CS 512, Spring 2018, Handout 22

Counterexamples and Probabilistic Model Checking

Assaf Kfoury

April 17, 2018

Counterexamples – in general

(material in this and later slides mostly due to Prof. J-P Katoen of Aachen Univ)

- Reminder: **model checking** = **bug hunting**,
  bugs are discovered by **counterexamples**,
  states that refute a given property (desirable or harmful).
- **Counterexamples** are (formally expressed) instances of system behavior that contradict a system's (formally expressed) specification.

Counterexamples – in general

- **Counterexamples in LTL** are typically finite execution paths:
  - To contradict $(\Box\,\varphi)$,
    we want a finite path ending in a $(\neg\varphi)$-state.
  - To contradict $(\Diamond\,\varphi)$,
    we want a finite $(\neg\varphi)$-path leading to a $(\neg\varphi)$-cycle.

  Methods of LTL model-checkers incorporate forms of **breadth-first search** for generating shortest counterexamples (*e.g.*, see Handout 13).

Counterexamples – in general

- **Counterexamples in LTL** are typically finite execution paths:
  - To contradict $(\Box\,\varphi)$,
    we want a finite path ending in a $(\neg\varphi)$-state.
  - To contradict $(\Diamond\,\varphi)$,
    we want a finite $(\neg\varphi)$-path leading to a $(\neg\varphi)$-cycle.

  Methods of LTL model-checkers incorporate forms of **breadth-first search** for generating shortest counterexamples (*e.g.*, see Handout 13).

- **Counterexamples in CTL** are typically finite trees of execution paths:
  - To contradict universal CTL,
    we want **all** paths in a tree of execution paths.
  - To contradict existential CTL,
    we want **one** path in a tree of execution paths.

  Methods of CTL model-checkers also incorporate some form of **breadth-first search**, combined with more advanced data structures.

Counterexamples – in PCTL (Probabilistic CTL)

- **Problem statement**:

  Given a WFF of PCTL of the form $\mathbb{P}_{\leqslant p}(\varphi)$
  – for example, in shorthand, $(p \,\mathbb{U}^{\leqslant 1/2}\, q)$ or $(\bigcirc^{\leqslant 2/3}\, p)$ –
  together with a Markov chain $\mathcal{M}$ and a state s in $\mathcal{M}$,
  we want to decide whether:

  $$\mathcal{M}, s \not\models \mathbb{P}_{\leqslant p}(\varphi) \quad \text{or, more succintly,} \quad s \not\models \mathbb{P}_{\leqslant p}(\varphi)$$

Counterexamples – in PCTL (Probabilistic CTL)

- **Problem statement**:

  Given a WFF of PCTL of the form $\mathbb{P}_{\leqslant p}(\varphi)$

  – for example, in shorthand, $(p \, \mathbb{U}^{\leqslant 1/2} \, q)$ or $(\bigcirc^{\leqslant 2/3} \, p)$ –

  together with a Markov chain $\mathcal{M}$ and a state s in $\mathcal{M}$,
  we want to decide whether:

  $$\mathcal{M}, s \not\models \mathbb{P}_{\leqslant p}(\varphi) \quad \text{or, more succintly,} \quad s \not\models \mathbb{P}_{\leqslant p}(\varphi)$$

- A **counterexample** C **for** $\mathbb{P}_{\leqslant p}(\varphi)$ **at state** s **in** $\mathcal{M}$
  is a set of finite paths (or **evidences**) in $\mathcal{M}$ satisfying:

  - if $\pi \in C$, then π starts at s and $\pi \models \varphi$, and
  - $\mathrm{Pr}(C) > p$ where $\mathrm{Pr}(C) \triangleq \sum_{\pi \in C} \mathrm{Pr}(\pi)$,
    i.e., the sum of the probabilities of the paths in C, exceeds p.

  If $\mathrm{Pr}(C) > p$, we conclude that $s \not\models \mathbb{P}_{\leqslant p}(\varphi)$.

