



Laboratoire
Méthodes
Formelles

université
PARIS-SACLAY



école
normale
supérieure
paris-saclay

Beyond Decisiveness of Infinite Markov Chains

Benoît Barbot

LACL
Université Paris-Est Créteil
France

Patricia Bouyer

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

Serge Haddad

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

Work partly supported by ANR projects MAVeriQ and BisoUS

Purpose of this work

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

Purpose of this work

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

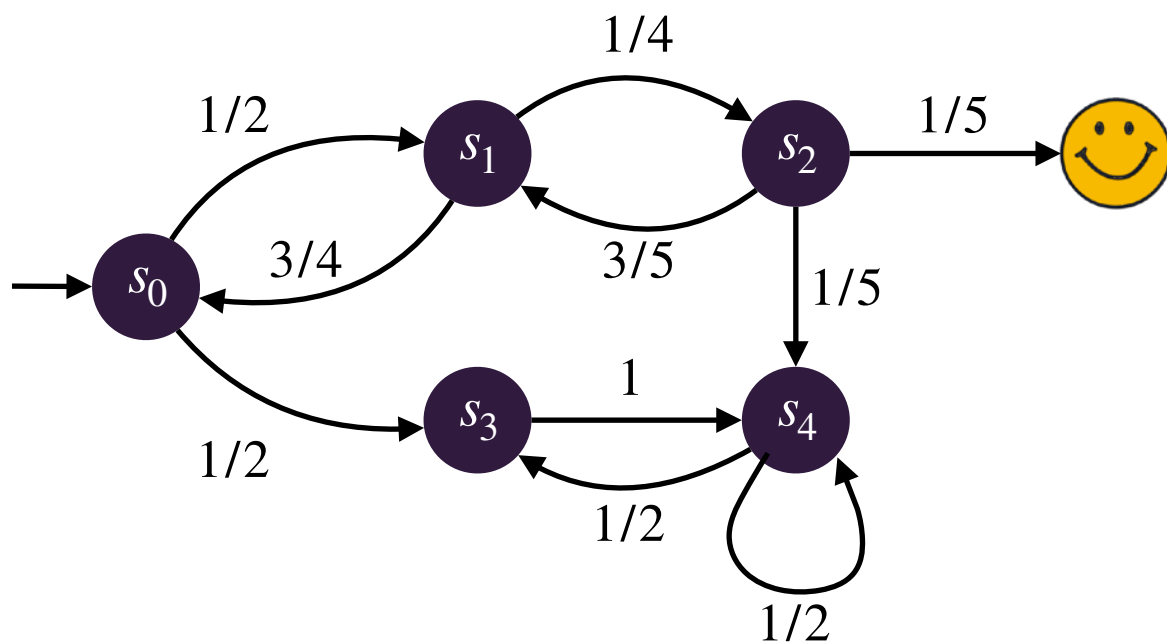
Discrete-time Markov chain (DTMC)

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

Discrete-time Markov chains

Discrete-time Markov chain (DTMC)

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

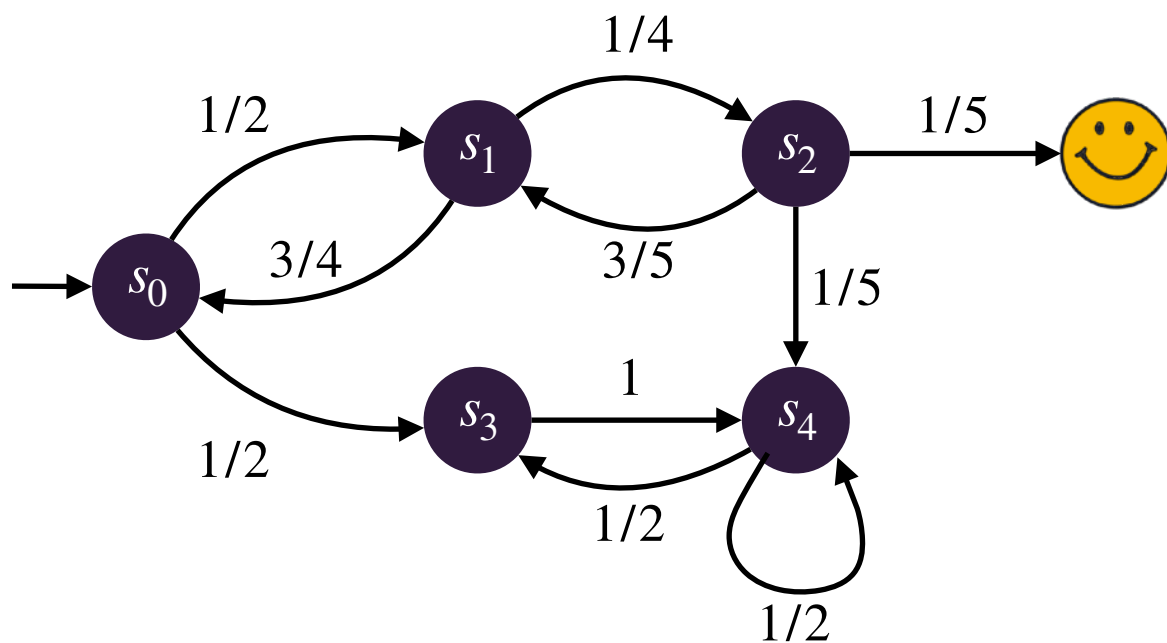


Finite Markov chain

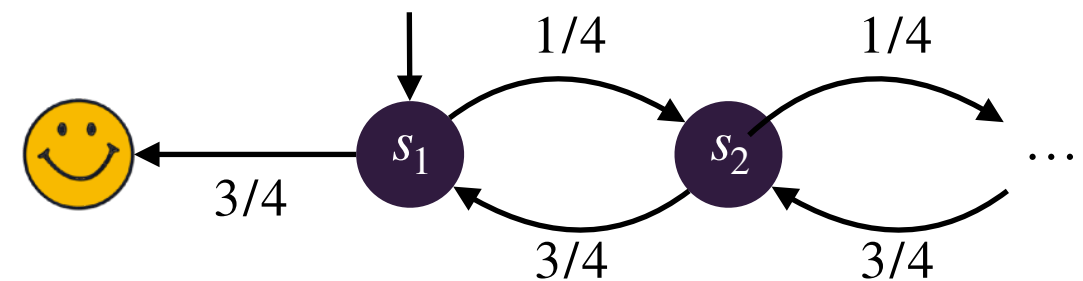
Discrete-time Markov chains

Discrete-time Markov chain (DTMC)

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



Finite Markov chain



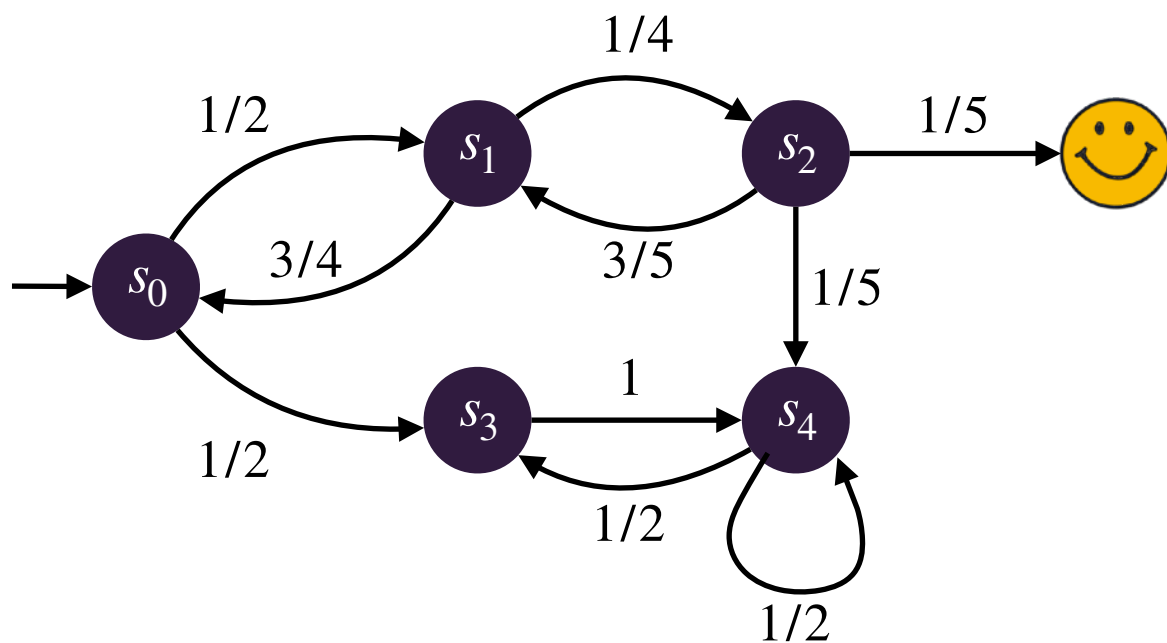
Denumerable Markov chain
(random walk of parameter 1/4)

Discrete-time Markov chains

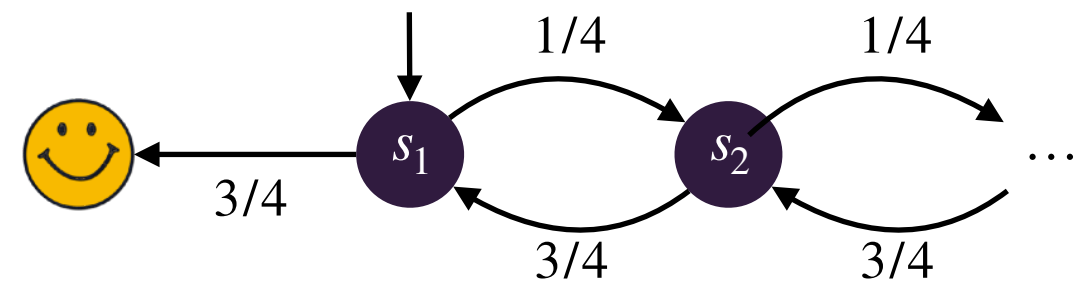
Discrete-time Markov chain (DTMC)

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

+ effectivity conditions...



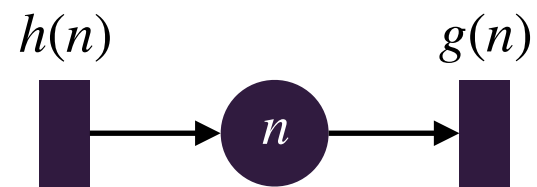
Finite Markov chain



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

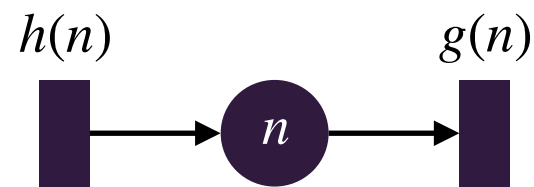
High-level models for (infinite) Markov chains

- ▶ Queues



High-level models for (infinite) Markov chains

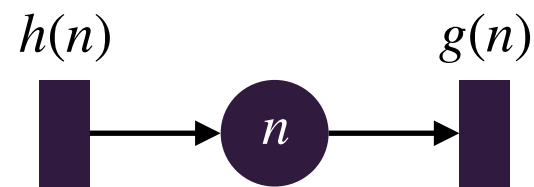
- Queues



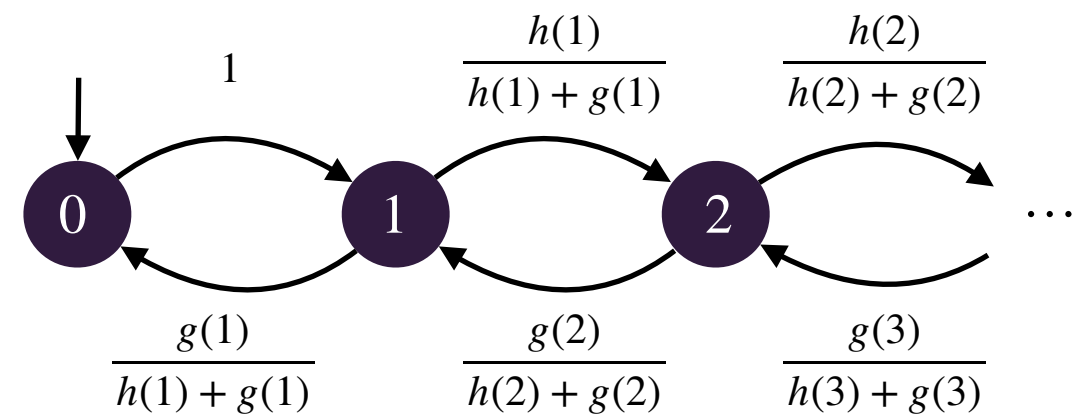
n is the
number of clients
(tokens)

High-level models for (infinite) Markov chains

► Queues

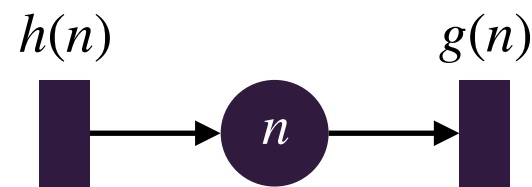


n is the number of clients (tokens)

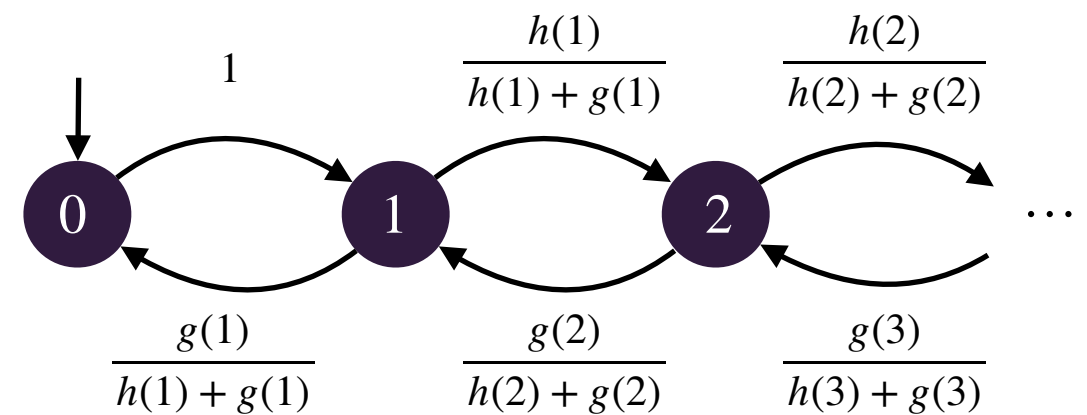


High-level models for (infinite) Markov chains

► Queues



n is the number of clients (tokens)

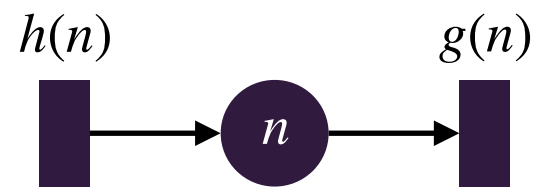


► Probabilistic pushdown automata

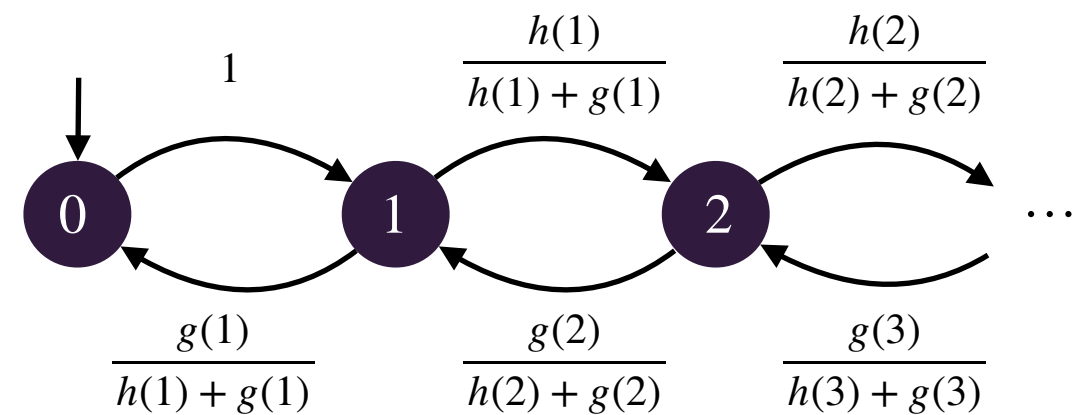
$$\begin{array}{lll} A \xrightarrow{1} C & A \xrightarrow{n} BB & B \xrightarrow{5} \varepsilon \\ B \xrightarrow{n} AA & C \xrightarrow{1} C & \end{array}$$

High-level models for (infinite) Markov chains

► Queues



n is the number of clients (tokens)



