Internal and External Logics of Abstract Interpretations

David Schmidt

Kansas State University

`www.cis.ksu.edu/~schmidt`

Motivations

How does a static analysis "connect" to the properties it is meant to prove?

- use data-flow analysis to compute available-expression sets to decide register allocation;
- use a state-space exploration to model check a temporal-logic safety property or program-transformation criterion
- apply predicate abstraction with counter-example-guided refinement (CEGAR) to generate an assertion set that proves a safety property

The value domain used by an analysis and the logic used for validation/transformation should be *one and the same* — the logic is *internal* to the value domain. If the values and logic differ, then the logic must be defined *externally*.

Developments from this paper

Let Σ be the program's state set; let A be the abstract domain; let $\gamma : A \rightarrow \mathcal{P}(\Sigma)$ be the *concretization function*.

1. γ defines a logic *internal* to A for Σ, where A's elements act *both* as computational values and as logical assertions. The model theory, $\models$, is defined by γ; the proof theory, $\vdash$, by $\sqsubseteq_A$.
2. The notion of (forwards) completeness from abstract interpretation theory *characterizes* the internal logic.
3. When a logic for Σ is proposed independently from γ, then an *external logic* must be fashioned from $\mathcal{P}_{\downarrow}(A)$. *But,* when γ preserves meets and joins, the external logic can be embedded within A^{op} (inverted).

In the last case, A has *two* interpretations: an overapproximating, *computational* interpretation, and an underapproximating *logical* interpretation (on A^{op}).

Abstract interpretation: computing on properties

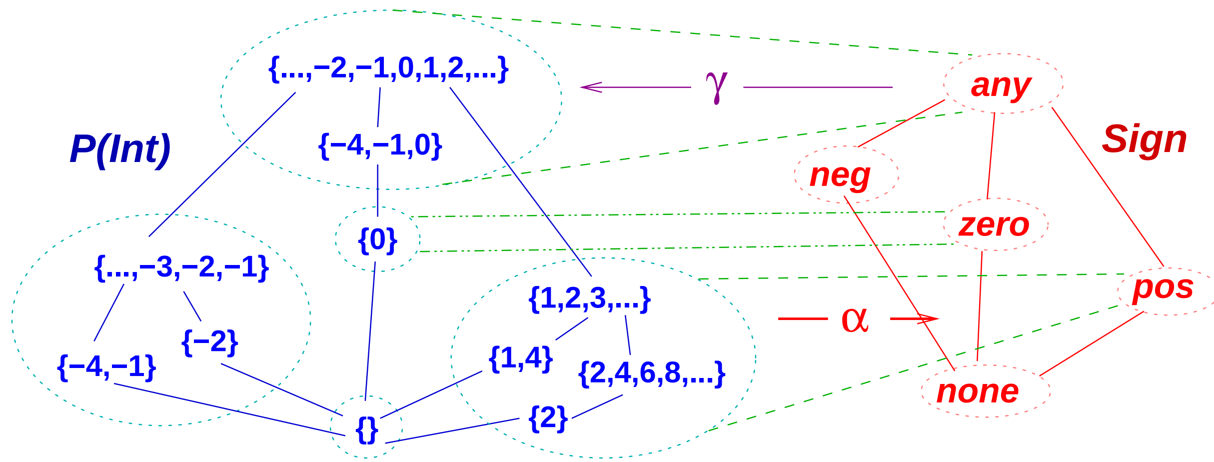
Example:

```
read(x)
if isPositive(x) :
    x:= pred(x)
x:= succ(x)
write(x)
```

Q: Is output pos?

A: abstractly interpret Int by

$Sign = \{neg, zero, pos, any, none\}$:



Standard, collecting interpretation:

$f : \mathcal{P}(Int) \to \mathcal{P}(Int)$:

$\mathrm{isPos}(S) = \{n \in S \mid n > 0\}$

$\mathrm{pred}(S) = \{n - 1 \mid n \in S\}$

$\mathrm{succ}(S) = \{n + 1 \mid n \in S\}$

Abstract interpretation: $f^\sharp : A \to A$:

$\mathrm{isPos}^\sharp(pos) = pos$

$\mathrm{isPos}^\sharp(neg) = none$

$\mathrm{isPos}^\sharp(any) = pos$, etc.

