Separability in probabilistic λ-calculus

Thomas LEVENTIS

I2M (Institut de Mathématiques de Marseille)

September 3, 2016

1. Introduction : deterministic separability
2. Call-by-name probabilistic λ-calculus
   - The calculus
   - Probabilistic Böhm trees
3. Probabilistic separation
   - Reading trees
   - Processing probabilities

λ-calculus

Definition

We define the λ-terms by

$$M, N := x \mid \lambda x.M \mid M\ N.$$

The contexts of terms are given by

$$C := [\,] \mid \lambda x.C \mid C\ M \mid M\ C.$$

The β-reduction is

$$(\lambda x.M)\ N \to_\beta M\left[{}^{N}/_{x}\right]$$

extended to contexts:

$$\text{if } M \to_\beta N \text{ then } C[M] \to_\beta C[N].$$

Strong separability

Definition

Two terms M and N are *separable* if for any two terms P and Q there is a context C such that

$$C[M] \twoheadrightarrow_\beta P$$
$$C[N] \twoheadrightarrow_\beta Q.$$

Strong separability

Definition

Two terms M and N are *separable* if for any two terms P and Q there is a context C such that

$$C[M] \twoheadrightarrow_\beta P$$
$$C[N] \twoheadrightarrow_\beta Q.$$

This is a strong property. For instance the term $\Omega = (\lambda x.x\ x)\ (\lambda x.x\ x)$ is not separable with any other term.

Head reduction

Definition

The *head reduction* is

$$\lambda x_1 ... x_n.(\lambda y.M)\ N\ P_1\ ...\ P_m \to_h \lambda x_1 ... x_n.M\left[{}^{N}/_{x}\right]\ P_1\ ...\ P_n.$$

Proposition

A term M is normal for the head reduction if and only if it is of the form

$$M = \lambda x_1 ... x_n.y\ P_1\ ...\ P_m.$$

Definition

A term is *solvable* if its head reduction normalizes, and it is *unsolvable* otherwise.

Observational equivalence

Definition

Two terms M and N are *semi-separable* if there is a context C such that $C[M]$ is solvable if and only if $C[N]$ is not.

Observational equivalence

Definition

Two terms M and N are *semi-separable* if there is a context C such that $C[M]$ is solvable if and only if $C[N]$ is not.

Definition

We say that two terms M and N are *observationally equivalent* if for every context C, $C[M]$ is solvable if and only if $C[N]$ is solvable. In other words M and N are observationally equivalent if and only if they are not semi-separable.

Observational equivalence

Definition

Two terms M and N are *semi-separable* if there is a context C such that $C[M]$ is solvable if and only if $C[N]$ is not.

Definition

We say that two terms M and N are *observationally equivalent* if for every context C, $C[M]$ is solvable if and only if $C[N]$ is solvable. In other words M and N are observationally equivalent if and only if they are not semi-separable.

Theorem (Hyland, Wadsworth)

Two terms are observationally equivalent if and only if they have the same infinitely extensional Böhm tree.

Probabilistic λ-calculus

Definition

We define the probabilistic λ-terms by

$$M, N := x \mid \lambda x.M \mid M\ N \mid M +_p N, p \in [0;1].$$

Probabilistic λ-calculus

Definition

We define the probabilistic λ-terms by

$$M, N := x \mid \lambda x.M \mid M\ N \mid M +_p N, p \in [0;1].$$

It is well know that call-by-name and call-by-value probabilistic λ-calculi are different: given a redex $(\lambda x.M)\ (N_1 +_p N_2)$ we must decide whether we duplicate the choice $N_1 +_p N_2$ or we choose once and duplicate the result.

Probabilistic λ-calculus

Definition

We define the probabilistic λ-terms by

$$M, N := x \mid \lambda x.M \mid M\ N \mid M +_p N, p \in [0;1].$$

It is well know that call-by-name and call-by-value probabilistic λ-calculi are different: given a redex $(\lambda x.M)\ (N_1 +_p N_2)$ we must decide whether we duplicate the choice $N_1 +_p N_2$ or we choose once and duplicate the result.