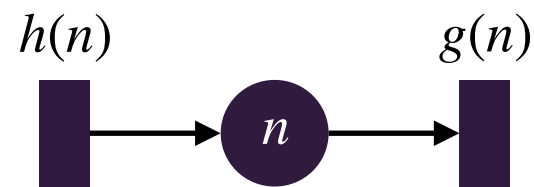
► Probabilistic pushdown automata

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

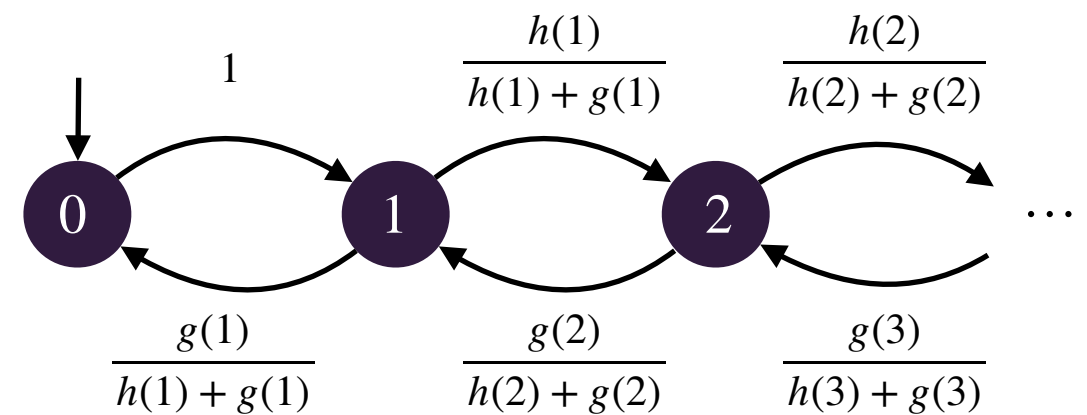
n is the height of the stack

High-level models for (infinite) Markov chains

► Queues



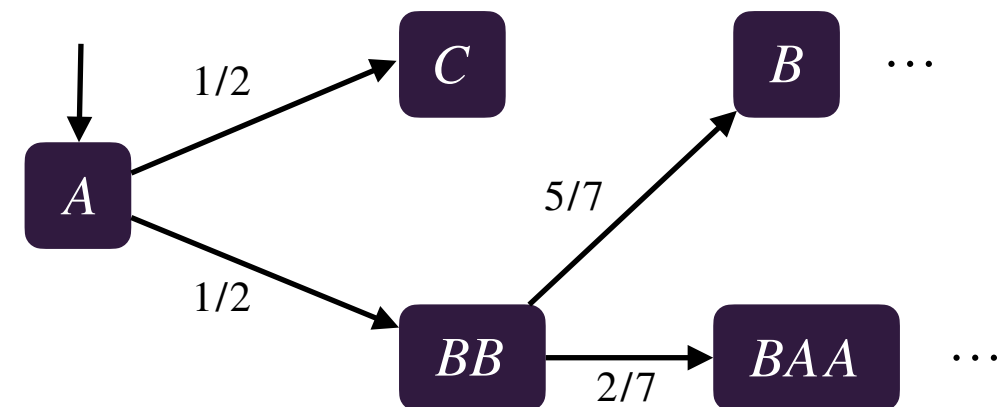
n is the number of clients (tokens)



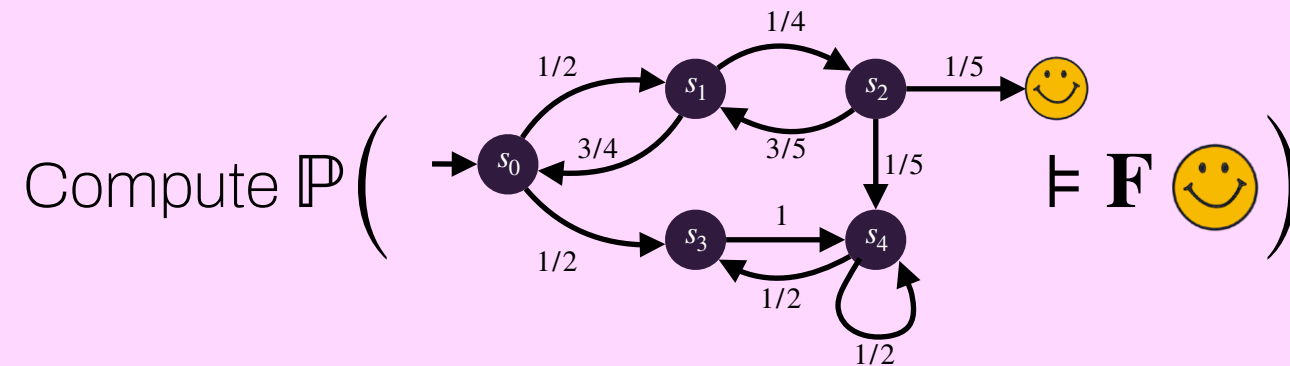
► Probabilistic pushdown automata

$A \xrightarrow{1} C$ $A \xrightarrow{n} BB$ $B \xrightarrow{5} \epsilon$
 $B \xrightarrow{n} AA$ $C \xrightarrow{1} C$

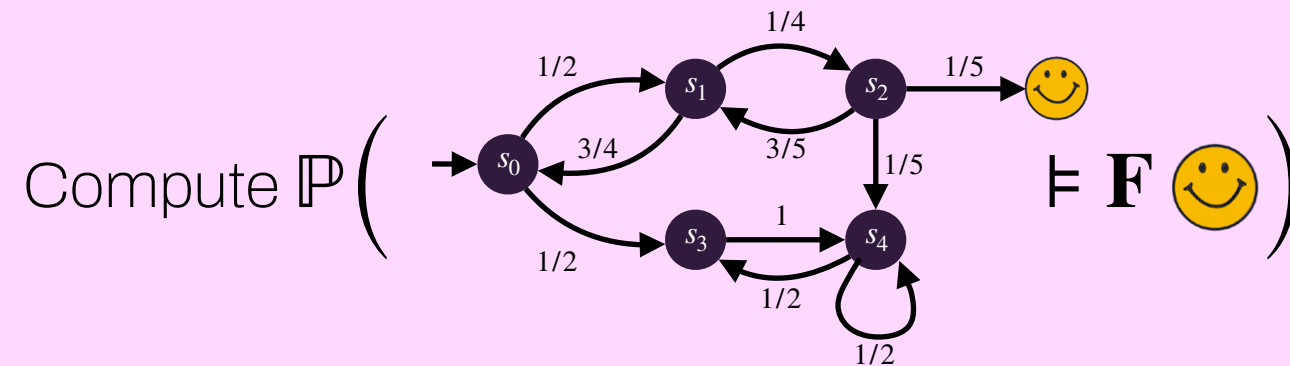
n is the height of the stack



Quantitative analysis of Markov chains



Quantitative analysis of Markov chains

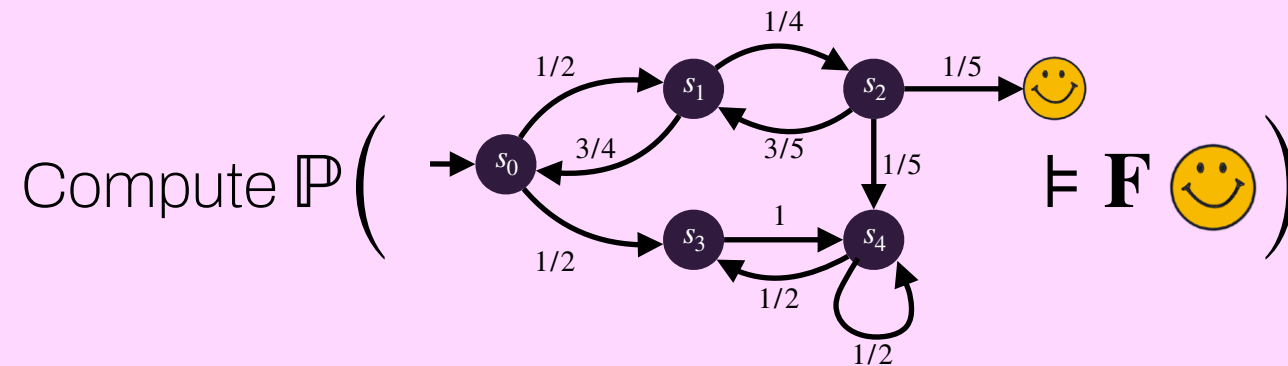


Closed-form solution

- ▶ Random walk of parameter $p > 1/2$:

$$\mathbb{P}_{s_n}(\mathbf{F} \text{ 😊 }) = \kappa^n, \text{ where } \kappa = \frac{1-p}{p}$$
- ▶ Does not always exist

Quantitative analysis of Markov chains



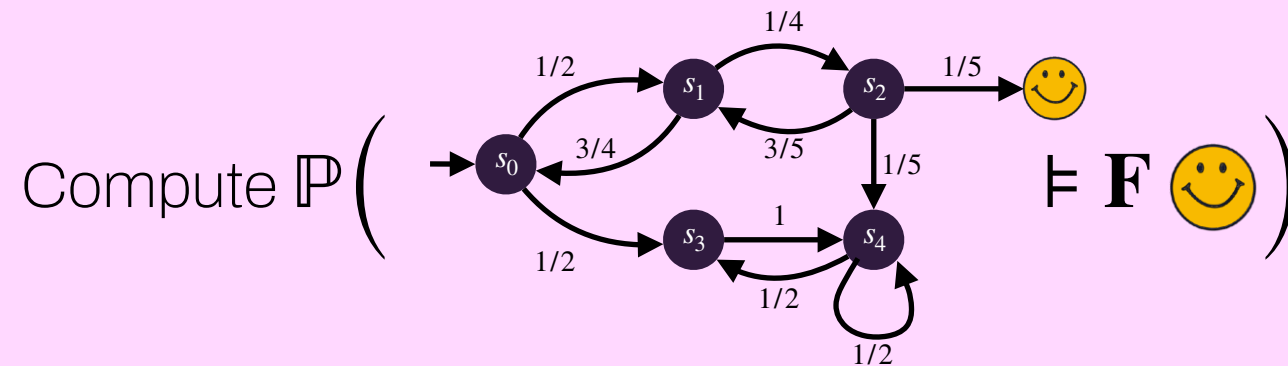
Closed-form solution

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

Apply a numerical method [RKPN04]

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

Quantitative analysis of Markov chains



Closed-form solution

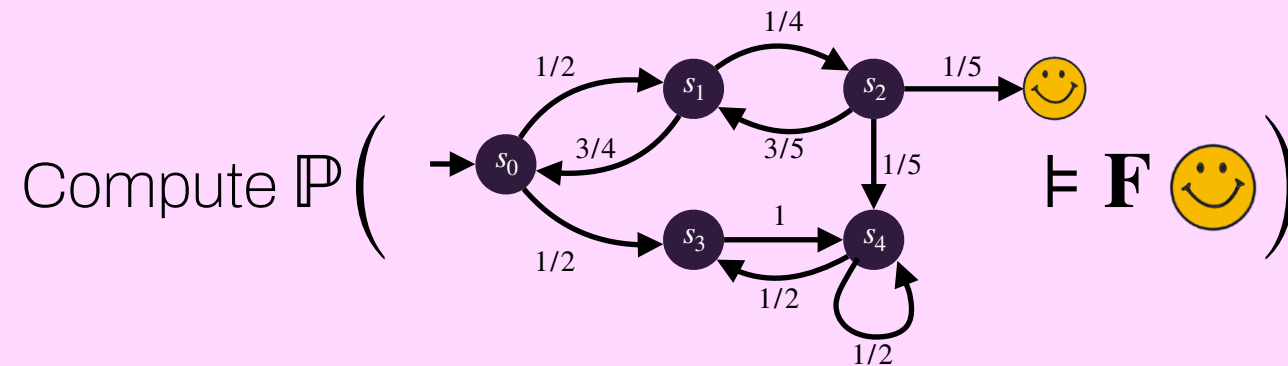
- ▶ Random walk of parameter $p > 1/2$:
 $\mathbb{P}_{s_n}(\mathbf{F} \text{ 😊}) = \kappa^n$, where $\kappa = \frac{1-p}{p}$
- ▶ Does not always exist

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

Apply a numerical method [RKPN04]

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

- ▶ No general method exists for infinite Markov chains
- ▶ Specific approaches for **decisive** Markov chains

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. ☺️ 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. ☺️ 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, ...

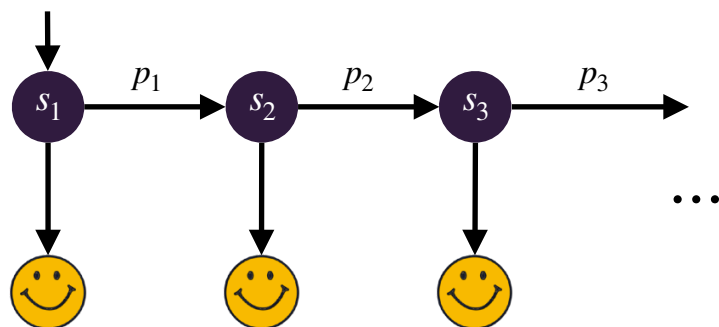
Decisiveness

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

Decisiveness

A DTMC \mathcal{C} is **decisive** from s w.r.t. ☺️ 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:



$$\bullet \quad \mathbb{P}(\mathbf{G} \neg \text{☺️}) = \prod_{i \geq 1} p_i$$

- Decisive iff this product equals 0

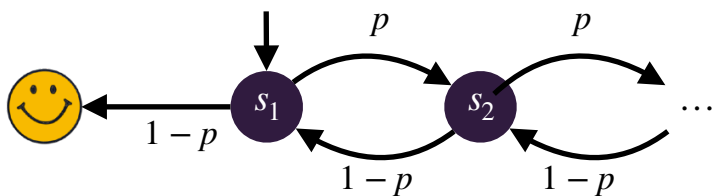
Decisiveness

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

Decisiveness

A DTMC \mathcal{C} is **decisive** from s w.r.t. ☺️ 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

Approximation scheme

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

Approximation scheme

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

Approximation scheme

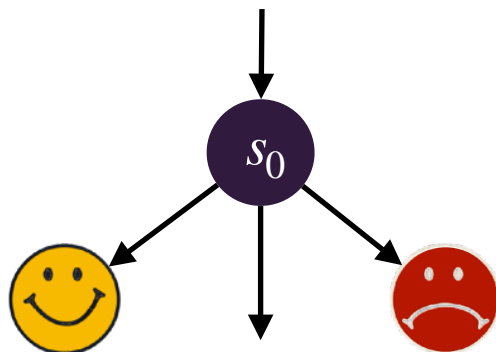
Given $\varepsilon > 0$, for every n , compute:

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

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

Approximation scheme

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



Approximation scheme

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

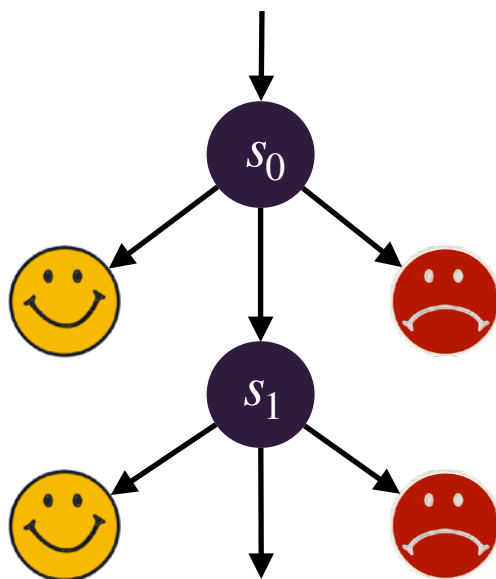
$$\begin{cases} p_n^{\text{yes}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😊}) \\ p_n^{\text{no}} &= \mathbb{P}(\mathbf{F}_{\leq n} \text{ 😞}) \end{cases}$$

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

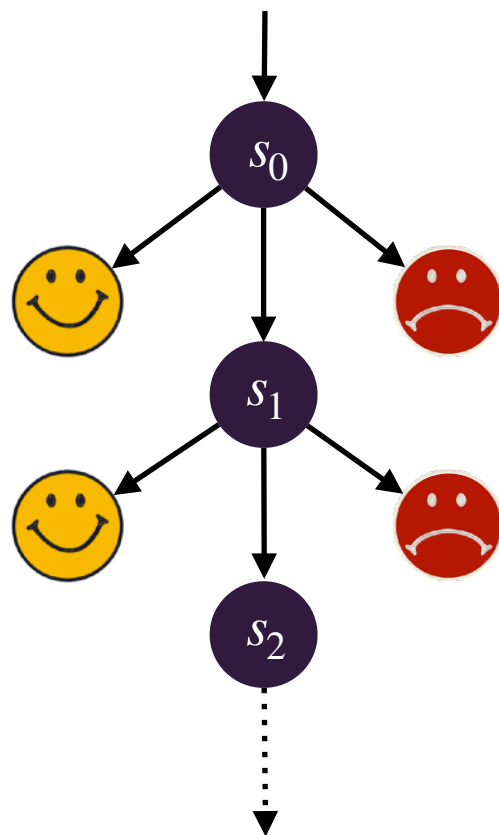
$$p_1^{\text{yes}} \leq \mathbb{P}(\mathbf{F} \text{ 😊}) \leq 1 - p_1^{\text{no}}$$

$$\text{I} \wedge \qquad \qquad \qquad \text{V} \text{I}$$

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

Approximation scheme

- ▶ Aim: compute probability of **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 $p_n^{\text{yes}} + p_n^{\text{no}} \geq 1 - \varepsilon$

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

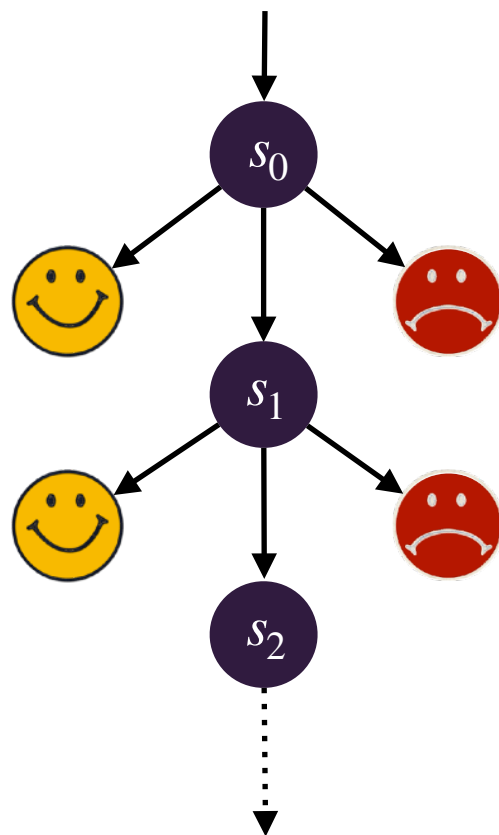
$$\bigwedge \qquad \qquad \qquad \bigvee$$

