Program Analysis and Verification

0368-4479

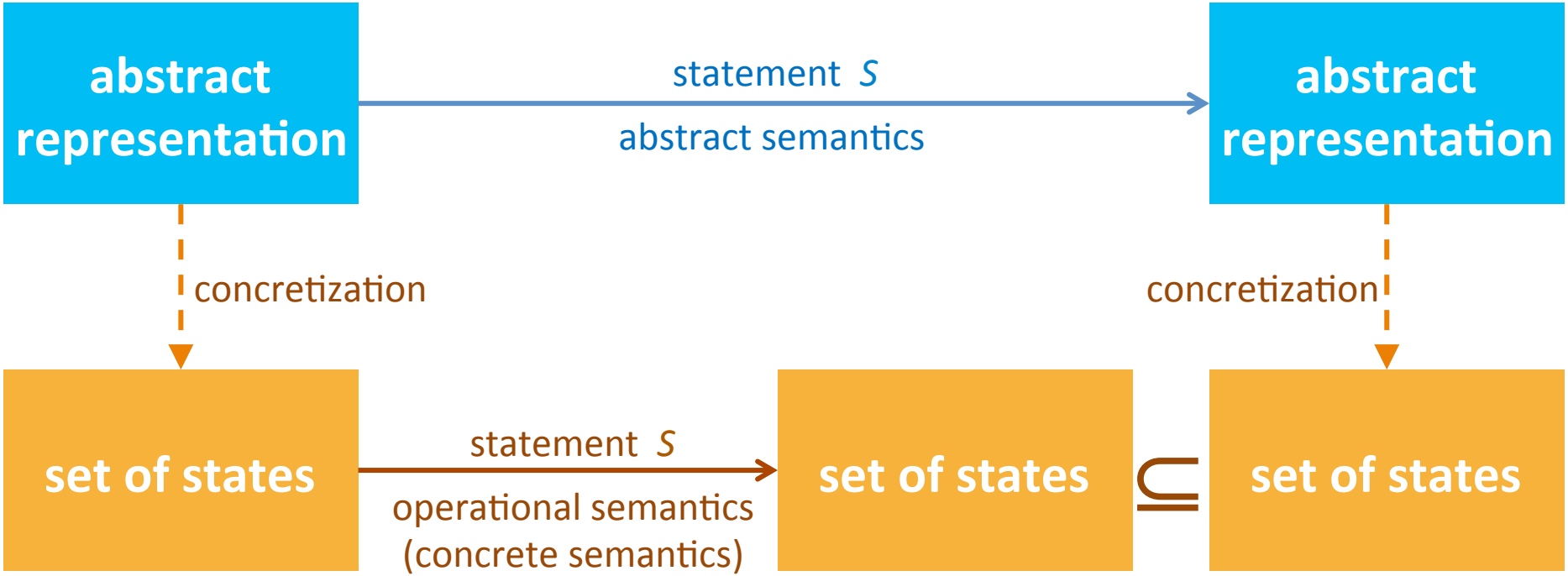
http://www.cs.tau.ac.il/~maon/teaching/2013-2014/paav/paav1314b.html

Noam Rinetzky

Lecture 12: Interprocedural Analysis

Slides credit: Roman Manevich, Mooly Sagiv, Eran Yahav

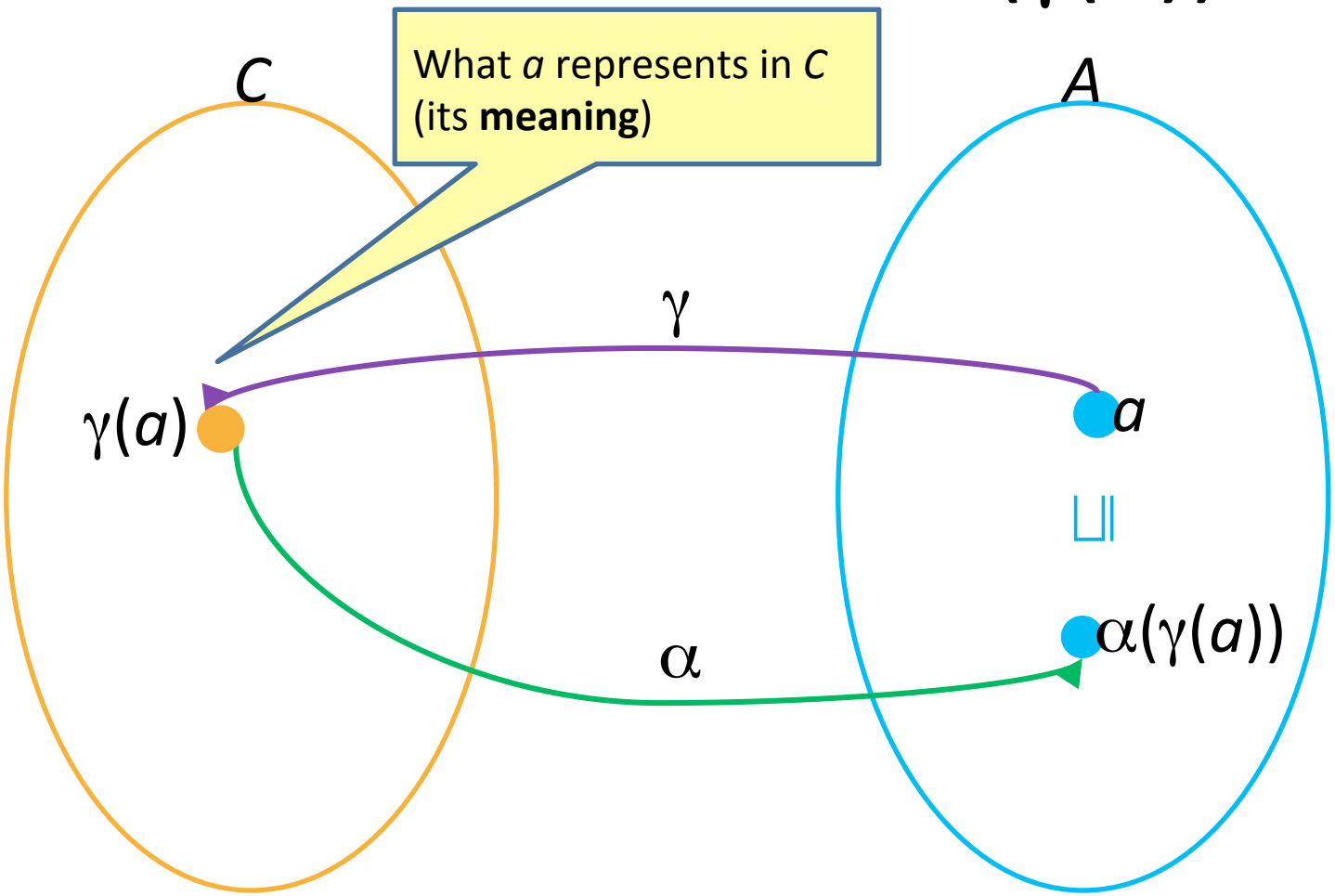
Abstract (conservative) interpretation



The collecting lattice

- Lattice for a given control-flow node v:
  $L_v = (2^{\mathbf{State}}, \subseteq, \cup, \cap, \varnothing, \mathbf{State})$
- Lattice for entire control-flow graph with nodes V:
  $$L_{CFG} = \text{Map}(V, L_v)$$
- We will use this lattice as a baseline for static analysis and define abstractions of its elements

Galois Connection: $\alpha(\gamma(a)) \sqsubseteq a$



Resulting algorithm

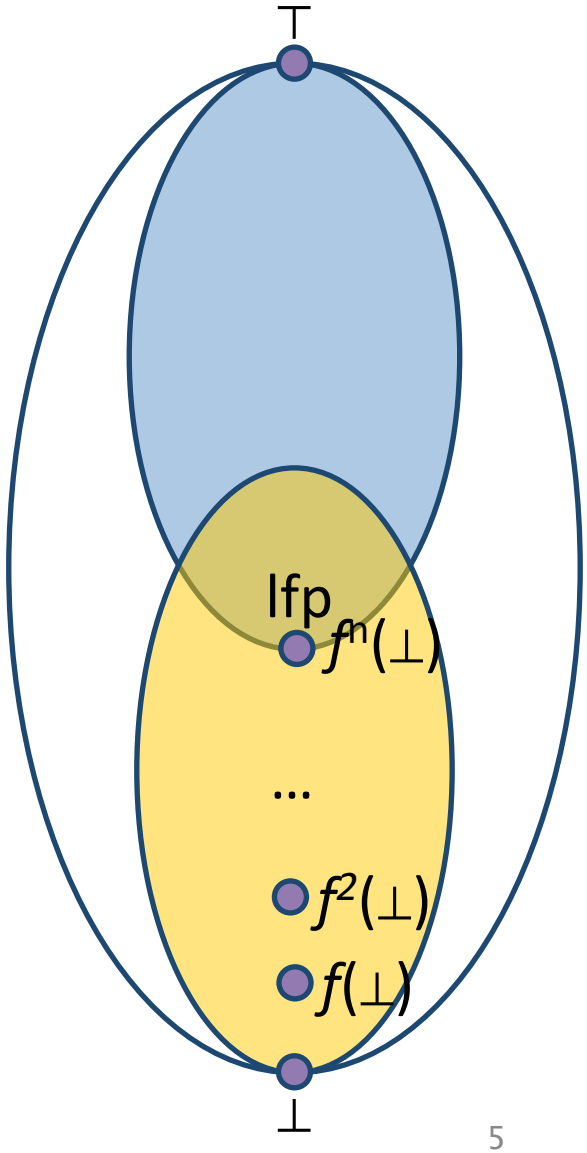
- Kleene's fixed point theorem gives a constructive method for computing the lfp

Mathematical definition

$$\text{lfp}(f) = \bigsqcup_{n \in N} f^n(\bot)$$

Algorithm

$d := \bot$
while $f(d) \neq d$ **do**
  $d := d \sqcup f(d)$
return d



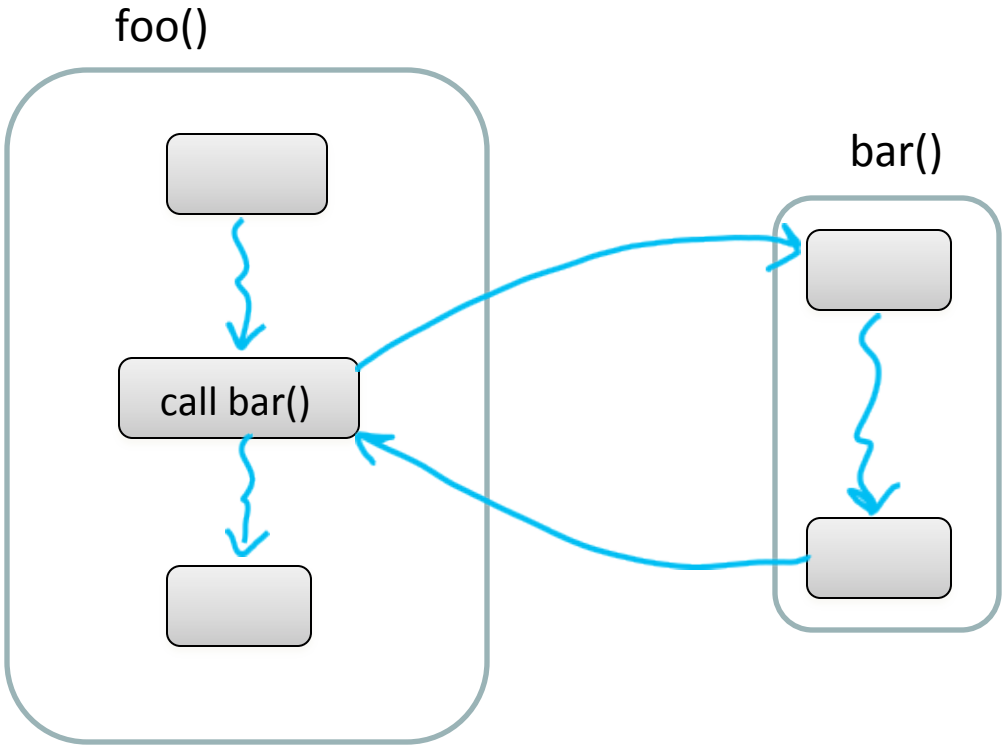
What about procedures?

