

# Introduction to relational verification: probabilistic programs

Gilles Barthe  
MPI-SP, Germany  
IMDEA Software Institute, Spain

August 2, 2019

# Probabilistic programs

Programs with

- ▶ random assignments  
(randomized algorithms, cryptography)
- ▶ conditioning (machine learning)

Focusing on the former

$c ::=$	$x \leftarrow e$	deterministic assignment
	$x \overset{\$}{\leftarrow} d$	probabilistic assignment
	$c; c$	sequencing
	if $e$ then $c$ else $c$	conditional
	while $e$ do $c$	while loop
	$x \leftarrow \mathcal{F}(e)$	procedure call

# Semantics of probabilistic programs

Denotational semantics (Kozen):

- ▶ **Mem** =  $\text{Var} \rightarrow \text{Val}$
- ▶  $\llbracket c \rrbracket : \mathbf{Mem} \rightarrow \text{Distr}(\mathbf{Mem})$

where  $\text{Distr}(A)$  is the set of functions  $\mu : A \rightarrow [0, 1]$  s.t.

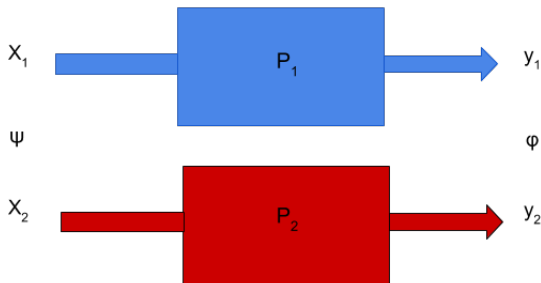
- ▶  $\text{supp}(\mu) = \{a \in A \mid \mu(a) > 0\}$  of  $\mu$  is discrete;
- ▶  $|\mu| = \sum_{a \in A} \mu(a)$  of  $\mu$  is defined and verifies  $|\mu| \leq 1$ .

Operational semantics: Markov chains

## Semantics of higher-order languages

- ▶ Aumann's theorem: measurable spaces do not form a ccc
- ▶ Quasi-Borel spaces do

## 2-properties: program equivalence

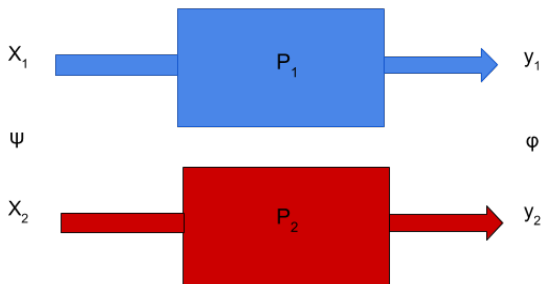


Precondition:  $x_1 = x_2$

Postcondition:  $y_1 = y_2$

$$v_1 \stackrel{\$}{\leftarrow} B_p; w_1 \stackrel{\$}{\leftarrow} B_{p'}; y_1 \leftarrow v_1 \wedge w_1$$
$$y_2 \stackrel{\$}{\leftarrow} B_{p \cdot p'}$$

## 2-properties: non-interference



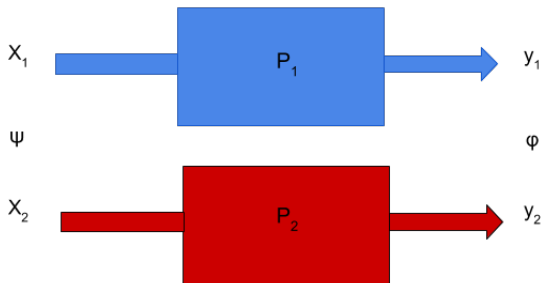
Precondition:  $x_1 = x_2$

Postcondition:  $y_1 = y_2$

$$k_1 \stackrel{\$}{\leftarrow} \mathcal{U}_{Z_p}; y_1 \leftarrow x_1 + k_1$$

$$k_2 \stackrel{\$}{\leftarrow} \mathcal{U}_{Z_p}; y_2 \leftarrow x_2 + k_2$$

## 2-properties: monotonicity



Precondition:  $x_1 \leq x_2$

Postcondition:  $y_1 \leq y_2$

$$y_1 \stackrel{\$}{\leftarrow} \mathcal{B}_{x_1}$$
$$y_2 \stackrel{\$}{\leftarrow} \mathcal{B}_{x_2}$$

