



Laboratoire  
Méthodes  
Formelles

université  
PARIS-SACLAY



école  
normale  
supérieure  
paris-saclay

# On the Probabilistic and Statistical Verification of Infinite Markov Chains

Patricia Bouyer

LMF, Université Paris-Saclay,  
CNRS, ENS Paris-Saclay  
France

Joint work with Benoît Barbot (LACL) and Serge Haddad (LMF)

Work partly supported by ANR projects MAVeriQ and BisoUS

# General purpose

Design algorithms to estimate probabilities in some **infinite-state** Markov chains, **with guarantees**

# General purpose

Design algorithms to estimate probabilities in some **infinite-state** Markov chains, **with guarantees**

## Our contributions

- ▶ Review two existing approaches (approximation algorithm and estimation algorithm) and specify the required hypothesis for correctness
- ▶ Propose an approach based on **importance sampling** and **abstraction** to partly relax the hypothesis
- ▶ Analyze empirically the approaches

# Discrete-time Markov chains

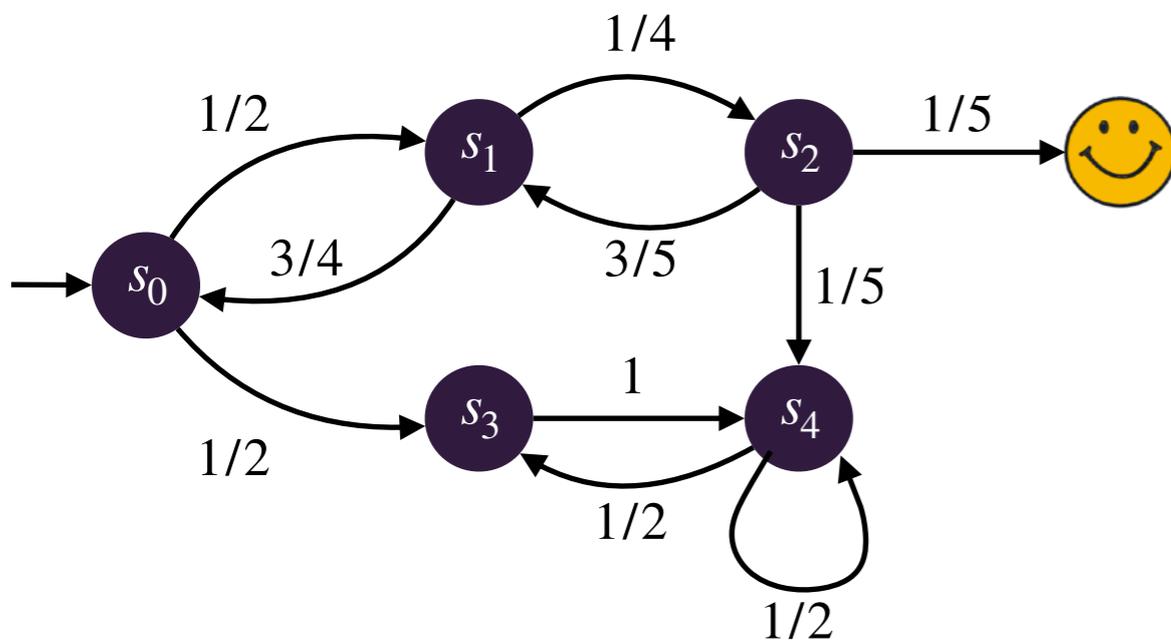
Discrete-time Markov chain (DTMC)

$\mathcal{C} = (\mathcal{S}, s_0, \delta)$  with  $\mathcal{S}$  at most denumerable,  $s_0 \in \mathcal{S}$  and  $\delta : \mathcal{S} \rightarrow \text{Dist}(\mathcal{S})$

# Discrete-time Markov chains

Discrete-time Markov chain (DTMC)

$\mathcal{C} = (\mathcal{S}, s_0, \delta)$  with  $\mathcal{S}$  at most denumerable,  $s_0 \in \mathcal{S}$  and  $\delta : \mathcal{S} \rightarrow \text{Dist}(\mathcal{S})$

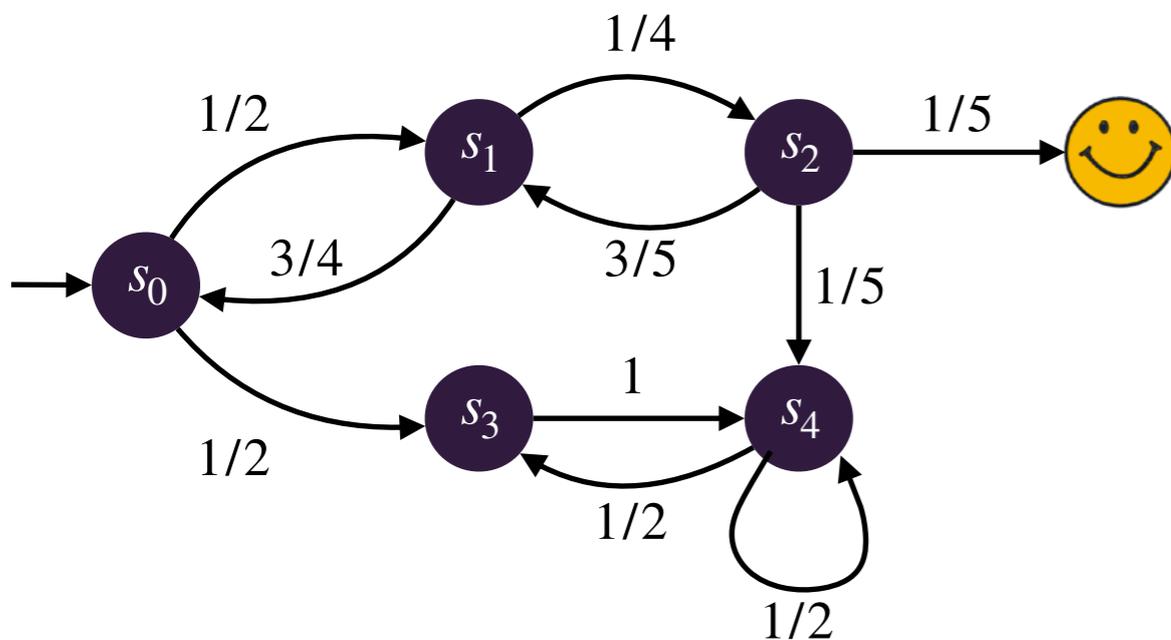


Finite Markov chain

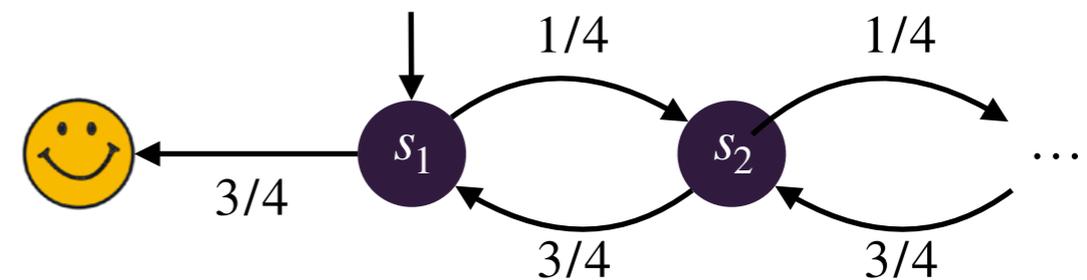
# Discrete-time Markov chains

## Discrete-time Markov chain (DTMC)

$\mathcal{C} = (\mathcal{S}, s_0, \delta)$  with  $\mathcal{S}$  at most denumerable,  $s_0 \in \mathcal{S}$  and  $\delta : \mathcal{S} \rightarrow \text{Dist}(\mathcal{S})$



Finite Markov chain



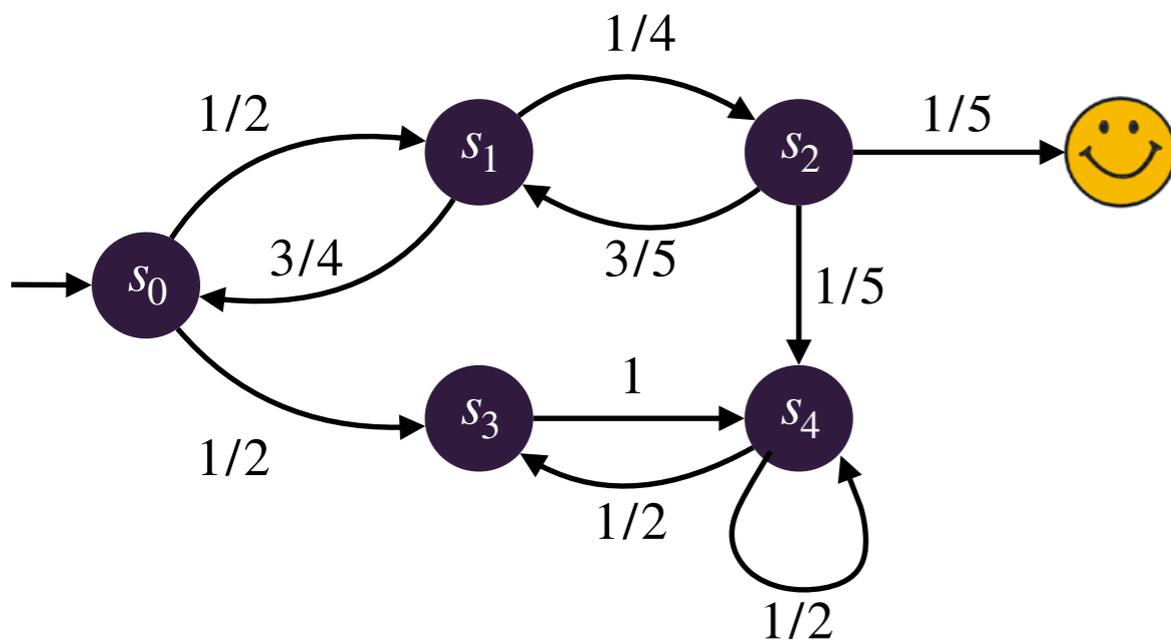
Countable Markov chain  
(random walk of parameter 1/4)

# Discrete-time Markov chains

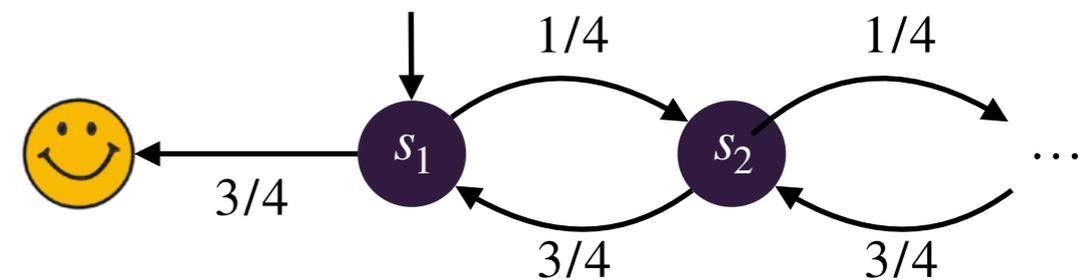
## Discrete-time Markov chain (DTMC)

$\mathcal{C} = (\mathcal{S}, s_0, \delta)$  with  $\mathcal{S}$  at most denumerable,  $s_0 \in \mathcal{S}$  and  $\delta : \mathcal{S} \rightarrow \text{Dist}(\mathcal{S})$

+ effectivity conditions...



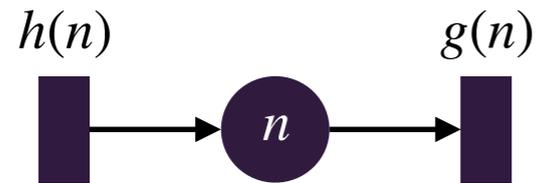
Finite Markov chain



Countable Markov chain  
(random walk of parameter 1/4)

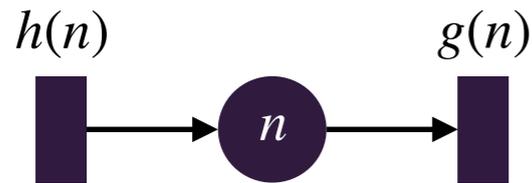
# High-level models for (infinite) Markov chains

- ▶ Queues



# High-level models for (infinite) Markov chains

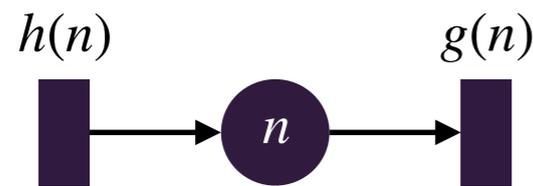
- ▶ Queues



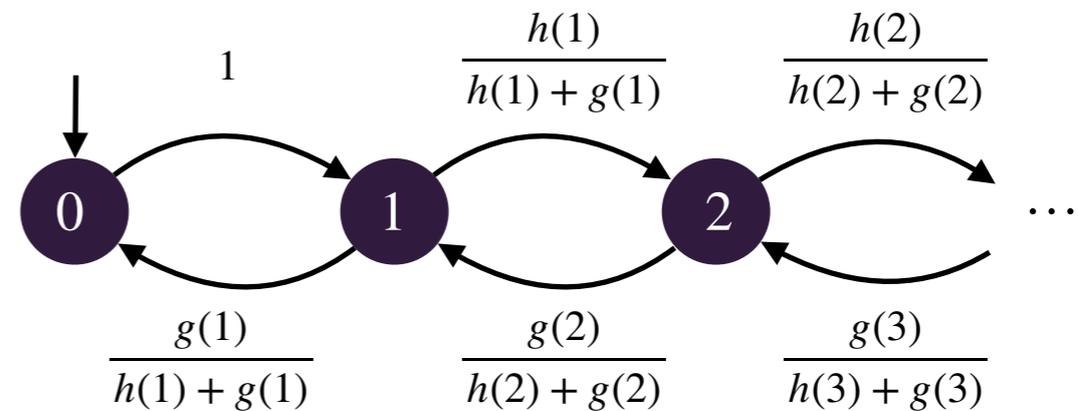
$n$  is the number of clients (tokens)

# High-level models for (infinite) Markov chains

► Queues

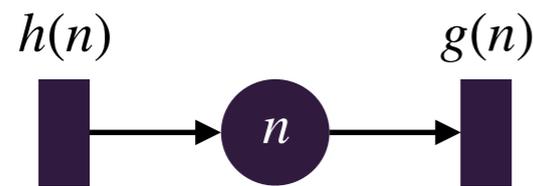


$n$  is the number of clients (tokens)

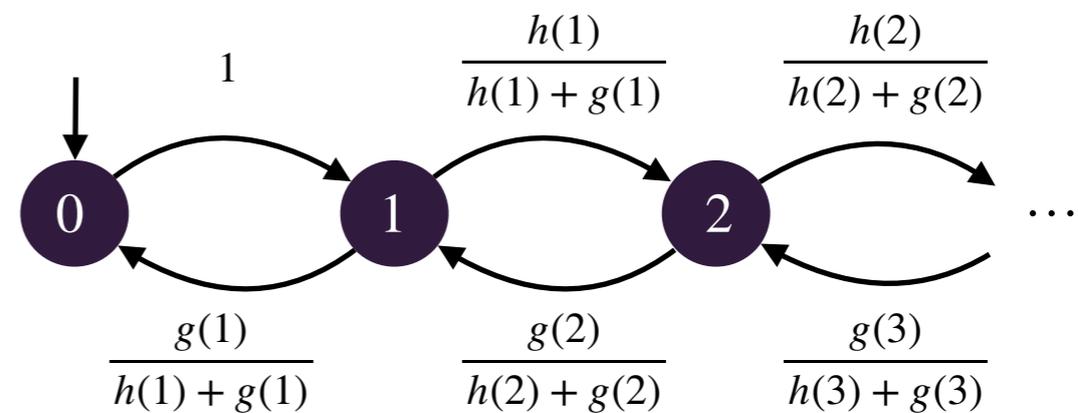


# High-level models for (infinite) Markov chains

- ▶ Queues



$n$  is the number of clients (tokens)

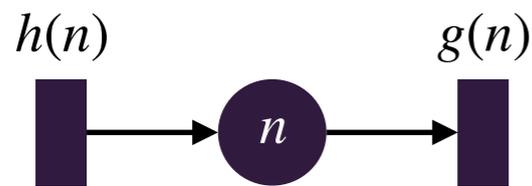


- ▶ Probabilistic pushdown automata

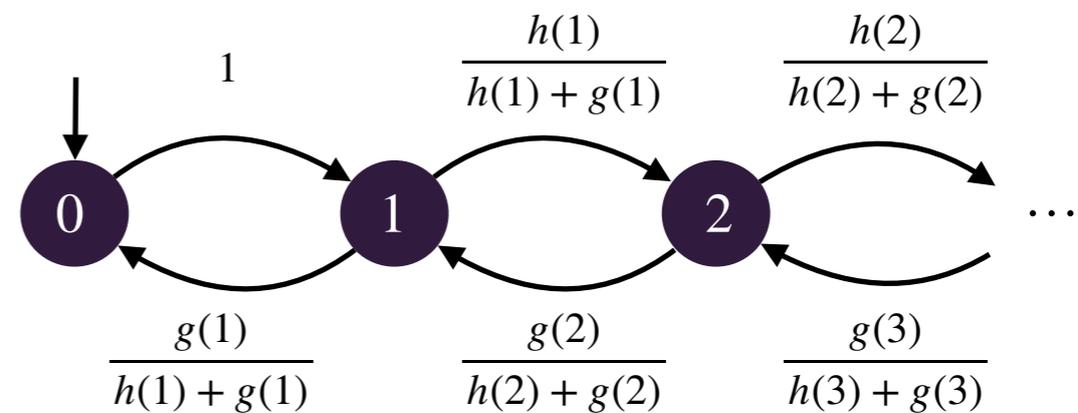
$$\begin{array}{l}
 A \xrightarrow{1} C \quad A \xrightarrow{n} BB \quad B \xrightarrow{5} \varepsilon \\
 B \xrightarrow{n} AA \quad C \xrightarrow{1} C
 \end{array}$$

# High-level models for (infinite) Markov chains

► Queues



$n$  is the number of clients (tokens)



► Probabilistic pushdown automata

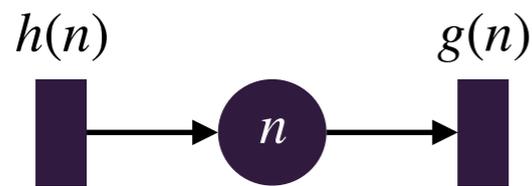
$$A \xrightarrow{1} C \quad A \xrightarrow{n} BB \quad B \xrightarrow{5} \varepsilon$$

$$B \xrightarrow{n} AA \quad C \xrightarrow{1} C$$

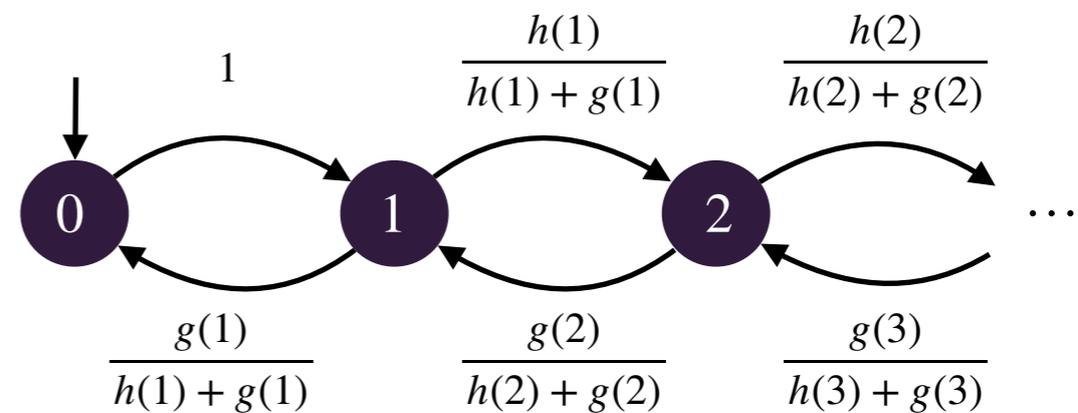
$n$  is the height of the stack

# High-level models for (infinite) Markov chains

► Queues



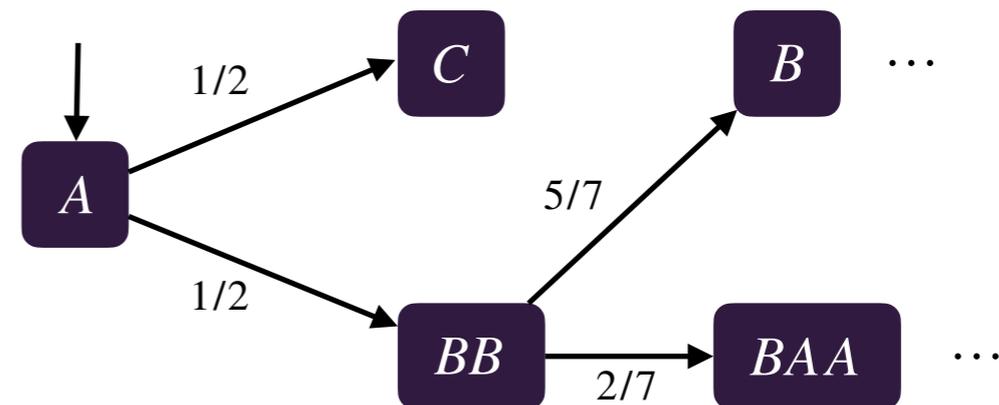
$n$  is the number of clients (tokens)



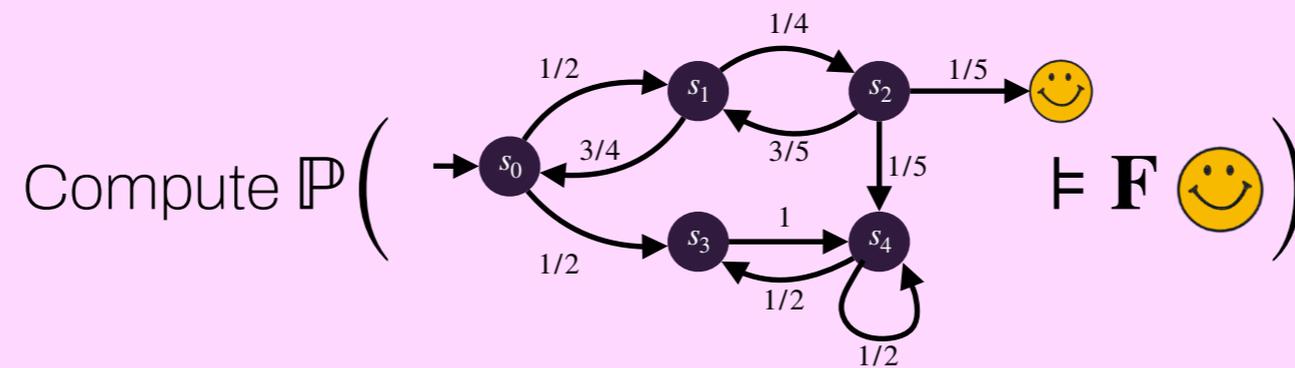
► Probabilistic pushdown automata

$$\begin{array}{l}
 A \xrightarrow{1} C \quad A \xrightarrow{n} BB \quad B \xrightarrow{5} \varepsilon \\
 B \xrightarrow{n} AA \quad C \xrightarrow{1} C
 \end{array}$$