Procedural program

```
void main() {
   int x;
   x = p(7);
   x = p(9);
}
```

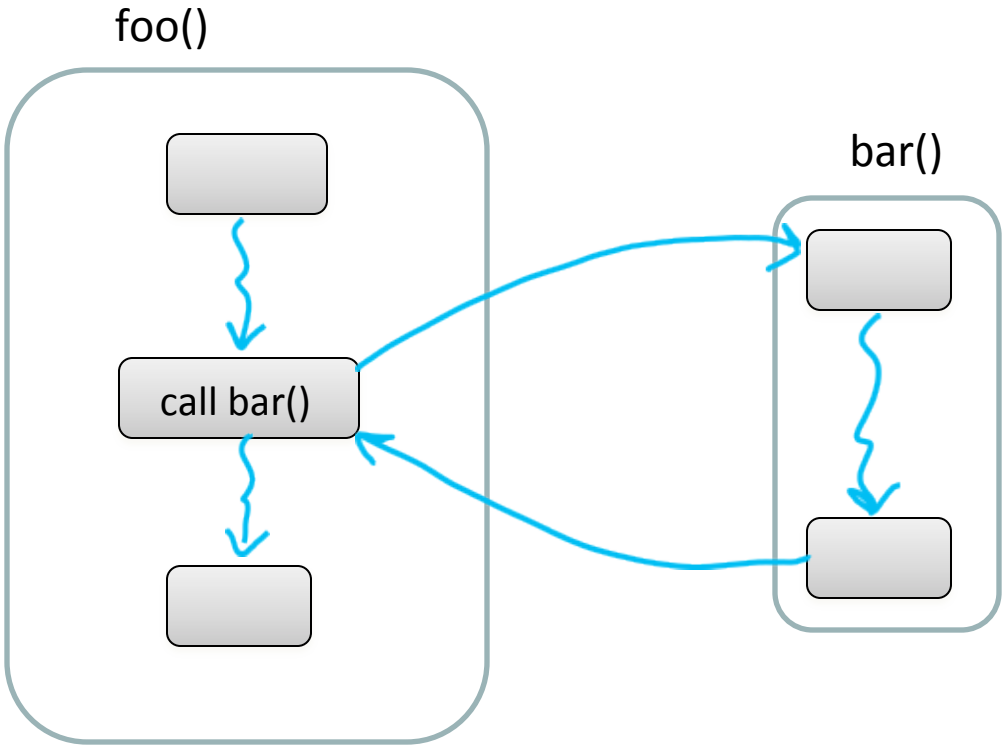
```
int p(int a) {
  return a + 1;
}
```

Effect of procedures



The effect of calling a procedure is the effect of executing its body

Interprocedural Analysis

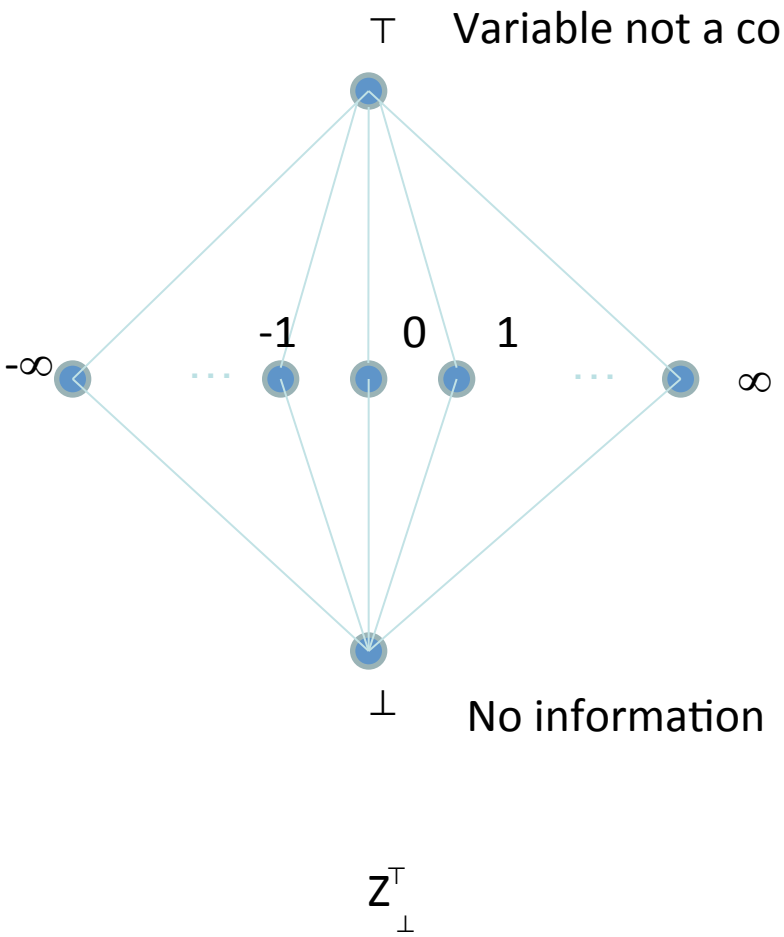


goal: compute the abstract effect of calling a procedure

Reduction to intraprocedural analysis

- Procedure inlining
- Naive solution: call-as-goto

Reminder: Constant Propagation



Reminder: Constant Propagation

- $L = (Var \rightarrow Z, \overset{\top}{\underset{\bot}{\sqsubseteq}})$
- $\sigma_1 \sqsubseteq \sigma_2$ iff $\forall x: \sigma_1(x) \sqsubseteq' \sigma_2(x)$
  - $\sqsubseteq'$ ordering in the Z $\overset{\top}{\underset{\bot}{}}$ lattice

- Examples:
  - $[x \mapsto \bot, y \mapsto 42, z \mapsto \bot] \sqsubseteq [x \mapsto \bot, y \mapsto 42, z \mapsto 73]$
  - $[x \mapsto \bot, y \mapsto 42, z \mapsto 73] \sqsubseteq [x \mapsto \bot, y \mapsto 42, z \mapsto \top]$

Reminder: Constant Propagation

- Conservative Solution
  - Every detected constant is indeed constant
    - But may fail to identify some constants
  - Every potential impact is identified
    - Superfluous impacts

Procedure Inlining

```
void main() {
   int x;
   x = p(7);
   x = p(9);
}
```

```
int p(int a) {
  return a + 1;
}
```

Procedure Inlining

```
void main() {
   int x;
   x = p(7);
   x = p(9);
}
```

```
int p(int a) {
  return a + 1;
}
```

```
void main() {
   int a, x, ret;
   [a ↦ ⊥ , x ↦ ⊥ , ret ↦ ⊥ ]
   a = 7; ret = a+1; x = ret;
   [a ↦7, x ↦8, ret ↦8]
   a = 9; ret = a+1; x = ret;
   [a ↦9, x ↦10, ret ↦10]
}
```

Procedure Inlining

- Pros
  - Simple
- Cons
  - Does not handle recursion
  - Exponential blow up
  - Reanalyzing the body of procedures

```
p1 {
  call p2
  …
  call p2
}
```

```
p2 {
  call p3
  …
  call p3
}
```

```
p3{
}
```

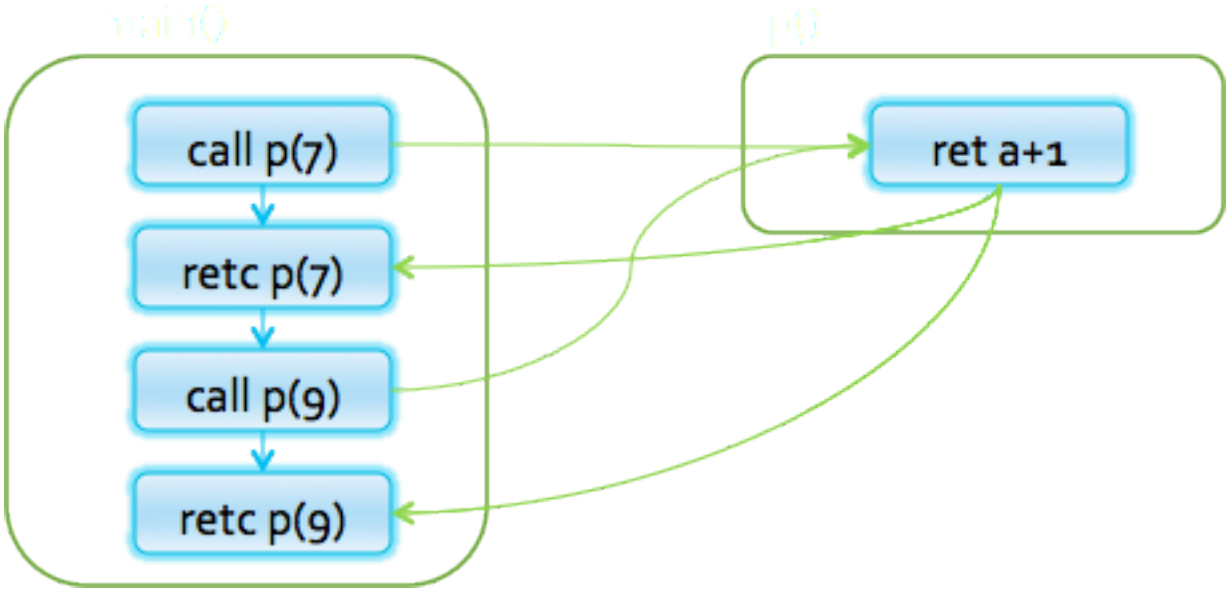
A Naive Interprocedural solution

- Treat procedure calls as gotos

Simple Example

```
void main() {
    int x ;
    x = p(7);
    x = p(9) ;
}
```

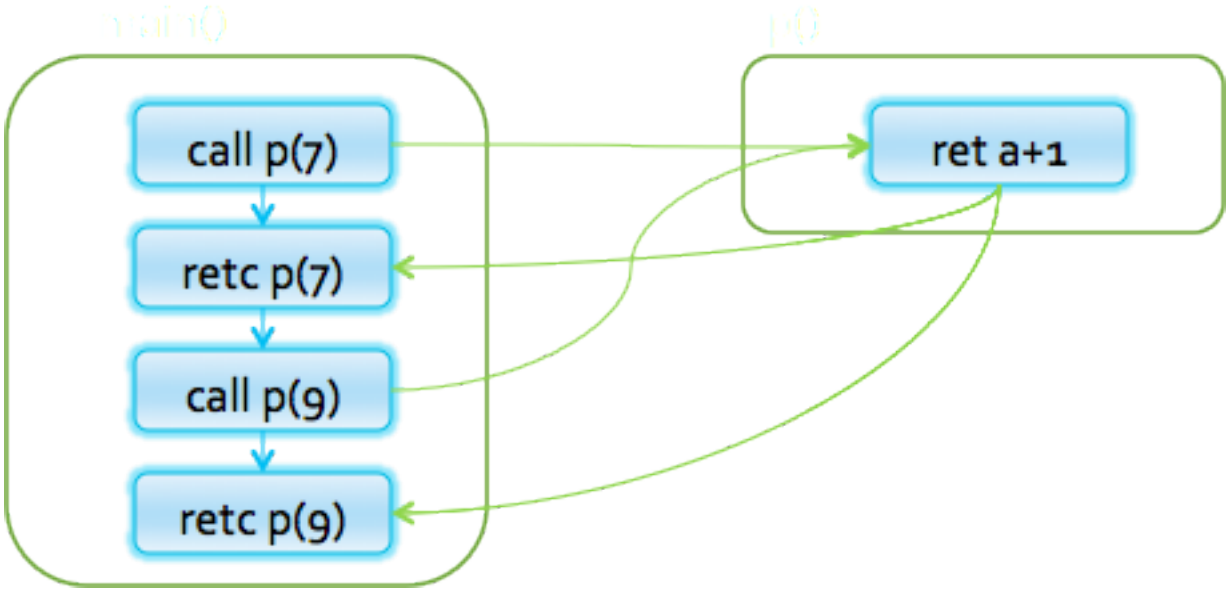
```
int p(int a) {
  return a + 1;
  }
```



Simple Example

```
void main() {
    int x ;
    x = p(7);
    x = p(9) ;
}
```

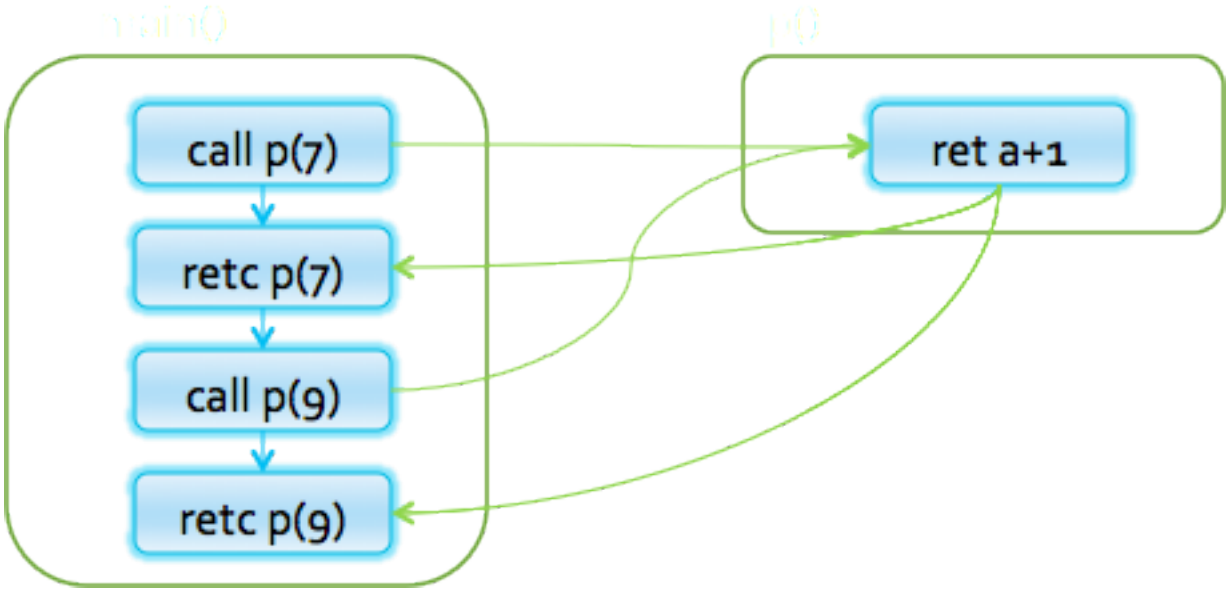
```
int p(int a) {
    [a ↦7]
    return a + 1;
}
```



