A Relational Definition of flow-sensitive May-Happen-in-Parallel Analysis

I. Hayes, D. Wainwright, K. Winter, C. Zhang

School of ITEE, The University of Queensland, Brisbane, Australia
Oracle Labs, Brisbane, Australia

Brisbane November 2014

Analysis of Concurrent Code

May-Happen-in-Parallel (MHP)

Analysis of Concurrent Code

May-Happen-in-Parallel (MHP)

- determines pairs of program locations that are potentially in parallel

Analysis of Concurrent Code

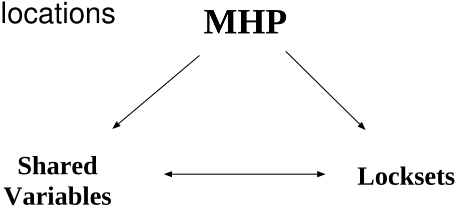
May-Happen-in-Parallel (MHP)

- determines pairs of program locations that are potentially in parallel
- increase precision

Analysis of Concurrent Code

May-Happen-in-Parallel (MHP)

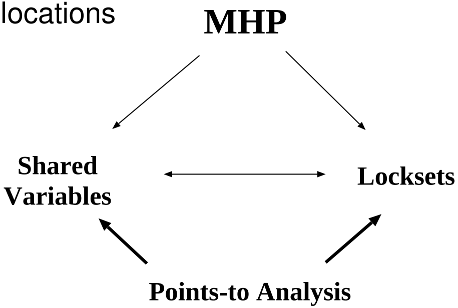
- determines pairs of program locations that are potentially in parallel
- increase precision



Analysis of Concurrent Code

May-Happen-in-Parallel (MHP)

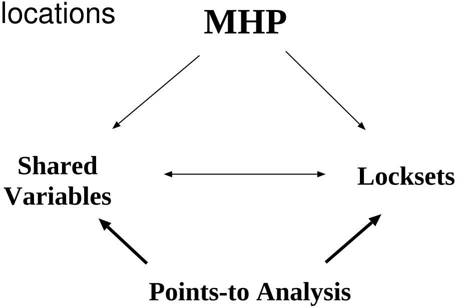
- determines pairs of program locations that are potentially in parallel
- increase precision



Analysis of Concurrent Code

May-Happen-in-Parallel (MHP)

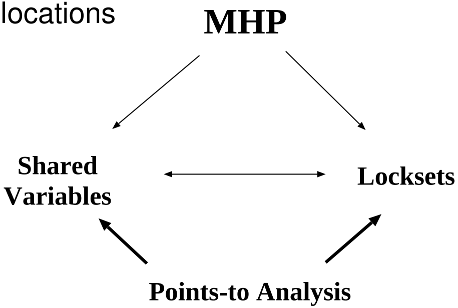
- determines pairs of program locations that are potentially in parallel
- increase precision
- reduce complexity



Analysis of Concurrent Code

May-Happen-in-Parallel (MHP)

- determines pairs of program locations that are potentially in parallel
- increase precision
- reduce complexity



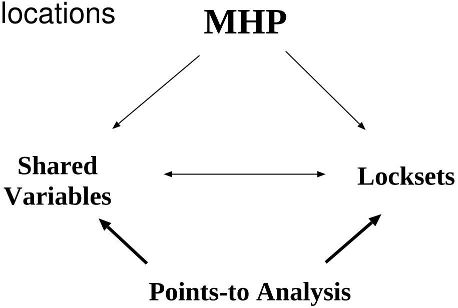
- flow-, context-sensitive

Analysis of Concurrent Code

May-Happen-in-Parallel (MHP)

- determines pairs of program locations that are potentially in parallel
- increase precision
- reduce complexity

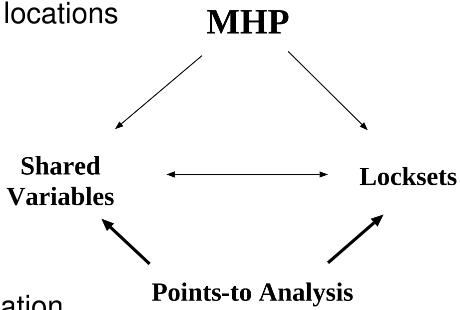
- flow-, context-sensitive
- definition



Analysis of Concurrent Code

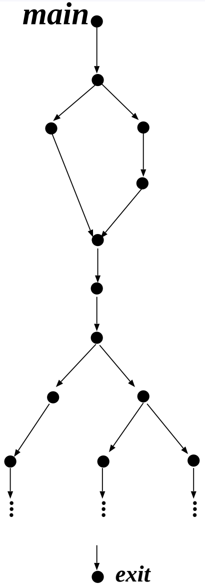
May-Happen-in-Parallel (MHP)

