

# The hierarchy of analytic semantics of weakly consistent parallelism

Jade Alglave (MSR-Cambridge, UCL, UK)

Patrick Cousot (NYU, Emer. ENS, PSL)

IMDEA seminar

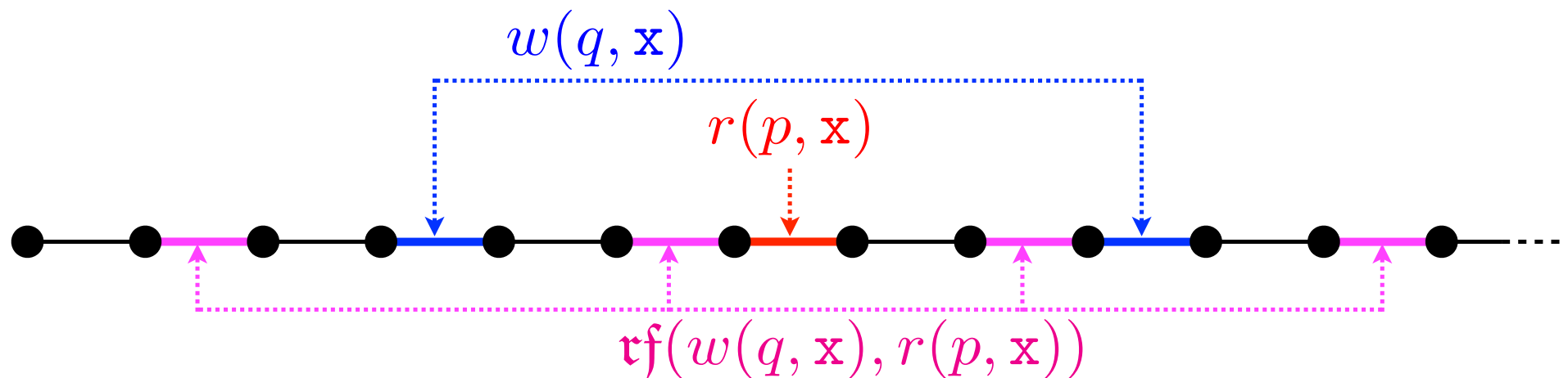
Madrid

Tuesday, May 24<sup>th</sup>, 2016 — 11:00 AM

# Analytic semantics

# Weak consistency models (WCM)

- Sequential consistency:  
reads  $r(p, x)$  are *implicitly coordinated* with writes  $w(q, x)$
- WCM:  
*No implicit coordination* (depends on architecture, program dependencies, and explicit fences)



# Analytic semantic specification

- **Anarchic semantics:**  
describes computations, no constraints on communications
- **cat specification (Jade Alglave & Luc Maranget):**  
imposes architecture-dependent communication constraints
- **Hierarchy of anarchic semantics:**  
many different styles to describe the same computations (e.g. stateless/stateful, interleaved versus true parallelism)

# Example: load buffer (LB)

● Program:  $\{ x = 0; y = 0; \}$

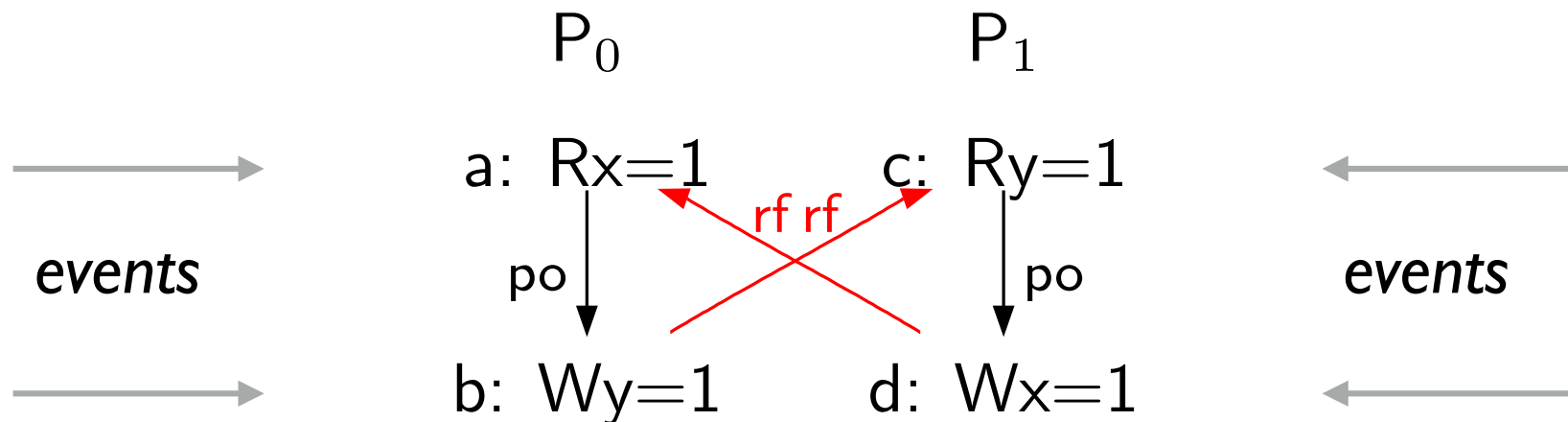
P0		P1	;
r[] r1 x		r[] r2 y	;
w[] y 1		w[] x 1	;