Simple Example

```
void main() {
    int x ;
    x = p(7);
    x = p(9) ;
}
```

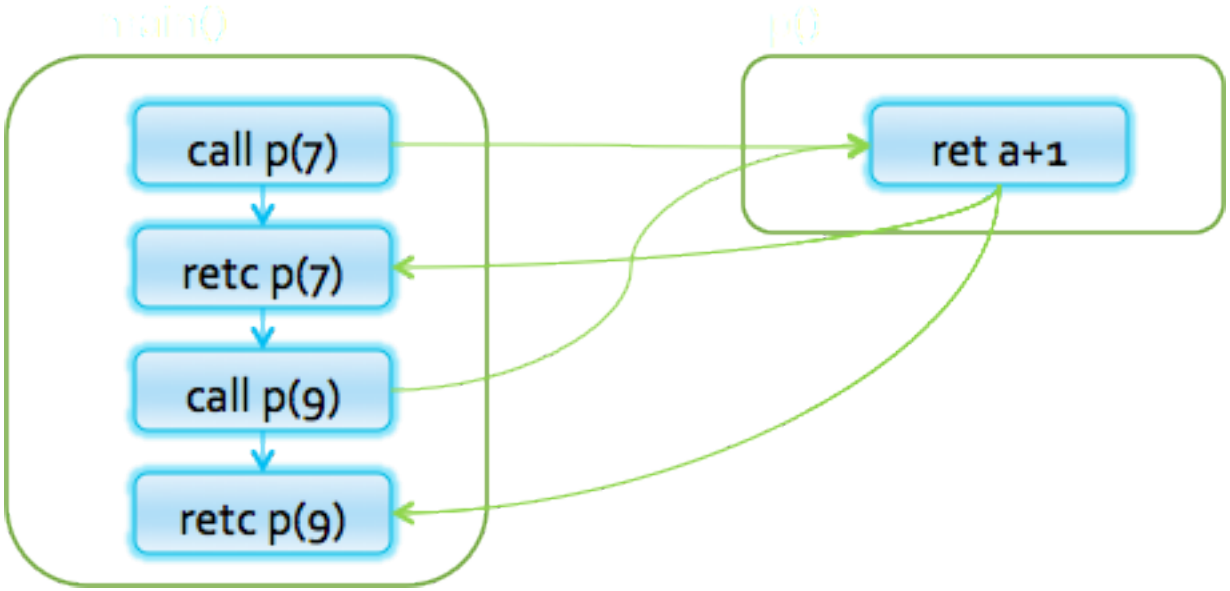
```
int p(int a) {
    [a ↦7]
    return a + 1;
    [a ↦7, $$ ↦8]
}
```



Simple Example

```
void main() {
    int x ;
    x = p(7);
    [x ↦8]
    x = p(9) ;
    [x ↦8]
}
```

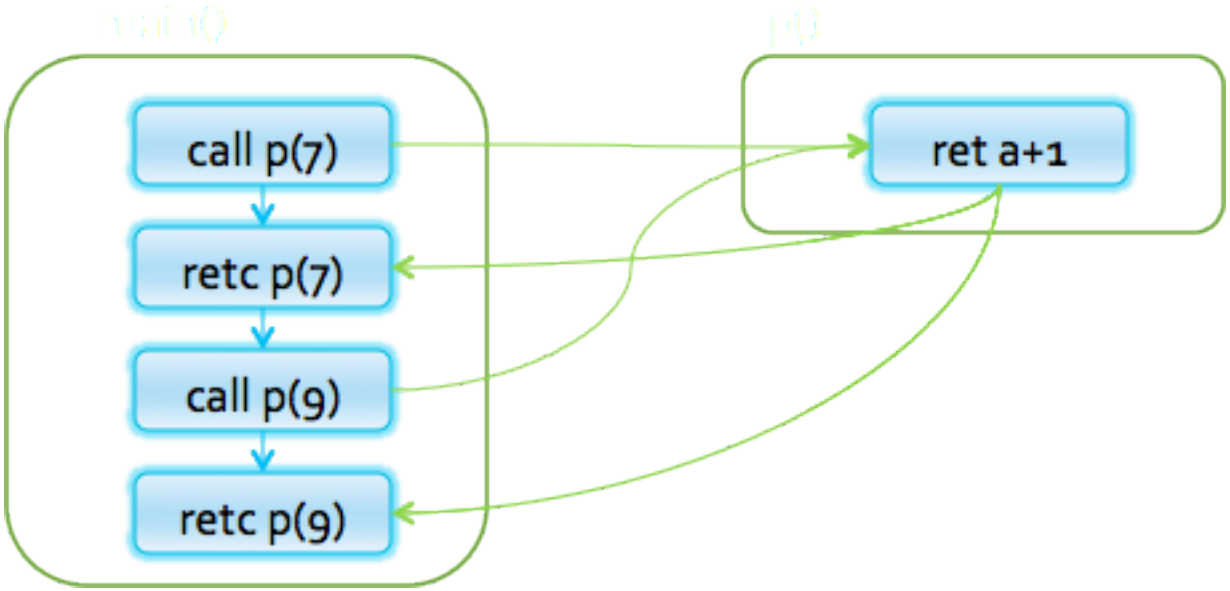
```
int p(int a) {
  [a ↦7]
    return a + 1;
  [a ↦7, $$ ↦8]

}
```



Simple Example

```
void main() {
    int x ;
    x = p(7);
     [x ↦8]
    x = p(9) ;
    [x ↦8]
}
```

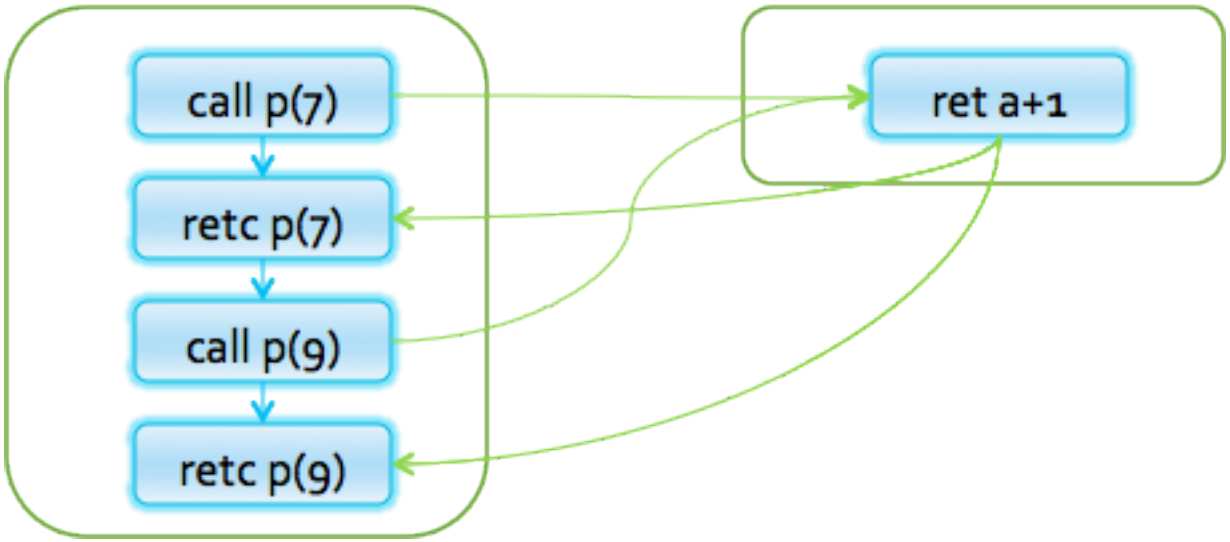
```
int p(int a) {
 [a ↦7]
   return a + 1;
 [a ↦7, $$ ↦8]

}
```



Simple Example

```
void main() {
    int x ;
    x = p(7);
     [x ↦8]
    x = p(9) ;
    [x ↦8]
}
```

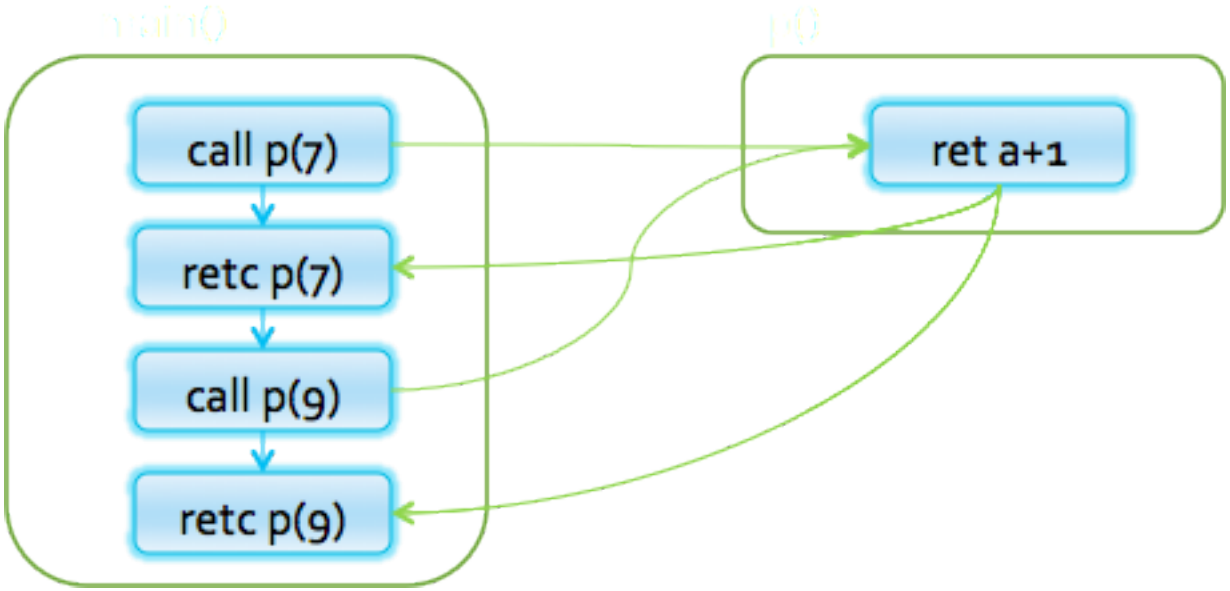
```
int p(int a) {
    [a ↦7]  [a ↦9]
     return a + 1;
    [a ↦7, $$ ↦8]

}
```



Simple Example

```
void main() {
    int x ;
    x = p(7);
    [x ↦8]
    x = p(9) ;
    [x ↦8]
}
```

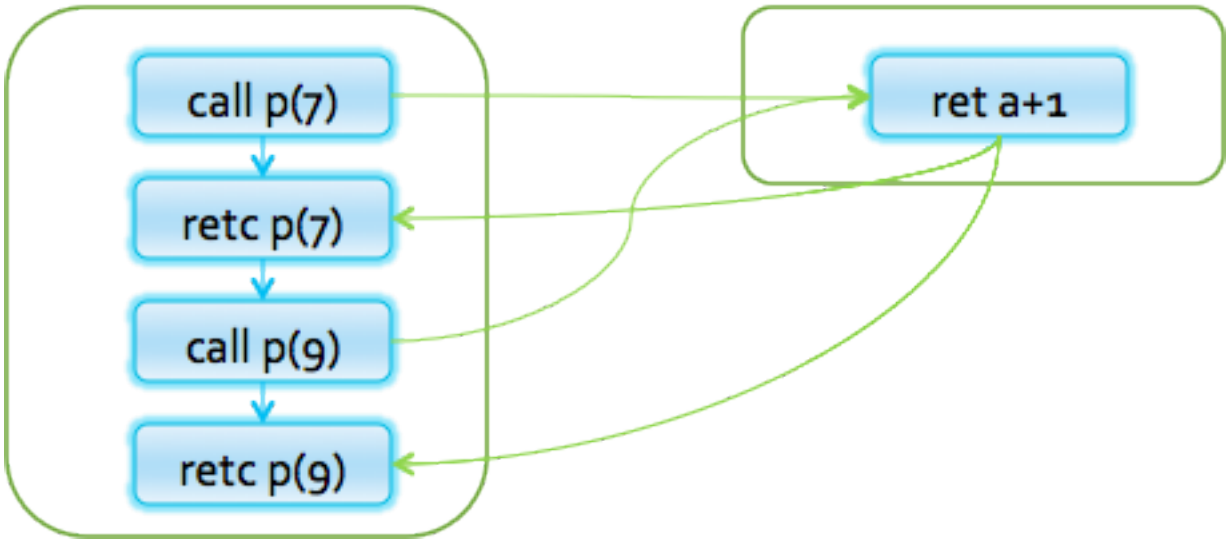
```
int p(int a) {
    [a ↦τ]
    return a + 1;
    [a ↦7, $$ ↦8]
}
```



Simple Example

```
void main() {
    int x ;
    x = p(7);
    [x ↦8]
    x = p(9);
    [x ↦8]

}
```