Counterexamples – in PCTL (Probabilistic CTL)

- **Problem statement**:

  Given a WFF of PCTL of the form $\mathbb{P}_{\leqslant p}(\varphi)$

  – for example, in shorthand, $(p\ \mathbb{U}^{\leqslant 1/2}\ q)$ or $(\bigcirc^{\leqslant 2/3}\ p)$ –
  together with a Markov chain $\mathcal{M}$ and a state s in $\mathcal{M}$,
  we want to decide whether:

  $$\mathcal{M}, s \not\models \mathbb{P}_{\leqslant p}(\varphi) \quad \text{or, more succintly,} \quad s \not\models \mathbb{P}_{\leqslant p}(\varphi)$$

- A **counterexample** C **for** $\mathbb{P}_{\leqslant p}(\varphi)$ **at state** s **in** $\mathcal{M}$
  is a set of finite paths (or **evidences**) in $\mathcal{M}$ satisfying:

  - if $\pi \in C$, then π starts at s and $\pi \models \varphi$, and
  - $\Pr(C) > p$ where $\Pr(C) \triangleq \sum_{\pi \in C} \Pr(\pi)$,
    i.e., the sum of the probabilities of the paths in C, exceeds p.

  If $\Pr(C) > p$, we conclude that $s \not\models \mathbb{P}_{\leqslant p}(\varphi)$.

- In this handout, we limit attention to discrete-time Markov chains –
  we delay work done on continuous-time Markov chains till next year (!).

Counterexamples – in PCTL (Probabilistic CTL)

- A counterexample C for $\mathbb{P}_{\leqslant p}(\varphi)$ is **minimal** if $\left|C\right| \leqslant \left|C'\right|$ for any counterexample C' for $\mathbb{P}_{\leqslant p}(\varphi)$.

Counterexamples – in PCTL (Probabilistic CTL)

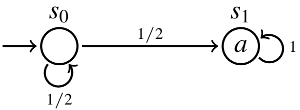
- A counterexample C for $\mathbb{P}_{\leqslant p}(\varphi)$ is **minimal** if $\left|C\right| \leqslant \left|C'\right|$ for any counterexample C' for $\mathbb{P}_{\leqslant p}(\varphi)$.
- A counterexample C for $\mathbb{P}_{\leqslant p}(\varphi)$ is **smallest** if C is minimal and $\mathrm{Pr}(C) \geqslant \mathrm{Pr}(C')$ for any minimal counterexample C' for $\mathbb{P}_{\leqslant p}(\varphi)$.

Counterexamples – in PCTL (Probabilistic CTL)

- A counterexample C for $\mathbb{P}_{\leqslant p}(\varphi)$ is **minimal** if $\left|C\right| \leqslant \left|C'\right|$ for any counterexample C' for $\mathbb{P}_{\leqslant p}(\varphi)$.
- A counterexample C for $\mathbb{P}_{\leqslant p}(\varphi)$ is **smallest** if C is minimal and $\Pr(C) \geqslant \Pr(C')$ for any minimal counterexample C' for $\mathbb{P}_{\leqslant p}(\varphi)$.
- **Fact**: Counterexamples for non-strict probability bounds (*i.e.*, bounds of the form "$\leqslant p$", not "$< p$") are **finite**.

Counterexamples – in PCTL (Probabilistic CTL)

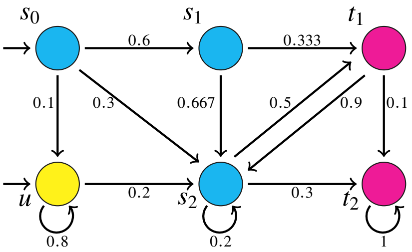
- A counterexample C for $\mathbb{P}_{\leqslant p}(\varphi)$ is **minimal** if $\left|C\right| \leqslant \left|C'\right|$ for any counterexample C' for $\mathbb{P}_{\leqslant p}(\varphi)$.
- A counterexample C for $\mathbb{P}_{\leqslant p}(\varphi)$ is **smallest** if C is minimal and $\Pr(C) \geqslant \Pr(C')$ for any minimal counterexample C' for $\mathbb{P}_{\leqslant p}(\varphi)$.
- **Fact**: Counterexamples for non-strict probability bounds (*i.e.*, bounds of the form "$\leqslant p$", not "$< p$") are **finite**.
- **Infinite** counterexamples may be needed for WFF's with strict probability bounds.
- For example, an **infinite** counterexample is needed for $s_0 \not\models \mathbb{P}_{<1}(\Diamond\, a)$, *i.e.*, for $s_0 \not\models (\Diamond^{<1}\, a)$ in the following Markov chain:



Example showing how to handle "until" WFF's in PCTL[1]

[1] Partly inspired by Example 10.41 in [PMC, page 786].

