Abstract Interpretation from a Topological Perspective

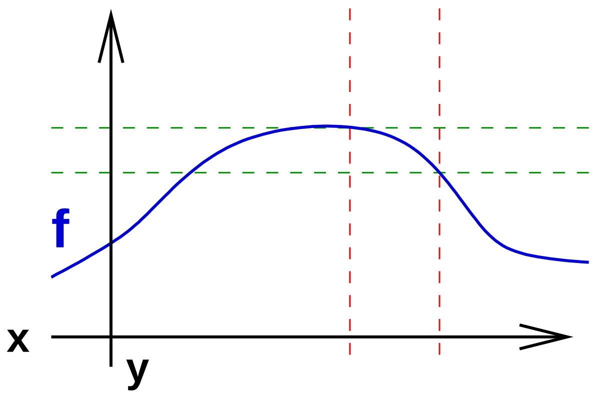
David Schmidt

Kansas State University

www.cis.ksu.edu/∼schmidt

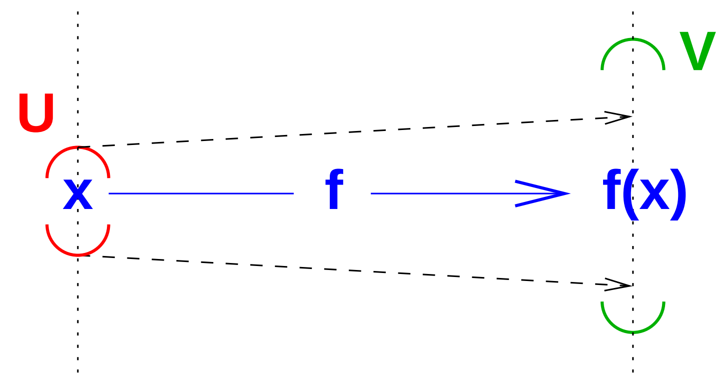
Motivation and overview of results

Topology *studies convergent mappings*



As argument x converges within smaller and smaller open invervals, so does answer $f(x)$.

In topology, "open interval" (a property of interest) generalizes to *open set*; "convergent function" generalizes to *continuous function*:



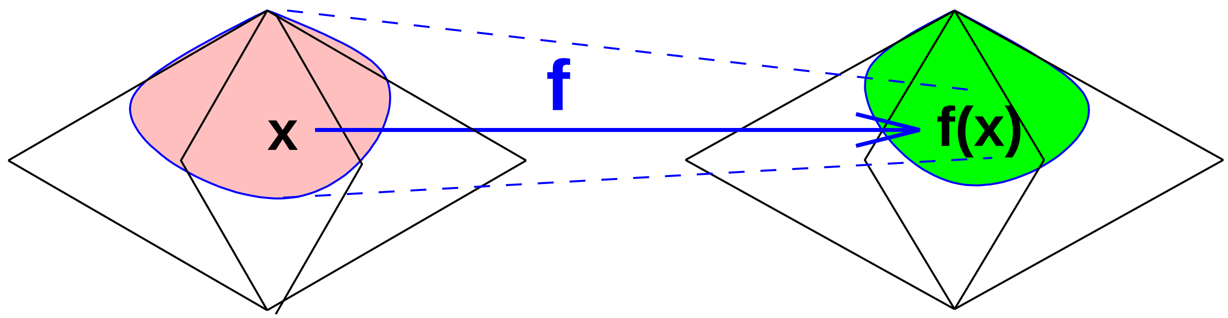
Let $\mathcal{O}_\Sigma$ be the open sets. $f : \Sigma \to \Sigma$ is *(topologically) continuous* iff for all $x \in \Sigma$ and $V \in \mathcal{O}_\Sigma$, if $f(x) \in V$, then there exists some $U \in \mathcal{O}_\Sigma$ such that $x \in U$ and $f[U] \subseteq V$.

Domain theory uses the Scott topology

For an algebraic lattice, $(\Sigma, \sqsubseteq)$, the Scott-open sets are those $U \subseteq \Sigma$ such that U is

- *upwards closed:* if $c \in U, c \sqsubseteq d$, then $d \in U$ also.
- *closed under tails of chains:* for every chain $C \subseteq \Sigma$, if $\sqcup C \in U$, then $\exists c \in C$ such that $c \in U$ also.

(A Scott-open U is like an interval, $(c, +\infty]$, on the real line.)



