Counterexamples and Probabilistic Model Checking

(this handout should be read after Handout 25)

Assaf Kfoury

April 14, 2016 (adjusted April 19, 2016)
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 \( (G \varphi) \),
    we want a finite path ending in a \( (\neg \varphi) \)-state.
  - To contradict \( (F \varphi) \),
    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 23).
Counterexamples – in general

- **Counterexamples in LTL** are typically finite execution paths:
  - To contradict $(\mathsf{G} \varphi)$, we want a finite path ending in a $(\neg \varphi)$-state.
  - To contradict $(\mathsf{F} \varphi)$, 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 23).

- **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.
Problem statement:

Given a WFF of PCTL of the form $P \leq_p \varphi$

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

Together with a Markov chain $M$ and a state $s$ in $M$, we want to decide whether:

$M, s \not\models P \leq_p \varphi$ or, more succinctly, $s \not\models P \leq_p \varphi$
Problem statement:

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

- for example, in shorthand, $(p \mathbf{U}^{\leq 1/2} q)$ or $(X^{\leq 2/3} p)$

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

  $$\mathcal{M}, s \not\models \mathcal{P}_{\leq p}(\varphi)$$

or, more succinctly,

  $$s \not\models \mathcal{P}_{\leq p}(\varphi)$$

A counterexample $C$ for $\mathcal{P}_{\leq p}(\varphi)$ at state $s$ in $\mathcal{M}$

is a set of finite paths (or evidences) in $\mathcal{M}$ satisfying:

- if $\pi \in C$, then $\pi$ 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 $s \not\models \mathcal{P}_{\leq p}(\varphi)$. 

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)

Problem statement:

Given a WFF of PCTL of the form $P \leq p (\varphi)$
-- for example, in shorthand, $(p \ U \leq 1/2 \ q)$ or $(X \leq 2/3 \ p)$ --
together with a Markov chain $M$ and a state $s$ in $M$,
we want to decide whether:

$M, s \not{\models} P \leq p (\varphi)$
or, more succintly, $s \not{\models} P \leq p (\varphi)$

A counterexample $C$ for $P \leq p (\varphi)$ at state $s$ in $M$
is a set of finite paths (or evidences) in $M$ satisfying:

- if $\pi \in C$, then $\pi$ 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 $s \not{\models} P \leq p (\varphi)$.

In this handout, we limit attention to discrete-time Markov chains --
we delay work done on continuous-time Markov chains till next year (!).
A counterexample $C$ for $P_{\leq p}(\varphi)$ is **minimal** if $|C| \leq |C'|$ for any counterexample $C'$ for $P_{\leq p}(\varphi)$. 

Fact: Counterexamples for non-strict probability bounds (i.e., bounds of the form $\leq$, not $<$) are **finite**. Infinite counterexamples may be needed for WFF's with strict probability bounds. For example, an infinite counterexample is needed for $s_0 \not\models P_{< 1}(F_a)$, i.e., for $s_0 \not\models (F_{< 1}a)$ in the following Markov chain:
Counterexamples – in PCTL (Probabilistic CTL)

▶ A counterexample $C$ for $\mathcal{P}_{\leq p}(\varphi)$ is **minimal** if $|C| \leq |C'|$ for any counterexample $C'$ for $\mathcal{P}_{\leq p}(\varphi)$.

▶ A counterexample $C$ for $\mathcal{P}_{\leq p}(\varphi)$ is **smallest** if $C$ is minimal and $\Pr(C) \geq \Pr(C')$ for any minimal counterexample $C'$ for $\mathcal{P}_{\leq p}(\varphi)$.
Counterexamples – in PCTL (Probabilistic CTL)

- A counterexample $C$ for $P_{\leq p}(\varphi)$ is **minimal** if $|C| \leq |C'|$ for any counterexample $C'$ for $P_{\leq p}(\varphi)$.

- A counterexample $C$ for $P_{\leq p}(\varphi)$ is **smallest** if $C$ is minimal and $Pr(C) \geq Pr(C')$ for any minimal counterexample $C'$ for $P_{\leq p}(\varphi)$.

- **Fact**: Counterexamples for non-strict probability bounds (i.e., bounds of the form “$\leq p$”, not “$< p$”) are **finite**.
A counterexample $C$ for $\mathcal{P}_{\leq p}(\varphi)$ is **minimal** if $|C| \leq |C'|$ for any counterexample $C'$ for $\mathcal{P}_{\leq p}(\varphi)$.

A counterexample $C$ for $\mathcal{P}_{\leq p}(\varphi)$ is **smallest** if $C$ is minimal and $\Pr(C) \geq \Pr(C')$ for any minimal counterexample $C'$ for $\mathcal{P}_{\leq p}(\varphi)$.

