02246 - Model Checking

$$M \models \Phi ?$$

Lecture 9 – Statistical Model Checking

Technical University of Denmark

System

refine/implement

Requirements

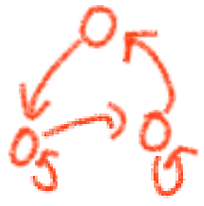
abstract/model

$M \models \Phi$?

some logic here

some smart algorithm here

Finite-state transition System



yes ☺

maybe 😐

no ☹

System

refine/implement

Requirements

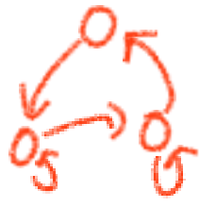
abstract/model

$M \models \Phi$?

some logic here

some algorithm here

Finite-state Transition System



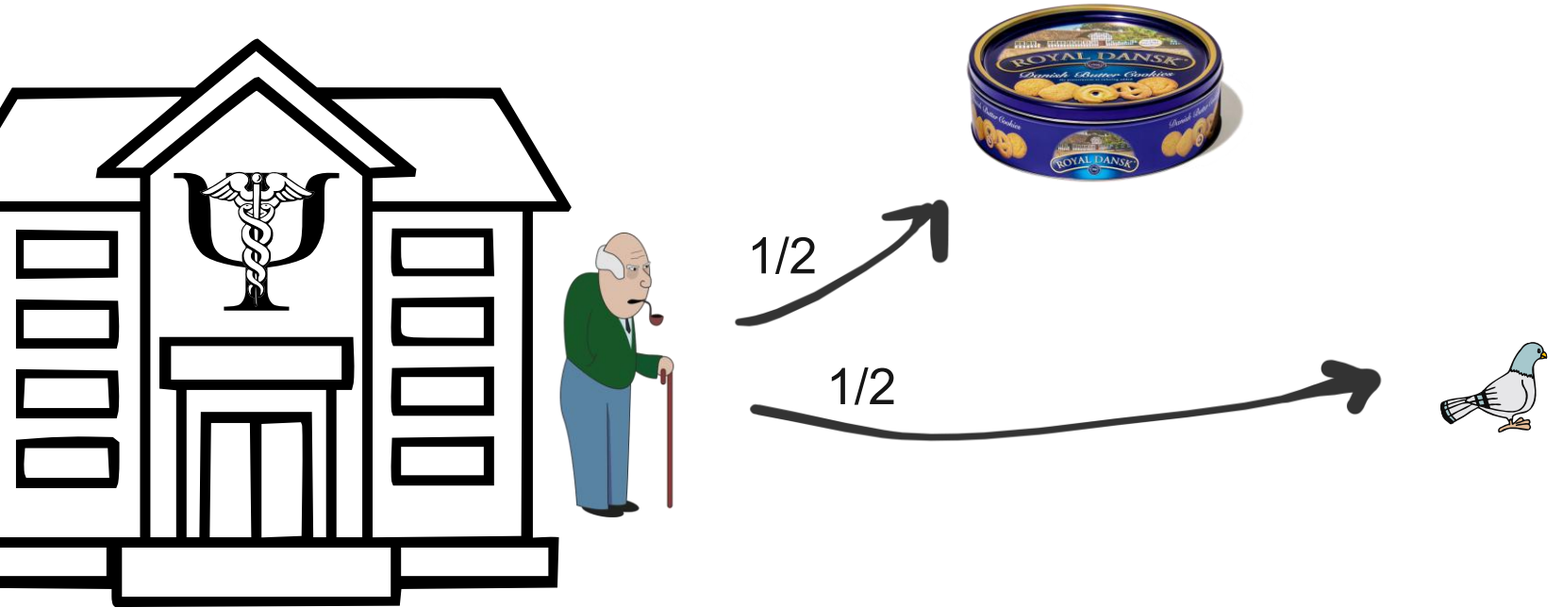
yes

maybe

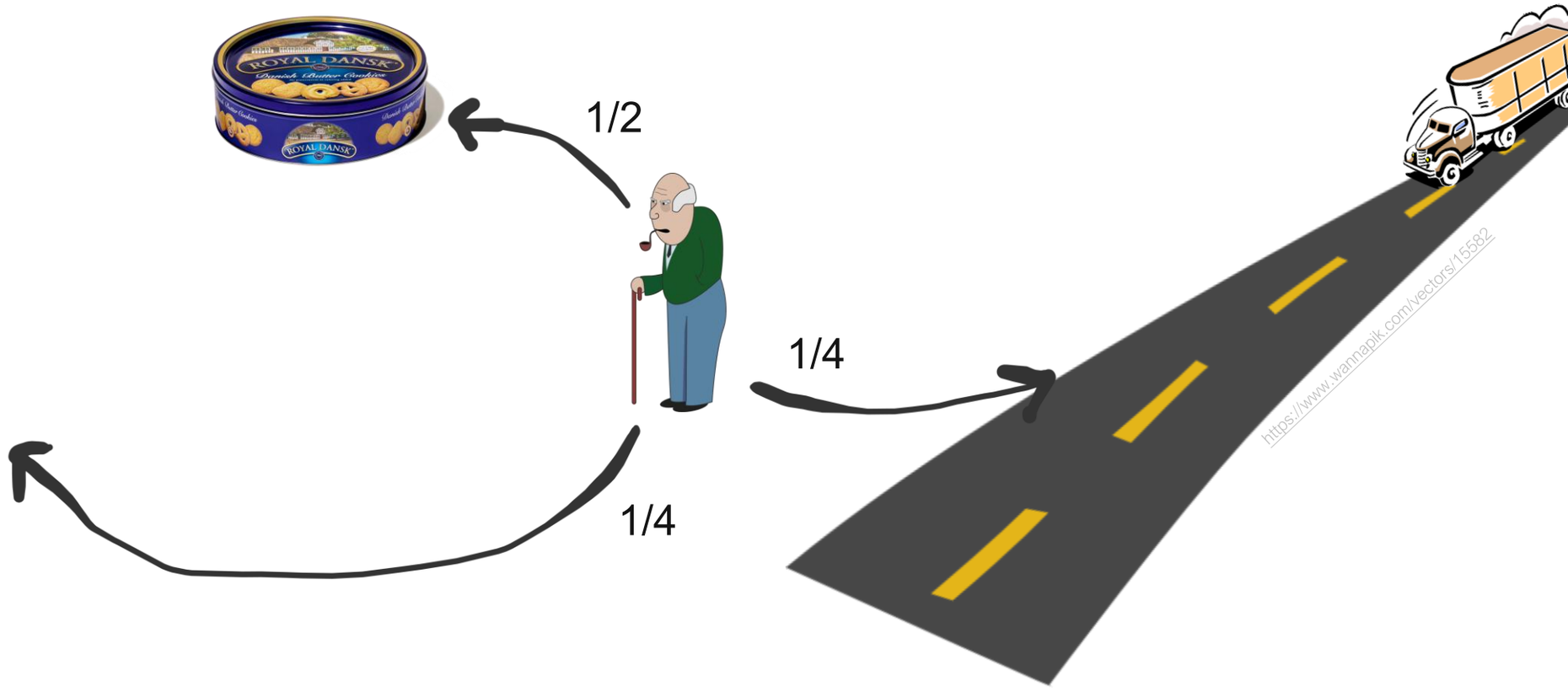
no?

