Program Analysis & Verification

Sriram K. Rajamani
MSR India
sriram@microsoft.com

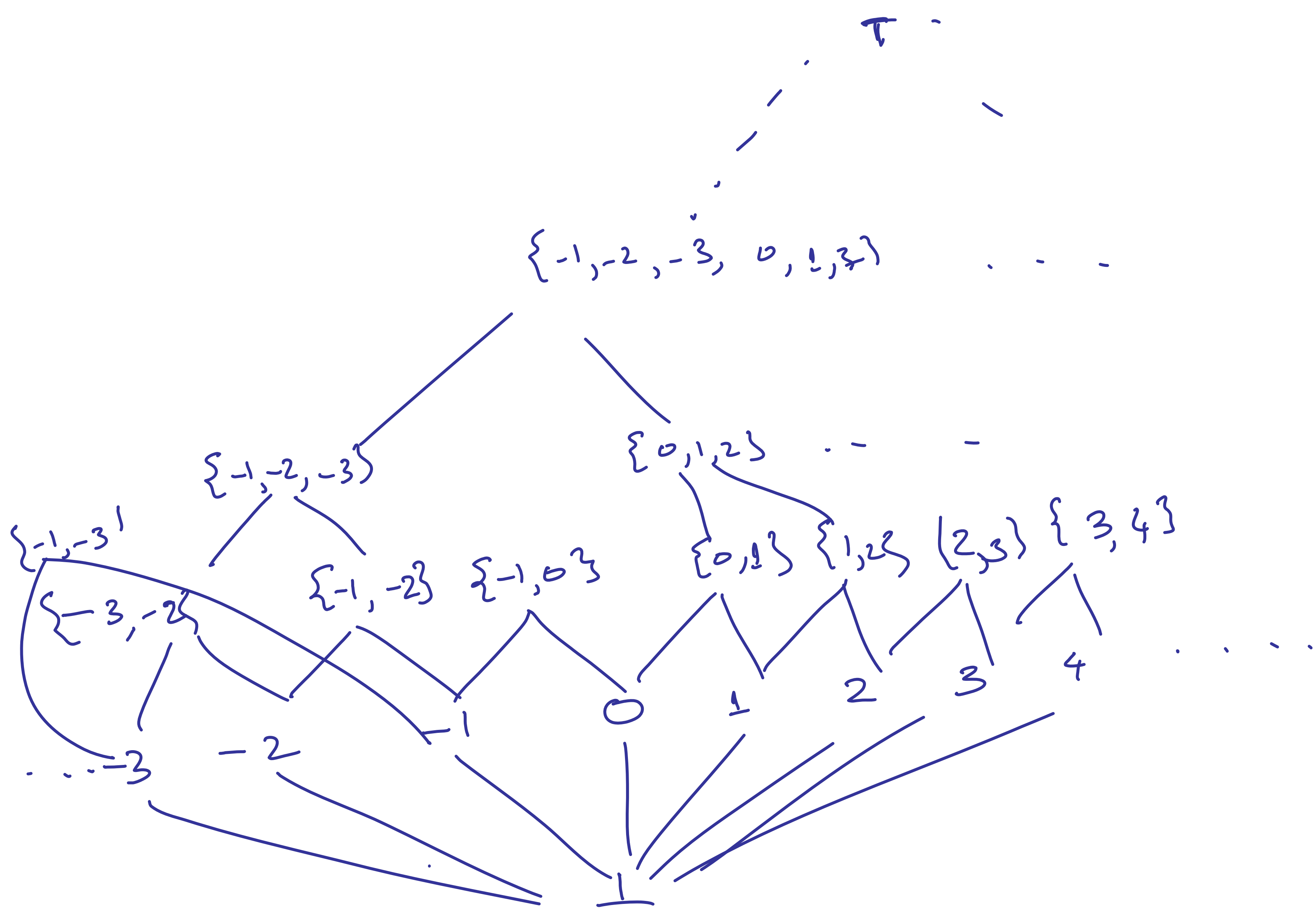
Lecture #2

General approach

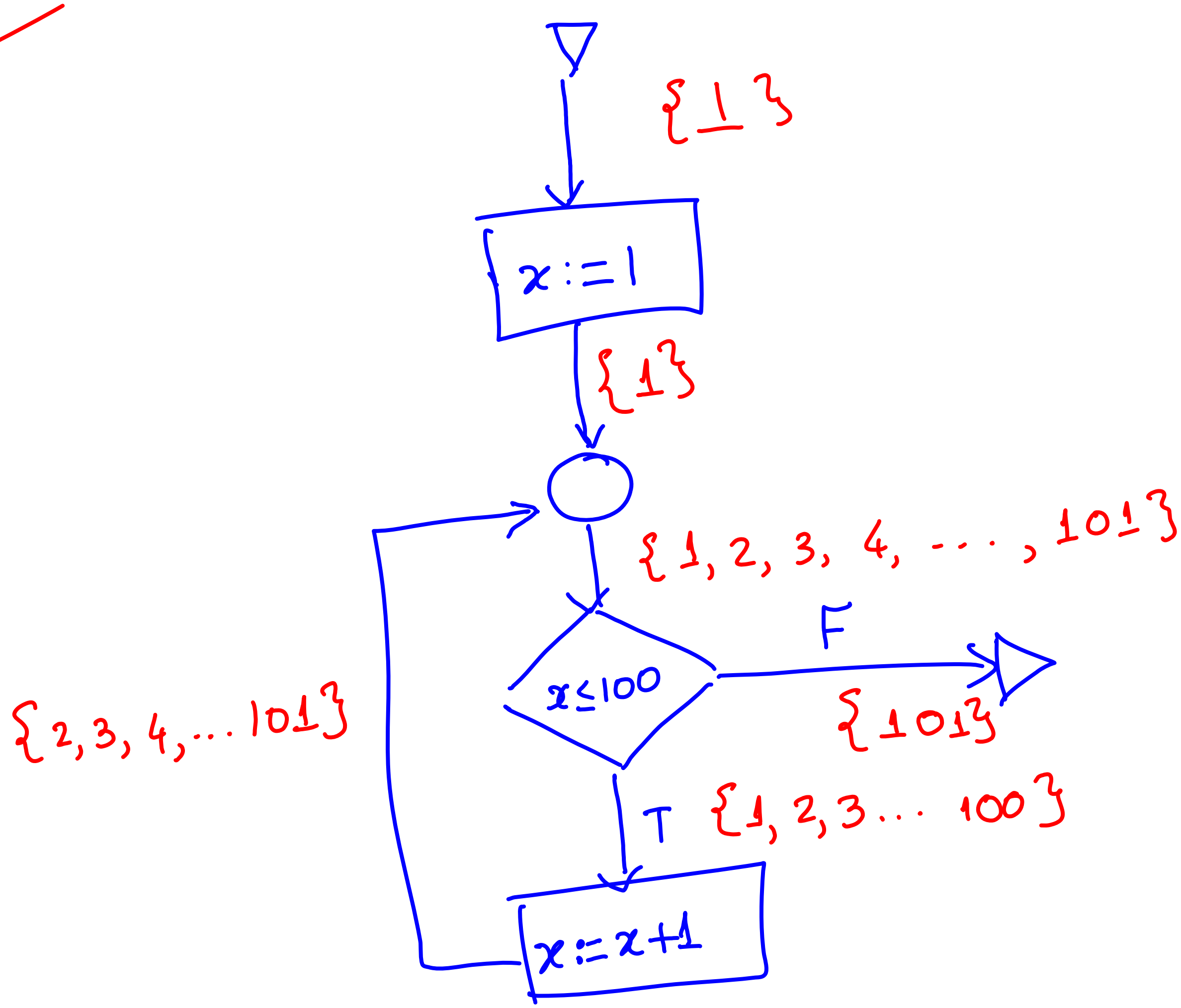
- Behaviors of program modelled as a semilattice
- Over-approx of program semantics $\div$ fixpoint computed over the semilattice