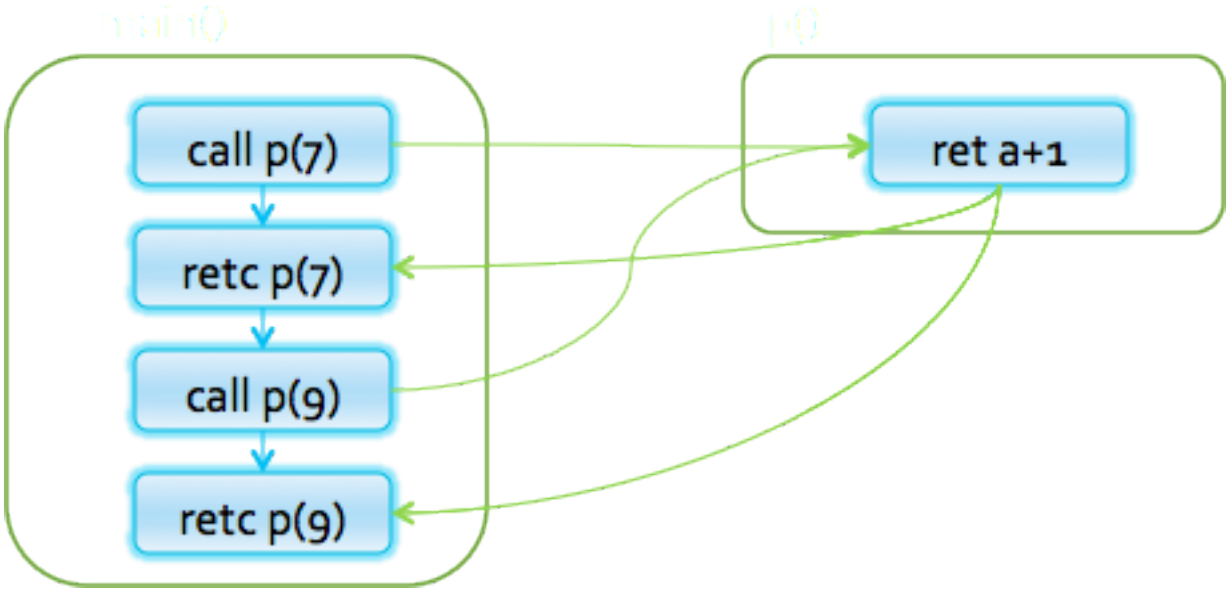
```
int p(int a) {
    [a ↦τ]
    return a + 1;
    [a ↦τ, $$ ↦τ]

}
```



Simple Example

```
void main() {
    int x ;
    x = p(7) ;
     [x ↦τ]
    x = p(9) ;
     [x ↦τ]

}
```

```
int p(int a) {
   [a ↦τ]
    return a + 1;
   [a ↦τ, $$ ↦τ]

}
```



A Naive Interprocedural solution

- Treat procedure calls as gotos
- Pros:
  - Simple
  - Usually fast
- Cons:
  - Abstract call/return correlations
  - Obtain a conservative solution

analysis by reduction

Call-as-goto

```
void main() {
   int x ;
   x = p(7) ;
   [x ↦⊤]
   x = p(9) ;
   [x ↦⊤]
}
```

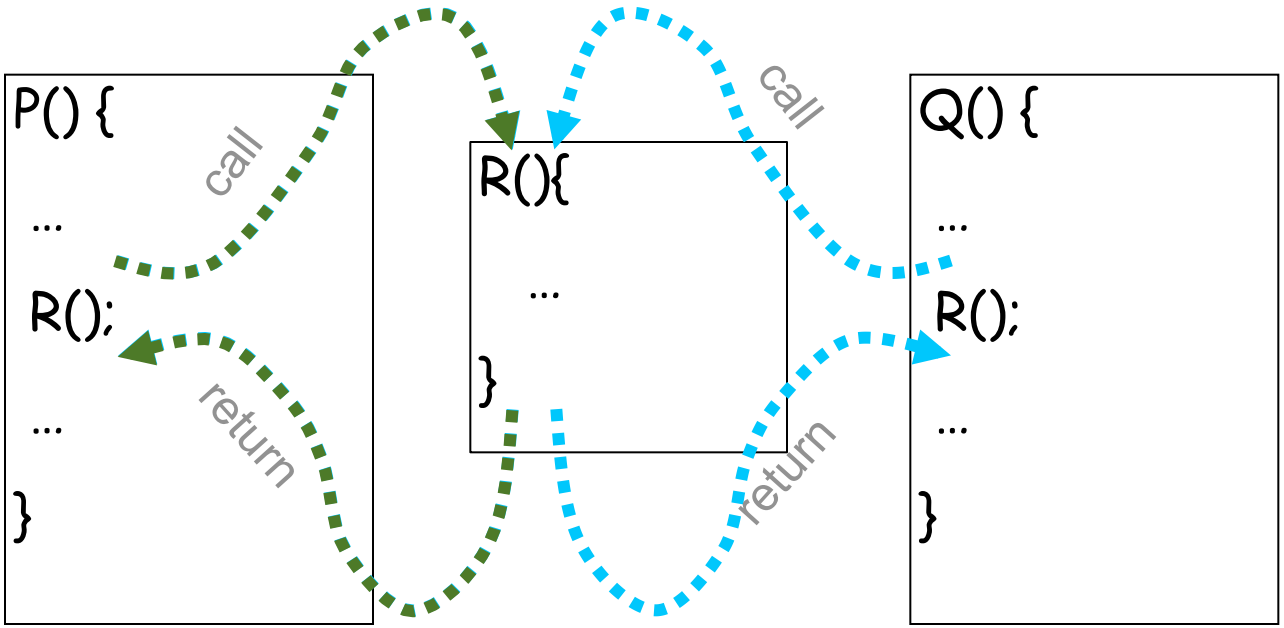
```
int p(int a) {
   [a ↦⊤]
   return a + 1;
   [a ↦⊤, $$ ↦⊤]
}
```

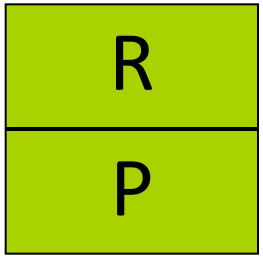
Procedure inlining

```
void main() {
   int a, x, ret;
   [a ↦ ⊥ , x ↦ ⊥ , ret ↦ ⊥ ]
   a = 7; ret = a+1; x = ret;
   [a ↦7, x ↦8, ret ↦8]
   a = 9; ret = a+1; x = ret;
   [a ↦9, x ↦10, ret ↦10]
}
```

why was the naive solution less precise?

Stack regime





Guiding light

- Exploit stack regime
  - ➔Precision
  - ➔Efficiency

Simplifying Assumptions

- Parameter passed by value
- No procedure nesting
- No concurrency

✓ Recursion is supported

Topics Covered

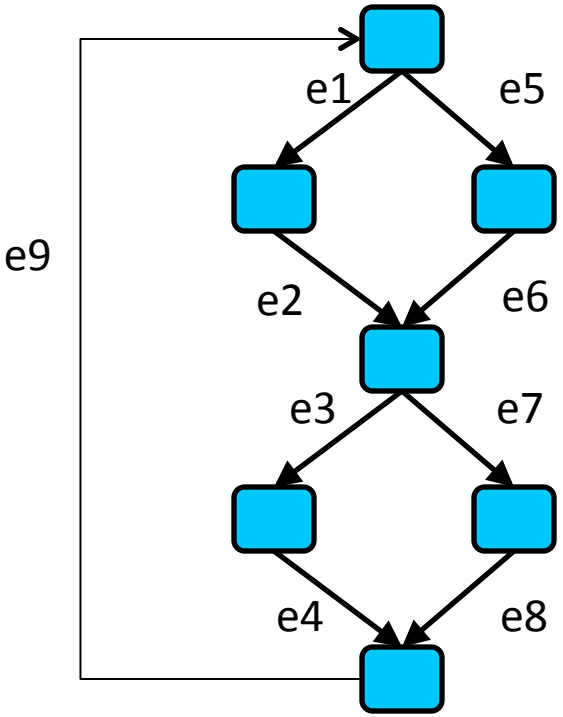
- ✓ Procedure Inlining
- ✓ The naive approach
- Valid paths
- The callstring approach
- The Functional Approach
- IFDS: Interprocedural Analysis via Graph Reachability
- IDE: Beyond graph reachability

- The trivial modular approach

Join-Over-All-Paths (JOP)

- Let paths(v) denote the potentially infinite set paths from start to v (written as sequences of edges)
- For a sequence of edges $[e_1, e_2, \ldots, e_n]$ define $f[e_1, e_2, \ldots, e_n]: L \rightarrow L$ by composing the effects of basic blocks
  $f[e_1, e_2, \ldots, e_n](l) = f(e_n)(\ldots(f(e_2)(f(e_1)(l)))\ldots)$
- $JOP[v] = \bigsqcup\{f[e_1, e_2, \ldots, e_n](\iota) \mid [e_1, e_2, \ldots, e_n] \in paths(v)\}$

Join-Over-All-Paths (JOP)



Paths transformers:
f[e1,e2,e3,e4]
f[e1,e2,e7,e8]
f[e5,e6,e7,e8]
f[e5,e6,e3,e4]
f[e1,e2,e3,e4,e9, e1,e2,e3,e4]
f[e1,e2,e7,e8,e9, e1,e2,e3,e4,e9,…]
…

JOP:
f[e1,e2,e3,e4](initial) ⊔
f[e1,e2,e7,e8](initial) ⊔
f[e5,e6,e7,e8](initial) ⊔
f[e5,e6,e3,e4](initial) ⊔ …

Number of program paths is unbounded due to loops

The lfp computation approximates JOP

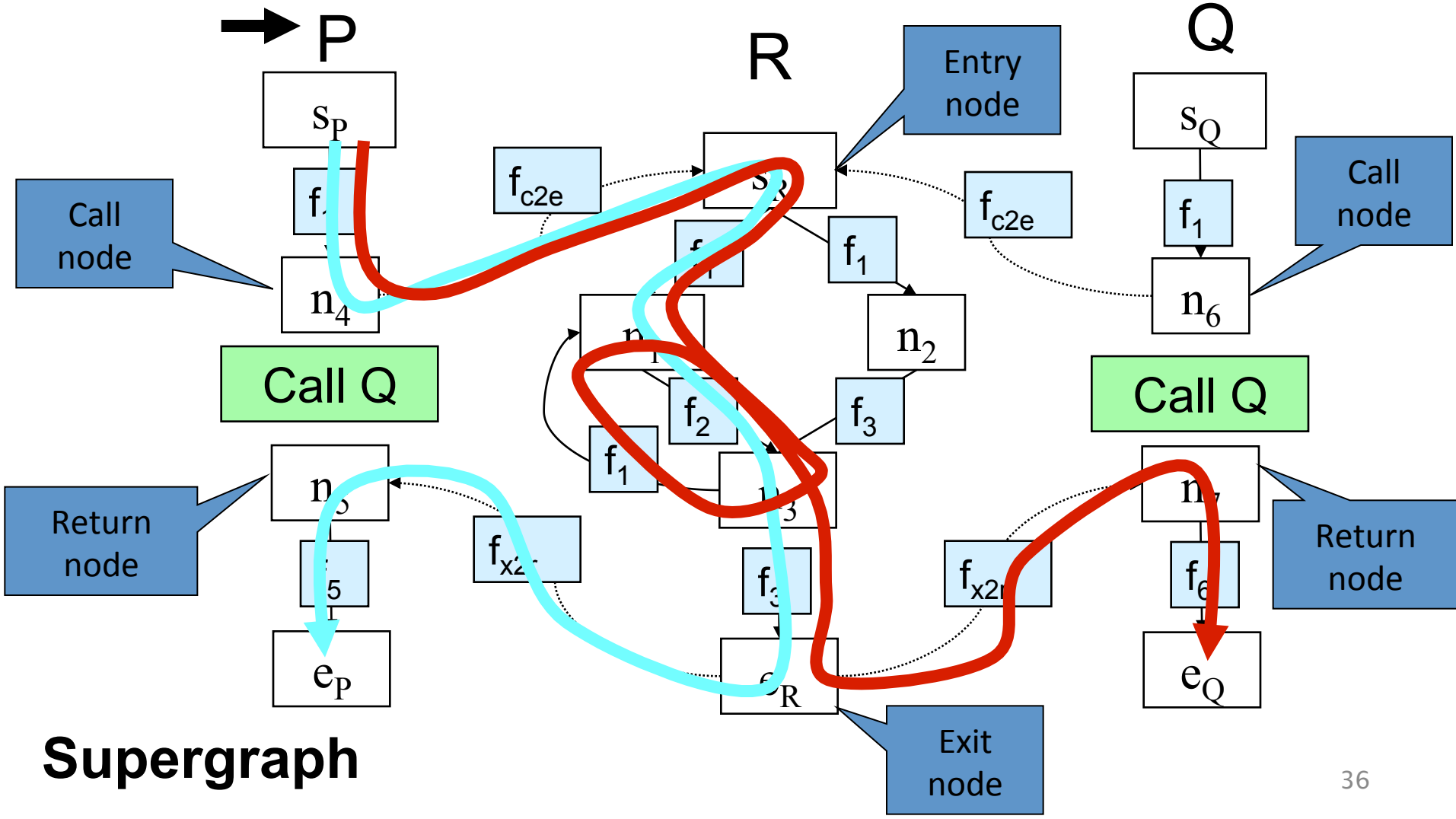
- JOP[v] = $\sqcup\{f\,[e_1, e_2, \ldots, e_n](\iota) \mid [e_1, e_2, \ldots, e_n] \in \text{paths}(v)\}$
- LFP[v] = $\sqcup\{f\,[e](\text{LFP}[v']) \mid e = (v', v)\}$

  LFP[v_0] = ι

- JOP $\sqsubseteq$ LFP - for a monotone function
  - $f(x \sqcup y) \sqsupseteq f(x) \sqcup f(y)$
- JOP = LFP - for a distributive function
  - $f(x \sqcup y) = f(x) \sqcup f(y)$

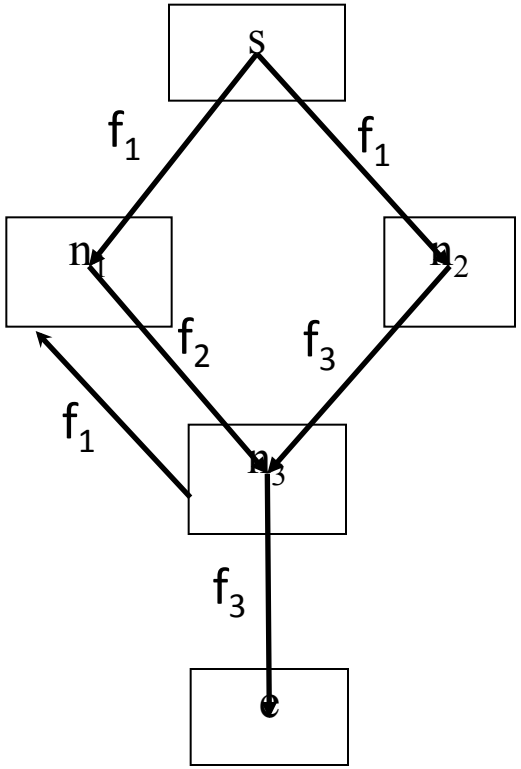
JOP may not be precise enough for interprocedural analysis!

