AAA616: Program Analysis

Lecture 9 — Control-Flow Analysis

Hakjoo Oh
2022 Fall

Control-Flow Analysis (CFA)

- In functional or object-oriented languages, the program's control flows are not explicit from the program syntax.
- Control-flow analysis is a static analysis that computes for each subexpression the set of functions that it could evaluate to.

Example 1

$$\texttt{((fn x => x}^{\mathbf{1}}\texttt{)}^{\mathbf{2}}\ \texttt{(fn y => y}^{\mathbf{3}}\texttt{)}^{\mathbf{4}}\texttt{)}^{\mathbf{5}}$$

1	$\{\ \texttt{fn y => y}^{\mathbf{3}}\ \}$
2	$\{\ \texttt{fn x => x}^{\mathbf{1}}\ \}$
3	$\emptyset$
4	$\{\ \texttt{fn y => y}^{\mathbf{3}}\ \}$
5	$\{\ \texttt{fn y => y}^{\mathbf{3}}\ \}$
x	$\{\ \texttt{fn y => y}^{\mathbf{3}}\ \}$
y	$\emptyset$

Example 2

$$\texttt{(let g = (fun f x => (f}^{1}\ \texttt{(fn y => y}^{2}\texttt{)}^{3}\texttt{)}^{4}\texttt{)}^{5}\ \texttt{in (g}^{6}\ \texttt{(fn z => z}^{7}\texttt{)}^{8}\texttt{)}^{9}\texttt{)}^{10}$$

1	{ `fun f x => (f (fn y => y))` }
2	$\emptyset$
3	{ `fn y => y` }
4	$\emptyset$
5	{ `fun f x => (f (fn y => y))` }
6	{ `fun f x => (f (fn y => y))` }
7	$\emptyset$
8	{ `fn z => z` }
9	$\emptyset$
10	$\emptyset$
f	{ `fun f x => (f (fn y => y))` }
g	{ `fun f x => (f (fn y => y))` }
x	{ `fn y => y`, `fn z => z` }
y	$\emptyset$
z	$\emptyset$

Language

$$
\begin{array}{rcll}
e & \rightarrow & t^l & \text{expressions (labelled terms)} \\
t & \rightarrow & n & \text{terms (unlabelled expressions)} \\
& \mid & x & \\
& \mid & \texttt{fn } x \texttt{ => } e_0 & \\
& \mid & \texttt{fun } f \ x \texttt{ => } e_0 & \\
& \mid & e_1 \ e_2 & \\
& \mid & \texttt{if } e_0 \texttt{ then } e_1 \texttt{ else } e_2 & \\
& \mid & \texttt{let } x \texttt{ = } e_1 \texttt{ in } e_2 & \\
& \mid & e_1 \ op \ e_2 &
\end{array}
$$

The 0-CFA Analysis

- 0-CFA aims to compute an abstract state of the form:

$$S \in \textbf{State} = (\textbf{Label} \cup \textbf{Var}) \rightarrow \mathbf{2}^{\textbf{Term}}$$

- Two steps:
  1. Generate equations over the state
  2. Solve the equations

Generating Equations

From the program

$$((\texttt{fn x => x}^1)^2\ (\texttt{fn y => y}^3)^4)^5$$

generate the equation as follows:

$$
\begin{array}{l}
\mathcal{C}(((\texttt{fn x => x}^1)^2\ (\texttt{fn y => y}^3)^4)^5) = \{\\
\quad \{\texttt{fn x => x}^1\} \subseteq S(2),\\
\quad S(x) \subseteq S(1),\\
\quad \{\texttt{fn y => y}^3\} \subseteq S(4),\\
\quad S(y) \subseteq S(3),\\
\quad \{\texttt{fn x => x}^1\} \subseteq S(2) \implies S(4) \subseteq S(x),\\
\quad \{\texttt{fn x => x}^1\} \subseteq S(2) \implies S(1) \subseteq S(5),\\
\quad \{\texttt{fn y => y}^3\} \subseteq S(2) \implies S(4) \subseteq S(y),\\
\quad \{\texttt{fn y => y}^3\} \subseteq S(2) \implies S(3) \subseteq S(5),\\
\}
\end{array}
$$

