SHOVEL: A SAT-based Tool for Information Flow Alarm Classification: Soundness and Completeness Proof

Jong-Gwon Kim, *Seoul Nat'l Univ., Korea*, *jgkim@ropas.snu.ac.kr*
Woosuk Lee, *Georgia Tech, USA*, *woosuk.lee@cc.gatech.edu*
Jaeseung Choi, *Seoul Nat'l Univ., Korea*, *jschoi@ropas.snu.ac.kr*
Chung-Kil Hur, *Seoul Nat'l Univ., Korea*, *gil.hur@sf.snu.ac.kr*
Kwangkeun Yi, *Seoul Nat'l Univ., Korea*, *kwang@ropas.snu.ac.kr*

I. PRELIMINARIES

Call-graphs For a program P, the call-graph $G = (N, E)$ is a *simple directed labeled graph* consisting of (i) N the set of nodes for functions in P; and (ii) $E \subseteq N \times N$ the set of call edges for possible function calls. We exclude self-loops (*i.e.,* edges of the form $f \to f$) from the graph for simplicity; their absence does not sacrifice the expressibility for user constraints.

We can also derive the set E^{-1} of return edges from the set E of call edges by reversing the direction of edges. We write $f \xrightarrow{\mathtt{c}} g$ for $(f, g) \in E$, and $f \xrightarrow{\mathtt{r}} g$ for $(f, g) \in E^{-1}$.

We will henceforth fix the program P and its call-graph G.

Call-Return Paths A call-return path p of length n is a sequence of n consecutive call and return edges, denoted $f_0 \xrightarrow{l_1} f_1 \xrightarrow{l_2} f_2 \cdots \xrightarrow{l_n} f_n$ for $l_i \in \{\mathtt{c}, \mathtt{r}\}$. A node f can be identified with the path f of length 0, and an edge $f \xrightarrow{l} g$ with the path $f \xrightarrow{l} g$ of length 1.

For a path $p = f_0 \xrightarrow{l_1} \cdots \xrightarrow{l_n} f_n$, we define $\mathrm{src}(p)$ the start node, $\mathrm{tgt}(p)$ the end node, $\mathrm{label}(p)$ the sequence of labels, and $\preceq$ the sub-path relation as follows:

$$\mathrm{src}(p) \stackrel{\text{def}}{=} f_0 \qquad \mathrm{tgt}(p) \stackrel{\text{def}}{=} f_n \qquad \mathrm{label}(p) \stackrel{\text{def}}{=} (l_1, \ldots, l_n)$$

$$s \preceq p \;\text{ iff }\; s = f_i \xrightarrow{l_{i+1}} \cdots \xrightarrow{l_j} f_j \text{ with } 0 \le i \le j \le n$$

For paths $p = f_0 \xrightarrow{l_1} \cdots \xrightarrow{l_n} f_n$ and $p' = f'_0 \xrightarrow{l'_1} \cdots \xrightarrow{l'_n} f'_m$, we define their composition $p \cdot p'$ as follows:

$$p \cdot p' \stackrel{\text{def}}{=} \begin{cases} f_0 \xrightarrow{l_1} \cdots \xrightarrow{l_n} f_n \xrightarrow{l'_1} \cdots \xrightarrow{l'_n} f'_m & \text{if } f_n = f'_0 \\ \text{undef} & \text{otherwise} \end{cases}$$

Well-formed Paths We define well-formed paths by excluding infeasible paths such as $f \xrightarrow{\mathtt{c}} g \xrightarrow{\mathtt{r}} h$ with $f \neq h$. To this end, we first define the set $[\![M]\!]$ of *matched* sequences of labels by the following context-free grammar:

$$\begin{array}{rcl} M & \to & \epsilon \\ & | & \mathtt{c}\ M\ \mathtt{r}\ M \end{array}$$

An example of matched labels is $(\mathtt{c}, \mathtt{c}, \mathtt{r}, \mathtt{c}, \mathtt{r}, \mathtt{r})$.

A path p is *well-formed* if every sub-path of it with matched labels has the same source and target node:

$$\forall s \preceq p.\ \ \mathrm{label}(s) \in [\![M]\!] \implies \mathrm{src}(s) = \mathrm{tgt}(s)\ .$$

A *matched* path is a well-formed path whose labels are matched. We define $\mathrm{WfPath}(f_{\mathrm{src}}, f_{\mathrm{snk}})$ as the set of all well-formed paths from f_{src} to f_{snk} in the graph G.

