

« Specification and Abstraction of Semantics »

Patrick Cousot

École normale supérieure
45 rue d'Ulm

75230 Paris cedex 05, France

Patrick.Cousot@ens.fr
www.di.ens.fr/~cousot

Radhia Cousot

CNRS & École polytechnique
Route de Saclay

91128 Palaiseau Cedex, France

Radhia.Cousot@polytechnique.fr
www.polytechnique.edu/Radhia.Cousot

A Tribute Workshop and Festival to Honor Neil D. Jones

Datalogisk Institut, Københavns Universitet, København,
Denmark— 25–26 August, 2007



Contents

Souvenir, Souvenir	3
Specification and abstraction of semantics	
Motivation	9
Bi-inductive structural definitions	13
Example: semantics of the eager λ -calculus	16
Abstraction	47
Conclusion	50



1. Souvenir, Souvenir



Neil D. Jones



*An explorer of automatic semantics-based program
manipulation*



A Long Common Professional Interest and Collaboration

- Semantique I;
- Semantique II;
- Atlantique;
- Daedalus;



Many more shared events

- Århus workshop in 81,
- ...
- POPL'97 in Paris,
- ...
- POPL'04 in Venice
- ...
- Decision to start **ASTRÉE**
- ...
- VMCAI'2009



Happy Souvenirs



Tribute to Neil, København, August 25th, 2007

2. Specification and abstraction of semantics



Motivation



Tribute to Neil, København, August 25th, 2007

Motivation

- We look for a formalism to **specify abstract program semantics**
 - from definitional semantics ...
 - to static program analysis algorithms
 - coping with **termination & non-termination**,
 - handling the many **different styles of presentations** found in the literature (rules, fixpoint, equations, constraints, ...) in a uniform way
- A simple **generalization of inductive definitions** from sets to posets seems adequate.



On the importance of defining both finite and infinite behaviors

- Example of the *choice operator* $E_1 \mid E_2$ where:

$$\begin{array}{l} E_1 \Rightarrow a \quad E_2 \Rightarrow b \quad \text{termination} \\ \text{or} \quad E_1 \Rightarrow \perp \quad E_2 \Rightarrow \perp \quad \text{non-termination} \end{array}$$

- The *finite behavior* of $E_1 \mid E_2$ is:

$$a \mid b \Rightarrow a \quad a \mid b \Rightarrow b \quad .$$



- But for the case $\perp \mid \perp \Rightarrow \perp$, the *infinite behaviors* of $E_1 \mid E_2$ depend on the choice method:

Non-deterministic	Parallel	Eager	Mixed left-to-right	Mixed right-to-left
$\perp \mid b \Rightarrow b$	$\perp \mid b \Rightarrow b$			$\perp \mid b \Rightarrow b$
$\perp \mid b \Rightarrow \perp$		$\perp \mid b \Rightarrow \perp$	$\perp \mid b \Rightarrow \perp$	$\perp \mid b \Rightarrow \perp$
$a \mid \perp \Rightarrow a$	$a \mid \perp \Rightarrow a$		$a \mid \perp \Rightarrow a$	
$a \mid \perp \Rightarrow \perp$		$a \mid \perp \Rightarrow \perp$	$a \mid \perp \Rightarrow \perp$	$a \mid \perp \Rightarrow \perp$

- Nondeterministic: an internal choice is made initially to evaluate E_1 or to evaluate E_2 ;
- Parallel: evaluate E_1 and E_2 concurrently, with an unspecified scheduling, and return the first available result a or b ;
- Mixed left-to-right: evaluate E_1 and then either return its result a or evaluate E_2 and return its result b ;
- Mixed right-to-left: evaluate E_2 and then either return its result b or evaluate E_1 and return its result a ;
- Eager: evaluate both E_1 and E_2 and return either results if both terminate.



Bi-inductive Structural Definitions

Over-simplified for the presentation!



Tribute to Neil, København, August 25th, 2007

Inductive definitions

Set-theoretic [Acz77]

$\langle \wp(\mathcal{U}), \subseteq \rangle$

$\frac{P}{c} \in \mathcal{R} \quad (P \in \wp(\mathcal{U}), c \in \mathcal{U})$

$F(X) \triangleq \left\{ c \mid \exists \frac{P}{c} \in \mathcal{R} : P \subseteq X \right\}$

$\text{lfp}^{\subseteq} F \in \wp(\mathcal{U})$

\subseteq -least $X : F(X) = X$

\subseteq -least $X : F(X) \subseteq X$

$\left\{ \frac{X}{c} \mid X \subseteq \mathcal{U} \wedge c \in F(X) \right\}$

universe

rules

transformer

fixpoint def.

equational def.

constraint def.

rules



Inductive definitions

Set-theoretic [Acz77]

$\langle \wp(\mathcal{U}), \subseteq \rangle$

$\frac{P}{c} \in \mathcal{R} \quad (P \in \wp(\mathcal{U}), c \in \mathcal{U})$

$F(X) \triangleq \left\{ c \mid \exists \frac{P}{c} \in \mathcal{R} : P \subseteq X \right\}$

$\text{lfp}^{\subseteq} F \in \wp(\mathcal{U})$

\subseteq -least $X : F(X) = X$

\subseteq -least $X : F(X) \subseteq X$

$\left\{ \frac{X}{c} \mid X \subseteq \mathcal{U} \wedge c \in F(X) \right\}$

Order-theoretic

$\langle \mathcal{D}, \sqsubseteq \rangle$

$\frac{P}{C} \in \mathcal{R} \quad (P, C \in \mathcal{D})$

$F(X) \triangleq \bigsqcup \left\{ C \mid \exists \frac{P}{c} \in \mathcal{R} : P \subseteq X \right\}$

$\text{lfp}^{\sqsubseteq} F \in \mathcal{D}$

\sqsubseteq -least $X : F(X) = X$

\sqsubseteq -least $X : F(X) \sqsubseteq X$

$\left\{ \frac{X}{F(X)} \mid X \in \mathcal{D} \right\}$

universe

rules

transformer

fixpoint def.

equational def.

constraint def.

rules



Semantics of the Eager λ -calculus



Tribute to Neil, København, August 25th, 2007

Syntax



Tribute to Neil, København, August 25th, 2007

Syntax of the Eager λ -calculus

$x, y, z, \dots \in \mathbb{X}$	variables
$c \in \mathbb{C}$	constants ($\mathbb{X} \cap \mathbb{C} = \emptyset$)
$c ::= 0 \mid 1 \mid \dots$	
$v \in \mathbb{V}$	values
$v ::= c \mid \lambda x. a$	
$e \in \mathbb{E}$	errors
$e ::= c a \mid e a$	
$a, a', a_1, \dots, b, \dots \in \mathbb{T}$	terms
$a ::= x \mid v \mid a a'$	



