

Principles of Program Analysis:

Abstract Interpretation

Transparencies based on Chapter 4 of the book: Flemming Nielson, Hanne Riis Nielson and Chris Hankin: [Principles of Program Analysis](#). Springer Verlag 2004. ©Flemming Nielson & Hanne Riis Nielson & Chris Hankin.

A Mundane Approach to Semantic Correctness

Semantics:

$$p \vdash v_1 \rightsquigarrow v_2$$

where $v_1, v_2 \in V$.

Note: \rightsquigarrow might be deterministic.

Program analysis:

$$p \vdash l_1 \triangleright l_2$$

where $l_1, l_2 \in L$.

Note: \triangleright should be deterministic:

$$f_p(l_1) = l_2.$$

What is the relationship between the semantics and the analysis?

Restrict attention to analyses where properties directly describe sets of values i.e. *“first-order” analyses* (rather than *“second-order” analyses*).

Example: Data Flow Analysis

Structural Operational Semantics:

Values: $V = \mathbf{State}$

Transitions:

$$S_{\star} \vdash \sigma_1 \rightsquigarrow \sigma_2$$

iff

$$\langle S_{\star}, \sigma_1 \rangle \rightarrow^* \sigma_2$$

Constant Propagation Analysis:

Properties: $L = \widehat{\mathbf{State}}_{\text{CP}} = (\mathbf{Var}_{\star} \rightarrow \mathbf{Z}^{\top})_{\perp}$

Transitions:

$$S_{\star} \vdash \hat{\sigma}_1 \triangleright \hat{\sigma}_2$$

iff

$$\hat{\sigma}_1 = \iota$$

$$\hat{\sigma}_2 = \sqcup \{ \text{CP}_{\bullet}(l) \mid l \in \mathit{final}(S_{\star}) \}$$

$$(\text{CP}_{\circ}, \text{CP}_{\bullet}) \models \text{CP}^{\bullet}(S_{\star})$$

Example: Control Flow Analysis

Structural Operational Semantics:

Values: $V = \mathbf{Val}$

Transitions:

$$e_{\star} \vdash v_1 \rightsquigarrow v_2$$

iff

$$[] \vdash (e_{\star} \ v_1^{l_1})^{l_2} \rightarrow^* v_2^{l_2}$$

Pure 0-CFA Analysis:

Properties: $L = \widehat{\mathbf{Env}} \times \widehat{\mathbf{Val}}$

Transitions:

$$e_{\star} \vdash (\hat{\rho}_1, \hat{v}_1) \triangleright (\hat{\rho}_2, \hat{v}_2)$$

iff

$$\hat{C}(l_1) = \hat{v}_1$$

$$\hat{C}(l_2) = \hat{v}_2$$

$$\hat{\rho}_1 = \hat{\rho}_2 = \hat{\rho}$$

$$(\hat{C}, \hat{\rho}) \models (e_{\star} \ c^{l_1})^{l_2}$$

for some place holder constant c

Correctness Relations

$$R : V \times L \rightarrow \{true, false\}$$

Idea: $v R l$ means that the value v is described by the property l .

Correctness criterion: R is preserved under computation:

$$\begin{array}{ccccc} p \vdash & v_1 & \rightsquigarrow & v_2 \\ & \vdots & & \vdots \\ & R & \Rightarrow & R \\ & \vdots & & \vdots \\ p \vdash & l_1 & \triangleright & l_2 \end{array}$$

logical relation:

$$(p \vdash \cdot \rightsquigarrow \cdot) (R \Rightarrow R) (p \vdash \cdot \triangleright \cdot)$$

Admissible Correctness Relations

$$v \mathbf{R} l_1 \wedge l_1 \sqsubseteq l_2 \Rightarrow v \mathbf{R} l_2$$

$$(\forall l \in L' \subseteq L : v \mathbf{R} l) \Rightarrow v \mathbf{R} (\bigsqcap L') \quad (\{l \mid v \mathbf{R} l\} \text{ is a Moore family})$$

Two consequences:

$$v \mathbf{R} \top$$

$$v \mathbf{R} l_1 \wedge v \mathbf{R} l_2 \Rightarrow v \mathbf{R} (l_1 \sqcap l_2)$$

Assumption: (L, \sqsubseteq) is a complete lattice.

Example: Data Flow Analysis

Correctness relation

$$R_{\text{CP}} : \text{State} \times \widehat{\text{State}}_{\text{CP}} \rightarrow \{\text{true}, \text{false}\}$$

is defined by

$$\sigma R_{\text{CP}} \hat{\sigma} \text{ iff } \forall x \in \text{FV}(S_{\star}) : (\hat{\sigma}(x) = \top \vee \sigma(x) = \hat{\sigma}(x))$$

Example: Control Flow Analysis

Correctness relation

$$R_{\text{CFA}} : \text{Val} \times (\widehat{\text{Env}} \times \widehat{\text{Val}}) \rightarrow \{true, false\}$$

is defined by

$$v R_{\text{CFA}} (\hat{\rho}, \hat{v}) \text{ iff } v \mathcal{V} (\hat{\rho}, \hat{v})$$

where \mathcal{V} is given by:

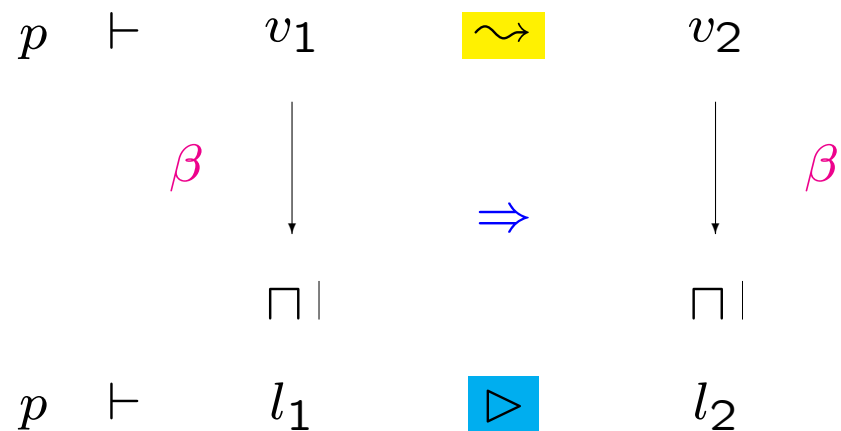
$$v \mathcal{V} (\hat{\rho}, \hat{v}) \text{ iff } \begin{cases} true & \text{if } v = c \\ t \in \hat{v} \wedge \forall x \in \text{dom}(\rho) : \rho(x) \mathcal{V} (\hat{\rho}, \hat{\rho}(x)) & \text{if } v = \text{close } t \text{ in } \rho \end{cases}$$

Representation Functions

$$\beta : V \rightarrow L$$

Idea: β maps a value to the *best* property describing it.

Correctness criterion:



Equivalence of Correctness Criteria

Given a representation function β we define a correctness relation R_β by

$$v R_\beta l \text{ iff } \beta(v) \sqsubseteq l$$

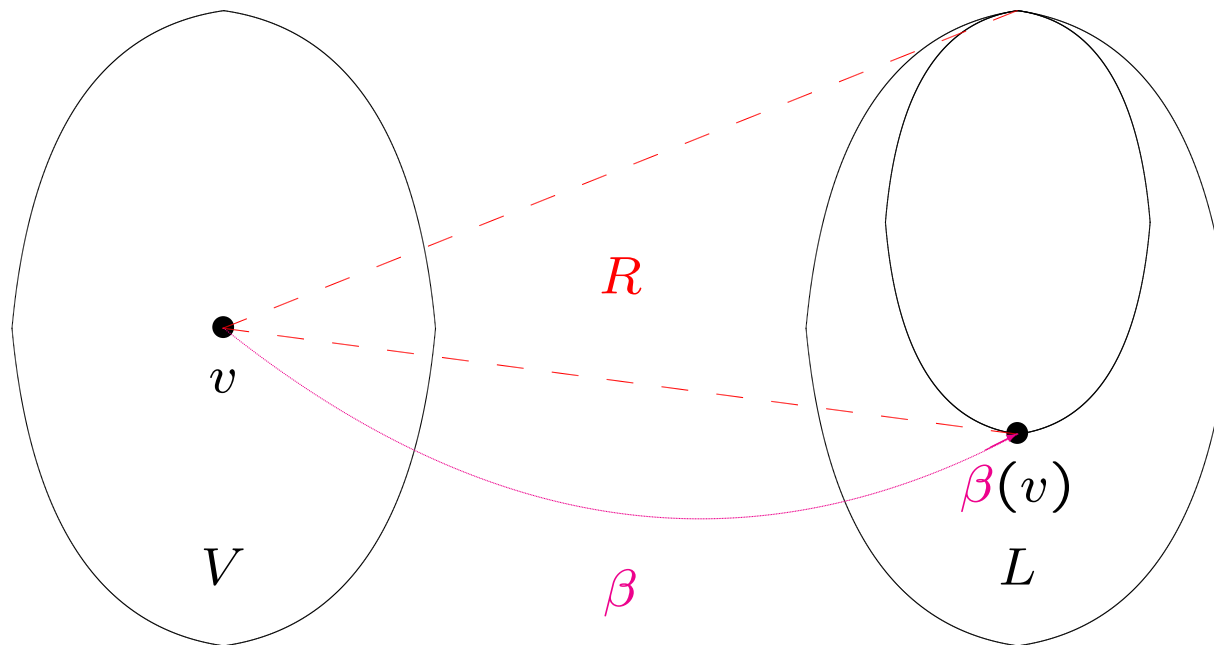
Given a correctness relation R we define a representation function β_R by

$$\beta_R(v) = \bigsqcap \{l \mid v R l\}$$

Lemma:

- (i) Given $\beta : V \rightarrow L$, then the relation $R_\beta : V \times L \rightarrow \{true, false\}$ is an admissible correctness relation such that $\beta_{R_\beta} = \beta$.
- (ii) Given an admissible correctness relation $R : V \times L \rightarrow \{true, false\}$, then β_R is well-defined and $R_{\beta_R} = R$.

Equivalence of Criteria: R is generated by β



Example: Data Flow Analysis

Representation function

$$\beta_{\text{CP}} : \text{State} \rightarrow \widehat{\text{State}}_{\text{CP}}$$

is defined by

$$\beta_{\text{CP}}(\sigma) = \lambda x. \sigma(x)$$

R_{CP} is generated by β_{CP} :

$$\sigma R_{\text{CP}} \hat{\sigma} \quad \underline{\text{iff}} \quad \beta_{\text{CP}}(\sigma) \sqsubseteq_{\text{CP}} \hat{\sigma}$$

Example: Control Flow Analysis

Representation function

$$\beta_{\text{CFA}} : \text{Val} \rightarrow \widehat{\text{Env}} \times \widehat{\text{Val}}$$

is defined by

$$\beta_{\text{CFA}}(v) = \begin{cases} (\lambda x. \emptyset, \emptyset) & \text{if } v = c \\ (\beta_{\text{CFA}}^E(\rho), \{t\}) & \text{if } v = \text{close } t \text{ in } \rho \end{cases}$$

$$\beta_{\text{CFA}}^E(\rho)(x) = \bigcup \{ \hat{\rho}_y(x) \mid \beta_{\text{CFA}}(\rho(y)) = (\hat{\rho}_y, \hat{v}_y) \text{ and } y \in \text{dom}(\rho) \} \\ \cup \begin{cases} \{ \hat{v}_x \} & \text{if } x \in \text{dom}(\rho) \text{ and } \beta_{\text{CFA}}(\rho(x)) = (\hat{\rho}_x, \hat{v}_x) \\ \emptyset & \text{otherwise} \end{cases}$$

R_{CFA} is generated by β_{CFA} :

$$v \ R_{\text{CFA}} \ (\hat{\rho}, \hat{v}) \quad \underline{\text{iff}} \quad \beta_{\text{CFA}}(v) \sqsubseteq_{\text{CFA}} (\hat{\rho}, \hat{v})$$

A Modest Generalisation

Semantics:

$$p \vdash v_1 \overset{\text{yellow}}{\rightsquigarrow} v_2$$

where $v_1 \in V_1, v_2 \in V_2$

Program analysis:

$$p \vdash l_1 \overset{\text{blue}}{\triangleright} l_2$$

where $l_1 \in L_1, l_2 \in L_2$

$$\begin{array}{cccc}
 p \vdash & v_1 & \overset{\text{yellow}}{\rightsquigarrow} & v_2 \\
 & \vdots & & \vdots \\
 & R_1 & \Rightarrow & R_2 \\
 & \vdots & & \vdots \\
 p \vdash & l_1 & \overset{\text{blue}}{\triangleright} & l_2
 \end{array}$$

logical relation:

$$(p \vdash \cdot \overset{\text{yellow}}{\rightsquigarrow} \cdot) (R_1 \Rightarrow R_2) (p \vdash \cdot \overset{\text{blue}}{\triangleright} \cdot)$$

Higher-Order Formulation

Assume that

- R_1 is an admissible correctness relation for V_1 and L_1 that is *generated by* the representation function $\beta_1 : V_1 \rightarrow L_1$
- R_2 is an admissible correctness relation for V_2 and L_2 that is *generated by* the representation function $\beta_2 : V_2 \rightarrow L_2$