$f : \Sigma \longrightarrow \Delta$ is Scott-topologically continuous iff it is chain continuous (i.e., for every chain $C \subseteq \Sigma$, $f(\sqcup C) = \sqcup\{f(c) \mid c \in C\}$).

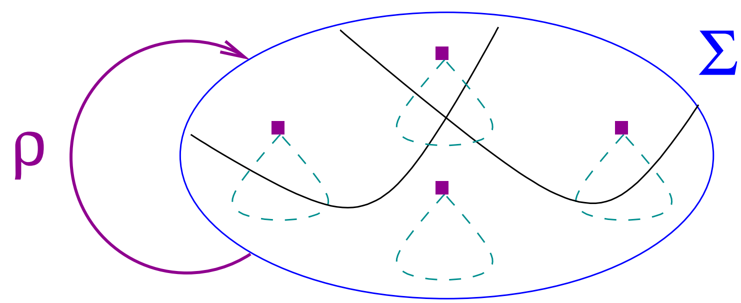
The Scott topology on algebraic lattice D *defines* D itself.

An application to abstract interpretation: Cousots' $\sqcup$-topology for a.i.

[Cousot²78] defines a Scott-like topology for complete lattices, where the basic open sets are up-closed, closed under tails of chains, and closed under binary $\sqcap$. They show equivalence of chain continuity to topological continuity.

Next, they show that an abstract interpretation on Σ, defined by an upper-closure map, $\rho : \Sigma \rightarrow \Sigma$, preserves convergence. That result follows from this key property:

Proposition: The $\sqcup$-topology on $\rho[\Sigma]$ is exactly the *relative $\sqcup$-topology* on Σ, that is, every $V \in \mathcal{O}_{\rho[\Sigma]}$ equals $U \cap \rho[\Sigma]$, for some $U \in \mathcal{O}_\Sigma$.

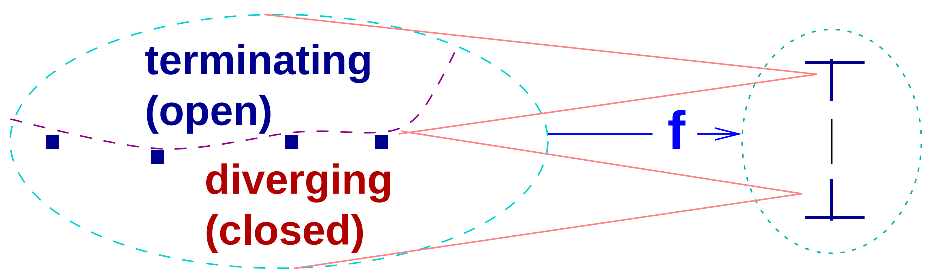


An application: Backwards strictness analysis

$f : D_\perp \to D_\perp$ is *strict*, if $f(\perp) = \perp$. This knowledge can help optimize code for f for lazy functional languages [Mycroft80] .

A backwards strictness analysis approximates $D_\perp$ by $2_D = \{\top, \perp\}$ and and f by $f^\sharp : D_\perp \to 2_D$, computing ${f^\sharp}^{-1}\{\top\}$ for termination information.

[Clack&PeytonJones85] showed how to use a finite set of minimal points (a *frontier*) to represent ${f^\sharp}^{-1}\{\top\}$.