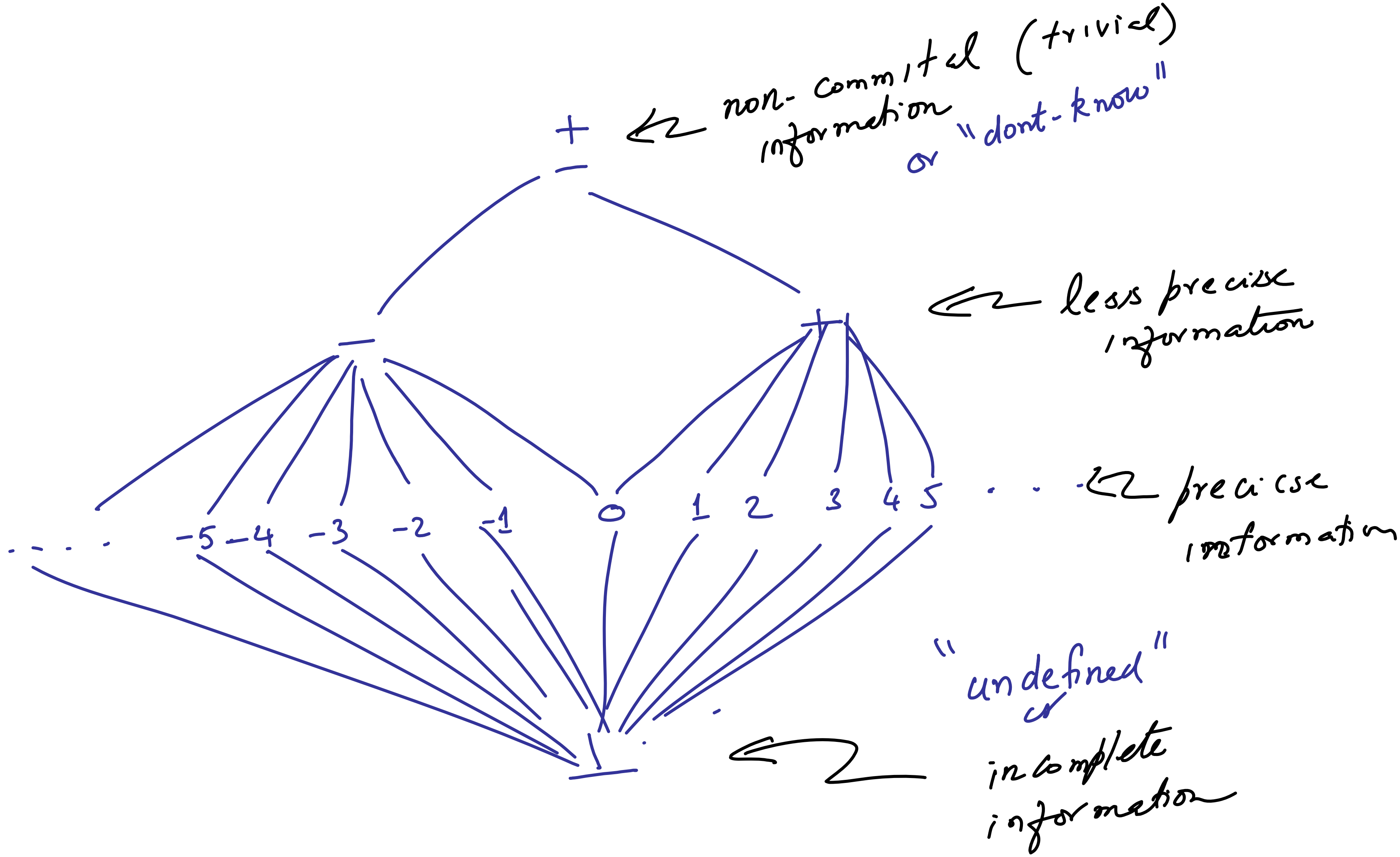
Semilattice $\div$ poset in which
* either every pair of elements has a supremum, or
* every pair of elements has an infimum

poset : set with a reflexive, anti-symmetric, transitive relation


⊤
{-1, -2, -3, 0, 1, 2}
{-1, -2, -3}
{0, 1, 2}
{-1, -3}
{-3, -2}
{-1, -2}
{-1, 0}
{0, 1}
{1, 2}
{2, 3}
{3, 4}
-3
-2
-1
0
1
2
3
4
⊥