Then the relation $R_1 \twoheadrightarrow R_2$ is an admissible correctness relation for $V_1 \rightarrow V_2$ and $L_1 \rightarrow L_2$

that is *generated by* the representation function $\beta_1 \twoheadrightarrow \beta_2$ defined by

$$(\beta_1 \twoheadrightarrow \beta_2)(\rightsquigarrow) = \lambda l_1. \bigsqcup \{ \beta_2(v_2) \mid \beta_1(v_1) \sqsubseteq l_1 \wedge v_1 \rightsquigarrow v_2 \}$$

Example:

Semantics:

$\text{plus} \vdash (z_1, z_2) \rightsquigarrow z_1 + z_2$

where $z_1, z_2 \in \mathbf{Z}$

Program analysis:

$\text{plus} \vdash ZZ \triangleright \{z_1 + z_2 \mid (z_1, z_2) \in ZZ\}$

where $ZZ \subseteq \mathbf{Z} \times \mathbf{Z}$

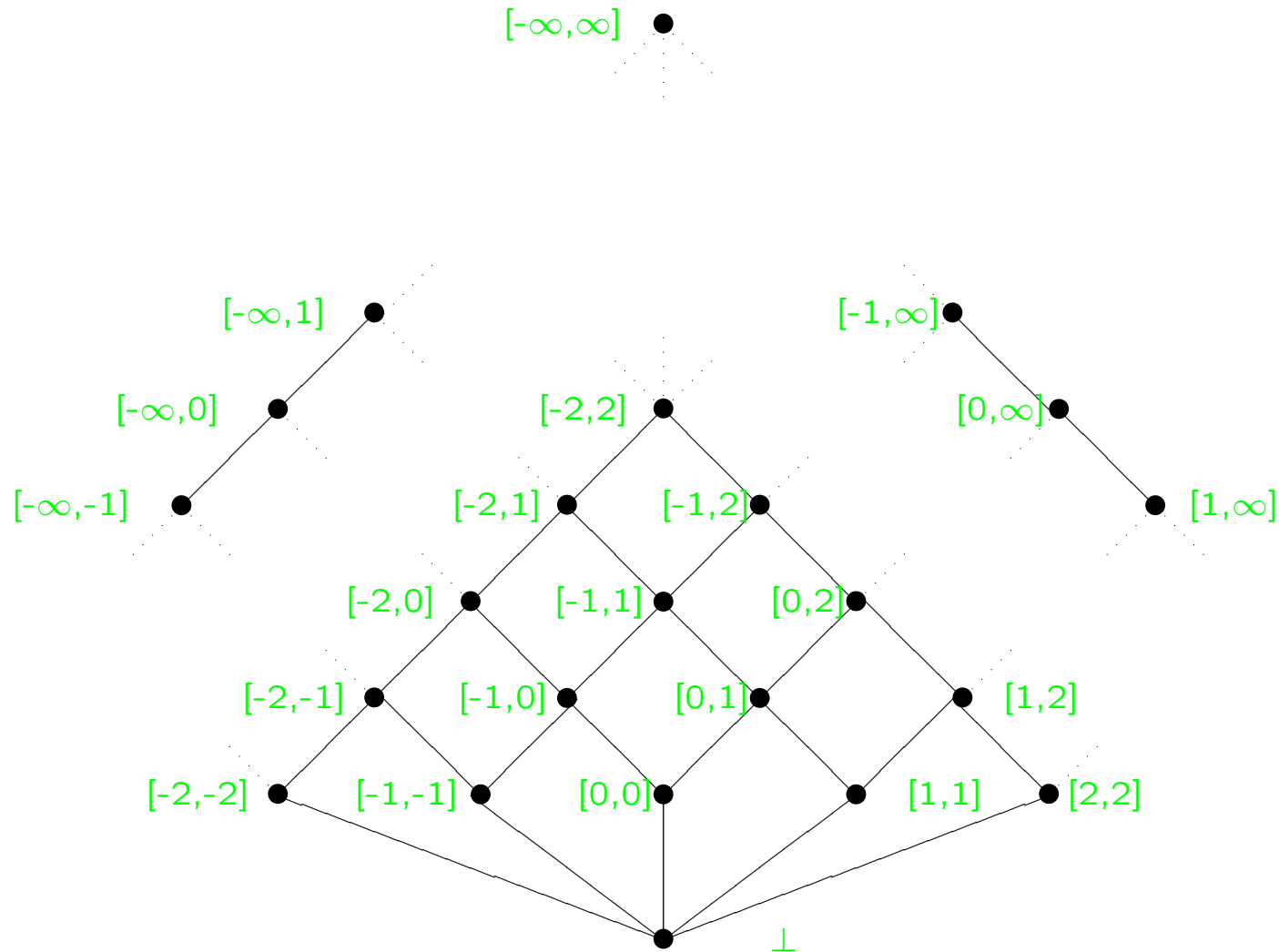
	Correctness relations	Representation functions
result	R_Z	$\beta_Z(z) = \{z\}$
argument	$R_{Z \times Z}$	$\beta_{Z \times Z}(z_1, z_2) = \{(z_1, z_2)\}$
plus	$(\text{plus} \vdash \cdot \rightsquigarrow \cdot)$ $(R_{Z \times Z} \twoheadrightarrow R_Z)$ $(\text{plus} \vdash \cdot \triangleright \cdot)$	$(\beta_{Z \times Z} \twoheadrightarrow \beta_Z)(\text{plus} \vdash \cdot \rightsquigarrow \cdot)$ $\sqsubseteq (\text{plus} \vdash \cdot \triangleright \cdot)$

Approximation of Fixed Points

- Fixed points
- Widening
- Narrowing

Example: lattice of intervals for *Array Bound Analysis*

The complete lattice $\mathbf{Interval} = (\mathbf{Interval}, \sqsubseteq)$



Fixed points

Let $f : L \rightarrow L$ be a *monotone function* on a complete lattice $L = (L, \sqsubseteq, \sqcup, \sqcap, \perp, \top)$.

l is a *fixed point* iff $f(l) = l$ $Fix(f) = \{l \mid f(l) = l\}$

f is *reductive* at l iff $f(l) \sqsubseteq l$ $Red(f) = \{l \mid f(l) \sqsubseteq l\}$

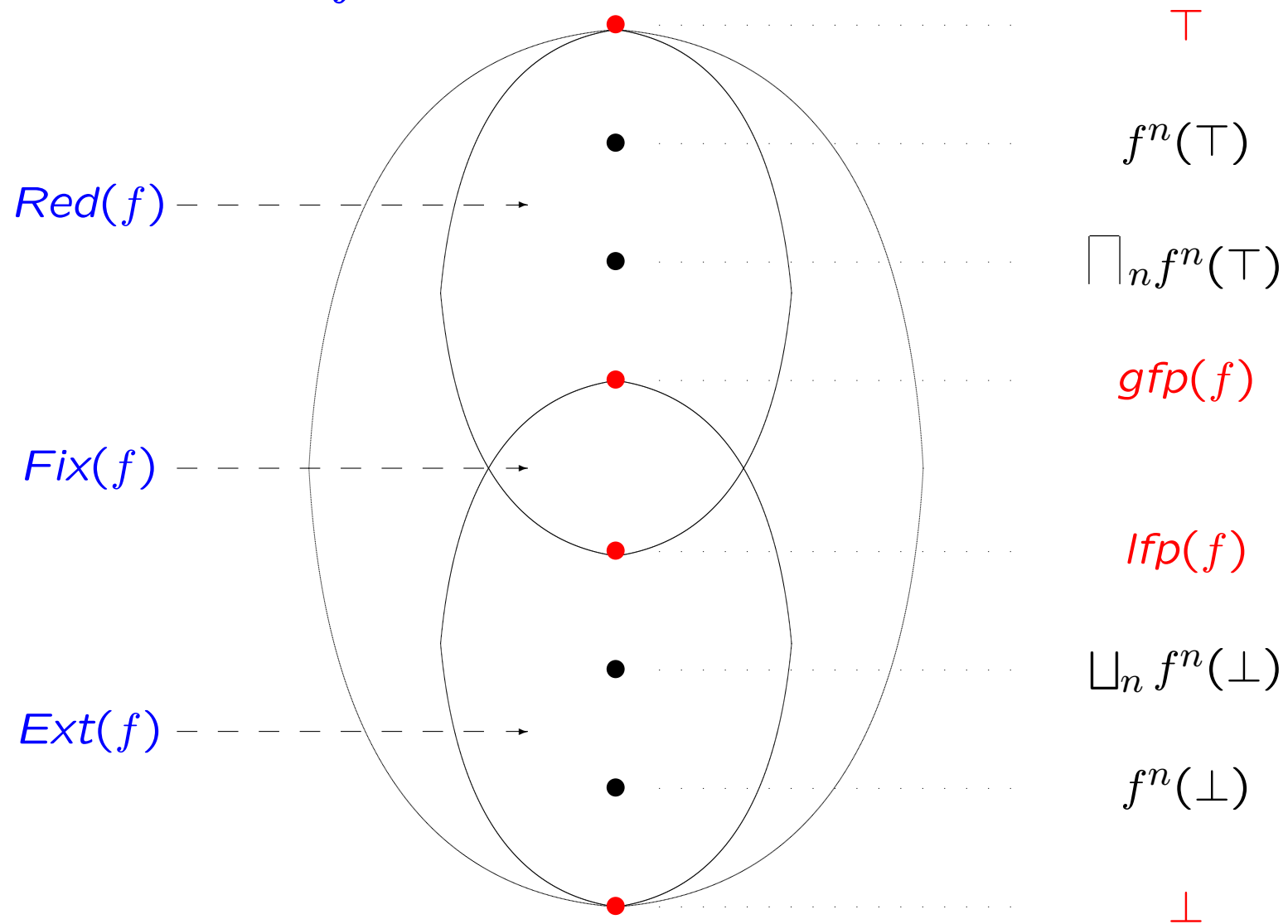
f is *extensive* at l iff $f(l) \sqsupseteq l$ $Ext(f) = \{l \mid f(l) \sqsupseteq l\}$

Tarski's Theorem ensures that

$$lfp(f) = \sqcap Fix(f) = \sqcap Red(f) \in Fix(f) \subseteq Red(f)$$

$$gfp(f) = \sqcup Fix(f) = \sqcup Ext(f) \in Fix(f) \subseteq Ext(f)$$

Fixed points of f



Widening Operators

Problem: We cannot guarantee that $(f^n(\perp))_n$ eventually stabilises nor that its least upper bound necessarily equals $\text{lfp}(f)$.

Idea: We replace $(f^n(\perp))_n$ by a new sequence $(f_{\nabla}^n)_n$ that is known to eventually stabilise and to do so with a value that is a safe (upper) approximation of the least fixed point.

The new sequence is parameterised on the widening operator ∇ : an upper bound operator satisfying a finiteness condition.

Upper bound operators

$\checkmark : L \times L \rightarrow L$ is an *upper bound operator* iff

$$l_1 \sqsubseteq l_1 \checkmark l_2 \sqsupseteq l_2$$

for all $l_1, l_2 \in L$.

Let $(l_n)_n$ be a sequence of elements of L . Define the sequence $(l_n^{\checkmark})_n$ by:

$$l_n^{\checkmark} = \begin{cases} l_n & \text{if } n = 0 \\ l_{n-1}^{\checkmark} \checkmark l_n & \text{if } n > 0 \end{cases}$$

Fact: If $(l_n)_n$ is a sequence and \checkmark is an upper bound operator then $(l_n^{\checkmark})_n$ is an ascending chain; furthermore $l_n^{\checkmark} \sqsupseteq \sqcup \{l_0, l_1, \dots, l_n\}$ for all n .

Example:

Let int be an arbitrary but fixed element of **Interval**.

An upper bound operator:

$$int_1 \sqcup^{int} int_2 = \begin{cases} int_1 \sqcup int_2 & \text{if } int_1 \sqsubseteq int \vee int_2 \sqsubseteq int_1 \\ [-\infty, \infty] & \text{otherwise} \end{cases}$$

Example: $[1, 2] \sqcup^{[0,2]} [2, 3] = [1, 3]$ and $[2, 3] \sqcup^{[0,2]} [1, 2] = [-\infty, \infty]$.

Transformation of: $[0, 0], [1, 1], [2, 2], [3, 3], [4, 4], [5, 5], \dots$

If $int = [0, \infty]$: $[0, 0], [0, 1], [0, 2], [0, 3], [0, 4], [0, 5], \dots$

If $int = [0, 2]$: $[0, 0], [0, 1], [0, 2], [0, 3], [-\infty, \infty], [-\infty, \infty], \dots$

Widening operators

An operator $\nabla : L \times L \rightarrow L$ is a *widening operator* iff

- it is an upper bound operator, and
- for all ascending chains $(l_n)_n$ the ascending chain $(l_n^\nabla)_n$ eventually stabilises.

