

Weak memory models using event structures

Simon Castellan, 17 mars 2016
Journée langages

Unexpected behaviours

A simple concurrent and imperative program:

x, y initialized to 0
 $x := 1 \parallel y := 2$
 $r \leftarrow y \parallel s \leftarrow x$
shared variable · local register

Expected outcome: $r \neq 0 \vee s \neq 0$.

Unexpected behaviours

A simple concurrent and imperative program:

x, y initialized to 0
 $r \leftarrow y \parallel s \leftarrow x$
 $x := 1 \parallel y := 2$
shared variable · local register

Expected outcome: $r \neq 0 \vee s \neq 0$.

Wrong on modern architectures (x86, ARM, ...).

Unexpected behaviours

Another simple program:

$$\begin{array}{l|l} x := 1 & y := 1 \\ r_1 \leftarrow x & s_1 \leftarrow y \\ r_2 \leftarrow y & s_2 \leftarrow x \end{array}$$

Expected outcome: $r_1 = s_1 = 1 \Rightarrow r_2 = s_2 = 1$

Unexpected behaviours

Another simple program:

$$\begin{array}{l|l} x := 1 & y := 1 \\ r_1 \leftarrow x & s_1 \leftarrow y \\ r_2 \leftarrow y & s_2 \leftarrow x \end{array}$$

Expected outcome: $r_1 = s_1 = 1 \Rightarrow r_2 = s_2 = 1$

Wrong even without read exchange (x86: *Read Own Write Early*).

A need to specify the behaviour

What are the expected behaviour of a concurrent programs?

→ It depends on the architectures.

Architectures need to be specified:

- ▶ what instructions can be reordered?
- ▶ how are writes propagated from one thread to the other?

A need to specify the behaviour

What are the expected behaviour of a concurrent programs?

→ It depends on the architectures.

Architectures need to be specified:

- ▶ what instructions can be reordered?
- ▶ how are writes propagated from one thread to the other?

To that end, manufacturers provide prosaic documents, but:

- ▶ *ambiguity*: behaviours that are not specified
- ▶ *inconsistent*: some observations may not be predicted.

Some architectures:

- ▶ SC (*Sequential consistency*): no reordering, sequential memory,
- ▶ ARM: reordering of instructions targeting different variables, write caches.
- ▶ x86: ...

Semantics saves the day

Semantics: Formalize mathematically the vendors specifications:

- ▶ get a (possibly computer-verified) proof of non-ambiguity,
- ▶ implement the specifications and mechanically **test it** against real life architectures.

Two main types of semantics among existing models:

- ▶ *operational semantics*: executions are described by the runs of an abstract machines,
- ▶ *axiomatic semantics*: the notion of valid execution is axiomatized.

Those models are called *weak memory models*.

Semantics and executions

The semantics generates from a program its possible *executions*:

Program	Some executions
$x := 1 \parallel y := 2$	$W_x^{(1)} \cdot W_y^{(2)} \cdot R_y^{(2)} \cdot R_x^{(1)}$
$r \leftarrow y \parallel s \leftarrow x$	$W_y^{(2)} \cdot R_x^{(0)} \cdot W_x^{(2)} \cdot R_y^{(1)}$

Executions can be formalized in different ways: traces, partial-order, ...

This talk

A semantics that is

- ▶ **denotational**: executions computed by induction
 - ▶ the semantics is thus *compositional*
- ▶ **compact**: based on event structures
 - ▶ no combinatorial explosion
- ▶ **extensible**: inspired from game semantics
 - ▶ it is easy to add loops, control operators, higher-order, ...

Outline of the talk:

1. **A semantics warm-up**: compute the SC semantics using *traces*.
2. Getting back the **causality**.
3. Our contribution: A **parametric** semantics using event structures.

I. A DENOTATIONAL SEMANTICS FOR SC

With traces of originality

Syntax precedes semantics

Our very simple programming language:

$$\begin{aligned} e, e' &::= \{ \textit{Expressions} \} \\ &\quad k \in \mathbb{N} \mid r \in \mathcal{R} \mid e + e' \\ \iota &::= \{ \textit{Instructions} \} \\ &\quad \mid a := e \quad \text{(Write on a variable)} \\ &\quad \mid r \leftarrow a \quad \text{(Read on a variable)} \\ t &::= \{ \textit{Threads} \} \\ &\quad \mid \iota; \dots; \iota \\ p &::= \{ \textit{Programs} \} \\ &\quad t_1 \parallel \dots \parallel t_n \end{aligned}$$

In real life: conditionals and barriers.

Denotational semantics

Goal: compute $\llbracket t \rrbracket \in E$ where E is some space of denotations.

Our space here: languages of traces.

$$\Sigma_a = \mathcal{V} \times \{R, W\} \quad (\text{Abstract memory event})$$

$$\Sigma_c = \Sigma_a \times \mathbb{N} \quad (\text{Concrete memory event})$$

$$E = \mathcal{P}(\Sigma_c^*)$$

Notations: $R_x^{(k)}$, $W_x^{(k)}$.

Two steps:

1. **Volatile semantics** $\llbracket t \rrbracket^O$: shared variables are considered *volatile*: $\llbracket x := 1; r \leftarrow x \rrbracket^O$ does not guarantee to read 1 in r .
2. **Closed semantics**: once $\llbracket t \rrbracket^O$ is calculated for the whole program, we restrict the scope of the variable $\llbracket x := 1; r \leftarrow x \rrbracket$ reads 1 in r .

Volatile semantics

Semantics of threads. Parametrized over $\rho : \mathcal{R} \rightarrow \mathbb{N}$.

$$\text{(Writes)} \quad \llbracket x := e; t \rrbracket \rho = W_x^{(\rho(e))} \cdot \llbracket t \rrbracket \rho$$

$$\text{(Reads)} \quad \llbracket r \leftarrow x; t \rrbracket \rho = \bigcup_{i \in \mathbb{N}} \left(R_x^{(i)} \cdot \llbracket t \rrbracket (\rho[r \leftarrow i]) \right)$$

Volatile semantics

Semantics of threads. Parametrized over $\rho : \mathcal{R} \rightarrow \mathbb{N}$.

$$\text{(Writes)} \quad \llbracket x := e; t \rrbracket \rho = W_x^{(\rho(e))} \cdot \llbracket t \rrbracket \rho$$

$$\text{(Reads)} \quad \llbracket r \leftarrow x; t \rrbracket \rho = \bigcup_{i \in \mathbb{N}} \left(R_x^{(i)} \cdot \llbracket t \rrbracket (\rho[r \leftarrow i]) \right)$$

Semantics of programs. Obtained by interleaving (\circledast):

$$\llbracket t_1 \parallel \dots \parallel t_n \rrbracket = \llbracket t_1 \rrbracket \emptyset \circledast \dots \circledast \llbracket t_n \rrbracket \emptyset$$

Volatile semantics

Semantics of threads. Parametrized over $\rho : \mathcal{R} \rightarrow \mathbb{N}$.