II. BACKBONE-BRANCH DECOMPOSITION

We define the backbone-branch decomposition.

Definition 1 (Decomposition)**.** A path p is decomposed into a *return backbone* $(r_1, .., r_n)$, a *call backbone* $(c_1, .., c_m)$ and *branches* $(b_0, .., b_{n+m})$ for $n, m \ge 0$ if the following hold:

1) $\forall i \in \{1, .., n\}.\ \ \mathrm{label}(r_i) = (\mathtt{r})$
2) $\forall i \in \{1, .., m\}.\ \ \mathrm{label}(c_i) = (\mathtt{c})$
3) $\forall i \in \{0, .., n+m\}.\ \ b_i$ well-matched path
4) $p = b_0 \cdot r_1 \cdot b_1 \cdots r_n \cdot b_n \cdot c_1 \cdot b_{n+1} \cdots c_m \cdot b_{n+m}$

Theorem 1. Every well-formed path p is uniquely decomposed into $(r_1, .., r_n), (c_1, .., c_m), (b_0, .., b_{n+m})$.

Proof: We prove the goal by induction on the length of p.

• Base case: $\mathrm{length}(p) = 0$:
The path p is uniquely decomposed into $(), (), (p)$.

• Inductive case: $\mathrm{length}(p) > 0$:
Suppose $p = e \cdot p'$ with $e = f \xrightarrow{l} g$. By the induction hypothesis applied to p', the path p' is uniquely decomposed into $(r_1, .., r_n), (c_1, .., c_m), (b_0, .., b_{n+m})$.

− When e is a return edge (*i.e.,* $l = \mathtt{r}$):
The path p is decomposed into

$$(e, r_1, .., r_n), (c_1, .., c_m), (f, b_0, .., b_{n+m}).$$

To show the uniqueness, consider any possible decomposition $(r'_1, .., r'_{n'}), (c'_1, .., c'_{m'}), (b'_0, .., b'_{n'+m'})$ of the path p. Since e is a return edge, we have $r'_1 = e$ and $b'_0 = f$. Thus the path p' is decomposed into $(r'_2, .., r'_{n'}), (c'_1, .., c'_{m'}), (b'_1, .., b'_{n'+m'})$. Since p' is uniquely decomposed, we have

$(r'_2,..,r'_{n'}) = (r_1,..,r_n)$, $(c'_1,..,c'_{m'}) = (c_1,..,c_m)$ and $(b'_1,..,b'_{n'+m'}) = (b_0,..,b_{n+m})$. Therefore, it follows that $(r'_1,..,r'_{n'}) = (e,r_1,..,r_n)$, $(c'_1,..,c'_{m'}) = (c_1,..,c_m)$ and $(b'_0,..,b'_{n'+m'}) = (f,b_0,..,b_{n+m})$.

– When e is a call edge (*i.e.,* $l = \mathtt{c}$):

1) If $n = 0$, then the path p is decomposed into
$$(),(e,c_1,..,c_m),(f,b_0,..,b_m).$$
2) If $n > 0$, then the path p is decomposed into
$$(r_2,..,r_n),(c_1,..,c_m),(e\cdot b_0\cdot r_1\cdot b_1,b_2,..,b_{n+m}).$$

To show the uniqueness, consider any possible decomposition $(r'_1,..,r'_{n'})$, $(c'_1,..,c'_{m'})$, $(b'_0,..,b'_{n'+m'})$ of the path p. Since e is a call edge, we have the following two cases.

1) When $n' = 0$, $c'_1 = e$ and $b'_0 = f$:
The path p' is decomposed into $()$, $(c'_2,..,c'_{m'})$, $(b'_1,..,b'_{m'})$. Since p' is uniquely decomposed, we have $() = (r_1,..,r_n)$, $(c'_2,..,c'_{m'}) = (c_1,..,c_m)$, $(b'_1,..,b'_{m'}) = (b_0,..,b_{n+m})$ and thus $n = 0$. Therefore, it follows that $(r'_1,..,r'_{n'}) = ()$, $(c'_1,..,c'_{m'}) = (e,c_1,..,c_m)$ and $(b'_0,..,b'_{n'+m'}) = (f,b_0,..,b_m)$.
2) When $b'_0 = e\cdot b''_0\cdot r'_0\cdot b'''_0$:
The path p' is decomposed into $(r'_0,r'_1,..,r'_{n'})$, $(c'_1,..,c'_{m'})$, $(b''_0,b'''_0,b'_1,..,b'_{n'+m'})$. Since p' is uniquely decomposed, we have $(r'_0,r'_1,..,r'_{n'}) = (r_1,..,r_n)$, $(c'_1,..,c'_{m'}) = (c_1,..,c_m)$, $(b''_0,b'''_0,b'_1,..,b'_{n'+m'}) = (b_0,..,b_{n+m})$ and thus $n > 0$. Therefore, it follows that $(r'_1,..,r'_{n'}) = (r_2,..,r_n)$, $(c'_1,..,c'_{m'}) = (c_1,..,c_m)$ and $(b'_0,..,b'_{n'+m'}) = (e\cdot b_0\cdot r_1\cdot b_1,b_2,..,b_{n+m})$.

