Data-flow Analysis: Lecture 4 & 5
Interprocedural Analysis: Sharir-Pnueli Technique

Deepak D’Souza

Department of Computer Science and Automation
Indian Institute of Science, Bangalore.
Outline of data-flow analysis part

- Overview
- Lattice Theory (mainly Knaster-Tarski fixpoint theorem)
- Some typical data-flow analyses (Kildall’s paper)
- Analysis using Datalog
Sharir-Pnueli framework

- Continuation of Kildall’s framework
- Finite lattice of data values
- Distributive data-flow functions
- Extend to handling programs with procedures
Example program with procedure calls

- We want MOP values at each program point.
- Ignoring procedurally “valid” paths loses too much data.
- Can we compute “MVP” (meet over valid paths) values instead?

```
    a := a-1
    F
    G
    H
    a := a*b
    t := a*b
    call p
    call p
    a := a-1
    print t
    ret
```

Data-flow Analysis: Lecture 4 & 5
Interprocedural Analysis: Sharir-Pnueli Technique – p.4
Overview of Sharir-Pnueli technique

- Goal: compute the MVP data value $x_n$ at each node $n$:
  \[
  x_n = \bigwedge_{\rho \in IVP(r_1,n)} f(0).
  \]

- Observe that $\rho \in IVP(r_1,n)$ can be decomposed into
  \[
  \rho = \sigma_1 || c_1 || \sigma_2 || c_2 || \sigma_3 || \cdots || c_{j-1} || \sigma_j.
  \]
  where each $\sigma_i$ is a “complete” path in $IVP_0(r_p, c_i)$.

- Set up data-fbw equations to compute $\phi_{r_p,n}$ (functions along complete IVP paths).

- Use this to set up data-fbw equations to compute $x_n$'s.
Balanced call-ret sequences

Paths in $IVP_0(r_p, n)$ for procedure $p$ and node $n$ in $p$: 

![Graph showing paths in $IVP_0(r_p, n)$ for procedure $p$ and node $n$ in $p$.]
Path decomposition

Paths in $IVP(r_1, n)$:

\[ \rho = \sigma_1 \parallel c_1 \parallel \sigma_2 \parallel c_2 \parallel \sigma_3 \parallel \cdots \parallel c_{j-1} \parallel \sigma_j. \]
Equations for $\phi_{r_p,n}$

\[
\phi_{r_p,r_p} \leq \text{id}_L \\
\phi_{r_p,n} = \bigwedge_{(m,n) \in E_p} (h_{m,n} \cdot \phi_{r_p,m})
\]

where

\[
h_{(m,n)} = \begin{cases} 
  f_{m,n} & \text{if } (m,n) \in E_p^0 \\
  \phi_{r_q,e_q} & \text{if } (m,n) \in E_p^1, \text{ and } m \text{ calls } q
\end{cases}
\]
Claim: The vector of values (functions)

\[
\bigwedge_{\sigma \in IVP_0(r_p,n)} f_\sigma
\]

for each procedure \(p\), is the greatest solution to above equations.
Basic correctness argument

\[ \bigwedge_{\sigma \in IVP_0(r_p, n)} f_\sigma = \bigwedge_{m} (\bigwedge_{\eta \in IVP_0(m, n)} f_\eta) \cdot (\bigwedge_{\delta \in IVP_0(r_p, m)} f_\delta). \]
Equations for $x_n$

\[
\begin{align*}
x_1 &= 0 \\
x_{r_p} &= \bigwedge_{\text{calls to } p \text{ in } q} \phi_{r_q, c}(x_{r_q}) \\
x_n &= \phi_{r_p, n}(x_{r_p}) \quad n \text{ in } N_p - \{r_p\}
\end{align*}
\]
Correctness

*Claim:* The vector of data values

\[
\bigwedge_{\text{paths } \rho \in \text{IVP} (r_1,n)} f_{\rho}(0)
\]

for each prog-pt \( n \), is the greatest solution to above equations for \( x_n \)'s.
Example

\[
\begin{align*}
\phi_{r_1,r_1} & \leq id \\
\phi_{r_1,c_1} & = 1 \cdot \phi_{r_1,r_1} \\
\phi_{r_1,n_1} & = \phi_{r_2,e_2} \cdot \phi_{r_1,c_1} \\
\phi_{r_1,e_1} & = 1 \cdot \phi_{r_1,n_2} \\
\phi_{r_2,r_2} & \leq id \\
\phi_{r_2,c_2} & = 0 \cdot \phi_{r_2,r_2} \\
\phi_{r_2,n_2} & = \phi_{r_2,e_2} \cdot \phi_{r_2,c_2} \\
\phi_{r_2,e_2} & = (id \cdot \phi_{r_2,r_2}) \land 1 \cdot \phi_{r_2,n_2}
\end{align*}
\]
Example: computing \( \phi \)'s round 1
Example: computing φ’s round 2

read a, b

read a, b

call p

call p

print t

A \rightarrow id

B \rightarrow f_{\Omega, id}

C \rightarrow f_{\Omega}

D \rightarrow f_{\Omega}

E \rightarrow f_{\Omega}

F \rightarrow id

G \rightarrow f_{\Omega, id}

H \rightarrow f_{\Omega}

I \rightarrow f_{\Omega}

J \rightarrow f_{\Omega}

K \rightarrow f_{\Omega, id}

L

M

N

6

r_1
c_1

n_1
e_1
t := a \cdot b

t := a \cdot b

a := a - 1

a := 0

t := a \cdot b

ret

print t

Data-flow Analysis: Lecture 4 & 5 Interprocedural Analysis: Sharir-Pnueli Technique – p.15
Example: computing $\phi$’s round 3
Example: computing $\phi$’s round 4

```
a := a^{-1}
F := a * b
print F
```

Data-flow Analysis: Lecture 4 & 5
Interprocedural Analysis: Sharir-Pnueli Technique – p.17
Example: computing $\phi$’s round 5

Data-flow Analysis: Lecture 4 & 5
Interprocedural Analysis: Sharir-Pnueli Technique – p.18
Example: computing $\phi$’s round 6

```
Example: computing $\phi$’s round 6

Data-flow Analysis: Lecture 4 & 5 Interprocedural Analysis: Sharir-Pnueli Technique – p.19
```
Example: computing $x_n$’s round 1

\[ x_{r_1} = 0 \]
\[ x_{r_2} = 1(x_{r_1}) \land 0(x_{r_2}). \]

This gives us

\[ x_{r_1} = x_{r_2} = 0. \]
Example: computing $x_n$’s final round

\[ x_{r_1} = 0 \]
\[ x_{r_2} = 1(x_{r_1}) \land 0(x_{r_2}) \]
\[ x_{n_1} = 1(x_{r_1}). \]

... 

This gives us

\[ x_{r_1} = x_{r_2} = 0. \]
\[ x_{n_1} = 1( \text{etc} ). \]