$$\text{(Writes)} \quad \llbracket x := e; t \rrbracket \rho = W_x^{(\rho(e))} \cdot \llbracket t \rrbracket \rho$$

$$\text{(Reads)} \quad \llbracket r \leftarrow x; t \rrbracket \rho = \bigcup_{i \in \mathbb{N}} \left(R_x^{(i)} \cdot \llbracket t \rrbracket (\rho[r \leftarrow i]) \right)$$

Semantics of programs. Obtained by interleaving (\circledast):

$$\llbracket t_1 \parallel \dots \parallel t_n \rrbracket = \llbracket t_1 \rrbracket \emptyset \circledast \dots \circledast \llbracket t_n \rrbracket \emptyset$$

Example. Define $\rho = (x := 1; y \leftarrow r \parallel y := 1; x \leftarrow s)$

- ▶ $W_x^{(1)} \cdot W_y^{(1)} \cdot R_y^{(3)} \cdot R_x^{(2)} \in \llbracket \rho \rrbracket$
- ▶ but $R_x^{(0)} \cdot R_y^{(0)} \cdot W_x^{(1)} \cdot W_y^{(1)} \notin \llbracket \rho \rrbracket$.

Closed semantics

Obtained by eliminating “inconsistent” traces (eg. $W_x^{(2)} \cdot R_x^{(3)}$)

Linear memory model. A language of “consistent” traces:

$$\begin{aligned} M(\mu : \mathcal{V} \rightarrow \mathbb{N}) ::= & \epsilon \\ & | R_x^{(\mu(x))} \cdot M(\mu) \\ & | W_x^{(k)} \cdot M(\mu[x \leftarrow k]) \\ M ::= & M(x \mapsto 0) \end{aligned}$$

Closed semantics: $\llbracket p \rrbracket = \llbracket p \rrbracket^O \cap M$.

Example. Write $p = (x := 1; r \leftarrow y) \parallel (y := 2; s \leftarrow x)$

- ▶ every trace of $\llbracket p \rrbracket$ ends with $R_x^{(1)}$ or a $R_y^{(2)}$.

Summary

Advantages.

- ▶ Easy to define semantics, by induction on programs.
- ▶ By making M more complex, complex cache schemes can be handled

Drawbacks.

- ▶ Combinatorial explosion due to interleavings.
- ▶ How to model reordering of instructions?

Towards partial-orders.

- ▶ Because of reorderings, threads are not totally ordered
- ▶ Our goal: compute fine precisely dependencies between the instructions, given an architecture.

II. EVENT STRUCTURES

Raiders of the lost causality

Replacing traces by partial-orders

Idea: volatile semantics should be a set of partial-orders.

Term:

$x := 1; y := 1;$

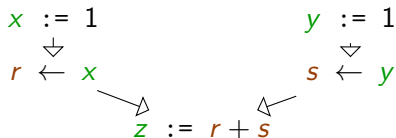
$r \leftarrow x; s \leftarrow y;$

$z := s + t$

Replacing traces by partial-orders

Idea: volatile semantics should be a set of partial-orders.

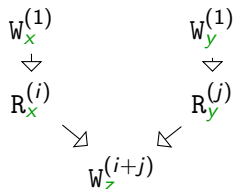
Dependencies (depends on the architecture):



Replacing traces by partial-orders

Idea: volatile semantics should be a set of partial-orders.

Executions (depends on the architecture):



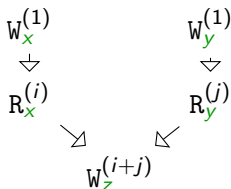
for $i, j \in \mathbb{N}^2$.

- ▶ traces on Σ_c becomes *partially ordered multisets* over Σ_c (pomsets)
- ▶ $\llbracket t \rrbracket^O$ becomes a set of such *pomsets*.

Replacing traces by partial-orders

Idea: volatile semantics should be a set of partial-orders.

Executions (depends on the architecture):

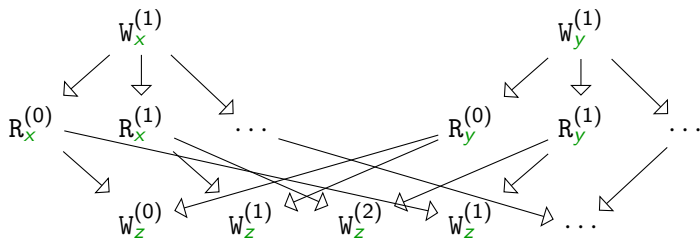


for $i, j \in \mathbb{N}^2$.

- ▶ traces on Σ_c becomes *partially ordered multisets* over Σ_c (pomsets)
- ▶ $\llbracket t \rrbracket^O$ becomes a set of such *pomsets*.
- ▶ **Problem:** lots of redundancies in the pomsets..

Can we sum up *all* executions in a single object?

Can we glue the executions all together in a partial-order? For instance:

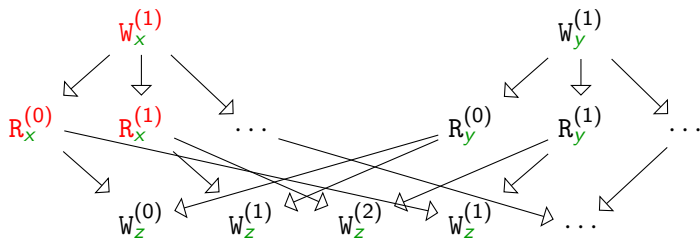


Which sets of events w are (partial) executions?

- ▶ w must be downward-closed for \rightarrow

Can we sum up *all* executions in a single object?

Can we glue the executions all together in a partial-order? For instance:

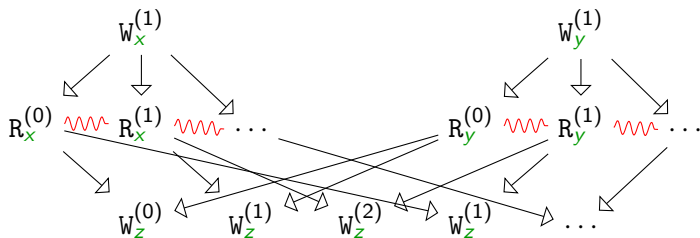


Which sets of events w are (partial) executions?

- ▶ w must be downward-closed for \rightarrow
- ▶ and ...? $\{W_x^{(1)}, R_x^{(0)}, R_x^{(1)}\}$ cannot be a valid execution.

Can we sum up *all* executions in a single object?

Can we glue the executions all together in a partial-order? For instance:



Which sets of events w are (partial) executions?

- ▶ w must be downward-closed for \rightarrow
- ▶ and ...? $\{W_x^{(1)}, R_x^{(0)}, R_x^{(1)}\}$ cannot be a valid execution.

\Rightarrow Need more structure than a partial-order: **conflicts**.

Event structures save the day

Definition (Event structures)

A set of event E with:

- ▶ A notion of **causality** represented by a *partial order* \leq_E
- ▶ A notion of **conflict** represented by a *relation* \sim_E
- ▶ A labelling $l : E \rightarrow \Sigma$.