Here we are interested in the call-by-name calculus.

Definition

The reductions of probabilistic terms is given by

$$\lambda x_1 ... x_n.(\lambda y.M)\ N\ P_1\ ...\ P_m \to^1 \lambda x_1 ... x_n.M\left[{}^N/_x\right]\ P_1\ ...\ P_n$$
$$\lambda x_1 ... x_n.(M +_p N)\ P_1\ ...\ P_m \to^p \lambda x_1 ... x_n.M\ P_1\ ...\ P_n$$
$$\lambda x_1 ... x_n.(M +_p N)\ P_1\ ...\ P_m \to^{1-p} \lambda x_1 ... x_n.N\ P_1\ ...\ P_n.$$

We have a probabilistic reduction: the reduction $M \to^p N$ occurs with probability p.

Definition

For every term M and head normal form H, the probability that M reduces into H is

$$\mathcal{P}(M \twoheadrightarrow H) = \sum_{M \to^{p_1} \dots \to^{p_n} H} \prod_{i=1}^{n} p_i .$$

We have a probabilistic reduction: the reduction $M \to^p N$ occurs with probability p.

Definition

For every term M and head normal form H, the probability that M reduces into H is

$$\mathcal{P}(M \twoheadrightarrow H) = \sum_{M \to^{p_1} \dots \to^{p_n} H} \prod_{i=1}^{n} p_i .$$

Definition

The *convergence probability* of a term M is

$$\mathcal{C}\mathrm{onv}(M) = \sum_{H \text{ head normal form}} \mathcal{P}(M \twoheadrightarrow H).$$

Probabilistic separability

Definition

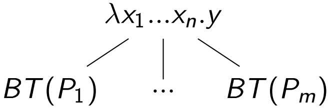
Two terms M and N are (semi-)separable if there is a context C such that $\mathcal{C}\mathrm{onv}(C[M]) \neq \mathcal{C}\mathrm{onv}(C[N])$.

Obviously if two deterministic terms (i.e. terms without sums) are semi-separable in a deterministic sense then they are separable in a probabilistic sense. What about the converse implication?

Deterministic Böhm trees

The deterministic Böhm tree $BT(M)$ of a deterministic term M is

Ω if M is unsolvable;

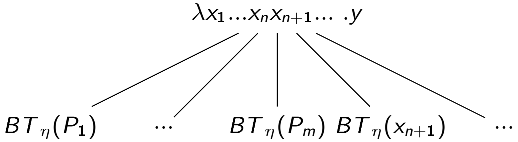


if $M \twoheadrightarrow_h \lambda x_1 \ldots x_n . y \; P_1 \; \ldots \; P_m$.

Deterministic Böhm trees

The infinitely extensional deterministic Böhm tree, or Nakajima tree, $BT_\eta(M)$ of a deterministic term M is

Ω if M is unsolvable;



if $M \twoheadrightarrow_h \lambda x_1 ... x_n.y\ P_1\ ...\ P_m$.

Probabilistic Böhm trees

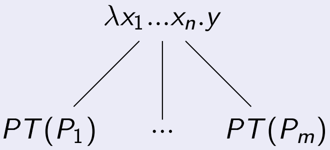
To extend the definition of Böhm trees to the probabilistic calculus we add a new layer.

Definition

The probabilistic Böhm tree $PT(M)$ of a term M is a subprobability distribution over value Böhm trees:

$$PT(M) : t \mapsto \sum_{H \text{ s.t. } VT(H)=t} \mathcal{P}(M \twoheadrightarrow H).$$

The value Böhm tree $VT(H)$ of a head normal form $H = \lambda x_1 ... x_n . y \; P_1 \; ... \; P_m$ is



