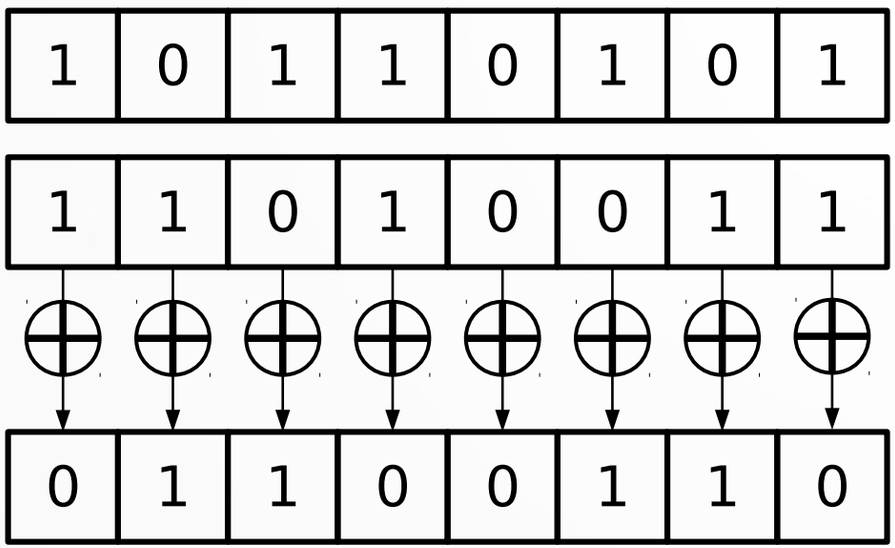
Relational Reasoning for Markov Chains in a Probabilistic Guarded Lambda Calculus

Alejandro Aguirre, Gilles Barthe, Lars Birkedal, Aleš Bizjak, Marco Gaboardi and Deepak Garg

Imdea Software, Aarhus University, University at Buffalo SUNY, MPI-SWS

one-time pad



message

key (uniformly sampled)

ciphertext

$$\mathrm{otp} : \mathsf{Msg} \to D(\mathsf{CT})$$

Key property: perfect secrecy

$$\forall m_1 m_2.\ \mathrm{otp}\ m_1 \approx \mathrm{otp}\ m_2$$

relational properties

- two executions of the same or different programs

$$\forall m_1 m_2.\ \mathrm{otp}\ m_1 \approx \mathrm{otp}\ m_2$$

two executions

- one could compute both and compare them

- is there a better way to reason?

relational logic

relational logic

A Relational Logic for Higher-Order Programs

ALEJANDRO AGUIRRE, IMDEA Software Institute, Spain
GILLES BARTHE, IMDEA Software Institute, Spain
MARCO GABOARDI, University at Buffalo, SUNY, USA
DEEPAK GARG, MPI-SWS, Germany
PIERRE-YVES STRUB, École Polytechnique, France

ICFP'17

relational logic

A Relational Logic for Higher-Order Programs

ALEJANDRO AGUIRRE, IMDEA Software Institute, Spain
GILLES BARTHE, IMDEA Software Institute, Spain
MARCO GABOARDI, University at Buffalo, SUNY, USA
DEEPAK GARG, MPI-SWS, Germany
PIERRE-YVES STRUB, École Polytechnique, France

ICFP'17

$$\Gamma \mid \Psi \vdash t_1 : \sigma_1 \sim t_2 : \sigma_2 \mid \phi$$

relational logic

A Relational Logic for Higher-Order Programs

ALEJANDRO AGUIRRE, IMDEA Software Institute, Spain
GILLES BARTHE, IMDEA Software Institute, Spain
MARCO GABOARDI, University at Buffalo, SUNY, USA
DEEPAK GARG, MPI-SWS, Germany
PIERRE-YVES STRUB, École Polytechnique, France

ICFP'17

$$\Gamma \mid \Psi \vdash t_1 : \sigma_1 \sim t_2 : \sigma_2 \mid \phi$$

Two terms

relational logic

A Relational Logic for Higher-Order Programs

ALEJANDRO AGUIRRE, IMDEA Software Institute, Spain
GILLES BARTHE, IMDEA Software Institute, Spain
MARCO GABOARDI, University at Buffalo, SUNY, USA
DEEPAK GARG, MPI-SWS, Germany
PIERRE-YVES STRUB, École Polytechnique, France

ICFP'17

$$\Gamma \mid \Psi \vdash t_1 : \sigma_1 \sim t_2 : \sigma_2 \mid \phi$$

Two terms

related by a property

relational logic

A Relational Logic for Higher-Order Programs

ALEJANDRO AGUIRRE, IMDEA Software Institute, Spain
GILLES BARTHE, IMDEA Software Institute, Spain
MARCO GABOARDI, University at Buffalo, SUNY, USA
DEEPAK GARG, MPI-SWS, Germany
PIERRE-YVES STRUB, École Polytechnique, France

ICFP'17

$$\Gamma \mid \Psi \vdash t_1 : \sigma_1 \sim t_2 : \sigma_2 \mid \phi$$

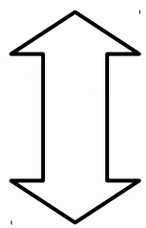
Two terms

related by a property

$$\vdash \mathrm{otp} : \mathsf{Msg} \to D(\mathsf{CT}) \sim \mathrm{otp} : \mathsf{Msg} \to D(\mathsf{CT}) \mid \forall m_1 m_2.\ \mathbf{r_1}\ m_1 \approx \mathbf{r_2}\ m_2$$

relational logic

$$\forall m_1 m_2.\ \mathrm{otp}\ m_1 \ \approx\ \mathrm{otp}\ m_2$$



$$\vdash \mathrm{otp} : \mathsf{Msg} \to D(\mathsf{CT}) \sim \mathrm{otp} : \mathsf{Msg} \to D(\mathsf{CT}) \mid \forall m_1 m_2.\ \mathbf{r_1}\ m_1 \approx \mathbf{r_2}\ m_2$$

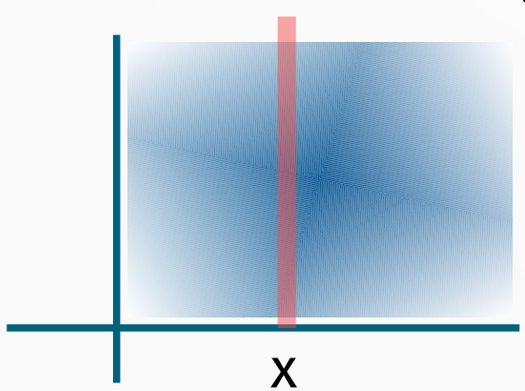
How to express this exactly?

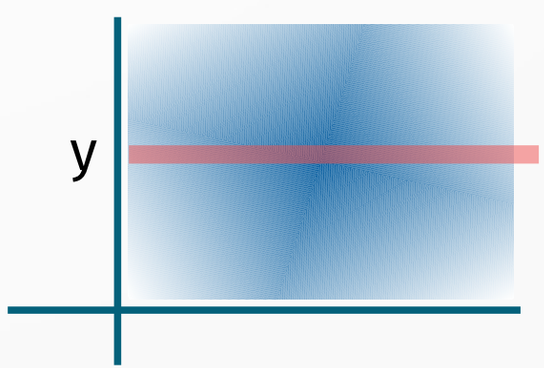
reminder: probability

(discrete) distribution $\mu : C \to [0,1]$ s.t. $\sum_{x \in C} \mu(x) = 1$

marginals for $\mu \in D(C_1 \times C_2)$

$$D(\pi_1)(\mu)(x) = \sum_{y \in C_2} \mu(x,y) \qquad D(\pi_2)(\mu)(y) = \sum_{x \in C_1} \mu(x,y)$$





couplings 101

- let $\mu_1 \in D(C_1),\ \mu_2 \in D(C_2)$

$\mu \in D(C_1 \times C_2)$
is a coupling
$\iff$ $\pi_1(\mu) = \mu_1 \quad \& \quad \pi_2(\mu) = \mu_2$

- let $R \subseteq C_1 \times C_2$ Useful cases: $R \in \{=, \leq, \geq, \dots\}$

$\mu \in D(C_1 \times C_2)$
is a R-coupling
$\iff$ $\pi_1(\mu) = \mu_1 \quad \& \quad \pi_2(\mu) = \mu_2$
$\& \quad \Pr_{(x_1,x_2)\sim\mu}[x_1\ R\ x_2] = 1$

(Denoted $\diamond_{\mu_1,\mu_2}(R)$)

example: coin flip

- Some ways of coupling two fair coins:

product $\diamond(\top)$

	H	T
H	1/4	1/4
T	1/4	1/4

equality $\diamond(=)$

	H	T
H	1/2	0
T	0	1/2

inequality $\diamond(\neq)$

	H	T
H	0	1/2
T	1/2	0

general

	H	T
H	p	½ - p
T	½ - p	p

fundamental lemma

If $\mu \in D(C_1 \times C_2)$ *is an R-coupling of* $\mu_1 \in D(C_1)$ *and* $\mu_2 \in D(C_2)$,