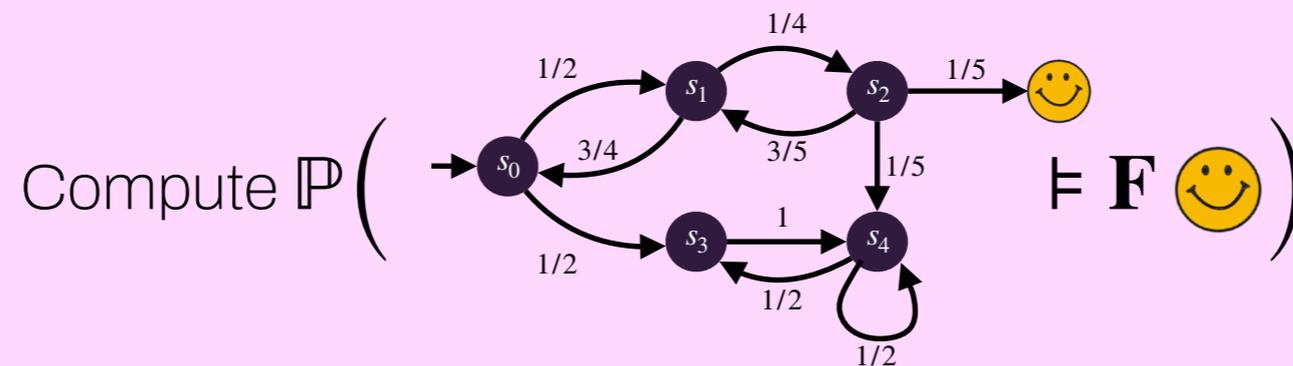
$n$  is the height of the stack



# Quantitative analysis of Markov chains



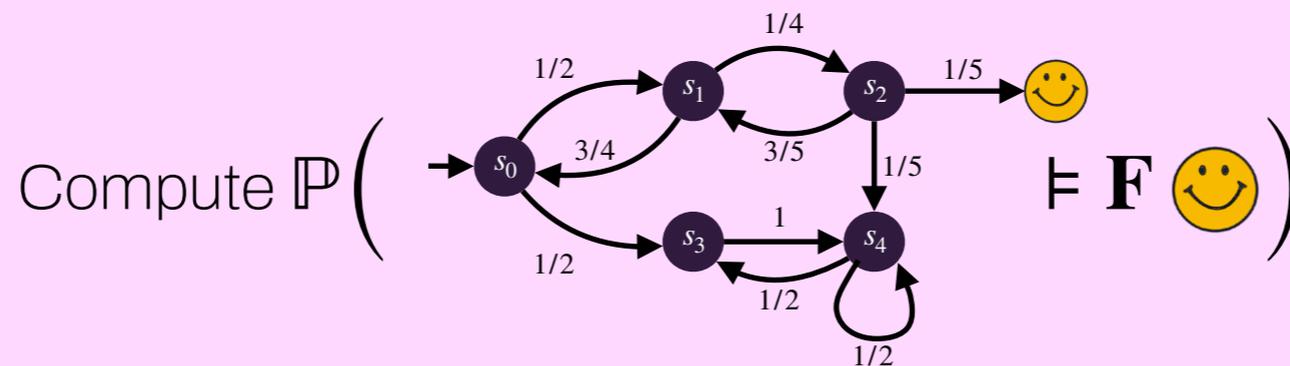
# Quantitative analysis of Markov chains



## Closed-form solution

- ▶ Random walk of parameter  $p > 1/2$ :  
 $\mathbb{P}_{s_n}(\mathbf{F} \text{ 😊}) = \kappa^n$ , where  $\kappa = \frac{1-p}{p}$
- ▶ Does not always exist

# Quantitative analysis of Markov chains



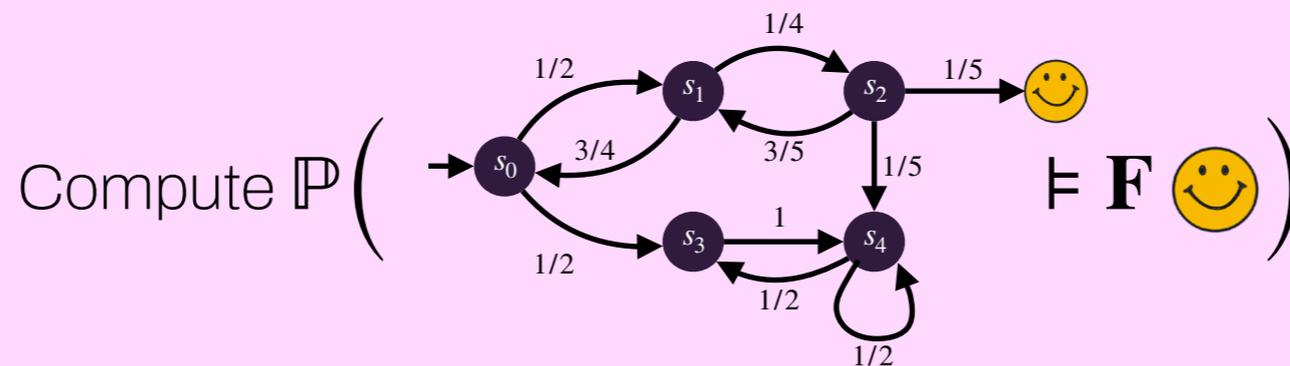
## Closed-form solution

- ▶ Random walk of parameter  $p > 1/2$ :  
 $\mathbb{P}_{s_n}(\mathbf{F} \text{ 😊}) = \kappa^n$ , where  $\kappa = \frac{1-p}{p}$
- ▶ Does not always exist

## Apply a numerical method [RKPN04]

- ▶ 
$$x_s = \begin{cases} 1 & \text{if } s = \text{😊} \\ 0 & \text{if } s \not\models \exists \mathbf{F} \text{ 😊} \\ \sum_t \mathbb{P}(s \rightarrow t) \cdot x_t & \text{otherwise} \end{cases}$$
- ▶  $\mathbb{P}_{s_0}(\mathbf{F} \text{ 😊}) = 1/19$
- ▶ System must be finite
- ▶ Prone to numerical error

# Quantitative analysis of Markov chains



## Closed-form solution

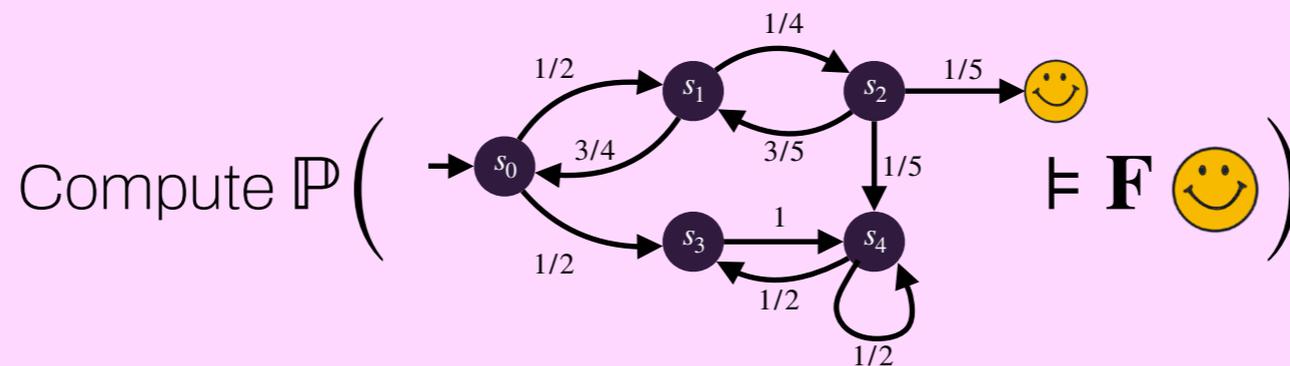
- ▶ Random walk of parameter  $p > 1/2$ :  
 $\mathbb{P}_{s_n}(\mathbf{F} \text{ 😊}) = \kappa^n$ , where  $\kappa = \frac{1-p}{p}$
- ▶ Does not always exist

## Apply a numerical method [RKPN04]

- ▶  $x_s = \begin{cases} 1 & \text{if } s = \text{😊} \\ 0 & \text{if } s \not\models \exists \mathbf{F} \text{ 😊} \\ \sum_t \mathbb{P}(s \rightarrow t) \cdot x_t & \text{otherwise} \end{cases}$
- ▶  $\mathbb{P}_{s_0}(\mathbf{F} \text{ 😊}) = 1/19$
- ▶ System must be finite
- ▶ Prone to numerical error

- ▶ No general method exists for infinite Markov chains

# Quantitative analysis of Markov chains



## Closed-form solution

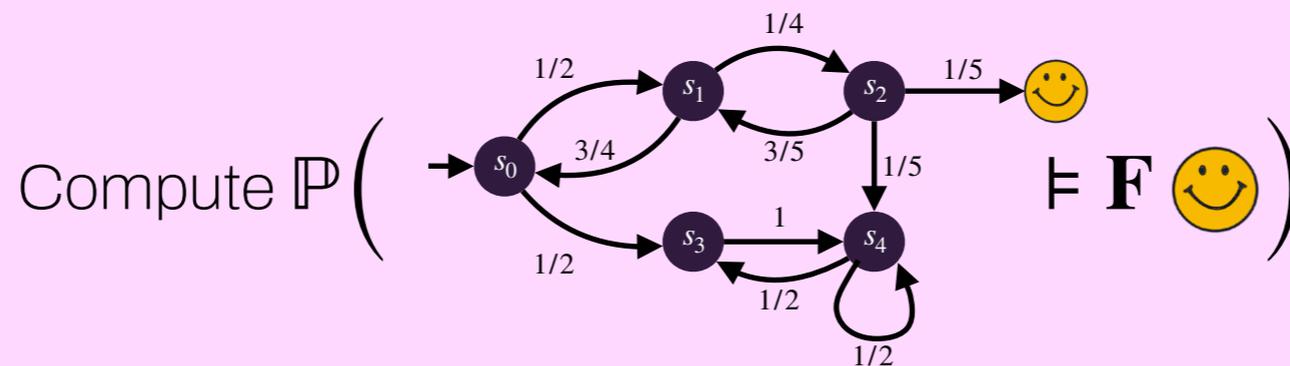
- ▶ Random walk of parameter  $p > 1/2$ :  
 $\mathbb{P}_{s_n}(\mathbf{F} \text{ 😊}) = \kappa^n$ , where  $\kappa = \frac{1-p}{p}$
- ▶ Does not always exist

## Apply a numerical method [RKPN04]

- ▶  $x_s = \begin{cases} 1 & \text{if } s = \text{😊} \\ 0 & \text{if } s \notin \exists \mathbf{F} \text{ 😊} \\ \sum_t \mathbb{P}(s \rightarrow t) \cdot x_t & \text{otherwise} \end{cases}$
- ▶  $\mathbb{P}_{s_0}(\mathbf{F} \text{ 😊}) = 1/19$
- ▶ System must be finite
- ▶ Prone to numerical error

- ▶ No general method exists for infinite Markov chains
- ▶ Ad-hoc methods in specific classes

# Quantitative analysis of Markov chains



## Closed-form solution

- ▶ Random walk of parameter  $p > 1/2$ :  
 $\mathbb{P}_{s_n}(\mathbf{F} \text{ 😊}) = \kappa^n$ , where  $\kappa = \frac{1-p}{p}$
- ▶ Does not always exist

## Apply a numerical method [RKPN04]

- ▶  $x_s = \begin{cases} 1 & \text{if } s = \text{😊} \\ 0 & \text{if } s \not\models \exists \mathbf{F} \text{ 😊} \\ \sum_t \mathbb{P}(s \rightarrow t) \cdot x_t & \text{otherwise} \end{cases}$
- ▶  $\mathbb{P}_{s_0}(\mathbf{F} \text{ 😊}) = 1/19$
- ▶ System must be finite
- ▶ Prone to numerical error

- ▶ No general method exists for infinite Markov chains
- ▶ Ad-hoc methods in specific classes
- ▶ Specific approaches for **decisive** Markov chains

# Decisiveness

$$\text{☹} = \{s \in S \mid s \not\equiv \exists \mathbf{F} \text{☺}\}$$

## Decisiveness

A DTMC  $\mathcal{C}$  is *decisive* from  $s$  w.r.t.  $\text{☺}$  if  $\mathbb{P}_s(\mathbf{F} \text{☺} \vee \mathbf{F} \text{☹}) = 1$

# Decisiveness

$$\text{☹} = \{s \in S \mid s \not\models \exists \mathbf{F} \text{☺}\}$$

## Decisiveness

A DTMC  $\mathcal{C}$  is **decisive** from  $s$  w.r.t.  $\text{☺}$  if  $\mathbb{P}_s(\mathbf{F} \text{☺} \vee \mathbf{F} \text{☹}) = 1$

- ▶ Examples of decisive Markov chains: finite Markov chains, probabilistic lossy channel systems, probabilistic VASS, noisy Turing machines, ...

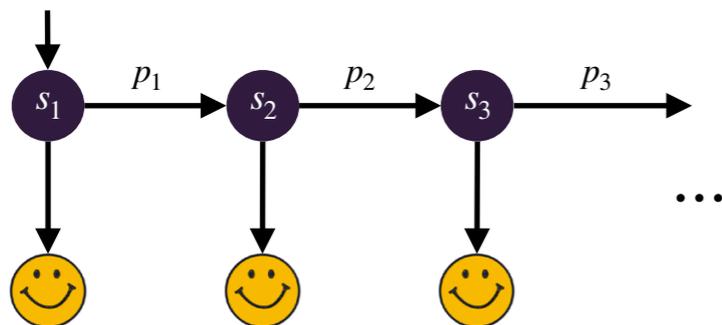
# Decisiveness

$$\text{☹} = \{s \in S \mid s \not\equiv \exists \mathbf{F} \text{☺}\}$$

## Decisiveness

A DTMC  $\mathcal{C}$  is **decisive** from  $s$  w.r.t.  $\text{☺}$  if  $\mathbb{P}_s(\mathbf{F} \text{☺} \vee \mathbf{F} \text{☹}) = 1$

- ▶ Examples of decisive Markov chains: finite Markov chains, probabilistic lossy channel systems, probabilistic VASS, noisy Turing machines, ...
- ▶ Example/counterexample:



- $\mathbb{P}(\mathbf{G} \neg \text{☺}) = \prod_{i \geq 1} p_i$
- Decisive iff this product equals 0

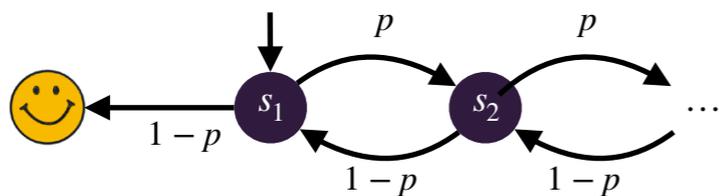
# Decisiveness

$$\text{☹} = \{s \in S \mid s \not\equiv \exists \mathbf{F} \text{☺}\}$$

## Decisiveness

A DTMC  $\mathcal{C}$  is **decisive** from  $s$  w.r.t.  $\text{☺}$  if  $\mathbb{P}_s(\mathbf{F} \text{☺} \vee \mathbf{F} \text{☹}) = 1$

- ▶ Examples of decisive Markov chains: finite Markov chains, probabilistic lossy channel systems, probabilistic VASS, noisy Turing machines, ...
- ▶ Example/counterexample:



- Recurrent random walk ( $p \leq 1/2$ ): decisive
- Transient random walk ( $p > 1/2$ ): not decisive

# Deciding decisiveness?

## Classes where decisiveness can be decided

- ▶ Probabilistic pushdown automata with constant weights [ABM07]
- ▶ Random walks with polynomial weights [FHY23]
- ▶ So-called probabilistic homogeneous one-counter machines with polynomial weights (this extends the model of quasi-birth death processes) [FHY23]

[ABM07] P.A. Abdulla, N. Ben Henda, R. Mayr. *Decisive Markov chains* (LMCS, 2007)

[FHY23] A. Finkel, S. Haddad, L. Yé. *About decisiveness of dynamic probabilistic models* (CONCUR'23)

# Approximation scheme

- ▶ Aim: compute probability of **F** 😊
- ▶ 😞 =  $\{s \in S \mid s \not\models \exists \mathbf{F} \text{ 😊}\}$

# Approximation scheme

- ▶ Aim: compute probability of  $\mathbf{F}$  😊
- ▶ 😞 =  $\{s \in S \mid s \not\models \exists \mathbf{F} \text{ 😊}\}$

## Approximation scheme

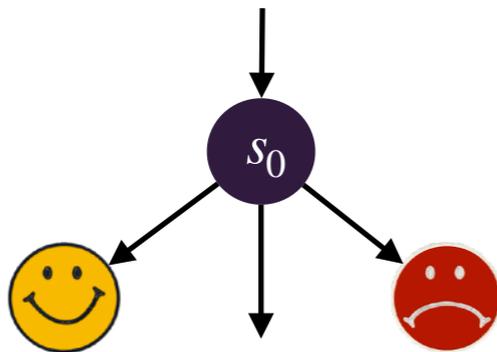
Given  $\varepsilon > 0$ , for every  $n$ , compute:

$$\begin{cases} p_n^{\text{yes}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😊}) \\ p_n^{\text{no}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😞}) \end{cases}$$

until  $p_n^{\text{yes}} + p_n^{\text{no}} \geq 1 - \varepsilon$

# Approximation scheme

- ▶ Aim: compute probability of  $\mathbf{F}$  😊
- ▶ 😞 =  $\{s \in S \mid s \not\models \exists \mathbf{F} \text{ 😊}\}$



## Approximation scheme

Given  $\varepsilon > 0$ , for every  $n$ , compute:

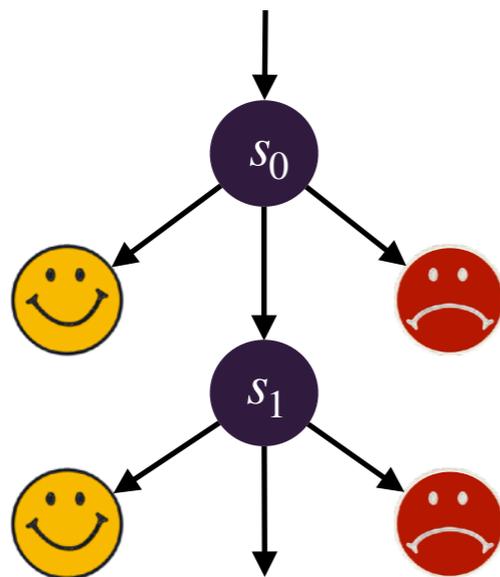
$$\begin{cases} p_n^{\text{yes}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😊}) \\ p_n^{\text{no}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😞}) \end{cases}$$

until  $p_n^{\text{yes}} + p_n^{\text{no}} \geq 1 - \varepsilon$

$$p_1^{\text{yes}} \leq \mathbb{P}(\mathbf{F} \text{ 😊}) \leq 1 - p_1^{\text{no}}$$

# Approximation scheme

- ▶ Aim: compute probability of  $\mathbf{F}$  😊
- ▶ 😞 =  $\{s \in S \mid s \not\models \exists \mathbf{F} \text{ 😊}\}$



## Approximation scheme

Given  $\varepsilon > 0$ , for every  $n$ , compute:

$$\begin{cases} p_n^{\text{yes}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😊}) \\ p_n^{\text{no}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😞}) \end{cases}$$

until  $p_n^{\text{yes}} + p_n^{\text{no}} \geq 1 - \varepsilon$

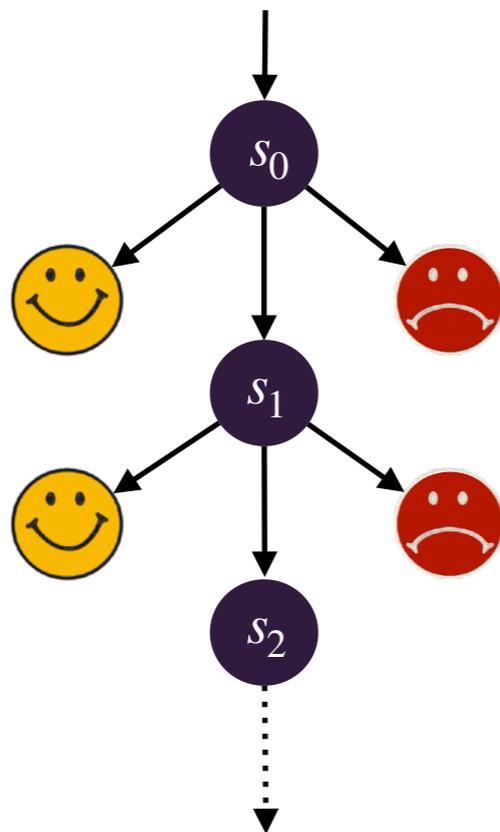
$$p_1^{\text{yes}} \leq \mathbb{P}(\mathbf{F} \text{ 😊}) \leq 1 - p_1^{\text{no}}$$

$\wedge$   $\vee$

$$p_2^{\text{yes}} \leq \mathbb{P}(\mathbf{F} \text{ 😊}) \leq 1 - p_2^{\text{no}}$$

# Approximation scheme

- ▶ Aim: compute probability of  $\mathbf{F}$  😊
- ▶ 😞 =  $\{s \in S \mid s \not\models \exists \mathbf{F} \text{ 😊}\}$



## Approximation scheme

Given  $\varepsilon > 0$ , for every  $n$ , compute:

$$\begin{cases} p_n^{\text{yes}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😊}) \\ p_n^{\text{no}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😞}) \end{cases}$$

until  $p_n^{\text{yes}} + p_n^{\text{no}} \geq 1 - \varepsilon$

$$p_1^{\text{yes}} \leq \mathbb{P}(\mathbf{F} \text{ 😊}) \leq 1 - p_1^{\text{no}}$$

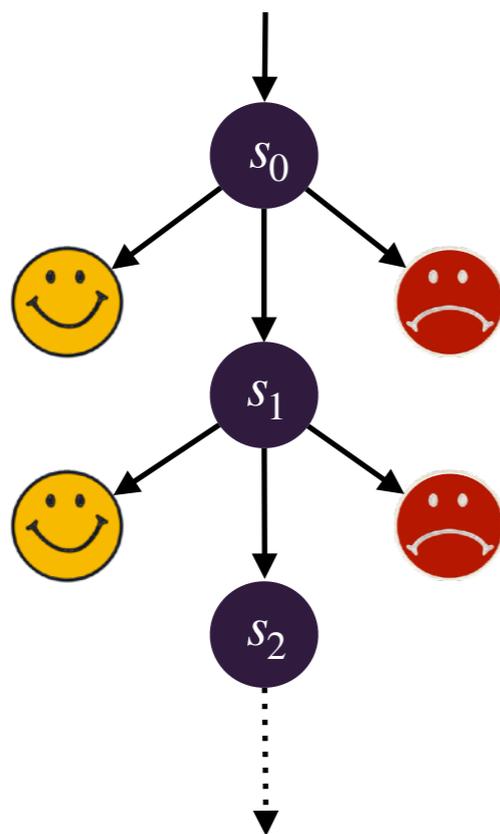
$\wedge$   $\forall$

$$p_2^{\text{yes}} \leq \mathbb{P}(\mathbf{F} \text{ 😊}) \leq 1 - p_2^{\text{no}}$$

$\wedge$   $\vdots$   $\forall$

# Approximation scheme

- ▶ Aim: compute probability of  $\mathbf{F}$  😊
- ▶ 😞 =  $\{s \in S \mid s \not\models \exists \mathbf{F} \text{ 😊}\}$



## Approximation scheme

Given  $\varepsilon > 0$ , for every  $n$ , compute:

$$\begin{cases} p_n^{\text{yes}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😊}) \\ p_n^{\text{no}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😞}) \end{cases}$$

until  $p_n^{\text{yes}} + p_n^{\text{no}} \geq 1 - \varepsilon$

$$p_1^{\text{yes}} \leq \mathbb{P}(\mathbf{F} \text{ 😊}) \leq 1 - p_1^{\text{no}}$$

$$\wedge \qquad \qquad \qquad \vee$$

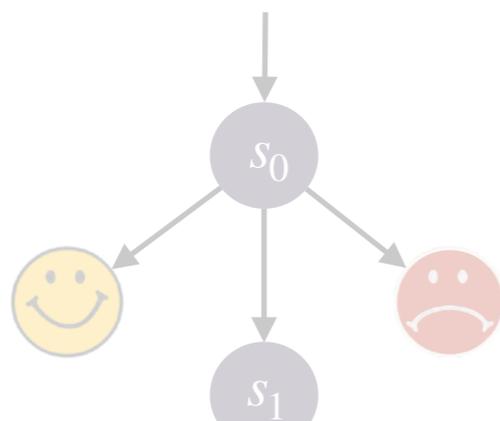
$$p_2^{\text{yes}} \leq \mathbb{P}(\mathbf{F} \text{ 😊}) \leq 1 - p_2^{\text{no}}$$

$$\wedge \qquad \qquad \qquad \vee$$

At the limit:  $\mathbb{P}(\mathbf{F} \text{ 😊}) \qquad \qquad \qquad 1 - \mathbb{P}(\mathbf{F} \text{ 😞})$

