

Kandinsky - Abstract Interpretation, 1925

Abstract Interpretation and Applications in Security, Data Science, and Machine Learning

OPLSS 2025

Caterina Urban

Inria & École Normale Supérieure | Université PSL

Static Analysis of Liveness Properties

The Art of Losing Precision

No Surprises, Please

What Could Possibly Go Right?

It's Complicated

Trace Properties



Liveness Properties = "Something Good Eventually Happens"

Example

• Termination: $T \stackrel{\text{def}}{=} \Sigma^*$

Liveness Property Verification

T cannot be verified by testing

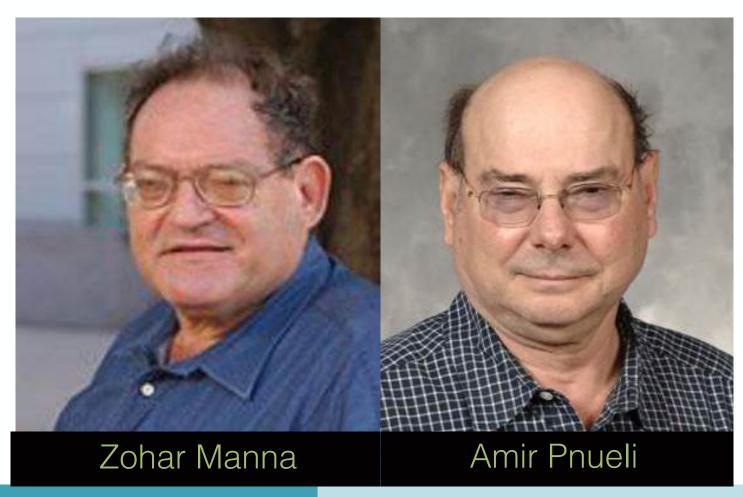




• falsifying T requires finding an infinite execution not in T

Liveness Properties

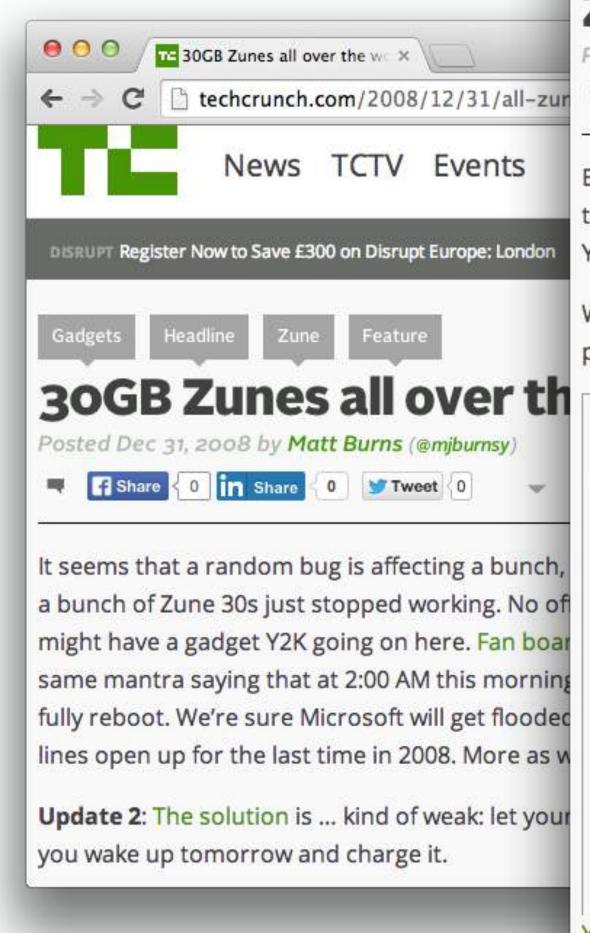
- Guarantee Properties
 "something good eventually happens at least once"
 - Example: Program Termination
- Recurrence Properties "something good eventually happens infinitely often"
 - Example: Starvation Freedom