Trace Semantics



Tribute to Neil, København, August 25th, 2007

Example I: Finite Computation

	function	argument
	$((\lambda x. x x) (\lambda y. y))$	$((\lambda z. z) 0)$
→		evaluate function
	$((\lambda y. y) (\lambda y. y))$	$((\lambda z. z) 0)$
→		evaluate function, cont'd
	$(\lambda y. y)$	$((\lambda z. z) 0)$
→		evaluate argument
	$(\lambda y. y)$	0
→		apply function to argument
	0	<i>a value!</i>



Example II: Infinite Computation

function argument

$(\lambda x. x x) (\lambda x. x x)$

→ apply function to argument

$(\lambda x. x x) (\lambda x. x x)$

→ apply function to argument

$(\lambda x. x x) (\lambda x. x x)$

→ apply function to argument

... *non termination!*



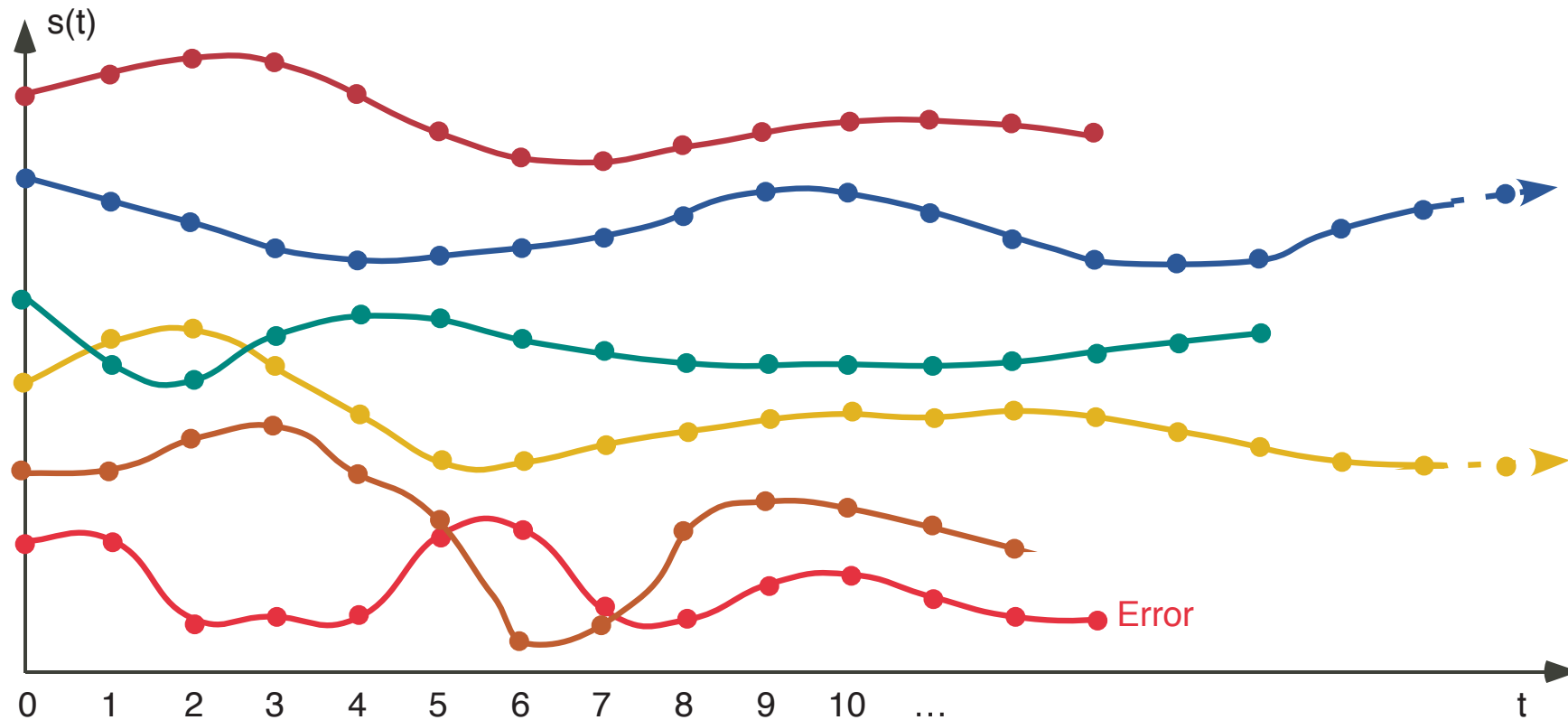
Example III: Erroneous Computation

	function	argument
	$((\lambda x. x x) ((\lambda z. z) 0))$	$((\lambda y. y) 0)$
→		evaluate argument
	$((\lambda x. x x) ((\lambda z. z) 0)) 0$	
→		evaluate function
	$((\lambda x. x x) 0) 0$	
→		evaluate function, cont'd
	$(0 0) 0$	

a runtime error!



Finite, Infinite and Erroneous Trace Semantics



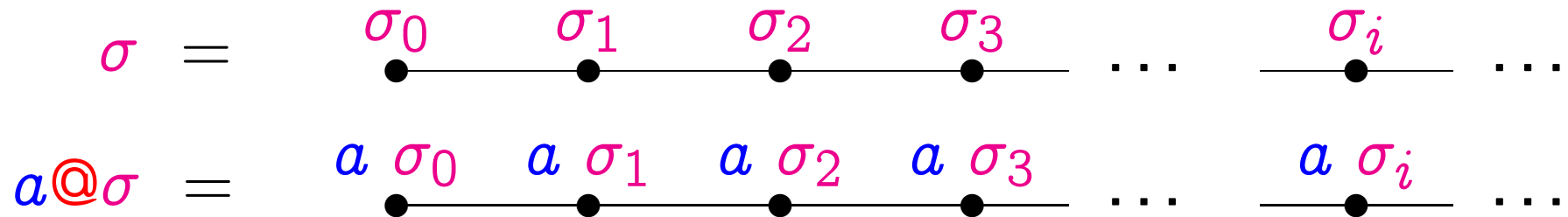
Traces

- \mathbb{T}^* (resp. \mathbb{T}^+ , \mathbb{T}^ω , \mathbb{T}^α and \mathbb{T}^∞) be the set of finite (resp. nonempty finite, infinite, finite or infinite, and nonempty finite or infinite) sequences of terms
- ϵ is the empty sequence $\epsilon \bullet \sigma = \sigma \bullet \epsilon = \sigma$.
- $|\sigma| \in \mathbb{N} \cup \{\omega\}$ is the length of $\sigma \in \mathbb{T}^\alpha$. $|\epsilon| = 0$.
- If $\sigma \in \mathbb{T}^+$ then $|\sigma| > 0$ and $\sigma = \sigma_0 \bullet \sigma_1 \bullet \dots \bullet \sigma_{|\sigma|-1}$.
- If $\sigma \in \mathbb{T}^\omega$ then $|\sigma| = \omega$ and $\sigma = \sigma_0 \bullet \dots \bullet \sigma_n \bullet \dots$