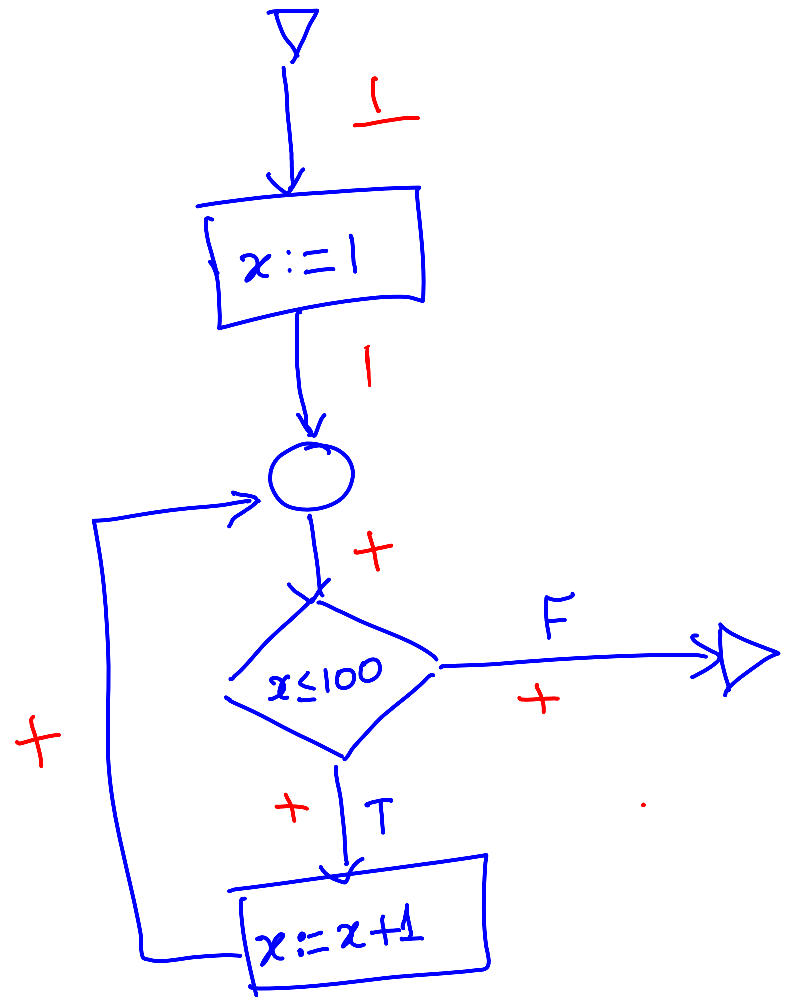
Fixpoint

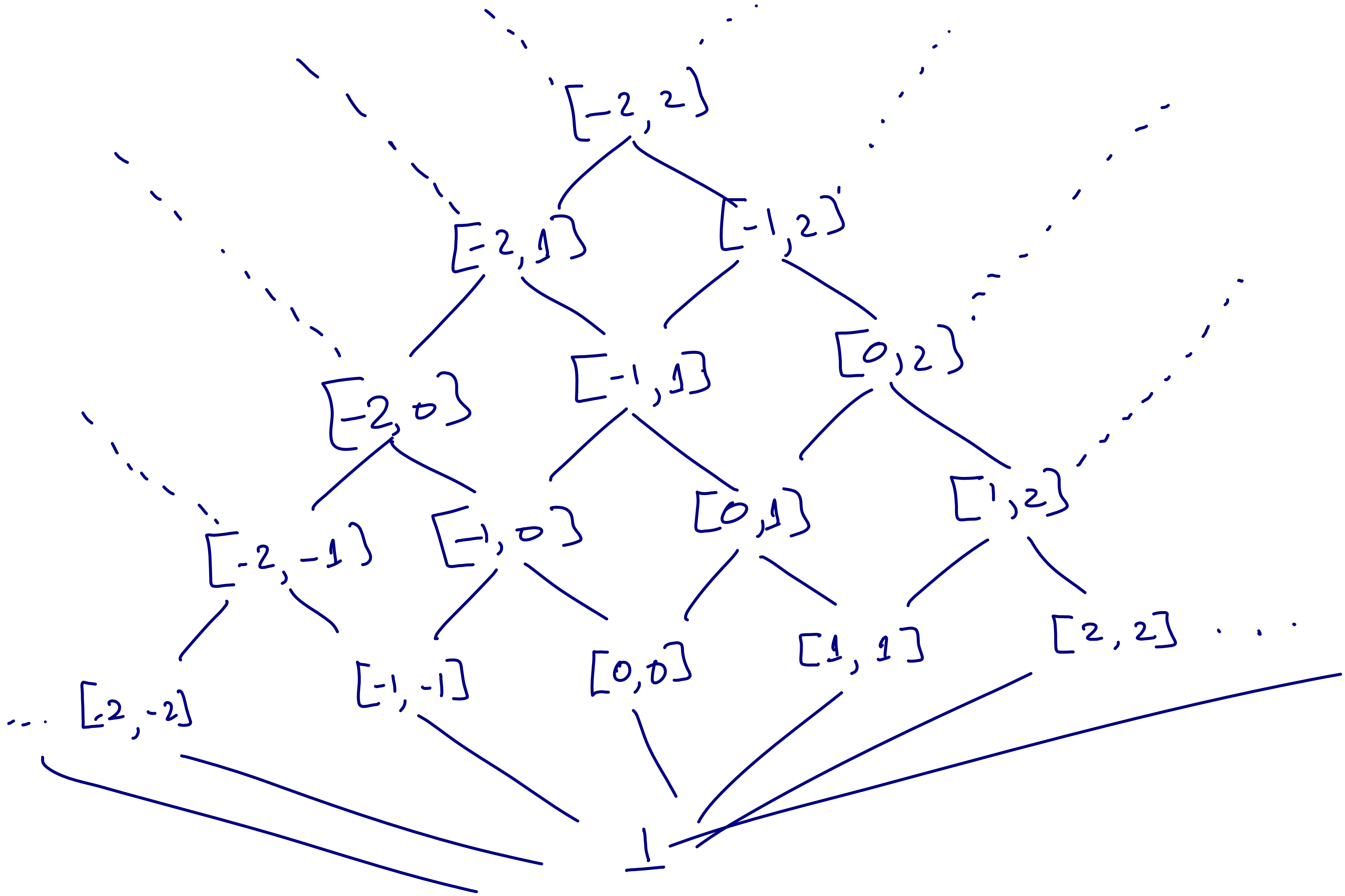




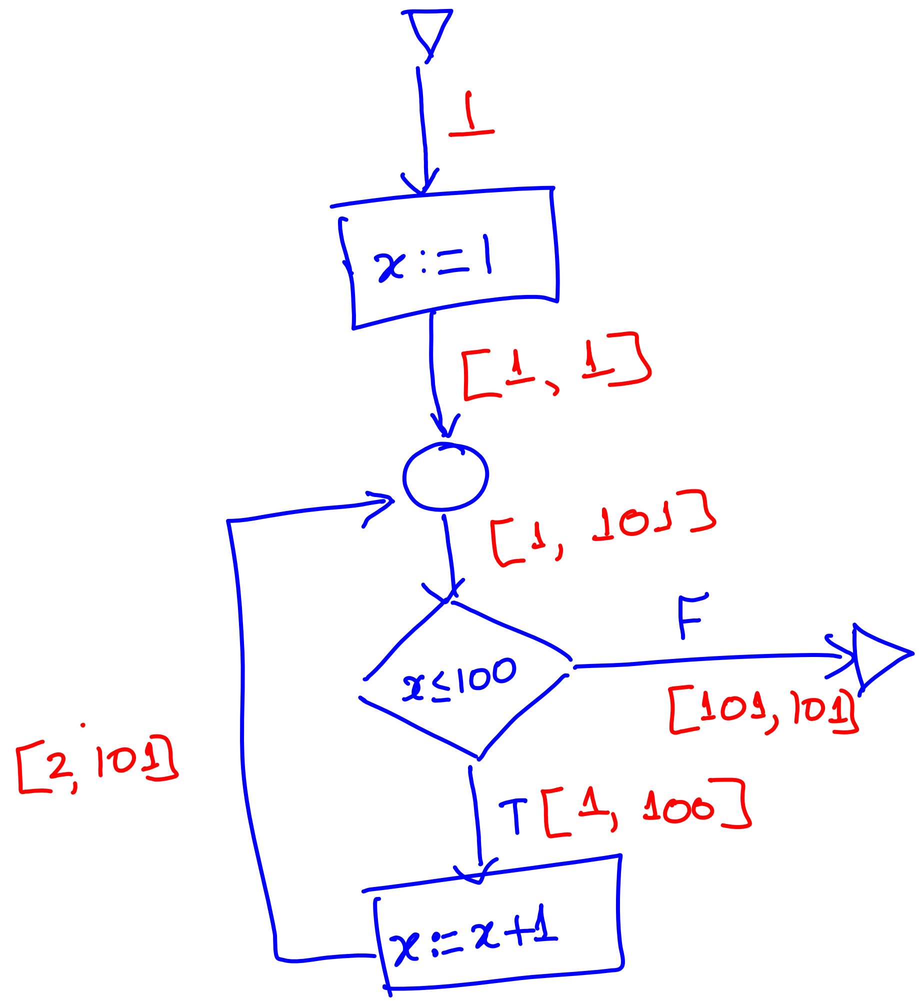
$\top$
non-commital (trivial) information or "dont-know"
$-$ $+$
less precise information
... -5 -4 -3 -2 -1 0 1 2 3 4 5 ...
precise information
$\perp$
"undefined" or incomplete information

Fixpoint




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

Fixpoint



So...

an abstract interpretation is really

$$\langle \underbrace{D}_{\text{domain}}, \underbrace{\circ}_{\text{meet}}, \underbrace{\leq}_{\text{p.o}}, \underbrace{\top}_{\text{Top}}, \underbrace{\bot}_{\text{bottom}}, \underbrace{I}_{\text{interpretation}} \rangle$$

$I : D \to D$

Science of Sound Abstract Interpretations

$$\langle D, \circ_D, \leq_D, \top_D, \bot_D, \sqcap_D \rangle \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} \langle A, \circ_A, \leq_A, \top_A, \bot_A, \sqcap_A \rangle$$

~~eg:~~

Sets of Integers $\rightleftarrows$ Signs

Sets of Integers $\rightleftarrows$ Intervals

Abstract interpretations <u>themselves</u> form a lattice!

Specifying an abstract interpretation

$$\langle D, \circ_D, \leq_D, \top_D, \bot_D, I_D \rangle \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} \langle A, \circ_A, \leq_A, \top_A, \bot_A, I_A \rangle$$

$$\alpha : D \longrightarrow A$$

$$\gamma : A \longrightarrow D$$

α, γ form a Galois connection iff

1. α, γ are order preserving
   $\forall d_1, d_2 \in D \quad d_1 \leq_D d_2 \Rightarrow \alpha(d_1) \leq_A \alpha(d_2)$
   $\forall a_1, a_2 \in A \quad a_1 \leq_A A_2 \Rightarrow \gamma(a_1) \leq_D \gamma(a_2)$
2. $\forall d \in D . d \leq \gamma(\alpha(d))$
3. $\forall a \in A \quad a = \alpha(\gamma(a))$

From Galois connection to abstract state transition fn.

$$\langle D, \circ_D, \leq_D, \top_D, \bot_D, I_D \rangle \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} \langle A, \circ_A, \leq_A, \top_A, \bot_A, I_A \rangle$$

Sp. $\langle \alpha, \gamma \rangle$ form a Galois connection

Can define $\underline{\underline{I_A}}$ in terms of I_D, α, γ.

$$I_A(a) = \alpha(I_D(\gamma(a)))$$

ie.. $I_A = \alpha \circ I_D \circ \gamma$