# Approximation scheme

- ▶ Aim: compute probability of  $\mathbf{F}$  😊
- ▶ 😞 =  $\{s \in S \mid s \not\models \exists \mathbf{F} \text{ 😊}\}$



The approximation scheme converges  
iff  
 $\mathcal{C}$  is decisive from  $s_0$  w.r.t. 😊

## Approximation scheme

Given  $\varepsilon > 0$ , for every  $n$ , compute:

$$\begin{cases} p_n^{\text{yes}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😊}) \\ p_n^{\text{no}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😞}) \end{cases}$$

until  $p_n^{\text{yes}} + p_n^{\text{no}} \geq 1 - \varepsilon$

$$p_1^{\text{yes}} \leq \mathbb{P}(\mathbf{F} \text{ 😊}) \leq 1 - p_1^{\text{no}}$$

$$\wedge \qquad \qquad \qquad \vee$$

$$p_2^{\text{yes}} \leq \mathbb{P}(\mathbf{F} \text{ 😊}) \leq 1 - p_2^{\text{no}}$$

$$\wedge \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vee$$

At the limit:  $\mathbb{P}(\mathbf{F} \text{ 😊}) \qquad \qquad \qquad 1 - \mathbb{P}(\mathbf{F} \text{ 😞})$

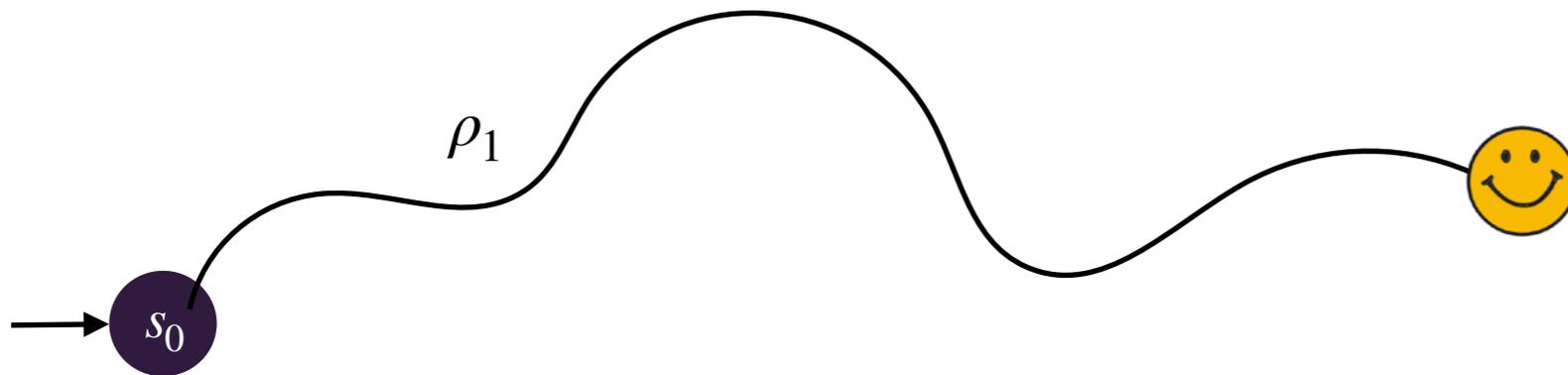
# Statistical model-checking

Sample  $N$  paths



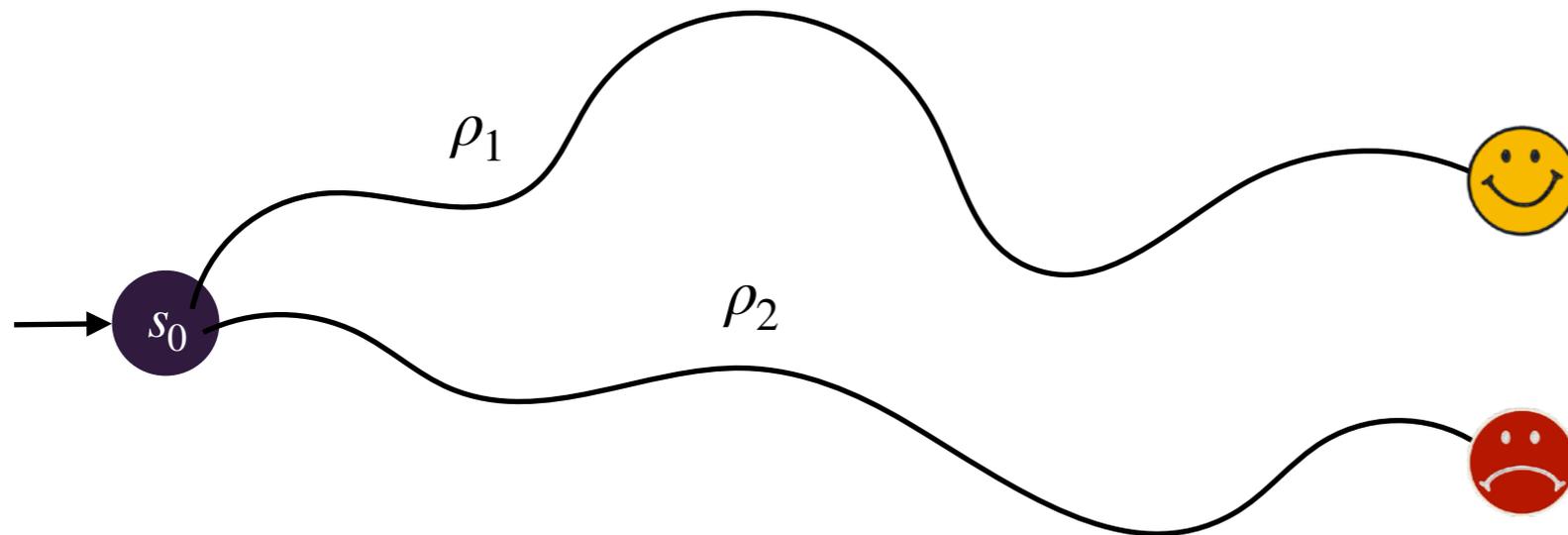
# Statistical model-checking

Sample  $N$  paths



# Statistical model-checking

Sample  $N$  paths

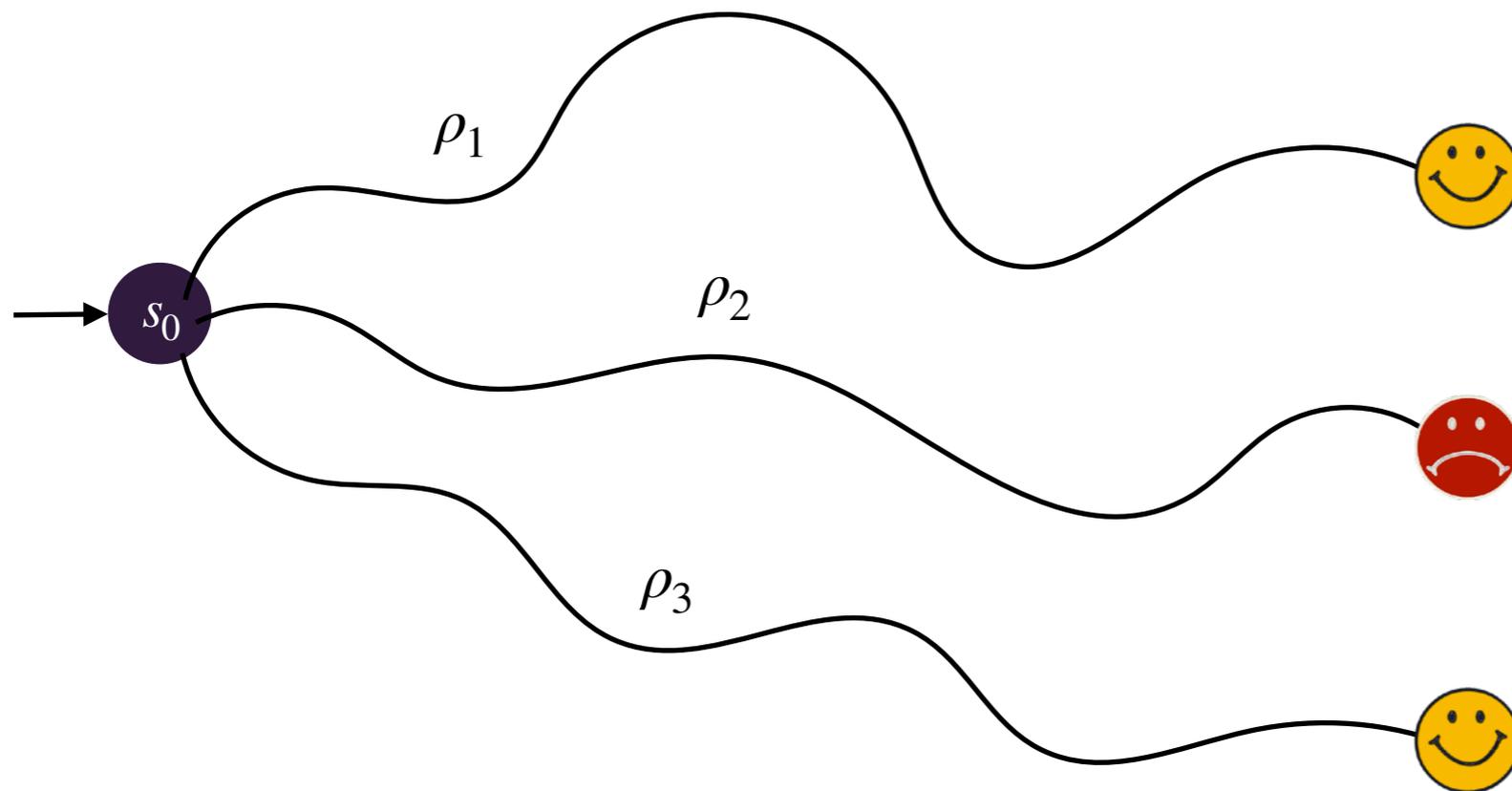


$$n_1 = 1$$

$$n_2 = n_1$$

# Statistical model-checking

Sample  $N$  paths



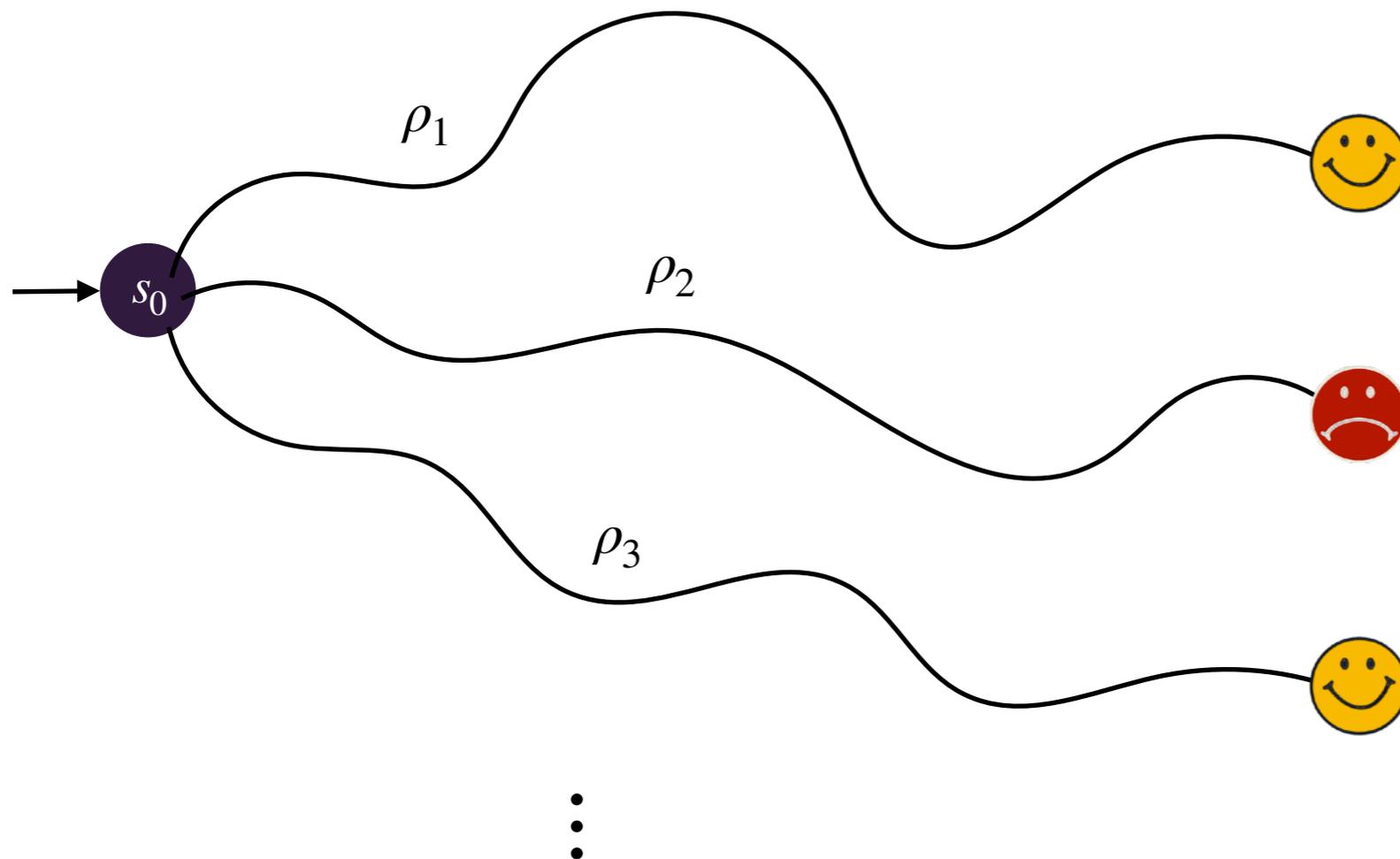
$$n_1 = 1$$

$$n_2 = n_1$$

$$n_3 = n_2 + 1$$

# Statistical model-checking

Sample  $N$  paths



$$n_1 = 1$$

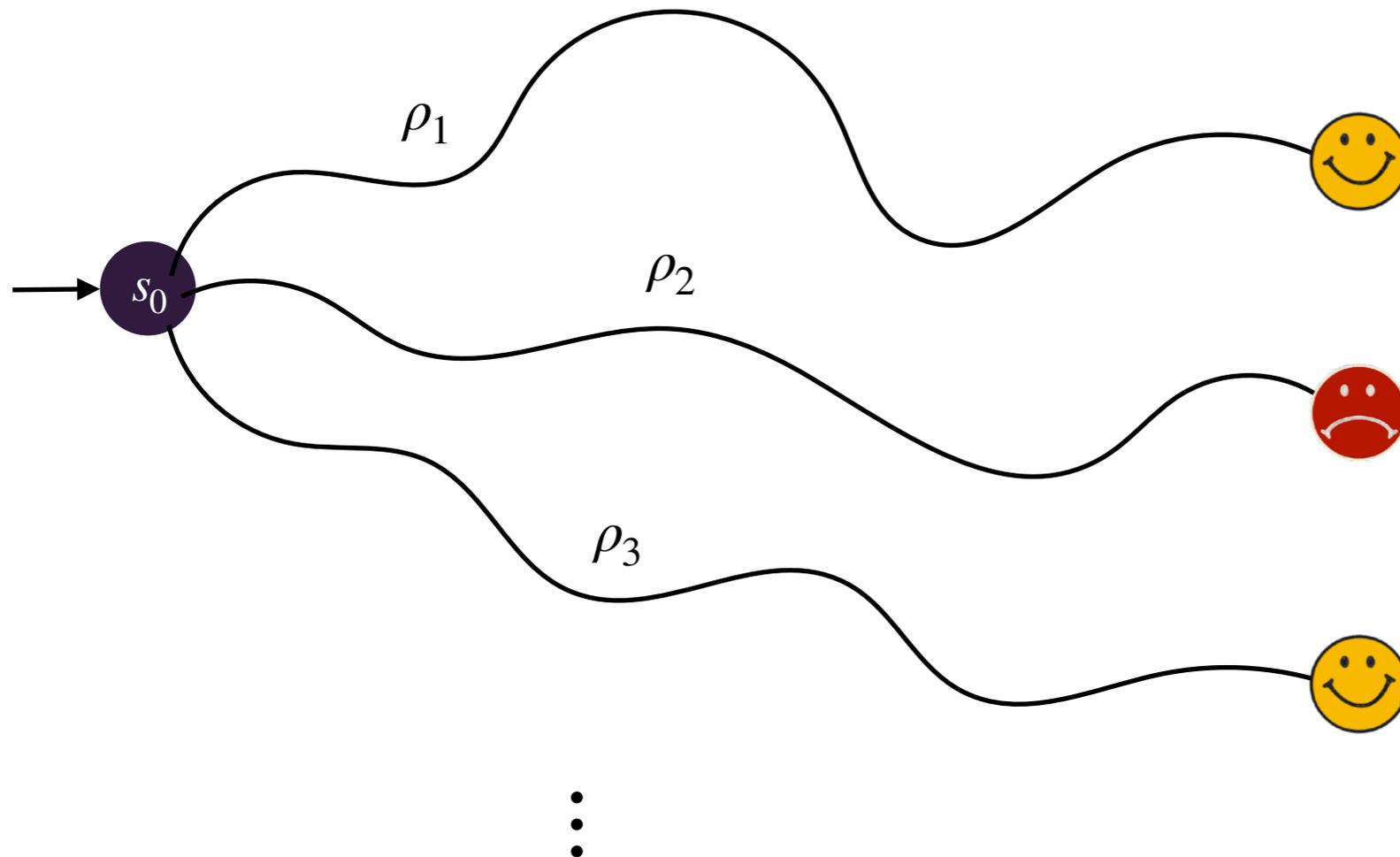
$$n_2 = n_1$$

$$n_3 = n_2 + 1$$

⋮

# Statistical model-checking

Sample  $N$  paths



$$n_1 = 1$$

$$n_2 = n_1$$

$$n_3 = n_2 + 1$$

⋮

Return  $\frac{n_N}{N}$  + some confidence interval  
(in the best case)

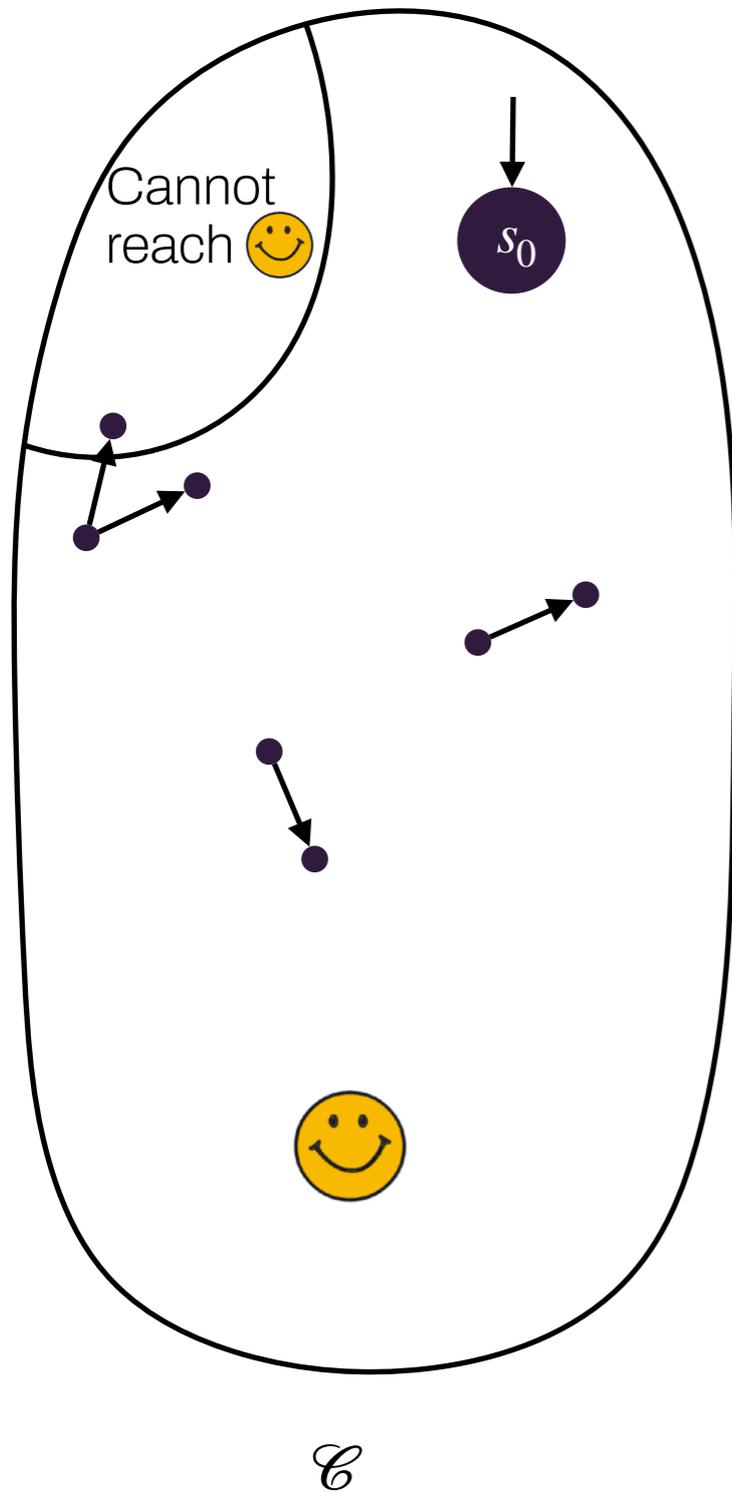
# Termination, efficiency and guarantees