kinda

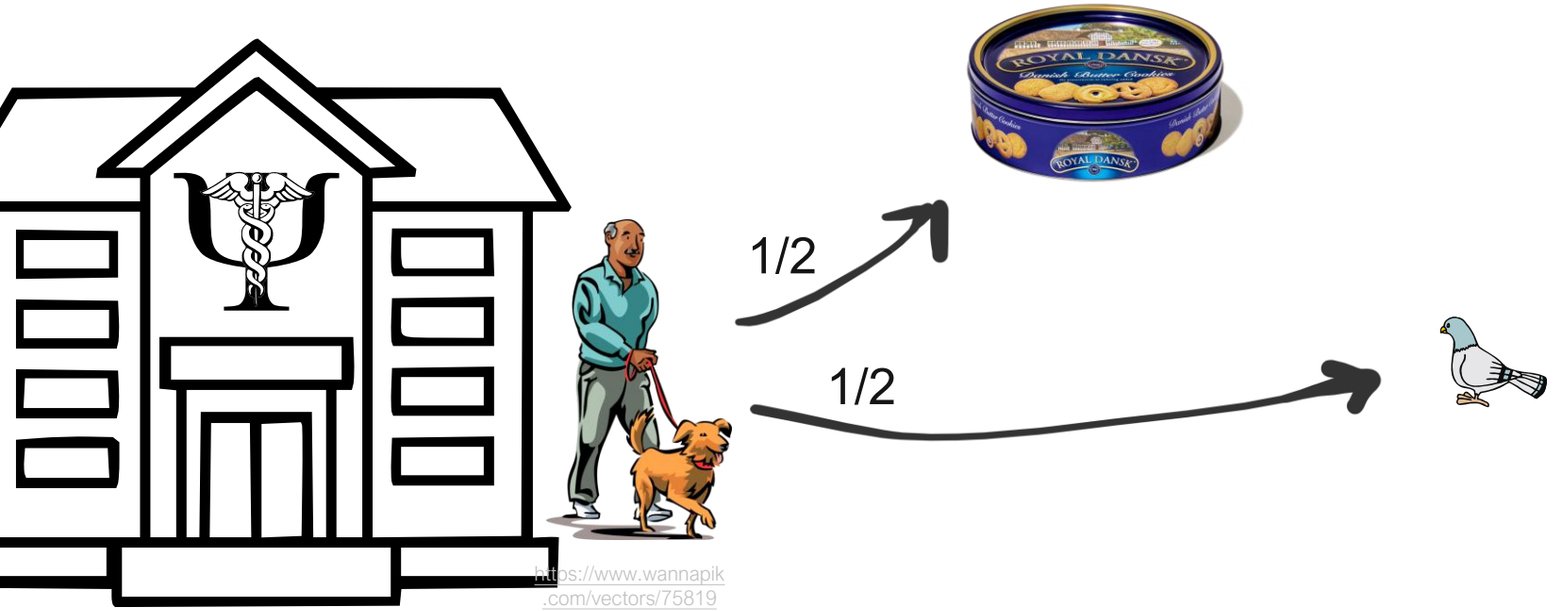
An appetiser



An appetiser



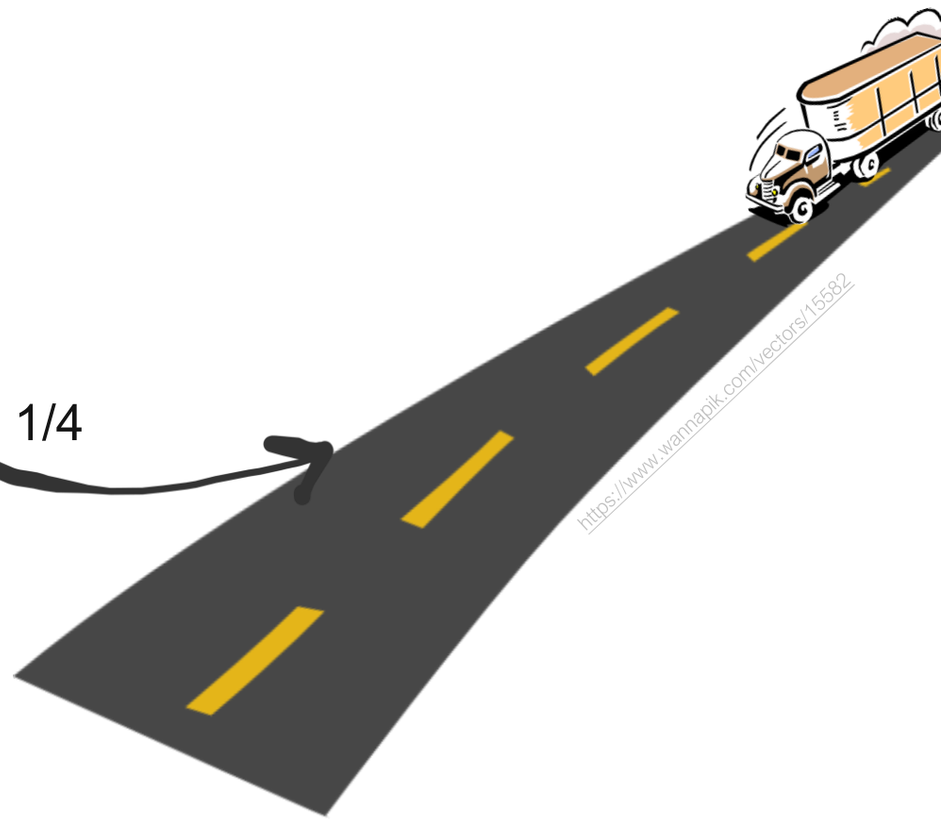
An appetiser



An appetiser

1/2

1/4

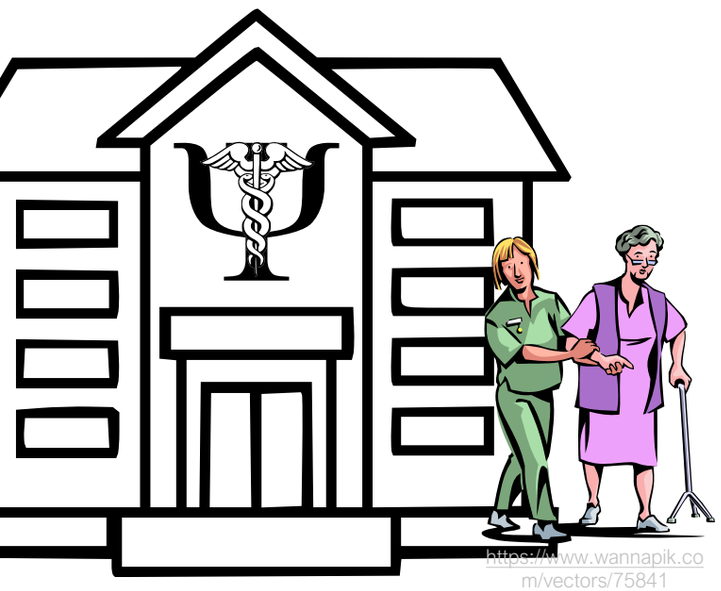


1/4

An appetiser

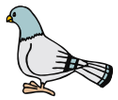
https://www.wannapik.com/vectors/75819

An appetiser



1/2

1/2



An appetiser

https://www.wannapik.com/vectors/75841

An appetiser

$$\mathbb{P}(🍪) = \frac{6}{7} \approx 0.86$$

1/2

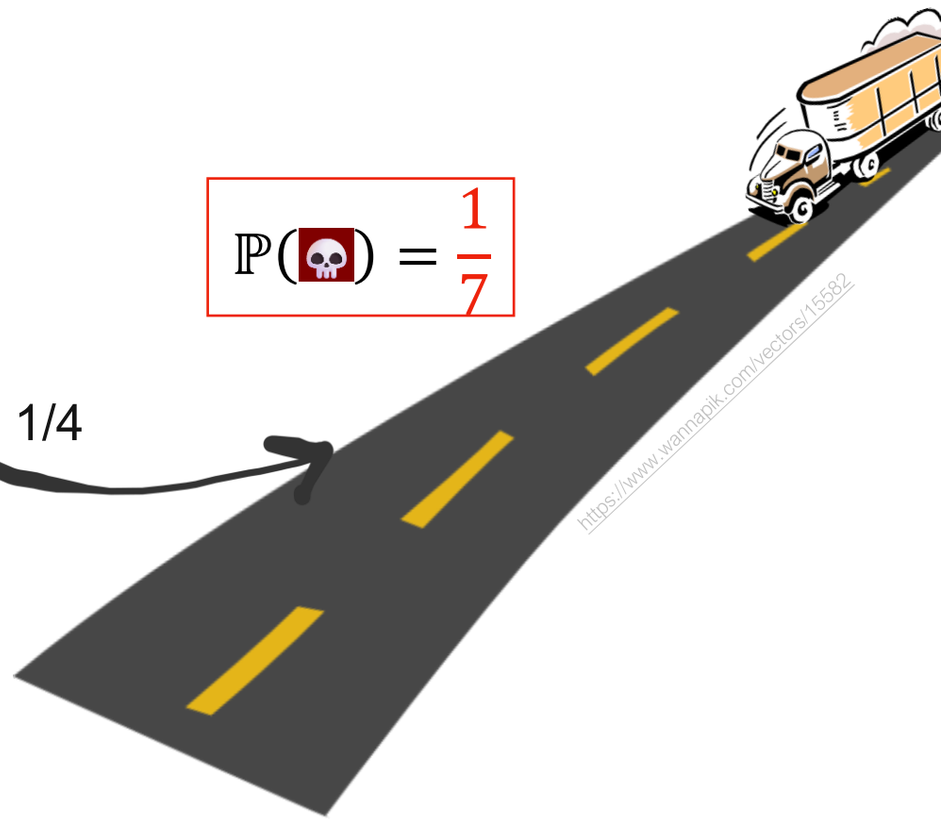
1/2

1/2

1/4

1/4

$$\mathbb{P}(☠) = \frac{1}{7}$$



$$\mathbb{P}(🍪) \approx \frac{\#succ}{\#tries} = \frac{2}{3} \approx 0.67$$

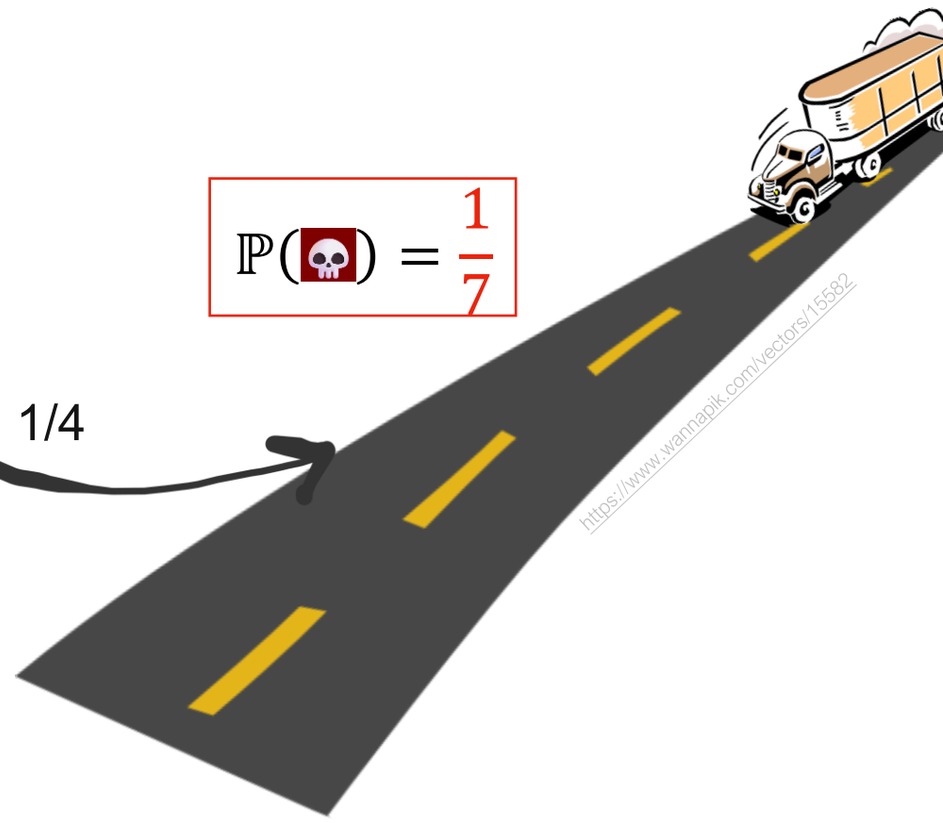
An appetiser

$$\mathbb{P}(🥨) = \frac{6}{7} \approx 0.86$$

```
dtmc

const COOKIES = 2;
const DED = 3;

module ASYLUM
   s:[0..3];
   [] s=0 -> 1/2: (s'=COOKIES)
           + 1/2: (s'=1);
   [] s=1 -> 1/2: (s'=COOKIES)
           + 1/4: (s'=0)
           + 1/4: (s'=DED);
endmodule
```

1/4

$$\mathbb{P}(☠) = \frac{1}{7}$$



$$\mathbb{P}(🥨) \approx \frac{\#succ}{\#tries} = \frac{2}{3} \approx 0.67$$

Key points of this lecture

Random Variables to generate traces from probabilistic models.

