Computing refinement

In order to automate CEGAR, we need an effective method for computing new predicates that result in the desired refinement.

Here, we discuss the following two methods for refinement of a predicate abstraction.

▶ Syntax based refinement
▶ Interpolation based refinement

Let us first define(remind) some notations.
Path constraints for spurious counterexample (reminder)

$V_i$ be the vector of variables obtained by adding subscript $i$ after each variable in $V$.

**Definition 19.1**

For a spurious counterexample $e_1 \ldots e_n$, path constraints $\text{pathCons}(e_1 \ldots e_n)$ is

$$\bigwedge_{i \in 1 \ldots n} e_i(V_{i-1}, V_i)$$

A path is **feasible** if corresponding path constraints is satisfiable.

**Note:** Path constraints are also known as “SSA formulas”. 
Syntax based refinement

\[ \text{core} = \text{unsatCore}(\text{pathCons}(e_1 \ldots e_n)). \]

\[ \text{preds} = \text{atoms of core} \text{ after erasing subscripts in its variables} \]

Add \text{preds} in the predicate domain to obtain the refined abstract domain
Interpolation

Definition 19.2
Let $A$ and $B$ be formulas such that $A \land B$ is unsat. An interpolant $I$ between $A$ and $B$ is a formula such that

1. $A \Rightarrow I$
2. $B \land I \Rightarrow \bot$
3. $\text{vars}(I) \subseteq \text{vars}(A) \cap \text{vars}(B)$

Theorem 19.1 (Craig interpolation theorem)
Interpolant always exists.

Example 19.1
Consider:

$A = x_1 + x_2 \leq 2 \land x_3 - x_2 \leq 0$
$B = 6x_4 - 2x_1 \leq -8 \land -3x_4 - x_3 \leq 0$

$\text{vars}(A) = \{x_1, x_2, x_3\} \quad \text{vars}(B) = \{x_1, x_3, x_4\} \quad \text{vars}(I) \subseteq \{x_1, x_3\}$
$I = x_1 + x_3 \leq 2$
Interpolation chain

We can extend the definition of interpolant to our setting

Definition 19.3
Consider unsat formula $\bigwedge_{i \in 1..m} e_i(V_{i-1}, V_i)$. An interpolant chain is a sequence of formulas such that $I_0 \ldots I_m$ such that

- $I_0 = \top$
- $\forall i \in 1..m I_{i-1} \land e_i(V_{i-1}, V_i) \Rightarrow I_i$
- $I_m = \bot$
- $\text{vars}(I_i) \subseteq V_i$
Interpolation for refinement

We compute interpolation chain $I_0 \ldots I_m$ for $\text{pathCons}(\text{cons})$

$preds = \text{atoms in } I_0 \ldots I_m \text{ after erasing subscripts in its variables}$

Add $preds$ in the predicate domain to obtain the refined abstract domain

Theorem 19.2

The new abstract domain eliminates spurious counterexample $e_1 \ldots e_n$
Example: interpolation for refinement

Program:

\[ \rho_1 : x := 0 \quad y := 0 \]
\[ \rho_2 : x := x + 1 \quad y := y + 1 \]
\[ \rho_3 : \text{skip} \]
\[ \rho_4 : x := x - 1 \quad y := y - 1 \]
\[ \rho_5 : x < 0 \land y > 0 \]

Spurious counterexample: \( \rho_1 \rho_2 \rho_3 \rho_4 \rho_4 \rho_5 \).

\[
\begin{align*}
\top & \quad \rho_1(x_0, y_0, x_1, y_1) = (x_1 = 0 \land y_1 = 0) \\
I_1 = y_1 & \leq 0 \\
\rho_2(x_1, y_1, x_2, y_2) = (x_2 = x_1 + 1 \land y_2 = y_1 + 1) \\
I_2 = y_2 & \leq 1 \quad \leftarrow \text{New predicate} \\
\rho_3(x_2, y_2, x_3, y_3) = (x_3 = x_2 \land y_3 = y_2) \\
I_3 = y_3 & \leq 0 \\
\rho_4(x_3, y_3, x_4, y_4) = (x_4 = x_3 - 1 \land y_4 = y_3 - 1) \\
I_4 = y_4 & \leq 0 \\
\rho_4(x_4, y_4, x_5, y_5) = (x_5 = x_4 - 1 \land y_5 = y_4 - 1) \\
I_5 = y_5 & \leq 0 \\
\rho_5(x_5, y_5, x_6, y_6) = (x_5 < 0 \land y_5 > 0) \\
\bot
\end{align*}
\]
Example: refined reachability graph

\[ P \text{reds} = \{ x \geq 0, y \leq 0, x \geq 1, y \leq 1 \} \]

Program:

\[ \rho_1 : x := 0 \]
\[ y := 0 \]
\[ \rho_2 : x := x + 1 \]
\[ y := y + 1 \]
\[ \rho_3 : \text{skip} \]
\[ \rho_4 : x := x - 1 \]
\[ y := y - 1 \]
\[ \rho_5 : x < 0 \land y > 0 \]

Exercise 19.1

Complete the ARG
Example: good refinement

Consider the earlier spurious counterexample again: $\rho_1\rho_2\rho_3\rho_4\rho_5$.

\[\emptyset\]

$\rho_1(x_0, y_0, x_1, y_1) = (x_1 = 0 \land y_1 = 0)$

$I_1 = y_1 \leq x_1$

$\rho_2(x_1, y_1, x_2, y_2) = (x_2 = x_1 + 1 \land y_2 = y_1 + 1)$

$I_2 = y_2 \leq x_2 \leftarrow \text{New predicate}$

$\rho_3(x_2, y_2, x_3, y_3) = (x_3 = x_2 \land y_3 = y_2)$

$I_3 = y_3 \leq x_3$

$\rho_4(x_3, y_3, x_4, y_4) = (x_4 = x_3 - 1 \land y_4 = y_3 - 1)$

$I_4 = y_4 \leq x_4$

$\rho_4(x_4, y_4, x_5, y_5) = (x_5 = x_4 - 1 \land y_5 = y_4 - 1)$

$I_5 = y_5 \leq x_5$

$\rho_5(x_5, y_5, x_6, y_6) = (x_5 < 0 \land y_5 > 0)$

$\bot$
Example: ARG without spurious counterexample

\[ Preds = \{x \geq 0, y \leq 0, x \geq 1, y \leq x\} \]

Program:

\[
\begin{align*}
\rho_1 &: \ x := 0 \\
& \quad y := 0 \\
\rho_2 &: \ x := x + 1 \\
& \quad y := y + 1 \\
\rho_3 &: \ skip; \quad (l_1, [1011]) \\
\rho_4 &: \ x := x - 1 \\
& \quad y := y - 1 \\
\rho_5 &: \ x < 0 \land y > 0
\end{align*}
\]

Exercise 19.2

Complete the ARG
End of Lecture 19