**Fact**: Counterexamples for non-strict probability bounds (i.e., bounds of the form “$\leq p$”, not “$< p$”) are **finite**.

**Infinite** counterexamples may be needed for WFF’s with strict probability bounds.

For example, an **infinite** counterexample is needed for $s_0 \not\models \mathcal{P}_{<1}(Fa)$, i.e., for $s_0 \not\models (F^{<1} a)$ in the following Markov chain:

```
      s0 1/2 □
        □
        □
    1/2 □
      s1 1 □
          □
```
Example showing how to handle “until” WFF’s in PCTL
Example showing how to handle “until” WFF’s in PCTL

Wanted:
counterexamples for $s_0 \not\models (\varphi \mathbf{U}^{\leq 1/2} \psi)$

blue states: only prop WFF $\varphi$ holds,
red states: only prop WFF $\psi$ holds,
yellow states: neither $\varphi$ nor $\psi$ hold.
Example showing how to handle “until” WFF’s in PCTL

Wanted:
counterexamples for $s_0 \not\models (\varphi \mathbf{U}^{1/2} \psi)$

<table>
<thead>
<tr>
<th>evidence</th>
<th>probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_1 \triangleq s_0 s_1 t_1$</td>
<td>0.2</td>
</tr>
<tr>
<td>$\pi_2 \triangleq s_0 s_1 s_2 t_1$</td>
<td>0.2</td>
</tr>
<tr>
<td>$\pi_3 \triangleq s_0 s_2 t_1$</td>
<td>0.15</td>
</tr>
<tr>
<td>$\pi_4 \triangleq s_0 s_1 s_2 t_2$</td>
<td>0.12</td>
</tr>
<tr>
<td>$\pi_5 \triangleq s_0 s_2 t_2$</td>
<td>0.09</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

blue states: only prop WFF $\varphi$ holds,
red states: only prop WFF $\psi$ holds,
yellow states: neither $\varphi$ nor $\psi$ hold.
Example showing how to handle “until” WFF’s in PCTL

Wanted:

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

<table>
<thead>
<tr>
<th>evidence</th>
<th>probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_1 \triangleq s_0 \ s_1 \ t_1$</td>
<td>0.2</td>
</tr>
<tr>
<td>$\pi_2 \triangleq s_0 \ s_1 \ s_2 \ t_1$</td>
<td>0.2</td>
</tr>
<tr>
<td>$\pi_3 \triangleq s_0 \ s_2 \ t_1$</td>
<td>0.15</td>
</tr>
<tr>
<td>$\pi_4 \triangleq s_0 \ s_1 \ s_2 \ t_2$</td>
<td>0.12</td>
</tr>
<tr>
<td>$\pi_5 \triangleq s_0 \ s_2 \ t_2$</td>
<td>0.09</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>counterexample</th>
<th>cardinality</th>
<th>probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>${\pi_1, \pi_2, \pi_3, \pi_4, \pi_5}$</td>
<td>5</td>
<td>0.76</td>
</tr>
<tr>
<td>${\pi_1, \pi_3, \pi_4, \pi_5}$</td>
<td>4</td>
<td>0.56</td>
</tr>
<tr>
<td>${\pi_2, \pi_3, \pi_4, \pi_5}$</td>
<td>4</td>
<td>0.76</td>
</tr>
<tr>
<td>${\pi_1, \pi_2, \pi_4}$</td>
<td>3</td>
<td>0.52</td>
</tr>
<tr>
<td>${\pi_1, \pi_2, \pi_3}$</td>
<td>3</td>
<td>0.55</td>
</tr>
</tbody>
</table>

blue states: only prop WFF $\varphi$ holds,
red states: only prop WFF $\psi$ holds,
yellow states: neither $\varphi$ nor $\psi$ hold.
Example showing how to handle “until” WFF’s in PCTL

\[ T_1 \]

\[ T_2 \]

\[ s_0 \]

\[ s_1 \]

\[ s_2 \]

\[ u \]

\[ \pi_1 \triangleq s_0 s_1 t_1 \]
\[ \pi_2 \triangleq s_0 s_1 s_2 t_1 \]
\[ \pi_3 \triangleq s_0 s_2 t_1 \]
\[ \pi_4 \triangleq s_0 s_1 s_2 t_2 \]
\[ \pi_5 \triangleq s_0 s_2 t_2 \]

\[ \{ \pi_1, \pi_2, \pi_3, \pi_4, \pi_5 \} \]
\[ \{ \pi_1, \pi_3, \pi_4, \pi_5 \} \]
\[ \{ \pi_2, \pi_3, \pi_4, \pi_5 \} \]
\[ \{ \pi_1, \pi_2, \pi_4 \} \]
\[ \{ \pi_1, \pi_2, \pi_3 \} \]

\[ \text{smallest} \]

\[ \text{blue states} : \text{only prop WFF } \varphi \text{ holds}, \]
\[ \text{red states} : \text{only prop WFF } \psi \text{ holds}, \]
\[ \text{yellow states} : \text{neither } \varphi \text{ nor } \psi \text{ hold.} \]
Step 1: Make all $\psi$-states and all $(\neg \varphi \land \neg \psi)$-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 $\psi$-states to it.
Step 3: Turn the Markov chain into a weighted digraph (directed graph), where:

\[ w(s, s') \equiv \log \left( \frac{1}{\Pr(s, s')} \right) \]