Operations on Traces

- For $a \in \mathbb{T}$ and $\sigma \in \mathbb{T}^\infty$, we define $a@ \sigma$ to be $\sigma' \in \mathbb{T}^\infty$ such that $\forall i < |\sigma| : \sigma'_i = a \sigma_i$



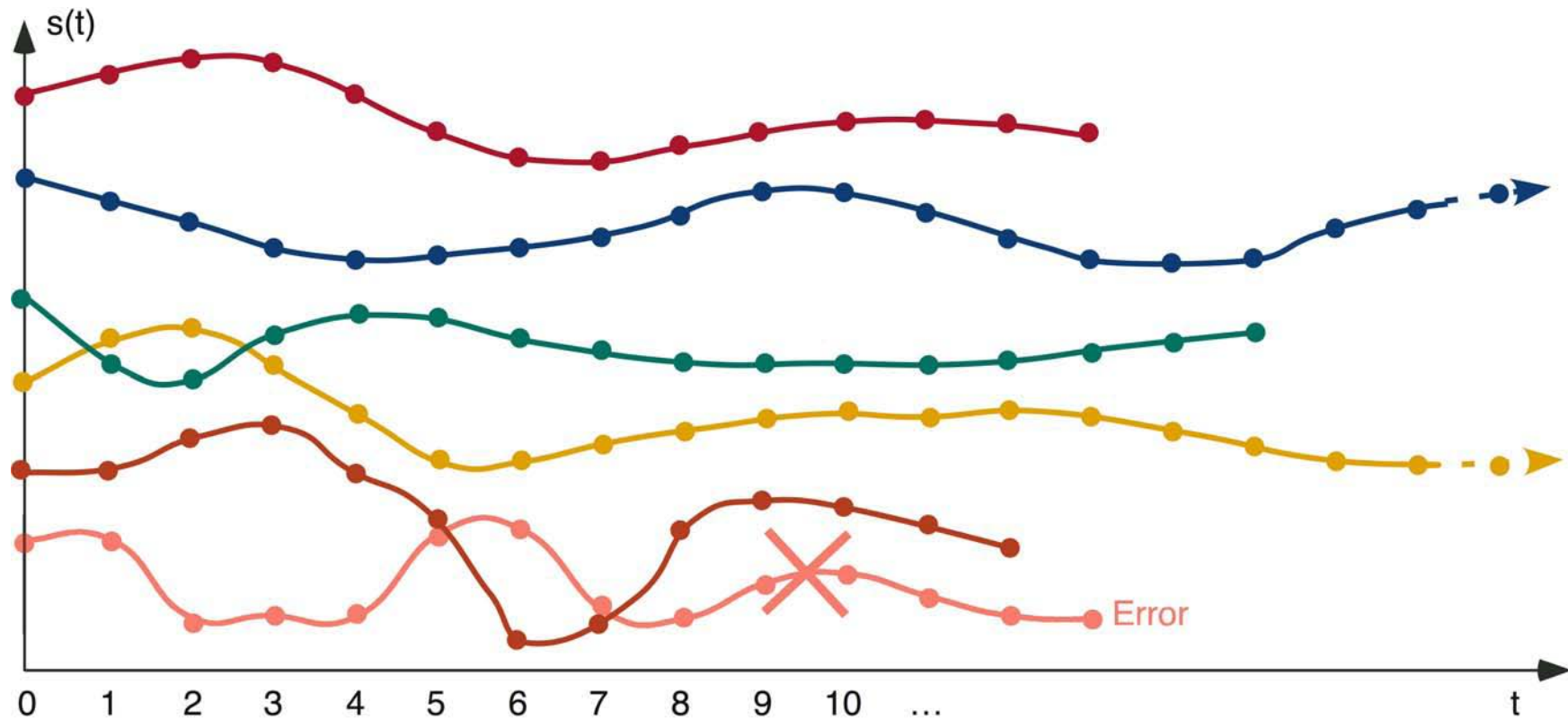
Operations on Traces (Cont'd)

- Similarly for $a \in \mathbb{T}$ and $\sigma \in \mathbb{T}^\infty$, $\sigma@a$ is σ' where $\forall i < |\sigma| : \sigma'_i = \sigma_i a$

$$\begin{array}{l} \sigma = \\ \sigma@a = \end{array} \begin{array}{ccccccc} \sigma_0 & \sigma_1 & \sigma_2 & \sigma_3 & \dots & \sigma_i & \dots \\ \bullet & \bullet & \bullet & \bullet & \dots & \bullet & \dots \\ \hline \sigma_0 a & \sigma_1 a & \sigma_2 a & \sigma_3 a & \dots & \sigma_i a & \dots \\ \bullet & \bullet & \bullet & \bullet & \dots & \bullet & \dots \\ \hline \end{array}$$



Finite and Infinite Trace Semantics \mathbb{S}



Tribute to Neil, København, August 25th, 2007

The Computational Lattice

Given $S, T \in \wp(\mathbb{T}^\infty)$, we define

- $S^+ \triangleq S \cap \mathbb{T}^+$ finite traces
- $S^\omega \triangleq S \cap \mathbb{T}^\omega$ infinite traces
- $S \sqsubseteq T \triangleq S^+ \subseteq T^+ \wedge S^\omega \supseteq T^\omega$ computational order
- $\langle \wp(\mathbb{T}^\infty), \sqsubseteq, \mathbb{T}^\omega, \mathbb{T}^+, \sqcup, \sqcap \rangle$ is a complete lattice