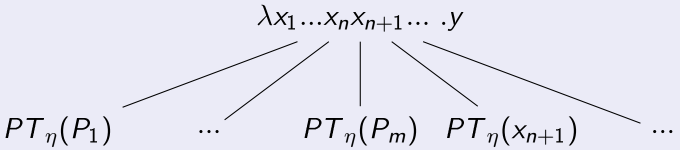
Probabilistic Böhm trees

Definition

The infinitely extensional probabilistic Böhm tree $PT_\eta(M)$ of a term M is:

$$PT_\eta(M) : t \mapsto \sum_{H \text{ s.t. } VT_\eta(H)=t} \mathcal{P}(M \twoheadrightarrow H).$$

The infinitely extensional value Böhm tree $VT_\eta(H)$ of a head normal form $H = \lambda x_1 \ldots x_n . y\ P_1\ \ldots\ P_m$ is



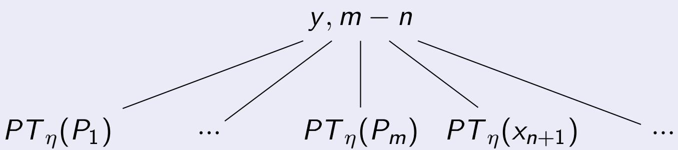
Probabilistic Böhm trees

Definition

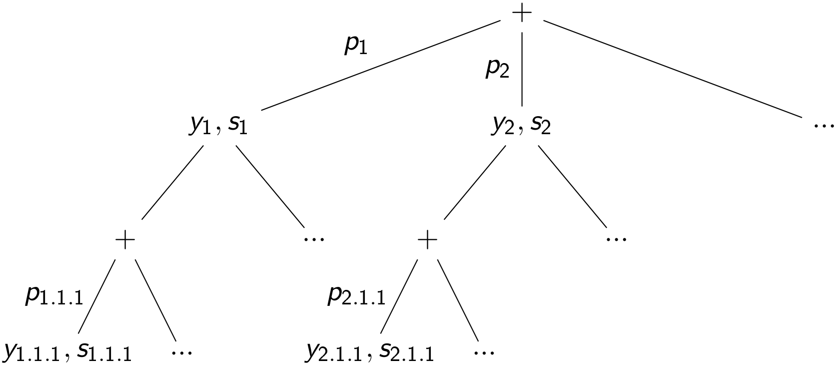
The infinitely extensional probabilistic Böhm tree $PT_\eta(M)$ of a term M is:

$$PT_\eta(M) : t \mapsto \sum_{H \text{ s.t. } VT_\eta(H)=t} \mathcal{P}(M \twoheadrightarrow H).$$

The infinitely extensional value Böhm tree $VT_\eta(H)$ of a head normal form $H = \lambda x_1 ... x_n . y \ P_1 \ ... \ P_m$ is



Infinitely extensional probabilistic Böhm trees are trees of the form:



Böhm trees form a model

Theorem

For any terms M and N and any context C:

- *if* $PT(M) = PT(N)$ *then* $PT(C[M]) = PT(C[N])$;
- *if* $PT_\eta(M) = PT_\eta(N)$ *then* $PT_\eta(C[M]) = PT_\eta(C[N])$.

Corollary

If $PT(M) = PT(N)$ *then* M *and* N *are observationally equivalent.*

Proof.

$$\mathcal{C}\mathrm{onv}(M) = \sum_t PT(M)(t).$$

Well-known deterministic techniques

Given

- finitely many head normal forms $H_i = \lambda x_1 ... x_{n_i}.y_i \; P_{i,1} \; ... \; P_{i,n_i+s_i}$,
- a family $(F_{y,s})_{(y,s)\in Var\times\mathbb{Z}}$ of terms
- and a family $(m_{y,s})_{(y,s)\in Var\times\mathbb{Z}}$ of integers,

we want to find a substitution $\sigma : Var \to \Lambda_+$ and terms $L_1,...,L_M$ such that

$$\sigma(H_i) \; L_1 \; ... \; L_M \twoheadrightarrow_h F_{y_i,s_i} \; \sigma(P'_{i,1}) \; ... \; \sigma(P'_{m_{y_i,s_i}})$$

with

$$\begin{array}{ll} P'_{i,j} = P_{i,j} & \text{for } j \leq n_i + s_i; \\ P'_{i,j} = x_{j-s_i} & \text{for } n_i + s_i < j \leq m_{y_i,s_i}. \end{array}$$

