Metric Reasonning About λ-Terms: The Affine Case

Raphaëlle Crubillé

FOCUS Working Group

Higher-Order Program Equivalence

When can two programs be considered **equivalent**?

- **Context equivalence** [Morris1968]:
  - Two terms M and N are context equivalent if their **observable behavior** is the same in **any** context.

Higher-Order Program Equivalence

When can two programs be considered **equivalent**?

- **Context equivalence** [Morris1968]:
  - Two terms M and N are context equivalent if their **observable behavior** is the same in **any** context.
  - Proving that two programs are **not** equivalent is relatively easy: just find **a** context that separates them.

Higher-Order Program Equivalence

When can two programs be considered **equivalent**?

- **Context equivalence** [Morris1968]:
  - Two terms M and N are context equivalent if their **observable behavior** is the same in **any** context.
  - Proving that two programs are **not** equivalent is relatively easy: just find **a** context that separates them.
  - Proving that two program are indeed **equivalent**, on the other hand, can be quite complicated.

Higher-Order Program Equivalence

When can two programs be considered **equivalent**?

- **Context equivalence** [Morris1968]:
  - Two terms M and N are context equivalent if their **observable behavior** is the same in **any** context.
  - Proving that two programs are **not** equivalent is relatively easy: just find **a** context that separates them.
  - Proving that two program are indeed **equivalent**, on the other hand, can be quite complicated.
- Other equivalence notion: Bisimilarity

Observables

- **Deterministic** setting:
  - Operationnal Semantics: $M \Downarrow V$.
  - Observables: Termination:

$$Obs(M) = \begin{cases} 1 \text{ if } M \Downarrow \\ 0 \text{ if } M \Uparrow \end{cases} \qquad \in \{0,1\}$$

- **Probabilistic** setting:
  - Operationnal Semantics: $[\![M]\!] \in \mathrm{Distr}\,(\mathbf{V})$.
  - Observables: Convergence Probability.

$$Obs(M) = \sum_{V \text{ a value}} [\![M]\!](V) \qquad \in [0,1]$$

Syntax and Operational Semantics of $\Lambda_\oplus$ [DLZorzi2012]

- **Terms**: $M, N ::= x \mid \lambda x.M \mid MM \mid M \oplus M$;
- **Values**: $V \in \mathbf{V} ::= \lambda x.M$;
- **Approximation (Big-Step) Semantics**:
  - $M \Downarrow \mathscr{D}$, where $\mathscr{D} : \mathbf{V} \to [0,1]$ sub-probability distribution.
  - Approximation from below: only finite distributions.

$$\frac{}{M \Downarrow \emptyset} \qquad \frac{}{V \Downarrow \{V^1\}} \qquad \frac{M \Downarrow \mathscr{D} \qquad N \Downarrow \mathscr{E}}{M \oplus N \Downarrow \frac{1}{2}\mathscr{D} + \frac{1}{2}\mathscr{E}}$$

$$\frac{M \Downarrow \mathscr{D} \qquad \{P\{x/N\} \Downarrow \mathscr{F}_P\}_{\lambda x.P \in \mathsf{S}(\mathscr{D})}}{MN \Downarrow \sum_{\lambda x.P \in \mathsf{S}(\mathscr{D})} \mathscr{D}(\lambda x.P) \cdot \mathscr{F}_P}$$

- **Semantics**: $[\![M]\!] = \sup_{M \Downarrow \mathscr{D}} \mathscr{D}$;

Syntax and Operational Semantics of $\Lambda_\oplus$ [DLZorzi2012]

- **Terms**: $M, N ::= x \;\big|\; \lambda x.M \;\big|\; MM \;\big|\; M \oplus M$;
- **Values**: $V \in \mathbf{V} ::= \lambda x.M$;
- **Approximation (Big-Step) Semantics**:
  - $M \Downarrow \mathscr{D}$, where $\mathscr{D} : \mathbf{V} \to [0,1]$ sub-probability distribution.
  - Approximation from below: only finite distributions.

$$\frac{}{M \Downarrow \emptyset} \qquad \frac{}{V \Downarrow \{V^1\}} \qquad \frac{M \Downarrow \mathscr{D} \qquad N \Downarrow \mathscr{E}}{M \oplus N \Downarrow \frac{1}{2}\mathscr{D} + \frac{1}{2}\mathscr{E}}$$

$$\frac{M \Downarrow \mathscr{D} \qquad \{P\{x/N\} \Downarrow \mathscr{F}_P\}_{\lambda x.P \in \mathsf{S}(\mathscr{D})}}{MN \Downarrow \sum_{\lambda x.P \in \mathsf{S}(\mathscr{D})} \mathscr{D}(\lambda x.P) \cdot \mathscr{F}_P}$$

- **Semantics**: $[\![M]\!] = \sup_{M \Downarrow \mathscr{D}} \mathscr{D}$;
- Variations: **Small-Step** Semantics, **Call-by-value** Evaluation.