Statistical Model Checking is RNG-fueled Monte Carlo simulation, that avoids the state-space explosion problem of Model Checking.

Depends on the law of large numbers—nondeterminism not supported.

Can give statistical guarantees, not absolute guarantees.

SMC can be good for bounded properties (safety),
but impossible or hard for steady-state properties (liveness).

Rare events affect convergence of SMC.

Statistical Model Checking

- Random variables for trace generation
- Statistics for model checking
- Confidence intervals
- How to use, and when not to use
- Exercises & homework

Probability measures

Let Q be a non-empty set of "cases" and $\mathcal{G}$ be a set of "events" such that

- $\mathcal{G} \subseteq 2^{\mathrm{Q}}$
- $\emptyset \in \mathcal{G}$
- $\mathcal{G}$ is closed under countable union and complement

σ-algebra

A function $\mu : \mathcal{G} \to [0,1]$ is a **probability measure** if it satisfies the following conditions

- $\mu(\bigcup_{i \geq 0} C_i) = \sum_{i \geq 0} \mu(C_i)$ for *pairwise disjoint* sets $C_0, C_1, \ldots$
- $\mu(\mathrm{Q} \setminus C) = 1 - \mu(C)$
- $\mu(\emptyset) = 0$

$(\mathrm{Q}, \mathcal{G})$ is called a ***measurable space***
$(\mathrm{Q}, \mathcal{G}, \mu)$ is called a ***probability space***

Random variables

Let $(\mathrm{Q}, \mathcal{G}, \mu)$ be a probability space → "sampling space"

and $(\mathrm{E}, \mathcal{E})$ be a measurable space → "events space"

A random variable is a measurable function $X: \mathrm{Q} \to \mathrm{E}$

That is, for any event $A \subseteq \mathrm{E}$, if $A \in \mathcal{E}$ then $X^{-1}(A) \in \mathcal{G}$

Example: betting with dice in the casino of Monte Carlo

- Sampling a fair die: $\mathrm{D} = \{1, 2, \dots, 6\}$ with $Pr(\{d\}) = \frac{1}{6}$ $\forall d \in \mathrm{D}$
- Sampling two dice: $\mathrm{Q} = D^2 = \{(1,1), (1,2), \dots, (6,6)\}$
  with $\mu(d_1, d_2) = \frac{1}{36}$
- Let $X(d_1, d_2) = d_1 + d_2$
- ^^^ possible events: $\mathrm{E} = \{2, 3, \dots, 12\}$

Random variables

Let $(\mathrm{Q}, \mathcal{G}, \mu)$ be a probability space → "sampling space"

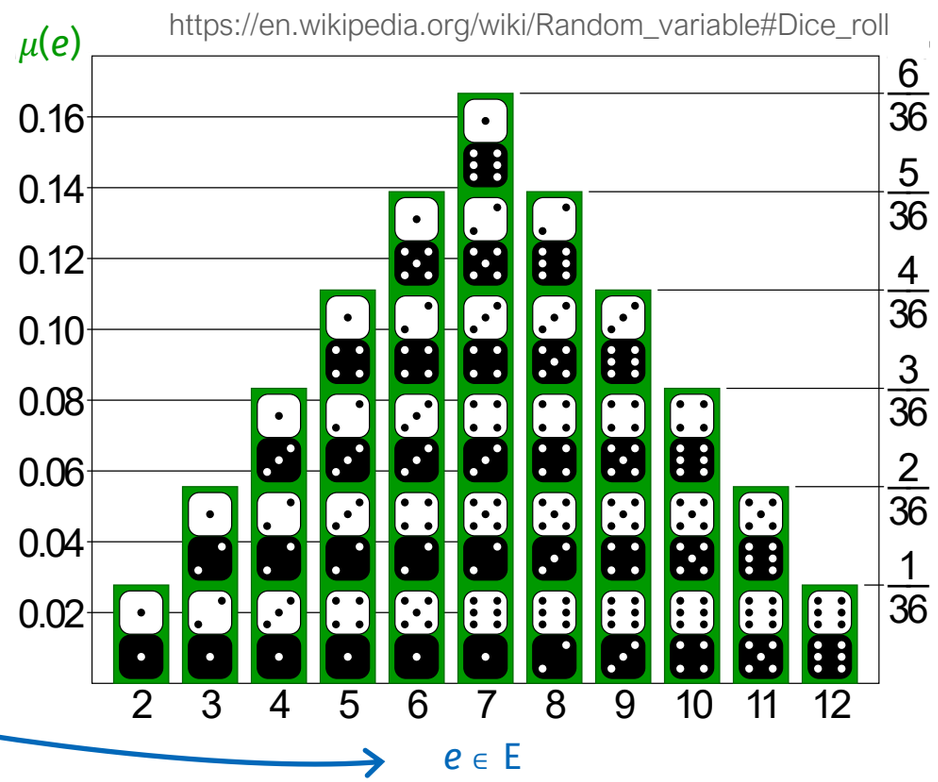
and $(\mathrm{E}, \mathcal{E})$ be a measurable space → "events space"

A random variable is a measurable function $X: \mathrm{Q} \to \mathrm{E}$

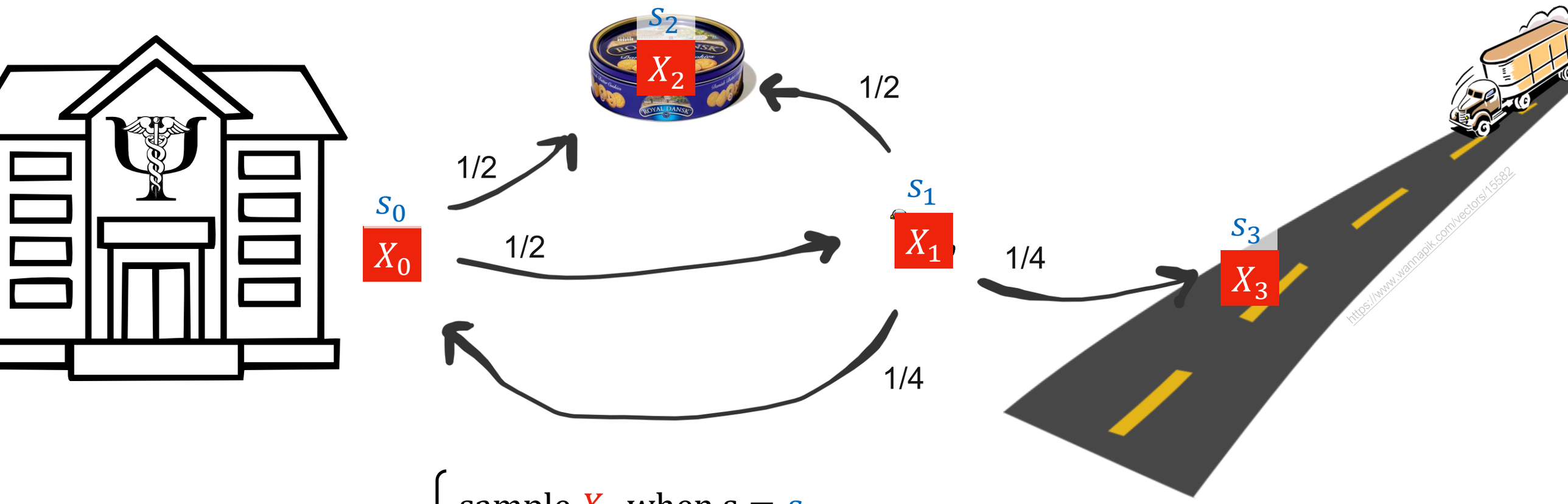
That is, for any event $A \subseteq \mathrm{E}$, if $A \in \mathcal{E}$ then $X^{-1}(A) \in \mathcal{G}$

Example: betting with dice in the casino of Monte Carlo

- Sampling a fair die: $\mathrm{D} = \{1, 2, \ldots, 6\}$ with $Pr(\{d\}) = \frac{1}{6}$ $\forall d \in \mathrm{D}$
- Sampling two dice: $\mathrm{Q} = D^2 = \{(1,1), (1,2), \ldots, (6,6)\}$
  with $\mu(d_1, d_2) = \frac{1}{36}$
- Let $X(d_1, d_2) = d_1 + d_2$
- ^^^ possible events: $\mathrm{E} = \{2, 3, \ldots, 12\}$