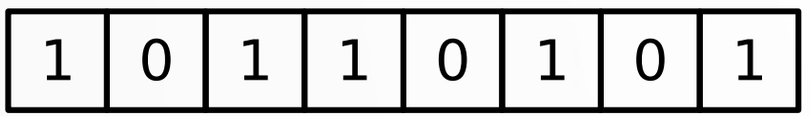
$$\Pr_{x_2 \sim \mu_2}[x_2 \in R(C_1)] = 1$$

Idea: To prove a relational property about distributions, we "sync randomness" to build the appropriate coupling

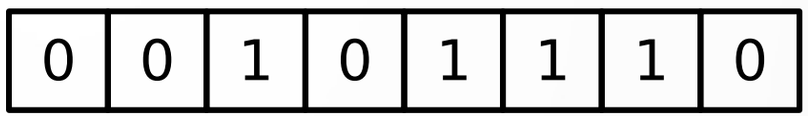
In particular:

- $\diamond_{\mu_1,\mu_2}(=) \Rightarrow \mu_1 = \mu_2$
- $\diamond_{\mu_1,\mu_2}(\leq) \Rightarrow \mu_1 \leq \mu_2$ (stochastic dominance)

one-time pad revisited

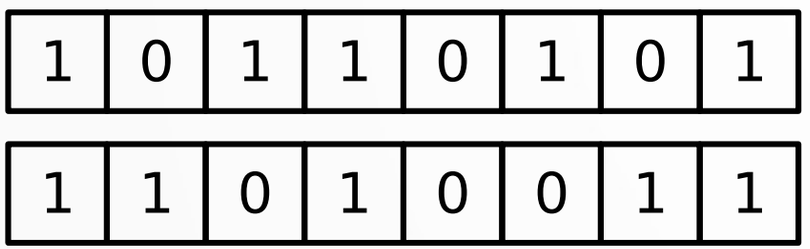


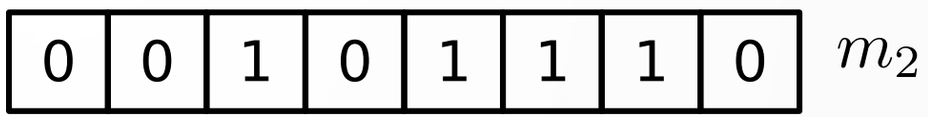
m_1



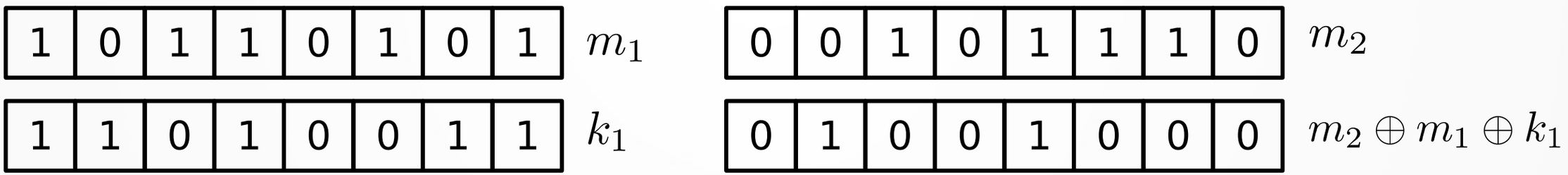
m_2

one-time pad revisited

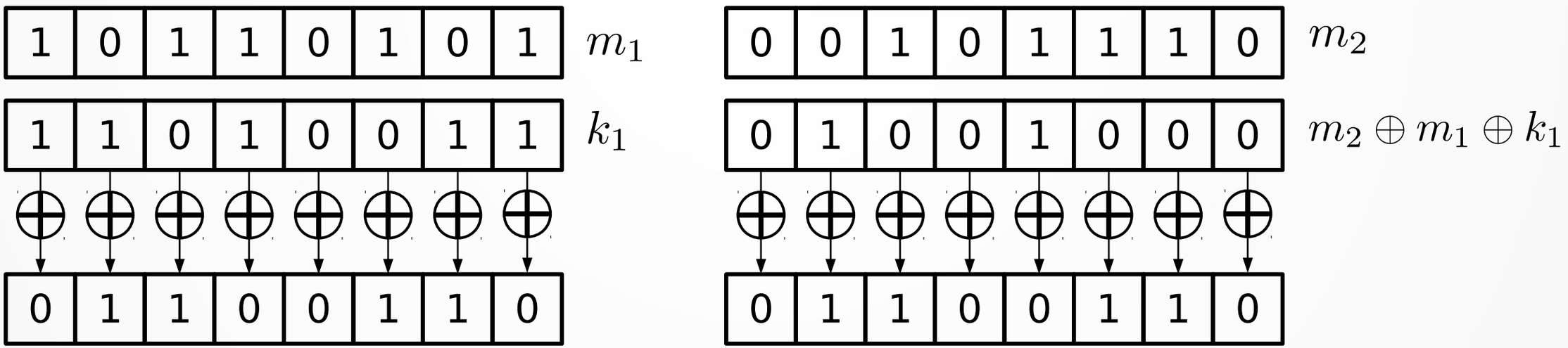




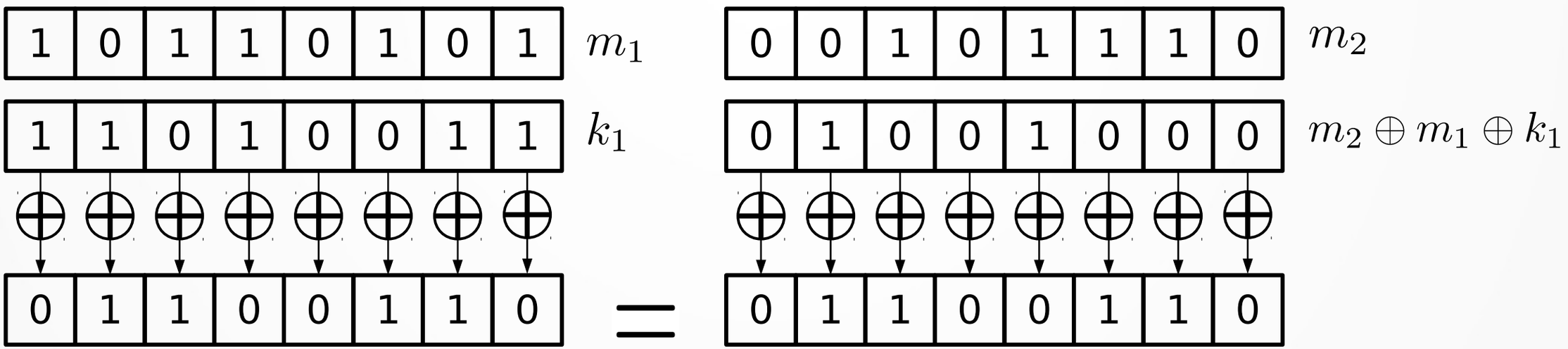
one-time pad revisited



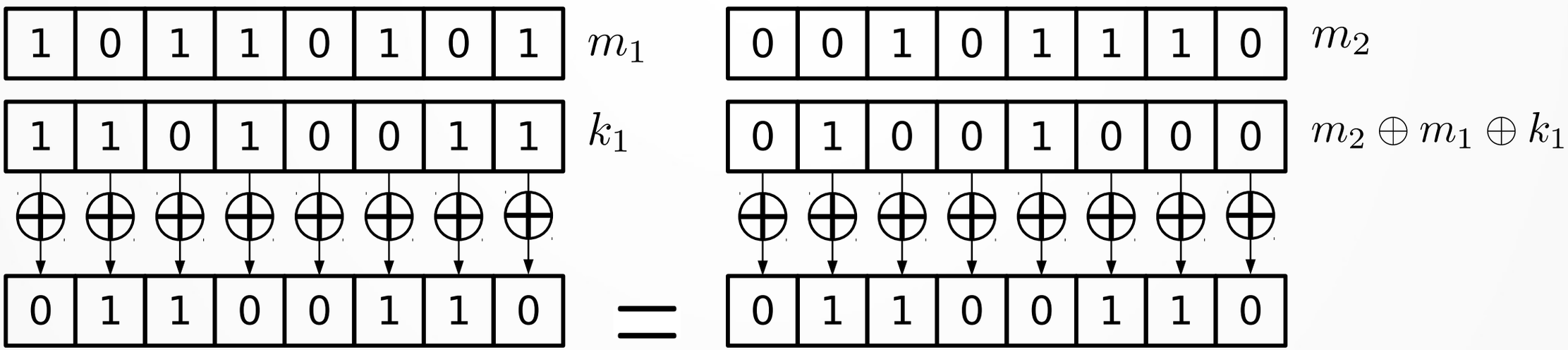
one-time pad revisited



one-time pad revisited



one-time pad revisited