$\text{exists}(0:r1=1 \wedge 1:r2=1)$

● Example of execution trace  $t \in S^\perp \llbracket P \rrbracket$ :

$t =$   $w(\text{start}, x, 0)$   $w(\text{start}, y, 0)$   $r(\text{P0}, x, 1)$   $\text{rf}[w(\text{P1}, x, 1), r(\text{P0}, x, 1)]$   $w(\text{P0}, y, 1)$   $r(\text{P1}, y, 1)$   
 $w(\text{P1}, x, 1)$   $\text{rf}[w(\text{P0}, y, 1), r(\text{P1}, y, 1)]$   $r(\text{finish}, x)$   $\text{rf}[w(\text{P1}, x, 1), r(\text{finish}, x, 1)]$   
 $r(\text{finish}, y, 1)$   $\text{rf}[w(\text{P0}, y, 1), r(\text{finish}, y, 1)]$

● Abstraction to cat *candidate execution*  $\alpha_\Xi(t)$ :



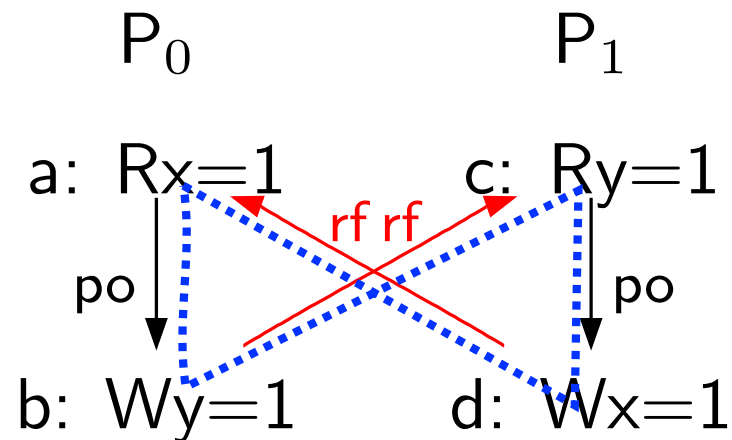
# Example: load buffer (LB), cont'd

- cat specification:

acyclic (po | rf)+

The cat semantics rejects this execution  $\alpha_{\Xi}(t)$  :

$$\text{cat} \llbracket \text{cat} \rrbracket (\alpha_{\Xi}(t)) = \text{false}$$



- The herd7 tool: [virginia.cs.ucl.ac.uk/herd/](http://virginia.cs.ucl.ac.uk/herd/)

# The WCM semantics

- Abstraction to a **candidate execution**:

$$\alpha_{\Xi}(t) \triangleq \langle \alpha_e(t), \alpha_{po}(t), \alpha_{rf}(t), \alpha_{iw}(t), \alpha_{fw}(t) \rangle$$

$$\alpha_{\Xi}(\mathcal{S}) \triangleq \{ \langle t, \alpha_{\Xi}(t) \rangle \mid t \in \mathcal{S} \}$$

- The **cat semantics**:

$$\alpha_{\text{cat}}(\mathcal{S}) \triangleq \{ t \mid \langle t, \Xi \rangle \in \mathcal{S} \wedge \text{cat}(\Xi) \}$$

- The **WCM semantics**:

$$\alpha_{\text{cat}} \circ \alpha_{\Xi}(\mathcal{S}[\mathcal{P}])$$

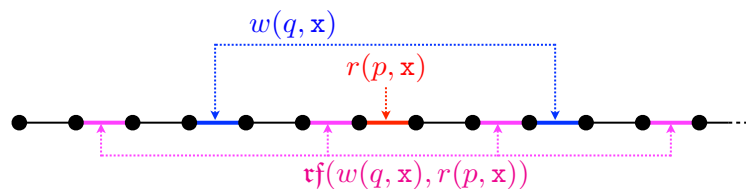
$$\text{GC: } \langle \wp(\mathcal{E}^{+\infty}), \sqsubseteq \rangle \xrightleftharpoons[\alpha_{\Xi}]{\gamma_{\Xi}} \langle \wp(\mathcal{E}^{+\infty} \times \Xi), \sqsubseteq \rangle \xrightleftharpoons[\alpha_{\text{cat}}]{\gamma_{\text{cat}}} \langle \wp(\mathcal{E}^{+\infty}), \sqsubseteq \rangle$$

# Definition of the anarchic semantics



# Axiomatic parameterized definition of the anarchic semantics

- The semantics  $S^\perp \llbracket P \rrbracket$  is a finite/infinite **sequence of interleaved events of processes** satisfying well-formedness conditions.
- Events:
  - local computations and tests on registers
  - start writing a shared variable  $w(q, \mathbf{x})$
  - start reading of shared variable  $r(p, \mathbf{x})$
  - communication event  $\text{rf}(w(q, \mathbf{x}), r(p, \mathbf{x}))$



# Axiomatic parameterized definition of the anarchic semantics

- Examples of language independent well-formedness conditions of a semantics  $S$ :

- uniqueness of events

$$\forall t \in \mathcal{S} . \forall t_1, t_2 \in \mathcal{E}^*, t_3 \in \mathcal{E}^{*\infty} . \forall e, e' \in \mathcal{E} . (t = t_1 e t_2 e' t_3) \implies (e \neq e') . \quad (\text{Wf}_1(S))$$