- determines pairs of program locations that are potentially in parallel
- increase precision
- reduce complexity

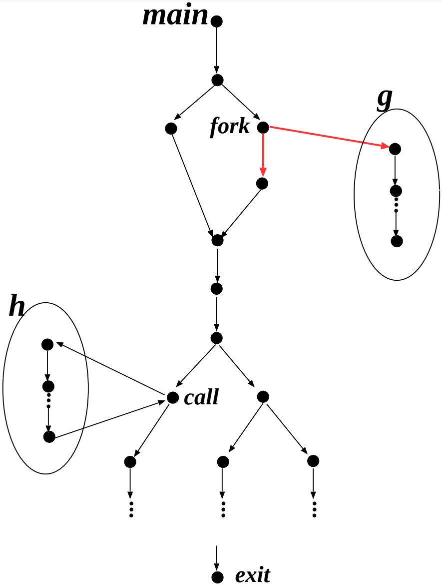


- flow-, context-sensitive
- definition
- relational → Datalog specification

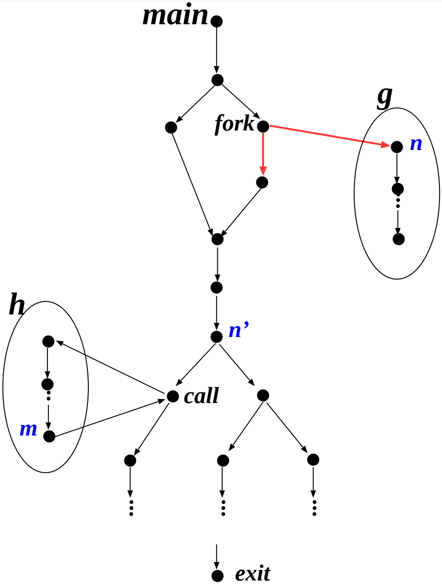
Parallel Execution



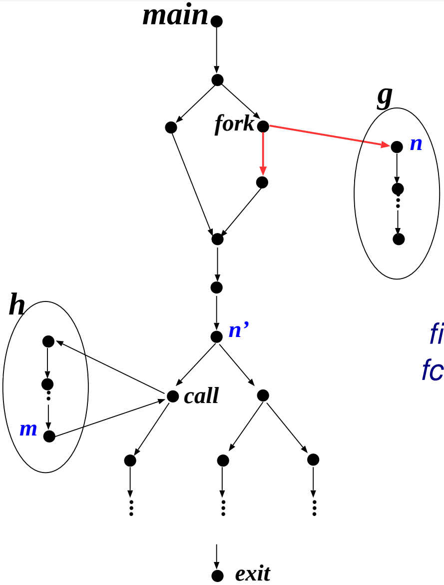
Parallel Execution



Parallel Execution



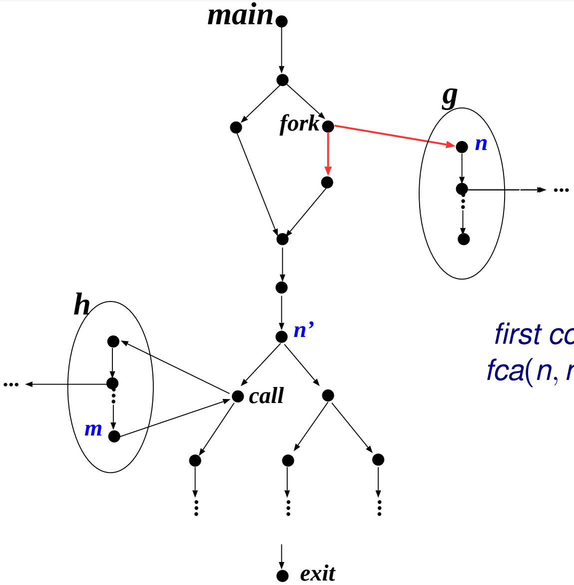
Parallel Execution



first common ancestor

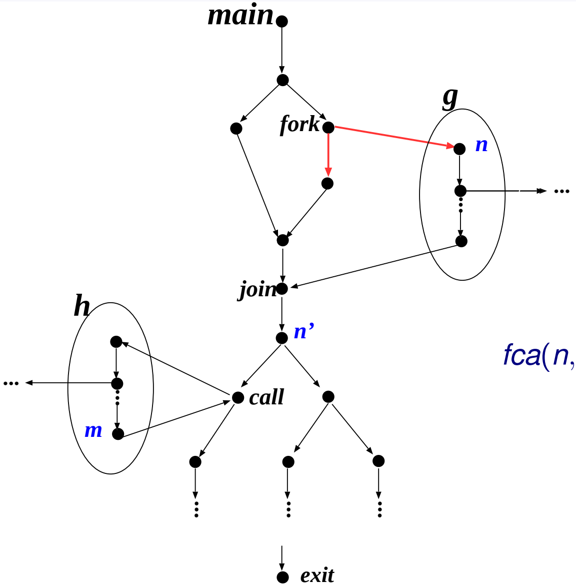
$fca(n, n') \in N^{fork}$

Parallel Execution



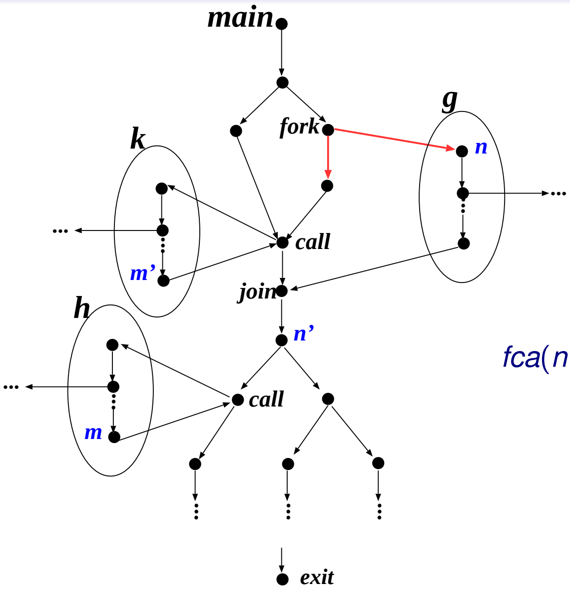
first common ancestor
$fca(n, n') \in N^{fork}$

Parallel Execution



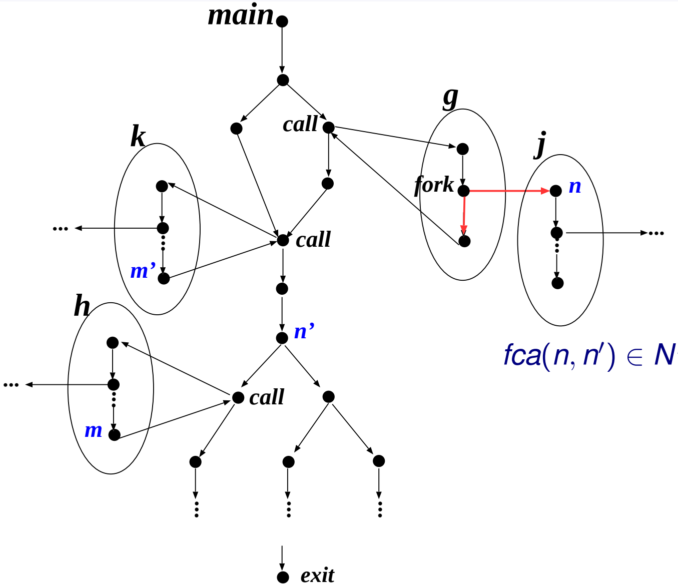
$$fca(n, n') \in N^{fork} \wedge (n', n) \notin anc$$