Bifinitary Trace Semantics $\vec{\mathbb{S}}$ of the Eager λ -calculus¹

$$\frac{v \in \vec{\mathbb{S}}, v \in \mathbb{V} \quad \frac{a[x \leftarrow v] \bullet \sigma \in \vec{\mathbb{S}}}{(\lambda x \cdot a) v \bullet a[x \leftarrow v] \bullet \sigma \in \vec{\mathbb{S}}} \sqsubseteq, v \in \mathbb{V}}$$

$$\frac{\sigma \in \vec{\mathbb{S}}^\omega \quad \frac{\sigma \bullet v \in \vec{\mathbb{S}}^+, (a v) \bullet \sigma' \in \vec{\mathbb{S}}}{(a@ \sigma) \bullet (a v) \bullet \sigma' \in \vec{\mathbb{S}}} \sqsubseteq, v, a \in \mathbb{V}}{a@ \sigma \in \vec{\mathbb{S}}} \sqsubseteq, a \in \mathbb{V}$$

$$\frac{\sigma \in \vec{\mathbb{S}}^\omega \quad \frac{\sigma \bullet v \in \vec{\mathbb{S}}^+, (v b) \bullet \sigma' \in \vec{\mathbb{S}}}{(\sigma@ b) \bullet (v b) \bullet \sigma' \in \vec{\mathbb{S}}} \sqsubseteq, v \in \mathbb{V}}{\sigma@ b \in \vec{\mathbb{S}}} \sqsubseteq$$

¹ Note: $a[x \leftarrow b]$ is the capture-avoiding substitution of b for all free occurrences of x within a . We let $\text{FV}(a)$ be the free variables of a . We define the call-by-value semantics of closed terms (without free variables) $\overline{\mathbb{T}} \triangleq \{a \in \mathbb{T} \mid \text{FV}(a) = \emptyset\}$.



Fixpoint big-step maximal trace semantics

The bifinitary trace semantics is

$$\vec{S} = \text{lfp}^{\sqsubseteq} \vec{F}$$

where $\vec{F} \in \wp(\overline{\mathbb{T}}^\infty) \mapsto \wp(\overline{\mathbb{T}}^\infty)$ is

$$\begin{aligned} \vec{F}(S) \triangleq & \{v \in \overline{\mathbb{T}}^\infty \mid v \in \mathbb{V}\} \cup \\ & \{(\lambda x \cdot a) v \cdot a[x \leftarrow v] \cdot \sigma \mid v \in \mathbb{V} \wedge a[x \leftarrow v] \cdot \sigma \in S\} \cup \\ & \{\sigma @ b \mid \sigma \in S^\omega\} \cup \\ & \{(\sigma @ b) \cdot (v b) \cdot \sigma' \mid \sigma \neq \epsilon \wedge \sigma \cdot v \in S^+ \wedge v \in \mathbb{V} \wedge (v b) \cdot \sigma' \in S\} \cup \\ & \{a @ \sigma \mid a \in \mathbb{V} \wedge \sigma \in S^\omega\} \cup \\ & \{(a @ \sigma) \cdot (a v) \cdot \sigma' \mid a, v \in \mathbb{V} \wedge \sigma \neq \epsilon \wedge \sigma \cdot v \in S^+ \wedge (a v) \cdot \sigma' \in S\} . \end{aligned}$$