■

III. Soundness Proof of Our Encoding

We formally prove the soundness of our encoding formula Φ.

Definition 2. The *image* $\alpha(p)$ of a well-formed path p is
$$(\{r_1,..,r_n\},\{c_1,..,c_m\},\mathrm{edge}^{\mathtt{c}}(b_0)\cup..\cup\mathrm{edge}^{\mathtt{c}}(b_{n+m}))$$
where $(r_1,..,r_n)$, $(c_1,..,c_m)$, $(b_0,..,b_{n+m})$ is the unique decomposition of p and $\mathrm{edge}^{\mathtt{c}}(b_i)$ denotes the set of call edges in b_i.

Theorem 2 (Soundness)**.** For any well-formed path p from f_{src} to f_{snk} and any node set V, we have $\alpha(p) \models \Phi_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$.

Proof: We prove the goal by induction on the length of p.

• Base case: $\mathrm{length}(p) = 0$:
The goal holds since $\alpha(p) = (\emptyset,\emptyset,\emptyset)$ and $f_{\mathrm{src}} = f_{\mathrm{snk}}$.

• Inductive case: $\mathrm{length}(p) > 0$:
For $(r_1,..,r_n)$, $(c_1,..,c_m)$, $(b_0,..,b_{n+m})$ the unique decomposition of p, we do case analysis on b_0.

$$\Phi(V) = \mathrm{Init}(V)\wedge\mathrm{IO}(V)\wedge\mathrm{OI}(V)\wedge\mathrm{BR}(V)$$

$$\mathrm{Init}(V) = \begin{cases} \overset{\bullet\to\circ}{\mathrm{R}}(V)\vee\overset{\bullet\to\circ}{\mathrm{C}}(V) & \text{if } f_{\mathrm{src}}\in V, f_{\mathrm{snk}}\notin V \quad (1)\\ \overset{\circ\to\bullet}{\mathrm{R}}(V)\vee\overset{\circ\to\bullet}{\mathrm{C}}(V) & \text{if } f_{\mathrm{src}}\notin V, f_{\mathrm{snk}}\in V \quad (2)\\ \mathtt{true} & \text{otherwise}\end{cases}$$

$$\mathrm{IO}(V) = \begin{cases} (\overset{\rightarrow\bullet}{\mathrm{R}}(V)\Rightarrow\overset{\bullet\to\circ}{\mathrm{R}}(V)\vee\overset{\bullet\to\circ}{\mathrm{C}}(V))\wedge & \text{if } f_{\mathrm{snk}}\notin V \quad (3)\\ \quad(\overset{\rightarrow\bullet}{\mathrm{C}}(V)\Rightarrow\overset{\bullet\to\circ}{\mathrm{C}}(V)) & \\ \mathtt{true} & \text{otherwise}\end{cases}$$

$$\mathrm{OI}(V) = \begin{cases} (\overset{\bullet\to}{\mathrm{R}}(V)\Rightarrow\overset{\circ\to\bullet}{\mathrm{R}}(V))\wedge & \text{if } f_{\mathrm{src}}\notin V \quad (4)\\ \quad(\overset{\bullet\to}{\mathrm{C}}(V)\Rightarrow\overset{\circ\to\bullet}{\mathrm{R}}(V)\vee\overset{\circ\to\bullet}{\mathrm{C}}(V)) & \\ \mathtt{true} & \text{otherwise}\end{cases}$$

$$\mathrm{BR}(V) = \begin{cases} \overset{\bullet\to}{\mathrm{B}}(V)\Rightarrow\overset{\circ\to\bullet}{\mathrm{B}}(V)\vee\mathrm{RC}(V) & \text{if } f_{\mathrm{src}},f_{\mathrm{snk}}\notin V \quad (5)\\ \mathtt{true} & \text{otherwise}\end{cases}$$

where $\mathrm{RC}(V) = \overset{\rightarrow\bullet}{\mathrm{R}}(V)\vee\overset{\bullet\to}{\mathrm{R}}(V)\vee\overset{\rightarrow\bullet}{\mathrm{C}}(V)\vee\overset{\bullet\to}{\mathrm{C}}(V)$

