

# On operational properties of quantitative extensions of $\lambda$ -calculus

Michele ALBERTI

Aix-Marseille Université  
Institut de Mathématiques de Marseille



Università di Bologna  
Dipartimento di Informatica



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

Marseille – December 5th, 2014

# Semantics

- Reasoning about program **behaviours** and **properties**,

## Operational Semantics

- Reasoning about program **behaviours** and **properties**,
- **Abstract** descriptions of program **executions**,

## Quantitative Operational Semantics

- Reasoning about program **behaviours** and **properties**,
- **Abstract** descriptions of program **executions**,
- Provide/Handle **fine-grained information** about computation.

## Some Motivations

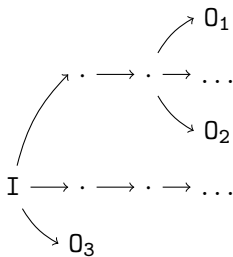
- **New computational models** with quantitative aspects,

Room for **fundamental**, **operationally based** investigations.

- Most of research focus is on:
  - Denotational semantics of programming languages,
  - Programming constructs and applications.

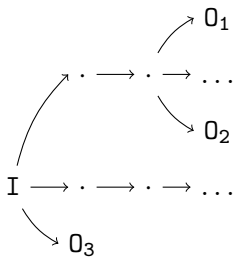
# Qualitative vs Quantitative Semantics

Non-deterministic computation:



# Qualitative vs Quantitative Semantics

Non-deterministic computation:

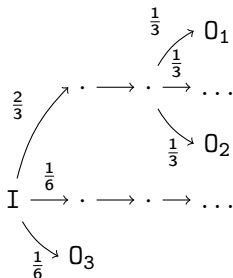


Qualitative is about what.

## Qualitative vs Quantitative Semantics

Quantitative semantics as a **refinement** of qualitative one.

**Probabilistic** computation:

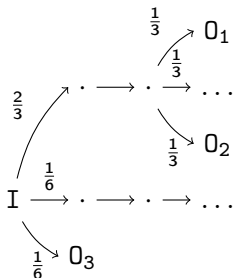




## Qualitative vs Quantitative Semantics

Quantitative semantics as a **refinement** of qualitative one.

**Probabilistic** computation:



Quantitative is about **how {much, many}**.

## Contributions: Algebraic $\lambda$ -calculus

- $\Lambda_{\Sigma}$  : untyped  $\Lambda$  with finite  $\mathbb{R}$ -valued linear combinations,
- Purpose:
  - Framework for {quantum, probabilistic, *etc.* } operational theory.

## Contributions: Algebraic $\lambda$ -calculus

- $\Lambda_{\Sigma}$  : untyped  $\Lambda$  with finite  $\mathbb{R}$ -valued linear combinations,
- Purpose:
  - Framework for {quantum, probabilistic, *etc.*} operational theory.
- **No normal forms** in presence of  $-1 \in \mathbb{R}$ 
  - No irreducible terms,
  - Every term reduces to every other term (**collapse** of  $\equiv$ ).

## Contributions: Algebraic $\lambda$ -calculus

- $\Lambda_{\Sigma}$  : untyped  $\Lambda$  with finite  $\mathbb{R}$ -valued linear combinations,
- Purpose:
  - Framework for {quantum, probabilistic, etc. } operational theory.
- **No normal forms** in presence of  $-1 \in \mathbb{R}$ 
  - No irreducible terms,
  - Every term reduces to every other term (**collapse** of  $\equiv$ ).

### Contributions:

- Normalisation scheme for the whole (**untyped**)  $\Lambda_{\Sigma}$   
(previously,  $V_{aux}$  for a simply typed setting)
  - First notion of normal form for  $\Lambda_{\Sigma}$ ,
  - $\cong$  does not collapse,
  - Reduction relation characterising these normal forms.
- Factorisation theorem.

## Contributions: Probabilistic $\lambda$ -calculus

- $\Lambda_{\oplus}$ : untyped  $\Lambda$  with probabilistic choice operator,
- Purpose:
  - Program equality in H-O probabilistic functional languages.

## Contributions: Probabilistic $\lambda$ -calculus

- $\Lambda_{\oplus}$ : untyped  $\Lambda$  with probabilistic choice operator,
- Purpose:
  - Program equality in H-O probabilistic functional languages.
- Bisimulation for proving observational equivalence
  - Bisimilarity need to be a **congruence** relation,
  - **Sound** w.r.t observational equivalence.

## Contributions: Probabilistic $\lambda$ -calculus

- $\Lambda_{\oplus}$ : untyped  $\Lambda$  with probabilistic choice operator,
- Purpose:
  - Program equality in H-O probabilistic functional languages.
- Bisimulation for proving observational equivalence
  - Bisimilarity need to be a **congruence** relation,
  - **Sound** w.r.t observational equivalence.

### Contributions:

- Probabilistic applicative bisimulation for  $\Lambda_{\oplus}$ ,
- Proof of congruence for probabilistic applicative bisimilarity
  - Along the lines of Howe's method,
  - Non-trivial “disentangling” properties for sets of real numbers.
- Full abstraction on pure  $\lambda$ -terms.

Part I

# Normal forms for the algebraic $\lambda$ -calculus



## Definition

$$M, N ::= x \mid \lambda x.M \mid (M)N \mid \sum_{i=1}^n a_i M_i$$

subject to (algebraic) linearity:

1.  $\lambda x. (\sum_{i=1}^n a_i M_i) = \sum_{i=1}^n a_i \lambda x. M_i$
2.  $(\sum_{i=1}^n a_i M_i) N = \sum_{i=1}^n a_i (M_i) N$

Definition ( $\Lambda_\Sigma$ ) $\{\text{Simple terms: } \Delta_R\} \quad s, t ::= x \mid \lambda x. s \mid (s) T$  $\{\text{Terms: } R\langle\Delta_R\rangle\} \quad S, T ::= \sum_{i=1}^n a_i s_i$ 

where  $a_i \in R$  and  $R$  a semiring.

Definition ( $\Lambda_\Sigma$ ) $\{\text{Simple terms: } \Delta_R\} \quad s, t ::= x \mid \lambda x.s \mid (s) T$  $\{\text{Terms: } R\langle\Delta_R\rangle\} \quad S, T ::= \sum_{i=1}^n a_i s_i$ 

where  $a_i \in R$  and  $R$  a semiring.

$\beta$ -reduction ( $\widetilde{\rightarrow}$ ):

- **revised  $\beta$ -rule:**  $(\lambda x.s) T \rightarrow s [T/x]$ ,
- extended to linear combinations:

$$au + V \widetilde{\rightarrow} aU' + V \text{ as soon as } a \neq 0 \text{ and } u \rightarrow U'.$$

Definition ( $\Lambda_\Sigma$ ) $\{\text{Simple terms: } \Delta_R\} \quad s, t ::= x \mid \lambda x.s \mid (s) T$  $\{\text{Terms: } R\langle\Delta_R\rangle\} \quad S, T ::= \sum_{i=1}^n a_i s_i$ 

where  $a_i \in R$  and  $R$  a semiring.