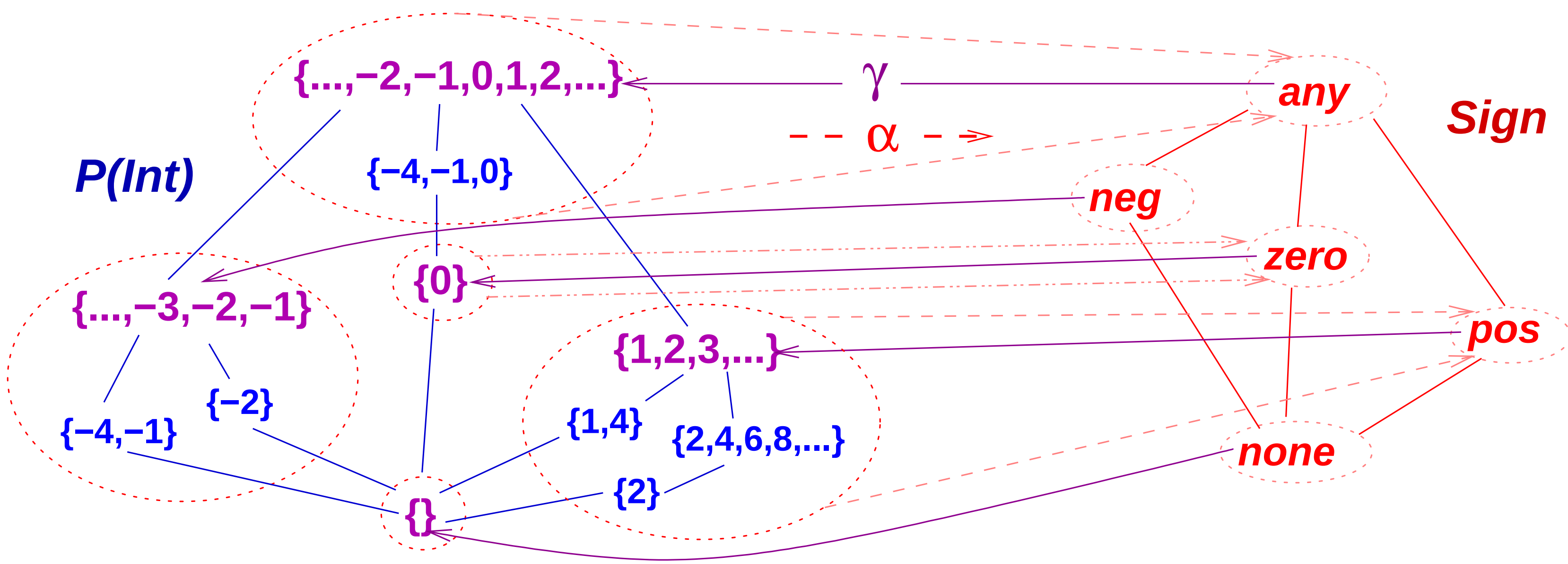
[Hunt89] noted that $f^\sharp$ is Scott-continuous, making ${f^\sharp}^{-1}\{\top\}$ a Scott-open set. Hunt defined frontier-based strictness analysis as calculation of open-set inverse images, showing how to lift efficiently to higher types, in the style of [BurnHankinAbramsky86] .

An application: abstract interpretation in logical form

[Abramsky91] applied frame theory (the axiomatization of the lattice of open sets) to domain theory, generating Scott domains from sets of atomic elements that act as primitive propositions in a domain logic, closing them under a set of frame axioms.

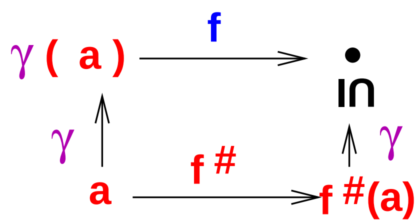
[Jensen92] observed one can use a *finite subset* of a domain's atomic elements to generate an *abstract domain* that approximates the concrete domain generated from *all* the atomic elements. Jensen called his methodology *abstract interpretation in logical form* and applied it to strictness analysis.

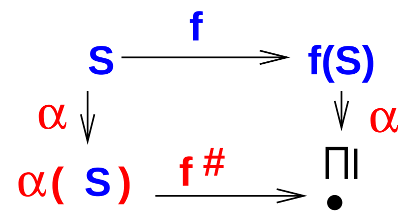
We step back from these applications and ask: *In what naive sense does an abstract domain define a "topology" on the concrete domain that it approximates?* What does it mean for a function to preserve and reflect the "open sets"? Do these notions define forwards and backwards static analyses and do they ensure soundness and completeness of the analyses?



Here, does $\gamma[Sign] = \{\{\}, \{\cdots, -2, -1\}, \{0\}, \{1, 2, \cdots\}, Int\}$ *define a "topology" on* Int*?*

For Galois connection, $\mathcal{P}(\Sigma)\langle\alpha,\gamma\rangle A$, for $f : \mathcal{P}(\Sigma) \rightarrow \mathcal{P}(\Sigma)$, an abstract function, $f^\sharp : A \rightarrow A$, is ***sound*** iff $f \circ \gamma \sqsubseteq \gamma \circ f^\sharp$ iff $\alpha \circ f \sqsubseteq f^\sharp \circ \alpha$:



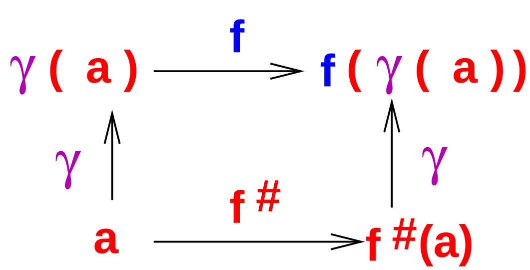


α and γ are semi-homomorphisms; $f^\sharp$ is a postcondition transformer on A; $f_0^\sharp = \alpha \circ f \circ \gamma$ is strongest.

