The Complexity of Finding Memoryless POMDP Policies

Sebastian Junges

Including work with:
Bernd Becker, Nils Jansen, Joost-Pieter Katoen,
Guillermo Perez, Tim Quatmann, Ralf Wimmer,
Leonore Winterer, Tobias Winkler

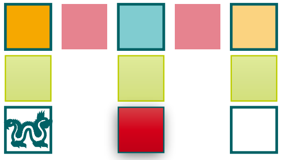
Radboud University, RWTH Aachen University,
University of Freiburg, University of Antwerp

Berkeley
UNIVERSITY OF CALIFORNIA

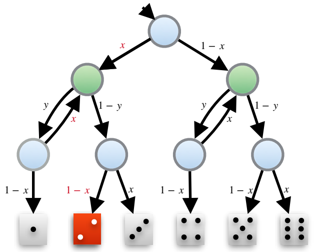
Outline

Step 1:

- Relate POMDPs + memoryless policies

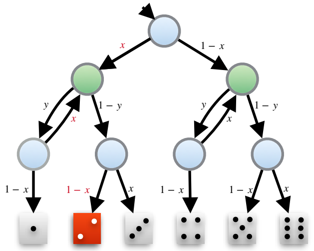


- to pMCs



Step 2:

- Discuss pMCs



- And relate them to the existential theory

$$-2x^2y + y \geqslant 5$$

$$\Updownarrow$$

$$2 \cdot \big((1-x)xy + (1-x)y + (1-y) - 1\big) + y \geqslant 5$$

$$\Updownarrow$$

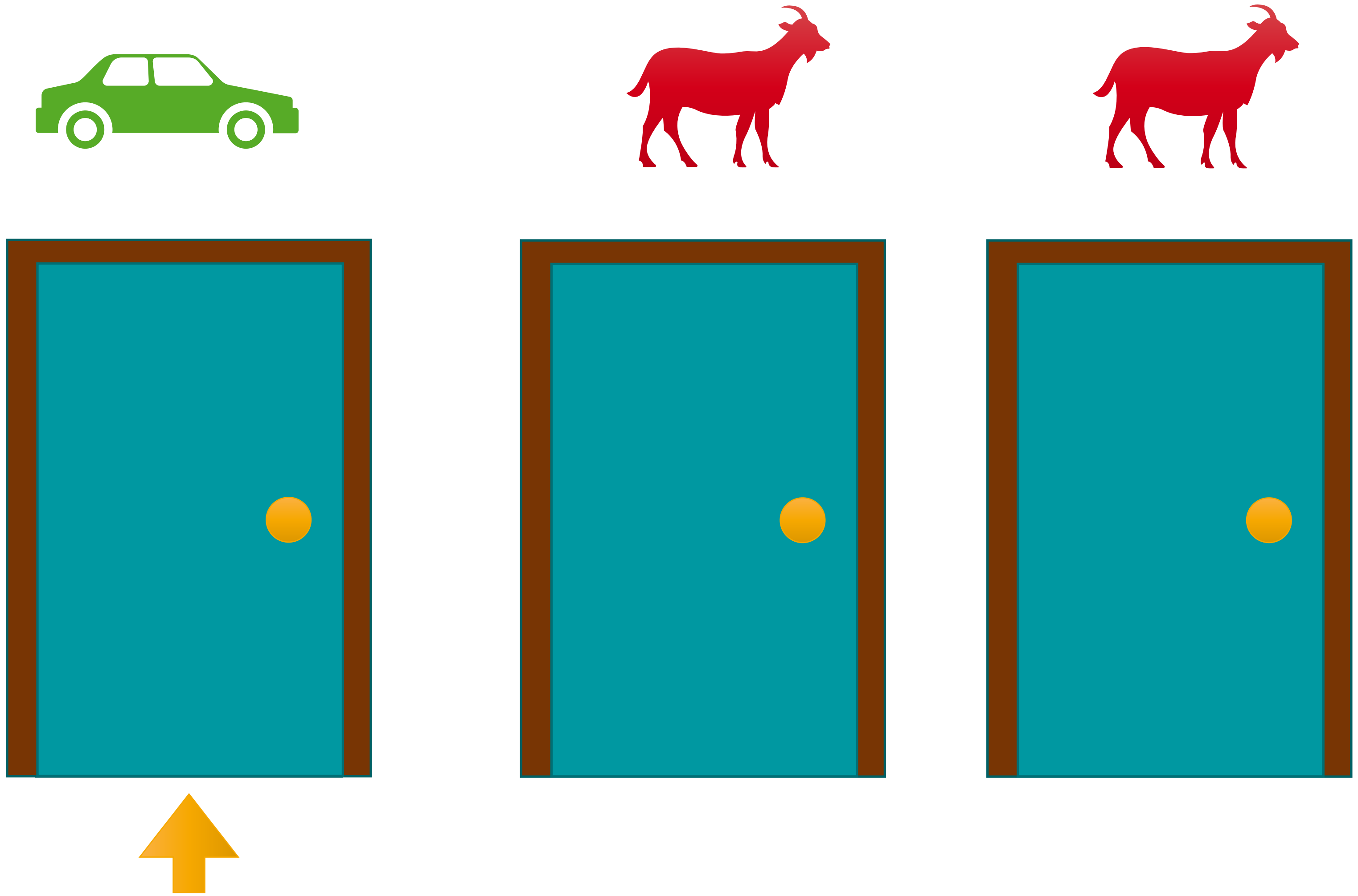
$$\frac{2 \cdot (1-x)xy + 2 \cdot (1-x)y + 2 \cdot (1-y) + y}{8} \geqslant \frac{7}{8}$$

For a formal treatment:

Sebastian Junges, Nils Jansen, Ralf Wimmer, Tim Quatmann, Leonore Winterer, Joost-Pieter Katoen, Bernd Becker: **Finite-State Controllers of POMDPs using Parameter Synthesis.** UAI 2018: 519-529

Sebastian Junges, Joost-Pieter Katoen, Guillermo A. Pérez, Tobias Winkler: **The Complexity of Reachability in Parametric Markov Decision Processes.** CoRR abs/2009.13128 (2020)

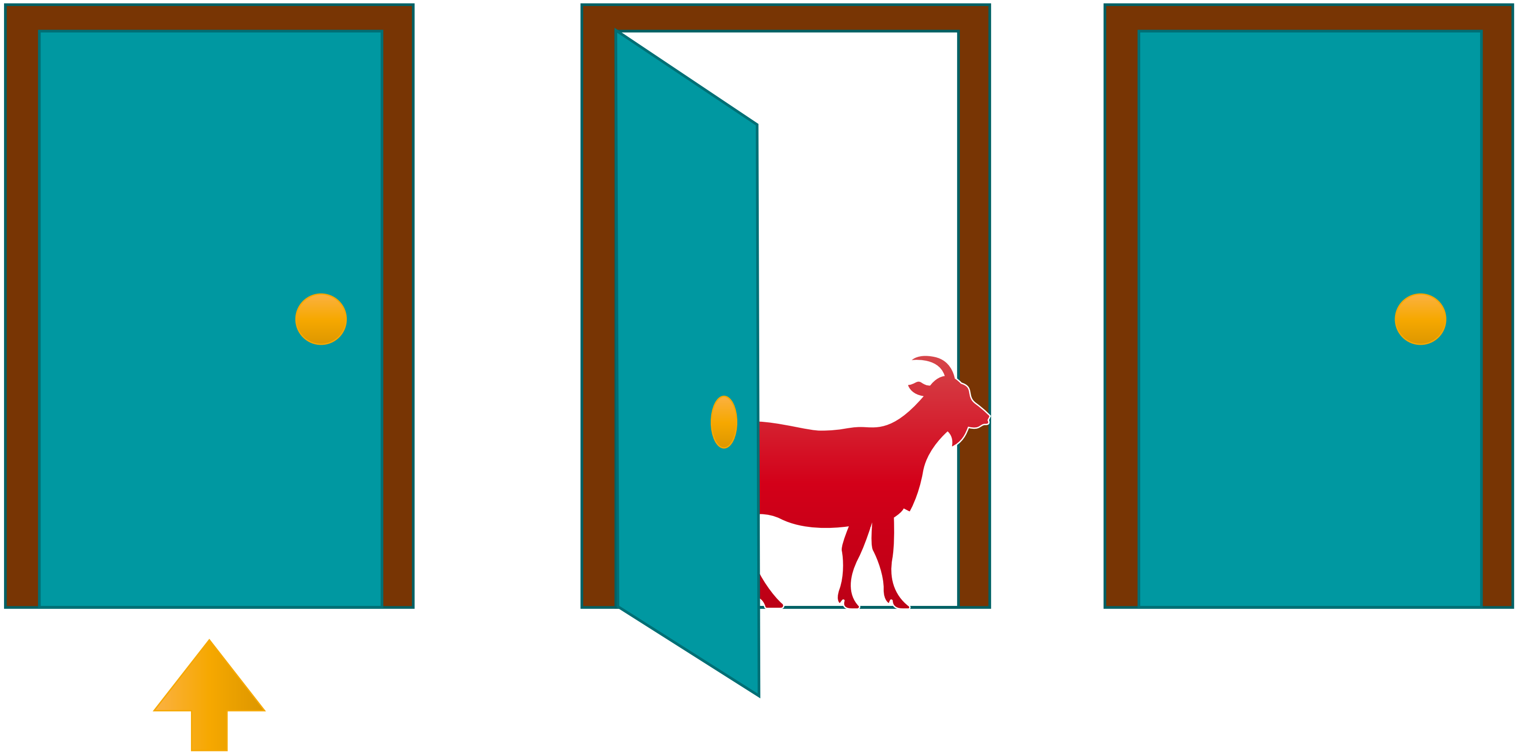
Monty Hall Problem



Monty Hall Problem

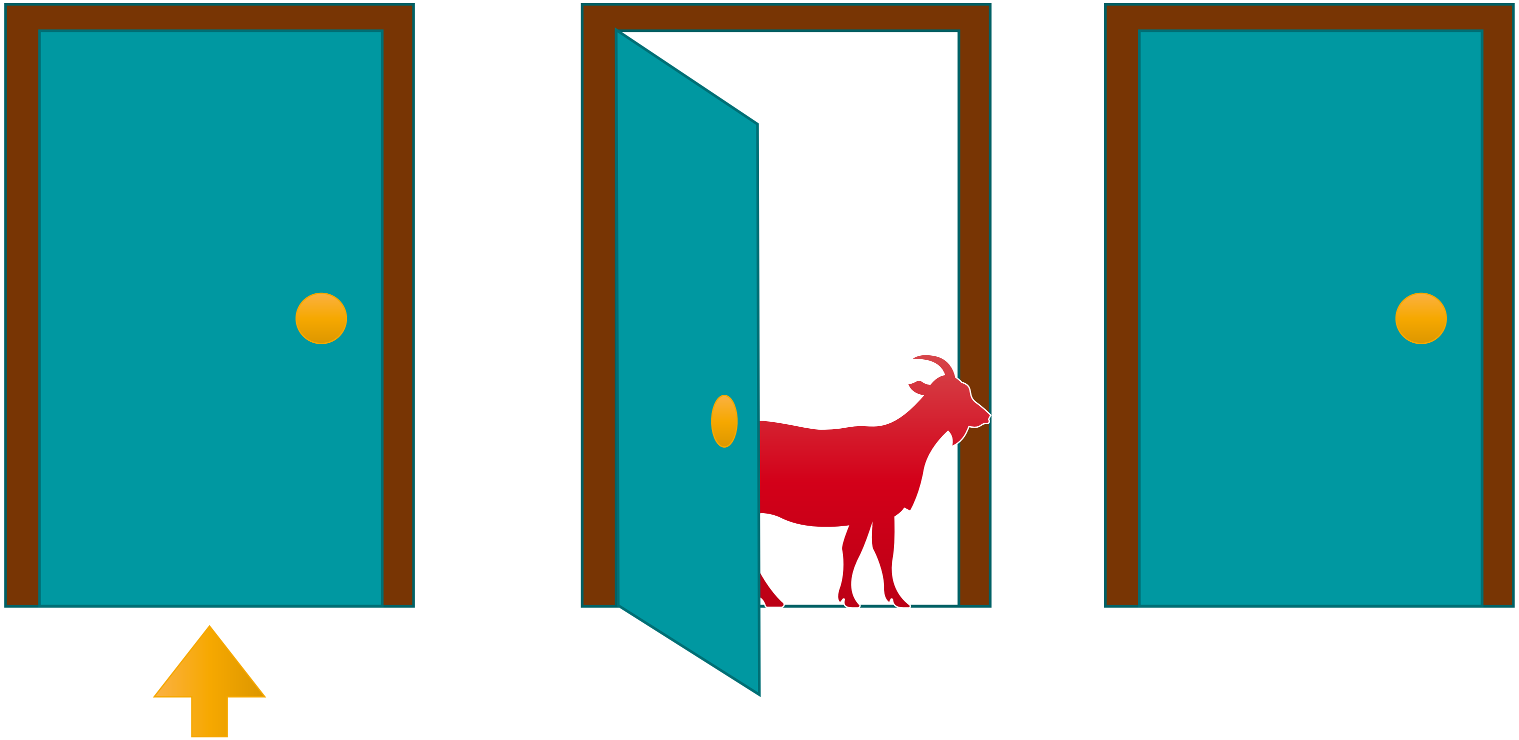
Proposal: Change if the car is behind the other door.