$\mathrm{succ}^\sharp(pos) = pos$

$\mathrm{succ}^\sharp(zero) = pos$

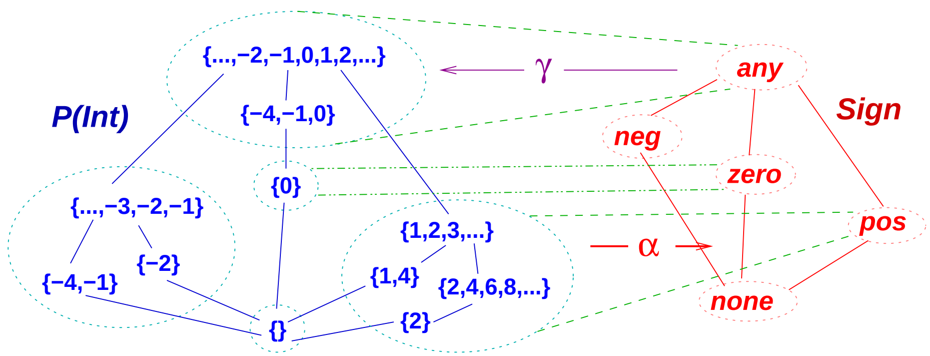
$\mathrm{succ}^\sharp(neg) = any$, etc.

$\mathrm{pred}^\sharp(neg) = neg$

$\mathrm{pred}^\sharp(zero) = neg$

$\mathrm{pred}^\sharp(pos) = any$, etc.

Abstract values = logical properties



Read computational values like $neg \in Sign$ as logical propositions, “isNegative”, etc.

For $S \subseteq \Sigma$, $a, a' \in A$, $\gamma : A \to \mathcal{P}(\Sigma)$, define

- $S \models a$ iff $S \subseteq \gamma(a)$ e.g., $\{-3, -1\} \models neg$
- $a \models a'$ iff $\gamma(a) \subseteq \gamma(a')$ e.g., $neg \models any$
- $a \vdash a'$ iff $a \sqsubseteq a'$ e.g., $neg \vdash any$

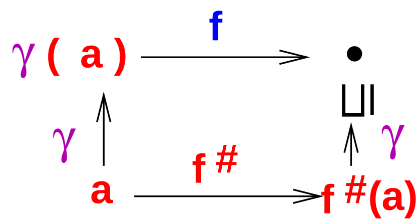
Proposition: *(soundness)* $a \vdash a'$ implies $a \models a'$.

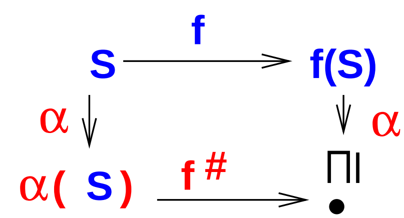
Proposition: *(completeness)* if γ is an upper adjoint of a Galois connection and is 1-1, then $a \models a'$ implies $a \vdash a'$.

Abstract transformers compute on properties

For $\mathsf{f} : \mathrm{PC} \to \mathrm{PC}$, $\mathsf{f}^\sharp : A \to A$ is *sound* iff

$$\mathsf{f} \circ \gamma \sqsubseteq \gamma \circ \mathsf{f}^\sharp \quad \text{iff} \quad \alpha \circ \mathsf{f} \sqsubseteq \mathsf{f}^\sharp \circ \alpha$$





This makes $\mathsf{f}^\sharp$ a *postcondition transformer*:

Proposition: *(soundness)* $S \models a$ implies $\mathsf{f}(S) \models \mathsf{f}^\sharp(a)$.

Example: For, $succ : \mathcal{P}(\mathrm{Int}) \to \mathcal{P}(\mathrm{Int})$, we have $succ\{0\} = \{1\}$, which is soundly mimicked by $succ^\sharp(zero) = pos$.

$\mathsf{f}^\sharp_{\mathrm{best}} = \alpha \circ \mathsf{f} \circ \gamma$ is the *strongest postcondition* transformer for A.