$$\langle D, \circ_D, \leq_D, \top_D, \bot_D, I_D \rangle \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} \langle A, \circ_A, \leq_A, \top_A, \bot_A, I_A \rangle$$

$$I_A = \alpha \circ I_D \circ \gamma$$

$$\boxed{\underline{\underline{\text{Theorem:}}} \quad \mathrm{Reach}(I_D) \leq_D \gamma(\mathrm{Reach}(I_A))}$$

Thus, any property proved on I_A carries over to I_D !!

Recipe for analysis:

Program's concrete interpretation : $\mathcal{C} = \langle D, \circ_D, \leq_D, \top_D, \bot_D, I_D \rangle$

Concrete semantics : Least Fix Point (I_D)

Difficulty : Least Fix Point (I_D) may be expensive to compute, or may not converge.

Solution: Come up with an abstract domain A and a Galois connection $D \underset{\gamma}{\overset{\alpha}{\rightleftarrows}} A$

Immediately get : $\mathcal{A} = \langle A, \circ_A, \leq_A, \top_A, \bot_A, I_A \rangle$

$I_A = \gamma \circ I_D \circ \alpha$

Abstract Semantics : Least Fix Point (I_A)

Hopefully, easier to compute!

Convergence of fixpoints

Consider $\mathcal{A} = \langle A, \circ_A, \leq_A, \top_A, \bot_A, \mathcal{I}_A \rangle$

If A has finite height then $LFP(\mathcal{I}_A)$ will converge even if A is infinite

If A has infinite height (or large finite height) then $LFP(\mathcal{I}_A)$ may not converge (or may be too expensive to compute)

Two techniques:

1) Widening: to accelerate and ensure convergence of fixpoints (even if A has infinite height)

2) Narrowing: to improve precision of fixpoint computed by widening

Consider $\mathcal{A} = \langle A, \cup, \leq, \top, \perp, I \rangle$

<u>Fixpoint computation:</u>

$$A_0 = \perp$$
$$A_1 = A_0 \cup I(A_0)$$
$$A_2 = A_1 \cup I(A_1)$$
$$\vdots$$
$$A_n = A_{n-1} \cup I(A_{n-1})$$
$$\vdots$$

(may not converge)

__Widening__ ∇ is a binary operator

$$\nabla : A \times A \xrightarrow[\text{partial}]{} A$$

$S_1 \nabla S_2$ __defined__ only if $S_1 \leq S_2$

Two properties of ∇:

__1__. Sp. $S_1 \nabla S_2 = S_3$, then $S_1 \leq S_2 \leq S_3$

2. For any infinite sequence

$S_0 \leq S_1 \leq S_2 \leq S_3 \ldots\ldots\ldots\ldots$

Sp. $R_0 = S_0$

$R_i = R_{i-1} \nabla (R_{i-1} \cup S_i)$, $i > 0$

__Then__

$R_0 \leq R_1 \leq R_2 \leq R_3 \ldots\ldots\ldots$ converges!!