Strategy depends on unobservable information

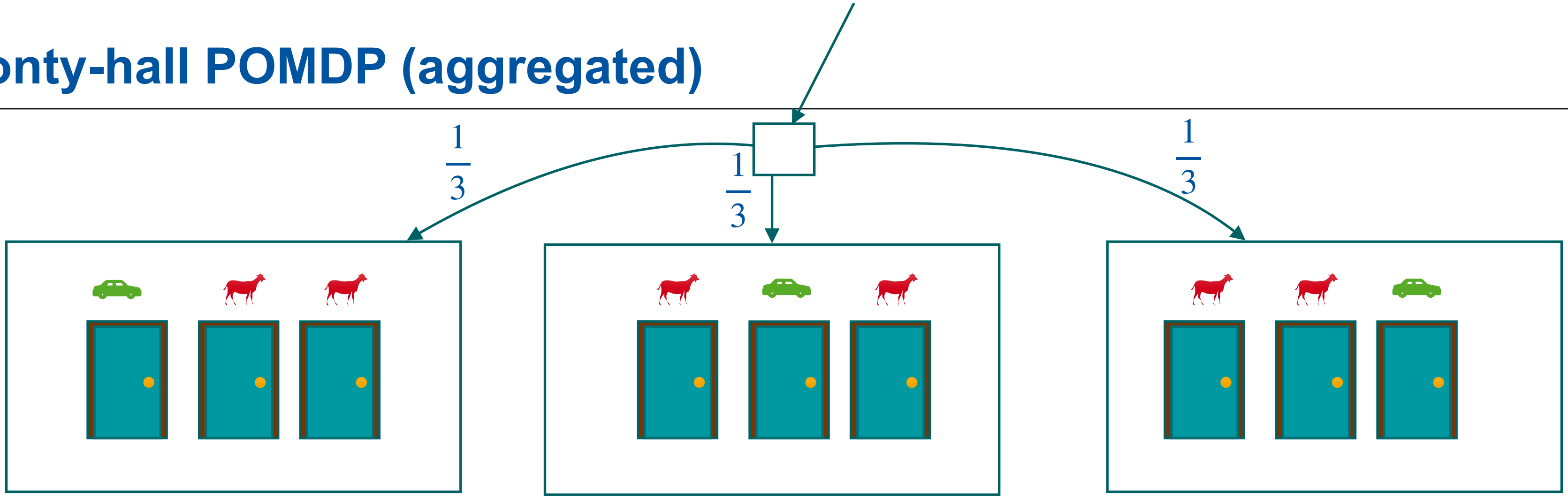


Monty Hall Problem: Humans are bad in reasoning under uncertainty

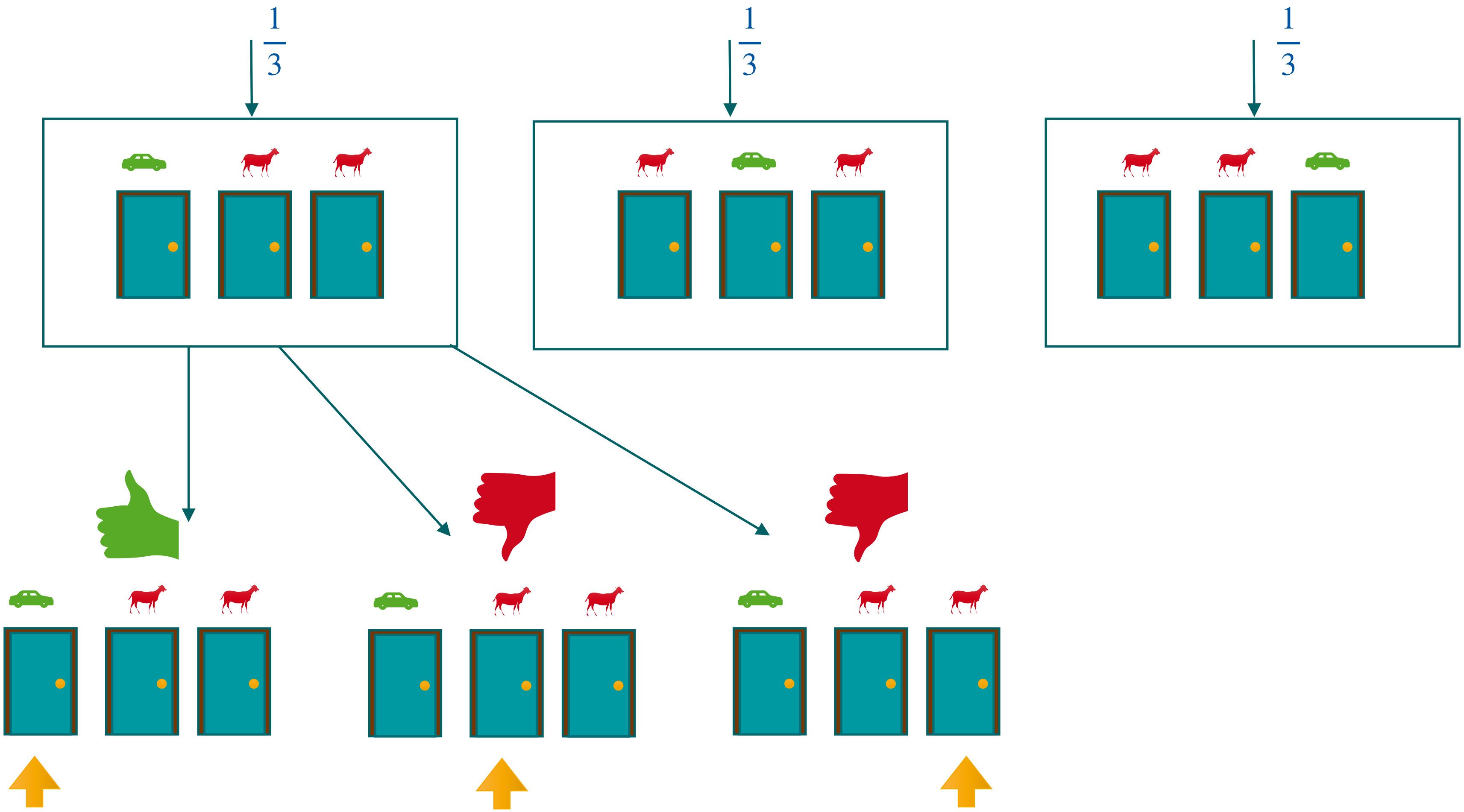
Should you change now?



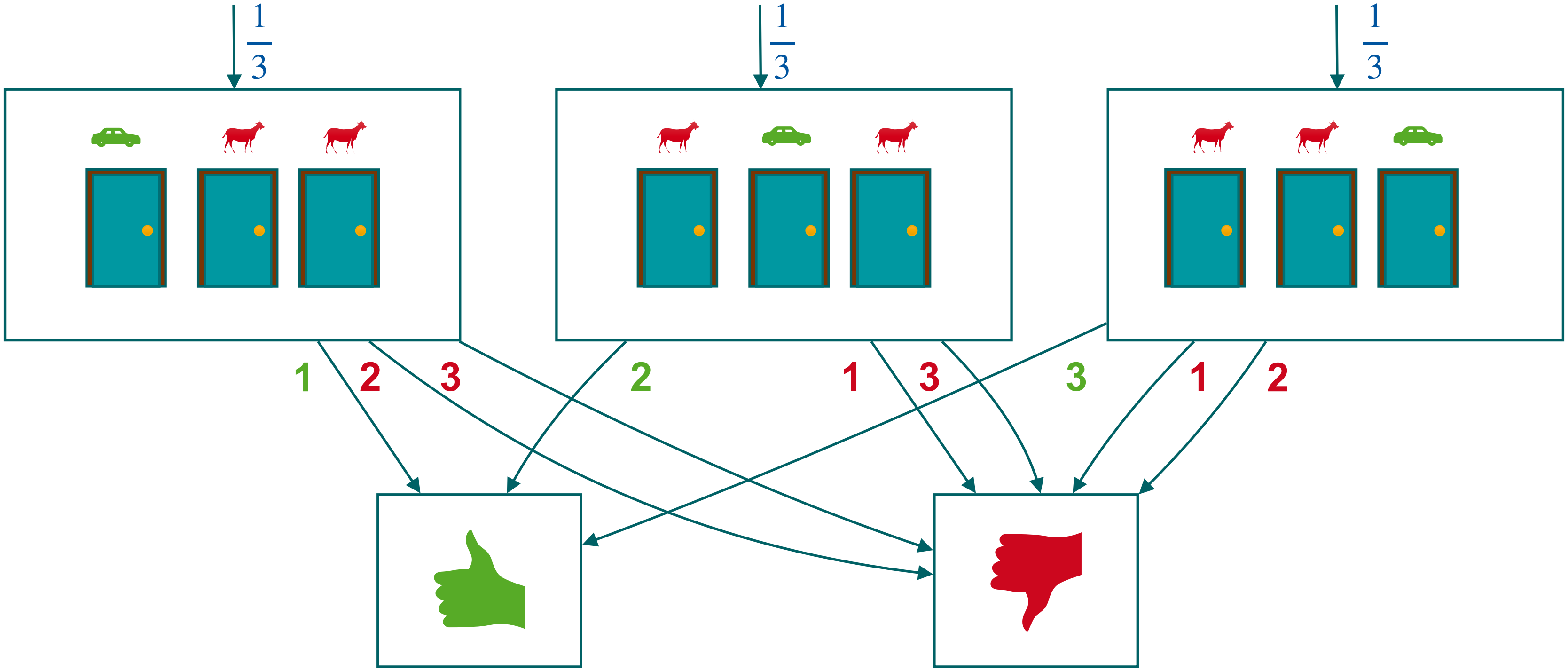
Monty-hall POMDP (aggregated)



Monty-hall POMDP (aggregated)

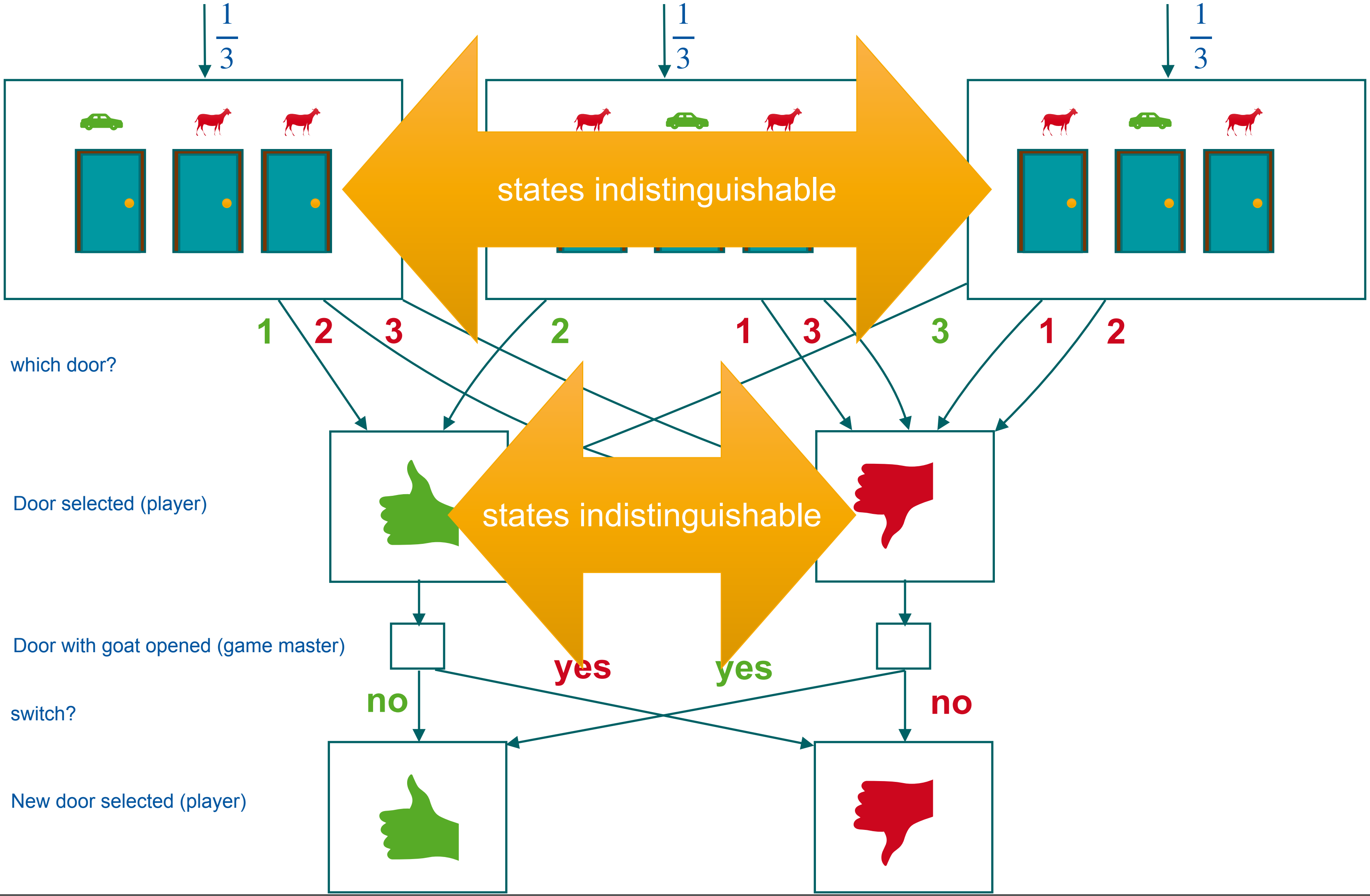


Monty-hall POMDP (aggregated)



Monty-hall POMDP (aggregated)

memory does not help



which door?

Door selected (player)

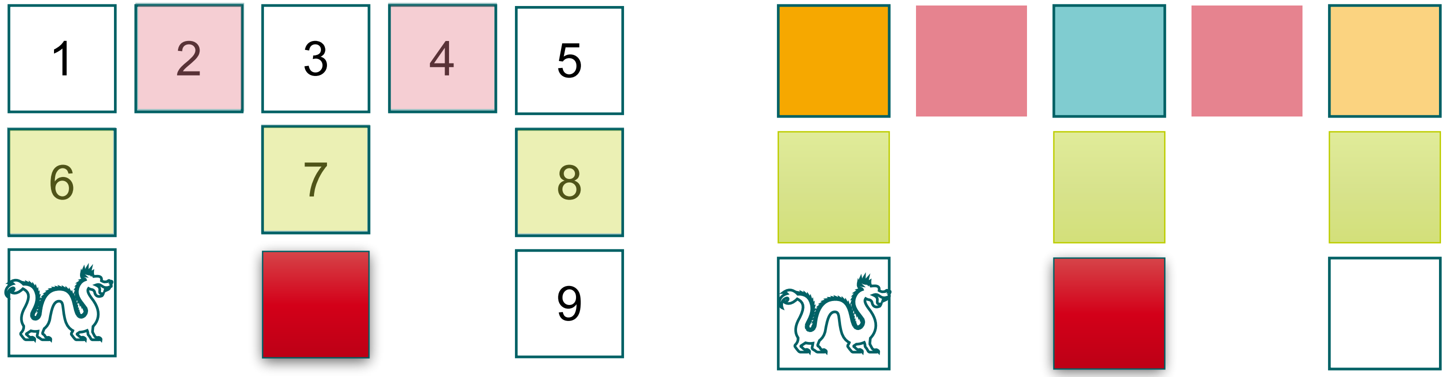
Door with goat opened (game master)

switch?

New door selected (player)

Randomisation and memory

POMDP: Reach red state without visiting the dragon.



same observations:
- {2,4}
- {6,7,8}

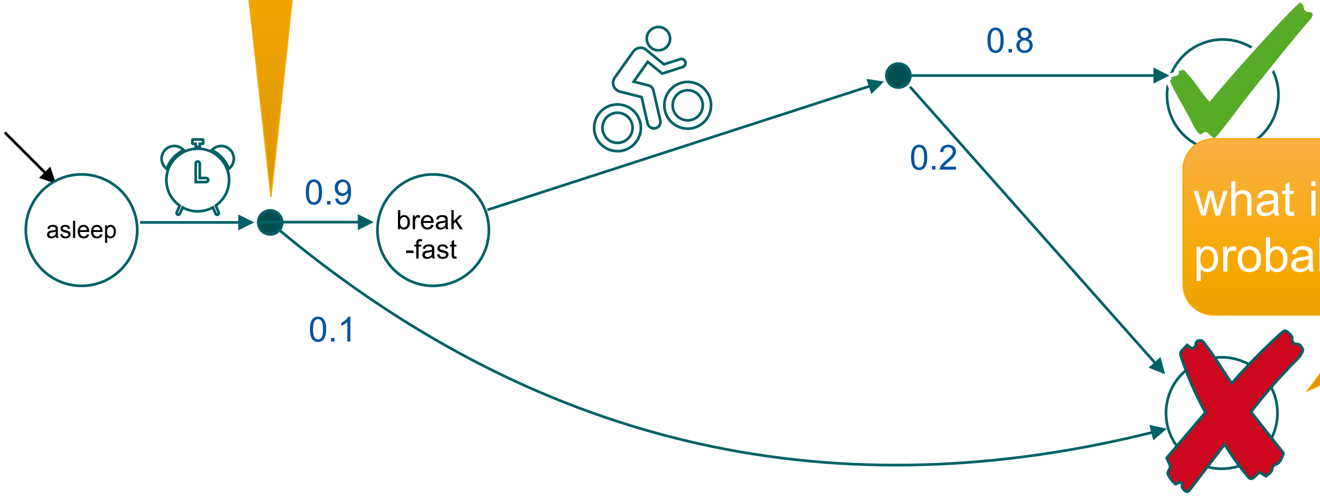
Start in 1 or 5:
Memoryless policy has to randomise in {2,4}

Start in 6 or 7:
no memoryless policy
store whether we have been in 3

Markov chain (MC): Arriving before 10am

in 2019!

every state/action maps to a distribution over successors

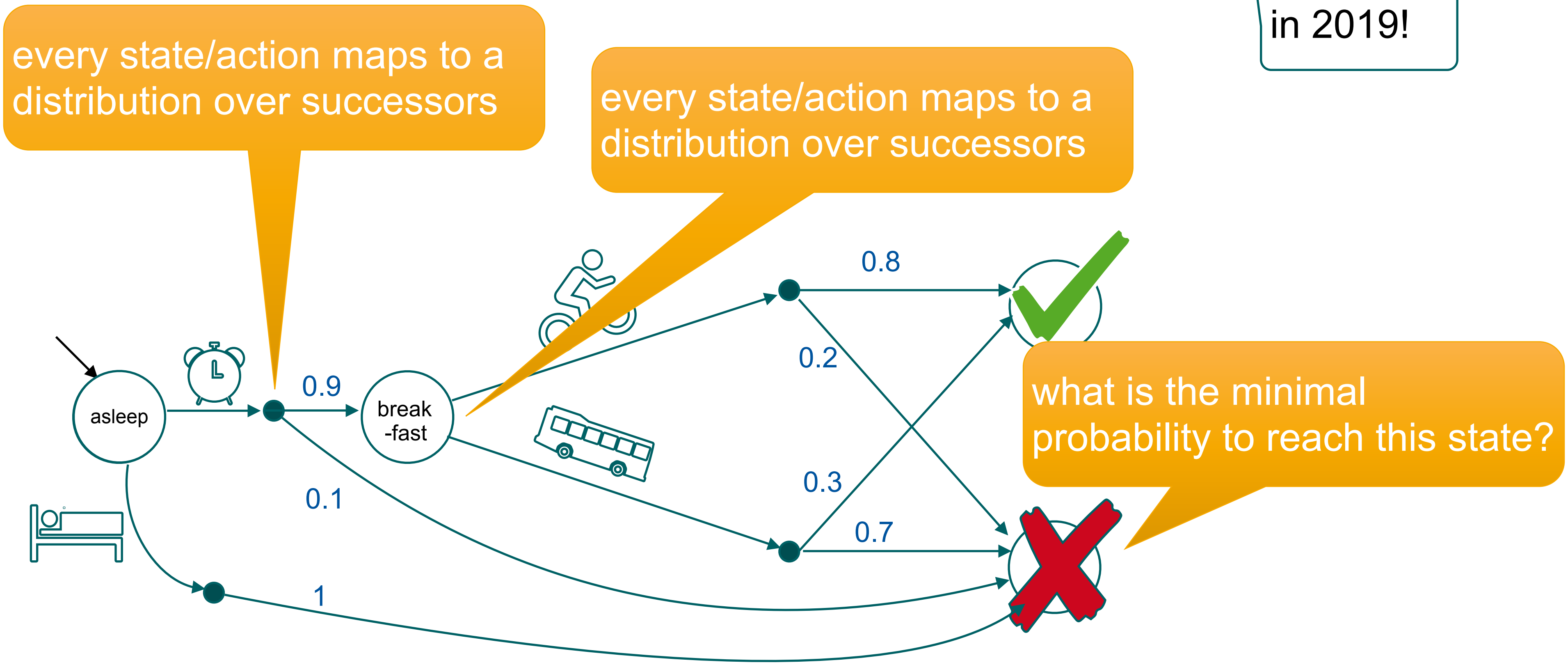


what is the probability to reach this state?