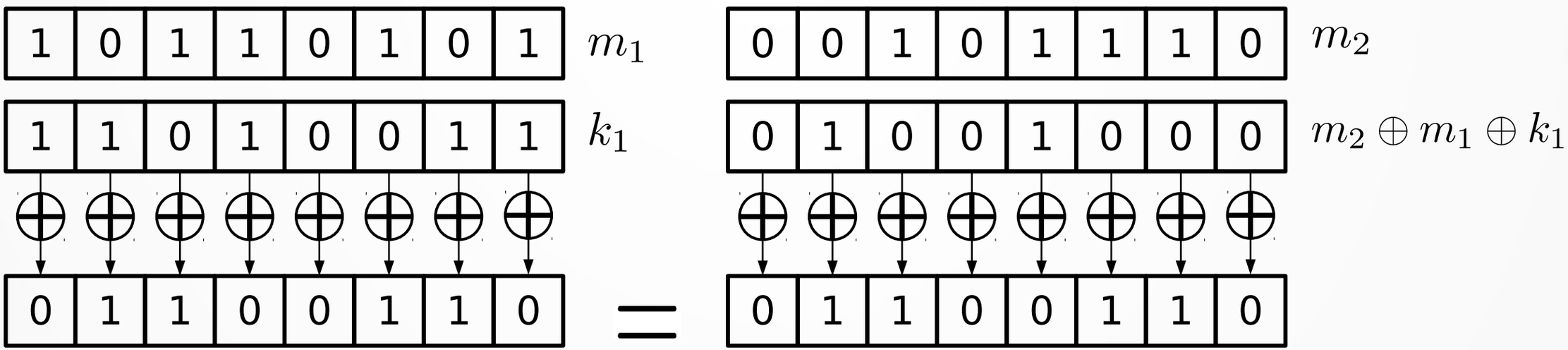


Therefore,

$$\forall m_1 m_2.\ \mathrm{otp}\ m_1 \ = \ \mathrm{otp}\ m_2$$

$$[\diamond_{\mathrm{otp}\ m_1, \mathrm{otp}\ m_2}(=)]$$

one-time pad revisited



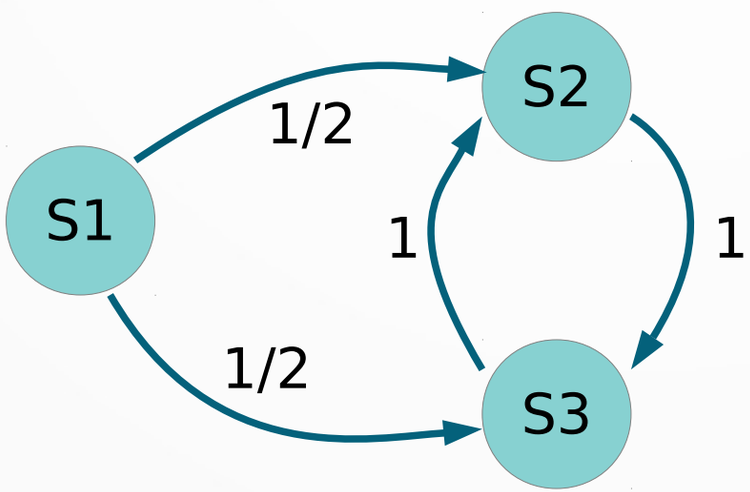
Therefore, $\forall m_1 m_2.\ \mathrm{otp}\ m_1 \ = \ \mathrm{otp}\ m_2$

$$[\diamond_{\mathrm{otp}\ m_1, \mathrm{otp}\ m_2}(=)]$$

One can also show, $\forall m_1.\ \mathrm{otp}\ m_1 \ = \ \mathsf{Unif}(\mathsf{Msg})$

(discrete time) markov chains

- discrete set of states S
- probabilistic transition function $S \to D(S)$

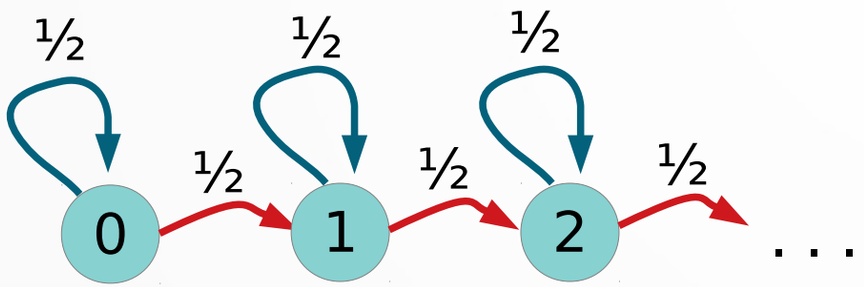


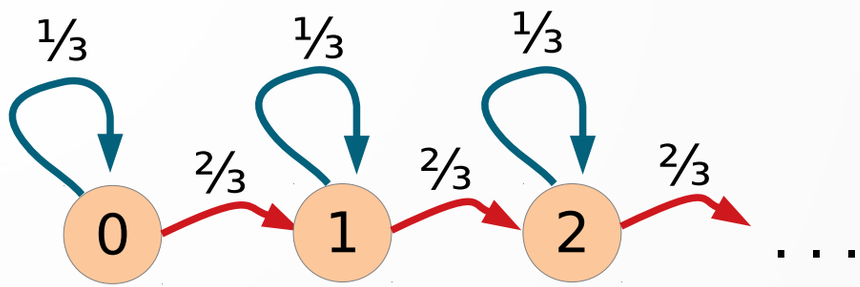
relational properties

relate two Markov chains or two runs of the same one

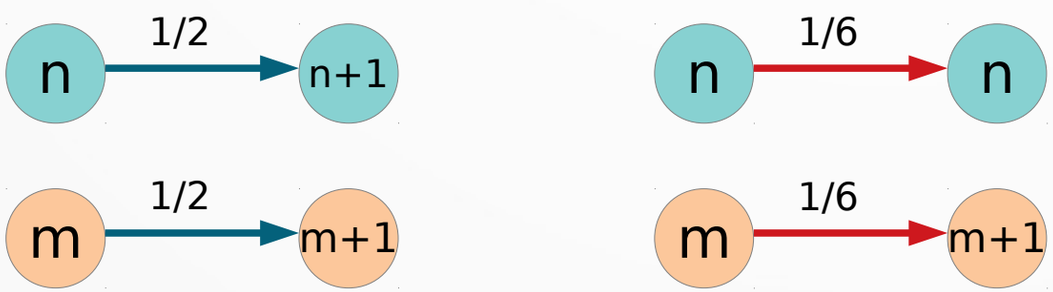
which one is faster?