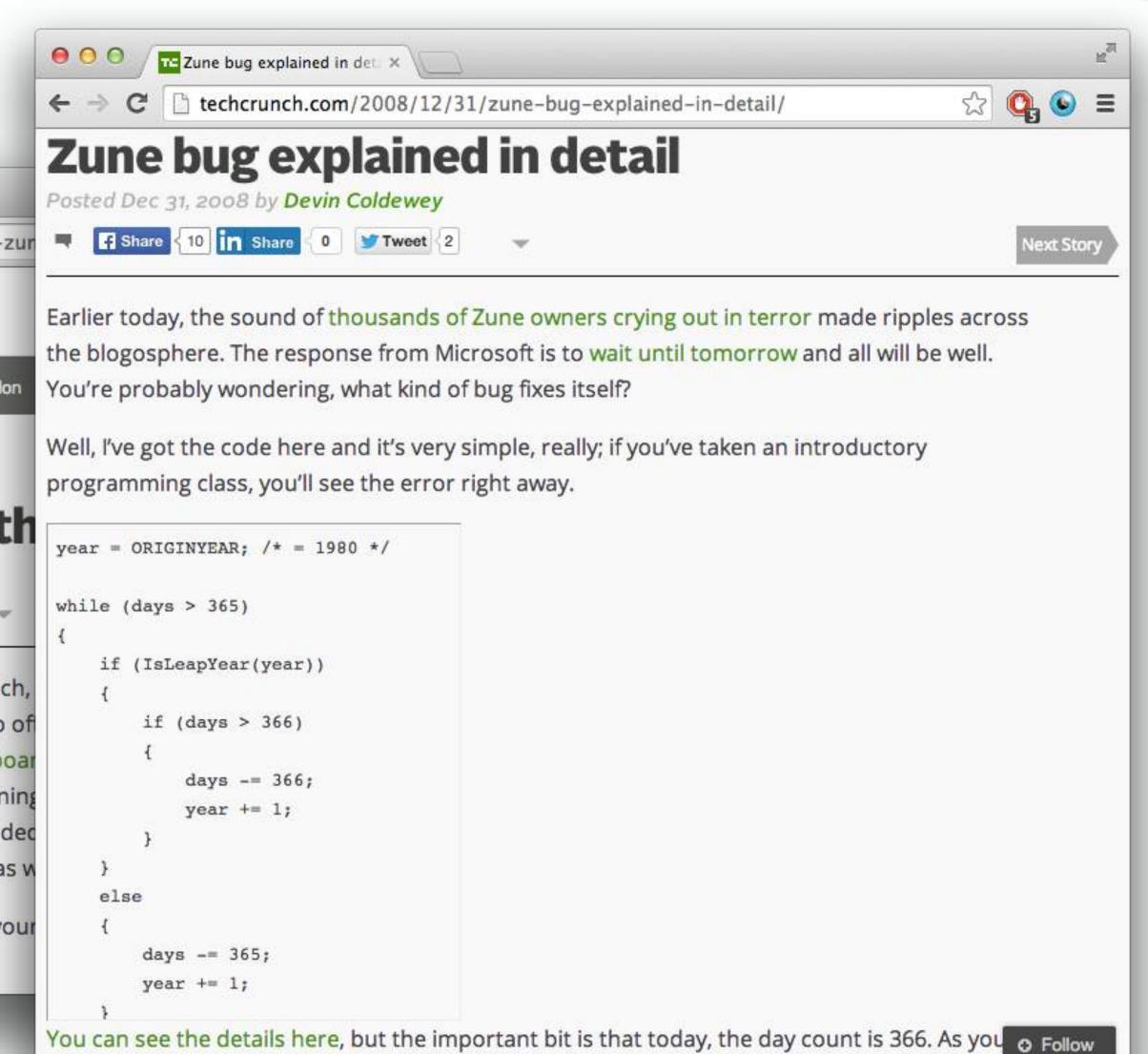


Program Termination

The Zune Bug

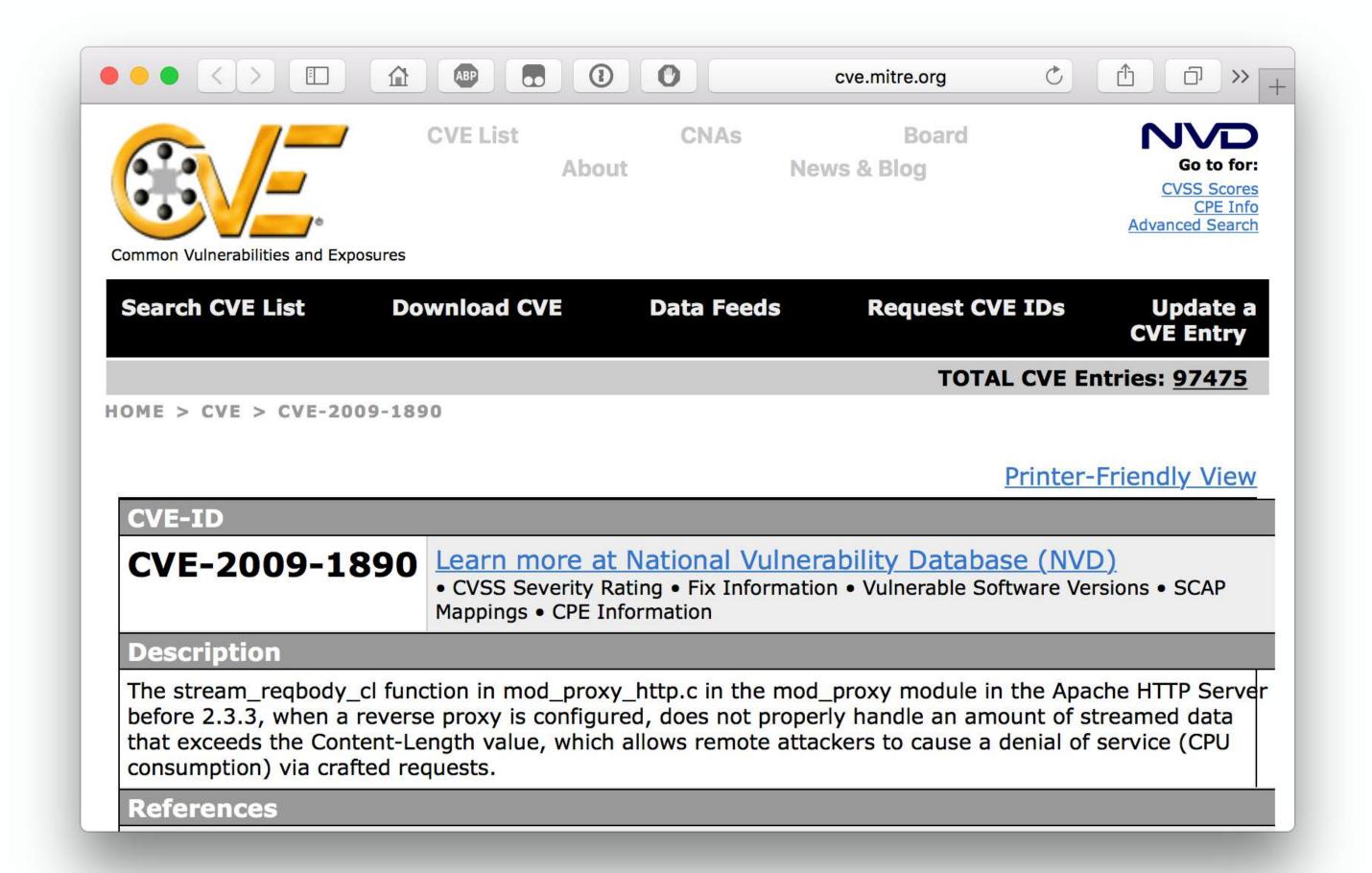
31 December 2008





Apache HTTP Server

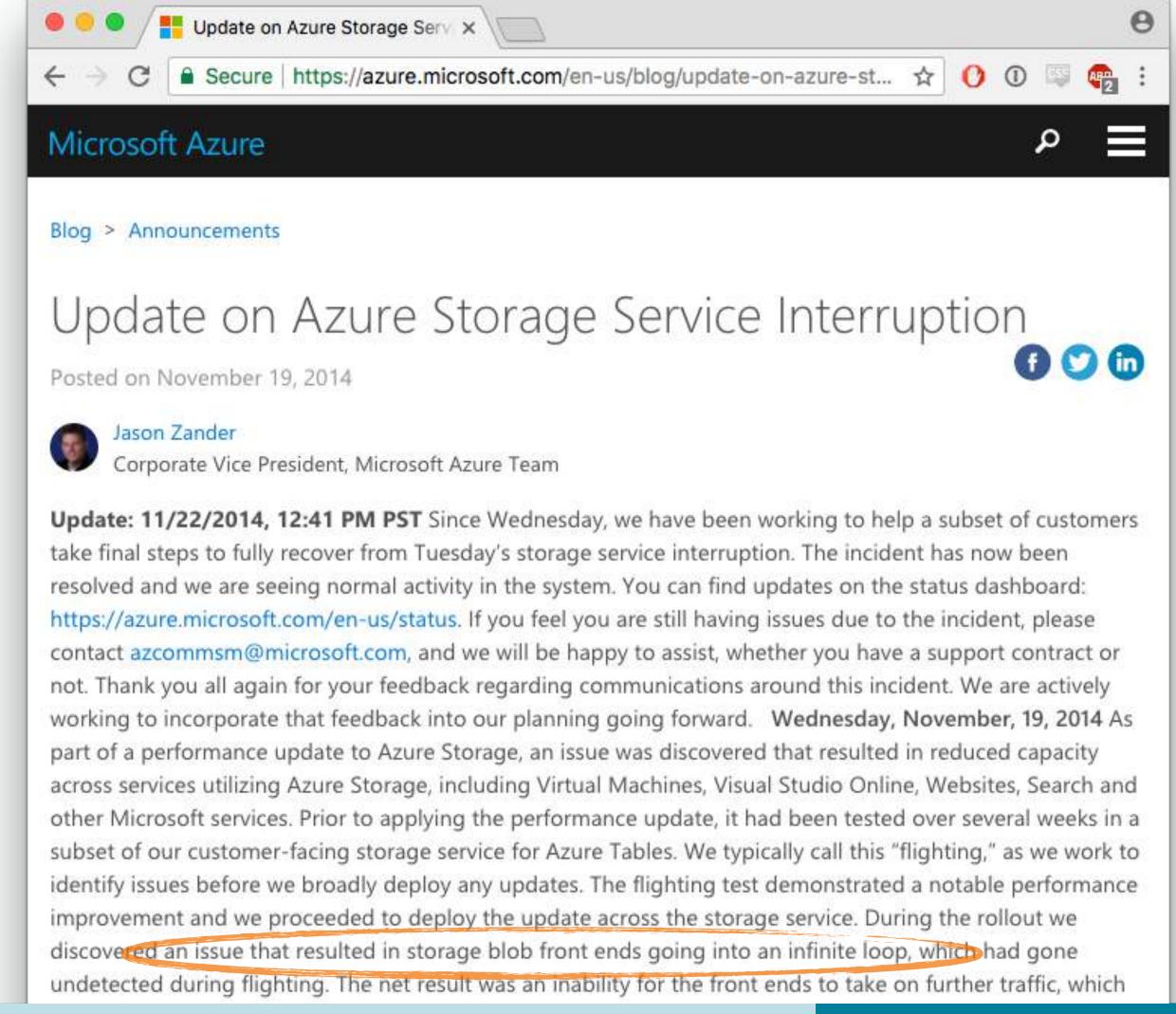
Versions < 2.3.3



denial-of-service attacks

Azure Storage Service

19 November 2014



Potential and Definite Termination

Potential Termination

Definition

A program with trace semantics $\mathscr{M} \in \mathscr{P}(\Sigma^{\infty})$ may terminate if and only if $\mathscr{M} \cap \Sigma^* \neq \varnothing$

Definite Termination

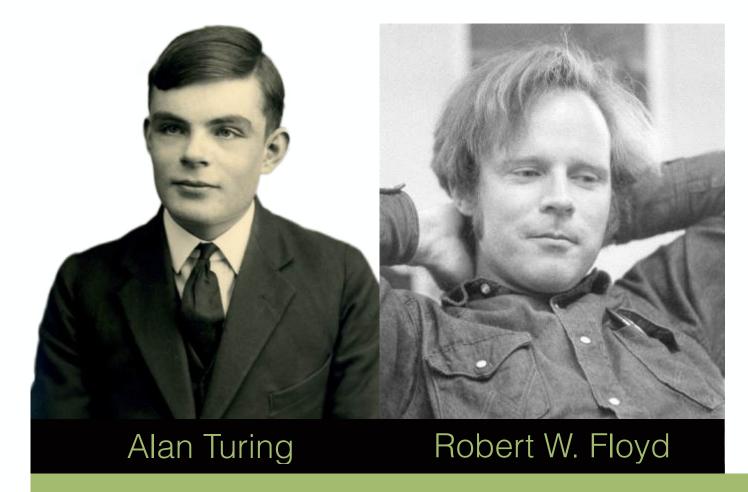
Definition

A program with trace semantics $\mathscr{M} \in \mathscr{S}(\Sigma^{\infty})$ must terminate if and only if $\mathscr{M} \subseteq \Sigma^{*}$