## Termination

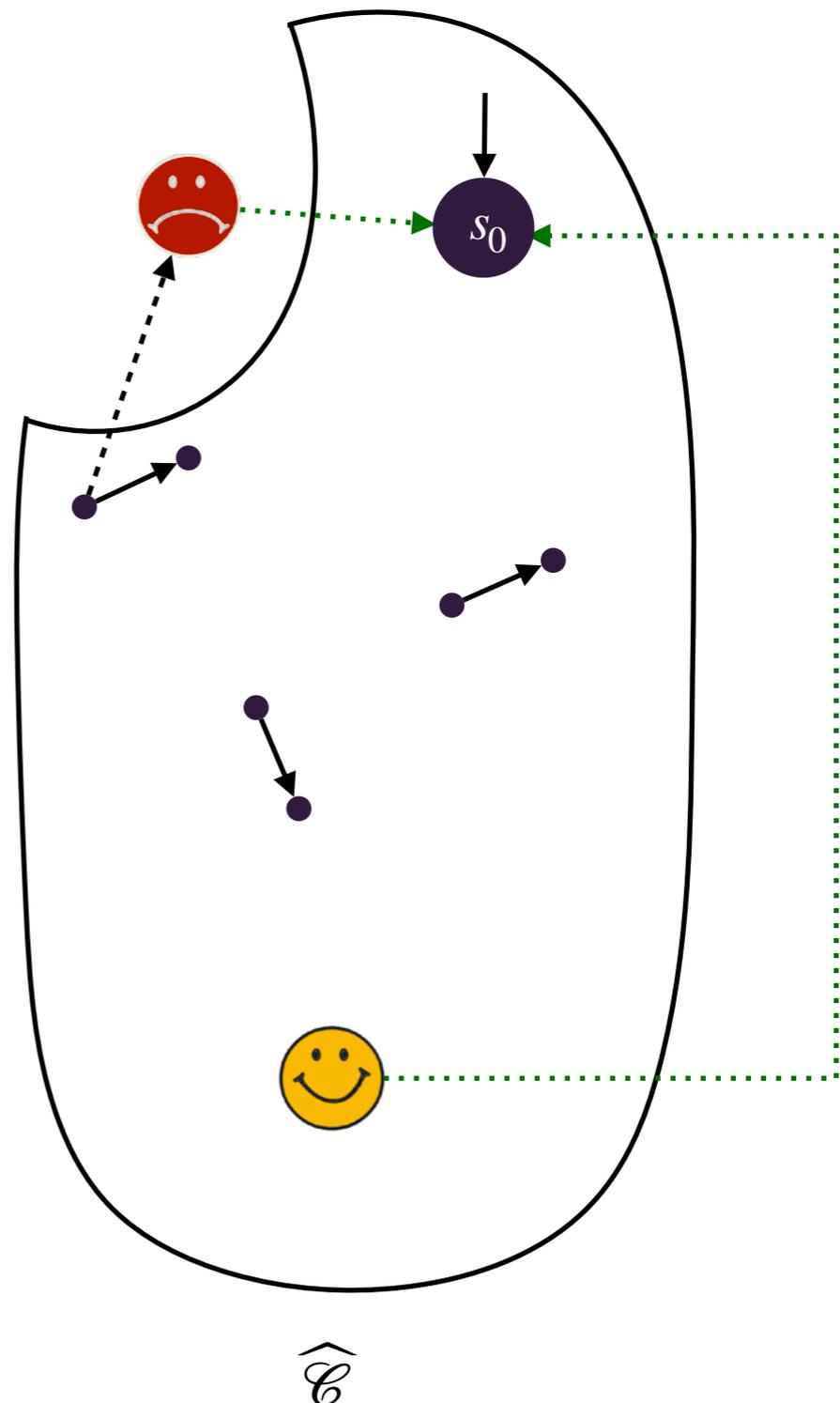
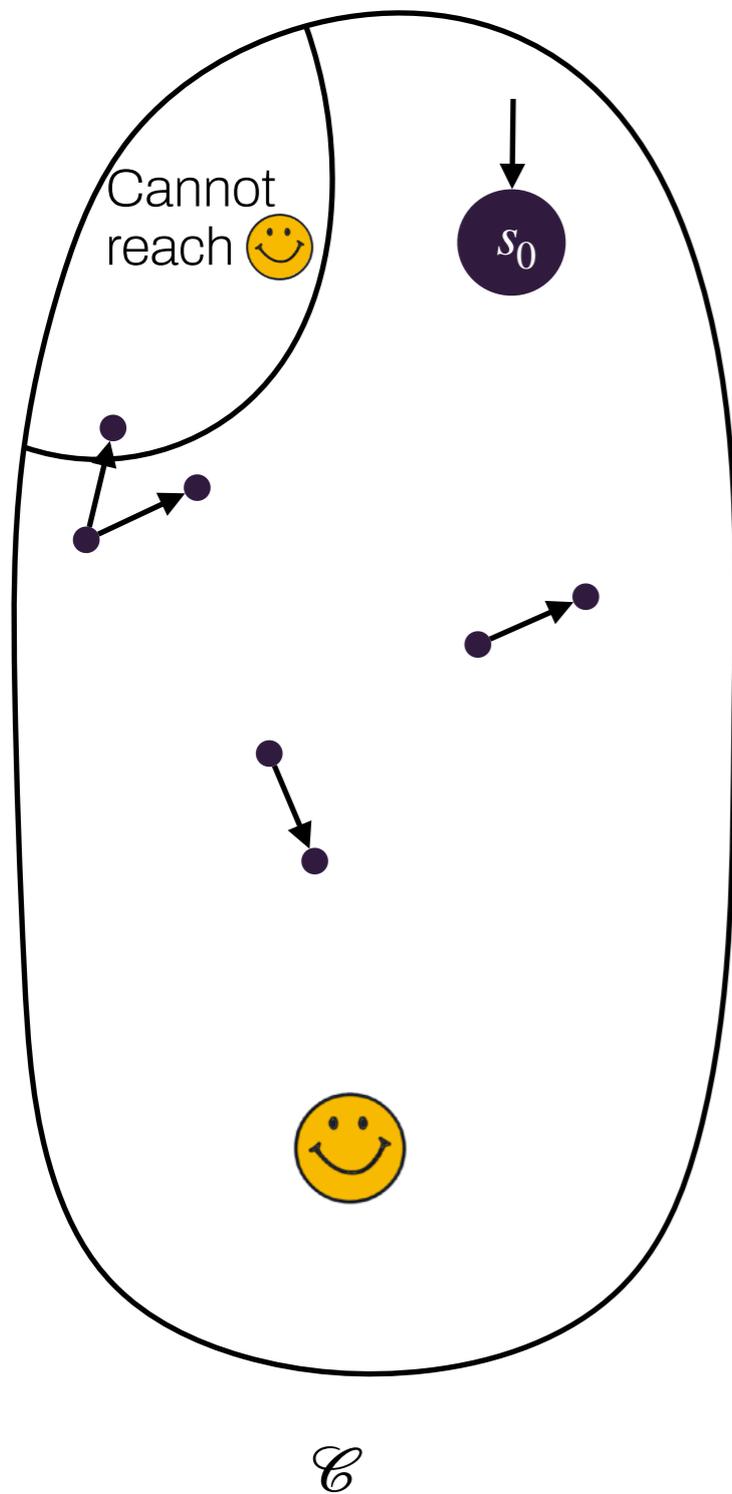
(To our knowledge, never expressed like this)

A sampled path starting at  $s_0$  almost-surely hits 😊 or 😞  
iff  
 $\mathcal{C}$  is decisive from  $s_0$  w.r.t. 😊

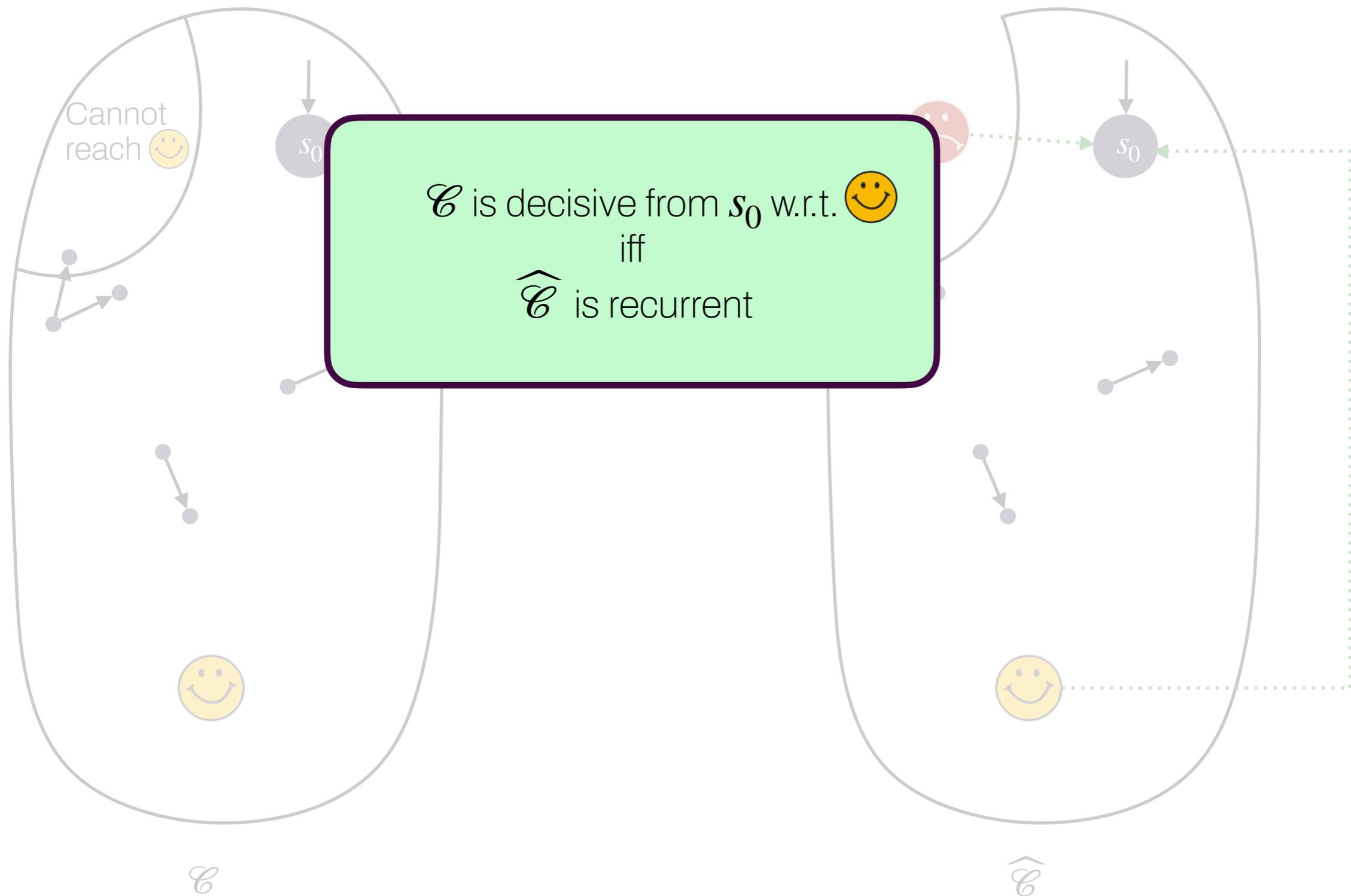
# Decisiveness vs recurrence



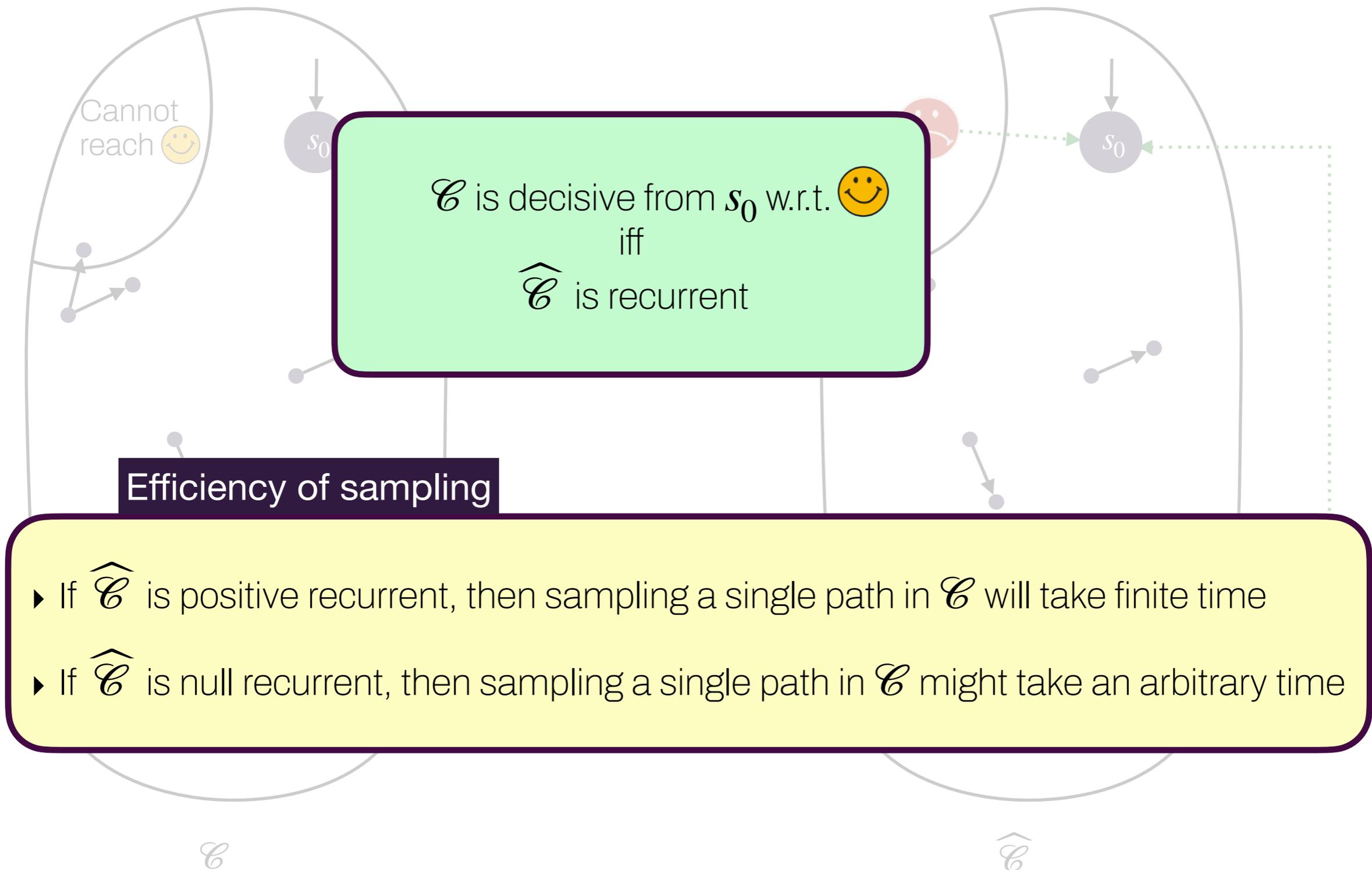
# Decisiveness vs recurrence



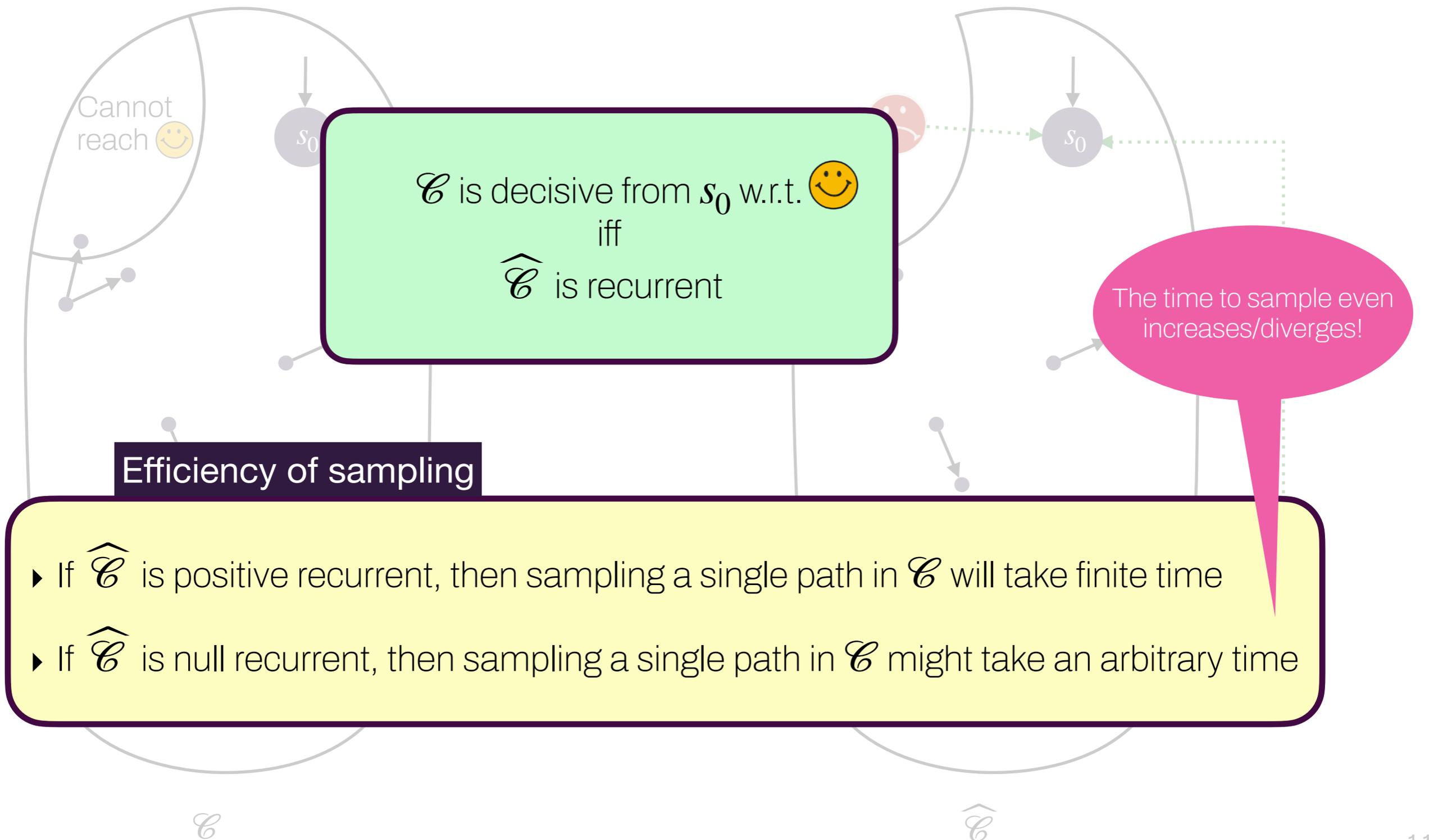
# Decisiveness vs recurrence



# Decisiveness vs recurrence



# Decisiveness vs recurrence



# Termination, efficiency and guarantees

## Termination

(To our knowledge, never expressed like this)

A sampled path starting at  $s_0$  almost-surely hits 😊 or 😞  
iff  
 $\mathcal{C}$  is decisive from  $s_0$  w.r.t. 😊

+ efficiency if finite return time  
(«  $\widehat{\mathcal{C}}$  is positive recurrent »)

# Termination, efficiency and guarantees

## Termination

(To our knowledge, never expressed like this)

A sampled path starting at  $s_0$  almost-surely hits 😊 or 😞  
iff  
 $\mathcal{C}$  is decisive from  $s_0$  w.r.t. 😊

+ efficiency if finite return time  
(«  $\widehat{\mathcal{C}}$  is positive recurrent »)

## Guarantees: Hoeffding's inequalities

Let  $\varepsilon, \delta > 0$ , let  $N \geq \frac{8}{\varepsilon^2} \log\left(\frac{2}{\delta}\right)$ . Then:

$$\mathbb{P}\left(\left|\frac{n_N}{N} - \mathbb{P}(\mathbf{F} \text{ 😊})\right| \geq \frac{\varepsilon}{2}\right) \leq \delta$$

# Termination, efficiency and guarantees

## Termination

(To our knowledge, never expressed like this)

A sampled path starting at  $s_0$  almost-surely hits 😊 or 😞  
iff  
 $\mathcal{C}$  is decisive from  $s_0$  w.r.t. 😊

+ efficiency if finite return time  
(«  $\widehat{\mathcal{C}}$  is positive recurrent »)

## Guarantees: Hoeffding's inequalities

Empirical average

Let  $\varepsilon, \delta > 0$ , let  $N \geq \frac{8}{\varepsilon^2} \log\left(\frac{2}{\delta}\right)$ . Then:

$$\mathbb{P}\left(\left|\frac{n_N}{N} - \mathbb{P}(\mathbf{F} \text{ 😊})\right| \geq \frac{\varepsilon}{2}\right) \leq \delta$$

# Termination, efficiency and guarantees

## Termination

(To our knowledge, never expressed like this)

A sampled path starting at  $s_0$  almost-surely hits 😊 or 😞  
iff  
 $\mathcal{C}$  is decisive from  $s_0$  w.r.t. 😊

+ efficiency if finite return time  
(«  $\widehat{\mathcal{C}}$  is positive recurrent »)

## Guarantees: Hoeffding's inequalities

Empirical average

Let  $\varepsilon, \delta > 0$ , let  $N \geq \frac{8}{\varepsilon^2} \log\left(\frac{2}{\delta}\right)$ . Then:

$$\mathbb{P}\left(\left|\frac{n_N}{N} - \mathbb{P}(\mathbf{F} \text{ 😊})\right| \geq \frac{\varepsilon}{2}\right) \leq \delta$$

Precision

# Termination, efficiency and guarantees

## Termination

(To our knowledge, never expressed like this)

A sampled path starting at  $s_0$  almost-surely hits 😊 or 😞  
iff  
 $\mathcal{C}$  is decisive from  $s_0$  w.r.t. 😊

+ efficiency if finite return time  
(«  $\widehat{\mathcal{C}}$  is positive recurrent »)

## Guarantees: Hoeffding's inequalities

Empirical average

Let  $\varepsilon, \delta > 0$ , let  $N \geq \frac{8}{\varepsilon^2} \log\left(\frac{2}{\delta}\right)$ . Then:

$$\mathbb{P}\left(\left|\frac{n_N}{N} - \mathbb{P}(\mathbf{F} \text{ 😊})\right| \geq \frac{\varepsilon}{2}\right) \leq \delta$$

Confidence level

Precision

# Termination, efficiency and guarantees

## Termination

(To our knowledge, never expressed like this)

A sampled path starting at  $s_0$  almost-surely hits 😊 or 😞  
iff  
 $\mathcal{C}$  is decisive from  $s_0$  w.r.t. 😊

+ efficiency if finite return time  
(«  $\widehat{\mathcal{C}}$  is positive recurrent »)

## Guarantees: Hoeffding's inequalities

Empirical average

Let  $\varepsilon, \delta > 0$ , let  $N \geq \frac{8}{\varepsilon^2} \log\left(\frac{2}{\delta}\right)$ . Then:

$$\mathbb{P}\left(\left|\frac{n_N}{N} - \mathbb{P}(\mathbf{F} \text{ 😊})\right| \geq \frac{\varepsilon}{2}\right) \leq \delta$$

Confidence level

Precision

$\left[\frac{n_N}{N} - \frac{\varepsilon}{2}; \frac{n_N}{N} + \frac{\varepsilon}{2}\right]$ : confidence interval

# Termination, efficiency and guarantees

## Termination

(To our knowledge, never expressed like this)

A sampled path starting at  $s_0$  almost-surely hits 😊 or 😞  
iff  
 $\mathcal{C}$  is decisive from  $s_0$  w.r.t. 😊

+ efficiency if finite return time  
(«  $\mathcal{C}$  is positive recurrent »)

## Guarantees: Hoeffding's inequalities

Let  $\varepsilon, \delta > 0$  s.t.  $N \geq \frac{8B^2}{\varepsilon^2} \log\left(\frac{2}{\delta}\right)$ . Then:

Empirical estimation

$$\mathbb{P}\left(\left|\frac{f_N}{N} - \mathbb{E}(f_L, \text{😊})\right| \geq \frac{\varepsilon}{2}\right) \leq \delta$$

$B$  bound on the function

$\left[\frac{f_N}{N} - \frac{\varepsilon}{2}; \frac{f_N}{N} + \frac{\varepsilon}{2}\right]$ : confidence interval

Value given by  $L$  for paths that stop at 😊

What can we do for  
non-decisive Markov chains??