$\beta$ -reduction ( $\rightsquigarrow$ ):

- **revised  $\beta$ -rule:**  $(\lambda x.s) T \rightarrow s [T/x]$ ,
- extended to linear combinations:

$$au + V \rightsquigarrow aU' + V \text{ as soon as } a \neq 0 \text{ and } u \rightarrow U'.$$

## Theorem ([Vau09])

Reduction  $\rightsquigarrow$  enjoys confluence.

## Factorisation in Pure $\lambda$ -calculus

**Head**  $\beta$ -reduction as the **essential** part of a computation.

Theorem ([Tak95])

*Any reduction  $s \rightarrow^* t$  can be reorganised so that  $s \rightarrow_h^* \rightarrow_{-h}^* t$ .*

**Takahashi's proof:** **decompose**  $\Rightarrow$  and then **swap** reductions.

# Factorisation in Pure $\lambda$ -calculus

**Head**  $\beta$ -reduction as the **essential** part of a computation.

Theorem ([Tak95])

*Any reduction  $s \rightarrow^* t$  can be reorganised so that  $s \rightarrow_{\mathbf{h}}^* \rightarrow_{\neg\mathbf{h}}^* t$ .*

**Takahashi's proof:** **decompose**  $\Rightarrow$  and then **swap** reductions.

Theorem ([Bar84])

- *Leftmost Reduction (hence, Head normalisability),*
- *Standardisation.*

## Factorisation in $\Lambda_\Sigma$

$\widetilde{\rightarrow}_h$  does not commute with  $\widetilde{\rightarrow}_{-h}$ .

## Factorisation in $\Lambda_\Sigma$

$\widetilde{\rightarrow}_h$  **does not** commute with  $\widetilde{\rightarrow}_{-h}$ .

### Function/Argument decomposition

$$\widetilde{\rightarrow} = \widetilde{\rightarrow}_f \cup \widetilde{\rightarrow}_a$$

- $\widetilde{\rightarrow}_a$  **argument** reduction:  $\widetilde{\rightarrow}$  in **argument** position,
- $\widetilde{\rightarrow}_f$  **function** reduction:  $\widetilde{\rightarrow}$  in **function** position.



## Factorisation in $\Lambda_\Sigma$

$\widetilde{\rightarrow}_h$  **does not** commute with  $\widetilde{\rightarrow}_{\neg h}$ .

### Function/Argument decomposition

$$\widetilde{\rightarrow} = \widetilde{\rightarrow}_f \cup \widetilde{\rightarrow}_a$$

- $\widetilde{\rightarrow}_a$  **argument** reduction:  $\widetilde{\rightarrow}$  in **argument** position,
- $\widetilde{\rightarrow}_f$  **function** reduction:  $\widetilde{\rightarrow}$  in **function** position.

### Theorem

*Any reduction  $S \widetilde{\rightarrow}^* T$  can be reorganised so that  $S \widetilde{\rightarrow}_f^* \widetilde{\rightarrow}_a^* T$ .*

**Proof.** Takahashi's with some technical adaptations.

## Head Normalisability

Simple terms exhibit the **structure** of pure  $\lambda$ -terms.

### Definition (Head normal form)

Let  $\text{HNF}_R$  the set of simple head normal forms:

$$\lambda x_1 \dots \lambda x_m. ((y) T_1) \dots T_n$$

with  $m, n \geq 0$  and  $T_1, \dots, T_n \in R\langle \Delta_R \rangle$ .

Hence,  $R\langle \text{HNF}_R \rangle$  is the set of head normal forms.

# Head Normalisability

Simple terms exhibit the **structure** of pure  $\lambda$ -terms.

## Definition (Head normal form)

Let  $\text{HNF}_R$  the set of simple head normal forms:

$$\lambda x_1 \dots \lambda x_m. ((y) T_1) \dots T_n$$

with  $m, n \geq 0$  and  $T_1, \dots, T_n \in R\langle\Delta_R\rangle$ .

Hence,  $R\langle\text{HNF}_R\rangle$  is the set of head normal forms.

Reduction  $\widetilde{\rightarrow}_f$  is head normalising:

## Theorem

*If  $S \in P\langle\Delta_P\rangle$  has head normal form, then  $S \widetilde{\rightarrow}_f^* \text{HNF}(S)$ .*

## Normalisability (issues)

- **Delicate** rewriting property in  $\Lambda_\Sigma$ : it **depends** on the semiring  $R$ .

## Normalisability (issues)

- **Delicate** rewriting property in  $\Lambda_\Sigma$ : it **depends** on the semiring  $R$ .
- If  $-1 \in R$ , so that  $1 + (-1) = 0$ , reduction **collapses**.

### Proposition [Vau09]

If  $R$  is not positive, *i.e.*  $a, b \in R^\bullet$  such that  $a + b = 0$ ,  $S$  reduces.

In general,  $R\langle\Delta_R\rangle$  has **no normal forms**.

## Normalisability (issues)

- **Delicate** rewriting property in  $\Lambda_{\Sigma}$ : it **depends** on the semiring  $R$ .
- If  $-1 \in R$ , so that  $1 + (-1) = 0$ , reduction **collapses**.

### Proposition [Vau09]

If  $R$  is not positive, *i.e.*  $a, b \in R^{\bullet}$  such that  $a + b = 0$ ,  $S$  reduces.

In general,  $R\langle\Delta_R\rangle$  has **no normal forms**.

However:

### Theorem ([Vau09])

$\mathbb{N}\langle\Delta_{\mathbb{N}}\rangle$  is conservative with respect to pure  $\lambda$ -calculus.

$\Rightarrow$  (strongly) **normalisable terms** and **normal forms**.

# The Semiring $\mathbb{P}$

## Definition

$\mathbb{P}$  is  $\mathbb{N}[\Xi]$ : the semiring of polynomials over a set of indeterminates  $\Xi$ , with non-negative integer coefficients.

- $\mathbb{P}$  exploitable as representation for every  $R$  (e.g.  $\mathbb{N}$ ,  $\mathbb{Q}$ , etc.),
- $\mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  inherits the good normalisability properties of  $\mathbb{N}\langle\Delta_{\mathbb{N}}\rangle$ .

# The Semiring $\mathbb{P}$

## Definition

$\mathbb{P}$  is  $\mathbb{N}[\Xi]$ : the semiring of polynomials over a set of indeterminates  $\Xi$ , with non-negative integer coefficients.