Well-known deterministic techniques

For $N \geq n_i$ we have:

$$H_i\ x_1\ \dots\ x_N = y_i\ P_{i,1}\ \dots\ P_{i,n_i+s_i}\ x_{n_i+1}\ \dots\ x_N$$

$$\sigma(H_i)\ \sigma(x_1)\ \dots\ \sigma(x_N) = \sigma(y_i)\ \sigma(P_{i,1})\ \dots\ \sigma(P_{i,n_i+s_i})\ \sigma(x_{n_i+1})\ \dots\ \sigma(x_N)$$

$$\sigma(H_i)\ \sigma(x_1)\ \dots\ \sigma(x_N)\ F_1\ \dots\ F_M = \sigma(y_i)\ \sigma(P_{i,1})\ \dots\ \sigma(P_{i,n_i+s_i})\ \sigma(x_{n_i+1})\ \dots\ \sigma(x_N)\ F_1\ \dots\ F_M.$$

Well-known deterministic techniques

For $N \geq n_i$ we have:

$$\begin{aligned} H_i\ x_1\ \dots\ x_N &= y_i\ P_{i,1}\ \dots\ P_{i,n_i+s_i}\ x_{n_i+1}\ \dots\ x_N \\ \sigma(H_i)\ \sigma(x_1)\ \dots\ \sigma(x_N) & \\ &= \sigma(y_i)\ \sigma(P_{i,1})\ \dots\ \sigma(P_{i,n_i+s_i})\ \sigma(x_{n_i+1})\ \dots\ \sigma(x_N) \\ \sigma(H_i)\ \sigma(x_1)\ \dots\ \sigma(x_N)\ F_1\ \dots\ F_M & \\ &= \sigma(y_i)\ \sigma(P_{i,1})\ \dots\ \sigma(P_{i,n_i+s_i})\ \sigma(x_{n_i+1})\ \dots\ \sigma(x_N)\ F_1\ \dots\ F_M. \end{aligned}$$

If $\sigma(y_i) = \lambda x_1 \dots x_{k_i}.x_{k_i}\ x_1\ \dots\ x_{k_i}$ with $1 \leq k_i - N - s_i \leq M$ then

$$\begin{aligned} &\sigma(H_i)\ \sigma(x_1)\ \dots\ \sigma(x_N)\ F_1\ \dots\ F_M \\ &\qquad \twoheadrightarrow_h F_{k_i-N-s_i}\ \sigma(P_{i,1})\ \dots\ \sigma(P_{i,n_i+s_i})\ \sigma(x_{n_i+1})\ \dots\ \sigma(x_N)\ F_1\ \dots\ F_M. \end{aligned}$$

Well-known deterministic techniques

For $n \in \mathbb{N}$ let $R_n = \lambda x_1 ... x_n x_{n+1}.x_{n+1}\ x_1\ ...\ x_n\ x_{n+1}$.
For $Y \subset Var$ a finite set of variables and $\delta \in \mathbb{N}$, $\Sigma_{Y,\delta}$ is the set of substitutions $\sigma : Var \to \Lambda^+$ such that:

- if $x \in Y$ then $\sigma(x) \neq x$;
- if $\sigma(x) \neq x$ then $\sigma(x) = R_n$ for some $n \geq \delta$;
- if $\sigma(x) = R_n$ and $\sigma(y) = R_m$ with $x \neq y$ then $|m - n| \geq \delta$.

