Parameter Synthesis in Markov Chains: The Why and How?

Sebastian Junges

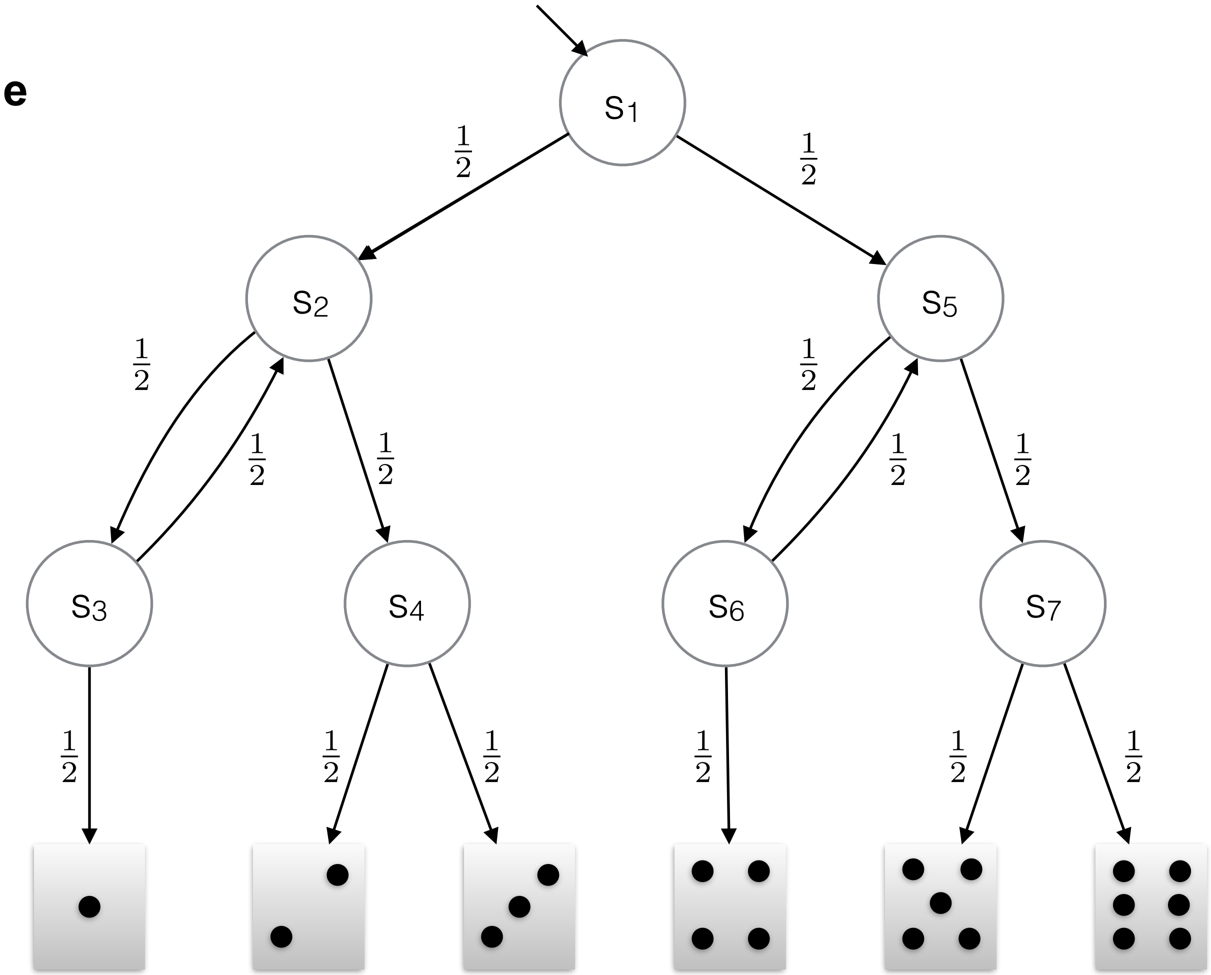
Last Lecture Modelling and Verification of Probabilistic Systems

Including work with and by:
Erika Abraham, Christel Baier, Bernd Becker, Harold Bruintjes, Florian Corzilius, Murat Cubuktepe, Christian Dehnert, Nils Jansen, Joost-Pieter Katoen, Joachim Klein, Lisa Hutschenreiter, Ufuk Topcu, Tim Quatmann, Matthias Volk, Leonore Winterer, Ralf Wimmer

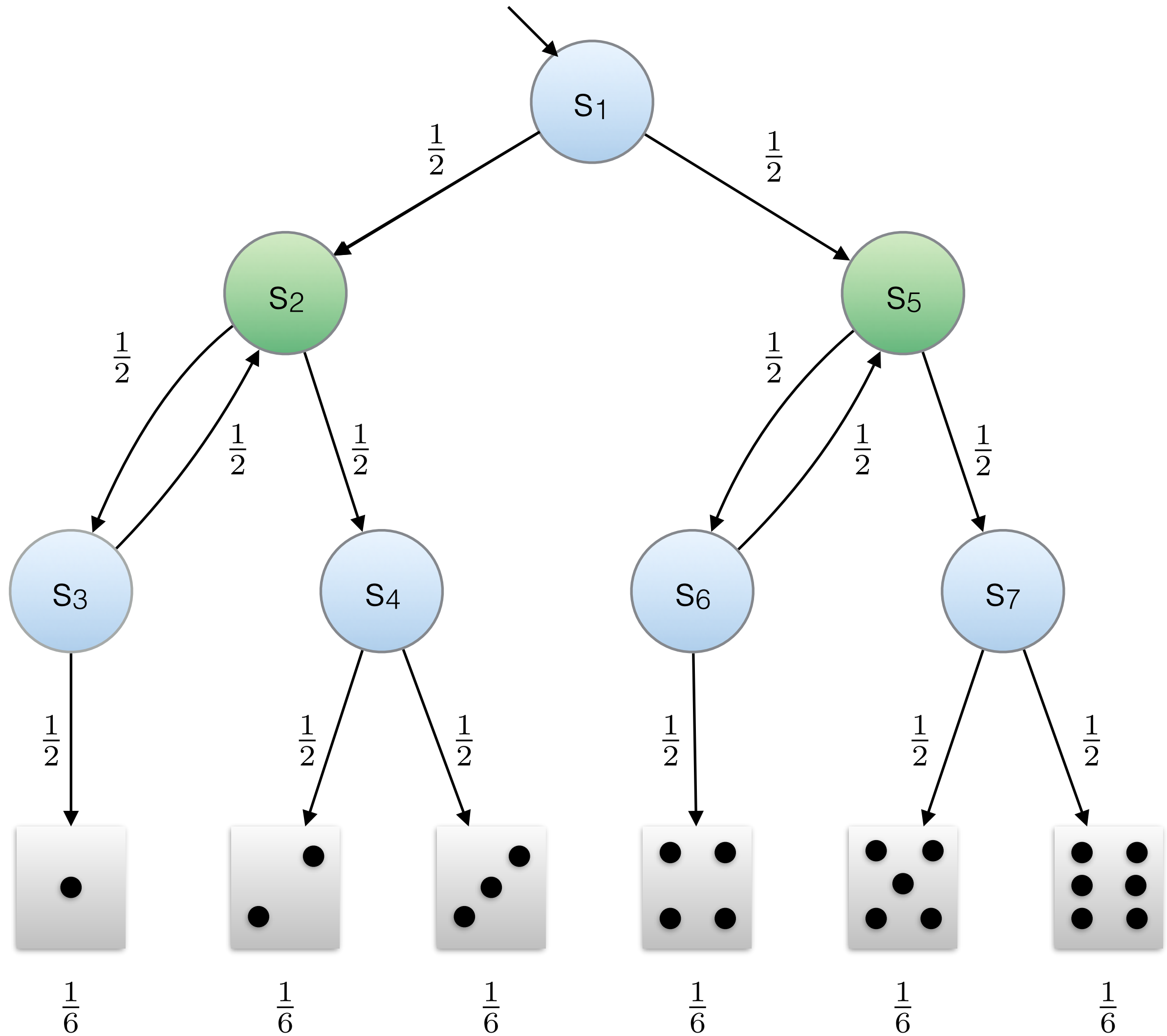
UnRAVeL Research Training Group – Uncertainty and Randomness in Algorithms, Verification, and Logic | RWTH AACHEN UNIVERSITY

parametric Markov Chains (pMCs)