- MCs are Markov Decision Processes with one action in every state

Markov decision processes (MDP): Arriving before 10am

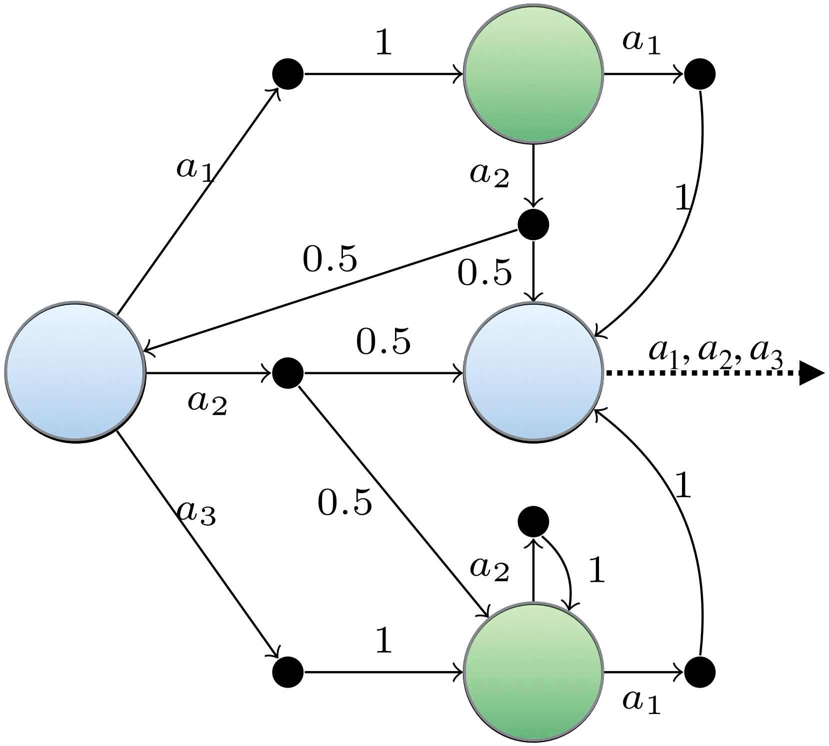
in 2019!

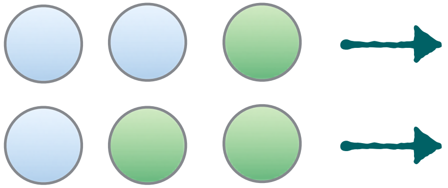


- MCs are Markov Decision Processes with one action in every state

POMDPs

MDPs with ‘observable colours’





Distribution over actions

Distribution over actions

Given **any** POMDP
is there an observation-based policy s.t.
the probability reaching ● > λ

Solving POMDPs is undecidable

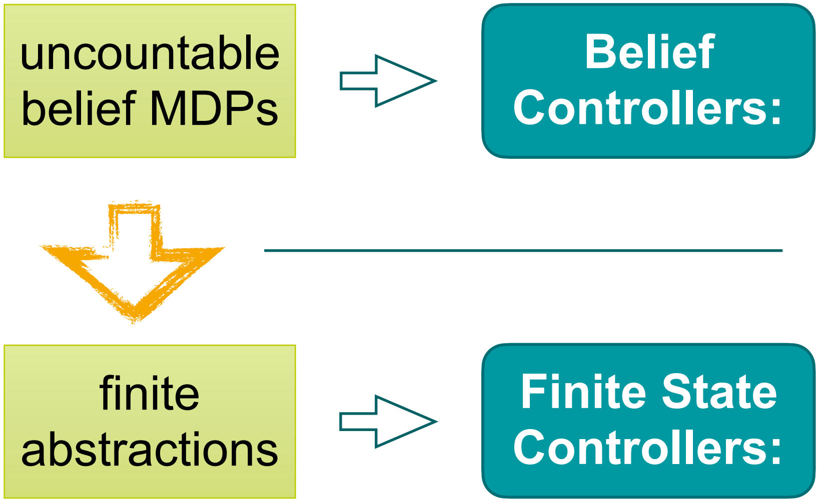
Given **any** POMDP
is there an observation-based policy s.t.
the probability reaching ● > λ

Undecidable!

But **cannot** be **avoided as** the **world**
is **a POMDP** most of the time

AI — A Modern Approach

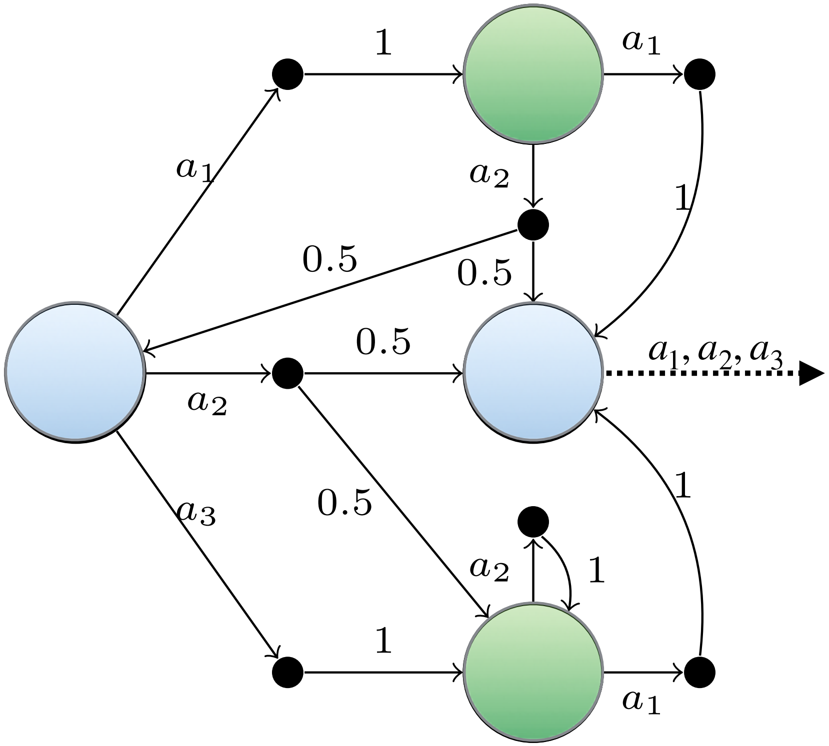
Approaches:



uncountable belief MDPs → **Belief Controllers:**

finite abstractions → **Finite State Controllers:**

Partially observable MDPs (POMDPs)



in PSPACE, NP-HARD
[Vlassis et al, 2012]

Given **any** POMDP
is there a memoryless strategy s.t. the
probability reaching ● >λ

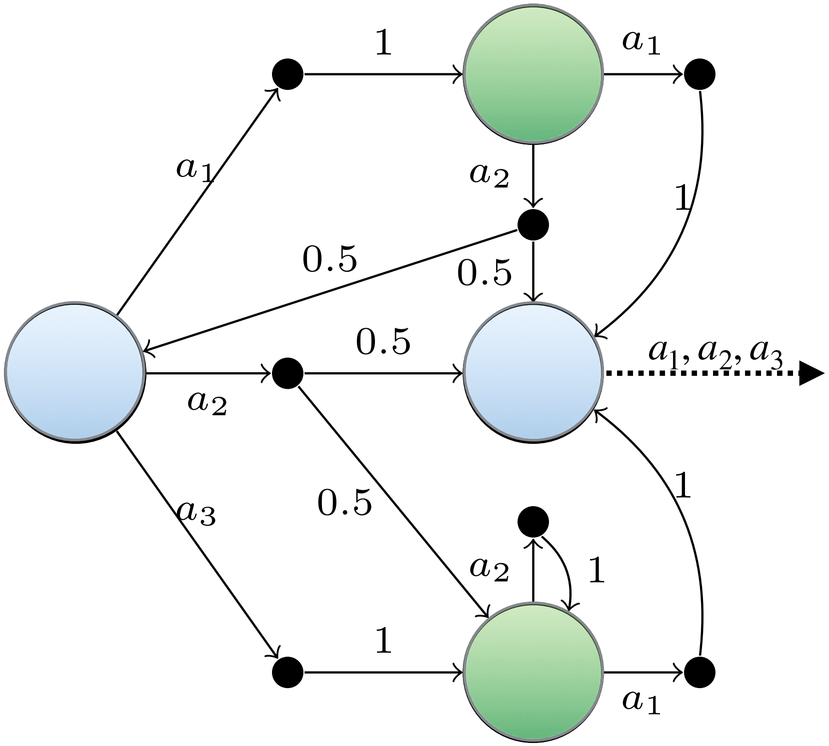
POMDP
memoryless strategy:
colours to distributions over actions

POMDPs with memoryless strategies

[UAI'18]

Maps observations to distributions over actions

maps observation/action pairs to probabilities



🟢 $a_1 \longmapsto 1 - x$

🟢 $a_2 \longmapsto x$

🔵 $a_1 \longmapsto y_1$

🔵 $a_2 \longmapsto y_2$

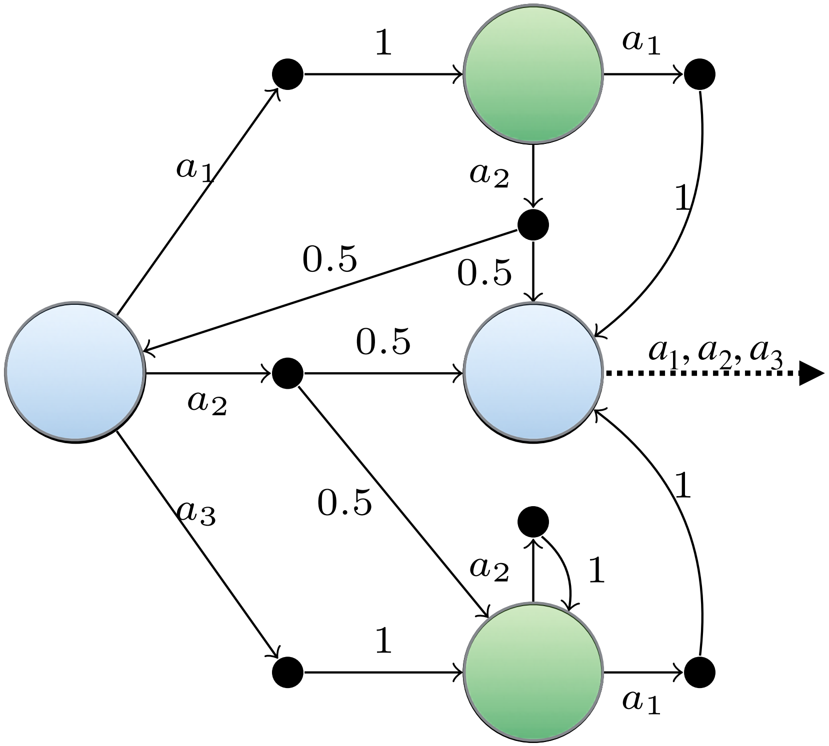
🔵 $a_3 \longmapsto 1 - y_1 - y_2$

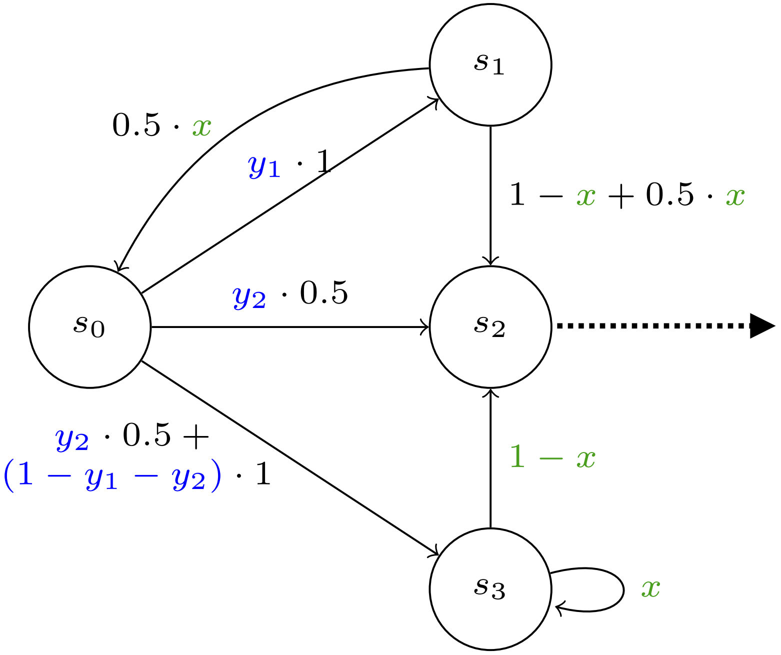
Strategy is uniquely described by values for x, y_1, y_2

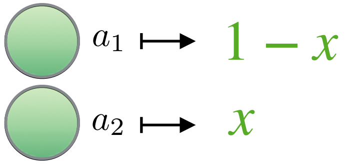
Applying the memoryless strategy

[UAI'18]

Induced Markov Chain with unknown probabilities

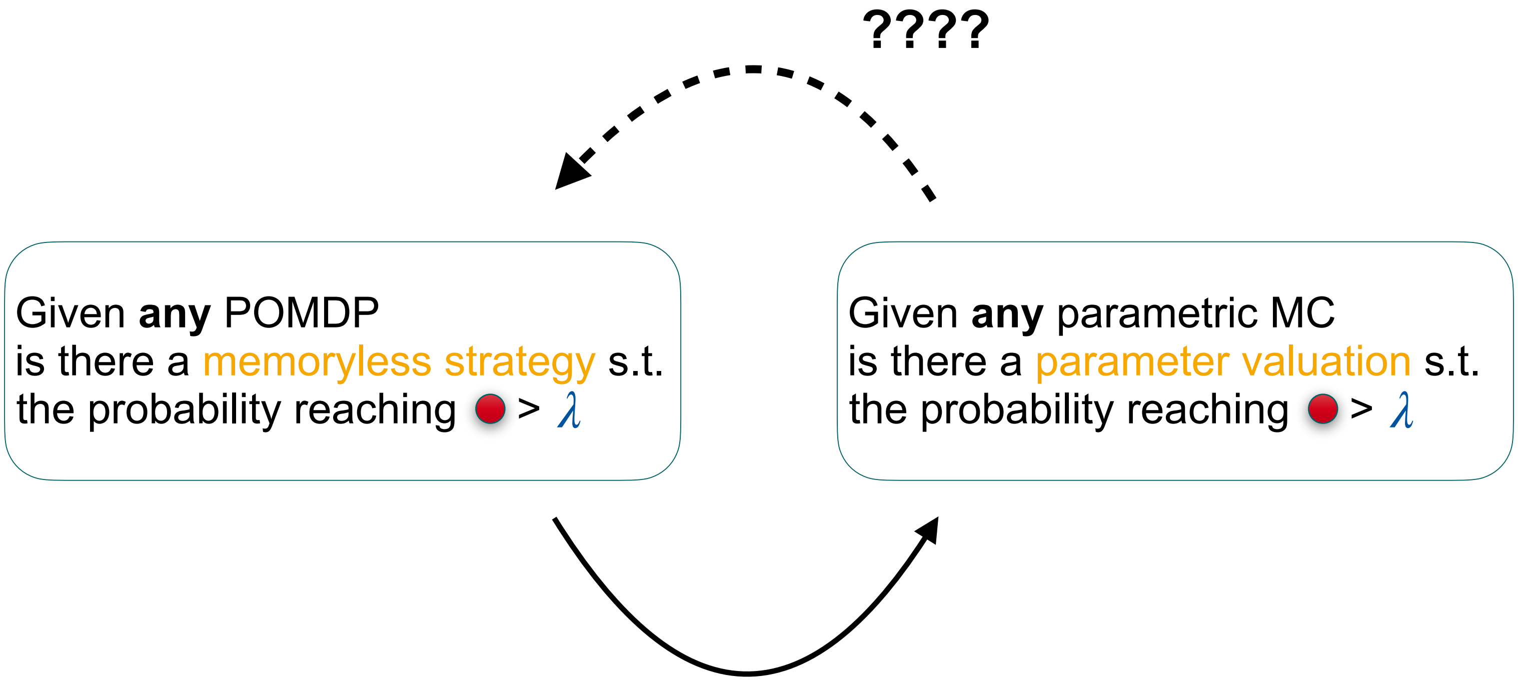






POMDPs and parametric Markov chains

[UAI'18]