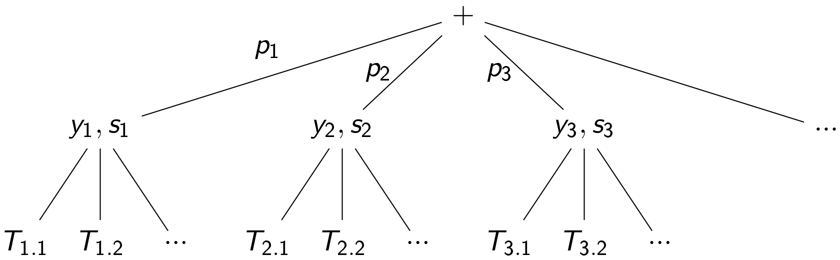
Proposition

Given finitely many head normal forms H_i,

$$\begin{gathered} \forall (m_{y,s}), \\ \exists Y \subset Var, \exists \delta \in \mathbb{N} : \\ \forall \sigma \in \Sigma_{Y,\delta}, \forall (F_{y,s}), \\ \exists L_1, ..., L_M : \\ \forall i,\ \sigma(H_i)\ L_1\ ...\ L_M \twoheadrightarrow_h F_{y_i,s_i}\ \sigma(P'_{i,1})\ ...\ \sigma(P'_{m_{y_i,s_i}}) \end{gathered}$$

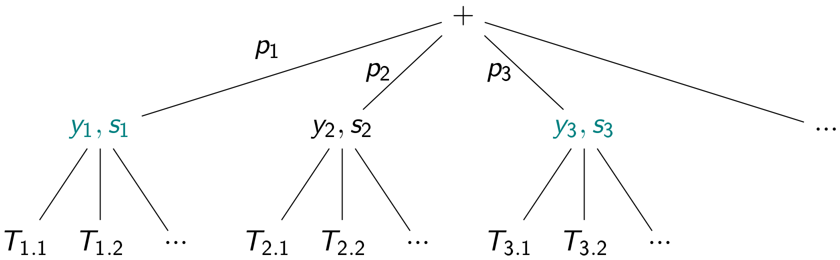
Given finitely many terms M_i and finitely many labels y_j, s_j along with terms F_j and integers m_j, we can replace the nodes labeled with (y_j, s_j) in $PT_\eta(M_i)$ by F_j applied to m_j arguments.

$$PT_\eta(M_i) =$$



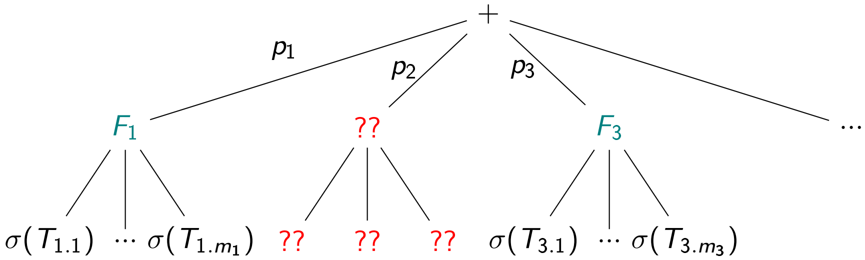
Given finitely many terms M_i and finitely many labels y_j, s_j along with terms F_j and integers m_j, we can replace the nodes labeled with (y_j, s_j) in $PT_\eta(M_i)$ by F_j applied to m_j arguments.

$PT_\eta(M_i) =$



Given finitely many terms M_i and finitely many labels y_j, s_j along with terms F_j and integers m_j, we can replace the nodes labeled with (y_j, s_j) in $PT_\eta(M_i)$ by F_j applied to m_j arguments.

$$"\sigma(M_i)\ L_1\ \dots\ L_M = "$$



This procedure can be done inductively.