Consider $\mathcal{A} = \langle A, \cup, \leq, \top, \bot, I \rangle$

Fixpoint computation:

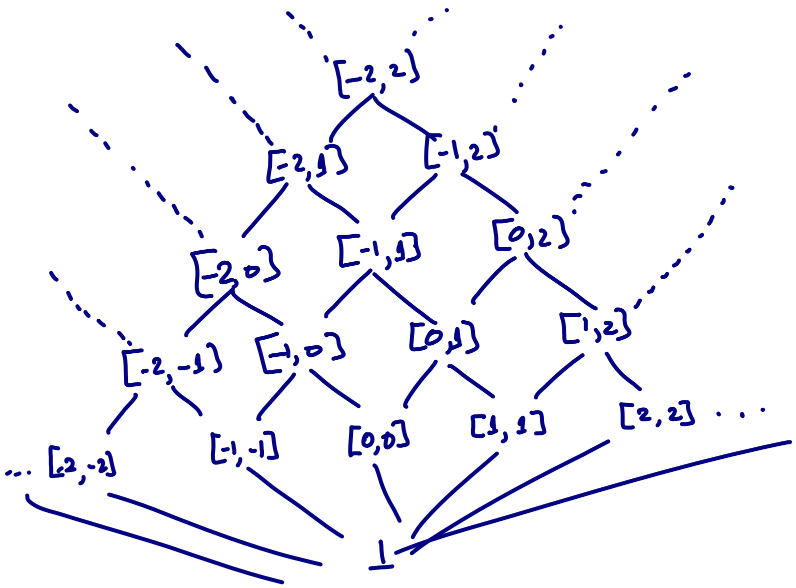
$$A_0 = \bot$$
$$A_1 = A_0 \cup I(A_0)$$
$$A_2 = A_1 \cup I(A_1)$$
$$\vdots$$
$$A_n = A_{n-1} \cup I(A_{n-1})$$
$$\vdots$$

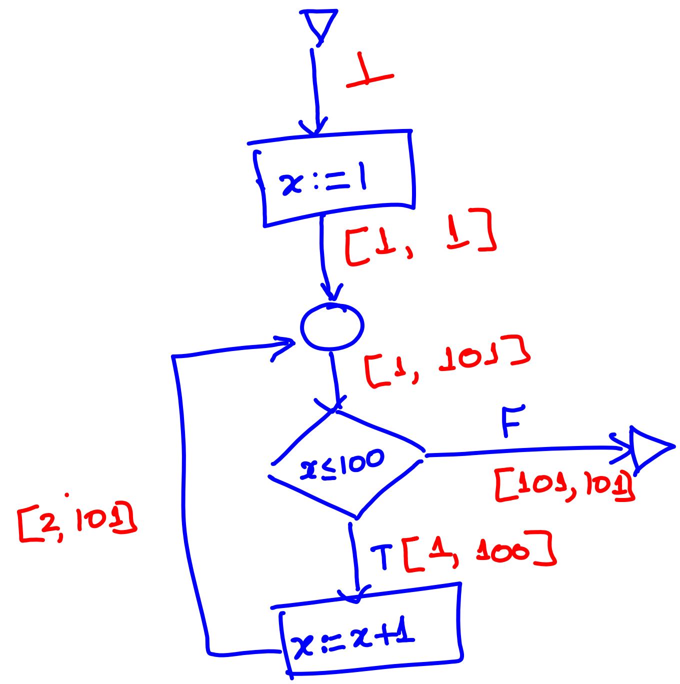
(may not converge)

$$R_0 = \bot$$
$$R_1 = R_0 \nabla (R_0 \cup I(R_0))$$
$$R_2 = R_1 \nabla (R_1 \cup I(R_1))$$
$$\vdots$$
$$R_n = R_{n-1} \nabla (R_{n-1} \cup I(R_{n-1}))$$
$$\vdots$$

Guaranteed to converge!