# Interpreting post-conditions

- ▶ What does the postcondition  $y_1 = y_2$  mean?

$$\forall a \in A, \Pr_{y_1 \stackrel{\$}{\leftarrow} \mu_1} [y_1 = a] = \Pr_{y_2 \stackrel{\$}{\leftarrow} \mu_2} [y_2 = a]$$

- ▶ What does the postcondition  $y_1 \leq y_2$  mean?

$$\forall a \in A, \Pr_{y_1 \stackrel{\$}{\leftarrow} \mu_1} [y_1 \geq a] \leq \Pr_{y_2 \stackrel{\$}{\leftarrow} \mu_2} [y_2 \geq a]$$

- ▶ What does the postcondition  $\phi$  mean?

$$\exists \mu \in \mathcal{C}_\phi(\mu_1, \mu_2)$$

# Couplings (Doeblin 1938)

Let  $\mu_1 \in \text{Distr}(A_1)$  and  $\mu_2 \in \text{Distr}(A_2)$ . The set  $\mathcal{C}(\mu_1, \mu_2)$  of couplings for  $\mu_1$  and  $\mu_2$  is defined as

$$\{\mu \in \text{Distr}(A_1 \times A_2) \mid \pi_1(\mu) = \mu_1 \wedge \pi_2(\mu) = \mu_2\}$$

## Marginals

Given  $\mu \in \text{Distr}(A_1 \times A_2)$  define  $\pi_1(\mu) \in \text{Distr}(A_1)$  and  $\pi_2(\mu) \in \text{Distr}(A_2)$  by

$$\pi_1(\mu)(a_1) = \sum_{a_2 \in A_2} \mu(a_1, a_2) \quad \pi_2(\mu)(a_2) = \sum_{a_1 \in A_1} \mu(a_1, a_2)$$

# Couplings and total variation distance

For every  $\mu \in \mathcal{C}(\mu_1, \mu_2)$ ,

$$TV(\mu_1, \mu_2) = \max_{X \subseteq A} |\Pr_{\mu_1}[X] - \Pr_{\mu_2}[X]| \leq \Pr_{\mu}[\lambda x_1, x_2. x_1 \neq x_2]$$

Moreover there is a “optimal” coupling  $\mu_0$  such that

$$TV(\mu_1, \mu_2) = \Pr_{\mu_0}[\lambda x_1, x_2. x_1 \neq x_2]$$

# $R$ -couplings

Let  $\mu_1 \in \text{Distr}(A_1)$  and  $\mu_2 \in \text{Distr}(A_2)$ . Let  $R \subseteq A_1 \times A_2$ . The set  $\mathcal{C}_R(\mu_1, \mu_2)$  of  $R$ -couplings for  $\mu_1$  and  $\mu_2$  is defined as

$$\{\mu \in \text{Distr}(A_1 \times A_2) \mid \pi_1(\mu) = \mu_1 \wedge \pi_2(\mu) = \mu_2 \wedge \text{supp}(\mu) \subseteq R\}$$

## Strassen Theorem

For every  $\mu_1 \in \text{Distr}(A_1)$  and  $\mu_2 \in \text{Distr}(A_2)$  s.t.  $|\mu_1| = |\mu_2| = 1$ , the following are equivalent:

- ▶  $\mathcal{C}_R(\mu_1, \mu_2) \neq \emptyset$
- ▶ for every  $X \subseteq A_1$ ,  $\mu_1(X) \leq \mu_2(R(X))$

# Consequences of Strassen's Theorem

Let  $A_1 = A_2 = A$ . Assume  $|\mu_1| = |\mu_2| = 1$ .