- $\mathbb{P}$  exploitable as representation for every  $R$  (e.g.  $\mathbb{N}, \mathbb{Q}$ , etc.),
- $\mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  inherits the good normalisability properties of  $\mathbb{N}\langle\Delta_{\mathbb{N}}\rangle$ .

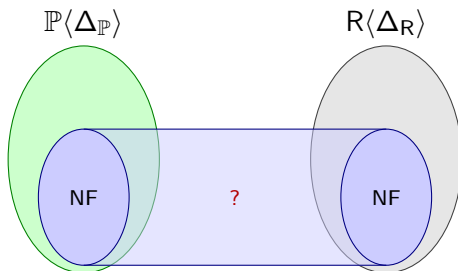
## Idea

Characterise normal forms in  $R\langle\Delta_R\rangle$  in terms of those in  $\mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$ .



## Abstract Construction

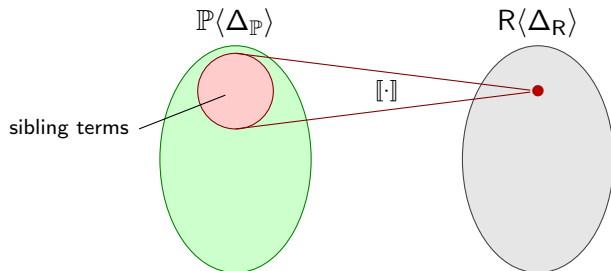
$\mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  as representation of any  $\mathbb{R}\langle\Delta_{\mathbb{R}}\rangle$ .



How can we **relate** the two?

## Abstract Construction

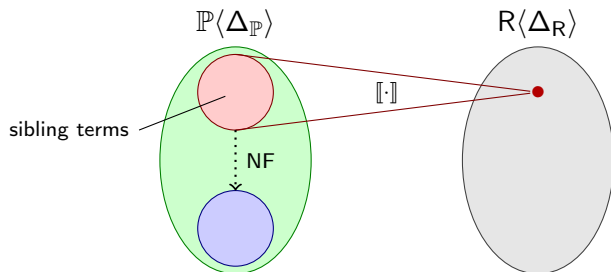
- **Term evaluation**  $\llbracket \cdot \rrbracket$ : morphism from  $\mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  to  $R\langle\Delta_{\mathbb{R}}\rangle$ ,



**Sibling terms** are notations for the same term •.

## Abstract Construction

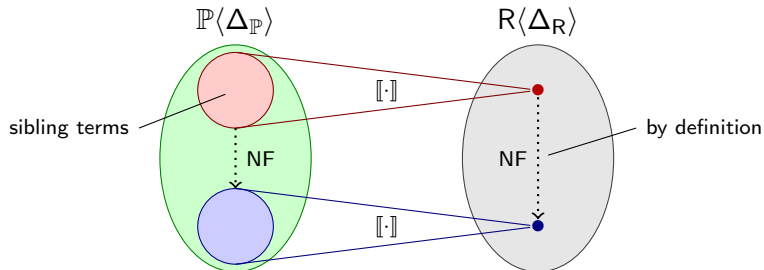
- **Term evaluation**  $\llbracket \cdot \rrbracket$ : morphism from  $\mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  to  $R\langle\Delta_{\mathbb{R}}\rangle$ ,



$\mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  exhibits normal forms.

# Abstract Construction

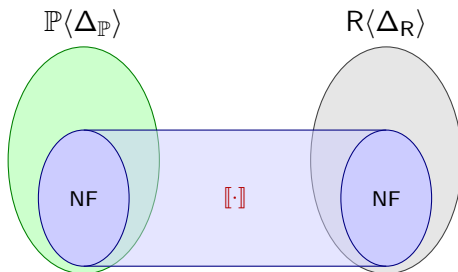
- **Term evaluation**  $\llbracket \cdot \rrbracket$ : morphism from  $\mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  to  $\mathbb{R}\langle\Delta_{\mathbb{R}}\rangle$ ,
- Normal forms in  $\mathbb{R}\langle\Delta_{\mathbb{R}}\rangle$ : terms **admitting** a redex-free writing.



NF(sibling terms) are notations for the same term:  $\bullet = \text{NF}(\bullet)$ .

## Abstract Construction

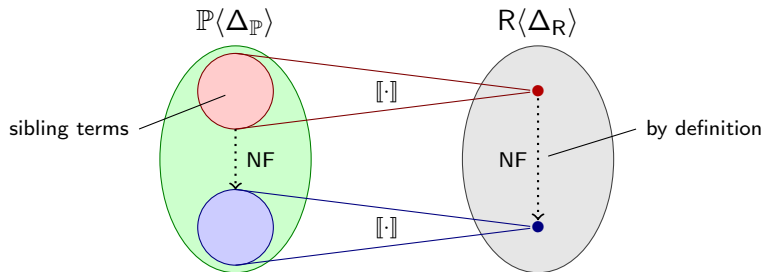
- **Term evaluation**  $\llbracket \cdot \rrbracket$ : morphism from  $\mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  to  $R\langle\Delta_{\mathbb{R}}\rangle$ ,
- Normal forms in  $R\langle\Delta_{\mathbb{R}}\rangle$ : terms **admitting** a redex-free writing.



NF of  $\mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  **induces** NF of  $R\langle\Delta_{\mathbb{R}}\rangle$  via  $\llbracket \cdot \rrbracket$ .

# Abstract Construction

- **Term evaluation**  $\llbracket \cdot \rrbracket$ : morphism from  $\mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  to  $\mathbb{R}\langle\Delta_{\mathbb{R}}\rangle$ ,
- Normal forms in  $\mathbb{R}\langle\Delta_{\mathbb{R}}\rangle$ : terms **admitting** a redex-free writing.



## Are Sibling Terms NF-Stable?

**Strong** normalisability:

### Theorem

*Let  $S, T \in \mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  be SN. If  $\llbracket S \rrbracket = \llbracket T \rrbracket$ , then  $\llbracket \text{NF}(S) \rrbracket = \llbracket \text{NF}(T) \rrbracket$ .*

## Are Sibling Terms NF-Stable?

~~Strong~~ Normalisability:

### Theorem

Let  $S, T \in \mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  be N. If  $\llbracket S \rrbracket = \llbracket T \rrbracket$ , then  $\llbracket \text{NF}(S) \rrbracket = \llbracket \text{NF}(T) \rrbracket$ .



## Are Sibling Terms NF-Stable?

~~Strong~~ Normalisability:

### Theorem

Let  $S, T \in \mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  be N. If  $\llbracket S \rrbracket = \llbracket T \rrbracket$ , then  $\llbracket \text{NF}(S) \rrbracket = \llbracket \text{NF}(T) \rrbracket$ .