Applicative Bisimulation [Abramsky93]

Terms

Applicative Bisimulation [Abramsky93]

Terms **Values**

Applicative Bisimulation [Abramsky93]

Terms	**Values**
M	
N	
L	
$\vdots$	

Applicative Bisimulation [Abramsky93]

Terms	**Values**
M	V
N	W
L	U
$\vdots$	$\vdots$

Applicative Bisimulation [Abramsky93]

Terms	**Values**
M	

Applicative Bisimulation [Abramsky93]

Terms **Values**

$$M \xrightarrow{eval} V$$

Applicative Bisimulation [Abramsky93]

Terms **Values**

$$M \xrightarrow{eval} V$$

$$\lambda x.N$$

Applicative Bisimulation [Abramsky93]

Terms		**Values**
M	$\xrightarrow{eval}$	V
$N\{L/x\}$	$\xleftarrow{L}$	$\lambda x.N$

Applicative Bisimulation [Abramsky93]: Deterministic Case

- **Simulation**

$$\begin{tikzcd}
 & & & \lambda x.L \arrow[dr, bend left, "T"] & & & \\
M \arrow[urrr, bend left, "eval"] & R & N \arrow[dr, bend right, "eval"'] & & L\{T/x\} & R & K\{T/x\} \\
 & & & \lambda x.K \arrow[urrr, bend right, "T"'] & & &
\end{tikzcd}$$

Applicative Bisimulation [Abramsky93]: Deterministic Case

- **Simulation**

$$
\begin{tikzcd}
 & & & \lambda x.L \arrow[dr, bend left, "T"] & & & \\
M \arrow[urrr, bend left, "eval"] & R & N \arrow[dr, bend right, "eval"'] & & L\{T/x\} & R & K\{T/x\} \\
 & & & \lambda x.K \arrow[urrr, bend right, "T"'] & & &
\end{tikzcd}
$$

- **Similarity**: union of all simulations, denoted $\precsim$;
- **Bisimilarity**: union of all bisimulations, denoted $\sim$.

Theorem

$M \equiv N$ *iff* $M \sim N$.

Applicative Bisimulation for a Probabilistic Language (1): Trace Equivalence

The weighted **Trace** LTS:

- states: $\mathrm{Distr}\,(\mathbf{V})$
- labels: @Terms
- weight function: $w(\mathscr{D}) = \sum_V \mathscr{D}(V)$
- Transition relation:

$$\begin{array}{l} \mathscr{D} \xrightarrow{@M} \sum_V \mathscr{D}(V) \cdot [\![V\,M]\!] \\ \mathscr{D} \xrightarrow{@N} \sum_V \mathscr{D}(V) \cdot [\![V\,N]\!] \\ \qquad\vdots \end{array}$$

Definition

$M \equiv^{tr} N$ if $[\![M]\!]\,R\,[\![N]\!]$ when R is a bisimulation with respect to the trace LTS such that $\mathscr{D}\,R\,\mathscr{E} \Rightarrow w(\mathscr{D}) = w(\mathscr{E})$.

Alternative Equivalente Definition: Evaluation Contexts

- Traces: contexts having a nice peculiar form

$$s ::= \epsilon \mid @V \cdot s.$$

- **Success probability of a trace**: $Pr(M, s)$.
  - $s \longrightarrow \mathcal{C}_s$:

$$Pr(M, s) = \sum_V [\![\mathcal{C}_s[M]]\!](V)$$

  - Examples: $M = \lambda x.(I \oplus \Omega) \oplus \Omega$.

$$s = \epsilon \qquad \mathcal{C}_s = [\cdot] \qquad Pr(M, s) = \frac{1}{2}$$

$$t = @I \cdot \epsilon \qquad \mathcal{C}_t = [\cdot]I \qquad Pr(M, t) = \frac{1}{4}$$

- Context equivalence versus trace equivalence:
  - Sound and complete for CBN [].
  - Unsound for CBV.

Alternative Equivalente Definition: Evaluation Contexts

- Traces: contexts having a nice peculiar form

$$s ::= \epsilon \mid @V \cdot s.$$

- **Success probability of a trace**: $Pr(M, s)$.

Sucess Probability of a Trace.

- $s = \epsilon \quad \longrightarrow \quad \mathcal{C}_s = [\cdot]$
- $s = @V \cdot t \quad \longrightarrow \quad \mathcal{C}_s = \mathcal{C}_t[[\cdot]V]$
- $s = \otimes L \cdot t \quad \longrightarrow \quad \mathcal{C}_s = \mathcal{C}_t[\texttt{let } \langle x, y\rangle = [\cdot] \texttt{ in } L]$

$$t = @I \cdot \epsilon \qquad \mathcal{C}_t = [\cdot]I \qquad Pr(M, t) = \frac{1}{4}$$

- Context equivalence versus trace equivalence:
  - Sound and complete for CBN [].
  - Unsound for CBV.

