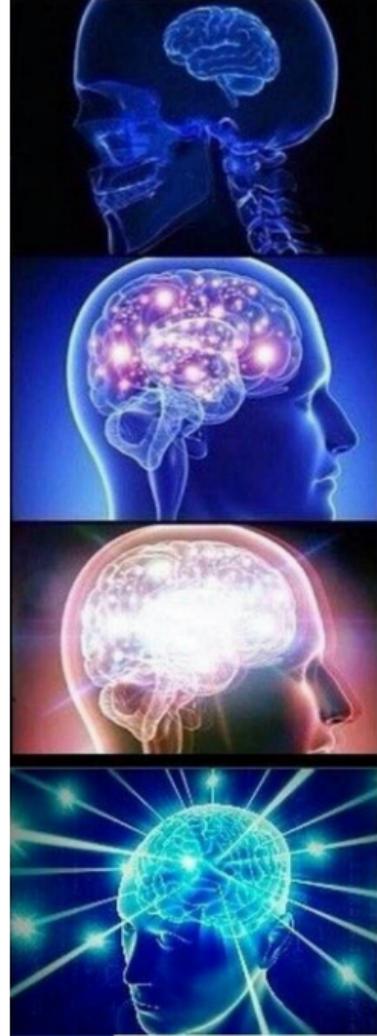
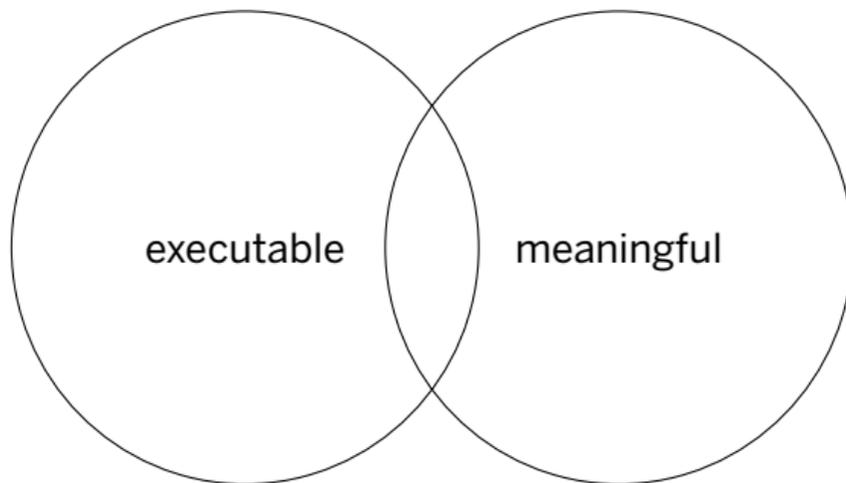


Calculating distributions

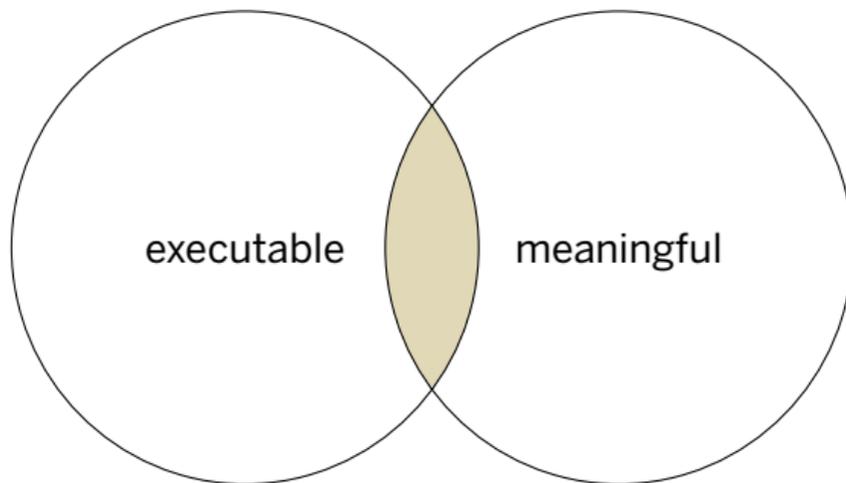
Chung-chieh Shan
Indiana University
2018-09-21



Calculating distributions



Calculating distributions



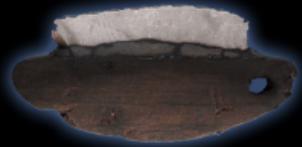
I'd also like to address this concept of being "fake" or "calculating."

If being "fake" means not thinking or feeling the same way in one moment than you thought or felt in a different moment, then lord help us all.

If being "calculating" is thinking through your words and actions and modeling the behavior you would like to see in the world, even when it is difficult, then **I hope more of you will become calculating.**

—BenDeLaCreme





**Creative definitions
and reasoning from first
principles**



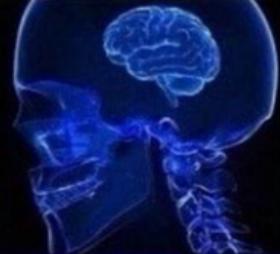
**Symbolic representations
of common definition
patterns**

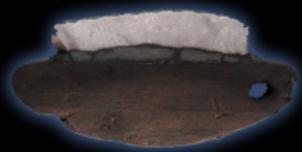


**Mechanical operations
for common reasoning
patterns**



**Virtuous cycle
of automation
and exploration
(Buchberger)**





**Creative definitions
and reasoning from first
principles**

natural
numbers



**Symbolic representations
of common definition
patterns**

unary,
binary



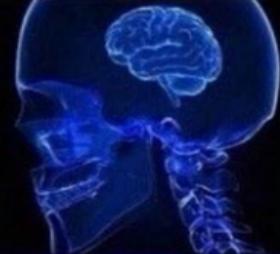
**Mechanical operations
for common reasoning
patterns**

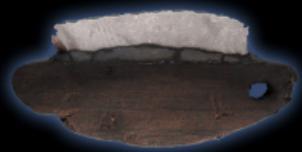
$<$, $+$, \div



**Virtuous cycle
of automation
and exploration
(Buchberger)**

rationals,
reals,
polynomials





**Creative definitions
and reasoning from first
principles**

natural
numbers

probability
distributions



**Symbolic representations
of common definition
patterns**

unary,
binary

table,
Bayes net,
probabilistic
program



**Mechanical operations
for common reasoning
patterns**

$<$, $+$, \div

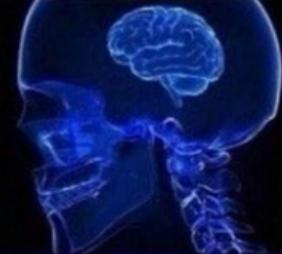
recognize,
integrate,
disintegrate

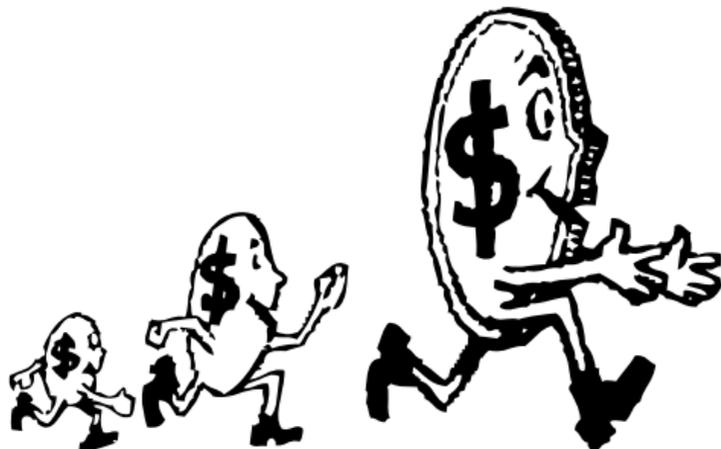


**Virtuous cycle
of automation
and exploration
(Buchberger)**

rationals,
reals,
polynomials

inference,
learning,
optimization



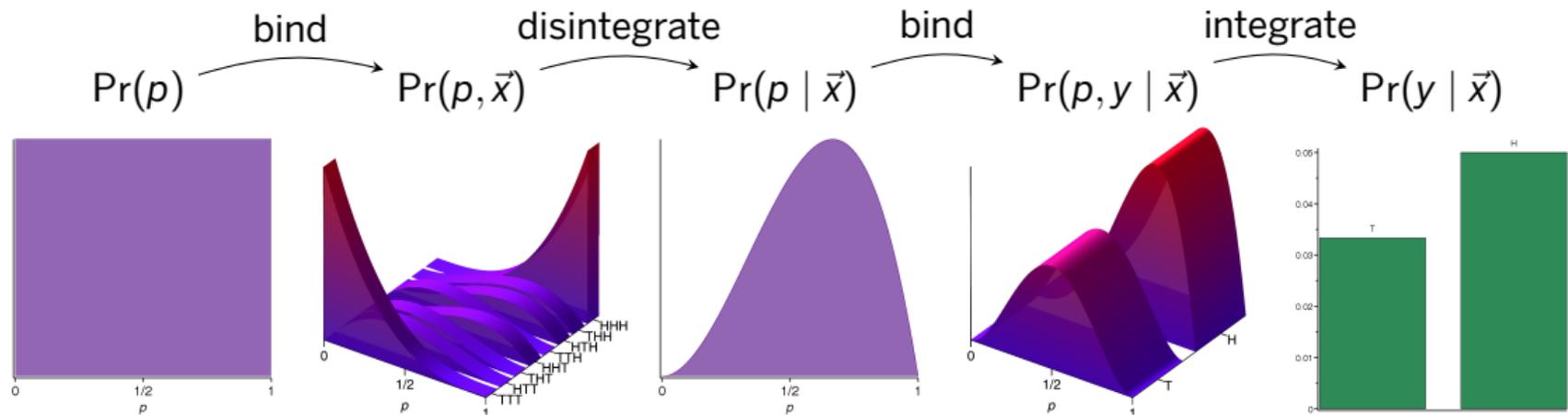


An unknown random process yields a stateless coin that can be flipped repeatedly to produce heads (H) or tails (T).

We assume that the probability p that the coin produces H each time is distributed uniformly between 0 and 1 by the process.

We flip the coin 3 times and observe THH.
What is the probability that the next flip produces H versus T?