## Equality couplings

The following are equivalent:

- ▶  $\mu_1 = \mu_2$
- ▶  $\mathcal{C}_=(\mu_1, \mu_2) \neq \emptyset$

## Stochastic dominance

Assume  $(A, \leq)$  is a partial order. Then the following are equivalent:

- ▶  $\mathcal{C}_\leq(\mu_1, \mu_2) \neq \emptyset$
- ▶ for every  $a$ ,  $\mu_1(\{x \in \mathbb{Z} \mid x \geq a\}) \leq \mu_2(\{x \in \mathbb{Z} \mid x \geq a\})$

Also Fundamental Theorem of  $R$ -couplings (and specializations for cryptography).

# Coupling coins

Let  $\mu_1, \mu_2$  be the uniform distribution over  $\{0, 1\}$ .

- ▶ trivial coupling:  $\mu(x, y) = \frac{1}{4}$
- ▶ equality coupling:  $\mu(x, x) = \frac{1}{2}$  and  $\mu(x, \neg x) = 0$
- ▶ inequality coupling:  $\mu(x, \neg x) = \frac{1}{2}$  and  $\mu(x, x) = 0$

Let  $\mu_1 = \mathcal{B}_{p_1}$  and  $\mu_2 = \mathcal{B}_{p_2}$  with  $p_1 < p_2$ . Define  $\mu$  such that

$$\begin{aligned}\mu(1, 1) &= p_1 \\ \mu(0, 1) &= p_2 - p_1 \\ \mu(0, 0) &= 1 - p_2\end{aligned}$$

Then  $\mu$  is a  $R$ -coupling for  $\mu_1$  and  $\mu_2$ , with  $R = \lambda_{x_1 x_2} \cdot x_1 \leq x_2$ .

# Balls and bins

- ▶ Input:  $x \in [0, 1]$
- ▶ Start with  $N$  balls and two empty bins (left and right)
- ▶ For each ball in left bin with probability  $x$  and right bin with probability  $1 - x$
- ▶ Count number  $y$  of balls in left bin

If  $x_1 \leq x_2$  then  $y_2$  stochastically dominates  $y_1$ . Formally

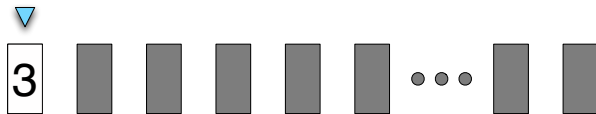
$$x_1 \leq x_2 \implies \mathcal{C}_{\lambda_{n_1, n_2}, n_1 \leq n_2}(y_1, y_2) \neq \emptyset$$

Proof: induction and coupling

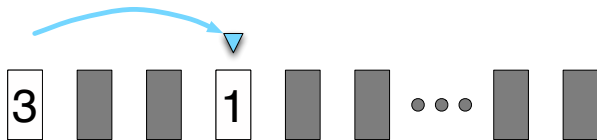
# Dynkin's card trick

- ▶ Input: position in  $\{0, \dots, 9\}$
- ▶ Repeat:
  - ▶ Draw uniformly random card  $\in \{1, \dots, 10\}$
  - ▶ Go forward that many steps
- ▶ Output last position before crossing 100

