Widening and Narrowing Approaches to Abstract Interpretation

Jaeho Shin

netj@ropas.snu.ac.kr

Programming Research Lab, Seoul National University

"Most of the examples and words here are from P. Cousot and R. Cousot’s work [1]."
Today’s Goal

- To understand the concept of widening and narrowing
- To see that the use of infinite abstract domains with widenings and narrowings is powerful than the Galois connection approaches with limited domains (finite or ones that satisfy chain condition)
- To take a look at several ideas/techniques for designing widening and narrowing operators
Contents

- Abstract Interpretation
- Galois Connection Approach
- Use of Infinite Abstract Domain
- Widening/Narrowing
- Comparison of Power
- Design of Widening/Narrowing
- Conclusion
Program Semantics

- Program semantics can be specified as \( \text{lfp}_{\bot_0}(F) \).

\[
\text{CPO} \quad L(\sqsubseteq, \sqcup) \\
\text{operator} \quad F \in L \xrightarrow{\text{con}} L \\
\bot_0 \sqsubseteq F(\bot_0), \bot_0 \in L
\]

- By Kleene’s fixpoint theorem,

\[
\text{lfp}_{\bot_0}(F) = \bigsqcup_{n \in \mathbb{N}} F^n(\bot_0)
\]
Abstract Interpretation

- "Abstract interpretation is formalized as an effective computation of an upper approximation $A$ of the program semantics"

- Approximation $A$ is sound in the sense that $\text{lfp}_{\perp_0}(F) \subseteq A$. 
Example 1

Let the collecting semantics of program $P$:

```plaintext
program P :
    var I : integer ;
begin
    I := 1;
    while I <= 100 do
        begin
            { I ∈ [ 1, 100 ] }
            I := I + 1;
        end;
    { I = 100 + 1 }
end;
```

be the set of possible values of integer variable $I$ when starting execution of the loop body.
Example 1 (cont’d)

Then, it is the least fixed point

$$\text{lfp}_\phi(F) = \{i \in \mathbb{Z} \mid 1 \leq i \leq 100\}$$

of:

$$F = \lambda X.(\{1\} \cup \{i + 1 \mid i \in X\}) \cap \{i \in \mathbb{Z} \mid i \leq 100\}$$

on the complete lattice $$L = \wp(\mathbb{Z})(\subseteq, \phi, \mathbb{Z}, \cap, \cup)$$.

$$A = \{i \in \mathbb{Z} \mid i \geq 0\}$$

is a sound upper approximation.
Galois Connection

**Definition**  If $L(\subseteq)$ and $\overline{L}(\subseteq)$ are posets, then $\langle \alpha, \gamma \rangle$ is a Galois connection, written $L \xleftarrow[\alpha]{} \overline{L}$, iff $\alpha \in L \xhookrightarrow{} \overline{L}$ and $\gamma \in \overline{L} \xhookrightarrow{} L$ such that:

$$\forall x \in L, \overline{y} \in \overline{L} : (\alpha(x) \subseteq \overline{y}) \iff (x \subseteq \gamma(\overline{y})).$$
Example 2

\( \phi(\mathbb{Z}) \) ordered by \( \subseteq \) is approximated by the lattice of intervals

\[ L = \{ \perp \} \cup \{ [l, u] \mid l \in \mathbb{Z} \cup \{ -\infty \} \land u \in \mathbb{Z} \cup \{ +\infty \} \land l \leq u \} \]

ordered by \( \sqsubseteq \) such that:

\[
\begin{align*}
\perp \sqsubseteq [l, u] & \overset{\text{def}}{=} \text{true} \\
[l_0, u_0] \sqsubseteq [l_1, u_1] & \overset{\text{def}}{=} l_1 \leq l_0 \leq u_0 \leq u_1
\end{align*}
\]

Galois connection is defined by:

\[
\begin{align*}
\gamma(\perp) & = \phi \\
\gamma([l, u]) & = \{ x \in \mathbb{Z} \mid l \leq x \leq u \} \\
\alpha(\phi) & = \perp \\
\alpha(X) & = [\min X, \max X]
\end{align*}
\]
Galois Connection Approach

**Definition** \( \langle \bar{L}, \bot_0, \bar{F} \rangle \) is an abstract interpretation of \( \langle L, \bot_0, F \rangle \), written \( \langle L, \bot_0, F \rangle \xleftarrow{\alpha} \langle \bar{L}, \bot_0, \bar{F} \rangle \), iff \( L \xleftarrow{\alpha} \bar{L} \), \( \alpha(\bot_0) \sqsubseteq \bot_0 \) and \( \alpha \circ F \sqsubseteq \bar{F} \circ \alpha \).

