Probabilistic Lambda Calculi

Research Project

Eric Jackson

We define two languages, $\lambda_{\mathbf{IID}}$ and $\lambda_{\mathbf{PC}}$, for modeling probabilistic behavior. The languages share the same syntax and the same static semantics, but differ in their treatment of the random expressions **coin** and **rand**.

Syntax

Expressions

$$\begin{aligned} e ::=\ & x \\ & |() \\ & |\lambda x : \tau.\, e \\ & |\mathbf{let}\ x = e_1\ \mathbf{in}\ e_2 \\ & |e_1\, e_2 \\ & |r \\ & |e_1 \oplus e_2 \\ & |\mathbf{coin} \\ & |\mathbf{rand} \\ & |(\mathbf{case}\ e_1\ \mathbf{of}\ e_2 \,|\, e_3) \\ & |\mathbf{inl}_{\tau_1+\tau_2} e \\ & |\mathbf{inr}_{\tau_1+\tau_2} e \\ & |\mathbf{\#1}\ e \\ & |\mathbf{\#2}\ e \\ & |(e_1, e_2) \end{aligned}$$

The languages include a variety of features familiar from standard extensions of the lambda calculus. There are two probabilistic constructs: **coin**, which evaluates randomly to 0 or 1, and **rand**, which evaluates randomly to a real number in $[0, 1]$. Finally, note that $\oplus$ is a placeholder which represents the binary operations: $+$, $-$, $\times$, and $/$.

Types

$$\begin{aligned} \tau ::=\ & \mathbf{unit} \\ & |\mathbb{R} \\ & |\tau_1 \to \tau_2 \\ & |\tau_1 + \tau_2 \\ & |\tau_1 \times \tau_2 \end{aligned}$$

The languages includes two base types, **unit** and $\mathbb{R}$—the type of real numbers. Additionally, there are types for functions, sums, and products.

Static Semantics

In addition to sharing the same syntax, the languages we define also share the same static semantics. Their typing rules are listed below.

$$\frac{\Gamma(x) = \tau}{\Gamma \vdash x : \tau}\ \textsc{Var}$$

$$\frac{}{\Gamma \vdash () : \mathbf{unit}}\ \textsc{Unit}$$

$$\frac{\Gamma \vdash e : \tau'}{\Gamma \vdash (\lambda x : \tau.\, e) : \tau \to \tau'}\ \textsc{Fun}$$

$$\frac{\Gamma \vdash e_1 : \tau \qquad \Gamma_{x:\tau} \vdash e_2 : \tau'}{\Gamma \vdash \mathbf{let}\ x = e_1\ \mathbf{in}\ e_2 : \tau'}\ \textsc{Let}$$

$$\frac{\Gamma \vdash e_1 : \tau \to \tau' \qquad \Gamma \vdash e_2 : \tau}{\Gamma \vdash e_1\, e_2 : \tau'}\ \textsc{App}$$

$$\frac{}{\Gamma \vdash r : \mathbb{R}}\ \textsc{Num}$$

$$\frac{\Gamma \vdash e_1 : \mathbb{R} \qquad \Gamma \vdash e_2 : \mathbb{R}}{\Gamma \vdash e_1 \oplus e_2 : \mathbb{R}}\ \textsc{Bop}$$

$$\frac{}{\Gamma \vdash \mathbf{coin} : \mathbb{R}}\ \textsc{Coin}$$

$$\frac{}{\Gamma \vdash \mathbf{rand} : \mathbb{R}}\ \textsc{Rand}$$

$$\frac{\Gamma \vdash e_1 : \tau_1 + \tau_2 \qquad \Gamma \vdash e_2 : \tau_1 \to \tau \qquad \Gamma \vdash e_3 : \tau_2 \to \tau}{\Gamma \vdash \mathbf{case}\ e_1\ \mathbf{with}\ e_2 \,|\, e_3 : \tau}\ \textsc{Case}$$

$$\frac{\Gamma \vdash e : \tau_1}{\Gamma \vdash \mathbf{inl}_{\tau_1+\tau_2} e : \tau_1 + \tau_2}\ \textsc{In-Left}$$

$$\frac{\Gamma \vdash e : \tau_2}{\Gamma \vdash \mathbf{inr}_{\tau_1+\tau_2} e : \tau_1 + \tau_2}\ \textsc{In-Right}$$

$$\frac{\Gamma \vdash e : \tau_1 \times \tau_2}{\Gamma \vdash \#\ \mathbf{1}\ e : \tau_1}\ \textsc{Proj-1}$$

$$\frac{\Gamma \vdash e : \tau_1 \times \tau_2}{\Gamma \vdash \#\ \mathbf{2}\ e : \tau_2}\ \textsc{Proj-2}$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \qquad \Gamma \vdash e_2 : \tau_2}{\Gamma \vdash (e_1, e_2) : \tau_1 \times \tau_2}\ \textsc{Pair}$$