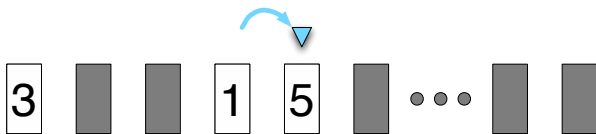
# In pictures



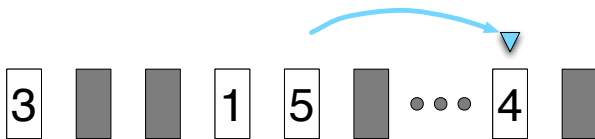
# In pictures



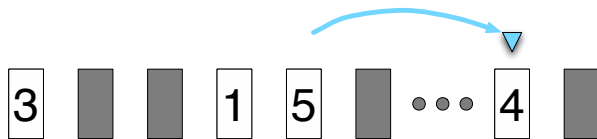
# In pictures



# In pictures

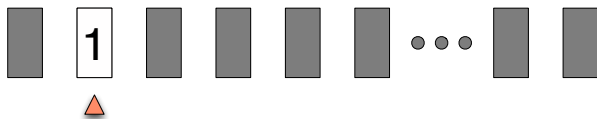


## In pictures

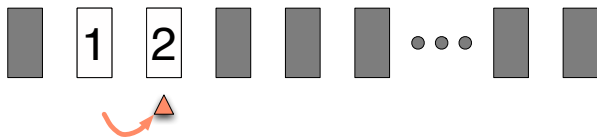


Output last position: 99

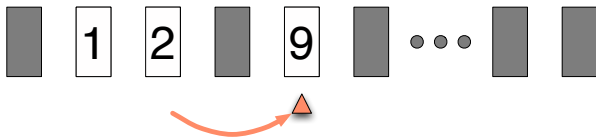
## Starting at a different position



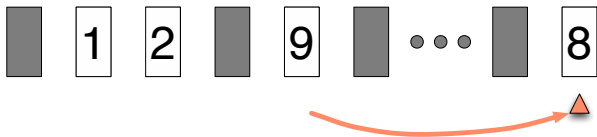
## Starting at a different position



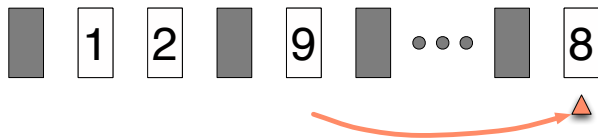
## Starting at a different position



## Starting at a different position

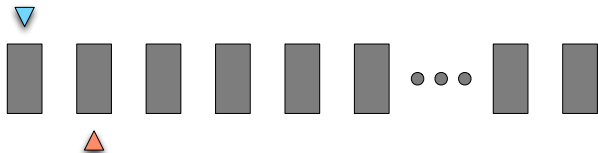


## Starting at a different position

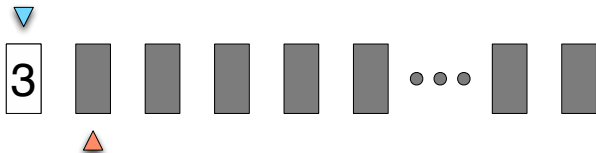


How close are the two output distributions?

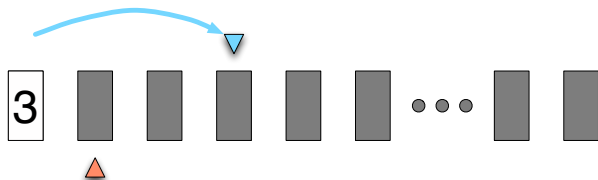
## Combine first process and second process



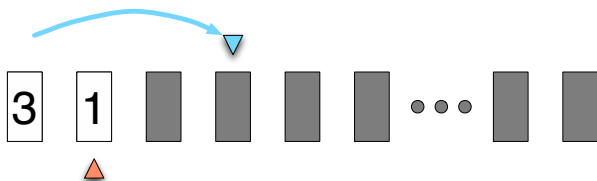
## Combine first process and second process



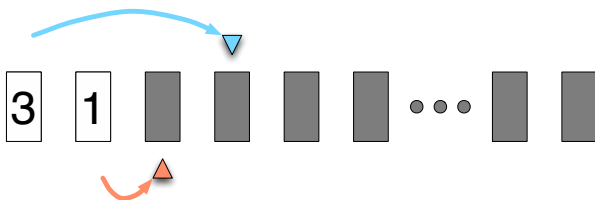
## Combine first process and second process



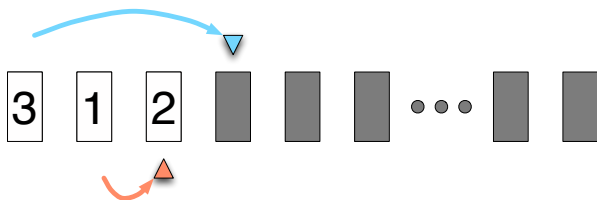
## Combine first process and second process



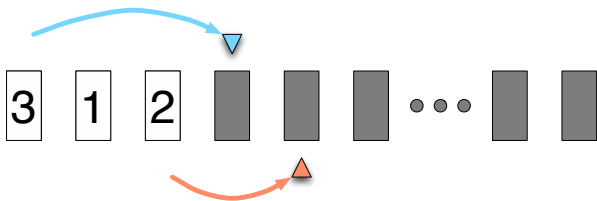
## Combine first process and second process



