Probabilistic Programming Languages

M2 MPRI 2021-2022

Guillaume Baudart (guillaume.baudart@ens.fr)
Christine Tasson (christine.tasson@lip6.fr)

Semantics of Probabilistic Programming

Discrete probability

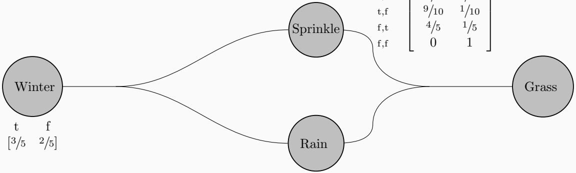
Bayesian Networks

Bayseian Network - Introduction

Example

$$\begin{array}{c|cc} W\backslash S & \mathrm{t} & \mathrm{f} \\ \hline \mathrm{t} & 1/5 & 4/5 \\ \mathrm{f} & 3/4 & 1/4 \end{array}$$

$$\begin{array}{c|cc} S,R\backslash G & \mathrm{t} & \mathrm{f} \\ \hline \mathrm{t,t} & 19/20 & 1/20 \\ \mathrm{t,f} & 9/10 & 1/10 \\ \mathrm{f,t} & 4/5 & 1/5 \\ \mathrm{f,f} & 0 & 1 \end{array}$$



$$\begin{array}{cc} \mathrm{t} & \mathrm{f} \\ \hline 3/5 & 2/5 \end{array}$$

$$\begin{array}{c|cc} W\backslash R & \mathrm{t} & \mathrm{f} \\ \hline \mathrm{t} & 4/5 & 1/5 \\ \mathrm{f} & 1/10 & 9/10 \end{array}$$

Purpose : compact representation of the joint distribution $\mathbb{P}(G, S, R, W)$

Variables X and their sample sets (or carrier or web) $|X|$

Conditional Probability Tables : $|\mathtt{Pa}(X)| \times |X| \to \mathbb{R}^+$

Dependancy to parents $\mathtt{Pa}(X)$ and independacy to other variables.

The function mass $\mathbb{P}(G = g, S = s, R = r, W = w)$ has dimension 2^4 !

Bayesian Network - Introduction

Definition A bayesian network is given by

A DAG

Labeled by variables and conditional probability tables (CPT)

Dependance and Independance

Parents : $\mathtt{Pa}(G) = \{S, R\}$ and $\mathtt{Pa}(S) = W$

The probability of X given variables depends only on $\mathtt{Pa}(X)$.

Joint distribution from conditional probability tables

$$\mathbb{P}(G, R, S, W) = \mathbb{P}(G|S, R)\mathbb{P}(S|W)\mathbb{P}(R|W)\mathbb{P}(W)$$

Conditional probability

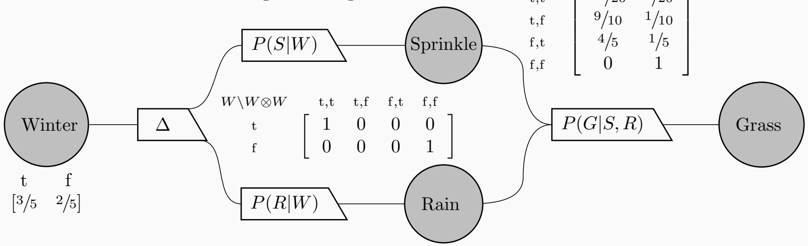
$$\mathbb{P}(G, R, S, W) = \mathbb{P}(G|S, R, W)\mathbb{P}(S, R, W)$$

Chain Rule

$$\mathbb{P}(G, R, S, W) = \mathbb{P}(G|S, R, W)\mathbb{P}(S|R, W)\mathbb{P}(R|W)\mathbb{P}(W)$$

Dependance $\mathbb{P}(G|S, R, W) = \mathbb{P}(G|S, R)$

Bayesian Network - Example



$$\begin{array}{c|cc} W \backslash S & \text{t} & \text{f} \\ \hline \text{t} & 1/5 & 4/5 \\ \text{f} & 3/4 & 1/4 \end{array}$$

$$\begin{array}{c|cc} S,R \backslash G & \text{t} & \text{f} \\ \hline \text{t,t} & 19/20 & 1/20 \\ \text{t,f} & 9/10 & 1/10 \\ \text{f,t} & 4/5 & 1/5 \\ \text{f,f} & 0 & 1 \end{array}$$

$$\begin{array}{c|cccc} W \backslash W\otimes W & \text{t,t} & \text{t,f} & \text{f,t} & \text{f,f} \\ \hline \text{t} & 1 & 0 & 0 & 0 \\ \text{f} & 0 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{cc} \text{t} & \text{f} \\ \hline 3/5 & 2/5 \end{array}$$

$$\begin{array}{c|cc} W \backslash R & \text{t} & \text{f} \\ \hline \text{t} & 4/5 & 1/5 \\ \text{f} & 1/10 & 9/10 \end{array}$$