Example showing how to handle "until" WFF's in PCTL[1]



Wanted:

counterexamples for $s_0 \not\models (\varphi \, \mathbb{U}^{\leqslant 1/2} \, \psi)$

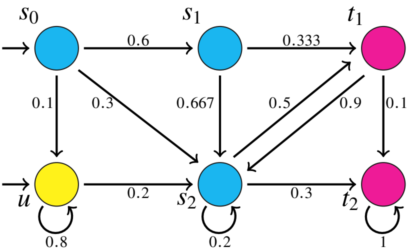
blue states : only prop WFF φ holds,

red states : only prop WFF ψ holds,

yellow states : neither φ nor ψ hold.

[1] Partly inspired by Example 10.41 in [PMC, page 786].

Example showing how to handle "until" WFF's in PCTL[1]



blue states: only prop WFF φ holds,

red states: only prop WFF ψ holds,

yellow states: neither φ nor ψ hold.

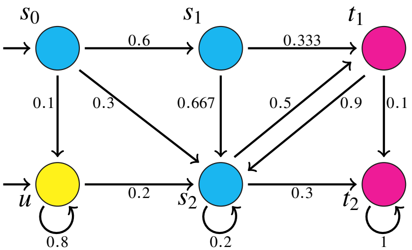
Wanted:

counterexamples for $s_0 \not\models (\varphi \; \mathbb{U}^{\leqslant 1/2} \; \psi)$

evidence	probability
$\pi_1 \triangleq s_0\, s_1\, t_1$	0.2
$\pi_2 \triangleq s_0\, s_1\, s_2\, t_1$	0.2
$\pi_3 \triangleq s_0\, s_2\, t_1$	0.15
$\pi_4 \triangleq s_0\, s_1\, s_2\, t_2$	0.12
$\pi_5 \triangleq s_0\, s_2\, t_2$	0.09
$\ldots$	$\ldots$

[1] Partly inspired by Example 10.41 in [PMC, page 786].

Example showing how to handle "until" WFF's in PCTL[1]



Wanted:

counterexamples for $s_0 \not\models (\varphi\ \mathbb{U}^{\leqslant 1/2}\ \psi)$

evidence	probability
$\pi_1 \triangleq s_0\, s_1\, t_1$	0.2
$\pi_2 \triangleq s_0\, s_1\, s_2\, t_1$	0.2
$\pi_3 \triangleq s_0\, s_2\, t_1$	0.15
$\pi_4 \triangleq s_0\, s_1\, s_2\, t_2$	0.12
$\pi_5 \triangleq s_0\, s_2\, t_2$	0.09
$\dots$	$\dots$

blue states: only prop WFF φ holds,

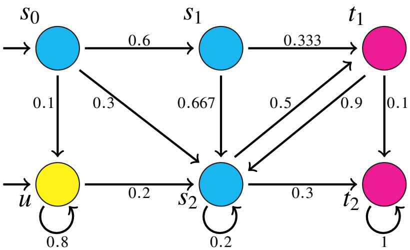
red states: only prop WFF ψ holds,

yellow states: neither φ nor ψ hold.

	counterexample	cardinality	probability
	$\{\pi_1, \pi_2, \pi_3, \pi_4, \pi_5\}$	5	0.76
	$\{\pi_1, \pi_3, \pi_4, \pi_5\}$	4	0.56
	$\{\pi_2, \pi_3, \pi_4, \pi_5\}$	4	0.76
minimal $\longrightarrow$	$\{\pi_1, \pi_2, \pi_4\}$	3	0.52
minimal $\longrightarrow$	$\{\pi_1, \pi_2, \pi_3\}$	3	0.55

[1] Partly inspired by Example 10.41 in [PMC, page 786].

Example showing how to handle "until" WFF's in PCTL[1]



blue states: only prop WFF φ holds,

red states: only prop WFF ψ holds,

yellow states: neither φ nor ψ hold.