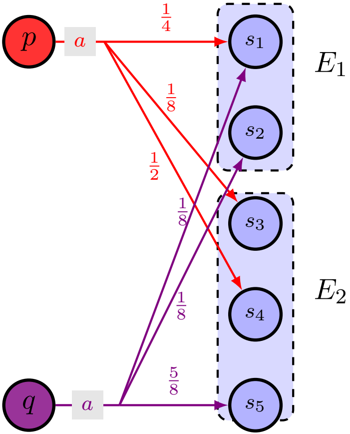
Applicative Bisimulation for a Probabilistic Language (2): Probabilistic Bisimulation

Labelled Markov Chain (LMC): a triple $\mathscr{M} = (\mathcal{S}, \mathcal{L}, \mathcal{P})$, where

- $\mathcal{S}$ is a countable set of *states*;
- $\mathcal{L}$ is a set of *labels*;
- $\mathcal{P}$ is a *transition probability matrix*, i.e., a function $\mathcal{P} : \mathcal{S} \times \mathcal{L} \times \mathcal{S} \to [0, 1]$ such that for every state s and for every label l, $\mathcal{P}(\mathcal{S}, l, t) = \sum_{t \in \mathcal{S}} \mathcal{P}(s, l, t) \leq 1$;

Bisimilarity (probabilistic case)

Let $(\mathcal{S}, \mathcal{L}, \mathcal{P})$ be a LMC (Labelled Markov Chain).



Bisimulation: R such that

- R equivalence relation on $\mathcal{S}$.
- $(p,q) \in R \Rightarrow$ for every equivalence class E, $a \in \mathcal{L}$,

$$\sum_{s \in E} \mathcal{P}(p,a,s) = \sum_{s \in E} \mathcal{P}(q,a,s)$$

.

A Labelled Markov Chain for $\Lambda_{\oplus}$

Terms **Values**

M

A Labelled Markov Chain for $\Lambda_{\oplus}$

Terms **Values**

$$
\begin{tikzcd}
M \arrow[r, "{eval,\ \llbracket M \rrbracket(V)}"] \arrow[dr, "{eval,\ \llbracket M \rrbracket(W)}"] \arrow[ddr, "{eval,\ \llbracket M \rrbracket(U)}"'] & V \\
& W \\
& U \\
& \vdots
\end{tikzcd}
$$

A Labelled Markov Chain for $\Lambda_{\oplus}$

Terms	**Values**
	$\lambda x.N$

A Labelled Markov Chain for $\Lambda_\oplus$

Terms **Values**

$$N\{W/x\} \xleftarrow{W,\ 1} \lambda x.N$$

Context Equivalence vs. Bisimulation

- **Contexts**:
  $\mathcal{C} ::= [] \mid \lambda x.\mathcal{C} \mid \mathcal{C}M \mid M\mathcal{C} \mid M \oplus \mathcal{C} \mid \mathcal{C} \oplus M.$
- **Context Equivalence**: $M \equiv N$ iff for every context C it holds that $\sum[\![\mathcal{C}[M]]\!] = \sum[\![\mathcal{C}[N]]\!]$.

Theorem

$\sim$ is included in $\equiv$.

Lemma

$\sim$ is a congruence.

- $M \sim N \implies C[M] \sim C[N]$
- Howe's technique.

Howe's Technique

$\mathcal{R}$

$\mathcal{R}^H$

Howe's Technique

$$\begin{tikzcd}
\mathcal{R} \arrow[rr, bend left, "\subseteq"] & & \mathcal{R}^H
\end{tikzcd}$$

Howe's Technique

$\mathcal{R}^H$ is a
Congruence
whenever $\mathcal{R}$ is
an equivalence

$$\begin{tikzcd} \mathcal{R} \arrow[rr, bend left, "\subseteq"] & & \mathcal{R}^H \end{tikzcd}$$

Howe's Technique

$\sim^H$ is a
Congruence

$$\begin{tikzcd} \sim \arrow[r, bend left, "\subseteq"] & \sim^H \end{tikzcd}$$

$\sim^H$ is a **Congruence**

$$\begin{tikzcd} \sim \arrow[rr, bend left, "\subseteq"] & & \sim^H \arrow[ll, bend left, "\supseteq"] \end{tikzcd}$$

Key Lemma

Full Abstraction?

- $\sim$ is a **sound** methodology for program equivalence.
- Is it also **complete**?
- **CBN**: **No** [DLSA2014]
  - Counterexample:

$$M = \lambda x.\lambda y.(\Omega \oplus I); \qquad N = \lambda x.(\lambda y.\Omega) \oplus (\lambda y.I).$$

Full Abstraction?

- $\sim$ is a **sound** methodology for program equivalence.
- Is it also **complete**?
- **CBN**: **No** [DLSA2014]
  - Counterexample:

$$M = \lambda x.\lambda y.(\Omega \oplus I); \qquad N = \lambda x.(\lambda y.\Omega) \oplus (\lambda y.I).$$

  - Of course, $I \not\sim \Omega$ and as a consequence

$$\lambda y.\Omega \not\sim \lambda y.I \not\sim \lambda y.(\Omega \oplus I) \implies M \not\sim N.$$

Full Abstraction?