In absence of non-determinism, potential and definite termination coincide

Definite Termination

Ranking Functions



Definition

Given a transition system $\langle \Sigma, \tau \rangle$, a **ranking function** is a partial function $f \colon \Sigma \rightharpoonup \mathcal{W}$ from the set of program states Σ into a well-ordered set $\langle \mathcal{W}, \leq \rangle$ whose value *strictly decreases* through transitions between states, that is, $\forall \sigma, \sigma' \in \text{dom}(f) \colon (\sigma, \sigma') \in \tau \Rightarrow f(\sigma') < f(\sigma)$

The best known well-ordered sets are naturals $\langle \mathbb{N}, \leq \rangle$ and ordinals $\langle \mathbb{O}, \leq \rangle$

Ranking Functions

Example

```
\begin{array}{l}
1x \leftarrow [-\infty, +\infty] \\
\text{while } 2(1 - x < 0) \text{ do} \\
3x \leftarrow x - 1 \\
\text{done}^4
\end{array}

\begin{array}{l}
\sum \stackrel{\text{def}}{=} \{1,2,3,4\} \times \mathscr{E}
```

$$\tau \stackrel{\text{def}}{=} \{ ((\mathbf{1}, \rho), (\mathbf{2}, \rho[X \mapsto v])) \mid \rho \in \mathcal{E}, v \in \mathbb{Z} \} \\
\cup \{ ((\mathbf{2}, \rho), (\mathbf{3}, \rho)) \mid |\rho \in \mathcal{E}, \exists v \in E[[1 - x]]\rho : v < 0 \} \\
\cup \{ ((\mathbf{3}, \rho), (\mathbf{2}, \rho[X \mapsto v])) \mid \rho \in \mathcal{E}, v \in E[[x - 1]]\rho \} \\
\cup \{ ((\mathbf{2}, \rho), (\mathbf{4}, \rho)) \mid |\rho \in \mathcal{E}, \exists v \in E[[1 - x]]\rho : v \not< 0 \}$$

Ranking Functions

Example

$$1x \leftarrow [-\infty, +\infty]$$
while $2(1 - x < 0)$ do
$$3x \leftarrow x - 1$$
done4

Most obvious ranking function: a mapping $f\colon \Sigma \to \mathbb{O}$ from each program state to (an upper bound on) the number of steps until termination

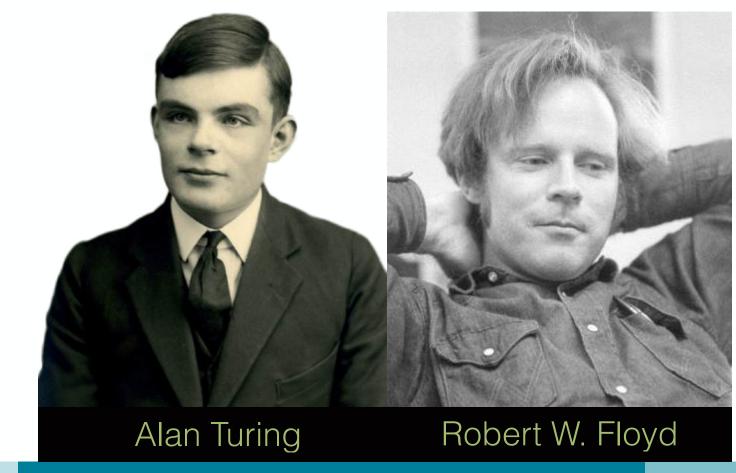
We define $f: \Sigma \to \mathbb{O}$ by partitioning with respect to the program control points, i.e., $f: \mathcal{L} \to (\mathcal{E} \to \mathbb{O})$

$$f(\mathbf{4}) \stackrel{\text{def}}{=} \lambda \rho.0$$

$$f(\mathbf{2}) \stackrel{\text{def}}{=} \lambda \rho. \begin{cases} 1 & 1 - \rho(x) \not< 0 \\ 2\rho(x) - 1 & 1 - \rho(x) < 0 \end{cases}$$

$$f(\mathbf{3}) \stackrel{\text{def}}{=} \lambda \rho. \begin{cases} 2 & 2 - \rho(x) \not< 0 \\ 2\rho(x) - 2 & 2 - \rho(x) < 0 \end{cases}$$

$$f(\mathbf{1}) \stackrel{\text{def}}{=} \lambda \rho. \omega$$



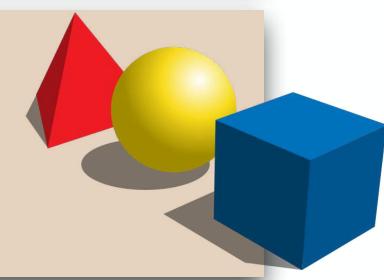
Static Termination Analysis

3-Step Recipe

practical tools targeting specific programs



abstract semantics, abstract domains algorithmic approaches to decide program properties



concrete semantics mathematical models of the program behavior



Static Termination Analysis

Program Termination Semantics

practical tools targeting specific programs

abstract semantics, abstract domains algorithmic approaches to decide program properties

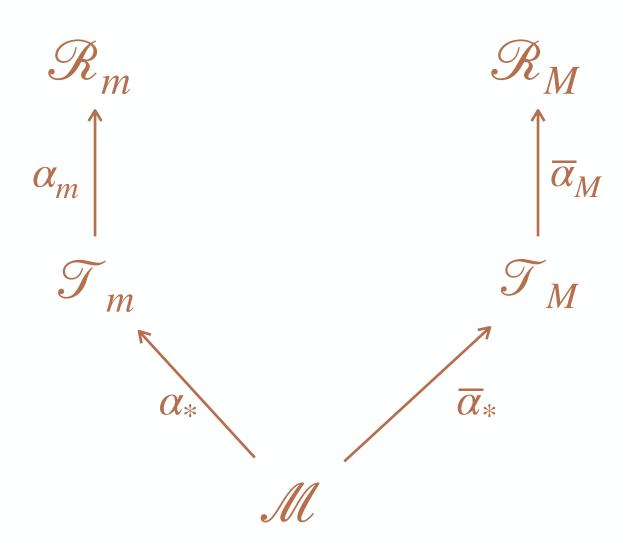




concrete semantics mathematical models of the program behavior



(Yet Another) Hierarchy of Semantics

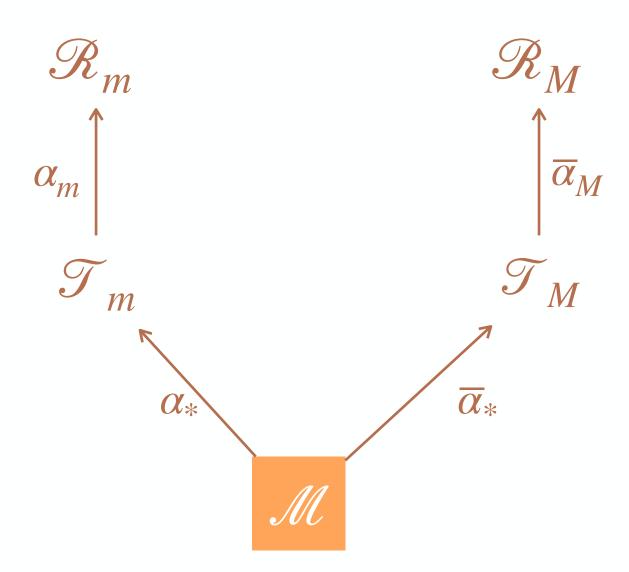


termination semantics

termination trace semantics

maximal trace semantics

(Yet Another) Hierarchy of Semantics



termination semantics

termination trace semantics

maximal trace semantics

Maximal Trace Semantics

Least Fixpoint Formulation

$$\mathcal{M} = \operatorname{lfp}_{\Sigma^{\omega}}^{\sqsubseteq} F$$

$$F(T) \stackrel{\mathsf{def}}{=} \mathcal{B} \cup \tau; T$$

•
$$F^0(\emptyset) = \Sigma^{\omega}$$

•
$$F^1(F^0) = \{c\} \cup \{ab\Sigma^\omega, bb\Sigma^\omega, bc\Sigma^\omega\}$$

$$F^0(\varnothing) = \Sigma^\omega$$

$$F^1(F^0) = \{c\} \cup \{ab\Sigma^\omega, bb\Sigma^\omega, bc\Sigma^\omega\}$$

$$F^2(F^1) = \{bc, c\} \cup \{abb\Sigma^\omega, bbb\Sigma^\omega, abc\Sigma^\omega, bbc\Sigma^\omega\}$$