(adapted from Eddy)



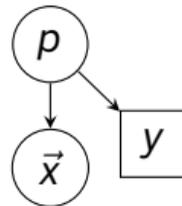
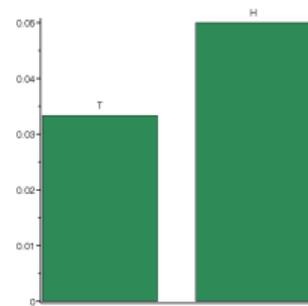
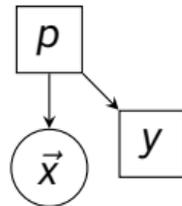
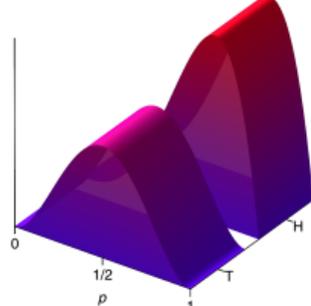
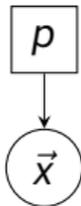
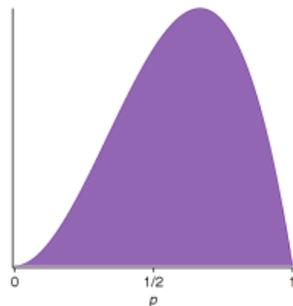
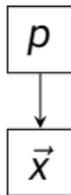
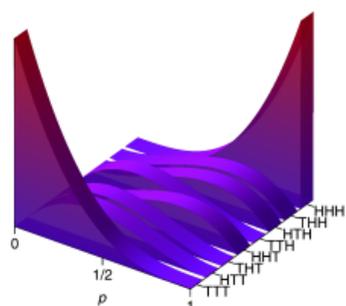
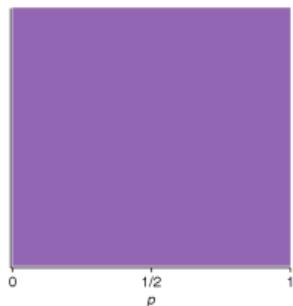
An unknown random process yields a stateless coin that can be flipped repeatedly to produce heads (H) or tails (T).

We assume that the probability p that the coin produces H each time is distributed uniformly between 0 and 1 by the process.

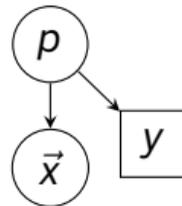
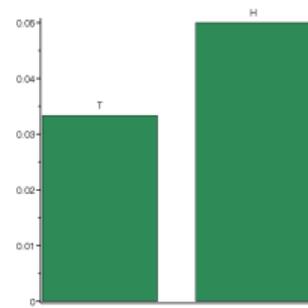
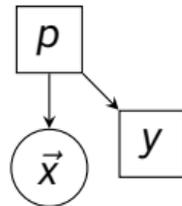
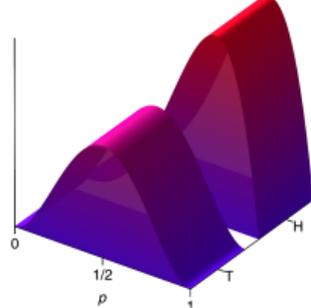
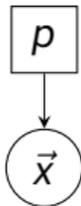
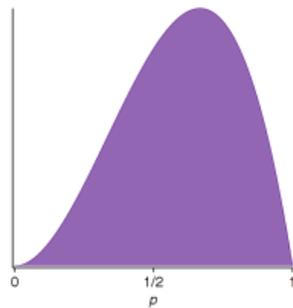
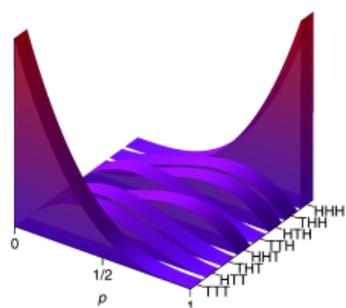
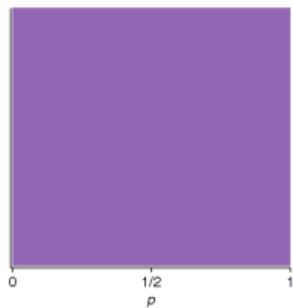
We flip the coin 3 times and observe THH. What is the probability that the next flip produces H versus T?

(adapted from Eddy)

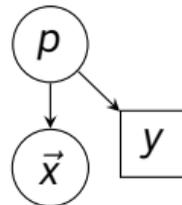
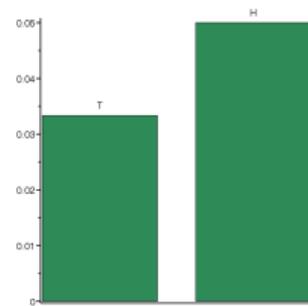
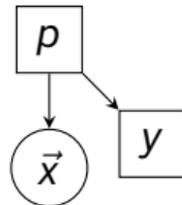
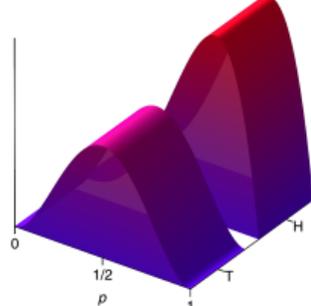
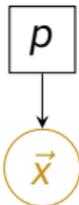
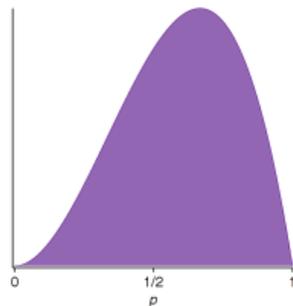
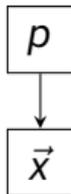
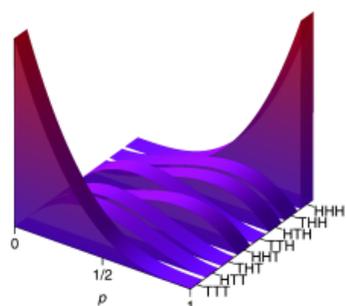
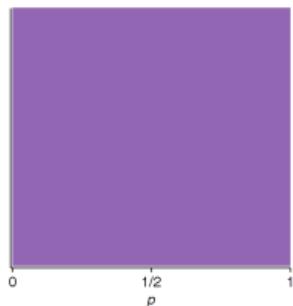
$$\Pr(p) \xrightarrow{\text{bind}} \Pr(p, \vec{x}) \xrightarrow{\text{disintegrate}} \Pr(p | \vec{x}) \xrightarrow{\text{bind}} \Pr(p, y | \vec{x}) \xrightarrow{\text{integrate}} \Pr(y | \vec{x})$$



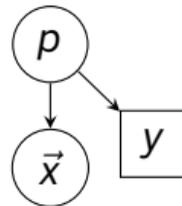
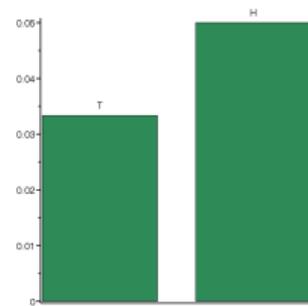
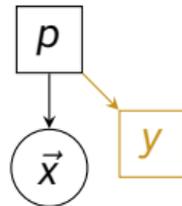
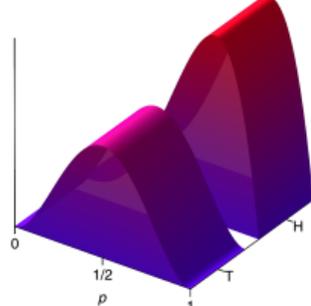
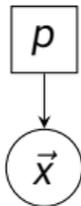
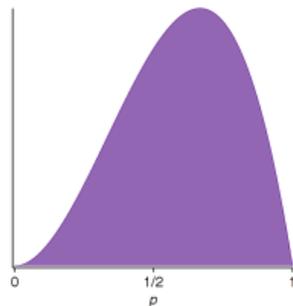
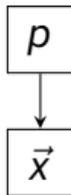
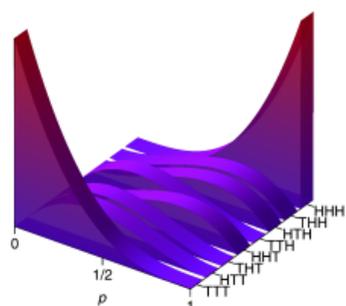
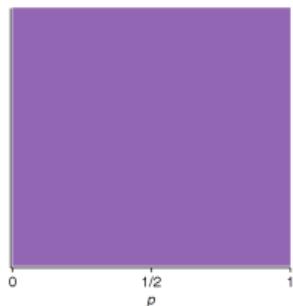
$$\Pr(p) \xrightarrow{\text{bind}} \Pr(p, \vec{x}) \xrightarrow{\text{disintegrate}} \Pr(p | \vec{x}) \xrightarrow{\text{bind}} \Pr(p, y | \vec{x}) \xrightarrow{\text{integrate}} \Pr(y | \vec{x})$$



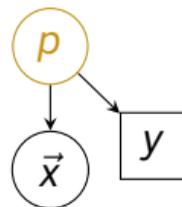
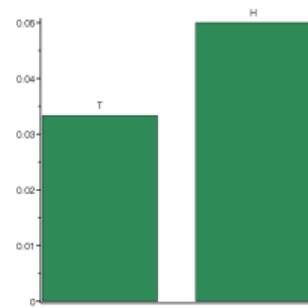
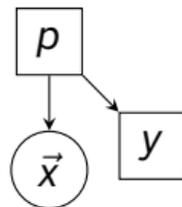
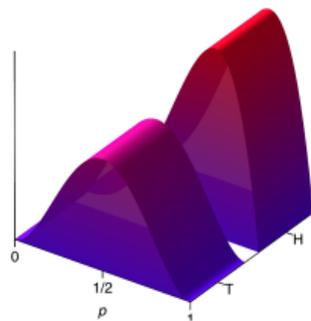
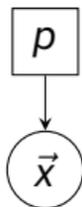
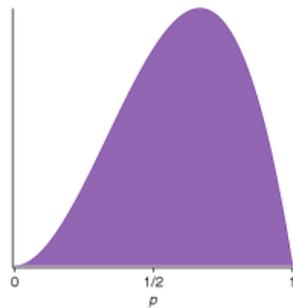
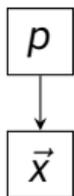
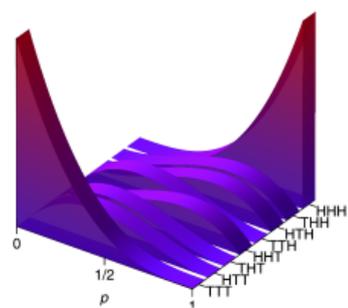
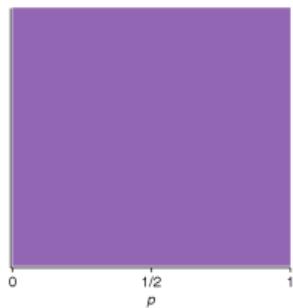
$\Pr(p)$ $\xrightarrow{\text{bind}}$ $\Pr(p, \vec{x})$ $\xrightarrow{\text{disintegrate}}$ $\Pr(p | \vec{x})$ $\xrightarrow{\text{bind}}$ $\Pr(p, y | \vec{x})$ $\xrightarrow{\text{integrate}}$ $\Pr(y | \vec{x})$