Wanted:

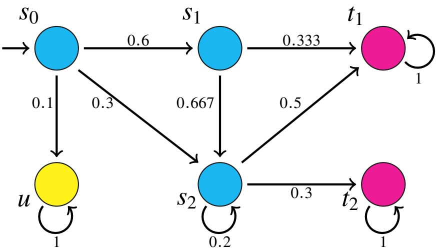
counterexamples for $s_0 \not\models (\varphi \; \mathbb{U}^{\leqslant 1/2} \; \psi)$

evidence	probability
$\pi_1 \triangleq s_0\, s_1\, t_1$	0.2
$\pi_2 \triangleq s_0\, s_1\, s_2\, t_1$	0.2
$\pi_3 \triangleq s_0\, s_2\, t_1$	0.15
$\pi_4 \triangleq s_0\, s_1\, s_2\, t_2$	0.12
$\pi_5 \triangleq s_0\, s_2\, t_2$	0.09
...	...

	counterexample	cardinality	probability
	$\{\pi_1, \pi_2, \pi_3, \pi_4, \pi_5\}$	5	0.76
	$\{\pi_1, \pi_3, \pi_4, \pi_5\}$	4	0.56
	$\{\pi_2, \pi_3, \pi_4, \pi_5\}$	4	0.76
	$\{\pi_1, \pi_2, \pi_4\}$	3	0.52
smallest $\longrightarrow$	$\{\pi_1, \pi_2, \pi_3\}$	3	0.55

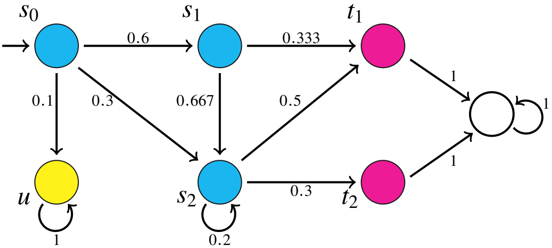
[1] Partly inspired by Example 10.41 in [PMC, page 786].

Obtaining smallest counterexamples



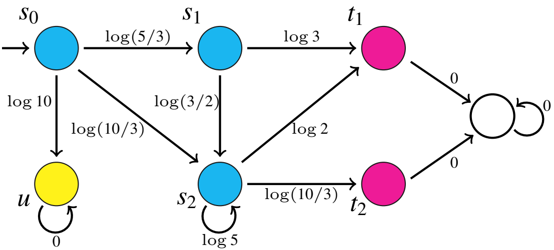
Step 1: Make all ψ-states and all $(\neg\varphi \wedge \neg\psi)$-states absorbing, which requires eliminating some transitions (*e.g.*, the transitions out of t_1 and u) and making the transition probability $= 1$ on all self-loops.

Adapting a bit more



Step 2: Insert a sink state and redirect all outgoing edges of ψ-states to it.

A weighted diagraph



Step 3: Turn the Markov chain into a weighted digraph (directed graph), where:

$$w(s, s') \triangleq \log\left(\frac{1}{\Pr(s, s')}\right)$$

for every pair of nodes/states s and s'. The logarithm can be base 10, or base e, or base 2 – it does not matter which base we choose.

A simple derivation

Given a finite path $\pi \triangleq s_0\, s_1\, s_2\ \cdots\ s_n$:

$$
\begin{aligned}
w(\pi) &= w(s_0, s_1) + w(s_1, s_2) + \cdots + w(s_{n-1}, s_n) \\
&= \log\left(\frac{1}{\Pr(s_0, s_1)}\right) + \log\left(\frac{1}{\Pr(s_1, s_2)}\right) + \cdots + \log\left(\frac{1}{\Pr(s_{n-1}, s_n)}\right) \\
&= \log\left(\frac{1}{\Pr(s_0, s_1) \cdot \Pr(s_1, s_2) \cdot\ \cdots\ \cdot \Pr(s_{n-1}, s_n)}\right) \\
&= \log\left(\frac{1}{\Pr(\pi)}\right)
\end{aligned}
$$