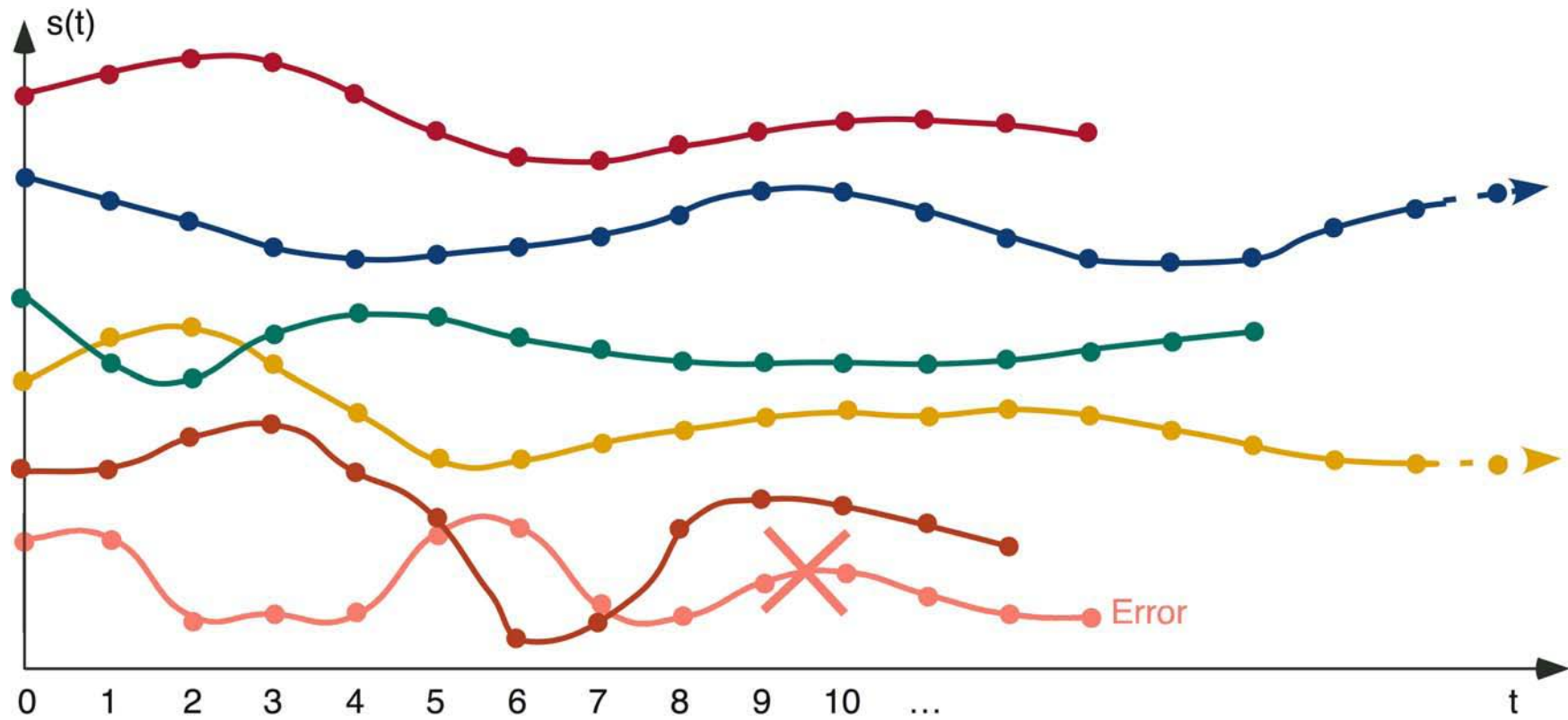


Relational Semantics

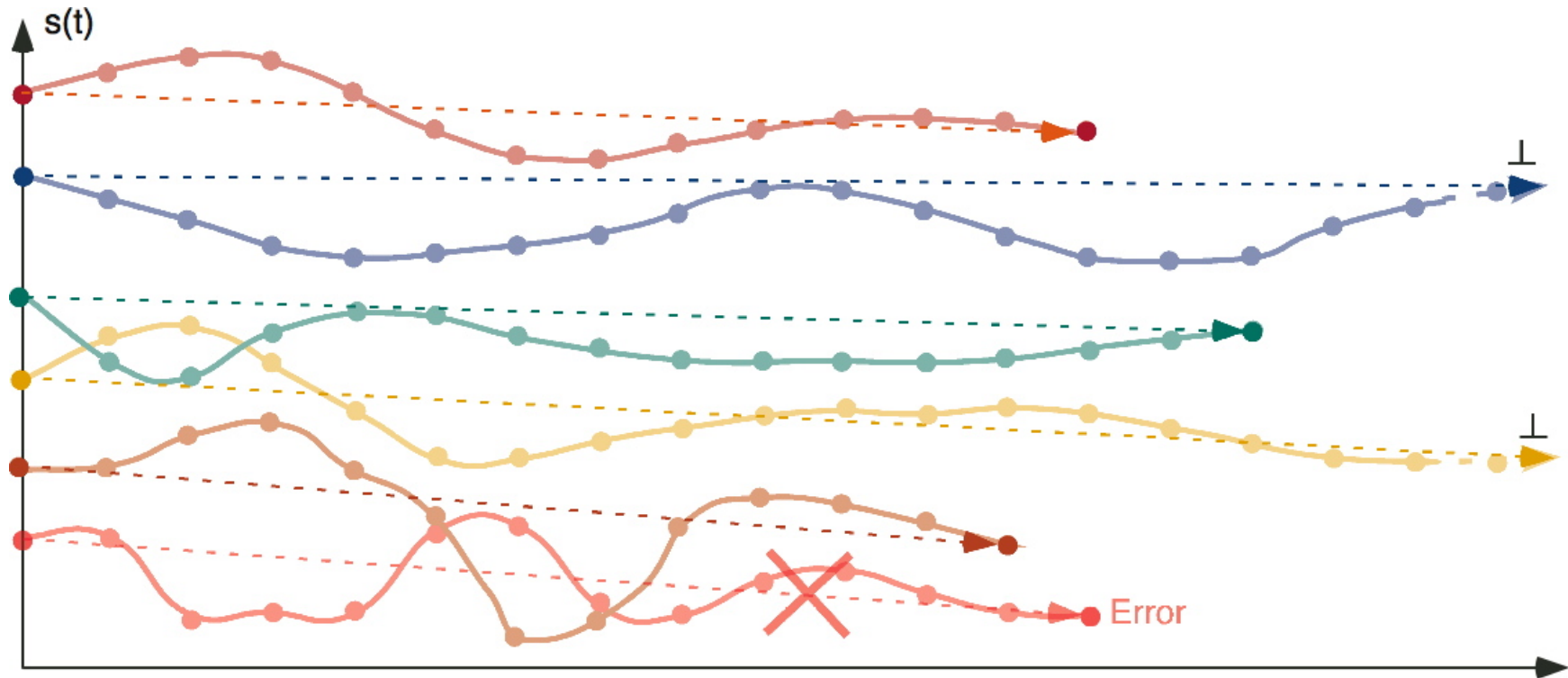


Tribute to Neil, København, August 25th, 2007

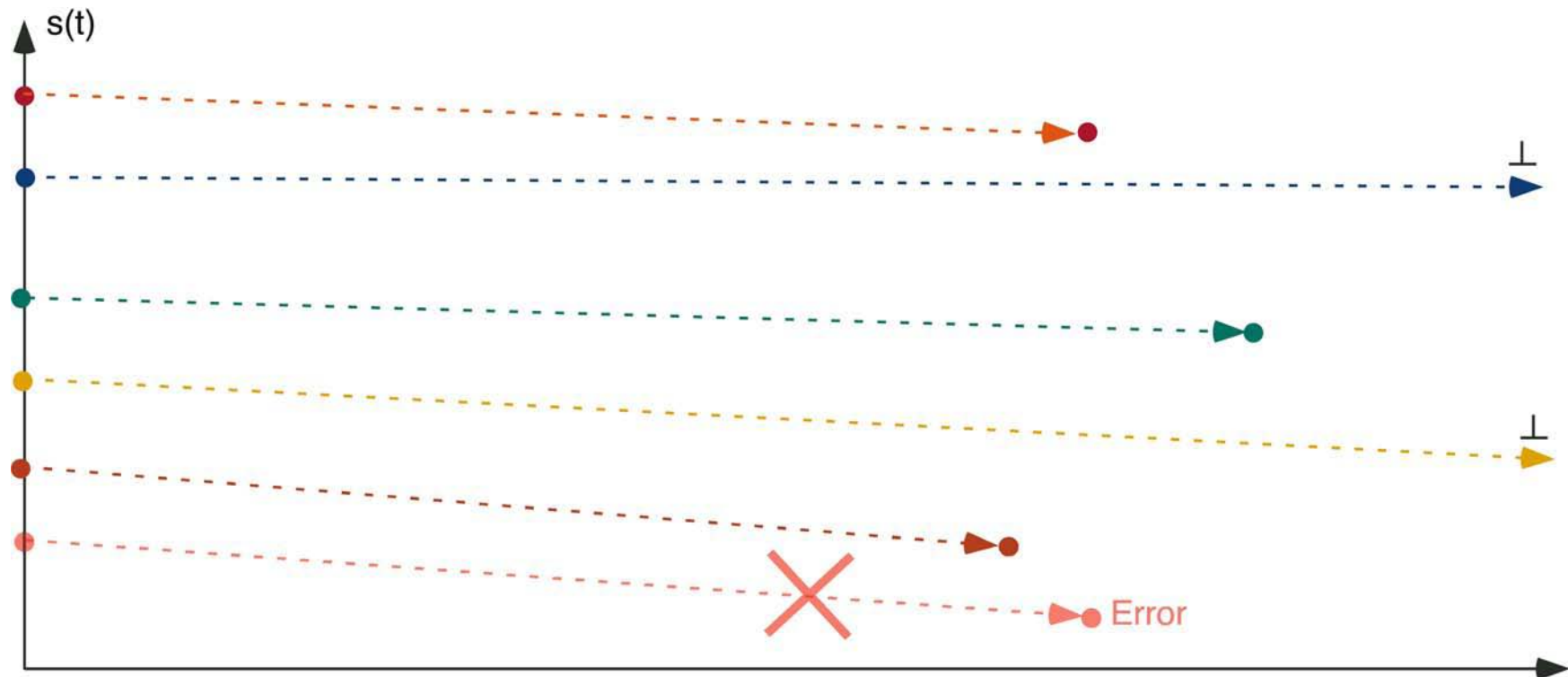
Trace Semantics



Relational Semantics = α (Trace Semantics)