Compute $\mathbb{P}(G)$ using :

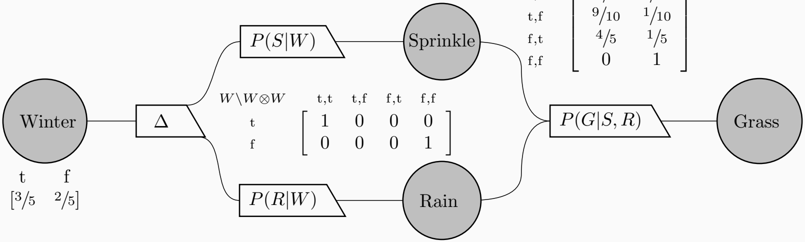
Joint dist. $\mathbb{P}(G,R,S,W) = \mathbb{P}(G|S,R)\ \mathbb{P}(S|W)\ \mathbb{P}(R|W)\ \mathbb{P}(W)$

Marginal $\mathbb{P}(G) = \sum_{(r,s,w)\in|R|\times|S|\times|W|} \mathbb{P}(G,R,S,W)$

Bayesian Network - Example

$$\begin{array}{c|cc} W\backslash S & \text{t} & \text{f} \\ \hline \text{t} & 1/5 & 4/5 \\ \text{f} & 3/4 & 1/4 \end{array}$$

$$\begin{array}{c|cc} S,R\backslash G & \text{t} & \text{f} \\ \hline \text{t,t} & 19/20 & 1/20 \\ \text{t,f} & 9/10 & 1/10 \\ \text{f,t} & 4/5 & 1/5 \\ \text{f,f} & 0 & 1 \end{array}$$



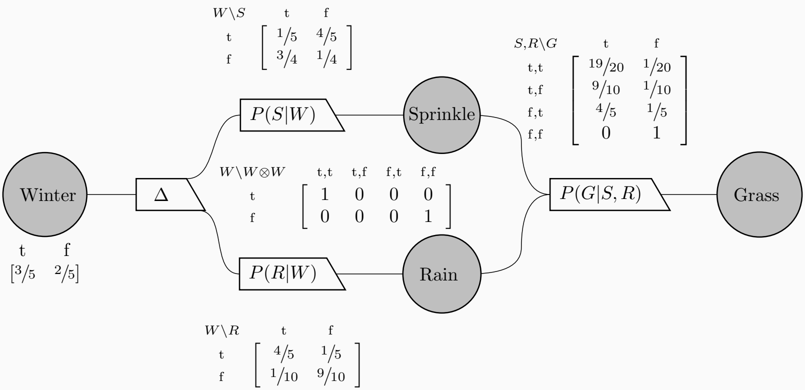
$$\begin{array}{c|cccc} W\backslash W\otimes W & \text{t,t} & \text{t,f} & \text{f,t} & \text{f,f} \\ \hline \text{t} & 1 & 0 & 0 & 0 \\ \text{f} & 0 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{cc} \text{t} & \text{f} \\ [3/5 & 2/5] \end{array}$$

$$\begin{array}{c|cc} W\backslash R & \text{t} & \text{f} \\ \hline \text{t} & 4/5 & 1/5 \\ \text{f} & 1/10 & 9/10 \end{array}$$

$$p(S) = \left(\sum_{a \in \{\text{t},\text{f}\}} P(S|W)_{a,b} \cdot p(W)_a \right)_{b \in \{\text{t},\text{f}\}}$$

Bayesian Network - Example



$$\begin{array}{c|cc} W\backslash S & \text{t} & \text{f} \\ \hline \text{t} & 1/5 & 4/5 \\ \text{f} & 3/4 & 1/4 \end{array}$$

$$\begin{array}{c|cc} S,R\backslash G & \text{t} & \text{f} \\ \hline \text{t,t} & 19/20 & 1/20 \\ \text{t,f} & 9/10 & 1/10 \\ \text{f,t} & 4/5 & 1/5 \\ \text{f,f} & 0 & 1 \end{array}$$