Widening operators

Given a monotone function $f : L \rightarrow L$ and a widening operator ∇ define the sequence $(f_{\nabla}^n)_n$ by

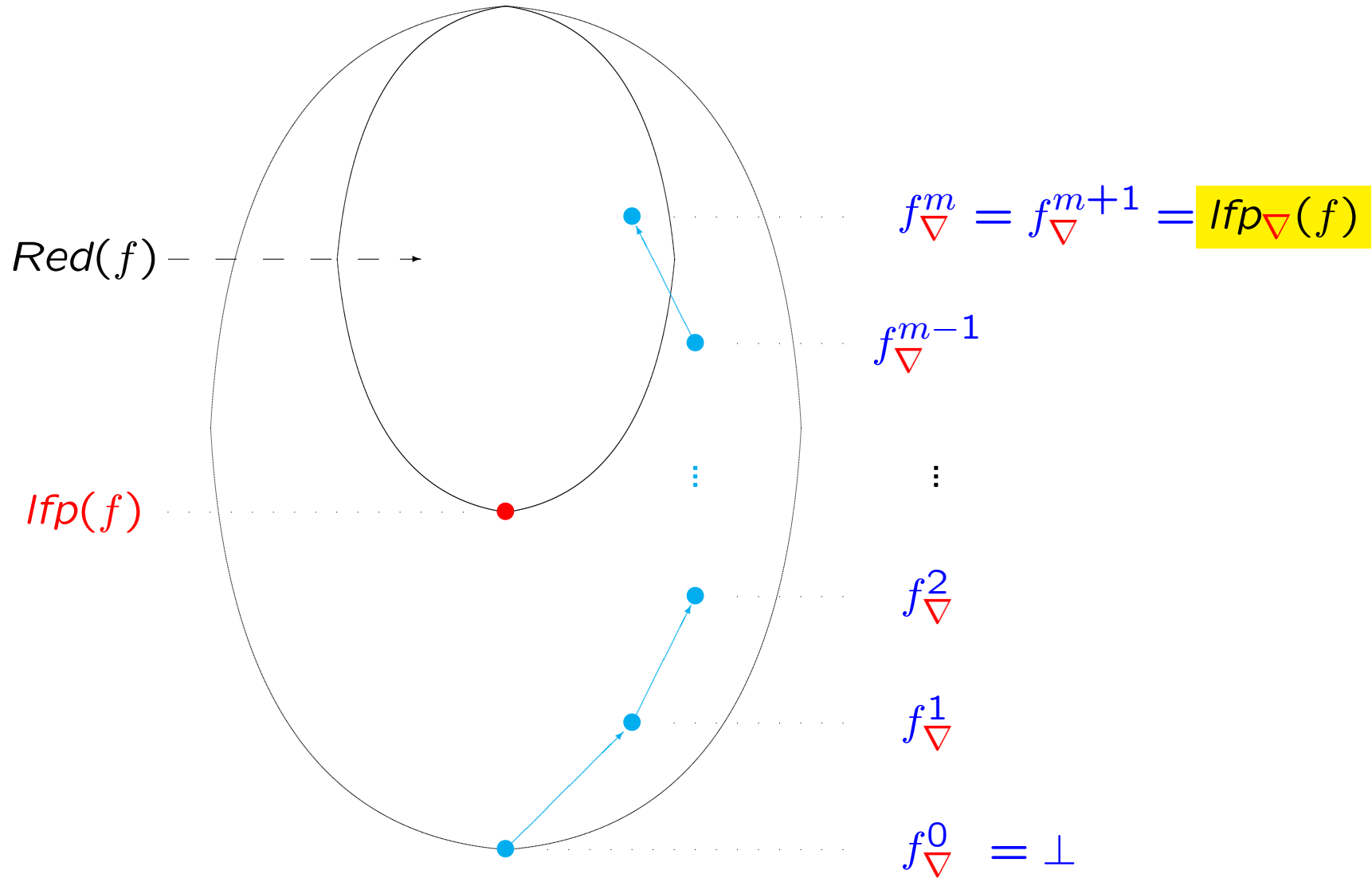
$$f_{\nabla}^n = \begin{cases} \perp & \text{if } n = 0 \\ f_{\nabla}^{n-1} & \text{if } n > 0 \wedge f(f_{\nabla}^{n-1}) \sqsubseteq f_{\nabla}^{n-1} \\ f_{\nabla}^{n-1} \nabla f(f_{\nabla}^{n-1}) & \text{otherwise} \end{cases}$$

One can show that:

- $(f_{\nabla}^n)_n$ is an ascending chain that eventually stabilises
- it happens when $f(f_{\nabla}^m) \sqsubseteq f_{\nabla}^m$ for some value of m
- Tarski's Theorem then gives $f_{\nabla}^m \sqsupseteq \text{lfp}(f)$

$$\text{lfp}_{\nabla}(f) = f_{\nabla}^m$$

The widening operator ∇ applied to f



Example:

Let K be a *finite* set of integers, e.g. the set of integers explicitly mentioned in a given program.

We shall define a widening operator ∇ based on K .

Idea: $[z_1, z_2] \nabla [z_3, z_4]$ is

$$[\text{LB}(z_1, z_3) , \text{UB}(z_2, z_4)]$$

where

- $\text{LB}(z_1, z_3) \in \{z_1\} \cup K \cup \{-\infty\}$ is the best possible lower bound, and
- $\text{UB}(z_2, z_4) \in \{z_2\} \cup K \cup \{\infty\}$ is the best possible upper bound.

The effect: a change in any of the bounds of the interval $[z_1, z_2]$ can only take place finitely many times – corresponding to the cardinality of K .

Example (cont.) — formalisation:

Let $z_i \in \mathbf{Z}' = \mathbf{Z} \cup \{-\infty, \infty\}$ and write:

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

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

$$int_1 \nabla int_2 = \begin{cases} \perp & \text{if } int_1 = int_2 = \perp \\ [\text{LB}_K(\text{inf}(int_1), \text{inf}(int_2)) , \text{UB}_K(\text{sup}(int_1), \text{sup}(int_2))] & \text{otherwise} \end{cases}$$

Example (cont.):

Consider the ascending chain $(int_n)_n$

$$[0, 1], [0, 2], [0, 3], [0, 4], [0, 5], [0, 6], [0, 7], \dots$$

and assume that $K = \{3, 5\}$.

Then $(int_n^\nabla)_n$ is the chain

$$[0, 1], [0, 3], [0, 3], [0, 5], [0, 5], [0, \infty], [0, \infty], \dots$$

which eventually stabilises.

Narrowing Operators

Status: Widening gives us an upper approximation $lfp_{\nabla}(f)$ of the least fixed point of f .

Observation: $f(lfp_{\nabla}(f)) \sqsubseteq lfp_{\nabla}(f)$ so the approximation can be improved by considering the iterative sequence $(f^n(lfp_{\nabla}(f)))_n$.

It will satisfy $f^n(lfp_{\nabla}(f)) \sqsupseteq lfp(f)$ for all n so we can stop at an arbitrary point.

The notion of **narrowing** is *one way* of encapsulating a termination criterion for the sequence.

Narrowing

An operator $\Delta : L \times L \rightarrow L$ is a *narrowing operator* iff

- $l_2 \sqsubseteq l_1 \Rightarrow l_2 \sqsubseteq (l_1 \Delta l_2) \sqsubseteq l_1$ for all $l_1, l_2 \in L$, and
- for all descending chains $(l_n)_n$ the sequence $(l_n^\Delta)_n$ eventually stabilises.

Recall: The sequence $(l_n^\Delta)_n$ is defined by:

$$l_n^\Delta = \begin{cases} l_n & \text{if } n = 0 \\ l_{n-1}^\Delta \Delta l_n & \text{if } n > 0 \end{cases}$$

Narrowing

We construct the sequence $([f]_{\Delta}^n)_n$

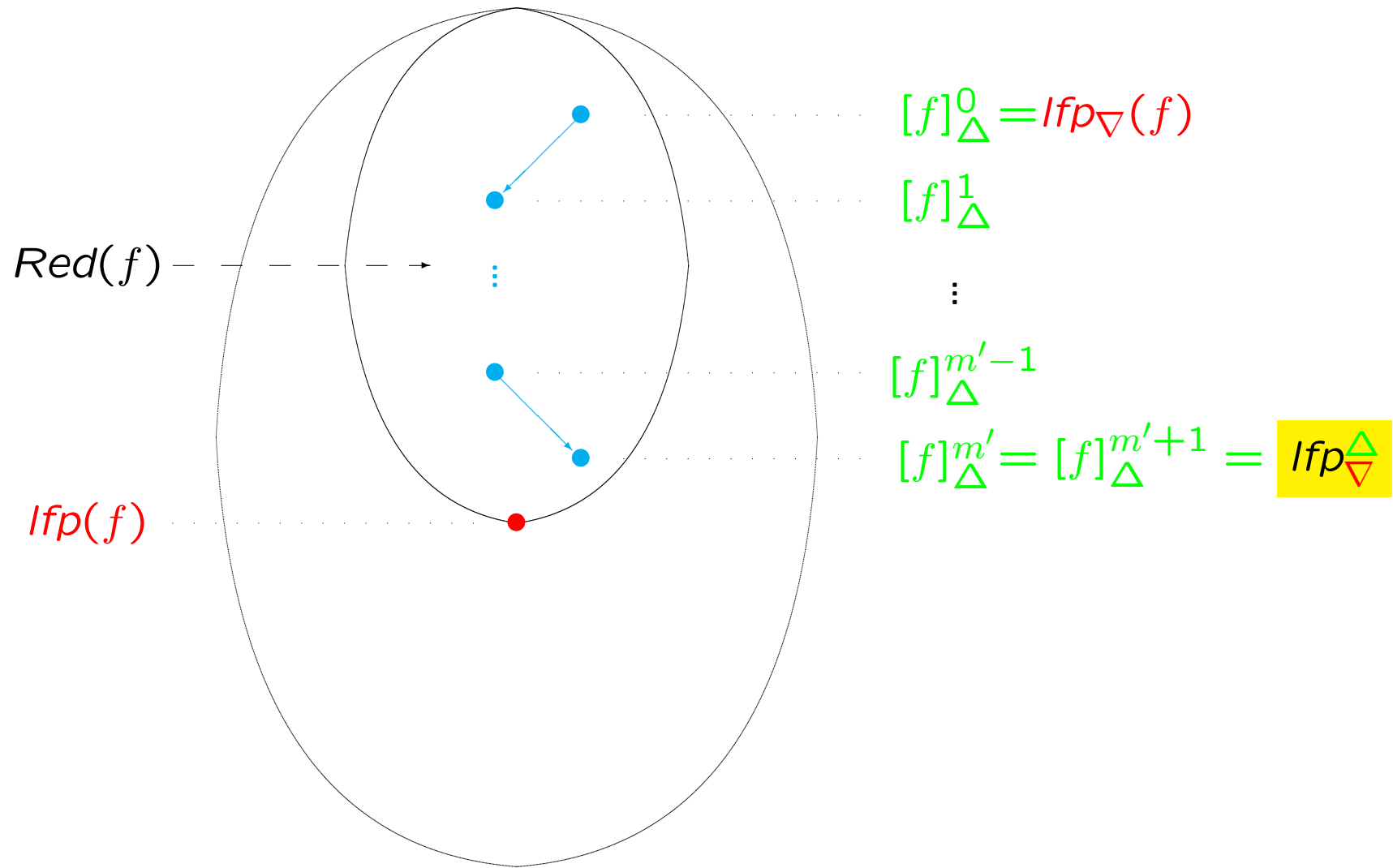
$$[f]_{\Delta}^n = \begin{cases} \text{Ifp}_{\nabla}(f) & \text{if } n = 0 \\ [f]_{\Delta}^{n-1} \triangle f([f]_{\Delta}^{n-1}) & \text{if } n > 0 \end{cases}$$

One can show that:

- $([f]_{\Delta}^n)_n$ is a descending chain where all elements satisfy $\text{Ifp}(f) \sqsubseteq [f]_{\Delta}^n$
- the chain eventually stabilises so $[f]_{\Delta}^{m'} = [f]_{\Delta}^{m'+1}$ for some value m'

$$\text{Ifp}_{\nabla}^{\Delta}(f) = [f]_{\Delta}^{m'}$$

The narrowing operator Δ applied to f



Example:

The complete lattice (**Interval**, \sqsubseteq) has two kinds of infinite descending chains:

- those with elements of the form $[-\infty, z]$, $z \in \mathbf{Z}$
- those with elements of the form $[z, \infty]$, $z \in \mathbf{Z}$

Idea: Given some fixed non-negative number N the narrowing operator Δ_N will force an infinite descending chain

$$[z_1, \infty], [z_2, \infty], [z_3, \infty], \dots$$

(where $z_1 < z_2 < z_3 < \dots$) to stabilise when $z_i > N$

Similarly, for a descending chain with elements of the form $[-\infty, z_i]$ the narrowing operator will force it to stabilise when $z_i < -N$

Example (cont.) — formalisation:

Define $\Delta = \Delta_N$ by

$$int_1 \Delta int_2 = \begin{cases} \perp & \text{if } int_1 = \perp \vee int_2 = \perp \\ [z_1, z_2] & \text{otherwise} \end{cases}$$

where

$$z_1 = \begin{cases} \inf(int_1) & \text{if } N < \inf(int_2) \wedge \sup(int_2) = \infty \\ \inf(int_2) & \text{otherwise} \end{cases}$$

$$z_2 = \begin{cases} \sup(int_1) & \text{if } \inf(int_2) = -\infty \wedge \sup(int_2) < -N \\ \sup(int_2) & \text{otherwise} \end{cases}$$

Example (cont.):

Consider the infinite descending chain $([n, \infty])_n$

$$[0, \infty], [1, \infty], [2, \infty], [3, \infty], [4, \infty], [5, \infty], \dots$$

and assume that $N = 3$.