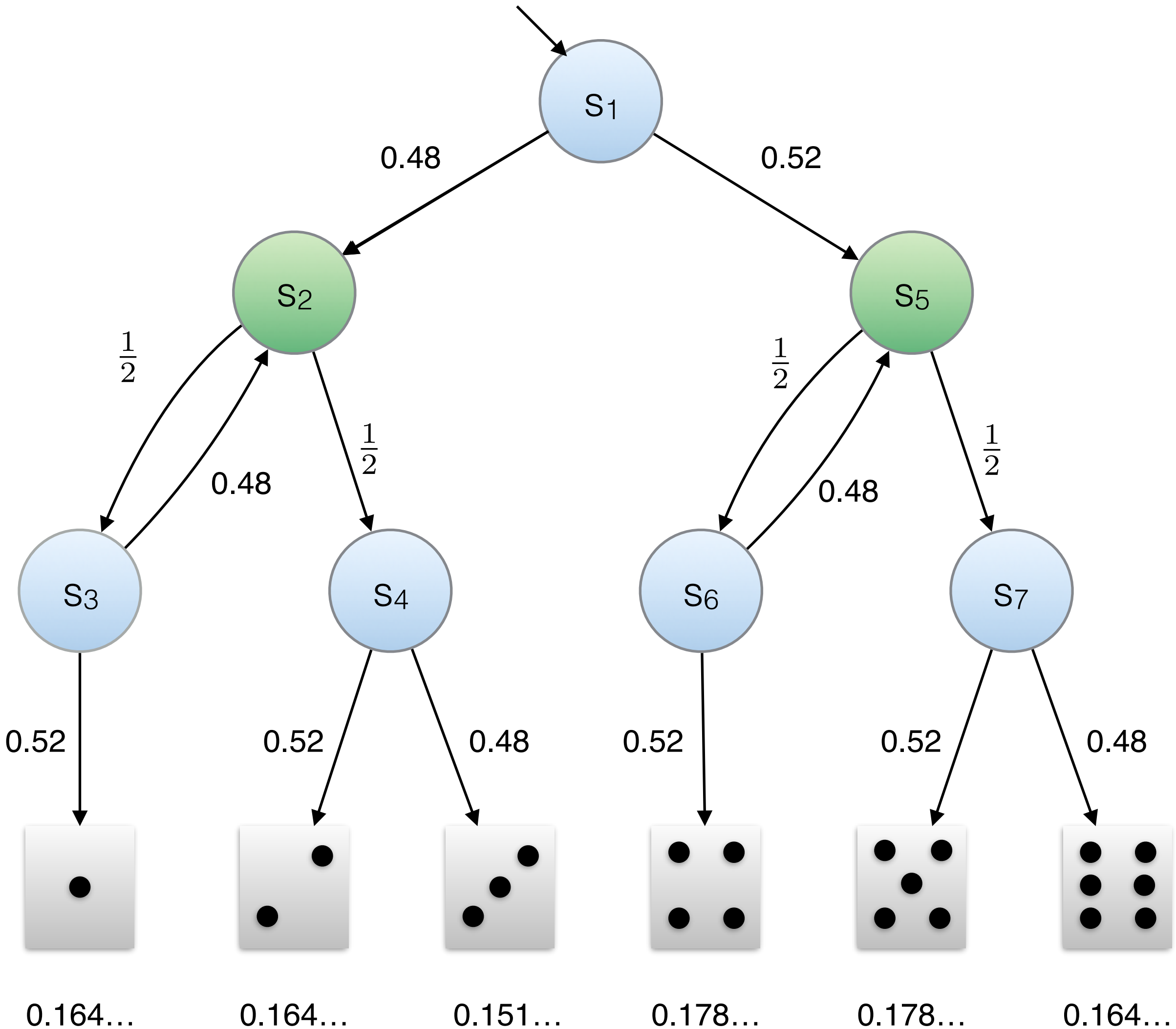
Knuth-Yao Die



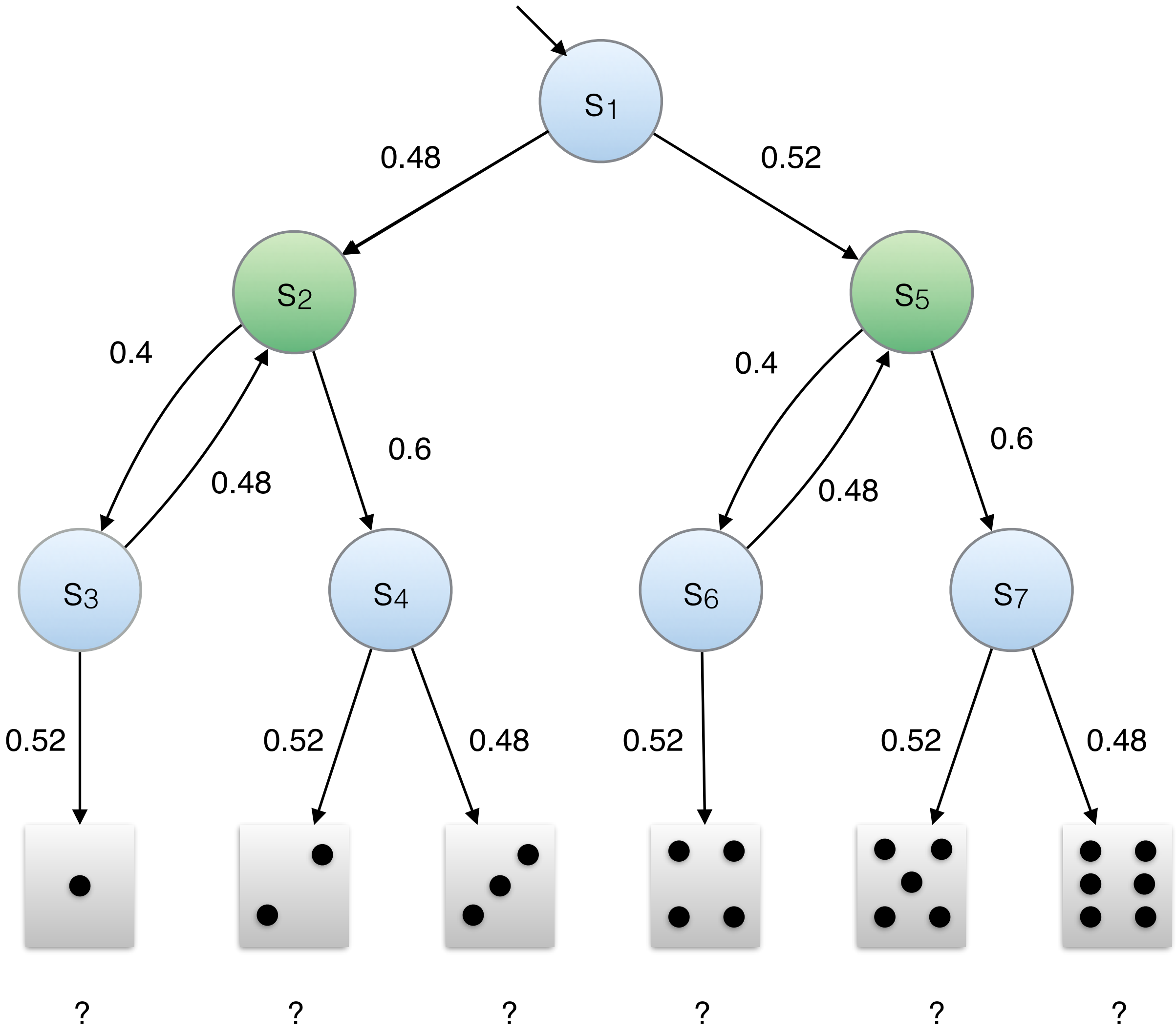
Knuth-Yao Die



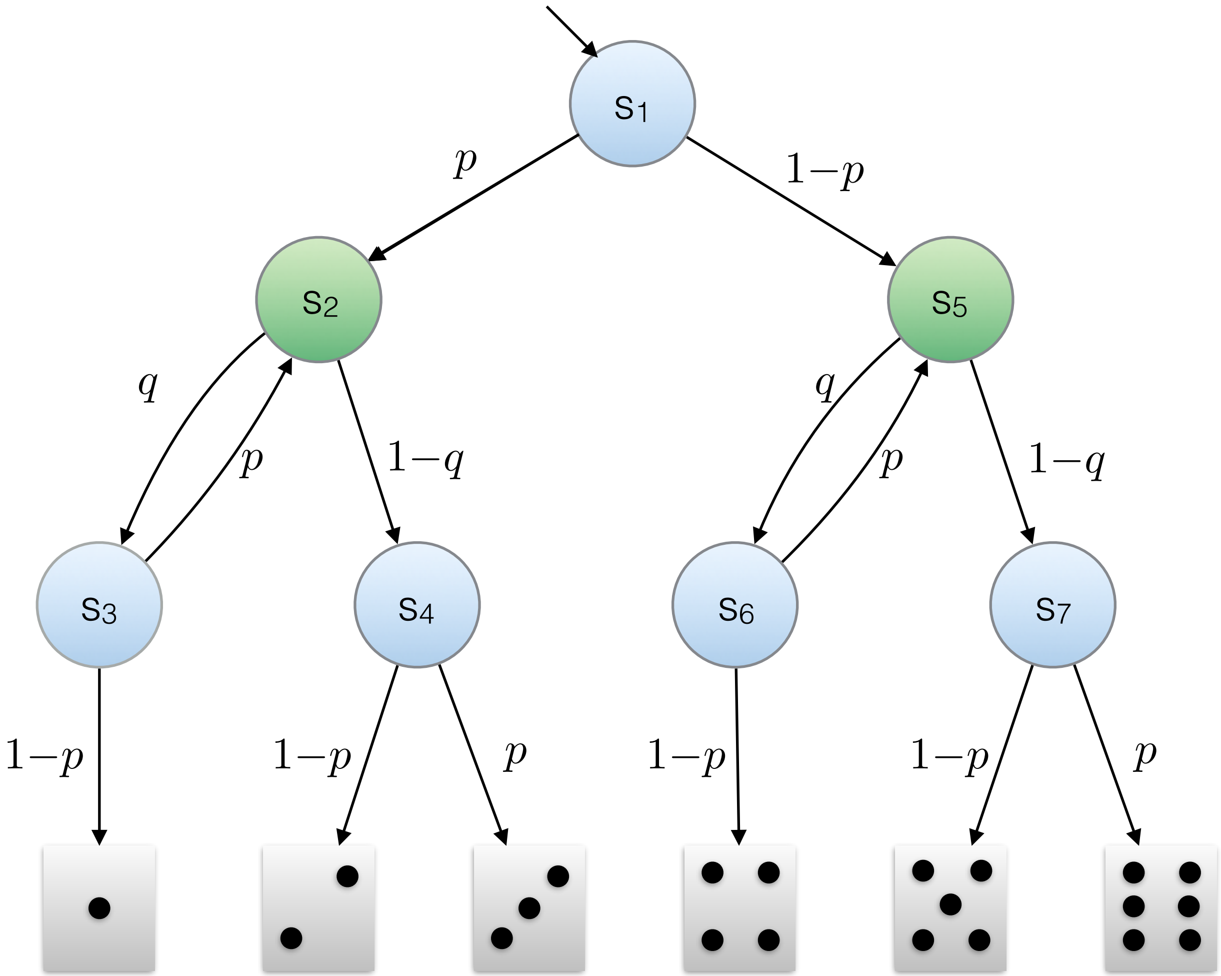
Knuth-Yao Die



Knuth-Yao Die



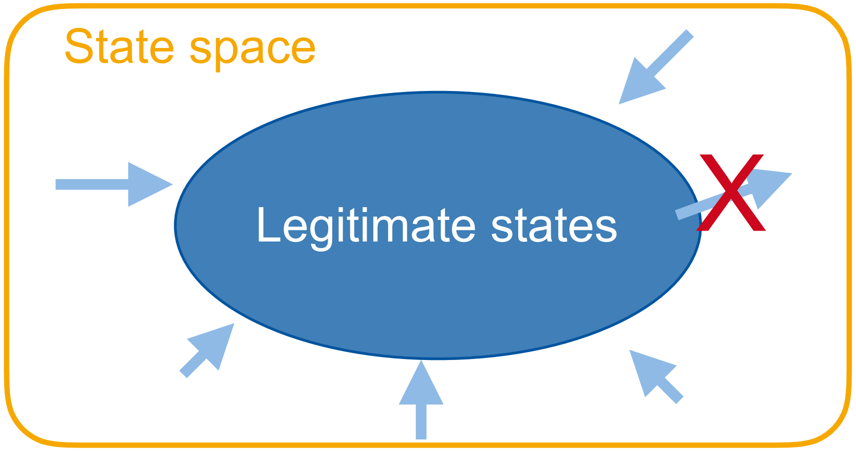
Knuth-Yao Die



Self-stabilizing algorithms

A distributed system is self-stabilizing iff it satisfies:

1. **Convergence (recovery):** Starting from any arbitrary state, it always reaches a legitimate state in a finite number of steps, and
2. **Closure:** It remains in legitimate states in the absence of faults

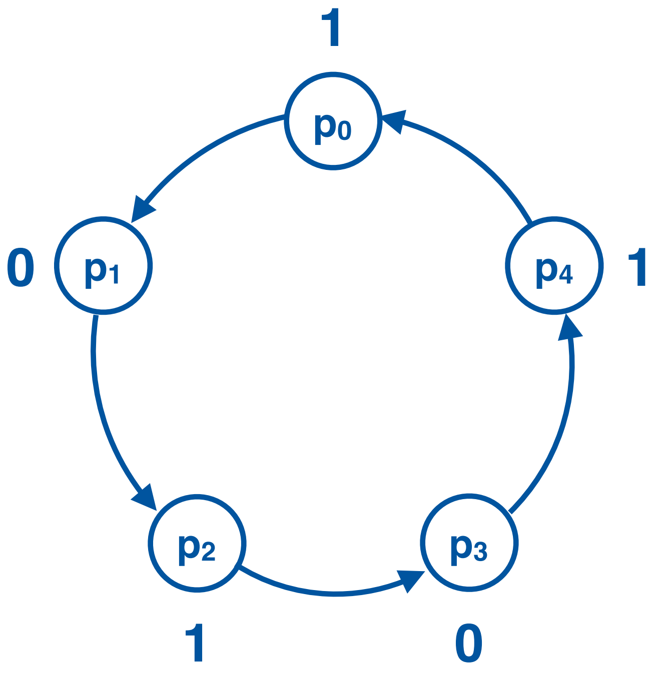


Herman's randomized self-stabilizing algorithm

Token in x_i, if $x_i = x_{i-1}$

$$x_i \neq x_{i-1} \rightarrow x_i := x_{i-1}$$

$$x_i = x_{i-1} \rightarrow \begin{cases} p: & x_i := 0 \\ 1-p: & x_i := 1 \end{cases}$$

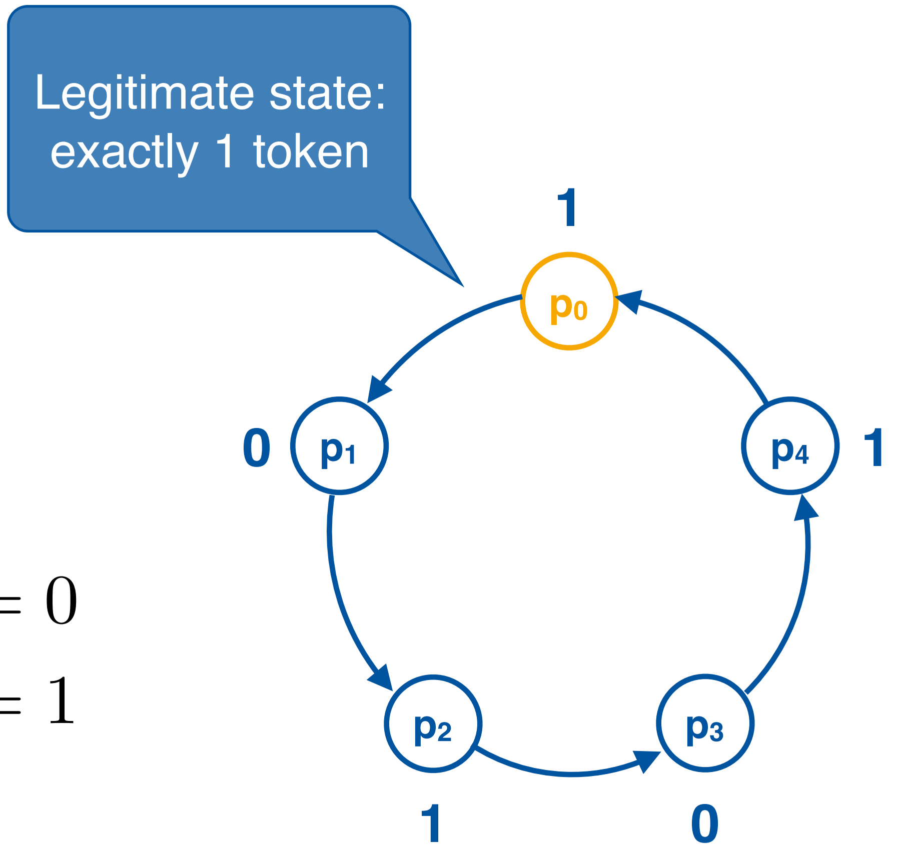


Herman's randomized self-stabilizing algorithm

Token in x_i, if $x_i = x_{i-1}$

$$x_i \neq x_{i-1} \rightarrow x_i := x_{i-1}$$

$$x_i = x_{i-1} \rightarrow \begin{cases} p: & x_i := 0 \\ 1-p: & x_i := 1 \end{cases}$$