- **Unique** normal form:

### Definition

Let  $T \in \mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  be N with  $\llbracket T \rrbracket = S \in \mathbb{R}\langle\Delta_{\mathbb{R}}\rangle$ :  $\text{NF}(S) = \llbracket \text{NF}(T) \rrbracket$ .

## Are Sibling Terms NF-Stable?

~~Strong~~ Normalisability:

### Theorem

Let  $S, T \in \mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  be N. If  $\llbracket S \rrbracket = \llbracket T \rrbracket$ , then  $\llbracket \text{NF}(S) \rrbracket = \llbracket \text{NF}(T) \rrbracket$ .

- **Unique** normal form:

### Definition

Let  $T \in \mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  be N with  $\llbracket T \rrbracket = S \in \mathbb{R}\langle\Delta_{\mathbb{R}}\rangle$ :  $\text{NF}(S) = \llbracket \text{NF}(T) \rrbracket$ .

- **Partial**, but **consistent**, term equivalence:

### Definition

$S \cong T$  if there are normalisable terms  $U, V \in \mathbb{P}\langle\Delta_{\mathbb{P}}\rangle$  such that  $\llbracket U \rrbracket = S$ ,  $\llbracket V \rrbracket = T$  and  $\llbracket \text{NF}(U) \rrbracket = \llbracket \text{NF}(V) \rrbracket$ .

## Canonical Reductions

Is it possible to **directly compute** these normal forms?

## Canonical Reductions

Is it possible to **directly compute** these normal forms?

- Parallel  $\beta$ -reduction  $\Rightarrow$ ,
- extended to **canonical** terms only:

$$\sum_{i=1}^n a_i s_i \widehat{\Rightarrow} \sum_{i=1}^n a_i T_i \text{ where, } s_i \neq s_j \text{ and } s_i \Rightarrow T_i.$$

## Canonical Reductions

Is it possible to **directly compute** these normal forms?

- Parallel  $\beta$ -reduction  $\Rightarrow$ ,
- extended to **canonical** terms only:

$$\sum_{i=1}^n a_i s_i \widehat{\Rightarrow} \sum_{i=1}^n a_i T_i \text{ where, } s_i \neq s_j \text{ and } s_i \Rightarrow T_i.$$

### Theorem

For all normalisable  $S \in \mathbb{P}\langle \Delta_{\mathbb{P}} \rangle$ ,  $\llbracket S \rrbracket \widehat{\Rightarrow}^* \llbracket \text{NF}(S) \rrbracket$ .

# Canonical Reductions

Is it possible to **directly compute** these normal forms?

- Parallel  $\beta$ -reduction  $\Rightarrow$ ,
- extended to **canonical** terms only:

$$\sum_{i=1}^n a_i s_i \widehat{\Rightarrow} \sum_{i=1}^n a_i T_i \text{ where, } s_i \neq s_j \text{ and } s_i \Rightarrow T_i.$$

## Theorem

For all normalisable  $S \in \mathbb{P}\langle \Delta_{\mathbb{P}} \rangle$ ,  $\llbracket S \rrbracket \widehat{\Rightarrow}^* \llbracket \text{NF}(S) \rrbracket$ .

- **Non-parallel** version characterises **strong** normalisability only,

# Canonical Reductions

Is it possible to **directly compute** these normal forms?

- Parallel  $\beta$ -reduction  $\Rightarrow$ ,
- extended to **canonical** terms only:

$$\sum_{i=1}^n a_i s_i \widehat{\Rightarrow} \sum_{i=1}^n a_i T_i \text{ where, } s_i \neq s_j \text{ and } s_i \Rightarrow T_i.$$

## Theorem

*For all normalisable  $S \in \mathbb{P}\langle \Delta_{\mathbb{P}} \rangle$ ,  $\llbracket S \rrbracket \widehat{\Rightarrow}^* \llbracket \text{NF}(S) \rrbracket$ .*

- **Non-parallel** version characterises **strong** normalisability only,
- Relations on canonical terms exhibit **bad** rewriting properties.

## Part II

# Coinductive equivalences in a probabilistic scenario



## Higher-Order Program Equivalence

How can we establish  
 $M$  equivalent to  $N$  ?

## Higher-Order Program Equivalence

How can we establish  
 $M$  equivalent to  $N$  ?

### Context equivalence

As soon as their **observable behaviour** is the same in **every** context:

if  $\text{Obs}(C \langle M \rangle, C \langle N \rangle)$  then  $M \simeq N$ .

# Higher-Order Program Equivalence

How can we establish  
 $M$  equivalent to  $N$  ?

## Context equivalence

As soon as their **observable behaviour** is the same in **every** context:

if  $\text{Obs}(C \langle M \rangle, C \langle N \rangle)$  then  $M \simeq N$ .

Observational equivalences:

- Subsume **compositional** reasonings (**congruence**),
- Involve **universal** quantification (**hard**),
- Amenable to **coinductive** formulations (**bisimulation**).

## Definition

{Terms:  $\Lambda$ }  $M, N ::= x \mid \lambda x.M \mid (M)N$

{Values:  $V$ }  $V, W ::= \lambda x.M$  (closed)

- **Lazy** operational semantics (**Weak** CbN):

$$\frac{}{V \Downarrow V} \text{ (val)} \qquad \frac{M \Downarrow \lambda x.L \quad L[N/x] \Downarrow V}{(M)N \Downarrow V} \text{ (term)}$$

## Definition

{Terms:  $\Lambda$ }  $M, N ::= x \mid \lambda x.M \mid (M) N$

{Values:  $V$ }  $V, W ::= \lambda x.M$  (closed)

- **Lazy** operational semantics (**Weak** CbN):

$$\overline{V \Downarrow V} \text{ (val)} \quad \frac{M \Downarrow \lambda x.L \quad L[N/x] \Downarrow V}{(M) N \Downarrow V} \text{ (term)}$$

- Labelled Transition System:

Terms	→	Values
$M$	$\xrightarrow{\text{eval}}$	$V$
$L[N/x]$	$\xleftarrow{N}$	$\lambda x.L$

# Applicative (Bi)Simulation

[Abr90]

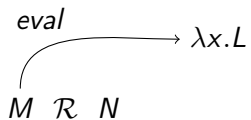
When is  $\mathcal{R}$  a simulation?

$M \mathcal{R} N$

# Applicative (Bi)Simulation

[Abr90]

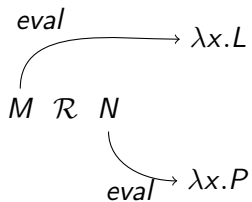
When is  $\mathcal{R}$  a simulation?



# Applicative (Bi)Simulation

[Abr90]

When is  $\mathcal{R}$  a *simulation*?