•
$$F_p^3(F_p^2) = \{abc, bbc, bc, c\} \cup \{abbb\Sigma^\omega, bbbb\Sigma^\omega, abbc\Sigma^\omega, bbbc\Sigma^\omega\}$$

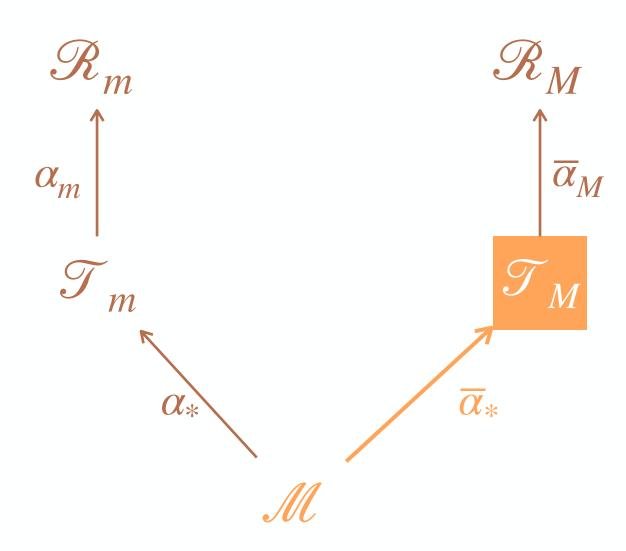
$$\mathcal{M} = \{ab^ic, b^ic, c \mid i \ge 1\} \cup \{ab^\omega, b^\omega\}$$

Maximal Trace Semantics

Example

```
while {}^{1}([-\infty, +\infty] \neq 0) do {}^{2}skip done<sup>3</sup> \Sigma \stackrel{\text{def}}{=} \{\mathbf{1}, \mathbf{2}, \mathbf{3}\} \times \mathscr{E} \tau \stackrel{\text{def}}{=} \{((\mathbf{1}, \rho), (\mathbf{2}, \rho)) \mid \rho \in \mathscr{E}\} \cup \{((\mathbf{2}, \rho), (\mathbf{1}, \rho)) \mid \rho \in \mathscr{E}\} \cup \{((\mathbf{1}, \rho), (\mathbf{3}, \rho)) \mid \rho \in \mathscr{E}\} \mathcal{M} \stackrel{\text{def}}{=} \{(\mathbf{1}, \rho)(\mathbf{2}, \rho)^{*}(\mathbf{3}, \rho) \mid \rho \in \mathscr{E}\} \cup \{(\mathbf{1}, \rho)(\mathbf{2}, \rho)^{\omega} \mid \rho \in \mathscr{E}\}
```

(Yet Another) Hierarchy of Semantics



termination semantics

definite termination trace semantics

maximal trace semantics

Definite Termination Abstraction

$$\langle \mathscr{P}(\Sigma^{\infty}), \sqsubseteq \rangle$$
 $\langle \mathscr{P}(\Sigma^{*}), \subseteq \rangle$ $\overline{\alpha}_{*}$

$$\overline{\alpha}_*(T) \stackrel{\mathrm{def}}{=} \{t \in T \cap \Sigma^* \mid \mathrm{nhdb}(t, T \cap \Sigma^\omega) = \varnothing\}$$
 where $\mathrm{nhdb}(t, T) \stackrel{\mathrm{def}}{=} \{t' \in T \mid \mathrm{pf}(t) \cap \mathrm{pf}(t') \neq \varnothing\}$
$$\mathrm{pf}(t) \stackrel{\mathrm{def}}{=} \{t' \in \Sigma^\infty \backslash \{\epsilon\} \mid \exists t'' \in \Sigma^\infty \colon t = t' \cdot t''\}$$

Example:

$$\alpha_*(\{ab, aba, bb, ba^{\omega}\}) = \{ab, aba\} \text{ since pf}(bb) \cap \text{pf}(ba^{\omega}) = \{b\} \neq \emptyset$$

Order Theory

Tarskian Fixpoint Transfer

Theorem

```
Let \langle C, \leq, \vee, \wedge, \perp, \top \rangle and \langle A, \sqsubseteq, \sqcup, \sqcap, \perp^{\#}, \top^{\#} \rangle be complete lattices, let f \colon C \to C and f^{\#} \colon A \to A be monotonic functions, and let \alpha \colon C \to A be an abstraction function that
```

- is a complete \land -morphism $(\forall S \subseteq C : f(\land S) = \sqcap \{f(s) \mid s \in S\}),$
- satisfies $f^{\#} \circ \alpha \sqsubseteq \alpha \circ f$,
- satisfies the post-fixpoint correspondence $\forall a^\# \in A : f^\#(a^\#) \sqsubseteq a^\# \Rightarrow \exists a \in C : f(a) \leq d \land \alpha(a) = a^\#$ (i.e., each abstract post-fixpoint of $f^\#$ is the abstraction by α of some concrete post-fixpoint of f).

Then, we have the fixpoint abstraction $\alpha(\operatorname{lfp}_c^{\leq} f) = \operatorname{lpf}_{\alpha(c)}^{\sqsubseteq} f^{\#}$

Maximal to Definite Termination Trace Semantics

Tarskian Fixpoint Transfer

Definite Termination Abstraction

$$\langle \mathscr{P}(\Sigma^\infty), \sqsubseteq \rangle$$
 $\langle \mathscr{P}(\Sigma^*), \subseteq \rangle$ complete lattice complete lattice $\overline{\alpha}_*$

Maximal Trace Semantics

$$\mathcal{M} = \operatorname{lfp}_{\Sigma^{\omega}}^{\sqsubseteq} F$$

$$F(T) \stackrel{\mathsf{def}}{=} \mathcal{B} \cup \tau; T$$

$$\overset{\mathsf{monotonic}}{=} \mathcal{B} \cup \tau; T$$

Definite Termination Trace Semantics

$$\mathcal{T}_{M} = \operatorname{lfp}_{\varnothing}^{\subseteq} \overline{F}_{*}$$

$$\overline{F}_{*}(T) \stackrel{\operatorname{def}}{=} \mathscr{B} \cup ((\tau \, ; \, T) \cap (\Sigma^{+} \backslash (\tau \, ; \, (\Sigma^{+} \backslash T)))))$$
 monotonic



 $\overline{\alpha}_*(\mathcal{M}) = \overline{\alpha}_*(\operatorname{Ifp}_{\Sigma^\omega}^{\sqsubseteq} F) = \operatorname{Ifp}_{\varnothing}^{\subseteq} \overline{F}_* = \mathcal{T}_M$

Exercise: prove this 🙂

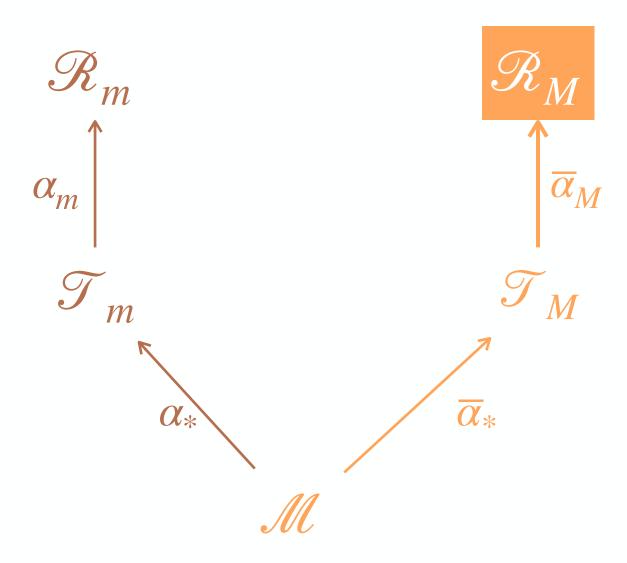
Example

```
while 1([-\infty, +\infty] \neq 0) do 2skip done<sup>3</sup>
```

$$\mathcal{M} \stackrel{\text{def}}{=} \{ (\mathbf{1}, \rho)(\mathbf{2}, \rho)^*(\mathbf{3}, \rho) \mid \rho \in \mathcal{E} \} \cup \{ (\mathbf{1}, \rho)(\mathbf{2}, \rho)^\omega \mid \rho \in \mathcal{E} \}$$