- $\sim$ is a **sound** methodology for program equivalence.
- Is it also **complete**?
- **CBN**: **No** [DLSA2014]
  - Counterexample:

$$M = \lambda x.\lambda y.(\Omega \oplus I); \qquad N = \lambda x.(\lambda y.\Omega) \oplus (\lambda y.I).$$

  - Of course, $I \not\sim \Omega$ and as a consequence

$$\lambda y.\Omega \not\sim \lambda y.I \not\sim \lambda y.(\Omega \oplus I) \implies M \not\sim N.$$

  - On the other hand, $M \equiv N$.
    - We need a CIU-Theorem for that.

Full Abstraction?

- $\sim$ is a **sound** methodology for program equivalence.
- Is it also **complete**?
- **CBN**: **No** [DLSA2014]
  - Counterexample:
    $$M = \lambda x.\lambda y.(\Omega \oplus I); \qquad N = \lambda x.(\lambda y.\Omega) \oplus (\lambda y.I).$$
  - Of course, $I \not\sim \Omega$ and as a consequence
    $$\lambda y.\Omega \not\sim \lambda y.I \not\sim \lambda y.(\Omega \oplus I) \implies M \not\sim N.$$
  - On the other hand, $M \equiv N$.
    - We need a CIU-Theorem for that.
- **CBV**: **Yes**[CDL2014]

Theorem

$\equiv$ *is fully abstract*

Proof:

- Bisimulation in a LMC is characterized by a test language: $t ::= \omega \mid a \cdot t \mid < t, t >$;
- Tests can be simulated by CBV-contexts.

Our Neighborhood

- Λ, where we observe **convergence**

	$\sim\ \subseteq\ \equiv$	$\equiv\ \subseteq\ \sim$	$\precsim\ \subseteq\ \leq$	$\leq\ \subseteq\ \precsim$
CBN	✓	✓	✓	✓
CBV	✓	✓	✓	✓

[Abramsky1990,Howe1993]

- $\Lambda_{\oplus}$ with nondeterministic semantics, where we observe **convergence**, in its **may** or **must** flavors.

	$\sim\ \subseteq\ \equiv$	$\equiv\ \subseteq\ \sim$	$\precsim\ \subseteq\ \leq$	$\leq\ \subseteq\ \precsim$
CBN	✓	×	✓	×
CBV	✓	×	✓	×

[Ong1993,Lassen1998]

Summing up equivalence result

	$\sim\ \subseteq\ \equiv$	$\equiv\ \subseteq\ \sim$	$\precsim\ \subseteq\ \leq$	$\leq\ \subseteq\ \precsim$
CBN	✓	×	✓	×
CBV	✓	✓	✓	×

Toward a notion of distance

- Two very similar probabilistic programs:

$$M = \Omega ::= (\lambda x.xx)(\lambda x.xx);$$
$$N = \Omega \oplus^{\epsilon} \lambda x.x \qquad \text{where } \epsilon \ll 1.$$

- We want to express the fact that M and N are **not** equivalent, but have a very **similar** behaviour.
- **Context Distance**:

$$\delta^{\mathrm{ctx}}(M,N) = \sup_{\mathcal{C} \text{ a context}} |Obs(\mathcal{C}[M]) - Obs(\mathcal{C}[N])|$$

- M and N are at context distance ϵ.

The Trivialization Phaenomenon

- Two other very similar programs:

$$M = I ::= \lambda x.x$$
$$N = I \oplus^{\epsilon} \Omega \qquad \text{with } \epsilon \ll 1$$

- We can construct a sequence of **amplification contexts** $\mathcal{C}_n$ such that:

$$Obs(\mathcal{C}_n[M]) = 1 \qquad Obs(\mathcal{C}_n[N]) = (1-\varepsilon)^n$$

$$\mathcal{C}_n = (\lambda x.\underbrace{(xI)\cdots(xI)}_{n})[\cdot]$$

- M and N are at context distance 1.
- In a language with **copying capabilities**, the contextual metric is itself too discriminating.

$\Lambda_\oplus$: An Affine λ-Calculus

- Every function uses its argument **at most once**.
- Low expressivity, but interesting structure of contextual metric.
- Syntax of the calculus:

$$M ::= x \;\Big|\; \lambda x.M \;\Big|\; MM \;\Big|\; M \oplus M \;\Big|\; \Omega$$
$$\Big|\; \texttt{let}\ \langle x, y\rangle = M\ \texttt{in}\ M \;\Big|\; \langle M, M\rangle.$$

We restrict ourselves to affine terms.