Parallel Execution



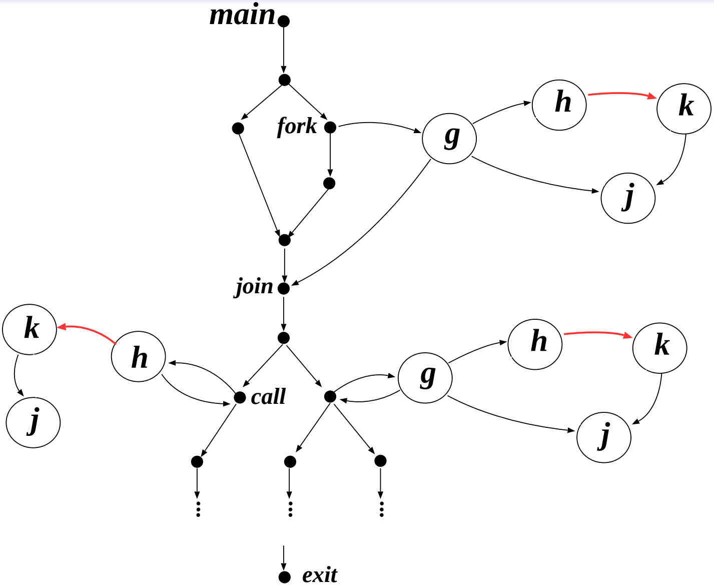
$$fca(n, n') \in N^{fork} \wedge (n', n) \notin anc$$