Independently and Identically Distributed ($\lambda_{\mathbf{IID}}$)

Values

$$\begin{aligned} v ::= {} & \lambda x : \tau .\, e \\ & |() \\ & |r \\ & |\mathbf{inl}_{\tau_1+\tau_2} v \\ & |\mathbf{inr}_{\tau_1+\tau_2} v \\ & |(e, e) \end{aligned}$$

There are six types of values, which are almost all familiar from traditional lambda calculus extensions. Note, however, that in $\lambda_{\mathbf{IID}}$ pairs of non-value expressions are values. The desirability of this construction can be demonstrated by the expression

$$\mathbf{let}\ x = (\mathbf{coin}, \mathbf{coin})\ \mathbf{in}\ (\mathbf{\#1}\ x) + (\mathbf{\#1}\ x)$$

Semantics

The small-step semantics can be modeled as a relation on $\mathbf{State} \times \mathbf{State}$ where $\mathbf{State} \triangleq e \times \{0,1\}^\omega \times [0,1]^\omega$. This construction allows us to model the behavior of **coin()** and **rand()** in terms of the random sequences in $\{0,1\}^\omega$ and $[0,1]^\omega$, respectively.

We will use a context to facilitate the presentation of the semantics.

$$\begin{aligned} E ::= {} & [\cdot] \\ & |E\, e \\ & |E \oplus e \\ & |v \oplus E \\ & |(\mathbf{case}\ E\ \mathbf{of}\ e_2 \,|\, e_3) \\ & |\mathbf{inl}_{\tau_1+\tau_2} E \\ & |\mathbf{inr}_{\tau_1+\tau_2} E \\ & |\#\ \mathbf{1}\ E \\ & |\#\ \mathbf{2}\ E \end{aligned}$$

Finally, the small-step semantics are given by the following inference rules.

$$\frac{\langle e,\, n,\, m\rangle \to \langle e',\, n',\, m'\rangle}{\langle E(e),\, n,\, m\rangle \to \langle E(e'),\, n',\, m'\rangle}\ \textsc{Context}$$

$$\frac{}{\langle \mathbf{rand},\, n,\, m\rangle \to \langle \mathbf{hd}\ m,\, n,\, \mathbf{tl}\ m\rangle}\ \textsc{Rand}$$

$$\frac{}{\langle (\lambda x : \tau .\, e)\, e_2,\, n,\, m\rangle \to \langle e\{e_2/x\},\, n,\, m\rangle}\ \beta\text{-}\textsc{Reduction}$$

$$\frac{}{\langle (\mathbf{case}\ \mathbf{inl}_{\tau_1+\tau_2} e\ \mathbf{of}\ e_2 \,|\, e_3),\, n,\, m\rangle \to \langle e_2\, e,\, n,\, m\rangle}\ \textsc{Case-Left}$$

$$\frac{}{\langle \mathbf{let}\ x = e_1\ \mathbf{in}\ e_2,\, n,\, m\rangle \to \langle e_2\{e_1/x\},\, n,\, m\rangle}\ \textsc{Let}$$

$$\frac{}{\langle (\mathbf{case}\ \mathbf{inr}_{\tau_1+\tau_2} e\ \mathbf{of}\ e_2 \,|\, e_3),\, n,\, m\rangle \to \langle e_3\, e,\, n,\, m\rangle}\ \textsc{Case-Right}$$

$$\frac{r_1 \bar{\oplus} r_2 = r}{\langle r_1 \oplus r_2,\, n,\, m\rangle \to \langle r,\, n,\, m\rangle}\ \textsc{Bop}$$

$$\frac{}{\langle \mathbf{\#1}\ (e_1, e_2),\, n,\, m\rangle \to \langle e_1,\, n,\, m\rangle}\ \textsc{Proj-1}$$

$$\frac{}{\langle \mathbf{coin},\, n,\, m\rangle \to \langle \mathbf{hd}\ n,\, \mathbf{tl}\ n,\, m\rangle}\ \textsc{Coin}$$

$$\frac{}{\langle \mathbf{\#2}\ (e_1, e_2),\, n,\, m\rangle \to \langle e_2,\, n,\, m\rangle}\ \textsc{Proj-2}$$

Example

$$\begin{aligned} \langle (\lambda x.\, x + x)\, \mathbf{coin},\, 1 :: 0 :: n,\, m\rangle & \to \langle \mathbf{coin} + \mathbf{coin},\, 1 :: 0 :: n,\, m\rangle \\ & \to \langle 1 + \mathbf{coin},\, 0 :: n,\, m\rangle \\ & \to \langle 1 + 0,\, n,\, m\rangle \\ & \to \langle 1,\, n,\, m\rangle \end{aligned}$$