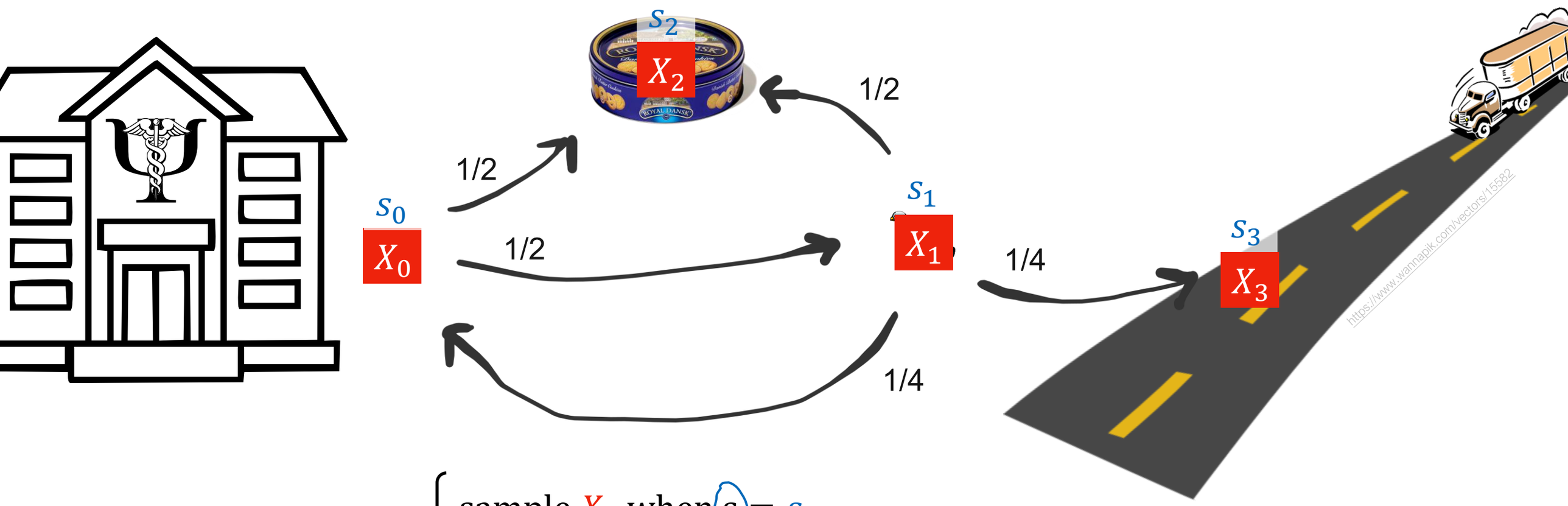


Random variables for trace generation



$$X = \{X_i \mid \mathrm{s} = s_\mathrm{i}\} = \begin{cases} \text{sample } X_0 \text{ when s} = s_0 \\ \quad \ldots \\ \text{sample } X_3 \text{ when s} = s_3 \end{cases}$$

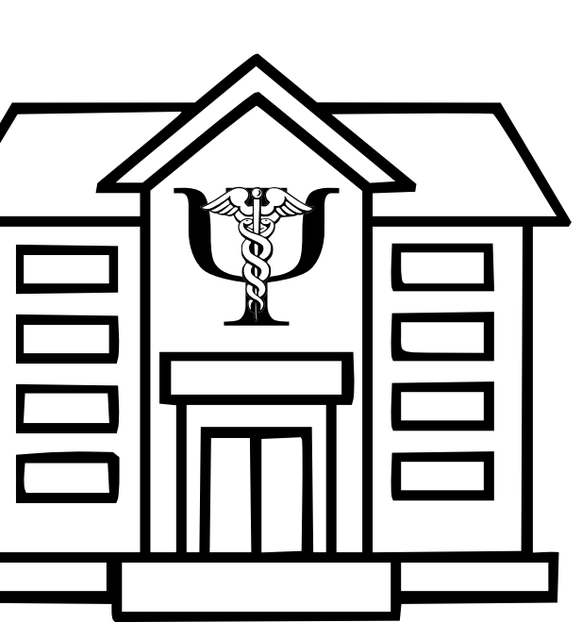
Random variables for trace generation

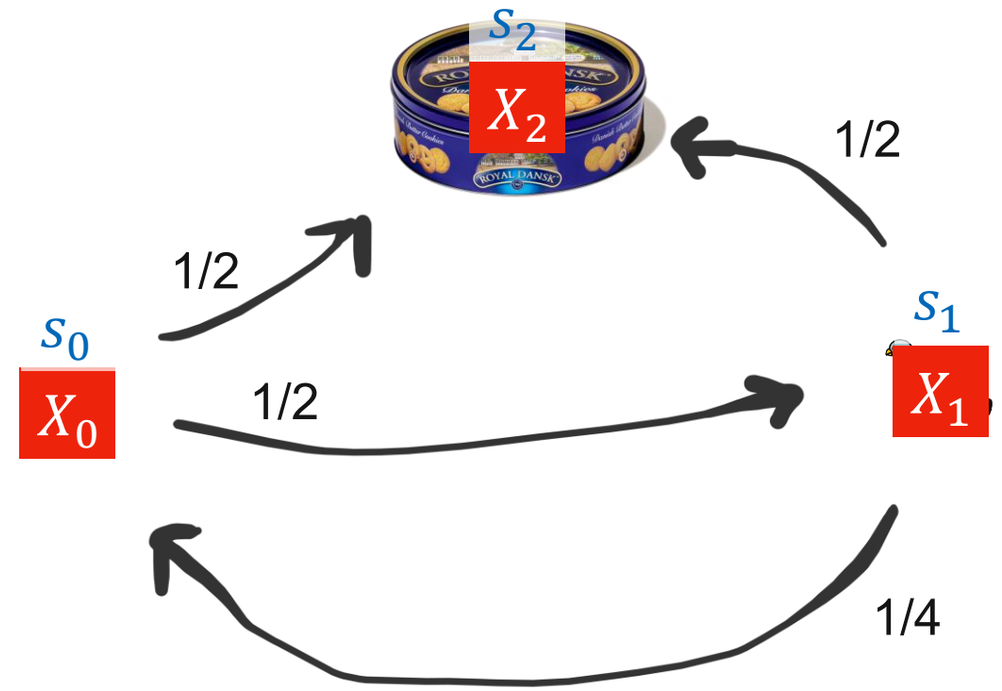


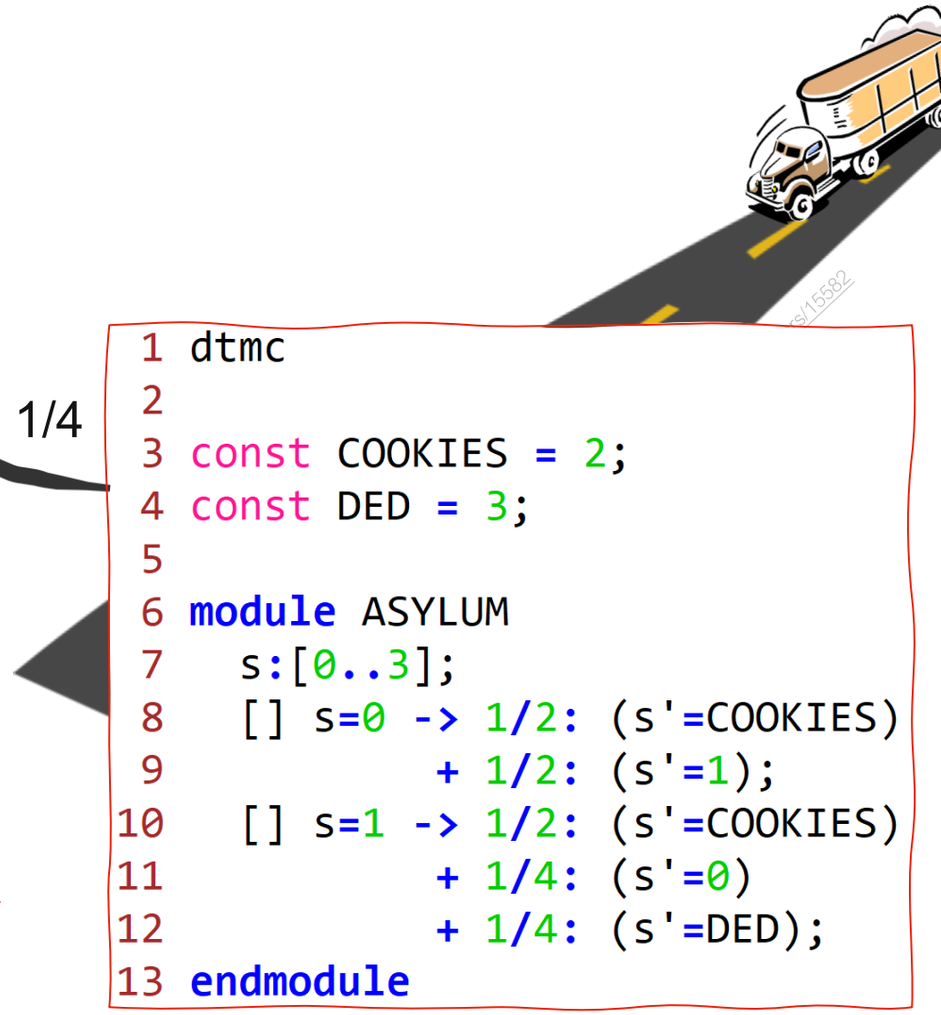
$$X = \{X_i \mid s = s_i\} = \begin{cases} \text{sample } X_0 \text{ when } s = s_0 \\ \quad \ldots \\ \text{sample } X_3 \text{ when } s = s_3 \end{cases}$$

Only keep current state (constant memory)

Random variables for trace generation







```
1  dtmc
2
3  const COOKIES = 2;
4  const DED = 3;
5
6  module ASYLUM
7    s:[0..3];
8    [] s=0 -> 1/2: (s'=COOKIES)
9            + 1/2: (s'=1);
10   [] s=1 -> 1/2: (s'=COOKIES)
11           + 1/4: (s'=0)
12           + 1/4: (s'=DED);
13 endmodule
```

