MPRI

Abstract Interpretation of Mobile Systems

Jérôme Feret
Département d'Informatique de l'École Normale Supérieure
INRIA, ÉNS, CNRS
http://www.di.ens.fr/~feret

January 14th, 2015

Overview

1. Overview
2. Mobile systems
3. Non standard semantics
4. Abstract Interpretation
5. Environment analyses
6. Occurrence counting analysis
7. Thread partitioning
8. Conclusion

Collecting semantics

$(\mathcal{C}, C_0, \rightarrow)$ is a transition system,
We restrict our study to its collecting semantics:
this is the set of the states that are reachable within a finite transition sequence.

$$\mathcal{S} = \{C \mid \exists i \in C_0,\ i \rightarrow^* C\}$$

It is also given by the least fixpoint of the following $\cup$-complete endomorphism $\mathbb{F}$:

$$\mathbb{F} = \begin{cases} \wp(\mathcal{C}) & \rightarrow \wp(\mathcal{C}) \\ X & \mapsto C_0 \cup \{C' \mid \exists C \in X,\ C \rightarrow C'\} \end{cases}$$

This fixpoint is usually not computable automatically.

Abstract domain

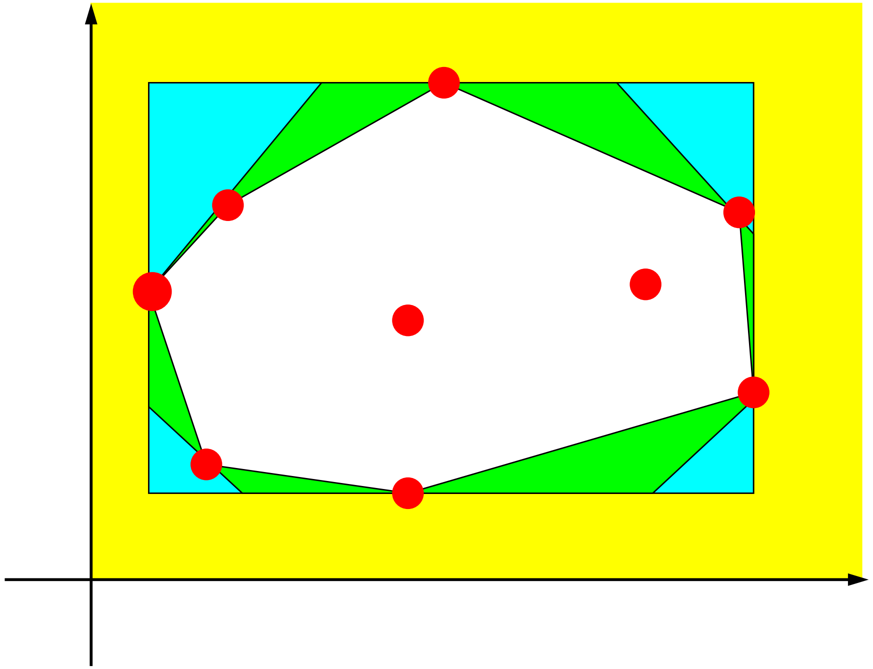
We introduce an abstract domain of properties:

- properties of interest;
- more complex properties used in calculating them.

This domain is often a lattice: $(\mathcal{D}^\sharp, \sqsubseteq, \sqcup, \bot, \sqcap, \top)$ and is related to the concrete domain $\wp(\mathcal{C})$ by a monotonic concretization function γ.

$\forall A \in \mathcal{D}^\sharp$, $\gamma(A)$ is the set of the elements which satisfy the property A.

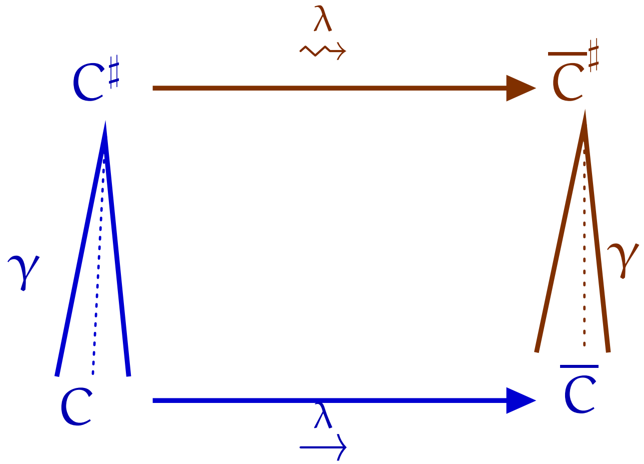
Numerical domains



- sign approximation;
- interval approximation;
- octagonal approximation;
- polyhedra approximation;
- concrete domain.

Abstract transition system

Let $C_0^\sharp$ be an abstraction of the initial states and $\rightsquigarrow$ be an abstract transition relation, which satisfies $C_0 \subseteq \gamma(C_0^\sharp)$ and the following diagram:



Then, $\mathcal{S} \subseteq \bigcup_{n \in \mathbb{N}} \gamma(\mathbb{F}^{\sharp n}(C_0^\sharp))$,

where $\mathbb{F}^\sharp(C^\sharp) = C_0^\sharp \sqcup C^\sharp \sqcup \left(\bigsqcup_{finite}\{\overline{C^\sharp} \mid C^\sharp \rightsquigarrow \overline{C^\sharp}\}\right)$.

Widening operator

We require a widening operator to ensure the convergence of the analysis:

$$\nabla : D^\sharp \times D^\sharp \to D^\sharp$$

such that:

- $\forall X_1^\sharp,\ X_2^\sharp \in D^\sharp,\ X_1^\sharp \sqcup X_2^\sharp \sqsubseteq X_1^\sharp \nabla X_2^\sharp$
- for all increasing sequence $(X_n^\sharp) \in \left(D^\sharp\right)^{\mathbb{N}}$, the sequence (X_n^∇) defined as
$$\begin{cases} X_0^\nabla = X_0^\sharp \\ X_{n+1}^\nabla = X_n^\nabla \ \nabla \ X_{n+1}^\sharp \end{cases}$$
is ultimately stationary.

Abstract iteration

The abstract iteration (C_n^∇) of $\mathbb{F}^\sharp$ defined as follows

$$\begin{cases} C_0^\nabla = C_0^\sharp \\ C_{n+1}^\nabla = \begin{cases} C_n^\nabla & \text{if } \mathbb{F}^\sharp(C_n^\nabla) \sqsubseteq C_n^\nabla \\ C_n^\nabla \nabla \mathbb{F}^\sharp(C_n^\nabla) & \text{otherwise} \end{cases} \end{cases}$$

is ultimately stationary and its limit C^∇ satisfies *lfp*$_\emptyset \mathbb{F} \subseteq \gamma(C^\nabla)$.

Example: Interval widening

We consider the complete $\mathcal{I}$ lattice of the natural number intervals.

$\mathcal{I}$ does not satisfy the increasing chain condition.

Given n a natural number, we use the following widening operator to ensure the convergence of the analyses based on the use of $\mathcal{I}$:

$$\begin{cases} [|a;b|] \ \nabla\ [|c;d|] = [|\mathit{min}\{a;c\};\infty|[& \text{if } d > \mathit{max}\{n;b\} \\ I \ \nabla\ J = I \sqcup J & \text{otherwise} \end{cases}$$

Composing two abstractions

Given two abstractions $(\mathcal{D}^\sharp, \gamma, C_0^\sharp, \leadsto, \nabla)$ and $(\mathcal{D}^\sharp, \gamma, C_0^\sharp, \leadsto, \nabla)$, and a reduction $\rho : \mathcal{D}^\sharp \times \mathcal{D}^\sharp \to \mathcal{D}^\sharp \times \mathcal{D}^\sharp$ which satisfy:

$$\forall (A, A) \in \mathcal{D}^\sharp \times \mathcal{D}^\sharp,\ \gamma(A) \cap \gamma(A) \subseteq \gamma(a) \cap \gamma(a) \text{ where } (a, a) = \rho(A, A).$$

Then $(\mathcal{D}^\sharp, \gamma, C_0^\sharp, \leadsto, \nabla)$ where:

- $\mathcal{D}^\sharp = \mathcal{D}^\sharp \times \mathcal{D}^\sharp$;
- ∇ is pair-wisely defined;
- $\gamma(A, A) = \gamma(A) \cap \gamma(A)$;
- $C_0^\sharp = \rho(C_0^\sharp, C_0^\sharp)$;
- $(A, A) \leadsto \rho(C, C)$
  if $B \leadsto C$ and $B \leadsto C$ and $(B, B) = \rho(A, A)$

is also an abstraction.

Overview

1. Overview
2. Mobile systems
3. Non standard semantics
4. Abstract Interpretation
5. Environment analyses
6. Occurrence counting analysis
7. Thread partitioning
8. Conclusion

Generic environment analysis

For each subset V of variables, we introduce a generic abstract domain $\mathcal{G}_V$ to describe the markers and the environments which may be associated to a syntactic component the free name of which is V:

$$\wp(\mathit{Id} \times (V \to (\mathit{Name} \times \mathit{Id}))) \xleftarrow{\gamma_V} \mathcal{G}_V.$$

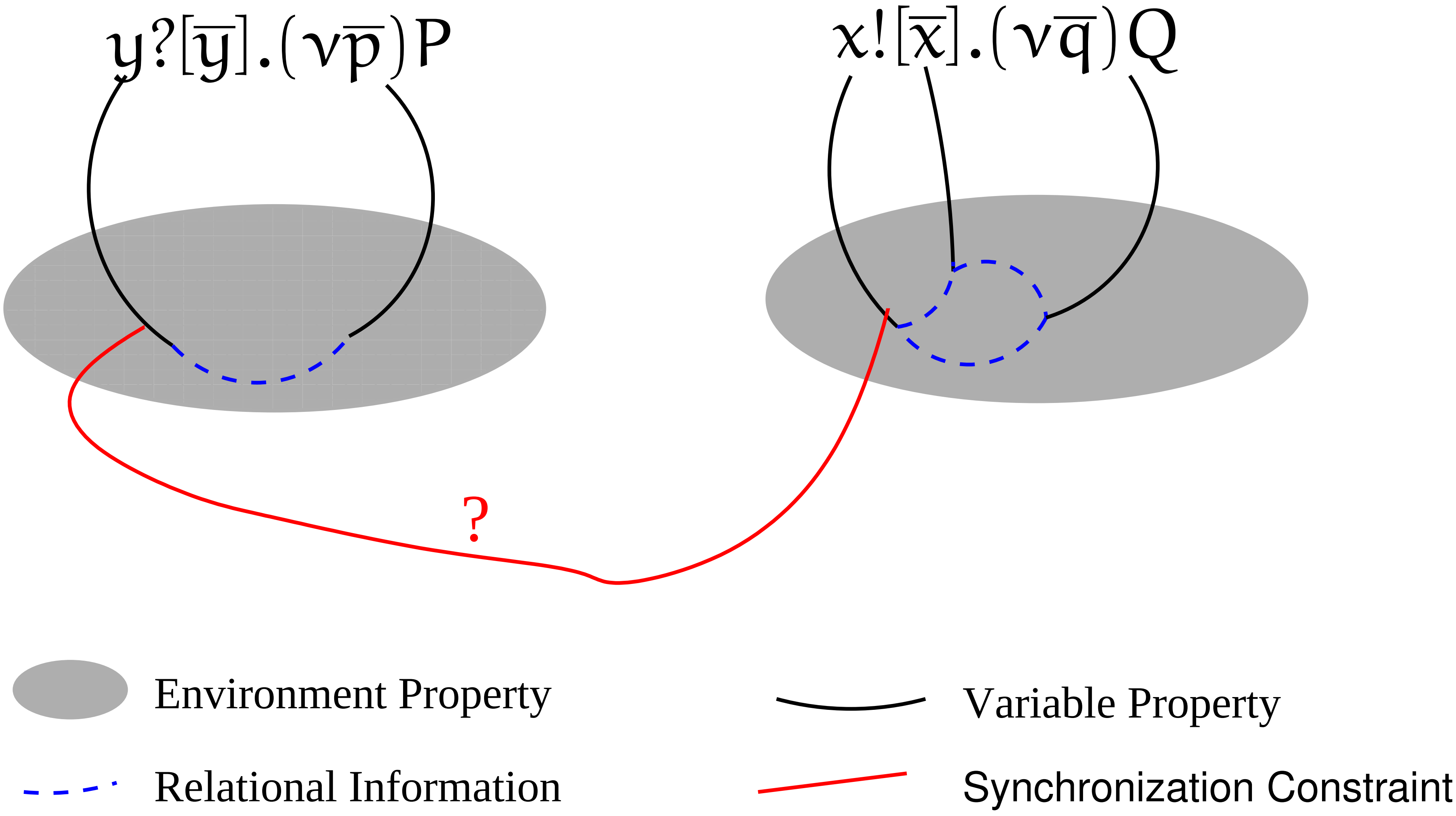
The abstract domain $C^\sharp$ is then the set:

$$C^\sharp = \prod_{p \in \mathcal{P}} \mathcal{G}_{fn(p)}$$

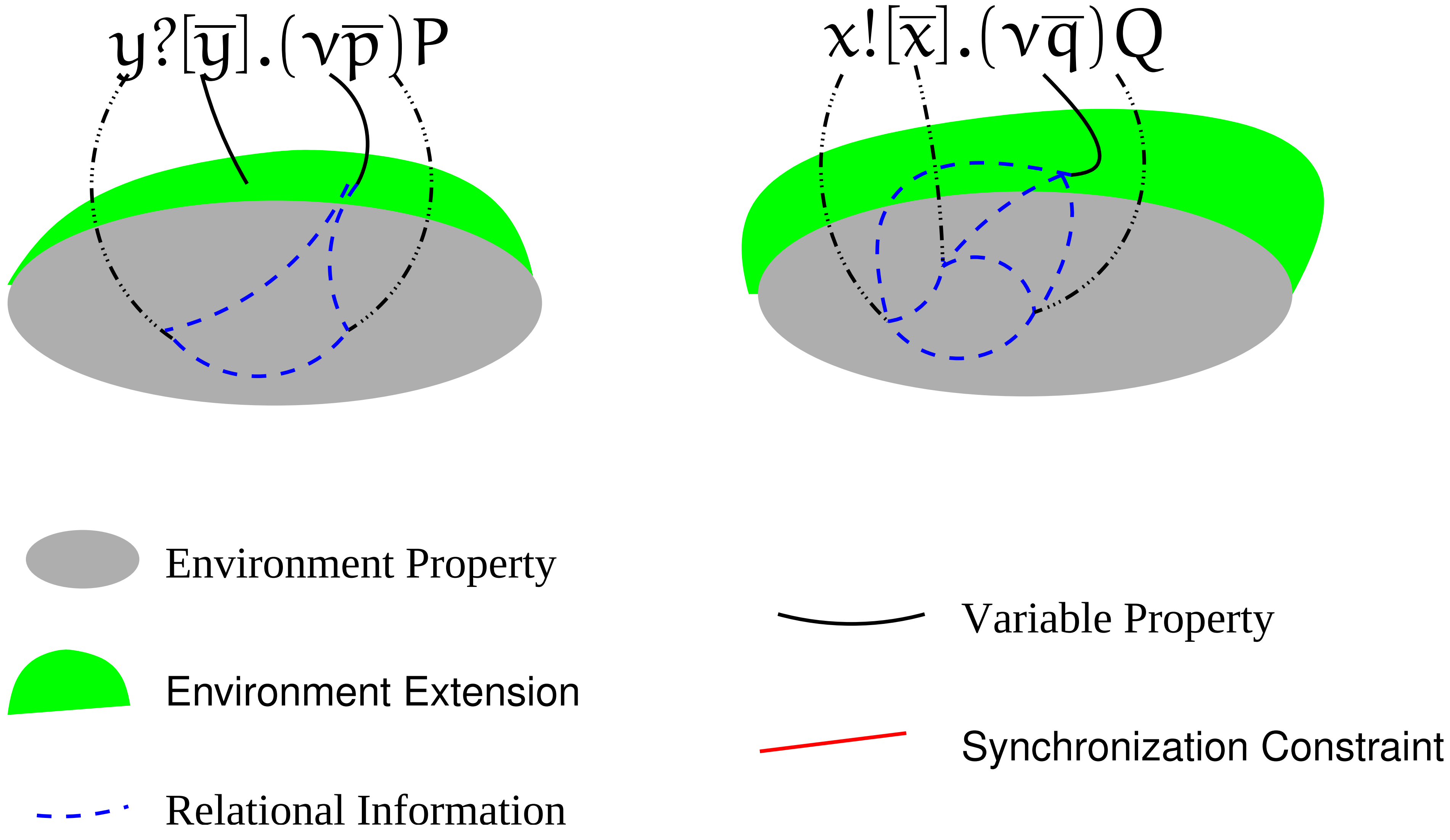
related to $\wp(C)$ by the concretization γ:

$$\gamma(f) = \{C \mid (p, \mathit{id}, E) \in C \implies (\mathit{id}, E) \in \gamma_{fn(p)}(f_p)\}.$$