$\Lambda_\oplus$: An Affine λ-Calculus

- Every function uses its argument **at most once**.
- Low expressivity, but interesting structure of contextual m
- S

W

Rules for Affinity–Selection:

$$\frac{}{\Gamma, x \vdash x} \qquad \frac{\Gamma, x \vdash M}{\Gamma \vdash \lambda x.M} \qquad \frac{\Gamma \vdash M \qquad \Delta \vdash N}{\Gamma, \Delta \vdash MN}$$

$$\frac{\Gamma \vdash M \qquad \Gamma \vdash N}{\Gamma \vdash M \oplus N} \qquad \frac{}{\Gamma \vdash \Omega}$$

$\Lambda_\oplus$: An Affine λ-Calculus

- Every function uses its argument **at most once**.
- Low expressivity, but interesting structure of contextual metric.
- Syntax of the calculus:

$$M ::= x \;\Big|\; \lambda x.M \;\Big|\; MM \;\Big|\; M \oplus M \;\Big|\; \Omega$$
$$\Big|\; \texttt{let}\ \langle x, y\rangle = M\ \texttt{in}\ M \;\Big|\; \langle M, M\rangle.$$

We restrict ourselves to affine terms.

$\Lambda_\oplus$: An Affine λ-Calculus

- Every function uses its argument **at most once**.
- Low expressivity, but interesting structure of contextual metric.
- Syntax of the calculus:

$$M ::= x \;\Big|\; \lambda x.M \;\Big|\; M M \;\Big|\; M \oplus M \;\Big|\; \Omega$$
$$\Big|\; \mathtt{let}\ \langle x, y\rangle = M\ \mathtt{in}\ M \;\Big|\; \langle M, M\rangle.$$

  We restrict ourselves to affine terms.
- We define an operationnal semantics.
- Values: $V ::= \lambda x.M \;\Big|\; \langle M, M\rangle$
  Result of a computation: **sub-distribution** of values:
  $M \Downarrow [\![M]\!]$

Context Distance for $\Lambda_\oplus$

- Context: affine term with a hole.
- **Context Distance:**

$$\delta^{\text{ctx}}(M, N) = \sup_{\mathcal{C} \text{ a context}} |\sum_V [\![\mathcal{C}[M]]\!](V) - \sum_V [\![\mathcal{C}[N]]\!](V)|.$$

Context Distance for $\Lambda_\oplus$

- Context: affine term with a hole.
- **Context Distance:**

$$\delta^{\mathrm{ctx}}(M,N) = \sup_{\mathcal{C}\text{ a context}} |\sum_V [\![\mathcal{C}[M]]\!](V) - \sum_V [\![\mathcal{C}[N]]\!](V)|.$$

- The Context Distance is a pseudo-metric.

Context Distance for $\Lambda_\oplus$

- Context: affine term with a hole.
- **Context Distance:**

Pseudo-Metric on $\Lambda_\oplus$:

$\mu : \Lambda_\oplus \times \Lambda_\oplus \to [0,1]$ such that:

- Symmetry;
- $\mu(M, M) = 0$;
- Triangular inequality.

- T

Context Distance for $\Lambda_{\oplus}$

- Context: affine term with a hole.
- **Context Distance:**

$$\delta^{\text{ctx}}(M, N) = \sup_{\mathcal{C} \text{ a context}} |\sum_{V} [\![\mathcal{C}[M]]\!](V) - \sum_{V} [\![\mathcal{C}[N]]\!](V)|.$$

- The Context Distance is a pseudo-metric.

Context Distance for $\Lambda_\oplus$

- Context: affine term with a hole.
- **Context Distance:**

$$\delta^{\mathrm{ctx}}(M,N) = \sup_{\mathcal{C} \text{ a context}} |\sum_V [\![\mathcal{C}[M]]\!](V) - \sum_V [\![\mathcal{C}[N]]\!](V)|.$$

- The Context Distance is a pseudo-metric.
- Running examples:

$$\begin{aligned}
M_1 &= \Omega \oplus I & N_1 &= \Omega \\
M_2 &= (\lambda x.\Omega) \oplus (\lambda x.I) & N_2 &= \lambda x.(\Omega \oplus I) \\
M_3 &= \langle \lambda x.I, \lambda x.I \rangle & N_3 &= \langle \lambda x.(I \oplus \Omega), \lambda x.(I \oplus \Omega) \rangle.
\end{aligned}$$

Context Distance for $\Lambda_\oplus$

- Context: affine term with a hole.
- C
- T
- R

Context distance for $M_1 = \Omega \oplus I$ and $N_1 = \Omega$ is $\frac{1}{2}$

- $\delta^{\mathrm{ctx}}(M_1, N_1) \geq \frac{1}{2}$

Context Distance for $\Lambda_\oplus$

- Context: affine term with a hole.
- C
- T
- R

Context distance for $M_1 = \Omega \oplus I$ and $N_1 = \Omega$ is $\frac{1}{2}$

- $\delta^{\mathrm{ctx}}(M_1, N_1) \geq \frac{1}{2}$
- $|\sum_V [\![\mathcal{C}[M_1]]\!](V) - \sum_V [\![\mathcal{C}[N_1]]\!](V)| \leq \frac{1}{2}$:

Proof: It holds that:

$$\begin{array}{rcl} [\![\mathcal{C}[M_1]]\!] & = & \sum p_i \cdot \mathscr{D}_i; \\ [\![\mathcal{C}[N_1]]\!] & = & \sum p_i \cdot \mathscr{E}_i; \end{array}$$

where

- either $\sum_V \mathscr{D}_i(V) \leq 1/2$ and $\mathscr{E}_i = \emptyset$,
- either $\mathscr{D}_i = \{\mathcal{E}_i[M_1]^1\}$ and $\mathscr{E}_i = \{\mathcal{F}_i[N_1]^1\}$.