Definition: $\mathsf{f}^\sharp$ is γ*-complete (forwards complete) for* f iff $\mathsf{f} \circ \gamma = \gamma \circ \mathsf{f}^\sharp$ [Giacobazzi01] . $\mathsf{f}^\sharp$ is α*-complete (backwards complete) for* f iff $\alpha \circ \mathsf{f} = \mathsf{f}^\sharp \circ \alpha$ [Cousots00] .

A *has an* internal logic *that* γ *preserves*

First, treat all $a \in A$ as primitive propositions (*isNeg*, *isPos*, etc.).

A *has conjunction* when

$$S \models \phi_1 \sqcap \phi_2 \text{ iff } S \models \phi_1 \text{ and } S \models \phi_2, \text{ for all } S \subseteq \Sigma.$$

That is, $\gamma(\phi \sqcap \psi) = \gamma(\phi) \cap \gamma(\psi)$, for all $\phi, \psi \in A$.

Proposition: When $\gamma : A \to \mathcal{P}(\Sigma)$ is an upper adjoint, then A has conjunction.

$Sign$ has conjunction; so do all predicate-abstraction analyses.

Proposition: When $\gamma(\phi \sqcup \psi) = \gamma(\phi) \cup \gamma(\psi)$, then A has *disjunction*: $S \models \phi \sqcup \psi$ iff $S \models \phi$ or $S \models \psi$.

$Sign$ lacks disjunction: $zero \models neg \sqcup pos$ (because $neg \sqcup pos = any$ but $zero \not\models neg$ and $zero \not\models pos$).

Complete lattice A is *distributive* if $a \sqcap (b \sqcup c) = (a \sqcap b) \sqcup (a \sqcap c)$, for all $a, b, c \in A$. When $\sqcap$ is Scott-continuous, then

$$\phi \Rightarrow \psi \;\equiv\; \bigsqcup\{a \in A \mid a \sqcap \phi \sqsubseteq \psi\}$$

satisfies the property, $a \vdash \phi \Rightarrow \psi$ iff $a \sqcap \phi \vdash \psi$.

Proposition: If A is a distributive complete lattice, $\sqcap$ is Scott-continuous, and upper adjoint γ is 1-1, then A has *Heyting implication*, $\phi \Rightarrow \psi$, such that

$$S \models \phi \Rightarrow \psi \text{ iff } \gamma(\alpha(S)) \cap \gamma(\phi) \subseteq \gamma(\psi).$$

That is, $\gamma(\phi \Rightarrow \psi) = \bigcup\{S \in \gamma[A] \mid S \cap \gamma(\phi) \subseteq \gamma(\psi)\}$.

Heyting implication is weaker than classical implication, where $S \models \phi \Rightarrow \psi$ iff $S \cap \gamma(\phi) \subseteq \gamma(\psi)$ iff for all $c \in S$, if $\{c\} \models \phi$, then $\{c\} \models \psi$.

The POS domain for groundness analysis of logic programs uses Heyting implication [Cortesi91,Marriott93] .

If $\gamma(\bot_A) = \emptyset \in \mathcal{P}(\Sigma)$, we have falsity ($\bot$); this yields the logic,

$$\phi \ ::= \ a \mid \phi_1 \sqcap \phi_2 \mid \phi_1 \sqcup \phi_2 \mid \phi_1 \Rightarrow \phi_2 \mid \bot$$

In particular, $\neg\phi$ abbreviates $\phi \Rightarrow \bot$ and defines the *refutation* of ϕ within A, as done in TVLA [Sagiv02] .

$\gamma : A \rightarrow \mathcal{P}(\Sigma)$ is the interpretation function for the internal logic:

$$\gamma(a) = \text{given}$$

$$\gamma(\phi \sqcap \psi) = \gamma(\phi) \cap \gamma(\psi)$$

$$\gamma(\phi \sqcup \psi) = \gamma(\phi) \cup \gamma(\psi)$$

$$\gamma(\phi \Rightarrow \psi) = \bigcup\{S \in \gamma[A] \mid S \cap \gamma(\phi) \subseteq \gamma(\psi)\}$$

$$\gamma(\bot) = \emptyset$$

γ-completeness characterizes the internal logic

The previous interpretation, e.g., for conjunction:

$$\gamma(\phi \sqcap \psi) = \gamma(\phi) \cap \gamma(\psi)$$

shows that γ-completeness is *exactly* the criterion for determining which connectives are embedded in A's internal logic:

Proposition: For $f : \mathcal{P}(\Sigma) \times \mathcal{P}(\Sigma) \times \cdots \to \mathcal{P}(\Sigma)$, A has connective $f^\sharp$ iff $f^\sharp$ is γ-complete for f:

$$\gamma(f^\sharp(\phi_1, \phi_2, \cdots)) = f(\gamma(\phi_1), \gamma(\phi_2), \cdots).$$

Example: For $Sign$, $\mathtt{negate}^\sharp$ is γ-complete for $\mathtt{negate}(S) = \{-n \mid n \in S\}$ (where $\mathtt{negate}^\sharp(pos) = neg$, $\mathtt{negate}^\sharp(neg) = pos$, etc.):

$$\phi ::= a \mid \phi_1 \sqcap \phi_2 \mid \mathtt{negate}^\sharp(\phi)$$

We can state "negate" assertions, e.g., $pos \models \mathtt{negate}^\sharp(neg \sqcap any)$.

Transition functions in the logic: predicate transformers

It is useful to know when $f(S) \models \phi$, *that is,* $S \models [f]\phi$.

Define $[\cdot]$ in terms of $\widetilde{pre}$:

$$[f](S) = \widetilde{pre}_f(S) = \bigcup\{S' \in \Sigma \mid f(S') \subseteq S\}, \qquad \text{for } S \subseteq \mathcal{P}(\Sigma)$$
$$[f^\sharp](a) = \widetilde{pre}_{f^\sharp}(a) = \{a' \in A \mid f^\sharp(a') \sqsubseteq a\}, \qquad \text{for } a \in A$$

When $f^\sharp$ is sound for f, then $\widetilde{pre}_{f^\sharp}$ is sound for $\widetilde{pre}_f$.

Proposition: Assume that γ is an upper adjoint and preserves joins. Then, $\widetilde{pre}_{f^\sharp}$ is γ*-complete* for $\widetilde{pre}_f$ iff $f^\sharp$ is α*-complete* for f.

In this case, we have that $[f^\sharp]$ exists in A's internal logic, where

$$\gamma([f^\sharp]\phi) = \widetilde{pre}_f(\gamma(\phi))$$

But, it is not so common that $f^\sharp$ is α-complete for f.

Logics not internal to the abstract domain

It is common to work with a logic "external" from A's internal logic (e.g., because the transition functions, $f^\sharp : A \to A$, lack α-completeness).

Example: We want this logic for reasoning about $Sign$:

$\phi ::= a \mid \phi_1 \wedge \phi_2 \mid \phi_1 \vee \phi_2 \mid [f]\phi$ for $a \in Sign$ and $f \in \{\mathrm{succ}, \mathrm{pred}\}$

where $[\![\cdot]\!] : \mathcal{L} \to \mathcal{P}(\Sigma)$ is defined

$$[\![a]\!] = \gamma(a) \qquad [\![\phi_1 \wedge \phi_2]\!] = [\![\phi_1]\!] \cap [\![\phi_2]\!]$$
$$[\![[f]\phi]\!] = \widetilde{pre}_f[\![\phi]\!] \qquad [\![\phi_1 \vee \phi_2]\!] = [\![\phi_1]\!] \cup [\![\phi_2]\!]$$

But this logic is not internal to $Sign$ (which lacks disjunction, and both $\mathrm{succ}^\sharp$ and $\mathrm{pred}^\sharp$ are not α-complete). *What do we do?*