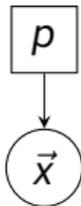
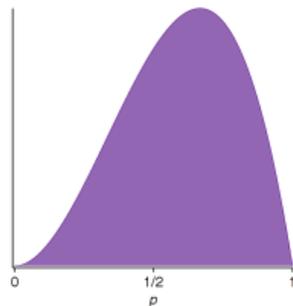
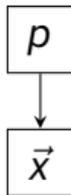
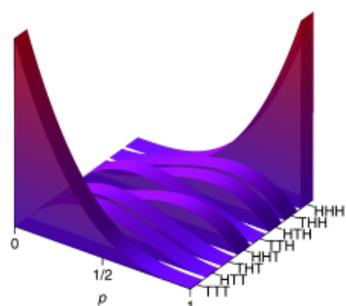
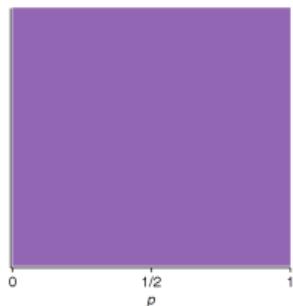
$$\Pr(p) \xrightarrow{\text{bind}} \Pr(p, \vec{x}) \xrightarrow{\text{disintegrate}} \Pr(p | \vec{x}) \xrightarrow{\text{bind}} \Pr(p, y | \vec{x}) \xrightarrow{\text{integrate}} \Pr(y | \vec{x})$$



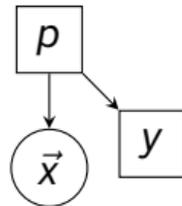
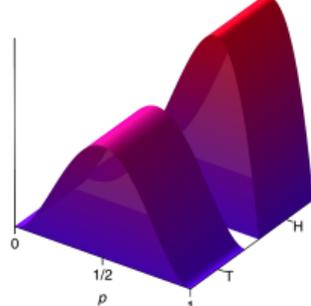
$$\Pr(p) \xrightarrow{\text{bind}} \Pr(p, \vec{x}) \xrightarrow{\text{disintegrate}} \Pr(p | \vec{x}) \xrightarrow{\text{bind}} \Pr(p, y | \vec{x}) \xrightarrow{\text{integrate}} \Pr(y | \vec{x})$$



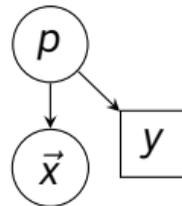
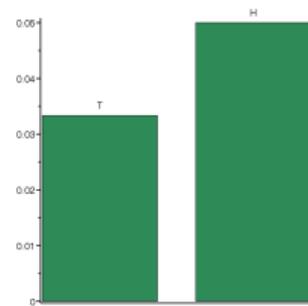
$$\Pr(p) \xrightarrow{\text{bind}} \Pr(p, \vec{x}) \xrightarrow{\text{disintegrate}} \Pr(p | \vec{x}) \xrightarrow{\text{bind}} \Pr(p, y | \vec{x}) \xrightarrow{\text{integrate}} \Pr(y | \vec{x})$$



||



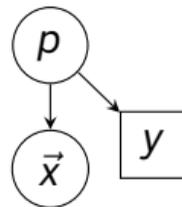
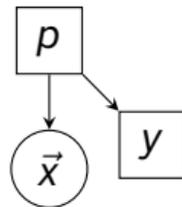
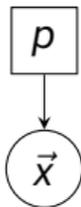
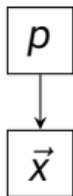
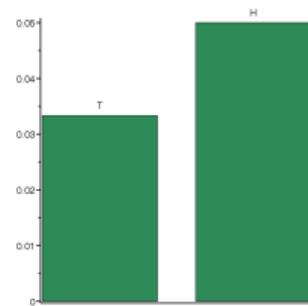
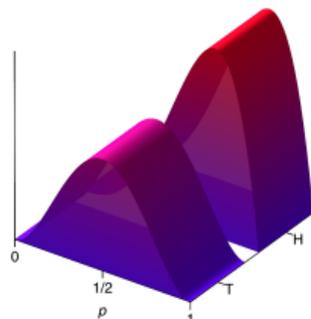
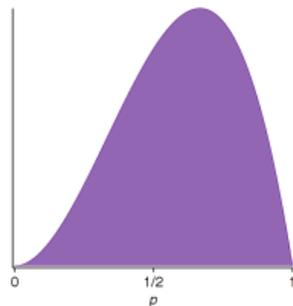
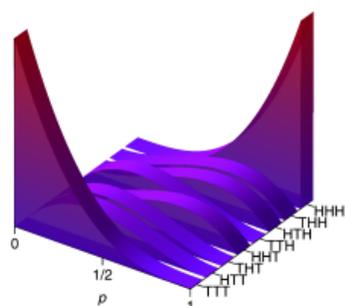
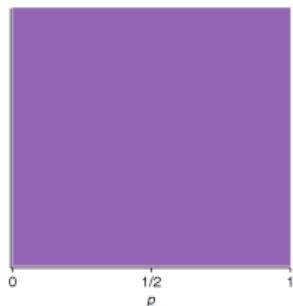
||



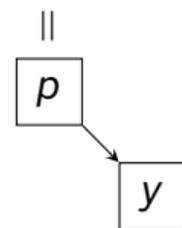
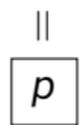
||

bind disintegrate bind integrate

$\Pr(p)$ $\Pr(p, \vec{x})$ $\Pr(p | \vec{x})$ $\Pr(p, y | \vec{x})$ $\Pr(y | \vec{x})$



simplify



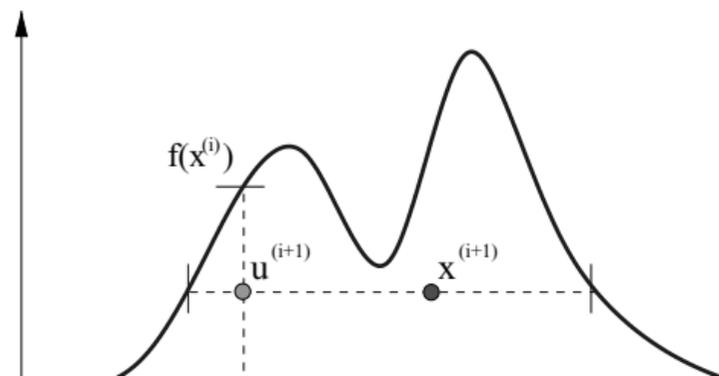
||

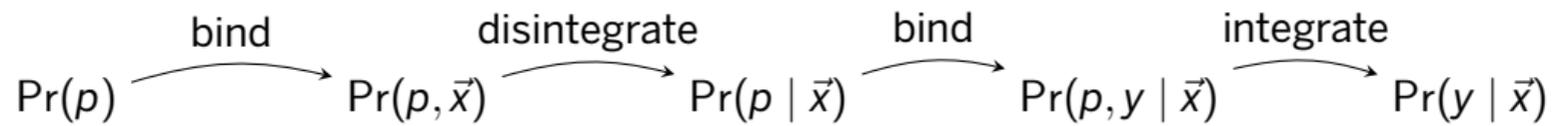


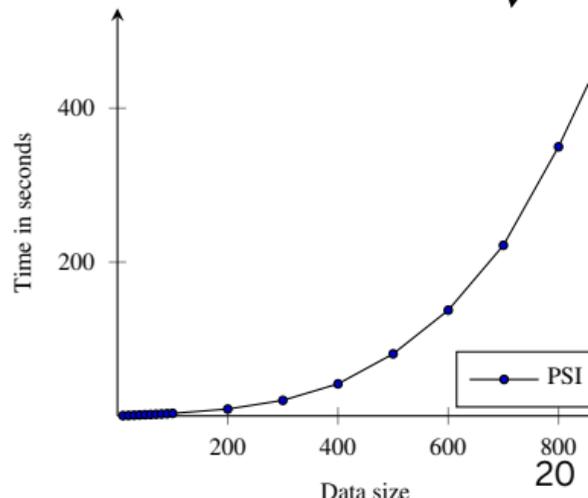
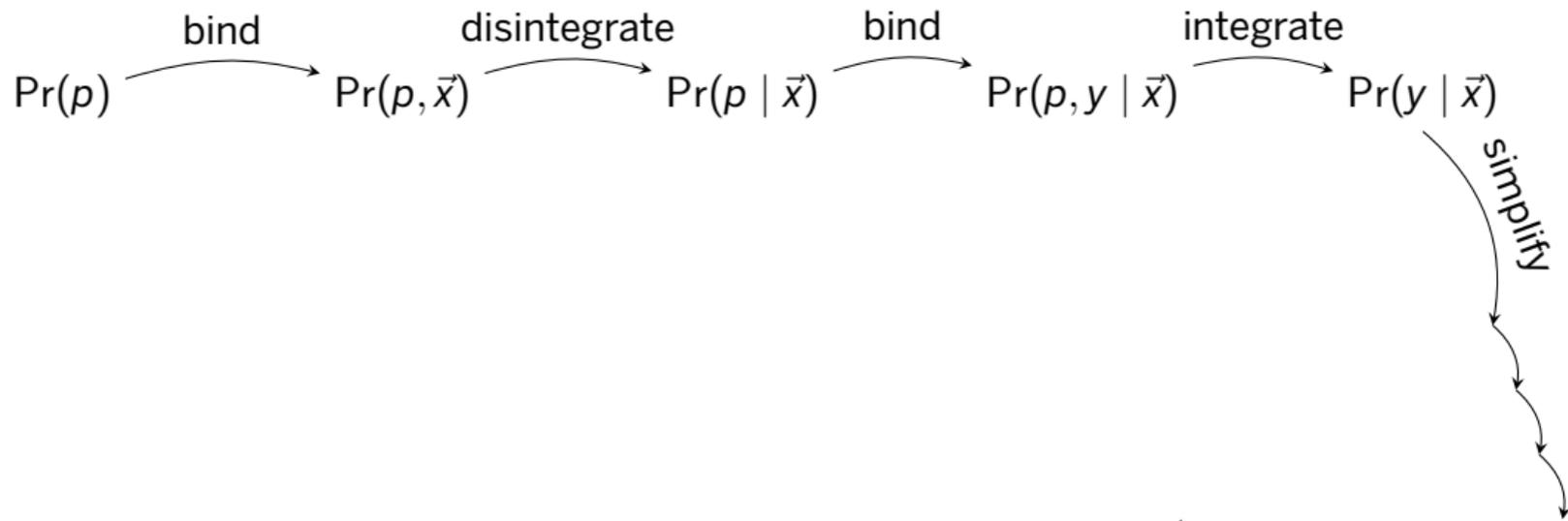
Approximations calculated exactly