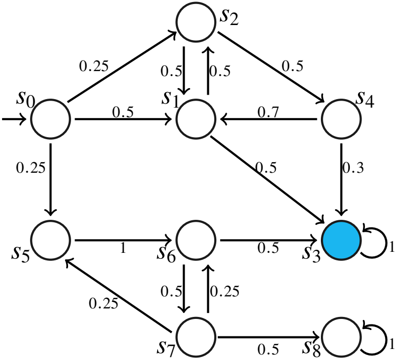
Conclusion 1: For all finite paths π and π' in the Markov chain, we have:

$$\underbrace{\Pr(\pi) \geqslant \Pr(\pi')}_{\text{in the Markov chain}} \quad \text{if and only if} \quad \underbrace{w(\pi) \leqslant w(\pi')}_{\text{in the weighted digraph}}$$

Conclusion 2: Finding a **strongest evidence** in the Markov chain is translated to a **shortest path problem** in the weighted digraph.

Another example: How to handle reachability properties

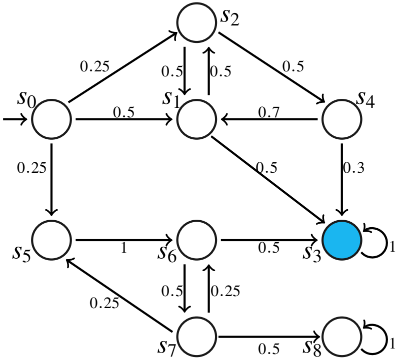
Wanted: counterexamples for $\mathbb{P}_{\leqslant 0.4}(\Diamond\,\varphi)$, or, in shorthand, $(\Diamond^{\leqslant 0.4}\,\varphi)$.



blue state : only one φ-state.

Another example: How to handle reachability properties

Wanted: counterexamples for $\mathbb{P}_{\leqslant 0.4}(\Diamond\,\varphi)$, or, in shorthand, $(\Diamond^{\leqslant 0.4}\,\varphi)$.



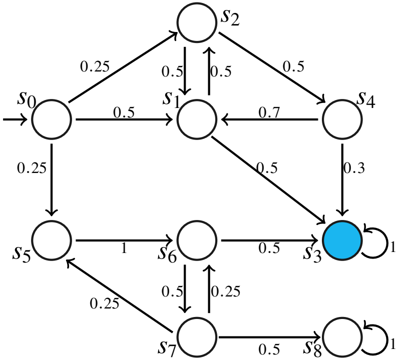
blue state : only one φ-state.

Approach 1, based on using the transition (right-stochastic) 9×9 matrix A:

$$
\begin{array}{c}
\\ s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \\ s_8
\end{array}
\begin{array}{c}
\begin{array}{ccccccccc} s_0 & s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \end{array} \\
\begin{bmatrix}
0 & .5 & .25 & 0 & 0 & .25 & 0 & 0 & 0 \\
0 & 0 & .5 & .5 & 0 & 0 & 0 & 0 & 0 \\
0 & .5 & 0 & 0 & .5 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & .7 & 0 & .3 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & .5 & 0 & 0 & 0 & .5 & 0 \\
0 & 0 & 0 & 0 & 0 & .25 & .25 & 0 & .5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\end{array}
$$

Another example: How to handle reachability properties

Wanted: counterexamples for $\mathbb{P}_{\leqslant 0.4}(\Diamond\,\varphi)$, or, in shorthand, $(\Diamond^{\leqslant 0.4}\,\varphi)$.



Approach 1, based on using the transition (right-stochastic) 9×9 matrix A:

$$
\begin{array}{c}
\\ s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \\ s_8
\end{array}
\begin{array}{c}
\begin{array}{ccccccccc} s_0 & s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \end{array} \\
\begin{bmatrix}
0 & .5 & .25 & 0 & 0 & .25 & 0 & 0 & 0 \\
0 & 0 & .5 & .5 & 0 & 0 & 0 & 0 & 0 \\
0 & .5 & 0 & 0 & .5 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & .7 & 0 & .3 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & .5 & 0 & 0 & 0 & .5 & 0 \\
0 & 0 & 0 & 0 & 0 & .25 & .25 & 0 & .5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\end{array}
$$

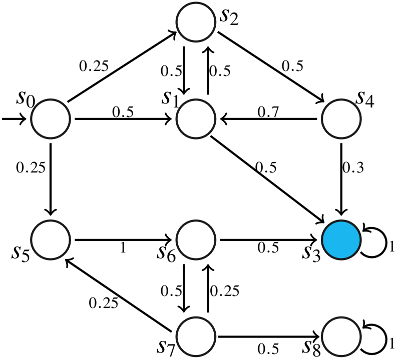
blue state : only one φ-state.