We can fashion an **external logic**

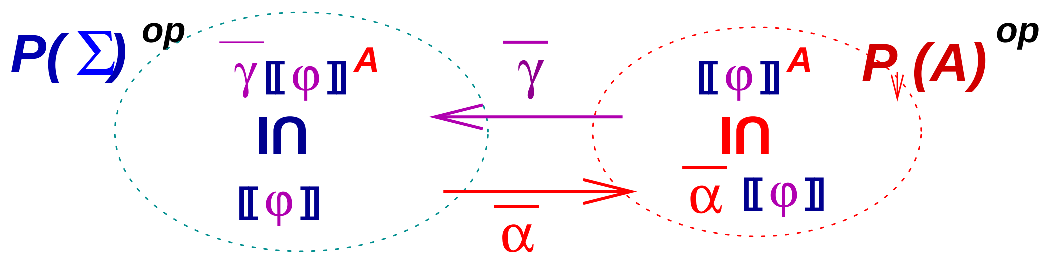
For each $[\![\phi]\!] \subseteq \Sigma$, define

$$[\![\phi]\!]^{\mathcal{A}} = \{a \in A \mid \gamma(a) \subseteq [\![\phi]\!]\}$$

Then, assert $a \vdash \phi$ iff $a \in [\![\phi]\!]^{\mathcal{A}}$.

This definition follows from a Galois connection whose abstract domain is $\mathcal{P}_{\downarrow}(A)^{\mathrm{op}}$ — downclosed subsets of A, ordered by superset:

$$\overline{\gamma}(T) = \bigcup\{\gamma(a) \mid a \in T\}$$
$$\overline{\alpha}(S) = \{a \mid \gamma(a) \subseteq S\}$$



That is, $[\![\phi]\!]^{A} = \overline{\alpha}[\![\phi]\!]$.

The inverted ordering gives *underapproximation*: $[\![\phi]\!] \supseteq \overline{\gamma}([\![\phi]\!]^{\mathcal{A}})$. This form of external logic is standard in "abstract model checking."

It is also standard to write an inductively defined approximation to $\overline{\alpha}[\![\phi]\!]$:

$$
\begin{aligned}
[\![a]\!]^{\mathcal{A}}_{\mathrm{ind}} &= \overline{\alpha}(\gamma(a)) \\
[\![\phi_1 \wedge \phi_2]\!]^{\mathcal{A}}_{\mathrm{ind}} &= [\![\phi_1]\!]^{\mathcal{A}}_{\mathrm{ind}} \cap [\![\phi_2]\!]^{\mathcal{A}}_{\mathrm{ind}} \\
[\![\phi_1 \vee \phi_2]\!]^{\mathcal{A}}_{\mathrm{ind}} &= [\![\phi_1]\!]^{\mathcal{A}}_{\mathrm{ind}} \cup [\![\phi_2]\!]^{\mathcal{A}}_{\mathrm{ind}} \\
[\![[f]\phi]\!]^{\mathcal{A}}_{\mathrm{ind}} &= \widetilde{pre}_{f^\sharp}[\![\phi]\!]^{\mathcal{A}}_{\mathrm{ind}} = \{a \in A \mid f^\sharp(a) \in [\![\phi]\!]^{\mathcal{A}}_{\mathrm{ind}}\}
\end{aligned}
$$

Entailment and provability are as expected: $a \models \phi$ iff $\gamma(a) \subseteq [\![\phi]\!]$, and $a \vdash \phi$ iff $a \in [\![\phi]\!]^{\mathcal{A}}_{\mathrm{ind}}$.

Soundness ($\vdash$ implies $\models$) is immediate, and completeness ($\models$ implies $\vdash$) follows when $\overline{\alpha} \circ [\![\,\cdot\,]\!] = [\![\,\cdot\,]\!]^{\mathcal{A}}_{\mathrm{ind}}$. This is called *logical best preservation* or *logical $\overline{\alpha}$-completeness* [Cousots00,Schmidt06] .

Embedding *the external logic within the internal*

Say that γ is an upper adjoint and that it *preserves joins*, that is,

$$\overline{\gamma}T = \bigcup_{a \in T} \gamma(a) = \gamma(\bigsqcup_{a \in T} a) = \gamma(\sqcup T)$$

So, $\overline{\gamma}[\mathcal{P}_{\downarrow}(A)] = \gamma[A]$ — their ranges are equal — and *there is no new expressivity gained by using sets of A-elements.*

Proposition: If A is a complete lattice and $\gamma : A \longrightarrow \mathcal{P}(\Sigma)$ preserves joins (as unions) *and* meets (as intersections), then

- γ is the upper adjoint of an *overapproximating* Galois connection between $(\mathcal{P}(\Sigma), \subseteq)$ and $(A, \sqsubseteq)$, where $\alpha_o(S) = \sqcap\{a \mid S \subseteq \gamma(a)\}$.