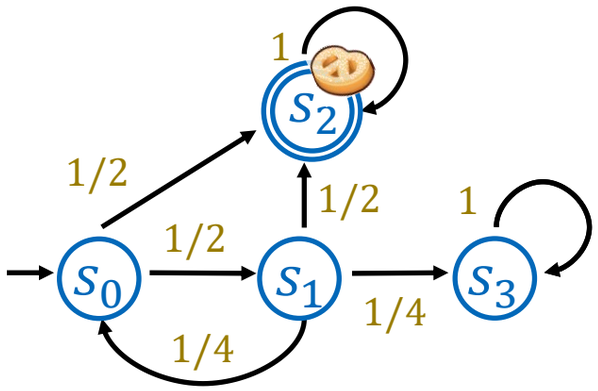
$$X = \{X_i \mid s = s_i\} = \begin{cases} \text{sample } X_0 \text{ when } s = s_0 \\ \quad \ldots \\ \text{sample } X_3 \text{ when } s = s_3 \end{cases}$$

- A probabilistic model can be seen as (very complex) random variable X
- Sampling X—generating traces—traverses the probability transition matrix of the model, one observation at a time
- Observing ***enough samples*** gives a fair idea of the general model behaviour

Statistical Model Checking

- Random variables for trace generation
- Statistics for model checking
- Confidence intervals
- How to use, and when not to use
- Exercises & homework

Counting for statistical estimation



$\mathbb{P}_{s_3}(🥨) = ?$ $\quad \mathbb{E}[X_3] = 1 \cdot s_3 \quad$ $\mathbb{E}[Y(X_3)] = 0$

$\mathbb{P}_{s_2}(🥨) = ?$ $\quad \mathbb{E}[X_2] = 1 \cdot s_2 \quad$ $\mathbb{E}[Y(X_2)] = 1$

$\mathbb{P}_{s_1}(🥨) = ?$ $\quad \mathbb{E}[X_1] = \frac{1}{4}s_3 + \frac{1}{2}s_2 + \frac{1}{4}s_0 \quad$ $\mathbb{E}[Y(X_1)] = \cdots$

$\mathbb{P}(🥨) = \mathbb{P}_{s_0}(🥨) = ?$

In terms of X, the expected observation if $s = s_3$ is $s' = s_3$

$$X = \{X_i \mid s = s_i\} = \begin{cases} \text{sample } X_0 \text{ if } s = s_0 \\ \quad \ldots \\ \text{sample } X_3 \text{ if } s = s_3 \end{cases}$$

$Y : S \to \{0,1\}$

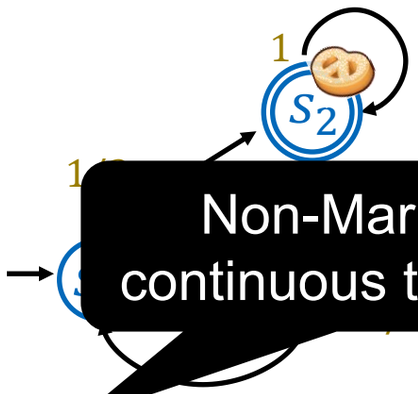
$Y(s) \doteq (s = s_2)\ ?\ 1 : 0$

$\mathbb{E}[Y(X_0)] = \frac{1}{2}1 + \frac{1}{2}\mathbb{E}[Y(X_1)]$

$\mathbb{E}[Y(X_1)] = \frac{1}{4}0 + \frac{1}{2}1 + \frac{1}{4}\mathbb{E}[Y(X_0)]$

~~Counting for statistical estimation~~

Good'ol probabilistic model checking



Non-Markovian continuous transitions

$\mathbb{P}_{s_3}(🥨) = ?$ $\quad \mathbb{E}[X_3] = 1 \cdot s_3 \quad \mathbb{E}[Y(X_3)] = 0$

$\mathbb{P}_{s_2}(🥨) = ?$ $\quad \mathbb{E}[X_2] = 1 \cdot s_2 \quad \mathbb{E}[Y(X_2)] = 1$

$\mathbb{P}_{s_1}(🥨) = ?$ $\quad \mathbb{E}[X_1] = \frac{1}{4} s_3 \quad \mathbb{E}[Y(X_1)] = \cdots$

Too-large (e.g. infinite) state space

$\mathbb{P}(🥨) = \mathbb{P}_{s_0}(🥨) = ?$

In terms of X, the expected observation if $s = s_3$ is $s' = s_3$

$$X = \{X_i \mid s = s_i\} = \begin{cases} \text{sample } X_0 \text{ if } s = s_0 \\ \quad \ldots \\ \text{sample } X_3 \text{ if } s = s_3 \end{cases}$$

$Y : S \to \{0,1\}$

$Y(s) \doteq (s = s_2)\ ?\ 1 : 0$

$\mathbb{E}[Y(X_0)] = \frac{1}{2} 1 + \frac{1}{2} \mathbb{E}[Y(X_1)]$

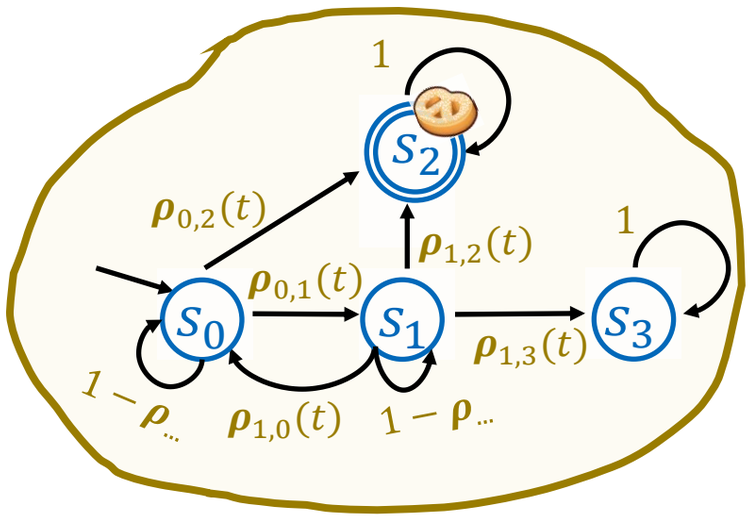
$\mathbb{E}[Y(X_1)] = \frac{1}{4} 0 + \frac{1}{2} 1 + \frac{1}{4} \mathbb{E}[Y(X_0)]$

Exact solution

$\mathbb{P}(🥨) = \frac{6}{7}$

Counting for statistical estimation

$\boldsymbol{\rho}_{\mathrm{i,j}}(t)$ is an aperiodic function on the number of transition steps taken, $t \in \mathbb{N}$



$\mathbb{P}_{s_3}(🥨) = ?$

$\mathbb{P}_{s_2}(🥨) = ?$

$\mathbb{P}_{s_1}(🥨) = ?$

$\mathbb{P}(🥨) = \mathbb{P}_{s_0}(🥨) = ?$

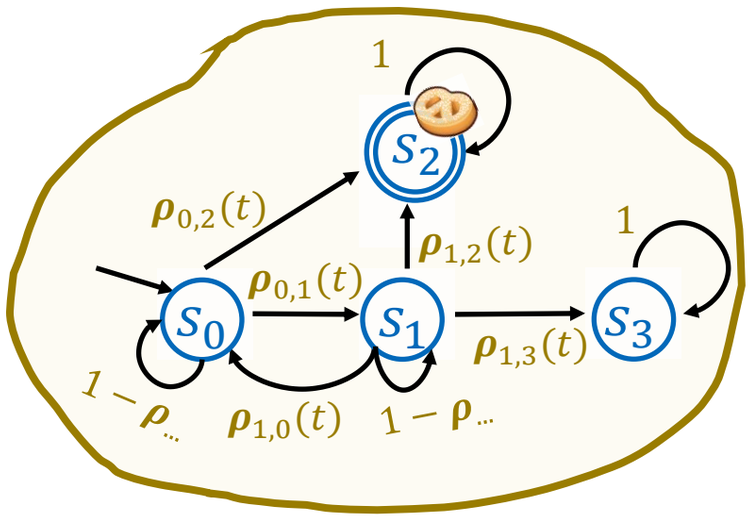
$$X = \{X_i \mid \mathrm{s} = s_{\mathrm{i}}\} = \begin{cases} \text{sample } X_0 \text{ if } \mathrm{s} = s_0 \\ \quad \ldots \\ \text{sample } X_3 \text{ if } \mathrm{s} = s_3 \end{cases}$$

$$Y : S \to \{0,1\}$$

$$Y(\mathrm{s}) \doteq (\mathrm{s} = s_2)\ ?\ 1 : 0$$

Counting for statistical estimation

➡ $\boldsymbol{\rho}_{\mathrm{i,j}}(t)$ is an aperiodic function on the number of transition steps taken, $t \in \mathbb{N}$ ⬅



$\mathbb{P}_{s_3}(🍪) = ?$

$\mathbb{P}_{s_2}(🍪) = ?$

$\mathbb{P}_{s_1}(🍪) = ?$

I wonder how does the probability transition matrix look like…

$\mathbb{P}(🍪) = \mathbb{P}_{s_0}(🍪) = ?$

Besides the current state, s, we must know the number of transitions steps taken, t

$$X = \{X_i \mid s = s_i\} = \begin{cases} \text{sample } X_0 \text{ if } s = s_0 \\ \quad \dots \\ \text{sample } X_3 \text{ if } s = s_3 \end{cases}$$

$$Y : S \to \{0,1\}$$

$$Y(s) \doteq (s = s_2)\ ?\ 1 : 0$$

$$X \to \text{sample } \boldsymbol{\rho}_{\mathrm{i},0}(t'), \dots, \boldsymbol{\rho}_{\mathrm{i},3}(t') \text{ if } s = s_i \text{ and } t = t'$$

Probabilities as expected values

The sample space (Π_M, cyl_M, μ) is built on the cylinder set of paths from model M,

and our r.v. Y maps paths to $(\{0,1\}, \mathcal{E})$ as follows:

$$Y(\pi) \doteq (\text{🥨} \in \pi)\ ?\ 1 : 0$$

Then $\mathbb{P}(\text{🥨}) = \mathbb{E}[Y]$.

$$\mathbb{E}[X] \doteq \sum_{i=0}^{\cancel{n-1}\ \infty} p_i\, x_i$$

$$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Law of large numbers (statistical estimation for model checking)

Let $\{\pi_i\}_{i=0}^{n-1}$ be a set of finite paths sampled ***independently*** from model M,

e.g. each π_i is a trace that starts in the initial state and has length 10000,

and let $Y_i \doteq Y(\pi_i) \in \{0,1\}$.