Using $\rho = \gamma \circ \alpha$, we can define $f_0^\sharp = \rho \circ f : \gamma[\Sigma] \rightarrow \gamma[\Sigma]$.

Forwards completeness

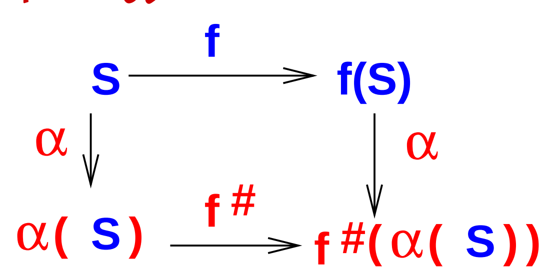
[Giacobazzi01] : $f \circ \gamma = \gamma \circ f^\sharp$



Backwards completeness

[Cousot²79,Giacobazzi00] :

$\alpha \circ f = f^\sharp \circ \alpha$



$f_0^\sharp$ is forwards complete for f iff $f \circ \rho = \rho \circ f \circ \rho$.

$f_0^\sharp$ is backwards complete for f iff $\rho \circ f = \rho \circ f \circ \rho$

A key characterization from *[GiacobazziQuintarelli01]*

The following result is the main result in [14] and it is the basis for a constructive characterization of the complete shell of an abstract domain, viz. the least (most abstract) domain which is $\mathcal{B}$-complete and includes a given domain. This result constructively characterizes the structure of $\mathcal{B}$-complete abstract domains for continuous functions. Recall that, if $f : C \to C$ is a unary function, then $f^{-1}(y) = \{\, x \mid f(x) = y \,\}$.

Theorem 1 ([14]). *Let $f : C \to C$ be continuous and $\rho \in uco(C)$. Then ρ is $\mathcal{B}$-complete for f iff $\bigcup_{y \in \rho(C)} max(f^{-1}(\downarrow y)) \subseteq \rho(C)$.*

This characterization of backwards completeness looks like the inverse-image definition of topological continuity, stated in a kind of frame theory.

Is a.i. completeness the same as topological continuity?

Where the current paper fits into this story...

Topology studies how functions compute on *properties* (open sets). This is exactly what abstract interpretation studies.

We proceed from these first principles: For concrete domain, Σ, abstract domain, A, and concretization map, $\gamma : A \rightarrow \mathcal{P}(\Sigma)$,

1. A defines a *property family*, $\mathcal{F}_\Sigma = \gamma[A] \subseteq \mathcal{P}(\Sigma)$.
2. if $\mathcal{F}_\Sigma$ is closed under intersection, it is a *closed family* (call it $\mathcal{C}_\Sigma$); if $\mathcal{F}_\Sigma$ is closed under union, it is an *open family* (call it $\mathcal{O}_\Sigma$).
3. For $f : \Sigma \rightarrow \Sigma$, we generalize the definition of continuity: f is $\mathcal{F}_\Sigma$-*continuous* iff for all $S \subseteq \Sigma$ and $U \in \mathcal{F}_\Sigma$,
   if $f[S] \subseteq U$, then there exists $V \in \mathcal{F}_\Sigma$ such that
   $S \subseteq V$ and $f[V] \subseteq U$.

We gain these results

1. Generalized continuity retains fundamental properties. In particular, f is continuous iff $f^{-1}[U] \in \mathcal{F}_\Sigma$ whenever $U \in \mathcal{F}_\Sigma$.
2. Closed families generate forwards abstract interpretations with best (strongest postcondition) precision (e.g., constant propagation), and open families generate backwards abstract interpretations with best (weakest precondition) precision (e.g., strictness analysis).
3. [Giacobazzi00 and 01] 's notions of forwards and backwards completeness are characterized as the topologically closed maps and topologically continuous maps upon a closed family. (There are analogous results for an open family).
4. [Smyth83] 's upper and lower topologies for powerdomains $\mathcal{P}(\Sigma)$ generate the abstract interpretations based on $\mathcal{F}_\Sigma$ for abstract-model checking of $\Box$ and $\Diamond$ in branching-time temporal logic.