Relational Semantics



Abstraction to the Bifinitary Relational Semantics of the Eager λ -calculus

remember the input/output behaviors,
forget about the intermediate computation steps

$$\alpha(T) \stackrel{\text{def}}{=} \{\alpha(\sigma) \mid \sigma \in T\}$$

$$\alpha(\sigma_0 \bullet \sigma_1 \bullet \dots \bullet \sigma_n) \stackrel{\text{def}}{=} \sigma_0 \Longrightarrow \sigma_n$$

$$\alpha(\sigma_0 \bullet \dots \bullet \sigma_n \bullet \dots) \stackrel{\text{def}}{=} \sigma_0 \Longrightarrow \perp$$



Bifinitary Relational Semantics of the Eager λ -calculus

$$\begin{array}{c}
 v \Rightarrow v, \quad v \in \mathbb{V} \\
 \hline
 a \Rightarrow \perp \\
 \hline
 a \ b \Rightarrow \perp \quad \sqsubseteq
 \end{array}
 \qquad
 \begin{array}{c}
 b \Rightarrow \perp \\
 \hline
 a \ b \Rightarrow \perp \quad \sqsubseteq, \quad a \in \mathbb{V}
 \end{array}$$

$$\begin{array}{c}
 a[x \leftarrow v] \Rightarrow r \\
 \hline
 (\lambda x \cdot a) \ v \Rightarrow r \quad \sqsubseteq, \quad v \in \mathbb{V}, \ r \in \mathbb{V} \cup \{\perp\}
 \end{array}$$

$$\begin{array}{c}
 a \Rightarrow v, \quad v \ b \Rightarrow r \\
 \hline
 a \ b \Rightarrow r \quad \sqsubseteq, \quad v \in \mathbb{V}, \ r \in \mathbb{V} \cup \{\perp\}
 \end{array}$$

$$\begin{array}{c}
 b \Rightarrow v, \quad a \ v \Rightarrow r \\
 \hline
 a \ b \Rightarrow r \quad \sqsubseteq, \quad a \in \mathbb{V}, \ v \in \mathbb{V}, \ r \in \mathbb{V} \cup \{\perp\}.
 \end{array}$$



On the computational ordering \sqsubseteq

- For the bifinitary trace semantics \mathcal{S} , we could replace the computational ordering \sqsubseteq by \supseteq (thus taking **greatest fixpoints** for \sqsubseteq);
- **Impossible** for the bifinitary relational semantics!
- Counter-example: the greatest fixpoint starts by assuming that we have the terminating execution

$$(\lambda x. x x)(\lambda x. x x) \implies (\lambda x. x x)(\lambda x. x x)$$

then the call rule $\frac{a[x \leftarrow v] \implies r}{(\lambda x. a) v \implies r} \Xi$, $v \in \mathbb{V}$, $r \in \mathbb{V} \cup \{\perp\}$ will preserve this invalid hypothesis!

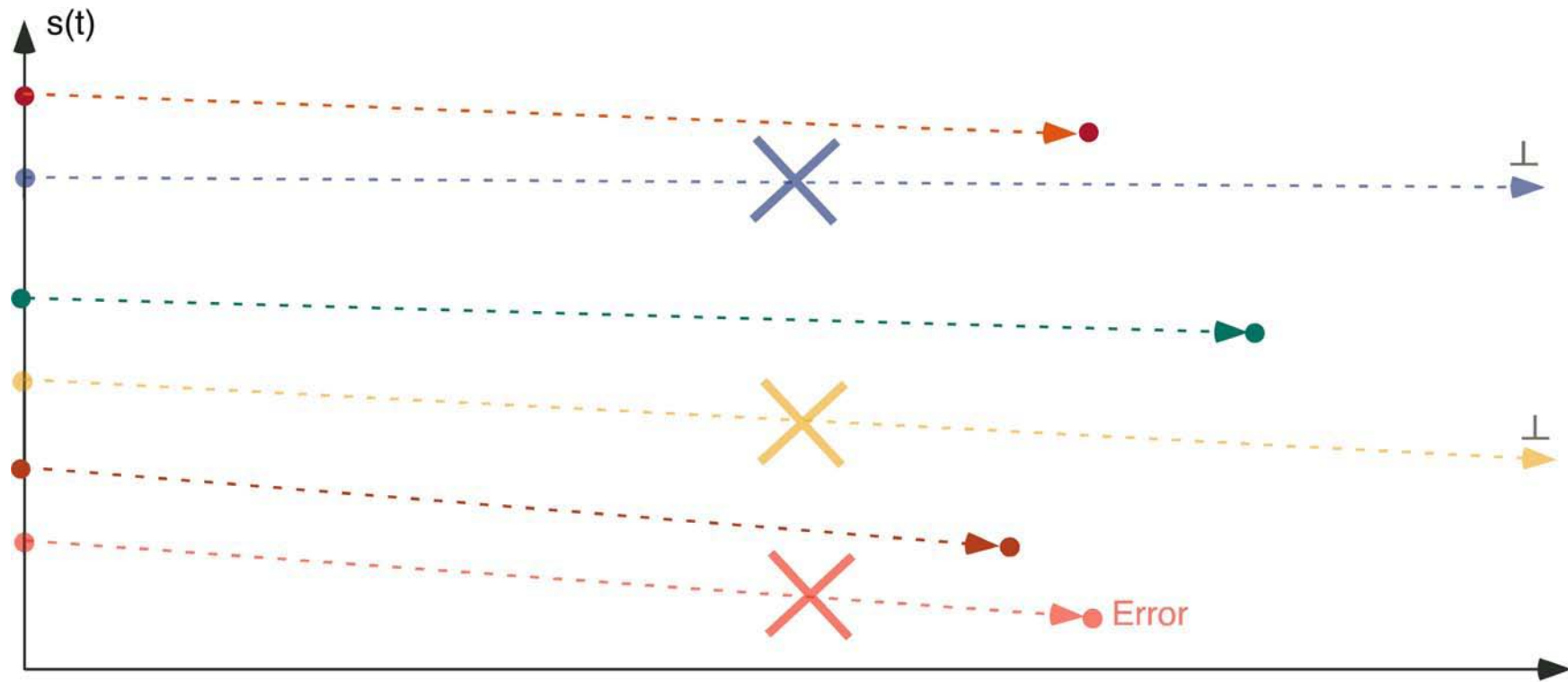


Natural Semantics



Tribute to Neil, København, August 25th, 2007

Natural Semantics = α (Relational Semantics)