**Principle** \( X = F(X) \) can be simplified into \( \bar{X} = \bar{F}(\bar{X}) \), and then solved iteratively starting from \( \bot_0 \).

**Restrictions** To ensure finite convergence of \( \bar{F}^n(\bot_0) \), (a) \( \bar{L} \) must be finite, or (b) \( \{ \bar{F}^n(\bot_0) \}_n \) must be an increasing chain and all strictly increasing chain in \( \bar{L} \) must be finite.
Example 3

Approximate equation \( \overline{X} = \overline{F(X)} \) corresponding to Ex.1’s program \( P \) using interval abstraction of Ex.2 is:

\[
\overline{F} = \lambda X.([1, 1] \sqcup (X \oplus [1, 1])) \sqcap [-\infty, 100]
\]

where

\[
\begin{align*}
\bot \sqcup X &= X \sqcup \bot = X \\
[l_0, u_0] \sqcup [l_1, u_1] &= [\min(l_0, l_1), \max(u_0, u_1)] \\
\bot \sqcap X &= X \sqcap \bot = \bot \\
[l_0, u_0] \sqcap [l_1, u_1] &= \bot \\
&= [\max(l_0, l_1), \min(u_0, u_1)] \text{ if } \max(l_0, l_1) > \min(u_0, u_1) \\
&= [\max(l_0, l_1), \min(u_0, u_1)] \text{ otherwise} \\
\bot \oplus X &= X \oplus \bot = \bot \\
[l_0, u_0] \oplus [l_1, u_1] &= [l_0 + l_1, u_0 + u_1]
\end{align*}
\]
Example 3 (cont’d)

This can be solved iteratively starting from $\bot$: 

$$
\bot, [1, 1], [1, 2], \cdots, [1, 100]
$$

But for nonterminating programs, this sequence might be infinite and strictly increasing.
Use of Infinite Abstract Domain

- To obtain more information
- To increase precision
- But, the interpretation may not be finitely computable
Widening

- Widening is another method for enforcing termination of the abstract interpretation.
- General idea is to eliminate unstable components through consecutive iterates.
- And find a more approximate but sound upper bound of the iteration sequence.
Definition of Widening

**Definition**  \( \text{widening} \; \sqcup \in L \times L \longrightarrow L \)

\[ \forall x, y \in L : x \sqsubseteq x \sqcup y \]
\[ \forall x, y \in L : y \sqsubseteq x \sqcup y \]

for all increasing chains \( x^0 \sqsubseteq x^1 \sqsubseteq \cdots \), the chain defined by \( y^0 = x^0, \cdots, y^{i+1} = y^i \sqcup x^{i+1}, \cdots \) is not strictly but increasing.
Upward Iteration Sequence

Upward iteration sequence \( \{ \hat{X}^n \}_n \)

\[
\begin{align*}
\hat{X}^0 &= \bot_0 \\
\hat{X}^{i+1} &= \hat{X}^i \\
&= \hat{X}^i \triangledown F(\hat{X}^i)
\end{align*}
\]

- is ultimately stationary.
- its limit \( \hat{A} \) is a sound upper approximate of \( \text{lfp}_{\bot_0}(F) \).
Narrowingimproves the precision of the upper approximate.

It improves the stabilized but over approximated extrapolation by widening,

while still maintaining the ultimately stationary property.
**Definition of Narrowing**

**Definition**  \(\text{narrowing } \triangle \in L \times L \rightarrow L\)

\[\forall x, y \in L : (y \sqsubseteq (x \triangle y) \subseteq x)\]

for all decreasing chains \(x^0 \sqsubseteq x^1 \sqsubseteq \cdots\), the chain defined by \(y^0 = x^0, \cdots, y^{i+1} = y^i \triangle x^{i+1}, \cdots\) is not strictly but decreasing.
Downward Iteration Sequence

Downward abstract iteration sequence \( \{ \hat{X}^n \}_n \)

\[
\begin{align*}
\hat{X}^0 &= \hat{A} \\
\hat{X}^{i+1} &= \hat{X}^i \triangleq F(\hat{X}^i)
\end{align*}
\]

- is ultimately stationary.
- its limit \( \hat{A} \) and each \( \hat{X}^i \) is a sound upper approximate of \( \text{lfp}_{\perp_0}(F) \).
Example 4