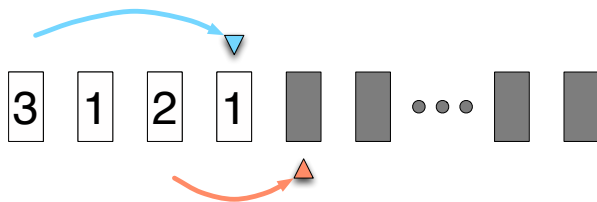
## Combine first process and second process



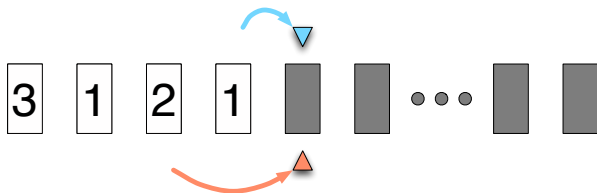
## Combine first process and second process



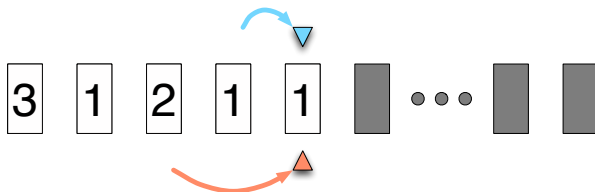
## Combine first process and second process



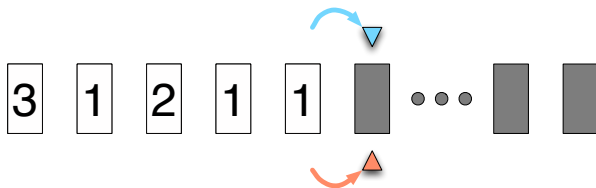
## Combine first process and second process



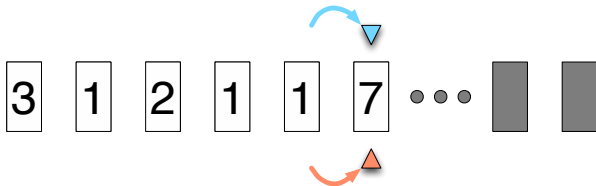
## Combine first process and second process



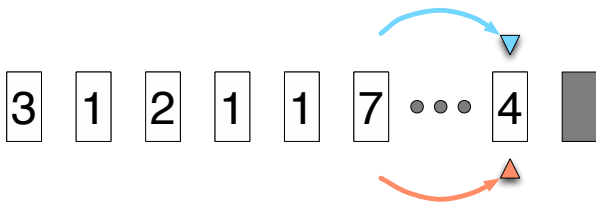
# Combine first process and second process



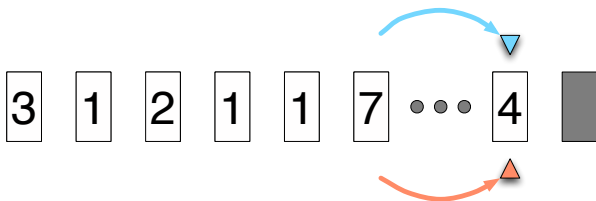
## Combine first process and second process



## Combine first process and second process



## Combine first process and second process



Consequence: for every number of steps  $N$ ,

$$TV(\text{dynkin}(x_1), \text{dynkin}(x_2)) \leq (9/10)^{N/5-2}$$

# Probabilistic Relational Hoare Logic

**Judgment:**

$$\models \{\Phi\}c_1 \sim c_2\{\Psi\}$$

where  $\Phi, \Psi \subseteq \mathbf{Mem} \times \mathbf{Mem}$

**Validity:** for every  $(m_1, m_2)$  s.t.  $(m_1, m_2) \in \Phi$ , there exists  $\mu$  in

$$\mathcal{C}_\Psi(\llbracket c_1 \rrbracket(m_1), \llbracket c_2 \rrbracket(m_2))$$