Abstraction to the Natural Big-Step Semantics of the Eager λ -calculus

remember the finite input/output behaviors,
forget about non-termination

$$\alpha(T) \stackrel{\text{def}}{=} \bigcup \{ \alpha(\sigma) \mid \sigma \in T \}$$

$$\alpha(\sigma_0 \Rightarrow \sigma_n) \stackrel{\text{def}}{=} \{ \sigma_0 \Rightarrow \sigma_n \}$$

$$\alpha(\sigma_0 \Rightarrow \perp) \stackrel{\text{def}}{=} \emptyset$$



Natural Big-Step Semantics of the Eager λ -calculus [Kah88]

$$v \Rightarrow v, \quad v \in \mathbb{V}$$

$$\frac{a[x \leftarrow v] \Rightarrow r}{(\lambda x \cdot a) \quad v \Rightarrow r} \subseteq, \quad v \in \mathbb{V}, r \in \mathbb{V}$$

$$\frac{a \Rightarrow v, \quad v \quad b \Rightarrow r}{a \quad b \Rightarrow r} \subseteq, \quad v \in \mathbb{V}, r \in \mathbb{V}$$

$$\frac{b \Rightarrow v, \quad a \quad v \Rightarrow r}{a \quad b \Rightarrow r} \subseteq, \quad a \in \mathbb{V}, v \in \mathbb{V}, r \in \mathbb{V}.$$

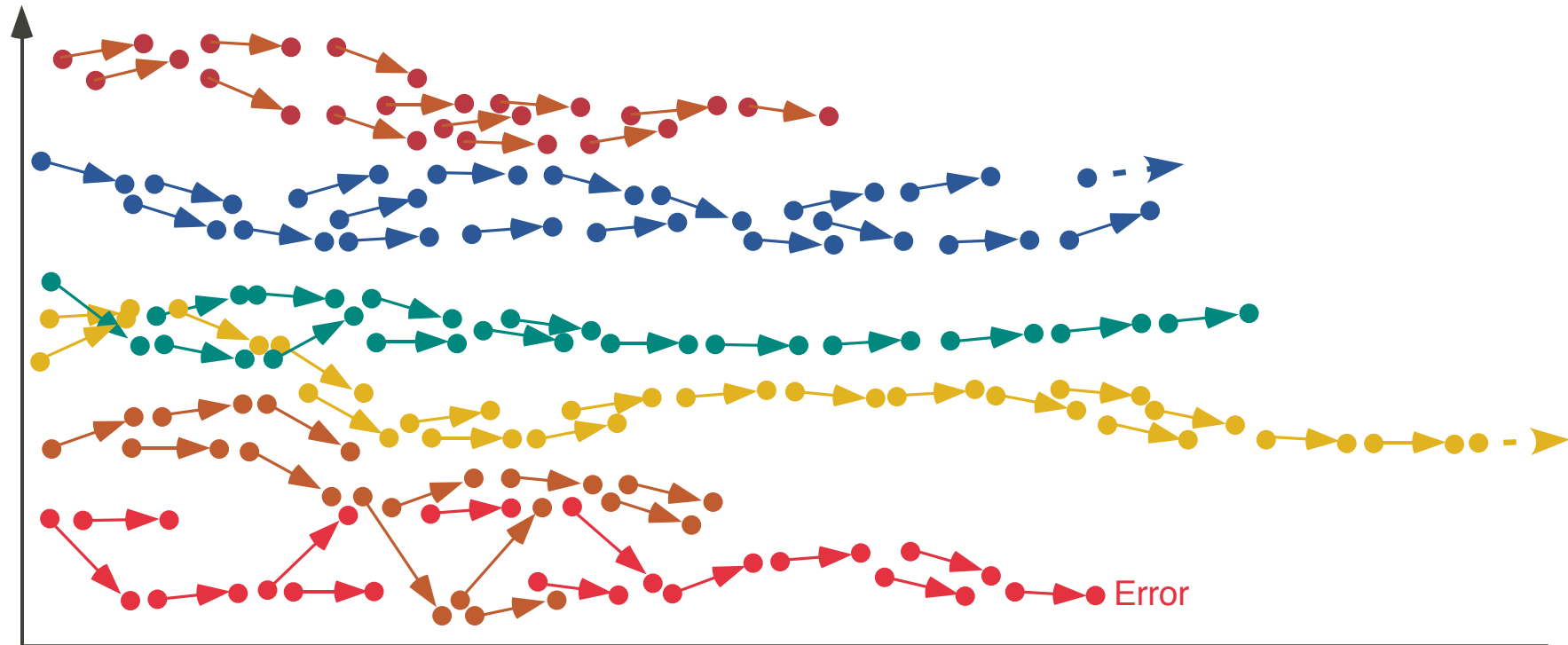


Transition Semantics



Tribute to Neil, København, August 25th, 2007

Transition Semantics = α (Trace Semantics)



Abstraction to the Transition Semantics of the Eager λ -calculus

remember execution steps,
forget about their sequencing

$$\alpha(T) \stackrel{\text{def}}{=} \bigcup \{ \alpha(\sigma) \mid \sigma \in T \}$$

$$\alpha(\sigma_0 \bullet \sigma_1 \bullet \dots \bullet \sigma_n) \stackrel{\text{def}}{=} \{ \sigma_i \longrightarrow \sigma_{i+1} \mid 0 \leq i \wedge i < n \}$$

$$\alpha(\sigma_0 \bullet \dots \bullet \sigma_n \bullet \dots) \stackrel{\text{def}}{=} \{ \sigma_i \longrightarrow \sigma_{i+1} \mid i \geq 0 \}$$



Transition Semantics of the Eager λ -calculus [Plo81]