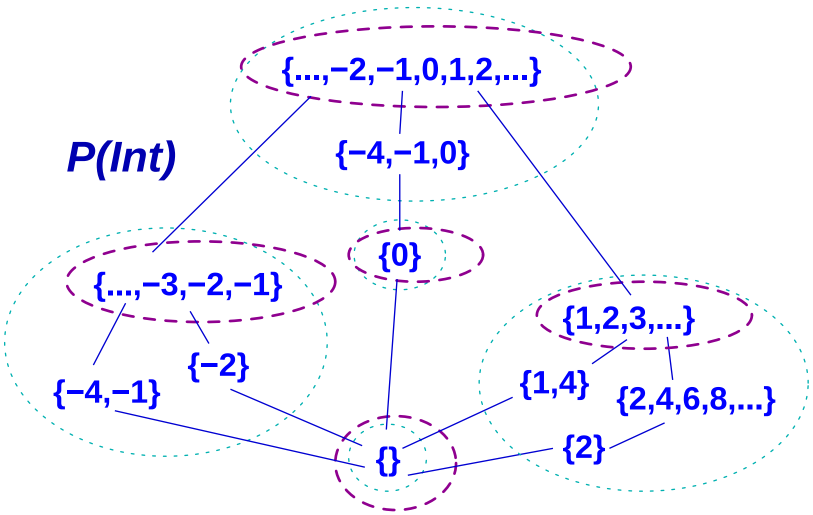
Technical details

Open sets are computable properties [Smyth83]

For an algebraic cpo, D, its Scott-basic-open sets are $\uparrow\mathrm{e}$, for each finite element, $\mathrm{e} \in \mathrm{D}$. *Read* $\mathrm{d} \in \uparrow\mathrm{e}$ *as "*d *has property* $\uparrow\mathrm{e}$*."*

But abstract intepretation is *finite computation on properties*; for an abstract domain, like $Sign$, the computable properties are $\gamma[Sign]$ (or, if you will, $\rho[\mathcal{P}(Sign)]$, where $\rho = \gamma \circ \alpha$).

Alas, $\rho[\mathcal{P}(Sign)]$ is closed under intersections (not necessarily unions). Also, there exist abstract domains A that possess *only* a γ but no α (and no ρ) [Cousot^{2}92] .



Let's weaken some definitions

For abstract domain A and $\gamma : A \rightarrow \mathcal{P}(\Sigma)$, define Σ's *property family* as $\mathcal{F}_\Sigma = \gamma[A]$.

For each $U \in \mathcal{F}_\Sigma$, its complement is $\sim U = \Sigma - U$; for $\mathcal{F}_\Sigma$, its *complement family*, $\sim\mathcal{F}_\Sigma$, is $\{\sim U \mid U \in \mathcal{F}_\Sigma\}$.

$\mathcal{F}_\Sigma$ is an *open family* if it is closed under unions; it has an interior operation, $\iota : \mathcal{P}(\Sigma) \rightarrow \mathcal{F}_\Sigma$. It is a *closed family* if it is closed under intersections; it has a closure operation, $\rho : \mathcal{P}(\Sigma) \rightarrow \mathcal{F}_\Sigma$. If $\mathcal{F}_\Sigma$ is an open family, then its complement is a closed family (and vice versa).

When γ is the upper adjoint of a Galois connection, then $\mathcal{F}_\Sigma$ is a closed family.

$f^\sharp : \mathcal{F}_\Sigma \rightarrow \mathcal{F}_\Sigma$ is *(overapproximating) sound for* $f : \Sigma \rightarrow \Sigma$ if for all $U \in \mathcal{F}_\Sigma$, $f[U] \subseteq f^\sharp[U]$.

When $\mathcal{F}_\Sigma$ is a closed family, $\rho \circ f$ is sound for f.

There are the obvious dual notions for *underapproximating soundness*.

If $\mathcal{C}_\Sigma$ is a closed family, its closure operator, ρ, defines a strongest-postcondition analysis:

For $f : \Sigma \to \Sigma$, define $f^\sharp : \mathcal{C}_\Sigma \to \mathcal{C}_\Sigma$ as $f^\sharp = \rho \circ f$. We have

$$\{\phi\}f\{f^\sharp(\phi)\}$$

holds true (where $\{\phi\}f\{\psi\}$ asserts $f[\phi] \subseteq \psi$, for $\phi, \psi \in \mathcal{C}_\Sigma$).

$f^\sharp(\phi) = \rho(f[\phi])$ defines the strongest postcondition of f and ϕ expressible in $\mathcal{C}_\Sigma$.