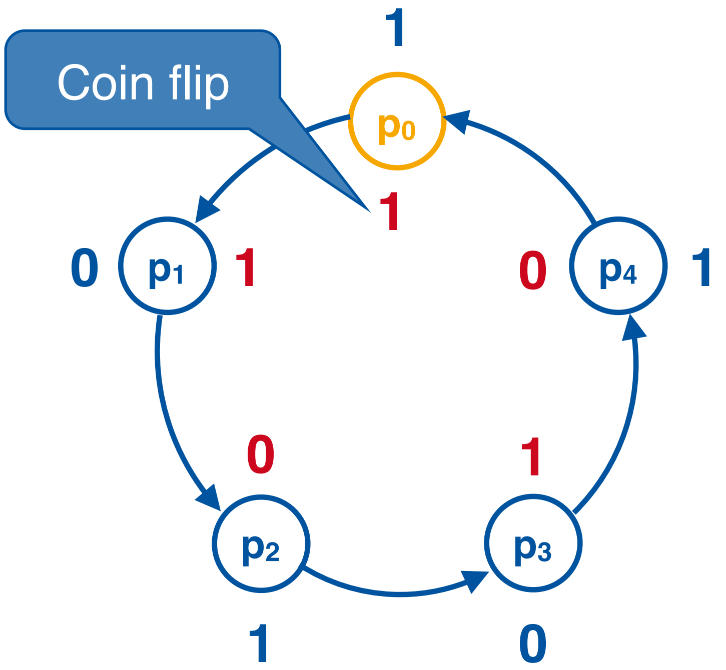


Herman's algorithm

Token in x_i, if $x_i = x_{i-1}$

$$x_i \neq x_{i-1} \rightarrow x_i := x_{i-1}$$

$$x_i = x_{i-1} \rightarrow \begin{cases} p: & x_i := 0 \\ 1-p: & x_i := 1 \end{cases}$$

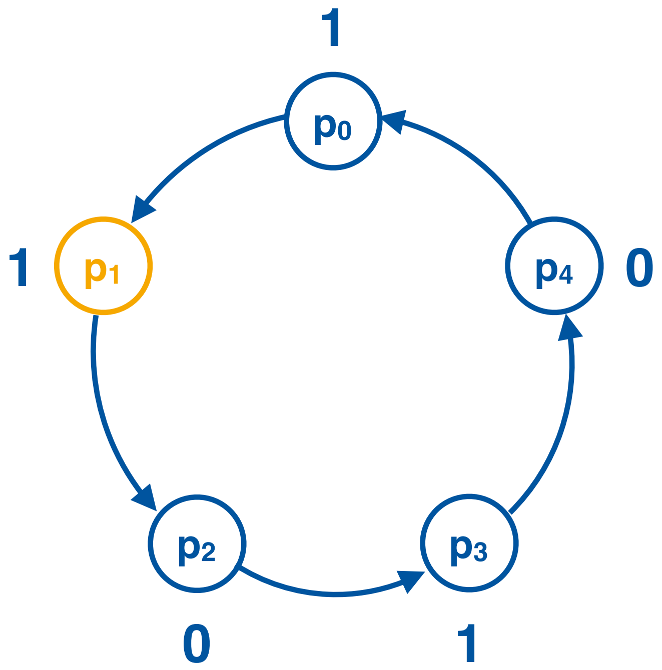


Herman's algorithm

$$\text{Token in } x_i, \text{ if } x_i = x_{i-1}$$

$$x_i \neq x_{i-1} \rightarrow x_i := x_{i-1}$$

$$x_i = x_{i-1} \rightarrow \begin{cases} p: & x_i := 0 \\ 1-p: & x_i := 1 \end{cases}$$

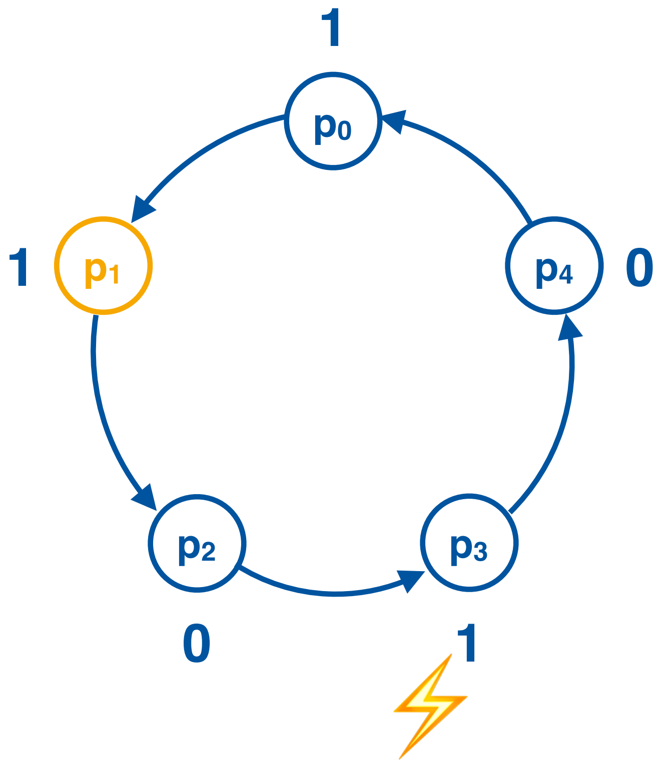


Recovery

Token in x_i, if $x_i = x_{i-1}$

$$x_i \neq x_{i-1} \rightarrow x_i := x_{i-1}$$

$$x_i = x_{i-1} \rightarrow \begin{cases} p: & x_i := 0 \\ 1-p: & x_i := 1 \end{cases}$$

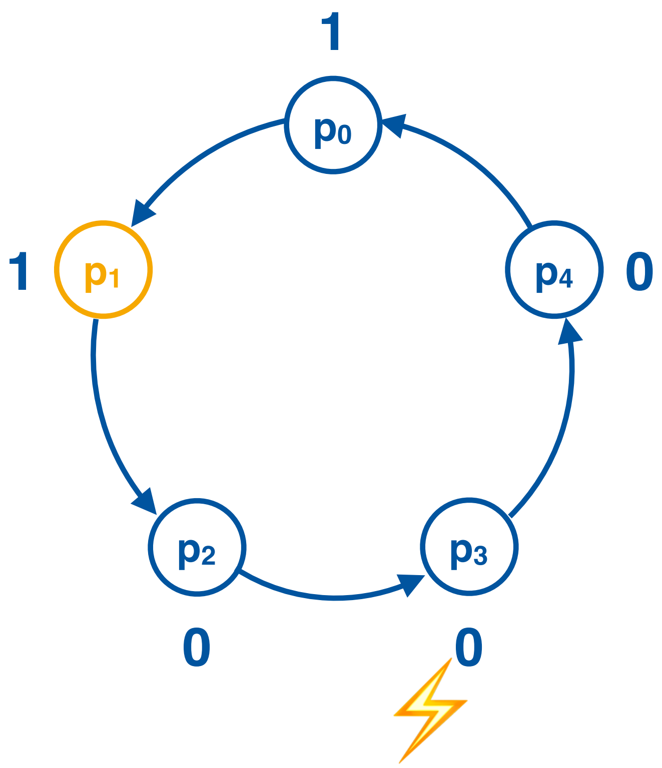


Recovery

Token in x_i, if $x_i = x_{i-1}$

$$x_i \neq x_{i-1} \rightarrow x_i := x_{i-1}$$

$$x_i = x_{i-1} \rightarrow \begin{cases} p: & x_i := 0 \\ 1-p: & x_i := 1 \end{cases}$$

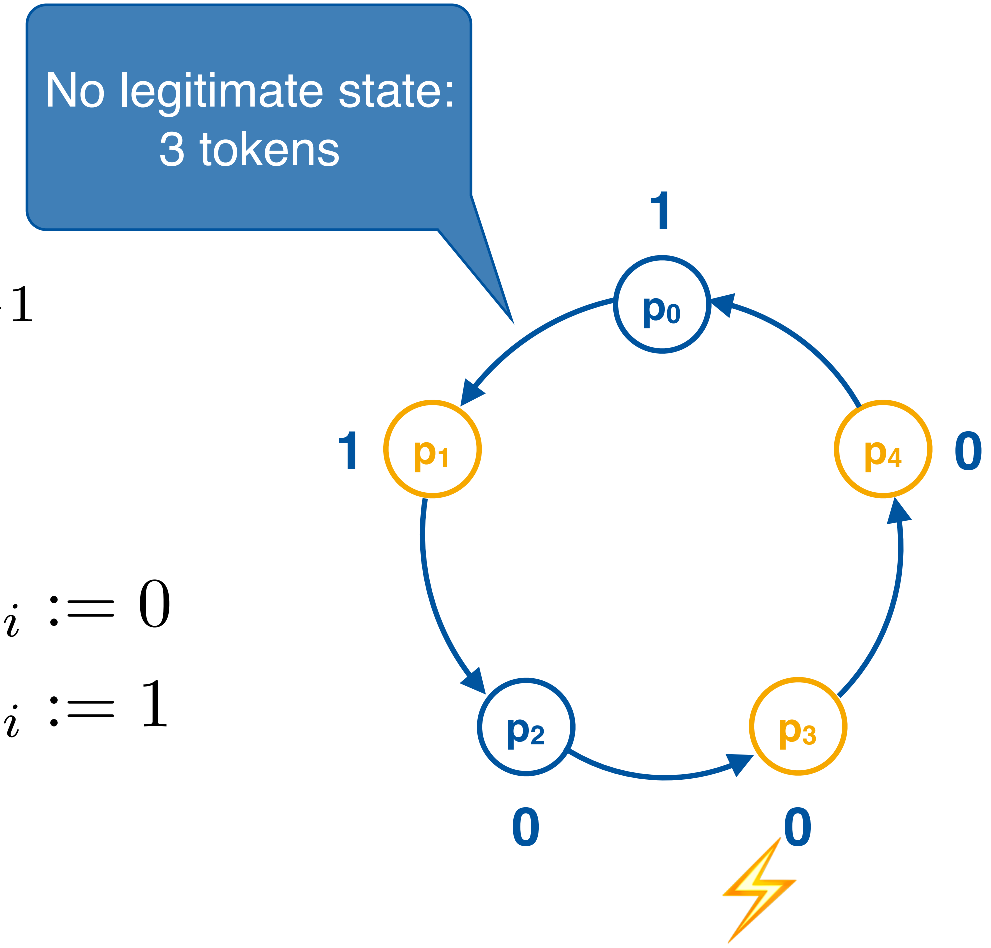


Recovery

Token in x_i, if $x_i = x_{i-1}$

$$x_i \neq x_{i-1} \rightarrow x_i := x_{i-1}$$

$$x_i = x_{i-1} \rightarrow \begin{cases} p: & x_i := 0 \\ 1-p: & x_i := 1 \end{cases}$$

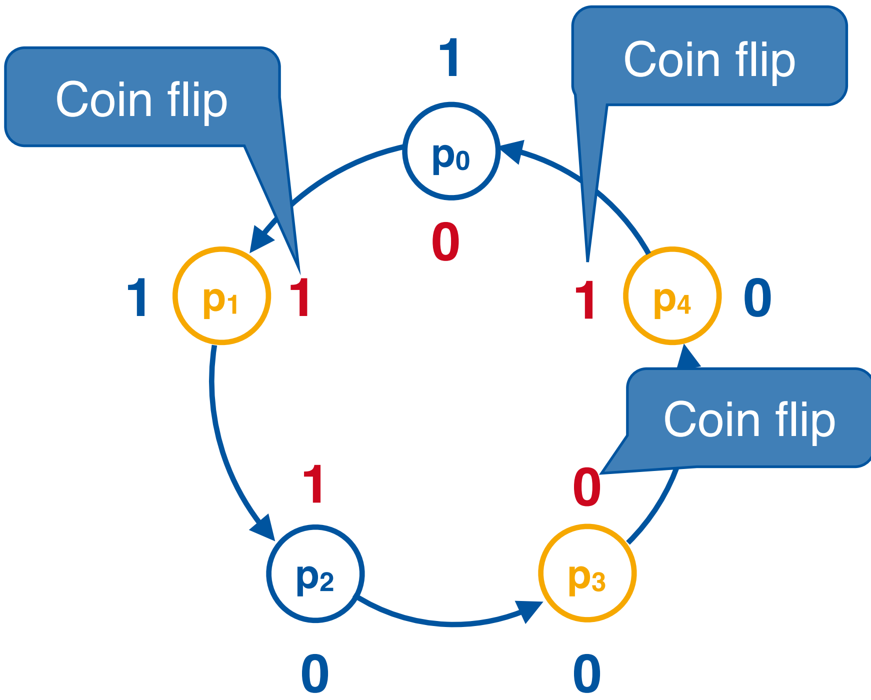


Recovery

Token in x_i, if $x_i = x_{i-1}$

$$x_i \neq x_{i-1} \rightarrow x_i := x_{i-1}$$

$$x_i = x_{i-1} \rightarrow \begin{cases} p: & x_i := 0 \\ 1-p: & x_i := 1 \end{cases}$$