Parallel Execution



$fca(n, n') \in N^{fork} \wedge (n', n) \notin anc$

Abstraction: Call Graphs

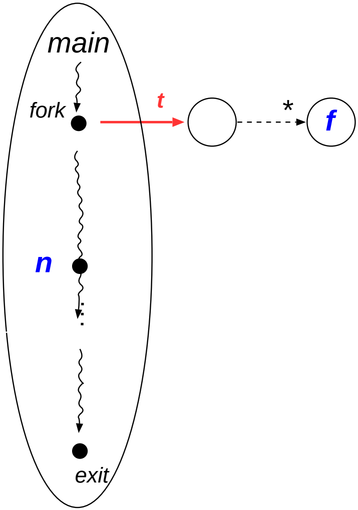


Four Types of MHP

- Descendant MHP
- Sibling MHP

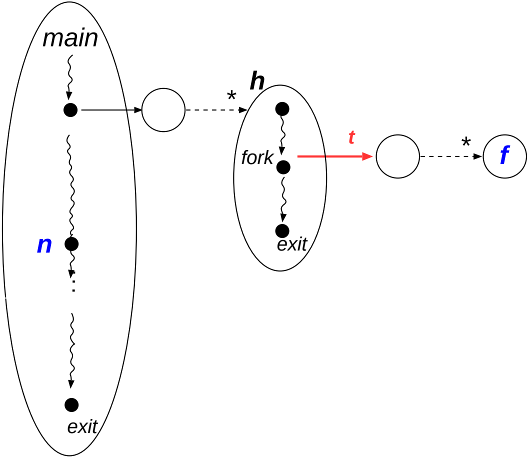
Four Types of MHP

- Descendant MHP
  - **direct**
- Sibling MHP



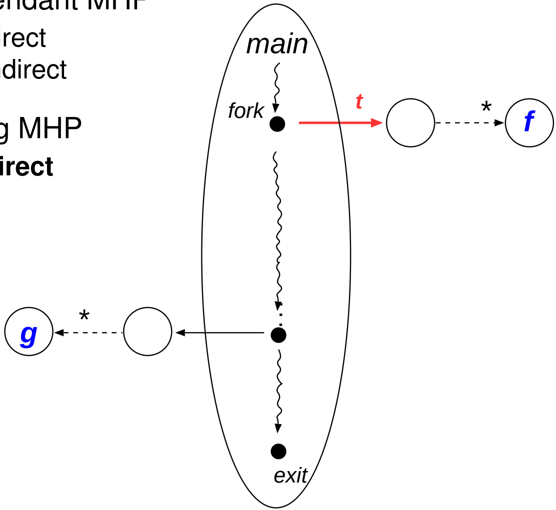
Four Types of MHP

- Descendant MHP
  - direct
  - **indirect**
- Sibling MHP



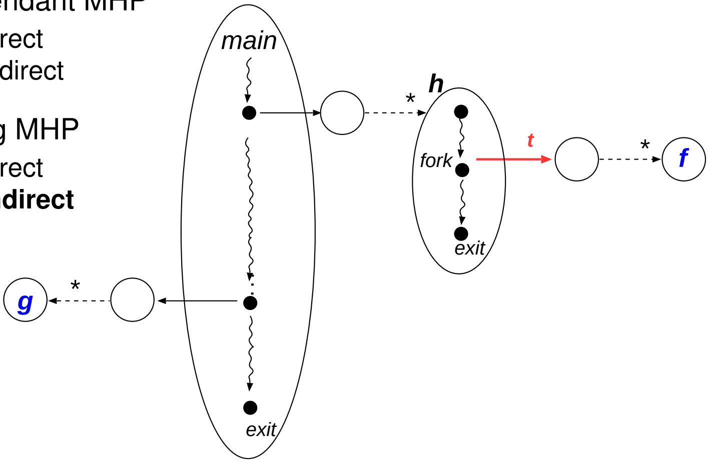
Four Types of MHP

- Descendant MHP
  - direct
  - indirect
- Sibling MHP
  - **direct**

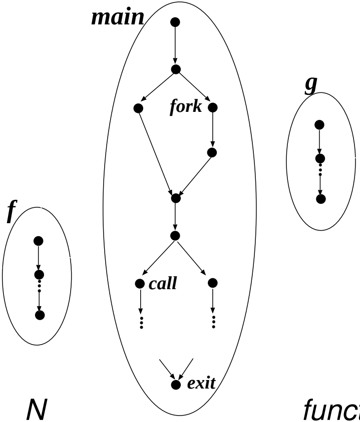


Four Types of MHP

- Descendant MHP
  - direct
  - indirect
- Sibling MHP
  - direct
  - **indirect**

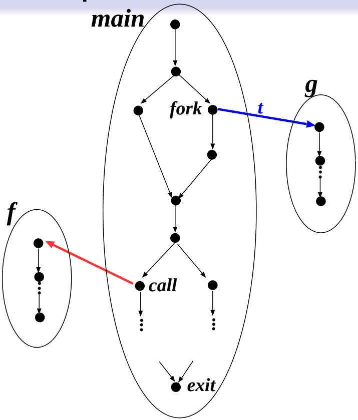


Control Flow Graph



nodes	N	*functions*	F
edges	$E \subseteq N \times N$		$NF \in N \to F$
label	$N \to S$		

Control Flow Graph



$$
\begin{aligned}
\mathit{Call} &= \{(n,f) \in N \times F \mid \mathit{label}(n) = \textbf{call } f\} \\
\mathit{Fork}_t &= \{(n,g) \in N \times F \mid \mathit{label}(n) = t \leftarrow \textbf{fork } g\} \\
\mathit{Fork} &= \bigcup_t \mathit{Fork}_t
\end{aligned}
$$

Relational Composition

with $R \subseteq A \times B$ and $S \subseteq B \times C$ and $Q \subseteq A \times C$

$$
\begin{aligned}
R \fatsemi S &= \{(a,c) \in A \times C \mid \exists b \cdot (a,b) \in R \land (b,c) \in S\} \\
R^{\sim} &= \{(b,a) \mid (a,b) \in R\} \\