Perfectly Correlated ($\lambda_{\mathbf{PC}}$)

Values

$$\begin{aligned} v ::= &\lambda x : \tau .\, e \\ &|() \\ &|r \\ &|\mathbf{inl}_{\tau_1+\tau_2} v \\ &|\mathbf{inr}_{\tau_1+\tau_2} v \\ &|(v, v) \end{aligned}$$

Again, there are six types of values. Note, however, that unlike $\lambda_{\mathbf{IID}}$ a pair is only a value if it is a pair of values.

Semantics

As with $\lambda_{\mathbf{IID}}$, the small-step semantics of $\lambda_{\mathbf{PC}}$ can be modeled as a relation on **State** $\times$ **State**.

We will again use a context to facilitate presentation, but one should notice several differences between the contexts of the two languages. For example, the $\lambda_{\mathbf{PC}}$ context includes pairs.

$$\begin{aligned} E ::= &[\cdot] \\ &|E\, e \\ &|v\, E \\ &|\mathbf{let}\ x = E\ \mathbf{in}\ e_2 \\ &|E \oplus e \\ &|v \oplus E \\ &|(\mathbf{case}\ E\ \mathbf{of}\ e_2 \,|\, e_3) \\ &|\mathbf{inl}_{\tau_1+\tau_2} E \\ &|\mathbf{inr}_{\tau_1+\tau_2} E \\ &|\#\ \mathbf{1}\ E \\ &|\#\ \mathbf{2}\ E \\ &|(E, e) \\ &|(v, E) \end{aligned}$$

The small-step semantics are given by the following inference rules.

$$\frac{\langle e,\, n,\, m\rangle \to \langle e',\, n',\, m'\rangle}{\langle E(e),\, n,\, m\rangle \to \langle E(e'),\, n',\, m'\rangle}\ \textsc{Context}$$

$$\frac{}{\langle \mathbf{rand},\, n,\, m\rangle \to \langle \mathbf{hd}\ m,\, n,\, \mathbf{tl}\ m\rangle}\ \textsc{Rand}$$

$$\frac{}{\langle (\lambda x : \tau .\, e)\, v,\, n,\, m\rangle \to \langle e\{v/x\},\, n,\, m\rangle}\ \beta\text{-}\textsc{Reduction}$$

$$\frac{}{\langle (\mathbf{case}\ \mathbf{inl}_{\tau_1+\tau_2} v\ \mathbf{of}\ e_2 \,|\, e_3),\, n,\, m\rangle \to \langle e_2\, v,\, n,\, m\rangle}\ \textsc{Case-Left}$$

$$\frac{}{\langle \mathbf{let}\ x = v\ \mathbf{in}\ e_2,\, n,\, m\rangle \to \langle e_2\{v/x\},\, n,\, m\rangle}\ \textsc{Let}$$

$$\frac{}{\langle (\mathbf{case}\ \mathbf{inr}_{\tau_1+\tau_2} v\ \mathbf{of}\ e_2 \,|\, e_3),\, n,\, m\rangle \to \langle e_3\, v,\, n,\, m\rangle}\ \textsc{Case-Right}$$

$$\frac{r_1 \bar{\oplus} r_2 = r}{\langle r_1 \oplus r_2,\, n,\, m\rangle \to \langle r,\, n,\, m\rangle}\ \textsc{Bop}$$

$$\frac{}{\langle \#\mathbf{1}\ (v_1, v_2),\, n,\, m\rangle \to \langle v_1,\, n,\, m\rangle}\ \textsc{Proj-1}$$

$$\frac{}{\langle \mathbf{coin},\, n,\, m\rangle \to \langle \mathbf{hd}\ n,\, \mathbf{tl}\ n,\, m\rangle}\ \textsc{Coin}$$

$$\frac{}{\langle \#\mathbf{2}\ (v_1, v_2),\, n,\, m\rangle \to \langle v_2,\, n,\, m\rangle}\ \textsc{Proj-2}$$

Example

$$\begin{aligned} \langle \mathbf{let}\ x = \mathbf{coin}\ \mathbf{in}\ x + x,\, 1 :: 0 :: n,\, m\rangle &\to \langle \mathbf{let}\ x = 1\ \mathbf{in}\ x + x,\, 0 :: n,\, m\rangle \\ &\to \langle 1 + 1,\, 0 :: n,\, m\rangle \\ &\to \langle 2,\, 0 :: n,\, m\rangle \end{aligned}$$