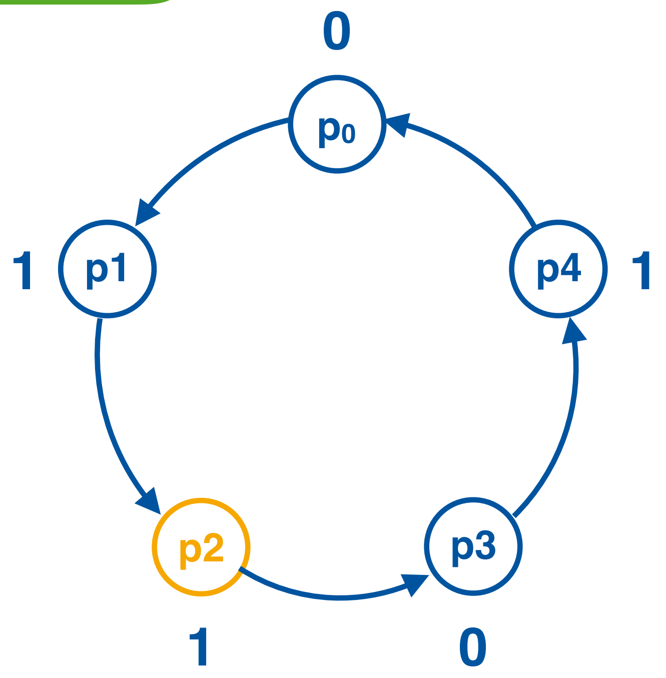
Recovery

Self-Stabilizing

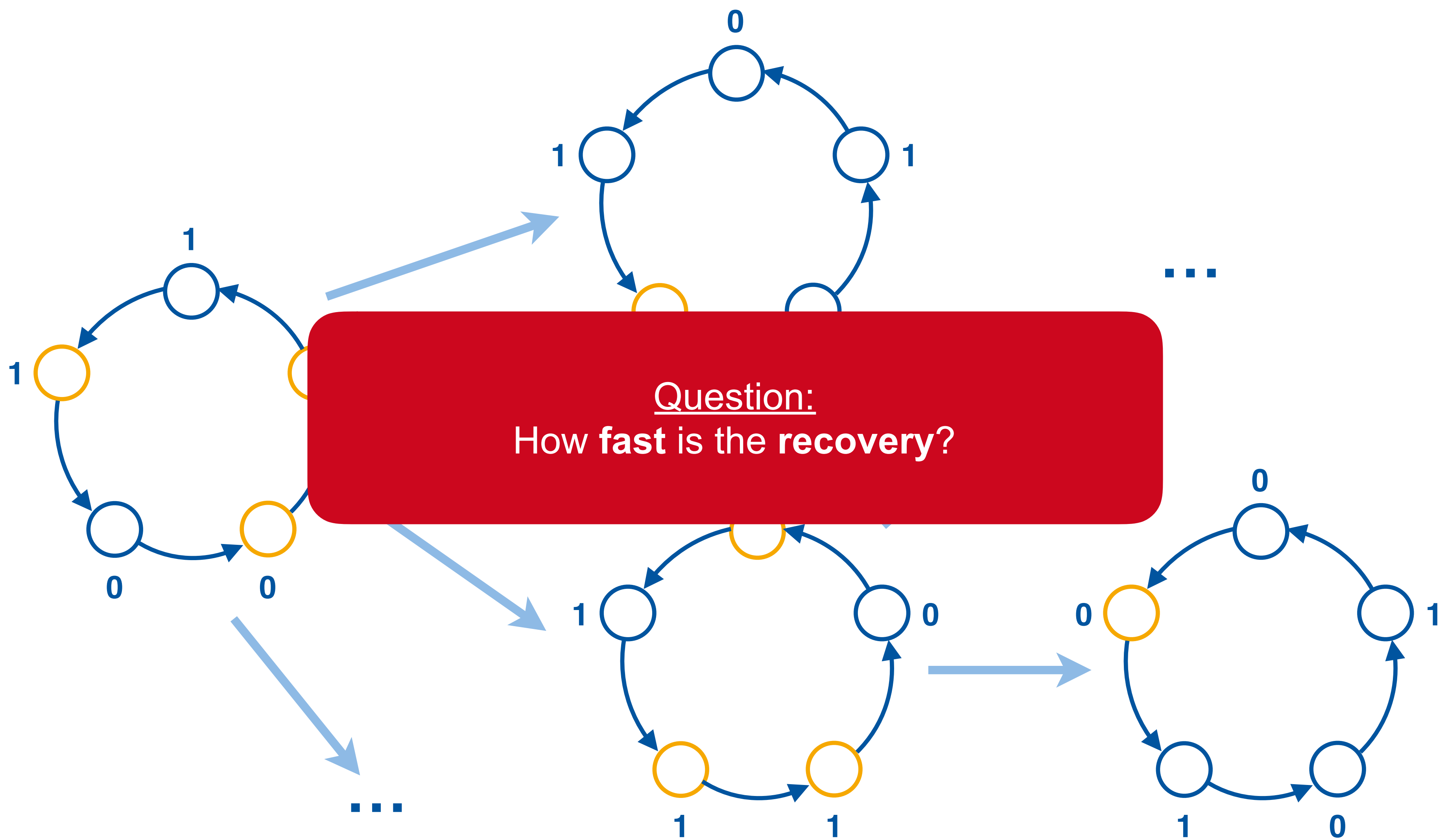
Token in x_i, if $x_i = x_{i-1}$

$$x_i \neq x_{i-1} \rightarrow x_i := x_{i-1}$$

$$x_i = x_{i-1} \rightarrow \begin{cases} p: & x_i := 0 \\ 1-p: & x_i := 1 \end{cases}$$



Recovery



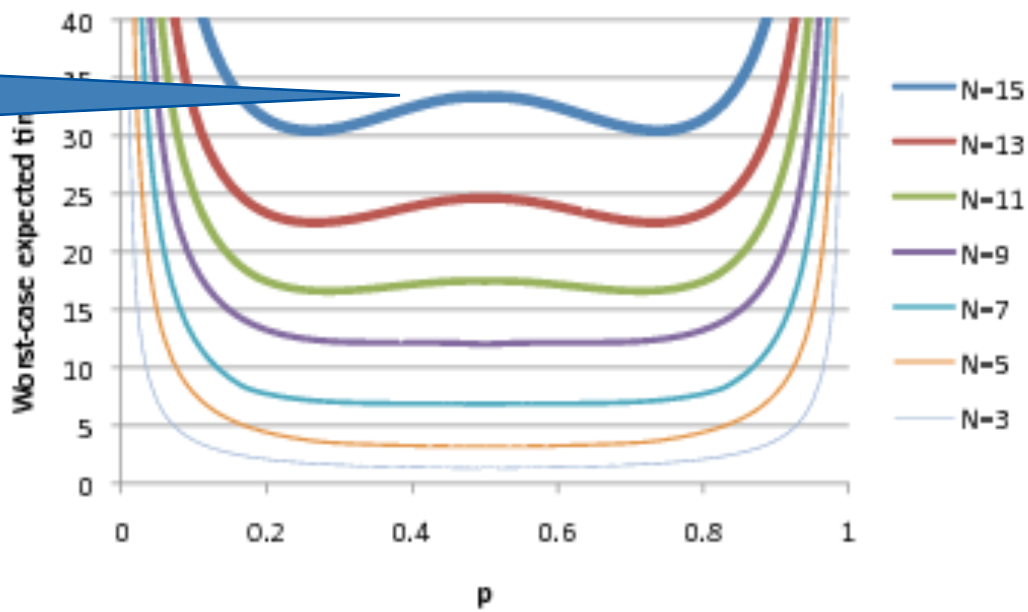
Question:
How **fast** is the **recovery**?

Research questions

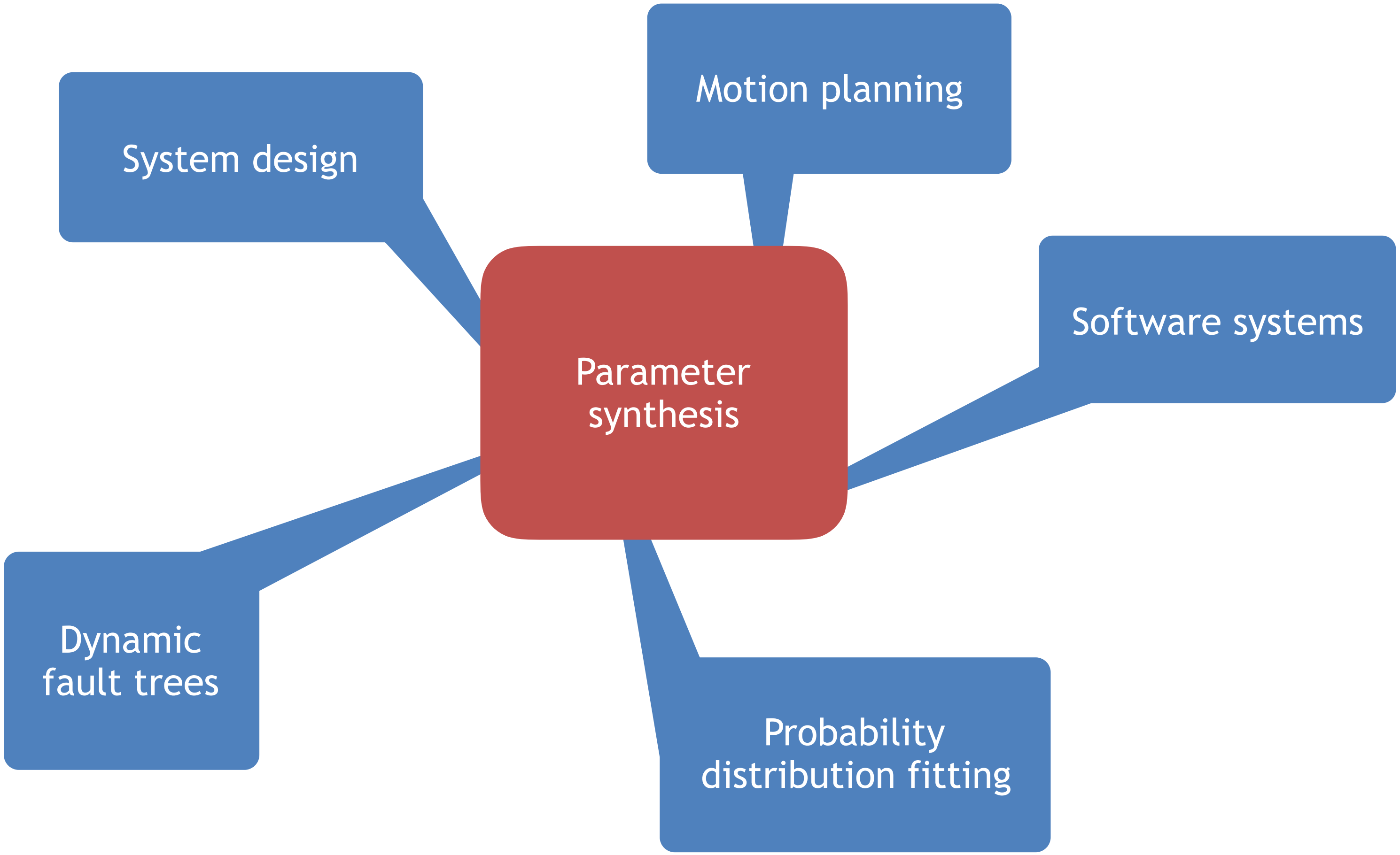
- How fast do these algorithms converge?
  - worst-case recovery time
  - average recovery time
- How does the randomization influence the recovery time?

use pMCs
[Aflaki et al. - 2017]

Fair coin not optimal



Other applications



What is a parametric Markov chain?

A parametric Markov chain is a DTMC over a set of parameters (or variables).

$$\mathcal{D} = (S, \mathsf{Var}, \mathbf{P}, s_\iota, \mathsf{AP}, L)$$

S (finite) set of states

s_ι initial state

AP, L atomic propositions & labelling

Var finite set of parameters or variables

$\mathbf{P} : S \times S \to \mathbb{Q}(\mathsf{Var})$ transition probabilities

Rational Functions

instantiation: $\mathsf{Var} \to \mathbb{R}$

Whiteboard: Well-defined, Graph-preserving, Simple

What is Parameter Synthesis?