If we desire preconditions from a closed family, then we must close it under unions, that is, perform a *disjunctive completion* of the abstract domain — *We use the closed family as a base for a topology* on Σ, namely, $\{\cup T \mid T \subseteq \mathcal{C}_\Sigma\}$, which is both an open *and* a closed family.

If we are truly interested in preconditions, we start with an *open* family of properties (one closed under unions), $\mathcal{O}_\Sigma \subseteq \mathcal{P}(\Sigma)$, so we have an interior operation, $\iota : \mathcal{P}(\Sigma) \to \mathcal{O}_\Sigma$.

We underapproximate the inverses of transition functions: For $f : \Sigma \to \Sigma$, define $f^{-o} : \mathcal{O}_\Sigma \to \mathcal{O}_\Sigma$ as $f^{-o} = \iota \circ f^{-1}$. This implies

$$\{f^{-o}(\psi)\}f\{\psi\}$$

holds true and $f^{-o}(\psi)$ is the *weakest precondition of* f *and* ψ *expressible in* $\mathcal{O}_\Sigma$.

Proposition: For closed family $\mathcal{C}_\Sigma$ and $\mathcal{O}_\Sigma = \sim\mathcal{C}_\Sigma$, $\widetilde{(f^{-1})^\sharp}(U) = f^{-o}(U)$, for all $U \in \mathcal{O}_\Sigma$. (Note: $\widetilde{(f^{-1})^\sharp} = \sim \circ (f^{-1})^\sharp \circ \sim$.)

That is, by using $\mathcal{C}_\Sigma$'s closure operator to define the overapproximating $(f^{-1})^\sharp$, we can compute an *underapproximating*, weakest-precondition analysis on $\mathcal{O}_\Sigma = \sim\mathcal{C}_\Sigma$ defined as $\widetilde{(f^{-1})^\sharp}$.

Property preservation by functions

For $f : \Sigma \to \Sigma$, define $f : \mathcal{P}(\Sigma) \to \mathcal{P}(\Sigma)$ as $f[S] = \{f(s) \mid s \in S\}$, and define $f^{-1} : \mathcal{P}(\Sigma) \to \mathcal{P}(\Sigma)$ as $f^{-1}(T) = \{s \in \Sigma \mid f(s) \in T\}$, as usual.

f is $\mathcal{F}_\Sigma$-*preserving* iff for all $U \in \mathcal{F}_\Sigma$, $f[U] \in \mathcal{F}_\Sigma$. In such a case, $f : \mathcal{F}_\Sigma \to \mathcal{F}_\Sigma$ is well defined.

This generalizes the notions of topologically open and closed maps.

Let $\mathcal{F}_\Sigma$ be a closed family, and let $\rho : \mathcal{P}(\Sigma) \to \mathcal{P}(\Sigma)$ be the associated closure operator.

For $f : \Sigma \to \Sigma$, define $f_0^\sharp : \mathcal{P}(\Sigma) \to \mathcal{P}(\Sigma)$ as $f_0^\sharp = \rho \circ f$, as usual.

Theorem: $f_0^\sharp$ is forwards complete for f iff f is $\mathcal{F}_\Sigma$-preserving, that is, iff f is a topologically closed map.

Property reflection (function continuity)

Let U_c (respectively, U_S) denote a member of $\mathcal{F}_\Sigma$ such that $c \in U_c$ (respectively, $S \subseteq U_S$):

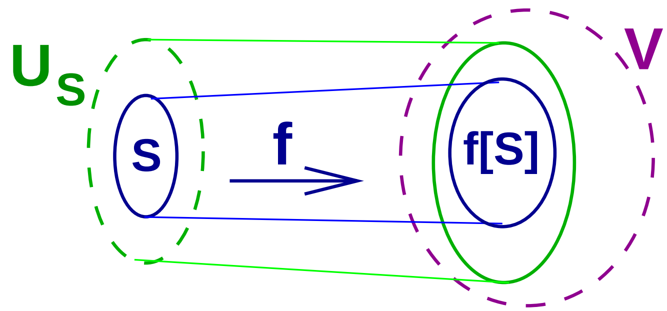
- For $c \in \Sigma$, $f : \Sigma \to \Sigma$ is *continuous at* c iff for all $V_{f(c)} \in \mathcal{F}_\Sigma$, there exists some $U_c \in \mathcal{F}_\Sigma$ such that $f[U_c] \subseteq V_{f(c)}$.
- For $S \subseteq \Sigma$, f is *continuous at* S iff for all $V_{f[S]} \in \mathcal{F}_\Sigma$, there exists some $U_S \in \mathcal{F}_\Sigma$ such that $f[U_S] \subseteq V_{f[S]}$.
- f is $\mathcal{F}_\Sigma$*-reflecting* iff for all $V \in \mathcal{F}_\Sigma$, $f^{-1}(V) \in \mathcal{F}_\Sigma$, that is, f^{-1} is $\mathcal{F}_\Sigma$-preserving.

