



Static Program Analysis

**Lecture 6: Dataflow Analysis V
(MOP vs. Fixpoint Solution)**

Winter Semester 2016/17

Thomas Noll

Software Modeling and Verification Group

RWTH Aachen University

<https://moves.rwth-aachen.de/teaching/ws-1617/spa/>

Recap: The MOP Solution

Outline of Lecture 6

Recap: The MOP Solution

Recap: Constant Propagation

Example of Constant Propagation Analysis

MOP vs. Fixpoint Solution

Coincidence of MOP and Fixpoint Solution

Undecidability of the MOP Solution

Recap: The MOP 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**
 - = **most precise** information for l (“reference solution”)

Definition (Paths)

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

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

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

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

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

Recap: The MOP Solution

The MOP Solution II

Definition (MOP solution)

Let $S = (Lab, E, F, (D, \sqsubseteq), \iota, \varphi)$ be a dataflow system where $Lab = \{l_1, \dots, l_n\}$.
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 $l \in Lab$,

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

Remark:

- $Path(l)$ is generally infinite
- ⇒ not clear how to compute $\text{mop}(l)$
- In fact: MOP solution generally undecidable (later)

Recap: Constant Propagation

Outline of Lecture 6

Recap: The MOP Solution

Recap: Constant Propagation

Example of Constant Propagation Analysis

MOP vs. Fixpoint Solution

Coincidence of MOP and Fixpoint Solution

Undecidability of the MOP Solution

Recap: Constant Propagation

Goal of Constant Propagation Analysis

Constant Propagation Analysis

The goal of **Constant Propagation Analysis** is to determine, for each program point, whether a variable has a constant value whenever execution reaches that point.

Recap: Constant Propagation

Goal of Constant Propagation Analysis

Constant Propagation Analysis

The goal of **Constant Propagation Analysis** is to determine, for each program point, whether a variable has a constant value whenever execution reaches that point.

Used for **Constant Folding**: replace reference to variable by constant value and evaluate constant expressions

Recap: Constant Propagation

Goal of Constant Propagation Analysis

Constant Propagation Analysis

The goal of **Constant Propagation Analysis** is to determine, for each program point, whether a variable has a constant value whenever execution reaches that point.

Used for **Constant Folding**: replace reference to variable by constant value and evaluate constant expressions

Example (Constant Propagation Analysis)

```
[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  
  end  
end  
end
```


Recap: Constant Propagation

Goal of Constant Propagation Analysis

Constant Propagation Analysis

The goal of **Constant Propagation Analysis** is to determine, for each program point, whether a variable has a constant value whenever execution reaches that point.

Used for **Constant Folding**: replace reference to variable by constant value and evaluate constant expressions

Example (Constant Propagation Analysis)

```
[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  
  end  
end  
end
```

- $y = z = 1$ at labels 4–7

Recap: Constant Propagation

Goal of Constant Propagation Analysis

Constant Propagation Analysis

The goal of **Constant Propagation Analysis** is to determine, for each program point, whether a variable has a constant value whenever execution reaches that point.

Used for **Constant Folding**: replace reference to variable by constant value and evaluate constant expressions

Example (Constant Propagation Analysis)

```
[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  
  end  
end  
end
```

- $y = z = 1$ at labels 4–7
- w, x not constant at labels 4–7

Recap: Constant Propagation

Goal of Constant Propagation Analysis

Constant Propagation Analysis

The goal of **Constant Propagation Analysis** is to determine, for each program point, whether a variable has a constant value whenever execution reaches that point.

Used for **Constant Folding**: replace reference to variable by constant value and evaluate constant expressions

Example (Constant Propagation Analysis)

```
[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  
  end  
end
```

- $y = z = 1$ at labels 4–7
- w, x not constant at labels 4–7
- possible optimisations:

```
[true]4  
[w := x+1]5  
[x := 3]7
```

Recap: Constant Propagation

Formalising Constant Propagation Analysis I

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

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

Example

$$Var_c = \{w, x, y, z\}, \delta_1 = (\underbrace{\perp}_w, \underbrace{1}_x, \underbrace{2}_y, \underbrace{\top}_z), \delta_2 = (\underbrace{3}_w, \underbrace{1}_x, \underbrace{4}_y, \underbrace{\top}_z)$$
$$\implies \delta_1 \sqcup \delta_2 = (\underbrace{3}_w, \underbrace{1}_x, \underbrace{\top}_y, \underbrace{\top}_z)$$