example: stochastic dominance

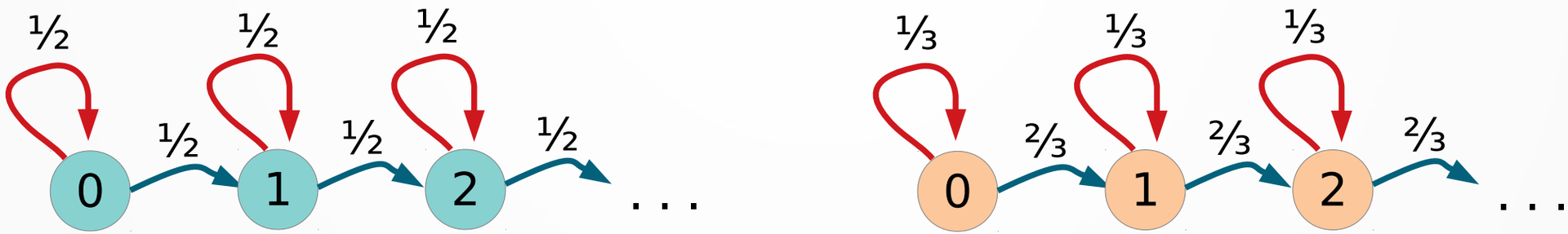


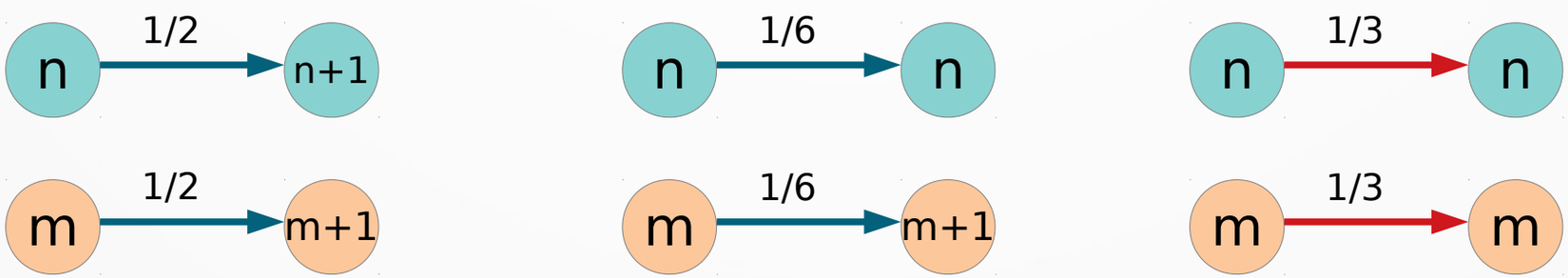


example: stochastic dominance

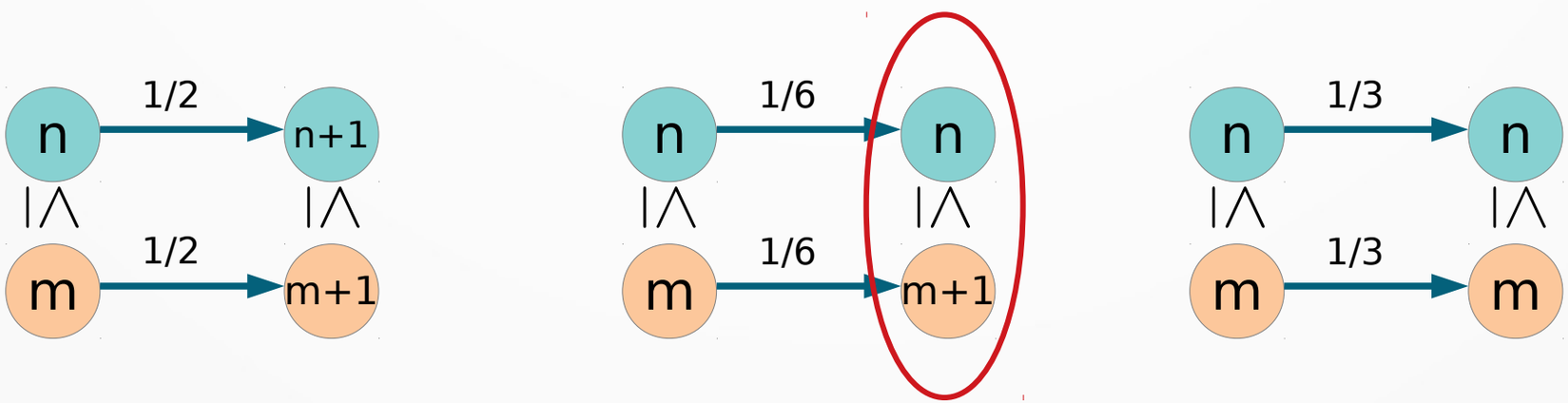


example: stochastic dominance





example: stochastic dominance



so far

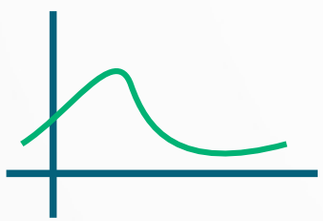
- Probabilities & Markov chains are a useful modeling tool
- Some interesting properties are relational
- These can be studied with couplings

representing Markov Chains

- infinite sequence (*stream*) of distributions

representing Markov Chains

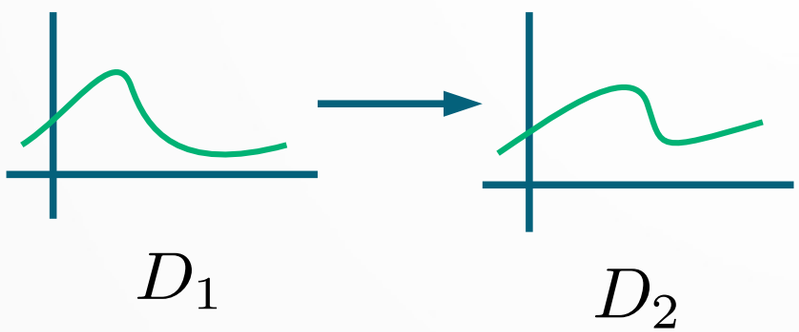
- infinite sequence (*stream*) of distributions



D_1

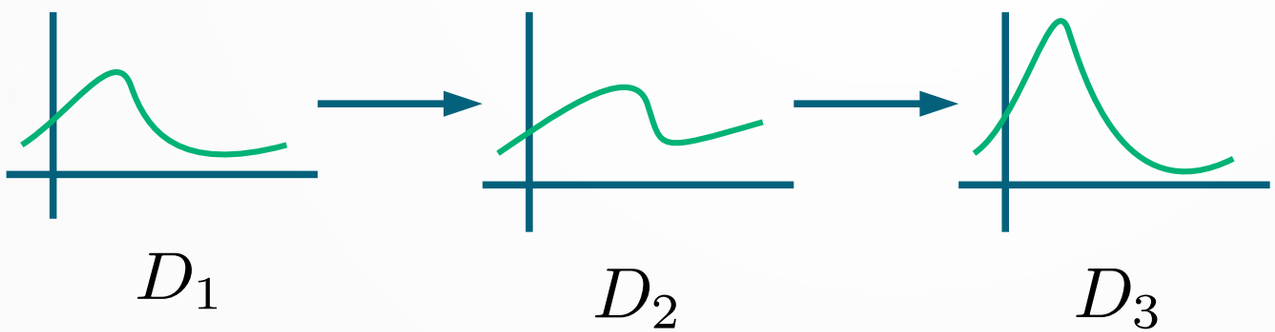
representing Markov Chains

- infinite sequence (*stream*) of distributions



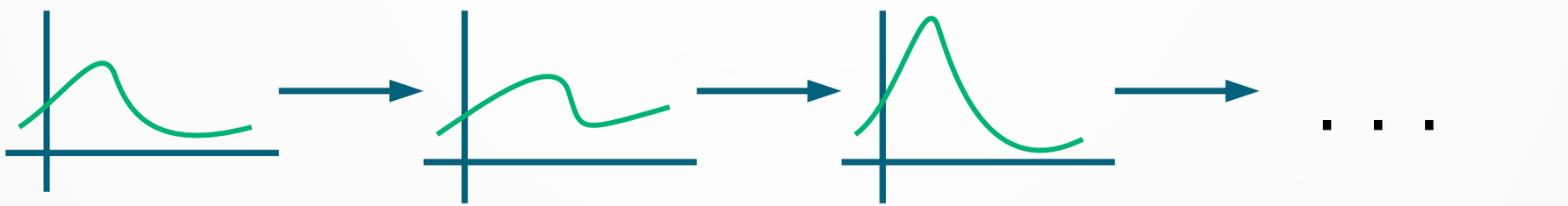
representing Markov Chains

- infinite sequence (*stream*) of distributions



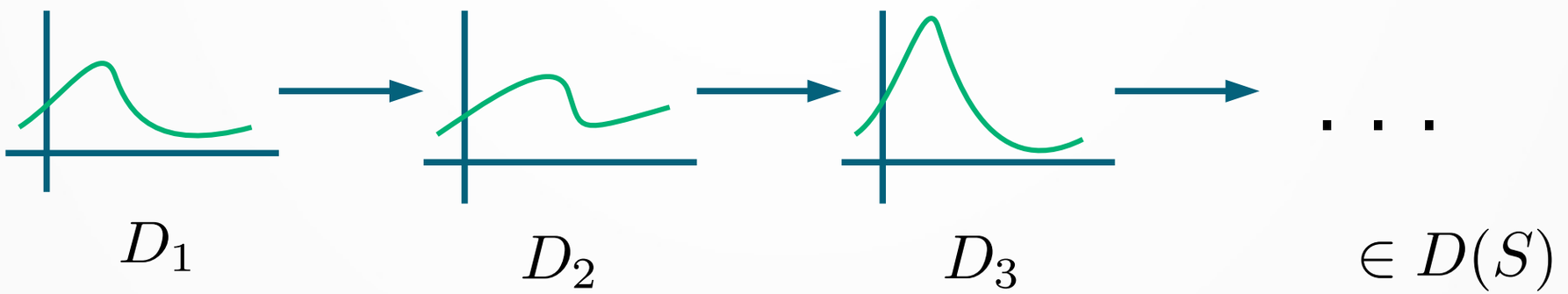
representing Markov Chains

- infinite sequence (*stream*) of distributions



representing Markov Chains

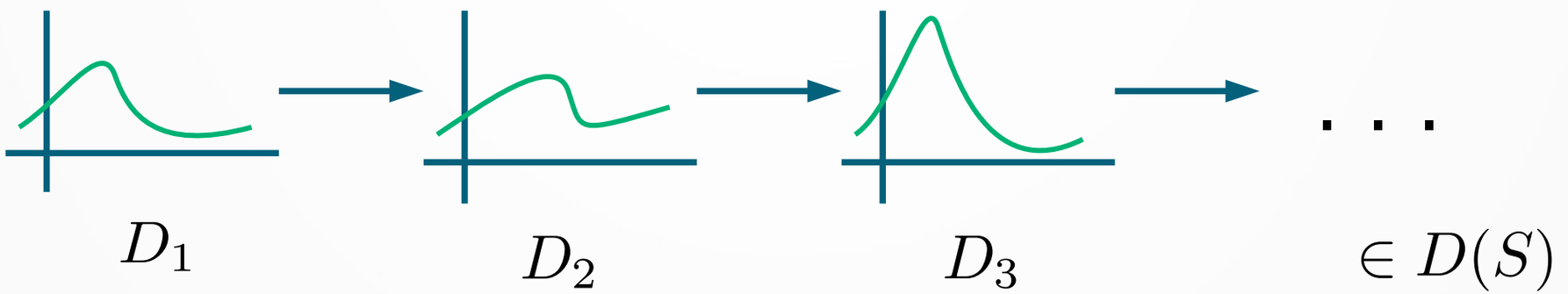
- infinite sequence (*stream*) of distributions



Problem : 1) Not expressive enough!

