

Semantics and Verification of Software

Lecture 20: Dataflow Analysis VII (MOP vs. Fixpoint 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/svsw08/`

Winter semester 2008/09

- 1 Repetition: MOP Solution
- 2 Repetition: Constant Propagation
- 3 Undecidability of the MOP Solution
- 4 MOP vs. Fixpoint Solution

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

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

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

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

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

Definition (MOP solution)

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

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

where, for every $l \in L$,

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

Remark:

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

⇒ not clear how to compute $\text{mop}(l)$

- In fact: MOP solution generally undecidable (later)

- 1 Repetition: MOP Solution
- 2 Repetition: Constant Propagation
- 3 Undecidability of the MOP Solution
- 4 MOP vs. Fixpoint Solution

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
 - $D := \{\delta \mid \delta : \text{Var}_c \rightarrow \mathbb{Z} \cup \{\perp, \top\}\}$
 - $\delta(x) = z \in \mathbb{Z}$: x has **constant value** z
 - $\delta(x) = \perp$: x **undefined**
 - $\delta(x) = \top$: x **overdefined** (i.e., different possible values)
 - $\sqsubseteq \subseteq D \times D$ defined by pointwise extension of $\perp \sqsubseteq z \sqsubseteq \top$ (for every $z \in \mathbb{Z}$)

Example

$$\begin{aligned}\text{Var}_c &= \{w, x, y, z\}, \\ \delta_1 &= (\underbrace{\perp}_w, \underbrace{1}_x, \underbrace{2}_y, \underbrace{\top}_z), \quad \delta_2 = (\underbrace{3}_w, \underbrace{1}_x, \underbrace{4}_y, \underbrace{\top}_z) \\ \implies \delta_1 \sqcup \delta_2 &= (\underbrace{3}_w, \underbrace{1}_x, \underbrace{\top}_y, \underbrace{\top}_z)\end{aligned}$$

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 $x \in \text{Var}_c$,
- transfer functions $\{\varphi_l \mid l \in L\}$ defined by

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

where

$$\begin{aligned} \mathfrak{A}[[x]]\delta &:= \delta(x) \\ \mathfrak{A}[[z]]\delta &:= z \end{aligned} \quad \mathfrak{A}[[a_1 \text{ op } a_2]]\delta := \begin{cases} z_1 \text{ op } z_2 & \text{if } z_1, z_2 \in \mathbb{Z} \\ \perp & \text{if } z_1 = \perp \text{ or } z_2 = \perp \\ \top & \text{otherwise} \end{cases}$$

if $z_1 := \mathfrak{A}[[a_1]]\delta$ and $z_2 := \mathfrak{A}[[a_2]]\delta$

Example

If $\delta = (\underbrace{\perp}_w, \underbrace{1}_x, \underbrace{2}_y, \underbrace{\top}_z)$, then

$$\varphi_l(\delta) = \begin{cases} (\underbrace{0}_w, \underbrace{1}_x, \underbrace{2}_y, \underbrace{\top}_z) & \text{if } B^l = (w := 0) \\ (\underbrace{3}_w, \underbrace{1}_x, \underbrace{2}_y, \underbrace{\top}_z) & \text{if } B^l = (w := y+1) \\ (\underbrace{\perp}_w, \underbrace{1}_x, \underbrace{2}_y, \underbrace{\top}_z) & \text{if } B^l = (w := w+x) \\ (\underbrace{\top}_w, \underbrace{1}_x, \underbrace{2}_y, \underbrace{\top}_z) & \text{if } B^l = (w := z+2) \end{cases}$$

- 1 Repetition: MOP Solution
- 2 Repetition: Constant Propagation
- 3 Undecidability of the MOP Solution
- 4 MOP vs. Fixpoint Solution

Undecidability of the MOP Solution

Theorem 20.1 (Undecidability of MOP solution)

The MOP solution for Constant Propagation is undecidable.

Undecidability of the MOP Solution

Theorem 20.1 (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 $u_1, \dots, u_n, v_1, \dots, v_n \in \Gamma^+$.

Does there exist $i_1, \dots, i_m \in \{1, \dots, n\}$ with $m \geq 1$ and $i_1 = 1$ such that $u_{i_1} u_{i_2} \dots u_{i_m} = v_{i_1} v_{i_2} \dots v_{i_m}$?