Context Distance for $\Lambda_\oplus$

- Context: affine term with a hole.
- **Context Distance:**

$$\delta^{\mathrm{ctx}}(M,N) = \sup_{\mathcal{C} \text{ a context}} |\sum_{V} [\![\mathcal{C}[M]]\!](V) - \sum_{V} [\![\mathcal{C}[N]]\!](V)|.$$

- The Context Distance is a pseudo-metric.
- Running examples:

$$\begin{aligned} M_1 &= \Omega \oplus I & N_1 &= \Omega \\ M_2 &= (\lambda x.\Omega) \oplus (\lambda x.I) & N_2 &= \lambda x.(\Omega \oplus I) \\ M_3 &= \langle \lambda x.I, \lambda x.I \rangle & N_3 &= \langle \lambda x.(I \oplus \Omega), \lambda x.(I \oplus \Omega) \rangle. \end{aligned}$$

Context Distance for $\Lambda_\oplus$

- Context: affine term with a hole.
- **Context Distance:**

$$\delta^{\mathrm{ctx}}(M,N) = \sup_{\mathcal{C} \text{ a context}} |\sum_V [\![\mathcal{C}[M]]\!](V) - \sum_V [\![\mathcal{C}[N]]\!](V)|.$$

- The Context Distance is a pseudo-metric.
- Running examples:

$$\begin{aligned} M_1 &= \Omega \oplus I & N_1 &= \Omega \\ M_2 &= (\lambda x.\Omega) \oplus (\lambda x.I) & N_2 &= \lambda x.(\Omega \oplus I) \\ M_3 &= \langle \lambda x.I, \lambda x.I \rangle & N_3 &= \langle \lambda x.(I \oplus \Omega), \lambda x.(I \oplus \Omega) \rangle. \end{aligned}$$

- **Unniversal quantification** over all contexts $\Longrightarrow$ Difficult to actually know distance between terms.

Trace Distance

- Generalisation of trace equivalence
- Generalisation of the Trace LTS: labels $=$ $@\mathbf{V} \cup \otimes L \cdot s \qquad x, y \vdash L$

$$w(\mathscr{D}) = \textstyle\sum_V \mathscr{D}(V)$$

$$\begin{tikzcd}
\mathscr{D} \arrow[r, "M"] \arrow[dr, "N"'] & \sum_V \mathscr{D}(V) \cdot [\![V M]\!] \\
 & \sum_V \mathscr{D}(V) \cdot [\![V N]\!] \\
 & \vdots
\end{tikzcd}$$

- R_ε: The biggest bisimulation R on this LTS such that $\mathscr{D} R \mathscr{E} \Rightarrow |w(\mathscr{D}) - w(\mathscr{E})| \leq \epsilon$.
- **Trace Distance**: $\delta^{\mathrm{tr}}(M, N) = \inf\{\epsilon \mid \{M^1\} R_\epsilon \{N^1\}\}$
- **Characterisation by traces**:

$$\delta^{\mathrm{tr}}(M, N) = \sup_{s \text{ a trace}} |Pr(M, s) - Pr(N, s)|$$

Trace Distance for $\Lambda_{\oplus}$

Theorem (Non-expansiveness)

For every M, N, and for every context $\mathcal{C}$:

$$\delta^{tr}(\mathcal{C}[M], \mathcal{C}[N]) \leq \delta^{tr}(M, N).$$

Trace Distance for $\Lambda_{\oplus}$

Theorem (Non-expansiveness)

For every M, N, and for every context $\mathcal{C}$:

$$\delta^{tr}(\mathcal{C}[M],\mathcal{C}[N]) \leq \delta^{tr}(M,N).$$

Theorem (Full Abstraction)

Trace distance and context distance coincide.

Trace Distance for $\Lambda_\oplus$

Theorem (Non-expansiveness)

For every M, N, and for every context $\mathcal{C}$:

$$\delta^{tr}(\mathcal{C}[M], \mathcal{C}[N]) \leq \delta^{tr}(M, N).$$

Theorem (Full Abstraction)

Trace distance and context distance coincide.

Proof:

- Adequacy: $\delta^{\mathrm{tr}}(M, N) \geq |\mathrm{Obs}(M) - \mathrm{Obs}(N)|$.
- **Soundness**: $\delta^{\mathrm{ctx}}(M, N) \leq \delta^{\mathrm{tr}}(M, N)$.