- traces start with an initialization of the shared variables  $(\text{Wf}_2(S))$

$$t = w(\text{start}, x, 0) w(\text{start}, y, 0) r(\text{P0}, x, 1) \text{rf}[w(\text{P1}, x, 1), r(\text{P0}, x, 1)] w(\text{P0}, y, 1) r(\text{P1}, y, 1) \\ w(\text{P1}, x, 1) \text{rf}[w(\text{P0}, y, 1), r(\text{P1}, y, 1)] r(\text{finish}, x) \text{rf}[w(\text{P1}, x, 1), r(\text{finish}, x, 1)] \\ r(\text{finish}, y, 1) \text{rf}[w(\text{P0}, y, 1), r(\text{finish}, y, 1)]$$

# Axiomatic parameterized definition of the anarchic semantics

- Examples of language independent well-formedness conditions of a semantics  $S$ :
  - finite traces are **maximal**

$$\forall t \in S \cap \mathcal{E}^+ . \nexists t' \in \mathcal{E}^{+\infty} . t t' \in S . \quad (\text{Wf}_3(S))$$

- the final value of shared variables in finite traces is known thanks to a **final read** ( $\text{Wf}_4(S)$ )

$t =$ 
 $w(\text{start}, x, 0)$ 
 $w(\text{start}, y, 0)$ 
 $r(\text{P0}, x, 1)$ 
 $\text{rf}[w(\text{P1}, x, 1), r(\text{P0}, x, 1)]$ 
 $w(\text{P0}, y, 1)$ 
 $r(\text{P1}, y, 1)$ 
  
 $w(\text{P1}, x, 1)$ 
 $\text{rf}[w(\text{P0}, y, 1), r(\text{P1}, y, 1)]$ 
 $r(\text{finish}, x)$ 
 $\text{rf}[w(\text{P1}, x, 1), r(\text{finish}, x, 1)]$ 
  
 $r(\text{finish}, y, 1)$ 
 $\text{rf}[w(\text{P0}, y, 1), r(\text{finish}, y, 1)]$

# Axiomatic parameterized definition of the anarchic semantics

- Examples of language independent well-formedness conditions of a semantics  $S$ :
  - **read events must be satisfied by a unique communication event**

$$\forall t \in S . \forall t_1 \in \mathcal{E}^*, t_2 \in \mathcal{E}^{*\infty} . (t = t_1 r(p, \mathbf{x}) t_2) \implies (\exists t_3 \in \mathcal{E}^*, t_4 \in \mathcal{E}^{*\infty} . t = t_3 \text{rf}[w(q, \mathbf{x}), r(p, \mathbf{x})] t_4) . \quad (\text{Wf}_5(S))$$

$$\forall t \in S . \forall t_1, t_2 \in \mathcal{E}^*, t_3 \in \mathcal{E}^{*\infty} . (t \neq t_1 \text{rf}[w(q, \mathbf{x}), r(p, \mathbf{x})] t_2 \text{rf}[w'(q', \mathbf{x}), r(p, \mathbf{x})] t_3) . \quad (\text{Wf}_6(S))$$

# Axiomatic parameterized definition of the anarchic semantics

- Examples of language independent well-formedness conditions of a semantics  $S$ :
  - communications cannot be **spontaneous** (must be originated by a read *and* a write)

$$\forall t \in S . \forall t_1 \in \mathcal{E}^*, t_2 \in \mathcal{E}^{*\infty} . (t = t_1 \text{ rf}[w(q, \mathbf{x}), r(p, \mathbf{x})] t_2) \implies (\exists t_3 \in \mathcal{E}^*, t_4 \in \mathcal{E}^{*\infty} . t = t_3 w(q, \mathbf{x}) t_4 \wedge \exists t_5 \in \mathcal{E}^*, t_6 \in \mathcal{E}^{*\infty} . t = t_5 r(p, \mathbf{x}) t_6) . \quad (\text{Wf}_7(S))$$


# Axiomatic parameterized definition of the anarchic semantics

- The language :

- Programs :  $\text{initialisation } \llbracket P_1 \parallel \dots \parallel P_n \rrbracket \text{ finalisation}$   


- Actions (labelled  $\ell \in \mathbb{L}(p)$ ) :

$a ::= m$	imperative actions	marker
$r := e$		assignment
$r := x$		read of shared variable x
$x := e$		write of shared variable x
$b \mid \neg b$	conditional actions	test

- Next action :  $\text{next}(p, \ell)$     $\text{nextt}(p, \ell)$     $\text{nextf}(p, \ell)$   


# Axiomatic parameterized definition of the anarchic semantics

- Example of language-dependent well-formedness condition: **computation** (markers: skip, fence, begin/end of rmw)

Any process  $p$     Any point  $k$   
in trace    Any label  $\ell$   
of  $p$

marker event by  
process  $p$  in trace  $\tau$

$$\forall p \in \mathbb{P}i . \forall k \in [1, 1 + |\tau|[ . \forall \ell \in \mathbb{L}(p) . \quad (\text{Wf}_{21}(\tau))$$

$$(\exists \theta \in \mathfrak{P}(p) . \bar{\tau}_k = \mathfrak{m}(\langle p, \ell, m, \theta \rangle))$$

$$\implies (\ell \in N^p(\tau, k) \wedge \text{action}(p, \ell) = m) .$$

(unique) event  
stamp  $\theta$

control of process  
 $p$  is at label  $\ell$

action of process  $p$   
is at label  $\ell$  is the  
marker action  $m$

# Axiomatic parameterized definition of the anarchic semantics

- Example of language-dependent well-formedness condition: **computation** (local variable assignment)

