Widening and Narrowing Approaches to Abstract Interpretation [a]

Jaeho Shin

netj@ropas.snu.ac.kr

Programming Research Lab, Seoul National University

[a] 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

Program Semantics

- Program semantics can be specified as $\mathrm{lfp}_{\perp_0}(F)$.

$$\begin{array}{rl} \text{CPO} & L(\sqsubseteq, \sqcup) \\ \text{operator} & F \in L \xmapsto{con} L \\ & \perp_0 \sqsubseteq F(\perp_0), \perp_0 \in L \end{array}$$

- By Kleene's fixpoint theorem,

$$\mathrm{lfp}_{\perp_0}(F) = \bigsqcup_{n \in \mathbb{N}} F^n(\perp_0)$$

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 $\mathrm{lfp}_{\perp_0}(F) \sqsubseteq A$.

Example 1

Let the collecting semantics of

```
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

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

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 $L(\sqsubseteq)$ and $\overline{L}(\overline{\sqsubseteq})$ are posets, then $\langle \alpha, \gamma \rangle$ is a *Galois connection*, written $L \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \overline{L}$, iff $\alpha \in L \longmapsto \overline{L}$ and $\gamma \in \overline{L} \longmapsto L$ such that:

$$\forall x \in L, \overline{y} \in \overline{L} : (\alpha(x) \mathrel{\overline{\sqsubseteq}} \overline{y}) \Longleftrightarrow (x \sqsubseteq \gamma(\overline{y})).$$

Example 2

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

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

ordered by $\overline{\sqsubseteq}$ such that:

$$\begin{array}{rcl} \perp \overline{\sqsubseteq} [l,u] & \stackrel{def}{=} & \text{true} \\ [l_0,u_0] \overline{\sqsubseteq} [l_1,u_1] & \stackrel{def}{=} & l_1 \leq l_0 \leq u_0 \leq u_1 \end{array}$$

Galois connection is defined by:

$$\begin{array}{rclcrcl} \gamma(\perp) & = & \phi & & \alpha(\phi) & = & \perp \\ \gamma([l,u]) & = & \{x \in \mathbb{Z} \mid l \leq x \leq u\} & & \alpha(X) & = & [\min X, \max X] \end{array}$$

Galois Connection Approach

Definition $\langle \overline{L}, \overline{\bot_0}, \overline{F} \rangle$ is an *abstract interpretation* of $\langle L, \bot_0, F \rangle$, written $\langle L, \bot_0, F \rangle \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \langle \overline{L}, \overline{\bot_0}, \overline{F} \rangle$, iff $L \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \overline{L}$, $\alpha(\bot_0) \mathrel{\overline{\sqsubseteq}} \overline{\bot_0}$ and $\alpha \circ F \mathrel{\overline{\sqsubseteq}} \overline{F} \circ \alpha$.

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

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

Example 3

Approximate equation $\overline{X} = \overline{F}(\overline{X})$ 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

$$\begin{array}{rcll}
\bot \sqcup X & = & X \sqcup \bot = X & \\
[l_0, u_0] \sqcup [l_1, u_1] & = & [\min(l_0, l_1), \max(u_0, u_1)] & \\
\bot \sqcap X & = & X \sqcap \bot = \bot & \\
[l_0, u_0] \sqcap [l_1, u_1] & = & \bot & \text{if } \max(l_0, l_1) > \min(u_0, u_1) \\
& = & [\max(l_0, l_1), \min(u_0, u_1)] & \text{otherwise} \\
\bot \oplus X & = & X \oplus \bot = \bot & \\
[l_0, u_0] \oplus [l_1, u_1] & = & [l_0 + l_1, u_0 + u_1] &
\end{array}$$

Example 3 (cont'd)

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

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

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 *widening* $\nabla \in L \times L \longmapsto L$

$$\forall x, y \in L : x \sqsubseteq x \nabla y$$
$$\forall x, y \in L : y \sqsubseteq x \nabla y$$

for all increasing chains $x^0 \sqsubseteq x^1 \sqsubseteq \cdots$, the chain defined by $y^0 = x^0, \cdots, y^{i+1} = y^i \nabla x^{i+1}, \cdots$ is not strictly but increasing.