- initial distribution over 9 states is $\boldsymbol{d}_0 = (1,0,0,0,0,0,0,0,0) = [1\;0\;0\;0\;0\;0\;0\;0\;0]$.
- distribution after 1 transition, 2 transitions, and 3 transitions, respectively:

$$\boldsymbol{d}_1 = \boldsymbol{d}_0 \cdot A = (0,\ .5,\ .25,\ \boxed{0},\ 0,\ .25,\ 0,\ 0,\ 0)$$
$$\boldsymbol{d}_2 = \boldsymbol{d}_0 \cdot A^2 = (0,\ .125,\ .25,\ \boxed{.25},\ .125,\ 0,\ .25,\ 0,\ 0)$$
$$\boldsymbol{d}_3 = \boldsymbol{d}_0 \cdot A^3 = (0,\ .2125,\ .0625,\ \boxed{.475},\ .125,\ 0,\ 0,\ .125,\ 0)$$

Another example: How to handle reachability properties

Wanted: counterexamples for $\mathbb{P}_{\leqslant 0.4}(\Diamond\,\varphi)$, or, in shorthand, $(\Diamond^{\leqslant 0.4}\,\varphi)$.



Approach 1, based on using the transition (right-stochastic) 9×9 matrix A:

$$
\begin{array}{c}
 \\ s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \\ s_8
\end{array}
\begin{array}{c}
\begin{array}{ccccccccc} s_0 & s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \end{array} \\
\begin{bmatrix}
0 & .5 & .25 & 0 & 0 & .25 & 0 & 0 & 0 \\
0 & 0 & .5 & .5 & 0 & 0 & 0 & 0 & 0 \\
0 & .5 & 0 & 0 & .5 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & .7 & 0 & .3 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & .5 & 0 & 0 & 0 & .5 & 0 \\
0 & 0 & 0 & 0 & 0 & .25 & .25 & 0 & .5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\end{array}
$$

blue state : only one φ-state.

- **Conclusion:** Starting from s_0, state s_3 is reached with probability $.475 > .4$ after 3 transitions.
- Hence, there is a counterexample C for $s_0 \not\models (\Diamond^{\leqslant .4}\,\varphi)$ consisting of finite paths, each with at most 3 transitions – but we have not determined the members of the counterexample C yet, nor do we know if it is **minimal** or **smallest** (cf. page 8)

Another example: How to handle reachability properties

Wanted: counterexamples for $\mathbb{P}_{\leqslant 0.4}(\Diamond\,\varphi)$, or, in shorthand, $(\Diamond^{\leqslant 0.4}\,\varphi)$.

- **Approach 2:** Let S be the set of states in the Markov chain, $s_0 \in S$ a single initial state, and Target $\subseteq S$ a non-empty set of target states.

  For every state s, we define the probability p_s of reaching the states in Target from s:

$$p_s \triangleq \begin{cases} 1 & \text{if } s \in \text{Target}, \\ 0 & \text{if no state in Target is reachable from } s, \\ \sum_{s' \in S} \mathrm{Pr}(s, s') \cdot p_{s'} & \text{otherwise.} \end{cases}$$

- This defines a system of linear equations over the variables $V \triangleq \{\, p_s \mid s \in S \,\}$ whose unique solution $\sigma : V \to [0, 1]$ assigns to each p_s the probability of reaching Target from s.

- Hence, $\mathcal{M} \models \mathbb{P}_{\leqslant \rho}(\Diamond\,\text{target})$ iff $\sigma(p_{s_0}) \leqslant \rho$, where "target" is an atomic proposition which labels every state in Target.

- **Advantage of Approach 2 over Approach 1**: Solving a system of linear equations instead of repeatedly multiplying stochastic matrices.

Another example: How to handle reachability properties

Wanted: counterexamples for $\mathbb{P}_{\leqslant 0.4}(\Diamond\, \varphi)$, or, in shorthand, $(\Diamond^{\leqslant 0.4}\, \varphi)$.