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

$$\bigwedge \qquad \vdots \qquad \qquad \bigvee$$

Approximation scheme

- ▶ Aim: compute probability of **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 $p_n^{\text{yes}} + p_n^{\text{no}} \geq 1 - \varepsilon$

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

$$\bigwedge \qquad \qquad \qquad \bigvee$$

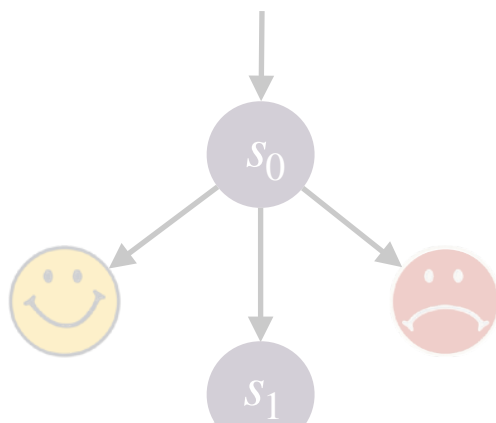
$$p_2^{\text{yes}} \leq \mathbb{P}(\mathbf{F} \text{ 😊}) \leq 1 - p_2^{\text{no}}$$

$$\bigwedge \qquad \qquad \qquad \bigvee$$

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

Approximation scheme

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



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

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

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

$$\text{I} \wedge \qquad \qquad \qquad \text{V} \text{I}$$

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

$$\text{I} \wedge \qquad \qquad \qquad \text{V} \text{I}$$

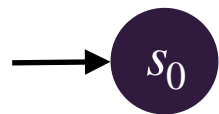
$$\mathbb{P}(\mathbf{F} \text{ 😊})$$

$$1 - \mathbb{P}(\mathbf{F} \text{ 😞})$$

At the limit:

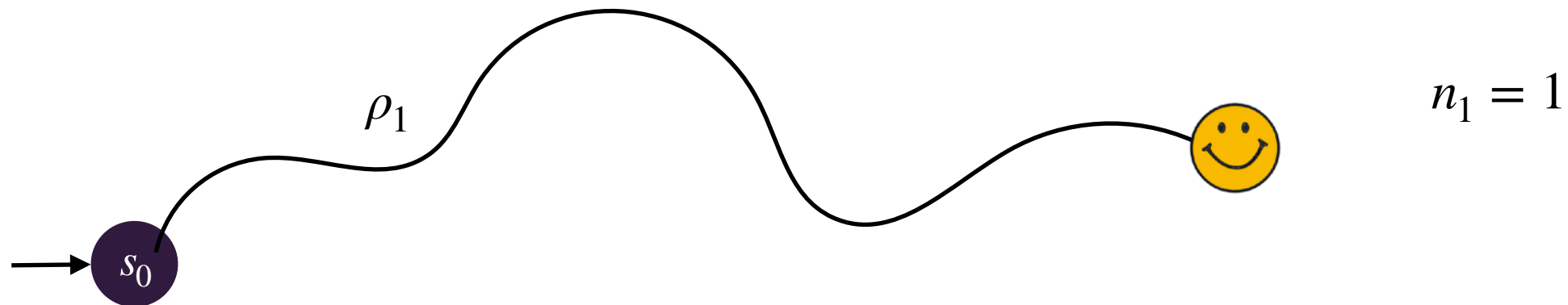
Statistical model-checking

Sample N paths



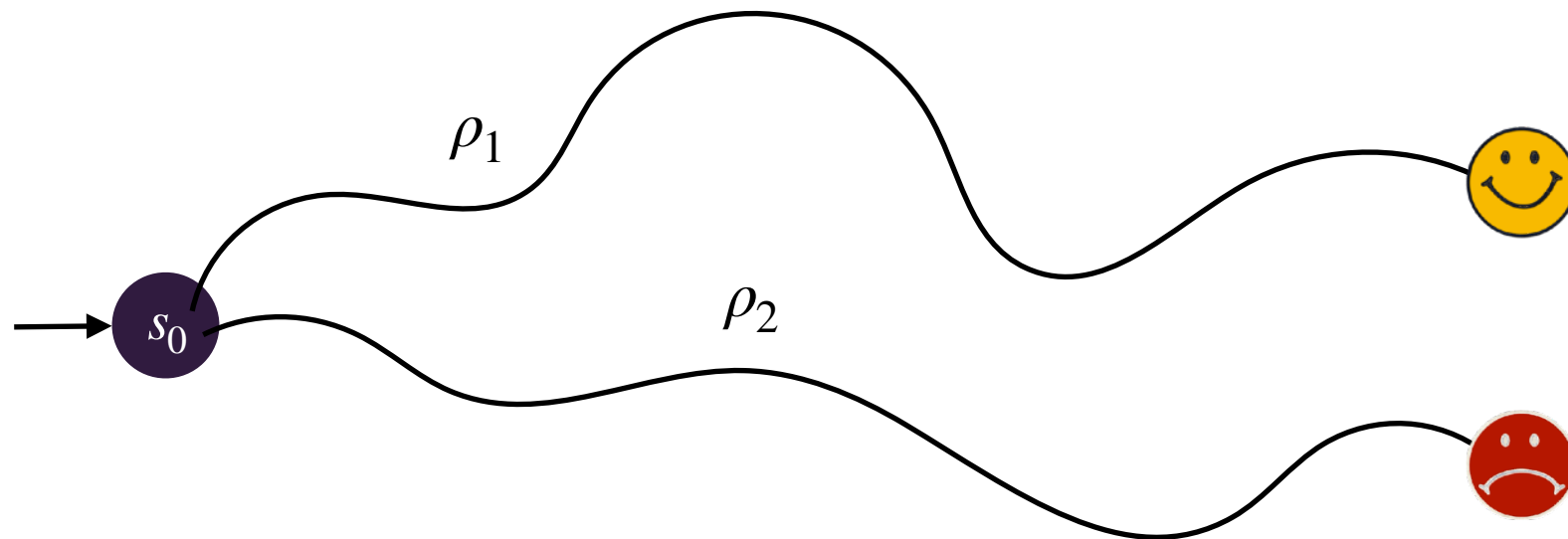
Statistical model-checking

Sample N paths



Statistical model-checking

Sample N paths

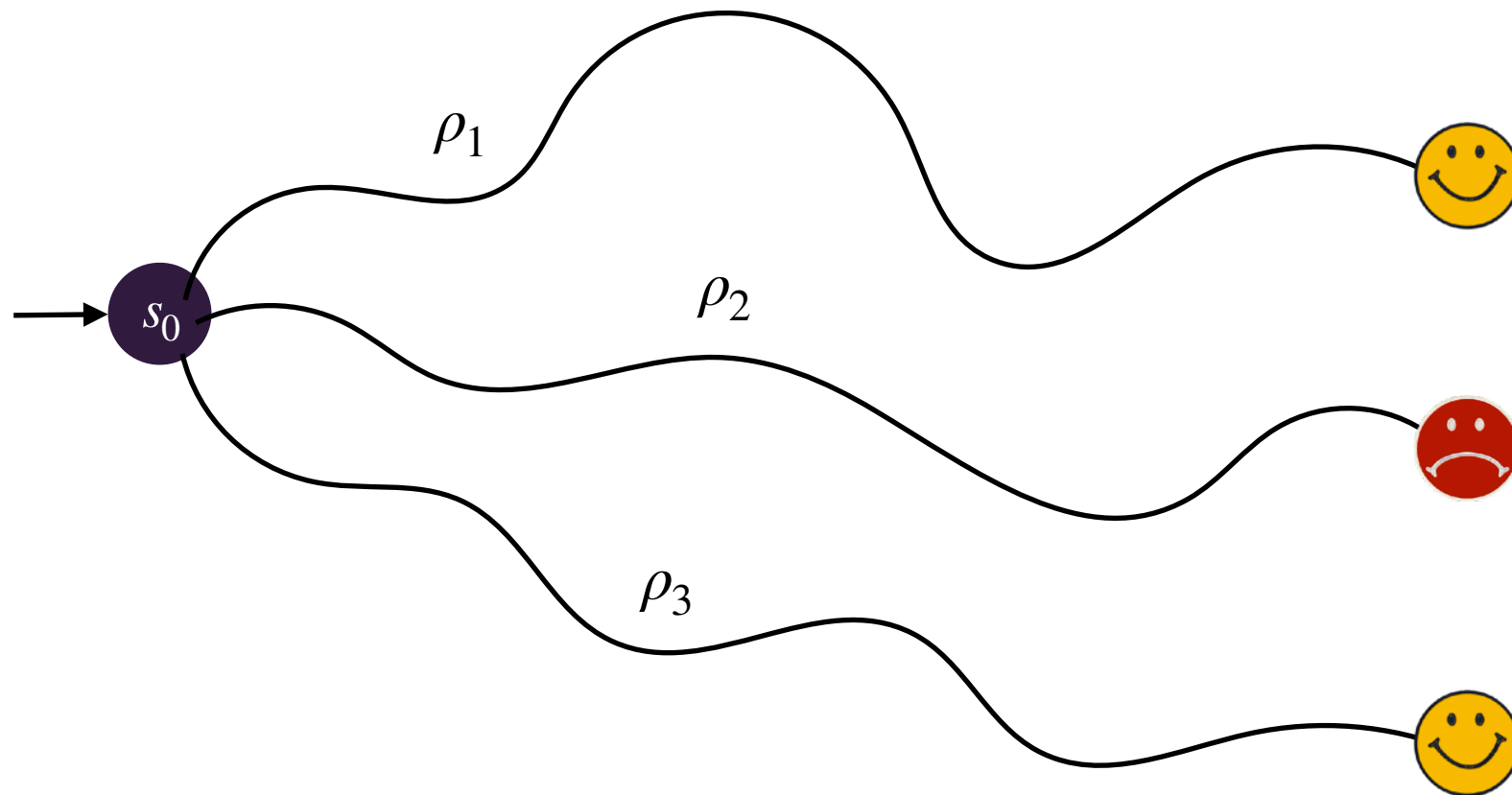


$$n_1 = 1$$

$$n_2 = n_1$$

Statistical model-checking

Sample N paths



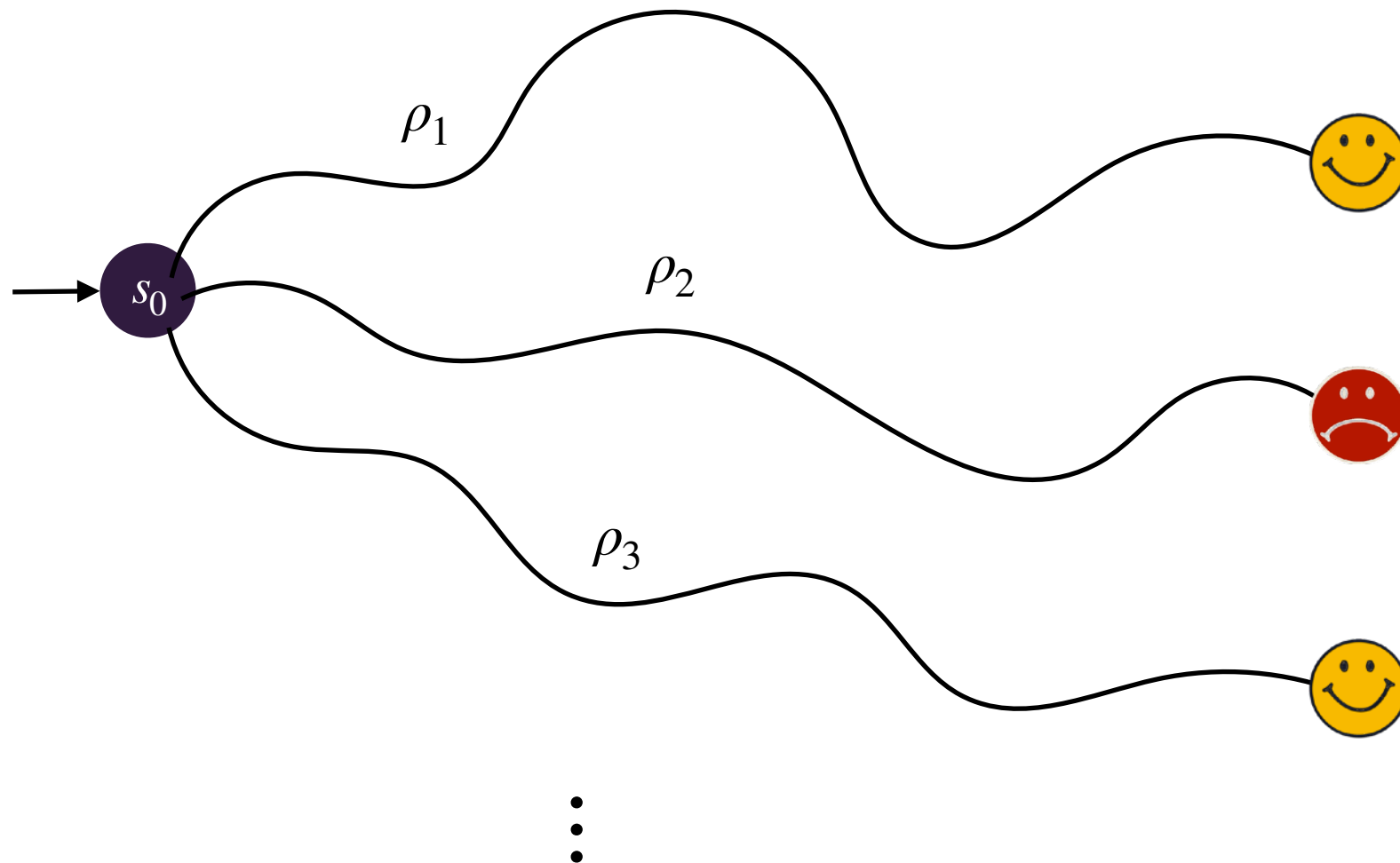
$$n_1 = 1$$

$$n_2 = n_1$$

$$n_3 = n_2 + 1$$

Statistical model-checking

Sample N paths



$$n_1 = 1$$

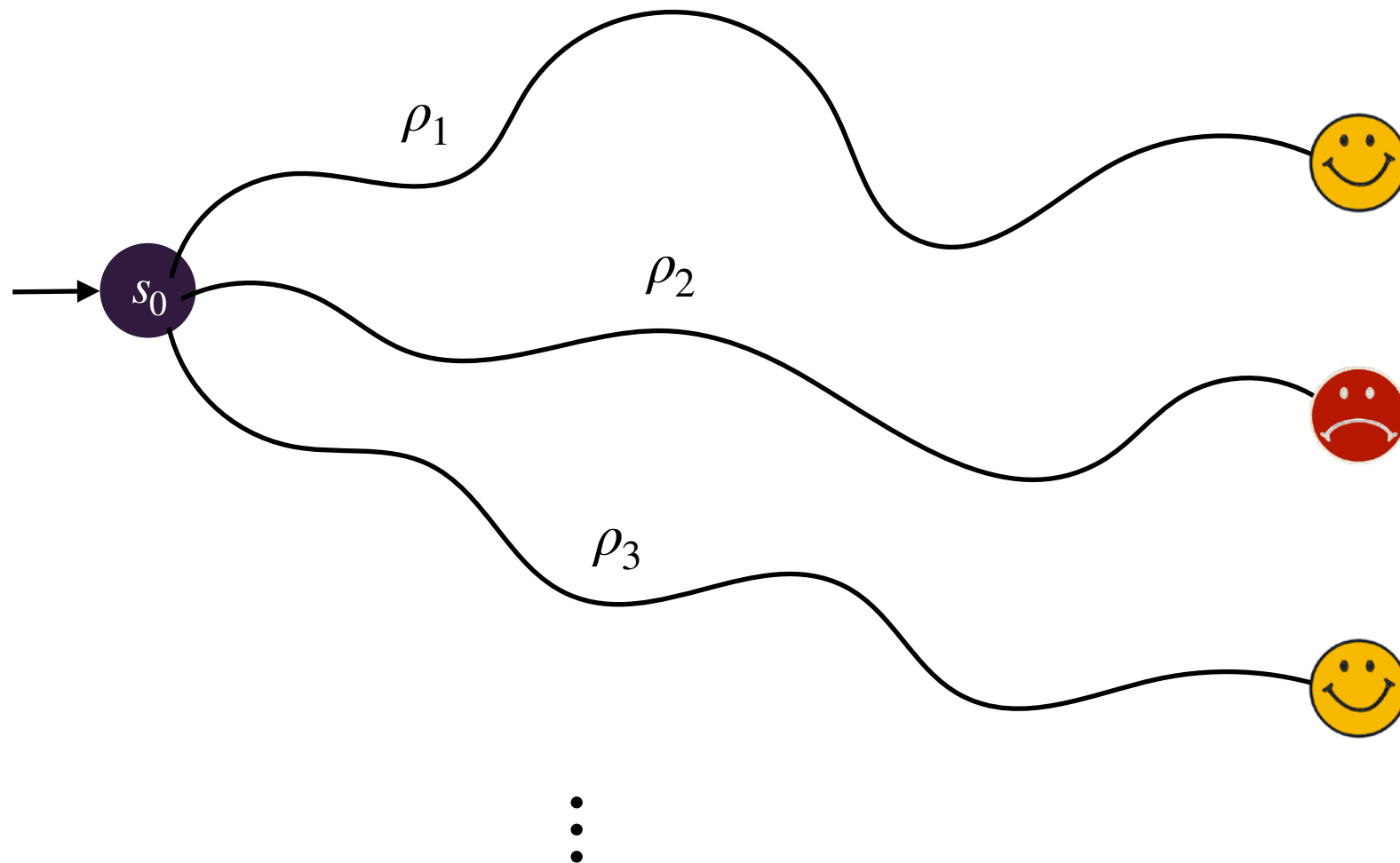
$$n_2 = n_1$$

$$n_3 = n_2 + 1$$

⋮

Statistical model-checking

Sample N paths



$$n_1 = 1$$

$$n_2 = n_1$$

$$n_3 = n_2 + 1$$

\vdots

Return $\frac{n_N}{N} + \text{some confidence interval}$

Termination, efficiency and guarantees

Termination

(To our knowledge, never expressed like this)