# Importance sampling

## [KH51]

[KH51] H. Kahn, T. E. Harris. *Estimation of particle transmission by random sampling* (National Bureau of Standards applied mathematics series, 1951)

[Bar14] B. Barbot. *Acceleration for statistical model checking* (PhD thesis)

[BHP12] B. Barbot, S. Haddad, C. Picaronny. *Coupling and Importance Sampling for Statistical Model Checking* (TACAS'12)

# Importance sampling

## [KH51]

- ▶ Originally used for rare events

[KH51] H. Kahn, T. E. Harris. *Estimation of particle transmission by random sampling* (National Bureau of Standards applied mathematics series, 1951)

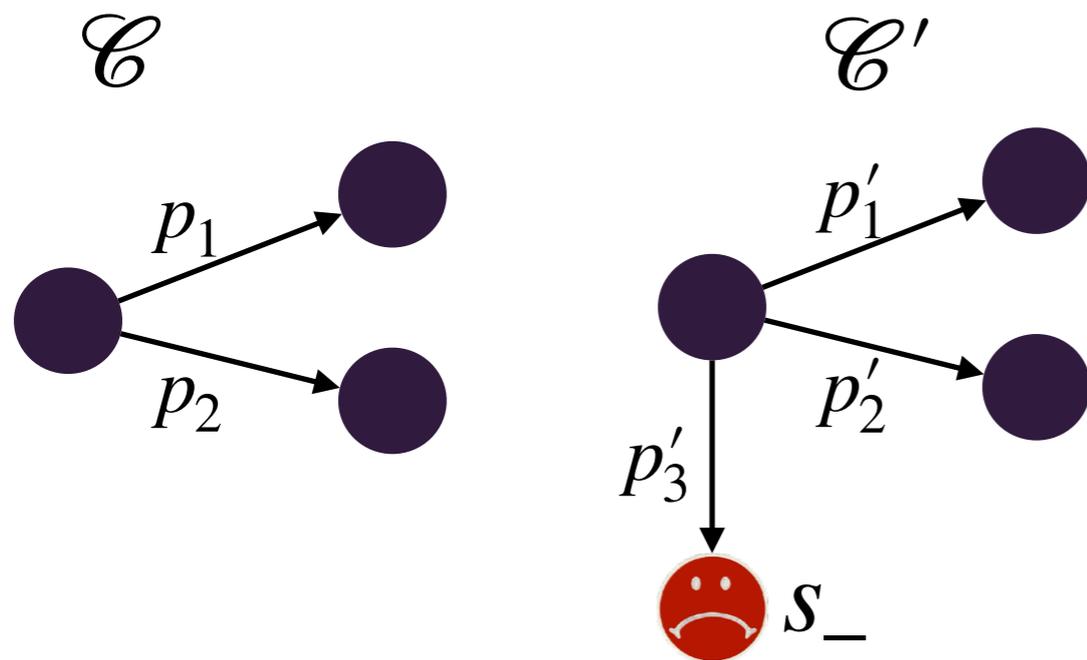
[Bar14] B. Barbot. *Acceleration for statistical model checking* (PhD thesis)

[BHP12] B. Barbot, S. Haddad, C. Picaronny. *Coupling and Importance Sampling for Statistical Model Checking* (TACAS'12)

# Importance sampling

## [KH51]

- ▶ Analyze a biased Markov chain  $\mathcal{C}'$



- ▶ Originally used for rare events

[KH51] H. Kahn, T. E. Harris. *Estimation of particle transmission by random sampling* (National Bureau of Standards applied mathematics series, 1951)

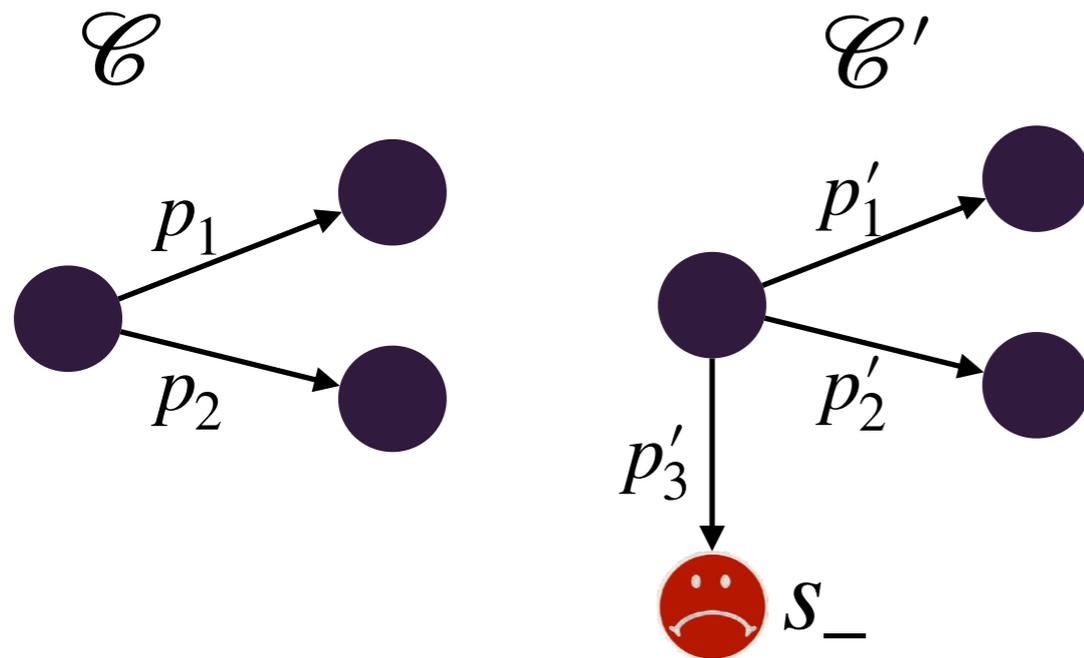
[Bar14] B. Barbot. *Acceleration for statistical model checking* (PhD thesis)

[BHP12] B. Barbot, S. Haddad, C. Picaronny. *Coupling and Importance Sampling for Statistical Model Checking* (TACAS'12)

# Importance sampling

## [KH51]

- ▶ Analyze a biased Markov chain  $\mathcal{C}'$



Correct the bias

$$\gamma(\rho) = \begin{cases} \frac{P(\rho)}{P'(\rho)} & \text{if } \rho \text{ ends in } \text{😊} \\ 0 & \text{otherwise} \end{cases}$$

- ▶ Originally used for rare events

[KH51] H. Kahn, T. E. Harris. Estimation of particle transmission by random sampling (National Bureau of Standards applied mathematics series, 1951)

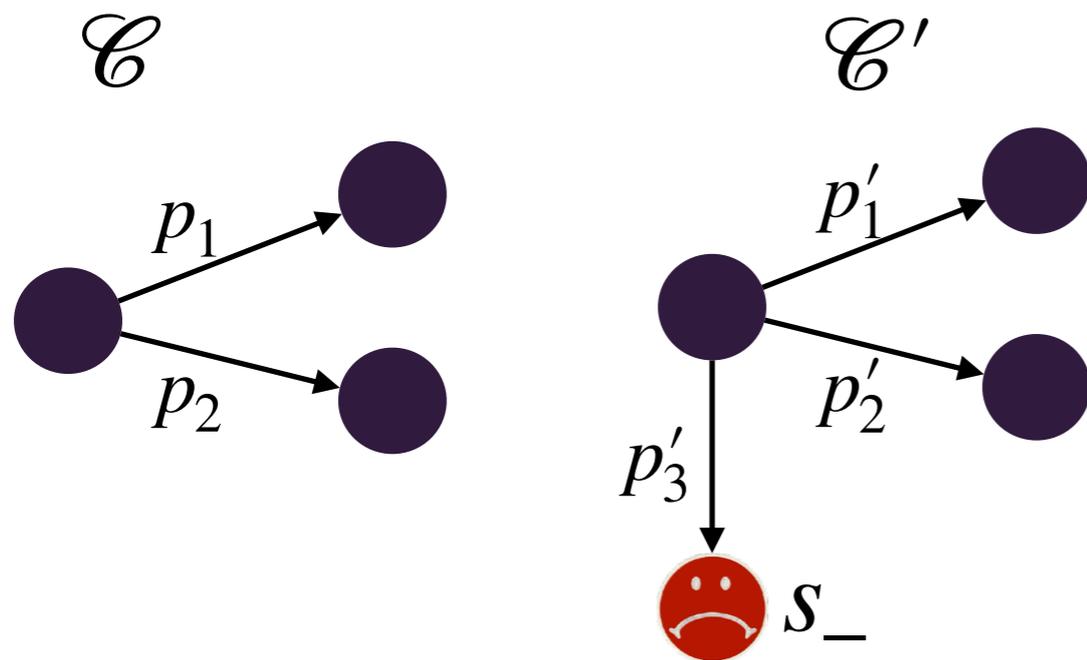
[Bar14] B. Barbot. Acceleration for statistical model checking (PhD thesis)

[BHP12] B. Barbot, S. Haddad, C. Picaronny. Coupling and Importance Sampling for Statistical Model Checking (TACAS'12)

# Importance sampling

## [KH51]

- ▶ Analyze a biased Markov chain  $\mathcal{C}'$



Correct the bias

$$\gamma(\rho) = \begin{cases} \frac{P(\rho)}{P'(\rho)} & \text{if } \rho \text{ ends in } \text{😊} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{P}_{\mathcal{C}}(\mathbf{F} \text{😊}) = \mathbb{E}_{\mathcal{C}'}(\gamma)$$

- ▶ Originally used for rare events

[KH51] H. Kahn, T. E. Harris. Estimation of particle transmission by random sampling (National Bureau of Standards applied mathematics series, 1951)

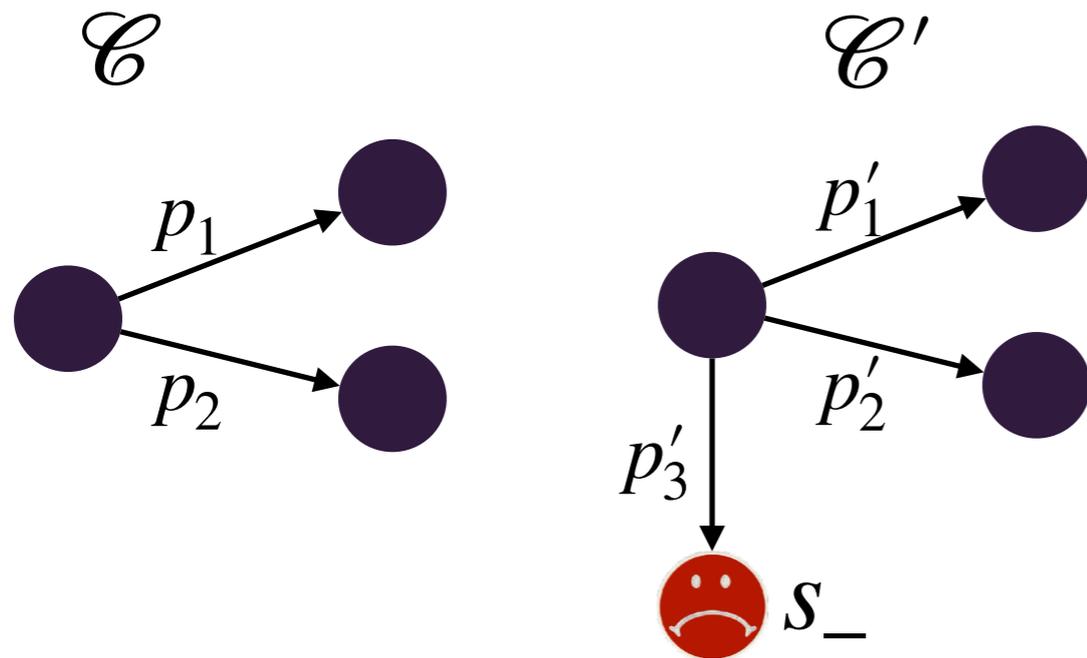
[Bar14] B. Barbot. Acceleration for statistical model checking (PhD thesis)

[BHP12] B. Barbot, S. Haddad, C. Picaronny. Coupling and Importance Sampling for Statistical Model Checking (TACAS'12)

# Importance sampling

## [KH51]

- ▶ Analyze a biased Markov chain  $\mathcal{C}'$



Correct the bias

$$\gamma(\rho) = \begin{cases} \frac{P(\rho)}{P'(\rho)} & \text{if } \rho \text{ ends in } \text{😊} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{P}_{\mathcal{C}}(\mathbf{F} \text{😊}) = \mathbb{E}_{\mathcal{C}'}(\gamma)$$

- ▶ Originally used for rare events

It is sufficient to compute  $\mathbb{E}_{\mathcal{C}'}(\gamma)$

[KH51] H. Kahn, T. E. Harris. Estimation of particle transmission by random sampling (National Bureau of Standards applied mathematics series, 1951)

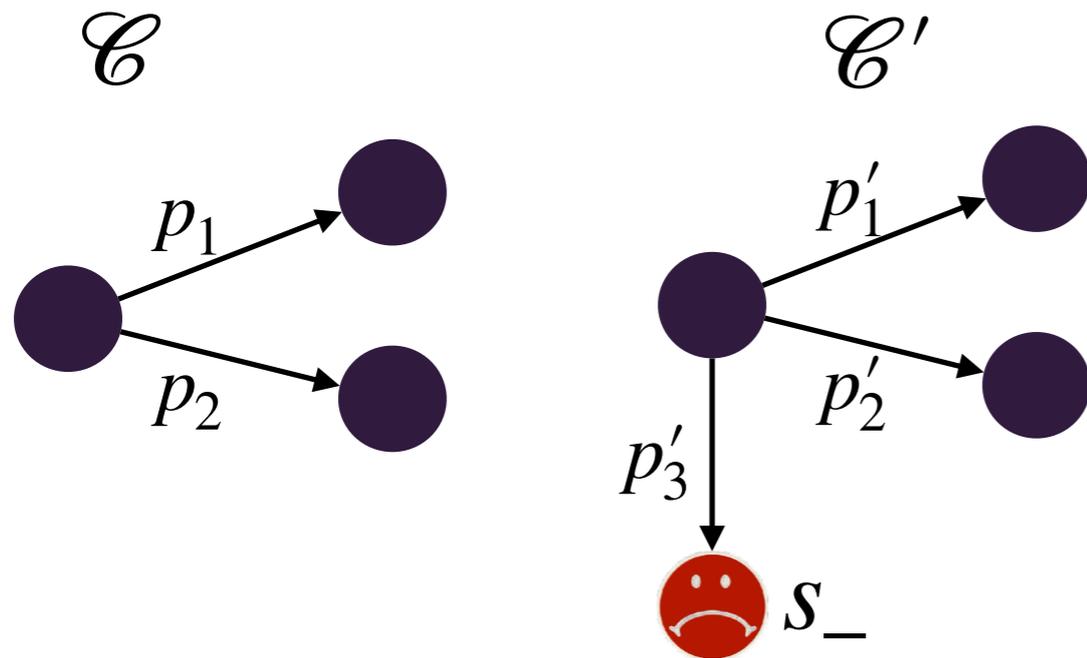
[Bar14] B. Barbot. Acceleration for statistical model checking (PhD thesis)

[BHP12] B. Barbot, S. Haddad, C. Picaronny. Coupling and Importance Sampling for Statistical Model Checking (TACAS'12)

# Importance sampling

## [KH51]

- ▶ Analyze a biased Markov chain  $\mathcal{C}'$



Correct the bias

$$\gamma(\rho) = \begin{cases} \frac{P(\rho)}{P'(\rho)} & \text{if } \rho \text{ ends in } \text{😊} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{P}_{\mathcal{C}}(\mathbf{F} \text{😊}) = \mathbb{E}_{\mathcal{C}'}(\gamma)$$

It is sufficient to compute  $\mathbb{E}_{\mathcal{C}'}(\gamma)$

- ▶ Originally used for rare events
- ▶ Setting giving statistical guarantees [BHP12, Bar14]

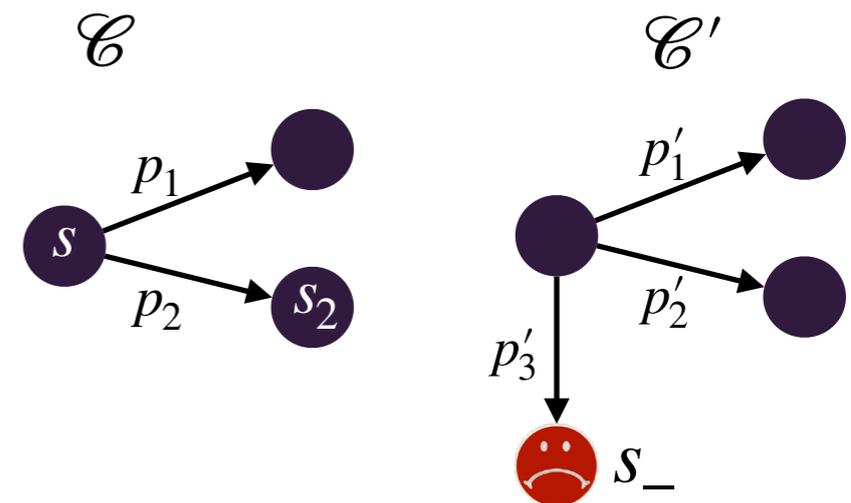
[KH51] H. Kahn, T. E. Harris. Estimation of particle transmission by random sampling (National Bureau of Standards applied mathematics series, 1951)

[Bar14] B. Barbot. Acceleration for statistical model checking (PhD thesis)

[BHP12] B. Barbot, S. Haddad, C. Picaronny. Coupling and Importance Sampling for Statistical Model Checking (TACAS'12)

# Properties of the biased Markov chain

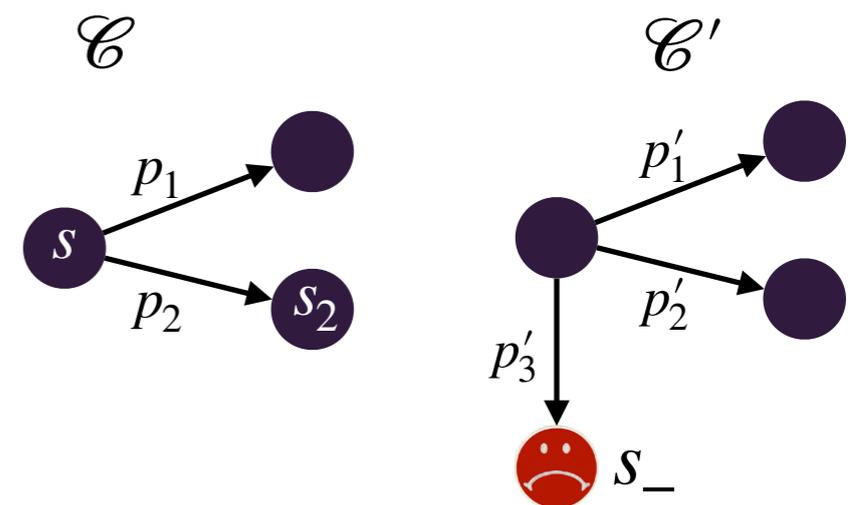
$$\mathbb{P}_{\mathcal{C}}(\mathbf{F} \text{ 😊}) = \mathbb{E}_{\mathcal{C}'}(\gamma)$$



# Properties of the biased Markov chain

$$\mathbb{P}_{\mathcal{C}}(\mathbf{F} \text{ 😊}) = \mathbb{E}_{\mathcal{C}'}(\gamma)$$

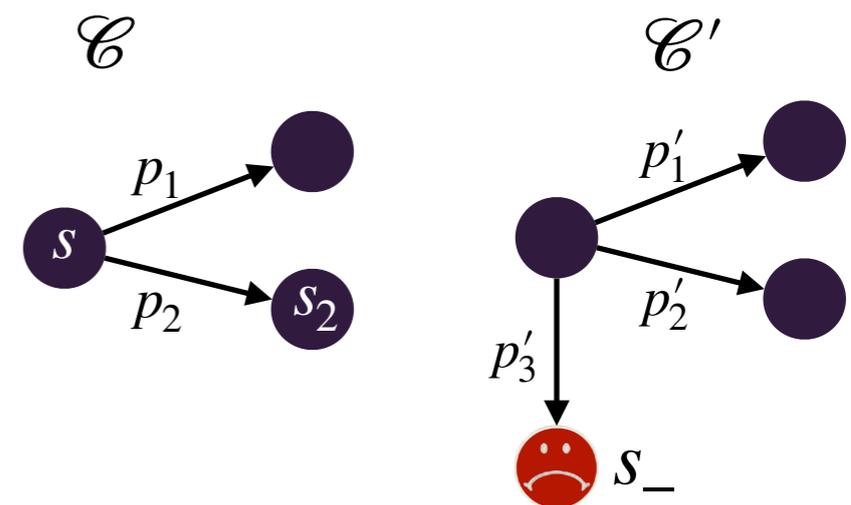
- ▶ The analysis of  $\mathcal{C}$  can be transferred to that of  $\mathcal{C}'$ , provided some conditions on  $\mathcal{C}'$



# Properties of the biased Markov chain

$$\mathbb{P}_{\mathcal{C}}(\mathbf{F} \text{ 😊}) = \mathbb{E}_{\mathcal{C}'}(\gamma)$$

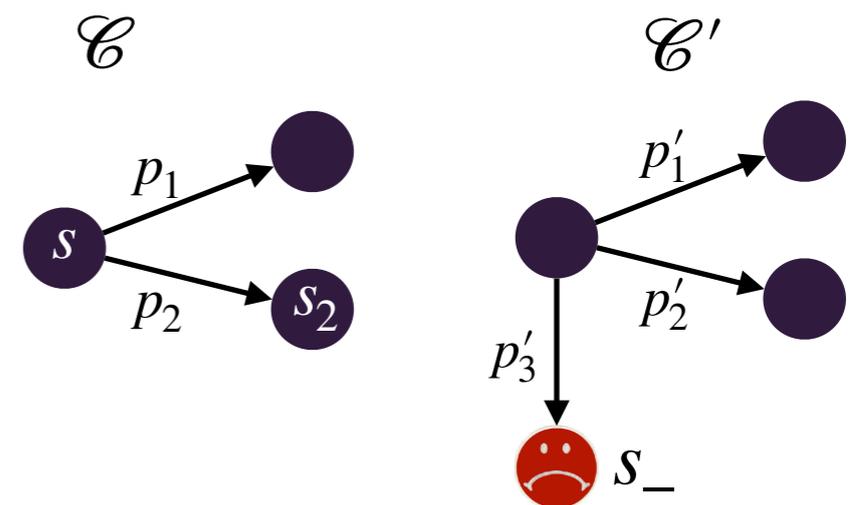
- ▶ The analysis of  $\mathcal{C}$  can be transferred to that of  $\mathcal{C}'$ , provided some conditions on  $\mathcal{C}'$ 
  - Decisiveness of  $\mathcal{C}'$  is required for both approx. and estim. methods



# Properties of the biased Markov chain

$$\mathbb{P}_{\mathcal{C}}(\mathbf{F} \text{ 😊}) = \mathbb{E}_{\mathcal{C}'}(\gamma)$$

- ▶ The analysis of  $\mathcal{C}$  can be transferred to that of  $\mathcal{C}'$ , provided some conditions on  $\mathcal{C}'$ 
  - Decisiveness of  $\mathcal{C}'$  is required for both approx. and estim. methods
  - Boundedness of  $\gamma$  is required as well

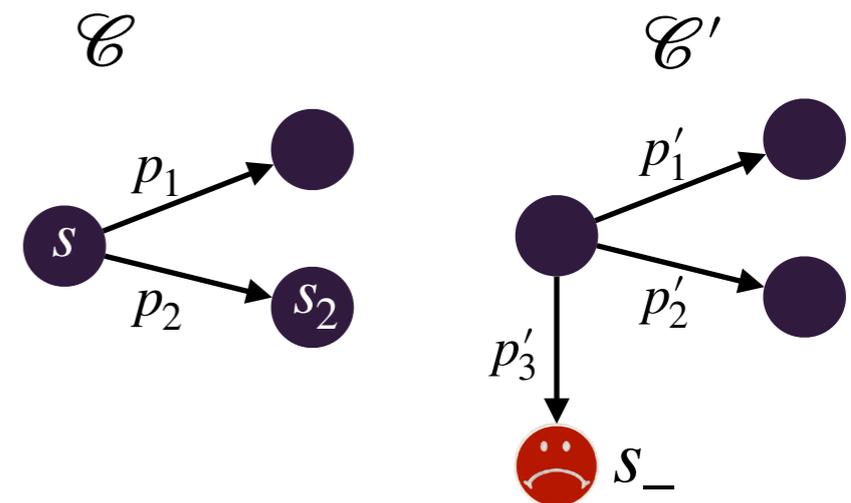


# Properties of the biased Markov chain

Define  $\mu(s)$  as  
 $\mathbb{P}_{\mathcal{C}}^s(\mathbf{F} \text{ 😊})$

$$\mathbb{P}_{\mathcal{C}}(\mathbf{F} \text{ 😊}) = \mathbb{E}_{\mathcal{C}'}(\gamma)$$

- ▶ The analysis of  $\mathcal{C}$  can be transferred to that of  $\mathcal{C}'$ , provided some conditions on  $\mathcal{C}'$ 
  - Decisiveness of  $\mathcal{C}'$  is required for both approx. and estim. methods
  - Boundedness of  $\gamma$  is required as well