Then **the empirical mean is an unbiased estimator for the expected value** of Y:

$$\overline{Y}_n \qquad \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} Y_i \longrightarrow \mathbb{E}[Y]$$

$$\forall\, \varepsilon > 0\ .\ \exists\, n \in \mathbb{N}\ .\ |\, \overline{Y}_n - \mathbb{E}[Y]| < \varepsilon$$

We can get arbitrarily close to $\mathbb{P}(\text{🥨})$, just keep sampling

Statistical Model Checking

- Random variables for trace generation
- Statistics for model checking
- Confidence intervals
- How to use, and when not to use
- Exercises & homework

Central limit theorem

By the law of large numbers $\forall\, \varepsilon > 0\,.\exists\, n \in \mathbb{N}\,.\, |\overline{Y}_n - \mathbb{E}[Y]| < \varepsilon.$

Which n do we need to get $\varepsilon = 0.01$ close to $\mathbb{E}[Y]$?

Point estimate 💩

Let $Z \doteq \frac{Y - \mathbb{E}[Y]}{\sqrt{\mathbb{V}[Y]}}$. Then $\mathbb{E}[Z] = 0$ and $\mathbb{V}[Z] = 1$.

Central Limit Theorem (for our statistical estimation)

Let $\overline{Y}_n$ be the empirical mean of sample size n for random variable Y, with positive finite variance $0 < \mathbb{V}[Y] = \sigma^2 < \infty$.

Known values

Then $\lim_{n\to\infty} \mathbb{P}\left(\frac{\overline{Y}_n - \mathbb{E}[Y]}{\sigma/\sqrt{n}} \le \frac{z}{\sigma}\right) \longrightarrow \Phi\left(\frac{z}{\sigma}\right)$, where $\Phi(\cdot)$ is the standard normal dist.

$$\mathbb{P}\left(-z_{\alpha/2} \le \frac{\overline{Y}_n - \mathbb{E}[Y]}{\sigma/\sqrt{n}} \le z_{\alpha/2}\right) \approx 1 - \alpha \quad \Leftrightarrow \quad \overline{Y}_n \pm z_{\alpha/2}\frac{S}{\sqrt{n}}$$ covers $\mathbb{E}[Y]$ with prob. $1 - \alpha$

$\mathbb{P}($🥯$)$

Central limit theorem

By the law of large numbers $\forall\, \varepsilon > 0\,.\, \exists\, n \in \mathbb{N}\,.\, |\overline{Y}_n - \mathbb{E}[Y]| < \varepsilon$.

Which n do we need to get $\varepsilon = 0.01$ close to $\mathbb{E}[Y]$?

Point estimate 💩

Let $Z \doteq \frac{Y - \mathbb{E}[Y]}{\sqrt{\mathbb{V}[Y]}}$. Then $\mathbb{E}[Z] = 0$ and $\mathbb{V}[Z] = 1$.

Central Limit Theorem (for our statistical estimation)

Let $\overline{Y}_n$ be the empirical mean of sample size n for random variable Y, with positive finite variance $0 < \mathbb{V}[Y] = \sigma^2 < \infty$.

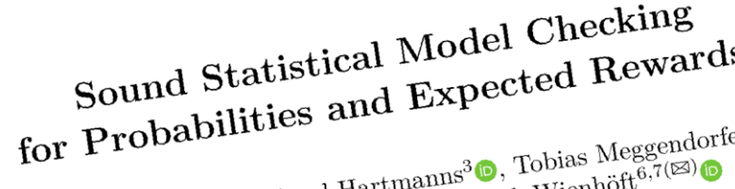
Known values

Then $\lim_{n\to\infty} \mathbb{P}\left(\frac{\overline{Y}_n - \mathbb{E}[Y]}{\sigma/\sqrt{n}} \leq \frac{z}{\sigma}\right) \longrightarrow \Phi\left(\frac{z}{\sigma}\right)$, where $\Phi(\cdot)$ is the standard normal dist.

$$\mathbb{P}\left(-z_{\alpha/2} \leq \frac{\overline{Y}_n - \mathbb{E}[Y]}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right) \approx 1 - \alpha \quad \Leftrightarrow \quad \overline{Y}_n \pm z_{\alpha/2}\frac{S}{\sqrt{n}} \text{ covers } \mathbb{E}[Y] \text{ with prob. } 1 - \alpha$$

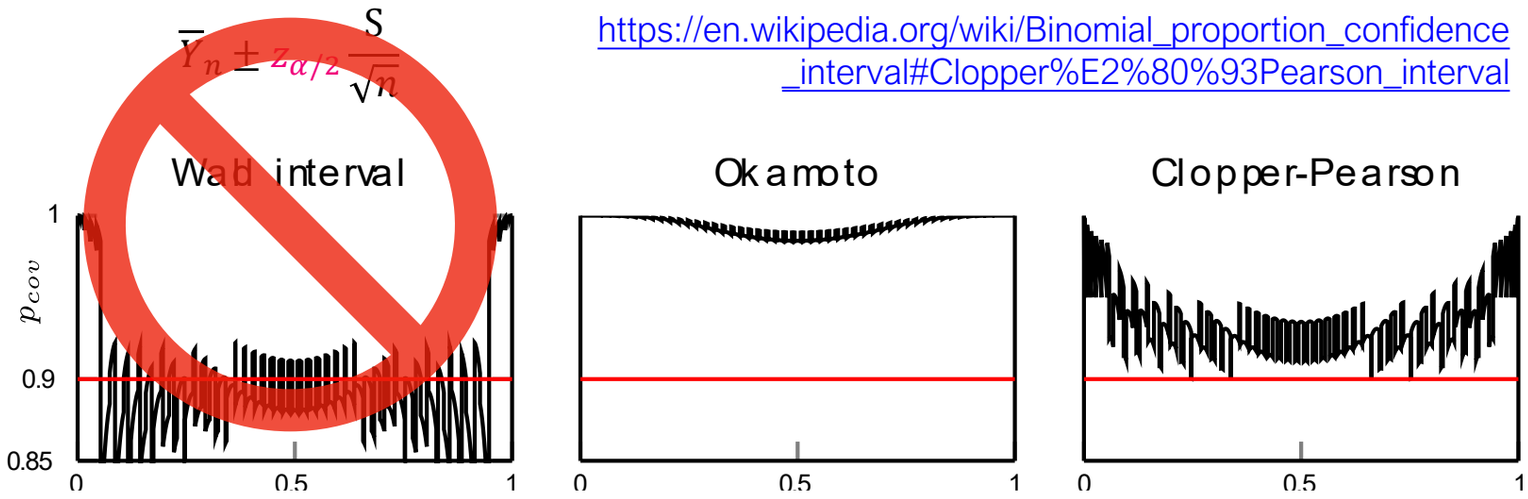
… but this way of building the interval is unsound 💩

$\mathbb{P}(🥨)$


Sound Statistical Model Checking
for Probabilities and Expected Rewards

Confidence intervals

- $\overline{Y}_n \doteq \frac{1}{n}\sum_{i=0}^{n-1} Y_i$
- $CI_{w,\alpha} = [l, u] \ni \overline{Y}_n$
- $w = u - l$ *smaller is better*
- $\alpha \doteq \mathbb{P}\left(CI_{w,\alpha} \not\ni \mathbb{E}[Y]\right)$

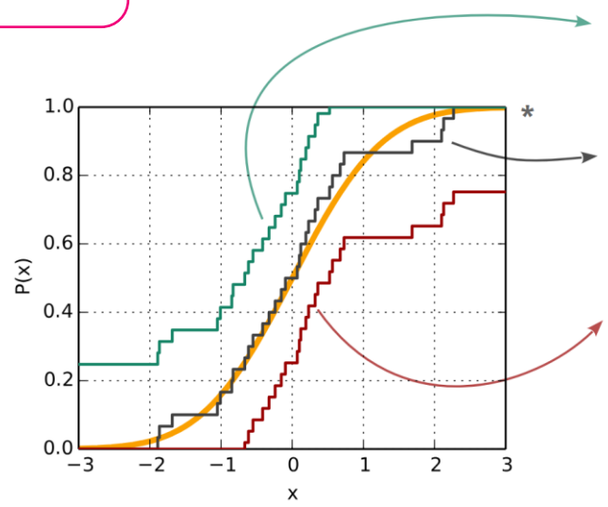


https://en.wikipedia.org/wiki/Binomial_proportion_confidence_interval#Clopper%E2%80%93Pearson_interval

Confidence interval for expected rewards

- $\overline{R}_n \doteq \frac{1}{n}\sum_{i=0}^{n-1} R_i$
- $CI_{w,\alpha}$ given by the DKW inequality

reward accumulated in i-th simulation run



$$\underline{F}(x) = \min\left\{1, \hat{F}(x) + \sqrt{\tfrac{1}{2k}\ln\tfrac{2}{\delta}}\right\}$$