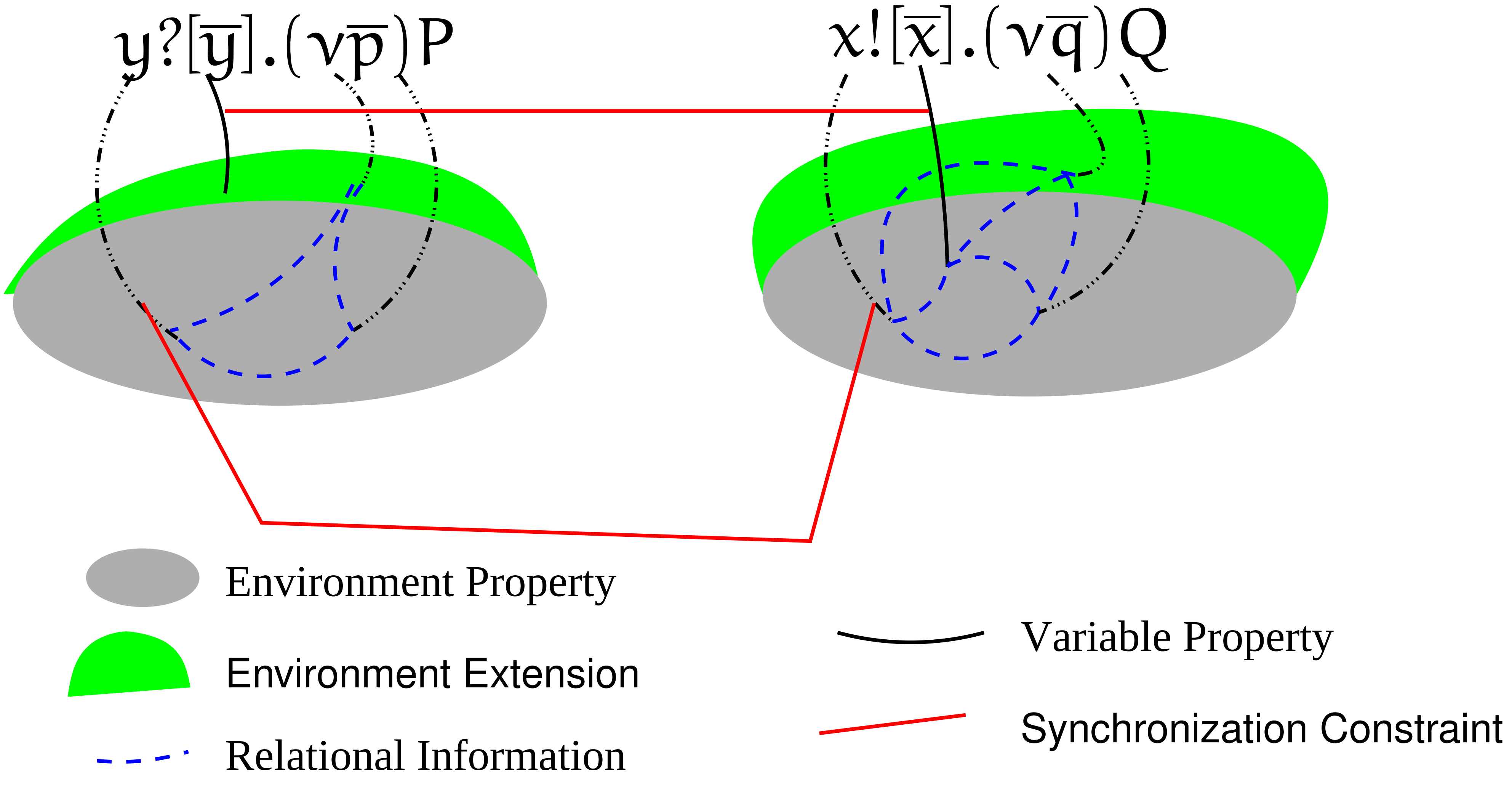
Abstract communication



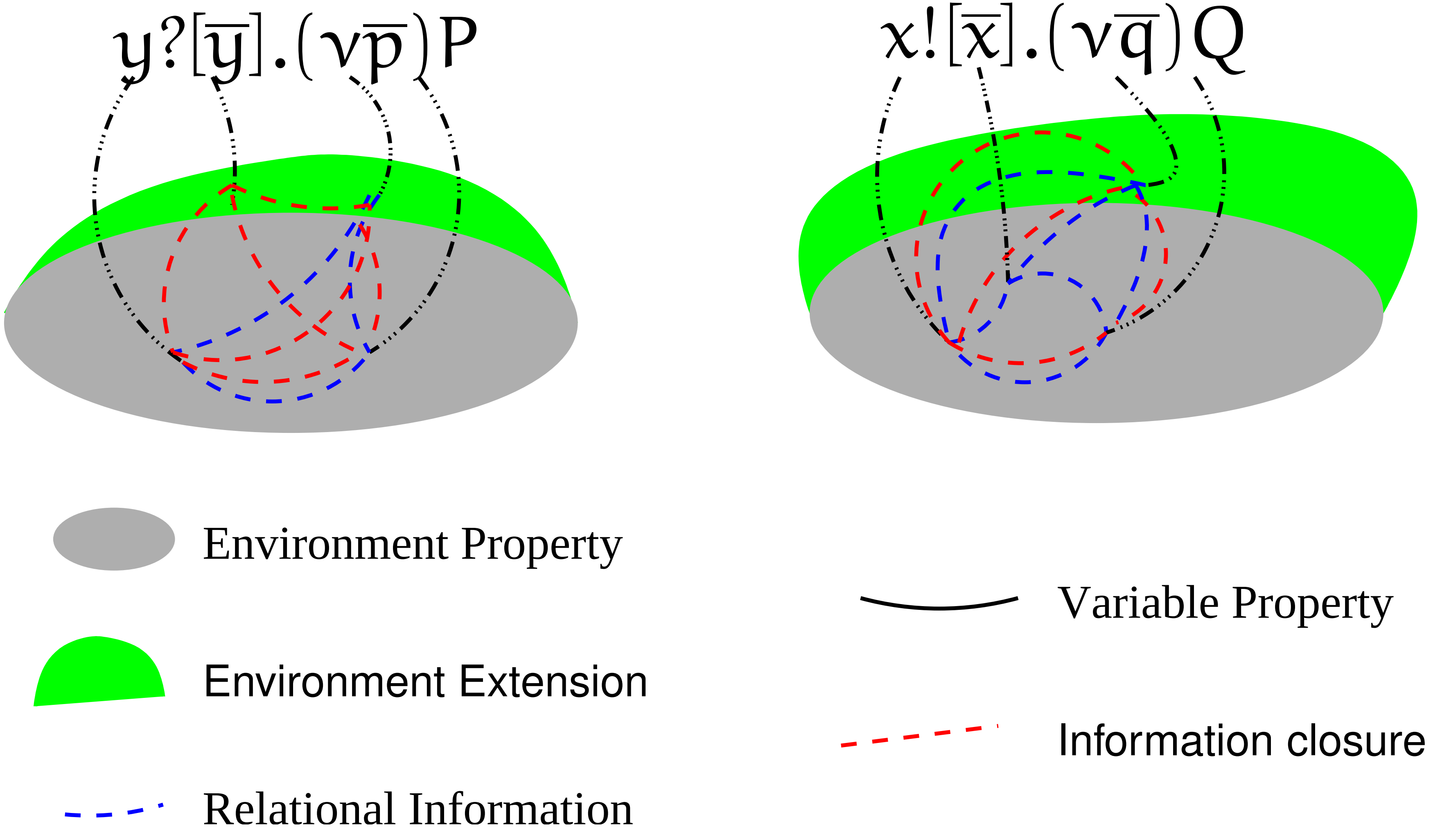
Extending environments



Synchronizing environments



Propagating information

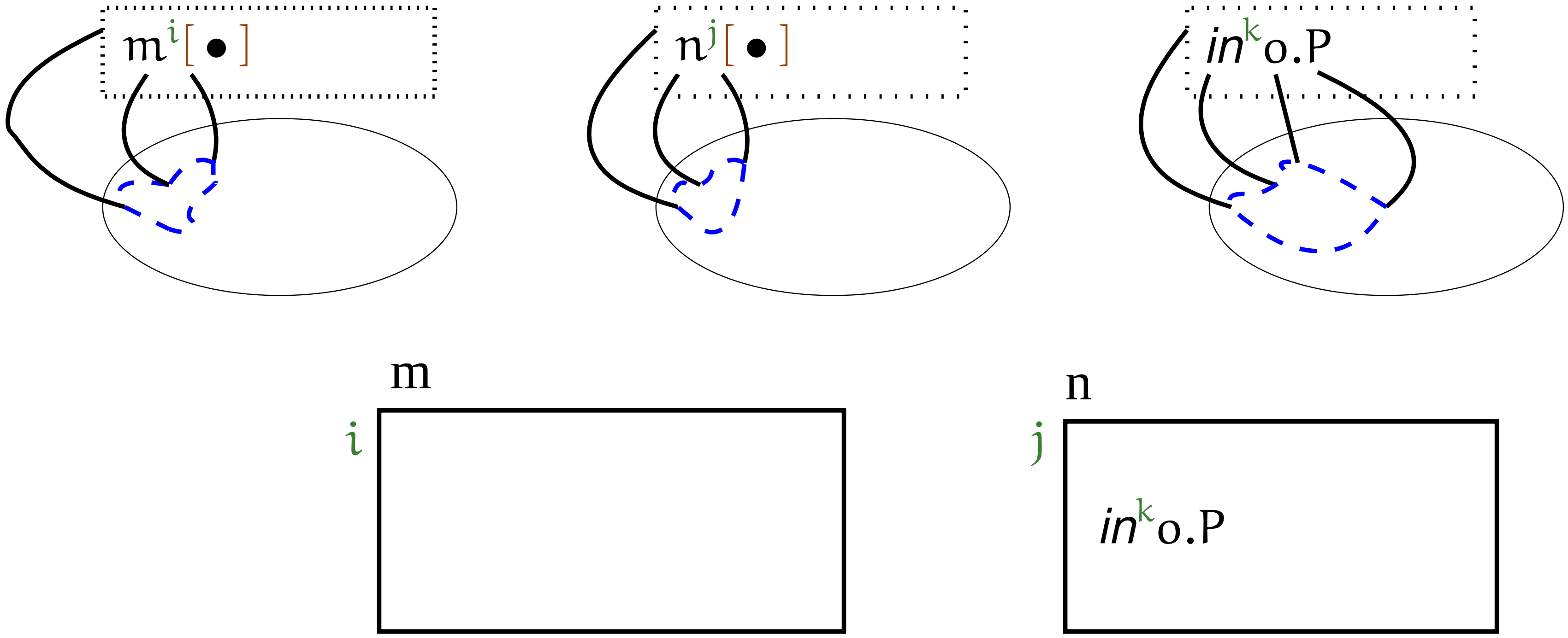


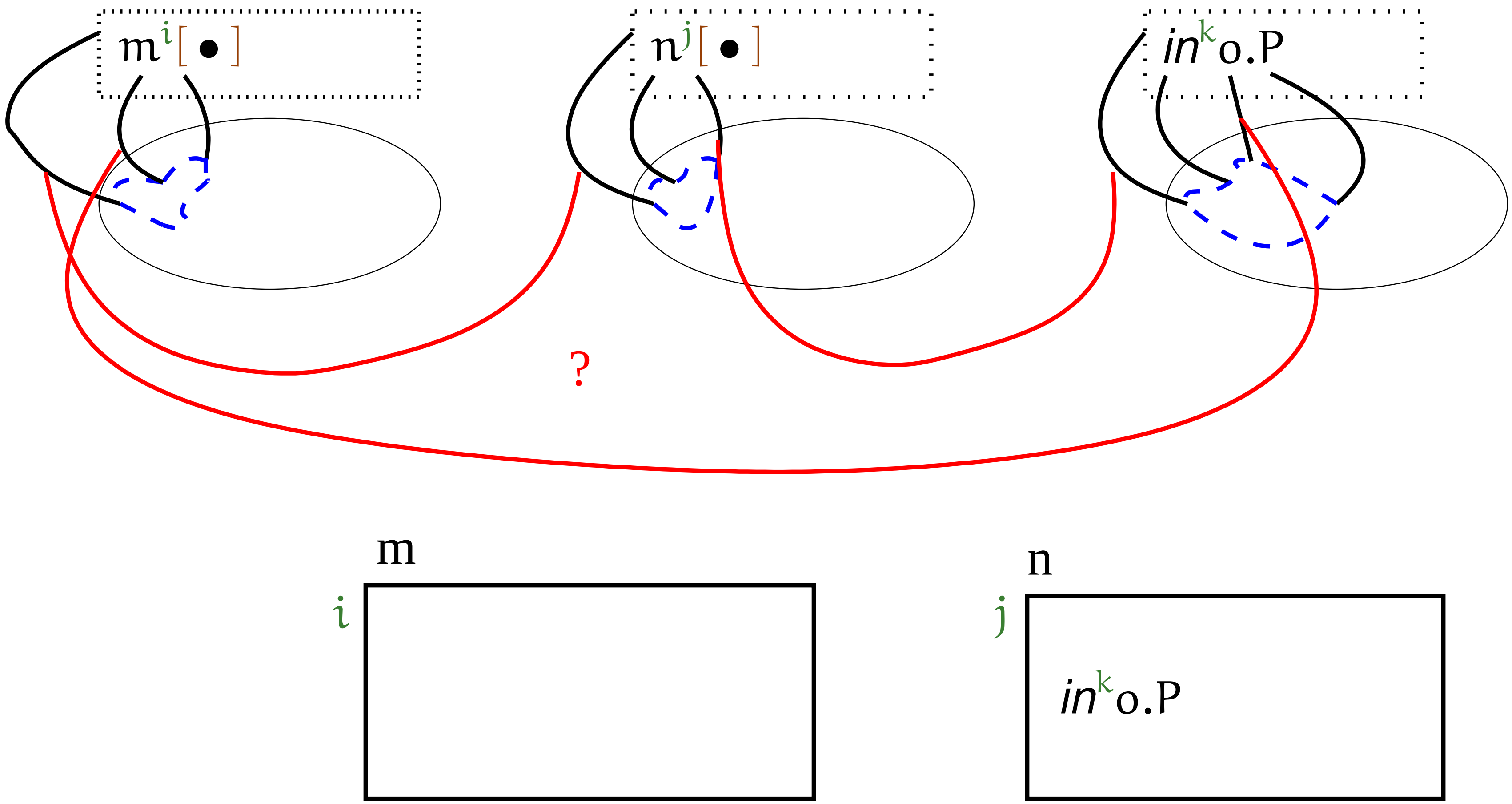
Generic primitives

We only require abstract primitives to:

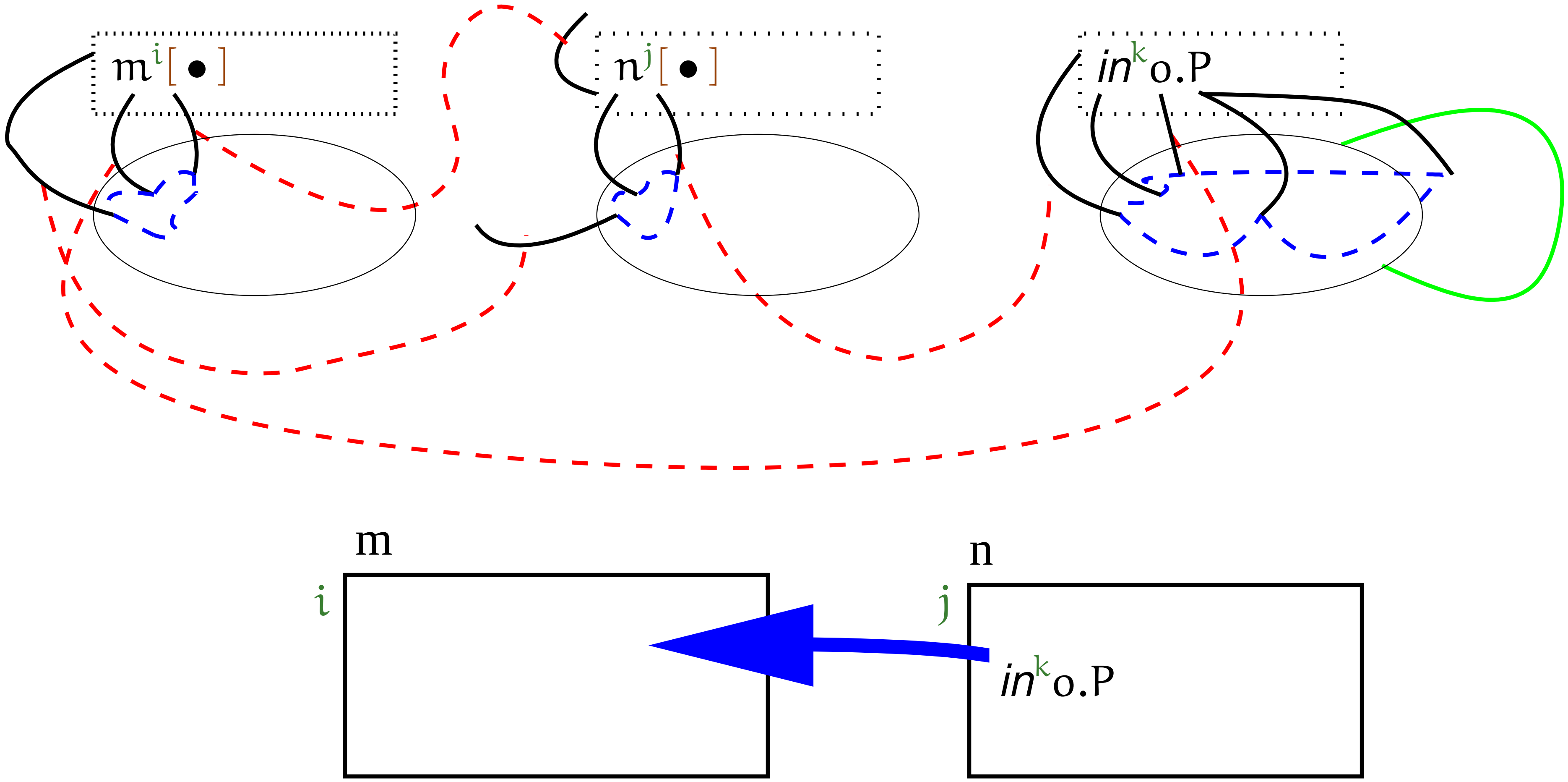
1. extend an environment domain,
2. gather the description of the linkage of two syntactic agents,
3. synchronize variables,
4. separate two descriptions,
5. restrict an environment domain.

About mobile ambients





$m^i[\bullet]$
$n^j[\bullet]$
$in^k o.P$
?
m
i
n
j
$in^k o.P$


$m^i[\bullet]$
$n^j[\bullet]$
$in^k o.P$
m
i
n
j
$in^k o.P$

Control flow analyses

We abstract for each variable x and each name restriction $\nu\, y$ the set of marker pairs $(\mathrm{id}_x, \mathrm{id}_y)$ such that the channel opened by the instance of the restriction $\nu\, y$ tagged with the marker id_y may be communicated to the variable x of a thread tagged by the marker id_x.

Let $Id^\sharp$ be an abstract domain of properties about marker pairs.

$$\gamma_{Id^2} : Id^\sharp \to \wp(Id^2)$$

$$\mathcal{G}_V = V \times Name \to Id^\sharp$$

$\gamma_V(a^\sharp)$ is the set of marker/environment pairs (id_x, E) such that:

$$\forall x \in V, E(x) = (y, id_y) \implies (id_x, id_y) \in \gamma_{Id^2}(a^\sharp(x, y)).$$

Regular approximation

We approximate the shape of the markers which may be associated to channel names linked to variables, and syntactic components, without relations among them.
We use the following abstract domain:

$$\wp(\Sigma) \times \wp(\Sigma) \times \wp(\Sigma \times \Sigma) \times \{\textit{true;false}\}.$$

$\gamma(I, F, T, b)$ is defined by $\gamma_1(I) \cap \gamma_2(F) \cap \gamma_3(T) \cap \gamma_4(b)$ where:

- $\gamma_1(I) = \{u \in \Sigma^* \mid |u| > 0 \Rightarrow u_1 \in I\}$,
- $\gamma_2(F) = \{u \in \Sigma^* \mid |u| > 0 \Rightarrow u_{|u|} \in F\}$,
- $\gamma_3(T) = \{u \in \Sigma^* \mid \forall a, b \in \Sigma^*, \lambda, \mu \in \Sigma,\ u = a.\lambda.\mu.b \Rightarrow (\lambda, \mu) \in T\}$,
- $\gamma_4(b) = \begin{cases} \Sigma^+ & \text{if } b = 0 \\ \Sigma^* & \text{otherwise.} \end{cases}$

Domain complexity is $O(n.|\Sigma|)$ and maximum iteration number is $O(n^4.|\Sigma|)$.

Comparison between channel and agent markers

We capture the difference between the occurrence number of letters in such two markers.

$$Id^2 = (\Sigma \to (\mathbb{Z} \cup \{\top\})) \cup \{\bot\}$$

γ_{Id^2} is defined as follows:

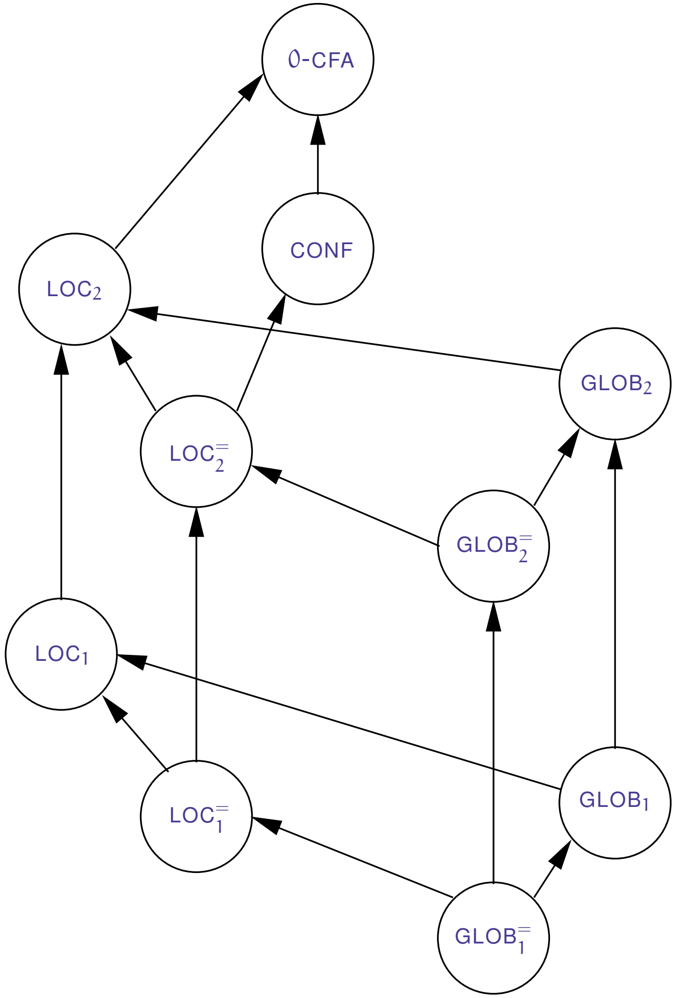
$$\gamma_{Id^2}(\bot) = \emptyset$$
$$\gamma_{Id^2}(f) = \{(u, v) \in (\Sigma^*)^2 \mid \forall \lambda,\ f(\lambda) \in \mathbb{Z} \implies |u|_\lambda - |v|_\lambda = f(n)\}.$$

Domain complexity is $O(|\Sigma|)$ and maximum iteration number is $O(n^3.|\Sigma|)$.

Several trade-offs

1. 0-cfa (0-CFA): $\mathit{Id}^\sharp = \{\bot; \top\}$,
   Cf [Nielson *et al.*:CONCUR'98], [Hennessy and Riely:HLCL'98].
2. Confinement (CONF): $\mathit{Id}^\sharp = \{\bot, =, \top\}$,
   Cf [Cardelli *et al.*:CONCUR'00].
3. Algebraic comparisons: we use the product between regular approximation and relational approximation.
   We can tune the complexity:
   - by capturing all numerical relations (GLOB_i), or only one relation per literal (LOC_i).
   - by choosing the set of literals among *Label* ($i = 2$)or Label^2 ($i = 1$).

Abstract semantics hierarchy



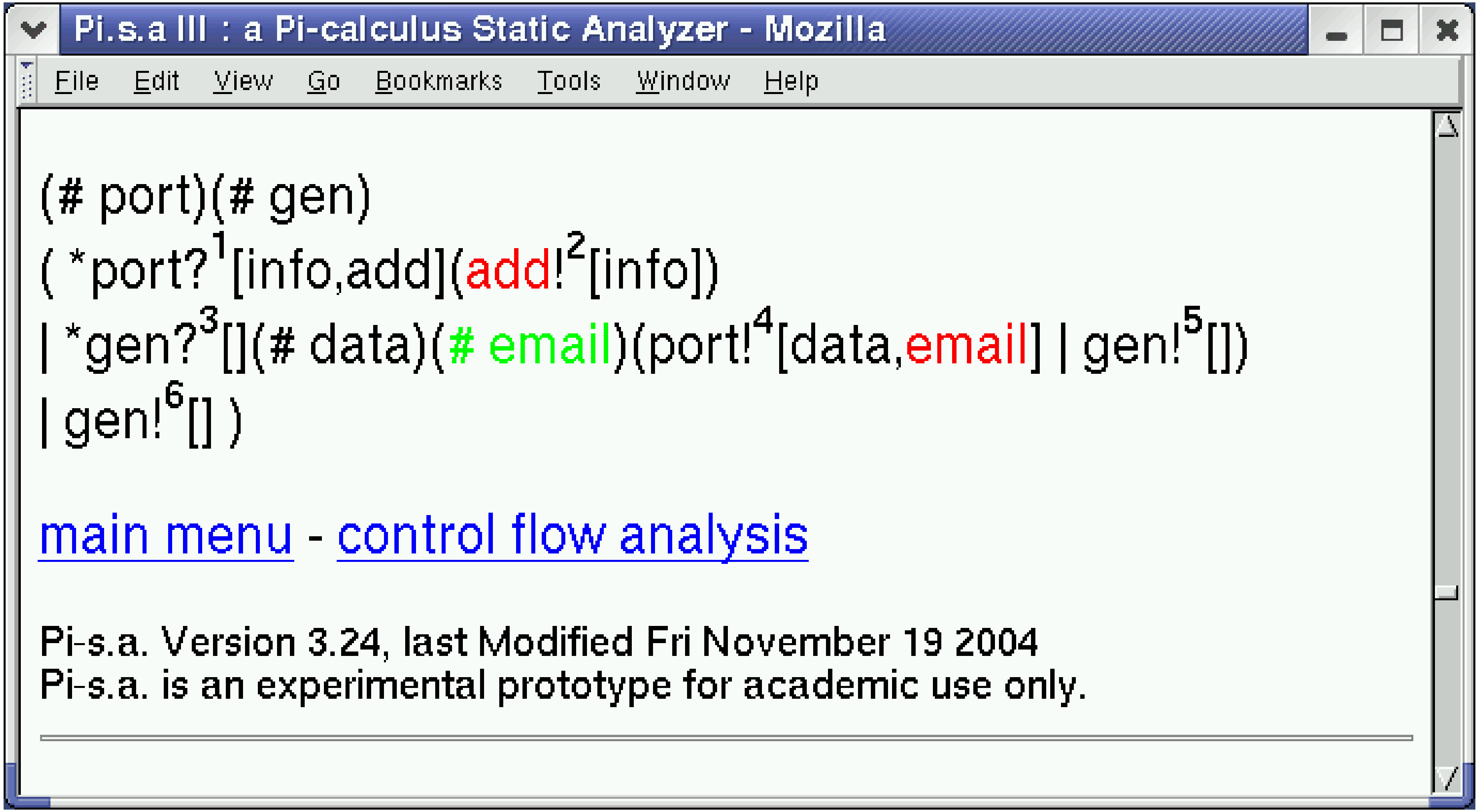
where

$$A \longrightarrow B$$

means that there exists $\alpha : A \longrightarrow B$, such that for any system $\mathcal{S}$,

$$\alpha(\llbracket \mathcal{S} \rrbracket_A^\sharp) \sqsubseteq_B \llbracket \mathcal{S} \rrbracket_B^\sharp.$$