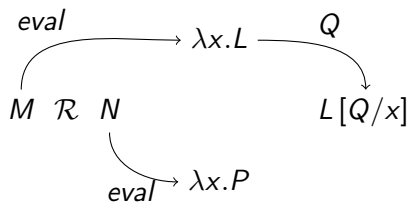




# Applicative (Bi)Simulation

[Abr90]

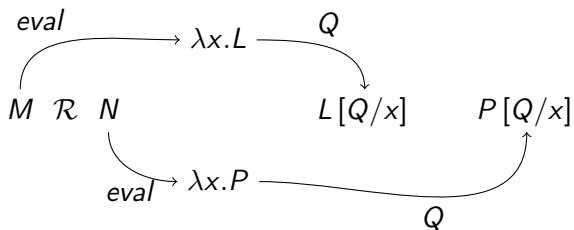
When is  $\mathcal{R}$  a **simulation**?



# Applicative (Bi)Simulation

[Abr90]

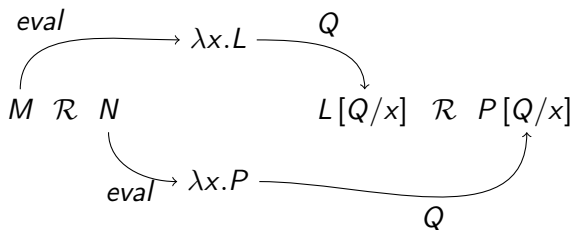
When is  $\mathcal{R}$  a simulation?



# Applicative (Bi)Simulation

[Abr90]

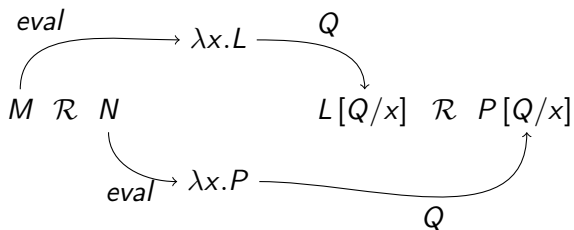
When is  $\mathcal{R}$  a simulation?



# Applicative (Bi)Simulation

[Abr90]

When is  $\mathcal{R}$  a simulation?



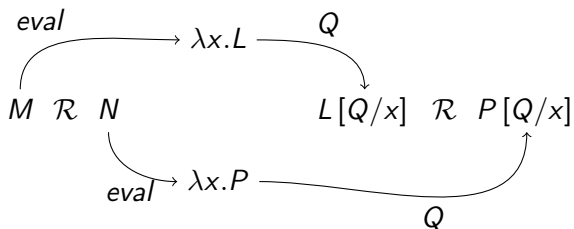
**Bisimulation:**  $\mathcal{R}$  and  $\mathcal{R}^{op}$  are simulations.

- **Similarity**  $\lesssim$ : union of all simulations,
- **Bisimilarity**  $\sim$ : union of all bisimulations.

# Applicative (Bi)Simulation

[Abr90]

When is  $\mathcal{R}$  a simulation?



**Bisimulation:**  $\mathcal{R}$  and  $\mathcal{R}^{op}$  are simulations.

- **Similarity**  $\lesssim$ : union of all simulations,
- **Bisimilarity**  $\sim$ : union of all bisimulations.

**Theorem (Applicative bisimilarity is context equivalence)**

$$M \sim N \iff M \simeq N$$

## $\Lambda_{\oplus}$ : Syntax and Probabilistic Operational Semantics

Same language of terms of the non-deterministic  $\lambda$ -calculus [dP95].

### Definition

{Terms:  $\Lambda_{\oplus}$ }      $M, N ::= x \mid \lambda x.M \mid (M) N \mid M \oplus N$

{Values:  $V\Lambda_{\oplus}$ }      $V, W ::= \lambda x.M$      (closed)

# $\Lambda_{\oplus}$ : Syntax and Probabilistic Operational Semantics

Same language of terms of the non-deterministic  $\lambda$ -calculus [dP95].

## Definition

{Terms:  $\Lambda_{\oplus}$ }      $M, N ::= x \mid \lambda x.M \mid (M) N \mid M \oplus N$

{Values:  $\forall \Lambda_{\oplus}$ }      $V, W ::= \lambda x.M$      (closed)

Non-deterministic meaning of a term:

- A set of values

$$M \mapsto \{V, W\},$$

- Potentially infinite ( $\Theta$  and  $\oplus$ ).

# $\Lambda_{\oplus}$ : Syntax and Probabilistic Operational Semantics

Same language of terms of the non-deterministic  $\lambda$ -calculus [dP95].

## Definition

{Terms:  $\Lambda_{\oplus}$ }  $M, N ::= x \mid \lambda x.M \mid (M)N \mid M \oplus N$

{Values:  $V\Lambda_{\oplus}$ }  $V, W ::= \lambda x.M$  (closed)

Probabilistic meaning of a term:

- A value distribution  $\mathcal{D} : V\Lambda_{\oplus} \rightarrow \mathbb{R}_{[0,1]}$

$$M \mapsto \{V^p, W^q\},$$

- Bounded by 1:  $\sum_{V \in V\Lambda_{\oplus}} \mathcal{D}(V) \leq 1$ .



# $\Lambda_{\oplus}$ : Syntax and Probabilistic Operational Semantics

Same language of terms of the non-deterministic  $\lambda$ -calculus [dP95].

## Definition

{Terms:  $\Lambda_{\oplus}$ }  $M, N ::= x \mid \lambda x.M \mid (M)N \mid M \oplus N$

{Values:  $V\Lambda_{\oplus}$ }  $V, W ::= \lambda x.M$  (closed)

Probabilistic meaning of a term:

- A value distribution  $\mathcal{D} : V\Lambda_{\oplus} \rightarrow \mathbb{R}_{[0,1]}$

$$M \mapsto \{V^p, W^q\},$$

- Bounded by 1:  $\sum_{V \in V\Lambda_{\oplus}} \mathcal{D}(V) \leq 1$ .

**Problem:** no finitary operational semantics suffices.

## $\Lambda_{\oplus}$ : Syntax and Probabilistic Operational Semantics

DL and Zorzi's *idea*: limit of approximated value distributions.

# $\Lambda_{\oplus}$ : Syntax and Probabilistic Operational Semantics

DL and Zorzi's **idea**: limit of approximated value distributions.