## Structural and 2-sided rules

$$\frac{\models \{\Phi'\}c_1 \sim c_2\{\Psi'\} \quad \Phi \implies \Phi' \quad \Psi' \implies \Psi}{\models \{\Phi\}c_1 \sim c_2\{\Psi\}} \text{ [CONSEQ]}$$

$$\frac{\models \{\Phi_1\}c_1 \sim c_2\{\Psi\} \quad \models \{\Phi_2\}c_1 \sim c_2\{\Psi\}}{\models \{\Phi_1 \vee \Phi_2\}c_1 \sim c_2\{\Psi\}} \text{ [CASE]}$$

$$\frac{\models \{\Phi\}c_1 \sim c_2\{\Theta\} \quad \models \{\Theta\}c'_1 \sim c'_2\{\Psi\}}{\models \{\Phi\}c_1; c'_1 \sim c_2; c'_2\{\Psi\}} \text{ [SEQ]}$$

$$\frac{}{\models \{\Psi[e_1/x_1][e_2/x_2]\}x_1 \leftarrow e_1 \sim x_2 \leftarrow e_2\{\Psi\}} \text{ [ASSN]}$$
$$\Phi \implies e_1 = e_2$$

$$\frac{\models \{\Phi \wedge e_1\}c_1 \sim c_2\{\Psi\} \quad \models \{\Phi \wedge \neg e_1\}c'_1 \sim c'_2\{\Psi\}}{\models \{\Phi\}\text{if } e_1 \text{ then } c_1 \text{ else } c'_1 \sim \text{if } e_2 \text{ then } c_2 \text{ else } c'_2\{\Psi\}} \text{ [COND]}$$

$$\frac{\Theta \implies e_1 = e_2 \quad \models \{\Theta \wedge e_1\}c_1 \sim c_2\{\Theta\}}{\models \{\Theta\}\text{while } e_1 \text{ do } c_1 \sim \text{while } e_2 \text{ do } c_2\{\Theta \wedge \neg e_1\}} \text{ [WHILE]}$$

# One-sided rules

$$\frac{}{\models \{\Psi[e_1/x_1]\}x_1 \leftarrow e_1 \sim \text{skip}\{\Psi\}} \text{[ASSG-L]}$$
$$\frac{\models \{\Phi \wedge e_1\}c_1 \sim c_2\{\Psi\} \quad \models \{\Phi \wedge \neg e_1\}c'_1 \sim c_2\{\Psi\}}{\models \{\Phi\}\text{if } e_1 \text{ then } c_1 \text{ else } c'_1 \sim c_2\{\Psi\}} \text{[COND-L]}$$
$$\frac{\models \{\Theta \wedge e_1\}c_1 \sim \text{skip}\{\Theta\} \quad \text{ast}(\text{while } e_1 \text{ do } c_1)}{\models \{\Theta\}\text{while } e_1 \text{ do } c_1 \sim \text{skip}\{\Theta \wedge \neg e_1\}} \text{[WHILE-L]}$$

# Random sampling

$$\frac{\mathcal{C}_\Theta(\llbracket \mu_1 \rrbracket, \llbracket \mu_2 \rrbracket) \neq \emptyset}{\Phi \triangleq \forall v_1 : T_1, v_2 : T_2, \Theta(v_1, v_2) \implies \Psi[v_1/x_1][v_2/x_2]} \text{ [RAND]}$$
$$\models \{\Phi\} x_1 \stackrel{\$}{\leftarrow} \mu_1 \sim x_2 \stackrel{\$}{\leftarrow} \mu_2 \{\Psi\}$$

$$\frac{h : T_1 \xrightarrow{1-1} T_2}{\Phi \triangleq \forall v_1 : T_1, v_2 : T_2, h(v_1) = v_2 \implies \Psi[v_1/x_1][v_2/x_2]} \text{ [RAND-UNIF]}$$
$$\models \{\Phi\} x_1 \stackrel{\$}{\leftarrow} \mathcal{U}_{T_1} \sim x_2 \stackrel{\$}{\leftarrow} \mathcal{U}_{T_2} \{\Psi\}$$

$$\frac{}{\models \{\forall v_1 \in \text{supp}(d_1), \Psi[v_1/x_1]\} x_1 \stackrel{\$}{\leftarrow} d_1 \sim \text{skip}\{\Psi\}} \text{ [RAND-L]}$$