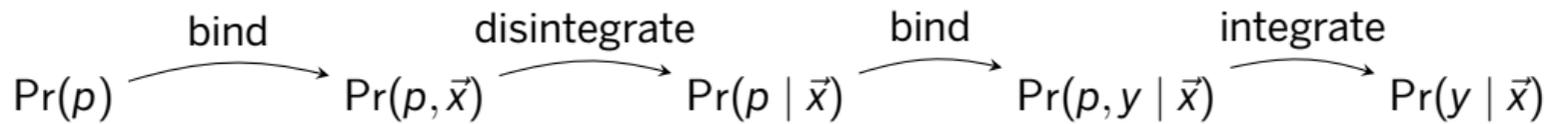
Approximations calculated exactly

$$\begin{aligned}x_1^{(t+1)} &\sim P(x_1|x_2^{(t)}, x_3^{(t)}, \dots, x_K^{(t)}) \\x_2^{(t+1)} &\sim P(x_2|x_1^{(t+1)}, x_3^{(t)}, \dots, x_K^{(t)}) \\x_3^{(t+1)} &\sim P(x_3|x_1^{(t+1)}, x_2^{(t+1)}, \dots, x_K^{(t)}), \text{ etc.}\end{aligned}$$

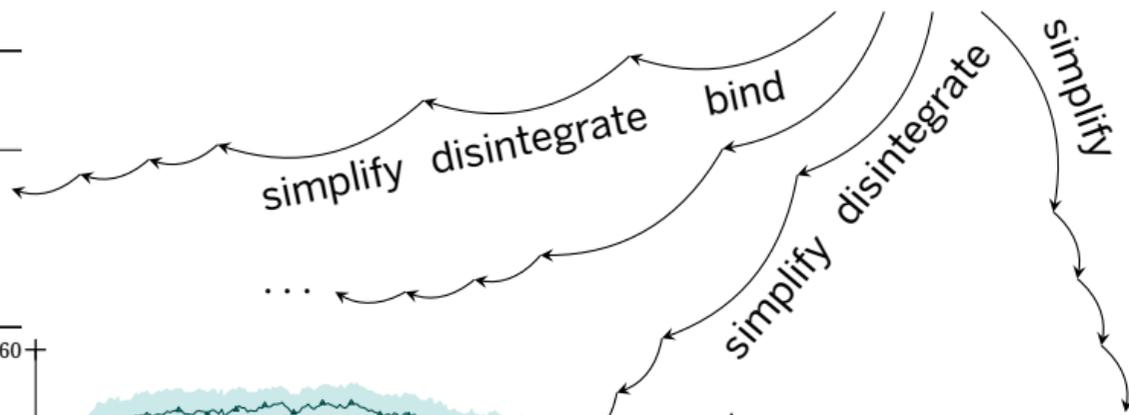




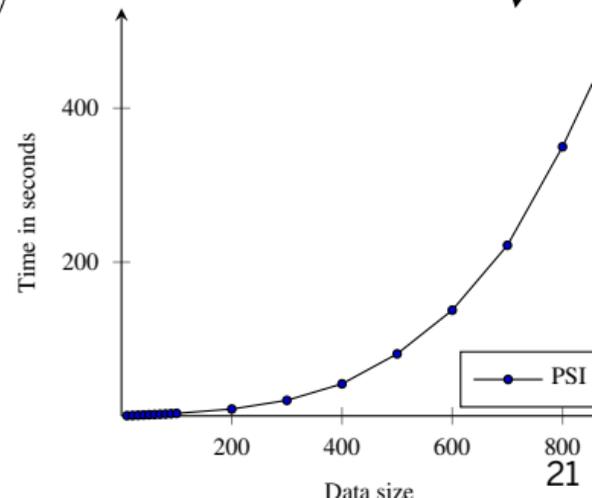
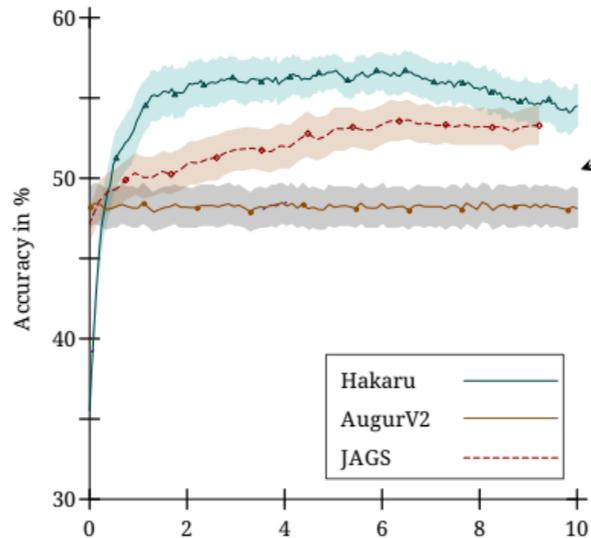




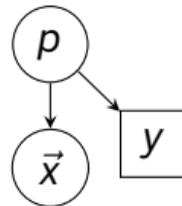
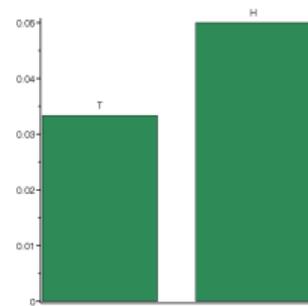
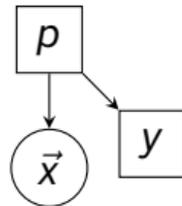
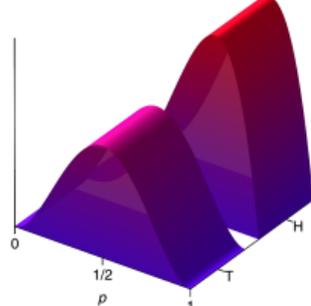
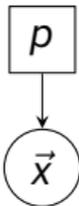
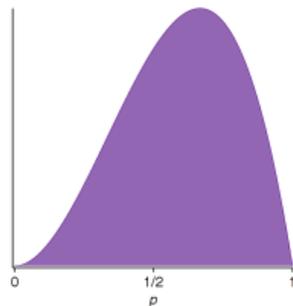
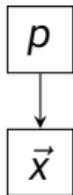
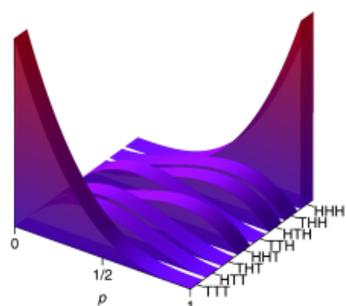
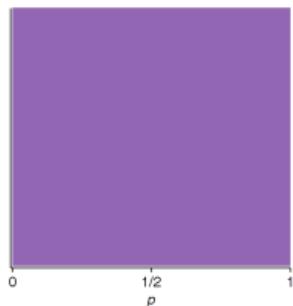
Inference method	Run time (msecs)	
	Mean	SD
WebPPL	1078	16
Hakaru without simplifications	1321	93
Hakaru with simplifications	269	10
Handwritten	207	4



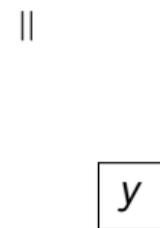
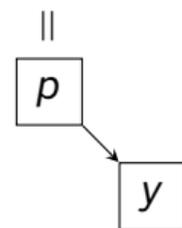
Put approximations
in the language!
(FLOPS 2016, UAI 2017)



$\Pr(p)$ $\xrightarrow{\text{bind}}$ $\Pr(p, \vec{x})$ $\xrightarrow{\text{disintegrate}}$ $\Pr(p | \vec{x})$ $\xrightarrow{\text{bind}}$ $\Pr(p, y | \vec{x})$ $\xrightarrow{\text{integrate}}$ $\Pr(y | \vec{x})$

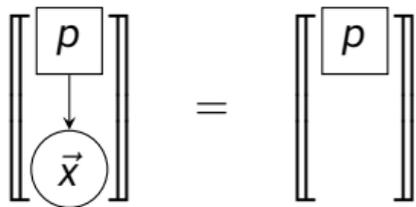


simplify $\left(\right.$ **recognize**
 \parallel



Recognizing a density function

Program denote measures:



Recognizing a density function

Program denote measures:

$$\left[\begin{array}{c} p \\ \downarrow \\ \vec{x} \end{array} \right] = \left[\begin{array}{c} p \\ \\ \phantom{\vec{x}} \end{array} \right] \in \mathbb{M}[0,1]$$

Recognizing a density function

Program denote measures:

$$\left[\begin{array}{c} \boxed{p} \\ \downarrow \\ \vec{x} \end{array} \right] = \left[\boxed{p} \right] \in \mathbb{M}[0,1] \subseteq ([0,1] \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