representing Markov Chains

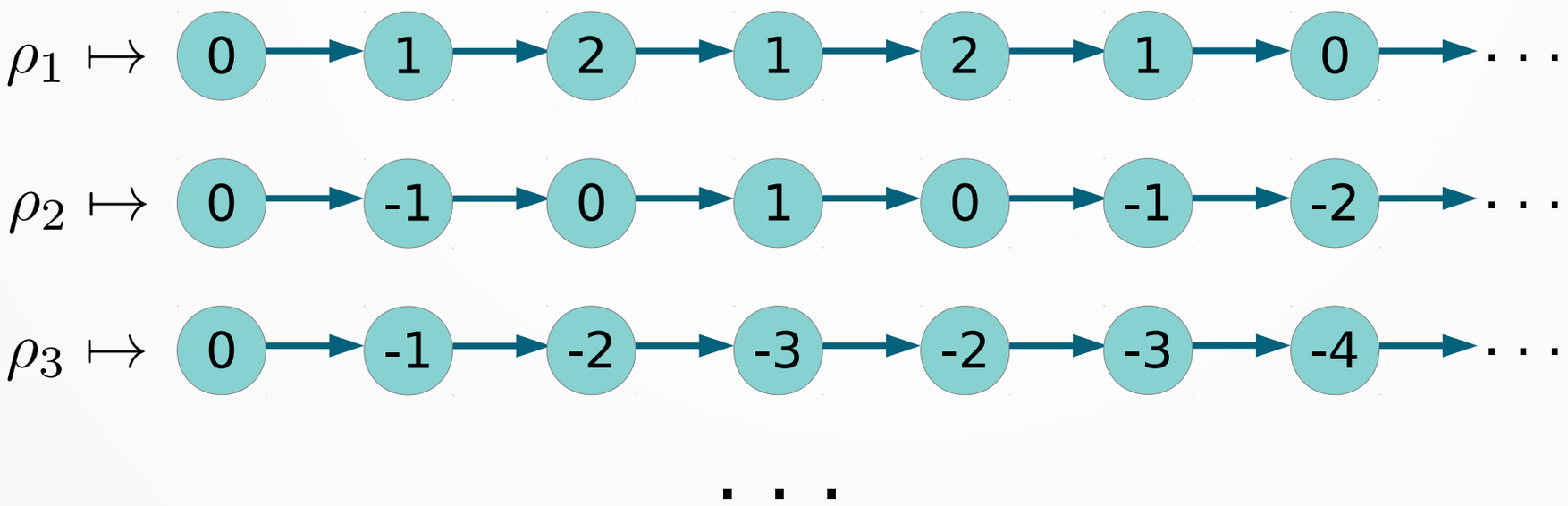
- infinite sequence (*stream*) of distributions



Problem : 1) Not expressive enough!
2) Probabilistic dependence

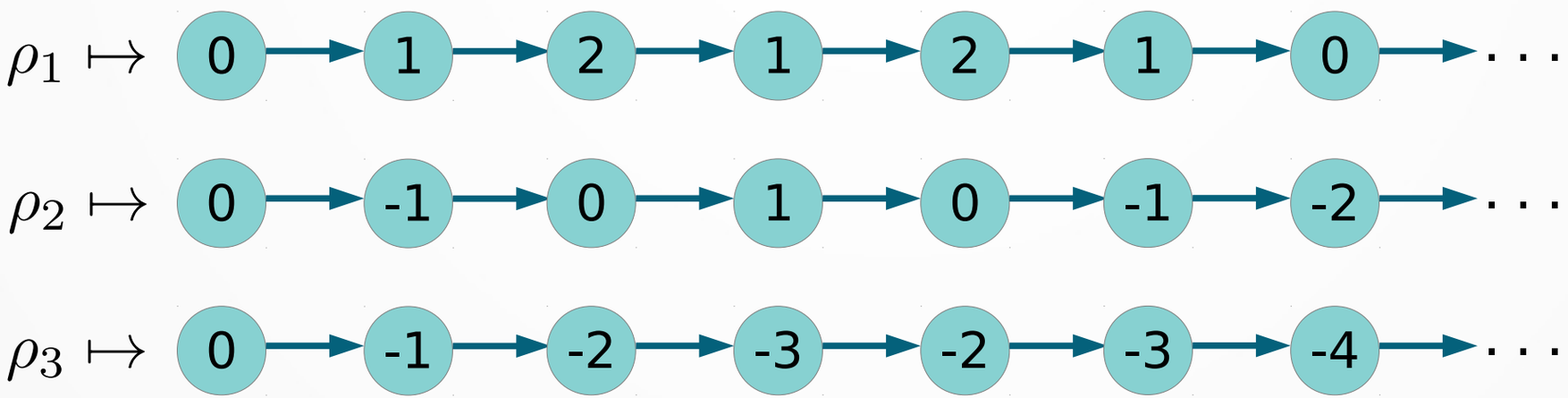
representing Markov Chains

- as distributions over streams



representing Markov Chains

- as distributions over streams

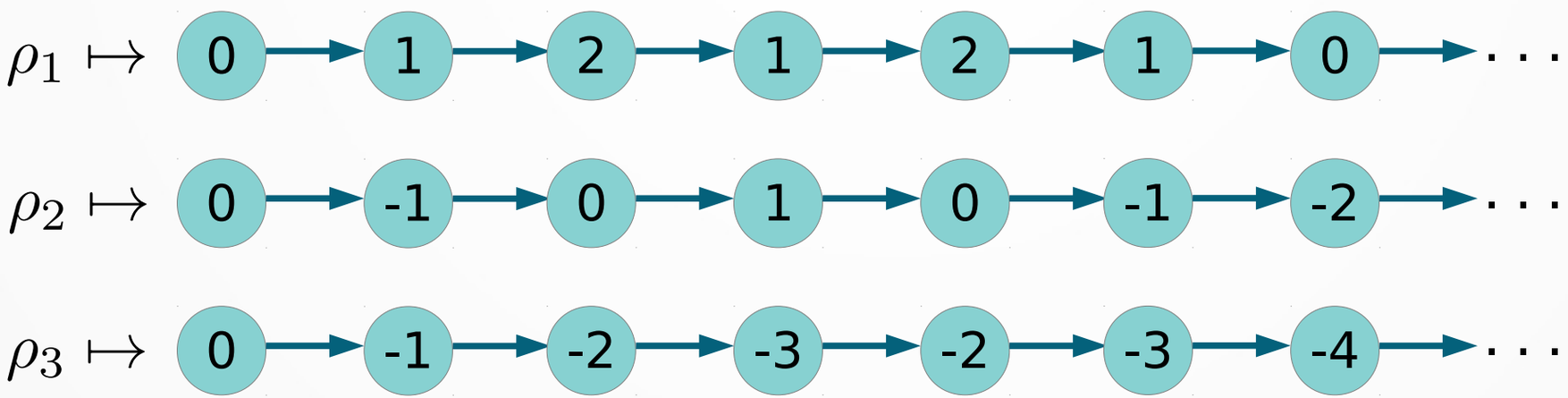


. . .

let $m\ s_0 \ = \ s_0 :: m\ (f\ s_0)$

representing Markov Chains

- as distributions over streams



. . .

$$\text{let } m\ s_0 \ = \ s_0 :: m\ (f\ s_0)$$

Problem : Not discrete!

the problem with stream definitions

- A stream is productive if it outputs every finite prefix in finite time

  For instance, $\text{let } s = 1 : s$

- Example of non-productive stream: $\text{let } s = 1 : \text{tl}(s)$

guarded lambda calculus

- Terms: $t, u ::= x \mid \lambda x.t \mid t\ u \mid \mathrm{fix}\ x.t \mid t :: u \mid \mathrm{tl}(u) \mid \ldots$
- Types: $A, B ::= b \mid N \mid A \to B \mid A \times B \mid \mathrm{Str}_A \mid \triangleright A$

$$\mathrm{Str}_A \cong A \times \triangleright \mathrm{Str}_A$$

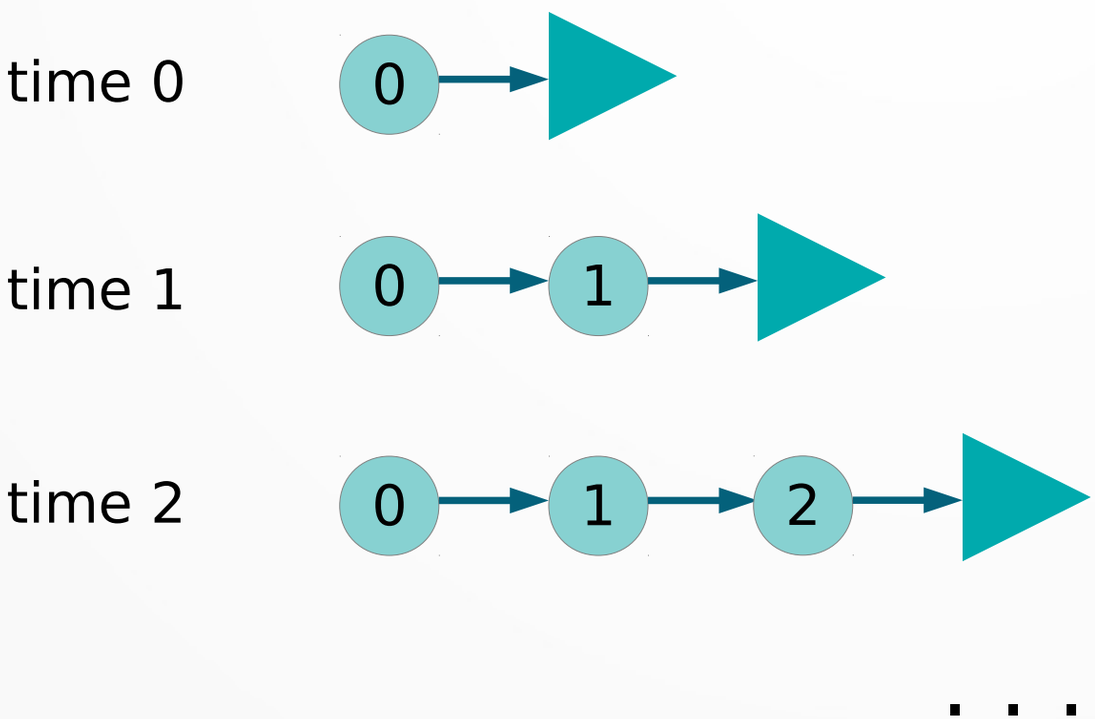
An element now

A stream “later”