Interprocedural analysis

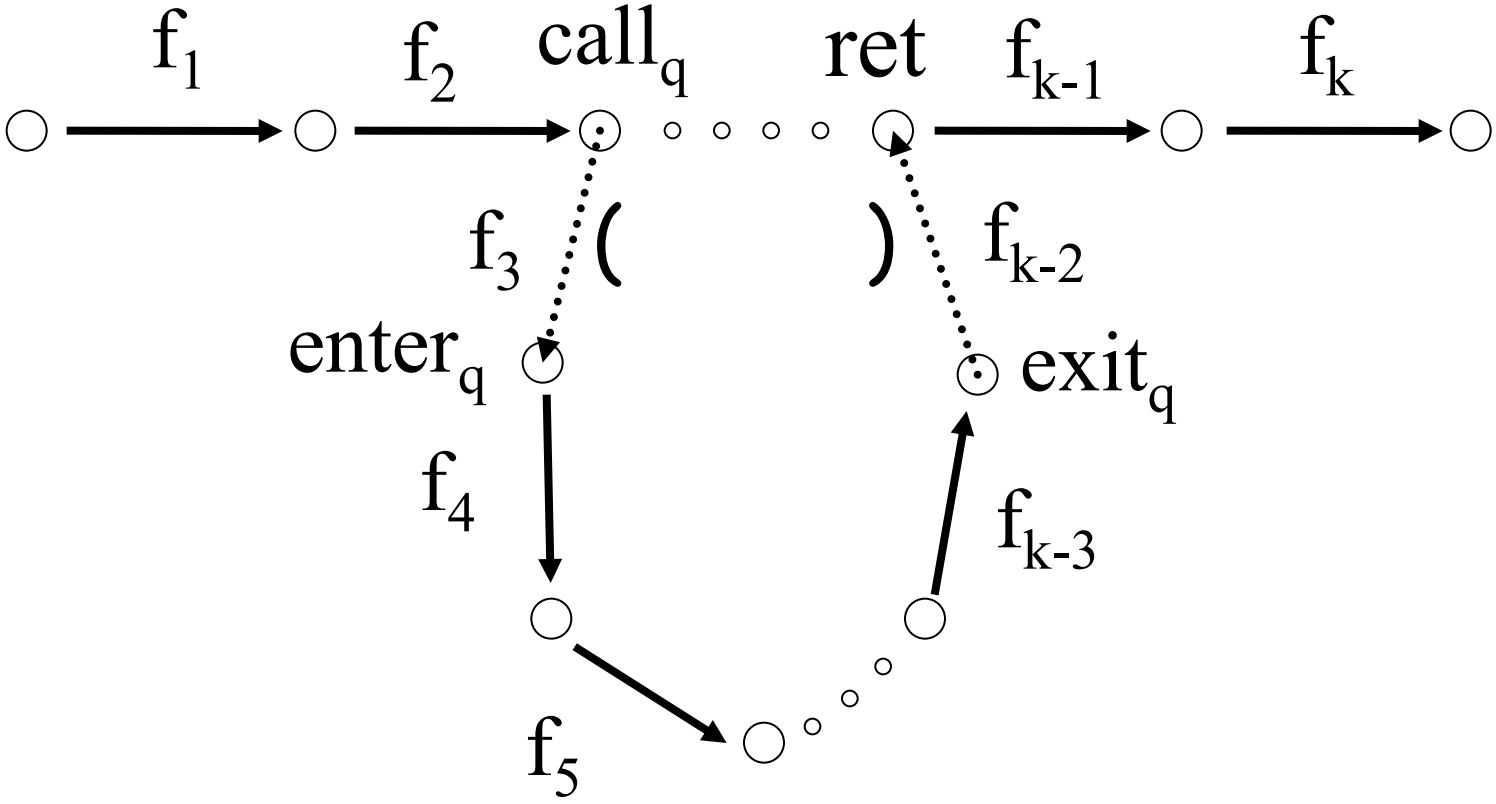


Paths

- paths(n) the set of paths from s to n
  - ((s,n_1), (n_1,n_3), (n_3,n_1))



Interprocedural Valid Paths

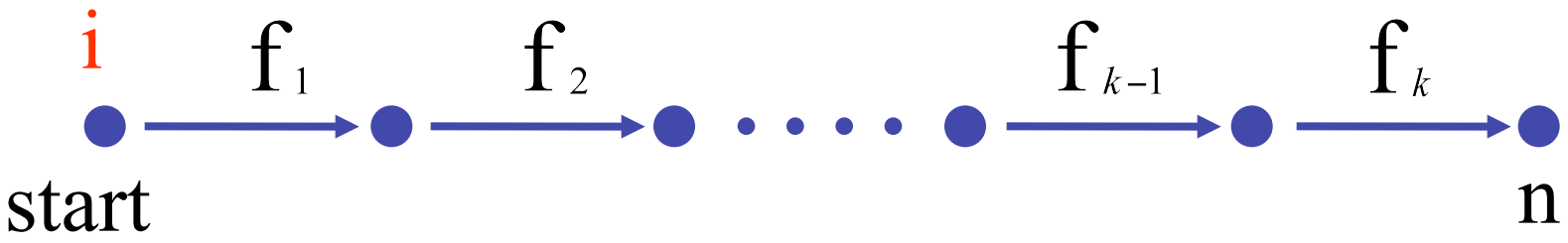


- IVP: all paths with matching calls and returns
  - And prefixes

Interprocedural Valid Paths

- IVP set of paths
  - Start at program entry
- Only considers matching calls and returns
  - aka, valid
- Can be defined via context free grammar
  - matched ::= matched $(_i$ matched $)_i$ | ε
  - valid ::= valid $(_i$ matched | matched
    - paths can be defined by a regular expression

Join Over All Paths (JOP)



$[\![fk \circ \ldots \circ f1]\!] \in L \rightarrow L$

- JOP[v] = $\sqcup\{[[e_1, e_2, \ldots, e_n]](\iota) \mid (e_1, \ldots, e_n) \in \text{paths}(v)\}$
- JOP $\sqsubseteq$ LFP
  - Sometimes JOP = LFP
    - precise up to "symbolic execution"
    - Distributive problem

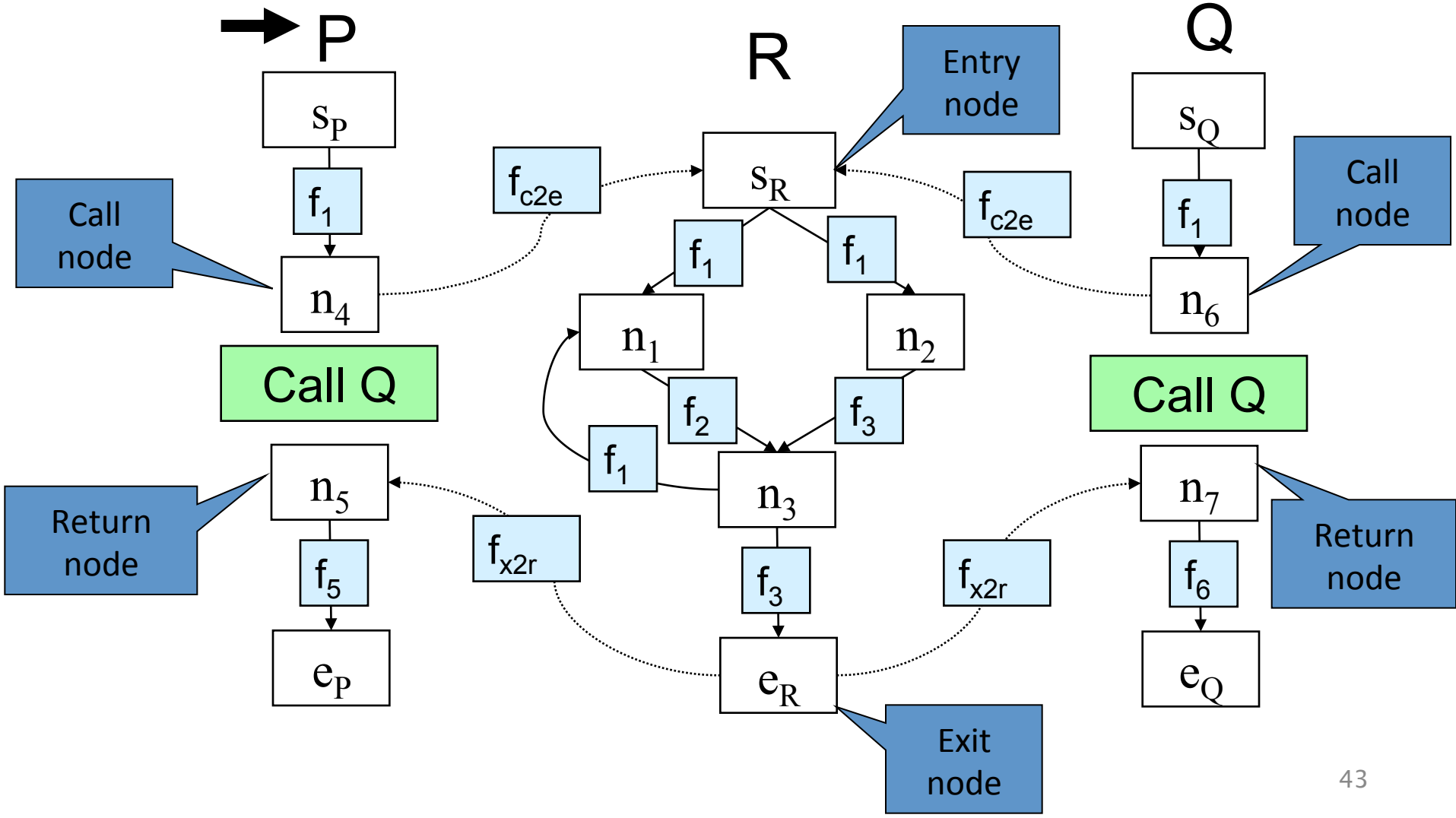
The Join-Over-Valid-Paths (JVP)

- vpaths(n) all valid paths from program start to n
- JVP[n] = $\sqcup\{[[e_1, e_2, \ldots, e]](\iota)$
  $(e_1, e_2, \ldots, e) \in \text{vpaths}(n)\}$
- JVP $\sqsubseteq$ JOP
  - In some cases the JVP can be computed
  - (Distributive problem)

The Call String Approach

- The data flow value is associated with sequences of calls (call string)
- Use Chaotic iterations over the supergraph

supergraph


P
R
Q
s_P
f_1
n_4
Call node
Call Q
n_5
Return node
f_5
e_P
f_{c2e}
s_R
Entry node
f_1
f_1
n_1
n_2
f_2
f_3
f_1
n_3
f_3
e_R
Exit node
f_{x2r}
f_{c2e}
s_Q
f_1
n_6
Call node
Call Q
n_7
Return node
f_6
e_Q
f_{x2r}

Simple Example

```
void main() {
    int x ;
    c1: x =  p(7);
    c2: x =  p(9) ;

}
```

```
int p(int a) {

   return a + 1;

}
```

Simple Example

```
void main() {
    int x ;
    c1: x =  p(7);
    c2: x =  p(9) ;

}
```

```
int p(int a) {
    c1: [a ↦7]
    return a + 1;

}
```

Simple Example

```
void main() {
   int x ;
   c1: x =  p(7);
   c2: x =  p(9) ;

}
```

```
int p(int a) {
   c1: [a ↦7]
   return a + 1;
   c1:[a ↦7, $$ ↦8]
}
```

Simple Example

```
void main() {
   int x ;
   c1: x =  p(7);
    ε: x ↦ 8
   c2: x =  p(9) ;

}
```

```
int p(int a) {
   c1: [a ↦7]
   return a + 1;
   c1:[a ↦7, $$ ↦8]
}
```

Simple Example

```
void main() {
    int x ;
    c1: x = p(7);
    ε: [x ↦ 8]
    c2: x = p(9) ;
}
```

```
int p(int a) {
   c1:[a ↦7]
    return a + 1;
   c1:[a ↦7, $$ ↦8]
}
```