(+ axioms)

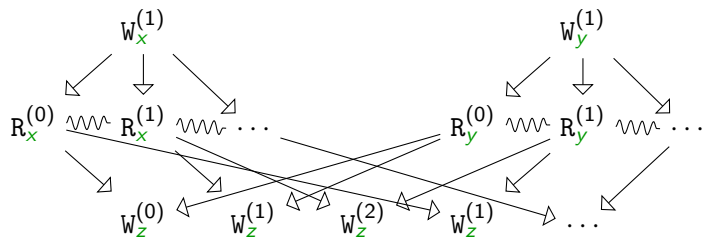
Definition (Configuration or partial execution)

A **configuration** of E is a subset w of E :

- ▶ downward-closed: $e \leq e' \in w \Rightarrow e \in w$.
- ▶ that does not contain two conflicting events

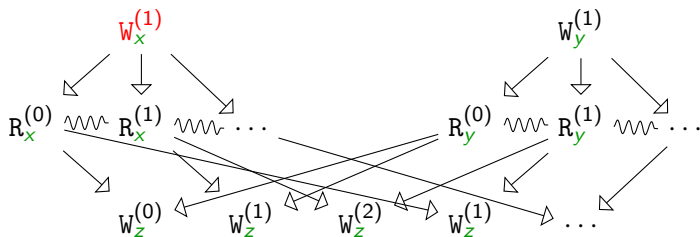
Event structures save the day

On the example:



Event structures save the day

On the example:

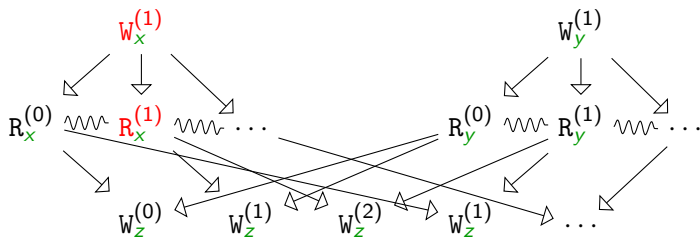


We have the configuration:

$W_x^{(1)}$

Event structures save the day

On the example:

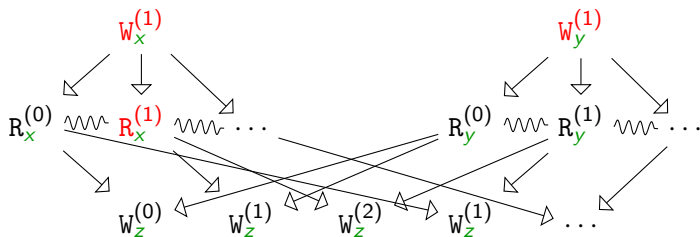


We have the configuration:

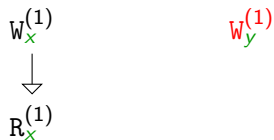


Event structures save the day

On the example:

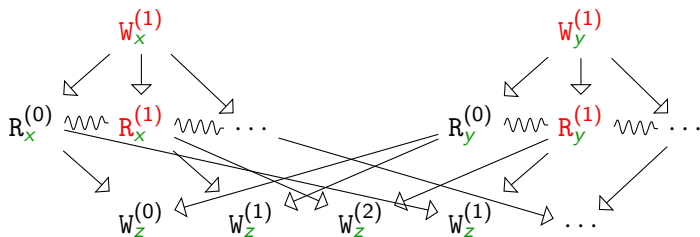


We have the configuration:

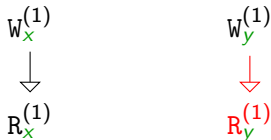


Event structures save the day

On the example:

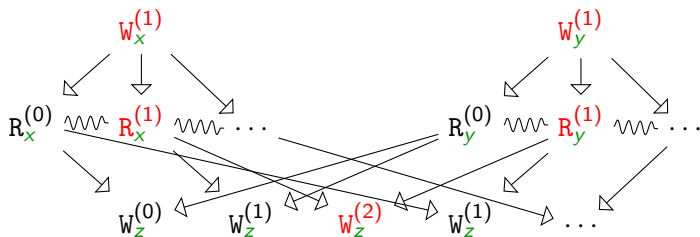


We have the configuration:

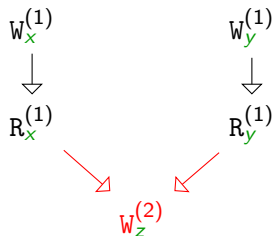


Event structures save the day

On the example:



We have the configuration:



III. DESIGNING A SEMANTICS WITH EVENT STRUCTURES

Dessine-moi une structure d'événements

Defining architectures

Now we define an architecture \mathcal{A} as a pair $(\rightarrow_{\mathcal{A}}, E)$:

- ▶ $\rightarrow_{\mathcal{A}} \subseteq \Sigma_a \times \Sigma_a$ indicates which causality cannot be erased.
- ▶ $E_{\mathcal{A}}$ is an event structure representing the memory model.

Examples for $\rightarrow_{\mathcal{A}}$:

- ▶ $\rightarrow_{\text{SC}} = \Sigma_a \times \Sigma_a$
- ▶ $\rightarrow_{\text{ARM}} = \{(e, e') \mid v(e) = v(e')\}$ ($v(x, _) = x$).
- ▶ $\rightarrow_{\text{x86}} = \dots$

Examples for $E_{\mathcal{A}}$ include all languages $M \subseteq \Sigma_c^*$ (they can be viewed as event structures).

Computing the semantics $\llbracket p \rrbracket_{\mathcal{A}}$

As previously, in two steps:

▶ **Volatile semantics:**

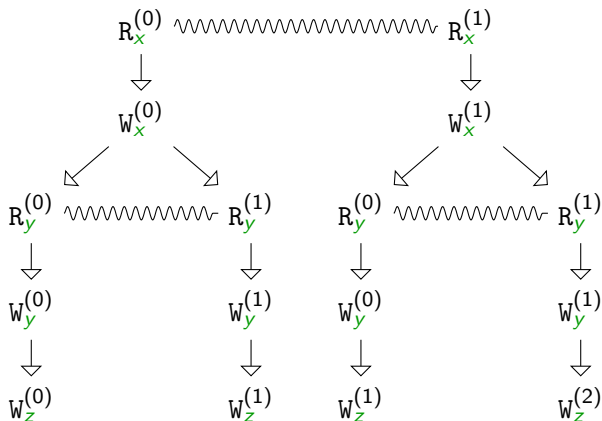
▶ *threads*: $\llbracket t \rrbracket_{\mathcal{A}}^O$ is defined as previously but where the causality outside $\rightarrow_{\mathcal{A}}$ are relaxed.

▶ *programs*: $\llbracket t_1 \parallel \dots \parallel t_n \rrbracket_{\mathcal{A}}^O = \llbracket t_1 \rrbracket_{\mathcal{A}}^O \parallel \dots \parallel \llbracket t_n \rrbracket_{\mathcal{A}}^O$
where \parallel is *parallel composition*.