R \parallel Q &= \{(b,c) \in B \times C \mid \exists a \cdot (a,b) \in R \land (a,c) \in Q\} \\
&= R^{\sim} \fatsemi Q
\end{aligned}
$$

Relational Composition

with $R \subseteq A \times B$ and $S \subseteq B \times C$ and $Q \subseteq A \times C$

$$
\begin{aligned}
R \fatsemi S &= \{(a, c) \in A \times C \mid \exists b \cdot (a, b) \in R \wedge (b, c) \in S\} \\
R^{\sim} &= \{(b, a) \mid (a, b) \in R\} \\
R \parallel Q &= \{(b, c) \in B \times C \mid \exists a \cdot (a, b) \in R \wedge (a, c) \in Q\} \\
&= R^{\sim} \fatsemi Q
\end{aligned}
$$

$$
\begin{aligned}
NF &\in N \to F \\
Call &= \{(n, f) \in N \times F | label(n) = \textbf{call } f\}
\end{aligned}
$$

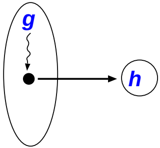
Relational Composition

with $R \subseteq A \times B$ and $S \subseteq B \times C$ and $Q \subseteq A \times C$

$$
\begin{aligned}
R \fatsemi S &= \{(a,c) \in A \times C \mid \exists b \cdot (a,b) \in R \wedge (b,c) \in S\} \\
R^{\sim} &= \{(b,a) \mid (a,b) \in R\} \\
R \,||\, Q &= \{(b,c) \in B \times C \mid \exists a \cdot (a,b) \in R \wedge (a,c) \in Q\} \\
&= R^{\sim} \fatsemi Q
\end{aligned}
$$

$$
\begin{aligned}
NF &\in N \to F \\
Call &= \{(n,f) \in N \times F | label(n) = \textbf{call } f\}
\end{aligned}
$$

$$
\begin{aligned}
FCall &= NF^{\sim} \fatsemi Call && F \times F \\
FFork &= NF^{\sim} \fatsemi Fork && F \times F \\
FCallFork &= FCall \cup FFork && F \times F
\end{aligned}
$$