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

Termination, efficiency and guarantees

Termination

(To our knowledge, never expressed like this)

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

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

Termination, efficiency and guarantees

Termination

(To our knowledge, never expressed like this)

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

+ efficiency if finite return time
(« \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)

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

+ efficiency if finite return time
(« \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 value

Precision

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

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)

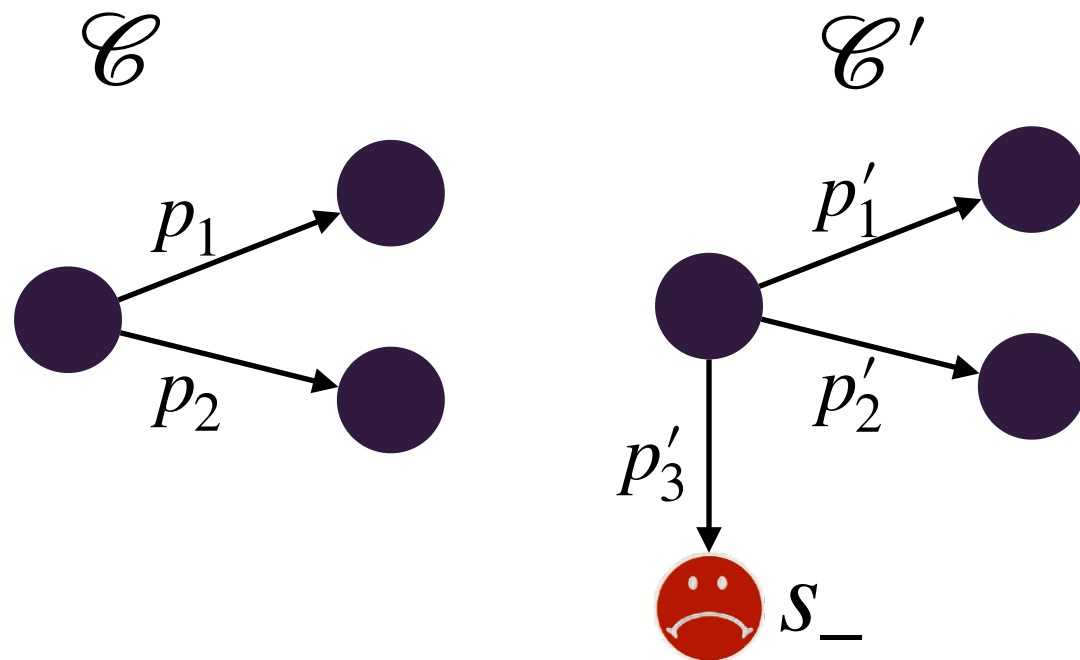
[Bar14] B. Barbot. *Acceleration for statistical model checking* (PhD thesis)

[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)

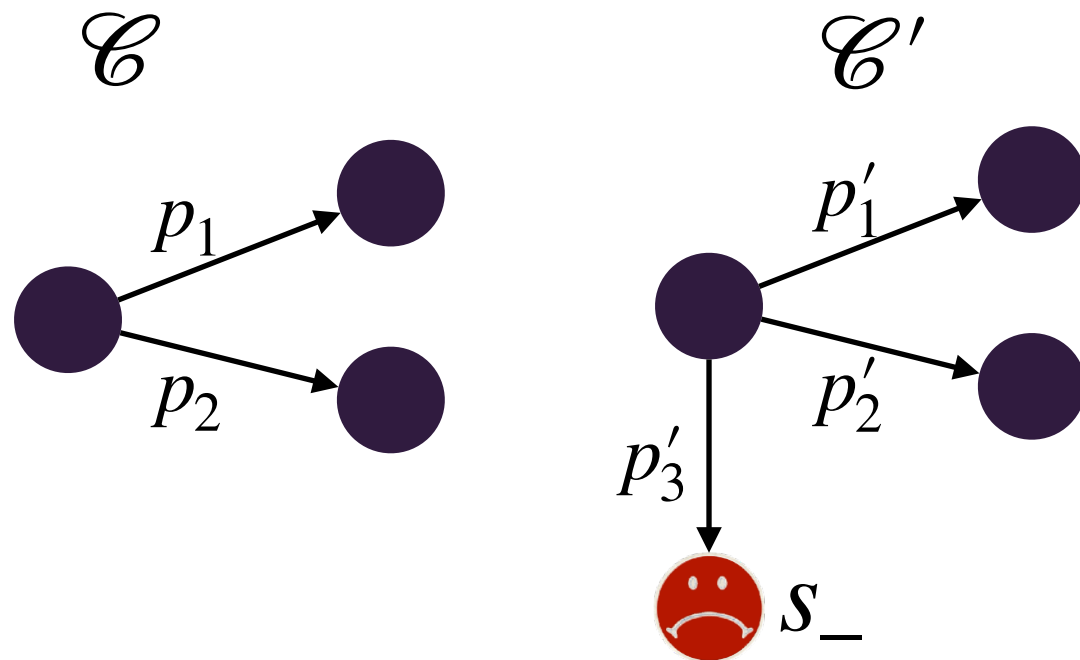
[Bar14] B. Barbot. Acceleration for statistical model checking (PhD thesis)

[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)

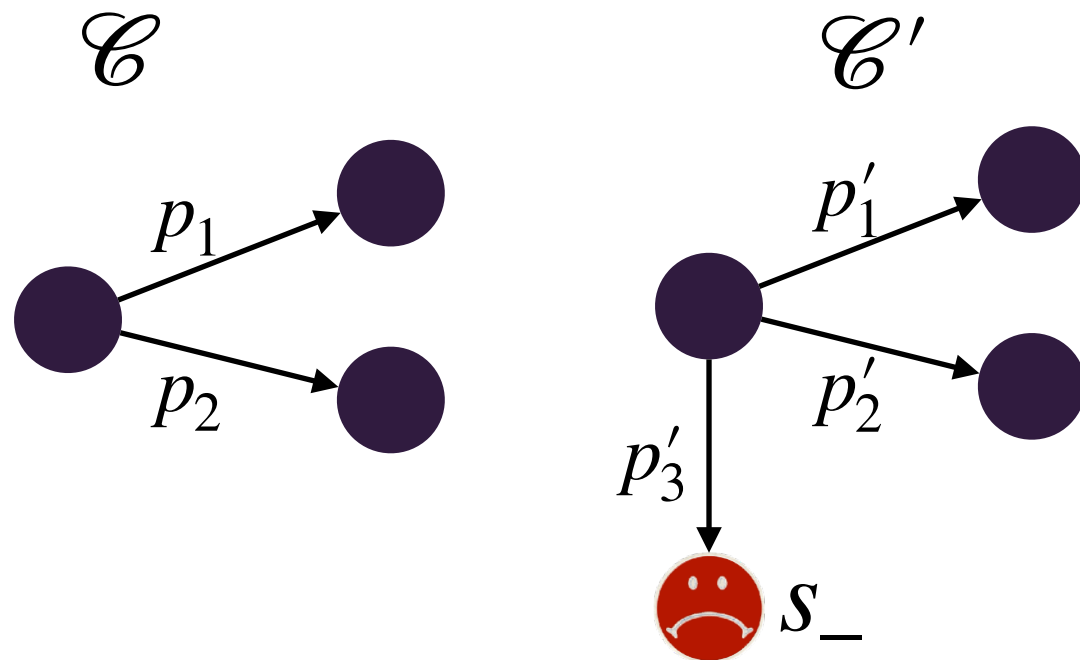
[Bar14] B. Barbot. Acceleration for statistical model checking (PhD thesis)

[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)

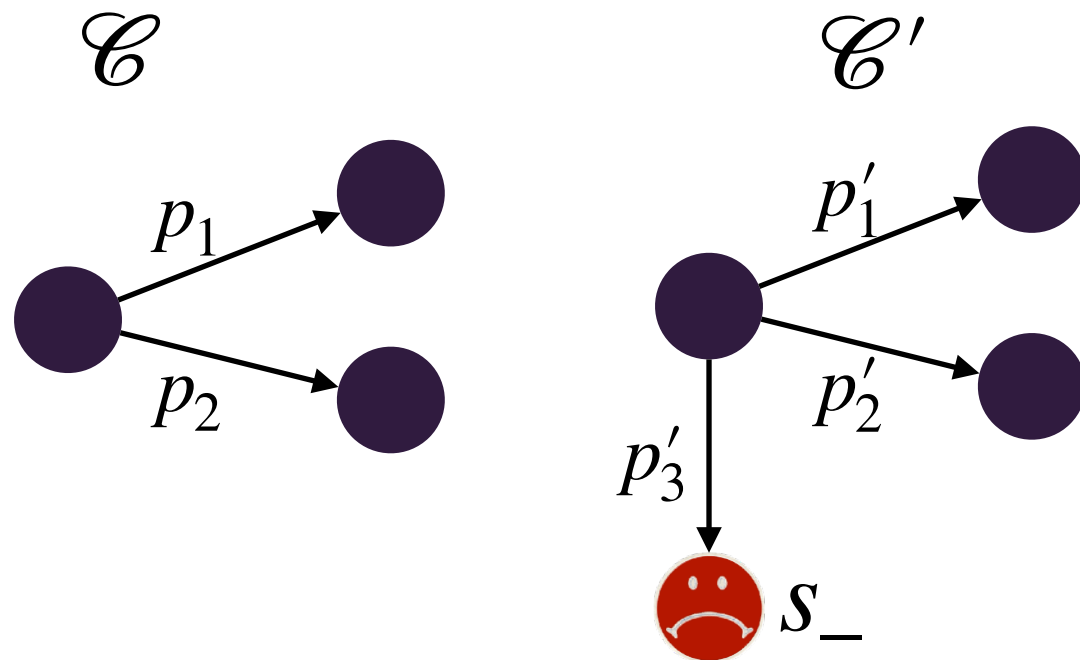
[Bar14] B. Barbot. Acceleration for statistical model checking (PhD thesis)

[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)

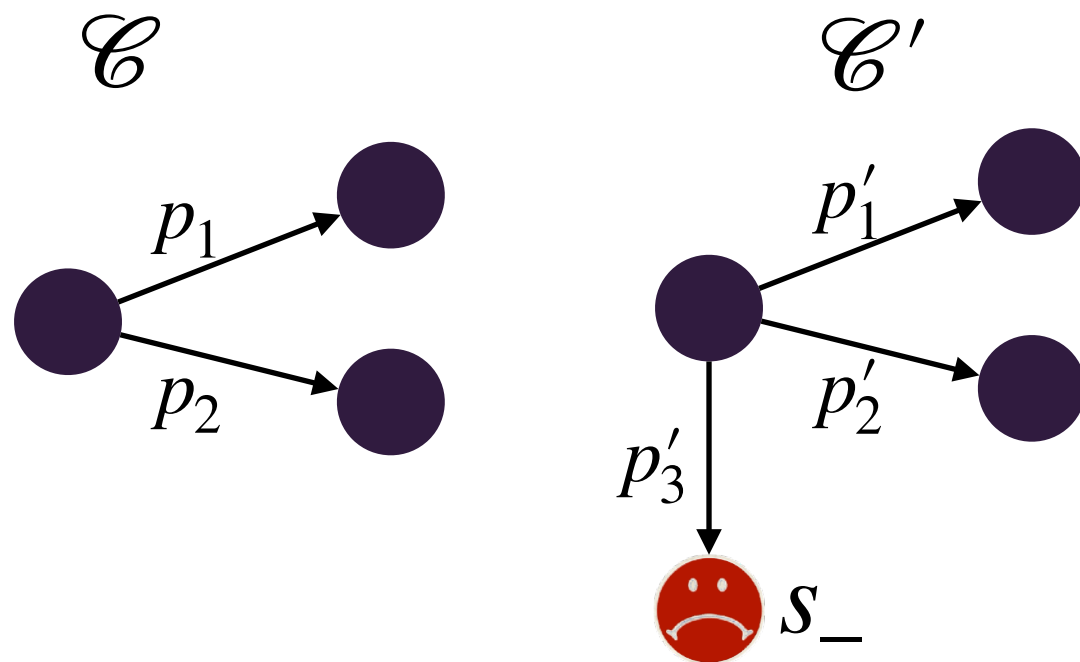
[Bar14] B. Barbot. Acceleration for statistical model checking (PhD thesis)

[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 $\mathbb{E}_{\mathcal{C}'}(\gamma)$

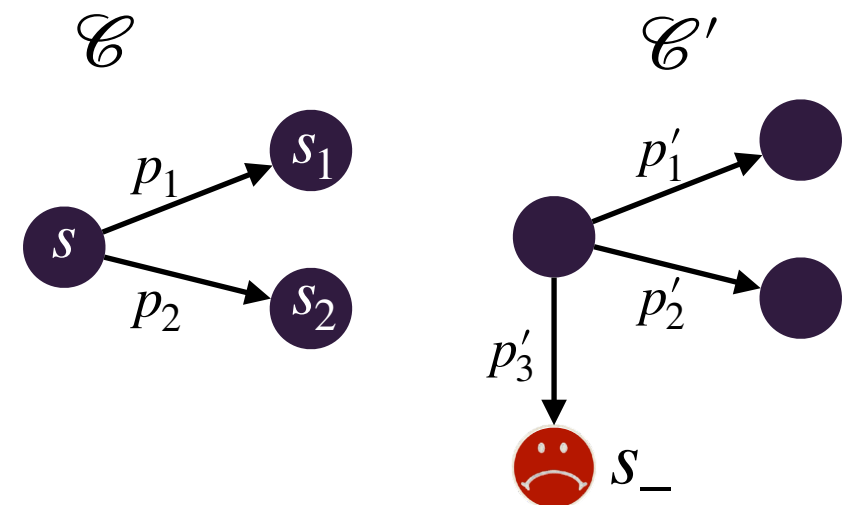
- Originally used for rare events
- Setting giving statistical guarantees [BHP12, Bar14]

[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)

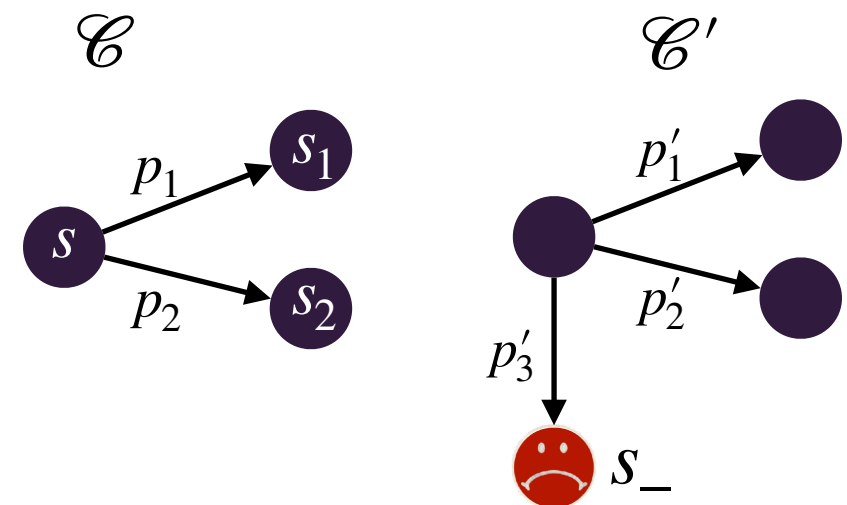
[BHP12] B. Barbot, S. Haddad, C. Picaronny. Coupling and Importance Sampling for Statistical Model Checking (TACAS'12)

Properties of the biased Markov chain



Properties of the biased Markov chain

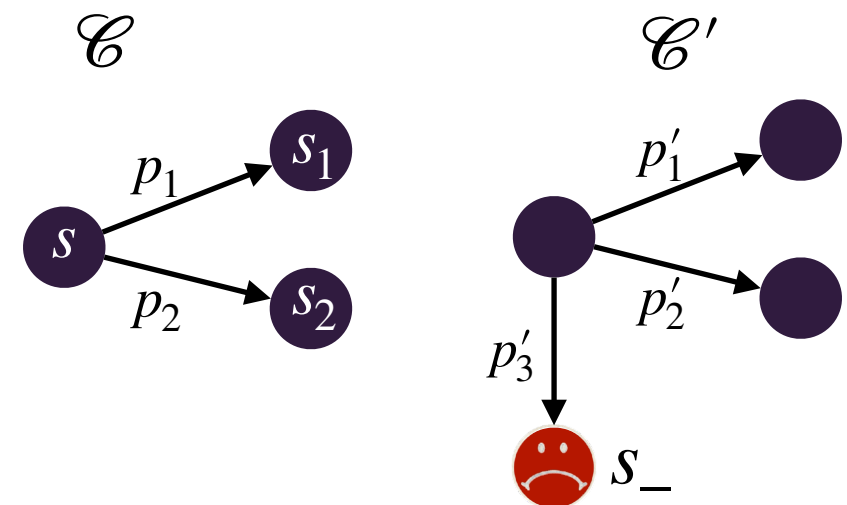
$$\mathbb{P}_{\mathcal{C}}(\mathbf{F} \text{ 😊 }) = \mathbb{E}_{\mathcal{C}'}(\gamma)$$