Example: 0-CFA



Analysis result

We detect that threads at program point 2 as the following shape:

$$\left(2, (3,6)(3,5)^{n}(1,4), \begin{cases} \textit{add} & \mapsto (\textit{email}, (3,6)(3,5)^{n}) \\ \textit{info} & \mapsto (\textit{data}, (3,6)(3,5)^{n}) \end{cases}\right)$$

Example: non-uniform result

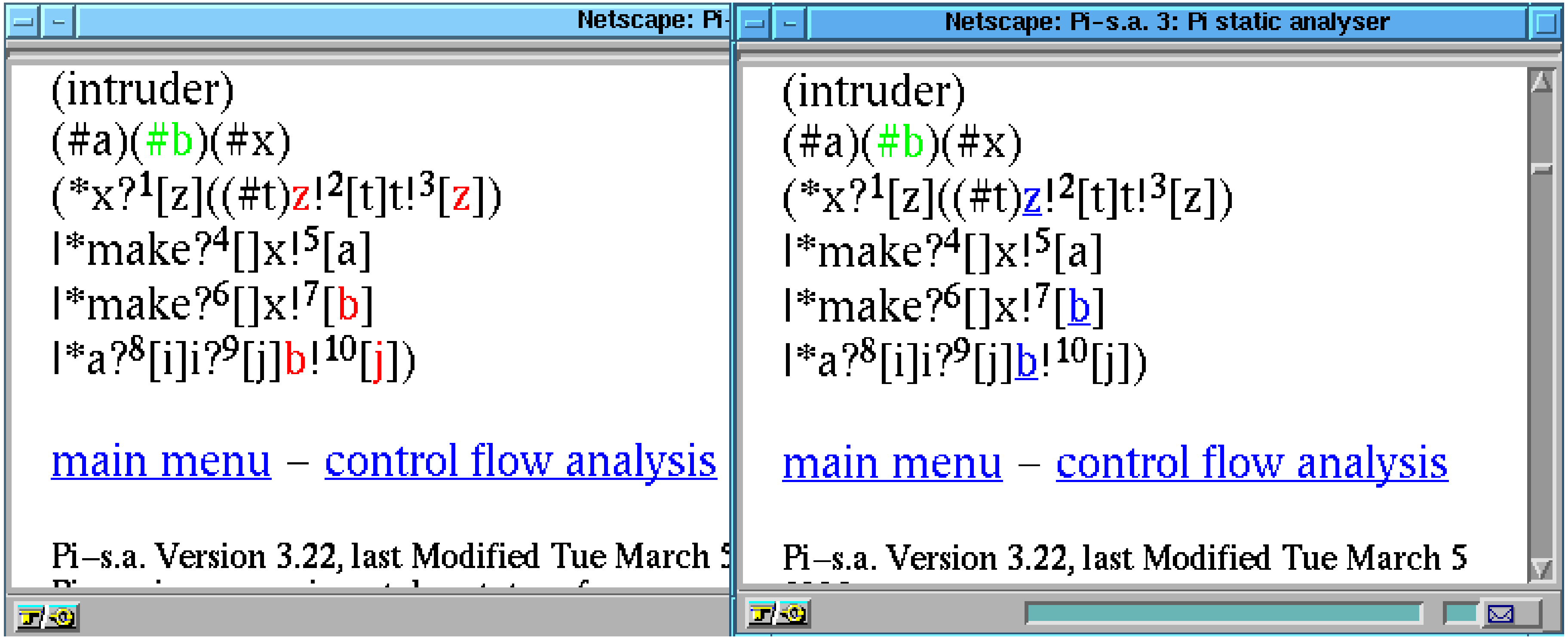
Pi.s.a III : a Pi-calculus Static Analyzer - Mozilla

(*port?^{1}[info,add](add!^{2}[info])
| *gen?^{3}[](# data)(# email)(port!^{4}[data,email] | gen!^{5}[])
| gen!^{6}[])

Start --> (3,6)A
A --> (3,5)A + (1,4)B
B --> END

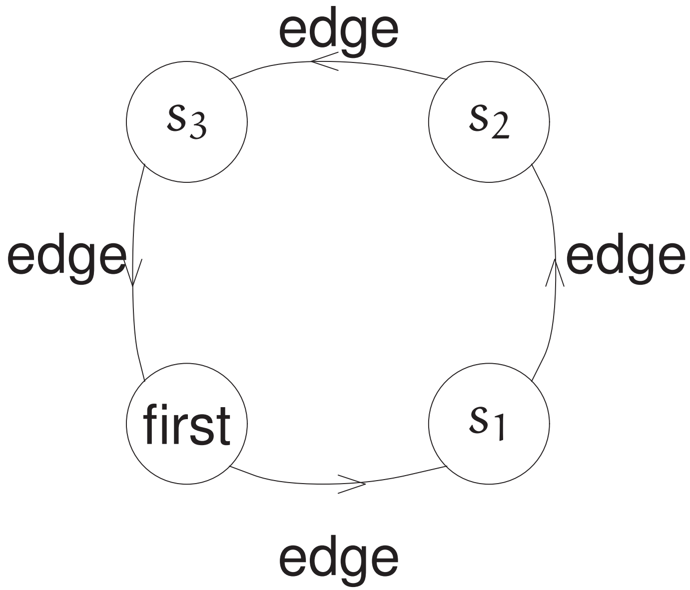
Start --> (3,6)A
A --> END + (3,5)A

(3,6) = (3,6)
(3,5) = (3,5)


Netscape: Pi-
(intruder)
(#a)(#b)(#x)
(*x?1[z]((#t)z!2[t]t!3[z])
|*make?4[]x!5[a]
|*make?6[]x!7[b]
|*a?8[i]i?9[j]b!10[j])
main menu – control flow analysis
Pi–s.a. Version 3.22, last Modified Tue March 5
Netscape: Pi–s.a. 3: Pi static analyser
(intruder)
(#a)(#b)(#x)
(*x?1[z]((#t)z!2[t]t!3[z])
|*make?4[]x!5[a]
|*make?6[]x!7[b]
|*a?8[i]i?9[j]b!10[j])
main menu – control flow analysis
Pi–s.a. Version 3.22, last Modified Tue March 5

Example: the ring of processes

$(\nu$ make$)(\nu$ edge$)(\nu$ first$)$
$(*$make$?^1[$*last*$](\nu$ *next*$)$
$(\text{edge}!^2[$*last*, *next*$]$
$|\ \text{make}!^3[$*next*$])$
$|\ *\text{make}?^4[$*last*$](\text{edge}!^5[$*last*, first$])$
$|\ \text{make}!^6[\text{first}])$



$$\underline{\sharp}(1,3)+1=\underline{\sharp}(1,3)$$

Example: Algebraic properties

Netscape: Pi-s.a. 3: Pi static analyser

((# make)(# mon)(# left0)
((*make?[1][left](# right)(mon![2][left,right]|make![3][right]))
|(*make?[4][left](mon![5][left,left0]))
|make![6][left0]))

Start --> (1,6)A
A --> (1,3)B
B --> END + (1,3)B

Start --> (1,6)A
A --> END + (1,3)A

(1,6) = (1,6)
(1,3) = (1,3) + 1

---

main menu – control flow analysis – (# right)

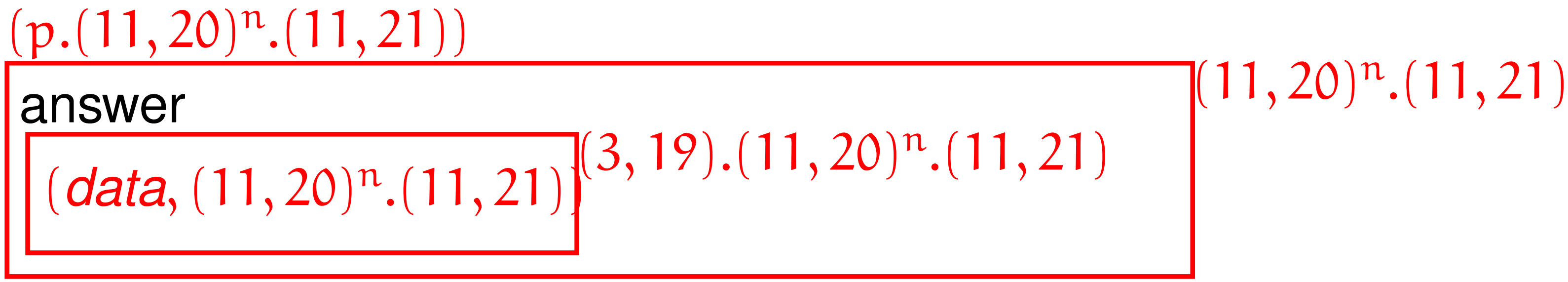
Pi–s.a. Version 3.16, last Modified Tue November 27 2001
Pi–s.a. is an experimental prototype for an academic use only.

Example

We detect that:

$$\begin{cases} (p^{12}[\bullet], (11,20)^m.(11,21), \_, [p \mapsto (p, (11,20)^m.(11,21))]) \\ (\text{answer}^8[\bullet], (3,19).(11,20)^n.(11,21), (12, (11,20)^n.(11,21), \_) \\ (\langle rep \rangle^9, \_, (8, (3,19).(11,20)^p.(11,21), [rep \mapsto (data, (11,20)^p.(11,21))])) \end{cases}$$

We deduce that each packet exiting the server has the following structure:



Limitations

Two main drawbacks:

1. we only prove equalities between Parrikh's vectors, some more work is needed in order to prove equalities of words;
2. we only capture properties involving comparison between channel name and agent markers:

$$
\begin{array}{l}
(\nu\ \text{make})(\nu\ \text{edge})(\nu\ \text{first})(\nu\ \text{first}) \\
\quad (*\text{make}?^1[\textit{last}](\nu\,\textit{next}) \\
\qquad\qquad (\text{edge}!^2[\textit{last},\textit{next}] \\
\qquad\qquad \mid\ \text{make}!^3[\textit{next}]) \\
\quad \mid\ *\text{make}?^6[\textit{last}](\text{edge}!^7[\textit{last},\text{first}]) \\
\quad \mid\ \text{make}!^8[\text{first}]) \\
\quad \mid\ \text{edge}?[x,y][x =^9 y][x \neq^{10} \text{first}]\text{Ok}!^{11}[]
\end{array}
$$

we cannot infer that 11 is unreachable.