- γ is the upper adjoint of an *underapproximating* Galois connection between $(\mathcal{P}(\Sigma), \supseteq)$ and $(A, \sqsupseteq)$, where $\alpha_u(S) = \sqcup\{a \mid S \supseteq \gamma(a)\}$.

1. For state-transition functions, $f : \mathcal{P}(\Sigma) \rightarrow \mathcal{P}(\Sigma)$, apply the first Galois connection, giving the *computational interpretation* of f: $f^{\sharp}_{best} = \alpha_o \circ f \circ \gamma$.

2. For logical connectives, $[\![\mathtt{f}(\phi_1, \phi_2, \cdots)]\!] = f([\![\phi_1]\!], [\![\phi_2]\!], \cdots)$, apply the second Galois connection, for the *logical interpretation* of f: $f^{\flat}_{best} = \alpha_u \circ f \circ (\gamma \times \gamma \times ...)$ and $[\![\mathtt{f}(\phi_1, \phi_2, \cdots)]\!]^{\mathcal{A}}_{ind} = f^{\flat}_{best}([\![\phi_1]\!]^{\mathcal{A}}_{ind}, [\![\phi_2]\!]^{\mathcal{A}}_{ind}, \cdots)$,

When the $f^{\flat}_{best}$s are α_u-complete, then $[\![\cdot]\!]^{\mathcal{A}}_{ind} = \alpha_u \circ [\![\cdot]\!]$, and the resulting internal logic *proves the same assertions* as the external logic:

- First, for $[\![\phi]\!]^{\mathcal{A}} = \overline{\alpha}[\![\phi]\!] \in \mathcal{P}_{\downarrow}(A)$, recall that $a \vdash \phi$ iff $a \in [\![\phi]\!]^{\mathcal{A}}$.
- Next, for $[\![\phi]\!]^{\mathcal{A}}_{ind} = \alpha_u[\![\phi]\!] \in A$, define $a \vdash \phi$ iff $a \sqsubseteq [\![\phi]\!]^{\mathcal{A}}_{ind}$.

Theorem: For all $a \in A$, $a \in [\![\phi]\!]^{\mathcal{A}}$ iff $a \sqsubseteq [\![\phi]\!]^{\mathcal{A}}_{ind}$.

Conclusions

- A static analysis is "logical" in that it *computes proofs* (via $\sqsubseteq$) in the abstract domain, A, that are *sound* (via $\models$, i.e., γ) in the concrete domain, Σ.

- γ-*completeness* (homomorphism property) characterizes the *internal logic* one can soundly validate on A-values, using $\sqsubseteq$. Assertions not in the internal logic can be approximated within an *external* logic defined with sets of A-values and checked using $\in$.

- When γ preserves joins and meets from A to $\mathcal{P}(\Sigma)$, the external logic can be *embedded* within the abstract domain, letting it *overapproximate* computations on Σ and *underapproximate* assertions on Σ.

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

1. A. Cortesi, G. Filé and W. Winsborough. Prop revisited: propositional formulas as an abstract domain for groundness analysis. LICS'91.

2. P. Cousot. Semantic foundations of program analysis. In *Program Flow Analysis*, S. Muchnick and N. Jones, eds. Prentice-Hall 1981.

3. P. Cousot and R. Cousot. Temporal abstract interpretation. POPL'00.

4. R. Giacobazzi and E. Quintarelli. Incompleteness, counterexamples, and refinements in abstract model checking. SAS'01, LNCS 2126.

5. K. Marriott and H. Sondergaard. Precise and efficient groundness analysis for logic programs. *ACM LOPLAS* 2 (1993).

6. F. Ranzato and F. Tapparo. Strong preservation of temporal fixpoint-based operators. VMCAI'06, LNCS 3855.

7. M. Sagiv, T. Reps, and R. Wilhelm. Parametric Shape Analysis via 3-Valued Logic. *ACM TOPLAS* (24) 2002.

8. D.A. Schmidt. Comparing completeness properties of static analyses and their logics. APLAS'06, LNCS 4279.