A Live, Multiple-Representation Probabilistic Programming Environment for Novices

Maria Gorinova, Advait Sarkar, Alan Blackwell, Don Syme
let Weather = Variable.Beta(2.0, 3.3)
let BadWeather = Variable.Bernoulli(Weather)
let BadRoads = Variable.New<bool>()
let HighNumberOfJourneys = Variable.New<bool>()
let HighSpeed = Variable.New<bool>()
let HighDangerLevel = Variable.New<bool>()
let HighNumberOfAccidents = Variable.New<bool>()
let IFBW = Variable.If(BadWeather)
BadRoads.SetTo(Variable.Bernoulli(0.7))
HighNumberOfJourneys.SetTo(Variable.Bernoulli(0.1))
IFBW.CloseBlock()
let IFNotBW = Variable.IfNot(BadWeather)
BadRoads.SetTo(Variable.Bernoulli(0.1))
HighNumberOfJourneys.SetTo(Variable.Bernoulli(0.8))
IFNotBW.CloseBlock()
let IFGR = Variable.If(BadRoads)
HighSpeed.SetTo(Variable.Bernoulli(0.4))

HighSpeed prior distribution is Bernoulli(0.5715)
Observing HighNumberOfAccidents to be true...

HighSpeed posterior distribution is Bernoulli(0.6305)
What is Probabilistic Programming?
What is a Probabilistic Graphical Model?

- Probabilistic graphical models:
  - Vertices are *random variables*
  - Edges are *direct dependencies*

\[
P(\text{Both Heads} = \text{true}) = P(\text{Coin1} = \text{true}) \times P(\text{Coin2} = \text{true})
\]

\[
= 0.5 \times 0.5
\]

\[
= 0.25
\]
What is a Probabilistic Graphical Model?

• What if $Coin1 = true$?

\[
P(BothHeads = true \mid Coin1 = true) = 0.5 \times 1 + 0.5 \times 0 = 0.5\]
What is a Probabilistic Graphical Model?

- What if $BothHeads = false$?
What is a Probabilistic Graphical Model?

• What if $\text{BothHeads} = \text{false}$?

$$P(\text{Coin1} = \text{true} \mid \text{BothHeads} = \text{false})$$
What is a Probabilistic Graphical Model?

- What if BothHeads = false?

\[
P(Coin1 = true \mid BothHeads = false) = \frac{P(BothHeads = false \mid Coin1 = true) \times P(Coin1 = true)}{P(BothHeads = false)}
\]

\[
= \frac{0.5 \times 0.5}{0.75}
\]

\[
= 0.33
\]
What is a Probabilistic Graphical Model?

• What if BothHeads = false?

\[
P(Coin1 = true \mid BothHeads = false) = \frac{P(BothHeads = false \mid Coin1 = true) \times P(Coin1 = true)}{P(BothHeads = false)}
\]

\[
= 0.5 \times 0.5/0.75
\]

\[
= 0.33
\]
What is Probabilistic Programming?

- Probabilistic programming language:

  ```
  let Coin1 = new Bernoulli(0.5)
  let Coin2 = new Bernoulli(0.5)
  let BothHeads = Coin1 && Coin2
  ```
What is Probabilistic Programming?

- Probabilistic programming language:

```csharp
let Coin1 = new Bernoulli(0.5)
let Coin2 = new Bernoulli(0.5)
let BothHeads = Coin1 && Coin2

BothHeads.Observe(false)
```
Learning a Gaussian

```plaintext
let data = [0.5, 0.01, ..., 0.7]
let N = data.Length
let X = new Gaussian[N]

let M = new Gaussian(0, 1)
let V = new Gamma(1, 1)

for (i = 0 to N)
    X[i] = new Gaussian(M, V)
    X[i].Observe(data[i])
```
Learning a Gaussian

let data = [0.5, 0.01, ..., 0.7]
let N = data.Length
let X = new Gaussian[N]

let M = new Gaussian(0, 1)
let V = new Gamma(1, 1)

X = Gaussian(M, V).ForEach(N)
X.Observe(data)
How is Probabilistic Programming different?
How is Probabilistic Programming different?

- **Conventional programming:** each variable has a single value – it holds a specific number, object, function

  vs.

- **Probabilistic programming:** each variable represents an entire probability distribution

![Graphs showing probability distributions for Coin1 and Coin2, as well as BothHeads, illustrating the differences in outcomes.](image-url)
How is Probabilistic Programming different?

• Conventional programming: each variable has a single value – it holds a specific number, object, function

  vs.

• Probabilistic programming: each variable represents an entire probability distribution
How is Probabilistic Programming different?

• Conventional programming: no side effects (in general)

  VS.

• Probabilistic programming: observations have side effects

```javascript
let Coin1 = new Boolean()
let Coin2 = new Boolean()
...
let BothHeads = Coin1 && Coin2
...
BothHeads = false
```
How is Probabilistic Programming different?

- **Conventional programming:** no side effects (in general)  
  vs.