▶ **Closed semantics**: $\llbracket p \rrbracket_{\mathcal{A}} = \llbracket p \rrbracket_{\mathcal{A}}^O \wedge E_{\mathcal{A}}$
where \wedge is the *synchronized product*: a generalization of intersection of languages to event structures.

Volatile semantics

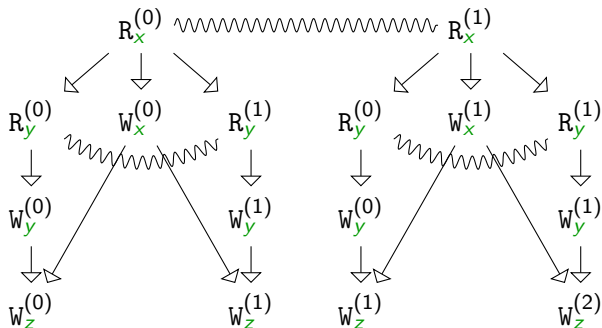
Pour $t = \begin{pmatrix} s \leftarrow x; x := s; \\ t \leftarrow y; y := t; \\ z := s + t \end{pmatrix}$, on a:



(SC)

Volatile semantics

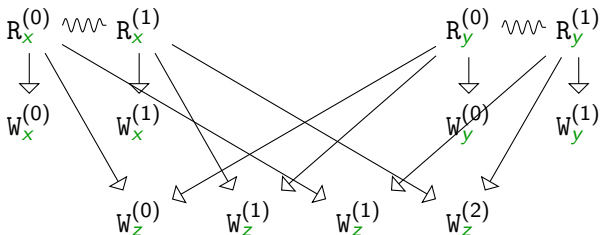
Pour $t = \begin{pmatrix} s \leftarrow x; x := s; \\ t \leftarrow y; y := t; \\ z := s + t \end{pmatrix}$, on a:



(x86)

Volatile semantics

Pour $t = \begin{pmatrix} s \leftarrow x; x := s; \\ t \leftarrow y; y := t; \\ z := s + t \end{pmatrix}$, on a:



(ARM)

Modelling caches

Goal: Model write caches.

Define an event structure \mathcal{E} whose configurations are all partial-orders (q, \leq_q) such that:

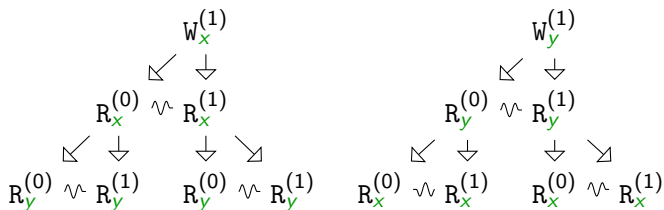
1. For all x , q restricted to writes on x is a total order.
2. For all $(R_x^{(k)}, _) \in q$, the set

$$\{W_x^{(l)} \mid W_x^{(l)} \in q\}$$

has a maximum for \leq_q and its value has the shape $W_x^{(k)}$.

Example

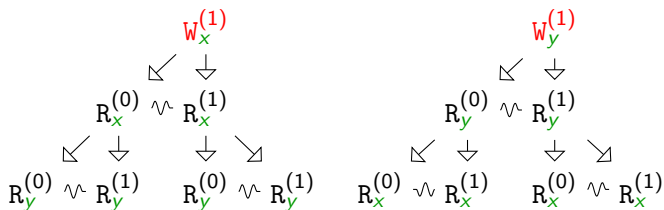
$$p = \begin{array}{l} x := 1 \\ r_1 \leftarrow x \\ r_2 \leftarrow y \end{array} \parallel \begin{array}{l} y := 1 \\ s_1 \leftarrow y \\ s_2 \leftarrow x \end{array}$$



(Volatile semantics for SC)

Example

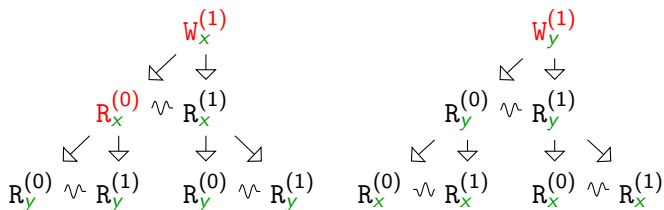
$$p = \begin{array}{l} x := 1 \\ r_1 \leftarrow x \\ r_2 \leftarrow y \end{array} \parallel \begin{array}{l} y := 1 \\ s_1 \leftarrow y \\ s_2 \leftarrow x \end{array}$$



(Computing $\llbracket p \rrbracket_{SC}^O \wedge \mathcal{E}$)

Example

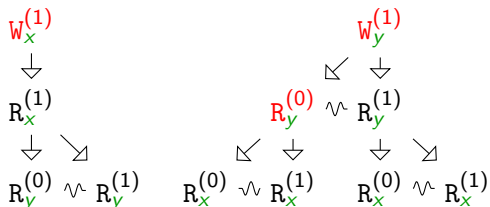
$$p = \begin{array}{l} x := 1 \\ r_1 \leftarrow x \\ r_2 \leftarrow y \end{array} \parallel \begin{array}{l} y := 1 \\ s_1 \leftarrow y \\ s_2 \leftarrow x \end{array}$$



(Computing $\llbracket p \rrbracket_{SC}^O \wedge \mathcal{E}$)

Example

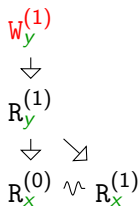
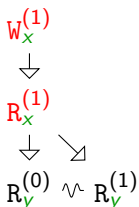
$$p = \begin{array}{l} x := 1 \\ r_1 \leftarrow x \\ r_2 \leftarrow y \end{array} \parallel \begin{array}{l} y := 1 \\ s_1 \leftarrow y \\ s_2 \leftarrow x \end{array}$$



(Computing $\llbracket p \rrbracket_{SC}^O \wedge \mathcal{E}$)

Example

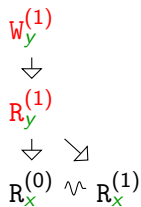
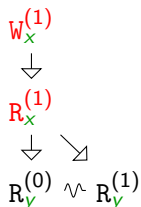
$$p = \begin{array}{l} x := 1 \\ r_1 \leftarrow x \\ r_2 \leftarrow y \end{array} \parallel \begin{array}{l} y := 1 \\ s_1 \leftarrow y \\ s_2 \leftarrow x \end{array}$$



(Computing $\llbracket p \rrbracket_{SC}^O \wedge \mathcal{E}$)

Example

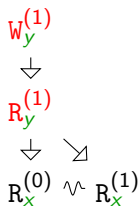
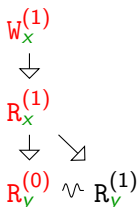
$$p = \begin{array}{l} x := 1 \\ r_1 \leftarrow x \\ r_2 \leftarrow y \end{array} \parallel \begin{array}{l} y := 1 \\ s_1 \leftarrow y \\ s_2 \leftarrow x \end{array}$$