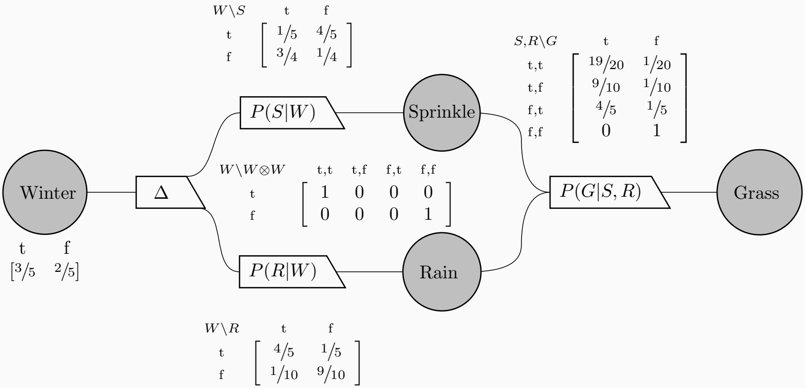
$$\begin{array}{c|cccc} W\backslash W\otimes W & \text{t,t} & \text{t,f} & \text{f,t} & \text{f,f} \\ \hline \text{t} & 1 & 0 & 0 & 0 \\ \text{f} & 0 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{cc} \text{t} & \text{f} \\ \hline 3/5 & 2/5 \end{array}$$

$$\begin{array}{c|cc} W\backslash R & \text{t} & \text{f} \\ \hline \text{t} & 4/5 & 1/5 \\ \text{f} & 1/10 & 9/10 \end{array}$$

$$\begin{aligned} p(W)\ P(S|W) &= p(S) \\ p(W)\ P(R|W) &= p(R) \end{aligned} \quad \text{and} \quad (p(S)\otimes p(R))\ P(G|S,R) = p(G)$$

Bayesian Network - Example



$$p(W)\ \Delta\ (P(S|W) \otimes P(R|W))\ P(G|S,R) \ = \ p(G)$$

Semantics of Functional Programming

Operational and Denotational

PCF, Programming Computable Functions (a functional typed programming language)

(D. Scott : A Type-Theoretical Alternative to ISWIM, CUCH, OWHY. Theor. Comput. Sci., 1993)

(G. Plotkin : LCF Considered as a Programming Language. Theor. Comput. Sci.,1977)

Syntax

$$M, N, P := \underbrace{x \mid \lambda x\, M \mid (M)\, N \mid (M, N)}_{\lambda\text{-calculus}} \mid \underbrace{\mathtt{fix}\, M}_{\text{Recursion}}$$

$$\mid \underbrace{0 \mid \mathtt{succ}\, M}_{\text{Arithmetics}} \mid \underbrace{\mathtt{true} \mid \mathtt{false} \mid \mathtt{if}\, M\, \mathtt{then}\, N\, \mathtt{else}\, P}_{\text{Conditionnal}}$$

Semantics

Operational : $M \to M'$ **Denotational :** $[\![\Gamma \vdash M : A]\!] : [\![\Gamma]\!] \to [\![A]\!]$

Wanted

Soundeness : If $M \to M'$, then $[\![\Gamma \vdash M : A]\!] = [\![\Gamma \vdash M' : A]\!]$.

Adequation Lemma : For every n, if $[\![M]\!] = n$ then $M \to^* n$.

PCF - Typing system

Types and contexts

$$A, B ::= \mathtt{unit} \mid \mathtt{int} \mid \mathtt{bool} \mid A \to B \qquad\qquad \Gamma ::= \top \mid \Gamma, x : A$$

Rules

$\Gamma, x : A \vdash x : A \qquad \Gamma \vdash () : \mathit{unit}$

if $\Gamma, x : A \vdash M : B$ then $\Gamma \vdash \lambda x.M : A \to B$

if $\Gamma \vdash M : A \to B$ and $\Gamma \vdash N : A$ then $\Gamma \vdash (M)N : B$

if $\Gamma \vdash M : A \to A$ then $\Gamma \vdash \mathtt{fix}\, M : A$

$\Gamma \vdash \mathtt{true} : \mathtt{bool}$ and $\Gamma \vdash \mathtt{false} : \mathtt{bool}$

if $\Gamma \vdash M : \mathtt{bool}$ and $\Gamma \vdash N : A$ and $\Gamma \vdash P : A$
then $\Gamma \vdash \mathtt{if}\, M\, \mathtt{then}\, N\, \mathtt{else}\, P : A$