Properties of the biased Markov chain

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

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

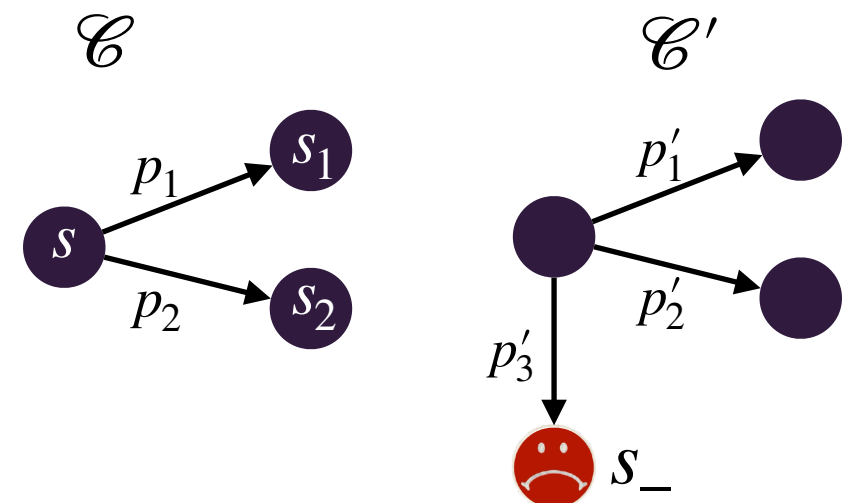


Properties of the biased Markov chain

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 r.v. in \mathcal{C} is distributed according to a Bernoulli distribution
The r.v. in \mathcal{C}' is distributed according to an unknown distribution

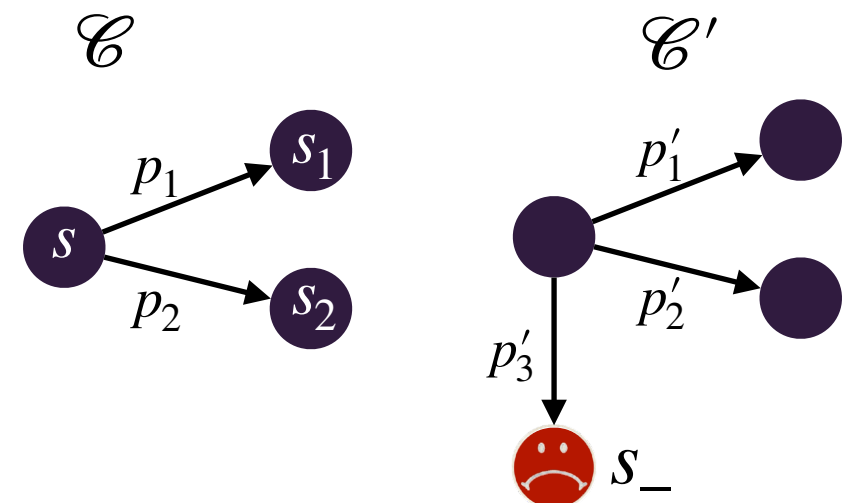


Properties of the biased Markov chain

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 r.v. in \mathcal{C} is distributed according to a Bernoulli distribution
The r.v. in \mathcal{C}' is distributed according to an unknown distribution
- ▶ The analysis of \mathcal{C} can be transferred to that of \mathcal{C}' , provided some conditions on \mathcal{C}'

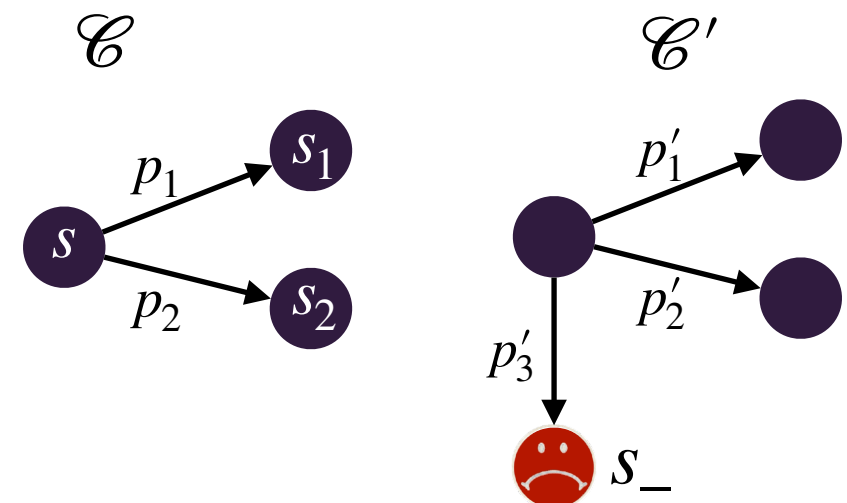


Properties of the biased Markov chain

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 r.v. in \mathcal{C} is distributed according to a Bernoulli distribution
The r.v. in \mathcal{C}' is distributed according to an unknown distribution
- ▶ 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

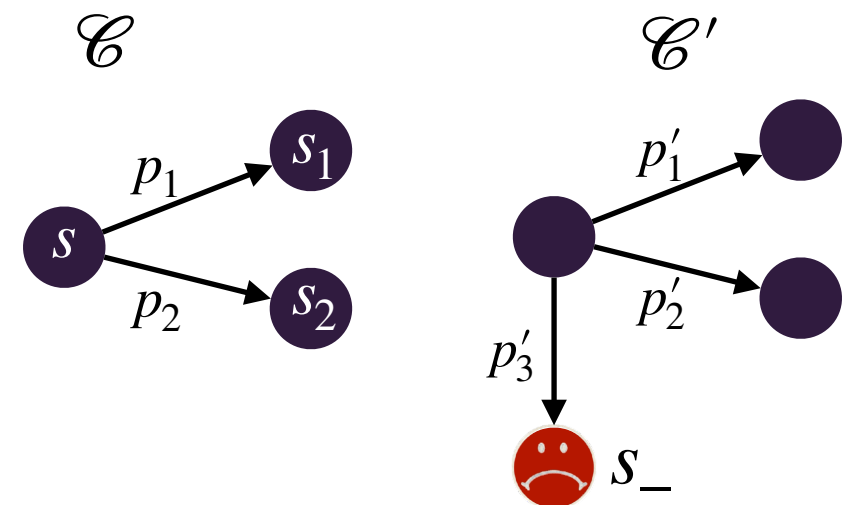


Properties of the biased Markov chain

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 r.v. in \mathcal{C} is distributed according to a Bernoulli distribution
The r.v. in \mathcal{C}' is distributed according to an unknown distribution
- ▶ 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 γ is required



Properties of the biased Markov chain

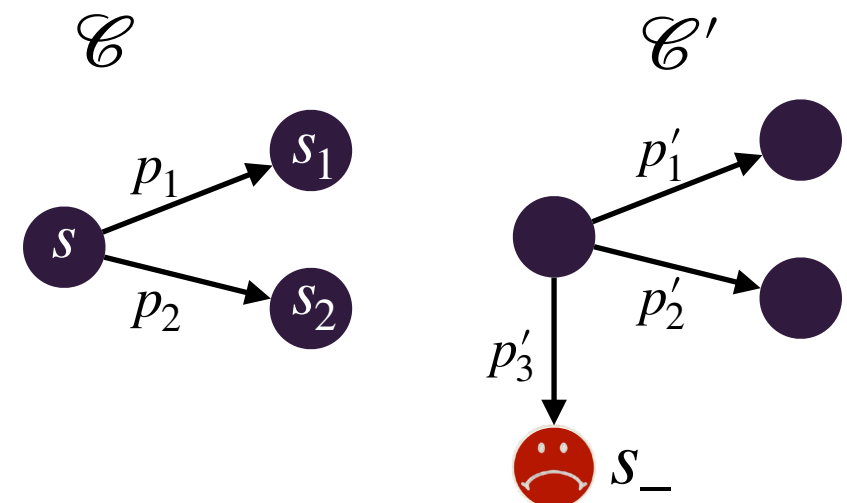
Define $\mu(s)$ as
 $\mathbb{P}_{\mathcal{C}}^s(\mathbf{F} \text{ 😊 })$

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

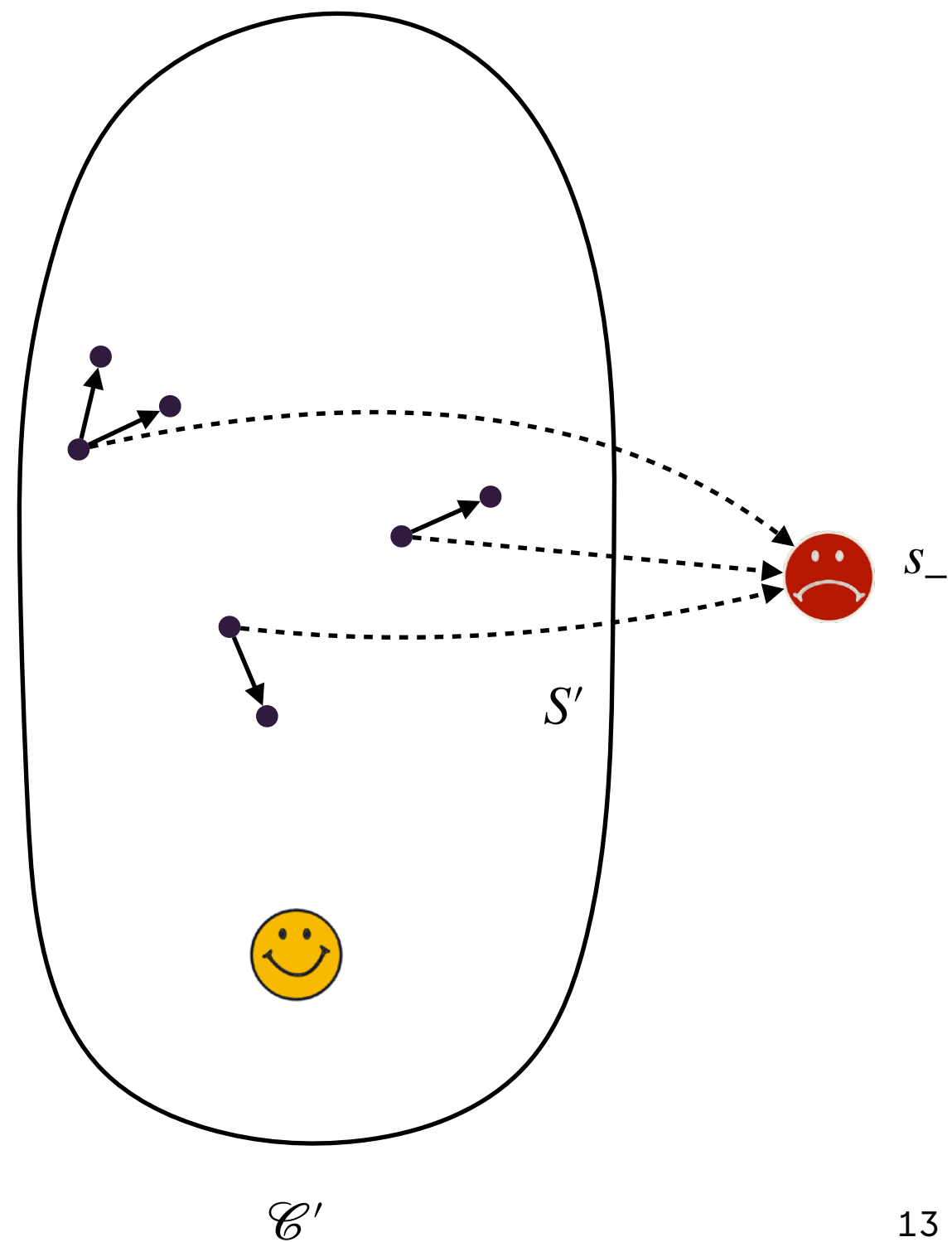
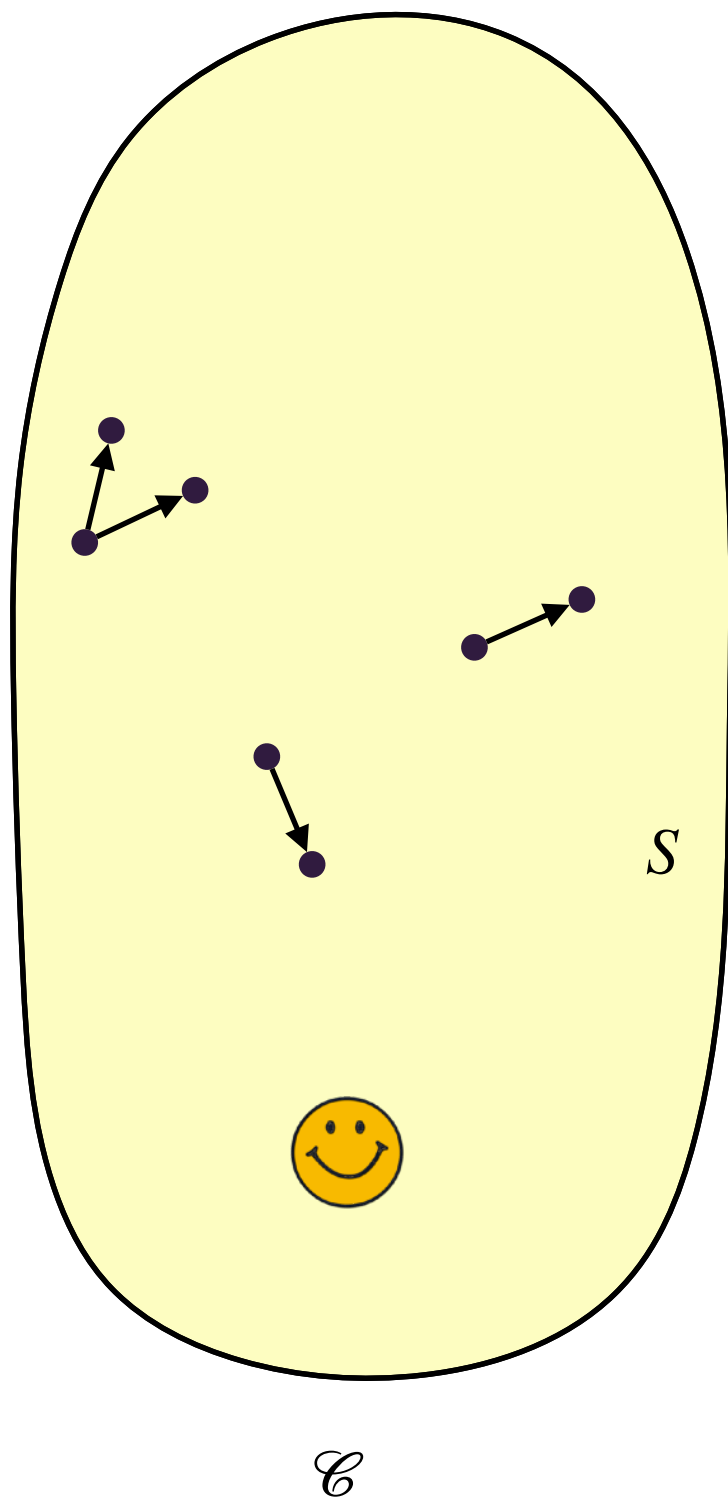
- ▶ The r.v. in \mathcal{C} is distributed according to a Bernoulli distribution
 The r.v. in \mathcal{C}' is distributed according to an unknown distribution
- ▶ The analysis of \mathcal{C} can be transferred to that of \mathcal{C}' , provided some conditions on \mathcal{C}'
 - Decisiveness of \mathcal{C}' is required for both approx. and estim. methods
 - Boundedness of γ is required

