$\mathbb{K}$ Summarizer: Foundations

Runtime Verification, Inc.

June 30, 2022

Purpose

Formalize the key concepts of the $\mathbb{K}$ summarizer in matching logic.

- $\mathbb{K}$ control-flow graphs
- Basic blocks
- Soundness and completeness

We work under the following assumptions:

- Formal semantics are deterministic (i.e., $\vdash \bullet\varphi \to \circ\varphi$)
- All patterns are constrained terms: $t \wedge p$

$\mathbb{K}$ 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 constrained terms;
- $E_r \subseteq V \times V$ 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.

Rewriting Edges

$t_1 \wedge p_1 \rightsquigarrow_r t_2 \wedge p_2$ means

- finite- and at-least-one-step rewriting

$$\vdash t_1 \wedge p_1 \to \bullet \diamond (t_2 \wedge p_2) \tag{1}$$

- $\diamond\varphi \equiv \mu X \,.\, \varphi \vee \bullet X$
- The next symbol $\bullet$ enforces "at-least-one-step"
- Thanks to determinism, we can just use the "one-path" operators $\bullet$ and $\diamond$.
- Equation (1) specifies a basic block.
  - All the concrete instances of $t_1 \wedge p_1$ are covered:

$$\vdash \forall \bar{x} \,.\, (t_1 \wedge p_1 \to \bullet \diamond (t_2 \wedge p_2))$$

  - Not getting stuck somewhere in the middle.
  - Determinism (by assumption)

Abstracting Edges

$t_1 \wedge p_1 \rightsquigarrow_a t_2 \wedge p_2$ means

- implication
$$\vdash t_1 \wedge p_1 \to \exists \bar{y} \,.\, t_2 \wedge p_2 \tag{2}$$
where $\bar{y} = FV(rhs) \setminus FV(lhs)$
- The most common case is when $t_1 \equiv t_2[\bar{y}]_{\bar{p}}$ where $\bar{p}$ are the positions of $\bar{y}$ in t_2
- It means that Equation (2) has a witness substitution
$$\pi = [t_{11}/y_1 \dots t_{1n}/y_n]$$
- $t_1 \wedge p_1 \rightsquigarrow_a^{\pi} t_2 \wedge p_2$

Splitting Edges

$t \wedge p \rightsquigarrow_s t \wedge (p \wedge q_i)$ for $i = 1, 2, \ldots, n$

- Complete Cases: $\vdash q_1 \vee \cdots \vee q_n$
- $t \wedge p \rightsquigarrow_s^{q_i} t \wedge (p \wedge q_i)$

Review

A KCFG $G = (V, E_r, E_a, E_s)$ has

- V: a set of constrained terms (nodes)
- $t_1 \wedge p_1 \rightsquigarrow_r t_2 \wedge p_2$: basic block ($\geq 1$ steps)
- $t \wedge p \rightsquigarrow_r^{\pi} t_2 \wedge p$ with a witness substitution π
- $t \wedge p \rightsquigarrow_r^{q_i} t \wedge (p \wedge q_i)$ with a condition q_i
  - $\vdash q_1 \vee \cdots \vee q_n$

Termination Condition

- Φ_T: a (user-provided) termination pattern
  if not provided, Φ_T is $\circ\bot$ (e.g., when `<k> . </k>`)
- $t \wedge p$ has no successors iff $\vdash t \wedge p \rightarrow \Phi_T$

Semantics Derived from a KCFG

Given a KCFG $G = (V, E_r, E_a, E_s)$, we derive a new semantics

- for every $\varphi_1 \rightsquigarrow_s^q \varphi_2 \rightsquigarrow_r \varphi_3$, add

$$\text{rule} \quad \varphi_1 \Rightarrow \varphi_3 \quad \text{requires } q$$

- Let Γ^G be the set of derived semantic rules.

Theorem

For any t and its KCFG G^t,

$$\Gamma^L \vdash t \Rightarrow t' \quad \textit{iff} \quad \Gamma^{G^t} \vdash t \Rightarrow t'$$