$\Gamma \vdash 0 : \mathtt{nat}$

if $\Gamma \vdash M : \mathtt{nat}$ then $\Gamma \vdash \mathtt{succ}\, M : \mathtt{nat}$

PCF - Operational Semantics

Call-by-name

$$\frac{}{(\lambda x\, M)N \to M\,[N/x]} \qquad \frac{M \to M'}{(M)N \to (M')N}$$

$$\frac{}{\texttt{fix } M \to (M)(\texttt{fix} M)}$$

$$\frac{}{\texttt{if}\,\texttt{true}\,\texttt{then}\,N\,\texttt{else}\,P \to N} \qquad \frac{}{\texttt{if}\,\texttt{false}\,\texttt{then}\,N\,\texttt{else}\,P \to P}$$

$$\frac{M \to M'}{\texttt{if } M\,\texttt{then}\,N\,\texttt{else}\,P \to \texttt{if } M'\,\texttt{then}\,N\,\texttt{else}\,P}$$

PCF - Denotational semantics

Category $\mathcal{C}$ SET, REL, MREL

Objects sets, sets, sets

Morphisms $(\mathcal{C}(A, B), id_A, \circ)$ function, relations, multirelations

Interpretation of types :

$[\![A]\!]$ is an object

$[\![G \vdash e : t]\!]$ is a morphism in $\mathcal{C}([\![G]\!], [\![t]\!])$

PCF - Denotational semantics

Cartesian $(\mathcal{C}, \&, \top)$ (SET, $\{*\}$, $\times$), (REL, $\emptyset$, $\uplus$), (MREL, $\emptyset$, $\uplus$)

Terminal object $\top$ such that there is a unique $A \to \top$
Cartesian product $A \& B$ together with projections $A \& B \to A$ and $A \& B \to B$ such that for any $f : Z \to A$ and $g : Z \to B$, there is a unique $< f, g >: Z \to A \& B$ compatible with projections.

Interpretation of contexts :

$$[\![\mathtt{unit}]\!] = \top$$
$$[\![x_1 : t_1, \ldots, x_k : t_k]\!] = [\![t_1]\!] \times \cdots \times [\![t_k]\!]$$

Interpretation of terms :

$$\left[\!\!\left[\frac{G \vdash e_1 : t_1 \qquad G \vdash e_2 : t_2}{G \vdash (e_1, e_2) : t_1 \times t_2}\right]\!\!\right] = \langle [\![G \vdash e_1 : t_1]\!], [\![G \vdash e_2 : t_2]\!] \rangle$$

PCF - Denotational semantics

Closed $A \Rightarrow B$ B^A, 💣, $\mathcal{M}_{fin}A \times B$

Evaluation $ev : (A \Rightarrow B) \times A \to B$
Curryfing if $f : C \times A \to B$ then there is a unique $\Lambda f : C \to A \Rightarrow B$ compatible with the evaluation.

Interpretation of terms :

$$\left[\!\!\left[\frac{G, x : t_1 \vdash e : t_2}{G \vdash \lambda x.e : t_1 \to t_2} \right]\!\!\right] = \Lambda(\llbracket G, x : t_1 \vdash e : t_2 \rrbracket$$

$$\left[\!\!\left[\frac{G \vdash e_1 : t_1 \to t_2 \qquad G \vdash e_2 : t_1}{G \vdash (e_1)e_2 : t_2} \right]\!\!\right] = ev \circ \langle \llbracket G \vdash e_1 : t_1 \to t_2 \rrbracket, \llbracket G \vdash e_2 : t_2 \rrbracket \rangle$$

PCF - Denotational semantics

CPO-enriched : CPO, REL, MREL

$\mathcal{C}(A, B)$ is a complete partial order and $\circ$ is continuous.

$$\left[\!\!\left[\frac{G \vdash e : t \to t}{G \vdash \mathit{fixe} : t} \right]\!\!\right] = \sup F^n \text{ with } \begin{cases} F^0 = \bot \in \mathcal{C}(G, t) \\ F^{n+1} = ev \circ \langle [\![G \vdash e : t \to t]\!], F^n \rangle \end{cases}$$

PCF - Denotational semantics

Theorem : Every cartesian closed category is a model of the λ-calculus.

Example : Compute $[\![\Gamma, x : A \vdash x : A]\!]$ and $[\![\Gamma \vdash (M)N : B]\!]$

Theorem : Every cartesian closed category CPO-enriched with booleans and natural objects is a model of PCF

Example : Compute $[\![\Gamma \vdash \texttt{fix}\, M : A]\!]$

Theorem : MREL is an adequate model of PCF.