Dependency analysis between names

We describe equality and inequality relations between the names linked to variables.

$$\mathcal{G}_V = \left\{ (A, R) \;\middle|\; \begin{array}{l} A \textit{ is a partition of } V \\ R \textit{ is a symetric anti-reflexive relation on } A \end{array} \right\}.$$

$\mathcal{G}_V$ is related to $\wp(\mathit{Id} \times (V \rightarrow (\mathit{Name} \times \mathit{Id})))$ by the following concretization function:

$$\gamma_V((A, R)) = \left\{ (\mathit{id}, E) \;\middle|\; \begin{array}{l} \forall \mathcal{X} \in A,\ \{x, y\} \subseteq \mathcal{X} \implies E(x) = E(y) \\ (\mathcal{X}, \mathcal{Y}) \in R \implies \forall x \in \mathcal{X}, y \in \mathcal{Y},\ E(x) \neq E(y) \end{array} \right\}$$

$\Longrightarrow$ implicit closure of relations and information propagation.

Dependency analysis between markers

We describe equality and inequality relations between the markers of threads and the names linked to variables.

$$\mathcal{G}_V = \left\{ (A, R) \;\middle|\; \begin{array}{l} A \textit{ is a partition of } V \uplus \{id_p\} \\ R \textit{ is a symetric anti-reflexive relation on } A \end{array} \right\}.$$

$\mathcal{G}_V$ is related to $\wp(\mathit{Id} \times (V \to (\mathit{Name} \times \mathit{Id})))$ by the following concretization function:

$$\gamma_V((A, R)) = \left\{ (id, E) \;\middle|\; \begin{array}{l} \forall \mathcal{X} \in A,\ x \in V,\ \{id_p, x\} \subseteq \mathcal{X} \implies id = \mathit{snd}(E(x)) \\ \forall \mathcal{X} \in A,\ x, y \in V,\ \{x, y\} \subseteq \mathcal{X} \implies \mathit{snd}(E(x)) = \mathit{snd}(E(y)) \\ \forall (\mathcal{X}, \mathcal{Y}) \in R,\ y \in V, \\ \quad id_p \in \mathcal{X} \textit{ and } y \in \mathcal{Y} \implies id \neq \mathit{snd}(E(y)) \\ \forall (\mathcal{X}, \mathcal{Y}) \in R,\ x, y \in V, \\ \quad x \in \mathcal{X} \textit{ and } y \in \mathcal{Y} \implies \mathit{snd}(E(x)) \neq \mathit{snd}(E(y)) \end{array} \right\}$$

$\implies$ implicit closure of relations and information propagation.

Global numerical analysis

We abstract relations between all the name markers and all the names linked to variables, and the thread markers:
For each $V \subseteq \mathit{Name}$, we introduce the set

$$\mathcal{X}_V = \{p^{\lambda} \mid \lambda \in \Sigma\} \cup \{c^{(\lambda,\nu)} \mid \lambda \in \Sigma \cup \mathit{Name},\ \nu \in V\}$$

The domain $\mathcal{G}_V$ is then the set of the affine relations system among $\mathcal{X}_V$ related to the concrete domain by the following concretization:

$$\gamma_V(\mathcal{K}) = \left\{ (\mathit{id}, E) \;\middle|\; \begin{pmatrix} p^{\lambda} \longrightarrow |\mathit{id}|_{\lambda} \\ x^{(y,\nu)} \longrightarrow (y = \mathit{first}(E(\nu))) \\ x^{(\lambda,\nu)} \longrightarrow |\mathit{snd}(E(\nu))|_{\lambda} \end{pmatrix} \ \mathit{satisfies}\ \mathcal{K} \right\}.$$

Pair-wise numerical analysis

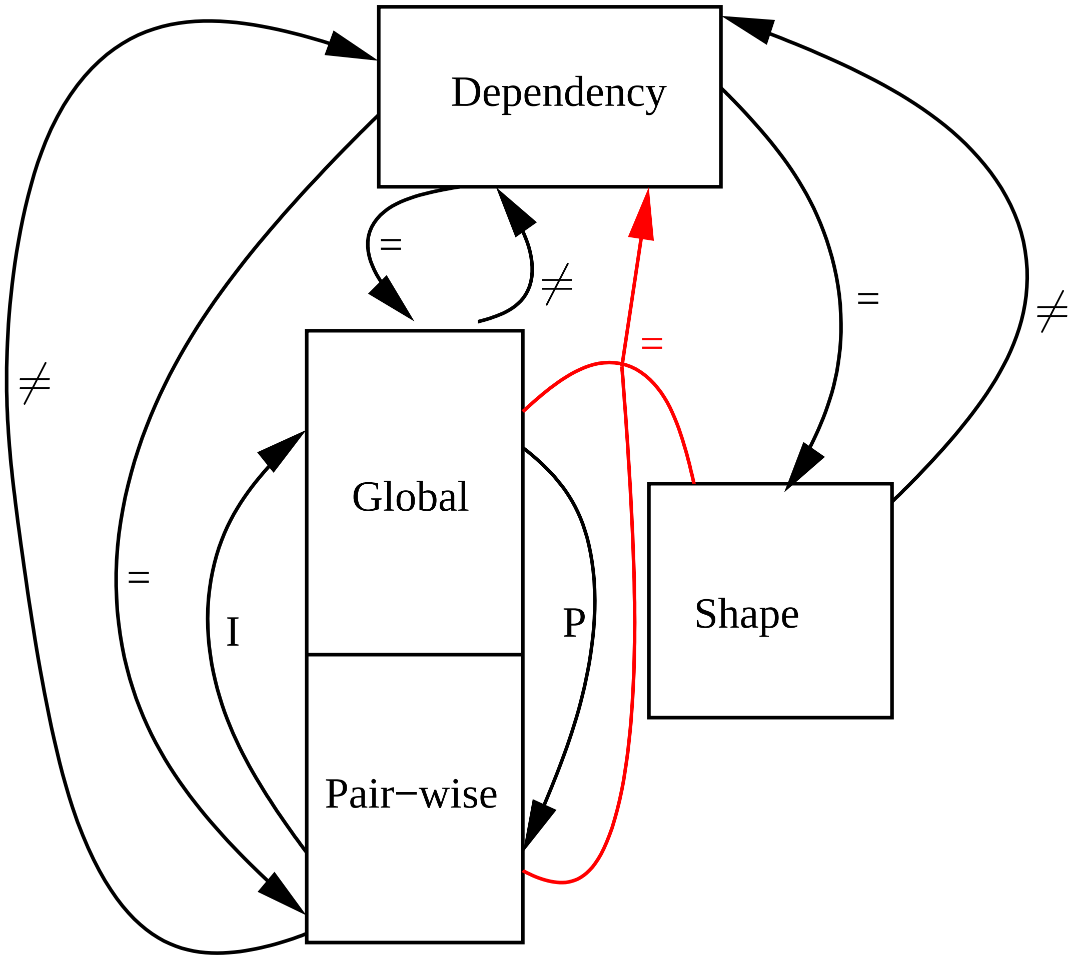
We compare pair-wisely markers, having partitioned in accordance with the name creations having created the names.
Let Φ be a linear form defined on $\mathbb{R}^{\Sigma}$, for each $V \subseteq \mathit{Name}$, the domain $\mathcal{G}_V$ is a pair of function (f, g):

$$\begin{array}{l} f \;:\; V \cup \mathit{Name} \rightarrow \{\ \mathit{Affine\ subspace\ of}\ \mathbb{R}^2\}, \\ g \;:\; (V \cup \mathit{Name})^2 \rightarrow \{\ \mathit{Affine\ subspace\ of}\ \mathbb{R}^2\}, \end{array}$$

the concretization $\gamma_V(f, g)$ is given by:

$$\left\{ (id, E) \;\middle|\; \begin{array}{l} E(x) = (y, id_y) \implies (\Phi((|id|_\lambda)_{\lambda\in\Sigma}), \Phi((|id_y|_\lambda)_{\lambda\in\Sigma})) \in f(x, y) \\ \begin{cases} E(x) = (y, id_y) \\ E(x') = (y', id'_y) \end{cases} \implies (\Phi((|id_y|_\lambda)_{\lambda\in\Sigma}), \Phi((|id'_y)|_\lambda)_{\lambda\in\Sigma}) \in g((x, y), (x', y')) \end{array} \right\}$$

Reduction



Example

$(\nu\ \text{make})(\nu\ \text{edge})(\nu\ \text{first})$
$(*\text{make}?^1[\textit{last}](\nu\textit{next})\ (\text{edge}!^2[\textit{last},\textit{next}] \mid \text{make}!^3[\textit{next}])$
$\mid *\text{make}?^6[\textit{last}](\text{edge}!^7[\textit{last},\text{first}])$
$\mid \text{make}!^8[\text{first}])$
$\mid \text{edge}?[x,y][x=^9 y][x \neq^{10} \text{first}]\text{Ok}!^{11}[]$

we first prove in global abstraction that:

$$f(2)\ \textit{satisfies}\ \begin{cases} c^{(1,3),\textit{next}} = c^{(1,3),\textit{last}} + c^{\textit{next},\textit{last}} \\ c^{\text{first},\textit{last}} + c^{\textit{next},\textit{last}} = 1 \end{cases}$$

$$f(7)\ \textit{satisfies}\ \begin{cases} c^{\textit{next},\textit{last}} + c^{\text{first},\textit{last}} = 1 \\ c^{\text{first},\text{first}} = 1 \end{cases}$$

Example

We then prove in pair-wise analysis that in process 9, x and y are respectively linked to names created by some instance of the restrictions :

1. (ν first) and (ν first),
2. (ν first) and (ν *next*),
3. (ν *next*) and (ν *next*) but distinct instances,
4. (ν *next*) and (ν first).

so, the matching pattern $[x = y]$ is satisfiable only in the first case !!!