**Approximation** CbN operational semantics [DLZ12]:

$$\overline{M \Downarrow \emptyset} \text{ (be)} \quad \overline{V \Downarrow \{V^1\}} \text{ (bv)} \quad \frac{M \Downarrow \mathcal{D} \quad N \Downarrow \mathcal{E}}{M \oplus N \Downarrow \frac{1}{2}\mathcal{D} + \frac{1}{2}\mathcal{E}} \text{ (bs)}$$

$$\frac{M \Downarrow \mathcal{D} \quad \{P[N/x] \Downarrow \mathcal{E}_{P,N}\}_{\lambda x. P \in \text{Supp}(\mathcal{D})}}{(M) N \Downarrow \sum_{\lambda x. P \in \text{Supp}(\mathcal{D})} \mathcal{D}(\lambda x. P) \cdot \mathcal{E}_{P,N}} \text{ (ba)}$$

# $\Lambda_{\oplus}$ : Syntax and Probabilistic Operational Semantics

DL and Zorzi's **idea**: limit of approximated value distributions.

**Approximation** CbN operational semantics [DLZ12]:

$$\overline{M \Downarrow \emptyset} \text{ (be)} \quad \overline{V \Downarrow \{V^1\}} \text{ (bv)} \quad \frac{M \Downarrow \mathcal{D} \quad N \Downarrow \mathcal{E}}{M \oplus N \Downarrow \frac{1}{2}\mathcal{D} + \frac{1}{2}\mathcal{E}} \text{ (bs)}$$

$$\frac{M \Downarrow \mathcal{D} \quad \{P[N/x] \Downarrow \mathcal{E}_{P,N}\}_{\lambda x. P \in \text{Supp}(\mathcal{D})}}{(M) N \Downarrow \sum_{\lambda x. P \in \text{Supp}(\mathcal{D})} \mathcal{D}(\lambda x. P) \cdot \mathcal{E}_{P,N}} \text{ (ba)}$$

## Definition (Semantics)

CbN **semantics** of  $M \in \Lambda_{\oplus}$ :  $\llbracket M \rrbracket = \sup_{M \Downarrow \mathcal{D}} \mathcal{D}$ .

- Probabilistic LTS: Labelled Markov Chains,

- Probabilistic LTS: Labelled Markov Chains,

## Definition

A **probabilistic bisimulation** is an equivalence relation  $\mathcal{R}$  on  $\mathcal{S}$  such that  $s \mathcal{R} t$  implies: for every  $\ell \in \mathcal{L}$  and  $\mathbf{E} \in \mathcal{S}/\mathcal{R}$ ,

$$\mathcal{P}(s, \ell, \mathbf{E}) = \mathcal{P}(t, \ell, \mathbf{E}).$$

**Probabilistic simulation** is required to be a preorder.

- Probabilistic LTS: Labelled Markov Chains,

## Definition

A **probabilistic bisimulation** is an equivalence relation  $\mathcal{R}$  on  $\mathcal{S}$  such that  $s \mathcal{R} t$  implies: for every  $\ell \in \mathcal{L}$  and  $\mathbf{E} \in \mathcal{S}/\mathcal{R}$ ,

$$\mathcal{P}(s, \ell, \mathbf{E}) = \mathcal{P}(t, \ell, \mathbf{E}).$$

**Probabilistic simulation** is required to be a preorder.

Similarity ( $\lesssim$ ) and Bisimilarity ( $\sim$ ) can always be formed.

**Proposition (Bisimilarity is similarity equivalence)**

$$\sim = \lesssim \cap \lesssim^{op}$$

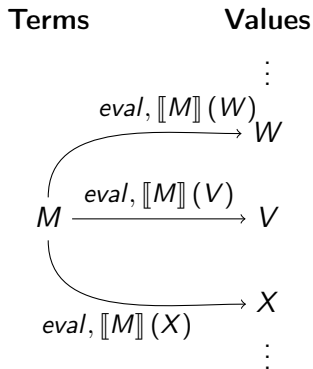
## $\Lambda_{\oplus}$ as a Labelled Markov Chain

- **States:** terms and values,



## $\Lambda_{\oplus}$ as a Labelled Markov Chain

- **States:** terms and values,
- **Transitions:**
  - $M$  evaluates to  $V$  with  $\llbracket M \rrbracket (V)$  probability,



## $\Lambda_{\oplus}$ as a Labelled Markov Chain

- **States:** terms and values,
- **Transitions:**
  - $M$  evaluates to  $V$  with  $\llbracket M \rrbracket(V)$  probability,
  - Values get a **term** in input.

Terms	Values
$L[N/x]$	$\lambda x.L$

$\xleftarrow{N, 1}$

## Probabilistic Applicative (Bi)Simulation

$\mathcal{R}$  **bisimulation** whenever equivalence relation and

- on **values**: for every  $L$ ,

$$\lambda x.M \mathcal{R} \lambda x.N$$

## Probabilistic Applicative (Bi)Simulation

$\mathcal{R}$  **bisimulation** whenever equivalence relation and

- on **values**: for every  $L$ ,

$$\begin{array}{ccc} & L & \\ & \curvearrowright & \\ \lambda x.M \ \mathcal{R} \ \lambda x.N & & M[L/x] \ \mathcal{R} \ N[L/x] \\ & \curvearrowleft & \\ & L & \end{array}$$

# Probabilistic Applicative (Bi)Simulation

$\mathcal{R}$  **bisimulation** whenever equivalence relation and

- on **values**: for every  $L$ ,



- on **terms**: for every  $E \in V\Lambda_{\oplus}/\mathcal{R}$ ,

$$M \mathcal{R} N$$

# Probabilistic Applicative (Bi)Simulation

$\mathcal{R}$  **bisimulation** whenever equivalence relation and

- on **values**: for every  $L$ ,

$$\begin{array}{ccc} & L & \\ & \curvearrowright & \\ \lambda x.M \ \mathcal{R} \ \lambda x.N & & M[L/x] \ \mathcal{R} \ N[L/x] \\ & \curvearrowleft & \\ & L & \end{array}$$

- on **terms**: for every  $E \in V\Lambda_{\oplus}/\mathcal{R}$ ,

$$\begin{array}{ccc} & eval & \\ & \curvearrowright & \\ M \ \mathcal{R} \ N & & \llbracket M \rrbracket \end{array}$$

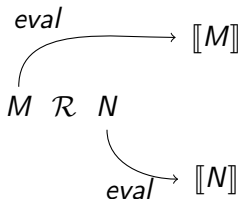
## Probabilistic Applicative (Bi)Simulation

$\mathcal{R}$  **bisimulation** whenever equivalence relation and

- on **values**: for every  $L$ ,



- on **terms**: for every  $E \in V\Lambda_{\oplus}/\mathcal{R}$ ,



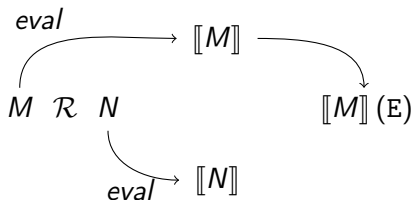
# Probabilistic Applicative (Bi)Simulation