Example : Compute $[\![\Gamma \vdash \texttt{if}\, M\, \texttt{then}\, N\, \texttt{else}\, P : A]\!]$

Probabilistic PCF

Syntax and Semantics

Probabilistic PCF

Syntax

$$M, N, P := \underbrace{x \mid \lambda x\, M \mid (M)\, N \mid (M, N)}_{\lambda\text{-calculus}} \mid \underbrace{\mathtt{fix}\, M}_{\text{Recursion}}$$

$$\mid \underbrace{0 \mid \mathtt{succ}\, M}_{\text{Arithmetics}} \mid \underbrace{\mathtt{true} \mid \mathtt{false} \mid \mathtt{if}\, M\, \mathtt{then}\, N\, \mathtt{else}\, P}_{\text{Conditionnal}}$$

$$\mid \underbrace{\mathtt{let\ x = sample(bernoulli}\ p\mathtt{)\ in}\ M}_{\text{Probability}}\ \forall p \in [0,1]$$

Types

$$\frac{\Gamma, x : \mathtt{bool} \vdash M : A}{\Gamma \vdash \mathtt{let\ x = sample(bernoulli}\ p\mathtt{)\ in}\ M : B}$$

Semantics

Operational : $M \xrightarrow{p} M'$ **Denotational :** $[\![\Gamma \vdash M : A]\!] : [\![\Gamma]\!] \to [\![A]\!]$

Wanted

Soundeness : $[\![\Gamma \vdash M : A]\!] = \sum_{M' \mid M \xrightarrow{p} M'} [\![\Gamma \vdash M' : A]\!]$.

Operational Semantics

$$\frac{}{(\lambda x\, M)N \xrightarrow{1} M\,[N/x]} \qquad \frac{M \xrightarrow{p} M'}{(M)N \xrightarrow{p} (M')N} \qquad \frac{}{\texttt{fix}\ M \xrightarrow{1} (M)(\texttt{fix}M)}$$

$$\frac{}{\texttt{if}\ \texttt{true}\ \texttt{then}\ N\ \texttt{else}\ P \xrightarrow{1} N} \qquad \frac{}{\texttt{if}\ \texttt{false}\ \texttt{then}\ N\ \texttt{else}\ P \xrightarrow{1} P}$$

$$\frac{M \xrightarrow{p} M'}{\texttt{succ}\ M \xrightarrow{p} \texttt{succ}\ M'}$$

$$\frac{M \xrightarrow{p} M'}{\texttt{if}\ M\ \texttt{then}\ N\ \texttt{else}\ P \xrightarrow{p} \texttt{if}\ M'\ \texttt{then}\ N\ \texttt{else}\ P}$$

$$\frac{}{\texttt{sample}(\texttt{bernoulli}\ p) \xrightarrow{p} \texttt{true}} \qquad \frac{}{\texttt{sample}(\texttt{bernoulli}\ p) \xrightarrow{1-p} \texttt{false}}$$

$$\frac{M \xrightarrow{p} M'}{\texttt{let}\ x = M\ \texttt{in}\ N \xrightarrow{p} \texttt{let}\ x = M'\ \texttt{in}\ N} \qquad \frac{\text{v is a constant}}{\texttt{let}\ x = v\ \texttt{in}\ N \xrightarrow{p} M\,[v/x]}$$

Operational Semantics

Transition Matrix :

$$\mathbf{Proba}(M, M') = \begin{cases} p & \text{if } M \xrightarrow{p} M' \\ 1 & \text{if } M \text{ normal and } M = M' \\ 0 & \text{otherwise.} \end{cases}$$

Iterated Transition Matrix :

$\mathbf{Proba}^k(M, N)$ is the probability that M reduces to N in at most k steps.
$\mathbf{Proba}^\infty(M, N)$ when N is normal is the probability that M reduces to N in any number of steps

Adequation Lemma (wanted) : If M closed term of type `nat`, then for every n, ~~if $[\![M]\!] = n$ then $M \to^* n$.~~

$$[\![M]\!]_n = \mathbf{Proba}^\infty(M, \underline{n})$$

Examples : Compute the operational semantics

Bernoulli :

```
let a = sample(bernoulli 0.5) in
let b = sample(bernoulli 0.5) in
let c = sample(bernoulli 0.5) in
a+b+c
```