Simple Example

```
void main() {
    int x ;
    c1: x = p(7);
    ε: [x ↦ 8]
    c2: x = p(9) ;
}
```

```
int p(int a) {
    c1:[a ↦7]
    c2:[a ↦9]
    return a + 1;
    c1:[a ↦7, $$ ↦8]
}
```

Simple Example

```
void main() {
   int x ;
   c1: x = p(7);
   ε: [x ↦ 8]
   c2: x = p(9) ;
 }
```

```
int p(int a) {
   c1:[a ↦7]
   c2:[a ↦9]
   return a + 1;
   c1:[a ↦7, $$ ↦8]
   c2:[a ↦9, $$ ↦10]
}
```

Simple Example

```
void main() {
   int x ;
   c1: x = p(7);
   ε: [x ↦ 8]
   c2: x = p(9) ;
   ε: [x ↦ 10]
}
```

```
int p(int a) {
   c1:[a ↦7]
   c2:[a ↦9]
   return a + 1;
   c1:[a ↦7, $$ ↦8]
   c2:[a ↦9, $$ ↦10]
}
```

The Call String Approach

- The data flow value is associated with sequences of calls (call string)
- Use Chaotic iterations over the supergraph

- To guarantee termination limit the size of call string (typically 1 or 2)
  - Represents tails of calls

- Abstract inline

Another Example (|cs|=2)

```
void main() {
    int x ;
    c1: x = p(7);
    ε: [x ↦ 16]
    c2: x = p(9) ;
    ε: [x ↦ 20]
}
```

```
int p(int a) {
  c1:[a ↦7]
  c2:[a ↦9]
   return c3: p1(a + 1);
  c1:[a ↦7, $$ ↦16]
  c2:[a ↦9, $$ ↦20]
}
```

```
int p1(int b) {
  c1.c3:[b ↦8]
  c2.c3:[b ↦10]
   return 2 * b;
  c1.c3:[b ↦8,$$↦16]
  c2.c3:[b ↦10,$$↦20]
}
```

Another Example (|cs|=1)

```
void main() {
   int x ;
   c1: x = p(7);
   ε: [x ↦ τ]
   c2: x = p(9) ;
   ε: [x ↦ τ]
}
```

```
int p(int a) {
  c1:[a ↦7]
  c2:[a ↦9]
   return c3: p1(a + 1);
  c1:[a ↦7, $$ ↦τ]
  c2:[a ↦9, $$ ↦τ]
}
```

```
int p1(int b) {
  (c1|c2)c3:[b ↦τ]
  return 2 * b;
  (c1|c2)c3:[b ↦τ, $$↦τ]
}
```

Handling Recursion

```
void main() {
   c1: p(7);
  ε: [x ↦τ]
}
```

```
int p(int a) {
   c1: [a ↦ 7]   c1.c2+: [a ↦ τ]
   if (…)  {
    c1: [a ↦ 7]  c1.c2+: [a ↦ τ]
    a = a -1 ;
    c1: [a ↦ 6]   c1.c2+: [a ↦ τ]
    c2: p (a);
    c1.c2*: [a ↦ τ]
    a = a + 1;
    c1.c2*: [a ↦ τ]
   }
  c1.c2*: [a ↦ τ]
   x = -2*a + 5;
  c1.c2*: [a ↦ τ, x↦τ]
}
```

Summary Call String

- Easy to implement
- Efficient for very small call strings
- Limited precision
  - Often loses precision for recursive programs
  - For finite domains can be precise even with recursion (with a bounded callstring)

- Order of calls can be abstracted
- Related method: procedure cloning

The Functional Approach

- The meaning of a procedure is mapping from states into states
- The abstract meaning of a procedure is function from an abstract state to abstract states
- Relation between input and output
- In certain cases can compute JVP

The Functional Approach

- Two phase algorithm
  - Compute the dataflow solution at the exit of a procedure as a function of the initial values at the procedure entry (functional values)
  - Compute the dataflow values at every point using the functional values

Phase 1

```
void main() {

   p(7);

}
```

```
int p(int a) {

 [a ↦a0, x ↦x0]
 if (…)  {

  [a ↦a0, x ↦x0]
   a = a -1 ;

  [a ↦a0-1, x ↦x0]
   p (a);

  [a ↦a0-1, x ↦-2a0+7]
   a = a + 1;

  [a ↦a0, x ↦-2a0+7]
 }

 [a ↦a0, x ↦x0]   [a ↦a0, x ↦T]
 x = -2*a + 5;

  [a ↦a0, x ↦ -2*a0+5]
}
```

$p(a_0,x_0) = [a \mapsto a_0, x \mapsto -2a_0 + 5]$

Phase 2

```
void main() {
    p(7);
   [x ↦ -9]
}
```

```
int p(int a) {
  [a ↦7, x ↦0]          [a ↦⊤, x ↦0]
  if (…)  {
    [a ↦7, x ↦0]        [a ↦⊤, x ↦0]
    a = a -1 ;
    [a ↦6, x ↦0]        [a ↦⊤, x ↦0]
    p (a);
     [a ↦6, x ↦-7]      [a ↦⊤, x ↦⊤]
    a = a + 1;
    [a ↦7, x ↦-7]       [a ↦⊤, x ↦⊤]
  }
  [a ↦7, x ↦0]          [a ↦⊤, x ↦⊤]
  x = -2*a + 5;
  [a ↦7, x ↦-9]         [a ↦⊤, x ↦⊤]
}
```

$p(a_0,x_0) = [a \mapsto a_0, x \mapsto -2a_0 + 5]$

Summary Functional approach

- Computes procedure abstraction
- Sharing between different contexts
- Rather precise
- Recursive procedures may be more precise/ efficient than loops
- But requires more from the implementation
  - Representing (input/output) relations
  - Composing relations

Issues in Functional Approach

- How to guarantee that finite height for functional lattice?
  - It may happen that L has finite height and yet the lattice of monotonic function from L to L do not
- Efficiently represent functions
  - Functional join
  - Functional composition
  - Testing equality

Tabulation

- Special case: L is finite
- Data facts: $d \in L \times L$
- Initialization:
  - $f_{start,start} = (\top,\top)$; otherwise $(\bot,\bot)$
  - $S[start, \top] = \top$

- Propagation of (x,y) over edge e = (n,n')
  - Maintain summary: $S[n',x] = S[n',x] \sqcup [\![n]\!](y))$
  - n intra-node: ➔ n' : $(x, [\![n]\!](y))$
  - n call-node:
    - ➔ n' : (y,y) if $S[n',y] = \bot$ and n' = entry node
    - ➔ n' : (x,z) if S[exit(call(n),y] = z and n' = ret-site-of n
  - n return-node: ➔ n' : (u,y) ; n_c = call-site-of n', $S[n_c,u]=x$

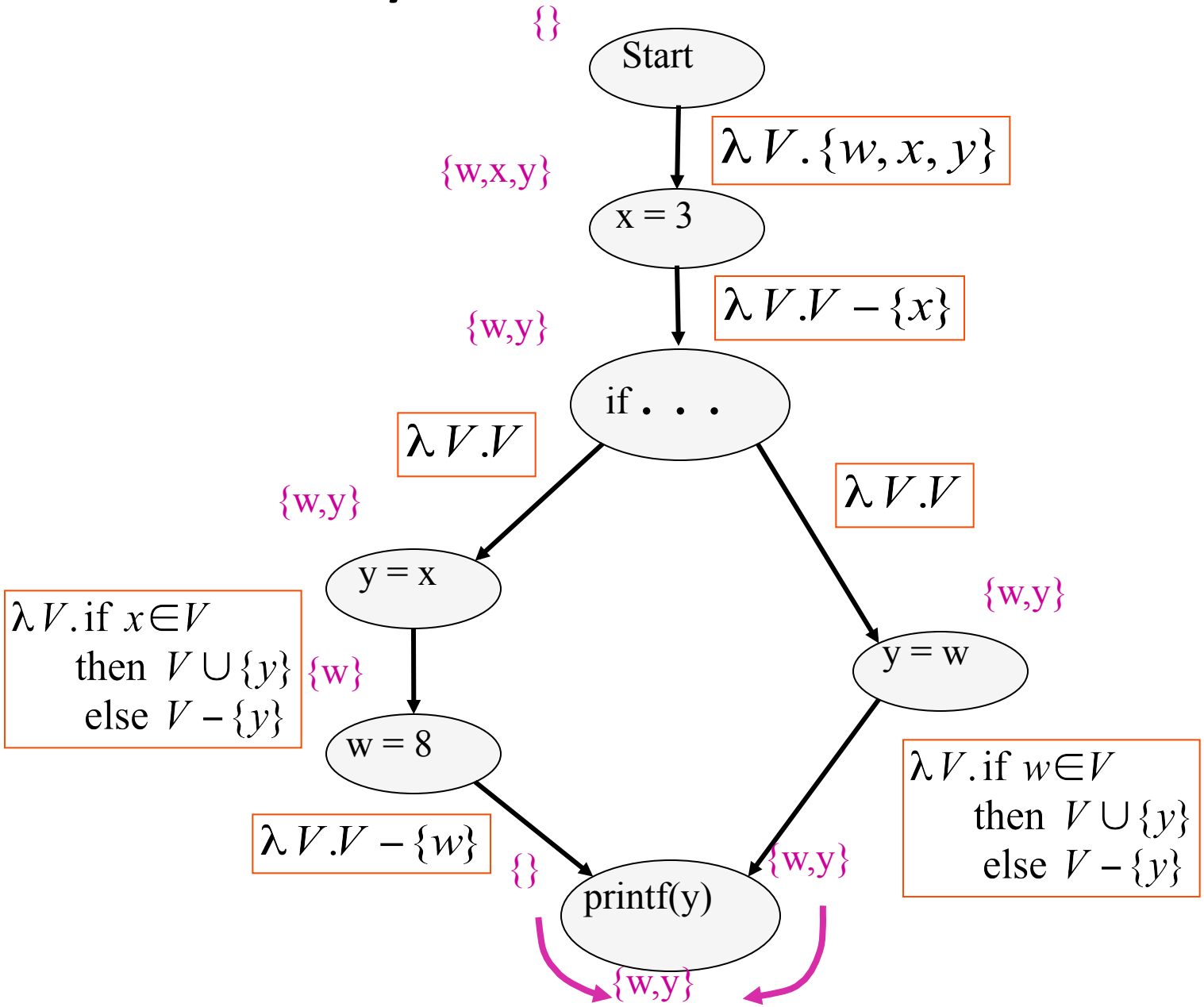
CFL-Graph reachability

- Special cases of functional analysis
- Finite distributive lattices
- Provides more efficient analysis algorithms
- Reduce the interprocedural analysis problem to finding context free reachability

IDFS / IDE

- IDFS Interprocedural Distributive Finite Subset
  Precise interprocedural dataflow analysis via graph reachability. Reps, Horowitz, and Sagiv, POPL'95
- IDE Interprocedural Distributive Environment
  Precise interprocedural dataflow analysis with applications to constant propagation. Reps, Horowitz, and Sagiv, FASE'95, TCS'96
  - More general solutions exist

Possibly Uninitialized Variables



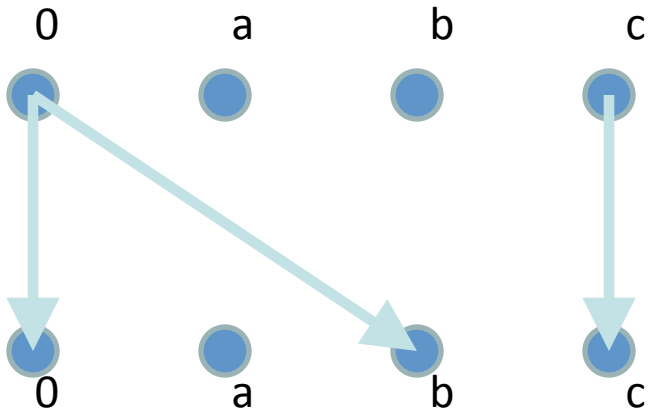
IFDS Problems

- Finite subset distributive
  - Lattice L = $\wp$(D)
  - $\sqsubseteq$ is $\subseteq$
  - $\sqcup$ is $\cup$
  - Transfer functions are distributive

- Efficient solution through formulation as CFL reachability

Encoding Transfer Functions

- Enumerate all input space and output space
- Represent functions as graphs with 2(D+1) nodes
- Special symbol "0" denotes empty sets (sometimes denoted Λ)
- Example: D = { a, b, c }
  f(S) = (S – {a}) U {b}



Efficiently Representing Functions

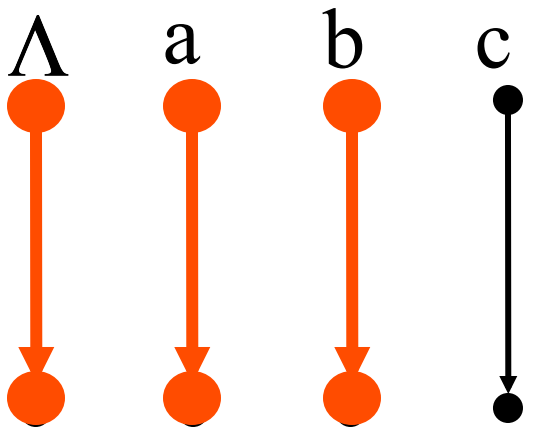
- Let $f:2^D \rightarrow 2^D$ be a distributive function
- Then:
  - $f(X) = f(\emptyset) \cup (\cup \{ f(\{z\}) \mid z \in X \})$
  - $f(X) = f(\emptyset) \cup (\cup \{ f(\{z\}) \setminus f(\emptyset) \mid z \in X \})$