(on the board)



- 1 Repetition: MOP Solution
- 2 Repetition: Constant Propagation
- 3 Undecidability of the MOP Solution
- 4 MOP vs. Fixpoint Solution

Theorem 20.2 (MOP vs. Fixpoint Solution)

Let $S = (L, E, F, (D, \sqsubseteq), \iota, \varphi)$ be a dataflow system. Then

$$\text{mop}(S) \sqsubseteq \text{fix}(\Phi_S)$$

MOP vs. Fixpoint Solution I

Theorem 20.2 (MOP vs. Fixpoint Solution)

Let $S = (L, E, F, (D, \sqsubseteq), \iota, \varphi)$ be a dataflow system. Then

$$\text{mop}(S) \sqsubseteq \text{fix}(\Phi_S)$$

Proof.

on the board



MOP vs. Fixpoint Solution I

Theorem 20.2 (MOP vs. Fixpoint Solution)

Let $S = (L, E, F, (D, \sqsubseteq), \iota, \varphi)$ be a dataflow system. Then

$$\text{mop}(S) \sqsubseteq \text{fix}(\Phi_S)$$

Proof.

on the board



The next example shows that both solutions can indeed be different.

Example 20.3 (Constant Propagation)

```
c := if [z > 0]1 then
    [x := 2;]2
    [y := 3;]3
else
    [x := 3;]4
    [y := 2;]5
    [z := x+y;]6
    [...]7
```

Example 20.3 (Constant Propagation)

```
c := if [z > 0]1 then
    [x := 2;]2
    [y := 3;]3
else
    [x := 3;]4
    [y := 2;]5
    [z := x+y;]6
    [...]7
```

Transfer functions

(for $\delta = (\delta(x), \delta(y), \delta(z)) \in D$):

$$\varphi_1((a, b, c)) = (a, b, c)$$

$$\varphi_2((a, b, c)) = (2, b, c)$$

$$\varphi_3((a, b, c)) = (a, 3, c)$$

$$\varphi_4((a, b, c)) = (3, b, c)$$

$$\varphi_5((a, b, c)) = (a, 2, c)$$

$$\varphi_6((a, b, c)) = (a, b, a + b)$$

Example 20.3 (Constant Propagation)

```
c := if [z > 0]1 then
    [x := 2;]2
    [y := 3;]3
else
    [x := 3;]4
    [y := 2;]5
    [z := x+y;]6
    [...]7
```

Transfer functions

(for $\delta = (\delta(x), \delta(y), \delta(z)) \in D$):

$$\varphi_1((a, b, c)) = (a, b, c)$$

$$\varphi_2((a, b, c)) = (2, b, c)$$

$$\varphi_3((a, b, c)) = (a, 3, c)$$

$$\varphi_4((a, b, c)) = (3, b, c)$$

$$\varphi_5((a, b, c)) = (a, 2, c)$$

$$\varphi_6((a, b, c)) = (a, b, a + b)$$

① Fixpoint solution:

$$CP_1 = \iota = (\top, \top, \top)$$

$$CP_2 = \varphi_1(CP_1) = (\top, \top, \top)$$

$$CP_3 = \varphi_2(CP_2) = (2, \top, \top)$$

$$CP_4 = \varphi_1(CP_1) = (\top, \top, \top)$$

$$CP_5 = \varphi_2(CP_2) = (3, \top, \top)$$

$$CP_6 = \varphi_3(CP_3) \sqcup \varphi_5(CP_5)$$

$$= (2, 3, \top) \sqcup (3, 2, \top) = (\top, \top, \top)$$

$$CP_7 = \varphi_6(CP_6) = (\top, \top, \top)$$

Example 20.3 (Constant Propagation)

```
c := if [z > 0]1 then
    [x := 2;]2
    [y := 3;]3
else
    [x := 3;]4
    [y := 2;]5
    [z := x+y;]6
    [...]7
```

Transfer functions

(for $\delta = (\delta(x), \delta(y), \delta(z)) \in D$):

$$\varphi_1((a, b, c)) = (a, b, c)$$

$$\varphi_2((a, b, c)) = (2, b, c)$$

$$\varphi_3((a, b, c)) = (a, 3, c)$$

$$\varphi_4((a, b, c)) = (3, b, c)$$

$$\varphi_5((a, b, c)) = (a, 2, c)$$

$$\varphi_6((a, b, c)) = (a, b, a + b)$$

① Fixpoint solution:

$$CP_1 = \iota = (\top, \top, \top)$$

$$CP_2 = \varphi_1(CP_1) = (\top, \top, \top)$$

$$CP_3 = \varphi_2(CP_2) = (2, \top, \top)$$

$$CP_4 = \varphi_1(CP_1) = (\top, \top, \top)$$

$$CP_5 = \varphi_2(CP_2) = (3, \top, \top)$$

$$CP_6 = \varphi_3(CP_3) \sqcup \varphi_5(CP_5)$$

$$= (2, 3, \top) \sqcup (3, 2, \top) = (\top, \top, \top)$$

$$CP_7 = \varphi_6(CP_6) = (\top, \top, \top)$$

② MOP solution:

$$\text{mop}(7) = \varphi_{[1,2,3,6]}(\top, \top, \top) \sqcup$$

$$\varphi_{[1,4,5,6]}(\top, \top, \top)$$

$$= (2, 3, 5) \sqcup (3, 2, 5)$$

$$= (\top, \top, 5)$$

A sufficient criterion for the coincidence of MOP and Fixpoint Solution is the distributivity of the transfer functions.