(Computing $\llbracket p \rrbracket_{SC}^O \wedge \mathcal{E}$)

Example

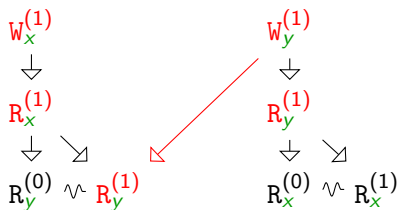
$$p = \begin{array}{l} x := 1 \\ r_1 \leftarrow x \\ r_2 \leftarrow y \end{array} \parallel \begin{array}{l} y := 1 \\ s_1 \leftarrow y \\ s_2 \leftarrow x \end{array}$$



(Computing $\llbracket p \rrbracket_{SC}^O \wedge \mathcal{E}$)

Example

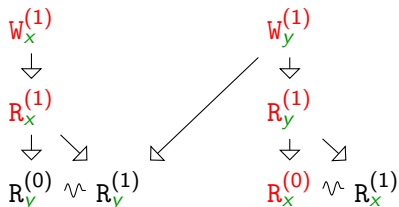
$$p = \begin{array}{l} x := 1 \\ r_1 \leftarrow x \\ r_2 \leftarrow y \end{array} \parallel \begin{array}{l} y := 1 \\ s_1 \leftarrow y \\ s_2 \leftarrow x \end{array}$$



(Computing $\llbracket p \rrbracket_{SC}^O \wedge \mathcal{E}$)

Example

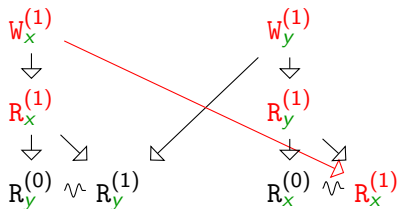
$$p = \begin{array}{l} x := 1 \\ r_1 \leftarrow x \\ r_2 \leftarrow y \end{array} \parallel \begin{array}{l} y := 1 \\ s_1 \leftarrow y \\ s_2 \leftarrow x \end{array}$$



(Computing $\llbracket p \rrbracket_{SC}^O \wedge \mathcal{E}$)

Example

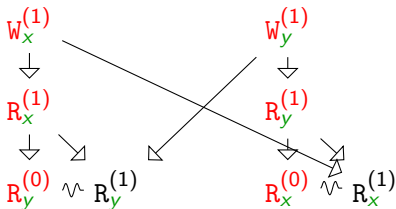
$$p = \begin{array}{l} x := 1 \\ r_1 \leftarrow x \\ r_2 \leftarrow y \end{array} \parallel \begin{array}{l} y := 1 \\ s_1 \leftarrow y \\ s_2 \leftarrow x \end{array}$$



(Computing $\llbracket p \rrbracket_{SC}^O \wedge \mathcal{E}$)

Example

$$p = \begin{array}{l} x := 1 \\ r_1 \leftarrow x \\ r_2 \leftarrow y \end{array} \parallel \begin{array}{l} y := 1 \\ s_1 \leftarrow y \\ s_2 \leftarrow x \end{array}$$



(Computing $\llbracket p \rrbracket_{sc}^O \wedge \mathcal{E}$)

We can see that we can observe $r_1 = s_1 = 1 \wedge r_2 = s_2 = 0$.

Conclusion

Extensions.

- ▶ We can define sub-models of \mathcal{E} corresponding to actual architectures.
- ▶ The model is inspired from a game semantics model and simplified in this first-order setting.

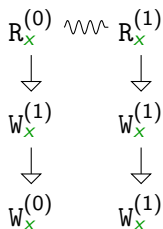
To go further.

- ▶ Look at barriers
- ▶ Compare that with axiomatic semantics (compare the executions)

Enrichir les labels

Considérons: $t = \begin{pmatrix} r \leftarrow x; \\ x := 1; \\ x := r. \end{pmatrix}$.

Sans réordonnement:



On a $W_x^{(1)} \rightarrow W_x^{(0)}$ et $W_x^{(1)} \rightarrow W_x^{(1)}$ selon le passé.

Enrichir les labels

Considérons: $t = \begin{pmatrix} r \leftarrow x; \\ x := 1; \\ x := r. \end{pmatrix}$.

Sans réordonnement:

$$\begin{array}{ccc} R_x^{(0),r:0} & \rightsquigarrow & R_x^{(1),r:1} \\ \downarrow & & \downarrow \\ W_x^{(1),r:0} & & W_x^{(1),r:1} \\ \downarrow & & \downarrow \\ W_x^{(0),r:0} & & W_x^{(1),r:1} \end{array}$$

On a $W_x^{(1)} \rightarrow W_x^{(0)}$ et $W_x^{(1)} \rightarrow W_x^{(1)}$ selon le passé.
Pour lever l'ambiguïté, on définit

$$\bar{\Sigma} = \Sigma \times (\mathcal{R} \rightarrow \mathbb{N}).$$

Dépendances entre labels

On peut définir maintenant $\xrightarrow{\text{dep}} \subseteq \bar{\Sigma} \times \bar{\Sigma}$ selon l'architecture:

- ▶ x86 (permuter écriture/lecture sur des zones différentes):

$$\begin{aligned} \xrightarrow{\text{dep}}_{\text{x86}} = \{ & (e, \rho), (e', \rho') \mid \\ & \rho \subseteq \rho' \\ & \wedge (\text{var}(e) = \text{var}(e') \vee \text{type}(\{e, e'\}) \neq \{\text{R}, \text{W}\}) \} \end{aligned}$$

- ▶ ARM (permuter les instructions sur des zones différentes):

$$\begin{aligned} \xrightarrow{\text{dep}}_{\text{ARM}} = \{ & (e, \rho), (e', \rho') \mid \\ & \rho \subseteq \rho' \\ & \wedge (\text{var}(e) = \text{var}(e')) \} \end{aligned}$$

avec $\text{type} : \Sigma \rightarrow \{\text{R}, \text{W}\}$, $\text{var} : \Sigma \rightarrow \mathcal{V}$.

Opérations sur les structures d'évènements

[noframenumbering]

- ▶ **Somme.** Pour $(I_x \in \bar{\Sigma})_{x \in X}$, on forme $\sum_{x \in X} u_x$:
 - ▶ évènements: X
 - ▶ causalité: \leq est l'égalité
 - ▶ conflit: deux évènements distincts sont en conflits:

$$\sum_{k \in \mathbb{N}} R_x^{(k)} = R_x^{(0)} \sim R_x^{(1)} \sim \dots$$

- ▶ **Mise en parallèle.** Soit E, F des structures d'évènements et $\rightarrow_d \subseteq E \times F$. On forme $E \parallel_{\rightarrow_d} F$:
 - ▶ évènements: union disjointe de E et F
 - ▶ causalité: plus petit ordre qui contient \leq_E, \leq_F et \rightarrow_d
 - ▶ conflit: union de \sim_E et \sim_F .