Fixpoint :

```
fix lambda b:bool
    let x = sample(bernoulli p) in
    let y = sample(bernoulli p) in
    if x then (if y then b else true)
         else (if y then false else b)
```

Geometric :

```
fix lambda f:bool->unit lambda d:bool
    let x = sample(d) in
    if x then (f)d else ()
```

Examples : Compute the operational semantics

Count - while

```
c = 0
x = 1
while (x>0)
        let y = (sample bernoulli p) in
        if y then x = 0 else c += 1
c
```

Count - fix

```
fix lambda f:bool->bool
    lambda c:bool
        let y = (sample bernoulli p) in
        if y then c else f(c+1)
```

Probabilistic PCF - Denotational Semantics

The category of Probabilistic Coherent spaces - Pcoh

Object : $(|A|, \mathrm{P}(A))$ with $|A|$ a set and $\mathrm{P}(A) \subset (\mathbb{R}^+)^{|A|}$ set of functions from web to positive reals.

Examples

$$|\texttt{unit}| = \{*\}$$
$$\mathrm{P}(\texttt{unit}) = [0,1]$$
$$|\texttt{int}| = \mathbb{N}$$
$$\mathrm{P}(\texttt{int}) = \{(x_n) | \sum x_i \leq n\}$$
$$|\texttt{bool}| = \{\texttt{true},\ \texttt{false}\}$$
$$\mathrm{P}(\texttt{bool}) = \{(x_{\texttt{true}}, x_{\texttt{false}}) \,|\, x_{\texttt{true}} + x_{\texttt{false}} \leq 1$$
$$|A \,\&\, B| = |A| \uplus |B|$$
$$\mathrm{P}(A \,\&\, B) = \{(x_i)_{i \in A \uplus B} \,|\, (x_i)_{i \in A} \in \mathrm{P}(A),\ (x_i)_{i \in B} \in \mathrm{P}(B)\}$$

Probabilistic PCF - Denotational Semantics

(Vincent Danos, Thomas Ehrhard, On probabilistic coherence spaces, 2011) **The category of Probabilistic Coherence spaces - Pcoh**

Object : $(|A|, \mathrm{P}(A))$ with $|A|$ a set and $\mathrm{P}(A) \subset (\mathbb{R}^+)^{|A|}$ set of functions from $|A|$ to $\mathbb{R}^+$

Morphism : $f : (|A|, \mathrm{P}(A)) \to (|B|, \mathrm{P}(B))$ a matrix $(f_{(m,b)})$ indexed by $\mathcal{M}_{\mathrm{fin}}(A) \times B$ such that

$$\forall x \in \mathrm{P}(A), f \cdot x : b \mapsto \sum_{m \in \mathcal{M}_{\mathrm{fin}}(A)} \prod_{a \in m} x(a)^{m(a)} f_{m,b} \in \mathrm{P}(B)$$

Examples

$f : [\![unit]\!] \to [\![unit]\!]$ such that $\forall x \in [0,1],\ f \cdot x = \sum_n f_n x^n \in [0,1]$

$f : [\![bool]\!] \to [\![bool]\!]$ such that $f = \begin{array}{c c} W\backslash S & \begin{array}{cc} \mathrm{t} & \mathrm{f} \end{array} \\ \begin{array}{c} \mathrm{t} \\ \mathrm{f} \end{array} & \begin{bmatrix} 1/5 & 4/5 \\ 3/4 & 1/4 \end{bmatrix} \end{array}$

$f : [\![bool]\!] \to [\![unit]\!]$ such that $f_{([\mathtt{true}^n, *], \mathtt{true})} = 1$ otherwise $f_{m,*} = 0$, then $f \cdot (p, 1-p) = \sum_n p^n (1-p)$ and $f \cdot (1,0) = 0$.

Probabilistic PCF - Denotational Semantics

A sound model

Pcoh is a cartesian closed category, CPO-enriched, hence a model of PCF, which is also a model of probabilistic PCF.