in the context of probabilistic model checking

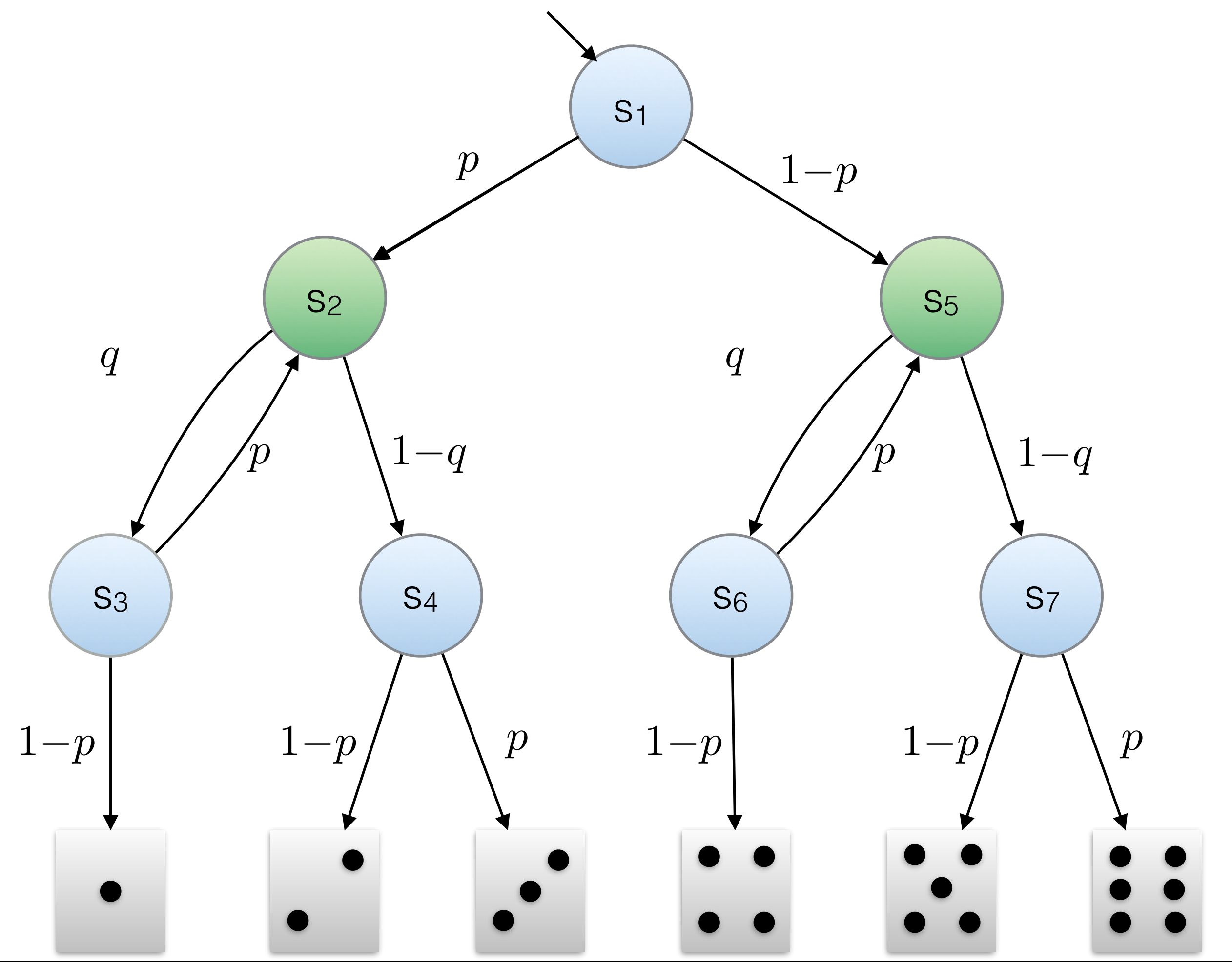
- parametric Markov chain (i.e. a family of induced Markov chains)
- reachability property

Check the property for every induced Markov chain in the **family**:
Learn something about the effect of the parameters

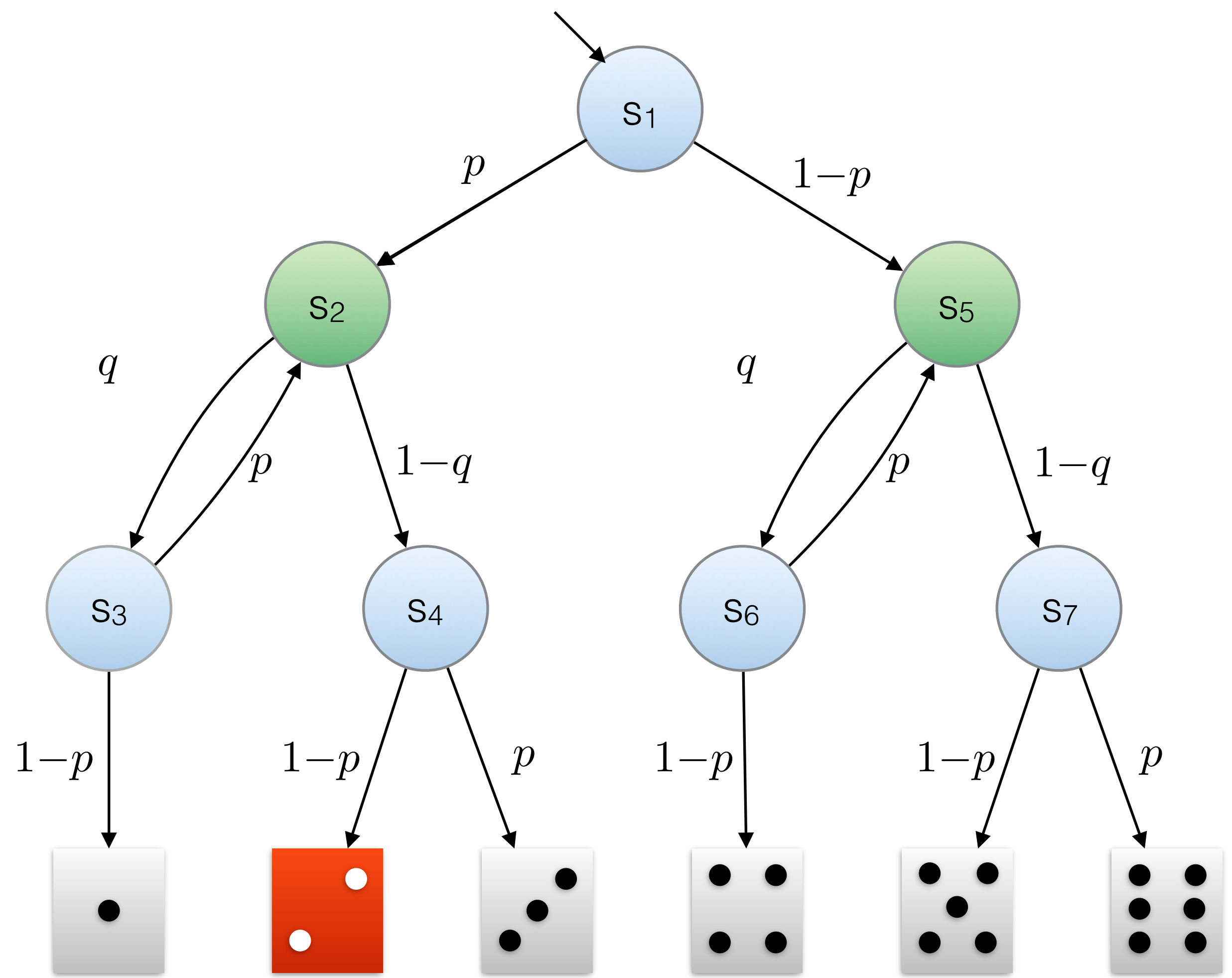
not touched today

- parametric MDPs and robust strategies
- richer properties
- connections to partially observable MDPs (prominent model in AI)

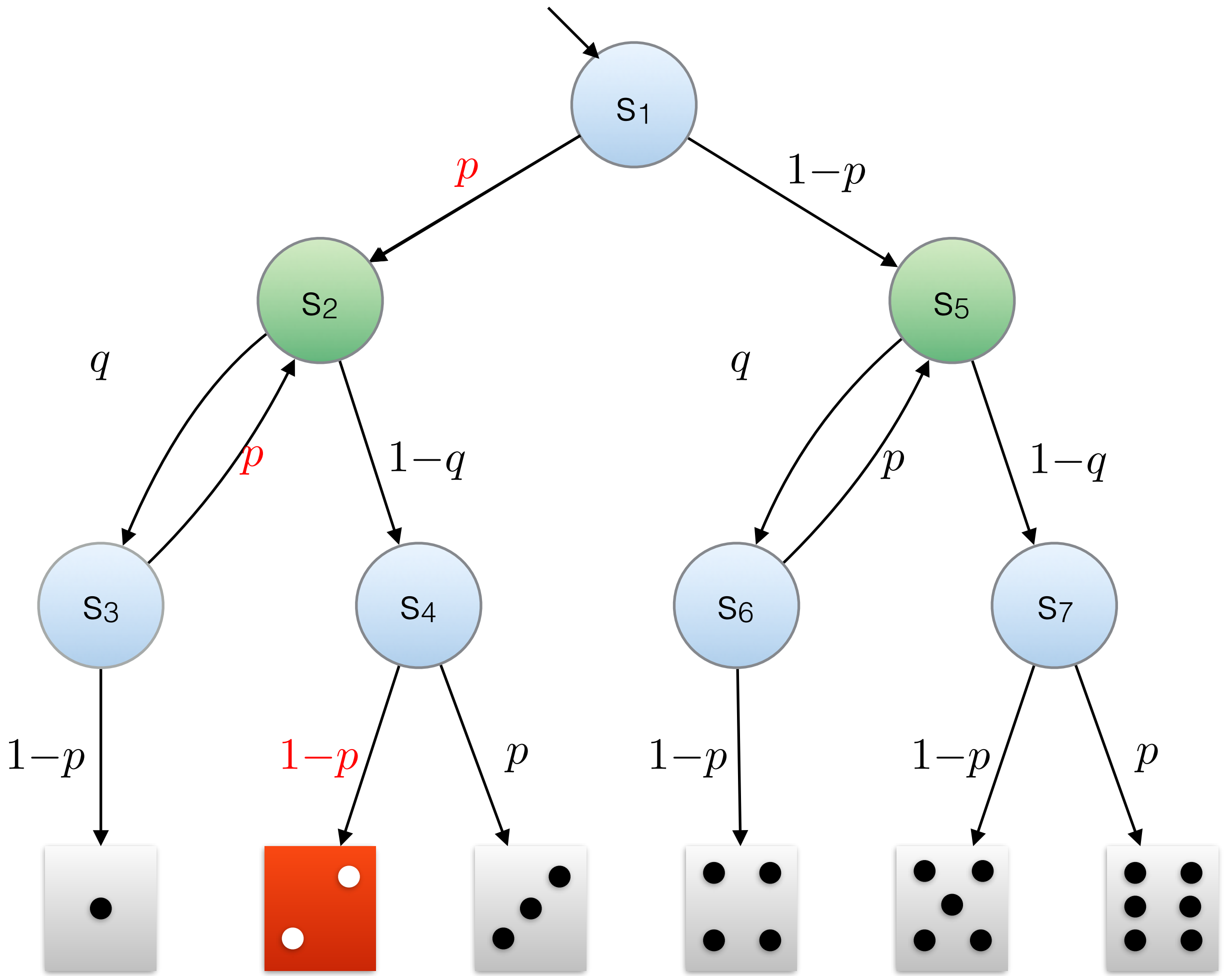
Knuth-Yao Die



Knuth-Yao Die



Knuth-Yao Die



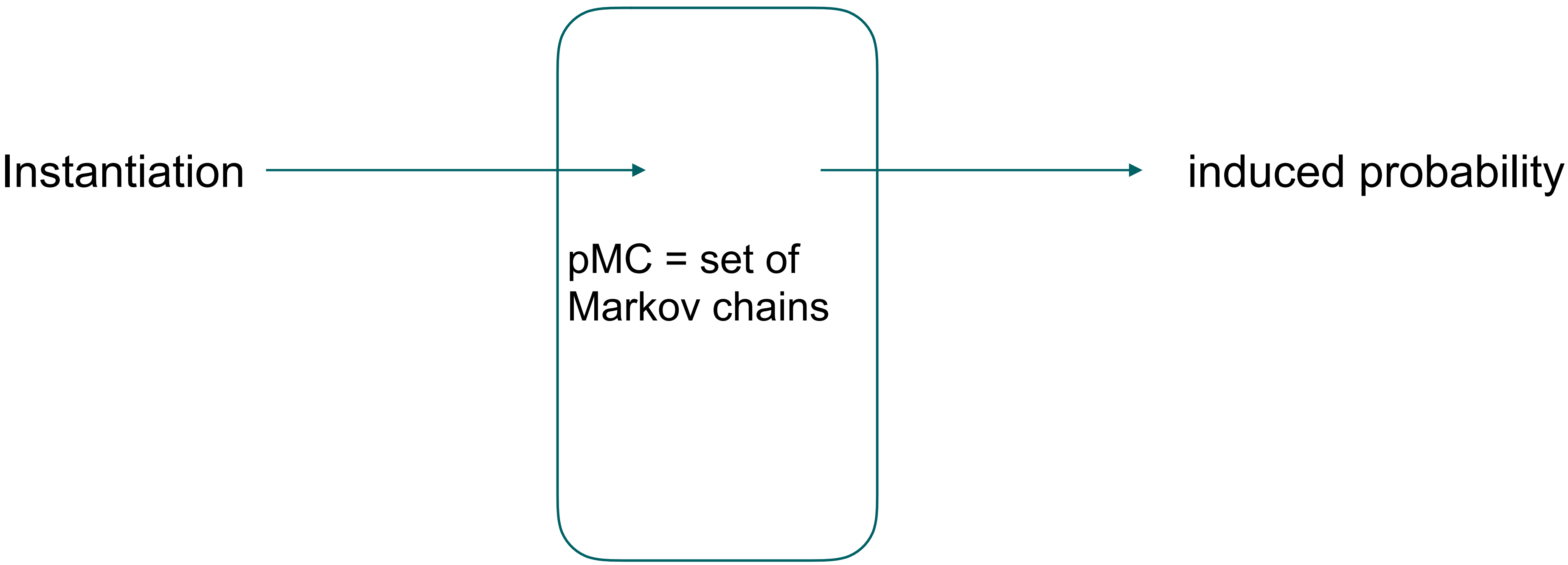
What is the Output of Parameter Synthesis?