- **Probabilistic programming:** observations have side effects
The IDE
Aims

• Design a tool that communicates the different conceptual basis of probabilistic programming

• Do that by visualising:
  • The graphical models
  • The posterior distributions of random variables
let Coin1 = Variable.Bernoulli(0.5)
let Coin2 = Variable.Bernoulli(0.5)
let BothHeads = Coin1 && Coin2
High-Level Overview

Checker Module

Source Code

Injection

Graph Module

Checking Phase

Injection Phase
Evaluation
User Study

• Aim: study the effect on learning PP concepts
• Structure:
  • **Learning Phase**: study theory with the help of a workbook. Complete 4 interleaved tasks
  • **Exercises Phase**: complete 8 additional exercises

<table>
<thead>
<tr>
<th>Part 1 (learning)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction to:</td>
</tr>
<tr>
<td>• F#'s syntax</td>
</tr>
<tr>
<td>• Probabilistic programming</td>
</tr>
<tr>
<td>• Infer.NET</td>
</tr>
<tr>
<td>Interleave 4 exercises</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part 2 (exercise)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 additional exercises</td>
</tr>
</tbody>
</table>
Exercises

• Debugging:
  “You will be presented with a graphical model that should represent the problem described below. There are several mistakes in the code. Find and correct all mistakes.”

• Observation:
  “Please explain out loud what problem the model describes. Are variables X and Y independent? What is the probability of X being true?”
Demo 2
Results: Learning

<table>
<thead>
<tr>
<th>Measurement</th>
<th>IDE Median</th>
<th>Text Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>191 s</td>
<td>220 s</td>
</tr>
<tr>
<td>Keystrokes</td>
<td>28.5</td>
<td>93</td>
</tr>
</tbody>
</table>
Results: Usability

• Increased performance in completing probabilistic programming exercises compared to a plain text editor:
  • IDE tasks vs plain-text tasks (*within user*)
    • Time difference (per **part 2** exercise): 80 s
    • Keystrokes difference (per **part 2** exercise): 22.5 keystrokes
Results: Confidence
Thank you!
Markov Blanket
More Results: Exercise Phase

<table>
<thead>
<tr>
<th>Measurement</th>
<th>V</th>
<th>p-value</th>
<th>IDE median</th>
<th>Text median</th>
<th>Difference median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>611</td>
<td>0.004162</td>
<td>288 sec</td>
<td>359 sec</td>
<td>80 sec</td>
</tr>
<tr>
<td>Keystrokes</td>
<td>302</td>
<td>0.002044</td>
<td>15.5</td>
<td>79</td>
<td>22.5</td>
</tr>
<tr>
<td>Backspaces</td>
<td>367.5</td>
<td>0.02397</td>
<td>3.5</td>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

(a) Results for all Part 2 exercises.

<table>
<thead>
<tr>
<th>Measurement</th>
<th>V</th>
<th>p-value</th>
<th>IDE median</th>
<th>Text median</th>
<th>Difference median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>205</td>
<td>0.2781</td>
<td>323 sec</td>
<td>337 sec</td>
<td>18 sec</td>
</tr>
<tr>
<td>Keystrokes</td>
<td>167</td>
<td>0.07112</td>
<td>34</td>
<td>90</td>
<td>64.5</td>
</tr>
<tr>
<td>Backspaces</td>
<td>188.5</td>
<td>0.2472</td>
<td>9</td>
<td>12</td>
<td>3</td>
</tr>
</tbody>
</table>

(b) Results for debugging exercises.

<table>
<thead>
<tr>
<th>Measurement</th>
<th>V</th>
<th>p-value</th>
<th>IDE median</th>
<th>Text median</th>
<th>Difference median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>101</td>
<td>0.001667</td>
<td>245 sec</td>
<td>369 sec</td>
<td>92 sec</td>
</tr>
<tr>
<td>Keystrokes</td>
<td>12</td>
<td>0.002448</td>
<td>0</td>
<td>1.5</td>
<td>0</td>
</tr>
<tr>
<td>Backspaces</td>
<td>31</td>
<td>0.0331</td>
<td>0</td>
<td>36.5</td>
<td>0</td>
</tr>
</tbody>
</table>

(c) Results for observation exercises.
Experiment Structure

<table>
<thead>
<tr>
<th>order of exercises</th>
<th>order of usage of editors</th>
</tr>
</thead>
<tbody>
<tr>
<td>straight order</td>
<td>IDE first - text editor first</td>
</tr>
<tr>
<td>Group ID: A/B 00</td>
<td>Group ID: A/B 01</td>
</tr>
<tr>
<td>reverse order</td>
<td>IDE first - text editor first</td>
</tr>
<tr>
<td>Group ID: A/B 10</td>
<td>Group ID: A/B 11</td>
</tr>
</tbody>
</table>

Group A
Learning with the IDE

Viewing examples and completing exercises using the IDE

<table>
<thead>
<tr>
<th>Group A.1</th>
<th>Group A.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDE</td>
<td>text editor</td>
</tr>
<tr>
<td>text editor</td>
<td>IDE</td>
</tr>
</tbody>
</table>

Group B
Learning without the IDE

Viewing examples and completing exercises using a plain text editor

<table>
<thead>
<tr>
<th>Group B.1</th>
<th>Group B.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDE</td>
<td>text editor</td>
</tr>
<tr>
<td>text editor</td>
<td>IDE</td>
</tr>
</tbody>
</table>

Between-subject study to evaluate as a learning tool

Within-subject study to evaluate usability