# Properties of the biased Markov chain

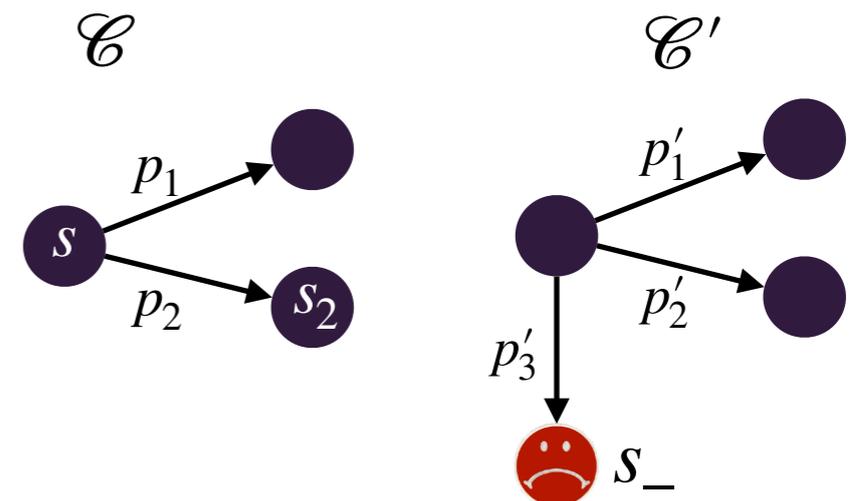
Define  $\mu(s)$  as  $\mathbb{P}_{\mathcal{C}}^s(\mathbf{F} \text{ 😊})$

$$\mathbb{P}_{\mathcal{C}}(\mathbf{F} \text{ 😊}) = \mathbb{E}_{\mathcal{C}'}(\gamma)$$

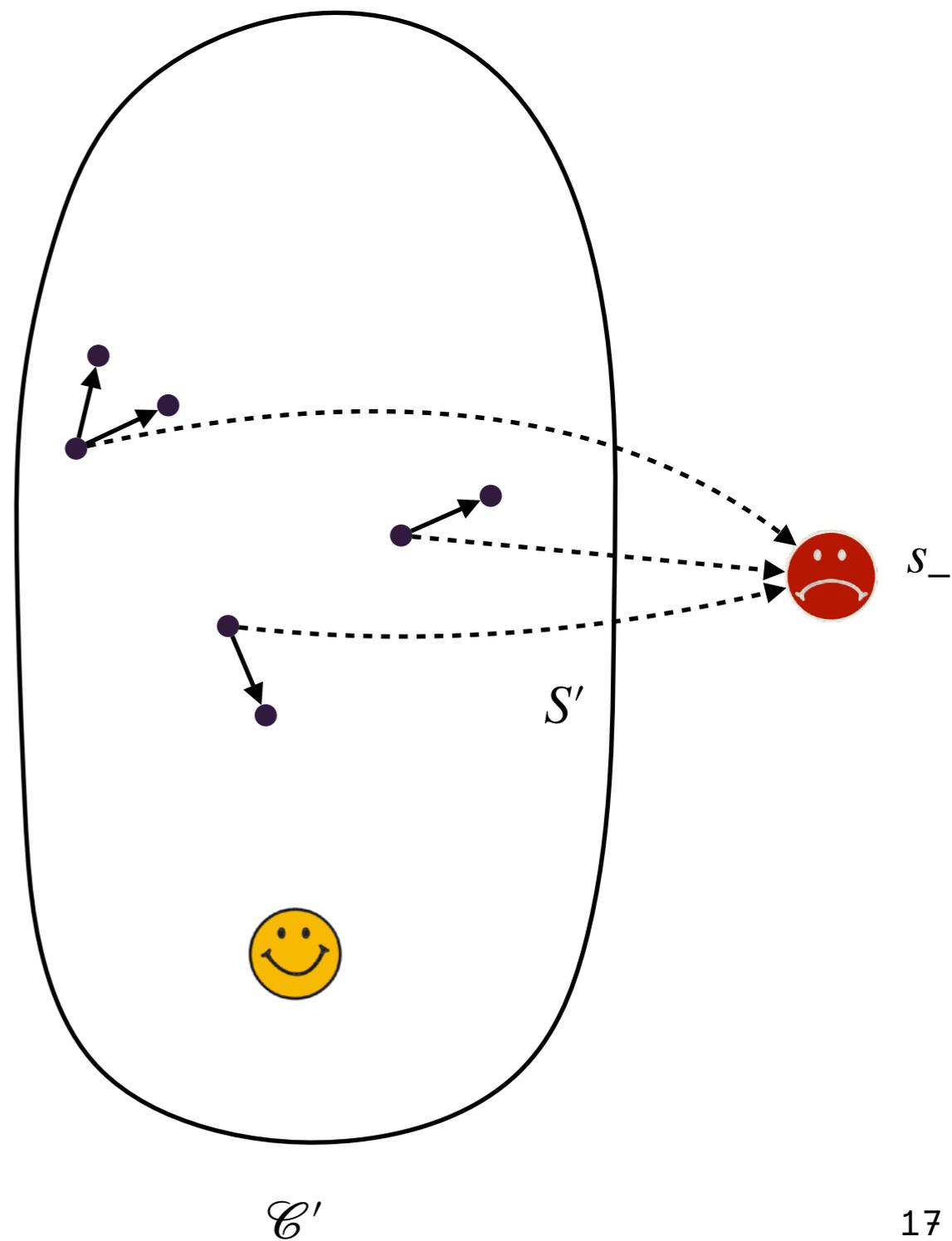
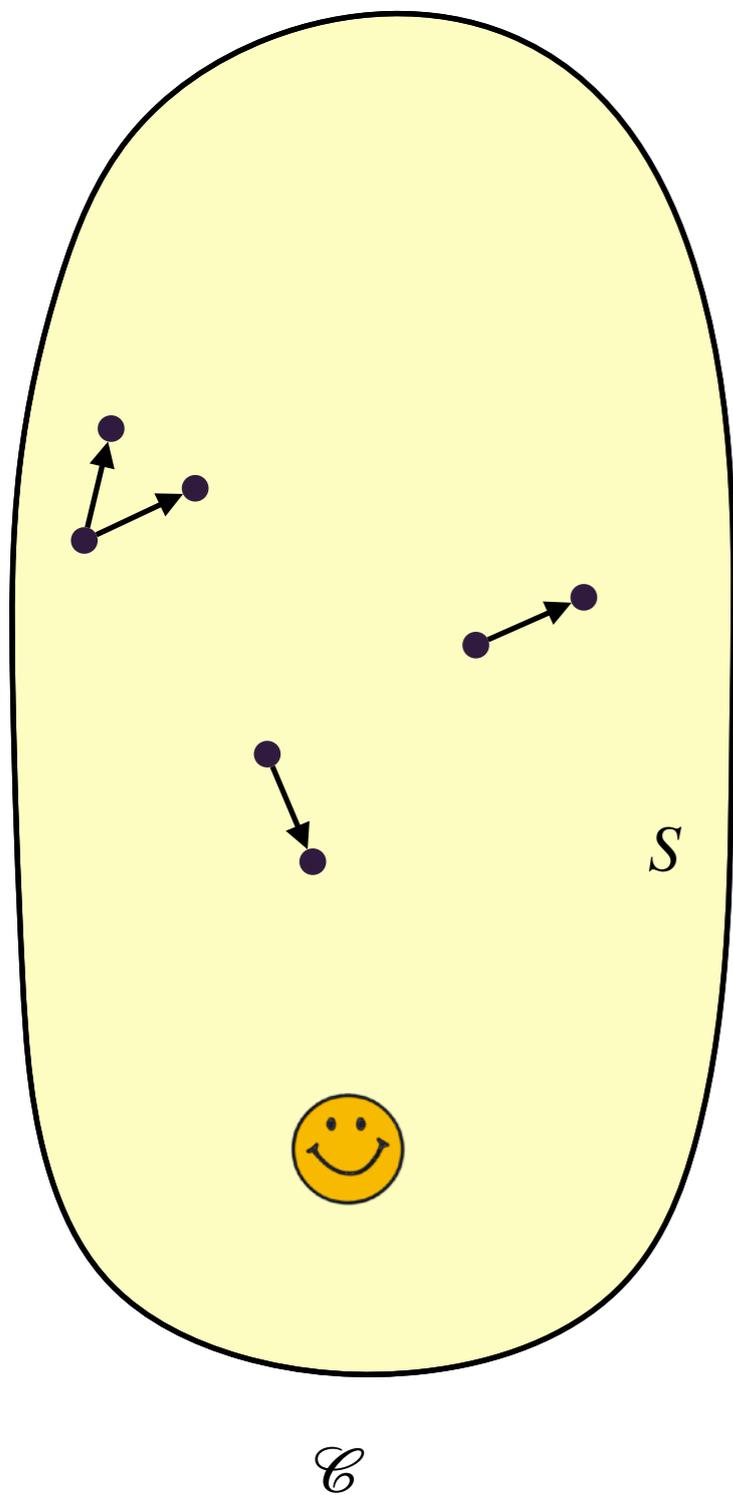
- ▶ The analysis of  $\mathcal{C}$  can be transferred to that of  $\mathcal{C}'$ , provided some conditions on  $\mathcal{C}'$ 
  - Decisiveness of  $\mathcal{C}'$  is required for both approx. and estim. methods
  - Boundedness of  $\gamma$  is required as well

There is a best choice:  $p'_i = \frac{\mu(s_i)}{\mu(s)} \cdot p_i$

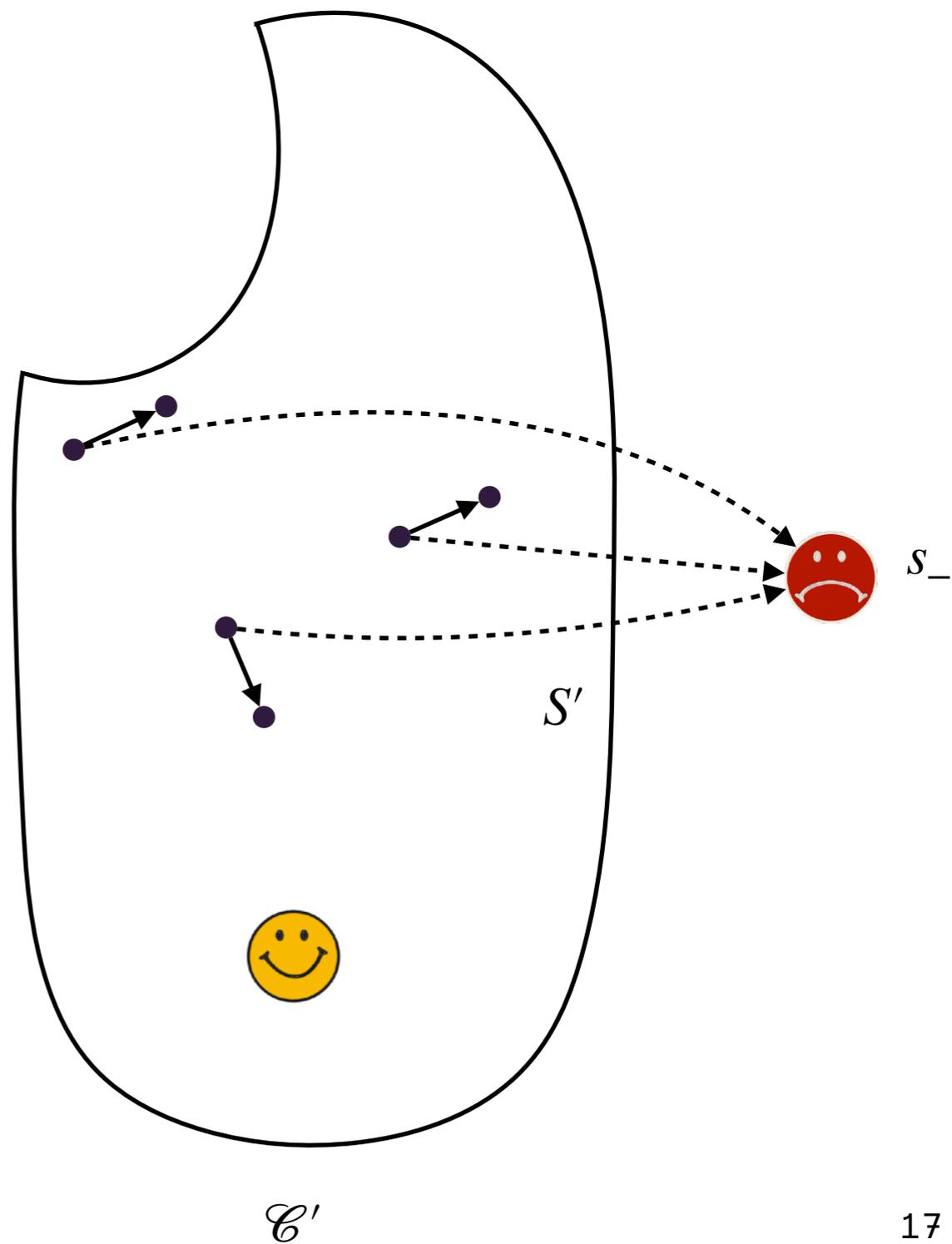
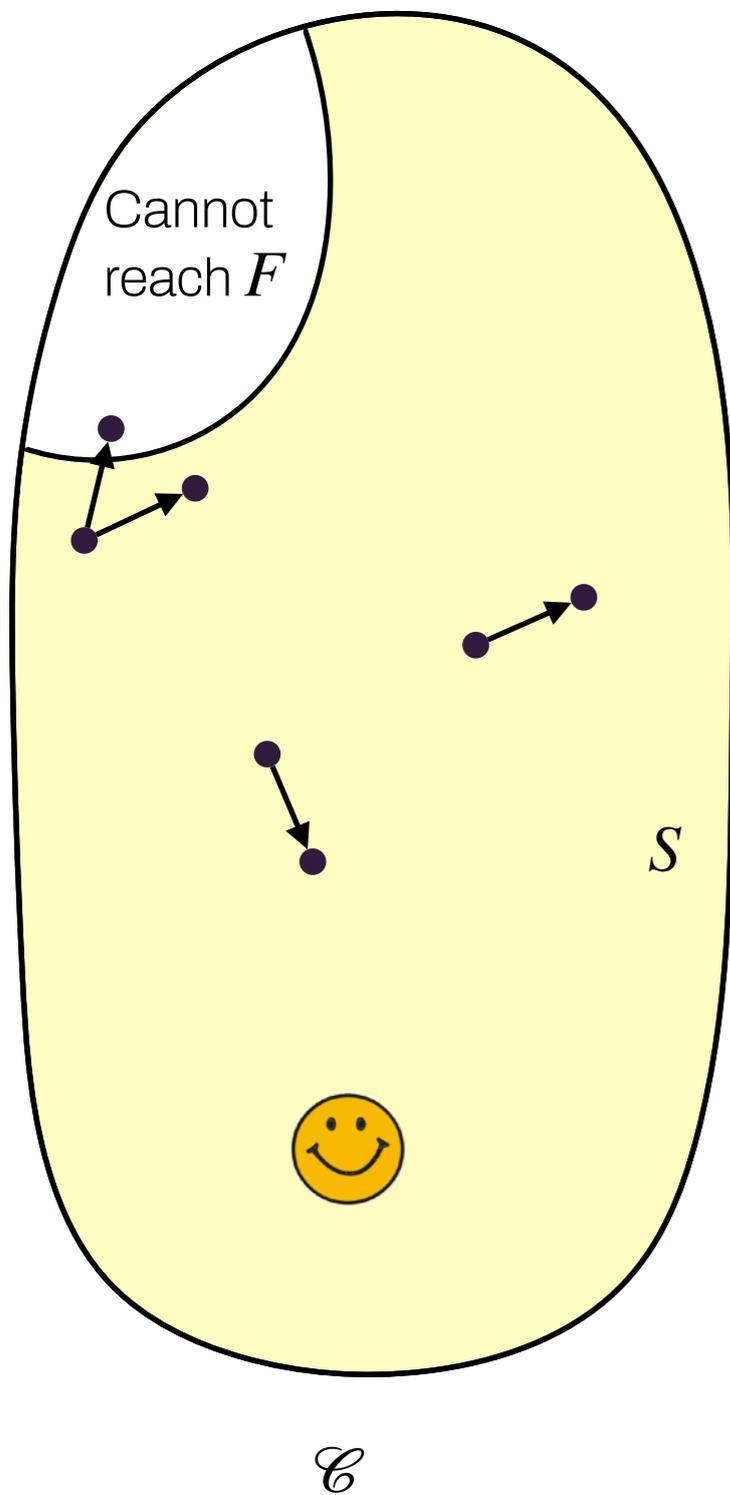
- ▶ The r.v. in  $\mathcal{C}'$  takes value  $\mu(s)$
- ▶ One needs to know  $\mu$ !



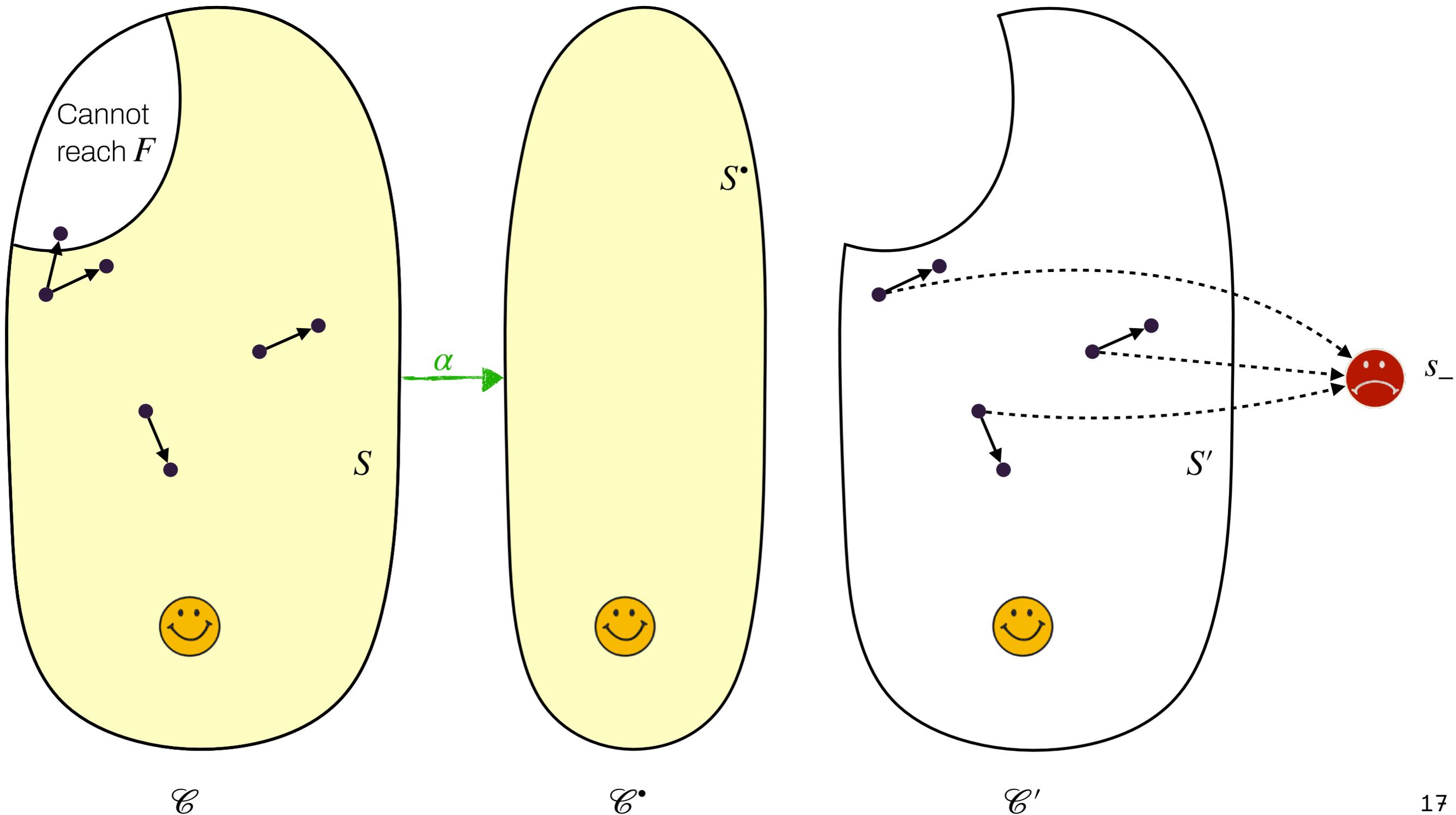
# Importance sampling via an abstraction



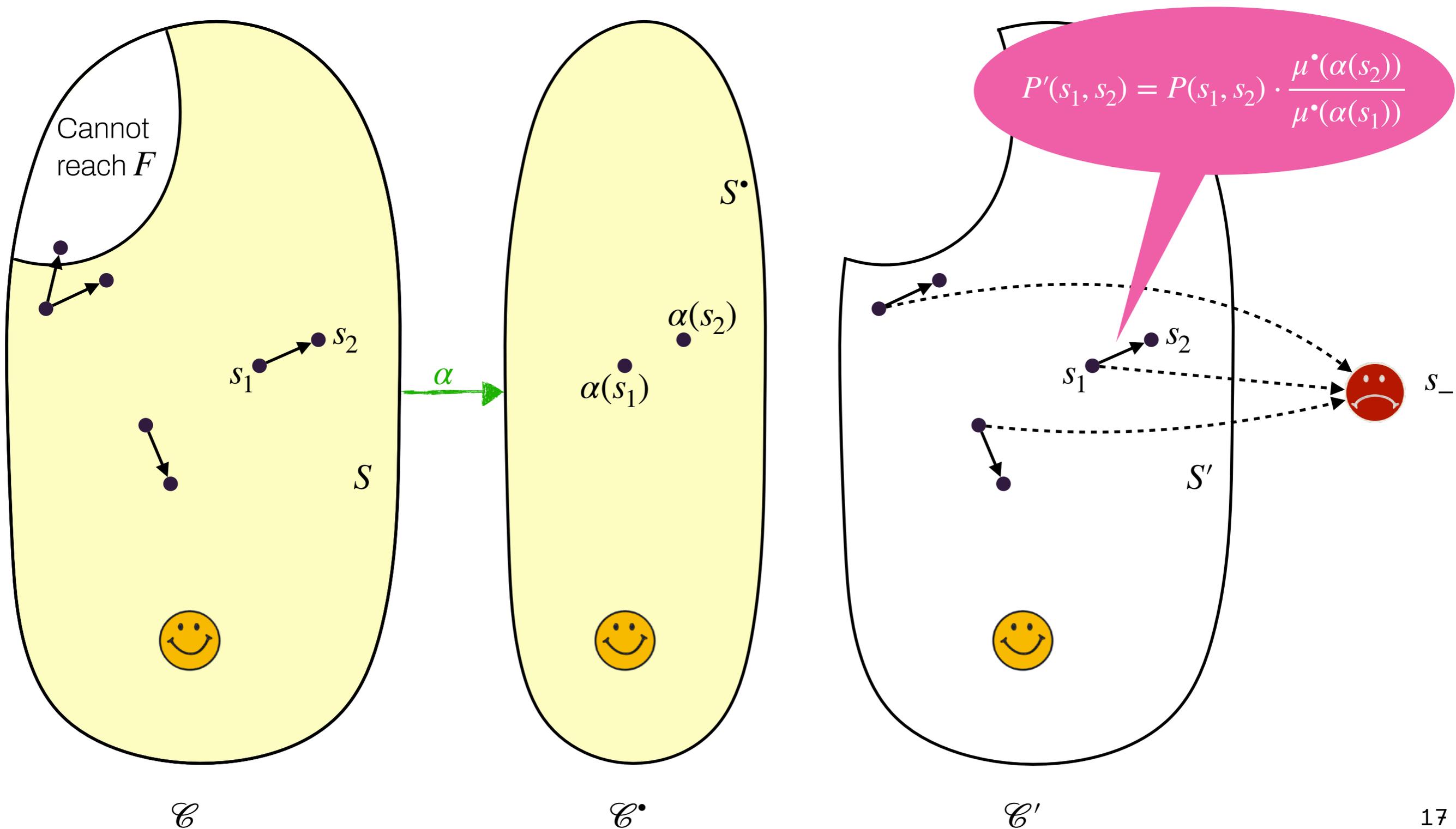
# Importance sampling via an abstraction