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

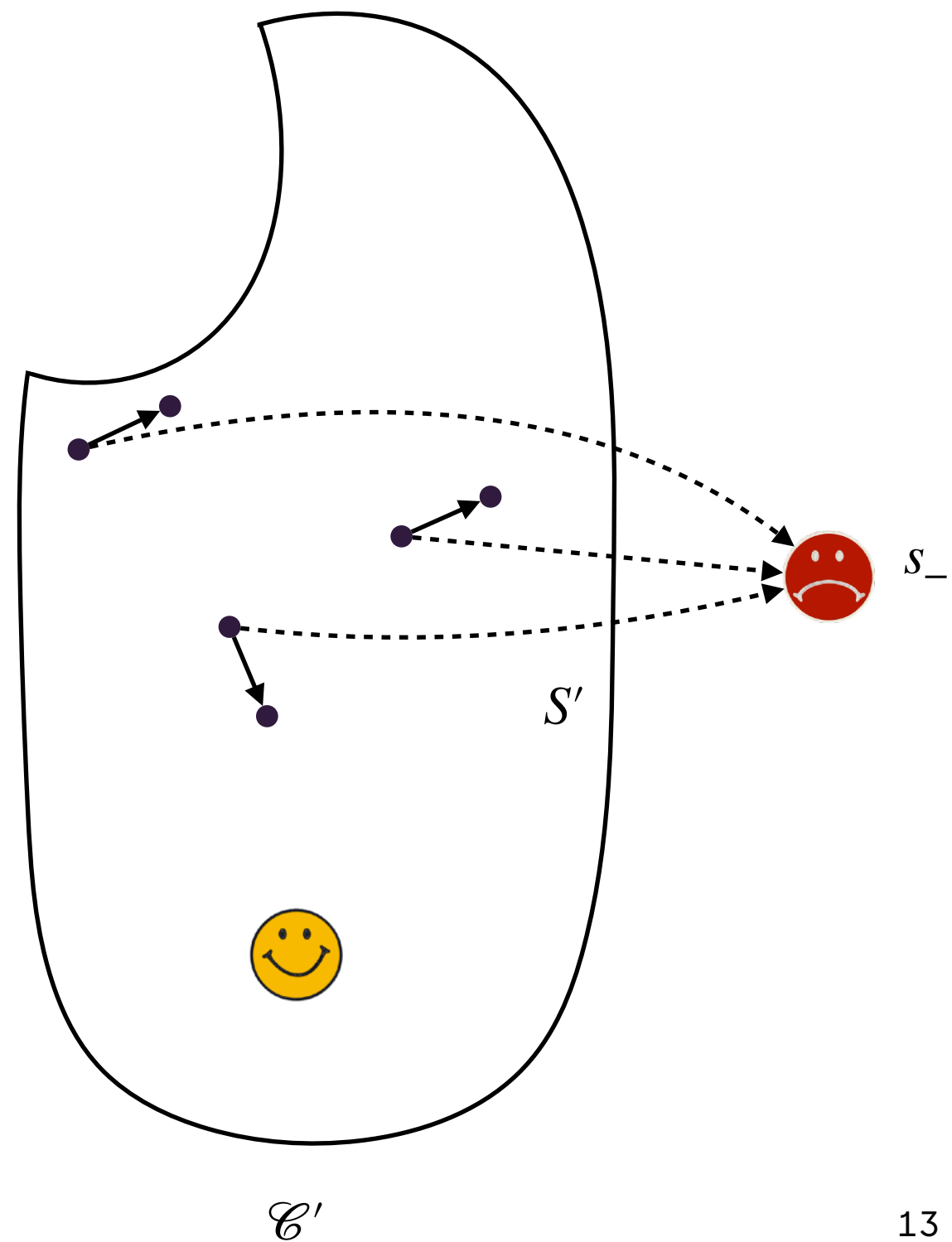
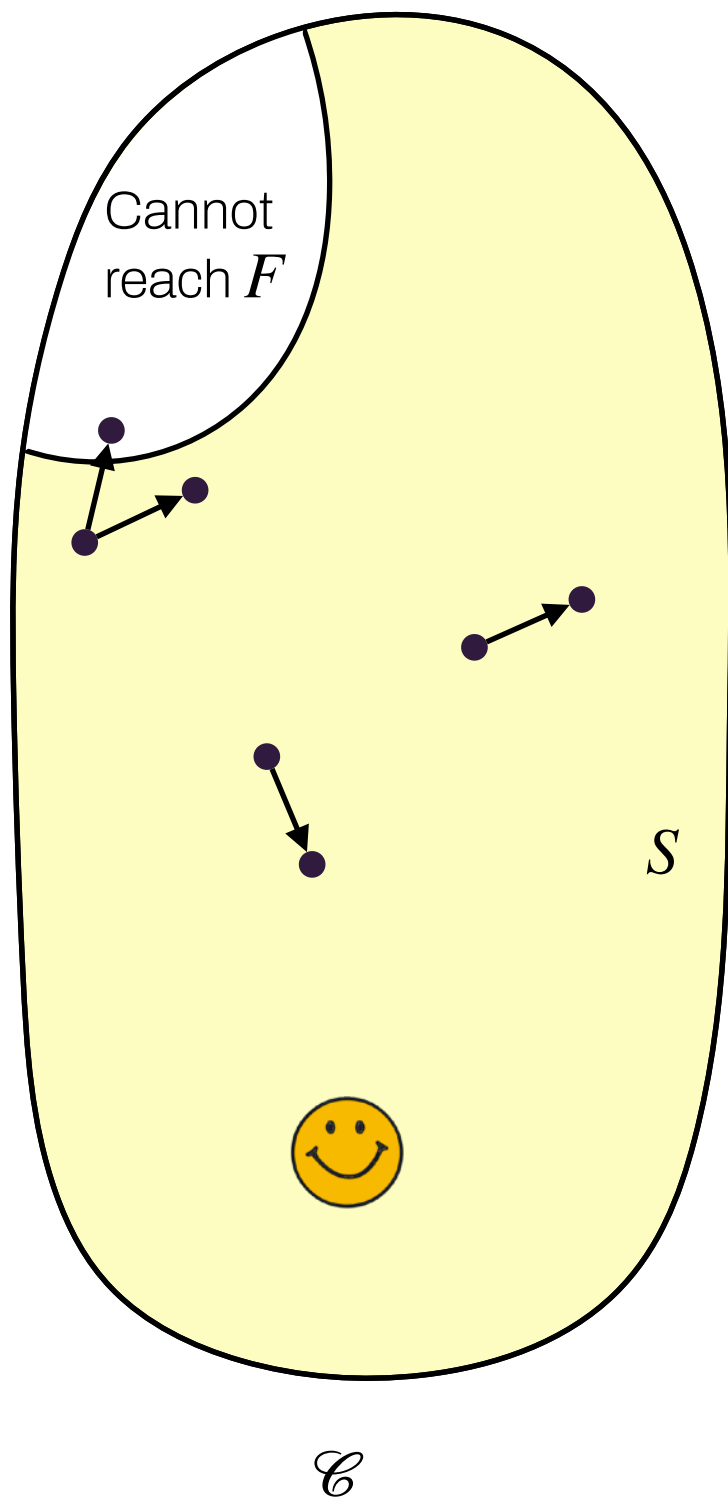
- ▶ The r.v. in \mathcal{C}' takes value $\mu(s)$
- ▶ One needs to know μ !