Representing Dataflow Functions

Identity Function

$$f = \lambda V.V$$

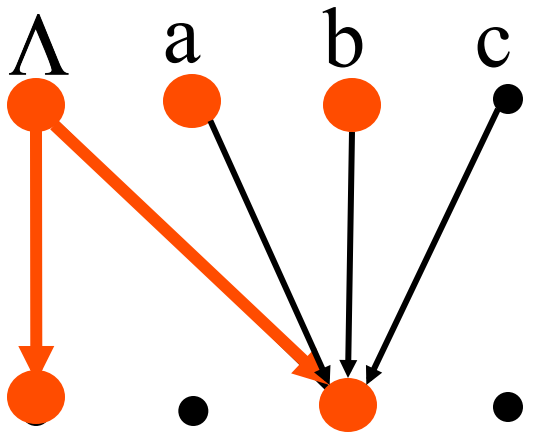
$$f(\{a,b\}) = \{a,b\}$$



Constant Function

$$f = \lambda V.\{b\}$$

$$f(\{a,b\}) = \{b\}$$

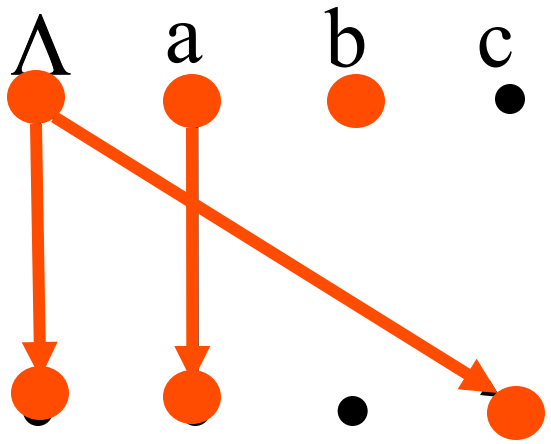


Representing Dataflow Functions

"Gen/Kill" Function

$$f = \lambda V.(V - \{b\}) \cup \{c\}$$

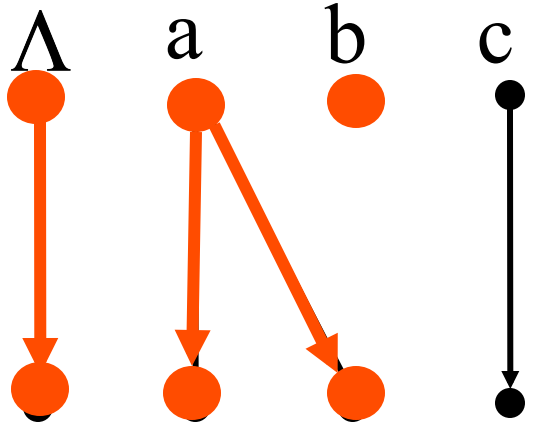
$$f(\{a,b\}) = \{a,c\}$$

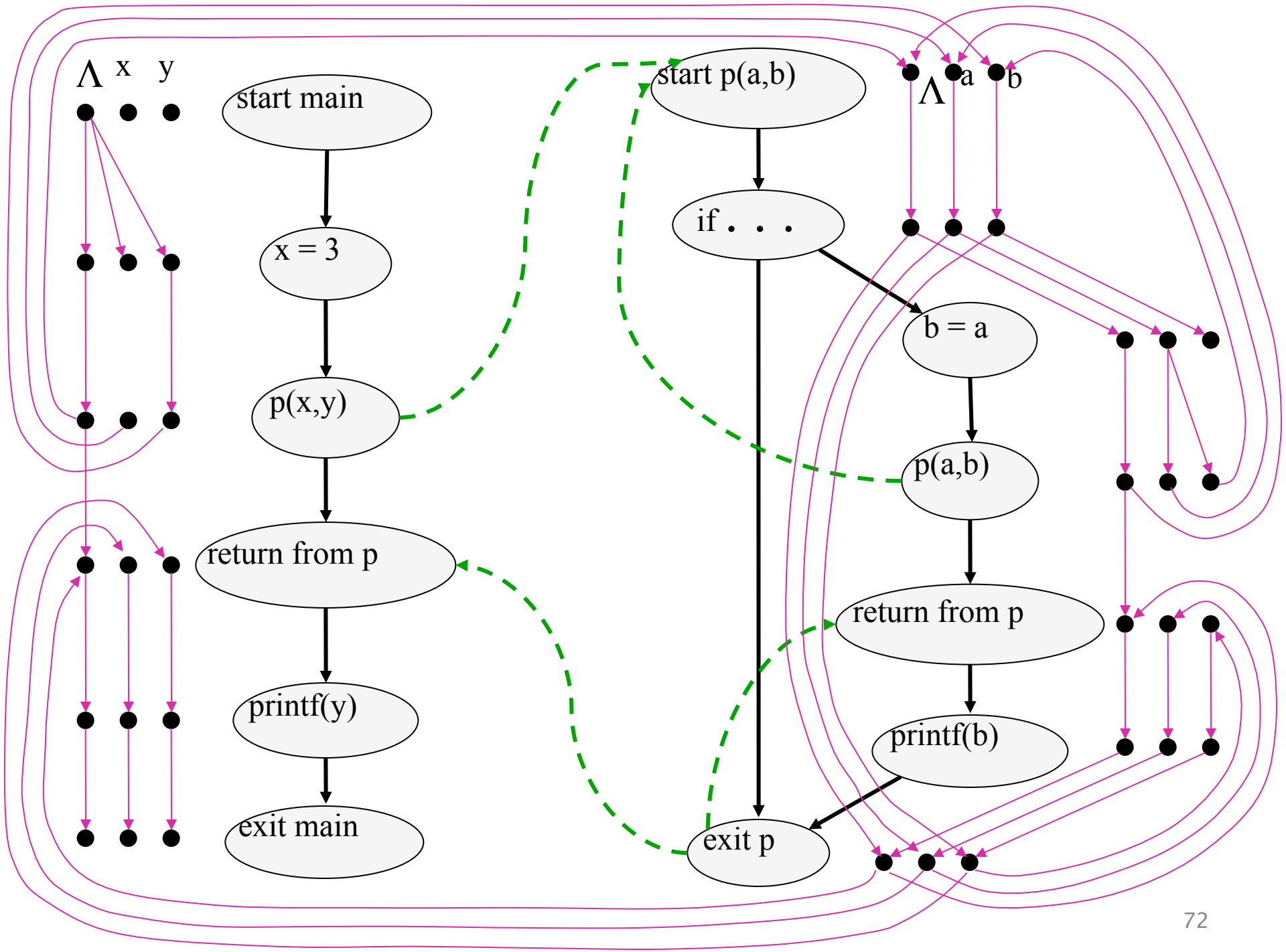


Non-"Gen/Kill" Function

$$f = \lambda V.\text{if } a \in V$$
$$\text{then } V \cup \{b\}$$
$$\text{else } V - \{b\}$$

$$f(\{a,b\}) = \{a,b\}$$

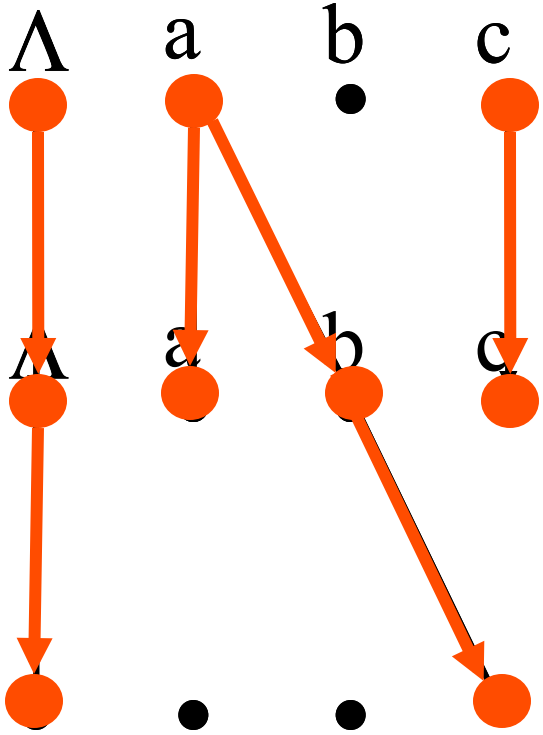



Λ x y
start main
x = 3
p(x,y)
return from p
printf(y)
exit main
start p(a,b)
if . . .
exit p
Λ a b
b = a
p(a,b)
return from p
printf(b)

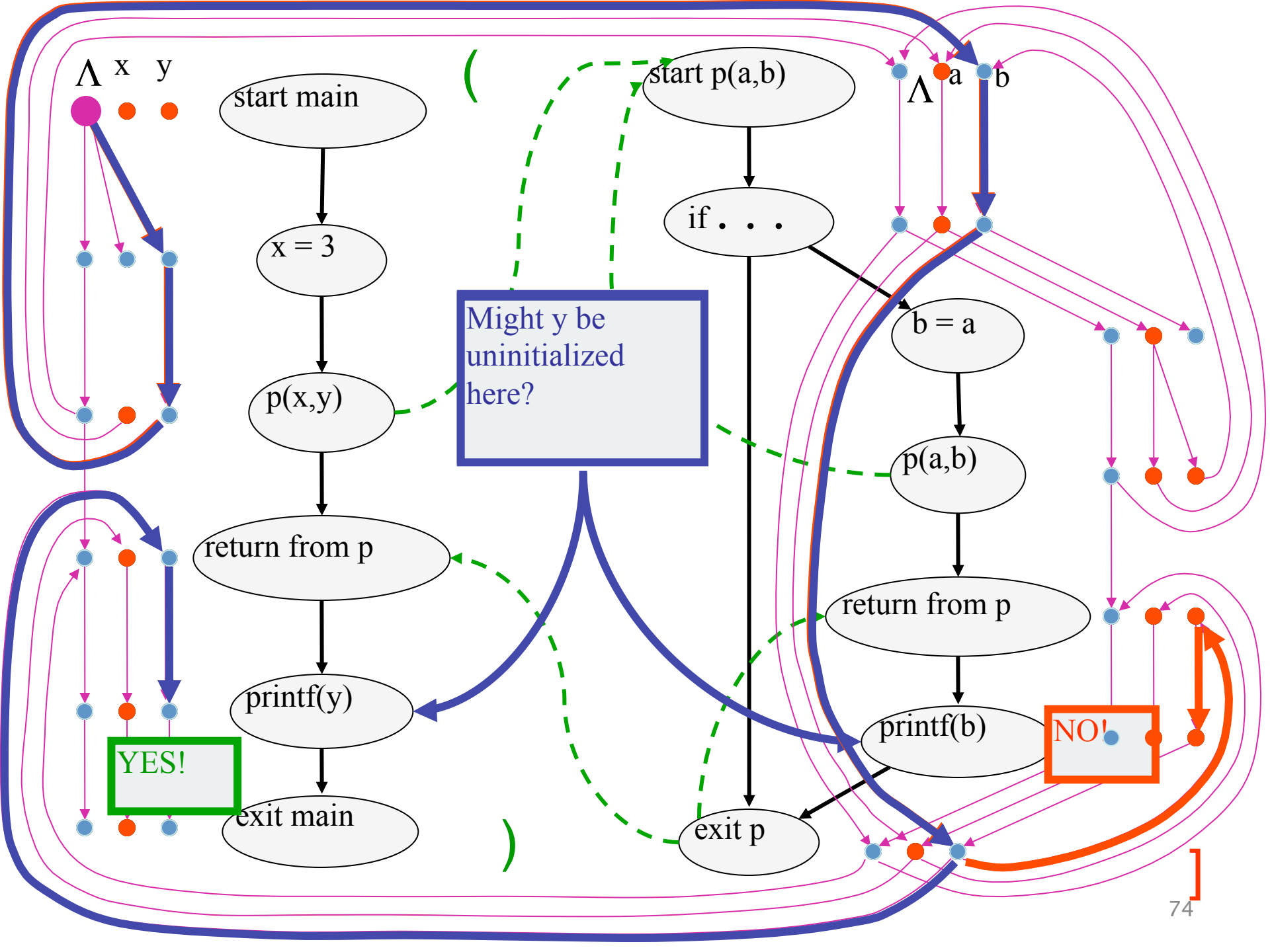
Composing Dataflow Functions

$$f_1 = \lambda V.\text{if } a \in V \text{ then } V \cup \{b\} \text{ else } V - \{b\}$$

$$f_2 = \lambda V.\text{if } b \in V \text{ then } \{c\} \text{ else } \phi$$



$$f_2 \circ f_1(\{a,c\}) = \boxed{\{c\}}$$


Λ x y
start main
x = 3
p(x,y)
return from p
printf(y)
exit main
YES!
start p(a,b)
if . . .
exit p
Might y be uninitialized here?
Λ a b
b = a
p(a,b)
return from p
printf(b)
NO!