- Given finitely many head normal forms $H_i = \lambda x_1 \ldots x_{n_i}.y_i \; P_{i,1} \; \ldots \; P_{i,n_i+s_i}$ and a family $(m_{y,s})$ we get conditions Y and δ.
- For every y_i, s_i and $k \leq m_{y_i,s_i}$, given finitely many head normal forms $H_{y_i,s_i,k,j}$ and a family $(m_{y,s,k,y',s'})$ we get conditions $Y_{y_i,s_i,k}$ and $\delta_{y_i,s_i,k}$.
- Let $Y' = Y \cup \left(\bigcup_{y_i,s_i,k} Y_{y_i,s_i,k}\right)$ and $\delta' = \max\,\{\delta\} \cup \left(\bigcup_{y_i,s_i,k}\{\delta_{y_i,s_i,k}\}\right)$, we have $\Sigma_{Y',\delta'} = \Sigma_{Y,\delta} \cap \left(\bigcap_{y_i,s_i,k_i} \Sigma_{Y_{y_i,s_i,k},\delta_{y_i,s_i,k}}\right)$.
- Let $\sigma \in \Sigma_{Y',\delta'}$, for every y_i, s_i and k and for every family of terms $(F_{y_i,s_i,k,y',s'})$ we are given a sequence of terms $\overrightarrow{L_{y_i,s_i,k}}$.
- Given a family $(F_{y,s})$ we consider a family $(F'_{y,s})$ such that $F'_{y_i,s_i} = \lambda z_1 \ldots z_{m_{y_i,s_i}}.F_{y_i,s_i} \; \left(z_1 \; \overrightarrow{L_{y_i,s_i,1}}\right) \; \ldots \; \left(z_{m_{y_i,s_i}} \; \overrightarrow{L_{y_i,s_i,m_{y_i,s_i}}}\right)$, we get a sequence of terms $\overrightarrow{L}$.
- Then given a term M, $\sigma(M) \; \overrightarrow{L}$ has the required form.

We can rewrite the nodes of a Böhm tree, but how do we extract a difference?

In the deterministic case we may consider only two terms, so we only need to use projections and constant functions:

- if a term is solvable and the other is not then they are already separated;
- if $M \twoheadrightarrow_h \lambda x_1 \ldots x_n . y \ P_1 \ \ldots \ P_{n+s}$ and $N \twoheadrightarrow_h \lambda x_1 \ldots x_{n'} . y' \ P'_1 \ \ldots \ P'_{n'+s'}$ with $(y, s) \neq (y', s')$ then we use constant functions to send M and N to any pair of terms;
- if $M \twoheadrightarrow_h \lambda x_1 \ldots x_n . y \ P_1 \ \ldots \ P_{n+s}$ and $N \twoheadrightarrow_h \lambda x_1 \ldots x_{n'} . y \ P'_1 \ \ldots \ P'_{n'+s}$ then $P_i \neq P'_i$ for some i (up to extensionality) and we use a projection to extract the difference.

But such a simple technique does not work in the probabilistic case. Consider for instance the terms

$$x \ y_1 \ z_1 +_{\frac{1}{2}} x \ y_2 \ z_2 \text{ and } x \ y_1 \ z_2 +_{\frac{1}{2}} x \ y_2 \ z_1 .$$

Goal

Let $I = \lambda x.x$, for $\epsilon > 0$ and $\alpha \in [0;1]$ we write $M \simeq_\epsilon I +_\alpha \Omega$ if

$$\alpha - \epsilon \leq \mathcal{P}(M \twoheadrightarrow_h I) \leq \mathcal{C}\mathrm{onv}(M) \leq \alpha + \epsilon.$$

We want to prove the following.

Theorem

Given finitely many terms M_i there are $\alpha_i \in [0;1]$ such that $\alpha_i = \alpha_j$ if and only if $PT_\eta(M_i) = PT_\eta(M_j)$, and such that

$$\forall \epsilon > 0, \exists C : \forall i, C[M_i] \simeq_\epsilon I +_{\alpha_i} \Omega.$$

Proposition

Given any evaluation function, i.e. any function $\varphi : [0;1]^n \to [0;1]$ *built as a composition of constant functions, projections, products and probabilistic sums, there is a term* $\underline{\varphi}$ *such that*

$$\underline{\varphi}\ (I +_{\alpha_1} \Omega) \ldots (I +_{\alpha_n} \Omega) \twoheadrightarrow_h I +_{\psi(\alpha_1,\ldots,\alpha_n)} \Omega.$$

Besides this representation is uniformly continuous: for all $\epsilon > 0$ *there is* $\delta > 0$ *such that for all* $\alpha_1, \ldots \alpha_n$ *if* $M_i \simeq_\delta I +_{\alpha_i} \Omega$ *for* $i \leq n$ *then*

$$\underline{\varphi}\ M_1 \ldots M_n \simeq_\epsilon I +_{\psi(\alpha_1,\ldots,\alpha_n)} \Omega.$$

Proof.