# Importance sampling via an abstraction

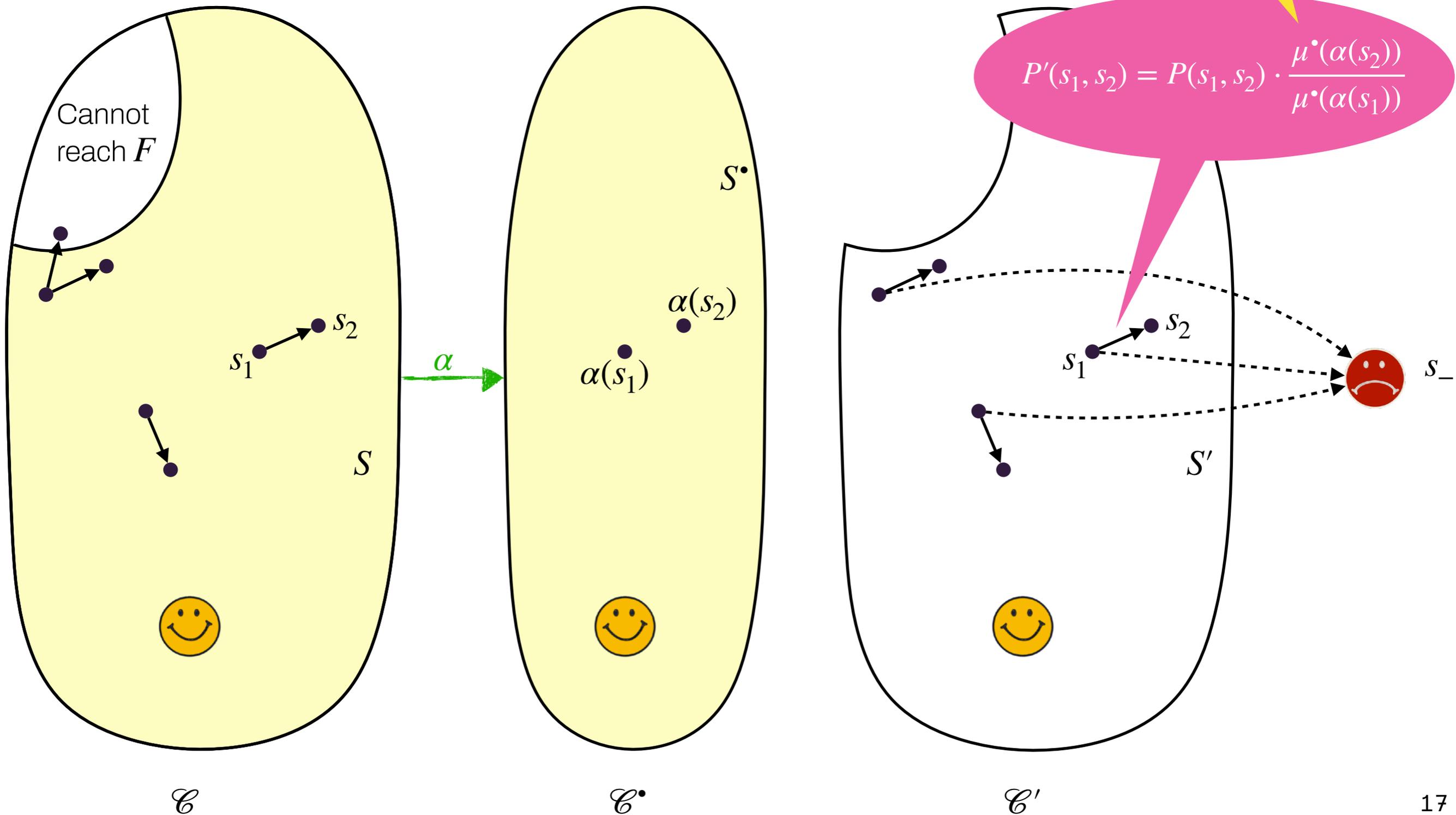


# Importance sampling via an abstraction

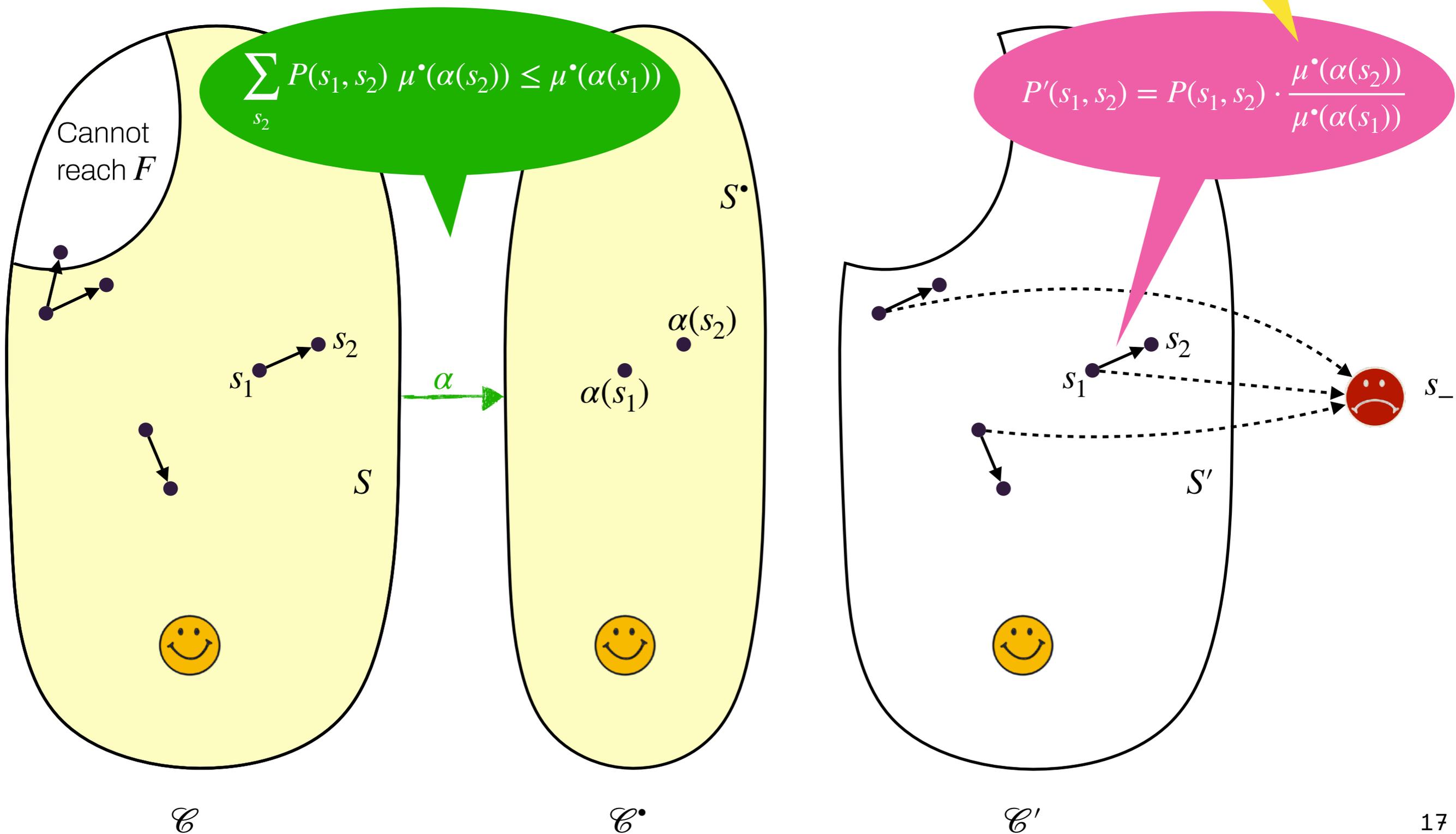


# Importance sampling via an abstraction

$\mu^\bullet$  is the probability to reach 😊 in  $\mathcal{C}^\bullet$

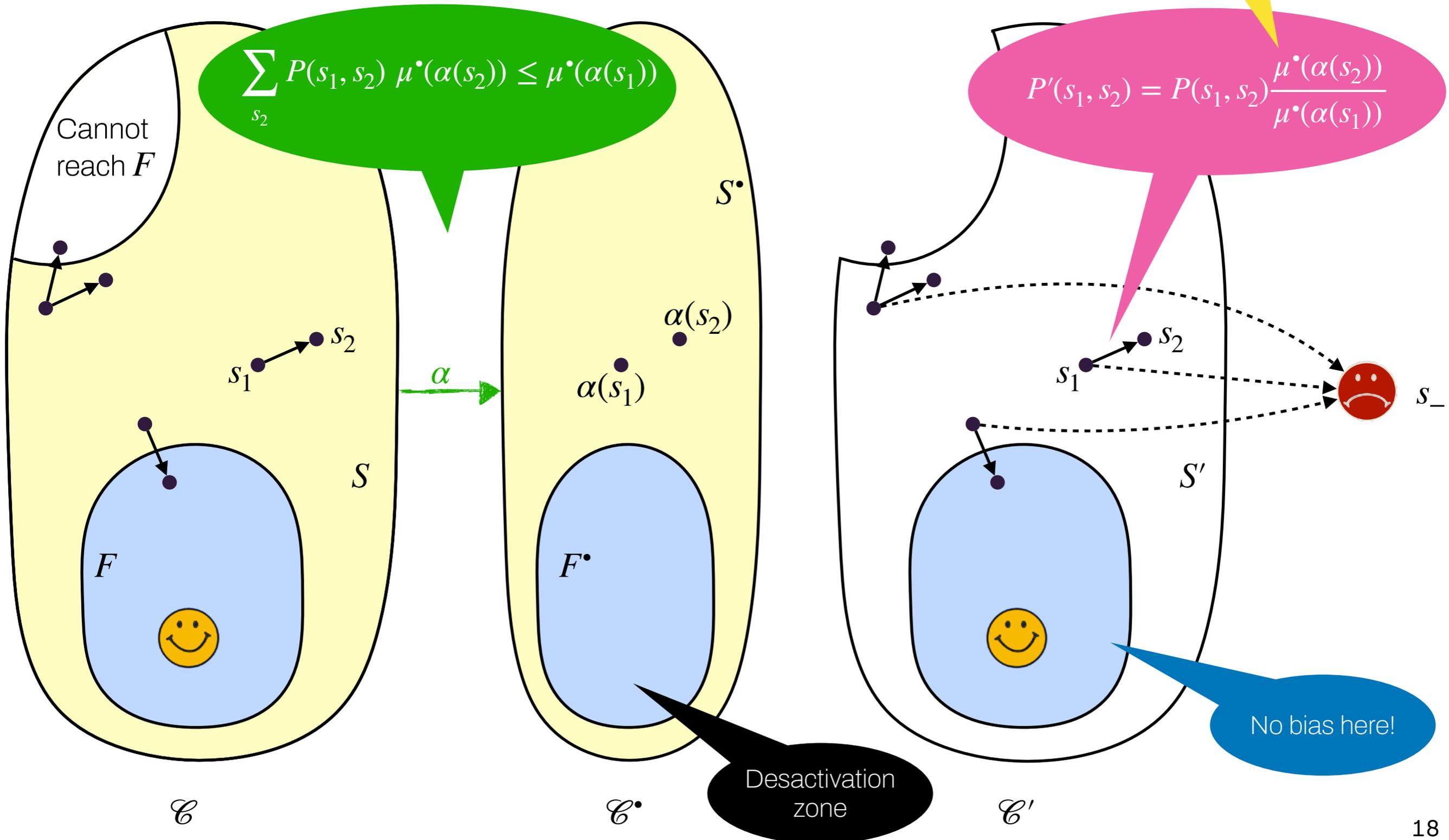


# Importance sampling via an abstraction

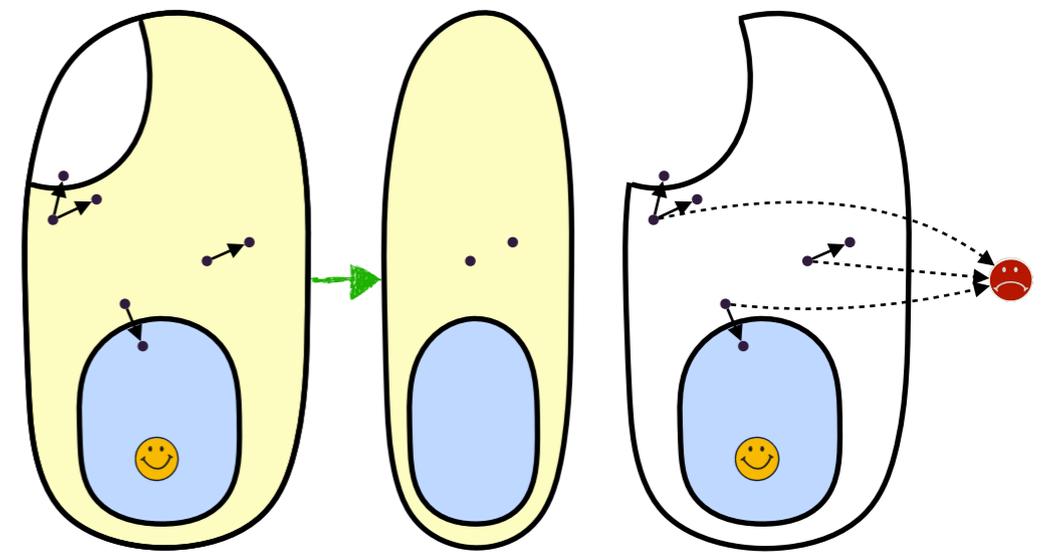


# Importance sampling via an abstraction

$\mu^\bullet$  is the probability to reach  $F^\bullet$  in  $\mathcal{C}^\bullet$



# Further properties of the biased Markov chain obtained via an abstraction

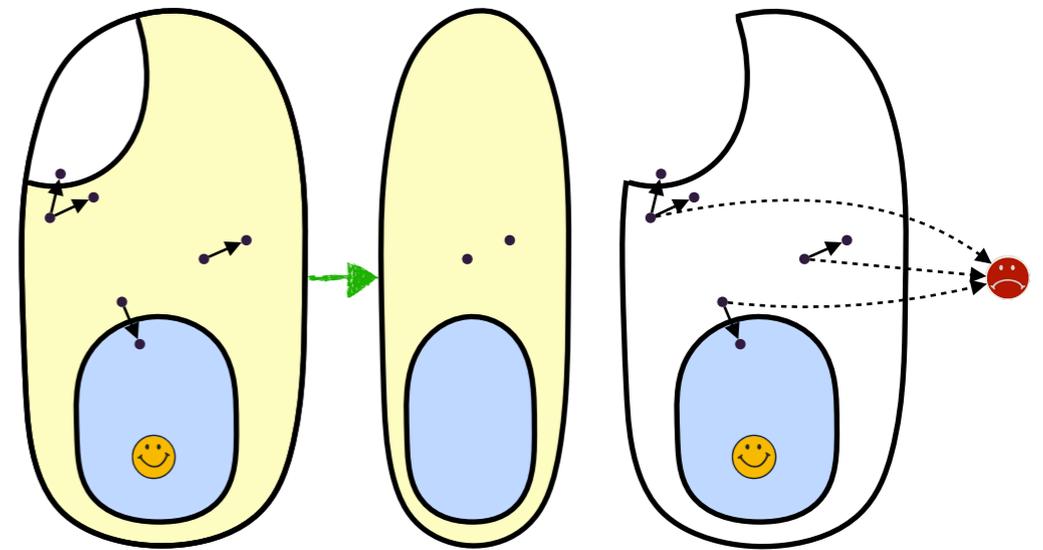


# Further properties of the biased Markov chain obtained via an abstraction

Property of the bias

$$\gamma(\rho) = \begin{cases} \mu^\bullet(\alpha(s_0)) & \text{if } \rho \text{ ends in } \text{😊} \\ 0 & \text{otherwise} \end{cases}$$

It is bivaluated, hence bounded

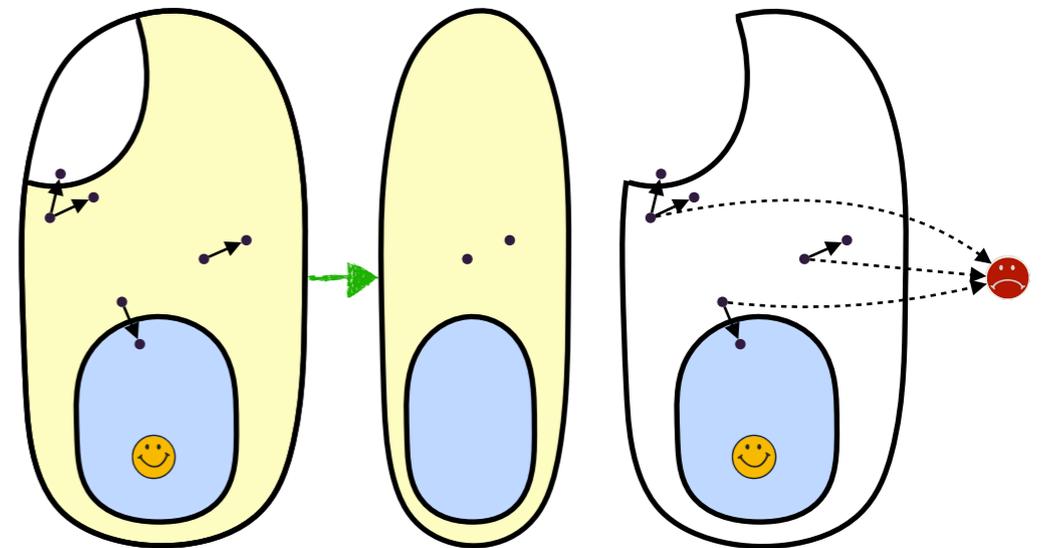


# Further properties of the biased Markov chain obtained via an abstraction

Property of the bias

$$\gamma(\rho) = \begin{cases} \mu^\bullet(\alpha(s_0)) & \text{if } \rho \text{ ends in } \text{😊} \\ 0 & \text{otherwise} \end{cases}$$

It is bivaluated, hence bounded



Theorem

If  $F$  is finite and for every  $0 \leq x < 1$ ,  $\{s \in S \mid \mu^\bullet(\alpha(s)) \geq x\}$  is finite, then  $\mathcal{C}'$  is decisive w.r.t. 😊.

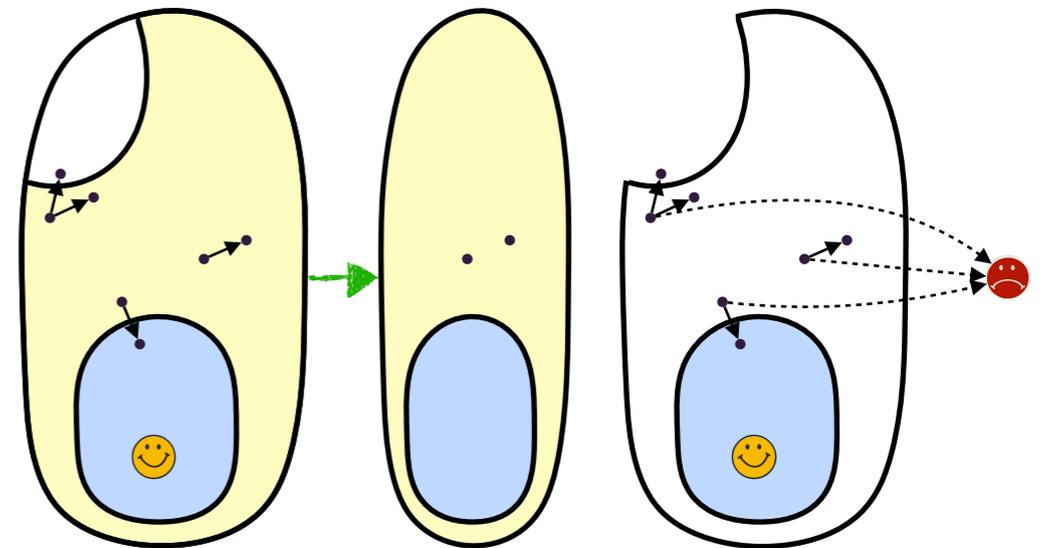
Proof using attractors, martingale theory

# Further properties of the biased Markov chain obtained via an abstraction

Property of the bias

$$\gamma(\rho) = \begin{cases} \mu^\bullet(\alpha(s_0)) & \text{if } \rho \text{ ends in } \text{😊} \\ 0 & \text{otherwise} \end{cases}$$

It is bivaluated, hence bounded



Theorem

If  $F$  is finite and for every  $0 \leq x < 1$ ,  $\{s \in S \mid \mu^\bullet(\alpha(s)) \geq x\}$  is finite, then  $\mathcal{C}'$  is decisive w.r.t. 😊.

Proof using attractors, martingale theory

- ▶ The analysis can be performed on  $\mathcal{C}'$ !

# And «concretely»?

# And «concretely»?

- ▶ Model = layered Markov chain (LMC)  $\mathcal{C}$ : there is a level function  $\lambda : \mathcal{S} \rightarrow \mathbb{N}$  s.t.
  - for every  $s_1 \rightarrow s_2$ ,  $\lambda(s_1) - \lambda(s_2) \leq 1$ , and
  - for every  $n$ ,  $\lambda^{-1}(n)$  is finite

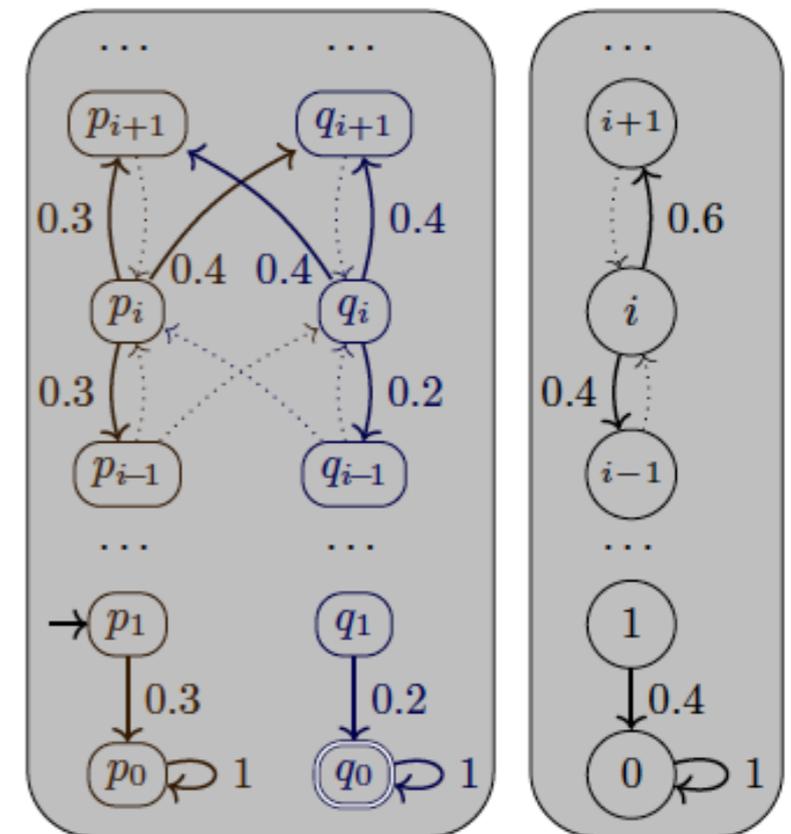
# And «concretely»?

