The $\mathbb{K}$ Summarizer

Runtime Verification, Inc.
slides composed by Xiaohong Chen

June 26, 2022

Table of Contents

Table of Contents

Matching Logic Patterns

$$\varphi \coloneqq x \mid X \mid \sigma(\varphi_1, \ldots, \varphi_n) \mid \bot \mid \varphi_1 \to \varphi_2 \mid \exists x \,.\, \varphi \mid \mu X \,.\, \varphi$$

Matching Logic Theories and Notations

- $\lceil \varphi \rceil$, definedness/ceiling
- $\lfloor \varphi \rfloor$, totality/flooring
- $\varphi_1 = \varphi_2$, equality
- $\varphi_1 \subseteq \varphi_2$, set inclusion
- $x \in \varphi$, membership
- $\neg_s \varphi$, sorted negation
- $\exists x : s \,.\, \varphi$ and $\forall x : s \,.\, \varphi$, sorted quantification
- $\mu X : s \,.\, \varphi$ and $\nu X : s \,.\, \varphi$, sorted fixpoints

Transition Systems

Let Cfg be a distinguished sort of configurations. Let us assume a configuration model where $\Rightarrow_{\mathsf{Cfg}}$ is a transition relation over configurations.

- $\bullet\varphi$, one-path next
  - $s \in \bullet\varphi$ if there exists s' such that $s \Rightarrow_{\mathsf{Cfg}} s'$ and $s' \in \varphi$.
  - NONSTOP $\equiv \bullet\top_{\mathsf{Cfg}}$, the set of non-terminating configurations
  - STOP $\equiv \neg_{\mathsf{Cfg}}$NONSTOP, the set of terminating configurations
- $\circ\varphi \equiv \neg_{\mathsf{Cfg}}\bullet\neg_{\mathsf{Cfg}}\varphi$, all-path next
  - $s \in \circ\varphi$ if for all s' such that $s \Rightarrow_{\mathsf{Cfg}} s'$, $s' \in \varphi$.
  - $\circ\bot$ is equivalent to STOP.
- Sometimes we want to enforce non-termination for all-path next.
- $\circ_s\varphi \equiv \circ\varphi \wedge$ NONSTOP, strong all-path next.
  - $\circ_s\varphi$ is equivalent to $\circ\varphi \wedge \bullet\varphi$.

Rewrite Rules

- $\varphi_1 \Rightarrow_1^{\exists} \varphi_2 \equiv \varphi_1 \to \bullet\varphi_2$, one-path one-step rewriting.
  - The superscript $\exists$ means "one-path"
  - The subscript 1 means "one-step"

A *rewrite rule* is an axiom of the form

$$\varphi_1 \Rightarrow_1^{\exists} \varphi_2$$

where φ_1 and φ_2 are patterns of sort Cfg.
Semantically, $\varphi_1 \Rightarrow_1^{\exists} \varphi_2$ holds iff for all $s \in \varphi_1$ there exists $s' \in \varphi_2$ such that $s \Rightarrow_{\mathsf{Cfg}} s'$.

One-Path Rewriting

- $\varphi_1 \Rightarrow^{\exists}_{2} \varphi_2 \equiv \varphi_1 \to \bullet\bullet\varphi_2$, one-path two-step rewriting.
- $\varphi_1 \Rightarrow^{\exists}_{3} \varphi_2 \equiv \varphi_1 \to \bullet\bullet\bullet\varphi_2$, one-path three-step rewriting.
- $\ldots$
- $\diamond\varphi_2 \equiv \mu X : \mathrm{Cfg} \,.\, \varphi_2 \vee \bullet X$
  - $s \in \diamond\varphi_2$ if there exists a finite execution trace $s \Rightarrow^{n}_{\mathrm{Cfg}} s'$ for $n \geq 0$ such that $s' \in \varphi_2$
- $\varphi_1 \Rightarrow^{\exists}_{*} \varphi_2 \equiv \varphi_1 \to \diamond\varphi_2$, one-path finite-step rewriting.
  - It holds iff for all $s \in \varphi_1$ there exists $s \Rightarrow^{n}_{\mathrm{Cfg}} s'$ for $n \geq 0$ such that $s' \in \varphi_2$.