register assignment event  
by process  $p$  in trace  $\tau$

(unique) event stamp  $\theta$

$$\forall p \in \mathbb{P}^i . \forall k \in ]1, 1 + |\tau|[ . \forall \ell \in \mathbb{L}(p) . \forall v \in \mathcal{D} . \quad (\text{Wf}_{22}(\tau))$$

$$(\exists \theta \in \mathfrak{P}(p) . \bar{\tau}_k = \mathbf{a}(\langle p, \ell, \mathbf{r} := e, \theta \rangle, v))$$

$$\implies (\ell \in N^p(\tau, k) \wedge \text{action}(p, \ell) = \mathbf{r} := e \wedge v = E^p[[e]](\tau, k - 1)) .$$

control of process  
 $p$  is at label  $\ell$

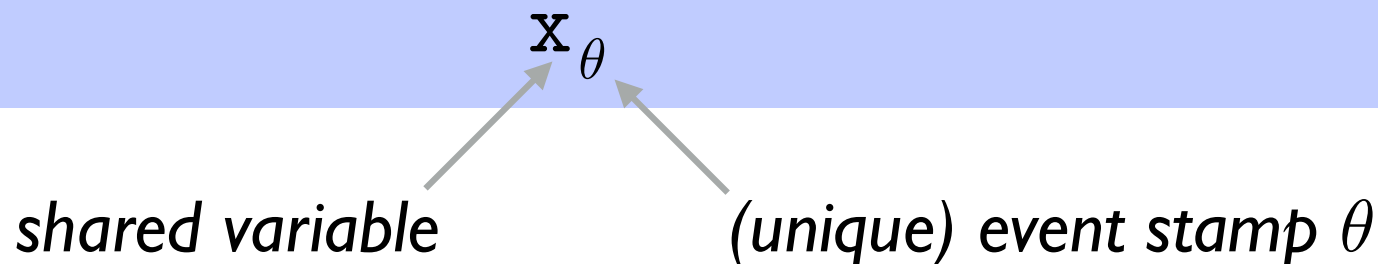
action of process  $p$   
is at label  $\ell$  is a  
register assignment

value  $v$  of  $e$  is  
evaluated by past-  
travel



# Media variables

- With WCM there is **no notion of** “*the current value of shared variable  $x$* ”
- At a given time each process may read a *different value* of the shared variable  $x$  (maybe guessed or unknown since a read may read from a future write)
- We use *media variables* (to record the values communicated between a write and read, whether the two accesses are on the same process or not)



# Axiomatic parameterized definition of the anarchic semantics

- Example: **communication**

- a read event is initiated by a read action:

*read event by  
process  $p$  in trace  $\tau$*

*unique media variable*

$$\forall p \in \mathbb{P}^i . \forall k \in ]1, 1 + |\tau|[ . \forall \ell \in \mathbb{L}(p) . \quad (\text{Wf}_{23}(\tau))$$

$$(\exists \theta \in \mathfrak{P}(p) . (\bar{\tau}_k = \mathbf{r}(\langle p, \ell, \mathbf{r} := \mathbf{x}, \theta \rangle, \mathbf{x}_\theta)))$$

$$\implies (\ell \in N^p(\tau, k) \wedge \text{action}(p, \ell) = \mathbf{r} := \mathbf{x}) .$$

- a read must read-from (**rf**) a write (weak fairness):

$$\forall p \in \mathbb{P}^i . \forall i \in ]1, 1 + |\tau|[ . \forall r \in \mathfrak{Rf}(p) . \quad (\text{Wf}_{26}(\tau))$$

$$(\bar{\tau}_i = r) \implies (\exists j \in ]1, 1 + |\tau|[ . \exists w \in \mathbb{W}^i . \bar{\tau}_j = \mathbf{rf}[w, r]) .$$

*communication (read-from) event*

# Axiomatic parameterized definition of the anarchic semantics

- **Predictive evaluation** of media variables:

$$V_{(32)}^p[[\mathbf{x}_\theta]](\tau, k) \triangleq v \text{ where } \exists! i \in [1, 1 + |\tau|[ \cdot (\bar{\tau}_i = \mathbf{r}(\langle p, \ell, \mathbf{r} := \mathbf{x}, \theta \rangle, \mathbf{x}_\theta)) \wedge \\ \exists! j \in [1, 1 + |\tau|[ \cdot (\bar{\tau}_j = \mathbf{rf}[\mathbf{w}(\langle p', \ell', \mathbf{x} := e', \theta' \rangle, v), \bar{\tau}_i])$$

- **Local past-travel** evaluation of an expression:

$$E_{(30)}^p[[\mathbf{r}]](\tau, k) \triangleq v \text{ if } k > 1 \wedge ((\bar{\tau}_k = \mathbf{a}(\langle p, \ell, \mathbf{r} := e, \theta \rangle, v)) \vee \\ (\bar{\tau}_k = \mathbf{r}(\langle p, \ell, \mathbf{r} := \mathbf{x}, \theta \rangle, \mathbf{x}_\theta) \wedge V^p[[\mathbf{x}_\theta]](\tau, k) = v))$$

$$E_{(30)}^p[[\mathbf{r}]](\tau, 1) \triangleq I[[0]] \quad \text{i.e. } \bar{\tau}_1 = \epsilon_{\text{start}} \text{ by } \mathbf{Wf}_{15}(\tau)$$

$$E_{(30)}^p[[\mathbf{r}]](\tau, k) \triangleq E_{(30)}^p[[\mathbf{r}]](\tau, k - 1) \quad \text{otherwise.}$$

