Abstracting Non-Linear Programs

MPRI 2–6: Abstract Interpretation,
application to verification and static analysis

Antoine Miné

year 2013–2014

course 05-B
18 October 2013

Abstraction framework

Issue:

Most relational domains can only deal with linear expressions.
How can we abstract non-linear assignments such as $\mathtt{X} := \mathtt{Y} \times \mathtt{Z}$?

Idea: replace $\mathtt{Y} \times \mathtt{Z}$ with a sound linear approximation.

Framework:

We define an approximation preorder $\preceq$ on expressions:

$$R \models e_1 \preceq e_2 \iff^{\text{def}} \forall \rho \in R,\ \mathsf{E}[\![e_1]\!] \rho \subseteq \mathsf{E}[\![e_2]\!] \rho.$$

Soundness properties if $\gamma(\mathcal{X}^\sharp) \models e \preceq e'$ then:

- $\mathsf{C}[\![\mathtt{V} := e]\!] \gamma(\mathcal{X}^\sharp) \subseteq \gamma(\mathsf{C}^\sharp[\![\mathtt{V} := e']\!] \mathcal{X}^\sharp)$
- $\mathsf{C}[\![e \bowtie 0]\!] \gamma(\mathcal{X}^\sharp) \subseteq \gamma(\mathsf{C}^\sharp[\![e' \bowtie 0]\!] \mathcal{X}^\sharp)$
- $\gamma(\mathcal{X}^\sharp) \cap (\overleftarrow{\mathsf{C}}[\![\mathtt{V} := e]\!] \gamma(\mathcal{R}^\sharp)) \subseteq \gamma(\mathsf{C}^\sharp[\![\overleftarrow{\mathtt{V} := e'}]\!]^\sharp(\mathcal{X}^\sharp, \mathcal{R}^\sharp))$

$\Longrightarrow$ we can now use e' in the abstract instead of e.

Linearization

In practice, we put expressions into affine interval form:

$$\texttt{exp}_\ell : [a_0, b_0] + \sum_k [a_k, b_k] \times \texttt{V}_\texttt{k}$$

Advantages:

- affine expressions are easy to manipulate,
- interval coefficients allow non-determinism in expressions, hence, the opportunity for abstraction,
- we can easily construct abstract transfer functions for affine interval expressions.

Linearization (cont.)

Operations on affine interval forms

- adding $\boxplus$ and subtracting $\boxminus$ two forms,
- multiplying $\boxtimes$ and dividing $\boxslash$ a form by an interval.

Noting i_k the interval $[a_k, b_k]$ and using interval operations $+^\sharp_b$, $-^\sharp_b$, $\times^\sharp_b$, $/^\sharp_b$ (e.g., $[a,b] +^\sharp_b [c,d] = [a+c, b+d]$):

- $(i_0 + \sum_k i_k \times \mathtt{V_k}) \boxplus (i'_0 + \sum_k i'_k \times \mathtt{V_k}) \stackrel{\text{def}}{=} (i_0 +^\sharp_b i'_0) + \sum_k (i_k +^\sharp_b i'_k) \times \mathtt{V_k}$
- $i \boxtimes (i_0 + \sum_k i_k \times \mathtt{V_k}) \stackrel{\text{def}}{=} (i \times^\sharp_b i_0) + \sum_k (i \times^\sharp_b i_k) \times \mathtt{V_k}$
- ...

Projection $\pi_k : \mathcal{D}^\sharp \to \mathtt{exp}_\ell$

We suppose we are given an abstract interval projection operator π_k such that:

$$\pi_k(\mathcal{X}^\sharp) = [a,b] \text{ such that } [a,b] \supseteq \{\, \rho(\mathtt{V_k}) \mid \rho \in \gamma(\mathcal{X}^\sharp) \,\}.$$

Linearization (cont.)