So far, three options have been considered in the literature (to the best of our knowledge)

- Option A:
  A generalisation of the output of non-parametric Markov Chain model checking
- Option B:
  A concise description of parameter values that yield satisfactory performance
- Option C:
  One parameter valuation that yields satisfactory performance

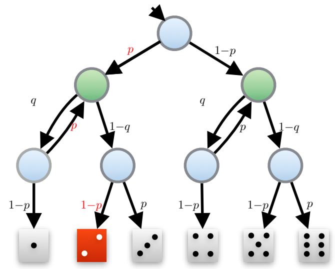
Generalised DTMC “Model Checking”

What is the probability that in a given DTMC T is eventually reached?



State Elimination

- Reachability in MCs = Linear Equation System
  states = rows



Analogous to computing regex from NFAs

[Daws - 2004]
[Hahn et al - 2011]

State Elimination:

Solving Linear Equation Systems while exploiting structure.

- Several heuristics - forward/backward/degree-based
- SCC-based approaches & efficient data structures

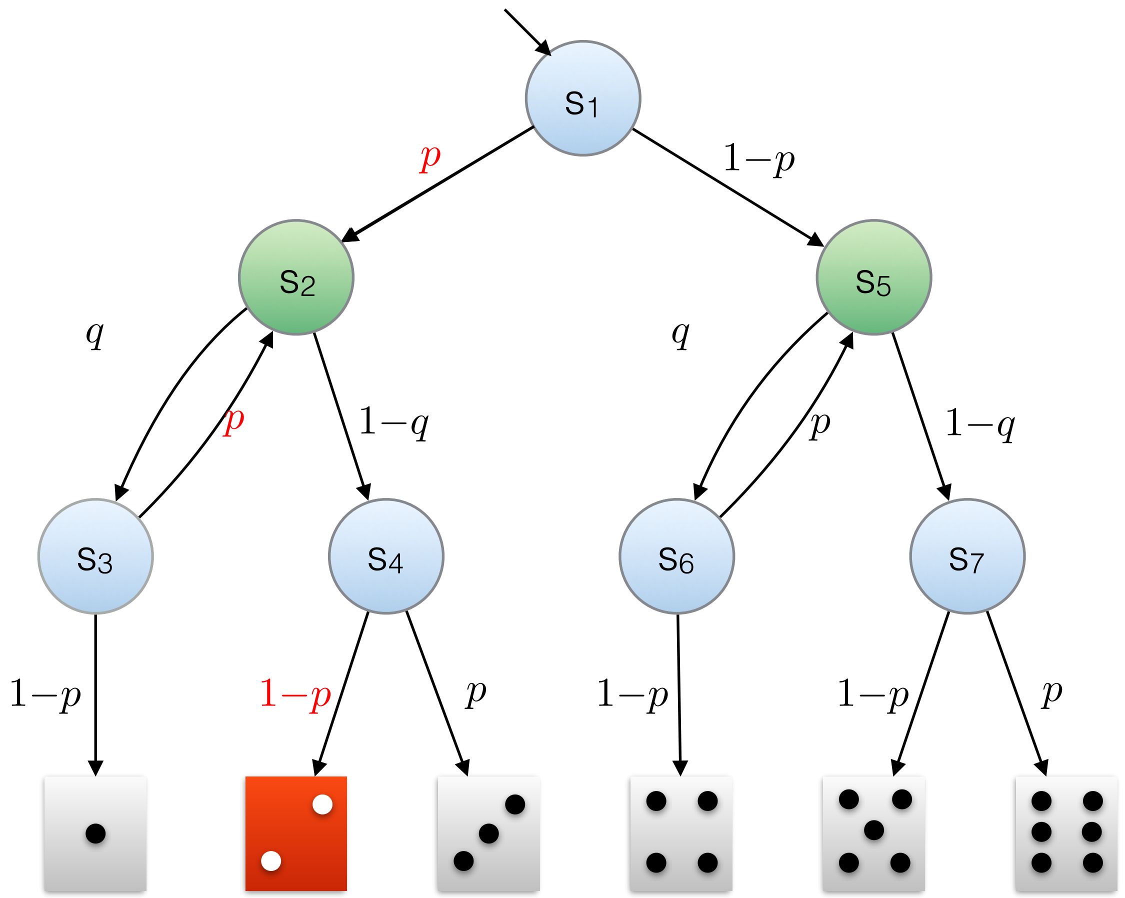
[Jansen et al. - 2014, CAV15]

- Borrow from Floyd-Warshall and Fraction-Free Gaussian

[Accepted - 2019, CAV17]

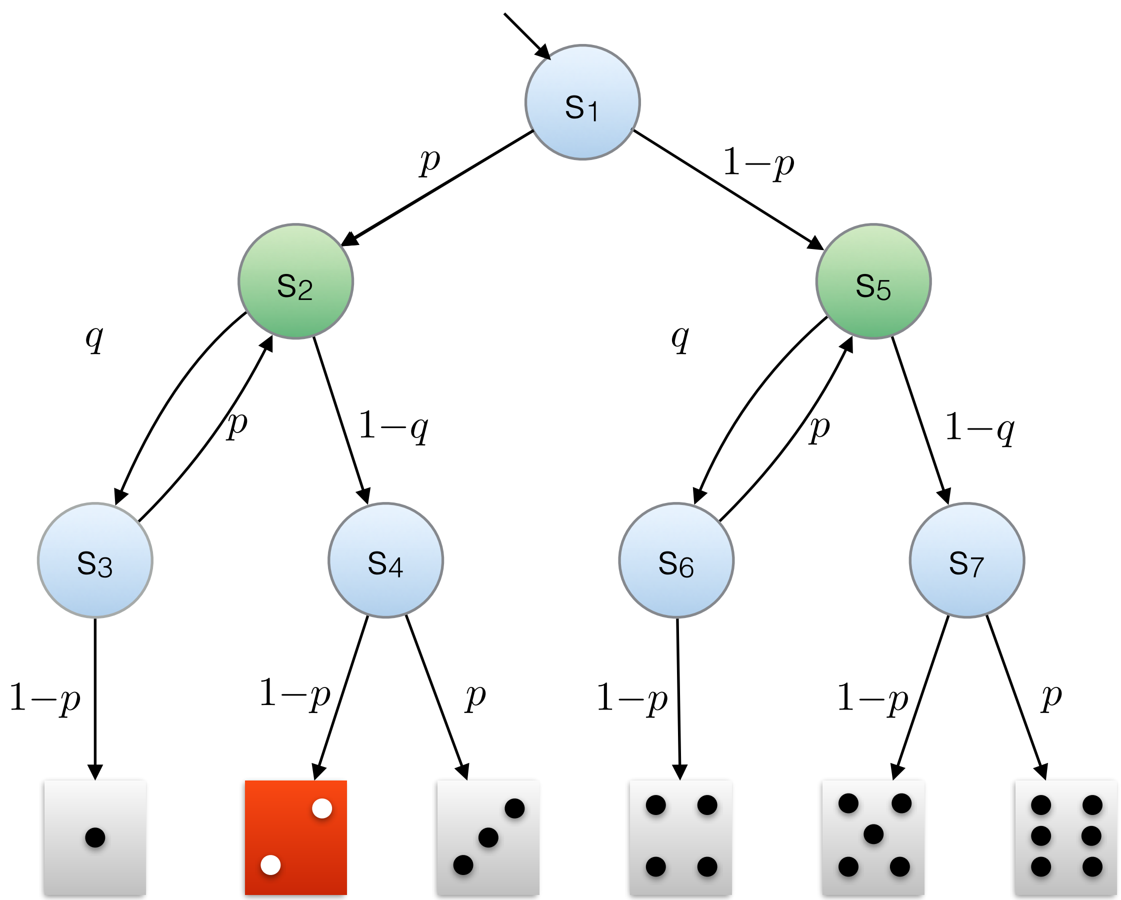
State Elimination

[Daws - 2004]
[Hahn et al - 2011]



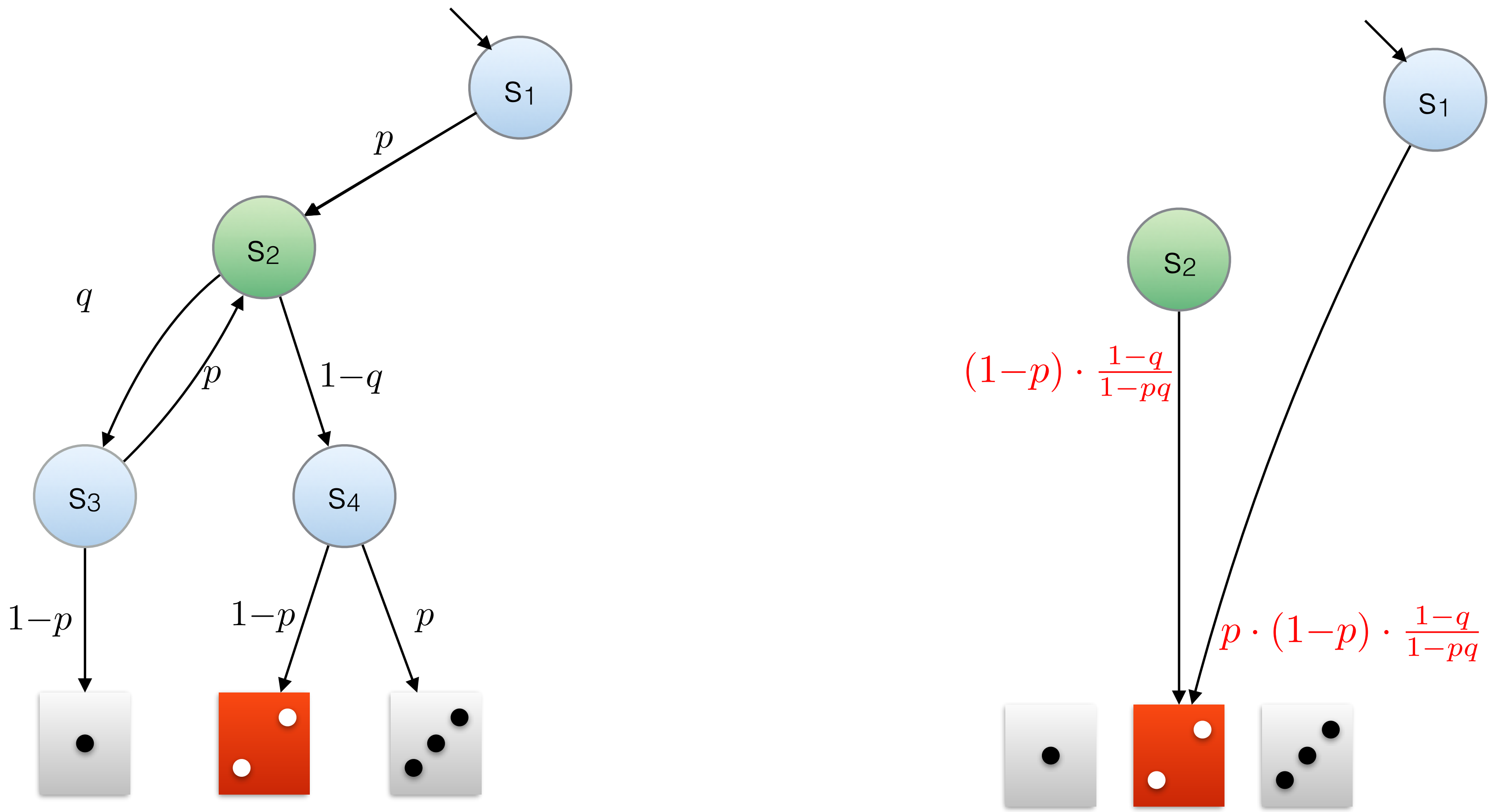
State Elimination

[Daws - 2004]
[Hahn et al - 2011]