Recap: Constant Propagation

Formalising Constant Propagation Analysis II

Dataflow system $S = (Lab, E, F, (D, \sqsubseteq), \iota, \varphi)$ (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 $\{\varphi_l \mid l \in Lab\}$ defined by

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

where

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

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

Recap: Constant Propagation

Formalising Constant Propagation Analysis III

Example

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

$$\varphi_1(\delta) = \begin{cases} (\underbrace{0}_w, \underbrace{1}_x, \underbrace{2}_y, \underbrace{\top}_z) & \text{if } B^l = (w := 0) \\ (\underbrace{3}_w, \underbrace{1}_x, \underbrace{2}_y, \underbrace{\top}_z) & \text{if } B^l = (w := y+1) \\ (\underbrace{\perp}_w, \underbrace{1}_x, \underbrace{2}_y, \underbrace{\top}_z) & \text{if } B^l = (w := w+x) \\ (\underbrace{\top}_w, \underbrace{1}_x, \underbrace{2}_y, \underbrace{\top}_z) & \text{if } B^l = (w := z+2) \end{cases}$$

Example of Constant Propagation Analysis

Outline of Lecture 6

Recap: The MOP Solution

Recap: Constant Propagation

Example of Constant Propagation Analysis

MOP vs. Fixpoint Solution

Coincidence of MOP and Fixpoint Solution

Undecidability of the MOP Solution

Example of Constant Propagation Analysis

An Example

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
    end
  end
```

$$\varphi_1(a, b, c, d) = (a, 1, c, d)$$

$$\varphi_2(a, b, c, d) = (a, b, 1, d)$$

$$\varphi_3(a, b, c, d) = (a, b, c, 1)$$

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

$$\varphi_5(a, b, c, d) = (b + c, b, c, d)$$

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

$$\varphi_7(a, b, c, d) = (a, c + 2, c, d)$$

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

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

MOP vs. Fixpoint Solution

Outline of Lecture 6

Recap: The MOP Solution

Recap: Constant Propagation

Example of Constant Propagation Analysis

MOP vs. Fixpoint Solution

Coincidence of MOP and Fixpoint Solution

Undecidability of the MOP Solution

MOP vs. Fixpoint Solution

MOP vs. Fixpoint Solution I

Example 6.2 (Constant Propagation)

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

MOP vs. Fixpoint Solution

MOP vs. Fixpoint Solution I

Example 6.2 (Constant Propagation)

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

Transfer functions

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

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

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

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

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

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

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

MOP vs. Fixpoint Solution

MOP vs. Fixpoint Solution I

Example 6.2 (Constant Propagation)

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

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

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

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

$$CP_4 = \varphi_1(CP_3) = (\top, \top, \top)$$

$$CP_5 = \varphi_4(CP_4) = (3, \top, \top)$$

$$CP_6 = \varphi_3(CP_3) \sqcup \varphi_5(CP_5) \\ = (2, 3, \top) \sqcup (3, 2, \top) = (\top, \top, \top)$$

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

MOP vs. Fixpoint Solution

MOP vs. Fixpoint Solution I

Example 6.2 (Constant Propagation)

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

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

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

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

$$CP_4 = \varphi_1(CP_3) = (\top, \top, \top)$$

$$CP_5 = \varphi_4(CP_4) = (3, \top, \top)$$

$$CP_6 = \varphi_3(CP_3) \sqcup \varphi_5(CP_5) \\ = (2, 3, \top) \sqcup (3, 2, \top) = (\top, \top, \top)$$

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

2. MOP solution:

$$\text{mop}(7) = \varphi_{[1,2,3,6]}(\top, \top, \top) \sqcup \\ \varphi_{[1,4,5,6]}(\top, \top, \top) \\ = (2, 3, 5) \sqcup (3, 2, 5) \\ = (\top, \top, 5)$$

MOP vs. Fixpoint Solution

MOP vs. Fixpoint Solution II

Theorem 6.3 (MOP vs. Fixpoint Solution)