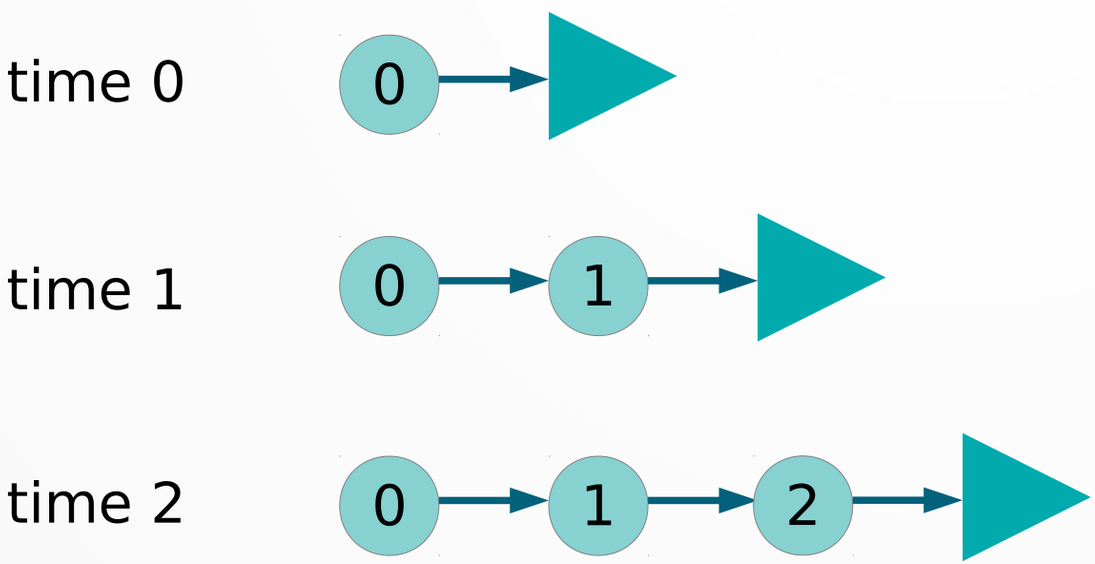
$$\frac{f : \triangleright A \to \triangleright \mathsf{Str}_B \vdash t : A \to \mathsf{Str}_B}{\vdash \mathrm{fix}\ f.t : A \to \mathsf{Str}_B}\ \mathsf{FIX_f}$$

[1] Clouston, R., Bizjak, A., Grathwohl, H.B., Birkedal, L.: The guarded lambda-calculus: Programming and reasoning with guarded recursion for coinductive types. (LMCS ‘16)

streams in glc



streams in glc



All prefixes are discrete!

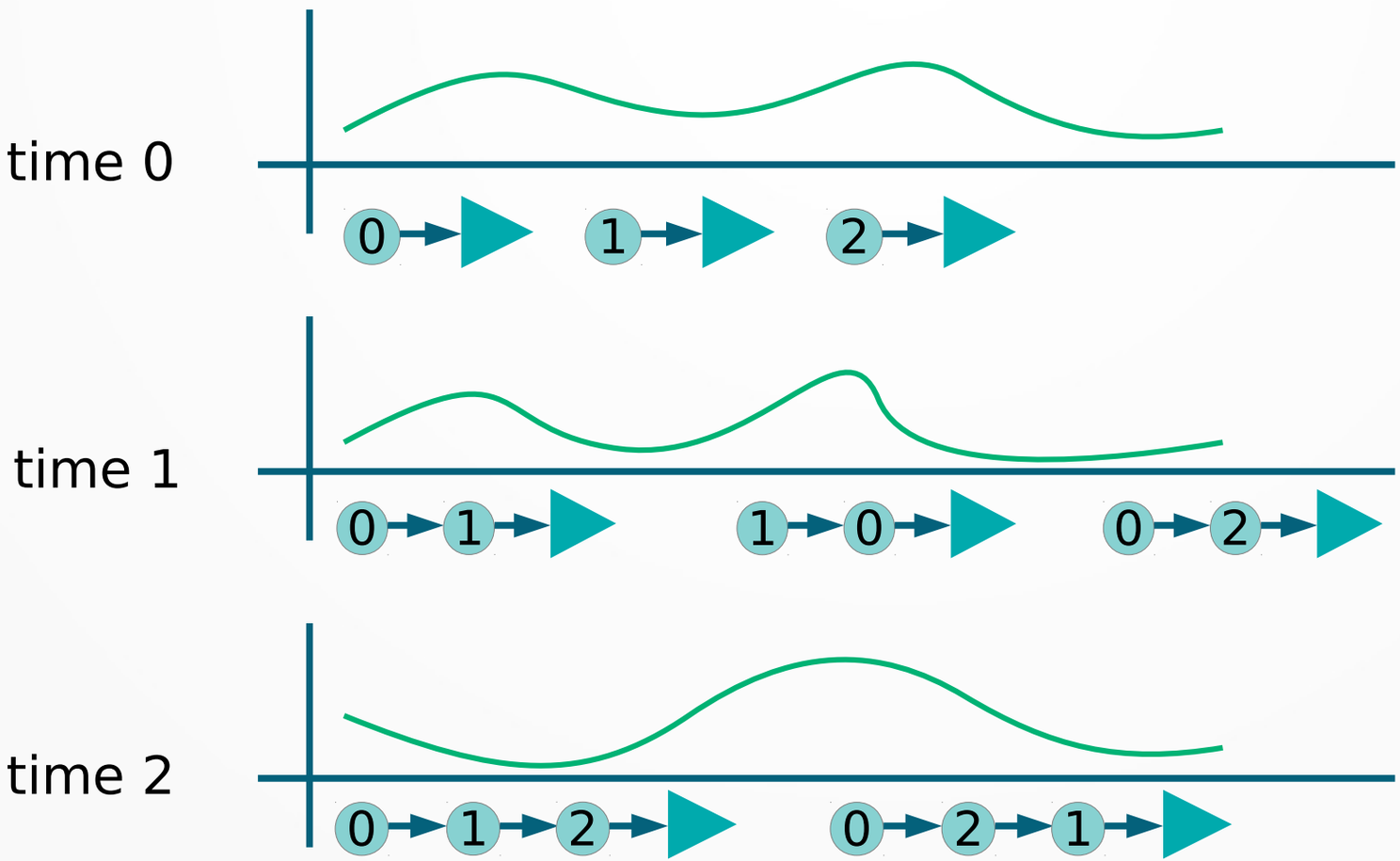
probabilistic glc

- Terms: $t, u ::= \cdots \mid \mathsf{munit}(t) \mid \mathsf{mlet}\ x = t\ \mathrm{in}\ u \mid \mathcal{B}(\rho) \mid \ldots$
- Types: $A, B ::= b \mid N \mid A \to B \mid A \times B \mid \mathrm{Str}_A \mid \triangleright A \mid D(C)$

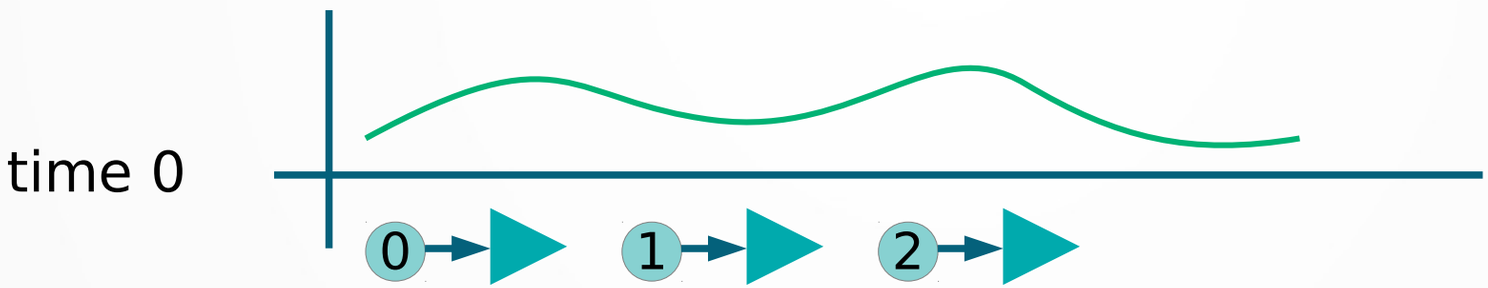
$$\frac{\vdash s : A \quad \vdash f : A \to D(A)}{\vdash \mathrm{markov}\ f\ s : D(\mathrm{Str}_A)}\ \mathsf{Markov}$$