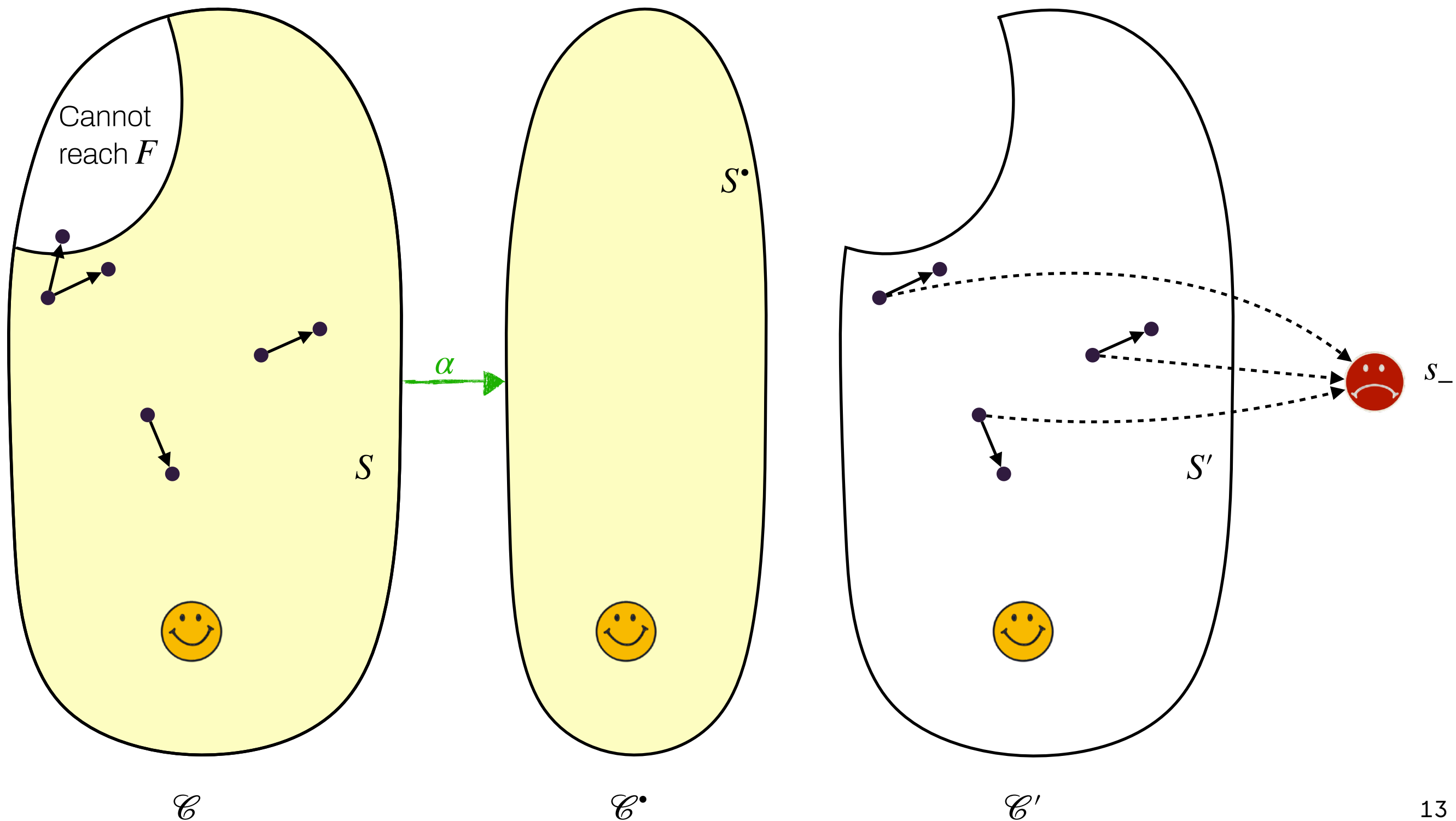
Importance sampling via an abstraction



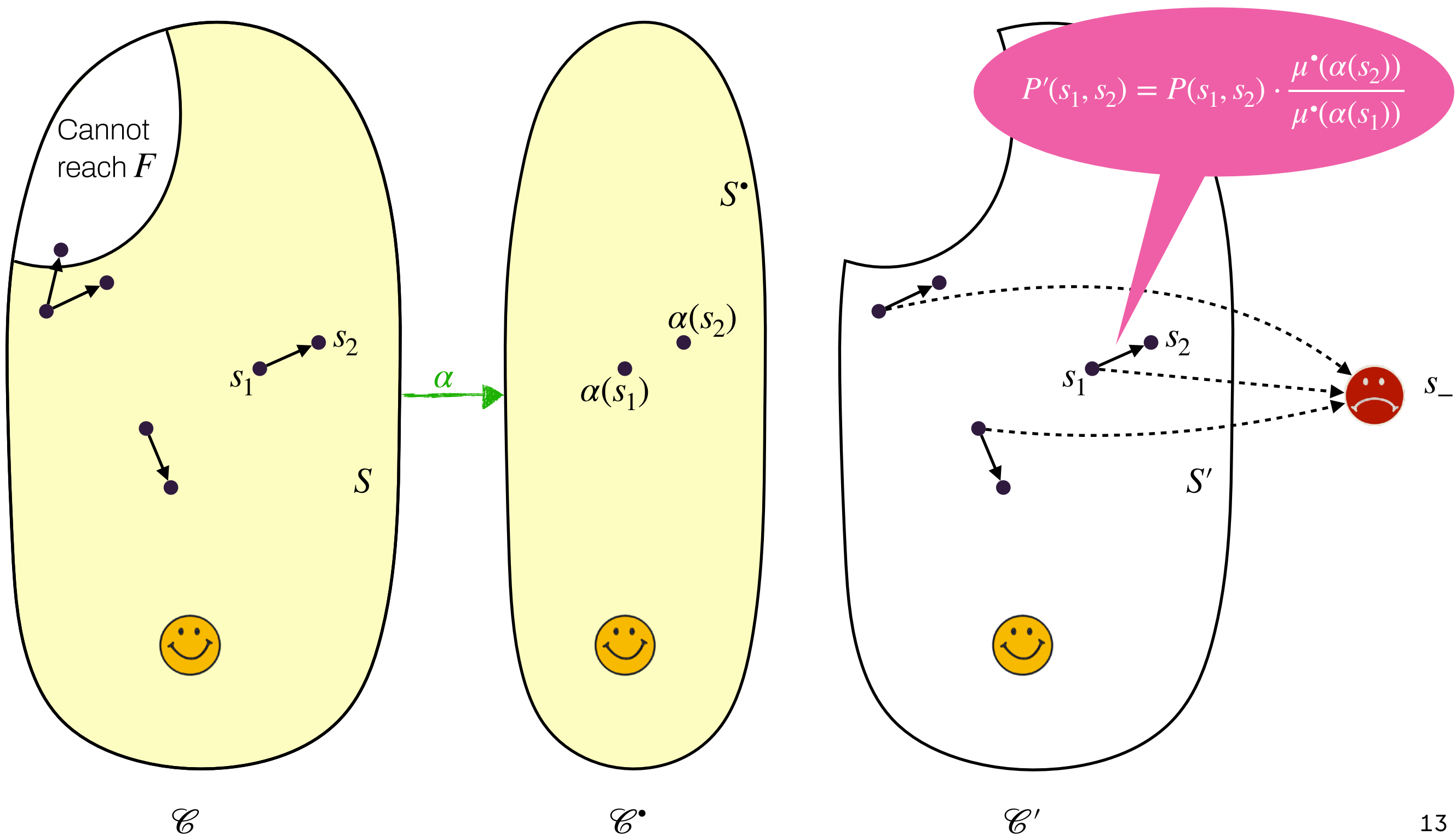
Importance sampling via an abstraction



Importance sampling via an abstraction

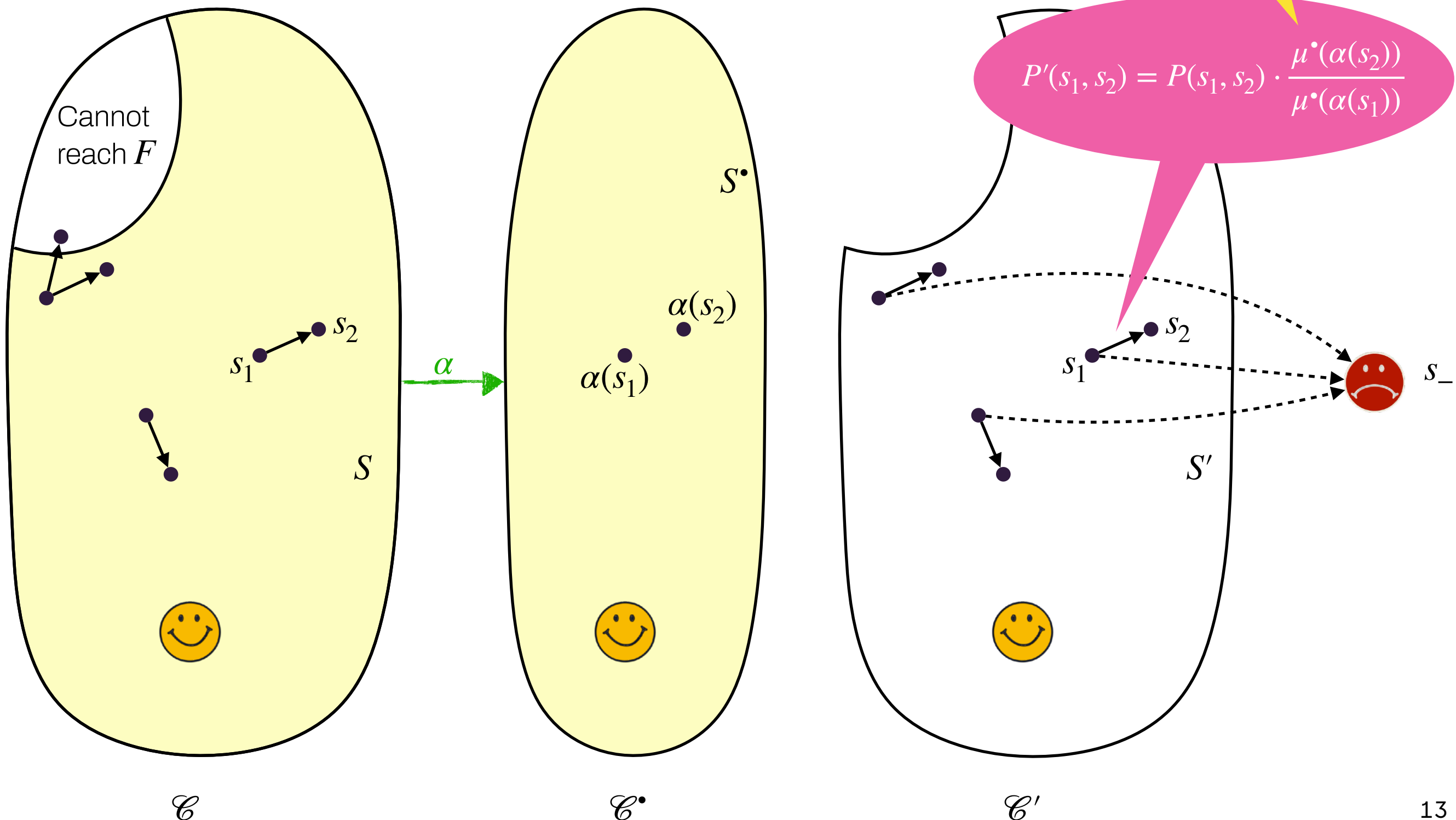


Importance sampling via an abstraction

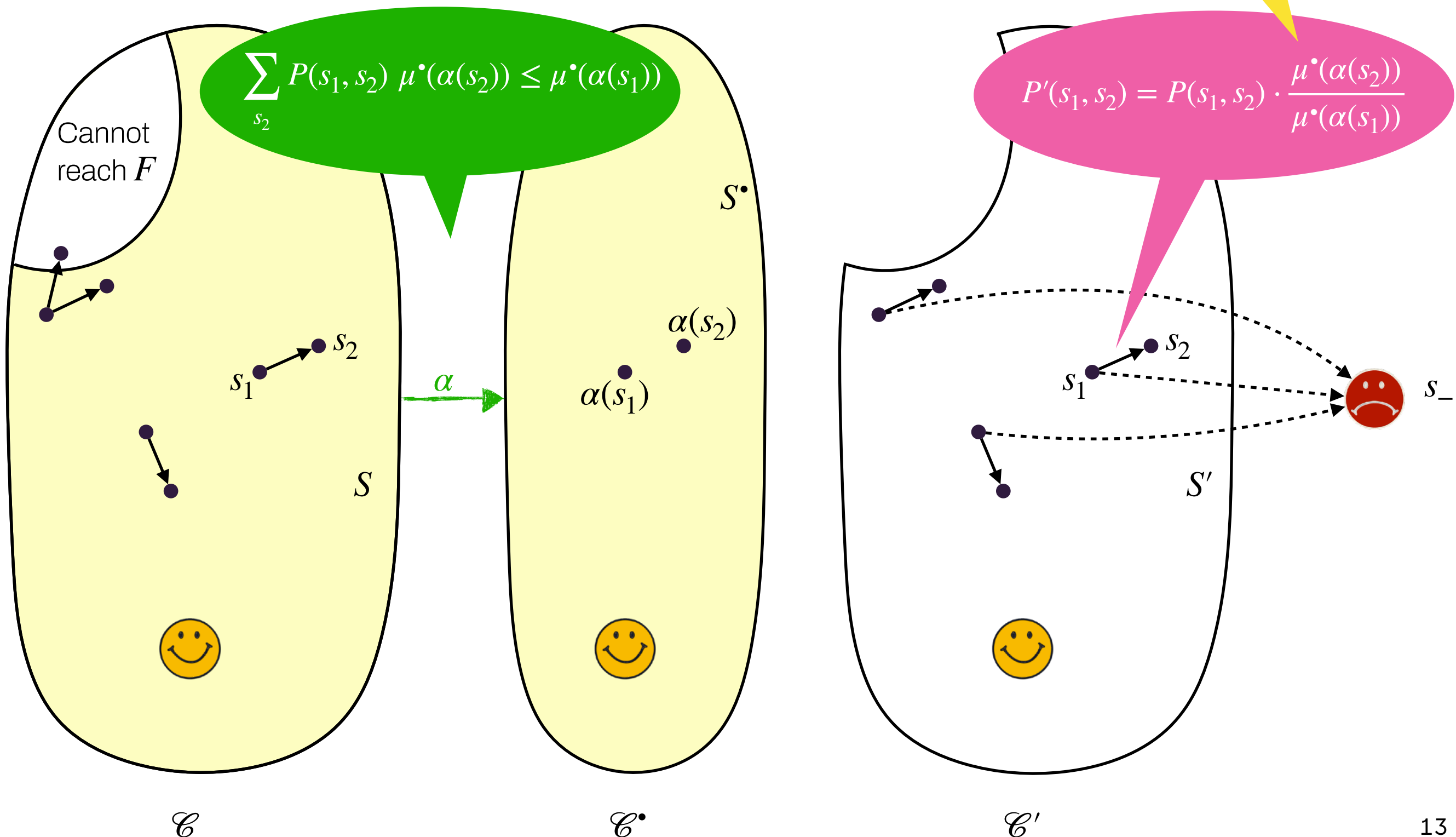


Importance sampling via an abstraction

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



Importance sampling via an abstraction



And «concretely»?

And «concretely»?

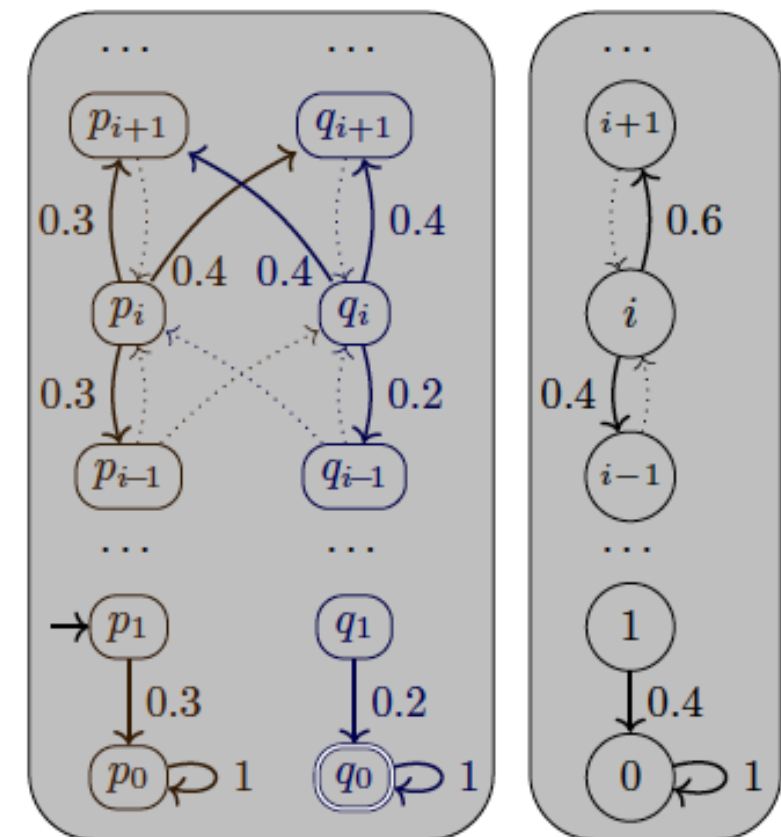
- ▶ Model = layered Markov chain (LMC) \mathcal{C} : there is a level function $\lambda : S \rightarrow \mathbb{N}$ s.t.
 - for every $s_1 \rightarrow s_2$, $\lambda(s_1) - \lambda(s_2) \leq 1$, and
 - for every n , $\lambda^{-1}(n)$ is finite

And «concretely»?

- ▶ Model = layered Markov chain (LMC) \mathcal{C} : there is a level function $\lambda : S \rightarrow \mathbb{N}$ 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^\bullet of parameter p

And «concretely»?

- ▶ Model = layered Markov chain (LMC) \mathcal{C} : there is a level function $\lambda : S \rightarrow \mathbb{N}$ 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^\bullet of parameter p



Correctness of the approach

Theorem

Let \mathcal{C} be an LMC with level function λ , \mathcal{C}_p^\bullet the random walk of parameter p . Assume there is N_0 s.t. $\frac{1}{2} < p < p_+ = \inf\{P^+(s) \mid \lambda(s) > N_0\}$. Then:

- ▶ \mathcal{C}_p^\bullet is an abstraction for \mathcal{C}
- ▶ The corresponding biased Markov chain \mathcal{C}' is decisive w.r.t. 😊
- ▶ The expected time to sample an execution is finite

Correctness of the approach

Theorem

Let \mathcal{C} be an LMC with level function λ , \mathcal{C}_p^\bullet the random walk of parameter p . Assume there is N_0 s.t. $\frac{1}{2} < p < p_+ = \inf\{P^+(s) \mid \lambda(s) > N_0\}$. Then:

- ▶ \mathcal{C}_p^\bullet is an abstraction for \mathcal{C}
- ▶ The corresponding biased Markov chain \mathcal{C}' is decisive w.r.t. 😊
- ▶ The expected time to sample an execution is finite

p -divergence

Correctness of the approach

Theorem

Let \mathcal{C} be an LMC with level function λ , \mathcal{C}_p^\bullet the random walk of parameter p . Assume there is N_0 s.t. $\frac{1}{2} < p < p_+ = \inf\{P^+(s) \mid \lambda(s) > N_0\}$. Then:

- ▶ \mathcal{C}_p^\bullet is an abstraction for \mathcal{C}
- ▶ The corresponding biased Markov chain \mathcal{C}' is decisive w.r.t. 😊
- ▶ The expected time to sample an execution is finite

p -divergence

- ▶ Argument based on a variation on Foster's theorem:

Correctness of the approach

Theorem

Let \mathcal{C} be an LMC with level function λ , \mathcal{C}_p^\bullet the random walk of parameter p . Assume there is N_0 s.t. $\frac{1}{2} < p < p_+ = \inf\{P^+(s) \mid \lambda(s) > N_0\}$. Then:

- ▶ \mathcal{C}_p^\bullet is an abstraction for \mathcal{C}
- ▶ The corresponding biased Markov chain \mathcal{C}' is decisive w.r.t. 😊
- ▶ The expected time to sample an execution is finite

p -divergence

- ▶ Argument based on a variation on Foster's theorem:
 - If there is $\varepsilon > 0$ and a non-negative Lyapunov function \mathcal{L} s.t. for every $s \notin A$,
$$\mathcal{L}(s) - \sum_{s'} P(s, s') \mathcal{L}(s') \geq \varepsilon,$$
 then for all $s \notin A$, the expected time to A is finite, implying that A is an attractor

Reached
almost-surely

Correctness of the approach

Theorem

Let \mathcal{C} be an LMC with level function λ , \mathcal{C}_p^\bullet the random walk of parameter p . Assume there is N_0 s.t. $\frac{1}{2} < p < p_+ = \inf\{P^+(s) \mid \lambda(s) > N_0\}$. Then:

- ▶ \mathcal{C}_p^\bullet is an abstraction for \mathcal{C}
- ▶ The corresponding biased Markov chain \mathcal{C}' is decisive w.r.t. 😊
- ▶ The expected time to sample an execution is finite

p -divergence

- ▶ Argument based on a variation on Foster's theorem:
 - If there is $\varepsilon > 0$ and a non-negative Lyapunov function \mathcal{L} s.t. for every $s \notin A$,
$$\mathcal{L}(s) - \sum_{s'} P(s, s') \mathcal{L}(s') \geq \varepsilon,$$
 then for all $s \notin A$, the expected time to A is finite, implying that A is an attractor
- ▶ Apply this theorem to \mathcal{C}'

Reached
almost-surely

Correctness of the approach

Theorem

Let \mathcal{C} be an LMC with level function λ , \mathcal{C}_p^\bullet the random walk of parameter p . Assume there is N_0 s.t. $\frac{1}{2} < p < p_+ = \inf\{P^+(s) \mid \lambda(s) > N_0\}$. Then:

- ▶ \mathcal{C}_p^\bullet is an abstraction for \mathcal{C}
- ▶ The corresponding biased Markov chain \mathcal{C}' is decisive w.r.t. 😊
- ▶ The expected time to sample an execution is finite

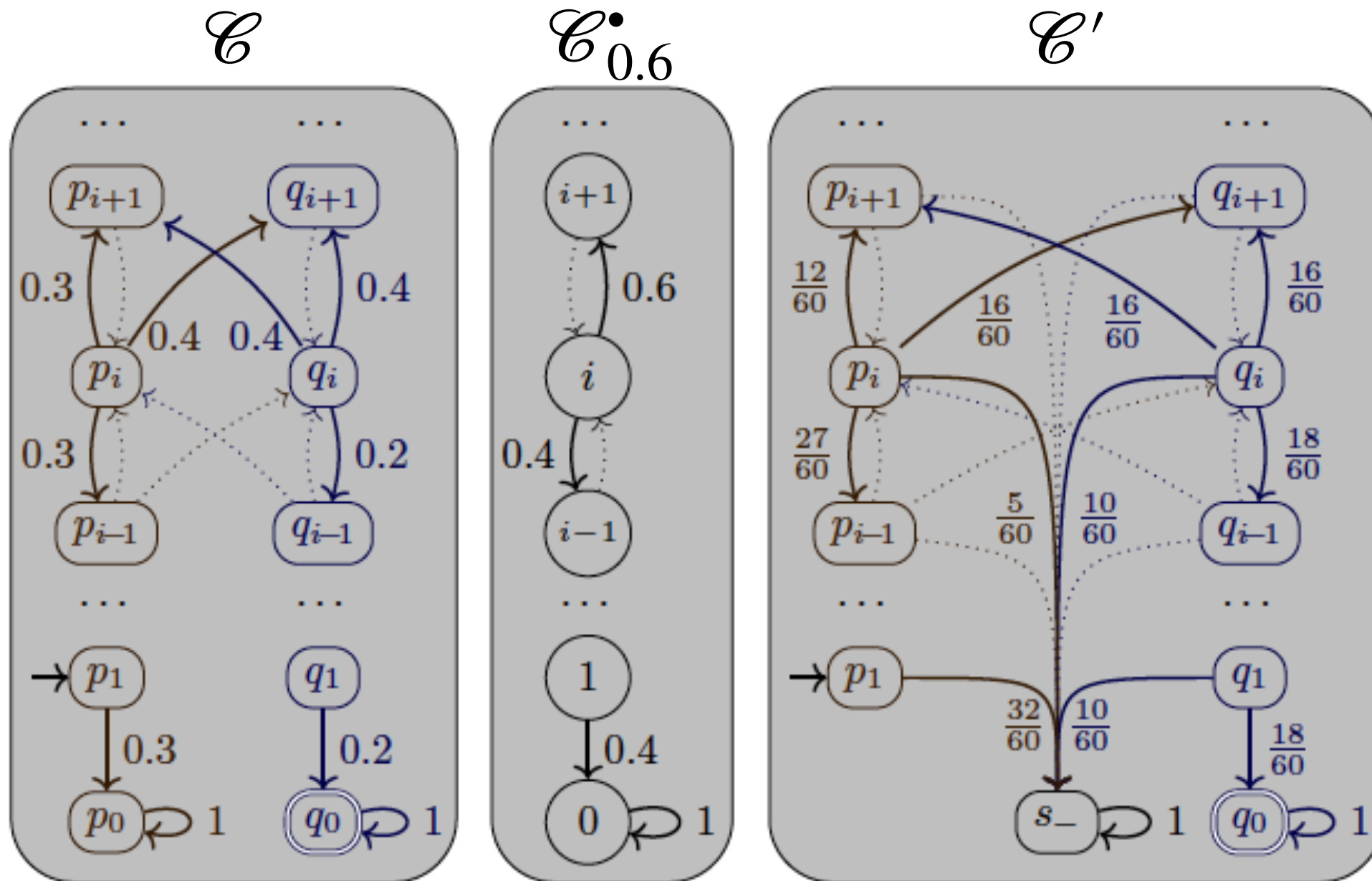
p -divergence

Congestion

- ▶ Argument based on a variation on Foster's theorem:
 - If there is $\varepsilon > 0$ and a non-negative Lyapunov function \mathcal{L} s.t. for every $s \notin A$,
$$\mathcal{L}(s) - \sum_{s'} P(s, s') \mathcal{L}(s') \geq \varepsilon,$$
 then for all $s \notin A$, the expected time to A is finite, implying that A is an attractor
- ▶ Apply this theorem to \mathcal{C}'

Reached
almost-surely

Example



\mathcal{C} is not decisive

\mathcal{C}' is decisive

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 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 p -divergent
 - Use the abstraction \mathcal{C}_p^\bullet

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 p -divergent
 - Use the abstraction \mathcal{C}_p^\bullet
 - 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 p -divergent
 - Use the abstraction \mathcal{C}_p^\bullet
 - Apply Approx and Estim on corresponding \mathcal{C}_p' (computed on-the-fly)

If \mathcal{C} is p -divergent, then \mathcal{C} is p' -divergent as soon as $1/2 < p' \leq p$

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 p -divergent
 - Use the abstraction \mathcal{C}_p^\bullet
 - Apply Approx and Estim on corresponding \mathcal{C}_p' (computed on-the-fly)

If \mathcal{C} is p -divergent, then \mathcal{C} is p' -divergent as soon as $1/2 < p' \leq p$

Is there a best 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 p -divergent
 - Use the abstraction \mathcal{C}_p^\bullet
 - Apply Approx and Estim on corresponding \mathcal{C}_p' (computed on-the-fly)

If \mathcal{C} is p -divergent, then \mathcal{C} is p' -divergent as soon as $1/2 < p' \leq p$

Is there a best p ?