On note $E \parallel F$ pour $E \parallel_{\emptyset} F$.

Définition de la sémantique

Note: $x := 1$ dépend des lectures sur x précédentes.

Définition de la sémantique

Note: $x := 1$ dépend des lectures sur x précédentes.

Notre sémantique est donc paramétrée par $\sigma : \mathcal{R} \rightarrow \mathcal{V}$.

1. Pour une instruction ι , D_ι : les registres dont ι dépend:

	Écriture	Lecture
x86	$D_x^\sigma := e = \sigma^{-1}(x) \cup \text{fv}(e)$	$D_r \leftarrow x = \{r\} \cup \text{dom}(\sigma)$
ARM		$D_r \leftarrow x = \{r\} \cup \sigma^{-1}(x)$

Définition de la sémantique

Note: $x := 1$ dépend des lectures sur x précédentes.

Notre sémantique est donc paramétrée par $\sigma : \mathcal{R} \rightarrow \mathcal{V}$.

1. Pour une instruction ι , D_ι : les registres dont ι dépend:

	Écriture	Lecture
x86 ARM	$D_x^\sigma := e = \sigma^{-1}(x) \cup \text{fv}(e)$	$D_r \leftarrow x = \{r\} \cup \text{dom}(\sigma)$ $D_r \leftarrow x = \{r\} \cup \sigma^{-1}(x)$

2. Sémantique des threads:

$$\llbracket x := e; t \rrbracket \sigma = \left(\sum_{\rho: (D_x^\sigma := e) \rightarrow \mathbb{N}} W_x^{(\rho(e), \rho)} \right) \parallel_{\text{dep}} \llbracket t \rrbracket \sigma$$
$$\llbracket r \leftarrow x; t \rrbracket \sigma = \left(\sum_{\rho: (D_r^\sigma \leftarrow x) \rightarrow \mathbb{N}} R_x^{(\rho(r), \rho)} \right) \parallel_{\text{dep}} \llbracket t \rrbracket (\sigma[r \leftarrow x])$$

Définition de la sémantique

Note: $x := 1$ dépend des lectures sur x précédentes.

Notre sémantique est donc paramétrée par $\sigma : \mathcal{R} \rightarrow \mathcal{V}$.

1. Pour une instruction ι , D_ι : les registres dont ι dépend:

	Écriture	Lecture
x86 ARM	$D_x^\sigma := e = \sigma^{-1}(x) \cup \text{fv}(e)$	$D_r \leftarrow x = \{r\} \cup \text{dom}(\sigma)$ $D_r \leftarrow x = \{r\} \cup \sigma^{-1}(x)$

2. Sémantique des threads:

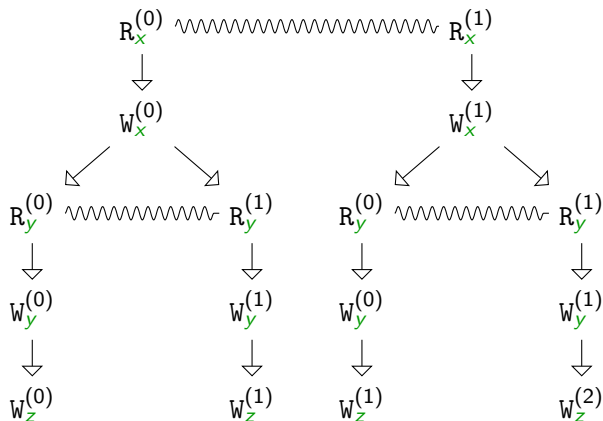
$$\llbracket x := e; t \rrbracket \sigma = \left(\sum_{\rho: (D_x^\sigma := e) \rightarrow \mathbb{N}} W_x^{(\rho(e)), \rho} \right) \parallel_{\text{dep}} \llbracket t \rrbracket \sigma$$
$$\llbracket r \leftarrow x; t \rrbracket \sigma = \left(\sum_{\rho: (D_r^\sigma \leftarrow x) \rightarrow \mathbb{N}} R_x^{(\rho(r)), \rho} \right) \parallel_{\text{dep}} \llbracket t \rrbracket (\sigma[r \leftarrow x])$$

3. Sémantique des programmes:

$$\llbracket t_1 \parallel \dots \parallel t_n \rrbracket = \llbracket t_1 \rrbracket \emptyset \parallel \dots \parallel \llbracket t_n \rrbracket \emptyset$$

Exemples

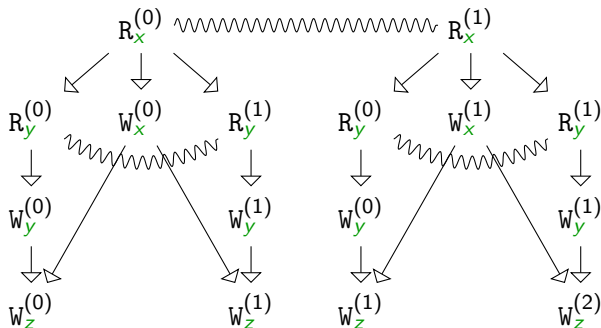
Pour $t = \begin{pmatrix} s \leftarrow x; x := s; \\ t \leftarrow y; y := t; \\ z := s + t \end{pmatrix}$, on a:



(SC)

Exemples

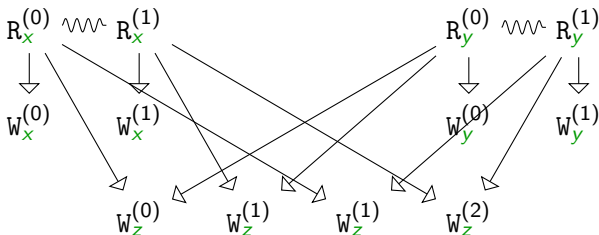
Pour $t = \begin{pmatrix} s \leftarrow x; x := s; \\ t \leftarrow y; y := t; \\ z := s + t \end{pmatrix}$, on a:



(x86)

Exemples

Pour $t = \begin{pmatrix} s \leftarrow x; x := s; \\ t \leftarrow y; y := t; \\ z := s + t \end{pmatrix}$, on a:



(ARM)

Et maintenant, la sémantique close

Pour calculer la sémantique close, on peut

1. Voir M comme une structure d'évènements E_M .
(Ses configurations sont les traces de M)
2. Calculer le produit synchronisé: $\llbracket p \rrbracket = \llbracket p \rrbracket^O \wedge E_M$.
(Opération combinatoire très compliquée)

Problème: l'ordre sur E_M est total donc $\llbracket p \rrbracket$ n'a pas de concurrence.

Solutions possibles:

1. Construire E_M comme $E_{M_x} \parallel E_{M_y} \parallel \dots$: concurrence inter-variable.
2. Construire E_M comme une collection d'ordre partiels: concurrence intra-variable.