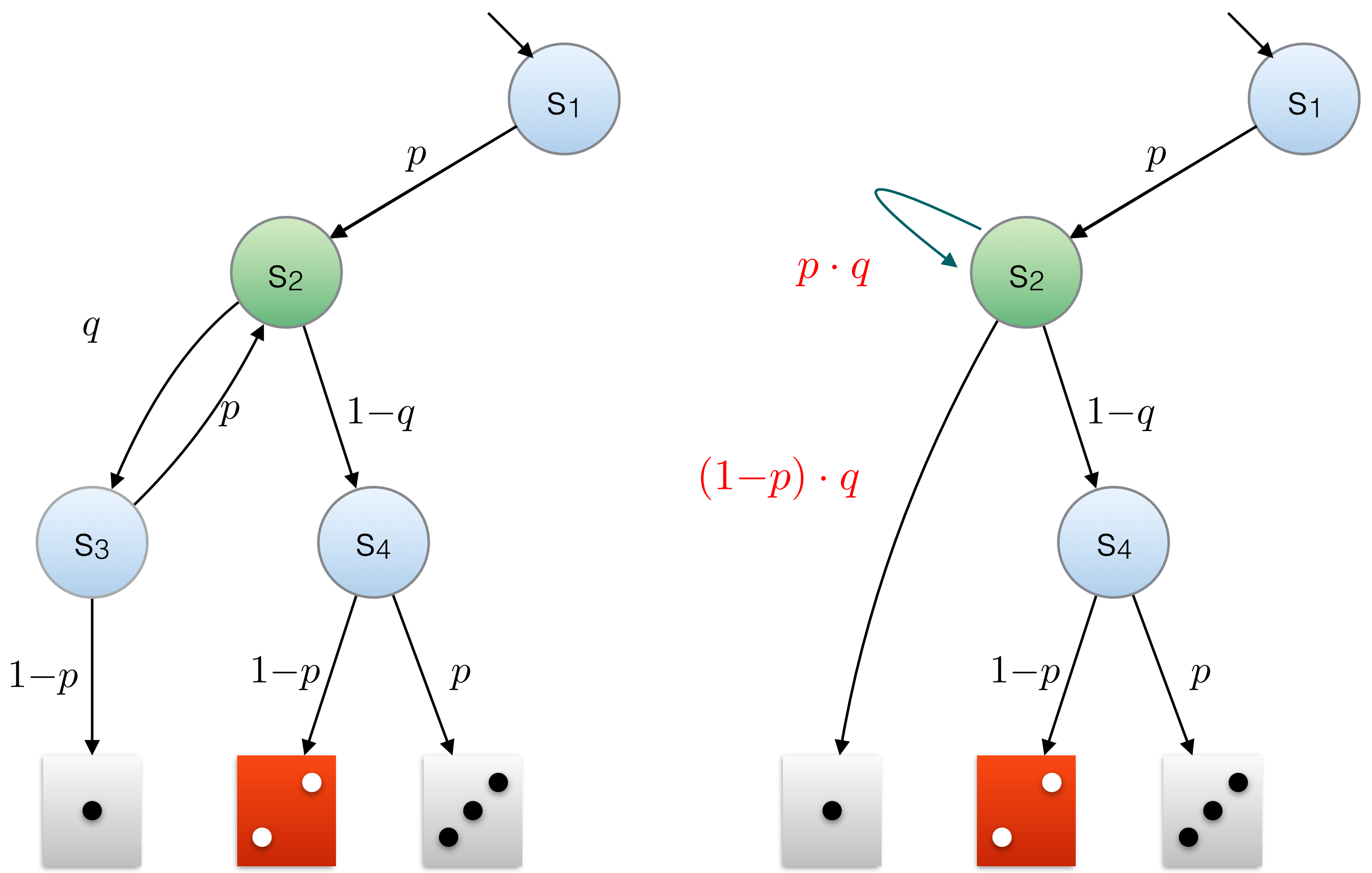


[Daws - 2004]
[Hahn et al - 2011]

State Elimination

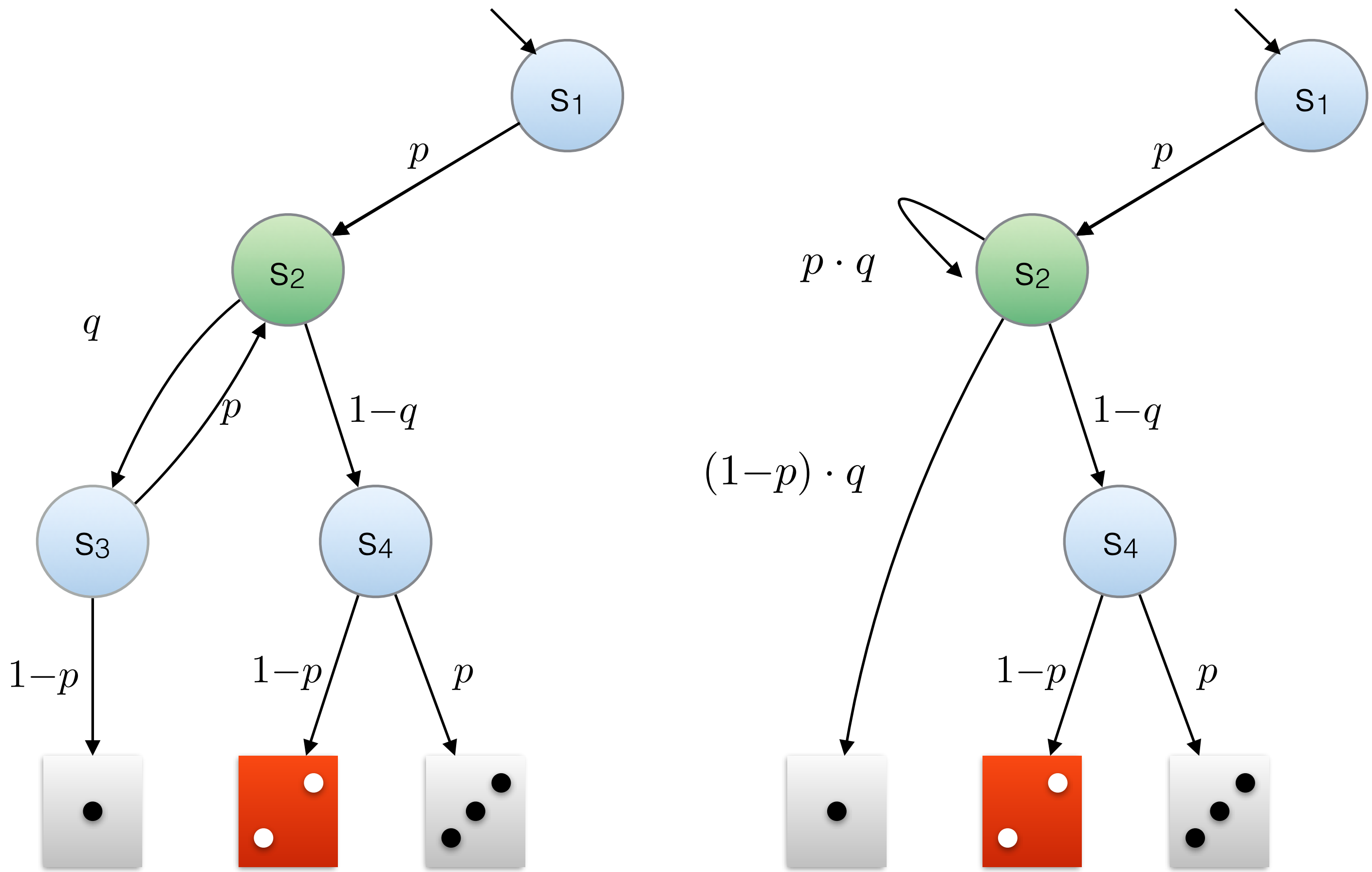


State Elimination



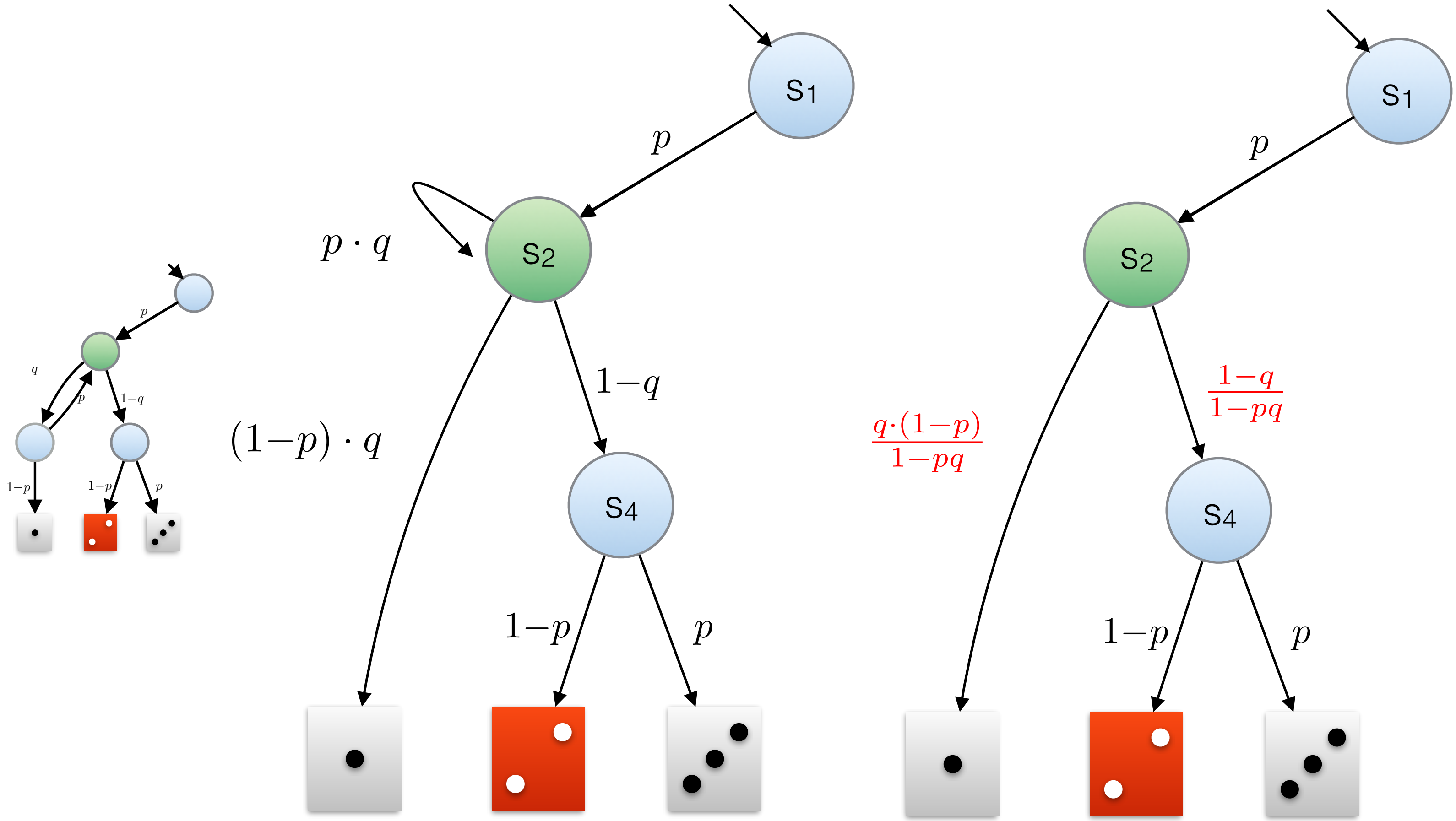
State Elimination

[Daws - 2004]
[Hahn et al - 2011]



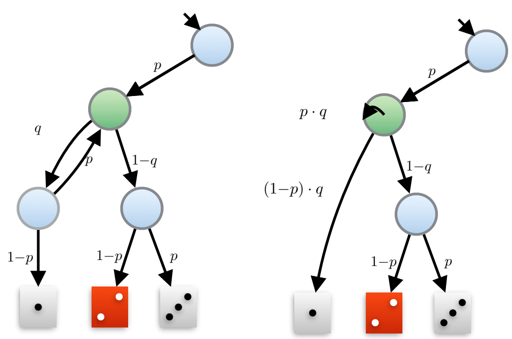
State Elimination

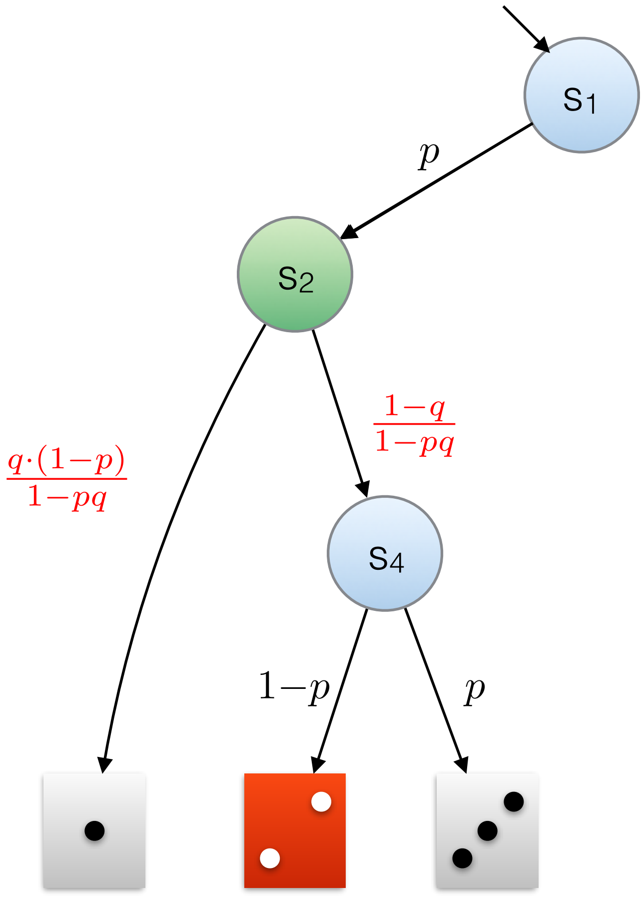
[Daws - 2004]
[Hahn et al - 2011]