$$\mathcal{T}_M \stackrel{\text{def}}{=} \emptyset$$

(Yet Another) Hierarchy of Semantics

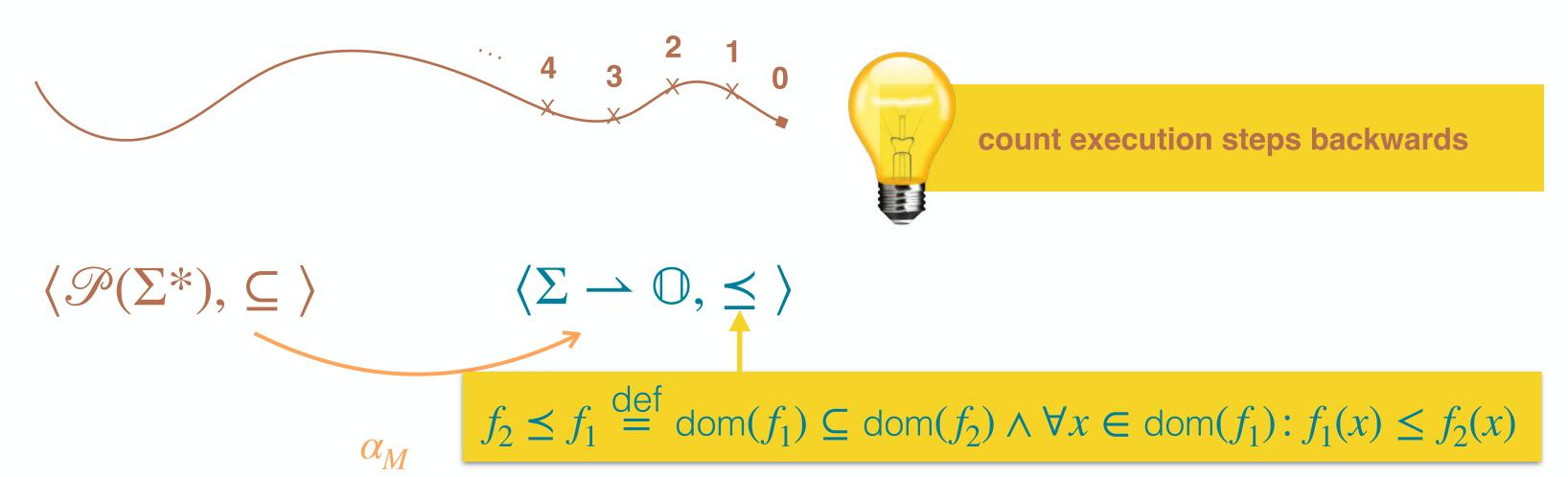


definite termination semantics

termination trace semantics

maximal trace semantics

Definite Ranking Abstraction



$$\begin{split} \overline{\alpha}_{M}(T) & \stackrel{\text{def}}{=} \overline{\alpha}_{V}(\overrightarrow{\alpha}(T)) \\ \text{where } \overline{\alpha}_{V}(\varnothing) & \stackrel{\text{def}}{=} \dot{\varnothing} \\ \overline{\alpha}_{V}(r)\sigma & \stackrel{\text{def}}{=} \begin{cases} 0 & \forall \sigma' \in \Sigma \colon (\sigma,\sigma') \not \in r \\ \sup\{\overline{\alpha}_{V}(r)\sigma' + 1 \mid \sigma' \in \text{dom}(\overline{\alpha}_{V}(r)) \land (\sigma,\sigma') \in r \} \end{cases} & \text{otherwise} \\ \overline{\alpha}(T) & \stackrel{\text{def}}{=} \{(\sigma,\sigma') \in \Sigma \times \Sigma \mid \exists t \in \Sigma^*, t' \in \Sigma^\infty \colon t\sigma\sigma't' \in T \} \end{split}$$

Least Fixpoint Formulation

$$\mathcal{R}_M \stackrel{\mathrm{def}}{=} \overline{\alpha}_M(\mathcal{T}_M) = \mathrm{lfp}^{\preceq} \overline{F}_M$$

$$\overline{F}_M(f) \sigma \stackrel{\mathrm{def}}{=} \begin{cases} 0 & \sigma \in \mathcal{B} \\ \sup\{f(\sigma') + 1 \mid (\sigma, \sigma') \in \tau\} & \sigma \in \mathrm{pre}_{\tau}(\mathrm{dom}(f)) \\ \mathrm{undefined} & \mathrm{otherwise} \end{cases}$$

Theorem

A program must terminate for traces starting from a set of initial state I if and only if $I \subseteq \text{dom}(\mathcal{R}_M)$

Denotational Formulation

We define the \mathscr{R}_M : $\Sigma \to \mathbb{O}$ by partitioning with respect to \mathscr{L} , i.e., \mathscr{R}_M : $\mathscr{L} \to (\mathscr{E} \to \mathbb{O})$.

Thus, for each program instruction stmt, we define a transformer $\mathcal{R}_M[[stmt]]: (\mathcal{E} \to \mathbb{O}) \to (\mathcal{E} \to \mathbb{O})$:

- $\mathcal{R}_M[X \leftarrow e]$
- \mathcal{R}_M [if $e \bowtie 0$ then s end]]
- \mathcal{R}_M [while $e \bowtie 0$ do s done]
- $\mathcal{R}_M[s_1; s_2]$

```
stmt ::= {}^{\ell}X \leftarrow expr^{\ell}
\mid \text{ if } {}^{\ell}expr \bowtie 0 \text{ then } stmt \text{ end}^{\ell}
\mid \text{ while } {}^{\ell}expr \bowtie 0 \text{ do } stmt \text{ done}^{\ell}
\mid stmt; stmt
```

$$\mathcal{R}_M[X \leftarrow e]$$

Example:

Let
$$\mathbb{V} = \{x\}$$
 and $f \colon \mathscr{C} \to \mathbb{O}$ defined as follows:
$$f(\rho) \stackrel{\text{def}}{=} \begin{cases} 2 & \rho(x) = 1 \\ 3 & \rho(x) = 2 \end{cases}$$
 undefined otherwise

We have

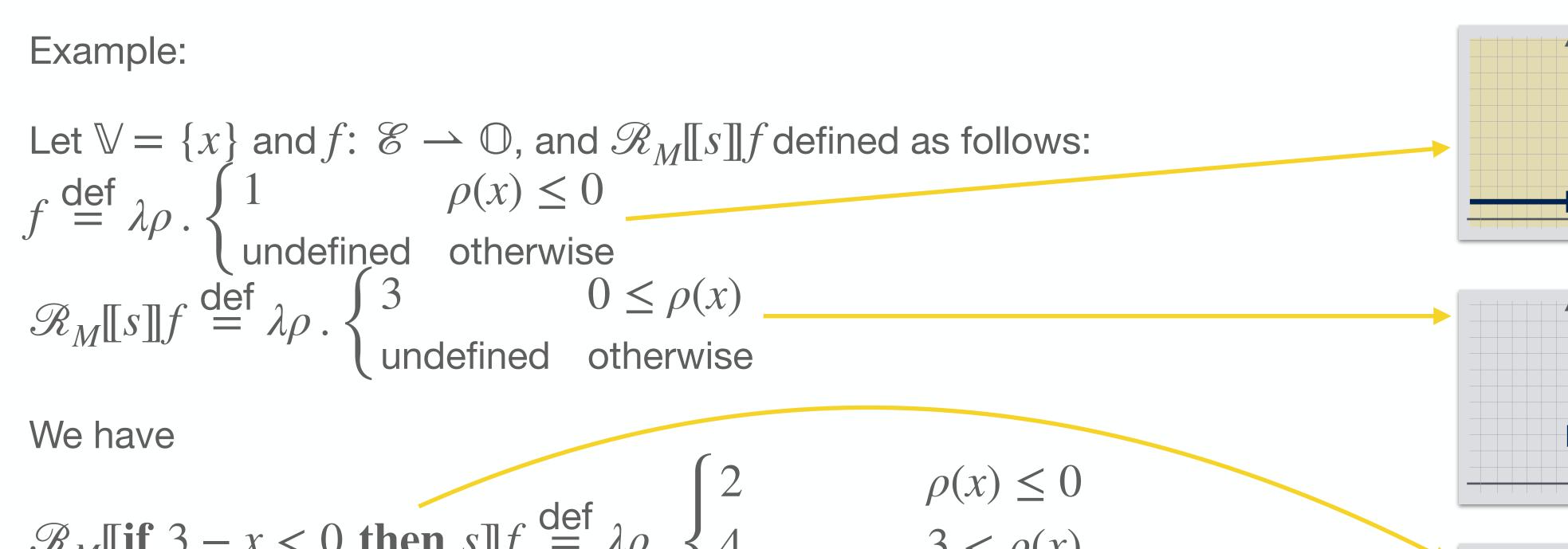
$$\mathcal{R}_{M}[[x \leftarrow x + [1,2]]]f \stackrel{\text{def}}{=} \lambda \rho . \begin{cases} 4 & \rho(x) = 0 \\ \text{undefined} & \text{otherwise} \end{cases}$$