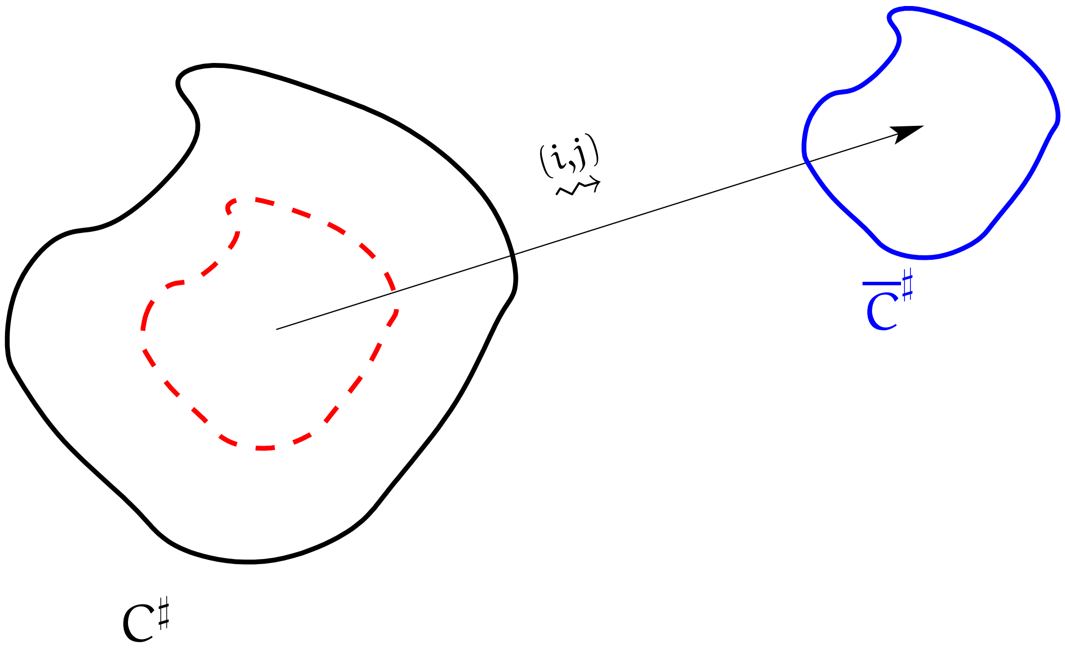
Overview

1. Overview
2. Mobile systems
3. Non standard semantics
4. Abstract Interpretation
5. Environment analyses
6. Occurrence counting analysis
7. Thread partitioning
8. Conclusion

Intuition

$$\left\{ \begin{array}{l} \left(1, \varepsilon, \left\{ \text{port} \mapsto (\text{port}, \varepsilon) \right.\right) \\ \left(3, \varepsilon, \left\{ \begin{array}{ll} \text{gen} & \mapsto (\text{gen}, \varepsilon) \\ \text{port} & \mapsto (\text{port}, \varepsilon) \end{array} \right.\right) \\ \left(2, id'_1, \left\{ \begin{array}{ll} add & \mapsto (email, id_1) \\ info & \mapsto (data, id_1) \end{array} \right.\right) \\ \left(2, id'_2, \left\{ \begin{array}{ll} add & \mapsto (email, id_2) \\ info & \mapsto (data, id_2) \end{array} \right.\right) \\ \left(5, id_2, \left\{ \text{gen} \mapsto (\text{gen}, \varepsilon) \right.\right) \end{array} \right\}$$

Abstract transition



Abstract domains

We design a domain for representing numerical constrains between

- the number of occurrences of processes $\sharp(i)$;
- the number of performed transitions $\underline{\sharp}(i,j)$.

We use the product of

- a non-relational domain:
  $\Longrightarrow$ the interval lattice;
- a relational domain:
  $\Longrightarrow$ the lattice of affine relationships.

Interval narrowing

An exact reduction is exponential.
We use:

- Gaus reduction: $\begin{cases} x+y+z=1 \\ x+y+t=2 \end{cases} \implies \begin{cases} x+y+z=1 \\ t-z=1 \end{cases}$
- Interval propagation: $\begin{cases} x+y+z=3 \\ x \in [\|0;\infty\|[\\ y \in [\|0;\infty\|[\\ z \in [\|0;\infty\|[\end{cases} \implies \begin{cases} x+y+z=3 \\ x \in [\|0;3\|] \\ y \in [\|0;\infty\|[\\ z \in [\|0;\infty\|[\end{cases}$
- Redundancy introduction: $\begin{cases} x+y-z=3 \\ x \in [\|1;2\|[\end{cases} \implies \begin{cases} x+y-z=3 \\ y-z \in [\|1;2\|] \\ x \in [\|1;2\|] \end{cases}$

to get a cubic approximated reduction.

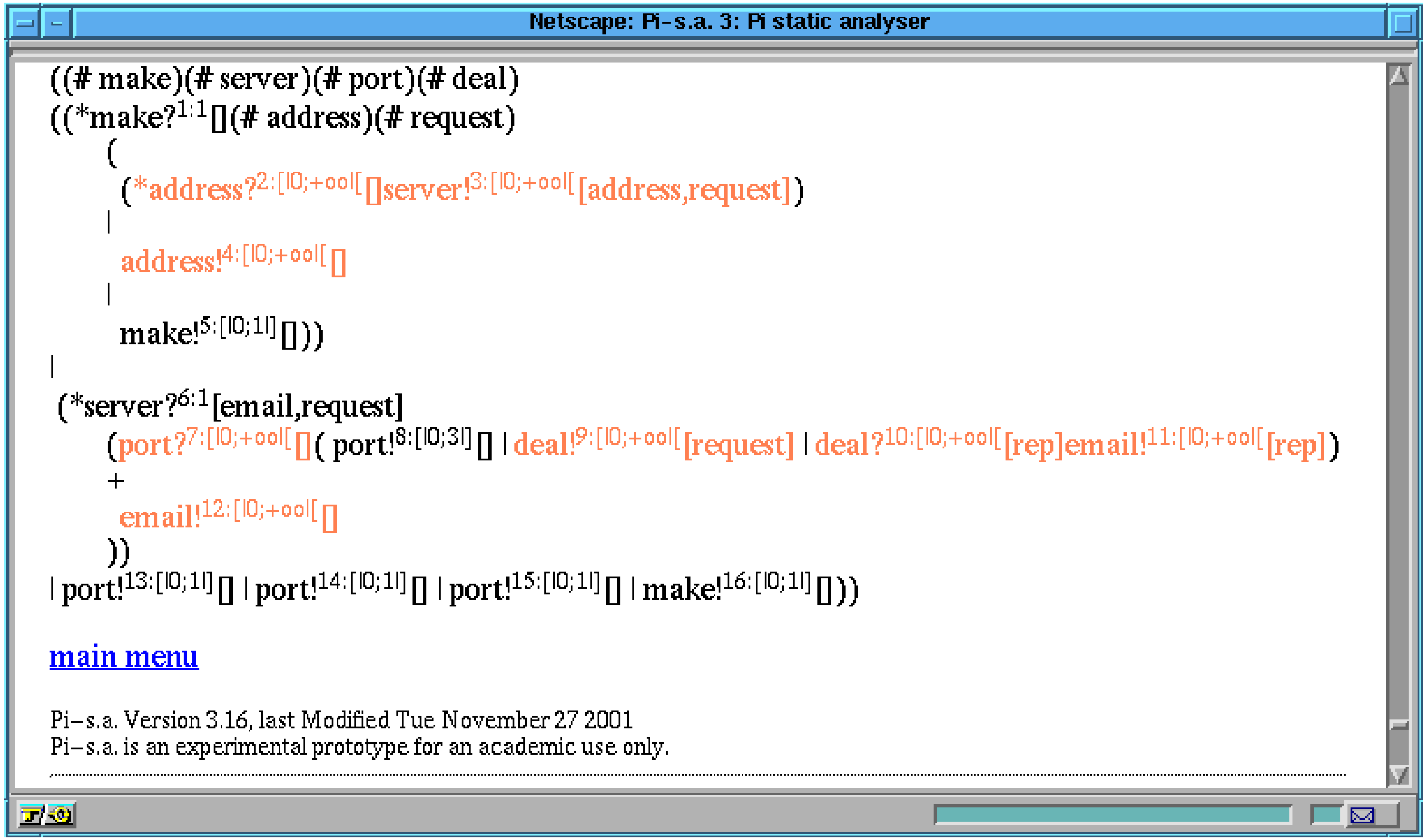
Example: non-exhaustion of resources

Netscape: Pi-s.a. 3: Pi static analyser

```
((# make)(# server)(# port)
((*make?^{1:1}[](# address)(# request)
     (
      (*address?^{2:[|0;+oo|[}[]server!^{3:[|0;+oo|[}[address,request])
      |
       address!^{4:[|0;+oo|[}[]
      |
       make!^{5:[|0;1|]}[]))
 |
  (*server?^{6:1}[email,data]
      (port?^{7:[|0;+oo|[}[](# deal)(
                 deal!^{8:[|0;3|]}[data]
                 |
                 deal?^{9:[|0;3|]}[rep](email!^{10:[|0;+oo|[}[rep] | port!^{11:[|0;3|]}[]))
      +
       email!^{12:[|0;+oo|[}[]))
| port!^{13:[|0;1|]}[] | port!^{14:[|0;1|]}[] | port!^{15:[|0;1|]}[]
| make!^{16:[|0;1|]}[]))
```

main menu

Example: exhaustion of resources

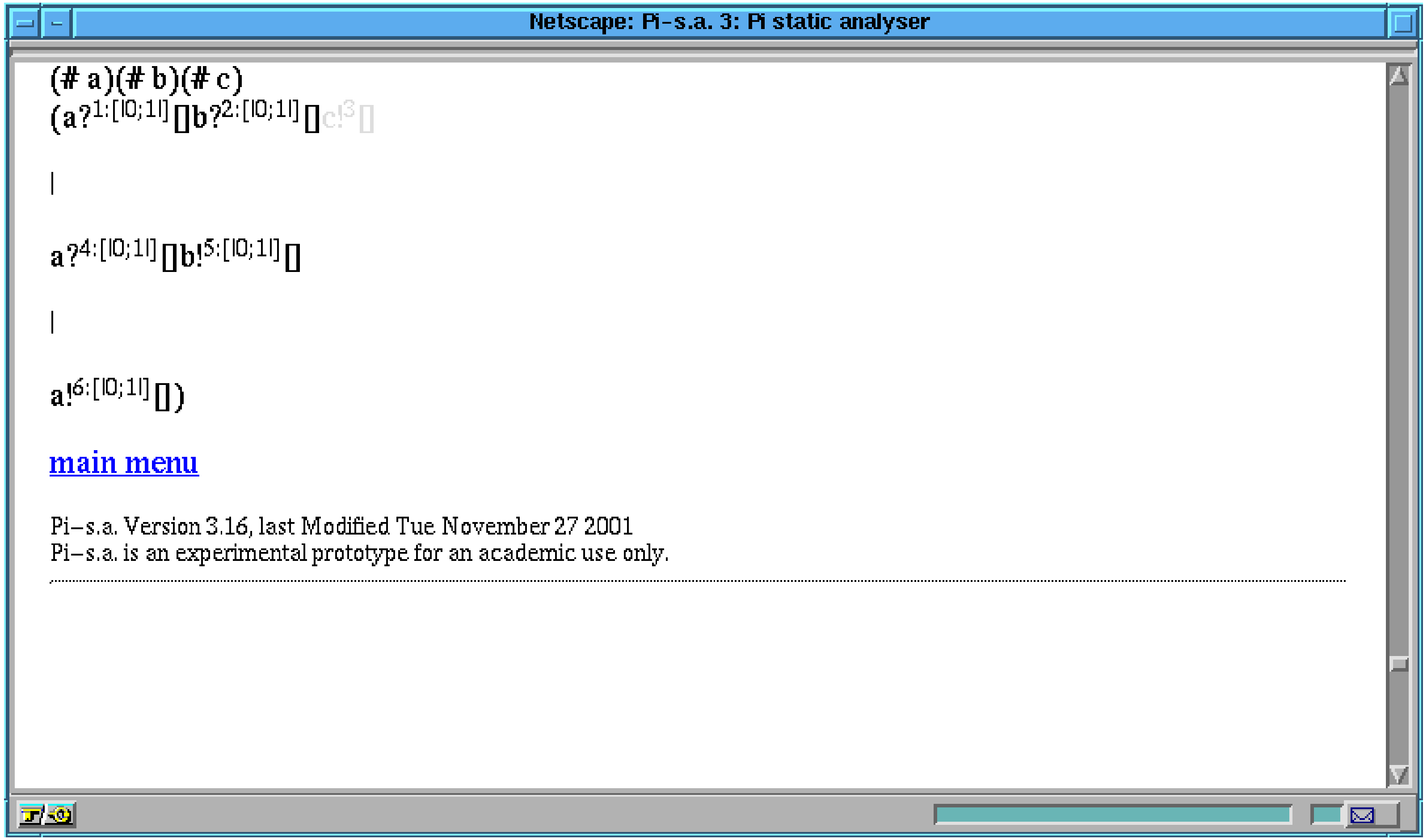


```
((# make)(# server)(# port)(# deal)
((*make?^{1:1}[](# address)(# request)
      (
       (*address?^{2:[|0;+∞|[}[]server!^{3:[|0;+∞|[}[address,request])
      |
       address!^{4:[|0;+∞|[}[]
      |
       make!^{5:[|0;1|]}[]))
|
 (*server?^{6:1}[email,request]
      (port?^{7:[|0;+∞|[}[]( port!^{8:[|0;3|]}[] | deal!^{9:[|0;+∞|[}[request] | deal?^{10:[|0;+∞|[}[rep]email!^{11:[|0;+∞|[}[rep])
      +
       email!^{12:[|0;+∞|[}[]
      ))
| port!^{13:[|0;1|]}[] | port!^{14:[|0;1|]}[] | port!^{15:[|0;1|]}[] | make!^{16:[|0;1|]}[]))
```