  Proof: $\left.\begin{array}{l}\text{Adequacy}\\ \text{Non-Expansiveness}\end{array}\right\} \Longrightarrow \text{Soundness}$.
- **Completness**:$\delta^{\mathrm{tr}}(M, N) \leq \delta^{\mathrm{ctx}}(M, N)$.

  Proof: Every trace can be simulated by a context.

Trace Distance for $\Lambda_\oplus$

Theorem (Non-expansiveness)

For every M, N, and for every context C:

Th

Tr

Pr

Proof Of Non-Expansiveness Theorem

- Plain induction on contexts fails.
- We need an operationnal semantics for terms in contexts:

$$\mathscr{D} \overset{s}{\Rightarrow}_{\mathbf{C}\times\mathbf{P}} \mathscr{E}$$

for $\mathscr{D}, \mathscr{E} \in \mathrm{Distr}\,(\mathrm{Terms} \times \mathbf{C})$

- Stronger notion of proximity: $\mathscr{D}\,\Delta_\varepsilon\,\mathscr{E}$
- $$\left.\begin{array}{l} \mathscr{D}\,\Delta_\varepsilon\,\mathscr{E} \\ \mathscr{D} \overset{s}{\Rightarrow}_{\mathbf{C}\times\mathbf{P}} \mathscr{F} \\ \mathscr{E} \overset{s}{\Rightarrow}_{\mathbf{C}\times\mathbf{P}} \mathscr{G} \end{array}\right\} \Longrightarrow \mathscr{F}\,\Delta_\varepsilon \mathscr{G}$$

Trace Distance on Examples

- $M_1 = \Omega \oplus I$ and $N_1 = \Omega$

For every trace s:

- $Pr(M_1, s) = \frac{1}{2} \cdot Pr(I, s)$
- $Pr(N_1, s) = 0$

The trace distance between M_1 and N_1 is $\frac{1}{2}$.

Trace Distance on Examples

- $M_1 = \Omega \oplus I$ and $N_1 = \Omega$

For every trace s:

- $Pr(M_1, s) = \frac{1}{2} \cdot Pr(I, s)$
- $Pr(N_1, s) = 0$

The trace distance between M_1 and N_1 is $\frac{1}{2}$.

- $M_2 = (\lambda x.\Omega) \oplus (\lambda x.I)$ and $N_2 = \lambda x.(\Omega \oplus I)$.
  - Traces don't take branching into account (evaluation context).
  - For every trace s, $Pr(M, s) = Pr(N, s)$.
  - The trace distance between M_2 and N_2 is 0.

Trace Distance on Examples

- $M_1 = \Omega \oplus I$ and $N_1 = \Omega$

For every trace s:

  - $Pr(M_1, s) = \frac{1}{2} \cdot Pr(I, s)$
  - $Pr(N_1, s) = 0$

The trace distance between M_1 and N_1 is $\frac{1}{2}$.

- $M_2 = (\lambda x.\Omega) \oplus (\lambda x.I)$ and $N_2 = \lambda x.(\Omega \oplus I)$.
  - Traces don't take branching into account (evaluation context).
  - For every trace s, $Pr(M, s) = Pr(N, s)$.
  - The trace distance between M_2 and N_2 is 0.

- $M_3 = \langle \lambda x.I, \lambda x.I \rangle$ and $N_3 = \langle \lambda x.(I \oplus \Omega), \lambda x.(I \oplus \Omega) \rangle$.
  - We have to consider traces of the form $\otimes L \cdot \epsilon$.
  - Not simpler than contextual distance.

Bisimulation Metric

- Generalization of bisimilarity to metrics.
- We define a Labelled Markov Chain: the $\Lambda_\oplus$ LMC.
- Operator F on metrics [DGJP2002]:

$$F(\mu)(s,t) = \sup_a \{\overline{\mu}(\mathscr{D},\mathscr{E}) \mid a \in Act,\ s \xrightarrow{a} \mathscr{D},\ t \xrightarrow{a} \mathscr{E}\}$$

Proposition

F has a greatest fixpoint, called bisimulation distance: δ^b

Bisimulation Metric

- Generalization of bisimilarity to metrics.
- We define a Labelled Markov Chain: the $\Lambda_\oplus$ LMC.
- Operator F on metrics [DGJP2002]:

$$F(\mu)(s,t) = \sup_a \{\overline{\mu}(\mathscr{D},\mathscr{E}) \mid a \in Act,\ s \xrightarrow{a} \mathscr{D},\ t \xrightarrow{a} \mathscr{E}\}$$

Proposition

F has a greatest fixpoint, called bisimulation distance: δ^b

Theorem

δ^b is non-expansive, thus sound with respect to context distance.