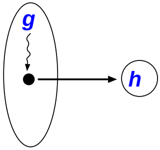
Relational Composition

with $R \subseteq A \times B$ and $S \subseteq B \times C$ and $Q \subseteq A \times C$

$$
\begin{aligned}
R \fatsemi S &= \{(a,c) \in A \times C \mid \exists b \cdot (a,b) \in R \wedge (b,c) \in S\} \\
R^{\sim} &= \{(b,a) \mid (a,b) \in R\} \\
R \mid\mid Q &= \{(b,c) \in B \times C \mid \exists a \cdot (a,b) \in R \wedge (a,c) \in Q\} \\
&= R^{\sim} \fatsemi Q
\end{aligned}
$$

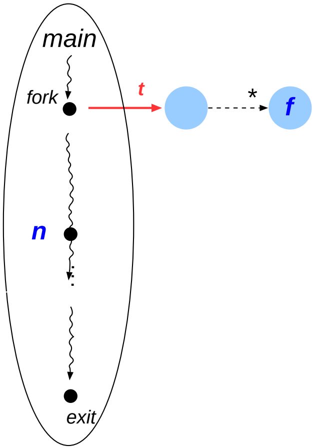
$$
\begin{aligned}
NF &\in N \rightarrow F \\
Call &= \{(n,f) \in N \times F | label(n) = \textbf{call } f\}
\end{aligned}
$$

$$
\begin{aligned}
FCall &= NF^{\sim} \fatsemi Call && F \times F \\
FFork &= NF^{\sim} \fatsemi Fork && F \times F \\
FCallFork &= FCall \cup FFork && F \times F
\end{aligned}
$$



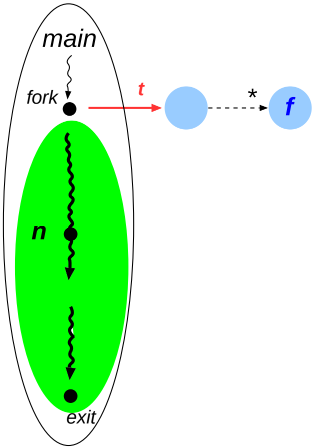
$FCallFork^*$ (Call Graph)

Direct Descendant MHP



$$FMHP_t \quad = \quad Fork_t \fatsemi FCallFork^* \fatsemi NF^{\sim}$$

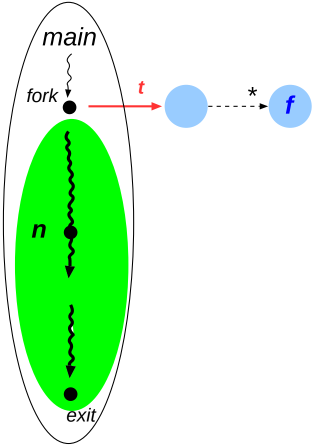
Direct Descendant MHP



$$FMHP_t \quad = \quad Fork_t \fatsemi FCallFork^* \fatsemi NF^{\sim}$$

$$\begin{aligned} NoJ_t &= (E - EJ_t)^* \\ EJ_t &= (N \times NJ_t) \cup (NJ_t \times N) \\ NJ_t &= \{n \in N | label(n) = \textbf{join}\ t\} \end{aligned}$$

Direct Descendant MHP

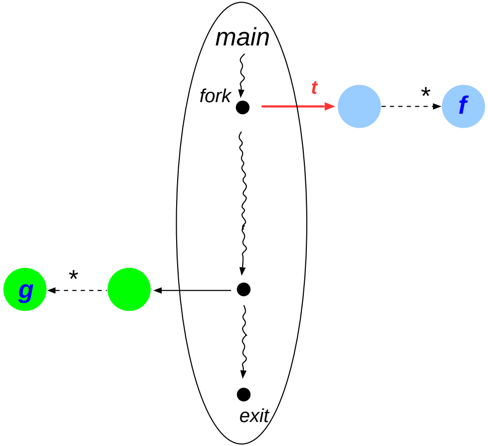


$$
\begin{aligned}
FMHP_t &= Fork_t \fatsemi FCallFork^* \fatsemi NF^{\sim}
\end{aligned}
$$

$$
\begin{aligned}
NoJ_t &= (E - EJ_t)^* \\
EJ_t &= (N \times NJ_t) \cup (NJ_t \times N) \\
NJ_t &= \{n \in N \mid label(n) = \textbf{join } t\}
\end{aligned}
$$