# Abstractions of the anarchic semantics

# Abstractions

- Anarchic semantics:

$$S^\perp \llbracket P \rrbracket \triangleq \lambda \langle \mathcal{B}, \text{sat}, \mathcal{D}, I, \mathcal{G}, V, E, N \rangle \cdot \{ \tau \in \mathcal{T} \llbracket P \rrbracket \mid_{\cong} \mid \text{Wf}_1(\tau) \wedge \dots \wedge \text{Wf}_{29}(\tau) \}$$

↑  
↑  
parameters of the semantics

↑  
↑  
trace well-formedness conditions

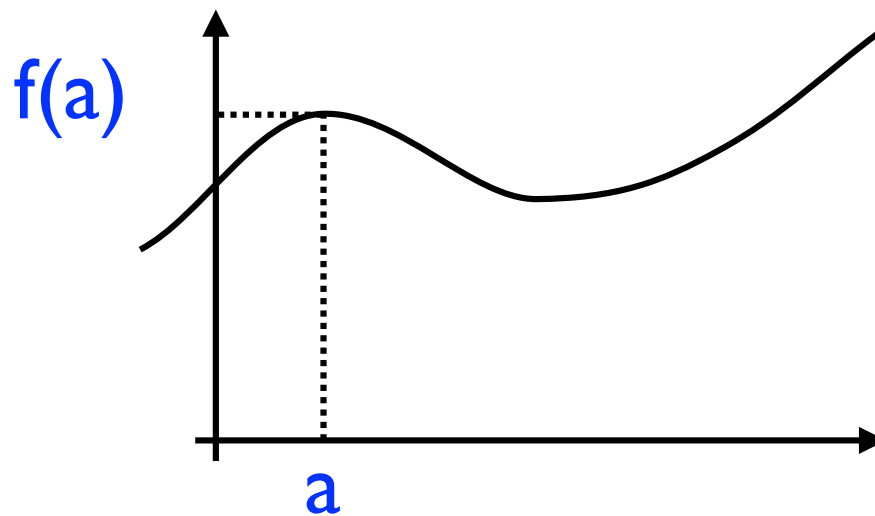
- Examples of **abstractions**:

- Choose data (e.g. ground values, uninterpreted symbolic expressions, interpreted symbolic expressions i.e. “symbolic guess”)
- Bind parameters (e.g. how expressions are evaluated)
- ...