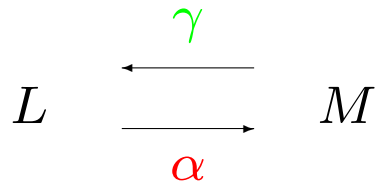
Then the narrowing operator Δ_N will give the sequence $([n, \infty]^{\Delta})_n$

$$[0, \infty], [1, \infty], [2, \infty], [3, \infty], [3, \infty], [3, \infty], \dots$$

Galois Connections

- Galois connections and adjunctions
- Extraction functions
- Galois insertions
- Reduction operators

Galois connections



α : *abstraction function*

γ : *concretisation function*

is a **Galois connection** if and only if

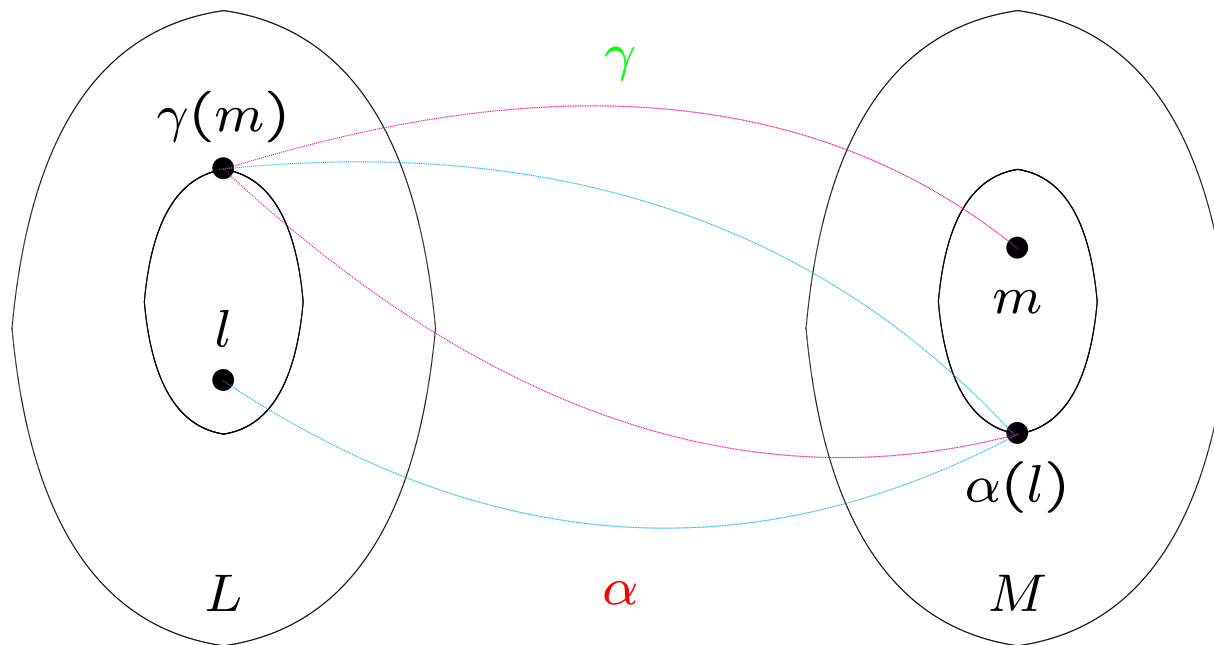
α and γ are monotone functions

that satisfy

$$\gamma \circ \alpha \sqsupseteq \lambda l.l$$

$$\alpha \circ \gamma \sqsubseteq \lambda m.m$$

Galois connections



$$\gamma \circ \alpha \sqsupseteq \lambda l.l$$

$$\alpha \circ \gamma \sqsubseteq \lambda m.m$$

Example:

Galois connection

$$(\mathcal{P}(\mathbf{Z}), \alpha_{\mathbf{ZI}}, \gamma_{\mathbf{ZI}}, \mathbf{Interval})$$

with concretisation function

$$\gamma_{\mathbf{ZI}}(int) = \{z \in \mathbf{Z} \mid \inf(int) \leq z \leq \sup(int)\}$$

and abstraction function

$$\alpha_{\mathbf{ZI}}(Z) = \begin{cases} \perp & \text{if } Z = \emptyset \\ [\inf'(Z), \sup'(Z)] & \text{otherwise} \end{cases}$$

Examples:

$$\begin{aligned} \gamma_{\mathbf{ZI}}([0, 3]) &= \{0, 1, 2, 3\} \\ \gamma_{\mathbf{ZI}}([0, \infty]) &= \{z \in \mathbf{Z} \mid z \geq 0\} \\ \alpha_{\mathbf{ZI}}(\{0, 1, 3\}) &= [0, 3] \\ \alpha_{\mathbf{ZI}}(\{2 * z \mid z > 0\}) &= [2, \infty] \end{aligned}$$

Adjunctions

$$L \begin{array}{c} \xleftarrow{\gamma} \\ \xrightarrow{\alpha} \end{array} M$$

is an *adjunction* if and only if

$\alpha : L \rightarrow M$ and $\gamma : M \rightarrow L$ are total functions

that satisfy

$$\alpha(l) \sqsubseteq m \quad \text{iff} \quad l \sqsubseteq \gamma(m)$$

for all $l \in L$ and $m \in M$.

Proposition: (α, γ) is an adjunction iff it is a Galois connection.

Galois connections from representation functions

A representation function $\beta : V \rightarrow L$ gives rise to a Galois connection

$$(\mathcal{P}(V), \alpha, \gamma, L)$$

where

$$\alpha(V') = \sqcup\{\beta(v) \mid v \in V'\}$$

$$\gamma(l) = \{v \in V \mid \beta(v) \sqsubseteq l\}$$

for $V' \subseteq V$ and $l \in L$.

This indeed defines an adjunction:

$$\begin{aligned} \alpha(V') \sqsubseteq l &\Leftrightarrow \sqcup\{\beta(v) \mid v \in V'\} \sqsubseteq l \\ &\Leftrightarrow \forall v \in V' : \beta(v) \sqsubseteq l \\ &\Leftrightarrow V' \subseteq \gamma(l) \end{aligned}$$

Galois connections from extraction functions

An *extraction function*

$$\eta : V \rightarrow D$$

maps the values of V to their best descriptions in D .

It gives rise to a representation function $\beta_\eta : V \rightarrow \mathcal{P}(D)$ (corresponding to $L = (\mathcal{P}(D), \subseteq)$) defined by

$$\beta_\eta(v) = \{\eta(v)\}$$

The associated Galois connection is

$$(\mathcal{P}(V), \alpha_\eta, \gamma_\eta, \mathcal{P}(D))$$

where

$$\alpha_\eta(V') = \cup\{\beta_\eta(v) \mid v \in V'\} = \{\eta(v) \mid v \in V'\}$$

$$\gamma_\eta(D') = \{v \in V \mid \beta_\eta(v) \subseteq D'\} = \{v \mid \eta(v) \in D'\}$$

Example:

Extraction function

$$\text{sign} : \mathbf{Z} \rightarrow \mathbf{Sign}$$

specified by

$$\text{sign}(z) = \begin{cases} - & \text{if } z < 0 \\ 0 & \text{if } z = 0 \\ + & \text{if } z > 0 \end{cases}$$

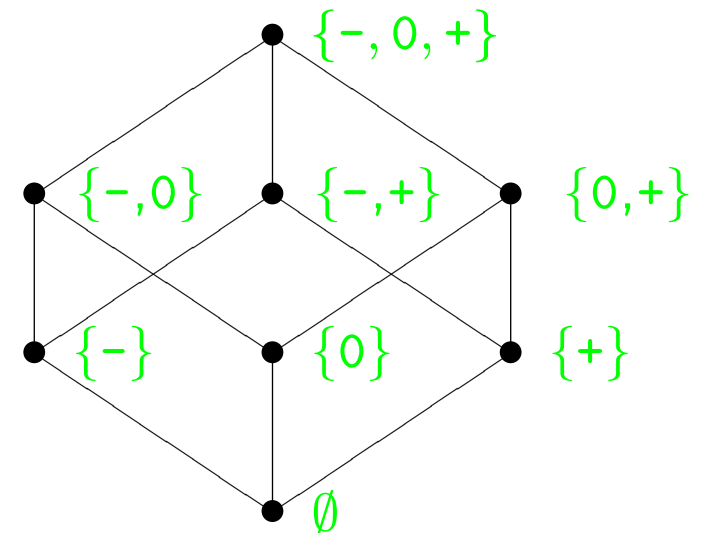
Galois connection

$$(\mathcal{P}(\mathbf{Z}), \alpha_{\text{sign}}, \gamma_{\text{sign}}, \mathcal{P}(\mathbf{Sign}))$$

with

$$\alpha_{\text{sign}}(Z) = \{\text{sign}(z) \mid z \in Z\}$$

$$\gamma_{\text{sign}}(S) = \{z \in \mathbf{Z} \mid \text{sign}(z) \in S\}$$



Properties of Galois Connections

Lemma: If (L, α, γ, M) is a Galois connection then:

- α uniquely determines γ by $\gamma(m) = \sqcup\{l \mid \alpha(l) \sqsubseteq m\}$
- γ uniquely determines α by $\alpha(l) = \sqcap\{m \mid l \sqsubseteq \gamma(m)\}$
- α is **completely additive** and γ is **completely multiplicative**

In particular $\alpha(\perp) = \perp$ and $\gamma(\top) = \top$.

Lemma:

- If $\alpha : L \rightarrow M$ is **completely additive** then there exists (an **upper adjoint**) $\gamma : M \rightarrow L$ such that (L, α, γ, M) is a Galois connection.
- If $\gamma : M \rightarrow L$ is **completely multiplicative** then there exists (a **lower adjoint**) $\alpha : L \rightarrow M$ such that (L, α, γ, M) is a Galois connection.

Fact: If (L, α, γ, M) is a Galois connection then

- $\alpha \circ \gamma \circ \alpha = \alpha$ and $\gamma \circ \alpha \circ \gamma = \gamma$

Example:

Define $\gamma_{\text{IS}} : \mathcal{P}(\text{Sign}) \rightarrow \text{Interval}$ by:

$$\begin{array}{ll} \gamma_{\text{IS}}(\{-, 0, +\}) = [-\infty, \infty] & \gamma_{\text{IS}}(\{-, 0\}) = [-\infty, 0] \\ \gamma_{\text{IS}}(\{-, +\}) = [-\infty, \infty] & \gamma_{\text{IS}}(\{0, +\}) = [0, \infty] \\ \gamma_{\text{IS}}(\{-\}) = [-\infty, -1] & \gamma_{\text{IS}}(\{0\}) = [0, 0] \\ \gamma_{\text{IS}}(\{+\}) = [1, \infty] & \gamma_{\text{IS}}(\emptyset) = \perp \end{array}$$

Does there exist an abstraction function

$$\alpha_{\text{IS}} : \text{Interval} \rightarrow \mathcal{P}(\text{Sign})$$

such that $(\text{Interval}, \alpha_{\text{IS}}, \gamma_{\text{IS}}, \mathcal{P}(\text{Sign}))$ is a Galois connection?

Example (cont.):

Is γ_{IS} completely multiplicative?

- if yes: then there exists a Galois connection
- if no: then there cannot exist a Galois connection

Lemma: If L and M are complete lattices and M is finite then $\gamma : M \rightarrow L$ is completely multiplicative if and only if the following hold:

- $\gamma : M \rightarrow L$ is monotone,
- $\gamma(\top) = \top$, and
- $\gamma(m_1 \sqcap m_2) = \gamma(m_1) \sqcap \gamma(m_2)$ whenever $m_1 \not\sqsubseteq m_2 \wedge m_2 \not\sqsubseteq m_1$

We calculate

$$\begin{aligned} \gamma_{IS}(\{-, 0\} \cap \{-, +\}) &= \gamma_{IS}(\{-\}) = [-\infty, -1] \\ \gamma_{IS}(\{-, 0\}) \sqcap \gamma_{IS}(\{-, +\}) &= [-\infty, 0] \sqcap [-\infty, \infty] = [-\infty, 0] \end{aligned}$$

showing that there is **no Galois connection** involving γ_{IS} .