$\mathcal{R}$  **bisimulation** whenever equivalence relation and

- on **values**: for every  $L$ ,



- on **terms**: for every  $E \in V\Lambda_{\oplus}/\mathcal{R}$ ,







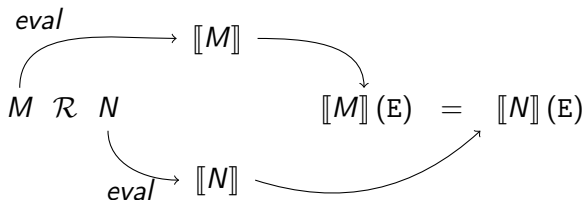
# Probabilistic Applicative (Bi)Simulation

$\mathcal{R}$  **bisimulation** whenever equivalence relation and

- on **values**: for every  $L$ ,



- on **terms**: for every  $E \in V\Lambda_{\oplus}/\mathcal{R}$ ,



# Context Equivalence vs. Bisimilarity

## Definition

Contexts  $C \Lambda_{\oplus}$ :

$$C ::= \langle \cdot \rangle \mid \lambda x. C \mid (C) M \mid (M) C \mid C \oplus M \mid M \oplus C$$

Context equivalence:  $M \simeq N \iff \sum [C \langle M \rangle] = \sum [C \langle N \rangle]$ .

Is  $\sim$  included in  $\simeq$ ?

# Context Equivalence vs. Bisimilarity

## Definition

Contexts  $\mathcal{C}\Lambda_{\oplus}$ :

$$C ::= \langle \cdot \rangle \mid \lambda x. C \mid (C) M \mid (M) C \mid C \oplus M \mid M \oplus C$$

Context equivalence:  $M \simeq N \iff \sum \llbracket C \langle M \rangle \rrbracket = \sum \llbracket C \langle N \rangle \rrbracket$ .

It is sufficient to show  $\sim$  congruence:

$$\begin{aligned} M \sim N &\implies C \langle M \rangle \sim C \langle N \rangle \\ &\implies \sum \llbracket C \langle M \rangle \rrbracket = \sum \llbracket C \langle N \rangle \rrbracket \\ &\implies M \simeq N. \end{aligned}$$

# Context Equivalence vs. Bisimilarity

## Definition

Contexts  $C \Lambda_{\oplus}$ :

$$C ::= \langle \cdot \rangle \mid \lambda x. C \mid (C) M \mid (M) C \mid C \oplus M \mid M \oplus C$$

Context equivalence:  $M \simeq N \iff \sum [C \langle M \rangle] = \sum [C \langle N \rangle]$ .

It is sufficient to show  $\sim$  congruence:

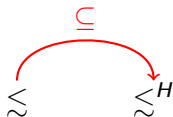
$$\begin{aligned} M \sim N &\implies C \langle M \rangle \sim C \langle N \rangle \\ &\implies \sum [C \langle M \rangle] = \sum [C \langle N \rangle] \\ &\implies M \simeq N. \end{aligned}$$

How to prove it? Direct proof **fails** due to application.

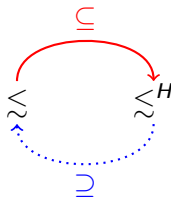
- **Howe's lifting:**  $\lesssim$  preorder  $\Rightarrow \lesssim^H$  **precongruence** and  $\lesssim \subseteq \lesssim^H$ ,

$$\lesssim \quad \lesssim^H$$

- **Howe's lifting:**  $\lesssim$  preorder  $\Rightarrow \lesssim^H$  precongruence and  $\lesssim \subseteq \lesssim^H$ ,



- **Howe's lifting:**  $\lesssim$  preorder  $\Rightarrow \lesssim^H$  **precongruence** and  $\lesssim \subseteq \lesssim^H$ ,
- **Key Lemma:**  $\lesssim^H$  is a **simulation**  $\Rightarrow \lesssim^H \subseteq \lesssim$  (to be proved!).





## Key Lemma

Key Lemma ( $\lesssim^H$  is a simulation)

If  $M \lesssim^H N$ , then for every  $X \subseteq \Lambda_{\oplus}(x)$  it holds that

$$\llbracket M \rrbracket (\lambda x. X) \leq \llbracket N \rrbracket (\lambda x. (\lesssim^H(X))).$$

**Proof sketch.**

- For all  $\mathcal{D}$  such that  $M \Downarrow \mathcal{D}$ ,  $\mathcal{D}(\lambda x. X) \leq \llbracket N \rrbracket (\lambda x. (\lesssim^H(X)))$ ,
- By induction on the (**finite**) derivation of  $M \Downarrow \mathcal{D}$ :
  - Value and probabilistic choice cases are easy,
  - Application case is hard: it requires probability assignments can always be **disentangled** (via max-flow min-cut theorem).

## Key Lemma

### Key Lemma ( $\lesssim^H$ is a simulation)

If  $M \lesssim^H N$ , then for every  $X \subseteq \Lambda_{\oplus}(x)$  it holds that

$$\llbracket M \rrbracket (\lambda x. X) \leq \llbracket N \rrbracket (\lambda x. (\lesssim^H(X))).$$

### Proof sketch.

- For all  $\mathcal{D}$  such that  $M \Downarrow \mathcal{D}$ ,  $\mathcal{D}(\lambda x. X) \leq \llbracket N \rrbracket (\lambda x. (\lesssim^H(X)))$ ,
- By induction on the (finite) derivation of  $M \Downarrow \mathcal{D}$ :
  - Value and probabilistic choice cases are easy,
  - Application case is hard: it requires probability assignments can always be disentangled (via max-flow min-cut theorem).

$\sim$  congruence, since  $\sim = \lesssim \cap \lesssim^{op}$  and  $\lesssim$  precongruence:

- $\lesssim \subseteq \lesssim^H$  (Howe's lifting),
- $\lesssim \supseteq \lesssim^H$  (Key Lemma).

## On Full Abstraction

- $\sim$  is **sound** w.r.t.  $\simeq$ .
- Is  $\sim$  also **complete**?

## On Full Abstraction

- $\sim$  is **sound** w.r.t.  $\simeq$ .
- Is  $\sim$  also **complete**? **No**, at least in **CbN**.

**Counterexample.** Consider:

$$M = \lambda x. \lambda y. (\Omega \oplus \mathbf{I}) \quad N = \lambda x. (\lambda y. \Omega) \oplus (\lambda y. \mathbf{I})$$

Since  $\mathbf{I} \not\sim \Omega$ , then

$$\lambda y. \Omega \not\sim \lambda y. \mathbf{I} \not\sim \lambda y. (\Omega \oplus \mathbf{I})$$

which implies  $M \not\sim N$ .

- Same counterexample to completeness in non-deterministic  $\Lambda$ ,

## On Full Abstraction

- $\sim$  is **sound** w.r.t.  $\simeq$ .
- Is  $\sim$  also **complete**? **No**, at least in **CbN**.