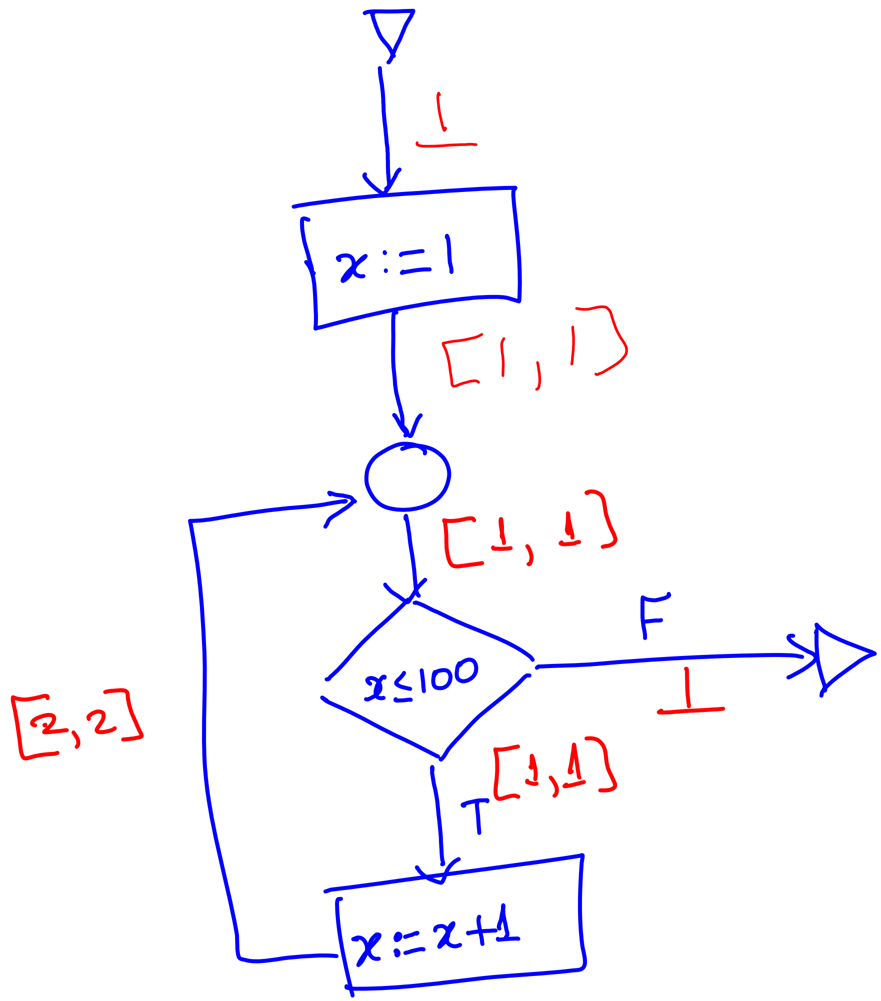
Widening over intervals



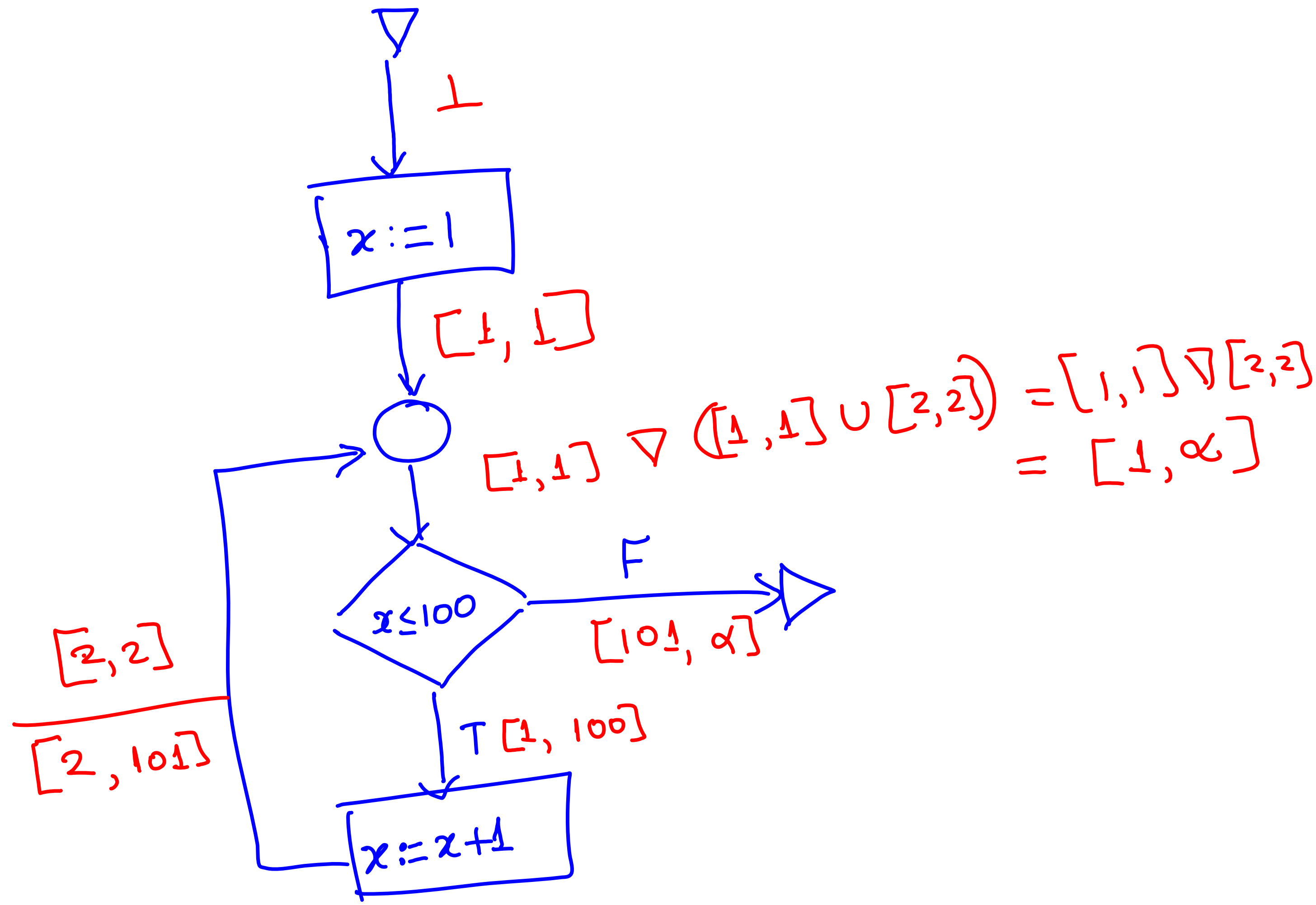


$$[i, j] \nabla [k, l] = [\text{if } k < i \text{ then } -\infty \text{ else } i, \; \text{if } l > j \text{ then } +\infty \text{ else } j]$$

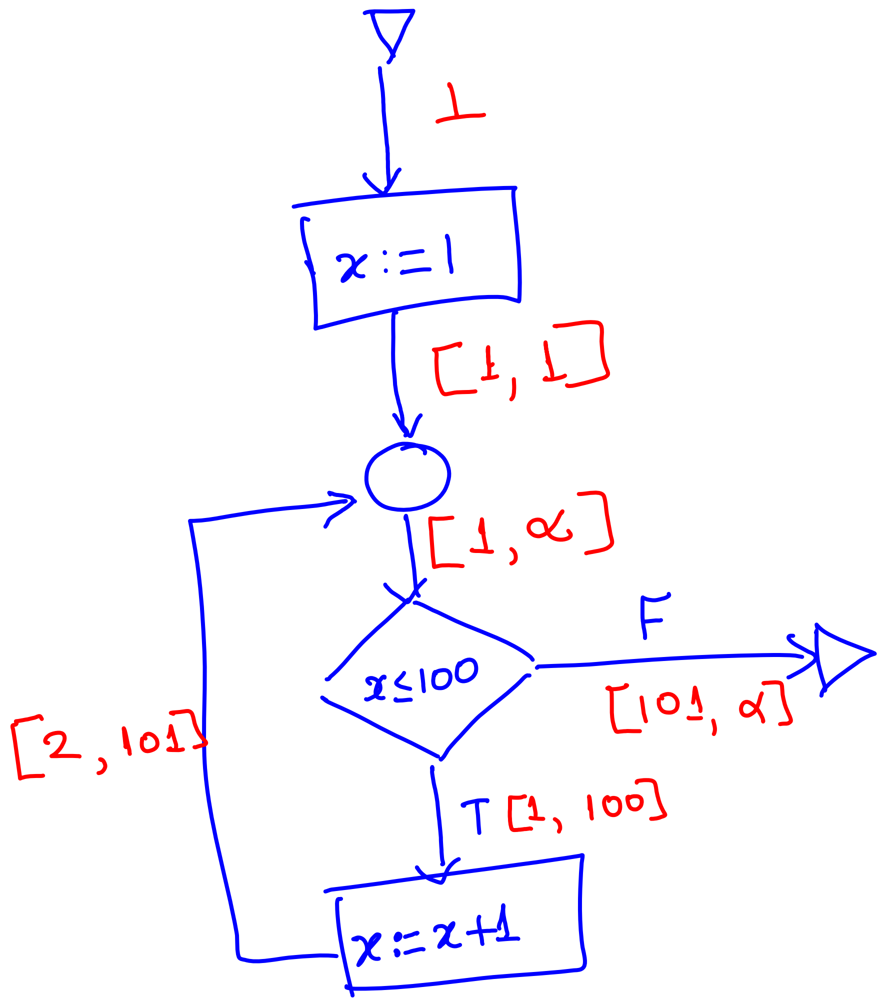
First iteration



Second iteration



Fixpoint



Narrowing: Technique to recover some of the precision lost in widening

Consider $\mathcal{A} = \langle A, \cup, \leq, \top, \bot, I \rangle$

$R_0 = \bot$

$R_1 = R_0 \nabla (R_0 \cup I(R_0))$

$R_2 = R_1 \nabla (R_1 \cup I(R_1))$

$\vdots$

$R_n = R_{n-1} \nabla (R_{n-1} \cup I(R_{n-1}))$

$\vdots$

Guaranteed to converge!