Upward Iteration Sequence

Upward iteration sequence $\{\hat{X}^n\}_n$

$$\begin{array}{rcll} \hat{X}^0 & = & \bot_0 & \\ \hat{X}^{i+1} & = & \hat{X}^i & \text{if } F(\hat{X}^i) \sqsubseteq \hat{X}^i \\ & = & \hat{X}^i \bigtriangledown F(\hat{X}^i) & \text{otherwise} \end{array}$$

- is ultimately stationary.
- its limit $\hat{A}$ is a sound upper approximate of $\mathrm{lfp}_{\bot_0}(F)$.

Narrowing

- Narrowing improves 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 *narrowing* $\triangle \in L \times L \longmapsto L$

$$\forall x, y \in L : (y \sqsubseteq (x \triangle y) \sqsubseteq x)$$

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 $\{\check{X}^n\}_n$

$$\begin{aligned} \check{X}^0 &= \hat{A} \\ \check{X}^{i+1} &= \check{X}^i \bigtriangleup F(\check{X}^i) \end{aligned}$$

- is ultimately stationary.
- its limit $\check{A}$ and each $\check{X}^i$ is a sound upper approximate of $\mathrm{lfp}_{\perp_0}(F)$.

Example 4

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

$$
\begin{aligned}
\bot \bigtriangledown X &= X \bigtriangledown \bot = X \\
[l_0, u_0] \bigtriangledown [l_1, u_1] &= [\text{if } l_1 < l_0 \text{ then } -\infty \text{ else } l_0, \\
&\quad\ \text{if } u_1 > u_0 \text{ then } +\infty \text{ else } u_0] \\
\bot \bigtriangleup X &= X \bigtriangleup \bot = \bot \\
[l_0, u_0] \bigtriangleup [l_1, u_1] &= [\text{if } l_0 = -\infty \text{ then } l_1 \text{ else } l_0, \\
&\quad\ \text{if } u_0 = +\infty \text{ then } u_1 \text{ else } u_0]
\end{aligned}
$$

Example 4 (cont'd)

Then the iteration sequence for resolving

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

becomes:

$$\begin{aligned} \hat{X}^0 &= \bot \\ \hat{X}^1 &= [1,1] \\ \hat{X}^2 &= [1,+\infty] \\ \hat{X}^3 &= \hat{X}^2 = \hat{A} \end{aligned} \qquad \begin{aligned} \check{X}^0 &= \hat{A} = [1,+\infty] \\ \check{X}^1 &= [1,100] \\ \check{X}^2 &= \check{X}^1 \end{aligned}$$

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.

Counter Example

program $P_{n_1 n_2}$:
  var I : integer ;
begin
  I := n_1;
  while I <= n_2 do
    begin
      { I $\in$ [n_1, n_2] }
      I := I + 1;
    end;
  { I = n_2 + 1 }
end;

For program $P_{n_1 n_2}$, it is impossible to choose a single limited abstract domain that has the equivalent precision to using interval abstraction domain with widening/narrowing.

Limited Domain Can't Do for ∇/Δ

- 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.

$\bigtriangledown/\bigtriangleup$ Can Do for Limited Domain

Whenever

- $L \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \overline{L}$, and
- $\overline{L}(\,\overline{\sqsubseteq}\,,\,\overline{\sqcup}\,)$ satisfies ascending chain condition,

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

$$x \bigtriangledown y = \gamma(\alpha(x) \,\overline{\sqcup}\, \alpha(y))$$

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

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

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:

$$
\begin{aligned}
x \bigtriangledown y &= \text{if } y \sqsubseteq x \text{ then } x \text{ else } \top \\
x \bigtriangleup y &= x
\end{aligned}
$$

Extrapolation Threshold

$$
\begin{aligned}
\langle x, i\rangle \overline{\triangledown} \langle y, i+1\rangle = \;& \text{if } y \sqsubseteq x \text{ then } \langle x, i+1\rangle \\
& \text{elsif } i \leq n \text{ then } \langle x \sqcup y, i+1\rangle \\
& \text{else } \langle x \triangledown y, i+1\rangle \\
\langle x, i\rangle \overline{\triangle} \langle y, i+1\rangle = \;& \text{if } i \leq n \text{ then } \langle x \sqcap y, i+1\rangle \\
& \text{else } \langle x \triangle y, i+1\rangle
\end{aligned}
$$

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"