- If $M \simeq_\epsilon I +_\alpha \Omega$ and $N \simeq_{\epsilon'} I +_\beta \Omega$ then $M\ N \simeq_{2\epsilon+\epsilon'} I +_{\alpha\beta} \Omega$.
- If $M \simeq_\epsilon I +_\alpha \Omega$ and $N \simeq_{\epsilon'} I +_\beta \Omega$ then $M +_p N \simeq_{p\epsilon+(1-p)\epsilon'} I +_{p\alpha+(1-p)\beta} \Omega$.

Evaluation of trees

Definition

An *evaluation structure* $\mathcal{E}$ is given by:

- finitely many pairs (y_i, s_i);
- evaluation functions φ_i with arities m_i;
- evaluations structures $\mathcal{S}_{i,k}$ for $k \leq m_i$.

Definition

The evaluation of a tree is

$$\begin{aligned}
\mathcal{S}(T) &= \sum_t T(t).\mathcal{S}(t) & \\
\mathcal{S}\left(\begin{array}{c} y_i, s_i \\ \swarrow \quad \downarrow \quad \searrow \\ T_1 \quad T_2 \quad \dots \end{array} \right) &= \varphi_i(\mathcal{S}_{i,1}(T_1), \dots, \mathcal{S}_{i,m_i}(T_{m_i})) & \\
\mathcal{S}(t) &= 0 & \text{otherwise.}
\end{aligned}$$

Separation

Proposition

Given finitely many terms M_i and an evaluations structure $\mathcal{S}$

$$\forall \epsilon, \exists C : \forall i, C[M_i] \simeq_\epsilon I +_{\mathcal{S}(PT_\eta(M_i))} \Omega.$$

Separation

Proposition

Given finitely many terms M_i and an evaluations structure $\mathcal{S}$

$$\forall\epsilon, \exists C : \forall i, C[M_i] \simeq_\epsilon I +_{\mathcal{S}(PT_\eta(M_i))} \Omega.$$

Proposition

Given finitely many trees T_i there is an evaluation structure $\mathcal{S}$ such that

$$\mathcal{S}(T_i) = \mathcal{S}(T_j) \text{ if and only if } T_i = T_j.$$

Separation

Proposition

Given finitely many terms M_i and an evaluations structure $\mathcal{S}$

$$\forall \epsilon, \exists C : \forall i, C[M_i] \simeq_\epsilon I +_{\mathcal{S}(PT_\eta(M_i))} \Omega.$$

Proposition

Given finitely many trees T_i there is an evaluation structure $\mathcal{S}$ such that

$$\mathcal{S}(T_i) = \mathcal{S}(T_j) \text{ if and only if } T_i = T_j.$$

We prove the second proposition by induction on the maximal depth of the differences.

Assume given finitely many pairwise distinct trees T_i.

- We can find finitely many value trees $t_j = \begin{array}{ccc} & y_j, s_j & \\ / & \mid & \backslash \\ T_{j,1} & T_{j,2} & \dots \end{array}$ for $j \in J$ such that for all $i \neq i'$ we have $T_i(t_j) \neq T_{i'}(t_j)$ for some j.
- We can find an integer m such that the sums

$$T_i^{(m)}(t_j) = \sum_{U_1, U_2, \dots} T_i \left(\begin{array}{cccccc} & & y_j, s_j & & & \\ / & / & \mid & \backslash & \backslash & \\ T_{j,1} & \dots & T_{j,m} & U_1 & \dots & \end{array} \right)$$

are distinct when the $T_i(t_j)$ are distinct.
- We assume w.l.o.g. that the t_j are pairwise distinct on the first m arguments.