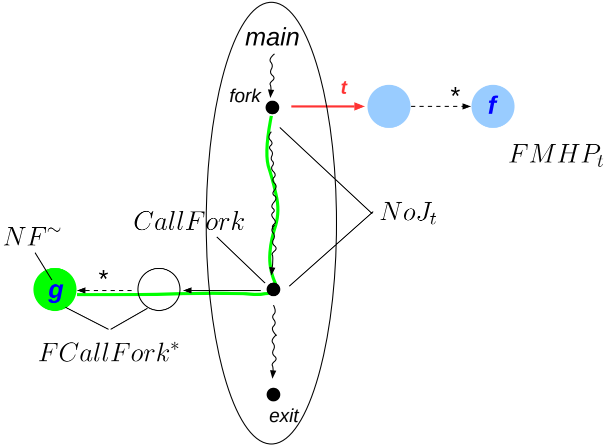
$$
directDMHP_t = NoJ_t \parallel FMHP_t
$$

Direct Sibling MHP



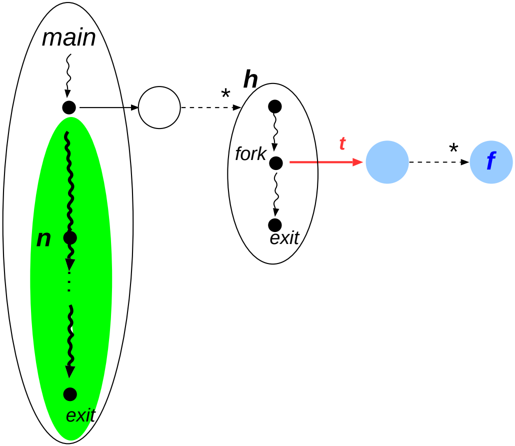
$$directSMHP_t \quad = \quad NoJ_t \,⨾\, CallFork \,⨾\, FCallFork^* \,⨾\, NF^{\sim} \;||\; FMHP_t$$

Direct Sibling MHP



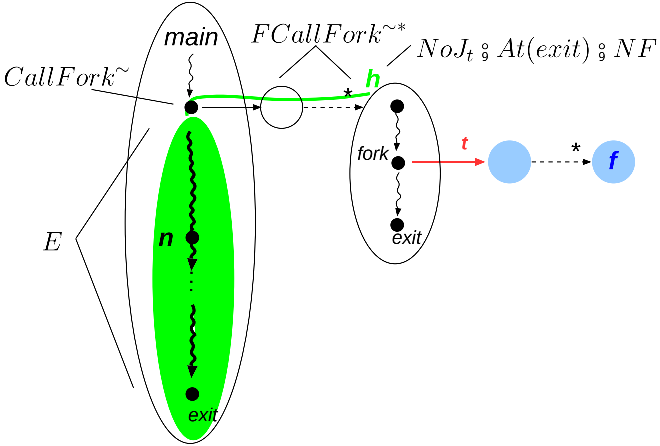
$$directSMHP_t = NoJ_t \fatsemi CallFork \fatsemi FCallFork^* \fatsemi NF^{\sim} \mid\mid FMHP_t$$

Indirect Descendant MHP



$$indirectDMHP_t \;=\; NoJ_t \fatsemi At(exit) \fatsemi NF \fatsemi FCallFork^{\sim *} \fatsemi CallFork^{\sim} \fatsemi E^* \;||\; FMHP_t$$