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

First example

- ▶ State-free proba. pushdown automaton $\mathcal{C}: \{A \xrightarrow{1} C; A \xrightarrow{n} BB; B \xrightarrow{5} \varepsilon; B \xrightarrow{n} AA; C \xrightarrow{1} C\}$
- ▶ Start from A , and target the empty stack

First example

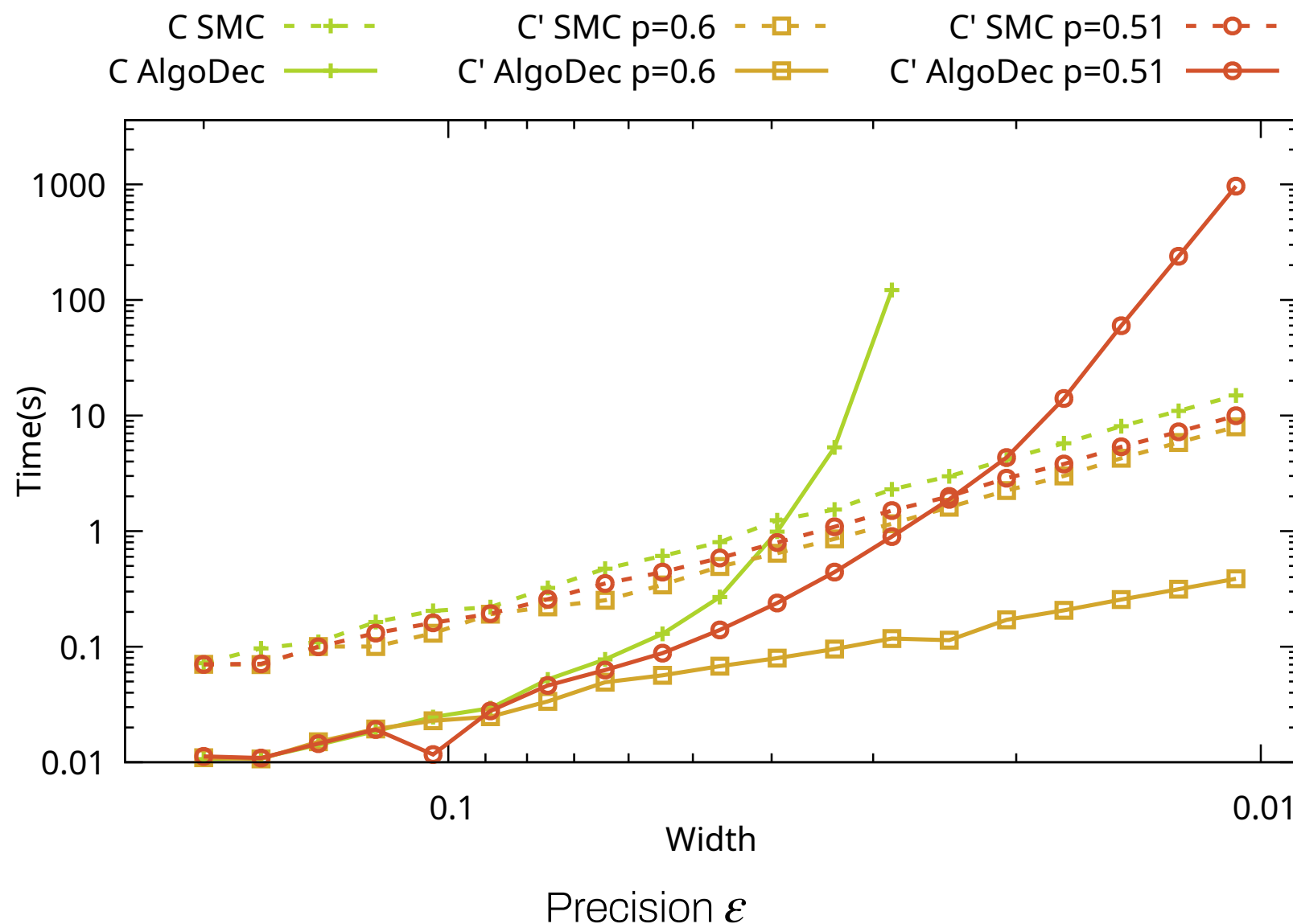
- ▶ State-free proba. pushdown automaton $\mathcal{C}: \{A \xrightarrow{1} C; A \xrightarrow{n} BB; B \xrightarrow{5} \varepsilon; B \xrightarrow{n} AA; C \xrightarrow{1} C\}$
- ▶ Start from A , and target the empty stack

It is decisive
It is p -divergent for every $1/2 < p < 1$

First example

- ▶ State-free proba. pushdown automaton $\mathcal{C}: \{A \xrightarrow{1} C; A \xrightarrow{n} BB; B \xrightarrow{5} \varepsilon; B \xrightarrow{n} AA; C \xrightarrow{1} C\}$
- ▶ Start from A , and target the empty stack

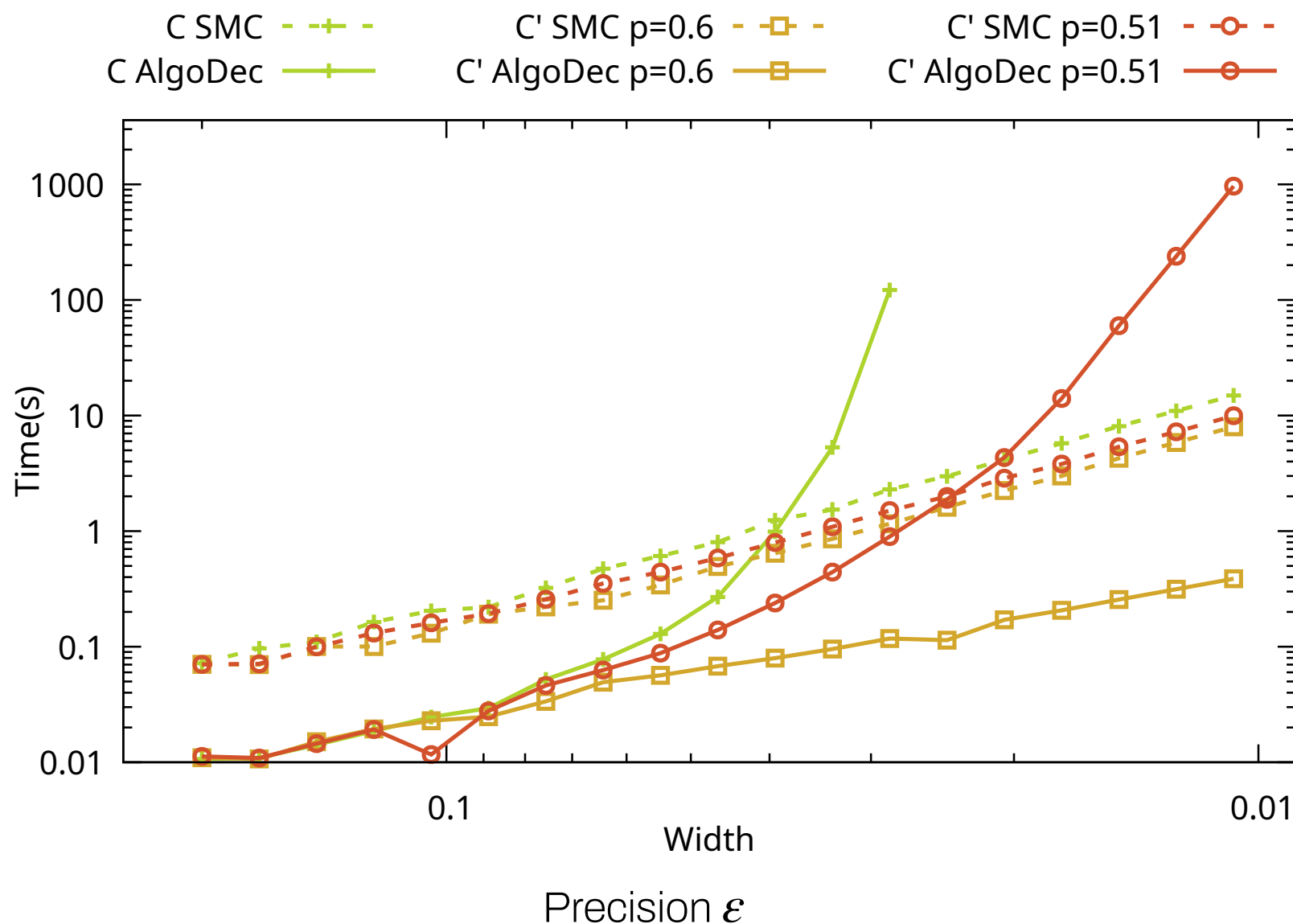
It is decisive
It is p -divergent for every $1/2 < p < 1$



First example

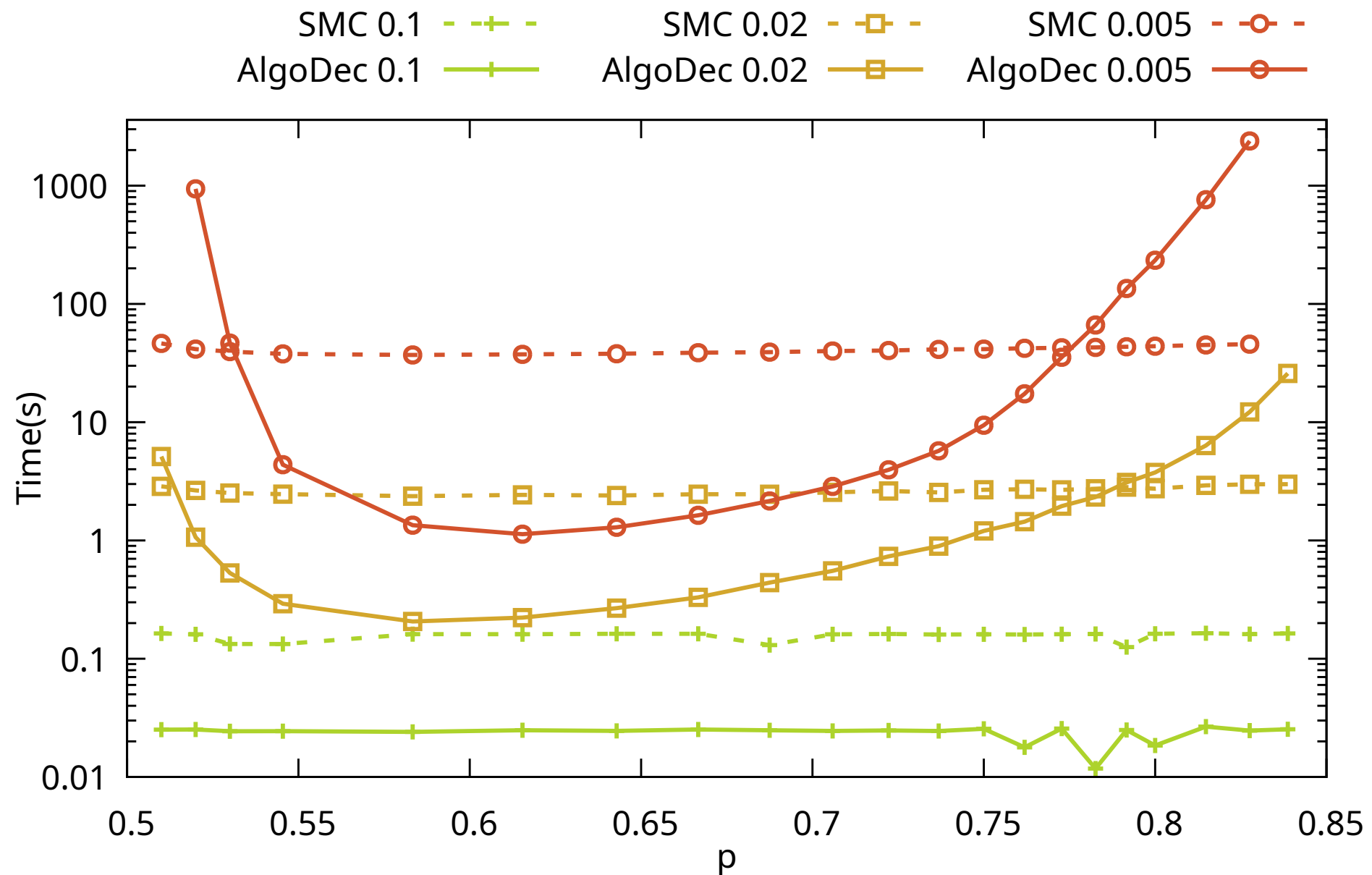
- ▶ State-free proba. pushdown automaton $\mathcal{C}: \{A \xrightarrow{1} C; A \xrightarrow{n} BB; B \xrightarrow{5} \varepsilon; B \xrightarrow{n} AA; C \xrightarrow{1} C\}$
- ▶ Start from A , and target the empty stack

It is decisive
It is p -divergent for every $1/2 < p < 1$



- ▶ In Estim (SMC): doubling the precision impacts in square on computation time (slope 2 in this log-log scale)
- ▶ Importance sampling seems to improve the analysis time, both for Approx and Estim (no formal guarantee, though)
- ▶ There seems to be « a best p » ($p = 0.6$ here)
- ▶ For that best p , Approx behaves very well!

First example — continued



Second example

Experimental results

- ▶ State-free proba. pushdown automaton \mathcal{C} :
 $\{A \xrightarrow{1} B; A \xrightarrow{1} C; B \xrightarrow{10} \varepsilon; B \xrightarrow{10+n} AA; C \xrightarrow{10} A; C \xrightarrow{10+n} BB\}$
- ▶ Start from A , and target the empty stack

Second example

Experimental results

- ▶ State-free proba. pushdown automaton \mathcal{C} :
 $\{A \xrightarrow{1} B; A \xrightarrow{1} C; B \xrightarrow{10} \varepsilon; B \xrightarrow{10+n} AA; C \xrightarrow{10} A; C \xrightarrow{10+n} BB\}$
- ▶ Start from A , and target the empty stack

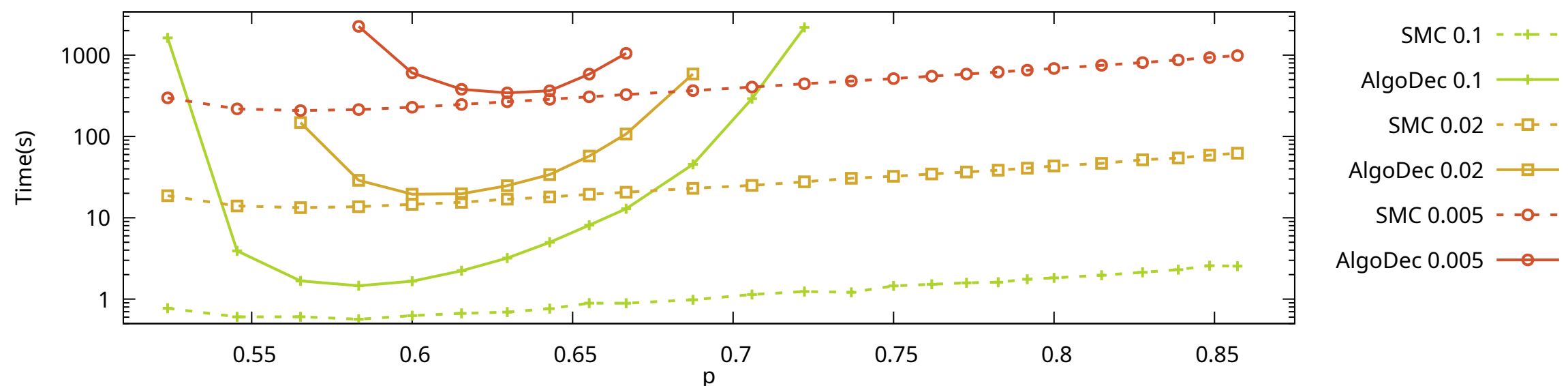
It is not decisive
It is p -divergent for every $1/2 < p < 1$

Second example

Experimental results

- State-free proba. pushdown automaton \mathcal{C} :
 $\{A \xrightarrow{1} B; A \xrightarrow{1} C; B \xrightarrow{10} \varepsilon; B \xrightarrow{10+n} AA; C \xrightarrow{10} A; C \xrightarrow{10+n} BB\}$
- Start from A , and target the empty stack

It is not decisive
 It is p -divergent for every $1/2 < p < 1$

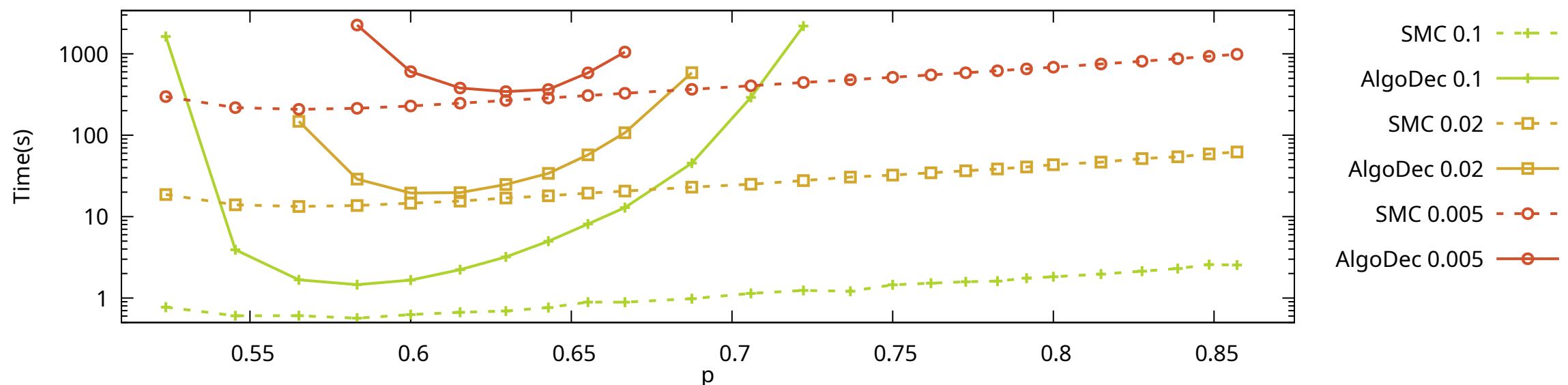


Second example

Experimental results

- ▶ State-free proba. pushdown automaton \mathcal{C} :
 $\{A \xrightarrow{1} B; A \xrightarrow{1} C; B \xrightarrow{10} \varepsilon; B \xrightarrow{10+n} AA; C \xrightarrow{10} A; C \xrightarrow{10+n} BB\}$
- ▶ Start from A , and target the empty stack

It is not decisive
 It is p -divergent for every $1/2 < p < 1$



- ▶ Estim-SMC not too sensitive to p
 - Nevertheless (log scale): clear bell effect on p
- ▶ Approx very sensitive to p

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
- ▶ Interesting **empirical results**

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?

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 »?

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 »?

Any theoretical justification for that?

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?