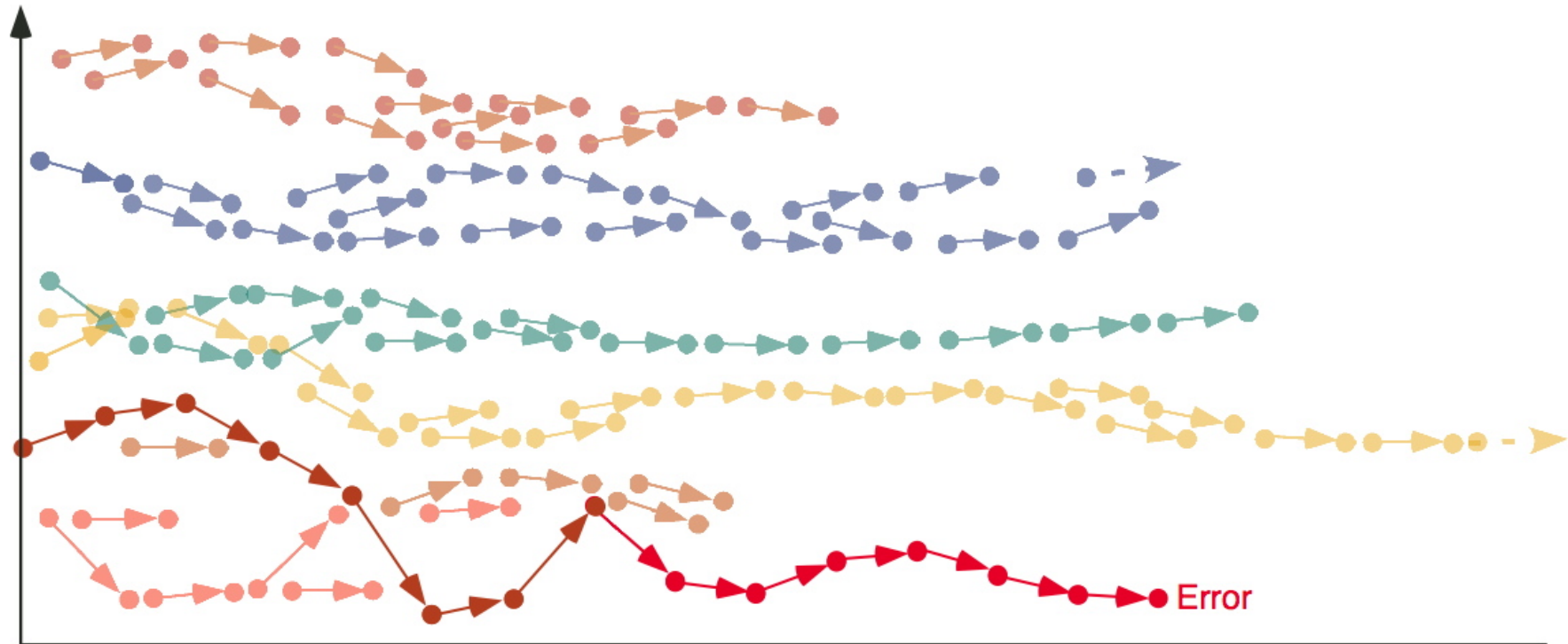
$$((\lambda x \cdot a) v) \longrightarrow a[x \leftarrow v]$$

$$\frac{a_0 \longrightarrow a_1}{a_0 b \longrightarrow a_1 b} \subseteq$$

$$\frac{b_0 \longrightarrow b_1}{v b_0 \longrightarrow v b_1} \subseteq .$$



Approximation



$$\begin{aligned}
 & ((\lambda x \cdot x \ x) ((\lambda z \cdot z) 0)) (\lambda y \cdot y) \rightarrow ((\lambda x \cdot x \ x) 0) (\lambda y \cdot y) \\
 & \rightarrow (0 \ 0) (\lambda y \cdot y) \quad \text{an error!}
 \end{aligned}$$



Abstraction



Tribute to Neil, København, August 25th, 2007

Kleenean abstraction

- $\langle \mathcal{D}, \sqsubseteq, \perp, \sqcup \rangle, \langle \mathcal{D}^\#, \sqsubseteq^\#, \perp^\#, \sqcup^\# \rangle$ dcpos
- $F \in \mathcal{D} \mapsto \mathcal{D}, F^\# \in \mathcal{D}^\# \mapsto \mathcal{D}^\#$ monotone
- $\alpha \in \mathcal{D} \mapsto \mathcal{D}^\#$ strict and continuous on chains of \mathcal{D}
- $\alpha \circ F = F^\# \circ \alpha$, commutation condition
 $\implies \alpha(\text{lfp}^{\sqsubseteq} F) = \text{lfp}^{\sqsubseteq^\#} F^\#$

OK for abstracting finite behaviors, not infinite ones



Tarskian abstraction

- $\langle \mathcal{D}, \sqsubseteq, \perp, \sqcup \rangle, \langle \mathcal{D}^\#, \sqsubseteq^\#, \perp^\#, \sqcup^\# \rangle$ dcpos
- $F \in \mathcal{D} \mapsto \mathcal{D}, F^\# \in \mathcal{D}^\# \mapsto \mathcal{D}^\#$ monotone
- $\alpha \in \mathcal{D} \mapsto \mathcal{D}^\#$ preserves meets
- $F^\# \circ \alpha \sqsubseteq^\# \alpha \circ F$, semi-commutation condition
- $\forall y \in \mathcal{D}^\# : (F^\#(y) \sqsubseteq^\# y) \implies (\exists x \in \mathcal{D} : \alpha(x) = y \wedge F(x) \sqsubseteq x)$
 $\implies \alpha(\text{lfp}^{\sqsubseteq} F) = \text{lfp}^{\sqsubseteq^\#} F^\#$

OK for abstracting infinite behaviors, not finite ones
 \implies abstract by parts.



Conclusion



Conclusion

- Both **finite and infinite semantics** are needed in **static analysis** (such as strictness, [Myc80]), typing [Cou97, Ler06], etc;
- Such static analyzes must be proved correct with respect to a semantics chosen at an **various level of abstraction** (small-step/big-step trace/relational/natural semantics);
- Static analyzes use various **equivalent presentations** (fixpoints, equational, constraints and inference rules)
- The bifinite extension of SOS *might* satisfy these needs.



THE END, THANK YOU

**Neil, for such a long friendship and
cooperation**

**Best wishes for your new constraintless
research career**



Bibliography

- [Acz77] P. Aczel. An introduction to inductive definitions. In J. Barwise, editor, *Handbook of Mathematical Logic*, volume 90 of *Studies in Logic and the Foundations of Mathematics*, pages 739–782. Elsevier, 1977.
- [Cou97] P. Cousot. Types as abstract interpretations, invited paper. In *24th POPL*, pages 316–331, Paris, FR, Jan. 1997. ACM Press.
- [Kah88] G. Kahn. Natural semantics. In K. Fuchi and M. Nivat, editors, *Programming of Future Generation Computers*, pages 237–258. Elsevier, 1988.
- [Ler06] X. Leroy. Coinductive big-step operational semantics. In P. Sestoft, editor, *Proc. 15th ESOP '2006*, Vienna, AT, LNCS 3924, pages 54–68. Springer, 27–28 Mar. 2006.
- [Myc80] A. Mycroft. The theory and practice of transforming call-by-need into call-by-value. In B. Robinet, editor, *Proc. 4th Int. Symp. on Programming*, Paris, FR, 22–24 Apr. 1980, LNCS 83, pages 270–281. Springer, 1980.
- [Plo81] G.D. Plotkin. A structural approach to operational semantics. Technical Report DAIMI FN-19, Aarhus University, DK, Sep. 1981.