Galois Connections are the Right Concept

We use the mundane approach to correctness to demonstrate this for:

- Admissible correctness relations
- Representation functions

The mundane approach: correctness relations

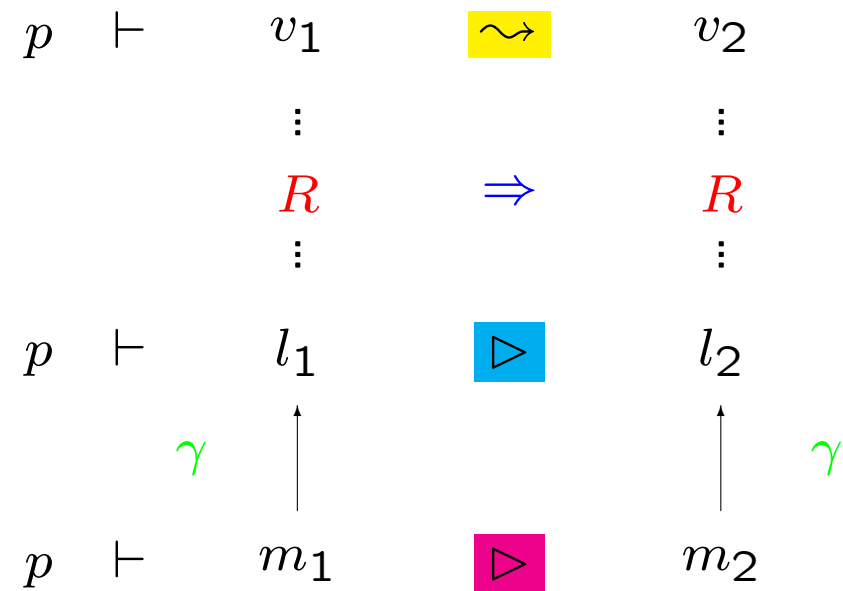
Assume

- $R : V \times L \rightarrow \{true, false\}$ is an admissible correctness relation
- (L, α, γ, M) is a Galois connection

Then $S : V \times M \rightarrow \{true, false\}$ defined by

$$v \ S \ m \quad \underline{\text{iff}} \quad v \ R \ (\gamma(m))$$

is an admissible correctness relation between V and M



The mundane approach: representation functions

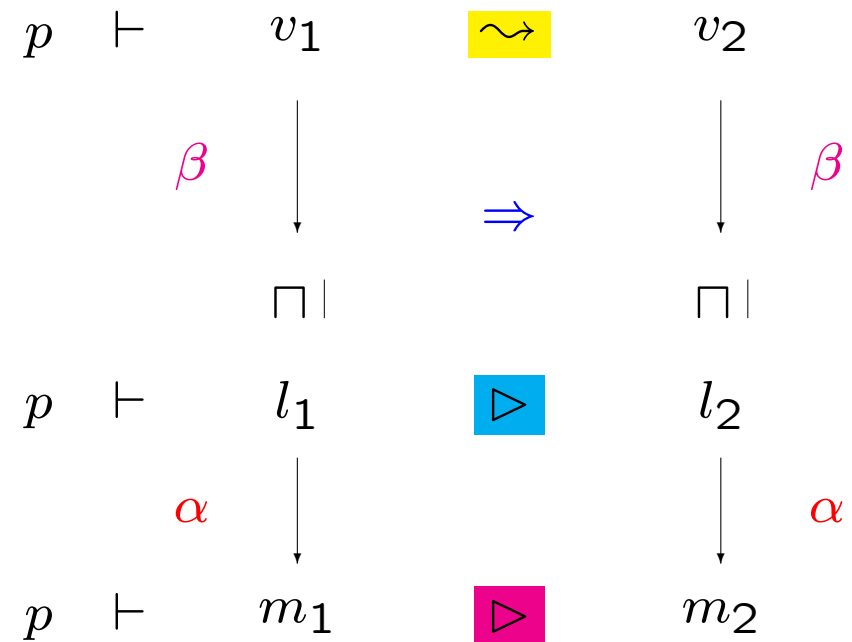
Assume

- $R : V \times L \rightarrow \{true, false\}$ is generated by $\beta : V \rightarrow L$
- (L, α, γ, M) is a Galois connection

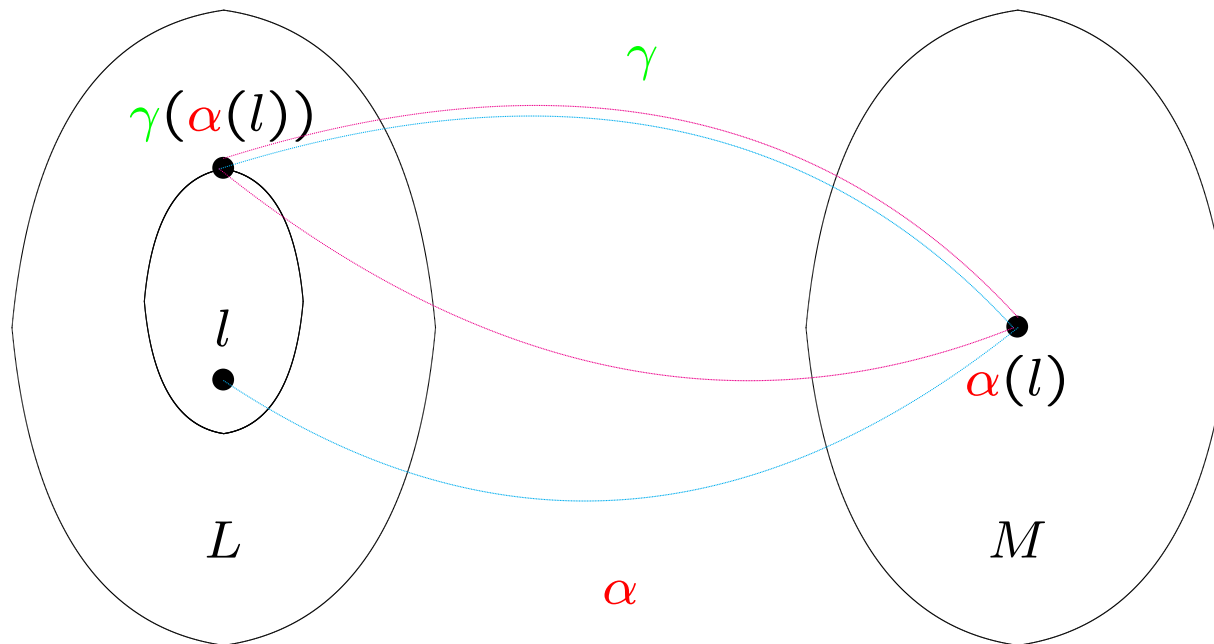
Then $S : V \times M \rightarrow \{true, false\}$ defined by

$$v \ S \ m \quad \text{iff} \quad v \ R \ (\gamma(m))$$

is generated by $\alpha \circ \beta : V \rightarrow M$



Galois Insertions



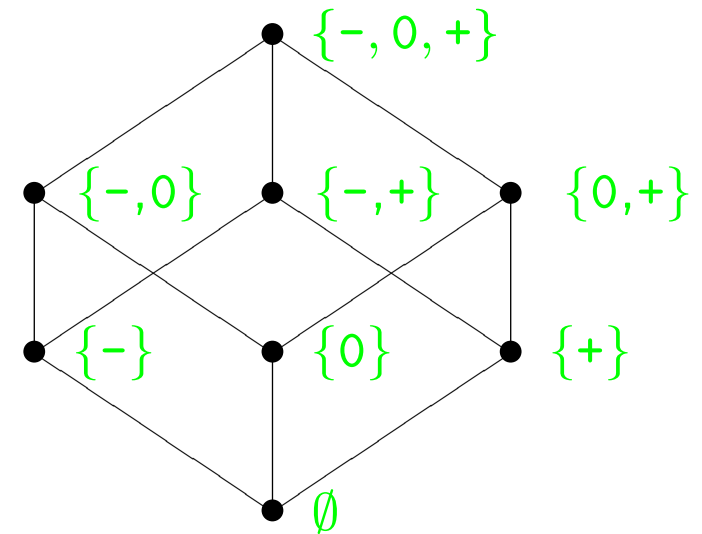
Monotone functions satisfying: $\gamma \circ \alpha \sqsupseteq \lambda l.l$ $\alpha \circ \gamma \sqsupseteq \lambda m.m$

Example (1):

$$(\mathcal{P}(\mathbf{Z}), \alpha_{\text{sign}}, \gamma_{\text{sign}}, \mathcal{P}(\mathbf{Sign}))$$

where $\text{sign} : \mathbf{Z} \rightarrow \mathbf{Sign}$ is specified by:

$$\text{sign}(z) = \begin{cases} - & \text{if } z < 0 \\ 0 & \text{if } z = 0 \\ + & \text{if } z > 0 \end{cases}$$



Is it a Galois insertion?

Example (2):

$$(\mathcal{P}(\mathbf{Z}), \alpha_{\text{signparity}}, \gamma_{\text{signparity}}, \mathcal{P}(\mathbf{Sign} \times \mathbf{Parity}))$$

where $\mathbf{Sign} = \{-, 0, +\}$ and $\mathbf{Parity} = \{\text{odd}, \text{even}\}$

and $\text{signparity} : \mathbf{Z} \rightarrow \mathbf{Sign} \times \mathbf{Parity}$:

$$\text{signparity}(z) = \begin{cases} (\text{sign}(z), \text{odd}) & \text{if } z \text{ is odd} \\ (\text{sign}(z), \text{even}) & \text{if } z \text{ is even} \end{cases}$$

Is it a Galois insertion?

Properties of Galois Insertions

Lemma: For a Galois connection (L, α, γ, M) the following claims are equivalent:

- (i) (L, α, γ, M) is a Galois insertion;
- (ii) α is surjective: $\forall m \in M : \exists l \in L : \alpha(l) = m$;
- (iii) γ is injective: $\forall m_1, m_2 \in M : \gamma(m_1) = \gamma(m_2) \Rightarrow m_1 = m_2$; and
- (iv) γ is an order-similarity: $\forall m_1, m_2 \in M : \gamma(m_1) \sqsubseteq \gamma(m_2) \Leftrightarrow m_1 \sqsubseteq m_2$.

Corollary: A Galois connection specified by an *extraction* function $\eta : V \rightarrow D$ is a Galois insertion if and only if η is surjective.

Example (1) reconsidered:

$$(\mathcal{P}(\mathbf{Z}), \alpha_{\text{sign}}, \gamma_{\text{sign}}, \mathcal{P}(\mathbf{Sign}))$$

$$\text{sign}(z) = \begin{cases} - & \text{if } z < 0 \\ 0 & \text{if } z = 0 \\ + & \text{if } z > 0 \end{cases}$$

is a Galois insertion because **sign** is surjective.

Example (2) reconsidered:

$$(\mathcal{P}(\mathbf{Z}), \alpha_{\text{signparity}}, \gamma_{\text{signparity}}, \mathcal{P}(\mathbf{Sign} \times \mathbf{Parity}))$$

$$\text{signparity}(z) = \begin{cases} (\text{sign}(z), \text{odd}) & \text{if } z \text{ is odd} \\ (\text{sign}(z), \text{even}) & \text{if } z \text{ is even} \end{cases}$$

is not a Galois insertion because **signparity** is not surjective.

Reduction Operators

Given a Galois connection (L, α, γ, M) it is **always** possible to obtain a Galois insertion by enforcing that the concretisation function γ is injective.

Idea: remove the superfluous elements from M using a *reduction operator*

$$\zeta : M \rightarrow M$$

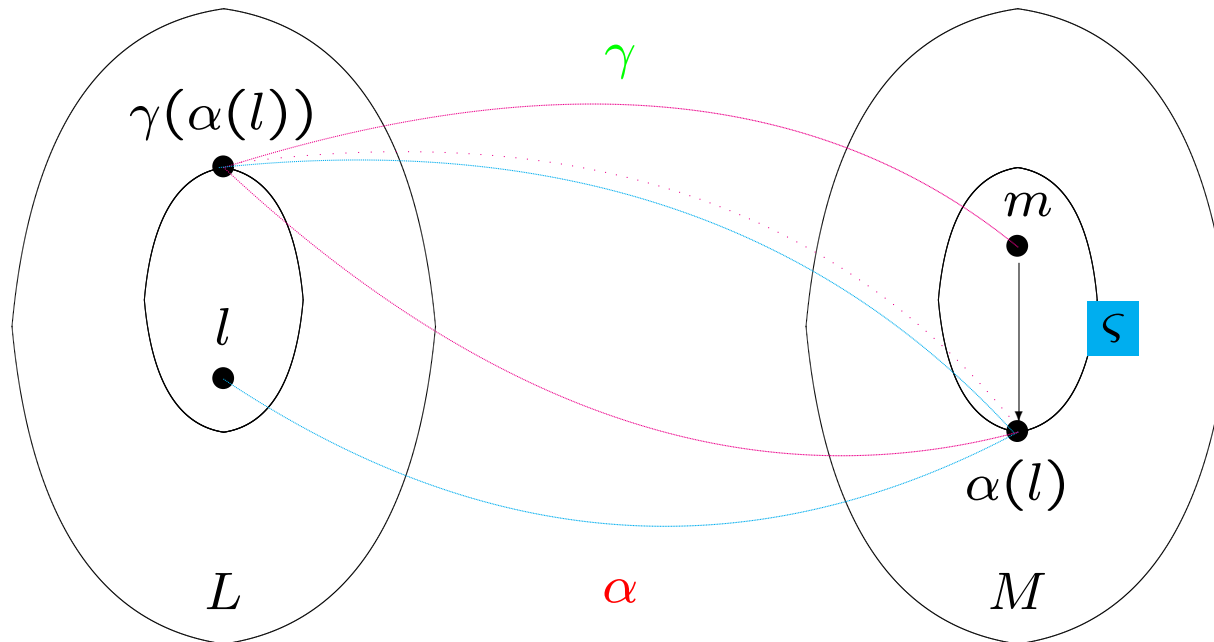
defined from the Galois connection.