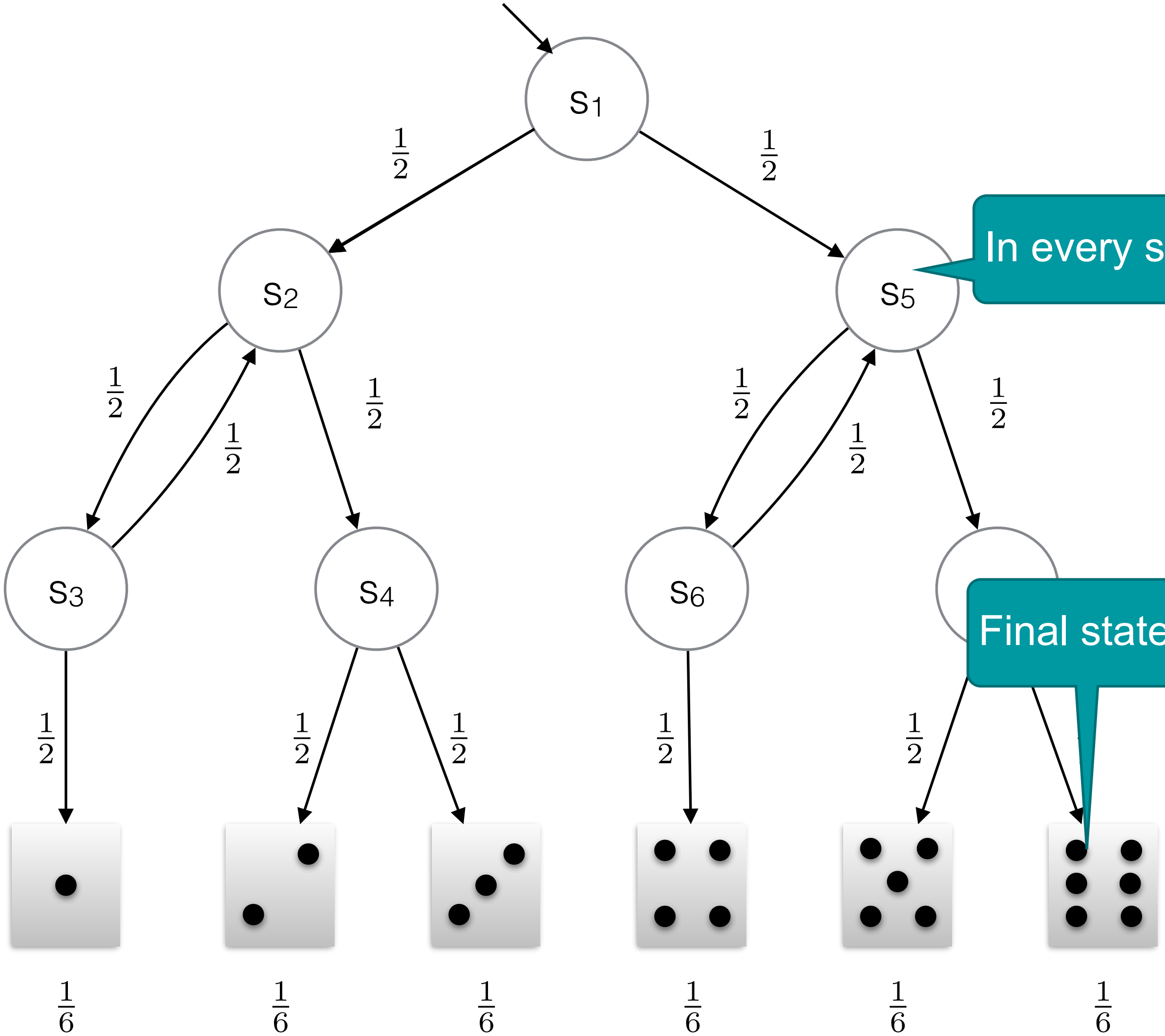
????

Given **any** POMDP
is there a memoryless strategy s.t.
the probability reaching ● > λ

Given **any** parametric MC
is there a parameter valuation s.t.
the probability reaching ● > λ

Markov chains

Knuth-Yao die

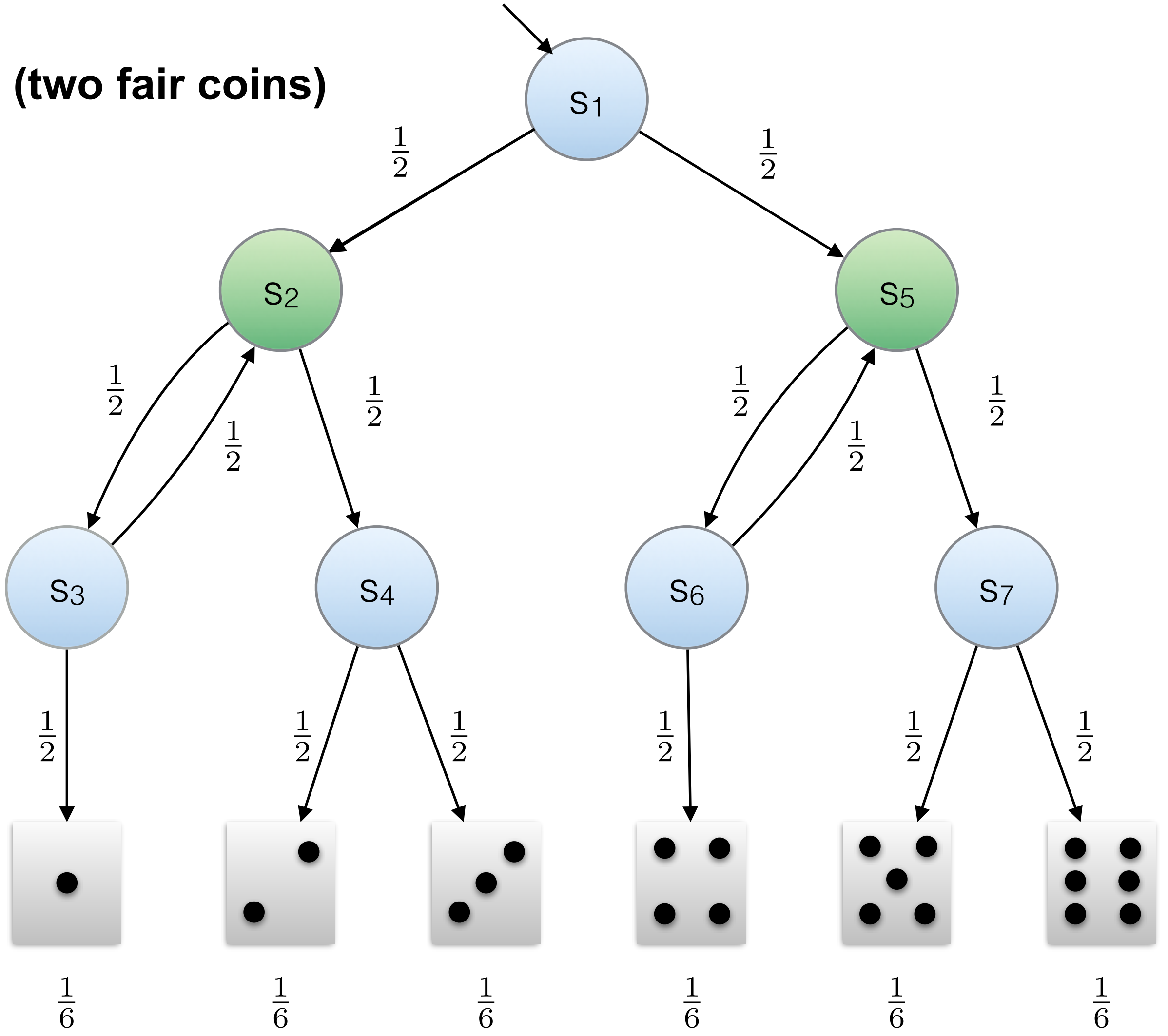


In every state, flip a coin

Final states: die outcomes

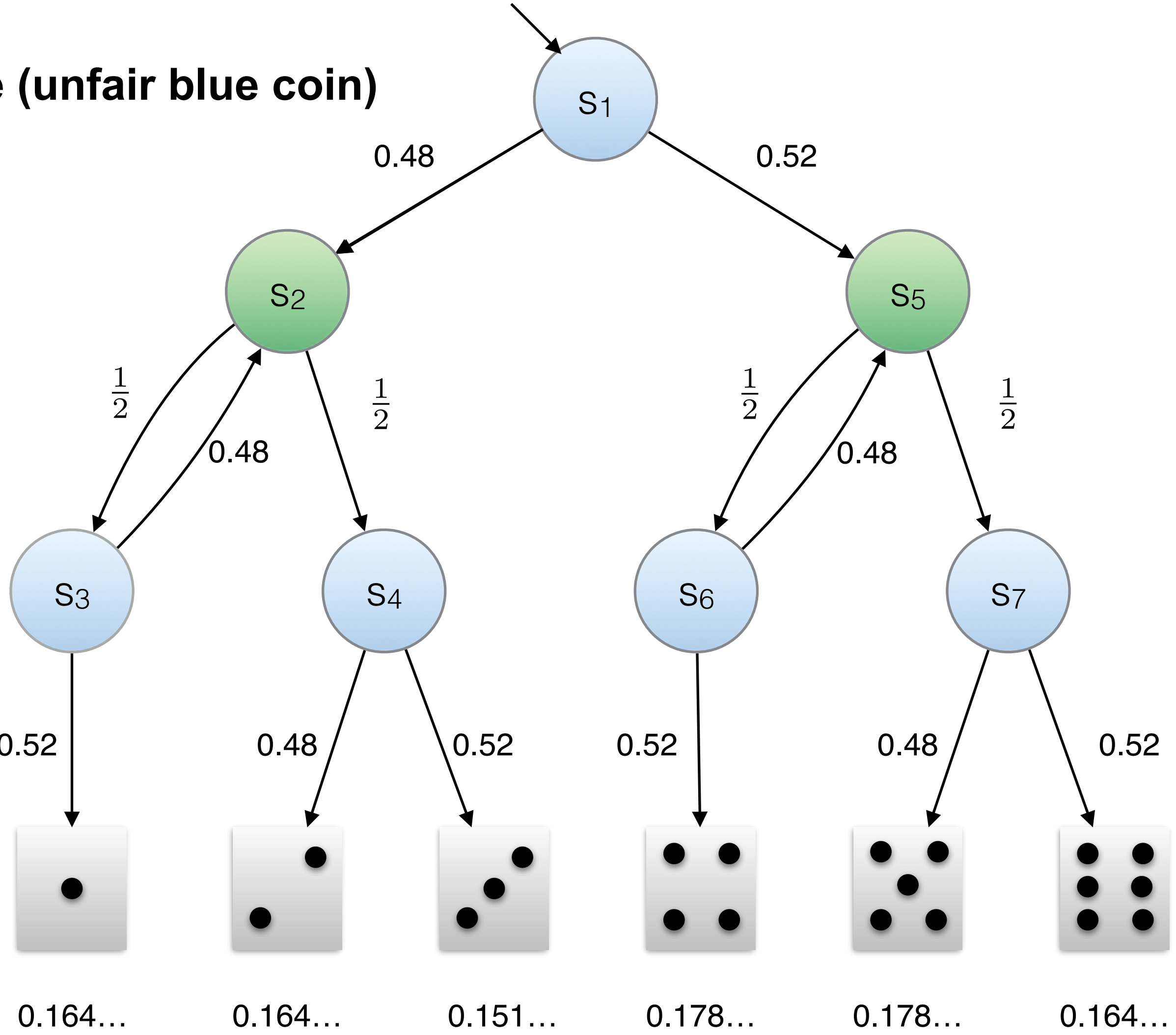
Markov chains

Knuth-Yao die (two fair coins)



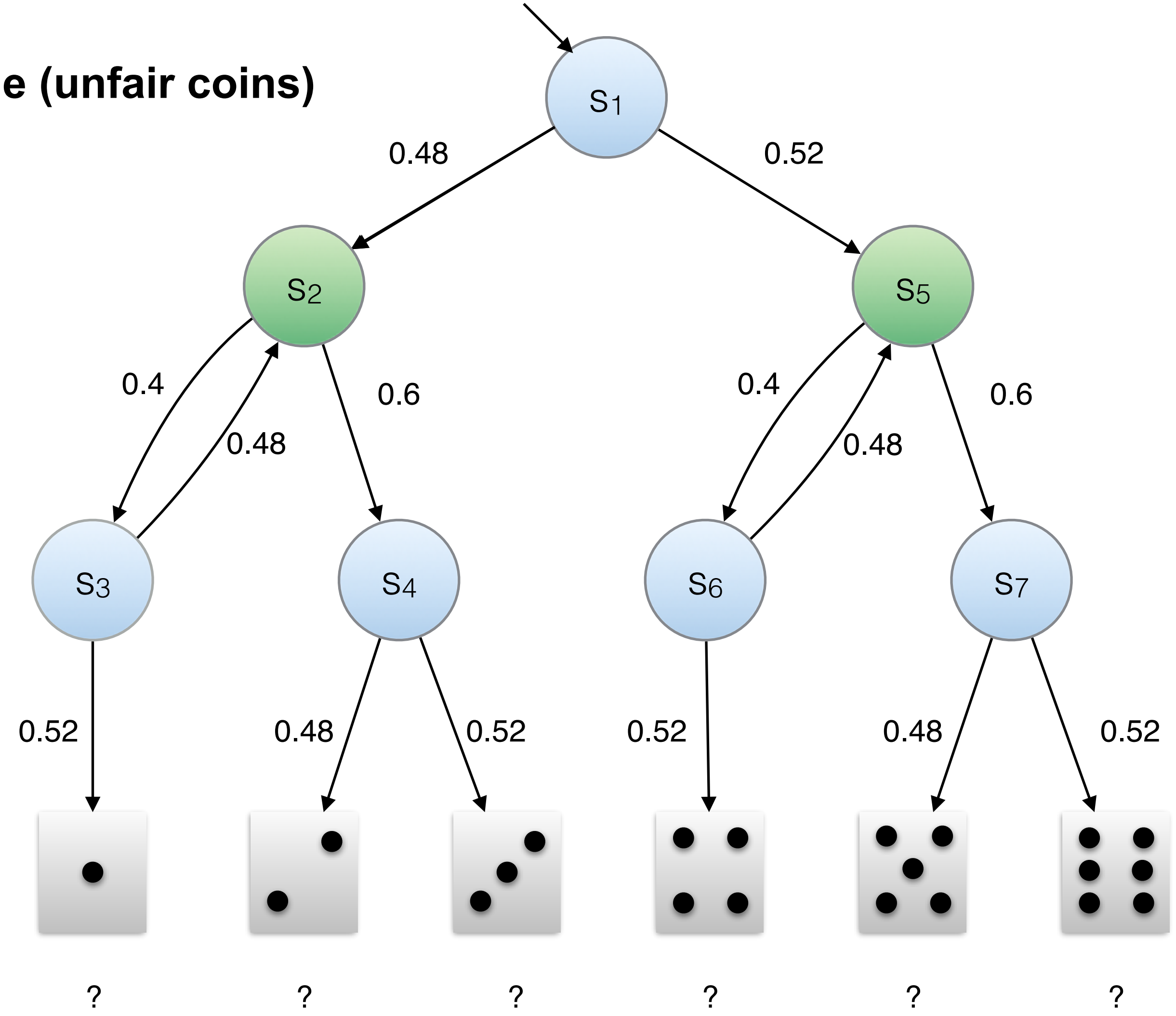
Markov chains

Knuth-Yao die (unfair blue coin)