# Binding a parameter of the semantics

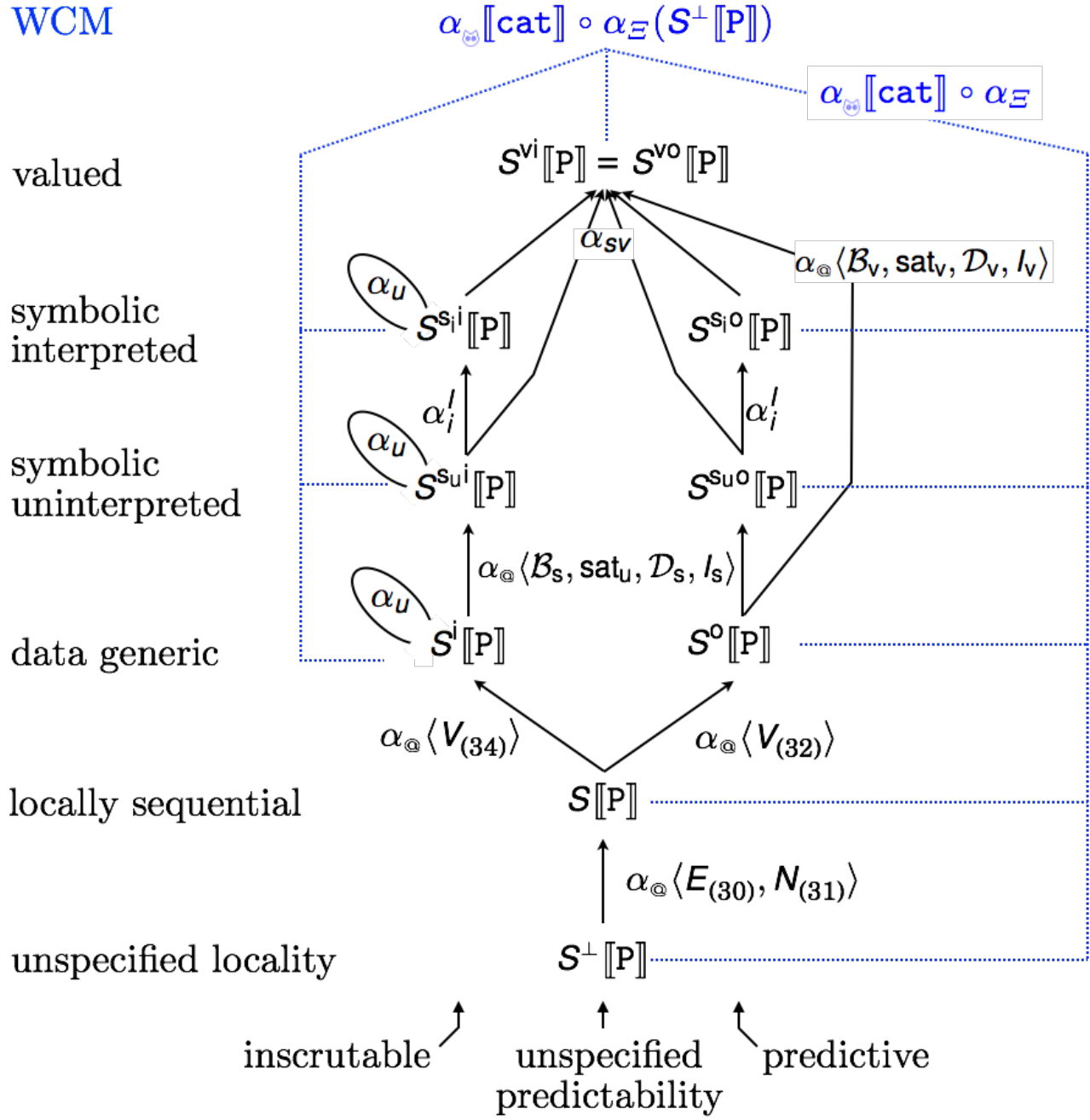
- The abstraction

$$\alpha_a(f) \stackrel{\text{def}}{=} f(a)$$



$$\langle \wp(A, B, \dots) \rightarrow \wp(R), \dot{\subseteq} \rangle \begin{array}{c} \xleftarrow{\gamma_a} \\ \xrightarrow{\alpha_a} \end{array} \langle \wp(B, \dots) \rightarrow \wp(R), \dot{\subseteq} \rangle$$

# The hierarchy of interleaved semantics



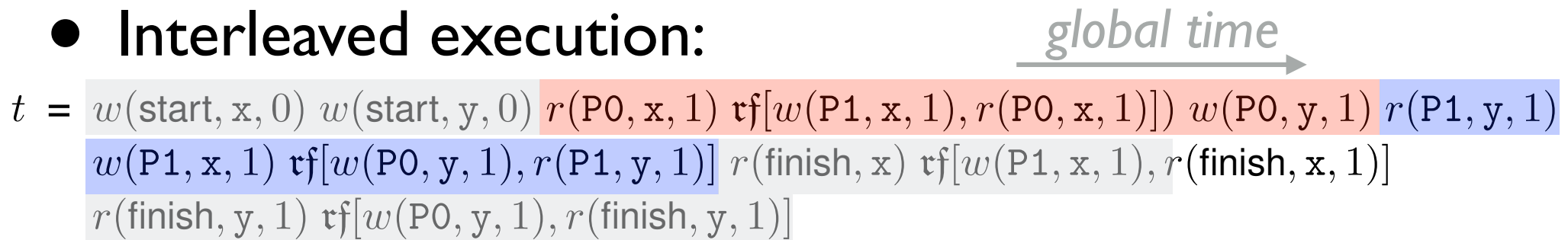
# True parallelism with local communications

- Extract from interleaved executions:
  - The **subtrace of each process** keeping **communications** in the process that read
    - ⇒ **no** more **global time** between processes
    - ⇒ **local time** between local actions and communications (a read can still tell when it is satisfied by which write)

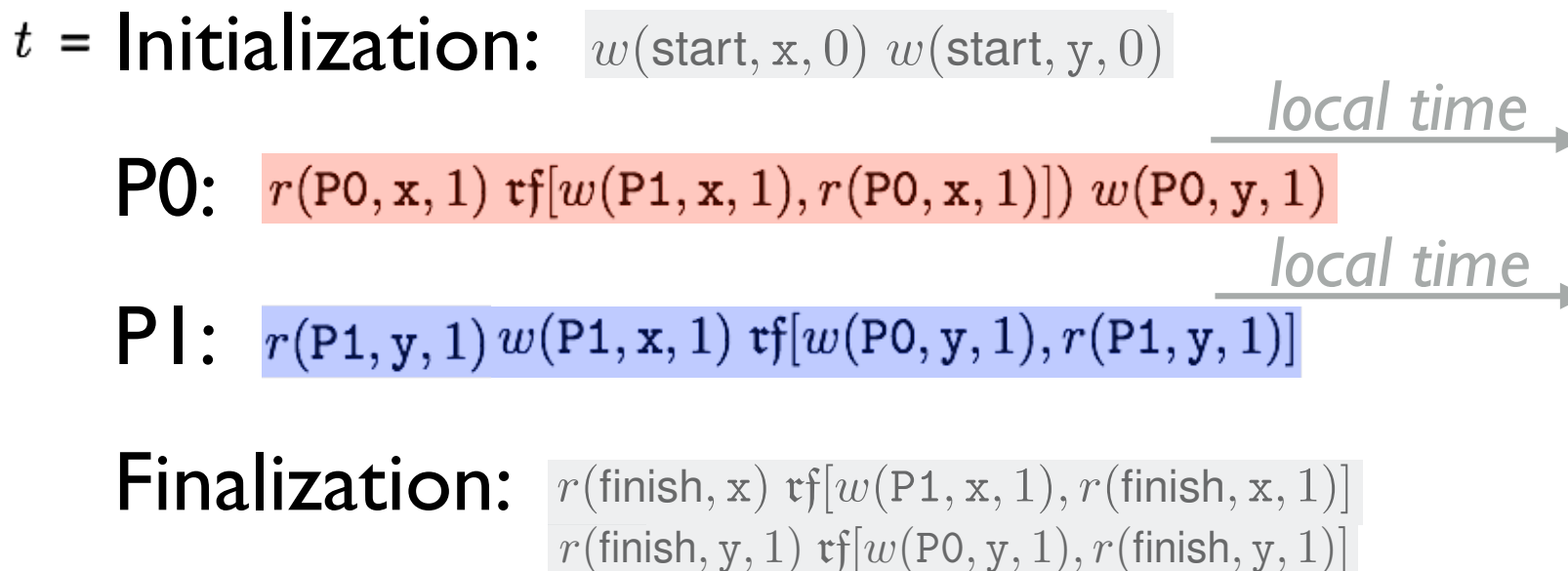


# True parallelism with local communications

- Interleaved execution:



- Parallel executions with interleaved communications:



# True parallelism of computations and communications

- Extract from interleaved executions:
  - The **subtrace of each process** (sequential execution of actions)
  - The **rf communication relation** (interactions between processes)
- ⇒ **no more global time** between processes
- ⇒ **no more global/local time** for communications

# True parallelism with separate communications

- Parallel executions with interleaved communications:

Initialization:  $w(\text{start}, x, 0) \ w(\text{start}, y, 0)$

P0:  $r(P0, x, 1) \ w(P0, y, 1)$  *local time*  $\rightarrow$

P1:  $r(P1, y, 1) \ w(P1, x, 1)$  *local time*  $\rightarrow$

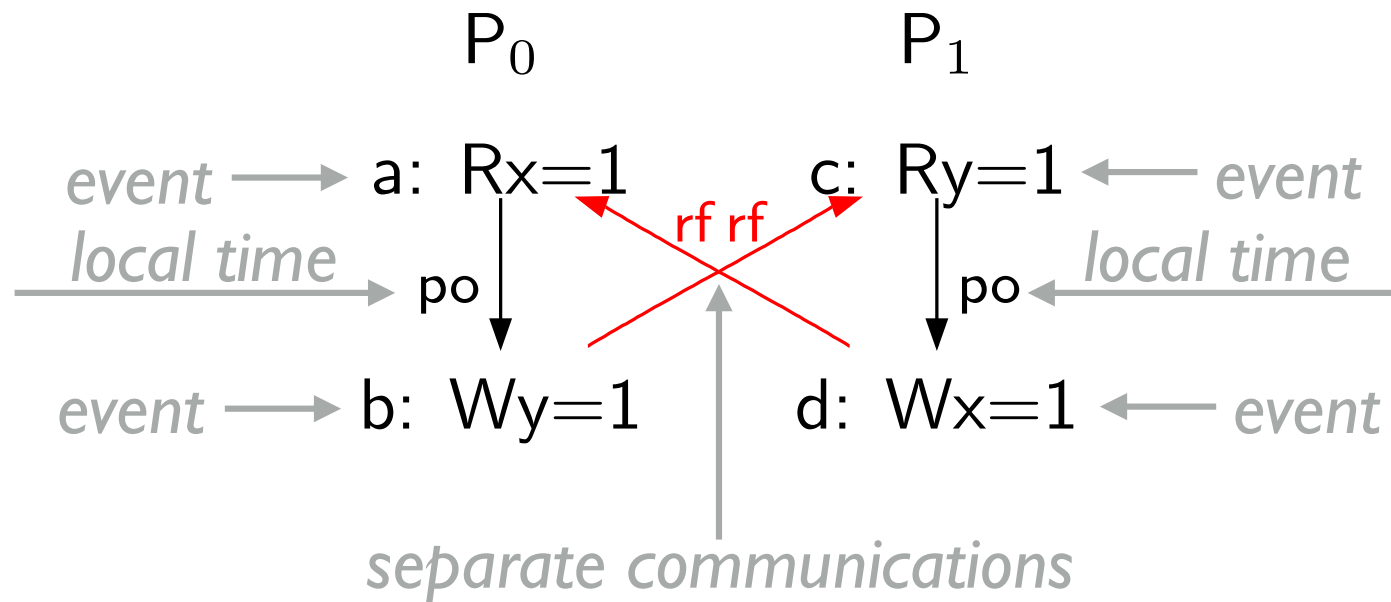
Finalization:  $r(\text{finish}, x) \ \text{rf}[w(P1, x, 1), r(\text{finish}, x, 1)]$   
 $r(\text{finish}, y, 1) \ \text{rf}[w(P0, y, 1), r(\text{finish}, y, 1)]$

Communications:

$\{ \text{rf}[w(P1, x, 1), r(P0, x, 1)] , \text{rf}[w(P0, y, 1), r(P1, y, 1)] \}$

# True parallelism with separate communications

- This is the semantics used by the herd7 tool:



+ interpreted symbolic expressions i.e. “symbolic guess”



# States

- At each point in a trace, **the state abstracts the past computation history** up to that point
- Example: classical **environment** (assigning values to register at each point  $k$  of the trace):

$$\rho^p(\tau, k) \triangleq \lambda \mathbf{r} \in \mathbb{R}(p) \cdot E^p[\mathbf{r}](\tau, k)$$

$$\nu^p(\tau, k) \triangleq \lambda \mathbf{x}_\theta \cdot V_{(32)}^p[\mathbf{x}_\theta](\tau, k)$$