For $(L,l)\in\{(R,\mathtt{r}),(C,\mathtt{c}),(B,\mathtt{b})\}$,
$$\begin{aligned}\overset{\bullet\to}{\mathrm{L}}(V) &= \bigvee\{x^l_{v,w}\mid v\in V\}\\ \overset{\rightarrow\bullet}{\mathrm{L}}(V) &= \bigvee\{x^l_{v,w}\mid w\in V\}\\ \overset{\bullet\to\circ}{\mathrm{L}}(V) &= \bigvee\{x^l_{v,w}\mid v\in V, w\notin V\}\\ \overset{\circ\to\bullet}{\mathrm{L}}(V) &= \bigvee\{x^l_{v,w}\mid v\notin V, w\in V\}\end{aligned}$$

Figure 1. Definition of $\Phi(V)$ for source f_{src} and sink f_{snk}.

– When $b_0 = f_{\mathrm{src}}$ (*i.e.,* of length 0):
Suppose $p = e\cdot p'$ with $e = f_{\mathrm{src}}\xrightarrow{l} g$. By the induction hypothesis applied to the path p', we have
$$\alpha(p')\models\Phi_{g,f_{\mathrm{snk}}}(V)\ .$$

We also have that
$$\begin{aligned}\alpha(p) &= \alpha(p')\cup(\emptyset,\{e\},\emptyset) \quad \text{if } n=0;\\ \alpha(p) &= \alpha(p')\cup(\{e\},\emptyset,\emptyset) \quad \text{otherwise.}\end{aligned}$$

We can prove this subgoal as follows. When $n = 0$, we have $p = c_1\cdot b_1\cdots c_m\cdot b_m$ and thus have $e = c_1$ and $p' = b_1\cdots c_m\cdot b_m$. Since p' is decomposed into $()$, $(c_2,..,c_m)$, $(b_1,..,b_m)$, the subgoal holds. When $n > 0$, we have $p = r_1\cdot b_1\cdots r_n\cdot b_n\cdot c_1\cdot b_{n+1}\cdots c_m\cdot b_{n+m}$ and thus have $e = r_1$ and $p' = b_1\cdots r_n\cdot b_n\cdot c_1\cdot b_{n+1}\cdots c_m\cdot b_{n+m}$. Since p' is decomposed into $(r_2,..,r_n)$, $(c_1,..,c_m)$, $(b_1,..,b_{n+m})$, the subgoal holds.

Then we show $\alpha(p)\models\Phi_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$ as follows.

1) $\alpha(p)\models\mathrm{Init}_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$ holds:
Suppose $f_{\mathrm{src}}\in V, f_{\mathrm{snk}}\notin V$.
If $g\notin V$, then $\alpha(p)\models\overset{\bullet\to\circ}{\mathrm{C}}(V)\vee\overset{\bullet\to\circ}{\mathrm{R}}(V)$ holds because of e. If $g\in V$, then $\alpha(p)\models\overset{\bullet\to\circ}{\mathrm{C}}(V)\vee\overset{\bullet\to\circ}{\mathrm{R}}(V)$ follows from $\alpha(p')\models\overset{\bullet\to\circ}{\mathrm{C}}(V)\vee\overset{\bullet\to\circ}{\mathrm{R}}(V)$, which follows from $\alpha(p')\models\mathrm{Init}_{g,f_{\mathrm{snk}}}(V)$.

Suppose $f_{\mathrm{src}}\notin V, f_{\mathrm{snk}}\in V$.

If $g \in V$, then $\alpha(p) \models \overset{\circ\to\bullet}{\mathrm{C}}(V) \vee \overset{\circ\to\bullet}{\mathrm{R}}(V)$ holds because of e. If $g \notin V$, then $\alpha(p) \models \overset{\circ\to\bullet}{\mathrm{C}}(V) \vee \overset{\circ\to\bullet}{\mathrm{R}}(V)$ follows from $\alpha(p') \models \overset{\circ\to\bullet}{\mathrm{C}}(V) \vee \overset{\circ\to\bullet}{\mathrm{R}}(V)$, which follows from $\alpha(p') \models \mathrm{Init}_{g,f_{\mathrm{snk}}}(V)$.