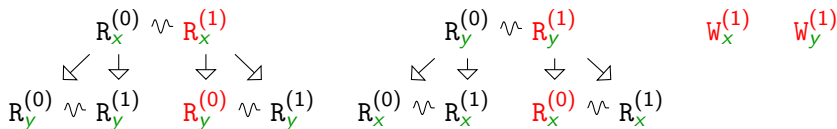
Un exemple

$$p = \begin{array}{l} r \leftarrow x \\ u \leftarrow y \end{array} \parallel \begin{array}{l} s \leftarrow y \\ v \leftarrow x \end{array} \parallel x := 1 \parallel y := 1$$

Un exemple

$$\rho = \begin{array}{l} r \leftarrow x \parallel s \leftarrow y \parallel x := 1 \parallel y := 1 \\ u \leftarrow y \parallel v \leftarrow x \end{array}$$

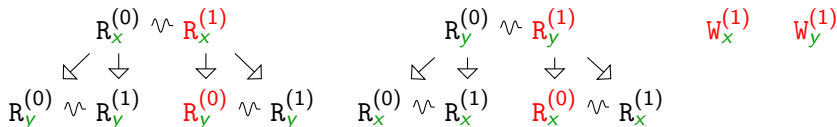
Sémantique ouverte (sur x86):



Un exemple

$$p = \begin{array}{l} r \leftarrow x \\ u \leftarrow y \end{array} \parallel \begin{array}{l} s \leftarrow y \\ v \leftarrow x \end{array} \parallel x := 1 \parallel y := 1$$

Sémantique ouverte (sur x86):



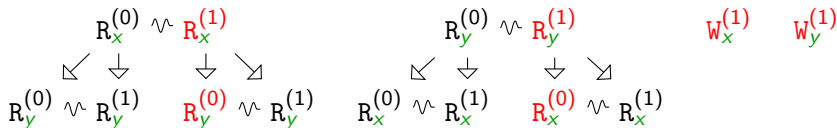
Sémantique close. (Mémoire par traces)

$$\begin{array}{l} W_x^{(1)} \rightarrow W_y^{(1)} \rightarrow R_x^{(1)} \rightarrow R_y^{(1)} \rightarrow R_y^{(0)} \rightarrow R_x^{(0)} \\ W_y^{(1)} \rightarrow W_x^{(1)} \rightarrow R_x^{(1)} \rightarrow R_y^{(1)} \rightarrow R_y^{(0)} \rightarrow R_x^{(0)} \\ W_x^{(1)} \rightarrow R_x^{(1)} \rightarrow W_y^{(1)} \rightarrow R_y^{(1)} \rightarrow R_y^{(0)} \rightarrow R_x^{(0)} \\ \dots \end{array} \in \mathcal{L}(\llbracket p \rrbracket)$$

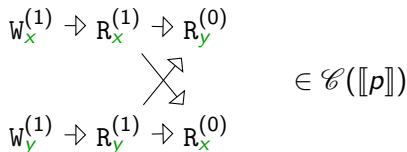
Un exemple

$$p = \begin{array}{l} r \leftarrow x \\ u \leftarrow y \end{array} \parallel \begin{array}{l} s \leftarrow y \\ v \leftarrow x \end{array} \parallel x := 1 \parallel y := 1$$

Sémantique ouverte (sur x86):



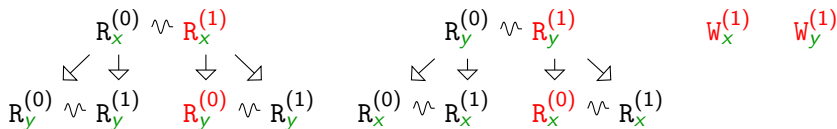
Sémantique close. (Mémoire avec concurrence inter-variable)



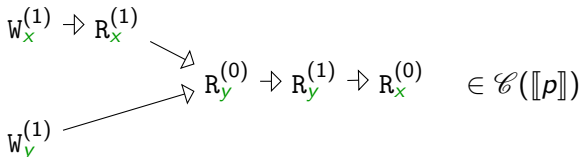
Un exemple

$$p = \begin{array}{l} r \leftarrow x \\ u \leftarrow y \end{array} \parallel \begin{array}{l} s \leftarrow y \\ v \leftarrow x \end{array} \parallel x := 1 \parallel y := 1$$

Sémantique ouverte (sur x86):



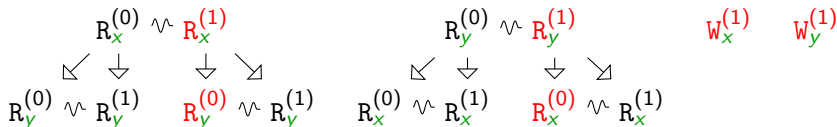
Sémantique close. (Mémoire avec concurrence inter-variable)



Un exemple

$$p = \begin{array}{l} r \leftarrow x \\ u \leftarrow y \end{array} \parallel \begin{array}{l} s \leftarrow y \\ v \leftarrow x \end{array} \parallel x := 1 \parallel y := 1$$

Sémantique ouverte (sur x86):



Sémantique close. (Mémoire avec concurrence intra-variable)

$$W_x^{(1)} \rightarrow R_x^{(1)} \rightarrow R_y^{(0)} \in \mathcal{C}(\llbracket p \rrbracket)$$

$$W_y^{(1)} \rightarrow R_y^{(1)} \rightarrow R_x^{(0)}$$

Un exemple de M non non séquentiel

On a la sémantique *volatile*, comment calculer la sémantique *close* ?

But: modéliser des écritures différées.

- ▶ Les labels indiquent le *thread* dont ils viennent: $\Sigma_{id} = \Sigma \times \mathbb{N}$
- ▶ À la mémoire globale $\mu : \mathcal{V} \rightarrow \mathbb{N}$, s'ajoute une mémoire locale $\lambda : \mathbb{N} \rightarrow (\mathcal{V} \rightarrow \mathbb{N})$.

Ensuite, on peut définir $M : (\mathbb{N} \rightarrow (\mathcal{V} \rightarrow \mathbb{N})) \rightarrow (\mathcal{V} \rightarrow \mathbb{N}) \rightarrow \Sigma_{id}^*$:

$$M_{\lambda, \rho} ::= \epsilon$$

Un exemple de M non non séquentiel

On a la sémantique *volatile*, comment calculer la sémantique *close* ?

But: modéliser des écritures différées.

- ▶ Les labels indiquent le *thread* dont ils viennent: $\Sigma_{id} = \Sigma \times \mathbb{N}$
- ▶ À la mémoire globale $\mu : \mathcal{V} \rightarrow \mathbb{N}$, s'ajoute une mémoire locale $\lambda : \mathbb{N} \rightarrow (\mathcal{V} \rightarrow \mathbb{N})$.

Ensuite, on peut définir $M : (\mathbb{N} \rightarrow (\mathcal{V} \rightarrow \mathbb{N})) \rightarrow (\mathcal{V} \rightarrow \mathbb{N}) \rightarrow \Sigma_{id}^*$:

$$M_{\lambda, \rho} ::= \epsilon$$
$$| (\mathbf{R}_x^{(\lambda(\iota)(x))}, \iota) \cdot M_{\lambda, \rho} \quad (\text{Lecture cache})$$

Un exemple de M non non séquentiel

On a la sémantique *volatile*, comment calculer la sémantique *close* ?