Interpretation of terms

$$\left[\!\left[\frac{}{G \vdash \mathtt{bernoulli}(p) : \mathtt{bool\ dist}}\right]\!\right] : \begin{cases} (\gamma, \mathtt{true}) \mapsto p \\ (\gamma, \mathtt{false}) \mapsto 1 - p \end{cases}$$

$$\left[\!\left[\frac{G \vdash d \ : \ t_1\ \mathtt{dist} \qquad G, x : t_1 \vdash e : t_2}{G \vdash \mathtt{let}\, x = d\, \mathtt{in}\, e \ : \ t_2}\right]\!\right] : (\gamma, a_2) \mapsto \sum_a [\![e]\!](\gamma, a, a_2)$$

$$\left[\!\left[\frac{}{G \vdash true : bool}\right]\!\right] : \begin{cases} (\gamma, \mathtt{true}) \mapsto 1 \\ (\gamma, \mathtt{false}) \mapsto 0 \end{cases}$$

$$[\![\mathtt{if}\, M\, \mathtt{then}\, N\, \mathtt{else}\, P]\!](\gamma, a) = [\![M]\!](\gamma, \mathtt{true})[\![N]\!](\gamma, a) + [\![M]\!](\gamma, \mathtt{false})[\![P]\!](\gamma, a)$$

Compute the denotational semantics

Bernoulli :

```
let a = sample(bernoulli 0.5) in
let b = sample(bernoulli 0.5) in
let c = sample(bernoulli 0.5) in
a+b+c
```

Compute the denotational semantics

Fixpoint :

```
fix lambda b:bool
    let x = sample(bernoulli p) in
    let y = sample(bernoulli p) in
    if x then (if y then b else true)
         else (if y then false else b)
```

Compute the denotational semantics

Geometric :

```
fix lambda f:bool->unit lambda d:bool
    let x = sample(d) in
    if x then (f)d else ()
```

Compute the denotational semantics

Count - while

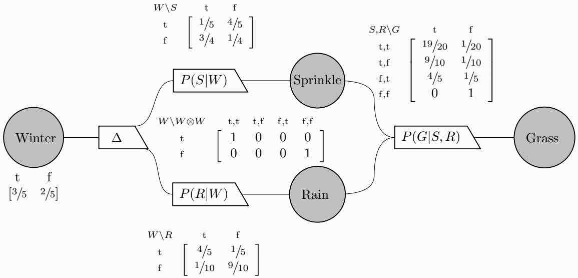
```
c = 0
x = 1
while (x>0)
        let y = (sample bernoulli p) in
        if y then x = 0 else c += 1
c
```

Count - fix

```
fix lambda f:bool->bool
    lambda c:bool
        let y = (sample bernoulli p) in
        if y then c else f(c+1)
```

Example of a Bayesian Network

Definition A bayesian network is given by a DAG whose nodes are labeled by types (variables) and edges by morphisms (CPT).

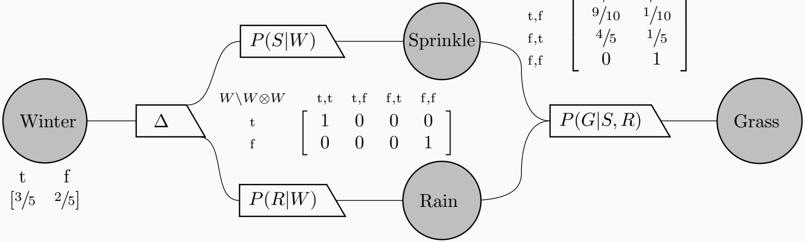


Tensor $[\![A \otimes B]\!]$ with $|A \otimes B| = |A| \times |B|$ and

$$\mathrm{P}(A \otimes B) = \{f : A \times B \to \mathbb{R}^+ \mid a \mapsto \sum_a f_{a,b} \in \mathrm{P}(A),$$
$$b \mapsto \sum_b f_{a,b} \in \mathrm{P}(B)\}.$$

Examples - Compute the semantics

A bayesian network is given by a DAG whose nodes are labeled by types (variables) and edges by morphisms (CPT).



$$\begin{array}{ccc} W\backslash S & \mathrm{t} & \mathrm{f} \\ \begin{array}{c}\mathrm{t}\\ \mathrm{f}\end{array} & \multicolumn{2}{c}{\begin{bmatrix} 1/5 & 4/5 \\ 3/4 & 1/4 \end{bmatrix}} \end{array}$$

$$\begin{array}{ccccc} W\backslash W\otimes W & \mathrm{t,t} & \mathrm{t,f} & \mathrm{f,t} & \mathrm{f,f} \\ \begin{array}{c}\mathrm{t}\\ \mathrm{f}\end{array} & \multicolumn{4}{c}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}} \end{array}$$

$$\begin{array}{ccc} W\backslash R & \mathrm{t} & \mathrm{f} \\ \begin{array}{c}\mathrm{t}\\ \mathrm{f}\end{array} & \multicolumn{2}{c}{\begin{bmatrix} 4/5 & 1/5 \\ 1/10 & 9/10 \end{bmatrix}} \end{array}$$