Bisimulation Metric

- Generalization of bisimilarity to metrics.
- We define a Labelled Markov Chain: the $\Lambda_{\oplus}$ LMC.
- Operator F on metrics [DGJP2002]:

$$F(\mu)(s,t) = \sup_a \{\overline{\mu}(\mathscr{D},\mathscr{E}) \mid a \in Act,\ s \xrightarrow{a} \mathscr{D},\ t \xrightarrow{a} \mathscr{E}\}$$

Proposition

F has a greatest fixpoint, called bisimulation distance: δ^b

Theorem

δ^b is non-expansive, thus sound with respect to context distance.

Proposition

δ^b is not complete.

Proof: $\delta^{\mathrm{b}}(M_2, N_2) = \frac{1}{2}$. However, M_2 and N_2 are at context distance 0.

Bisimulation Metric

- Generalization of bisimilarity to metrics.
- We define a Labelled Markov Chain: the $\Lambda_\oplus$ LMC.
- Operator F on metrics [DGJP2002]:

Proof Of Non-Expansiveness Theorem

- probabilistic and quantitative variation of Howe's method;
- uses Kantorovitch's duality in a crucial way.

Pr

F

Theorem

δ^b is non-expansive, thus sound with respect to context distance.

Proposition

δ^b is not complete.

Proof: $\delta^{\mathrm{b}}(M_2, N_2) = \frac{1}{2}$. However, M_2 and N_2 are at context distance 0.

The Tuple Distance

Extending the $\Lambda_{\oplus}$-LMC:

- Goal: Simplifying the handling of pairs in the $\Lambda_{\oplus}$-LMC.
- States: sequences of values (tuples).

The Tuple Distance

Extending the $\Lambda_{\oplus}$-LMC:

- Goal: Simplifying the handling of pairs in the $\Lambda_{\oplus}$-LMC.
- States: sequences of values (tuples).
- Actions: $\mathrm{unfold}^i \;\Big|\; @(\Gamma, V)^i$ with $i \in \mathbb{N}$, and $\Gamma \vdash V$.

$$\boxed{[\langle M, N\rangle, U]} \xrightarrow[\llbracket M \rrbracket(V) \cdot \llbracket N \rrbracket(W)]{\mathrm{unfold}^1} \boxed{[V, W, U]}$$

$$\boxed{[\lambda x.M, W, U]} \xrightarrow[\llbracket M\{V\{W/x_2\}\}/x\} \rrbracket(T)]{@(\{x_2\}, V)^1} \boxed{[T, U]}$$

- Tuple Distance:

$$\delta^{\mathrm{mul}}(M, N) = \sup_s |Pr([M], s) - Pr([N], s)|$$

Theorem

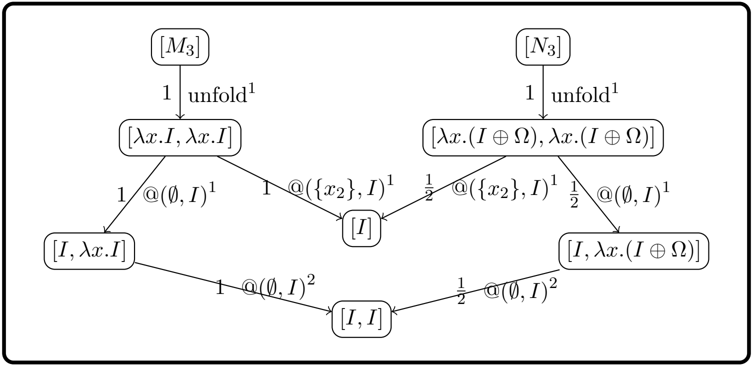
- *δ^{mul} is non-expansive.*
- *δ^{mul} is fully abstract with respect to context distance.*

Example for Tuple Distance

$$M_3 = \langle \lambda x.I, \lambda x.I \rangle \qquad \text{and} \qquad N_3 = \langle \lambda x.(I \oplus \Omega), \lambda x.(I \oplus \Omega) \rangle$$

Proposition

$\delta^{mul}(M_3, N_3) = \frac{3}{4}$



Conclusion

- **Contribution**:
  - The first account of behavioural metrics in higher-order languages.
  - Satisfactory proof techniques for pairs.

Conclusion

- **Contribution**:
  - The first account of behavioural metrics in higher-order languages.
  - Satisfactory proof techniques for pairs.
- **Related Work**:
  - Metrics in probabilistic systems [DGJP02, ...]
  - Metrics for process algebras [GT14]
  - Higher-order probabilistic languages [JP89,EPT14, ...]
    - $\Lambda_{\oplus}$ as a LMC [DLSA14,CDL14]

Conclusion

- **Contribution**:
  - The first account of behavioural metrics in higher-order languages.
  - Satisfactory proof techniques for pairs.
- **Related Work**:
  - Metrics in probabilistic systems [DGJP02, ...]
  - Metrics for process algebras [GT14]
  - Higher-order probabilistic languages [JP89,EPT14, ...]
    - $\Lambda_{\oplus}$ as a LMC [DLSA14,CDL14]
- **Further Work**:
  - The Non-Affine Case: Extending the tuple LMC to a calculus with copying capabilities.
  - Application to cryptography: computational indistinguishability.

Thank you.

Questions ?