- ▶ Model = layered Markov chain (LMC)  $\mathcal{C}$ : there is a level function  $\lambda : \mathcal{S} \rightarrow \mathbb{N}$  s.t.
  - for every  $s_1 \rightarrow s_2$ ,  $\lambda(s_1) - \lambda(s_2) \leq 1$ , and
  - for every  $n$ ,  $\lambda^{-1}(n)$  is finite
- ▶ Abstraction = random walk  $\mathcal{C}_p$  of parameter  $p > \frac{1}{2}$

# And «concretely»?

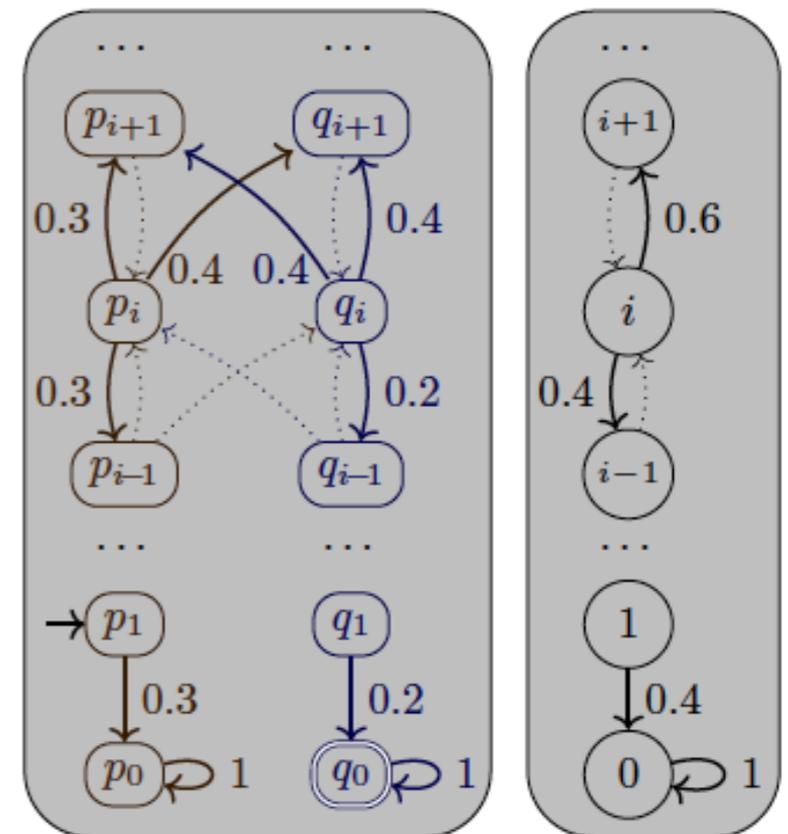
- ▶ Model = layered Markov chain (LMC)  $\mathcal{C}$ : there is a level function  $\lambda : \mathcal{S} \rightarrow \mathbb{N}$  s.t.
  - for every  $s_1 \rightarrow s_2$ ,  $\lambda(s_1) - \lambda(s_2) \leq 1$ , and
  - for every  $n$ ,  $\lambda^{-1}(n)$  is finite

- ▶ Abstraction = random walk  $\mathcal{C}_p$  of parameter  $p > \frac{1}{2}$



# And «concretely»?

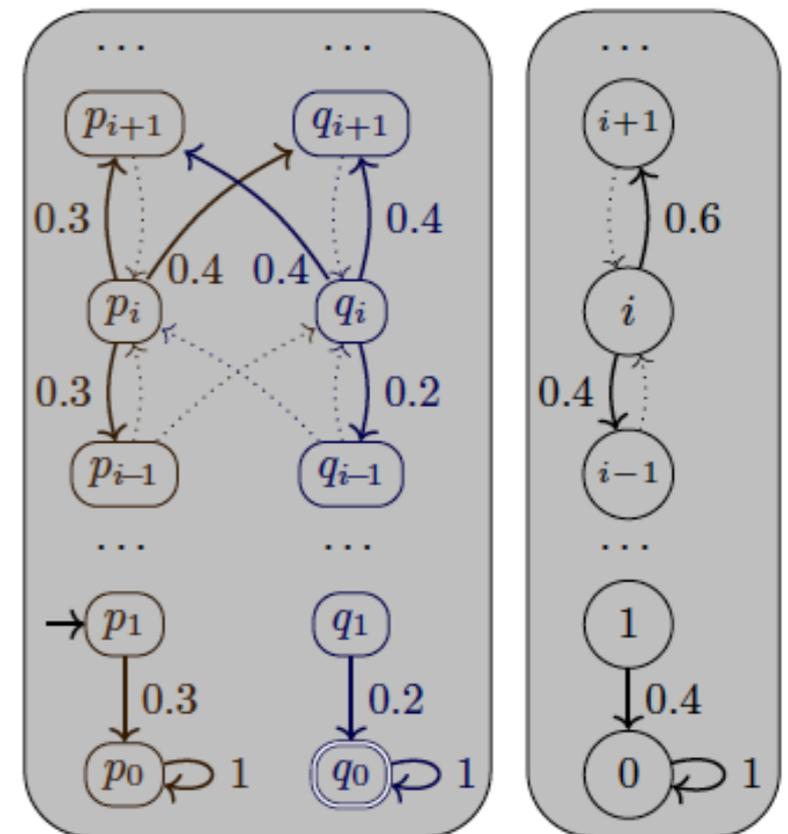
- ▶ Model = layered Markov chain (LMC)  $\mathcal{C}$ : there is a level function  $\lambda : \mathcal{S} \rightarrow \mathbb{N}$  s.t.
  - for every  $s_1 \rightarrow s_2$ ,  $\lambda(s_1) - \lambda(s_2) \leq 1$ , and
  - for every  $n$ ,  $\lambda^{-1}(n)$  is finite
- ▶ Abstraction = random walk  $\mathcal{C}_p$  of parameter  $p > \frac{1}{2}$



Only one condition needs to be satisfied...  
The **monotony** condition!

# And «concretely»?

- ▶ Model = layered Markov chain (LMC)  $\mathcal{C}$ : there is a level function  $\lambda : \mathcal{S} \rightarrow \mathbb{N}$  s.t.
  - for every  $s_1 \rightarrow s_2$ ,  $\lambda(s_1) - \lambda(s_2) \leq 1$ , and
  - for every  $n$ ,  $\lambda^{-1}(n)$  is finite
- ▶ Abstraction = random walk  $\mathcal{C}_p$  of parameter  $p > \frac{1}{2}$

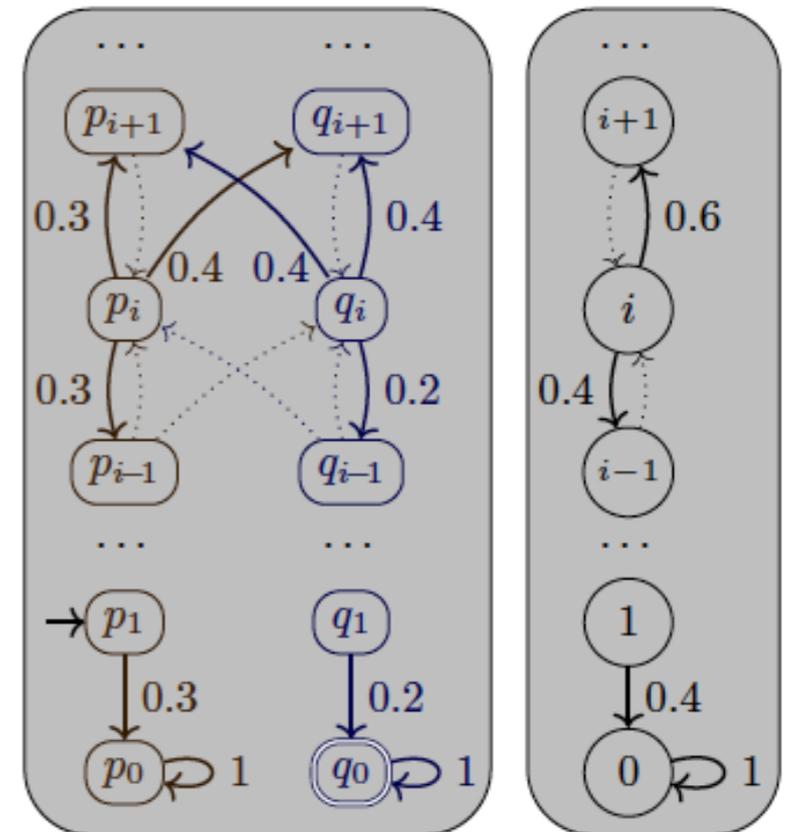


Only one condition needs to be satisfied...  
The **monotony** condition!

That will be ensured by a  
divergence property depending on  $p$ ,  
expressing a congestion  
phenomenon

# And «concretely»?

- ▶ Model = layered Markov chain (LMC)  $\mathcal{C}$ : there is a level function  $\lambda : \mathcal{S} \rightarrow \mathbb{N}$  s.t.
  - for every  $s_1 \rightarrow s_2$ ,  $\lambda(s_1) - \lambda(s_2) \leq 1$ , and
  - for every  $n$ ,  $\lambda^{-1}(n)$  is finite
- ▶ Abstraction = random walk  $\mathcal{C}_p$  of parameter  $p > \frac{1}{2}$

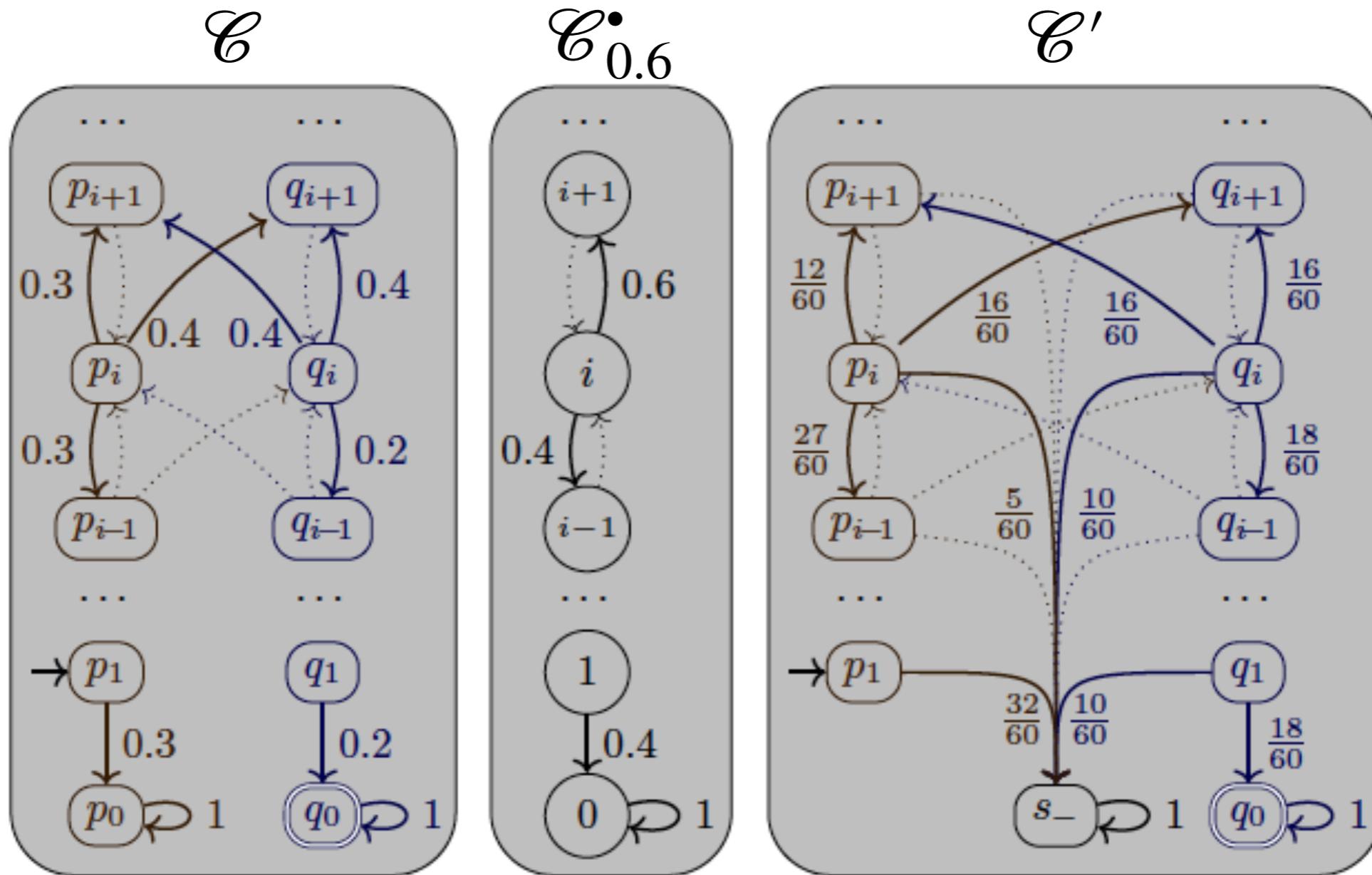


Only one condition needs to be satisfied...  
The **monotony** condition!

That will be ensured by a divergence property depending on  $p$ , expressing a congestion phenomenon

$p$  OK and  $\frac{1}{2} < p' < p \Rightarrow p'$  OK

# Example



$\mathcal{C}$  is not decisive

$\mathcal{C}'$  is decisive  
+ finite sampling time

# Implementation and experiments

<https://cosmos.lacl.fr/>

[BBDHP15] P. Ballarini, B. Barbot, M. Duflot, S. Haddad, N. Pekergin. *Hasl: A new approach for performance evaluation and model checking from concepts to experimentation* (Performance Evaluation)

# Implementation and experiments

- ▶ Implementation of the two approaches in tool **Cosmos**  
(development effort: Benoît Barbot)

# Implementation and experiments

- ▶ Implementation of the two approaches in tool **Cosmos** (development effort: Benoît Barbot)
- ▶ Application to probabilistic pushdown automata viewed as LMCs

# Implementation and experiments

- ▶ Implementation of the two approaches in tool **Cosmos** (development effort: Benoît Barbot)
- ▶ Application to probabilistic pushdown automata viewed as LMCs
- ▶ Methodology:

# Implementation and experiments

- ▶ Implementation of the two approaches in tool **Cosmos** (development effort: Benoît Barbot)
- ▶ Application to probabilistic pushdown automata viewed as LMCs
- ▶ Methodology:
  - If  $\mathcal{C}$  is decisive

# Implementation and experiments

- ▶ Implementation of the two approaches in tool **Cosmos** (development effort: Benoît Barbot)
- ▶ Application to probabilistic pushdown automata viewed as LMCs
- ▶ Methodology:
  - If  $\mathcal{C}$  is decisive
    - Apply Approx and Estim on  $\mathcal{C}$

# Implementation and experiments

- ▶ Implementation of the two approaches in tool **Cosmos** (development effort: Benoît Barbot)
- ▶ Application to probabilistic pushdown automata viewed as LMCs
- ▶ Methodology:
  - If  $\mathcal{C}$  is decisive
    - Apply Approx and Estim on  $\mathcal{C}$
  - If  $\mathcal{C}$  is «  $\hat{p}$ -divergent »

# Implementation and experiments

- ▶ Implementation of the two approaches in tool **Cosmos** (development effort: Benoît Barbot)
- ▶ Application to probabilistic pushdown automata viewed as LMCs
- ▶ Methodology:
  - If  $\mathcal{C}$  is decisive
    - Apply Approx and Estim on  $\mathcal{C}$
  - If  $\mathcal{C}$  is «  $\hat{p}$ -divergent »
    - Use the abstraction  $\mathcal{C}_p^\bullet$  with  $\frac{1}{2} < p < \hat{p}$

# Implementation and experiments

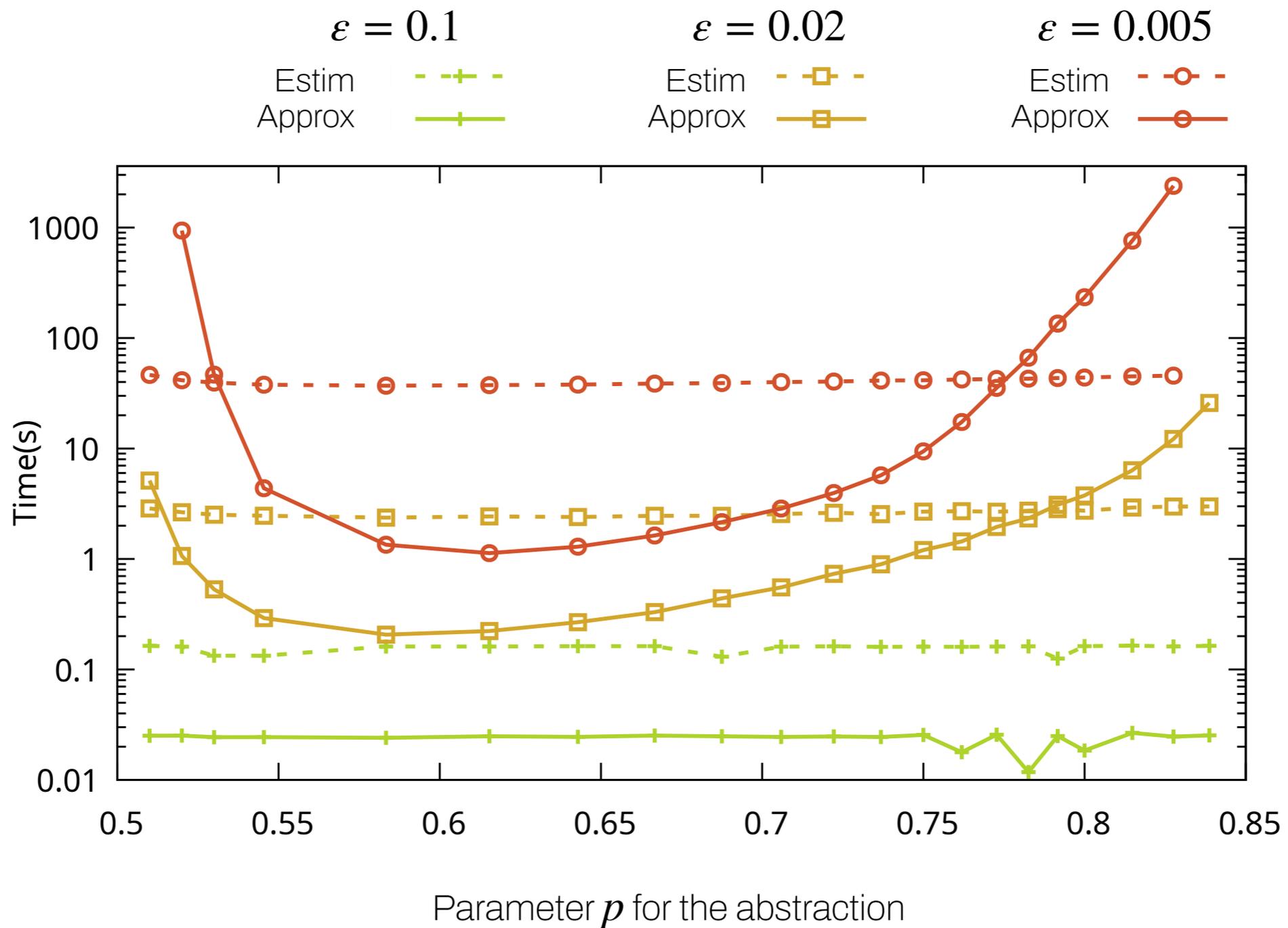
- ▶ Implementation of the two approaches in tool **Cosmos** (development effort: Benoît Barbot)
- ▶ Application to probabilistic pushdown automata viewed as LMCs
- ▶ Methodology:
  - If  $\mathcal{C}$  is decisive
    - Apply Approx and Estim on  $\mathcal{C}$
  - If  $\mathcal{C}$  is «  $\hat{p}$ -divergent »
    - Use the abstraction  $\mathcal{C}_p^\bullet$  with  $\frac{1}{2} < p < \hat{p}$
    - Apply Approx and Estim on corresponding  $\mathcal{C}'_p$  (computed on-the-fly)

# Implementation and experiments

- ▶ Implementation of the two approaches in tool **Cosmos** (development effort: Benoît Barbot)
- ▶ Application to probabilistic pushdown automata viewed as LMCs
- ▶ Methodology:
  - If  $\mathcal{C}$  is decisive
    - Apply Approx and Estim on  $\mathcal{C}$
  - If  $\mathcal{C}$  is «  $\hat{p}$ -divergent »
    - Use the abstraction  $\mathcal{C}_p^\bullet$  with  $\frac{1}{2} < p < \hat{p}$
    - Apply Approx and Estim on corresponding  $\mathcal{C}'_p$  (computed on-the-fly)

Note: in all experiments, the confidence is set to 99 %

# Examples of results



# Conclusion

# Conclusion

- ▶ Two approaches (numerical and statistical) for analysis of infinite Markov chains

# Conclusion

- ▶ Two approaches (numerical and statistical) for analysis of infinite Markov chains
  - Both require a **decisiveness** assumption!

# Conclusion

- ▶ Two approaches (numerical and statistical) for analysis of infinite Markov chains
  - Both require a **decisiveness** assumption!
- ▶ Use of **importance sampling** to handle some non-decisive Markov chains

# Conclusion

- ▶ Two approaches (numerical and statistical) for analysis of infinite Markov chains
  - Both require a **decisiveness** assumption!
- ▶ Use of **importance sampling** to handle some non-decisive Markov chains
  - Original application of the importance sampling idea

# Conclusion

- ▶ Two approaches (numerical and statistical) for analysis of infinite Markov chains
  - Both require a **decisiveness** assumption!
- ▶ Use of **importance sampling** to handle some non-decisive Markov chains
  - Original application of the importance sampling idea
  - Both approaches can be applied to the biased Markov chains (conditions for correctness are given)

# Conclusion

- ▶ Two approaches (numerical and statistical) for analysis of infinite Markov chains
  - Both require a **decisiveness** assumption!
- ▶ Use of **importance sampling** to handle some non-decisive Markov chains
  - Original application of the importance sampling idea
  - Both approaches can be applied to the biased Markov chains (conditions for correctness are given)
  - A general low-level model (LMC) + application to prob. pushdown automata

# Conclusion

- ▶ Two approaches (numerical and statistical) for analysis of infinite Markov chains
  - Both require a **decisiveness** assumption!
- ▶ Use of **importance sampling** to handle some non-decisive Markov chains
  - Original application of the importance sampling idea
  - Both approaches can be applied to the biased Markov chains (conditions for correctness are given)
  - A general low-level model (LMC) + application to prob. pushdown automata

Some more classes to be applied?

# Conclusion

- ▶ Two approaches (numerical and statistical) for analysis of infinite Markov chains
  - Both require a **decisiveness** assumption!
- ▶ Use of **importance sampling** to handle some non-decisive Markov chains
  - Original application of the importance sampling idea
  - Both approaches can be applied to the biased Markov chains (conditions for correctness are given)
  - A general low-level model (LMC) + application to prob. pushdown automata
- ▶ Interesting **empirical results**

Some more classes to be applied?

# Conclusion

- ▶ Two approaches (numerical and statistical) for analysis of infinite Markov chains
  - Both require a **decisiveness** assumption!
- ▶ Use of **importance sampling** to handle some non-decisive Markov chains
  - Original application of the importance sampling idea
  - Both approaches can be applied to the biased Markov chains (conditions for correctness are given)
  - A general low-level model (LMC) + application to prob. pushdown automata
- ▶ Interesting **empirical results**
  - Acceleration of the verification of decisive Markov chains in some cases?

Some more classes to be applied?

# Conclusion

- ▶ Two approaches (numerical and statistical) for analysis of infinite Markov chains
  - Both require a **decisiveness** assumption!
- ▶ Use of **importance sampling** to handle some non-decisive Markov chains
  - Original application of the importance sampling idea
  - Both approaches can be applied to the biased Markov chains (conditions for correctness are given)
  - A general low-level model (LMC) + application to prob. pushdown automata
- ▶ Interesting **empirical results**
  - Acceleration of the verification of decisive Markov chains in some cases?
  - Existence of a « best  $p$  »?

Some more classes to be applied?

# Conclusion

- ▶ Two approaches (numerical and statistical) for analysis of infinite Markov chains
  - Both require a **decisiveness** assumption!
- ▶ Use of **importance sampling** to handle some non-decisive Markov chains
  - Original application of the importance sampling idea
  - Both approaches can be applied to the biased Markov chains (conditions for correctness are given)
  - A general low-level model (LMC) + application to prob. pushdown automata
- ▶ Interesting **empirical results**
  - Acceleration of the verification of decisive Markov chains in some cases?
  - Existence of a « best  $p$  »?

Some more classes to be applied?

Any theoretical justification for that?