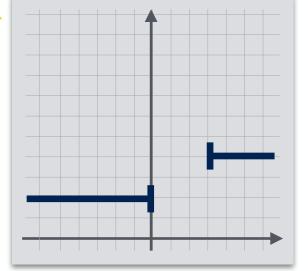
 \mathcal{R}_M [[if $e \bowtie 0$ then s end]]

$$\mathcal{R}_{M}$$
[if $e \bowtie 0$ then s end]] $f \stackrel{\text{def}}{=} \lambda \rho$. $\begin{cases} 0 \\ 0 \\ 3 \end{cases}$ undefined otherwise

- $\Im f(\rho) + 1$ $\rho \in \operatorname{dom}(f) \land \forall v \in E[[e]] \rho \colon v \bowtie 0$

 $\mathcal{R}_M[[if \ e \bowtie 0 \ then \ s \ end]]$ (continue)





 \mathcal{R}_M [while $e \bowtie 0$ do s done]]

$$\mathscr{R}_M$$
 [while $e \bowtie 0$ do s done] $f \stackrel{\text{def}}{=} \operatorname{lfp}_{\overset{\sim}{\varnothing}} \overline{F}_M$

where
$$F_M(x) \stackrel{\mathrm{def}}{=} \lambda \rho$$
 .
$$\begin{cases} 1 \\ 2 \\ 3 \end{cases}$$
 undefined otherwise

- $\textcircled{1} \quad \sup\{\mathscr{R}_M[\![s]\!]x(\rho)+1, f(\rho)+1\} \quad \rho \in \mathrm{dom}(\mathscr{R}_M[\![s]\!]x) \cap \mathrm{dom}(f) \land \mathbb{R}_M[\![s]\!]x(\rho) = 0$
 - $\rho \in \text{dom}(\mathcal{R}_M[s]x) \cap \text{dom}(f) \land \exists v_1, v_2 \in E[e]\rho \colon v_1 \bowtie 0 \land v_2 \bowtie 0$

 $\rho \in \text{dom}(\mathcal{R}_M[[s]]x) \land \\ \forall v \in E[[e]]\rho \colon v \bowtie 0$

 $\Im f(\rho) + 1$

 $\rho \in \mathrm{dom}(f) \land \forall v \in E[\![e]\!] \rho \colon v \bowtie 0$

$$\mathcal{R}_M[s_1;s_2]$$

$$\mathcal{R}_{M}[s_1; s_2]f \stackrel{\mathsf{def}}{=} \mathcal{R}_{M}[s_1](\mathcal{R}_{M}[s_2]f)$$

Denotational Formulation

Definition

The definite termination semantics $\mathcal{R}_M[[s^\ell]]: \mathcal{E} \to \mathbb{O}$ of a program s^ℓ is:

$$\mathcal{R}_{M}[\![\mathbf{s}^{\ell}]\!] \stackrel{\mathsf{def}}{=} \mathcal{R}_{M}[\![\mathbf{s}]\!](\lambda \rho.0)$$

where $\mathscr{R}_M[[s]]: (\mathscr{E} \to \mathbb{O}) \to (\mathscr{E} \to \mathbb{O})$ is the definite termination semantics of each instruction s

Theorem

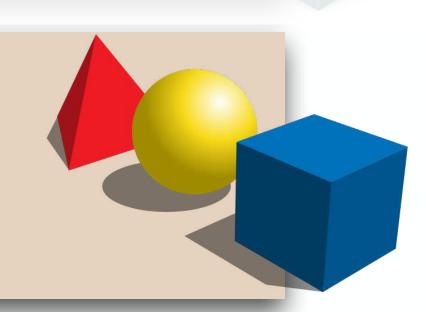
A program \mathbf{s}^ℓ must terminate starting from a set of initial states I if and only if $I \subseteq \mathrm{dom}(\mathscr{R}_M[\![\mathbf{s}^\ell]\!])$

Termination Static Analysis

Abstract Program Termination Semantics

practical tools
targeting specific program

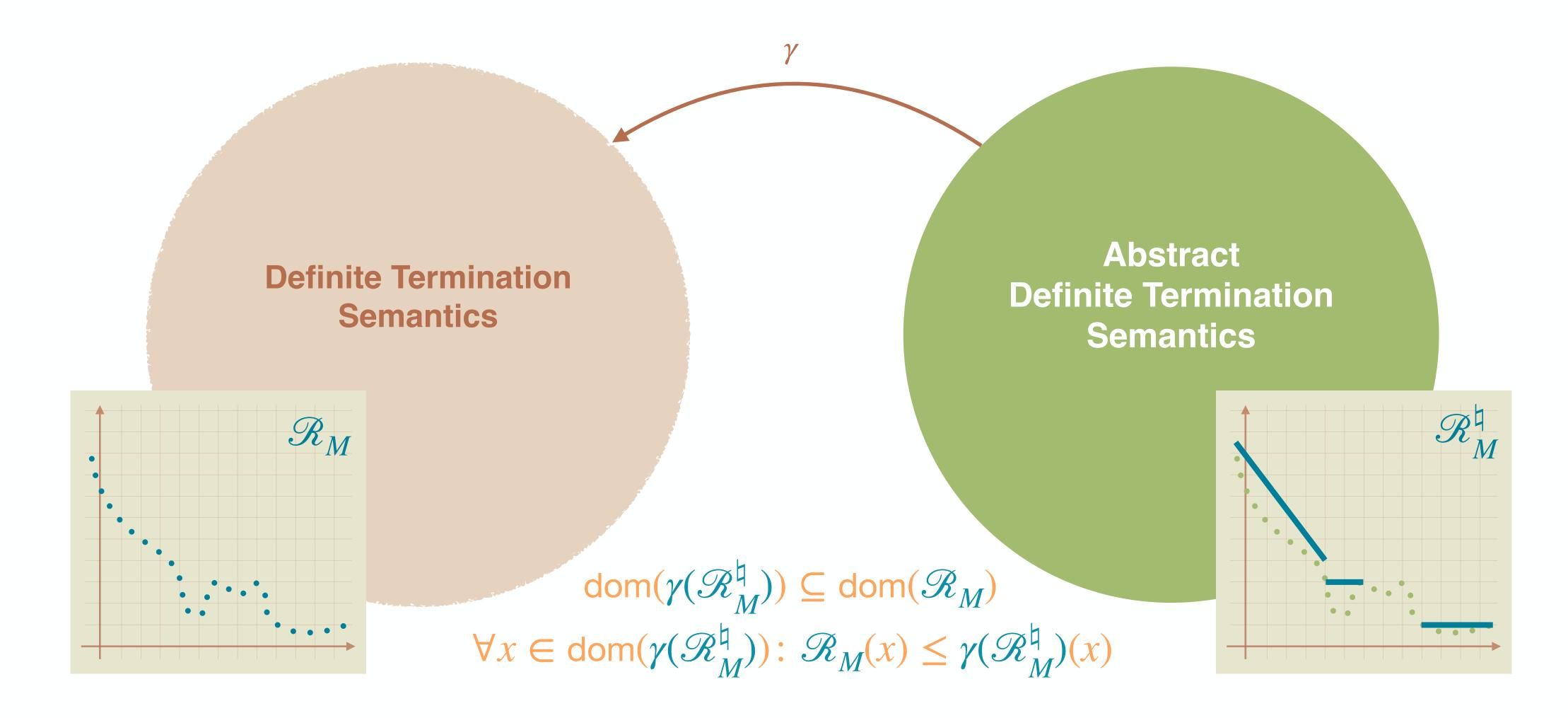
abstract semantics, abstract domains algorithmic approaches to decide program properties