Definition 20.4 (Distributivity)

- Let $(D, \sqsubseteq)$ and $(D', \sqsubseteq')$ be complete lattices, and let $F : D \rightarrow D'$. F is called **distributive (w.r.t. $(D, \sqsubseteq)$ and $(D', \sqsubseteq')$)** if, for every $d_1, d_2 \in D$,

$$F(d_1 \sqcup_D d_2) = F(d_1) \sqcup_{D'} F(d_2).$$

A sufficient criterion for the coincidence of MOP and Fixpoint Solution is the distributivity of the transfer functions.

Definition 20.4 (Distributivity)

- Let $(D, \sqsubseteq)$ and $(D', \sqsubseteq')$ be complete lattices, and let $F : D \rightarrow D'$. F is called **distributive (w.r.t. $(D, \sqsubseteq)$ and $(D', \sqsubseteq')$)** if, for every $d_1, d_2 \in D$,

$$F(d_1 \sqcup_D d_2) = F(d_1) \sqcup_{D'} F(d_2).$$

- A dataflow system $S = (L, E, F, (D, \sqsubseteq), \iota, \varphi)$ is called **distributive** if every $\varphi_l : D \rightarrow D$ is so.

Example 20.5

- ① The Available Expressions dataflow system is distributive:

$$\begin{aligned}\varphi_l(A_1 \sqcup A_2) &= ((A_1 \cap A_2) \setminus \text{kill}_{\text{AE}}(B^l)) \cup \text{gen}_{\text{AE}}(B^l) \\ &= ((A_1 \setminus \text{kill}_{\text{AE}}(B^l)) \cup \text{gen}_{\text{AE}}(B^l)) \cap \\ &\quad ((A_2 \setminus \text{kill}_{\text{AE}}(B^l)) \cup \text{gen}_{\text{AE}}(B^l)) \\ &= \varphi_l(A_1) \sqcup \varphi_l(A_2)\end{aligned}$$

Example 20.5

- ① The Available Expressions dataflow system is distributive:

$$\begin{aligned}\varphi_l(A_1 \sqcup A_2) &= ((A_1 \cap A_2) \setminus \text{kill}_{\text{AE}}(B^l)) \cup \text{gen}_{\text{AE}}(B^l) \\ &= ((A_1 \setminus \text{kill}_{\text{AE}}(B^l)) \cup \text{gen}_{\text{AE}}(B^l)) \cap \\ &\quad ((A_2 \setminus \text{kill}_{\text{AE}}(B^l)) \cup \text{gen}_{\text{AE}}(B^l)) \\ &= \varphi_l(A_1) \sqcup \varphi_l(A_2)\end{aligned}$$

- ② The Live Variables dataflow system is distributive (similar)

Example 20.5

- ❶ The Available Expressions dataflow system is distributive:

$$\begin{aligned}\varphi_l(A_1 \sqcup A_2) &= ((A_1 \cap A_2) \setminus \text{kill}_{\text{AE}}(B^l)) \cup \text{gen}_{\text{AE}}(B^l) \\ &= ((A_1 \setminus \text{kill}_{\text{AE}}(B^l)) \cup \text{gen}_{\text{AE}}(B^l)) \cap \\ &\quad ((A_2 \setminus \text{kill}_{\text{AE}}(B^l)) \cup \text{gen}_{\text{AE}}(B^l)) \\ &= \varphi_l(A_1) \sqcup \varphi_l(A_2)\end{aligned}$$

- ❷ The Live Variables dataflow system is distributive (similar)
- ❸ The Constant Propagation dataflow system is not distributive:

$$\begin{aligned}(\top, \top, \top) &= \varphi_{z:=x+y}((2, 3, \top) \sqcup (3, 2, \top)) \\ &\neq \varphi_{z:=x+y}((2, 3, \top)) \sqcup \varphi_{z:=x+y}((3, 2, \top)) \\ &= (\top, \top, 5)\end{aligned}$$

Theorem 20.6 (MOP vs. Fixpoint Solution)

Let $S = (L, E, F, (D, \sqsubseteq), \iota, \varphi)$ be a distributive dataflow system. Then

$$\text{mop}(S) = \text{fix}(\Phi_S)$$

Theorem 20.6 (MOP vs. Fixpoint Solution)

Let $S = (L, E, F, (D, \sqsubseteq), \iota, \varphi)$ be a distributive dataflow system. Then

$$\text{mop}(S) = \text{fix}(\Phi_S)$$

Proof.

- by showing that $\Phi_S(\text{mop}(S)) = \text{mop}(S)$...
(see [Nielson/Nielson/Hankin 2005, p. 81])
- ... and using $\text{mop}(S) \sqsubseteq \text{fix}(\Phi_S)$ (Theorem 20.2)