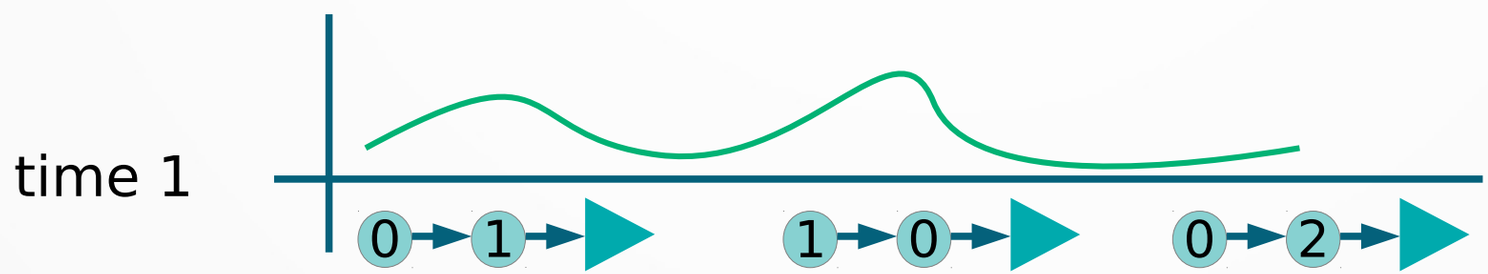
$$\begin{aligned}\mathrm{markov}\ f\ s \equiv \mathrm{fix}\ f.\lambda h.\lambda s.&\\ &\mathrm{mlet}\ xs = (f\ h\ (h\ s))\ \mathrm{in}\\ &\mathrm{munit}(s :: xs)\end{aligned}$$

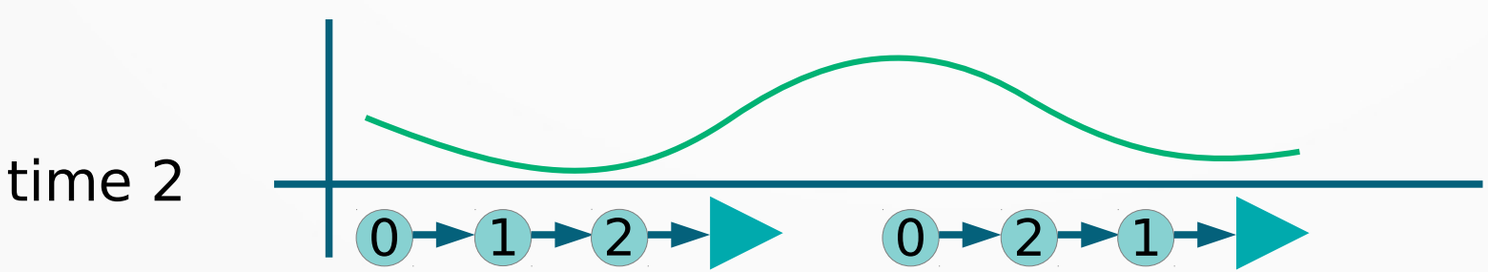
stream distributions in glc



stream distributions in glc







. . .

Discrete Probabilities!

(two) relational logics for glc

$$\Delta \mid \Sigma \mid \Gamma \mid \Psi \vdash t_1 : A_1 \sim t_2 : A_2 \mid \phi$$

Two terms

related by a property/coupling

$$\Delta \mid \Sigma \mid \Gamma \mid \Psi \vdash t_1 : \mathsf{D}(C_1) \sim t_2 : \mathsf{D}(C_2) \mid \diamond_{[x_1 \leftarrow \mathbf{r}_1, x_2 \leftarrow \mathbf{r}_2]} \phi$$

proving properties of Markov chains

$$\vdash t_1 \sim t_2 \mid \psi$$

$$\vdash f_1 \ \sim \ f_2 \ \mid \ \forall x_1 x_2.\ \psi(x_1, x_2) \ \Rightarrow \ \diamond\psi(\mathbf{r_1}\ x_1, \mathbf{r_2}\ x_2)$$

$$\frac{\vdash \forall x_1 x_2 s_1 s_2.\ \psi(x_1, x_2) \ \Rightarrow \ \triangleright\phi(s_1, s_2) \ \Rightarrow \ \phi(x_1 :: s_1, x_2 :: s_2)}{\vdash \text{markov}\ f_1\ t_1 : D(\text{Str}_A) \ \sim \ \text{markov}\ f_2\ t_2 : D(\text{Str}_A) \ \mid \ \diamond\phi}$$

proving properties of Markov chains

$$\text{1)}\quad \vdash t_1 \sim t_2 \mid \psi$$

$$\vdash f_1 \sim f_2 \mid \forall x_1 x_2.\ \psi(x_1, x_2) \Rightarrow \diamond\psi(\mathbf{r_1}\ x_1, \mathbf{r_2}\ x_2)$$

$$\vdash \forall x_1 x_2 s_1 s_2.\ \psi(x_1, x_2) \Rightarrow \triangleright\phi(s_1, s_2) \Rightarrow \phi(x_1 :: s_1, x_2 :: s_2)$$

$$\vdash \text{markov}\ f_1\ t_1 : D(\text{Str}_A) \sim \text{markov}\ f_2\ t_2 : D(\text{Str}_A) \mid \diamond\phi$$

1) Prove local property about pairs of states

proving properties of Markov chains

1) $\vdash t_1 \sim t_2 \mid \psi$

2) $\vdash f_1 \sim f_2 \mid \forall x_1 x_2.\ \psi(x_1, x_2) \Rightarrow \diamond\psi(\mathbf{r_1}\ x_1, \mathbf{r_2}\ x_2)$

$$\frac{\vdash \forall x_1 x_2 s_1 s_2.\ \psi(x_1, x_2) \Rightarrow \triangleright\phi(s_1, s_2) \Rightarrow \phi(x_1 :: s_1, x_2 :: s_2)}{\vdash \mathrm{markov}\ f_1\ t_1 : D(\mathrm{Str}_A) \sim \mathrm{markov}\ f_2\ t_2 : D(\mathrm{Str}_A) \mid \diamond\phi}$$

1) Prove local property about pairs of states
2) Show that the transition functions preserves it

proving properties of Markov chains

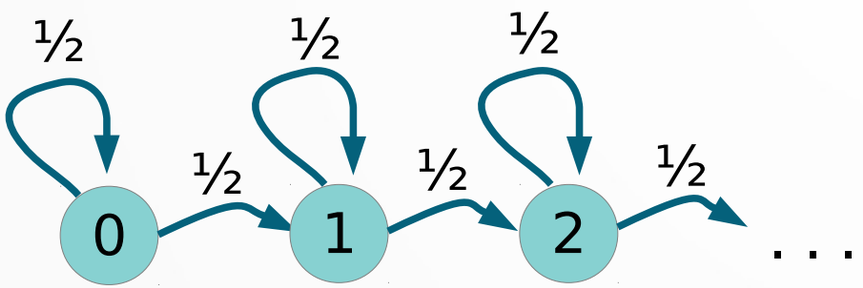
1) $\vdash t_1 \sim t_2 \mid \psi$

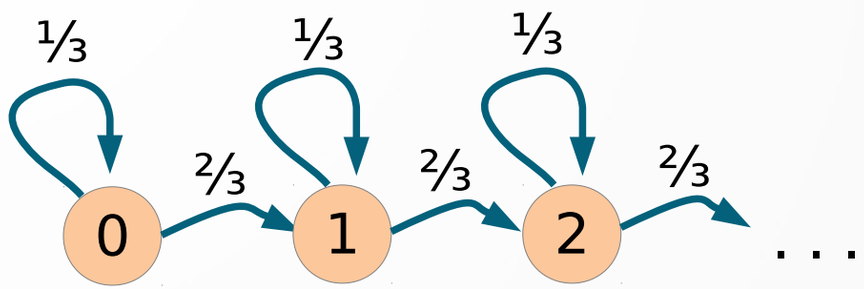
2) $\vdash f_1 \sim f_2 \mid \forall x_1 x_2.\ \psi(x_1, x_2) \Rightarrow \diamond\psi(\mathbf{r_1}\ x_1, \mathbf{r_2}\ x_2)$

3) $$\frac{\vdash \forall x_1 x_2 s_1 s_2.\ \psi(x_1, x_2) \Rightarrow \triangleright\phi(s_1, s_2) \Rightarrow \phi(x_1 :: s_1, x_2 :: s_2)}{\vdash \mathrm{markov}\ f_1\ t_1 : D(\mathrm{Str}_A) \sim \mathrm{markov}\ f_2\ t_2 : D(\mathrm{Str}_A) \mid \diamond\phi}$$

1) Prove local property about pairs of states
2) Show that the transition functions preserves it
3) Lift the property to a global property about pairs of streams

example: stochastic dominance





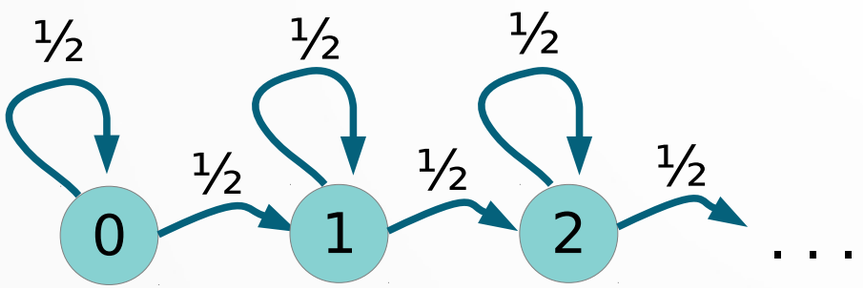
$h_1 ::= \lambda x.\text{mlet } b = \mathcal{B}(1/2) \text{ in } \text{munit}(x + b)$