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

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

Reminder: by Definition 4.9,

$$\Phi_S : D^n \rightarrow D^n : (d_1, \dots, d_n) \mapsto (d'_1, \dots, d'_n)$$

where $Lab = \{1, \dots, n\}$ and, for each $l \in Lab$,

$$d'_l := \begin{cases} \iota & \text{if } l \in E \\ \bigsqcup \{ \varphi_{l'}(d_{l'}) \mid (l', l) \in F \} & \text{otherwise} \end{cases}$$

MOP vs. Fixpoint Solution

MOP vs. Fixpoint Solution II

Theorem 6.3 (MOP vs. Fixpoint Solution)

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

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

Reminder: by Definition 4.9,

$$\Phi_S : D^n \rightarrow D^n : (d_1, \dots, d_n) \mapsto (d'_1, \dots, d'_n)$$

where $Lab = \{1, \dots, n\}$ and, for each $l \in Lab$,

$$d'_l := \begin{cases} \iota & \text{if } l \in E \\ \bigsqcup \{ \varphi_{l'}(d_{l'}) \mid (l', l) \in F \} & \text{otherwise} \end{cases}$$

Proof.

on the board □

Remark: as Example 6.2 shows, $\text{mop}(S) \neq \text{fix}(\Phi_S)$ is possible

Coincidence of MOP and Fixpoint Solution

Outline of Lecture 6

Recap: The MOP Solution

Recap: Constant Propagation

Example of Constant Propagation Analysis

MOP vs. Fixpoint Solution

Coincidence of MOP and Fixpoint Solution

Undecidability of the MOP Solution

Coincidence of MOP and Fixpoint Solution

Distributivity of Transfer Functions I

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

Definition 6.4 (Distributivity)

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

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

Coincidence of MOP and Fixpoint Solution

Distributivity of Transfer Functions I

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

Definition 6.4 (Distributivity)

- Let (D, \sqsubseteq) and (D', \sqsubseteq') be complete lattices. Function $F : D \rightarrow D'$ 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 $S = (Lab, E, F, (D, \sqsubseteq), \iota, \varphi)$ is called **distributive** if every $\varphi_l : D \rightarrow D$ ($l \in Lab$) is so.

Distributivity of Transfer Functions II

Example 6.5

1. The Available Expressions dataflow system is **distributive**:

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

Distributivity of Transfer Functions II

Example 6.5

1. The Available Expressions dataflow system is **distributive**:

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

2. The Live Variables dataflow system is **distributive**: similarly

Distributivity of Transfer Functions II

Example 6.5

1. The Available Expressions dataflow system is **distributive**:

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

2. The Live Variables dataflow system is **distributive**: similarly
3. The Constant Propagation dataflow system is **not distributive** (cf. Example 6.2):

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

Coincidence of MOP and Fixpoint Solution

Coincidence of MOP and Fixpoint Solution

Theorem 6.6 (MOP vs. Fixpoint Solution)

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

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

Coincidence of MOP and Fixpoint Solution

Coincidence of MOP and Fixpoint Solution

Theorem 6.6 (MOP vs. Fixpoint Solution)

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

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

Proof.

- $\text{mop}(S) \sqsubseteq \text{fix}(\Phi_S)$: Theorem 6.3
- $\text{fix}(\Phi_S) \sqsubseteq \text{mop}(S)$: as $\text{fix}(\Phi_S)$ is the *least* fixpoint of Φ_S , it suffices to show that $\Phi_S(\text{mop}(S)) = \text{mop}(S)$ (on the board)



Undecidability of the MOP Solution

Outline of Lecture 6

Recap: The MOP Solution

Recap: Constant Propagation

Example of Constant Propagation Analysis

MOP vs. Fixpoint Solution

Coincidence of MOP and Fixpoint Solution

Undecidability of the MOP Solution

Undecidability of the MOP Solution

Undecidability of the MOP Solution

Theorem 6.7 (Undecidability of MOP solution)

The MOP solution for Constant Propagation is undecidable.

Undecidability of the MOP Solution

Undecidability of the MOP Solution

Theorem 6.7 (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 $u_1, \dots, u_n, v_1, \dots, v_n \in \Gamma^+$.

Do 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)

