Static Program Analysis

Lecture 6: Dataflow Analysis V (Constant Propagation & Undecidability of MOP Solution)

Thomas Noll

Lehrstuhl für Informatik 2
(Software Modeling and Verification)

RWTH Aachen University

noll@cs.rwth-aachen.de

http://www-i2.informatik.rwth-aachen.de/i2/spa11/

Summer Semester 2011

Outline

The MOP Solution I

- Other solution method for dataflow systems
- MOP = Meet Over all Paths
- Analysis information for block B^l = least upper bound over all paths leading to l

Definition (Paths)

Let $S = (L, E, F, (D, \sqsubseteq), \iota, \varphi)$ be a dataflow system. For every $l \in L$, the set of paths up to l is given by

$$\begin{aligned} Path(l) := \{[l_1, \ldots, l_{k-1}] \mid\ & k \geq 1, l_1 \in E, \\ & (l_i, l_{i+1}) \in F \text{ for every } 1 \leq i \leq k, l_k = l\}. \end{aligned}$$

For a path $p = [l_1, \ldots, l_{k-1}] \in Path(l)$, we define the transfer function $\varphi_p : D \to D$ by

$$\varphi_p := \varphi_{l_{k-1}} \circ \ldots \circ \varphi_{l_1} \circ \mathrm{id}_D$$

(so that $\varphi_{[]} = \mathrm{id}_D$).

The MOP Solution II

Definition (MOP solution)

Let $S = (L, E, F, (D, \sqsubseteq), \iota, \varphi)$ be a dataflow system where $L = \{l_1, \ldots, l_n\}$. The MOP solution for S is determined by

$$\mathrm{mop}(S) := (\mathrm{mop}(l_1), \ldots, \mathrm{mop}(l_n)) \in D^n$$

where, for every $l \in L$,

$$\mathrm{mop}(l) := \bigsqcup \{\varphi_p(\iota) \mid p \in \mathit{Path}(l)\}.$$

Remark:

- $\mathit{Path}(l)$ is generally infinite

$\Longrightarrow$ not clear how to compute $\mathrm{mop}(l)$

- In fact: MOP solution generally undecidable (later)

Outline

Formalizing Constant Propagation Analysis I

The dataflow system $S = (L, E, F, (D, \sqsubseteq), \iota, \varphi)$ is given by

- set of labels $L := L_c$,
- extremal labels $E := \{\text{init}(c)\}$ (forward problem),
- flow relation $F := \text{flow}(c)$ (forward problem),
- complete lattice $(D, \sqsubseteq)$ where
  - $D := \{\delta \mid \delta : Var_c \to \mathbb{Z} \cup \{\bot, \top\}\}$
    - $\delta(x) = z \in \mathbb{Z}$: x has constant value z
    - $\delta(x) = \bot$: x undefined
    - $\delta(x) = \top$: x overdefined (i.e., different possible values)
  - $\sqsubseteq \subseteq D \times D$ defined by pointwise extension of $\bot \sqsubseteq z \sqsubseteq \top$ (for every $z \in \mathbb{Z}$)

Example

$Var_c = \{\texttt{w}, \texttt{x}, \texttt{y}, \texttt{z}\}$,

$\delta_1 = (\underbrace{\bot}_{\texttt{w}}, \underbrace{1}_{\texttt{x}}, \underbrace{2}_{\texttt{y}}, \underbrace{\top}_{\texttt{z}})$, $\delta_2 = (\underbrace{3}_{\texttt{w}}, \underbrace{1}_{\texttt{x}}, \underbrace{4}_{\texttt{y}}, \underbrace{\top}_{\texttt{z}})$

$\implies \delta_1 \sqcup \delta_2 = (\underbrace{3}_{\texttt{w}}, \underbrace{1}_{\texttt{x}}, \underbrace{\top}_{\texttt{y}}, \underbrace{\top}_{\texttt{z}})$

Dataflow system $S = (L, E, F, (D, \sqsubseteq), \iota, \varphi)$ (continued):

- extremal value $\iota := \delta_\top \in D$ where $\delta_\top(x) := \top$ for every $x \in Var_c$ (i.e., every x has (unknown) default value)
- transfer functions $\{\varphi_l \mid l \in L\}$ defined by

$$\varphi_l(\delta) := \begin{cases} \delta & \text{if } B^l = \texttt{skip} \text{ or } B^l \in BExp \\ \delta[x \mapsto val_\delta(a)] & \text{if } B^l = (x \texttt{ := } a) \end{cases}$$

where

$$\begin{aligned} val_\delta(x) &:= \delta(x) \\ val_\delta(z) &:= z \end{aligned} \qquad val_\delta(a_1 \; op \; a_2) := \begin{cases} z_1 \; op \; z_2 & \text{if } z_1, z_2 \in \mathbb{Z} \\ \bot & \text{if } z_1 = \bot \text{ or } z_2 = \bot \\ \top & \text{otherwise} \end{cases}$$

for $z_1 := val_\delta(a_1)$ and $z_2 := val_\delta(a_2)$

Outline

Example of Constant Propagation Analysis

Example 6.1

Constant Propagation Analysis for

```
c := [x := 1]^1;
     [y := 1]^2;
     [z := 1]^3;
     while [z > 0]^4 do
       [w := x+y]^5;
       if [w = 2]^6 then
         [x := y+2]^7
```

$$
\begin{aligned}
\varphi_1((a,b,c,d)) &= (a,1,c,d)\\
\varphi_2((a,b,c,d)) &= (a,b,1,d)\\
\varphi_3((a,b,c,d)) &= (a,b,c,1)\\
\varphi_4((a,b,c,d)) &= (a,b,c,d)\\
\varphi_5((a,b,c,d)) &= (b+c,b,c,d)\\
\varphi_6((a,b,c,d)) &= (a,b,c,d)\\
\varphi_7((a,b,c,d)) &= (a,c+2,c,d)
\end{aligned}
$$

1. Fixpoint solution (on the board)
2. MOP solution (on the board)

Outline

Undecidability of the MOP Solution

Theorem 6.2 (Undecidability of MOP solution)

The MOP solution for Constant Propagation is undecidable.

Undecidability of the MOP Solution

Theorem 6.2 (Undecidability of MOP solution)

The MOP solution for Constant Propagation is undecidable.

Proof.

Based on undecidability of Modified Post Correspondence Problem:
Let Γ be some alphabet, $n \in \mathbb{N}$, and $u_1, \ldots, u_n, v_1, \ldots, v_n \in \Gamma^+$.
Do there exist $i_1, \ldots, i_m \in \{1, \ldots, n\}$ with $m \geq 1$ and $i_1 = 1$ such that $u_{i_1} u_{i_2} \ldots u_{i_m} = v_{i_1} v_{i_2} \ldots v_{i_m}$?

(on the board) $\square$