# Product programs

- ▶ Every proof in pRHL builds a product program
- ▶ Product programs can be made explicit

$$\models \{\Phi\}c_1 \sim c_2\{\Psi\} \rightsquigarrow c$$

Example:

$$\frac{\models \{\Phi \wedge \Phi'\}c_1 \sim c_2\{\Psi\} \rightsquigarrow c \quad \models \{\Phi \wedge \neg\Phi'\}c_1 \sim c_2\{\Psi\} \rightsquigarrow c_{\neg}}{\models \{\Phi\}c_1 \sim c_2\{\Psi\} \rightsquigarrow \text{if } \Phi' \text{ then } c^{\times} \text{ else } c_{\neg}^{\times}}$$

Application to Dynkin's trick

- ▶ product program simulates two programs
- ▶ bound probability of coinciding in product program

# Balls and bins

```
y := 0; i := 0;  
while i ≤ N do  
  b ←$ Bx;  
  y ← y + b;  
  i ← i + 1;
```

Goal:

$$\models \{x_1 \leq x_2\} c_1 \sim c_2 \{y_1 \leq y_2\}$$

Proof: by [WHILE], [SEQ], [ASS] and [FRAME] suffices to show

$$\models \{y_1 \leq y_2\} b_1 \overset{\$}{\leftarrow} B_{x_1} \sim b_2 \overset{\$}{\leftarrow} B_{x_2} \{y_1 + b_1 \leq y_2 + b_2\}$$

# Eager sampling

$$\models \{z_1 = z_2\} c_1 \sim c_2 \{x_1 = x_2\}$$

where

$$c_1 \triangleq x_1 \stackrel{\$}{\leftarrow} \mu; \text{ if } z_1 = 0 \text{ then } z_1 \leftarrow z_1 + x_1 \text{ else } x_1 \leftarrow z_1$$

$$c_2 \triangleq \text{ if } z_2 = 0 \text{ then } x_2 \stackrel{\$}{\leftarrow} \mu; z_2 \leftarrow z_2 + x_2 \text{ else } x_2 \leftarrow z_2$$

# One-time pad

$$\begin{aligned}k &\leftarrow^{\$} \mathcal{U}_{\{0,1\}^n}; \\y &\leftarrow x \oplus k\end{aligned}$$

Goal:

$$\models \{\top\} c_1 \sim c_2 \{y_1 = y_2\}$$

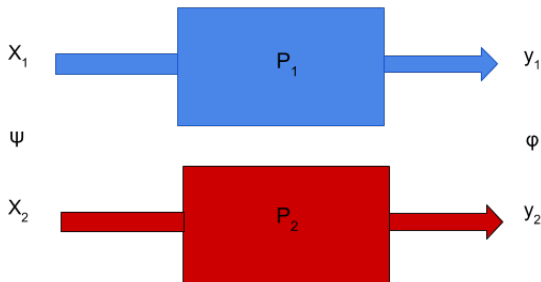
Proof: by [ASS], we must show

$$\models \{\top\} k_1 \leftarrow^{\$} \mathcal{U}_{\{0,1\}^n} \sim k_2 \leftarrow^{\$} \mathcal{U}_{\{0,1\}^n} \{x_1 \oplus k_1 = x_2 \oplus k_2\}$$

By [RAND] with  $\mu(k_1, k_2) = \frac{\mathbb{1}_{x_1 \oplus k_1 = x_2 \oplus k_2}}{2^n}$ , must show

$$\forall x_1 \ x_2, \ x_1 \oplus k_1 = x_2 \oplus k_2 \implies x_1 \oplus k_1 = x_2 \oplus k_2$$

## 2-properties: uniformity



Precondition:  $x_1 = x_2$

Postcondition:  $y_1 = a_1 \implies y_2 = a_2$

- ▶ By the Fundamental Theorem of Couplings,  
 $\Pr_\mu[y = a_1] \leq \Pr_\mu[y = a_2]$
- ▶ By quantifying over  $a_1$  and  $a_2$ ,  $y$  is uniform