All-Path Rewriting

- $\varphi_1 \Rightarrow_*^{\exists} \varphi_2 \equiv \varphi_1 \to \diamond\varphi_2$, one-path finite-step rewriting.
  - It holds iff for all $s \in \varphi_1$ there exists $s \Rightarrow_{\mathsf{Cfg}}^{n} s'$ for $n \geq 0$ such that $s' \in \varphi_2$.
- Question: what should "all-path symbolic execution" $\varphi_1 \Rightarrow_*^{\forall} \varphi_2$ mean?
- (All-path reachability): for all $s \in \varphi_1$ and all complete (i.e., finite and maximum) paths

$$s = s_0 \Rightarrow_{\mathsf{Cfg}} s_1 \Rightarrow_{\mathsf{Cfg}} \ldots \Rightarrow_{\mathsf{Cfg}} s_n, \quad s_n \text{ terminating}$$

  there exists $0 \leq m \leq n$ such that $s_m \in \varphi_2$.
  - Problem is that behaviors on infinite paths are entirely ignored.
- (More Reasonable?): for all $s \in \varphi_1$ and all complete or infinite paths

$$s = s_0 \Rightarrow_{\mathsf{Cfg}} s_1 \Rightarrow_{\mathsf{Cfg}} \cdots$$

  there exists $m \geq 0$ such that $s_m \in \varphi_2$.

All-Path Rewriting

- $\varphi_1 \Rightarrow_1^{\forall} \varphi_2 \equiv \varphi_1 \to \circ\varphi_2$, all-path one-step rewriting.
  - It holds iff for all $s \in \varphi_1$ and $s \Rightarrow_{\mathsf{Cfg}} s'$, $s' \in \varphi_2$. Note that s is allowed to be terminating.
- $\varphi_1 \Rightarrow_1^{s,\forall} \varphi_2 \equiv \varphi_1 \to \circ_s\varphi_2$, strong all-path one-step rewriting.
  - In addition to $\varphi_1 \Rightarrow_1^{\forall} \varphi_2$, it requires all $s \in \varphi_1$ to be non-terminating.
- $\varphi_1 \Rightarrow_2^{\forall} \varphi_2 \equiv \varphi_1 \to \circ\circ\varphi_2$, one-path two-step rewriting.
- $\varphi_1 \Rightarrow_2^{s,\forall} \varphi_2 \equiv \varphi_1 \to \circ_s\circ_s\varphi_2$, strong one-path two-step rewriting.
- $\varphi_1 \Rightarrow_*^{\forall} \varphi_2 \equiv \varphi_1 \to \mu X \,.\, \varphi_2 \vee \circ_s X$, all-path rewriting.
- $\varphi_1 \Rightarrow_{\mathsf{reach}}^{\forall} \varphi_2 \equiv \varphi_1 \to \nu X \,.\, \varphi_2 \vee \circ_s X$, all-path reachability.

All-Path Rewriting

Lemma

The following statements about s are equivalent:

1. *For all complete or infinite paths $s = s_0 \Rightarrow_{\mathsf{Cfg}} s_1 \Rightarrow_{\mathsf{Cfg}} \cdots$ there exists $m \geq 0$ such that $s_m \in \varphi$;*
2. $s \in |\mu X \,.\, \varphi_2 \vee \circ_s X|_M$

Proof (TODO).

Let ξ and η denote the set of all configurations that satisfy (1) and (2), respectively. Prove both $\xi \subseteq \eta$ and $\eta \subseteq \xi$. Hint: use the Knaster-Tarski theorem to convert η into a big intersection. $\square$

Lemma

All-path rewriting implies all-path reachability: $\vdash \varphi_1 \Rightarrow_*^\forall \varphi_2$ *implies* $\vdash \varphi_1 \Rightarrow_{\text{reach}}^\forall \varphi_2$.

Proof.

Simply observe that all-path rewriting

$$\varphi_1 \Rightarrow_*^\forall \varphi_2 \equiv \varphi_1 \rightarrow \mu X \,.\, \varphi_2 \vee \circ_s X$$

is the least fixpoint while all-path reachability

$$\varphi_1 \Rightarrow_{\text{reach}}^\forall \varphi_2 \equiv \varphi_1 \rightarrow \nu X \,.\, \varphi_2 \vee \circ_s X$$

is the greatest fixpoint. $\square$

Table of Contents

$\mathbb{K}$ Control-Flow Graphs

Definition

A $\mathbb{K}$ control-flow graph (abbreviated KCFG) $G = (V, E_r, E_a, E_s)$ is a finite directed graph with three types of edges where

- the vertex set V is a set of patterns;
- $E_r \subseteq V \times V$ is called the *rewriting relation*;
- $E_a \subseteq V \times V$ is called the *abstracting relation*;
- $E_s \subseteq V \times V$ is called the *splitting relation*.