Markov chains

Knuth-Yao die (unfair coins)



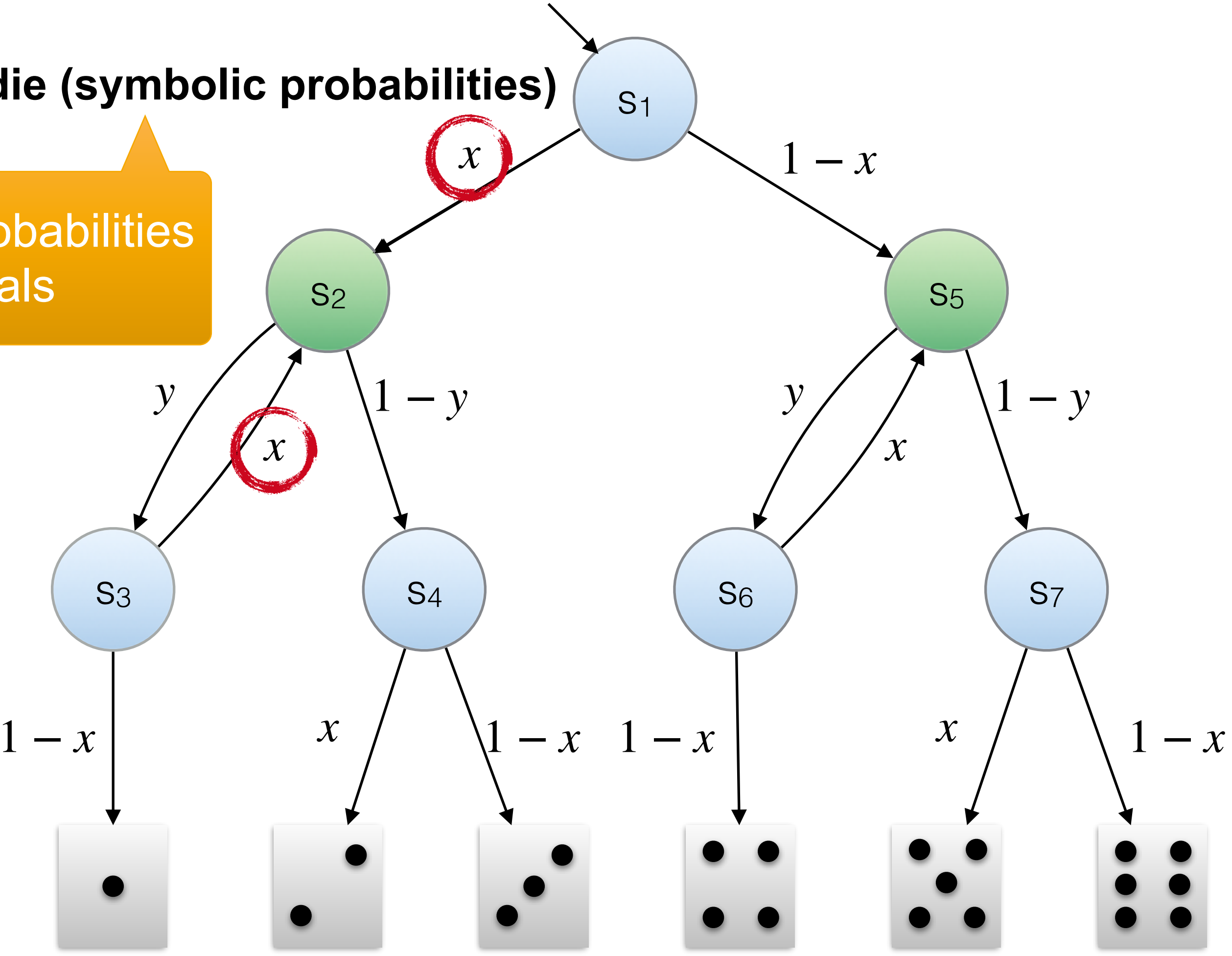
Parametric Markov chains (pMCs)

Knuth-Yao die (symbolic probabilities)

Transition probabilities are polynomials

in this talk:

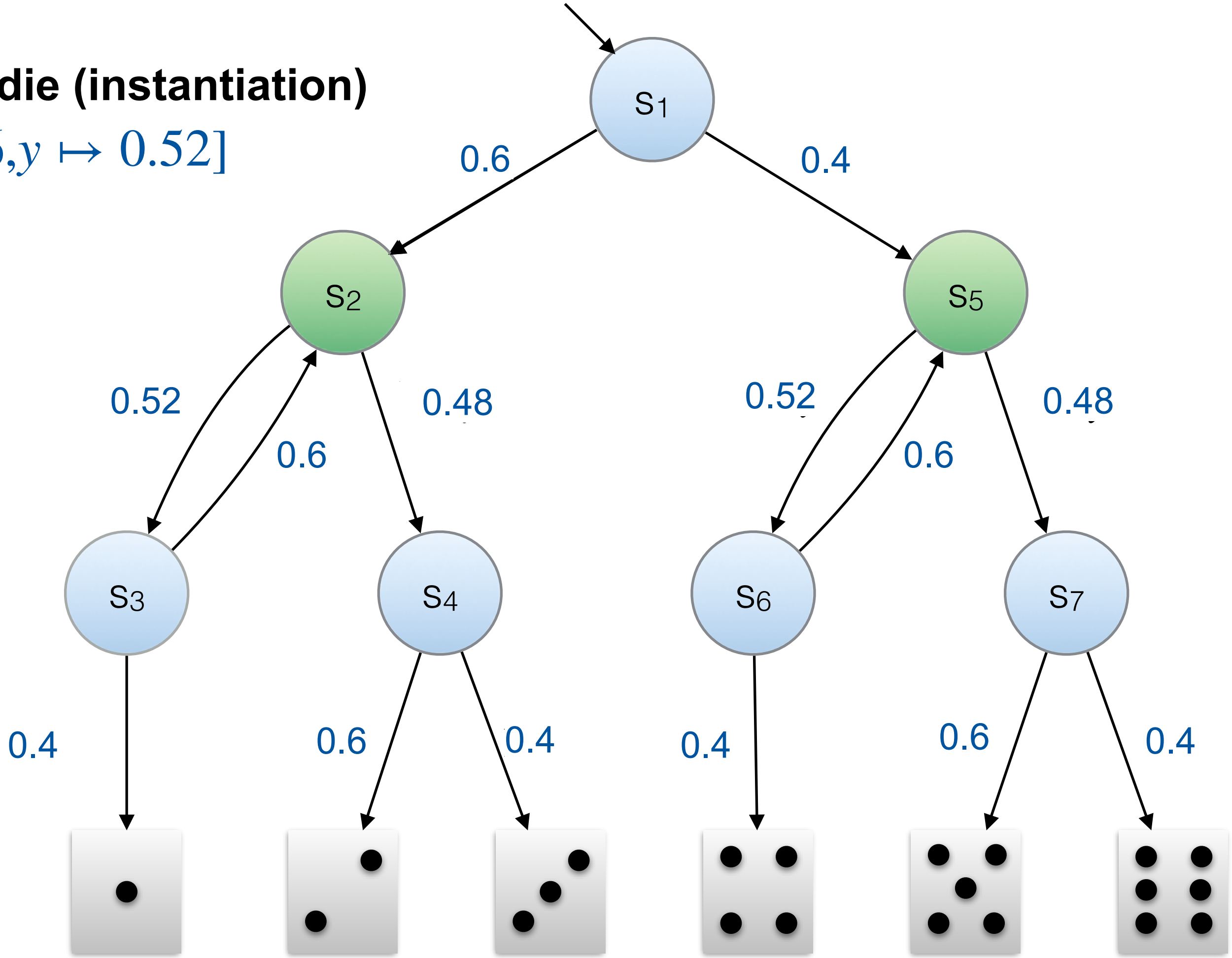
$\{x, 1-x\}$



Parametric Markov chains (pMCs)

Knuth-Yao die (instantiation)

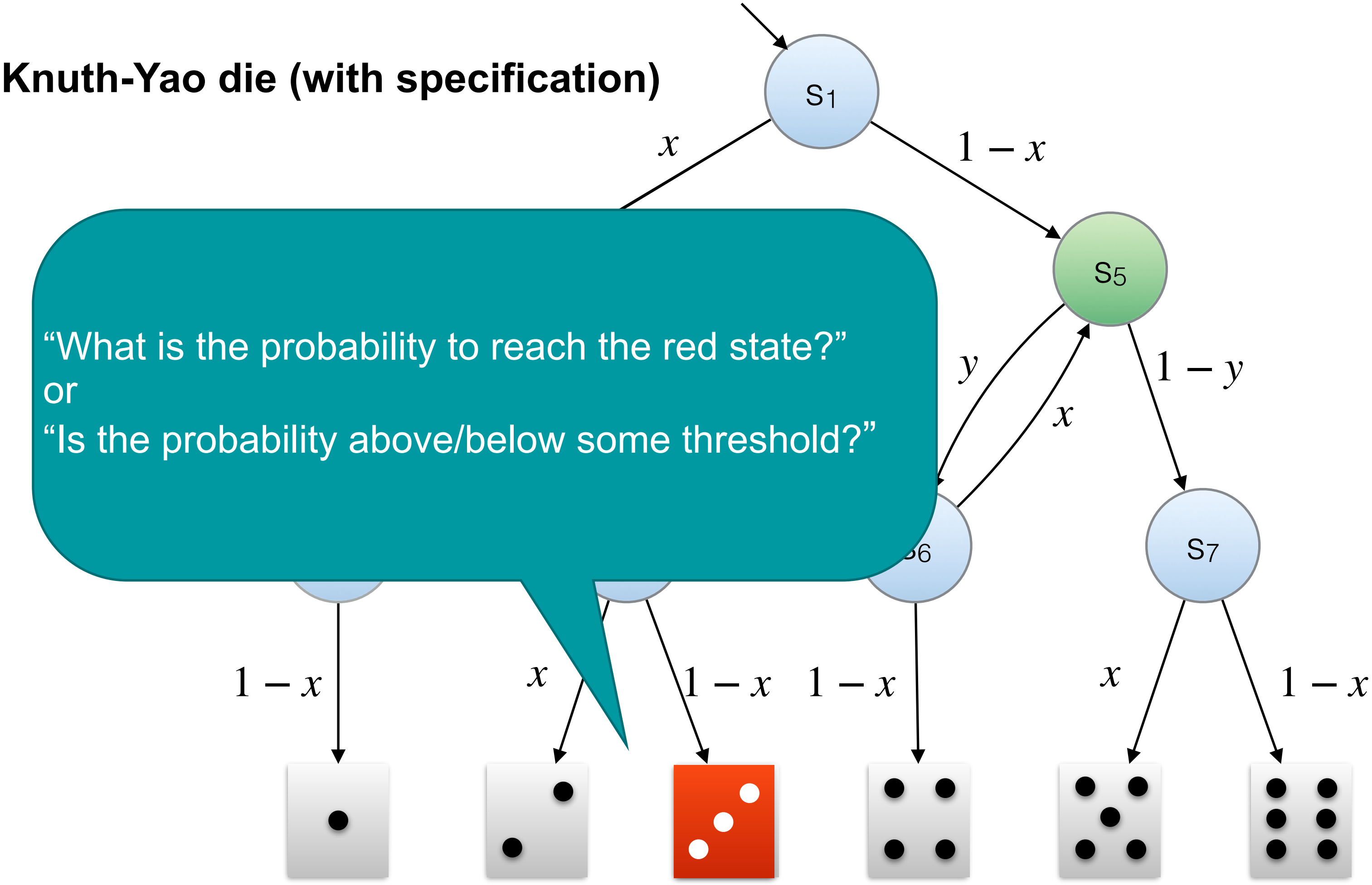
$\mathscr{M}[x \mapsto 0.6, y \mapsto 0.52]$



Berkeley
UNIVERSITY OF CALIFORNIA

Parametric Markov chains (pMCs)

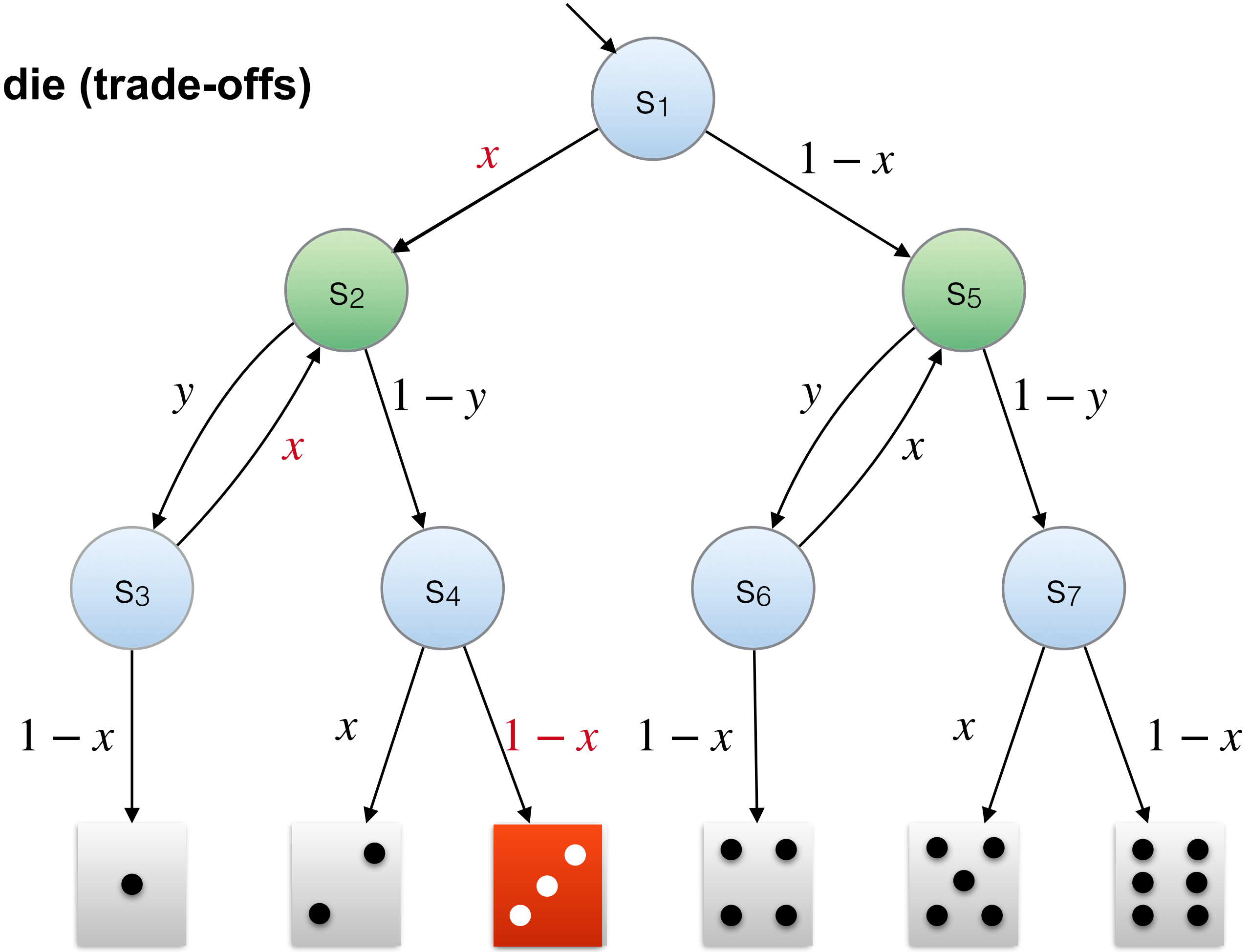
Knuth-Yao die (with specification)



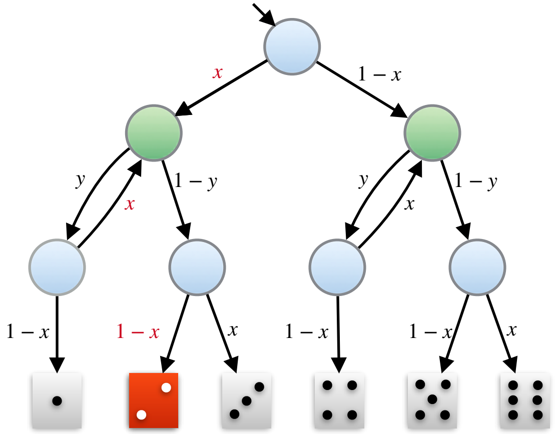
"What is the probability to reach the red state?"
or
"Is the probability above/below some threshold?"