# One-time pad, revisited

$$k \stackrel{\$}{\leftarrow} \mathcal{U}_{\{0,1\}^n}; y \leftarrow x \oplus k$$

Uniformity of one-time pad

$$\models \{\top\} c_1 \sim c_2 \{y_1 = a_1 \Rightarrow y_2 = a_2\}$$

Proof: by [ASS] and [RAND] with  $\mu(k_1, k_2) = \frac{\mathbb{1}_{x_1 \oplus k_1 \oplus a_1 = x_2 \oplus k_2 \oplus a_2}}{2^n}$ ,  
must show

$$x_1 \oplus k_1 \oplus a_1 = x_2 \oplus k_2 \oplus a_2 \implies x_1 \oplus k_1 = a_1 \implies x_2 \oplus k_2 = a_2$$

# One-time pad, revisited

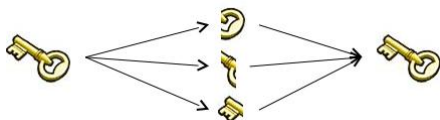
Uniformity implies secrecy

$$\frac{\forall a_1, a_2, \models \{\Phi\} c_1 \sim c_2 \{y_1 = a_1 \implies y_2 = a_2\}}{\models \{\Phi\} c_1 \sim c_2 \{y_1 = y_2\}} \text{ [UNIF-EQUAL]}$$

Derived rule: pointwise equality

$$\frac{\forall a, \models \{\Phi\} c_1 \sim c_2 \{y_1 = a \implies y_2 = a\}}{\models \{\Phi\} c_1 \sim c_2 \{y_1 = y_2\}} \text{ [UNIF-EQUAL]}$$

# Secret sharing



Value  $x$  encoded by  $t + 1$ -tuple of prob. values  $(x_0 \dots x_t)$  s.t.

- ▶  $x_0, \dots, x_t$  are independent and uniformly distributed
- ▶  $x = x_0 + \dots + x_t$

We can prove:

$$\forall v, w. \models \{\top\} \text{share}(a) \sim \text{share}(b) \{a_i = b_i\}$$

and:

$$\forall v, w. \models \{\top\} \text{share}(a) \sim \text{share}(b) \{a_i = v \Rightarrow b_i = w\}$$

and also:

$$\models \{\top\} \text{share}(a) \sim \text{share}(b) \{\wedge_{i \neq j} a_i = v_i \Rightarrow b_i = w_i\}$$

# van Neumann's trick

## Goal

Generate a fair coin flip, using coin flips with bias  $p \in (0, 1)$ .

## Procedure

1. Flip two coins with bias  $p$
2. Re-flip as long as they are equal
3. Return the first coin flip the first time they differ

# Proving uniformity

```
x ← 0;  
y ← 0;  
while x = y do  
  x ←$ Bp;  
  y ←$ Bp;
```

To prove:

$$\forall b_1, b_2. \models \{T\} c_1 \sim c_2 \{x_1 = b_1 \implies x_2 = b_2\}$$

Interesting case:  $b_1 \neq b_2$ , so we show:

$$\models \{T\} c_1 \sim c_2 \{x_1 \neq x_2\}$$

# Proof

By [WHILE] with loop invariant

$\Theta \triangleq x_2 = (\text{if } x_1 = y_1 \text{ then } y_2 \text{ else } \neg x_1)$ , suffices to show

$$\models \{\top\} x_1 \stackrel{\$}{\leftarrow} B_p; y_1 \stackrel{\$}{\leftarrow} B_p \sim x_2 \stackrel{\$}{\leftarrow} B_p; y_2 \stackrel{\$}{\leftarrow} B_p \{\Theta\}$$

By code motion, suffices to show

$$\models \{\top\} x_1 \stackrel{\$}{\leftarrow} B_p; y_1 \stackrel{\$}{\leftarrow} B_p \sim y_2 \stackrel{\$}{\leftarrow} B_p; x_2 \stackrel{\$}{\leftarrow} B_p \{\Theta\}$$

Applying [RAND] with identity couplings, we must show:

$$x_1 = y_2 \implies x_2 = y_1 \implies \Theta$$

# Extensions

- ▶ Approximate pRHL and applications to differential privacy
- ▶ Expectation pRHL and applications to algorithmic stability

Higher-order languages