Recognizing a density function

Program denote measures:

$$\left[\begin{array}{c} \boxed{p} \\ \downarrow \\ \vec{x} \end{array} \right] = \left[\begin{array}{c} \boxed{p} \\ \\ \phantom{\vec{x}} \end{array} \right] \in \mathbb{M}[0,1] \subseteq ([0,1] \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$f \mapsto$$

Recognizing a density function

Program denote measures:

$$\left[\begin{array}{c} \boxed{p} \\ \downarrow \\ \vec{x} \end{array} \right] = \left[\begin{array}{c} \boxed{p} \\ \\ \phantom{\vec{x}} \end{array} \right] \in \mathbb{M}[0,1] \subseteq ([0,1] \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$f \mapsto \int_0^1 \sum_{\vec{x} \in \{H,T\}^3} p^{\sum_i \langle x_i=H \rangle} (1-p)^{\sum_i \langle x_i=T \rangle} \langle \vec{x} = THH \rangle \cdot f(p) dp$$

Recognizing a density function

Program denote measures:

$$\left[\begin{array}{c} \boxed{p} \\ \downarrow \\ \vec{x} \end{array} \right] = \left[\begin{array}{c} \boxed{p} \\ \\ \phantom{\vec{x}} \end{array} \right] \in \mathbb{M}[0,1] \subseteq ([0,1] \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$f \mapsto \int_0^1 \sum_{\vec{x} \in \{H,T\}^3} p^{\sum_i \langle x_i = H \rangle} (1-p)^{\sum_i \langle x_i = T \rangle} \langle \vec{x} = THH \rangle \cdot f(p) dp$$

Recognizing a density function

Program denote measures:

$$\left[\begin{array}{c} \boxed{p} \\ \downarrow \\ \textcircled{\vec{x}} \end{array} \right] = \left[\begin{array}{c} \boxed{p} \\ \\ \phantom{\textcircled{\vec{x}}} \end{array} \right] \in \mathbb{M}[0,1] \subseteq ([0,1] \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$f \mapsto \int_0^1 \sum_{\vec{x} \in \{H,T\}^3} p^{\sum_i \langle x_i=H \rangle} (1-p)^{\sum_i \langle x_i=T \rangle} \langle \vec{x} = THH \rangle \cdot f(p) dp$$

Recognizing a density function

Program denote measures:

$$\left[\begin{array}{c} \boxed{p} \\ \downarrow \\ \vec{x} \end{array} \right] = \left[\begin{array}{c} \boxed{p} \\ \\ \phantom{\vec{x}} \end{array} \right] \in \mathbb{M}[0,1] \subseteq ([0,1] \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$f \mapsto \int_0^1 \sum_{\vec{x} \in \{H,T\}^3} p^{\sum_i \langle x_i=H \rangle} (1-p)^{\sum_i \langle x_i=T \rangle} \langle \vec{x} = THH \rangle \cdot f(p) dp$$

Recognizing a density function

Program denote measures:

$$\left[\begin{array}{c} \boxed{p} \\ \downarrow \\ \vec{x} \end{array} \right] = \left[\begin{array}{c} \boxed{p} \\ \\ \phantom{\vec{x}} \end{array} \right] \in \mathbb{M}[0,1] \subseteq ([0,1] \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$\begin{aligned} f &\mapsto \int_0^1 \sum_{\vec{x} \in \{H,T\}^3} p^{\sum_i \langle x_i=H \rangle} (1-p)^{\sum_i \langle x_i=T \rangle} \langle \vec{x} = THH \rangle \cdot f(p) dp \\ &= \int_0^1 p^2(1-p) \cdot f(p) dp \end{aligned}$$

Recognizing a density function

Program denote measures:

$$\left[\begin{array}{c} \boxed{p} \\ \downarrow \\ \vec{x} \end{array} \right] = \left[\begin{array}{c} \boxed{p} \\ \\ \phantom{\vec{x}} \end{array} \right] \in \mathbb{M}[0,1] \subseteq ([0,1] \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$\begin{aligned} f &\mapsto \int_0^1 \sum_{\vec{x} \in \{H,T\}^3} p^{\sum_i \langle x_i = H \rangle} (1-p)^{\sum_i \langle x_i = T \rangle} \langle \vec{x} = THH \rangle \cdot f(p) dp \\ &= \int_0^1 p^2 (1-p) \cdot f(p) dp \end{aligned}$$

Need to recognize simplified denotation as Beta distribution ...

Recognizing a density function

Goal: recognize $h(p) = p^2(1 - p)$ as the *density* of **beta 3 2**

Robustness challenge: many equivalent ways to write $p^2(1 - p)$ arise

Modularity challenge: many distribution families (**beta, normal, ...**) known

Recognizing a density function

Goal: recognize $h(p) = p^2(1 - p)$ as the *density* of **beta 3 2**

Robustness challenge: many equivalent ways to write $p^2(1 - p)$ arise

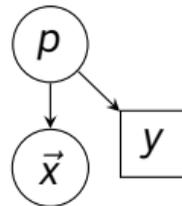
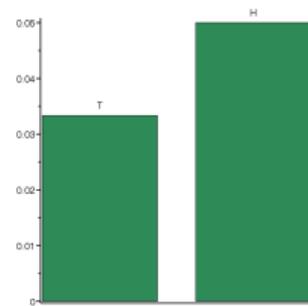
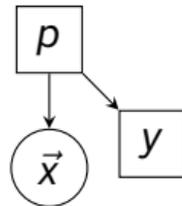
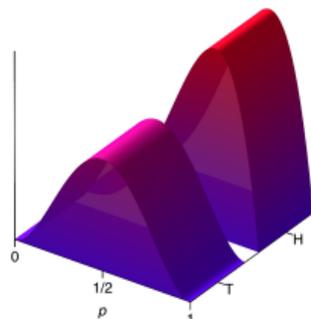
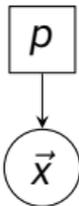
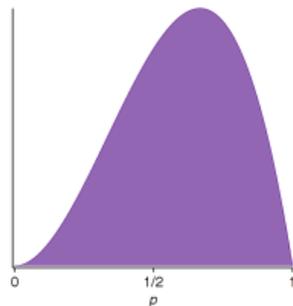
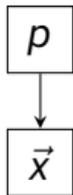
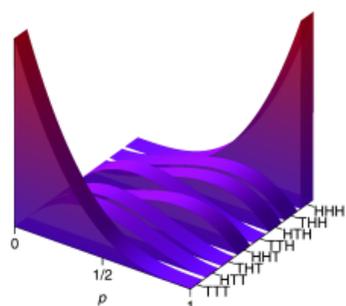
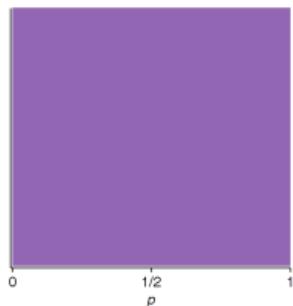
Modularity challenge: many distribution families (**beta, normal, ...**) known

Solution: characterize density functions by their **holonomic representation**,
a *homogeneous linear differential equation* such as

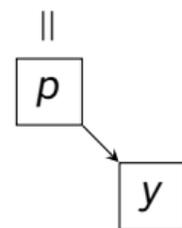
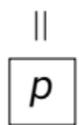
$$p(1 - p) \cdot h'(p) + (p - 2(1 - p)) \cdot h(p) = 0$$

computed *compositionally!*

$\Pr(p)$ $\xrightarrow{\text{bind}}$ $\Pr(p, \vec{x})$ $\xrightarrow{\text{disintegrate}}$ $\Pr(p | \vec{x})$ $\xrightarrow{\text{bind}}$ $\Pr(p, y | \vec{x})$ $\xrightarrow{\text{integrate}}$ $\Pr(y | \vec{x})$

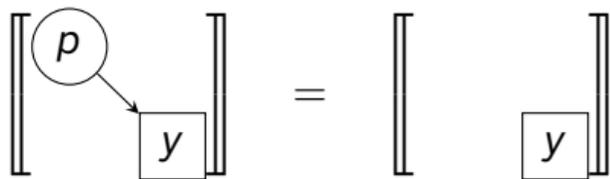


simplify $\left\{ \begin{array}{l} \text{recognize} \\ \text{integrate} \end{array} \right.$



Eliminating a random variable

Program denote measures:



Eliminating a random variable

Program denote measures:

$$\left[\begin{array}{c} p \\ \downarrow \\ y \end{array} \right] = \left[\begin{array}{c} y \end{array} \right] \in \mathcal{M}\{\mathbf{H}, \mathbf{T}\} \subseteq (\{\mathbf{H}, \mathbf{T}\} \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$f \mapsto \int_0^1 p^2(1-p) \sum_{y \in \{\mathbf{H}, \mathbf{T}\}} p^{\langle y=\mathbf{H} \rangle} (1-p)^{\langle y=\mathbf{T} \rangle} \cdot f(y) dp$$

Eliminating a random variable

Program denote measures:

$$\left[\begin{array}{c} p \\ \swarrow \\ y \end{array} \right] = \left[\begin{array}{c} y \end{array} \right] \in \mathbb{M}\{H, T\} \subseteq (\{H, T\} \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$f \mapsto \int_0^1 p^2(1-p) \sum_{y \in \{H, T\}} p^{\langle y=H \rangle} (1-p)^{\langle y=T \rangle} \cdot f(y) dp$$

Eliminating a random variable

Program denote measures:

$$\left[\begin{array}{c} p \\ \downarrow \\ y \end{array} \right] = \left[\begin{array}{c} y \end{array} \right] \in \mathcal{M}\{H, T\} \subseteq (\{H, T\} \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$f \mapsto \int_0^1 p^2(1-p) \sum_{y \in \{H, T\}} p^{\langle y=H \rangle} (1-p)^{\langle y=T \rangle} \cdot f(y) dp$$

Eliminating a random variable

Program denote measures:

$$\left[\begin{array}{c} p \\ \downarrow \\ y \end{array} \right] = \left[\begin{array}{c} y \end{array} \right] \in \mathcal{M}\{H, T\} \subseteq (\{H, T\} \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$f \mapsto \int_0^1 p^2(1-p) \sum_{y \in \{H, T\}} p^{\langle y=H \rangle} (1-p)^{\langle y=T \rangle} \cdot f(y) dp$$