The Tabulation Algorithm

- Worklist algorithm, start from entry of "main"
- Keep track of
  - Path edges: matched paren paths from procedure entry
  - Summary edges: matched paren call-return paths
- At each instruction
  - Propagate facts using transfer functions; extend path edges
- At each call
  - Propagate to procedure entry, start with an empty path
  - If a summary for that entry exits, use it
- At each exit
  - Store paths from corresponding call points as summary paths
  - When a new summary is added, propagate to the return node

Interprocedural Dataflow Analysis via CFL-Reachability

- Graph: Exploded control-flow graph
- L: L(unbalLeft)
  - unbalLeft = valid
- Fact d holds at n iff there is an L(unbalLeft)-path from $\langle start_{main}, \Lambda \rangle$ to $\langle n, d \rangle$

Asymptotic Running Time

- CFL-reachability
  - Exploded control-flow graph: ND nodes
  - Running time: $O(N^3D^3)$
- Exploded control-flow graph → Special structure

Running time: $O(ED^3)$

Typically: $E \approx N$, hence $O(ED^3) \approx O(ND^3)$

"Gen/kill" problems: $O(ED)$

IDE

- Goes beyond IFDS problems
  - Can handle unbounded domains
- Requires special form of the domain
- Can be much more efficient than IFDS

Example Linear Constant Propagation

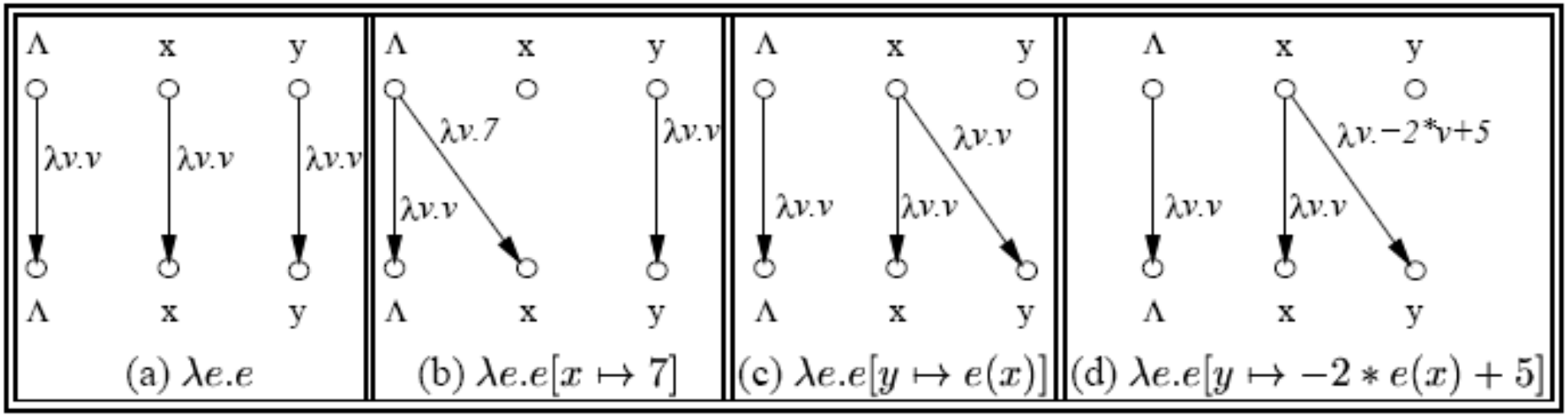
- Consider the constant propagation lattice
- The value of every variable y at the program exit can be represented by:

$$y = \bigsqcup \{(a_x x + b_x) \mid x \in Var_* \} \bigsqcup c$$

$$a_x, c \in Z \cup \{\bot, \top\} \quad b_x \in Z$$

- Supports efficient composition and "functional" join
  - [z := a * y + b]
  - What about [z:=x+y]?

Linear constant propagation

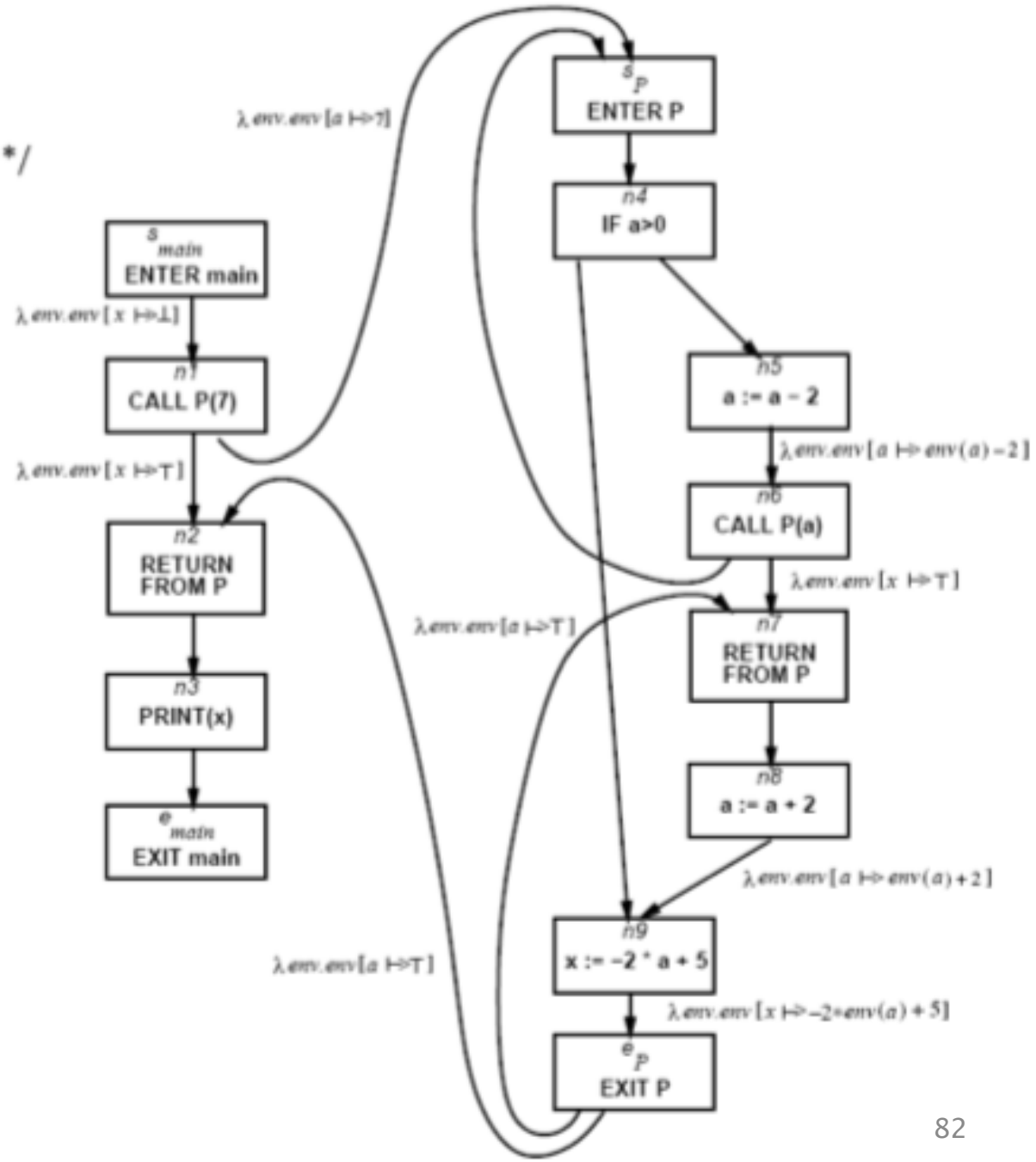


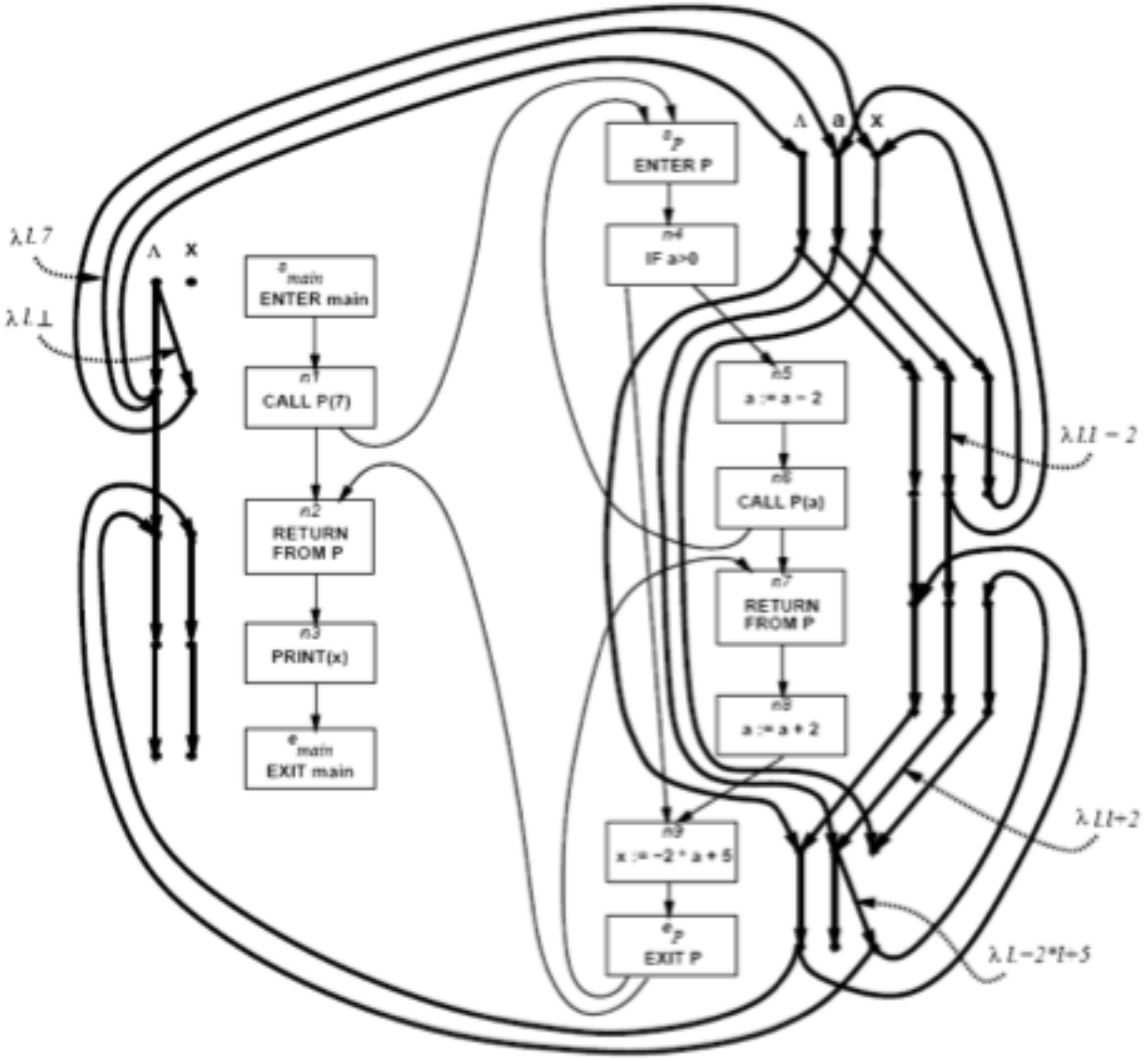
Point-wise representation of environment transformers

IDE Analysis

- Point-wise representation closed under composition
- CFL-Reachability on the exploded graph
- Compose functions

```
declare x: integer
program main
begin
        call P(7)
        print (x) /* x is a constant here */
end

procedure P (value a : integer)
begin /* a is not a constant here */
        if a > 0 then
           a := a − 2
           call P (a)
           a := a + 2
        fi
        x := −2 * a + 5
        /* x is not a constant here */
end
```




λ.L.7
λ.l.⊥
Λ
x
s main
ENTER main
n1
CALL P(7)
n2
RETURN
FROM P
n3
PRINT(x)
e main
EXIT main
s P
ENTER P
n4
IF a>0
n5
a := a - 2
n6
CALL P(a)
n7
RETURN
FROM P
n8
a := a + 2
n9
x := -2 * a + 5
e P
EXIT P
Λ
a
x
λ.l.l − 2
λ.l.l+2
λ.L−2*l+5

Costs

- $O(ED^3)$
- Class of value transformers $F \subseteq L \rightarrow L$
  - $id \in F$
  - Finite height
- Representation scheme with (efficient)
    - Application
    - Composition
    - Join
    - Equality
    - Storage

Conclusion

- Handling functions is crucial for abstract interpretation
- Virtual functions and exceptions complicate things
- But scalability is an issue
  - Small call strings
  - Small functional domains
  - Demand analysis

Challenges in Interprocedural Analysis

- Respect call-return mechanism
- Handling recursion
- Local variables
- Parameter passing mechanisms
- The called procedure is not always known
- The source code of the called procedure is not always available

A trivial treatment of procedure

- Analyze a single procedure
- After every call continue with conservative information
  - Global variables and local variables which “may be modified by the call” have unknown values
- Can be easily implemented
- Procedures can be written in different languages
- Procedure inline can help

Disadvantages of the trivial solution

- Modular (object oriented and functional) programming encourages small frequently called procedures
- Almost all information is lost

Bibliography

- Textbook 2.5
- Patrick Cousot & Radhia Cousot. Static determination of dynamic properties of recursive procedures In IFIP Conference on Formal Description of Programming Concepts, E.J. Neuhold, (Ed.), pages 237-277, St-Andrews, N.B., Canada, 1977. North-Holland Publishing Company (1978).
- Two Approaches to interprocedural analysis by Micha Sharir and Amir Pnueli
- IDFS Interprocedural Distributive Finite Subset Precise interprocedural dataflow analysis via graph reachability. Reps, Horowitz, and Sagiv, POPL'95
- IDE Interprocedural Distributive Environment Precise interprocedural dataflow analysis with applications to constant propagation. Sagiv, Reps, Horowitz, and TCS'96