**Counterexample.** Consider:

$$M = \lambda x. \lambda y. (\Omega \oplus \mathbf{I}) \quad N = \lambda x. (\lambda y. \Omega) \oplus (\lambda y. \mathbf{I})$$

Since  $\mathbf{I} \not\sim \Omega$ , then

$$\lambda y. \Omega \not\sim \lambda y. \mathbf{I} \not\sim \lambda y. (\Omega \oplus \mathbf{I})$$

which implies  $M \not\sim N$ .

- Same counterexample to completeness in non-deterministic  $\Lambda$ ,
- Nonetheless,  $M \simeq N$  (via CIU-equivalence).

## Full Abstraction on Pure $\lambda$ -terms

What about  $\sim$  vs.  $\simeq$  on pure, **deterministic**,  $\lambda$ -terms?

## Full Abstraction on Pure $\lambda$ -terms

What about  $\sim$  vs.  $\simeq$  on pure, **deterministic**,  $\lambda$ -terms?

- Lévy-Longo trees are the **lazy variant** of Böhm trees:
  - $LT(\lambda x.\Omega) = \lambda x.\perp$  whereas  $LT(\Omega) = \perp$ ,

## Full Abstraction on Pure $\lambda$ -terms

What about  $\sim$  vs.  $\simeq$  on pure, **deterministic**,  $\lambda$ -terms?

- Lévy-Longo trees are the **lazy variant** of Böhm trees:
  - $LT(\lambda x.\Omega) = \lambda x.\perp$  whereas  $LT(\Omega) = \perp$ ,
  - Induce the finest equivalence in a **lazy** regime.

Definition (Lévy-Longo equality)

$$M =_{LL} N \iff LT(M) = LT(N).$$



## Full Abstraction on Pure $\lambda$ -terms

What about  $\sim$  vs.  $\simeq$  on pure, **deterministic**,  $\lambda$ -terms?

- Lévy-Longo trees are the **lazy variant** of Böhm trees:
  - $LT(\lambda x.\Omega) = \lambda x.\perp$  whereas  $LT(\Omega) = \perp$ ,
  - Induce the finest equivalence in a **lazy** regime.

Definition (Lévy-Longo equality)

$$M =_{LL} N \iff LT(M) = LT(N).$$

- $\simeq \subseteq =_{LL}$  via a variation of Böhm-out technique,

## Full Abstraction on Pure $\lambda$ -terms

What about  $\sim$  vs.  $\simeq$  on pure, **deterministic**,  $\lambda$ -terms?

- Lévy-Longo trees are the **lazy variant** of Böhm trees:
  - $LT(\lambda x.\Omega) = \lambda x.\perp$  whereas  $LT(\Omega) = \perp$ ,
  - Induce the finest equivalence in a **lazy** regime.

### Definition (Lévy-Longo equality)

$$M =_{LL} N \iff LT(M) = LT(N).$$

- $\simeq \subseteq =_{LL}$  via a variation of Böhm-out technique,
- $=_{LL} \subseteq \sim$  via a notion of strict applicative bisimulation.

## Full Abstraction on Pure $\lambda$ -terms

What about  $\sim$  vs.  $\simeq$  on pure, **deterministic**,  $\lambda$ -terms?

- Lévy-Longo trees are the **lazy variant** of Böhm trees:
  - $LT(\lambda x.\Omega) = \lambda x.\perp$  whereas  $LT(\Omega) = \perp$ ,
  - Induce the finest equivalence in a **lazy** regime.

### Definition (Lévy-Longo equality)

$$M =_{\text{LL}} N \iff LT(M) = LT(N).$$

- $\simeq \subseteq =_{\text{LL}}$  via a variation of Böhm-out technique,
- $=_{\text{LL}} \subseteq \sim$  via a notion of strict applicative bisimulation.

### Theorem

*On pure  $\lambda$ -terms:  $\sim$ ,  $\simeq$  and  $=_{\text{LL}}$  all coincide.*

Thanks!

$(Q \& A^*)^*$  time.



# Labelled Markov Chains

- Probabilistic labelled transition systems,
- Discrete state space and time.

## Definition

A **Labelled Markov Chain** is a triple  $(\mathcal{S}, \mathcal{L}, \mathcal{P})$  such that:

- $\mathcal{S}$  is a countable set of **states**;
- $\mathcal{L}$  is set of **labels**;
- $\mathcal{P}$  is a **transition probability matrix**, i.e. a function  $\mathcal{P} : \mathcal{S} \times \mathcal{L} \times \mathcal{S} \rightarrow \mathbb{R}_{[0,1]}$  such that for every  $s \in \mathcal{S}$  and  $l \in \mathcal{L}$ :

$$\mathcal{P}(s, l, \mathcal{S}) = \sum_{t \in \mathcal{S}} \mathcal{P}(s, l, t) \leq 1.$$

# Howe's Technique

[Pit11]

- Construct  $\mathcal{R}^H$  from  $\mathcal{R}$  such that:
  - $\mathcal{R}^H$  is a **precongruence**, whenever  $\mathcal{R}$  is a preorder,
  - $\mathcal{R} \subseteq \mathcal{R}^H$  and (via **Key Lemma**)  $\mathcal{R} \supseteq \mathcal{R}^H$ .

# Howe's Technique

[Pit11]

- Construct  $\mathcal{R}^H$  from  $\mathcal{R}$  such that:
  - $\mathcal{R}^H$  is a **precongruence**, whenever  $\mathcal{R}$  is a preorder,
  - $\mathcal{R} \subseteq \mathcal{R}^H$  and (via **Key Lemma**)  $\mathcal{R} \supseteq \mathcal{R}^H$ .
- Howe's lifting for  $\Lambda_{\oplus}$ :

$$\frac{\bar{x} \vdash x \mathcal{R} M}{\bar{x} \vdash x \mathcal{R}^H M} \quad \frac{\bar{x} \cup \{x\} \vdash M \mathcal{R}^H L \quad \bar{x} \vdash \lambda x. L \mathcal{R} N \quad x \notin \bar{x}}{\bar{x} \vdash \lambda x. M \mathcal{R}^H N}$$

$$\frac{\bar{x} \vdash M \mathcal{R}^H P \quad \bar{x} \vdash N \mathcal{R}^H Q \quad \bar{x} \vdash (P) Q \mathcal{R} L}{\bar{x} \vdash (M) N \mathcal{R}^H L}$$

$$\frac{\bar{x} \vdash M \mathcal{R}^H P \quad \bar{x} \vdash N \mathcal{R}^H Q \quad \bar{x} \vdash P \oplus Q \mathcal{R} L}{\bar{x} \vdash M \oplus N \mathcal{R}^H L}$$