Let $\underline{\underline{R_m}}$ be the fixpoint

$\boxed{\text{Narrowing: Technique to recover some of the precision lost in widening}}$

Consider $\mathcal{A} = \langle A, \cup, \leq, \top, \bot, I \rangle$ $\quad T(X) = X \cup I(X)$

$R_0 = \bot$

$R_1 = R_0 \nabla (R_0 \cup I(R_0))$

$R_2 = R_1 \nabla (R_1 \cup I(R_1))$

$\vdots$

$R_n = R_{n-1} \nabla (R_{n-1} \cup I(R_{n-1}))$

$\vdots$

Guaranteed to converge!

Let $\underline{\underline{R_m}}$ be the fixpoint

Let $S_0 = R_m$

Note: $lfp(T) \leq R_m$

Let $S_1 = I(S_0)$

$S_1 \leq S_0$

$S_2 = I(S_1)$

$S_2 \leq S_1 \leq S_0$

[due to monotonicity of I]

Get $S_k \ldots \ldots S_2 \leq S_1 \leq S_0$

fixpoint

$\boxed{lfp(I) \leq S_k}$

but this may not converge either!

Narrow operator : Δ

A technique to enforce convergence of decreasing sequences...

let $S_0 \geq S_1 \geq S_2 \dots\dots$

$\Delta : A \longrightarrow A$ is a <u>narrow</u> operator if it satisfies:

1. $\forall\ s, s' \in A$. $\quad s \geq s' \Rightarrow s \geq s \Delta s' \geq s'$

2. for any decreasing sequence

$$S_0 \geq I(S_0) \geq I^2(S_0) \dots\dots$$

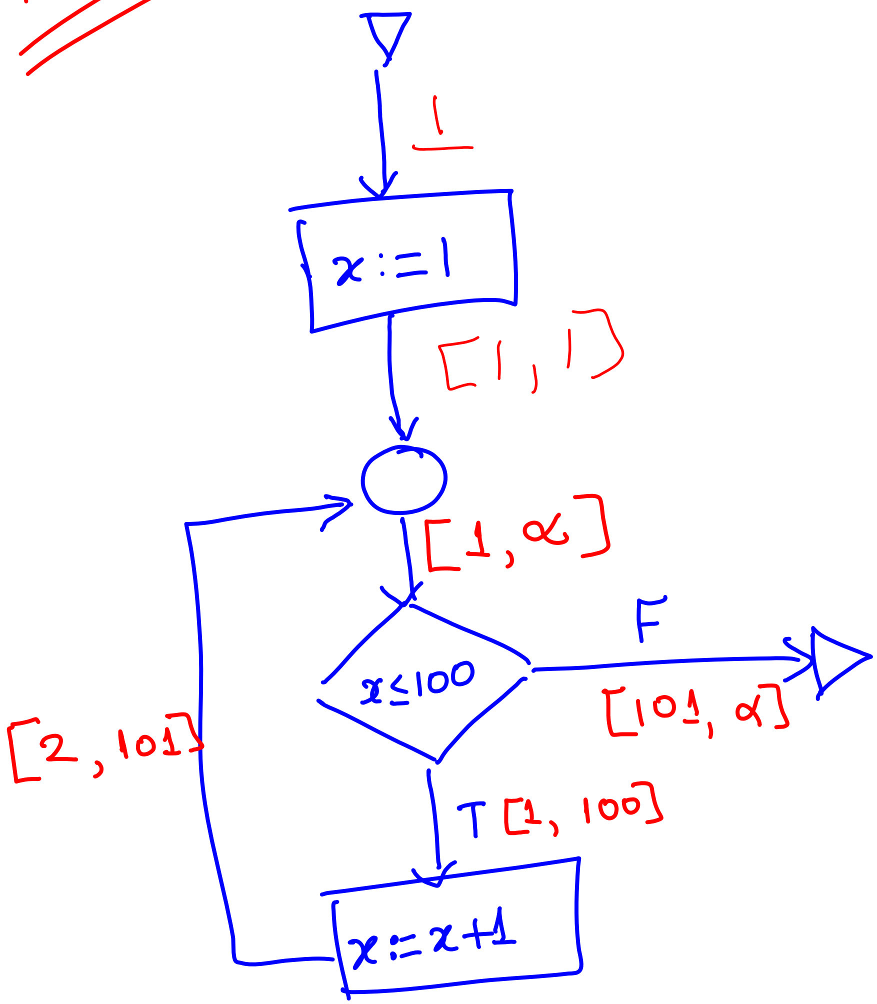
the sequence

$$t_0 = S_0, \quad t_n = t_{n-1} \nabla I(t_n)$$

i.e. $t_0 \geq t_1 \geq t_2 \dots\dots$,

<u>Converges</u>

Recall ...