2) $\alpha(p) \models \mathrm{IO}_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$ holds:
Suppose $f_{\mathrm{snk}} \notin V$.
If $g \notin V$, then $\alpha(p) \models \mathrm{IO}_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$ follows from $\alpha(p') \models \mathrm{IO}_{g,f_{\mathrm{snk}}}(V)$ because we have $\mathrm{tgt}(e) = g \notin V$ and thus $\alpha(p) \models \overset{\bullet\to}{\mathrm{C}}(V) \iff \alpha(p') \models \overset{\bullet\to}{\mathrm{C}}(V)$ and $\alpha(p) \models \overset{\bullet\to}{\mathrm{R}}(V) \iff \alpha(p') \models \overset{\bullet\to}{\mathrm{R}}(V)$.
If $g \in V$, then $\alpha(p) \models \overset{\bullet\to}{\mathrm{R}}(V) \Rightarrow \overset{\bullet\to\circ}{\mathrm{C}}(V) \vee \overset{\bullet\to\circ}{\mathrm{R}}(V)$ follows from $\alpha(p) \models \overset{\bullet\to\circ}{\mathrm{C}}(V) \vee \overset{\bullet\to\circ}{\mathrm{R}}(V)$, which follows from $\alpha(p') \models \overset{\bullet\to\circ}{\mathrm{C}}(V) \vee \overset{\bullet\to\circ}{\mathrm{R}}(V)$, which follows from $\alpha(p') \models \mathrm{Init}_{g,f_{\mathrm{snk}}}(V)$. Now we show the subgoal $\alpha(p) \models \overset{\bullet\to}{\mathrm{C}}(V) \Rightarrow \overset{\bullet\to\circ}{\mathrm{C}}(V)$ by case analysis on n. If $n = 0$, then the subgoal follows from $\alpha(p) \models \overset{\bullet\to\circ}{\mathrm{C}}(V)$, which follows from $\alpha(p) \models \overset{\bullet\to\circ}{\mathrm{C}}(V) \vee \overset{\bullet\to\circ}{\mathrm{R}}(V)$ since the return backbone is empty and thus $\overset{\bullet\to\circ}{\mathrm{R}}(V) = \texttt{false}$. If $n > 0$, then e is a return edge and thus the subgoal is equivalent to $\alpha(p') \models \overset{\bullet\to}{\mathrm{C}}(V) \Rightarrow \overset{\bullet\to\circ}{\mathrm{C}}(V)$, which follows from $\alpha(p') \models \mathrm{IO}_{g,f_{\mathrm{snk}}}(V)$.

3) $\alpha(p) \models \mathrm{OI}_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$ holds:
Suppose $f_{\mathrm{src}} \notin V$.
If $g \notin V$, then $\alpha(p) \models \mathrm{OI}_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$ follows from $\alpha(p') \models \mathrm{OI}_{g,f_{\mathrm{snk}}}(V)$ because we have $\mathrm{src}(e) = f_{\mathrm{src}} \notin V$ and thus $\alpha(p) \models \overset{\bullet\!\to}{\mathrm{R}}(V) \iff \alpha(p') \models \overset{\bullet\!\to}{\mathrm{R}}(V)$ and $\alpha(p) \models \overset{\bullet\!\to}{\mathrm{C}}(V) \iff \alpha(p') \models \overset{\bullet\!\to}{\mathrm{C}}(V)$.
If $g \in V$, then $\alpha(p) \models \overset{\bullet\!\to}{\mathrm{C}}(V) \Rightarrow \overset{\circ\to\bullet}{\mathrm{C}}(V) \vee \overset{\circ\to\bullet}{\mathrm{R}}(V)$ follows from $\alpha(p) \models \overset{\circ\to\bullet}{\mathrm{C}}(V) \vee \overset{\circ\to\bullet}{\mathrm{R}}(V)$, which follows from the fact that $\mathrm{src}(e) = f_{\mathrm{src}} \notin V$ and $\mathrm{tgt}(e) = g \in V$. Now we show the subgoal $\alpha(p) \models \overset{\bullet\!\to}{\mathrm{R}}(V) \Rightarrow \overset{\circ\to\bullet}{\mathrm{R}}(V)$ by case analysis on n. If $n = 0$, then the subgoal holds trivially since the return backbone is empty and thus $\overset{\bullet\!\to}{\mathrm{R}}(V) = \overset{\circ\to\bullet}{\mathrm{R}}(V) = \texttt{false}$. If $n > 0$, then the subgoal follows from $\alpha(p) \models \overset{\circ\to\bullet}{\mathrm{R}}(V)$, which follows from the fact that e is a return edge with $\mathrm{src}(e) = f_{\mathrm{src}} \notin V$ and $\mathrm{tgt}(e) = g \in V$.