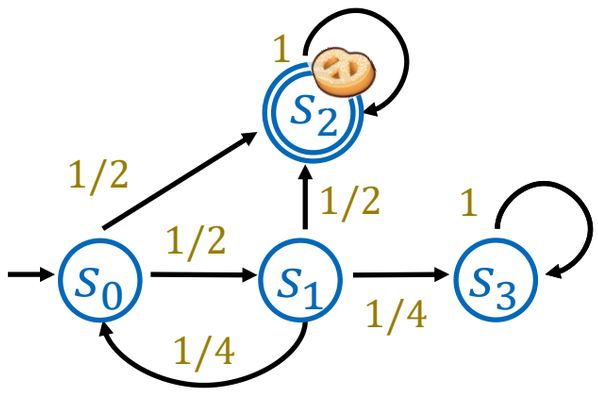
$$\hat{F}(x) \doteq \frac{1}{k}\sum_{i=1}^{k} \mathbb{1}_{X_i \le x} \qquad \underline{F}(x) \le F(x) \le \overline{F}(x)$$

$$\overline{F}(x) = \max\left\{0, \hat{F}(x) - \sqrt{\tfrac{1}{2k}\ln\tfrac{2}{\delta}}\right\}$$

Statistical Model Checking

- Random variables for trace generation
- Statistics for model checking
- Confidence intervals
- How to use, and when not to use
- Exercises & homework

Transient probability estimation ✅



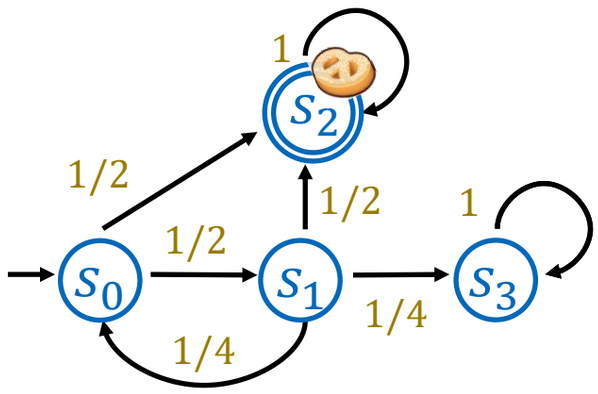
- ✅ $\mathbb{P}(\Diamond^{\leq 2}$ 🍪$)$
- ✅ $\mathbb{P}\,(\Box^{\leq 1}\,(s_0 \vee s_2))$
- ✅ $\mathbb{P}\,(\neg s_3\ \mathrm{U}\ s_2)$

User chooses number of runs for desired statistical conf.

All runs terminate with probability 1

```
1  dtmc
2
3  const COOKIES = 2;
4  const DED = 3;
5
6  module ASYLUM
7    s:[0..3];
8    [] s=0 -> 1/2: (s'=COOKIES)
9              + 1/2: (s'=1);
10   [] s=1 -> 1/2: (s'=COOKIES)
11             + 1/4: (s'=0)
12             + 1/4: (s'=DED);
13 endmodule
```

Unbounded probability estimation ✅❌



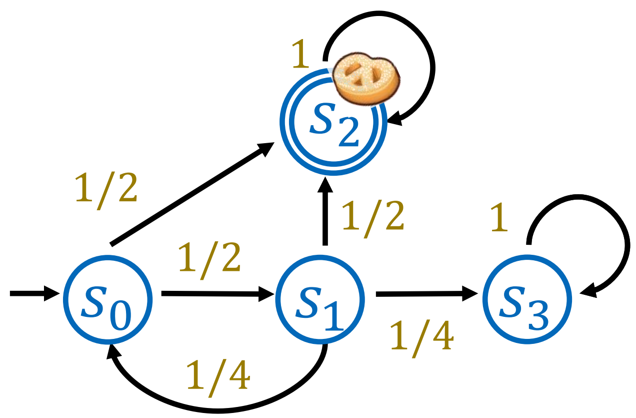
- ✅ $\mathbb{P}(\Diamond$ 🍪$)$
- ✅ $\mathbb{P}\ (\Box\ (s_0 \vee s_2))$
- ✅ $\mathbb{P}(\Diamond^{\geq 2}$ 🍪$)$

Besides number of runs, user must choose max path length

All runs are forced to terminate

```
1  dtmc
2
3  const COOKIES = 2;
4  const DED = 3;
5
6  module ASYLUM
7    s:[0..3];
8    [] s=0 -> 1/2: (s'=COOKIES)
9              + 1/2: (s'=1);
10   [] s=1 -> 1/2: (s'=COOKIES)
11             + 1/4: (s'=0)
12             + 1/4: (s'=DED);
13 endmodule
```

Steady state



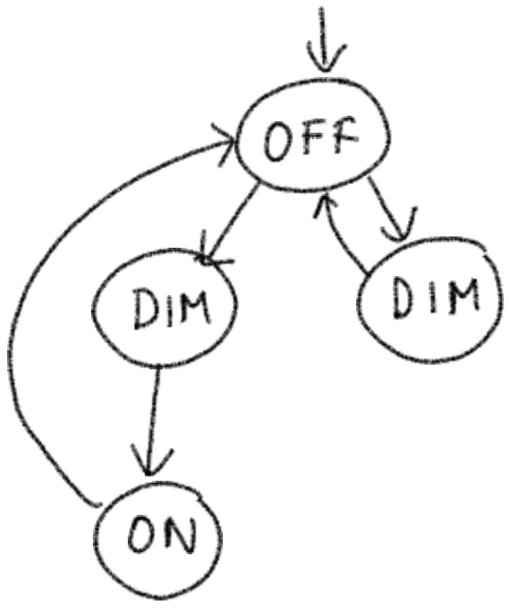
S(🍪)

```
dtmc

const COOKIES = 2;
const DED = 3;

module ASYLUM
  s:[0..3];
  [] s=0 -> 1/2: (s'=COOKIES)
          + 1/2: (s'=1);
  [] s=1 -> 1/2: (s'=COOKIES)
          + 1/4: (s'=0)
          + 1/4: (s'=DED);
endmodule
```

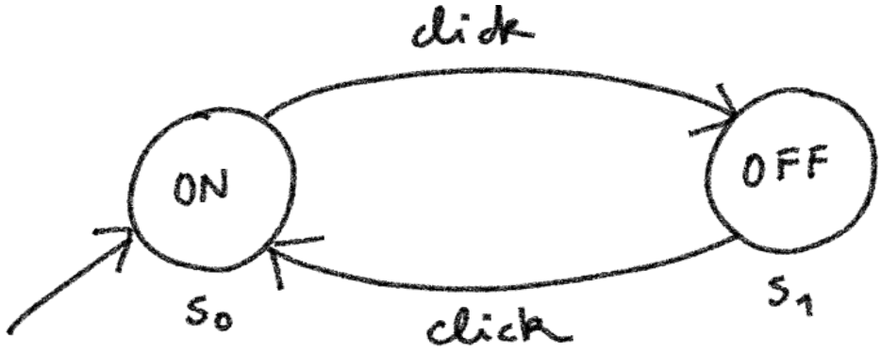
Two standard ways:

- Find a **regeneration point**, and perform transient analysis in the strongly connected component from the initial state to that point;
- Run for a very very very very very very very very very very very very very very very very very very very very very very very very very very very very very very very long run, discard that transient phase, keep the state, and perform "transient batches" always continuing from the last state reached every time you generated a new batch.