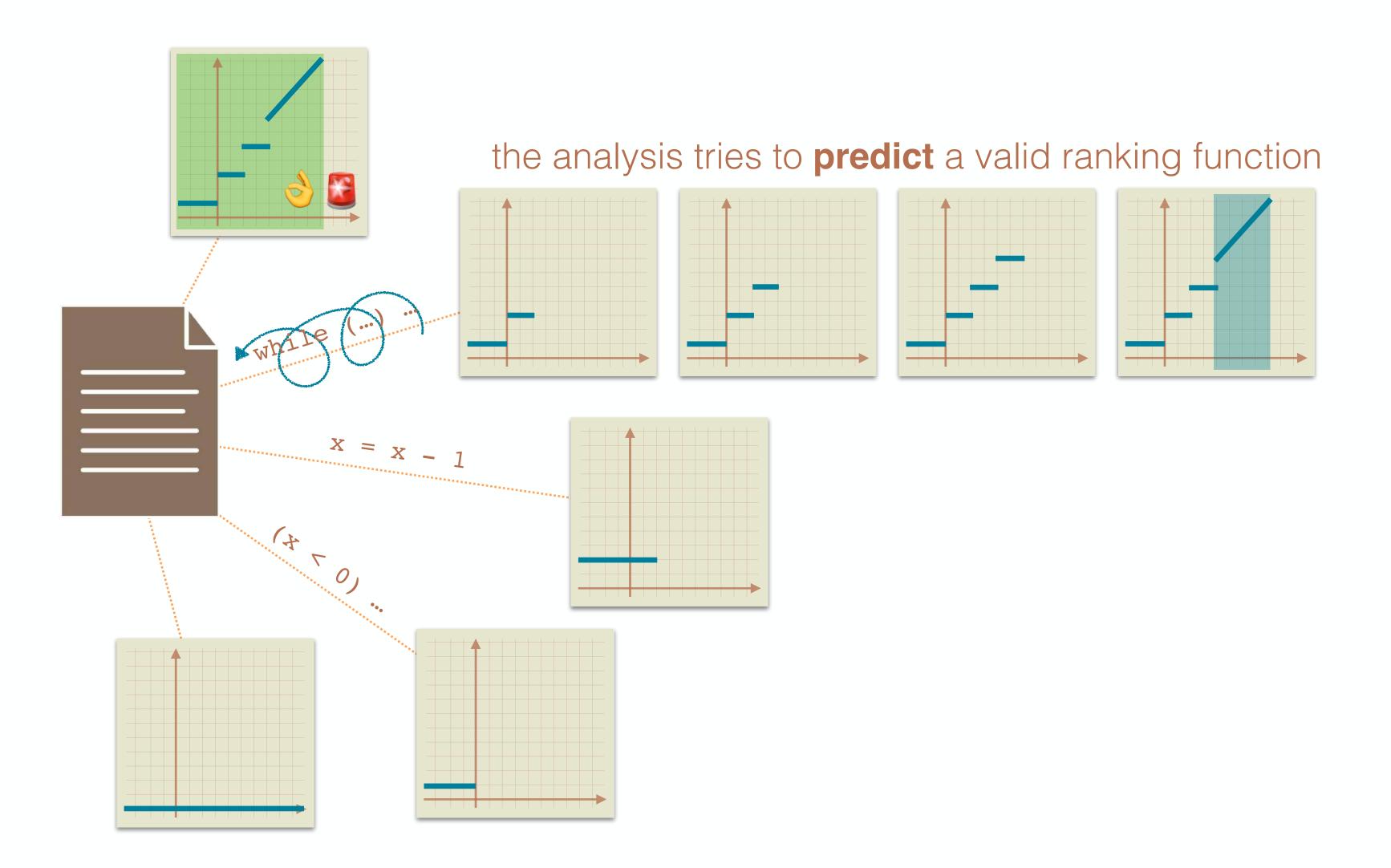
concrete semantics mathematical models of the program behavior



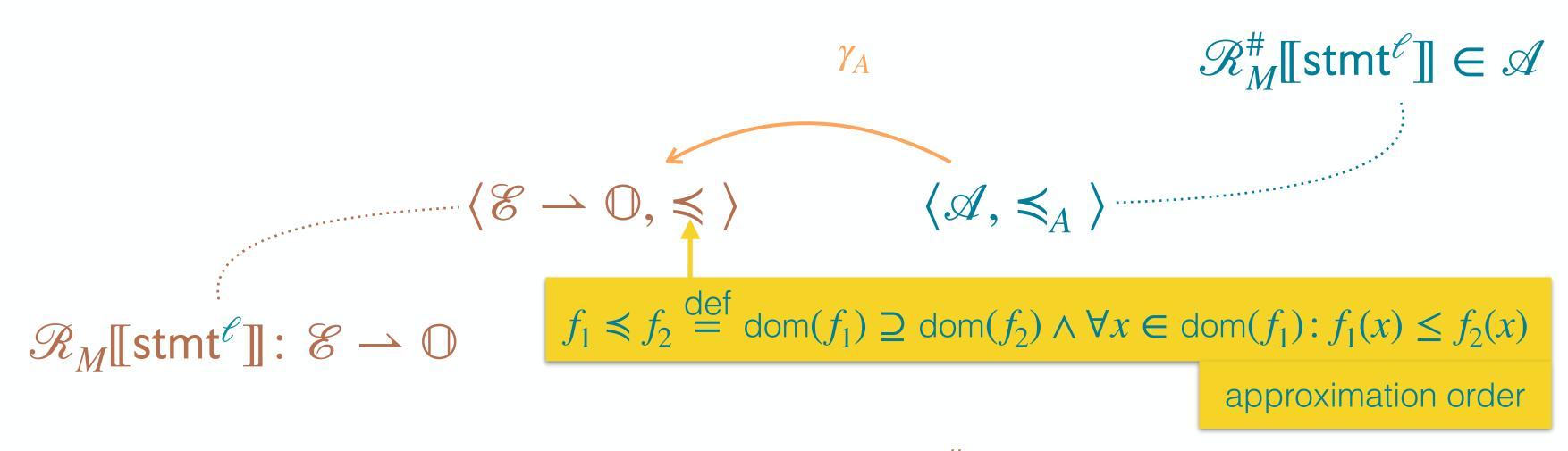
Piecewise-Defined Ranking Functions



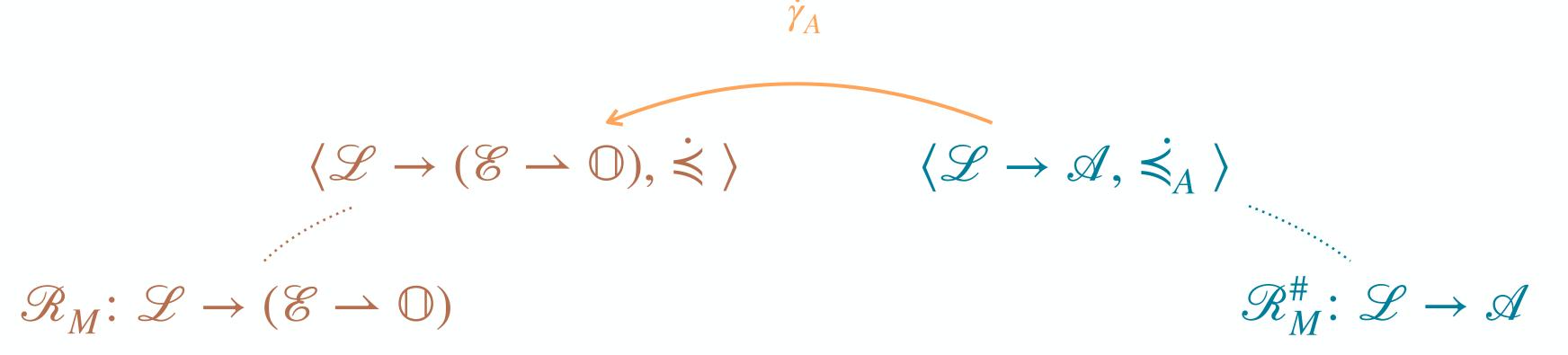
Termination Static Analysis



Piecewise-Defined Function Abstraction



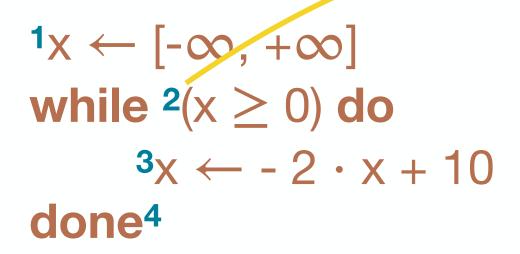
By pointwise lifiting we obtain an abstraction $\mathcal{R}_{M}^{\#}$ of \mathcal{R}_{M} :