For the lattice of intervals $\overline{L}$ of Ex.3, widening and narrowing can be defined as:

$$\bot \triangledown X = X \triangledown \bot = X$$

$$[l_0, u_0] \triangledown [l_1, u_1] = \begin{cases} \text{if } l_1 < l_0 \text{ then } -\infty \text{ else } l_0, \\ \text{if } u_1 > u_0 \text{ then } +\infty \text{ else } u_0 \end{cases}$$

$$\bot \triangle X = X \triangle \bot = \bot$$

$$[l_0, u_0] \triangle [l_1, u_1] = \begin{cases} \text{if } l_0 = -\infty \text{ then } l_1 \text{ else } l_0, \\ \text{if } u_0 = +\infty \text{ then } u_1 \text{ else } u_0 \end{cases}$$
Example 4 (cont’d)

Then the iteration sequence for resolving

\[ X = \overline{F}(X) = ([1, 1] \sqcup (X \oplus [1, 1])) \cap [-\infty, 100] \]

becomes:

\[
\begin{align*}
\hat{X}^0 &= \bot \\
\hat{X}^1 &= [1, 1] \\
\hat{X}^2 &= [1, +\infty] \\
\hat{X}^3 &= \hat{X}^2 = \hat{A} \\
\end{align*}
\]

\[
\begin{align*}
\check{X}^0 &= \check{A} = [1, +\infty] \\
\check{X}^1 &= [1, 100] \\
\check{X}^2 &= \check{X}^1 \\
\end{align*}
\]
Aren’t both the same?

- Unappreciated conjecture: “Given an infinite abstract domain with specific widening/narrowing operators, it is possible to find a finite lattice which will give the same results.”
Aren’t both the same?

- Unappreciated conjecture: “Given an infinite abstract domain with specific widening/narrowing operators, it is possible to find a finite lattice which will give the same results.”

- No! There is a counter example.
program \(P_{n_1n_2}\) :
    var I : integer ;
begin
    I := \(n_1\);  
    while I \(\leq\) \(n_2\) do 
    begin 
    \{ I \in [ n_1, n_2 ] \} 
    I := I + 1; 
    end; 
\{ I = n_2 + 1 \} 
end;

For program \(P_{n_1n_2}\), it is impossible to choose a single limited abstract domain that has the equivalent precision to using interval abstraction domain with widening/narrowing.
As we can see in the previous example, using limited domain is not powerful enough as infinite domain with widening/narrowing.

Furthermore, for a particular program it is not possible to infer the set of needed abstract values by a simple inspection of the text of the program.
Can Do for Limited Domain

Whenever

- $L \xleftarrow{\gamma} \xrightarrow{\alpha} \bar{L}$, and
- $\bar{L}(\sqsubseteq, \square)$ satisfies ascending chain condition,

then widening $\triangledown \in L \times L \longmapsto L$ can be defined as:

$$x \triangledown y = \gamma(\alpha(x) \square \alpha(y))$$

and narrowing $\triangle \in L \times L \longmapsto L$ as:

$$x \triangle y = x \sqcap \gamma \circ \alpha(y)$$

that have the same cost and precision up to $\langle \alpha, \gamma \rangle$. 
Design of Widening/Narrowing

- When a Galois connection to a limited domain is given, widening/narrowing is automatically defined.
- Try to use least upper bounds/greatest lower bounds as long as the iterates follow finite chains and extrapolate when some iterate belongs to an infinite one.
- However, widening/narrowing always exist:

\[
x \triangledown y = \begin{cases} 
y \sqsubseteq x & \text{if } y \sqsubseteq x \text{ then } x \\
\top & \text{else}
\end{cases}
\]

\[
x \triangledown y = x
\]
Extrapolation Threshold

\[
\langle x, i \rangle \bar{\triangledown} \langle y, i + 1 \rangle =
\begin{cases}
    \langle x, i + 1 \rangle & \text{if } y \subseteq x \\
    \langle x \bigtriangledown y, i + 1 \rangle & \text{if } i \leq n \\
    \langle x \bigtriangledown y, i + 1 \rangle & \text{else}
\end{cases}
\]

\[
\langle x, i \rangle \bar{\triangle} \langle y, i + 1 \rangle =
\begin{cases}
    \langle x \bigtriangleup y, i + 1 \rangle & \text{if } i \leq n \\
    \langle x \bigtriangleup y, i + 1 \rangle & \text{else}
\end{cases}
\]

Limit the number of iterations to some given positive integer \( n \) until extrapolation.
Conclusion

- Widening/narrowing approach is more powerful than the Galois connection approach with limited domains.
- Widening/narrowing approach can improve,  
  ♦ the precision  
  ♦ the speed of convergence  
  of the analyses significantly.
- Combination of the two approaches using infinite abstract domain is worthwhile, practical.
References

[1] Patrick Cousot and Radhia Cousot, “Comparing the Galois Connection and Widening/Narrowing Approaches to Abstract Interpretation”