Nondeterminism



$$\mathbb{P}(\Diamond^{\leq 2}\ \mathrm{ON})$$



$S = \{s_0, s_1\}$

$\rightarrow = \{(s_0, \text{click}, s_1), (s_1, \text{click}, s_0)\}$

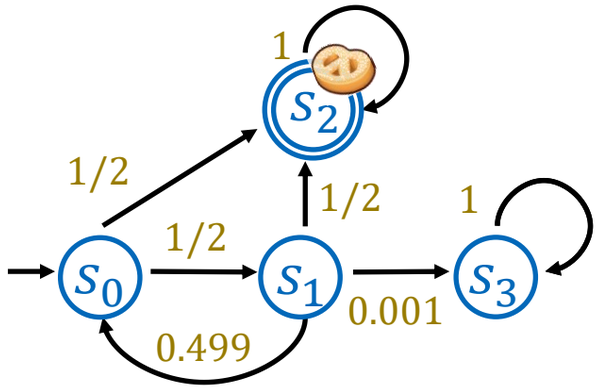
$A = \{\text{click}\}$

$L = \{s_0 \mapsto \{\text{ON}\},\ s_1 \mapsto \{\text{OFF}\}\}$

$AP = \{\text{ON}, \text{OF}\}$

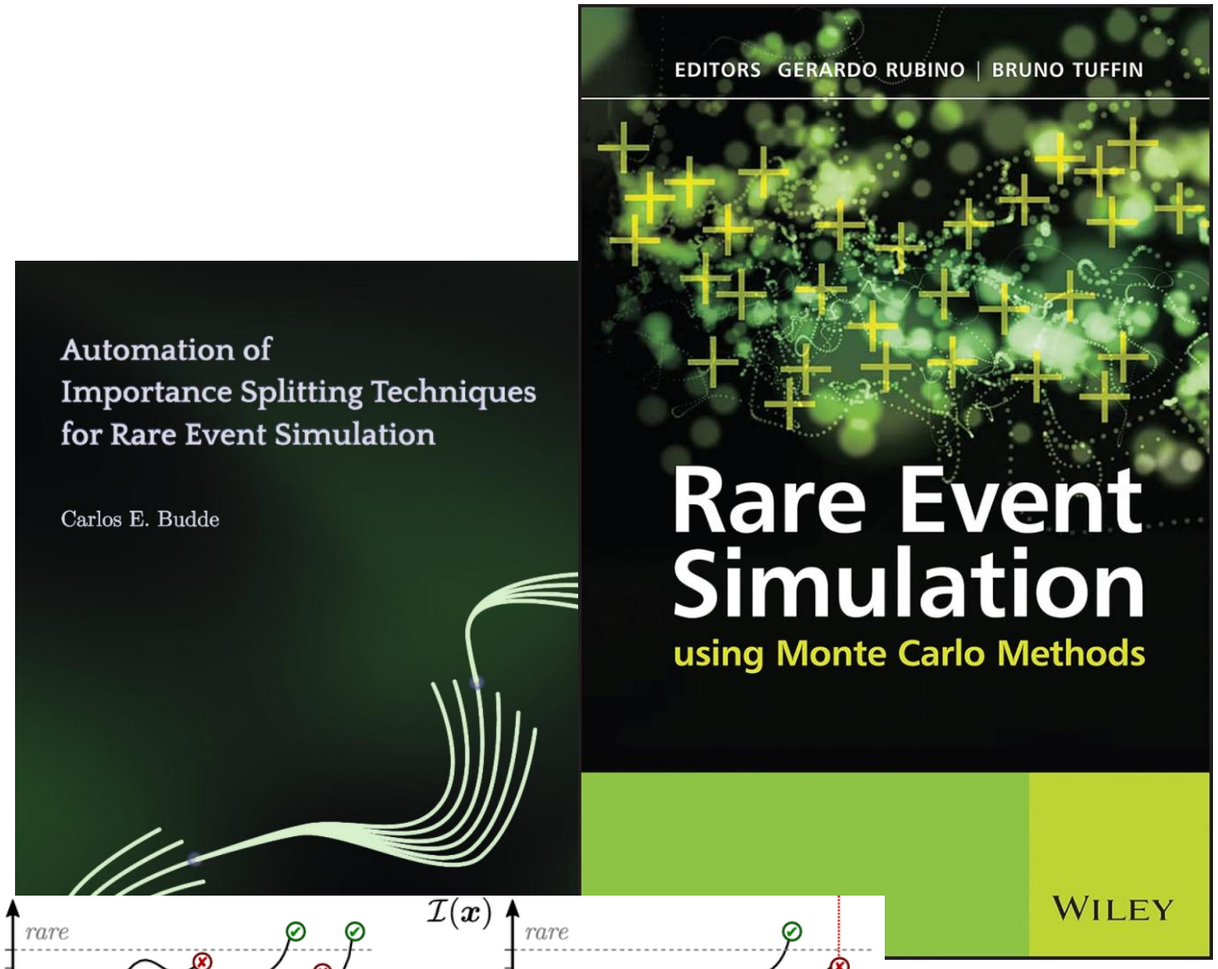
$s_0 = \{s_0\}$

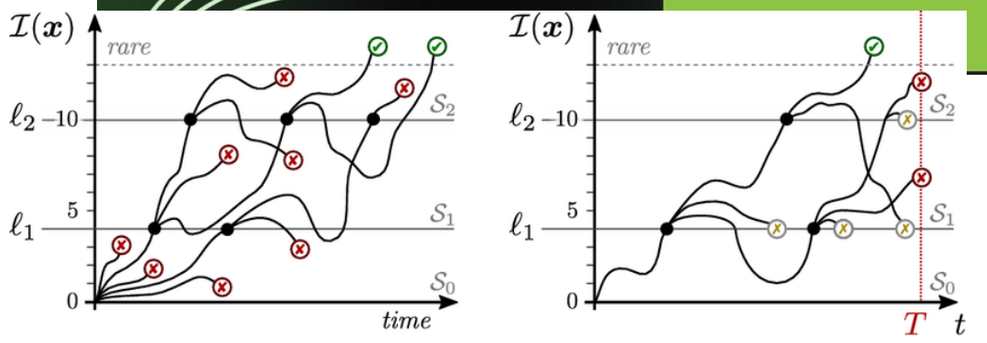
Rare events



$$\mathbb{P}(\Diamond^{\leq 2} s_3)$$

```
1  dtmc
2
3  const COOKIES = 2;
4  const DED = 3;
5
6  module ASYLUM
7    s:[0..3];
8    [] s=0 -> 1/2: (s'=COOKIES)
9              + 1/2: (s'=1);
10   [] s=1 -> 1/2: (s'=COOKIES)
11             +.499: (s'=0)
12             +.001: (s'=DED);
13 endmodule
```





(a) FE$_5$ for *Prob*(¬✘U✔) **(b)** RST$_{\text{ES}}$ for UNREL$_T$

Statistical Model Checking

- Random variables for trace generation
- Statistics for model checking
- Confidence intervals
- How to use, and when not to use
- Exercises & homework

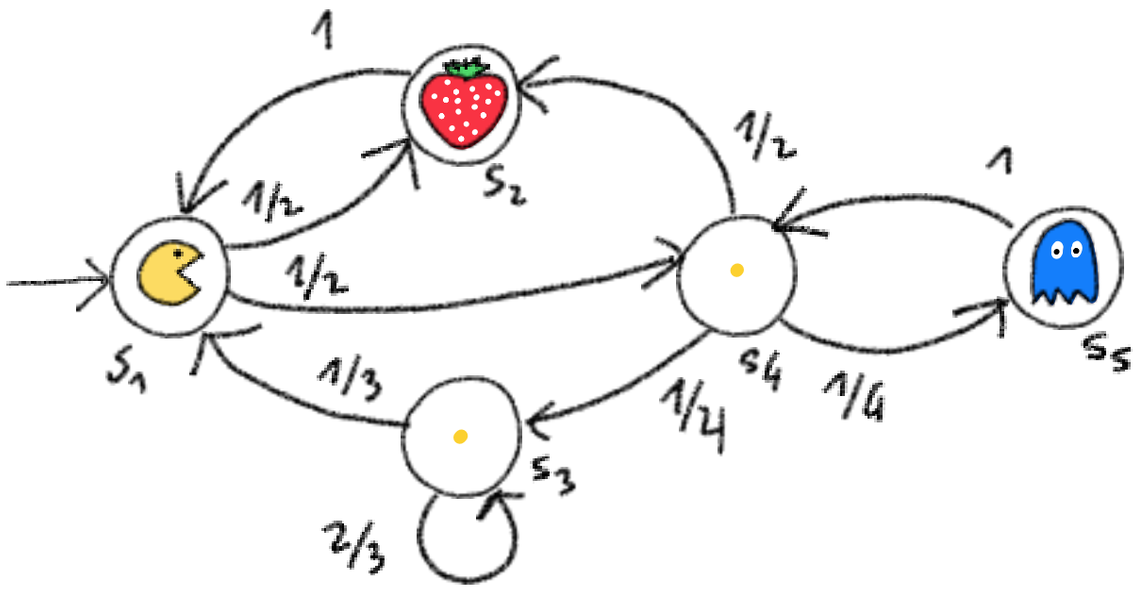
Statistical Model Checking

- Random variables for trace generation
- Statistics for model checking
- Confidence intervals
- How to use, and when not to use
- Exercises & homework

APPENDIX: Exercises

Exercise 9.1

Using PRISM, apply statistical model checking (SMC, i.e. the simulation option) to try to approximate the values of these properties. Good luck.



a) $\mathbb{P}_{=?}[\ \bigcirc$ 🍓 $]$

b) $\mathbb{P}_{=?}[\ \bigcirc$ (🍓 $\vee \bigcirc$ 🍓)$]$

c) $\mathbb{P}_{=?}[\Diamond$ 👻 $]$

d) $\mathbb{P}_{=?}[\neg$ 👻 U 🍓 $]$

e) $\mathbb{P}_{=?}[\ \Box \neg$ 👻 $]$

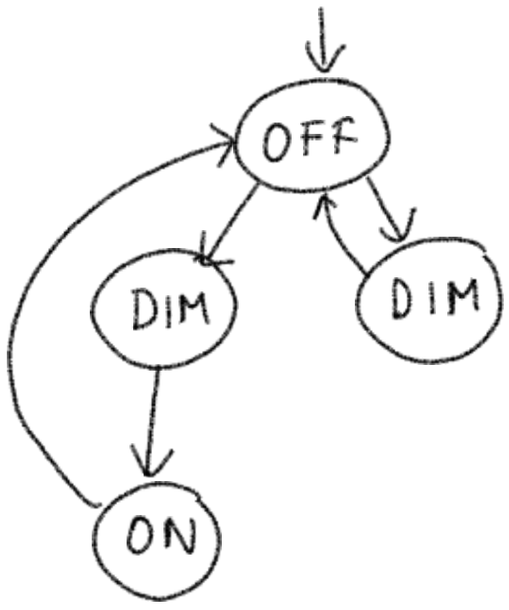
1. First use the default simulation parameters of PRISM.

2. Then experiment with the number of runs, until the approximated values coincide with the numerical (“true”) values up to the third decimal point.

3. Same as above, but also play with the “Simulation method”, until the relative error is lower than 0.001

Exercise 9.2

Define this transition system in PRISM, and use SMC to estimate the values of these properties.



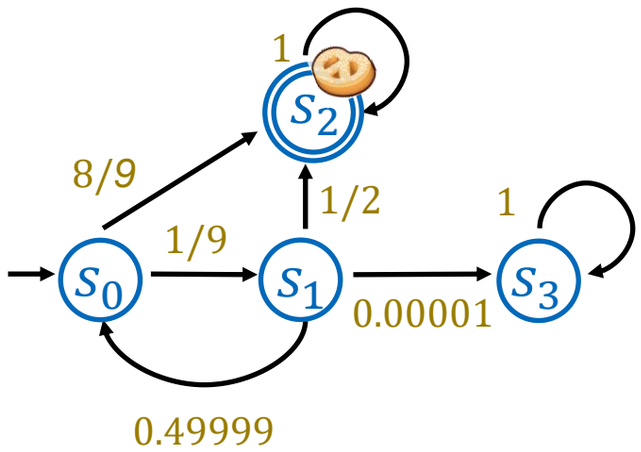
a) $\mathbb{P}(\Diamond^{\leq 2}\ \mathbf{ON})$

b) $\mathbb{P}(\Diamond\ \mathbf{ON})$

Explain to the TA / professor what
is going on here: how does PRISM simulate nondeterminism?

Exercise 9.3

Define this transition system in PRISM, and use SMC to estimate the values of these properties.



$$\mathbb{P}(\Diamond^{\leq 2}\ s_3)$$

Obtain a relative error < 0.001 by any means possible. Explain your solution.