Two types of equations:

$$lhs \subseteq rhs, \qquad \{t\} \subseteq rhs' \implies lhs \subseteq rhs$$

Generating Equations

$$
\begin{array}{rcl}
\mathcal{C}(n^l) & = & \emptyset \\
\mathcal{C}(x^l) & = & \{S(x) \subseteq S(l)\} \\
\mathcal{C}((\texttt{fn } x \texttt{ => } e_0)^l) & = & \{\{\texttt{fn } x \texttt{ => } e_0\} \subseteq S(l)\} \cup \mathcal{C}(e_0) \\
\mathcal{C}((\texttt{fun } f \; x \texttt{ => } e_0)^l) & = & \{\{\texttt{fun } f \; x \texttt{ => } e_0\} \subseteq S(l)\} \cup \mathcal{C}(e_0) \\
 & & \cup\{\{\texttt{fun } f \; x \texttt{ => } e_0\} \subseteq S(f)\} \\
\mathcal{C}((t_1^{l_1} \; t_2^{l_2})^l) & = & \mathcal{C}(t_1^{l_1}) \cup \mathcal{C}(t_2^{l_2}) \\
 & & \cup\{\{t\} \subseteq S(l_1) \implies S(l_2) \subseteq S(x) \\
 & & \quad \mid t = (\texttt{fn } x \texttt{ => } t_0^{l_0}) \in \textsf{Term}\} \\
 & & \cup\{\{t\} \subseteq S(l_1) \implies S(l_0) \subseteq S(l) \\
 & & \quad \mid t = (\texttt{fn } x \texttt{ => } t_0^{l_0}) \in \textsf{Term}\} \\
 & & \cup\{\{t\} \subseteq S(l_1) \implies S(l_2) \subseteq S(x) \\
 & & \quad \mid t = (\texttt{fun } f \; x \texttt{ => } t_0^{l_0}) \in \textsf{Term}\} \\
 & & \cup\{\{t\} \subseteq S(l_1) \implies S(l_0) \subseteq S(l) \\
 & & \quad \mid t = (\texttt{fun } f \; x \texttt{ => } t_0^{l_0}) \in \textsf{Term}\} \\
\mathcal{C}((\texttt{if } t_0^{l_0} \texttt{ then } t_1^{l_1} \texttt{ else } t_2^{l_2})^l) & = & \\
\mathcal{C}((\texttt{let } x\texttt{=}t_1^{l_1} \texttt{ in } t_2^{l_2})^l) & = & \\
\mathcal{C}((t_1^{l_1} \; op \; t_2^{l_2})^l) & = &
\end{array}
$$

Solving the Equations

$$
\begin{array}{l}
\mathcal{C}(((\texttt{fn x => x}^1)^2\ (\texttt{fn y => y}^3)^4)^5) = \{ \\
\quad \{\texttt{fn x => x}^1\} \subseteq S(2), \\
\quad S(x) \subseteq S(1), \\
\quad \{\texttt{fn y => y}^3\} \subseteq S(4), \\
\quad S(y) \subseteq S(3), \\
\quad \{\texttt{fn x => x}^1\} \subseteq S(2) \implies S(4) \subseteq S(x), \\
\quad \{\texttt{fn x => x}^1\} \subseteq S(2) \implies S(1) \subseteq S(5), \\
\quad \{\texttt{fn y => y}^3\} \subseteq S(2) \implies S(4) \subseteq S(y), \\
\quad \{\texttt{fn y => y}^3\} \subseteq S(2) \implies S(3) \subseteq S(5), \\
\}
\end{array}
$$