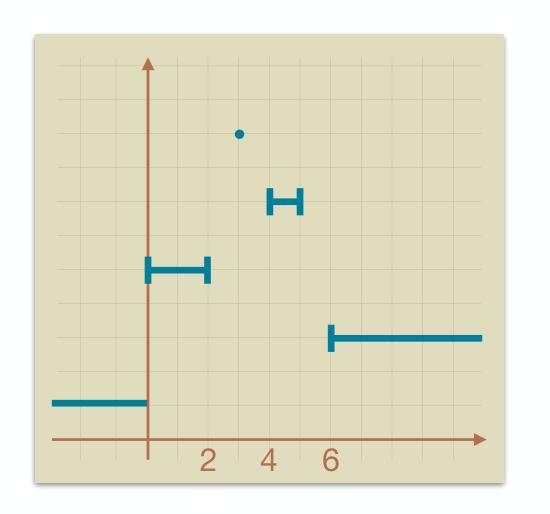


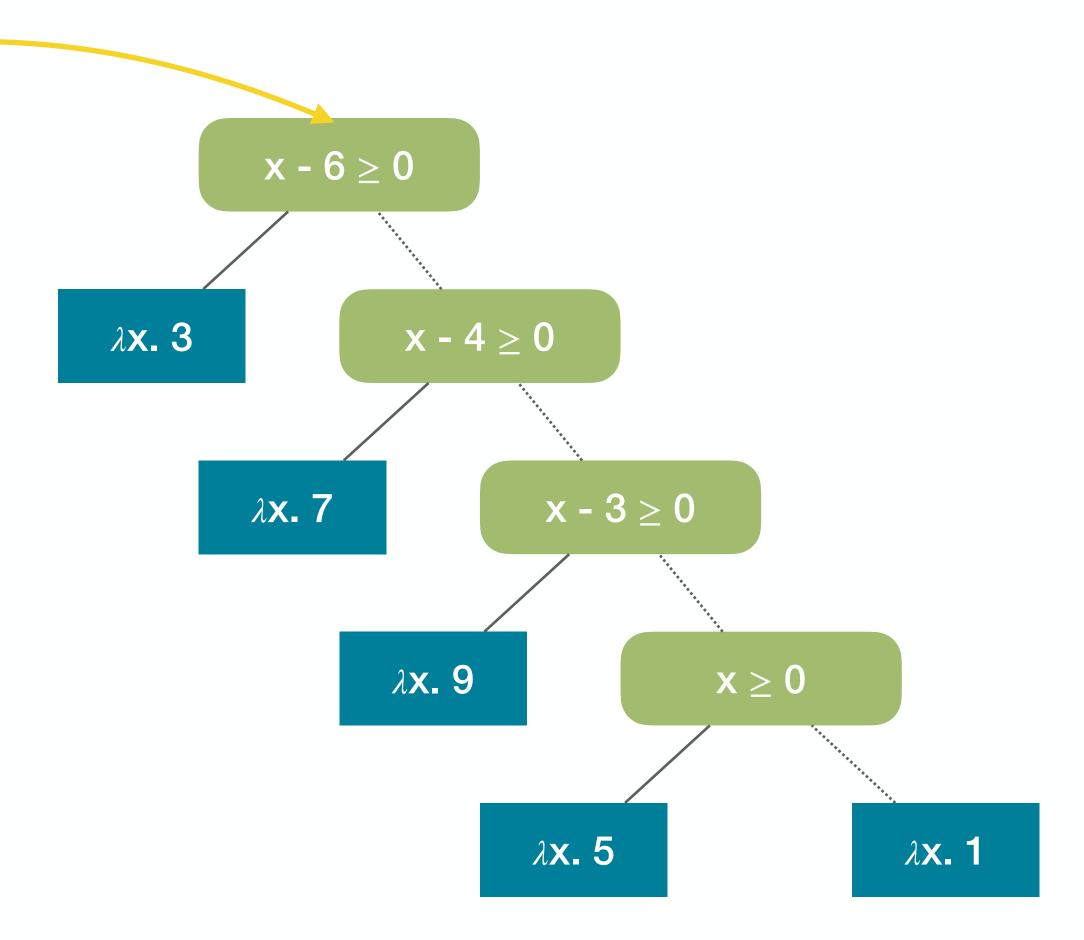
Piecewise-Defined Function Domain

 $\langle \mathcal{A}, \preccurlyeq_A \rangle$





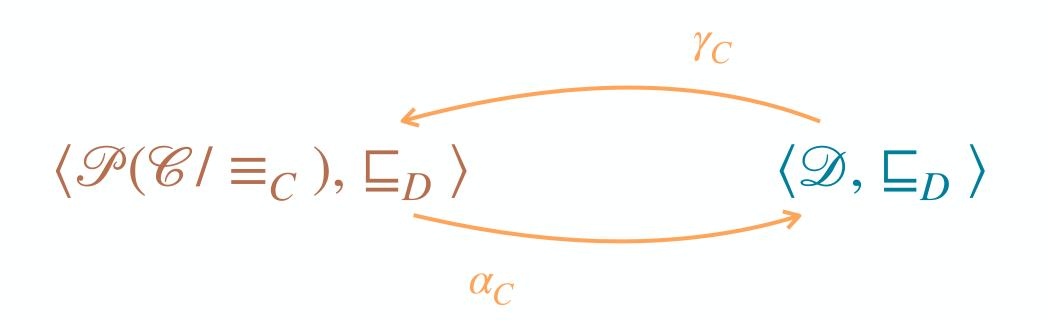




Piecewise-Defined Function Domain

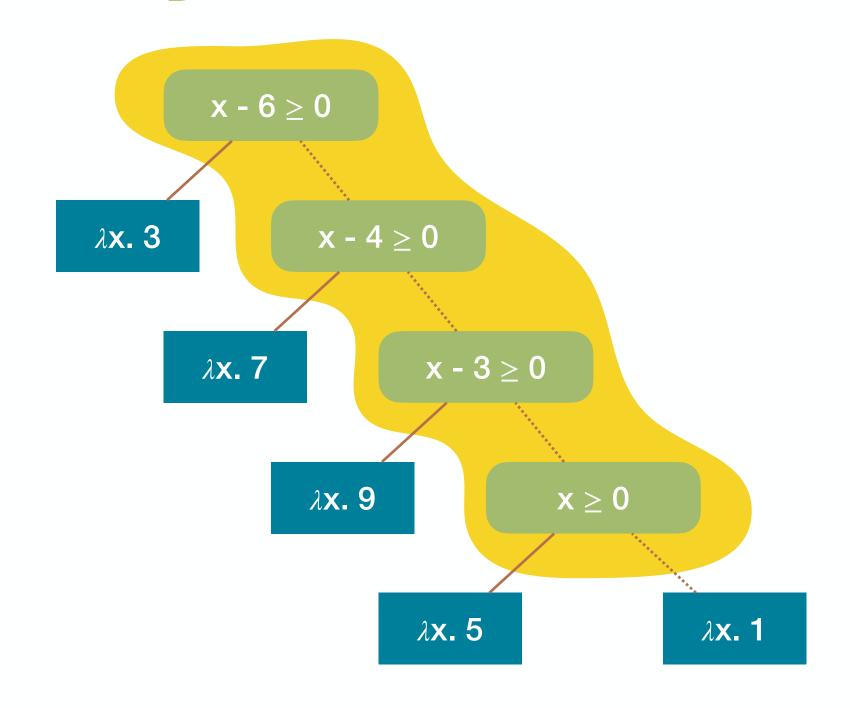
Linear Constraints Auxiliary Abstract Domain

• Parameterized by an *underlying numerical abstract domain* $\langle \mathcal{D}, \sqsubseteq_D \rangle$ (e.g., intervals, polyhedra):



Example:

$$X \to [-\infty, 3], Y \to [0, \infty] \xrightarrow{\gamma_C} \{3 - X \ge 0, Y \ge 0\}$$



• \mathscr{C} is a set of linear constraints in canonical form, equipped with a total order \leq_C :

$$\mathscr{C} \stackrel{\mathsf{def}}{=} \{c_1 \cdot X_1 + c_k \cdot X_k + c_{k+1} \geq 0 \mid X_1, \dots, X_k \in \mathbb{V} \land c_1, \dots, c_{k+1} \in \mathbb{Z} \land \gcd(|c_1|, \dots, |c_{k+1}|) = 1\}$$

Natural-Valued Ranking Functions

Piecewise-Defined Function Domain

Functions Auxiliary Abstract Domain

• Parameterized by an *underlying numerical abstract domain* $\langle \mathcal{D}, \sqsubseteq_D \rangle$

•
$$\mathscr{F} \stackrel{\mathsf{def}}{=} \{ \perp_F \} \cup (\mathbb{Z}^{|\mathbb{V}|} \to \mathbb{N}) \cup \{ \mathsf{T}_F \}$$

We consider affine functions:

$$\mathcal{F}_A \stackrel{\text{def}}{=} \{ \perp_F \} \cup \{ f \colon \mathbb{Z}^{|\mathbb{N}|} \to \mathbb{N} \mid$$

$$f(X_1, \dots, X_k) = \sum_{i=1}^k m_i \cdot X_i + q$$

$$\} \cup \{ \top_F \}$$