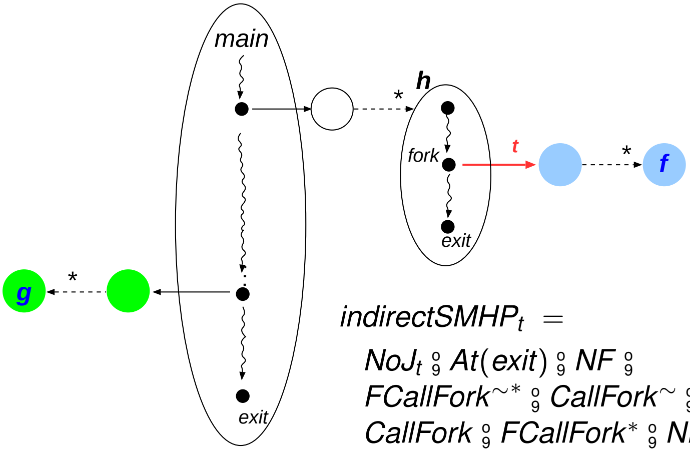
Indirect Descendant MHP



$$indirectDMHP_t =$$

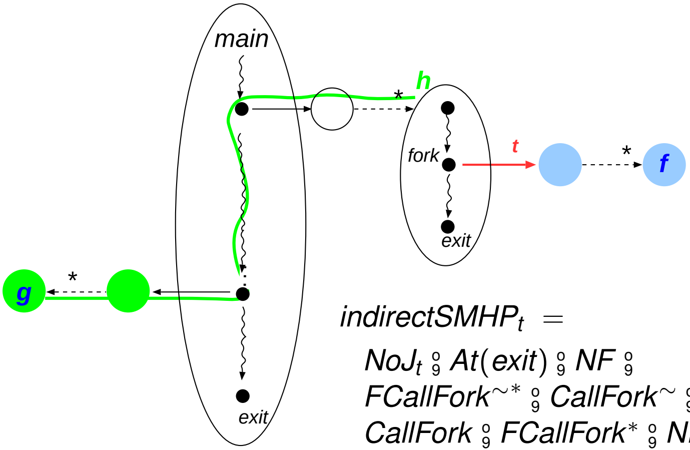
$$NoJ_t \fatsemi At(exit) \fatsemi NF \fatsemi FCallFork^{\sim *} \fatsemi CallFork^{\sim} \fatsemi E^{*} \mid\mid FMHP_t$$

Indirect Sibling MHP



$$
\begin{aligned}
&\mathit{indirectSMHP}_t = \\
&\quad \mathit{NoJ}_t \fatsemi \mathit{At}(\mathit{exit}) \fatsemi \mathit{NF} \fatsemi \\
&\quad \mathit{FCallFork}^{\sim *} \fatsemi \mathit{CallFork}^{\sim} \fatsemi E^{*} \fatsemi \\
&\quad \mathit{CallFork} \fatsemi \mathit{FCallFork}^{*} \fatsemi \mathit{NF}^{\sim} \\
&\quad || \\
&\quad \mathit{FMHP}_t
\end{aligned}
$$

Indirect Sibling MHP



$$
\begin{aligned}
&\mathit{indirectSMHP}_t \;=\\
&\quad \mathit{NoJ}_t \fatsemi \mathit{At}(\mathit{exit}) \fatsemi \mathit{NF} \fatsemi\\
&\quad \mathit{FCallFork}^{\sim *} \fatsemi \mathit{CallFork}^{\sim} \fatsemi E^{*} \fatsemi\\
&\quad \mathit{CallFork} \fatsemi \mathit{FCallFork}^{*} \fatsemi \mathit{NF}^{\sim}\\
&\quad \|\\
&\quad \mathit{FMHP}_t
\end{aligned}
$$

Conclusion and Outlook

- Definiton of May-Happen-in-Parallel:
  - as Relations
  - (non-relational)

Conclusion and Outlook

- Definiton of May-Happen-in-Parallel:
  - as Relations
  - (non-relational)

- Datalog implementation

Conclusion and Outlook

- Definiton of May-Happen-in-Parallel:
  - as Relations
  - (non-relational)

- Datalog implementation
- Compare with current implementation

Conclusion and Outlook

- Definiton of May-Happen-in-Parallel:
  - as Relations
  - (non-relational)

- Datalog implementation
- Compare with current implementation
- Adapt relational definitions to improve efficiency

Conclusion and Outlook

- Definiton of May-Happen-in-Parallel:
  - as Relations
  - (non-relational)

- Datalog implementation
- Compare with current implementation
- Adapt relational definitions to improve efficiency
- Improve other work on data races using Datalog (e.g, [Naik, Aiken, Whaley, PDLI'06])

Conclusion and Outlook

- Definiton of May-Happen-in-Parallel:
  - as Relations
  - (non-relational)

- Datalog implementation
- Compare with current implementation
- Adapt relational definitions to improve efficiency
- Improve other work on data races using Datalog (e.g, [Naik, Aiken, Whaley, PDLI'06])

- Lockset and Shared Variable Analysis
- Points-to Analysis