Proposition: Let (L, α, γ, M) be a Galois connection and define the reduction operator $\zeta : M \rightarrow M$ by

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

Then $\zeta[M] = (\{\zeta(m) \mid m \in M\}, \sqsubseteq_M)$ is a complete lattice and $(L, \alpha, \gamma, \zeta[M])$ is a Galois insertion.

The reduction operator $\zeta : M \rightarrow M$



Reduction operators from extraction functions

Assume that the Galois connection $(\mathcal{P}(V), \alpha_\eta, \gamma_\eta, \mathcal{P}(D))$ is given by an extraction function $\eta : V \rightarrow D$.

Then the reduction operator ς_η is given by

$$\varsigma_\eta(D') = D' \cap \eta[V]$$

where $\eta[V] = \{d \in D \mid \exists v \in V : \eta(v) = d\}$.

Since $\varsigma_\eta[\mathcal{P}(D)]$ is isomorphic to $\mathcal{P}(\eta[V])$ the resulting Galois insertion is isomorphic to

$$(\mathcal{P}(V), \alpha_\eta, \gamma_\eta, \mathcal{P}(\eta[V]))$$

Systematic Design of Galois Connections

The “functional composition” (or “sequential composition”) of two Galois connections is also a Galois connection:

$$L_0 \begin{array}{c} \xleftarrow{\gamma_1} \\ \xrightarrow{\alpha_1} \end{array} L_1 \begin{array}{c} \xleftarrow{\gamma_2} \\ \xrightarrow{\alpha_2} \end{array} L_2 \begin{array}{c} \xleftarrow{\gamma_3} \\ \xrightarrow{\alpha_3} \end{array} \cdots \begin{array}{c} \xleftarrow{\gamma_k} \\ \xrightarrow{\alpha_k} \end{array} L_k$$

A catalogue of techniques for combining Galois connections:

- independent attribute method
- direct product
- reduced product
- total function space
- relational method
- direct tensor product
- reduced tensor product
- monotone function space

Running Example: Array Bound Analysis

Approximation of the difference in magnitude between two numbers (typically the index and the bound):

- a Galois connection for approximating pairs (z_1, z_2) of integers by their difference $|z_1| - |z_2|$
- a Galois connection for approximating integers using a finite lattice $\{<-1, -1, 0, +1, >+1\}$
- a Galois connection for their functional composition

Example: Difference in Magnitude

$$(\mathcal{P}(\mathbf{Z} \times \mathbf{Z}), \alpha_{\text{diff}}, \gamma_{\text{diff}}, \mathcal{P}(\mathbf{Z}))$$

where the extraction function $\text{diff} : \mathbf{Z} \times \mathbf{Z} \rightarrow \mathbf{Z}$ calculates the difference in magnitude:

$$\text{diff}(z_1, z_2) = |z_1| - |z_2|$$

The abstraction and concretisation functions are

$$\alpha_{\text{diff}}(ZZ) = \{|z_1| - |z_2| \mid (z_1, z_2) \in ZZ\}$$

$$\gamma_{\text{diff}}(Z) = \{(z_1, z_2) \mid |z_1| - |z_2| \in Z\}$$

for $ZZ \subseteq \mathbf{Z} \times \mathbf{Z}$ and $Z \subseteq \mathbf{Z}$.

Example: Finite Approximation

$$(\mathcal{P}(\mathbf{Z}), \alpha_{\text{range}}, \gamma_{\text{range}}, \mathcal{P}(\mathbf{Range}))$$

where $\mathbf{Range} = \{<-1, -1, 0, +1, >+1\}$

and the extraction function $\text{range} : \mathbf{Z} \rightarrow \mathbf{Range}$ is

$$\text{range}(z) = \begin{cases} <-1 & \text{if } z < -1 \\ -1 & \text{if } z = -1 \\ 0 & \text{if } z = 0 \\ +1 & \text{if } z = 1 \\ >+1 & \text{if } z > 1 \end{cases}$$

The abstraction and concretisation functions are

$$\alpha_{\text{range}}(Z) = \{\text{range}(z) \mid z \in Z\}$$

$$\gamma_{\text{range}}(R) = \{z \mid \text{range}(z) \in R\}$$

for $Z \subseteq \mathbf{Z}$ and $R \subseteq \mathbf{Range}$.

Example: Functional Composition

$$(\mathcal{P}(\mathbf{Z} \times \mathbf{Z}), \alpha_{\mathbf{R}}, \gamma_{\mathbf{R}}, \mathcal{P}(\mathbf{Range}))$$

where

$$\alpha_{\mathbf{R}} = \alpha_{\text{range}} \circ \alpha_{\text{diff}}$$

$$\gamma_{\mathbf{R}} = \gamma_{\text{diff}} \circ \gamma_{\text{range}}$$

The explicit formulae for the abstraction and concretisation functions

$$\alpha_{\mathbf{R}}(ZZ) = \{\text{range}(|z_1| - |z_2|) \mid (z_1, z_2) \in ZZ\}$$

$$\gamma_{\mathbf{R}}(R) = \{(z_1, z_2) \mid \text{range}(|z_1| - |z_2|) \in R\}$$

correspond to the extraction function $\text{range} \circ \text{diff}$.

Approximation of Pairs

Independent Attribute Method

Let $(L_1, \alpha_1, \gamma_1, M_1)$ and $(L_2, \alpha_2, \gamma_2, M_2)$ be Galois connections.

The *independent attribute method* gives a Galois connection

$$(L_1 \times L_2, \alpha, \gamma, M_1 \times M_2)$$

where

$$\alpha(l_1, l_2) = (\alpha_1(l_1), \alpha_2(l_2))$$

$$\gamma(m_1, m_2) = (\gamma_1(m_1), \gamma_2(m_2))$$

Example: Detection of Signs Analysis

Given

$$(\mathcal{P}(\mathbf{Z}), \alpha_{\text{sign}}, \gamma_{\text{sign}}, \mathcal{P}(\mathbf{Sign}))$$

using the extraction function `sign`.

The independent attribute method gives

$$(\mathcal{P}(\mathbf{Z}) \times \mathcal{P}(\mathbf{Z}), \alpha_{SS}, \gamma_{SS}, \mathcal{P}(\mathbf{Sign}) \times \mathcal{P}(\mathbf{Sign}))$$

where

$$\alpha_{SS}(Z_1, Z_2) = (\{\text{sign}(z) \mid z \in Z_1\}, \{\text{sign}(z) \mid z \in Z_2\})$$

$$\gamma_{SS}(S_1, S_2) = (\{z \mid \text{sign}(z) \in S_1\}, \{z \mid \text{sign}(z) \in S_2\})$$

Motivating the Relational Method

The independent attribute method often leads to imprecision!

Semantics: The expression $(x, -x)$ may have a value in

$$\{(z, -z) \mid z \in \mathbf{Z}\}$$

Analysis: When we use $\mathcal{P}(\mathbf{Z}) \times \mathcal{P}(\mathbf{Z})$ to represent sets of pairs of integers we cannot do better than representing $\{(z, -z) \mid z \in \mathbf{Z}\}$ by

$$(\mathbf{Z}, \mathbf{Z})$$

Hence the best property describing it will be

$$\alpha_{SS}(\mathbf{Z}, \mathbf{Z}) = (\{-, 0, +\}, \{-, 0, +\})$$

Relational Method

Let $(\mathcal{P}(V_1), \alpha_1, \gamma_1, \mathcal{P}(D_1))$ and $(\mathcal{P}(V_2), \alpha_2, \gamma_2, \mathcal{P}(D_2))$ be Galois connections.

The *relational method* will give rise to the Galois connection

$$(\mathcal{P}(V_1 \times V_2), \alpha, \gamma, \mathcal{P}(D_1 \times D_2))$$

where

$$\alpha(VV) = \bigcup \{ \alpha_1(\{v_1\}) \times \alpha_2(\{v_2\}) \mid (v_1, v_2) \in VV \}$$

$$\gamma(DD) = \{ (v_1, v_2) \mid \alpha_1(\{v_1\}) \times \alpha_2(\{v_2\}) \subseteq DD \}$$

Generalisation to arbitrary complete lattices: use *tensor products*.

Relational Method from Extraction Functions

Assume that the Galois connections $(\mathcal{P}(V_i), \alpha_i, \gamma_i, \mathcal{P}(D_i))$ are given by *extraction functions* $\eta_i : V_i \rightarrow D_i$ as in

$$\alpha_i(V_i') = \{\eta_i(v_i) \mid v_i \in V_i'\}$$

$$\gamma_i(D_i') = \{v_i \mid \eta_i(v_i) \in D_i'\}$$

Then the Galois connection $(\mathcal{P}(V_1 \times V_2), \alpha, \gamma, \mathcal{P}(D_1 \times D_2))$ has

$$\alpha(VV) = \{(\eta_1(v_1), \eta_2(v_2)) \mid (v_1, v_2) \in VV\}$$

$$\gamma(DD) = \{(v_1, v_2) \mid (\eta_1(v_1), \eta_2(v_2)) \in DD\}$$

which also can be obtained directly from the extraction function $\eta : V_1 \times V_2 \rightarrow D_1 \times D_2$ defined by

$$\eta(v_1, v_2) = (\eta_1(v_1), \eta_2(v_2))$$

Example: Detection of Signs Analysis

Using the relational method we get a Galois connection

$$(\mathcal{P}(\mathbf{Z} \times \mathbf{Z}), \alpha_{SS'}, \gamma_{SS'}, \mathcal{P}(\mathbf{Sign} \times \mathbf{Sign}))$$

where

$$\begin{aligned}\alpha_{SS'}(ZZ) &= \{(\text{sign}(z_1), \text{sign}(z_2)) \mid (z_1, z_2) \in ZZ\} \\ \gamma_{SS'}(SS) &= \{(z_1, z_2) \mid (\text{sign}(z_1), \text{sign}(z_2)) \in SS\}\end{aligned}$$

corresponding to an extraction function `twosigns` : $\mathbf{Z} \times \mathbf{Z} \rightarrow \mathbf{Sign} \times \mathbf{Sign}$ defined by

$$\text{twosigns}(z_1, z_2) = (\text{sign}(z_1), \text{sign}(z_2))$$

Advantages of the Relational Method

Semantics: The expression $(x, -x)$ may have a value in

$$\{(z, -z) \mid z \in \mathbf{Z}\}$$

In the present setting $\{(z, -z) \mid z \in \mathbf{Z}\}$ is an element of $\mathcal{P}(\mathbf{Z} \times \mathbf{Z})$.

Analysis: The best “relational” property describing it is

$$\alpha_{SS'}(\{(z, -z) \mid z \in \mathbf{Z}\}) = \{(-, +), (0, 0), (+, -)\}$$

whereas the best “independent attribute” property was

$$\alpha_{SS}(\mathbf{Z}, \mathbf{Z}) = (\{-, 0, +\}, \{-, 0, +\})$$

Function Spaces

Total Function Space

Let (L, α, γ, M) be a Galois connection and let S be a set.

The Galois connection for the *total function space*

$$(S \rightarrow L, \alpha', \gamma', S \rightarrow M)$$

is defined by

$$\alpha'(f) = \alpha \circ f \qquad \gamma'(g) = \gamma \circ g$$

Do we need to assume that S is non-empty?

Monotone Function Space

Let $(L_1, \alpha_1, \gamma_1, M_1)$ and $(L_2, \alpha_2, \gamma_2, M_2)$ be Galois connections.