- For the Markov chain $\mathcal{M}$ shown on slide 21, we obtain:

$$
\begin{aligned}
p_{s_0} &= 0.5\, p_{s_1} + 0.25\, p_{s_2} + 0.25\, p_{s_5} & p_{s_1} &= 0.5\, p_{s_2} + 0.5\, p_{s_3} \\
p_{s_2} &= 0.5\, p_{s_1} + 0.5\, p_{s_4} & p_{s_3} &= 1 \\
p_{s_4} &= 0.7\, p_{s_1} + 0.3\, p_{s_3} & p_{s_5} &= 1\, p_{s_6} \\
p_{s_6} &= 0.5\, p_{s_3} + 0.5\, p_{s_7} & p_{s_7} &= 0.25\, p_{s_5} + 0.25\, p_{s_6}
\end{aligned}
$$

  We can remove all states from $\mathcal{M}$ which do not reach states in Target. In this example, we remove s_8, thus also removing equation $p_{s_8} = 0$.

- Solving the system of linear equations (by hand or by using Matlab or Octave), we obtain a solution $\sigma : \{p_{s_0}, p_{s_1}, \ldots, p_{s_7}\} \to [0, 1]$ such that:

$$
\begin{aligned}
\sigma(p_{s_0}) &= 11/12 & \sigma(p_{s_1}) &= \sigma(p_{s_2}) = \sigma(p_{s_3}) = \sigma(p_{s_4}) = 1 \\
\sigma(p_{s_5}) &= \sigma(p_{s_6}) = 2/3 & \sigma(p_{s_7}) &= 1/3
\end{aligned}
$$

- **Conclusion:** Starting from s_0, state s_3 is reached with probability $\dfrac{11}{12} > .4$

  Hence, there is a counterexample C for $s_0 \not\models (\Diamond^{\leqslant .4}\, \varphi)$, though we do not know the members of C yet!!

Another example: How to handle reachability properties

Wanted: counterexamples for $\mathbb{P}_{\leqslant 0.4}(\Diamond\,\varphi)$, or, in shorthand, $(\Diamond^{\leqslant 0.4}\,\varphi)$.

- **Approach 3**, most efficient and most direct, repeats the steps carried out to find counterexamples for $s_0 \not\models (\varphi\ \mathbb{U}^{\leqslant 1/2}\ \psi)$, from slide 12 to slide 20.
- We obtain, in order of decreasing probabilities:

evidence	weight (rounded)	probability
$\pi_1 \triangleq s_0\,s_1\,s_3$	1.39	0.25
$\pi_2 \triangleq s_0\,s_5\,s_6\,s_3$	2.08	0.125
$\pi_3 \triangleq s_0\,s_2\,s_1\,s_3$	2.77	0.0625
$\pi_4 \triangleq s_0\,s_1\,s_2\,s_1\,s_3$	2.77	0.0625
$\pi_5 \triangleq s_0\,s_2\,s_4\,s_1\,s_3$	3.13	0.04375
$\pi_6 \triangleq s_0\,s_1\,s_2\,s_4\,s_1\,s_3$	3.13	0.04375
$\pi_7 \triangleq s_0\,s_2\,s_4\,s_3$	3.28	0.03750
$\pi_8 \triangleq s_0\,s_1\,s_2\,s_4\,s_3$	3.28	0.03750
$\ldots$	$\ldots$	$\ldots$

  where we take weight $w(s,s') \triangleq -\ln(\Pr(s,s'))$ for all states $s,s' \in S$.

- $\sum_{i\in\{1,2,3\}} \Pr(\pi_i) = \sum_{i\in\{1,2,4\}} \Pr(\pi_i) = 0.4375 > 0.4$

  (but why not $\{\pi_1, \pi_2, s_0 s_2 s_4\}$ or $\{\pi_1, \pi_2, s_0 s_1 s_2 s_4\}$???)

  implies both $\{\pi_1, \pi_2, \pi_3\}$ and $\{\pi_1, \pi_2, \pi_4\}$ are smallest counterexamples.

(THIS PAGE INTENTIONALLY LEFT BLANK)