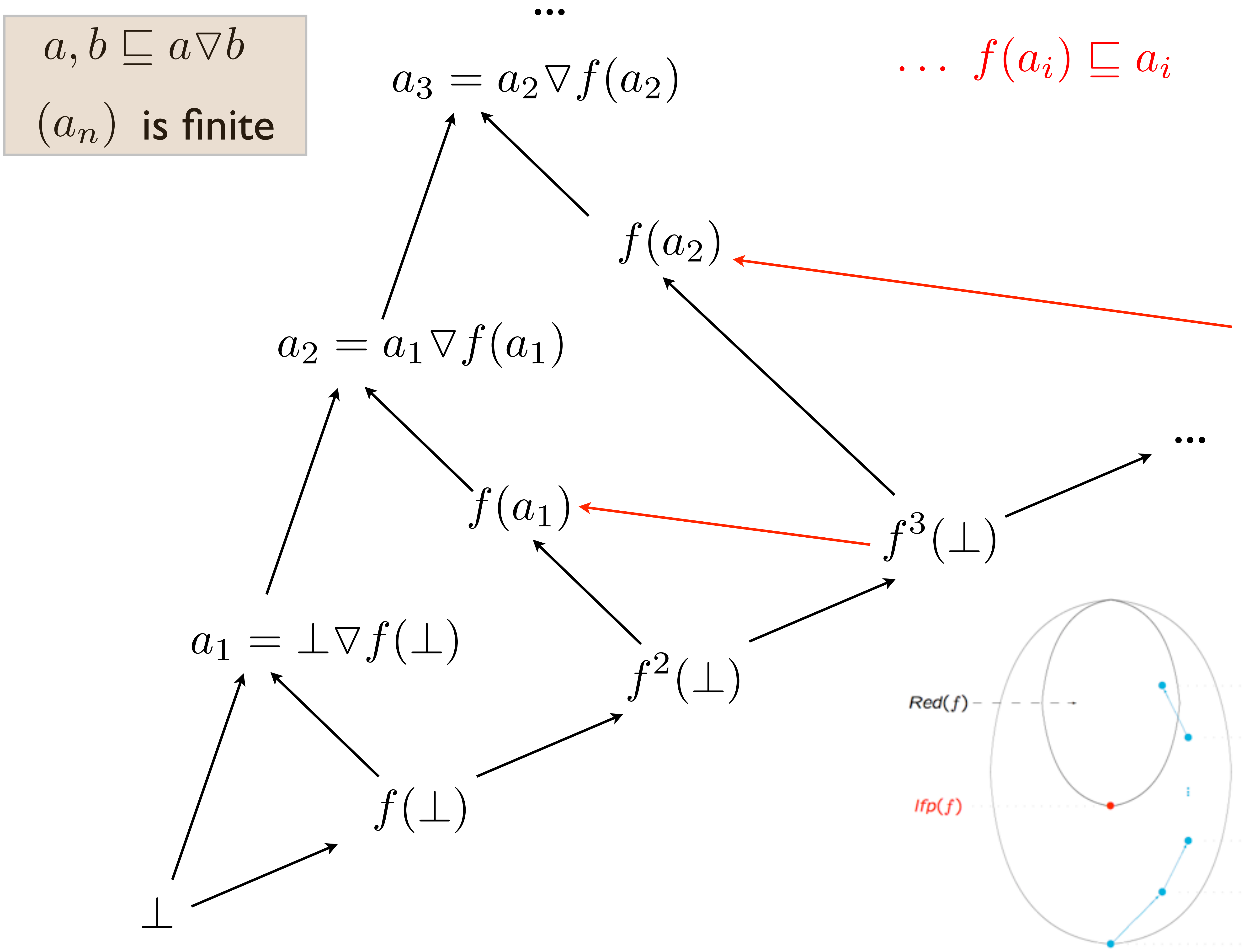
Fixed points



Widening

$f(x) \sqsubseteq x$




$a, b \sqsubseteq a \nabla b$
(a_n) is finite
$\ldots \; f(a_i) \sqsubseteq a_i$
...
$a_3 = a_2 \nabla f(a_2)$
$f(a_2)$
$a_2 = a_1 \nabla f(a_1)$
$f(a_1)$
$f^3(\bot)$
...
$a_1 = \bot \nabla f(\bot)$
$f^2(\bot)$
$f(\bot)$
$\bot$
Red(f)
lfp(f)

Example

K - numerical constants that appear in the source code

$$[z_1, z_2] \nabla [z_3, z_4] \quad = \quad [\ \mathrm{LB}(z_1, z_3)\ ,\ \mathrm{UB}(z_2, z_4)\ ]$$