Parametric Markov chains (pMCs)

Knuth-Yao die (trade-offs)



Problem statement: Parameter synthesis



Given:

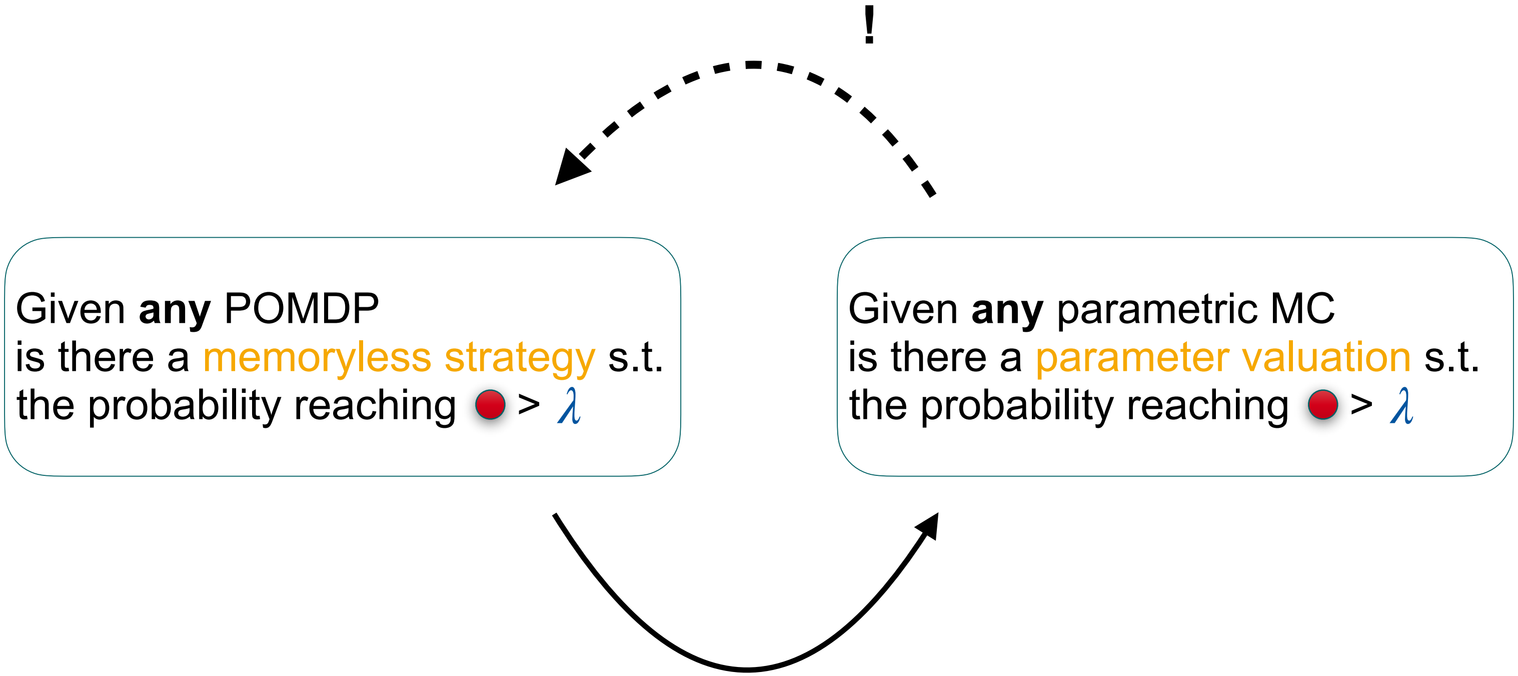
a parametric MC $\mathscr{M}$
with parameters $\mathbf{x}$

Find: $\mathrm{val}\colon \mathbf{x} \to [0,1]$

such that: $\mathscr{M}[\mathrm{val}] \vDash \varphi$, i.e., a red state is reached with probability at least/at most λ

POMDPs and parametric Markov chains

[UAI’18]

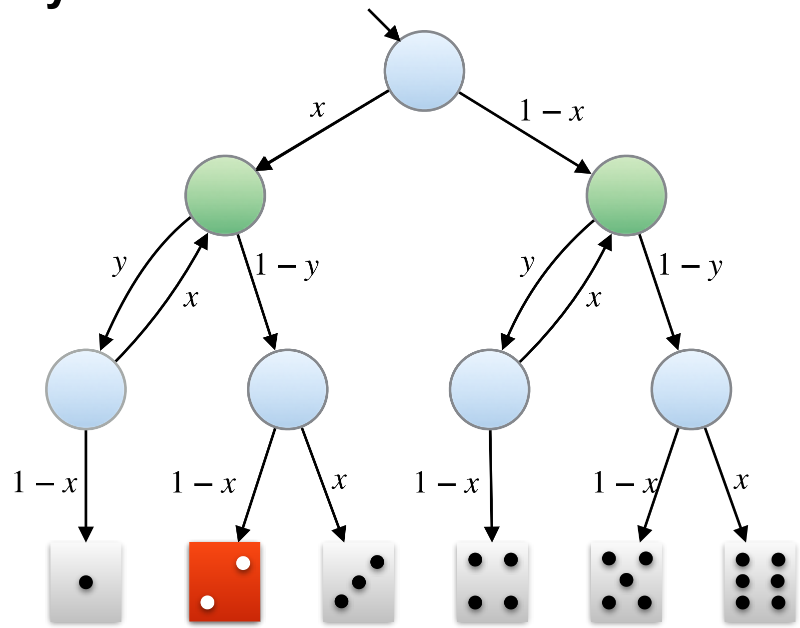


Given **any** POMDP
is there a memoryless strategy s.t.
the probability reaching ● > λ

Given **any** parametric MC
is there a parameter valuation s.t.
the probability reaching ● > λ

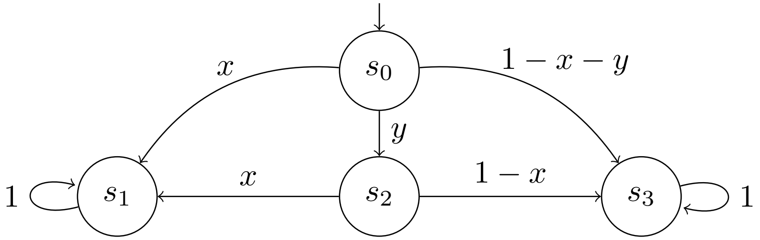
Simple pMC

Easily translated



Does not work in general

Counterexample:



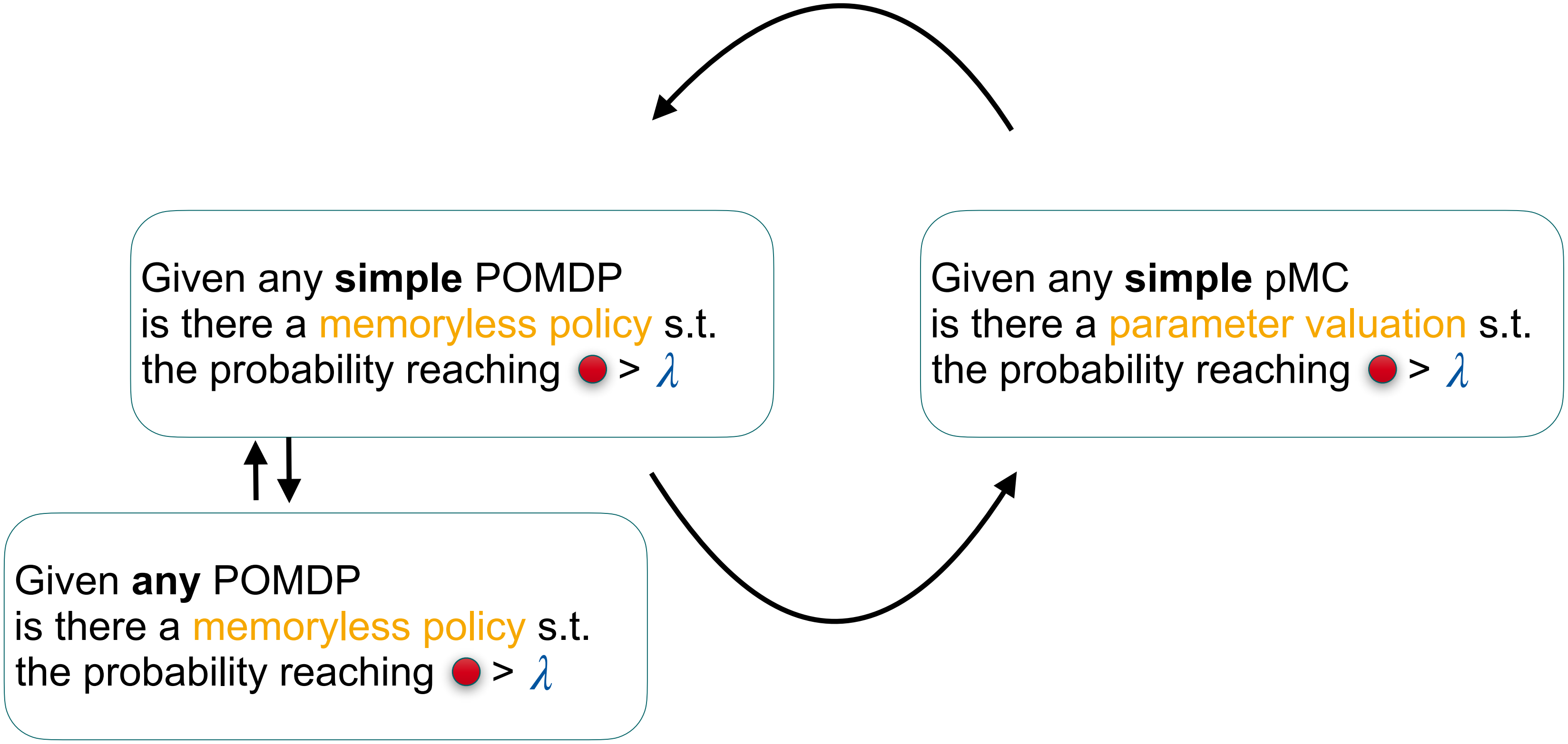
A pMC is simple iff

(1) $P(s,s') \in \{x, 1-x \mid x \text{ parameter}\} \cup \mathbb{Q}$ for all s, s'

(2) $\sum_{s'} P(s,s') = 1$ for all s .

POMDPs and parametric Markov chains

[UAI'18]



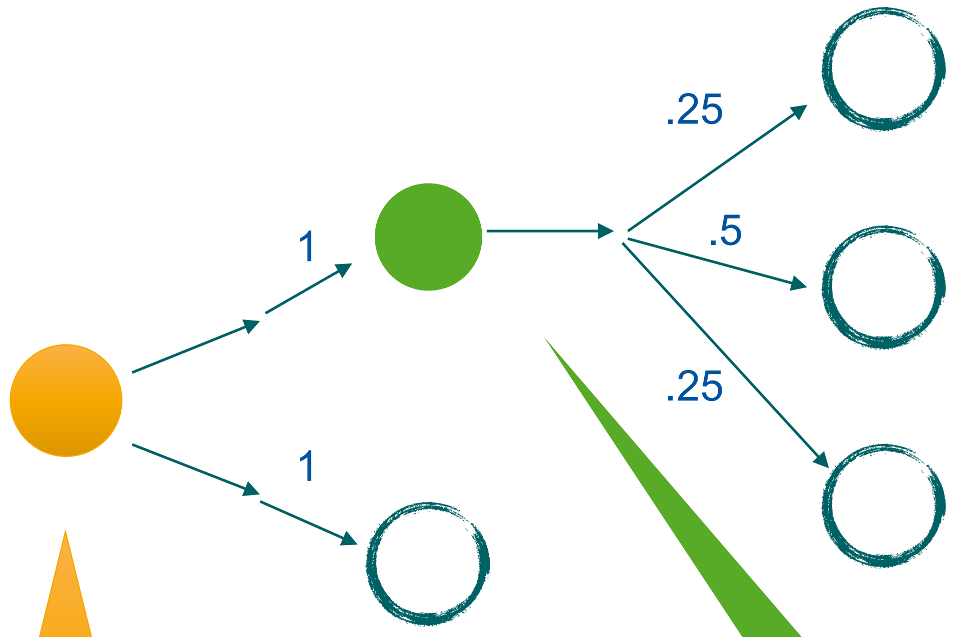
Given any **simple** POMDP
is there a memoryless policy s.t.
the probability reaching ● > λ

Given any **simple** pMC
is there a parameter valuation s.t.
the probability reaching ● > λ

Given **any** POMDP
is there a memoryless policy s.t.
the probability reaching ● > λ

Simple POMDPs

Every state is of either type 1 or type 2



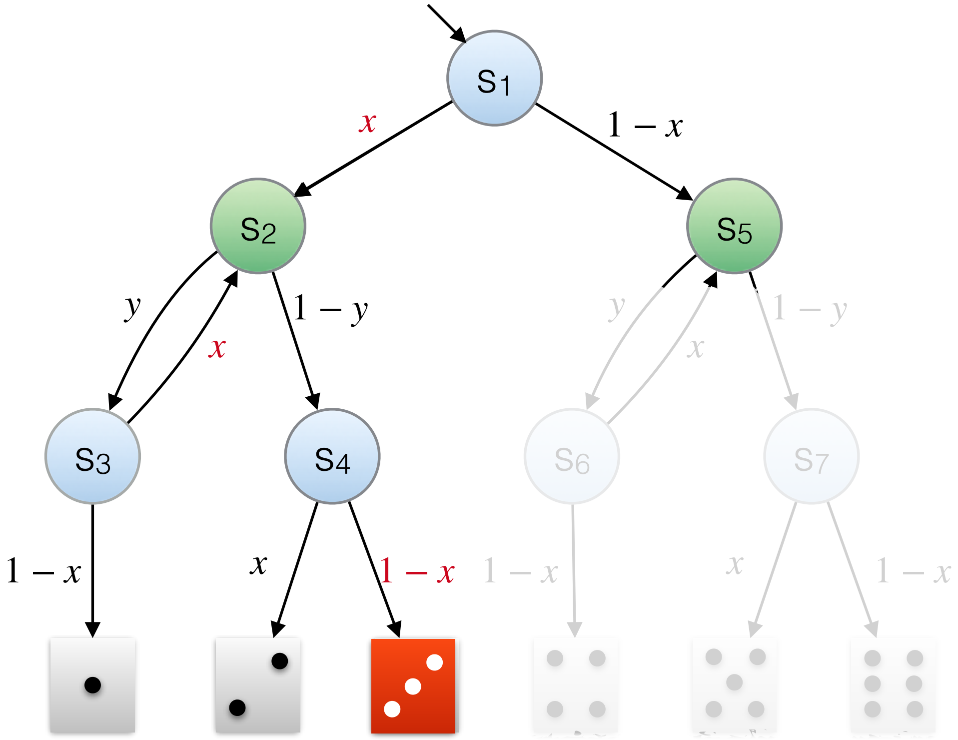
Nondeterministic choice,
Dirac Distributions

Unique choice,
Distribution over
successor states

Encoding feasibility in Existential Theory of the Reals (ETR)

Does a valuation exist s.t. a red state is reached with probability is more than 1/6?

yes, iff the constraints are satisfiable



$$\exists p_i \exists x, y :$$

$$0 < x < 1,\ 0 < y < 1$$

$$p_{\text{⚂}} = 1$$

$$p_5 = 0 \quad p_{\text{⚀}} = 0 \quad p_{\text{⚁}} = 0$$

$$p_4 = x \cdot p_{\text{⚁}} + (1-x) \cdot p_{\text{⚂}}$$

$$p_3 = x \cdot p_2 + (1-x) \cdot p_{\text{⚀}}$$

$$p_2 = y \cdot p_3 + (1-y) \cdot p_4$$

$$p_1 = x \cdot x_2 + (1-x) \cdot p_5$$

$$p_1 > 1/6$$

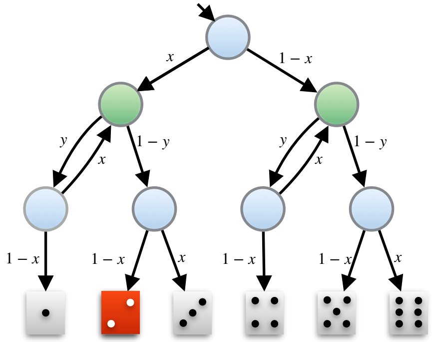
Efficiency?