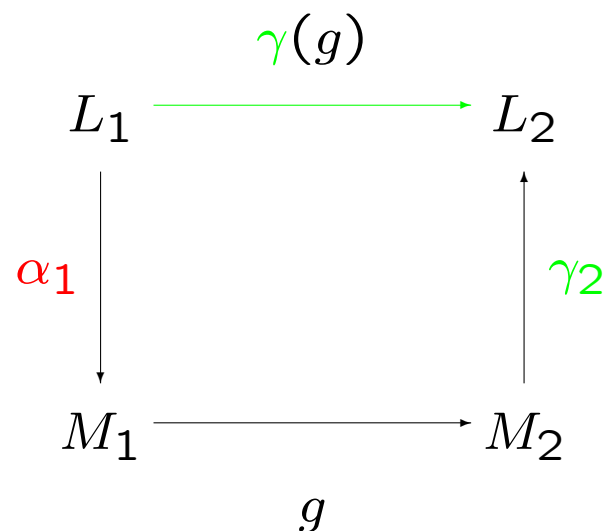
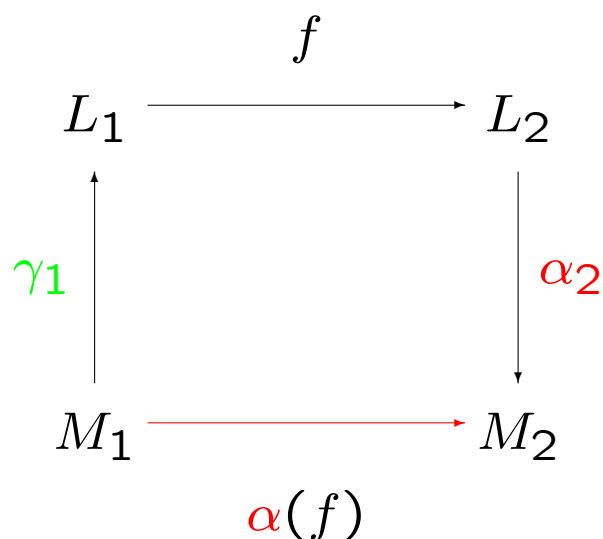
The Galois connection for the *monotone function space*

$$(L_1 \rightarrow L_2, \alpha, \gamma, M_1 \rightarrow M_2)$$

is defined by

$$\alpha(f) = \alpha_2 \circ f \circ \gamma_1$$

$$\gamma(g) = \gamma_2 \circ g \circ \alpha_1$$



Performing Analyses Simultaneously

Direct Product

Let $(L, \alpha_1, \gamma_1, M_1)$ and $(L, \alpha_2, \gamma_2, M_2)$ be Galois connections.

The *direct product* is the Galois connection

$$(L, \alpha, \gamma, M_1 \times M_2)$$

defined by

$$\begin{aligned}\alpha(l) &= (\alpha_1(l), \alpha_2(l)) \\ \gamma(m_1, m_2) &= \gamma_1(m_1) \sqcap \gamma_2(m_2)\end{aligned}$$

Example:

Combining the [detection of signs analysis](#) for pairs of integers with the [analysis of difference in magnitude](#).

We get the Galois connection

$$(\mathcal{P}(\mathbf{Z} \times \mathbf{Z}), \alpha_{SSR}, \gamma_{SSR}, \mathcal{P}(\mathbf{Sign} \times \mathbf{Sign}) \times \mathcal{P}(\mathbf{Range}))$$

where

$$\alpha_{SSR}(ZZ) = (\{(\text{sign}(z_1), \text{sign}(z_2)) \mid (z_1, z_2) \in ZZ\}, \\ \{\text{range}(|z_1| - |z_2|) \mid (z_1, z_2) \in ZZ\})$$

$$\gamma_{SSR}(SS, R) = \{(z_1, z_2) \mid (\text{sign}(z_1), \text{sign}(z_2)) \in SS\} \\ \cap \{(z_1, z_2) \mid \text{range}(|z_1| - |z_2|) \in R\}$$

Motivating the Direct Tensor Product

The expression $(x, 3*x)$ may have a value in

$$\{(z, 3 * z) \mid z \in \mathbf{Z}\}$$

which is described by

$$\alpha_{SSR}(\{(z, 3 * z) \mid z \in \mathbf{Z}\}) = (\{(-, -), (0, 0), (+, +)\}, \{0, <-1\})$$

But

- any pair described by $(0, 0)$ will have a difference in magnitude described by 0
- any pair described by $(-, -)$ or $(+, +)$ will have a difference in magnitude described by <-1

and the analysis cannot express this.

Direct Tensor Product

Let $(\mathcal{P}(V), \alpha_1, \gamma_1, \mathcal{P}(D_1))$ and $(\mathcal{P}(V), \alpha_2, \gamma_2, \mathcal{P}(D_2))$ be Galois connections.

The *direct tensor product* is the Galois connection

$$(\mathcal{P}(V), \alpha, \gamma, \mathcal{P}(D_1 \times D_2))$$

defined by

$$\begin{aligned}\alpha(V') &= \bigcup \{ \alpha_1(\{v\}) \times \alpha_2(\{v\}) \mid v \in V' \} \\ \gamma(DD) &= \{ v \mid \alpha_1(\{v\}) \times \alpha_2(\{v\}) \subseteq DD \}\end{aligned}$$

Direct Tensor Product from Extraction Functions

Assume that the Galois connections $(\mathcal{P}(V), \alpha_i, \gamma_i, \mathcal{P}(D_i))$ are given by *extraction functions* $\eta_i : V \rightarrow D_i$ as in

$$\alpha_i(V') = \{\eta_i(v) \mid v \in V'\}$$

$$\gamma_i(D'_i) = \{v \mid \eta_i(v) \in D'_i\}$$

The Galois connection $(\mathcal{P}(V), \alpha, \gamma, \mathcal{P}(D_1 \times D_2))$ has

$$\alpha(V') = \{(\eta_1(v), \eta_2(v)) \mid v \in V'\}$$

$$\gamma(DD) = \{v \mid (\eta_1(v), \eta_2(v)) \in DD\}$$

corresponding to the extraction function $\eta : V \rightarrow D_1 \times D_2$ defined by

$$\eta(v) = (\eta_1(v), \eta_2(v))$$

Example:

Using the direct tensor product to combine the **detection of signs analysis** for pairs of integers with the **analysis of difference in magnitude**.

$$(\mathcal{P}(\mathbf{Z} \times \mathbf{Z}), \alpha_{SSR'}, \gamma_{SSR'}, \mathcal{P}(\mathbf{Sign} \times \mathbf{Sign} \times \mathbf{Range}))$$

is given by

$$\begin{aligned}\alpha_{SSR'}(ZZ) &= \{(\mathbf{sign}(z_1), \mathbf{sign}(z_2), \mathbf{range}(|z_1| - |z_2|)) \mid (z_1, z_2) \in ZZ\} \\ \gamma_{SSR'}(SSR) &= \{(z_1, z_2) \mid (\mathbf{sign}(z_1), \mathbf{sign}(z_2), \mathbf{range}(|z_1| - |z_2|)) \in SSR\}\end{aligned}$$

corresponding to **twosignsrange** : $\mathbf{Z} \times \mathbf{Z} \rightarrow \mathbf{Sign} \times \mathbf{Sign} \times \mathbf{Range}$ given by

$$\mathbf{twosignsrange}(z_1, z_2) = (\mathbf{sign}(z_1), \mathbf{sign}(z_2), \mathbf{range}(|z_1| - |z_2|))$$

Advantages of the Direct Tensor Product

The expression $(x, 3*x)$ may have a value in $\{(z, 3 * z) \mid z \in \mathbf{Z}\}$ which in the direct tensor product can be described by

$$\alpha_{SSR'}(\{(z, 3 * z) \mid z \in \mathbf{Z}\}) = \{(-, -, <-1), (0, 0, 0), (+, +, <-1)\}$$

compared to the direct product that gave

$$\alpha_{SSR}(\{(z, 3 * z) \mid z \in \mathbf{Z}\}) = (\{(-, -), (0, 0), (+, +)\}, \{0, <-1\})$$

Note that the Galois connection is *not* a Galois insertion because

$$\gamma_{SSR'}(\emptyset) = \emptyset = \gamma_{SSR'}(\{(0, 0, <-1)\})$$

so $\gamma_{SSR'}$ is not injective and hence we do not have a Galois insertion.

From Direct to Reduced

Reduced Product

Let $(L, \alpha_1, \gamma_1, M_1)$ and $(L, \alpha_2, \gamma_2, M_2)$ be Galois connections.

The *reduced product* is the Galois *insertion*

$$(L, \alpha, \gamma, \varsigma[M_1 \times M_2])$$

defined by

$$\alpha(l) = (\alpha_1(l), \alpha_2(l))$$

$$\gamma(m_1, m_2) = \gamma_1(m_1) \sqcap \gamma_2(m_2)$$

$$\varsigma(m_1, m_2) = \bigsqcap \{(m'_1, m'_2) \mid \gamma_1(m_1) \sqcap \gamma_2(m_2) = \gamma_1(m'_1) \sqcap \gamma_2(m'_2)\}$$

Reduced Tensor Product

Let $(\mathcal{P}(V), \alpha_1, \gamma_1, \mathcal{P}(D_1))$ and $(\mathcal{P}(V), \alpha_2, \gamma_2, \mathcal{P}(D_2))$ be Galois connections.

The *reduced tensor product* is the Galois *insertion*

$$(\mathcal{P}(V), \alpha, \gamma, \varsigma[\mathcal{P}(D_1 \times D_2)])$$

defined by

$$\begin{aligned}\alpha(V') &= \bigcup \{ \alpha_1(\{v\}) \times \alpha_2(\{v\}) \mid v \in V' \} \\ \gamma(DD) &= \{ v \mid \alpha_1(\{v\}) \times \alpha_2(\{v\}) \subseteq DD \} \\ \varsigma(DD) &= \bigcap \{ DD' \mid \gamma(DD) = \gamma(DD') \}\end{aligned}$$

Example: Array Bounds Analysis

The superfluous elements of $\mathcal{P}(\mathbf{Sign} \times \mathbf{Sign} \times \mathbf{Range})$ will be removed when we use a reduced tensor product:

The reduction operator $\mathcal{S}_{SSR'}$ amounts to

$$\mathcal{S}_{SSR'}(SSR) = \bigcap \{SSR' \mid \gamma_{SSR'}(SSR) = \gamma_{SSR'}(SSR')\}$$

where $SSR, SSR' \subseteq \mathbf{Sign} \times \mathbf{Sign} \times \mathbf{Range}$.

The singleton sets constructed from the following 16 elements

$$\begin{array}{l} (-, 0, <-1), \quad (-, 0, -1), \quad (-, 0, 0), \\ (0, -, 0), \quad (0, -, +1), \quad (0, -, >+1), \\ (0, 0, <-1), \quad (0, 0, -1), \quad (0, 0, +1), \quad (0, 0, >+1), \\ (0, +, 0), \quad (0, +, +1), \quad (0, +, >+1), \\ (+, 0, <-1), \quad (+, 0, -1), \quad (+, 0, 0) \end{array}$$

will be mapped to the empty set (as they are useless).

Example (cont.): Array Bounds Analysis

The remaining 29 elements of $\mathbf{Sign} \times \mathbf{Sign} \times \mathbf{Range}$ are

$(-, -, <-1), (-, -, -1), (-, -, 0), (-, -, +1), (-, -, >+1),$
 $(-, 0, +1), (-, 0, >+1),$
 $(-, +, <-1), (-, +, -1), (-, +, 0), (-, +, +1), (-, +, >+1),$
 $(0, -, <-1), (0, -, -1), (0, 0, 0), (0, +, <-1), (0, +, -1),$
 $(+, -, <-1), (+, -, -1), (+, -, 0), (+, -, +1), (+, -, >+1),$
 $(+, 0, +1), (+, 0, >+1),$
 $(+, +, <-1), (+, +, -1), (+, +, 0), (+, +, +1), (+, +, >+1)$

and they describe disjoint subsets of $\mathbf{Z} \times \mathbf{Z}$.

Any collection of properties can be described in 4 bytes.

Summary

The Array Bound Analysis has been designed from three simple Galois connections specified by **extraction functions**:

- (i) an analysis approximating integers by their sign,
- (ii) an analysis approximating pairs of integers by their difference in magnitude, and
- (iii) an analysis approximating integers by their closeness to 0, 1 and -1 .

These analyses have been combined using:

- (iv) the **relational product** of analysis (i) with itself,
- (v) the **functional composition** of analyses (ii) and (iii), and
- (vi) the **reduced tensor product** of analyses (iv) and (v).

Induced Operations

Given: Galois connections $(L_i, \alpha_i, \gamma_i, M_i)$ so that M_i is more approximate than (i.e. is coarser than) L_i .

Aim: Replace an existing analysis over L_i with an analysis making use of the coarser structure of M_i .

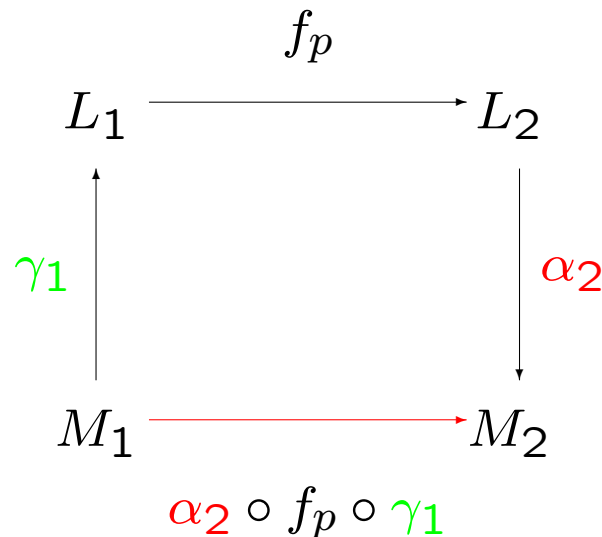
Methods:

- **Inducing along the abstraction function:** move the computations from L_i to M_i .
- Application to Data Flow Analysis.
- **Inducing along the concretisation function:** move a widening from M_i to L_i .