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 $\pi \triangleq s_0 s_1 s_2 \cdots s_n$:

\[
    w(\pi) = w(s_0, s_1) + w(s_1, s_2) + \cdots + w(s_{n-1}, s_n)
    = \log \left( \frac{1}{\Pr(s_0, s_1)} \right) + \log \left( \frac{1}{\Pr(s_1, s_2)} \right) + \cdots + \log \left( \frac{1}{\Pr(s_{n-1}, s_n)} \right)
    = \log \left( \frac{1}{\Pr(s_0, s_1) \cdot \Pr(s_1, s_2) \cdot \cdots \cdot \Pr(s_{n-1}, s_n)} \right)
    = \log \left( \frac{1}{\Pr(\pi)} \right)
\]

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

\[
    \Pr(\pi) \geq \Pr(\pi') \quad \text{if and only if} \quad w(\pi) \leq w(\pi')
\]

in the Markov chain and 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 $\mathcal{P}_{\leq 0.4}(F \varphi)$, or, in shorthand, $(F^{\leq 0.4} \varphi)$.

![Diagram of a state transition graph]

- **Blue state**: only one $\varphi$-state.
Another example: How to handle reachability properties

Wanted: counterexamples for $\mathcal{P}_{\leq 0.4} (F \varphi)$, or, in shorthand, $(F \leq 0.4 \varphi)$.

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

\[
\begin{bmatrix}
  s_0 & s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\
  s_0 & 0 & .5 & .25 & 0 & 0 & .25 & 0 & 0 & 0 \\
  s_1 & 0 & 0 & .5 & .5 & 0 & 0 & 0 & 0 & 0 \\
  s_2 & 0 & .5 & 0 & 0 & .5 & 0 & 0 & 0 & 0 \\
  s_3 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
  s_4 & 0 & .7 & 0 & .3 & 0 & 0 & 0 & 0 & 0 \\
  s_5 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
  s_6 & 0 & 0 & 0 & .5 & 0 & 0 & 0 & .5 & 0 \\
  s_7 & 0 & 0 & 0 & 0 & .25 & .25 & 0 & .5 & 0 \\
  s_8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 
\end{bmatrix}
\]

blue state: only one $\varphi$-state.
Another example: How to handle reachability properties

Wanted: counterexamples for $\mathcal{P}_{\leq 0.4}(F \varphi)$, or, in shorthand, $(F^{\leq 0.4} \varphi)$.

![State transition diagram]

**Approach 1**, based on using the transition (right-stochastic) $9 \times 9$ matrix $A$:

\[
\begin{bmatrix}
    s_0 & s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\
    s_0 & 0 & .5 & .25 & 0 & 0 & .25 & 0 & 0 & 0 \\
    s_1 & 0 & 0 & .5 & .5 & 0 & 0 & 0 & 0 & 0 \\
    s_2 & 0 & .5 & 0 & 0 & .5 & 0 & 0 & 0 & 0 \\
    s_3 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
    s_4 & 0 & .7 & 0 & .3 & 0 & 0 & 0 & 0 & 0 \\
    s_5 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
    s_6 & 0 & 0 & 0 & .5 & 0 & 0 & 0 & .5 & 0 \\
    s_7 & 0 & 0 & 0 & 0 & .25 & .25 & 0 & .5 & 0 \\
    s_8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

- blue state: only one $\varphi$-state.
- initial distribution over 9 states is $d_0 = (1, 0, 0, 0, 0, 0, 0, 0, 0) = [1 0 0 0 0 0 0 0 0]$.
- distribution after 1 transition, 2 transitions, and 3 transitions, respectively:
  \[
  d_1 = d_0 \cdot A = (0, .5, .25, 0, 0, .25, 0, 0, 0)
  \]
  \[
  d_2 = d_0 \cdot A^2 = (0, .125, .25, .25, .125, 0, .25, 0, 0)
  \]
  \[
  d_3 = d_0 \cdot A^3 = (0, .2125, .0625, .475, .125, 0, 0, .125, 0)
  \]
Another example: How to handle reachability properties