The second item is needed because $\mathcal{F}_\Sigma$ might not be an open family.

If $\mathcal{F}_\Sigma$ is a topology, then all three notions are equivalent.

reflection, cont.

f is continuous at $S \subseteq \Sigma$:



If $f[S] \subseteq V \in \mathcal{F}_\Sigma$, then there exists $U_S \in \mathcal{F}_\Sigma$ such that $f[U_S] \subseteq V$.

Proposition:

1. f is $\mathcal{F}_\Sigma$-reflecting iff f is continuous at S, for all $S \subseteq \Sigma$.
2. If $\mathcal{F}_\Sigma$ is an open family, then f is $\mathcal{F}_\Sigma$-reflecting iff f is continuous at c, for all $c \in \Sigma$.
3. $f : \Sigma \to \Sigma$ is $\sim\mathcal{F}_\Sigma$-reflecting iff f is $\mathcal{F}_\Sigma$-reflecting.

reflection, concl.

For $S, S' \subseteq \Sigma$, write $S \leq_{\mathcal{F}_\Sigma} S'$ iff for all $K \in \mathcal{F}_\Sigma, S \subseteq K$ implies $S' \subseteq K$. Write $S \equiv_{\mathcal{F}_\Sigma} S'$ iff $S \leq_{\mathcal{F}_\Sigma} S'$ and $S' \leq_{\mathcal{F}_\Sigma} S$. That is, S and S' share the same properties.

Definition: $f : \Sigma \to \Sigma$ is *backwards-$\mathcal{F}_\Sigma$-complete* iff for all $S, S' \subseteq \Sigma$, $S \equiv_{\mathcal{F}_\Sigma} S'$ implies $f[S] \equiv_{\mathcal{F}_C} f[S']$ cf. Slide 12.

Proposition: If f is $\mathcal{F}_\Sigma$-reflecting, then it is backwards-$\mathcal{F}_\Sigma$-complete.

Lemma: If $\mathcal{F}_\Sigma$ is a closed family, then TFAE:
(i) f is backwards-$\mathcal{F}_\Sigma$-complete;
(ii) for all $S \subseteq \Sigma$, $f[S] \equiv_{\mathcal{F}_\Sigma} f[\rho(S)]$;
(iii) $\rho \circ f = \rho \circ f \circ \rho$

Theorem: For closed family, $\mathcal{F}_\Sigma$, f is backwards-$\mathcal{F}_\Sigma$-complete iff it is $\mathcal{F}_\Sigma$-reflecting.

So, abstract-interpretation backwards completeness is topological continuity.

What about open families?

Let $\mathcal{F}_\Sigma$ be open (closed under unions) and $\iota : \mathcal{P}(\Sigma) \to \mathcal{F}_\Sigma$ be its interior map.

We use an open family to perform an underapproximating *precondition analysis*: for $f : \Sigma \to \Sigma$, define $f^{-1} : \mathcal{P}(\Sigma) \to \mathcal{P}(\Sigma)$ as $f^{-1}(S) = \{s \in \Sigma \mid f(s) \in S\}$, as usual.

The strongest (*weakest precondition*) abstract function for f^{-1} is $\iota \circ f^{-1} : \mathcal{F}_\Sigma \to \mathcal{F}_\Sigma$.

Define *forwards-$\mathcal{F}_\Sigma$-completeness:* $f^{-1} \circ \iota = \iota \circ f^{-1} \circ \iota$
backwards-$\mathcal{F}_\Sigma$-completeness: $\iota \circ f^{-1} = \iota \circ f^{-1} \circ \iota$

f^{-1} is $\mathcal{F}_\Sigma$-preserving iff f^{-1} is forwards-$\mathcal{F}_\Sigma$-complete iff f is $\sim\mathcal{F}_\Sigma$-reflecting iff f is $\mathcal{F}_\Sigma$-reflecting.

This is the classic pre- post-condition duality of predicate transformers.