We write $\varphi \rightsquigarrow_r \psi$ ($\varphi \rightsquigarrow_a \psi$ and $\varphi \rightsquigarrow_s \psi$, resp.) for the three types of edges.

Definition

A KCFG $G = (V, E_r, E_a, E_s)$ is *sound* if

1. (rewriting edges) $\vdash \varphi \Rightarrow^{\forall}_{*} \psi$ for all $\varphi \rightsquigarrow_r \psi$;
2. (abstracting edges) $\vdash \varphi \to \exists \bar{x} \,.\, \psi$ for all $\varphi \rightsquigarrow_a \psi$, where $\bar{x} = FV(\psi) \setminus FV(\varphi)$.
3. (splitting edges) $\vdash \varphi \leftrightarrow \psi_1 \vee \cdots \vee \psi_n$ for all $\varphi, \psi_1, \ldots, \psi_n$ such that $\psi_1, \ldots, \psi_n$ are all the E_s-successors of φ in G.

In the current implementation of the $\mathbb{K}$ summarizer, every pattern φ has the form $t \wedge p$, called a constrained term, where t is a term and p is a predicate pattern.
We call such KCFGs *regular*.

Definition

A regular KCFG is one that satisfies the following conditions:

1. It is sound.
2. All patterns are constrained terms.
3. For all splitting edges $t_1 \wedge p_1 \rightsquigarrow_s t_2 \wedge p_2$ we have $t_1 \equiv t_2$ and $\vdash p_2 \rightarrow p_1$.

A key property about a regular KCFG is that it is a complete summary of all possible concrete executions.

Definition

Given a KCFG, a flow π is a sequence of the vertex patterns following the three types of edges. A flow instance (π, τ) is a flow π associated with a substitution τ with $\mathrm{dom}(\tau) = FV(\pi)$.

Theorem (Tentative for now)

For any concrete execution of the original transition system, there exists a corresponding flow instance (π, τ) in the KCFG.

Table of Contents

Substitution Patterns

Definition

Given a substitution

$$\tau \equiv [\varphi_1/x_1, \ldots, \varphi_n/x_n]$$

we define a corresponding *substitution pattern*

$$\varphi^\tau \equiv (x_1 = \varphi_1) \wedge \cdots \wedge (x_n = \varphi_n)$$

Lemma

$\vdash \varphi^\tau \to (\psi = \psi\tau)$.

Proof.

This (derived) proof rule is called (Equality Elimination). $\square$

Matching

Definition

Let φ and ψ be two patterns. We say that φ *matches* ψ if

$$\vdash \varphi \to \exists FV(\psi) \,.\, \psi$$

We say that $\{\sigma_1, \ldots, \sigma_n\}$ is a *complete solution* to the matching problem

$$\varphi_1 \lhd_? \psi_1, \ldots, \varphi_m \lhd_? \psi_m$$

if

$$\vdash \left(\bigwedge_{i=1}^{m} \varphi_i \subseteq \psi_i \right) \leftrightarrow \varphi^{\tau_1} \vee \cdots \vee \varphi^{\tau_n}$$

Matching

Lemma

For terms t and s, the matching logic definition of matching coincides with the classical term matching.

Lemma

φ matches ψ if and only if

$$\vdash (\exists FV(\varphi) \,.\, \varphi) \subseteq (\exists FV(\psi) \,.\, \psi)$$

Unification

Definition

Let φ and ψ be two patterns. We say that φ *unifies* with ψ if

$$\vdash \lceil (\exists FV(\varphi) \,.\, \varphi) \wedge (\exists FV(\psi) \,.\, \psi) \rceil$$

We say that $\{\sigma_1, \ldots, \sigma_n\}$ is a *complete solution* to the unification problem

$$\varphi_1 =_? \psi_1, \ldots, \varphi_m =_? \psi_m$$

if

$$\vdash \left(\bigwedge_{i=1}^{m} \lceil \varphi_i \wedge \psi_i \rceil \right) \leftrightarrow \varphi^{\tau_1} \vee \cdots \vee \varphi^{\tau_n}$$

Unification

Lemma

For terms t and s, the matching logic definition of unification coincides with classical term unification.

Both definitions of matching and unification in matching logic work with underlying theories, in which case we obtain the classical matching/unification modulo theories.