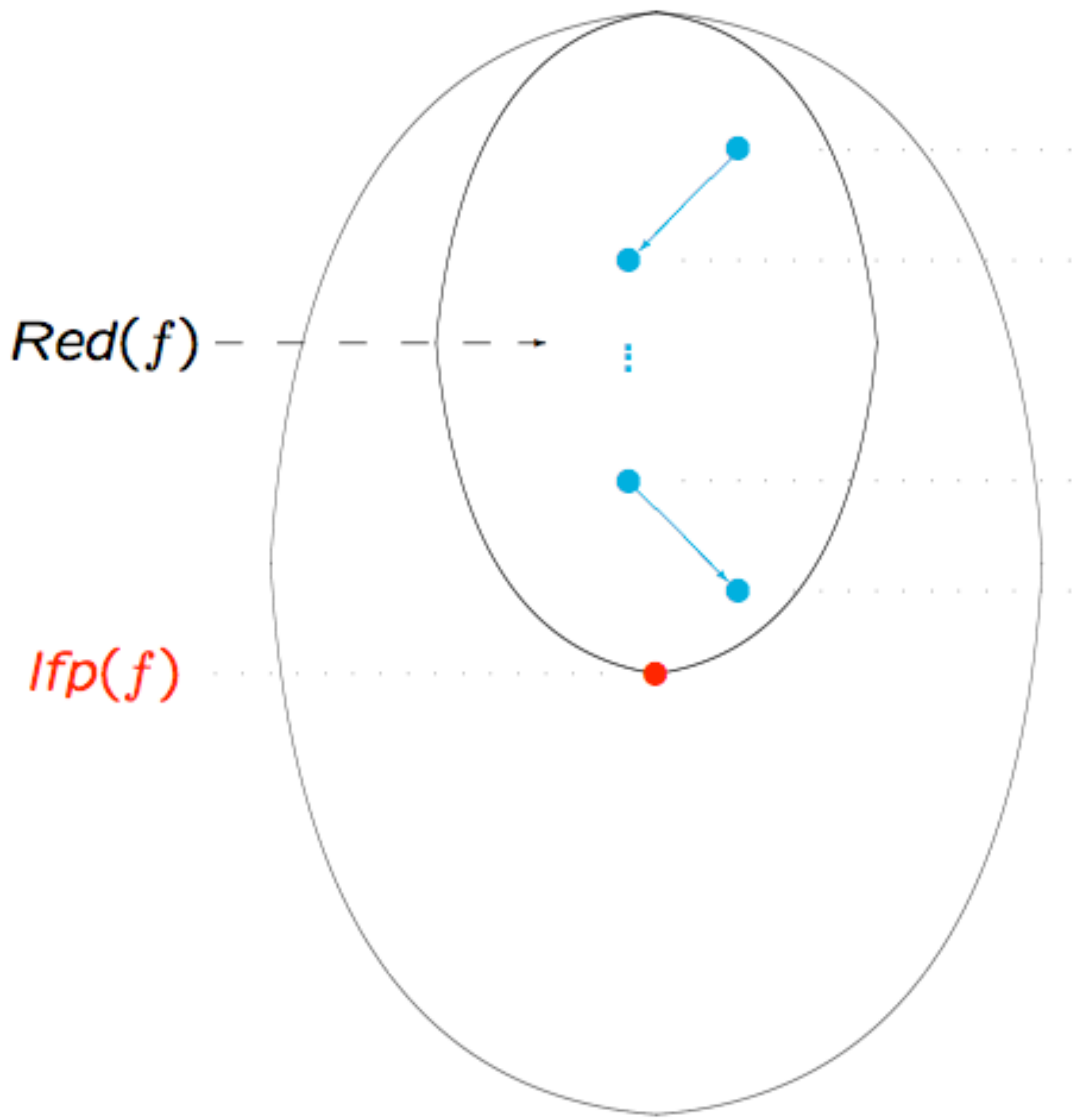
$$\mathrm{LB}(z_1, z_3) \in \{z_1\} \cup K \cup \{-\infty\}$$

$$\mathrm{UB}(z_2, z_4) \in \{z_2\} \cup K \cup \{\infty\}$$

$$\mathrm{LB}_K(z_1, z_3) = \begin{cases} z_1 & \text{if } z_1 \leq z_3 \\ k & \text{if } z_3 < z_1 \ \wedge \ k = \max\{k \in K \mid k \leq z_3\} \\ -\infty & \text{if } z_3 < z_1 \ \wedge \ \forall k \in K : z_3 < k \end{cases}$$

$$\mathrm{UB}_K(z_2, z_4) = \begin{cases} z_2 & \text{if } z_4 \leq z_2 \\ k & \text{if } z_2 < z_4 \ \wedge \ k = \min\{k \in K \mid z_4 \leq k\} \\ \infty & \text{if } z_2 < z_4 \ \wedge \ \forall k \in K : k < z_4 \end{cases}$$

Narrowing



$[z_1, \infty], [z_3, \infty], [z_3, \infty], \ldots$ $\qquad$ $z_1 < z_2 < z_3 < \ldots$