Solving systems of polynomials — in general — is exponential in number of variables

states + parameters

Eliminating state-variables

Results in a rational function $f(\mathbf{x})$ over the parameters $\mathbf{x}$



$$x \cdot (1-x) \cdot \frac{1-y}{1-xy}$$

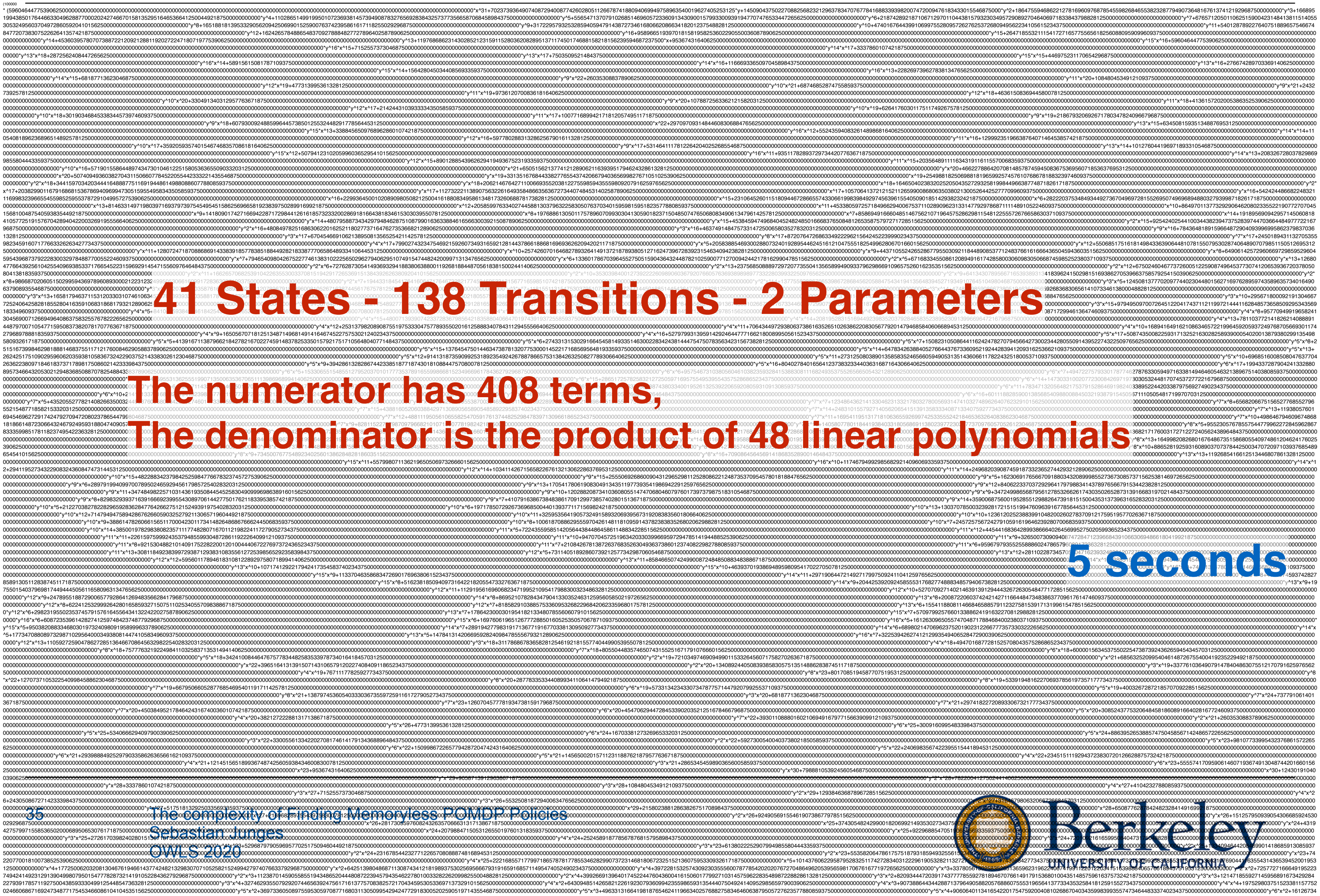
State elimination (as in NFAs) or Gaussian elimination w/ polynomials

[Daws’04]
[Hahn et al.’11]
[Delgado et al.’11]
[Jansen et al.’14]
[CAV’2015]
[Fillieri et al.’17]
[Hutschenreiter et al.’17]
[INFOCOMP’20]

For a pMC with k parameters, n states and linear polynomials as probabilities:

- The rational function can be exponential in k (even for acyclic pMCs)
- For any fixed k, the computation can be done in polynomial time in n

Result of state elimination



41 States - 138 Transitions - 2 Parameters

The numerator has 408 terms,

The denominator is the product of 48 linear polynomials

5 seconds

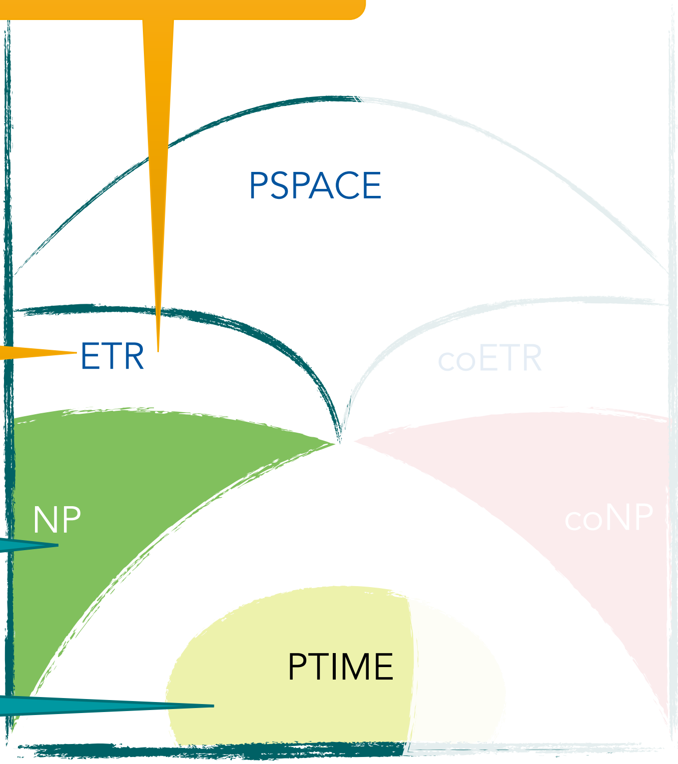
Efficiency?

exponential in parameters

Solving polynomial inequality — in general — is exponential in number of variables

parameters

Recap: Complexity theory



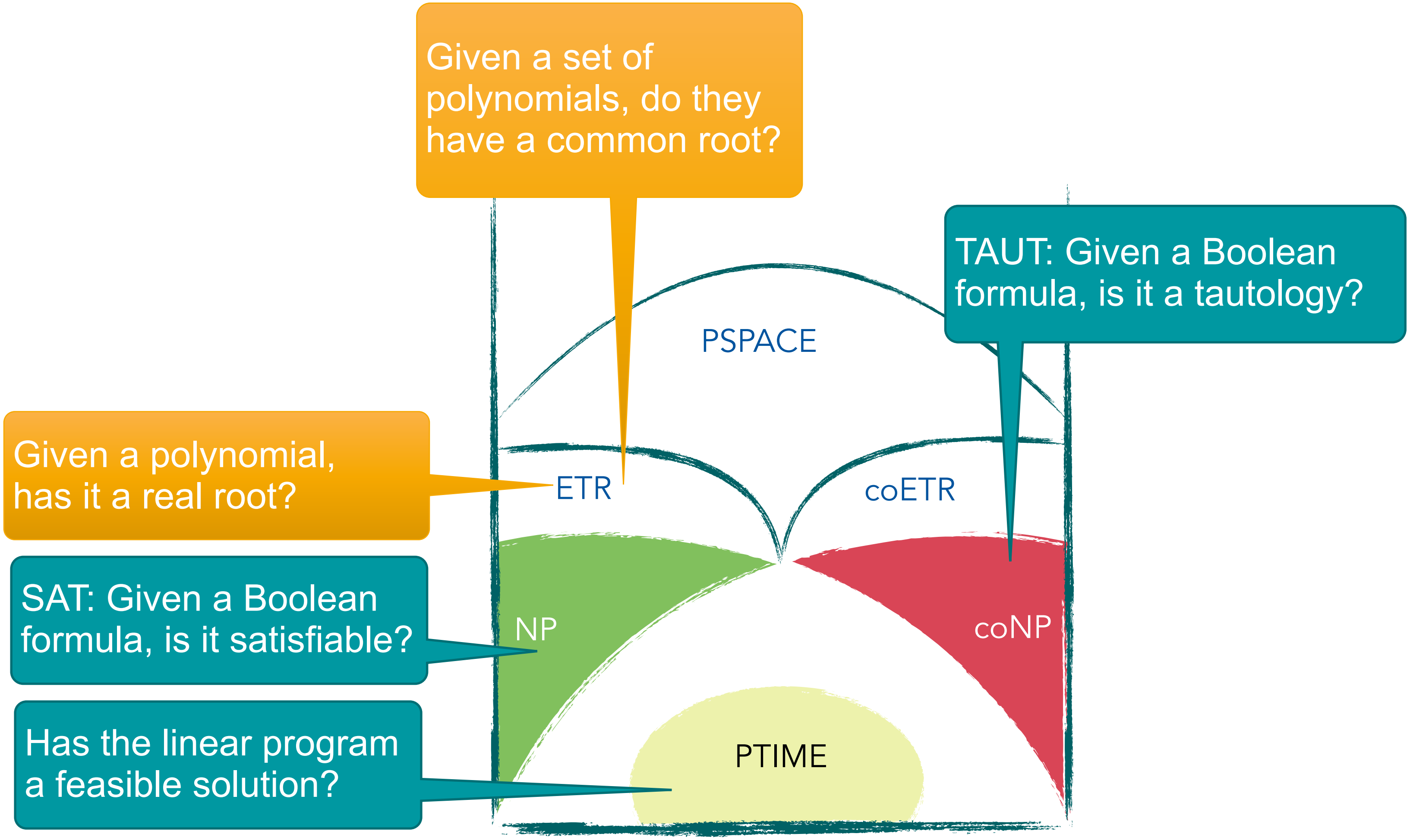
Given a set of polynomials, do they have a common root?

Given a polynomial, has it a real root?

SAT: Given a Boolean formula, is it satisfiable?

Has the LP a feasible solution?

Recap: Complexity theory



Given a set of polynomials, do they have a common root?

TAUT: Given a Boolean formula, is it a tautology?

Given a polynomial, has it a real root?

SAT: Given a Boolean formula, is it satisfiable?

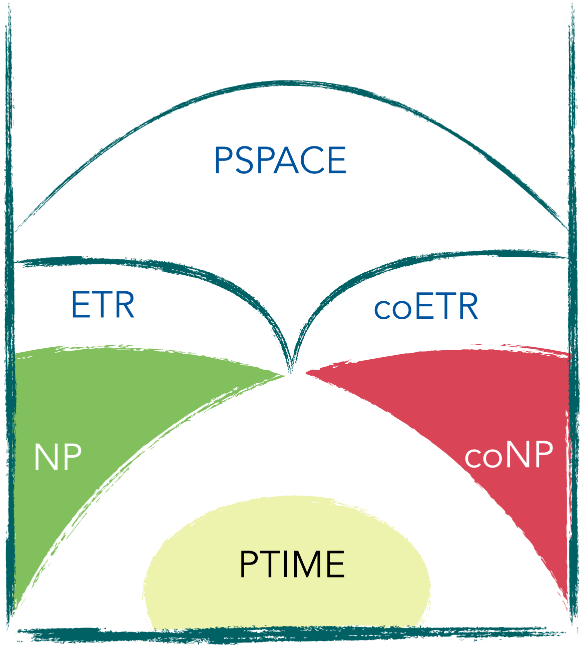
Has the linear program a feasible solution?

How difficult is parameter synthesis?

[CONCUR’19]

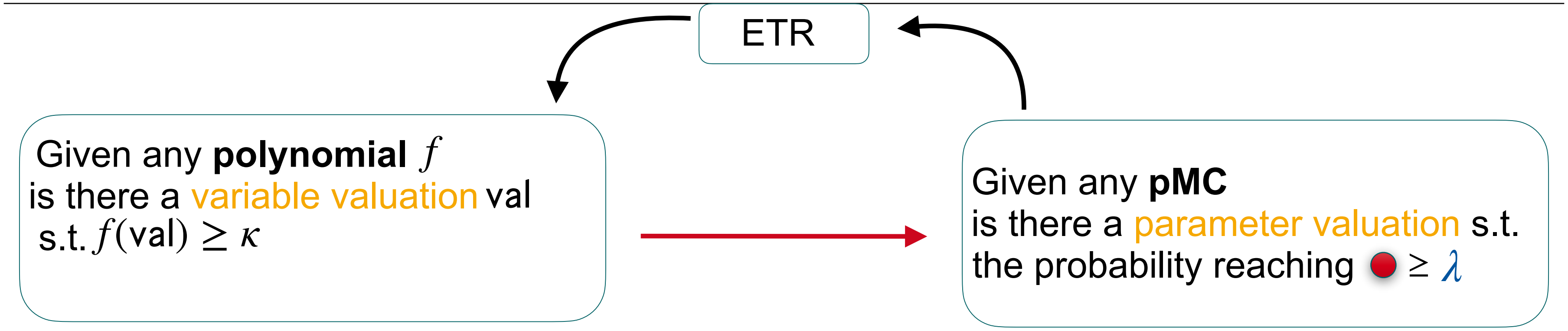
Given: a parametric MC $\mathscr{M}$ with parameters $\mathbf{x}$ **exists:** $\mathsf{val}\colon \mathbf{x} \to [0,1]$ **s.t.:** in $\mathscr{M}[\mathsf{val}]$ a red state is reached with probability [relation] λ

model	relation	
pMC	$\leq$ $\geq$	ETR-complete
	$<$ $>$	NP-hard in ETR



Encoding polynomial inequalities as pMC

[Chonev'19]
[CONCUR'19]



ETR

Given any **polynomial** f
is there a variable valuation val
s.t. $f(\text{val}) \geq \kappa$

Given any **pMC**
is there a parameter valuation s.t.
the probability reaching ● $\geq \lambda$

$$-2x^2y + y \geqslant 5$$

$$\Updownarrow$$

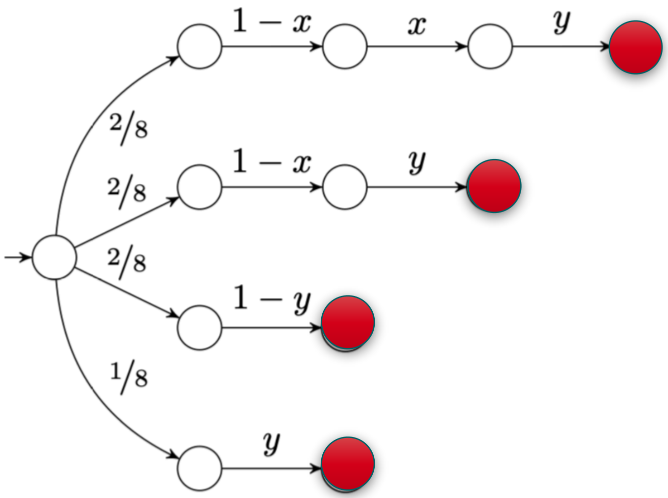
$$2\cdot\big((1-x)xy + (1-x)y + (1-y) - 1\big) + y \geqslant 5$$

$$\Updownarrow$$

$$\frac{2\cdot(1-x)xy + 2\cdot(1-x)y + 2\cdot(1-y) + y}{8} \geqslant \frac{7}{8}$$

$$\Longleftrightarrow$$

Probability of reaching ● at least 7/8



How difficult is parameter synthesis?

[CONCUR'19]

Given: a parametric MC $\mathscr{M}$ with parameters $\mathbf{x}$ **exists:** $\mathsf{val}\colon \mathbf{x} \to [0,1]$ **s.t.:** in $\mathscr{M}[\mathsf{val}]$ a red state is reached with probability [relation] λ

model

a-priori fixed

pMC

in P

utschenreiter et al. 2017]

exponential in parameters

Solving polynomial inequality — in general — is exponential in number of variables

parameters

What about parametric MDPs?