**Wanted:** counterexamples for \( P_{\leq 0.4}(F \varphi) \), or, in shorthand, \( (F^{\leq 0.4} \varphi) \).

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

\[
\begin{bmatrix}
    s_0 & s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\
    s_0 & 0 & .5 & .25 & 0 & 0 & .25 & 0 & 0 & 0 \\
    s_1 & 0 & 0 & .5 & .5 & 0 & 0 & 0 & 0 & 0 \\
    s_2 & 0 & .5 & 0 & 0 & .5 & 0 & 0 & 0 & 0 \\
    s_3 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
    s_4 & 0 & .7 & 0 & .3 & 0 & 0 & 0 & 0 & 0 \\
    s_5 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
    s_6 & 0 & 0 & 0 & .5 & 0 & 0 & 0 & .5 & 0 \\
    s_7 & 0 & 0 & 0 & 0 & 0 & .25 & .25 & 0 & .5 \\
    s_8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 
\end{bmatrix}
\]

**Conclusion:** Starting from \( s_0 \), state \( s_3 \) is reached with probability \(.475 > .4\) after 3 transitions.

- Hence, there is a counterexample \( C \) for \( s_0 \not\models (F^{\leq 0.4} \varphi) \) 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)

**blue state**: only one \( \varphi \)-state.
Another example: How to handle reachability properties

**Wanted:** counterexamples for $P_{\leq 0.4}(F \varphi)$, or, in shorthand, $(F^{\leq 0.4} \varphi)$.

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

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

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

  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, $M \models P_{\leq \rho}(F \text{ target})$ iff $\sigma(p_{s_0}) \leq \rho$, 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 $P_{\leq 0.4}(F\varphi)$, or, in shorthand, $(F_{\leq 0.4}\varphi)$.

- For the Markov chain $M$ shown on slide 21, we obtain:

\[
\begin{align*}
    p_{s_0} &= 0.5 p_{s_1} + 0.25 p_{s_2} + 0.25 p_{s_5} & p_{s_1} &= 0.5 p_{s_2} + 0.5 p_{s_3} \\
    p_{s_2} &= 0.5 p_{s_1} + 0.5 p_{s_4} & p_{s_3} &= 1 \\
    p_{s_4} &= 0.7 p_{s_1} + 0.3 p_{s_3} & p_{s_5} &= 1 p_{s_6} \\
    p_{s_6} &= 0.5 p_{s_3} + 0.5 p_{s_7} & p_{s_7} &= 0.25 p_{s_5} + 0.25 p_{s_6}
\end{align*}
\]

We can remove all states from $M$ 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 $\sigma : \{p_{s_0}, p_{s_1}, \ldots, p_{s_7}\} \rightarrow [0, 1]$ such that:

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

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

Hence, there is a counterexample $C$ for $s_0 \nmodels (F_{\leq 0.4}\varphi)$, though we do not know the members of $C$ yet!!
Another example: How to handle reachability properties

**Wanted:** counterexamples for $\mathcal{P}_{\leq 0.4}(F \varphi)$, or, in shorthand, $(F^{\leq 0.4} \varphi)$.

- **Approach 3**, most efficient and most direct, repeats the steps carried out to find counterexamples for $s_0 \not\models (\varphi U^{1/2} \psi)$, from slide 12 to slide 20.

- We obtain, in order of decreasing probabilities:

<table>
<thead>
<tr>
<th>evidence</th>
<th>weight (rounded)</th>
<th>probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_1 \triangleq s_0 s_1 s_3$</td>
<td>1.39</td>
<td>0.25</td>
</tr>
<tr>
<td>$\pi_2 \triangleq s_0 s_5 s_6 s_3$</td>
<td>2.08</td>
<td>0.125</td>
</tr>
<tr>
<td>$\pi_3 \triangleq s_0 s_2 s_1 s_3$</td>
<td>2.77</td>
<td>0.0625</td>
</tr>
<tr>
<td>$\pi_4 \triangleq s_0 s_1 s_2 s_1 s_3$</td>
<td>2.77</td>
<td>0.0625</td>
</tr>
<tr>
<td>$\pi_5 \triangleq s_0 s_2 s_4 s_1 s_3$</td>
<td>3.13</td>
<td>0.04375</td>
</tr>
<tr>
<td>$\pi_6 \triangleq s_0 s_1 s_2 s_4 s_1 s_3$</td>
<td>3.13</td>
<td>0.04375</td>
</tr>
<tr>
<td>$\pi_7 \triangleq s_0 s_2 s_4 s_3$</td>
<td>3.28</td>
<td>0.03750</td>
</tr>
<tr>
<td>$\pi_8 \triangleq s_0 s_1 s_2 s_4 s_3$</td>
<td>3.28</td>
<td>0.03750</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

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.