4) $\alpha(p) \models \mathrm{BR}_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$ holds:
Suppose $f_{\mathrm{src}} \notin V, f_{\mathrm{snk}} \notin V$.
If $g \in V$, then $\alpha(p) \models \mathrm{BR}_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$ follows from $\alpha(p) \models \overset{\circ\to\bullet}{\mathrm{R}}(V) \vee \overset{\circ\to\bullet}{\mathrm{C}}(V)$, which follows from the fact that $\mathrm{src}(e) = f_{\mathrm{src}} \notin V, \mathrm{tgt}(e) = g \in V$. If $g \notin V$, then $\alpha(p) \models \mathrm{BR}_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$ follows from $\alpha(p') \models \mathrm{BR}_{g,f_{\mathrm{snk}}}(V)$ because $\alpha(p) \models \overset{\bullet\!\to}{\mathrm{B}}(V) \iff \alpha(p') \models \overset{\bullet\!\to}{\mathrm{B}}(V)$.

– When $b_0 = c_0 \cdot b_0' \cdot r_0 \cdot b_0''$:

1) $\alpha(p) \models X_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$ for $X \in \{\mathrm{Init}, \mathrm{IO}, \mathrm{OI}\}$ holds:
Suppose $p = b_0 \cdot p'$. By the induction hypothesis applied to the path p', we have $\alpha(p') \models \Phi_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$. Since p and p' have the same backbone, $\alpha(p) \models X_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$ is equivalent to $\alpha(p') \models X_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$, which follows from $\alpha(p') \models \Phi_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$.

2) $\alpha(p) \models \mathrm{BR}_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$ holds:
Suppose $f_{\mathrm{src}} \notin V, f_{\mathrm{snk}} \notin V$.
Suppose $c_0 = f_{\mathrm{src}} \xrightarrow{\mathtt{c}} g$. Then we have $r_0 = g \xrightarrow{\mathtt{r}} f_{\mathrm{src}}$.
If $g \in V$, then $\alpha(p) \models \mathrm{BR}_{f_{\mathrm{src}},f_{\mathrm{snk}}}(V)$ follows from $\alpha(p) \models \overset{\circ\to\bullet}{\mathrm{B}}(V)$, which follows from the fact that c_0 is a call edge in the branch b_0 with $\mathrm{src}(c_0) = f_{\mathrm{src}} \notin V$ and $\mathrm{tgt}(c_0) = g \in V$.
If $g \notin V$, suppose $p = c_0 \cdot p'$. By the induction hypothesis applied to the path p', we have $\alpha(p') \models \Phi_{g,f_{\mathrm{snk}}}(V)$. We also have $\alpha(p) \cup (\{r_0\}, \emptyset, \emptyset) = \alpha(p') \cup (\emptyset, \emptyset, \{c_0\})$. Since the source and target nodes of c_0 and r_0 are not in V, $\alpha(p) \models \overset{\bullet\!\to}{\mathrm{B}}(V) \Rightarrow \overset{\circ\to\bullet}{\mathrm{B}}(V) \vee \mathrm{RC}(V)$ is equivalent to $\alpha(p') \models \overset{\bullet\!\to}{\mathrm{B}}(V) \Rightarrow \overset{\circ\to\bullet}{\mathrm{B}}(V) \vee \mathrm{RC}(V)$, which follows from $\alpha(p') \models \Phi_{g,f_{\mathrm{snk}}}(V)$.

■

IV. Sound and Completeness of Shovel

The correctness of Shovel is a direct consequence of the soundness of the encoding formula Φ.

Theorem 3. Shovel(μ) returns a minimal element in the set $\{\, \alpha(p) \mid p \in \mathrm{WfPath}(f_{\mathrm{src}}, f_{\mathrm{snk}}) \wedge \alpha(p) \models \mu \,\}$ if the set is not empty; otherwise, returns UNSAT.

Proof: This directly follows from Theorem 2 because (i) Shovel iterates until it finds a minimal solution using the sound encoding Φ; and (ii) Shovel always terminates since it rules out at least one path at each iteration. ■