Powerdomains

[Smyth83] showed, for algebraic Scott-domain D, that the lower powerdomain, $\mathcal{P}_L(D)$, and the upper powerdomain, $\mathcal{P}_U(D)$, are generated from D's Scott-topology, $\mathcal{O}_D$, as follows:

- *lower topology,* $\mathcal{O}^L_{\mathcal{O}_D}$: generated from the base $\mathcal{B}^L_{\mathcal{O}_D} = \{\exists U \mid U \in \mathcal{O}_D\}$, where $\exists U = \{S \subseteq D \mid S \cap U \neq \emptyset\}$ ("all sets that meet U")
- *upper topology,* $\mathcal{O}^U_{\mathcal{O}_D}$: generated from the base $\mathcal{B}^U_{\mathcal{O}_D} = \{\forall U \mid U \in \mathcal{O}_D\}$, where $\forall U = \{S \subseteq D \mid S \subseteq U\}$ ("all sets covered by U")

We can show that when Σ is abstractly interpreted by closed family, $\mathcal{C}_\Sigma$, the abstract interpretation for ACTL checking that proves the most $\mathcal{C}_\Sigma$-properties is generated from the (co)base $\mathcal{B}^U_{\mathcal{C}_\Sigma}$.

Similarly, the best abstract interpretation for ECTL checking is generated from the (co)base $\mathcal{B}^L_{\mathcal{C}_\Sigma}$.

For $f : \Sigma \to \mathcal{P}(\Sigma)$, there are two preimage maps:

1. $\widetilde{pre}_f(S) = \{c \in \Sigma \mid f(c) \subseteq S\}$
2. $pre_f(S) = \{c \in \Sigma \mid f(c) \cap S \neq \emptyset\}$

Abstract model checking [CleavelandIyerYankelevich95, DamsGerthGrumberg97] starts from $\mathcal{C}_\Sigma$ to generate an a.i. for $\mathcal{P}(\Sigma)$:

Let $\mathcal{C}^{U}_{\mathcal{C}_\Sigma}$ be the closed family generated from $\mathcal{B}^{U}_{\mathcal{C}_\Sigma}$:

Theorem: For $\mathcal{C}^{U}_{\mathcal{C}_\Sigma}$, $\widetilde{pre}_f$ is a $\mathcal{C}_\Sigma$-preserving map iff f is $\mathcal{C}_\Sigma\mathcal{C}^{U}_{\mathcal{C}_\Sigma}$-reflecting.

That is, $[f]$ can be precisely model checked exactly when f is $\mathcal{C}_\Sigma\mathcal{C}^{U}_{\mathcal{C}_\Sigma}$-continuous.

Let $\mathcal{C}^{L}_{\mathcal{C}_\Sigma}$ be the closed family generated from $\mathcal{B}^{L}_{\mathcal{C}_\Sigma}$:

Theorem: For $\mathcal{C}^{L}_{\mathcal{C}_\Sigma}$, pre_f is a $\mathcal{C}_\Sigma$-preserving map iff f is $\mathcal{C}_\Sigma\mathcal{C}^{L}_{\mathcal{C}_\Sigma}$-reflecting.

That is, $\langle f \rangle$ can be precisely model checked exactly when f is $\mathcal{C}_\Sigma\mathcal{C}^{L}_{\mathcal{C}_\Sigma}$-continuous.

This is the origin of Dams's mixed Kripke structures.

Conclusion

...many key notions and theorems from abstract interpretation theory appear as definitions and corollaries of “naive topology”...

References **This talk:** `www.cis.ksu.edu/~schmidt/papers`

1. S. Abramsky. Domain theory in logical form. *Ann.Pure Appl.Logic* (51) 1991.
2. P. Cousot and R. Cousot. Static determination of dynamic properties of recursive procedures. In *Formal Description of Programming Concepts*, North-Holland 1978.
3. R. Giacobazzi, F. Ranzato, and F. Scozzari. Making abstract interpretations complete. *J. ACM* (47) 2000.
4. R. Giacobazzi and E. Quintarelli. Incompleteness, counterexamples, and refinements in abstract model checking. Proc. SAS'01, LNCS 2126, 2001.
5. S. Hunt. Frontiers and open sets in abstract intepretation. Proc. ACM Conf. Functional Prog. and Comp. Architecture, 1989.
6. T. Jensen. Abstract Interpretation in Logical Form. PhD Thesis, Imperial College, 1992.
7. M.B. Smyth. Powerdomains and predicate transformers: a topological view. Proc. ICALP'83, LNCS 154, 1983.