Eliminating a random variable

Program denote measures:

$$\left[\begin{array}{c} p \\ \downarrow \\ y \end{array} \right] = \left[\begin{array}{c} y \end{array} \right] \in \mathcal{M}\{H, T\} \subseteq (\{H, T\} \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$f \mapsto \int_0^1 p^2(1-p) \sum_{y \in \{H, T\}} p^{\langle y=H \rangle} (1-p)^{\langle y=T \rangle} \cdot f(y) dp$$

Eliminating a random variable

Program denote measures:

$$\left[\begin{array}{c} p \\ \swarrow \\ y \end{array} \right] = \left[\begin{array}{c} y \end{array} \right] \in \mathbb{M}\{\text{H}, \text{T}\} \subseteq (\{\text{H}, \text{T}\} \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$f \mapsto \int_0^1 p^2(1-p) \sum_{y \in \{\text{H}, \text{T}\}} p^{\langle y=\text{H} \rangle} (1-p)^{\langle y=\text{T} \rangle} \cdot f(y) dp$$

Eliminating a random variable

Program denote measures:

$$\left[\begin{array}{c} \textcircled{p} \\ \downarrow \\ \boxed{y} \end{array} \right] = \left[\boxed{y} \right] \in \mathbb{M}\{H, T\} \subseteq (\{H, T\} \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$\begin{aligned} f &\mapsto \int_0^1 p^2(1-p) \sum_{y \in \{H, T\}} p^{\langle y=H \rangle} (1-p)^{\langle y=T \rangle} \cdot f(y) dp \\ &= \sum_{y \in \{H, T\}} \left(\int_0^1 p^2(1-p) p^{\langle y=H \rangle} (1-p)^{\langle y=T \rangle} dp \right) \cdot f(y) \end{aligned}$$

Eliminating a random variable

Program denote measures:

$$\left[\begin{array}{c} \textcircled{p} \\ \downarrow \\ \boxed{y} \end{array} \right] = \left[\boxed{y} \right] \in \mathbb{M}\{H, T\} \subseteq (\{H, T\} \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

Measures compute expectations:

$$\begin{aligned} f &\mapsto \int_0^1 p^2(1-p) \sum_{y \in \{H, T\}} p^{\langle y=H \rangle} (1-p)^{\langle y=T \rangle} \cdot f(y) dp \\ &= \sum_{y \in \{H, T\}} \left(\int_0^1 p^2(1-p) p^{\langle y=H \rangle} (1-p)^{\langle y=T \rangle} dp \right) \cdot f(y) \\ &= \sum_{y \in \{H, T\}} \left(\frac{1}{20} \right)^{\langle y=H \rangle} \left(\frac{1}{30} \right)^{\langle y=T \rangle} \cdot f(y) \end{aligned}$$

Eliminating a random variable

Program denote measures:

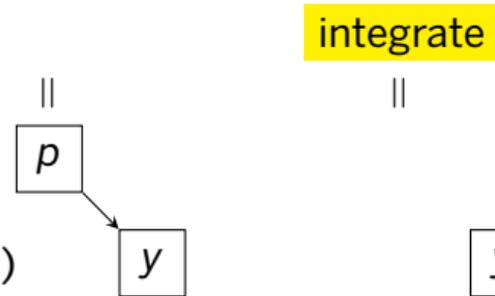
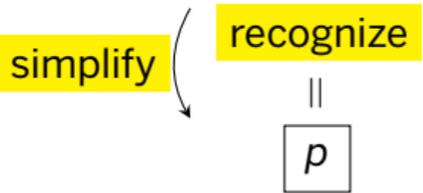
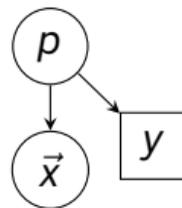
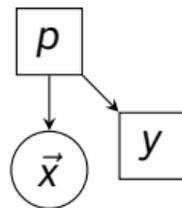
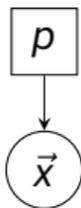
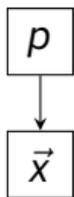
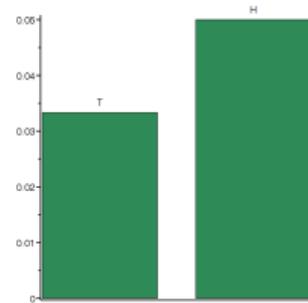
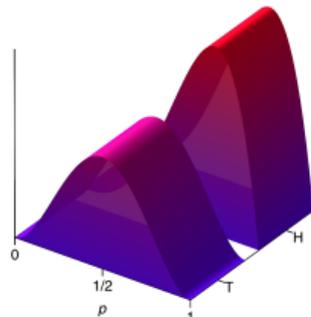
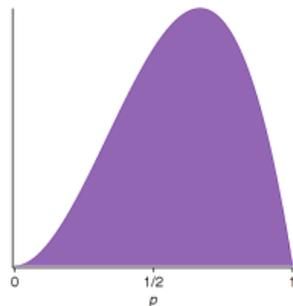
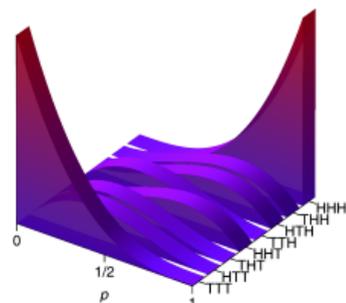
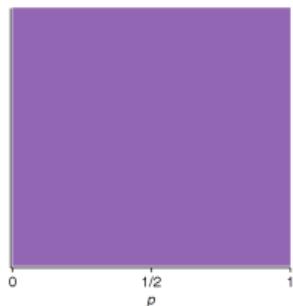
$$\left[\begin{array}{c} \textcircled{p} \\ \downarrow \\ \boxed{y} \end{array} \right] = \left[\begin{array}{c} \\ \boxed{y} \end{array} \right] \in \mathbb{M}\{H, T\} \subseteq (\{H, T\} \rightarrow \mathbb{R}_+) \rightarrow \mathbb{R}_+$$

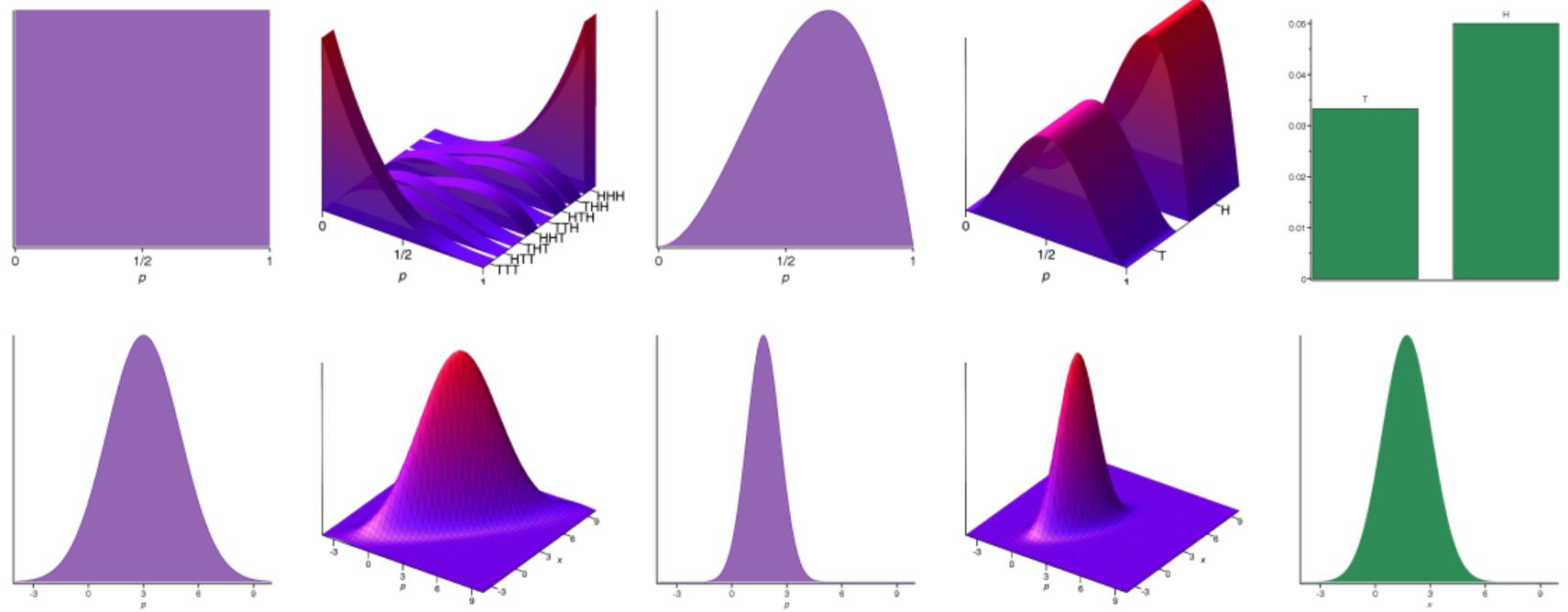
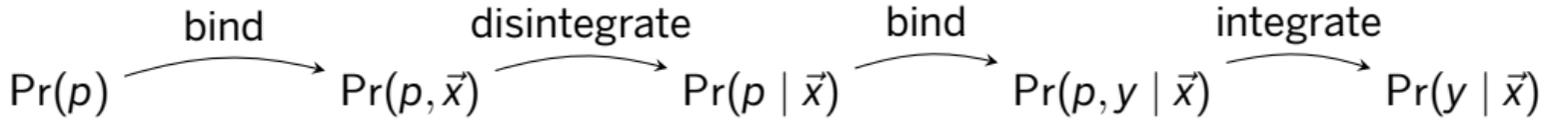
Measures compute expectations:

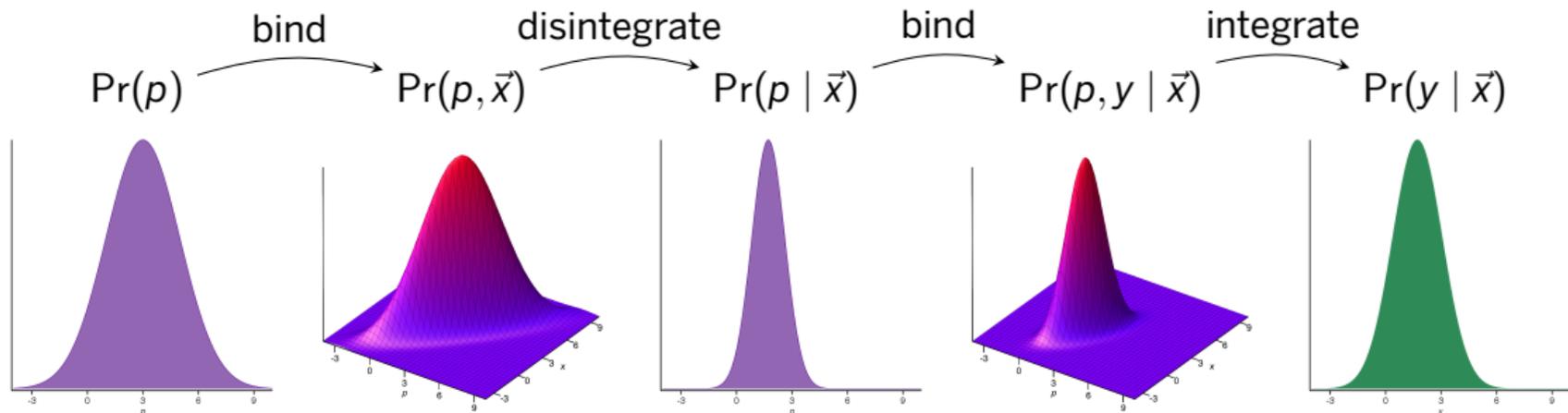
$$\begin{aligned} f &\mapsto \int_0^1 p^2(1-p) \sum_{y \in \{H, T\}} p^{\langle y=H \rangle} (1-p)^{\langle y=T \rangle} \cdot f(y) dp \\ &= \sum_{y \in \{H, T\}} \left(\int_0^1 p^2(1-p) p^{\langle y=H \rangle} (1-p)^{\langle y=T \rangle} dp \right) \cdot f(y) \\ &= \sum_{y \in \{H, T\}} \left(\frac{1}{20} \right)^{\langle y=H \rangle} \left(\frac{1}{30} \right)^{\langle y=T \rangle} \cdot f(y) \end{aligned}$$

bind disintegrate bind integrate

$\Pr(p)$ $\Pr(p, \vec{x})$ $\Pr(p | \vec{x})$ $\Pr(p, y | \vec{x})$ $\Pr(y | \vec{x})$







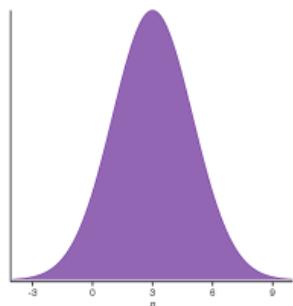
An unknown random process yields a stateless particle whose one-dimensional position can be measured repeatedly to produce a real number.

We assume that the position p of the particle is distributed normally with mean 3 and standard deviation 2.

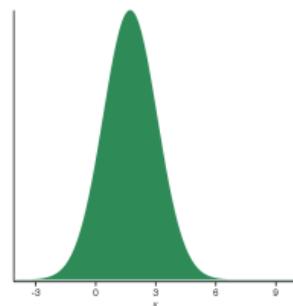
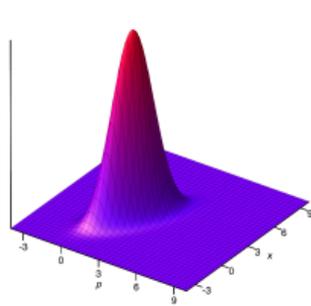
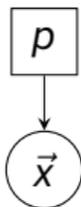
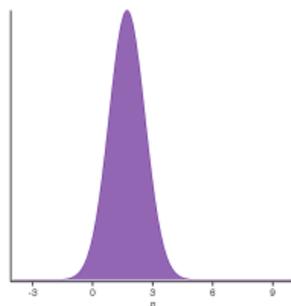
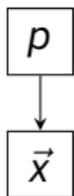
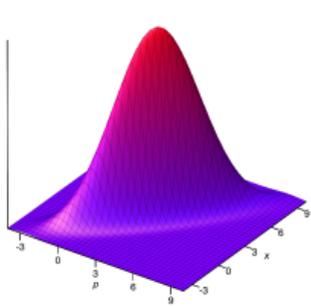
We measure the particle 3 times, each time drawing independently from the normal distribution with mean p and standard deviation 1, and observe $-1.4, +1.0, -0.2$.

What is the distribution of the next measurement?

$$\Pr(p) \xrightarrow{\text{bind}} \Pr(p, \vec{x}) \xrightarrow{\text{disintegrate}} \Pr(p | \vec{x}) \xrightarrow{\text{bind}} \Pr(p, y | \vec{x}) \xrightarrow{\text{integrate}} \Pr(y | \vec{x})$$



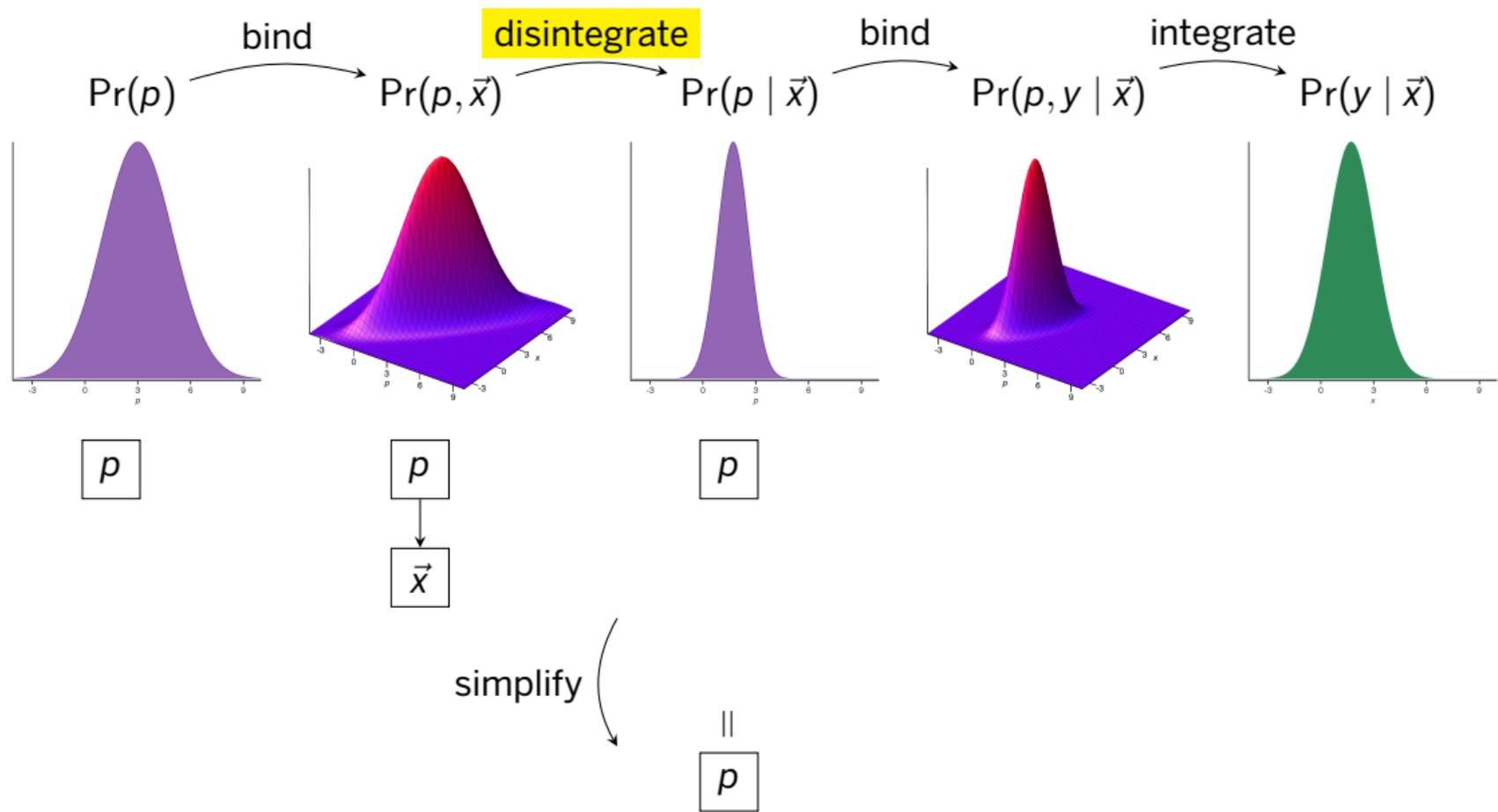
p



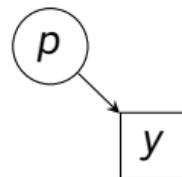
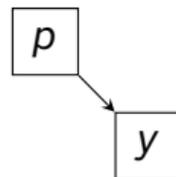
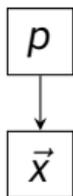
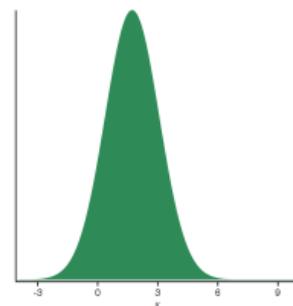
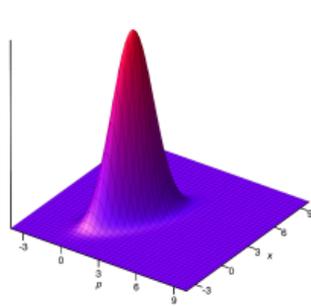
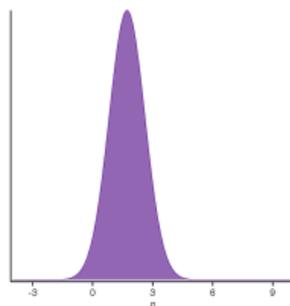
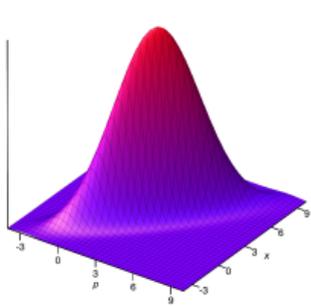
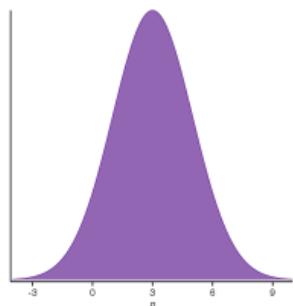
simplify

||

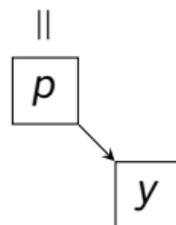
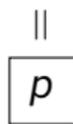
0



$$\Pr(p) \xrightarrow{\text{bind}} \Pr(p, \vec{x}) \xrightarrow{\text{disintegrate}} \Pr(p | \vec{x}) \xrightarrow{\text{bind}} \Pr(p, y | \vec{x}) \xrightarrow{\text{integrate}} \Pr(y | \vec{x})$$