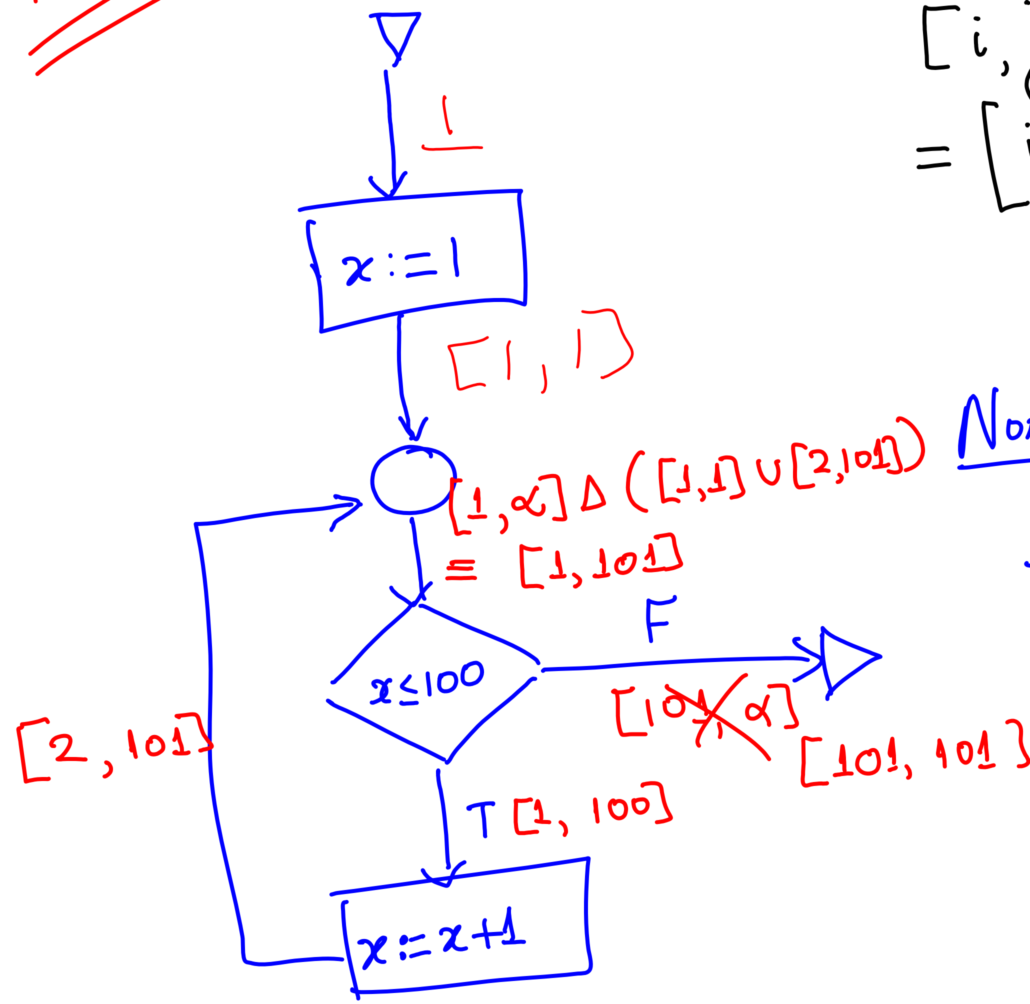
Define:

$$[i, j] \,\Delta\, [k, l] = \left[\begin{array}{l} \text{if } i = -\infty \text{ then } k \text{ else } i, \\ \text{if } j = +\infty \text{ then } l \text{ else } j \end{array}\right]$$

Note:

Δ satisfies both conditions for a narrow operator

Recall …



Define:

$$[i, j] \,\Delta\, [k, l]$$
$$= [\text{if } i = -\infty \text{ then } k \text{ else } i, \; \text{if } j = +\infty \text{ then } l \text{ else } j]$$

Note:

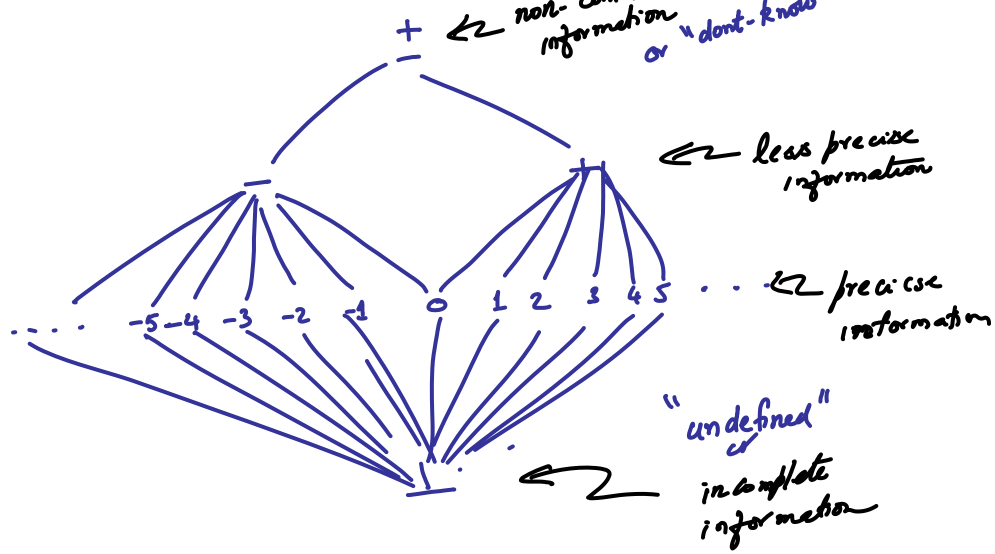
Δ satisfies both conditions for a narrow operator

Assignment Questions:

1

Recall the "signs" abstract domain

This domain keeps track of the sign of an integer



1) Design an abstract domain with finite height that keeps track of the **remainder when divided by 3** (mod 3)

2) What are α & γ? Prove that they form a Galois connection

3) What are the abstract interpretations of the usual operators:

a) assignments

b) conditionals

(2) Recall the definition of the Narrow operator

$\mathcal{A} : \langle A, \oplus, \leq, \top, \bot, I \rangle$

$T(X) = X \cup I(X)$

Narrow operator : Δ

A technique to enforce convergence of decreasing sequences...

let $s_0 \geq s_1 \geq s_2 \ldots\ldots$

$\Delta : A \longrightarrow A$ is a narrow operator if it satisfies:

(1) $\forall\ s, s' \in A$. $s \geq s' \Rightarrow s \geq s \Delta s' \geq s'$

(2) for any decreasing sequence $s_0 \geq I(s_0) \geq I^2(s_0) \ldots\ldots$

the sequence $t_0 = s_0$, $t_n = t_{n-1} \nabla I(t_n)$

i.e. $t_0 \geq t_1 \geq t_2 \ldots\ldots$, converge.

Suppose you start the decreasing sequence

$s_0 \geq I(s_0) \ldots\ldots$

from a widened fixpoint that is above the least fix point of a system

i.e. $s_0 \geq lfp(T)$

PROVE THAT THE FIX POINT OF:

$t_0 \geq t_1 \geq t_2 \ldots$

IS A SOUND fixpoint (i.e... it is "above" the least fix point)

i.e... $lfp(t_0, t_1, t_2, \ldots) \geq lfp(T)$

HINT: SEE 9.3.1 IN COUSOT-COUSOTS' PAPER

3. Consider two variants of the narrow operator

V1: $[i, j] \Delta [k, l] =$
$[\text{if } i = -\infty \text{ then } k \text{ else } \min(i, k) \text{ fi},$
$\text{if } j = \infty \text{ then } l \text{ else } \max(j, l)]$

V2: $[i, j] \Delta [k, l] = [\min(i, k), \max(j, l)]$

For both these variants:
Do they satisfy the <u>two</u> conditions for a <u>narrow</u> operator
(ie.. soundness & convergence) ?