$h_2 ::= \lambda x.\text{mlet } b = \mathcal{B}(2/3) \text{ in } \text{munit}(x + b)$

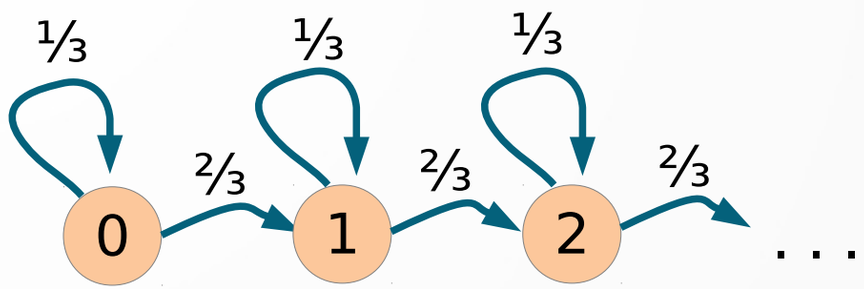
$$\vdash \text{markov } 0\ h_1 : D(\mathsf{Str}_{\mathbb{N}}) \sim \text{markov } 0\ h_2 : D(\mathsf{Str}_{\mathbb{N}}) \mid \diamond(\mathbf{r_1} \leq \mathbf{r_2})$$

Idea: $\diamond_{\mathcal{B}(1/2),\mathcal{B}(2/3)}(\leq)$

example: stochastic dominance

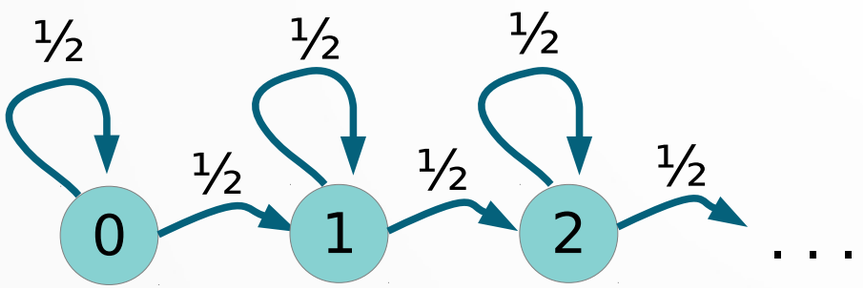


$h_1 ::= \lambda x.\text{mlet } b = \mathcal{B}(1/2) \text{ in } \text{munit}(x + b)$

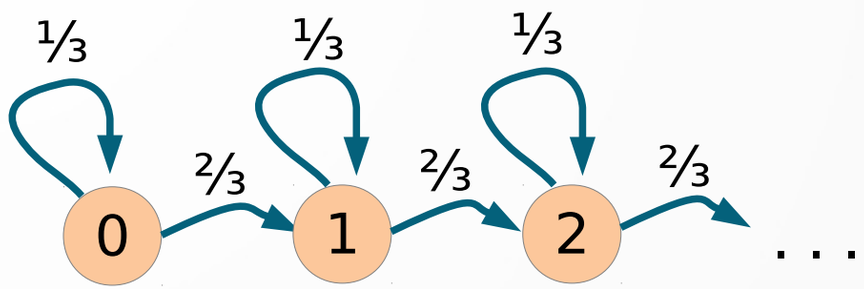


$h_2 ::= \lambda x.\text{mlet } b = \mathcal{B}(2/3) \text{ in } \text{munit}(x + b)$

example: stochastic dominance



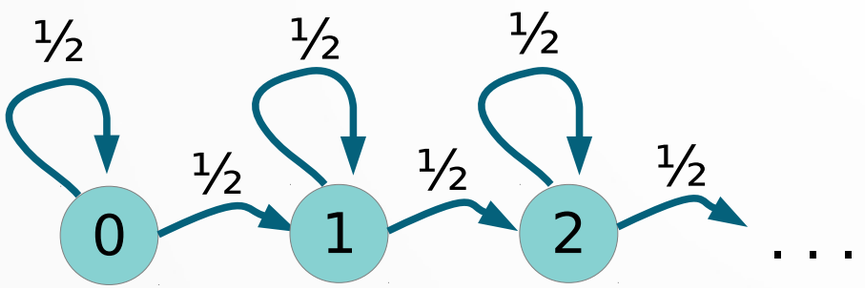
$h_1 ::= \lambda x.\text{mlet } b = \mathcal{B}(1/2) \text{ in } \text{munit}(x + b)$



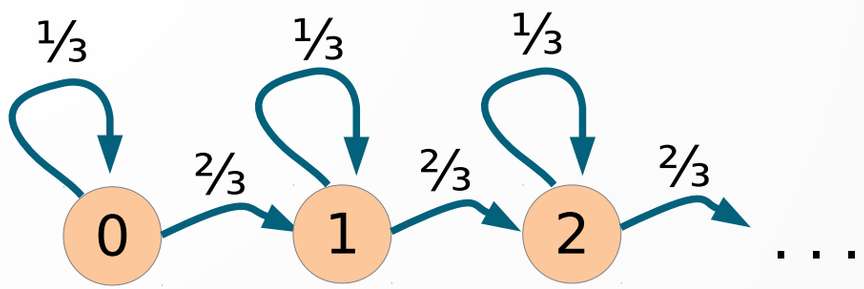
$h_2 ::= \lambda x.\text{mlet } b = \mathcal{B}(2/3) \text{ in } \text{munit}(x + b)$

1) $0 \leq 0$

example: stochastic dominance



$h_1 ::= \lambda x.\text{mlet } b = \mathcal{B}(1/2) \text{ in } \text{munit}(x + b)$

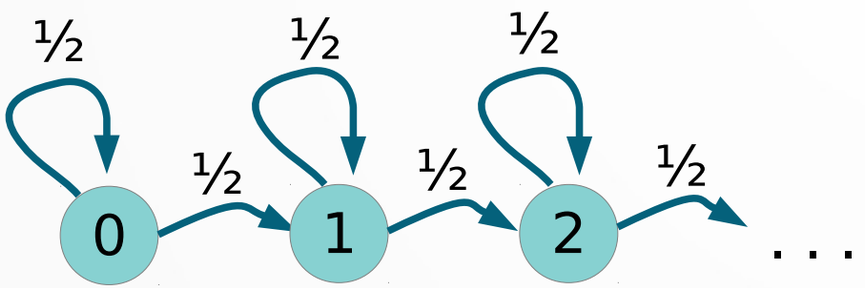


$h_2 ::= \lambda x.\text{mlet } b = \mathcal{B}(2/3) \text{ in } \text{munit}(x + b)$

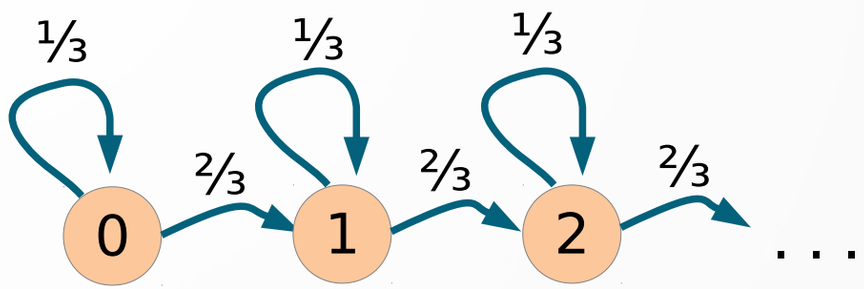
1) $0 \leq 0$

2) $x_1 \leq x_2 \Rightarrow \diamond(x_1 + \mathcal{B}(1/2) \leq x_2 + \mathcal{B}(2/3))$

example: stochastic dominance



$h_1 ::= \lambda x.\text{mlet } b = \mathcal{B}(1/2) \text{ in } \text{munit}(x + b)$



$h_2 ::= \lambda x.\text{mlet } b = \mathcal{B}(2/3) \text{ in } \text{munit}(x + b)$

1) $0 \leq 0$

2) $x_1 \leq x_2 \Rightarrow \diamond(x_1 + \mathcal{B}(1/2) \leq x_2 + \mathcal{B}(2/3))$

3) $\forall i. \diamond (\text{markov } h_1\ 0 \leq \text{markov } h_2\ 0)$

in the paper

- denotational semantics
- rest of the proof rules
- relational logic for deterministic streams
- more examples
  - Random walks (lazy vs non-lazy, 3D vs 4D)
  - Approximation series

conclusion

- a logic to reason about probabilistic infinite data structures
    - markov chains are represented as distribution over streams
    - properties are proven via couplings
- future work:
    - extend the language
    - continuous distributions

Thanks!

Relational Reasoning for Markov Chains in a Probabilistic Guarded Lambda Calculus

Alejandro Aguirre, Gilles Barthe, Lars Birkedal, Aleš Bizjak, Marco Gaboardi and Deepak Garg

Imdea Software, Aarhus University, University at Buffalo SUNY, MPI-SWS