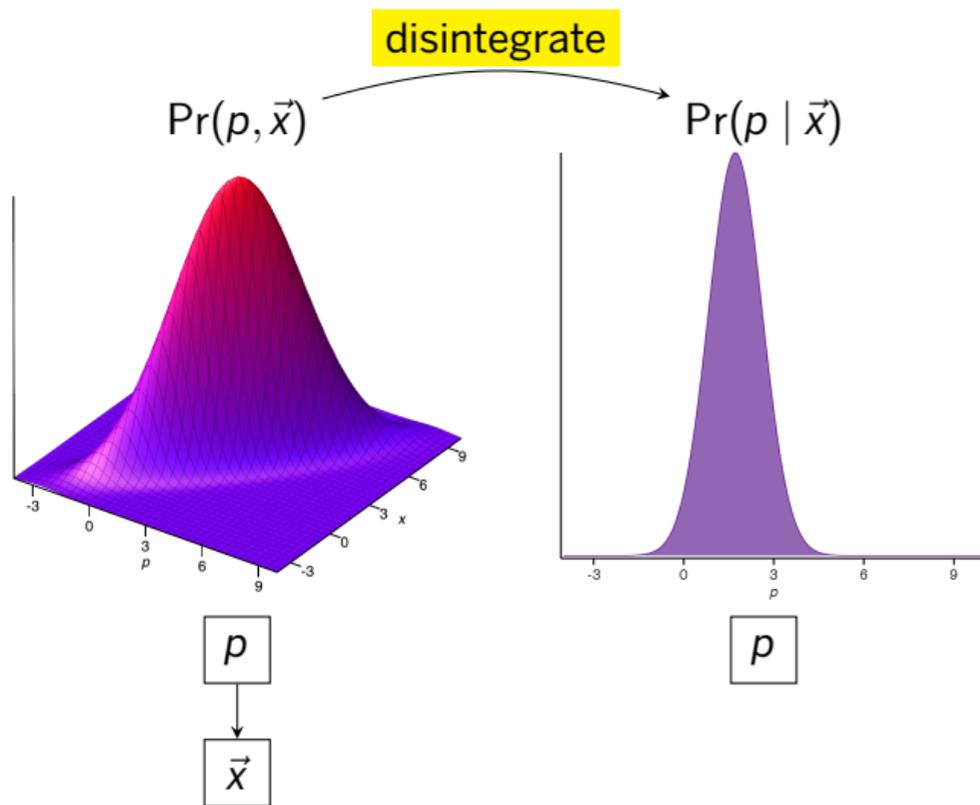
simplify



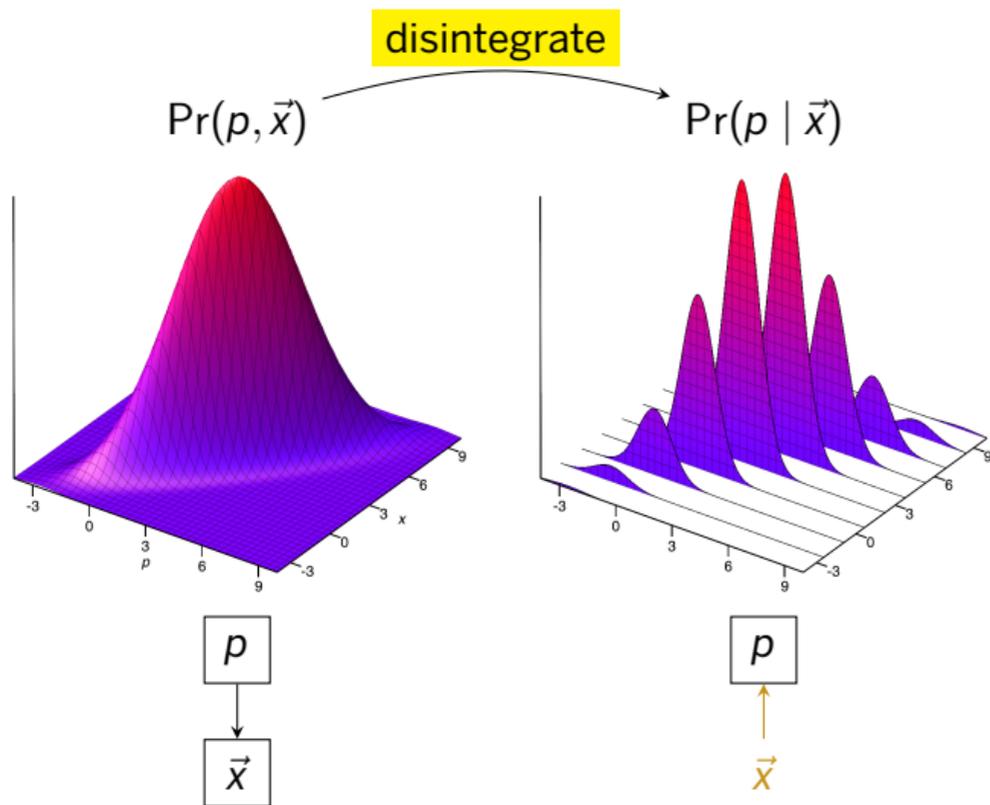
||



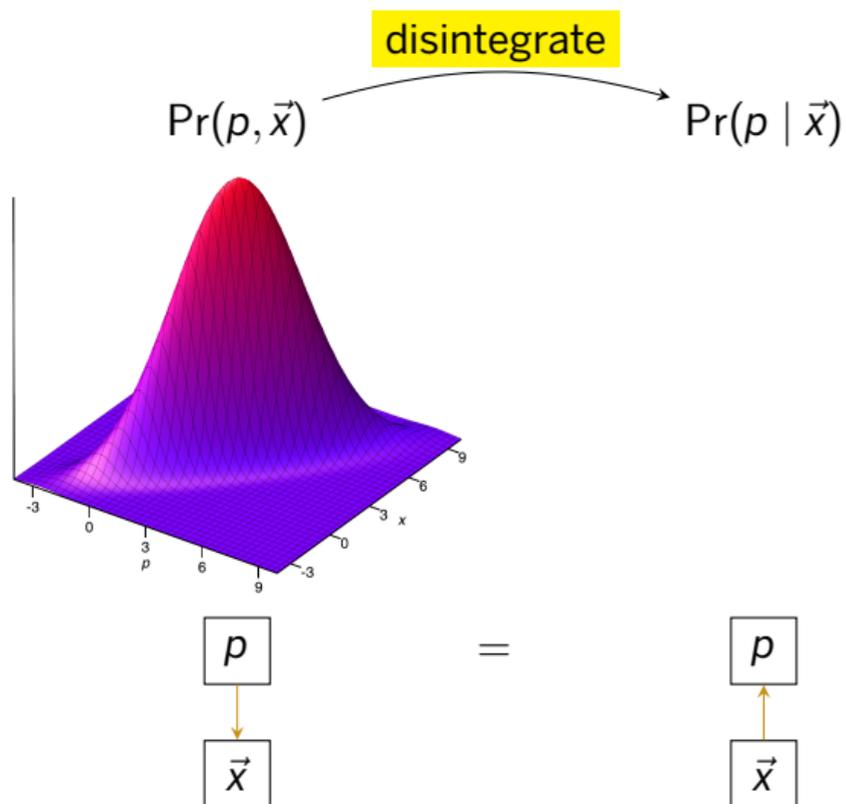
Disintegrating a joint measure



Disintegrating a joint measure



Disintegrating a joint measure

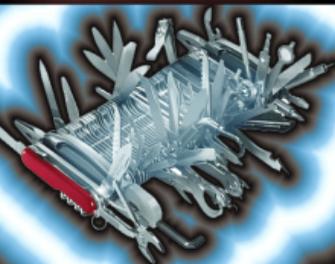
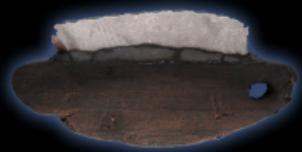


Program transformation
derived from semantics.

Tricky when \vec{x} is not just drawn
from a primitive distribution:

- ▶ total momentum
- ▶ loop over array
- ▶ clamped measurement
- ▶ coordinate-wise MCMC

Addressed in recent work.
(ICFP 2016, POPL 2017,
ICFP 2017)



7

8

9

÷

4

5

6

×

1

2

3

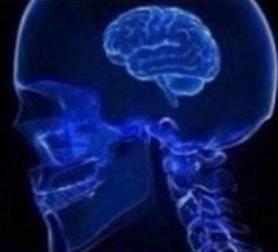
-

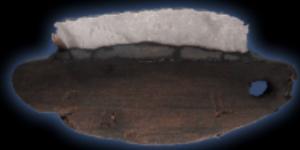
0

.

=

+





7
plate

8
lambda

9
apply

÷
disintegrate



4
pair

5
fst

6
snd

×
bind



1
dirac

2
beta

3
normal

—
gradient



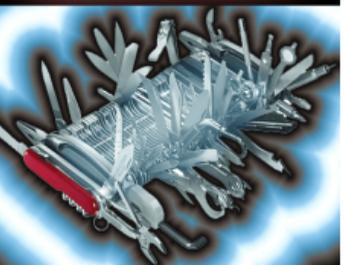
0
mzero

·
factor

=
simplify

+
plus





Thanks!

Jacques Carette

Oleg Kiselyov

Wazim Mohammed Ismail

Praveen Narayanan

Norman Ramsey

Wren Romano

Sam Tobin-Hochstadt

Rajan Walia

Robert Zinkov

÷

disintegrate

×

bind

—

gradient

0

mzero

.

factor

=

simplify

+

plus