main menu

Pi–s.a. Version 3.16, last Modified Tue November 27 2001
Pi–s.a. is an experimental prototype for an academic use only.

Example: mutual exclusion



Example: token ring

Pi.s.a III : a Pi-calculus Static Analyzer - Mozilla

```
(# make)(# mon)(# left0)
( (*make?^1:1[left](# right)(mon!^2:[|0;+oo|[[left,right] | make!^3:[|0;1|][right]))
| (*make?^4:1[left](mon!^5:[|0;1|][left,left0]))
| make!^6:[|0;1|][left0]

| (*mon?^7:1[prev,next]
      (*prev?^8:[|0;+oo|[[](# crit)
            ( crit!^9:[|0;1|][] | (crit?^10:[|0;1|][]next!^11:[|0;1|][]))))
| left0!^12:[|0;1|][])
```

Comparison

- Non relational analyses.
  [Levi and Maffeis: SAS'2001]

- Syntactic criteria.
  [Nielson *et al.*:SAS'2004]

- Abstract multisets.
  [Nielson *et al.*:SAS'1999,POPL'2000]

- Finite control systems.
  [Dam:IC'96],[Charatonik *et al.*:ESOP'02]

Overview

1. Overview
2. Mobile systems
3. Non standard semantics
4. Abstract Interpretation
5. Environment analyses
6. Occurrence counting analysis
7. Thread partitioning
8. Conclusion

Computation unit

Gather threads inside an unbounded number of dynamically created computation units.
Then detect mutual exclusion inside each computation unit.

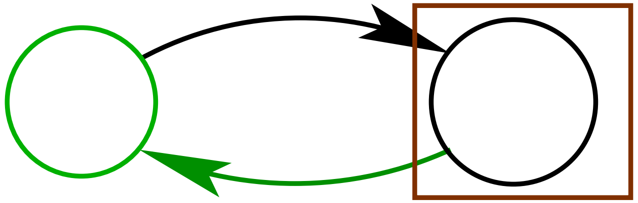
Each thread is associated with a computation unit, which is left as a parameter of:

- the model
- and the properties of interest.

For instance:

- in the π-calculus, the channel on which the input/output action is performed;
- in ambients, agent location and the location of its location [Nielson:POPL'2000].

Thread partitioning



Thread partitioning

We gather threads according to their computation unit.
We count the occurrence number of threads inside each computation unit.

To simulate a computation step, we require:

- to relate the computation units of:
  1. the threads that are consumed;
  2. the threads that are spawned.

  This may rely on the model structure (ambients) or on a precise environment analysis (other models).
- an occurrence counting analysis:
  to count occurrence of threads inside each computation unit.

Concrete partitioning

B: a finite set of indice.
We define the set of computation units as:

$$\textit{unit} \stackrel{\Delta}{=} B \rightarrow \textit{Label} \times \textit{Id}.$$

give-index maps each program point p to a function $\textit{give-index}(p) \in B \rightarrow fn(p)$.

Given a thread $t = (p, \textit{id}, E)$, we define its computation unit $\textit{give-unit}(t)$ as:

$$\textit{give-unit}(t) = [b \in B \rightarrow E(\textit{give-index}(p)(b))].$$

Abstract computation unit

There may be an unbounded number of computation units.

To get a decidable abstraction, we merge the description of the computation units that have the same labels.

We define:

$$\mathsf{UNIT}^\sharp \stackrel{\Delta}{=} \mathcal{B} \rightarrow \mathit{Label}.$$

The abstraction function:

$$\Pi_{unit} \in \begin{cases} \mathit{unit} & \longrightarrow \mathsf{UNIT}^\sharp \\ [b \in \mathcal{B} \mapsto (l_b, \_)] & \longmapsto [b \mapsto l_b]; \end{cases}$$

maps each computation unit to an abstract one.

Abstract domain

Our main domain is a Cartesian product:

$$\mathcal{C}^{\sharp}_{part} \stackrel{\Delta}{=} \left(\Pi_{p \in \mathcal{L}_p} \mathcal{G}_{fn(p)}\right) \times \left(\text{UNIT}^{\sharp} \rightarrow \mathcal{N}_{\mathcal{L}_p}\right).$$

The set $\gamma_{part}(\text{ENV}, \text{CU})$ contains any configuration $(v, C) \in \Sigma^* \times \mathcal{S}$ that satisfies:

1. $(v, C) \in \gamma_{\text{ENV}}(\text{ENV})$;
2. for any computation unit $u \in unit$, there exists a function

   $$t \in \{(0) \in \mathbb{N}^{\mathcal{L}_p}\} \cup \left(\gamma_{\mathcal{N}_{\mathcal{L}_p}}(\text{CU}(\Pi_{unit}(u)))\right)$$

   such that:

   $$t(p) = Card(\{(p, id, E) \in C \mid give\text{-}unit(p, id, E) = u\}).$$

Balance molecule

To simulate an abstract computation step,

we compute an abstract molecule that describes:

- both the n threads that are interacting;
- and the m threads that are launched;

we also collect any information about the values in computation units:

- each thread is launched in a computation unit. Each value occurring in this computation unit may either be fresh, or may come from interacting threads;

  (we take into account these constraints in the abstract molecule).

Admissible relations

Then, we consider any potential choice for:

1. the equivalence relation among the computation unit of the $(n + m)$ threads involved in the computation step;
2. abstract computation units associated to each thread.

Each choice induces some constraints about:

- the control flow;
- the number of threads inside computation units;

We use these constraints to:

1. check that this choice is possible;
2. refine control flow and occurrence counting information;

Then, we simulate the computation step.

Shared-memory example

A memory cell will be denoted by three channel names, *cell*, *read*, *write*:

- the channel name *cell* describes the content of the cell:

  the process *cell*![*data*] means that the cell *cell* contains the information *data*, this name is internal to the memory (not visible by the user).

- the channel name *read* allows reading requests:

  the process *read*![*port*] is a request to read the content of the cell, and send it to the port *port*,

- the channel name *write* allows writing requests:

  the process *write*![*data*] is a request to write the information *data* inside the cell.

Implementation

$\text{System} := (\nu\ \text{create})(\nu\ \text{null})(*\text{create}?[\text{d}].\text{Allocate}(d))$

$\text{Allocate}(\text{d}) :=$

$(\nu\ cell)(\nu\ write)(\nu\ read)$

$\text{init}(cell) \mid \text{read}(read,cell) \mid \text{write}(write,cell) \mid \text{d}![read;write]$

where

- $\text{init}(cell) := cell![\text{null}]$
- $\text{read}(read,cell) := *read?[port].cell?[u](cell![u] \mid port![u])$
- $\text{write}(write,cell) := *write?[data,ack].cell?[u].(cell![data] \mid ack![])$

Absence of race conditions

The computation unit of a thread is the name of the channel on which it performs its i/o action.

We detect that there is never two simultaneous outputs on a channel opened by an instance of a (ν *cell*) restriction.

Other Applications

By choosing appropriate settings for the computation unit, it can be used to infer the following causality properties:

- authentication in cryptographic protocols;
- absence of race conditions in dynamically allocated memories;
- update integrity in reconfigurable systems.

Overview

1. Overview
2. Mobile systems
3. Non standard semantics
4. Abstract Interpretation
5. Environment analyses
6. Occurrence counting analysis
7. Thread partitioning
8. Conclusion

Conclusion

We have designed generic analyses:

- automatic, sound, terminating, approximate,
- model independent (META-language),
- context independent.

We have captured:

- dynamic topology properties:
  absence of communication leak between recursive agents,
- concurrency properties:
  mutual exclusion, non-exhaustion of resources,
- combined properties:
  absence of race conditions, authentication (non-injective agreement).

Future Work I
Enriching the META-language

- term defined up to an equational theory (applied pi),
  $\Longrightarrow$ analyzing cryptographic protocols with XOR;
- higher order communication;
  $\Longrightarrow$ agents may communicate running programs;
  $\Longrightarrow$ agents may duplicate running programs;
- Using our framework to describe and analyze mobility in industrial applications (ERLANG).

Future works II
High level properties

Fill the gap between:

- low level properties captured by our analyses;
- high level properties specified by end-users.

Our goal:

- check some formula in a logic [Caires and Cardelli:IC'2003/TCS'2004]
- still distinguishing recursive instances
  $\neq$ [Kobayashi:POPL'2001]

Future works III
Analyzing probabilistic semantics

In a biological system, a cell may die or duplicate itself. The choice between these two opposite behaviors is controlled by the concentration of components in the system.

$\Longrightarrow$ a reachability analysis is useless.

- Using a semantics where the transitions are chosen according to probabilistic distributions:

  $\Longrightarrow$ (e.g token-based abstract machines [Palamidessi:FOSSACS'00])

- Existing analyses consider finite control systems [Logozzo:SAVE'2001,Degano *et al.*:TSE'2001]
- We want to design an analysis for capturing the probabilistic behavior of unbounded systems.