Inducing along the Abstraction Function

Given Galois connections $(L_i, \alpha_i, \gamma_i, M_i)$ so that M_i is more approximate than L_i .

Replace an existing analysis $f_p : L_1 \rightarrow L_2$ with a new and more approximate analysis $g_p : M_1 \rightarrow M_2$: take $g_p = \alpha_2 \circ f_p \circ \gamma_1$.



The analysis $\alpha_2 \circ f_p \circ \gamma_1$ is *induced* from f_p and the Galois connections.

Example:

A very precise analysis for `plus` based on $\mathcal{P}(\mathbf{Z})$ and $\mathcal{P}(\mathbf{Z} \times \mathbf{Z})$:

$$f_{\text{plus}}(\mathbf{Z}\mathbf{Z}) = \{z_1 + z_2 \mid (z_1, z_2) \in \mathbf{Z}\mathbf{Z}\}$$

Two Galois connections

$$(\mathcal{P}(\mathbf{Z}), \alpha_{\text{sign}}, \gamma_{\text{sign}}, \mathcal{P}(\mathbf{Sign}))$$

$$(\mathcal{P}(\mathbf{Z} \times \mathbf{Z}), \alpha_{\text{SS}'}, \gamma_{\text{SS}'}, \mathcal{P}(\mathbf{Sign} \times \mathbf{Sign}))$$

An approximate analysis for `plus` based on $\mathcal{P}(\mathbf{Sign})$ and $\mathcal{P}(\mathbf{Sign} \times \mathbf{Sign})$:

$$g_{\text{plus}} = \alpha_{\text{sign}} \circ f_{\text{plus}} \circ \gamma_{\text{SS}'}$$

Example (cont.):

We calculate

$$\begin{aligned} g_{\text{plus}}(SS) &= \alpha_{\text{sign}}(f_{\text{plus}}(\gamma_{SS'}(SS))) \\ &= \alpha_{\text{sign}}(f_{\text{plus}}(\{(z_1, z_2) \in \mathbf{Z} \times \mathbf{Z} \mid (\text{sign}(z_1), \text{sign}(z_2)) \in SS\})) \\ &= \alpha_{\text{sign}}(\{z_1 + z_2 \mid z_1, z_2 \in \mathbf{Z}, (\text{sign}(z_1), \text{sign}(z_2)) \in SS\}) \\ &= \{\text{sign}(z_1 + z_2) \mid z_1, z_2 \in \mathbf{Z}, (\text{sign}(z_1), \text{sign}(z_2)) \in SS\} \\ &= \bigcup \{s_1 \oplus s_2 \mid (s_1, s_2) \in SS\} \end{aligned}$$

where $\oplus : \mathbf{Sign} \times \mathbf{Sign} \rightarrow \mathcal{P}(\mathbf{Sign})$ is the “addition” operator on signs (so e.g. $+ \oplus + = \{+\}$ and $+ \oplus - = \{-, 0, +\}$).

The Mundane Correctness of f_p carries over to g_p

The correctness relation R_i for V_i and L_i :

$$R_i : V_i \times L_i \rightarrow \{true, false\} \text{ is generated by } \beta_i : V_i \rightarrow L_i$$

Correctness of f_p means

$$(p \vdash \cdot \rightsquigarrow \cdot) (R_1 \twoheadrightarrow R_2) f_p$$

(with $R_1 \twoheadrightarrow R_2$ being generated by $\beta_1 \twoheadrightarrow \beta_2$).

The correctness relation S_i for V_i and M_i :

$$S_i : V_i \times M_i \rightarrow \{true, false\} \text{ is generated by } \alpha_i \circ \beta_i : V_i \rightarrow M_i$$

One can prove that

$$(p \vdash \cdot \rightsquigarrow \cdot) (R_1 \twoheadrightarrow R_2) f_p \wedge \alpha_2 \circ f_p \circ \gamma_1 \sqsubseteq g_p$$

$$\Rightarrow (p \vdash \cdot \rightsquigarrow \cdot) (S_1 \twoheadrightarrow S_2) g_p$$

with $S_1 \twoheadrightarrow S_2$ being generated by $(\alpha_1 \circ \beta_1) \twoheadrightarrow (\alpha_2 \circ \beta_2)$.

Fixed Points in the Induced Analysis

Let $f_p = \text{Ifp}(F)$ for a monotone function $F : (L_1 \rightarrow L_2) \rightarrow (L_1 \rightarrow L_2)$.

The Galois connections $(L_i, \alpha_i, \gamma_i, M_i)$ give rise to a Galois connection $(L_1 \rightarrow L_2, \alpha, \gamma, M_1 \rightarrow M_2)$.

Take $g_p = \text{Ifp}(G)$ where $G : (M_1 \rightarrow M_2) \rightarrow (M_1 \rightarrow M_2)$ is an “upper approximation” to F : we demand that $\alpha \circ F \circ \gamma \sqsubseteq G$.

Then for all $m \in M_1 \rightarrow M_2$:

$$G(m) \sqsubseteq m \Rightarrow F(\gamma(m)) \sqsubseteq \gamma(m)$$

and $\text{Ifp}(F) \sqsubseteq \gamma(\text{Ifp}(G))$ and $\alpha(\text{Ifp}(F)) \sqsubseteq \text{Ifp}(G)$

Application to Data Flow Analysis

A *generalised Monotone Framework* consists of:

- the property space: a complete lattice $L = (L, \sqsubseteq)$;
- the set \mathcal{F} of monotone functions from L to L .

An *instance* **A** of a generalised Monotone Framework consists of:

- a finite flow, $F \subseteq \mathbf{Lab} \times \mathbf{Lab}$;
- a finite set of extremal labels, $E \subseteq \mathbf{Lab}$;
- an extremal value, $\iota \in L$; and
- a mapping f . from the labels \mathbf{Lab} of F and E to monotone transfer functions from L to L .

Application to Data Flow Analysis

Let (L, α, γ, M) be a Galois connection.

Consider an instance **B** of the generalised Monotone Framework M that satisfies

- the mapping g from the labels **Lab** of F and E to monotone transfer functions of $M \rightarrow M$ satisfies $g_\ell \sqsupseteq \alpha \circ f_\ell \circ \gamma$ for all ℓ ; and
- the extremal value j satisfies $\gamma(j) = \iota$;

and otherwise **B** is as **A**.

One can show that a solution to the **B**-constraints gives rise to a solution to the **A**-constraints:

$$(B_\circ, B_\bullet) \models \mathbf{B}^\sqsupseteq \text{ implies } (\gamma \circ B_\circ, \gamma \circ B_\bullet) \models \mathbf{A}^\sqsupseteq$$

The Mundane Approach to Semantic Correctness

Here $F = \text{flow}(S_*)$ and $E = \{\text{init}(S_*)\}$.

Correctness of every solution to $A \sqsupseteq$ amounts to:

Assume $(A_\circ, A_\bullet) \models A \sqsupseteq$ and $\langle S_*, \sigma_1 \rangle \rightarrow^* \sigma_2$.

Then $\beta(\sigma_1) \sqsubseteq \iota$ implies $\beta(\sigma_2) \sqsubseteq \sqcup\{A_\bullet(\ell) \mid \ell \in \text{final}(S_*)\}$.

where $\beta : \text{State} \rightarrow L$.

One can then prove the correctness result for B:

Assume $(B_\circ, B_\bullet) \models B \sqsupseteq$ and $\langle S_*, \sigma_1 \rangle \rightarrow^* \sigma_2$.

Then $(\alpha \circ \beta)(\sigma_1) \sqsubseteq \jmath$ implies $(\alpha \circ \beta)(\sigma_2) \sqsubseteq \sqcup\{B_\bullet(\ell) \mid \ell \in \text{final}(S_*)\}$.

Sets of States Analysis

Generalised Monotone Framework over $(\mathcal{P}(\text{State}), \subseteq)$.

Instance **SS** for S_* :

- the flow F is $\text{flow}(S_*)$;
- the set E of extremal labels is $\{\text{init}(S_*)\}$;
- the extremal value ι is State ; and
- the transfer functions are given by f_l^{SS} :

$$[x := a]^\ell \quad f_l^{\text{SS}}(\Sigma) = \{\sigma[x \mapsto \mathcal{A}[[a]]\sigma] \mid \sigma \in \Sigma\}$$

$$[\text{skip}]^\ell \quad f_l^{\text{SS}}(\Sigma) = \Sigma$$

$$[b]^\ell \quad f_l^{\text{SS}}(\Sigma) = \Sigma$$

where $\Sigma \subseteq \text{State}$.

Correctness: Assume $(\text{SS}_0, \text{SS}_\bullet) \models \text{SS}^\supseteq$ and $\langle S_*, \sigma_1 \rangle \rightarrow^* \sigma_2$.
Then $\sigma_1 \in \text{State}$ implies $\sigma_2 \in \bigcup \{\text{SS}_\bullet(\ell) \mid \ell \in \text{final}(S_*)\}$.

Constant Propagation Analysis

Generalised Monotone Framework over $\widehat{\text{State}}_{\text{CP}} = ((\text{Var} \rightarrow \mathbf{Z}^\top)_\perp, \sqsubseteq)$.

Instance **CP** for S_\star :

- the flow F is $\text{flow}(S_\star)$;
- the set E of extremal labels is $\{\text{init}(S_\star)\}$;
- the extremal value ι is $\lambda x. \top$; and
- the transfer functions are given by the mapping f_l^{CP} :

$$[x := a]^l : f_l^{\text{CP}}(\hat{\sigma}) = \begin{cases} \perp & \text{if } \hat{\sigma} = \perp \\ \hat{\sigma}[x \mapsto \mathcal{A}_{\text{CP}}[[a]]\hat{\sigma}] & \text{otherwise} \end{cases}$$

$$[\text{skip}]^l : f_l^{\text{CP}}(\hat{\sigma}) = \hat{\sigma}$$

$$[b]^l : f_l^{\text{CP}}(\hat{\sigma}) = \hat{\sigma}$$

Galois Connection

The representation function $\beta_{\text{CP}} : \text{State} \rightarrow \widehat{\text{State}}_{\text{CP}}$ is defined by

$$\beta_{\text{CP}}(\sigma) = \sigma$$

This gives rise to a Galois connection

$$(\mathcal{P}(\text{State}), \alpha_{\text{CP}}, \gamma_{\text{CP}}, \widehat{\text{State}}_{\text{CP}})$$

where $\alpha_{\text{CP}}(\Sigma) = \sqcup\{\beta_{\text{CP}}(\sigma) \mid \sigma \in \Sigma\}$ and $\gamma_{\text{CP}}(\hat{\sigma}) = \{\sigma \mid \beta_{\text{CP}}(\sigma) \sqsubseteq \hat{\sigma}\}$.

One can show that for all labels ℓ

$$f_{\ell}^{\text{CP}} \sqsupseteq \alpha_{\text{CP}} \circ f_{\ell}^{\text{SS}} \circ \gamma_{\text{CP}} \text{ as well as } \gamma_{\text{CP}}(\lambda x. \top) = \text{State}$$

It follows that **CP** is an upper approximation to the analysis induced from **SS** and the Galois connection; therefore it is correct.

Inducing along the Concretisation Function

Given an upper bound operator

$$\nabla_M : M \times M \rightarrow M$$

and a Galois connection (L, α, γ, M) .

Define an upper bound operator

$$\nabla_L : L \times L \rightarrow L$$

by

$$l_1 \nabla_L l_2 = \gamma(\alpha(l_1) \nabla_M \alpha(l_2))$$

It defines a widening operator if one of the following conditions holds:

- (i) M satisfies the Ascending Chain Condition, or
- (ii) (L, α, γ, M) is a Galois insertion and $\nabla_M : M \times M \rightarrow M$ is a widening.

Precision of the Induced Widening Operator

Lemma: Let (L, α, γ, M) be a Galois insertion such that $\gamma(\perp_M) = \perp_L$ and let $\nabla_M : M \times M \rightarrow M$ be a widening operator.

Then the widening operator $\nabla_L : L \times L \rightarrow L$ defined by

$$l_1 \nabla_L l_2 = \gamma(\alpha(l_1) \nabla_M \alpha(l_2))$$

satisfies

$$\text{Ifp}_{\nabla_L}(f) = \gamma(\text{Ifp}_{\nabla_M}(\alpha \circ f \circ \gamma))$$

for all monotone functions $f : L \rightarrow L$.

Precision of the Induced Widening Operator

Corollary: Let M be of finite height, let (L, α, γ, M) be a Galois insertion (such that $\gamma(\perp_M) = \perp_L$), and let ∇_M equal the least upper bound operator \sqcup_M .

Then the above lemma shows that $\text{Ifp}_{\nabla_L}(f) = \gamma(\text{Ifp}(\alpha \circ f \circ \gamma))$.

This means that $\text{Ifp}_{\nabla_L}(f)$ equals the result we would have obtained if we decided to work with $\alpha \circ f \circ \gamma : M \rightarrow M$ instead of the given $f : L \rightarrow L$; furthermore the number of iterations needed turn out to be the same. However, for all other operations the increased precision of L is available.