1	$\emptyset$
2	$\emptyset$
3	$\emptyset$
4	$\emptyset$
5	$\emptyset$
x	$\emptyset$
y	$\emptyset$

1	$\emptyset$
2	{ fn x => x^1 }
3	$\emptyset$
4	{ fn y => y^3 }
5	$\emptyset$
x	$\emptyset$
y	$\emptyset$

1	{ fn y => y^3 }
2	{ fn x => x^1 }
3	$\emptyset$
4	{ fn y => y^3 }
5	$\emptyset$
x	{ fn y => y^3 }
y	$\emptyset$

1	{ fn y => y^3 }
2	{ fn x => x^1 }
3	$\emptyset$
4	{ fn y => y^3 }
5	{ fn y => y^3 }
x	{ fn y => y^3 }
y	$\emptyset$

Solving the Equation

$$
\begin{array}{l}
\textbf{solve}(C, S) = \\
\quad \text{let } S' = \textbf{update}(C, S) \\
\quad \text{if } \forall a. S'(a) \subseteq S(a) \text{ then } S \\
\quad \text{else } \textbf{solve}(C, S')
\end{array}
$$

$$
\begin{array}{l}
\textbf{update}(C, S) = \\
\quad \text{for } c \text{ in } C: \\
\quad\quad \text{if } c = (\{t\} \subseteq S(a)): \\
\quad\quad\quad S(a) := S(a) \cup \{t\} \\
\quad\quad \text{if } c = (S(a_1) \subseteq S(a_2)): \\
\quad\quad\quad S(a_2) := S(a_2) \cup S(a_1) \\
\quad\quad \text{if } c = (\{t\} \subseteq S(a_1) \implies S(a_2) \subseteq S(a_3)): \\
\quad\quad\quad \text{if } t \in S(a_1) \text{ then } S(a_3) := S(a_3) \cup S(a_2) \\
\quad \text{return } S
\end{array}
$$

Limitation

(let f = (fn x => $\mathtt{x}^{\mathbf{1}}$)$^{\mathbf{2}}$ in ((f$^{\mathbf{3}}$ f$^{\mathbf{4}}$)$^{\mathbf{5}}$ (fn y => $\mathtt{y}^{\mathbf{6}}$)$^{\mathbf{7}}$)$^{\mathbf{8}}$)$^{\mathbf{9}}$

1	{ fn x => $\mathtt{x}^{\mathbf{1}}$, fn y => $\mathtt{y}^{\mathbf{6}}$ }
2	{ fn x => $\mathtt{x}^{\mathbf{1}}$ }
3	{ fn x => $\mathtt{x}^{\mathbf{1}}$ }
4	{ fn x => $\mathtt{x}^{\mathbf{1}}$ }
5	{ fn x => $\mathtt{x}^{\mathbf{1}}$, fn y => $\mathtt{y}^{\mathbf{6}}$ }
6	{ fn y => $\mathtt{y}^{\mathbf{6}}$ }
7	{ fn y => $\mathtt{y}^{\mathbf{6}}$ }
8	{ fn x => $\mathtt{x}^{\mathbf{1}}$, fn y => $\mathtt{y}^{\mathbf{6}}$ }
9	{ fn x => $\mathtt{x}^{\mathbf{1}}$, fn y => $\mathtt{y}^{\mathbf{6}}$ }
f	{ fn x => $\mathtt{x}^{\mathbf{1}}$ }
x	{ fn x => $\mathtt{x}^{\mathbf{1}}$, fn y => $\mathtt{y}^{\mathbf{6}}$ }
y	{ fn y => $\mathtt{y}^{\mathbf{6}}$ }

The result says that the overall expression (label 9) may evaluate to two functions but only fun y => $\mathtt{y}^{\mathbf{6}}$ is possible in the real execution.

Summary

- 0-CFA: context-insensitive control-flow analysis.
  1. Derive a set of equations
  2. Solve the equations
- Possible extension: k-CFA