# Prefixes, transitions, ...

- Abstract traces by their prefixes:

$$\overleftarrow{\alpha}(S) \triangleq \bigcup \{ \overleftarrow{\alpha}(\tau) \mid \tau \in S \}$$

$$\overleftarrow{\alpha}(\tau) \triangleq \{ \tau \langle j \rangle \mid j \in [1, 1 + |\tau|] \}$$

$$\tau \langle j \rangle \triangleq \langle \xrightarrow{\overline{\tau}_i} \underline{\tau}_i \mid i \in [1, 1 + j] \rangle$$

- and transitions: extract transitions from traces

⇒ communication fairness is lost, inexact abstraction,

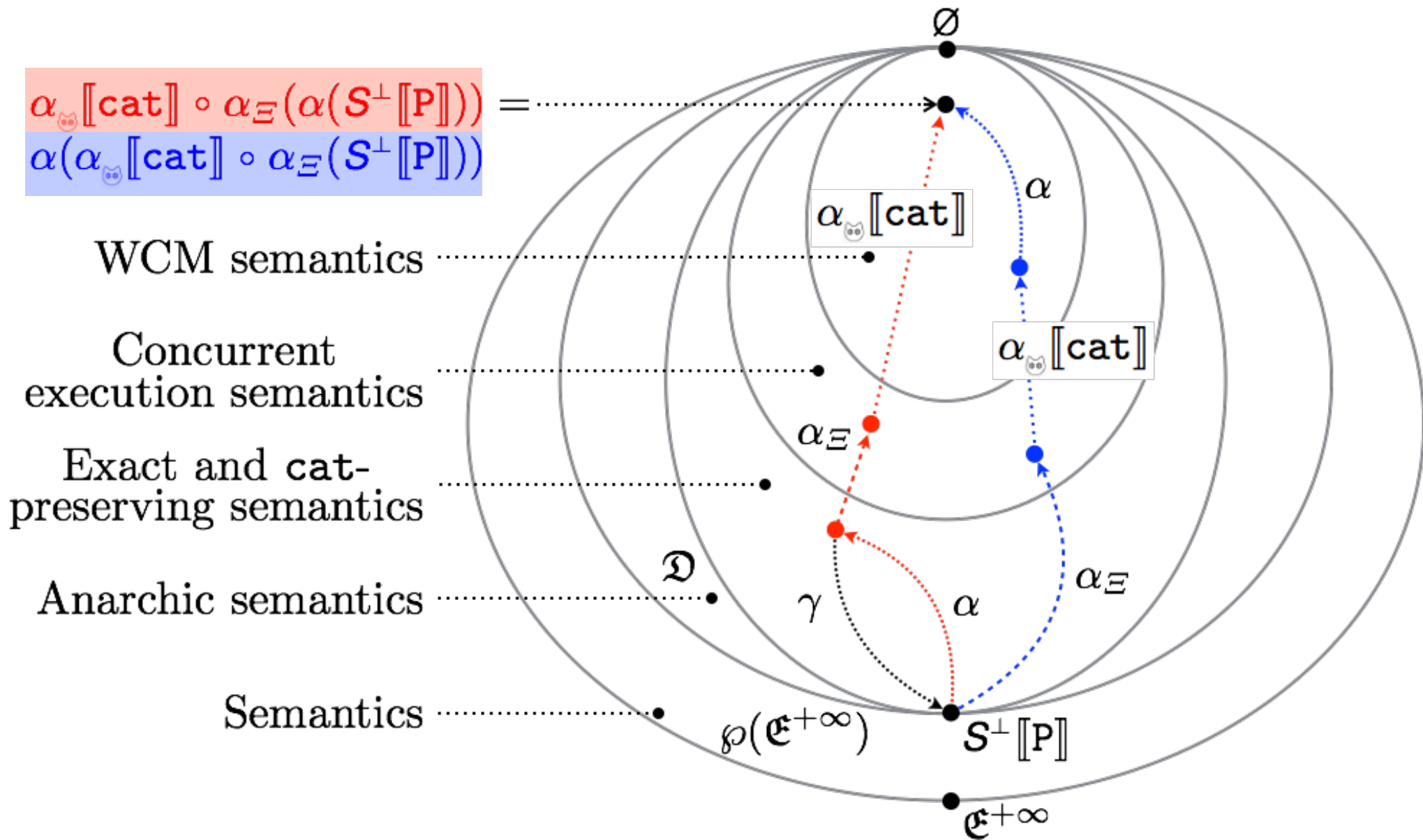
⇒ add fairness condition

⇒ impossible to implement with a scheduler ( $\neq$  process fairness)

# Effect of the cat specification on the hierarchy



# Exactness and cat preservation



# The cat abstraction

- The same cat specification  $\alpha_{\text{cat}} \llbracket \text{cat} \rrbracket$  applies equally to any concurrent execution abstraction  $\alpha_{\Xi}$  of any interleaved/truly parallel semantics in the hierarchy
- The appropriate level of abstraction to specify WCM:
  - No states, only marker (e.g. fence), r, w, rf(w,r) events
  - No values in events
  - No global time (only po order of events per process)
  - Time of communications forgotten (only rf of who communicates with whom)

# Conclusion

# Conclusion

- **Analytic semantics**: a new style of semantics
- The hierarchy of **anarchic semantics** describes the same computations and potential communications in very different styles
- The **cat semantics** restricts communications to a machine/network architecture in the same way for all semantics in the hierarchy
- This idea of **parameterized semantics at various levels of abstraction** is useful for
  - **Verification**
  - **Static analysis**

The End