Selecting an action in every state

Given: a parametric MDP $\mathscr{M}$ with parameters $\mathbf{x}$

exists: $\mathsf{val}\colon \mathbf{x} \to [0,1]$ **such that for all** $\sigma\colon S \to \mathsf{Act}$: $\mathscr{M}_\sigma[\mathsf{val}] \vDash \varphi$

How difficult is parameter synthesis?

[CONCUR'19]

The complexity landscape for parameter synthesis (simplified)

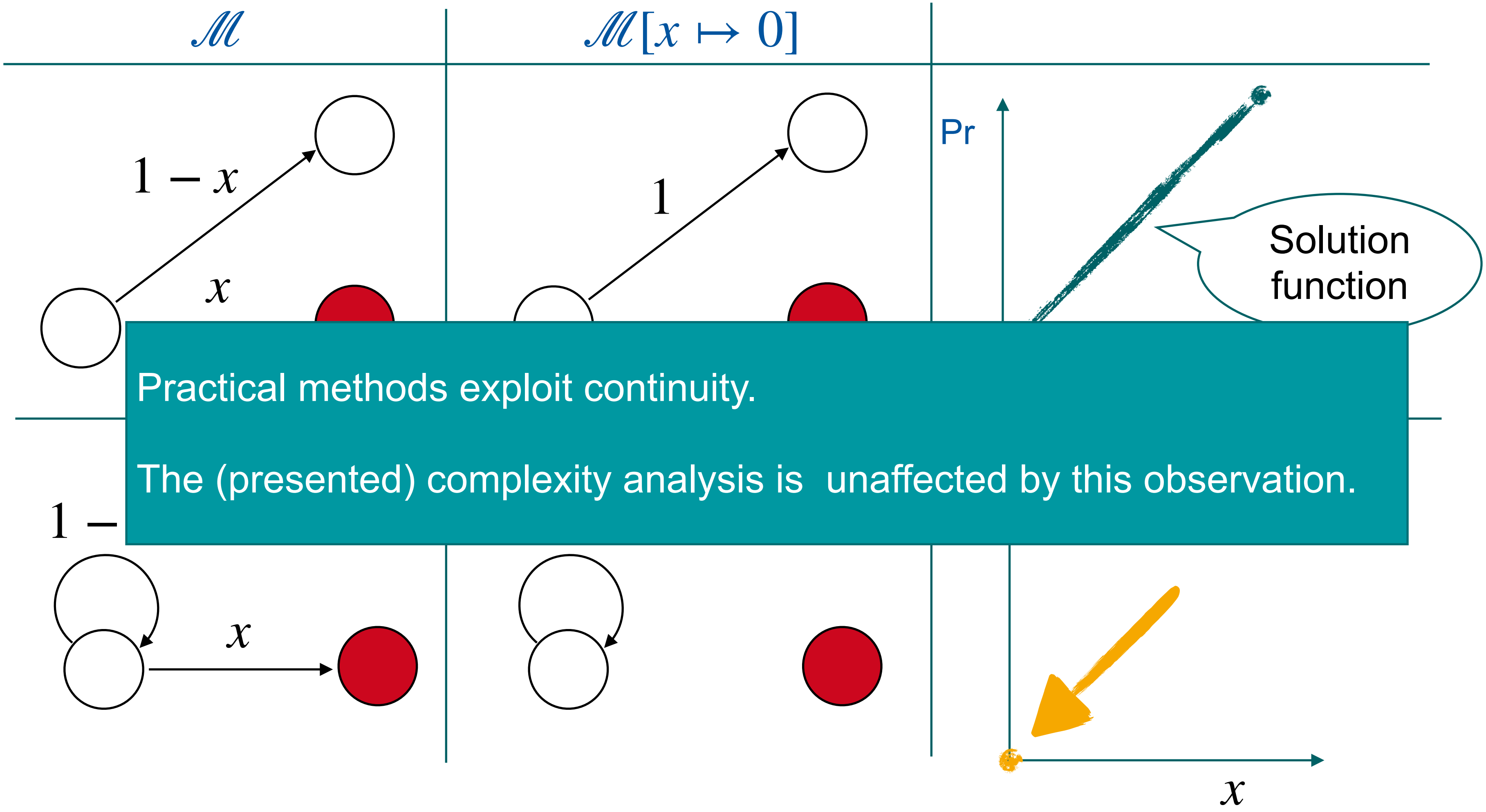
model	parameters relation	arbitrarily many	a-priori fixed
pMC	$\leq$ $\geq$	ETR-complete	in P [Hutschenreiter et al. 2017]
	$<$ $>$	NP-hard in ETR	
pMDP	$<$ $\leq$ $>$ $\geq$	ETR-complete	in NP

ETR encoding as extension of the standard LP for MDPs

How to eliminate state variables?

Graph preservation

$x \mapsto 0$ is **not** graph preserving

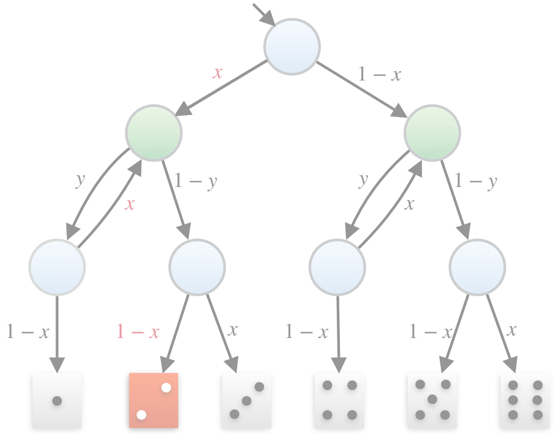


Practical methods exploit continuity.

The (presented) complexity analysis is unaffected by this observation.

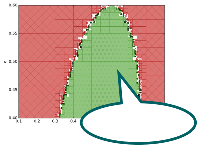
Problem statement: Parameter synthesis

Given:



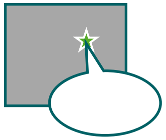
a parametric MDP $\mathscr{M}$
with parameters $\mathbf{x}$

Find:



all/
many

some

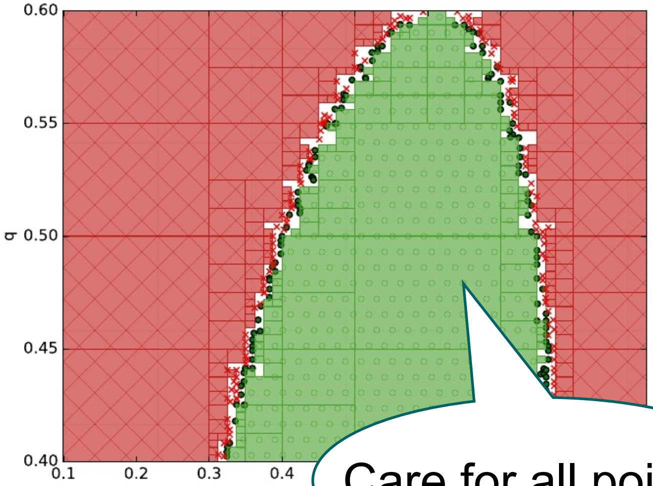


$\text{val} : \mathbf{x} \rightarrow [0,1]$

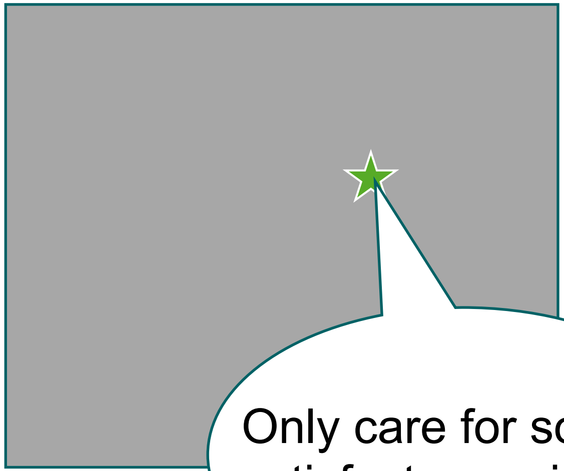
such that: $\mathscr{M}_\sigma[\text{val}] \models \varphi$, i.e., a red state is reached with probability at least/at most λ

Practical parameter synthesis

Two settings

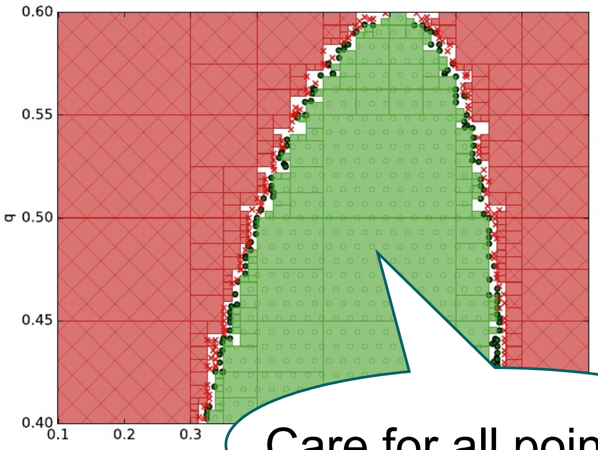


Care for all points

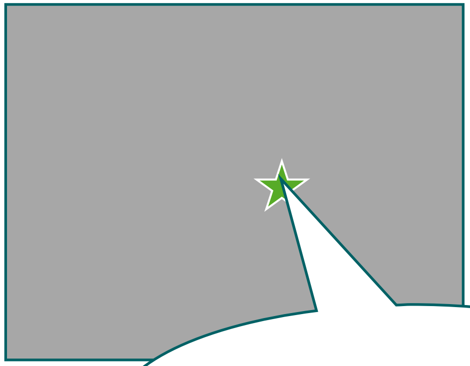


Only care for some satisfactory point

Practical Parameter Synthesis



Care for all points



Only care for some satisfactory point

Several variants of encoding
via SMT so

Parameter
abstraction

Monotonicity [ATVA'19] |

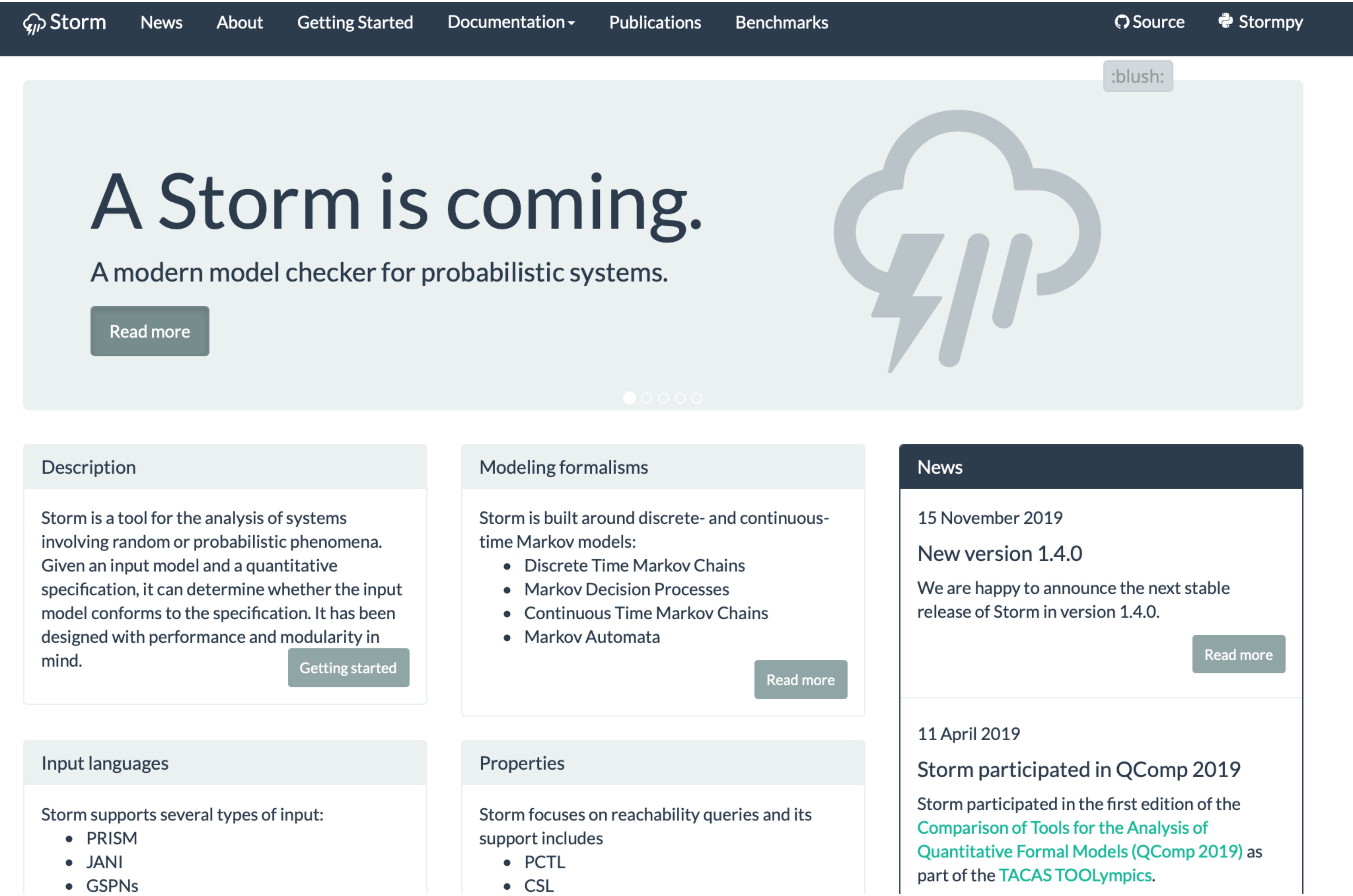
methods
swarm
[Chen et al.'14]

optimisation schemes [TACAS'17]
[ATVA'18]

surveyed in [arXiv'19]

Methods assume and exploit (to some extent) that **the graph structure is fixed.**

Storm (www.stormchecker.org)



A Storm is coming.

A modern model checker for probabilistic systems.

Description

Storm is a tool for the analysis of systems involving random or probabilistic phenomena. Given an input model and a quantitative specification, it can determine whether the input model conforms to the specification. It has been designed with performance and modularity in mind.

Modeling formalisms

Storm is built around discrete- and continuous-time Markov models:

- Discrete Time Markov Chains
- Markov Decision Processes
- Continuous Time Markov Chains
- Markov Automata

News

15 November 2019

New version 1.4.0

We are happy to announce the next stable release of Storm in version 1.4.0.

11 April 2019

Storm participated in QComp 2019

Storm participated in the first edition of the Comparison of Tools for the Analysis of Quantitative Formal Models (QComp 2019) as part of the TACAS TOOLympics.

Input languages

Storm supports several types of input:

- PRISM
- JANI
- GSPNs

Properties

Storm focuses on reachability queries and its support includes

- PCTL
- CSL

Related work … necessarily incomplete here.

Parameter Synthesis

Infinite state systems: e.g., Chakarov et al., Esparza et al., Kaminski et al., Zuck et al., etc.

Various Applications: e.g., Aflaki et al., Calinescu et al., Fillieri et al., Polgreen et al., Rosenblum et al.

Modal transition systems: e.g., Benes et al., Delahaye et al.

POMDPs with small strategies: e.g., Chatterjee et al., Amato et al.

Interval/Constraint MDPs: e.g., Delahaye et al, Chatterjee et al., Chen et al., Hahn et al., Larsen et al.

Quantitative Verification of Software Product Lines, e.g., Ghezzi et al, Terbeek et al,

parametric Continuous-Time MCs: e.g., Ceska et al., Han et al.

Future challenges

The complexity of feasibility in pMDPs with one parameter

Robust strategies instead of (parameter) feasibility

Parameter Synthesis

feasibility = **exists:** $\mathsf{val} : \mathbf{x} \to [0,1]$ **such that for all** $\sigma : S \to \mathsf{Act}$

robust strategies = **exists:** $\sigma : S \to \mathsf{Act}$ **such that for all** $\mathsf{val} : \mathbf{x} \to [0,1]$

New challenges for verification:
Expensive (but powerful) abstraction techniques &
Symbolic probabilistic model checking

Want to know more?

sjunges@berkeley.edu

For a formal treatment:

Sebastian Junges, Nils Jansen, Ralf Wimmer, Tim Quatmann, Leonore Winterer, Joost-Pieter Katoen, Bernd Becker:
Finite-State Controllers of POMDPs using Parameter Synthesis. UAI 2018: 519-529

Sebastian Junges, Joost-Pieter Katoen, Guillermo A. Pérez, Tobias Winkler:
The Complexity of Reachability in Parametric Markov Decision Processes. CoRR abs/2009.13128 (2020)