$$\begin{array}{ccc} S,R\backslash G & \mathrm{t} & \mathrm{f} \\ \begin{array}{c}\mathrm{t,t}\\ \mathrm{t,f}\\ \mathrm{f,t}\\ \mathrm{f,f}\end{array} & \multicolumn{2}{c}{\begin{bmatrix} 19/20 & 1/20 \\ 9/10 & 1/10 \\ 4/5 & 1/5 \\ 0 & 1 \end{bmatrix}} \end{array}$$

$$\begin{array}{cc} \mathrm{t} & \mathrm{f} \\ [3/5 & 2/5] \end{array}$$

$$p(G) = P(G|S,R) \cdot (P(S|W) \otimes P(R|W)) \cdot \Delta \cdot p(W)$$

Examples - Compute the semantics

Funny Bernoulli by rejection sampling :

```
fix lambda f:bool
  let a = sample(bernoulli 0.5) in
  let b = sample(bernoulli 0.5) in
  let c = sample(bernoulli 0.5) in
  if (a==0 or b==0) then a+b+c
                         else f
```

Examples - Compute the semantics

Funny Bernoulli :

```
let FB =
  let a = sample(bernoulli 0.5) in
  let b = sample(bernoulli 0.5) in
  let c = sample(bernoulli 0.5) in
  assume(a==0 or b==0) ;
  a+b+c
```

Reminder on `enumeration` algorithm.

Probabilistic Language - Type and Semantics

$$[\![G \vdash e \ : \ t]\!] : [\![G]\!] \times [\![t]\!] \to \mathbb{R}^+ \text{ and } [\![G \vdash e \ : \ t \ \mathtt{dist}]\!](\gamma, \mathtt{a}) \in \mathrm{P}\,(\mathtt{G} \to \mathtt{t})$$

$$\left[\!\!\left[\frac{}{G \vdash \mathtt{bernoulli}(p) : \mathtt{bool\ dist}}\right]\!\!\right] : \begin{cases} (\gamma, \mathtt{true}) \mapsto p \\ (\gamma, \mathtt{false}) \mapsto 1 - p \end{cases}$$

$$\left[\!\!\left[\frac{G \vdash d \ : \ t_1 \ \mathtt{dist} \qquad G, x : t_1 \vdash e : t_2}{G \vdash \mathtt{let}\, x = d \,\mathtt{in}\, e \ : \ t_2}\right]\!\!\right] : (\gamma, a_2) \mapsto \sum_a [\![e]\!](\gamma, a, a_2)$$

$$\left[\!\!\left[\frac{G \vdash e_1 \ : \ \mathtt{bool} \qquad G \vdash e_2 \ : \ t}{G \vdash \mathtt{assume}(e_1) \ ; \ e_2 \ : \ t}\right]\!\!\right](\gamma, a) = \begin{cases} [\![e_2]\!](\gamma, a) & \text{if } [\![e_1]\!](g, true) = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\left[\!\!\left[\frac{G \vdash e \ : \ t}{G \vdash \mathtt{infer}(e) \ : \ t \ \mathtt{dist}}\right]\!\!\right](\gamma, a) = \frac{[\![G \vdash e \ : \ t]\!](\gamma, a)}{\sum_{a \in |t|} [\![G \vdash e \ : \ t]\!](\gamma, a)}$$

Examples

Dice :

```
let Dice =
  let a = sample(unif(1,6)) in
  let b = sample(unif(1,6)) in
  assume(a - b== 1) ;
  a+b
```

Cannabis :

```
let Cannabis =
    let b = sample(bernoulli 0.5) in
    if b then let c = sample(bernoulli 0.6) in
                assume c; c
          else true
```

Take home

Semantics for (discrete) probabilistic programs

operational semantics and denotational semantics

Probabilistic coherent spaces (be able to compute semantics)

a model of Bayesian Networks
Probabilistic PCF
First Order Probabilistic language with conditionning

References

Semantics of probabilistic programs by Dexter Kozen
An Introduction to Probabilistic Programming. Jan-Willem van de Meent, Brooks Paige, Hongseok Yang and Frank Wood
Probabilistic Coherent Spaces by Vincent Danos and Thomas Ehrhard
A predicate/state transformer for Bayesian learning by Bart Jacobs and Fabio Zanasi.
Bayesian Networks and Proof Nets Talk by Claudia Faggian.