- Let $\delta = \min\{|T_i^{(m)}(t_j) - T_{i'}^{(m)}(t_j)| \neq 0 \mid i, i', j\}$. We can extend the family of t_j for $j \in J$ to finitely many value trees t_j for $j \in J' \supset J$ such that the t_j for $j \in J'$ are pairwise distinct on the first m arguments and such that for all i,

$$\sum_{t \text{ not equal on the first } m \text{ arguments to a } t_j, j \in J'} T_i(t) < \delta.$$

- For $k \leq m$ we apply the induction hypothesis to the $T_{j,k}$ with $j \in J'$ and we get structures $\mathcal{S}_k$. With these we can turn T_i into a subprobability distribution

$$\rho_i : y, s, \alpha_1, \dots \alpha_m \mapsto \sum_{\mathcal{S}_k(T_k) = \alpha_k} \sum_{U_1, \dots} T_i \left(\begin{array}{c} y, s \\ T_1 \quad \dots \quad T_m \quad U_1 \quad \dots \end{array} \right).$$

- For all j we have

$$T^{(m)}(t_j) \leq \rho_i\left(y_j, s_j, \mathcal{S}_1(T_{j,1}), \dots, \mathcal{S}_m(T_{j,m})\right) < T^{(m)}(t_j) + \delta.$$

- The ρ_i are pairwise distinct, with $\rho_i(y_j, s_j, \alpha_1, ..., \alpha_m) \neq \rho_{i'}(y_j, s_j, \alpha_1, ..., \alpha_m)$ for some j when $i \neq i'$.
- Let $\mathcal{S}$ be an evaluation structure with finitely many (y_l, s_l) including at least the labels (y_j, s_j), evaluation functions φ_l of arity m, and substructures $\mathcal{S}_{l,k} = \mathcal{S}_k$. We have

$$\mathcal{S}(T_i) = \sum_l \sum_{\alpha_1,...,\alpha_m} \rho_i(y_l, s_l, \alpha_1, ..., \alpha_m) \times \varphi_l(\alpha_1, ..., \alpha_m).$$

Question

Given pairwise distinct subprobability distributions $\rho_i : L \times [0;1]^m \to [0;1]$, are there evaluation functions φ_l such that the $\sum_l \sum_{\alpha_1,...,\alpha_m} \rho_i(l, \alpha_1, ..., \alpha_m) \times \varphi_l(\alpha_1, ..., \alpha_m)$ are pairwise distinct?

Question

Given pairwise distinct subprobability distributions $\rho_i : L \times [0;1]^m \to [0;1]$, are there evaluation functions φ_l such that the $\sum_l \sum_{\alpha_1,...,\alpha_m} \rho_i(l, \alpha_1, ..., \alpha_m) \times \varphi_l(\alpha_1, ..., \alpha_m)$ are pairwise distinct?

- First step: using projections and probabilistic sums we can contract the variables α_k and turn the ρ_i into pairwise distinct subprobability distributions $\tau_i : [0;1] \to [0;1]$.
- Second step: for $i \neq i'$ we can find $n \in \mathbb{N}$ such that $\sum_\beta \tau_i(\beta) \times \beta^n \neq \sum_\beta \tau_{i'}(\beta) \times \beta^n$.
- Third step: we combine the previous result to find the required evaluation functions.

Going back to where we came from:

- We can separate subprobability distributions over $L \times [0;1]^m$ using evaluation functions.
- With this we can build evaluations structures to separate trees.
- Then we can find contexts to separate terms.

Conclusion

We have natural definitions for the observational equivalence and the Böhm trees, and these definitions preserves Hyland and Wadsworth's result.

We can get stronger results if we have additional hypothesis (almost sure normalization or normalization of the terms, etc.).

We can also define sensible theories in the probabilistic λ-calculus in such a way that the observational equivalence is also the largest coherent sensible theory.