But: modéliser des écritures différées.

- ▶ Les labels indiquent le *thread* dont ils viennent: $\Sigma_{id} = \Sigma \times \mathbb{N}$
- ▶ À la mémoire globale $\mu : \mathcal{V} \rightarrow \mathbb{N}$, s'ajoute une mémoire locale $\lambda : \mathbb{N} \rightarrow (\mathcal{V} \rightarrow \mathbb{N})$.

Ensuite, on peut définir $M : (\mathbb{N} \rightarrow (\mathcal{V} \rightarrow \mathbb{N})) \rightarrow (\mathcal{V} \rightarrow \mathbb{N}) \rightarrow \Sigma_{id}^*$:

$$M_{\lambda, \rho} ::= \epsilon$$

$$| (R_x^{(\lambda(\iota)(x))}, \iota) \cdot M_{\lambda, \rho} \quad \text{(Lecture cache)}$$

$$| (R_x^{(\mu(x))}, \iota) \cdot M_{\lambda[(\iota, x) \leftarrow n], \mu} \quad \text{(Lecture globale)}$$

Un exemple de M non non séquentiel

On a la sémantique *volatile*, comment calculer la sémantique *close* ?

But: modéliser des écritures différées.

- ▶ Les labels indiquent le *thread* dont ils viennent: $\Sigma_{id} = \Sigma \times \mathbb{N}$
- ▶ À la mémoire globale $\mu : \mathcal{V} \rightarrow \mathbb{N}$, s'ajoute une mémoire locale $\lambda : \mathbb{N} \rightarrow (\mathcal{V} \rightarrow \mathbb{N})$.

Ensuite, on peut définir $M : (\mathbb{N} \rightarrow (\mathcal{V} \rightarrow \mathbb{N})) \rightarrow (\mathcal{V} \rightarrow \mathbb{N}) \rightarrow \Sigma_{id}^*$:

$M_{\lambda, \rho} ::= \epsilon$

| $(R_x^{(\lambda(\iota)(x))}, \iota) \cdot M_{\lambda, \rho}$ (Lecture cache)

| $(R_x^{(\mu(x))}, \iota) \cdot M_{\lambda[(\iota, x) \leftarrow n], \mu}$ (Lecture globale)

| $(W_\iota^{(n)}, \iota) \cdot M_{\lambda[(\iota, x) \leftarrow n], \mu[x \leftarrow n]}$ (Écriture)

Un exemple de M non non séquentiel

On a la sémantique *volatile*, comment calculer la sémantique *close* ?

But: modéliser des écritures différées.

- ▶ Les labels indiquent le *thread* dont ils viennent: $\Sigma_{id} = \Sigma \times \mathbb{N}$
- ▶ À la mémoire globale $\mu : \mathcal{V} \rightarrow \mathbb{N}$, s'ajoute une mémoire locale $\lambda : \mathbb{N} \rightarrow (\mathcal{V} \rightarrow \mathbb{N})$.

Ensuite, on peut définir $M : (\mathbb{N} \rightarrow (\mathcal{V} \rightarrow \mathbb{N})) \rightarrow (\mathcal{V} \rightarrow \mathbb{N}) \rightarrow \Sigma_{id}^*$:

$$M_{\lambda, \rho} ::= \epsilon$$

$$\mid (\mathbf{R}_x^{(\lambda(\ell)(x))}, \ell) \cdot M_{\lambda, \rho} \quad (\text{Lecture cache})$$

$$\mid (\mathbf{R}_x^{(\mu(x))}, \ell) \cdot M_{\lambda[(\ell, x) \leftarrow n], \mu} \quad (\text{Lecture globale})$$

$$\mid (\mathbf{W}_\ell^{(n)}, \ell) \cdot M_{\lambda[(\ell, x) \leftarrow n], \mu[x \leftarrow n]} \quad (\text{Écriture})$$

$$M ::= M(\ell \mapsto (x \mapsto 0))(x \mapsto 0)$$

Traces d'une structure d'évènements

Definition (Trace d'une configuration)

Une trace de $w \in \mathcal{C}(E)$ est une linéarisation de l'ordre partiel w .
L'ensemble des traces de w est dénoté $\text{tr}(w)$.

Exemple: $\text{tr}(R_x^{(1)} \quad W_y^{(2)}) = \{R_x^{(1)} \cdot W_y^{(2)}, W_y^{(2)} \cdot R_x^{(1)}\}$.

Traces d'une structure d'évènements

Definition (Trace d'une configuration)

Une trace de $w \in \mathcal{C}(E)$ est une linéarisation de l'ordre partiel w .
L'ensemble des traces de w est dénoté $\text{tr}(w)$.

Exemple: $\text{tr}(R_x^{(1)} \quad W_y^{(2)}) = \{R_x^{(1)} \cdot W_y^{(2)}, W_y^{(2)} \cdot R_x^{(1)}\}$.

On en déduit une nouvelle sémantique close par traces:

$$\llbracket p \rrbracket = \bigcup \{ \text{tr}(w) \cap M \mid w \in \mathcal{C}(\llbracket p \rrbracket^0) \}$$

Programme	$(x := 1; r \leftarrow y) \parallel (y := 1; s \leftarrow x)$
Sémantique volatile	$W_x^{(1)} \quad R_y^{(0)} \wedge R_y^{(1)} \quad W_y^{(1)} \quad R_x^{(0)} \wedge R_x^{(2)}$
Éléments dans $\llbracket p \rrbracket$	$R_x^{(0)} \cdot R_y^{(0)} \cdot W_x^{(1)} \cdot W_y^{(1)}$ $R_x^{(0)} \cdot W_x^{(1)} \cdot R_y^{(0)} \cdot W_y^{(1)}$

Et avec une structure d'évènements

- ▶ On doit séquentialiser *toute l'exécution* → explosion
- ▶ Or, seulement l'histoire d'**une** variable doit l'être.

Et avec une structure d'évènements

- ▶ On doit séquentialiser *toute l'exécution* → explosion
- ▶ Or, seulement l'histoire d'**une** variable doit l'être.
- ▶ Tous nos modèles mémoires ont une forme particulière:

$$M = M_x \circledast M_y \circledast \dots \quad (\text{avec } \text{var}(M_x) = \{x\} \text{ pour } x \in \mathcal{V})$$

- ▶ Cela mène à une notion d'exécutions partiellement ordonnées: paires $(w, (t_x))$
 - ▶ $w \in \mathcal{C}(\llbracket p \rrbracket^0)$
 - ▶ pour $x, t_x \in M_x$ avec $\text{tr}(w) \cap (t_x \circledast t_y \circledast \dots) \neq \emptyset$
- ▶ **Théorème.** Il existe une structure d'évènements dont les configurations sont les exécutions partiellement ordonnées.