Intervalization $\iota : (\texttt{exp}_\ell \times \mathcal{D}^\sharp) \to \texttt{exp}_\ell$

Flattens the expression into a single interval:

$$\iota(i_0 + \sum_k (i_k \times \texttt{V}_\texttt{k}),\, \mathcal{X}^\sharp) \stackrel{\text{def}}{=} i_0 \;+^\sharp_b\; \sum^\sharp_{b,\,k} (i_k \times^\sharp_b \pi_k(\mathcal{X}^\sharp)).$$

Linearization $\ell : (\texttt{exp} \times \mathcal{D}^\sharp) \to \texttt{exp}_\ell$

Defined by induction on the syntax of expressions:

- $\ell(\texttt{V}, \mathcal{X}^\sharp) \stackrel{\text{def}}{=} [1,1] \times \texttt{V}$,
- $\ell([a,b], \mathcal{X}^\sharp) \stackrel{\text{def}}{=} [a,b]$,
- $\ell(e_1 + e_2, \mathcal{X}^\sharp) \stackrel{\text{def}}{=} \ell(e_1, \mathcal{X}^\sharp) \boxplus \ell(e_2, \mathcal{X}^\sharp)$,
- $\ell(e_1 - e_2, \mathcal{X}^\sharp) \stackrel{\text{def}}{=} \ell(e_1, \mathcal{X}^\sharp) \boxminus \ell(e_2, \mathcal{X}^\sharp)$,
- $\ell(e_1 / e_2, \mathcal{X}^\sharp) \stackrel{\text{def}}{=} \ell(e_1, \mathcal{X}^\sharp) \boxslash \iota(\ell(e_2, \mathcal{X}^\sharp), \mathcal{X}^\sharp)$,
- $\ell(e_1 \times e_2, \mathcal{X}^\sharp) \stackrel{\text{def}}{=}$ can be $\begin{cases} \text{either} & \iota(\ell(e_1, \mathcal{X}^\sharp), \mathcal{X}^\sharp) \boxtimes \ell(e_2, \mathcal{X}^\sharp), \\ \text{or} & \iota(\ell(e_2, \mathcal{X}^\sharp), \mathcal{X}^\sharp) \boxtimes \ell(e_1, X^\sharp). \end{cases}$

Linearization application

Property soundness of the linearization:

For any abstract domain $\mathcal{D}^\sharp$, any $\mathcal{X}^\sharp \in \mathcal{D}^\sharp$ and $e \in \texttt{exp}$, we have:

$$\gamma(\mathcal{X}^\sharp) \models e \preceq \ell(e, \mathcal{X}^\sharp)$$

Remarks:

- ℓ results in a loss of precision,
- ℓ is not monotonic for $\preceq$.
  (e.g., $\ell(\texttt{V}/\texttt{V}, \texttt{V} \mapsto [1, +\infty]) = [0, 1] \times \texttt{V} \not\preceq 1$)

Application to the octagon domain

```
Y:=[0,+∞];
T:=[-1,1];
X:=T×Y
```

- $\texttt{T} \times \texttt{Y}$ is linearized as $[-1, 1] \times \texttt{Y}$,
- we can prove that $|\texttt{X}| \leq \texttt{Y}$.

Linearization application (cont.)

Application to the interval domain

$C^\sharp[\![\, \mathtt{V} := \ell(e, \mathcal{X}^\sharp) \,]\!]\, \mathcal{X}^\sharp$ is always more precise than $C^\sharp[\![\, \mathtt{V} := e \,]\!]\, \mathcal{X}^\sharp$

ℓ simplifies symbolically variables occurring several times.

Example: $\mathtt{X} := 2 \times \mathtt{V} - \mathtt{V}$, where $\mathtt{V} \in [a, b]$:

- using vanilla intervals:
  $E^\sharp[\![\, 2 \times \mathtt{V} - \mathtt{V} \,]\!]\, (\mathcal{X}^\sharp) = 2 \times^\sharp_b [a, b] -^\sharp_b [a, b] = [2a - b, 2b - a]$,
- after linearization $\ell(2 \times \mathtt{V} - \mathtt{V}, \mathcal{X}^\sharp) = \mathtt{V}$, so
  $$E^\sharp[\![\, \ell(2 \times \mathtt{V} - \mathtt{V}, \mathcal{X}^\sharp) \,]\!]\, \mathcal{X}^\sharp = [a, b]$$
  strictly more precise than $[2a - b, 2b - a]$ when $a \neq b$.