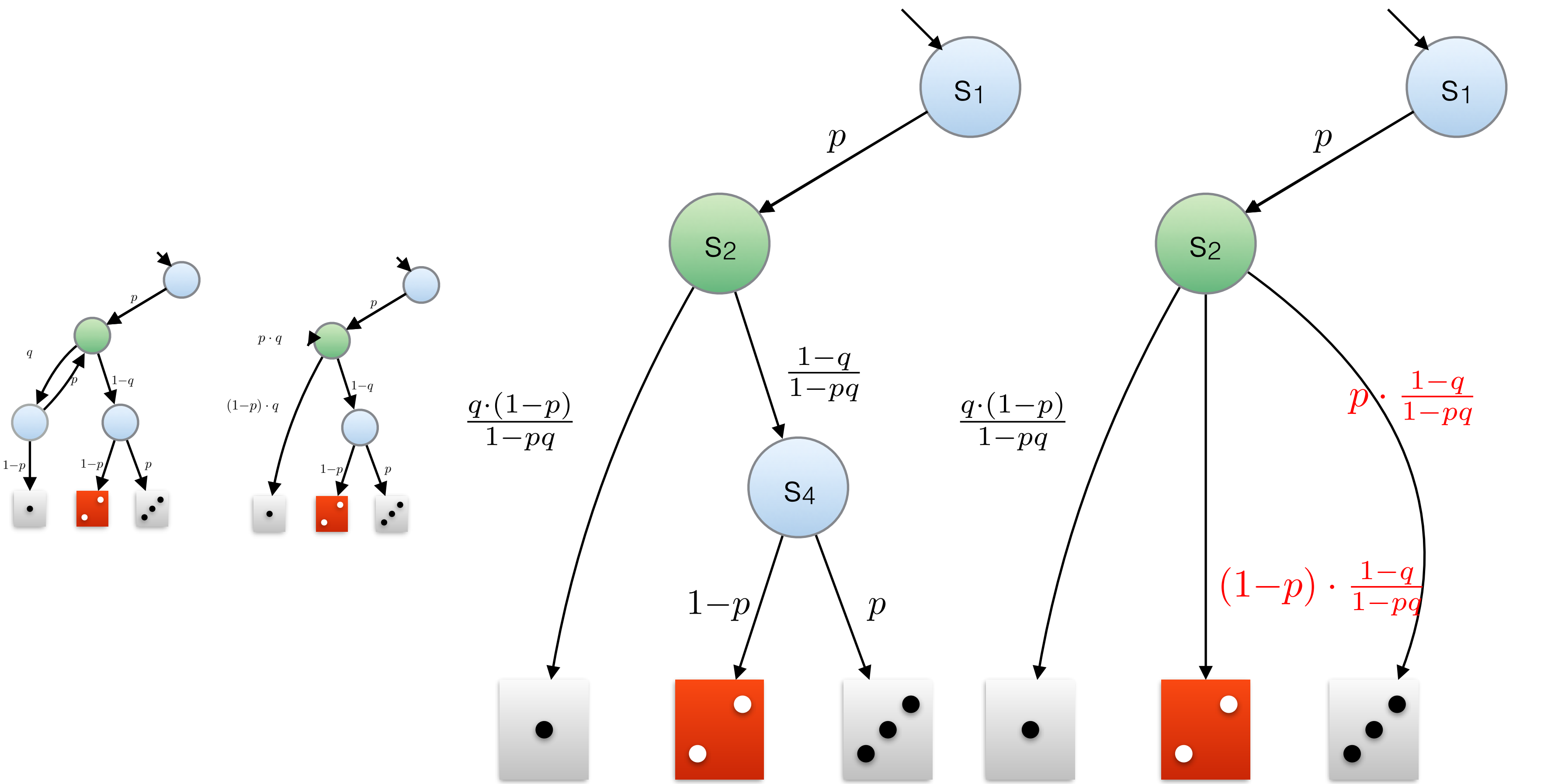
State Elimination

[Daws - 2004]
[Hahn et al - 2011]

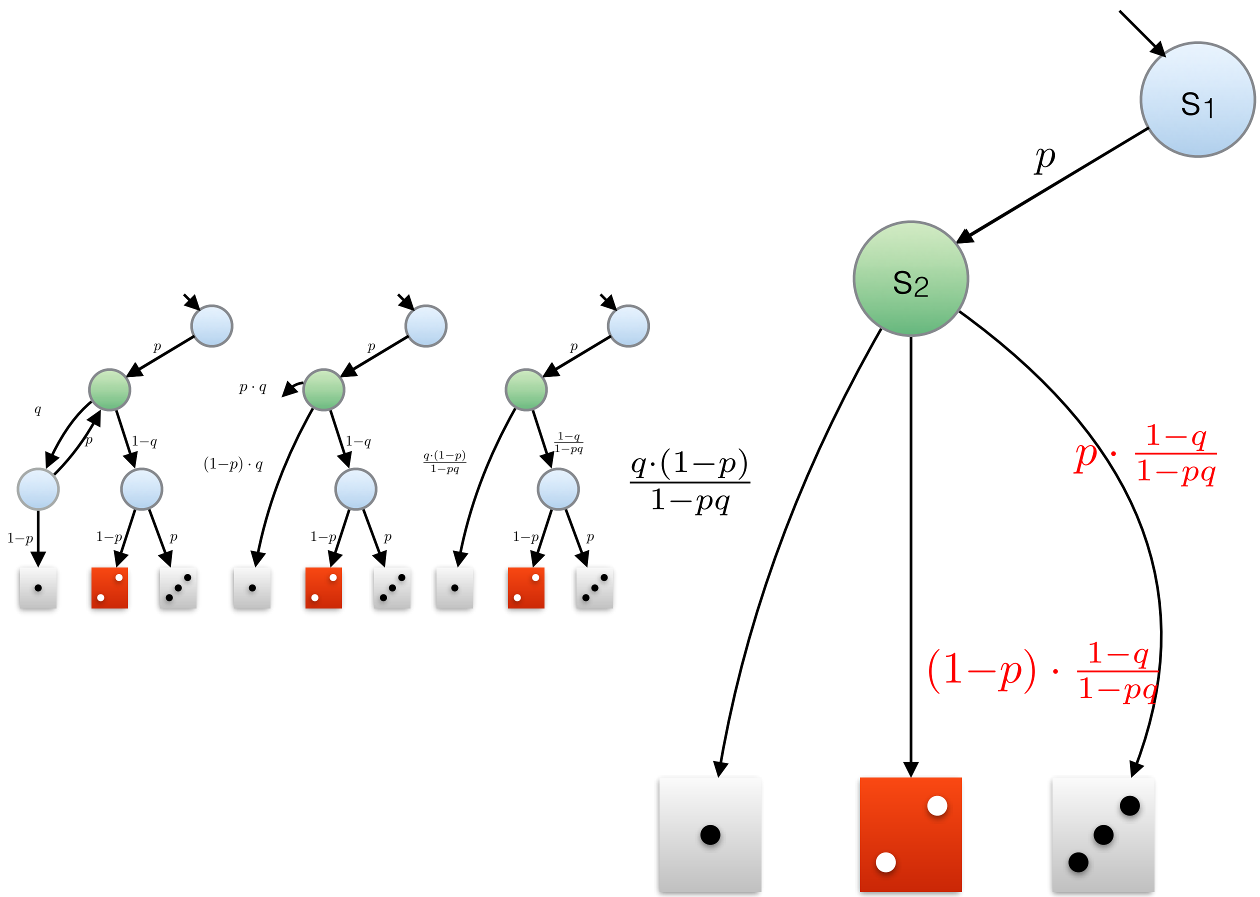




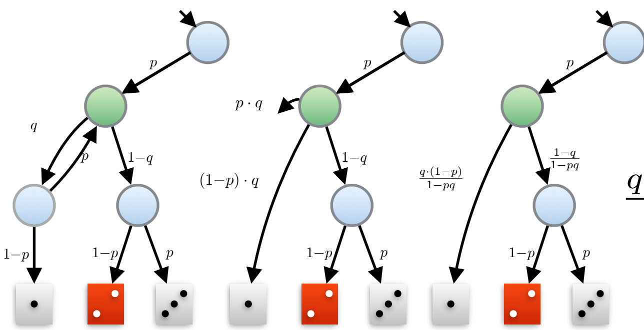
State Elimination

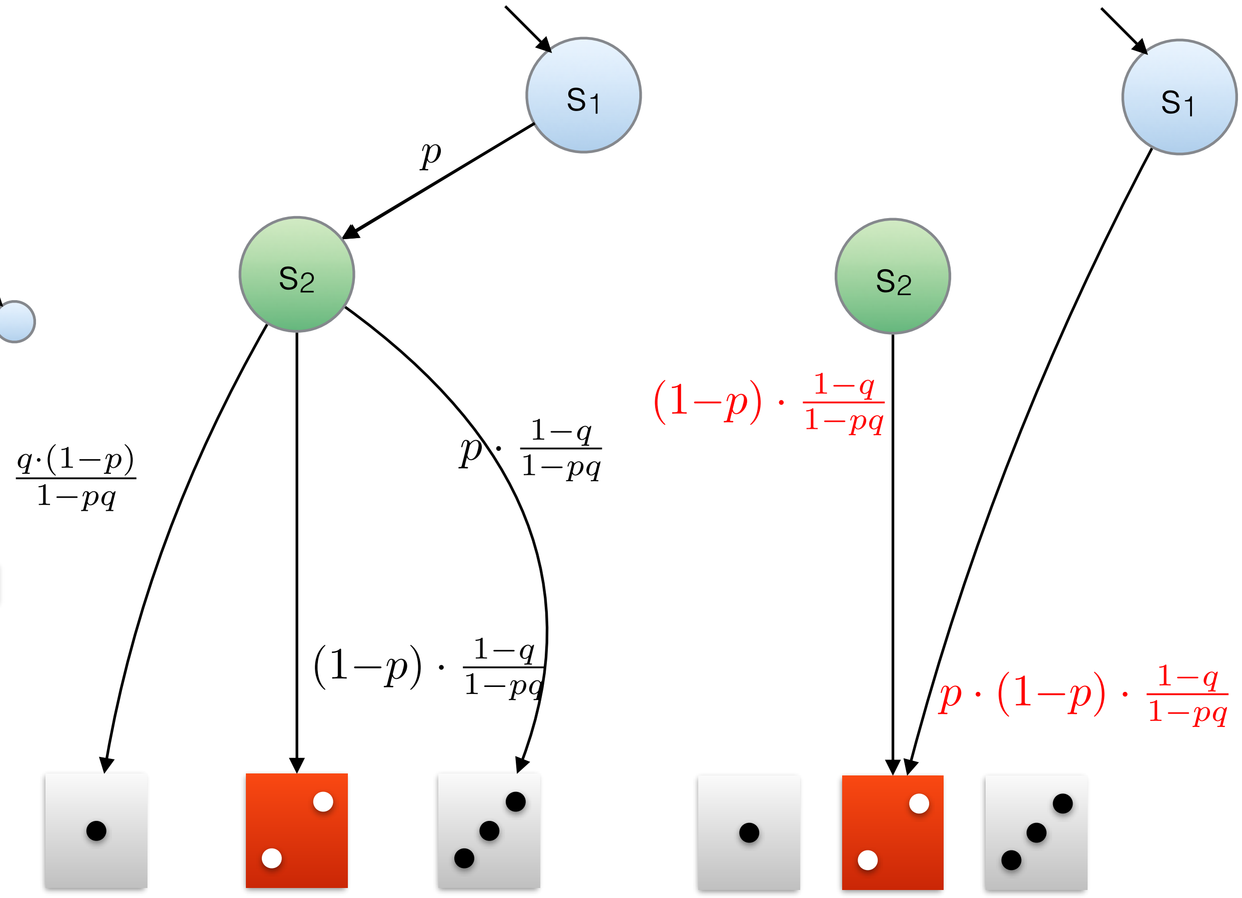


State Elimination



State Elimination





The size of the solution

[Information & Computation, 2019]

Facts

For a pMC with k parameters, n states and polynomial probabilities of degree d.

- The rational function (actually a polynomial) can be exponential in k (even for acyclic pMCs)
- Hence, computation of the function is exponential in the worst case
- For any fixed k, the computation can be done in polynomial time in n and d.

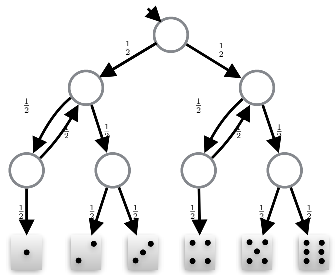
Computing the rational function

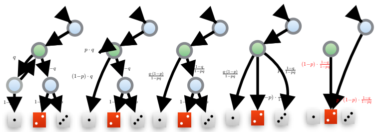
Using Storm

Tool support for state elimination

		instance	#states	#trans	PRISM verif.	PRISM total	PARAM verif.	PARAM total	PROPhESY verif.	PROPhESY total
reachability	brp	(128, 5)	10376	13827	215	218	5	7	2	3
		(256, 5)	20744	27651	1237	1242	32	33	8	10
	crowds	(15, 5)	592060	1754860	TO	TO	18*	48*	1	46
		(20, 5)	2061951	7374951	TO	TO	75*	194*	4	165
	nand	(20, 2)	154942	239832	886	901	44	48	16	22
		(20, 5)	384772	594792	TO	TO	319	328	89	104
exp. reward	egl	(5, 4)	74750	75773	5	11			< 1	5
		(8, 4)	7536638	7602173	543	910	–	–	7	607
	nand	(20, 2)	154942	239832	TO	TO	264	2033	5	12
		(20, 5)	384772	594792	TO	TO	TO	TO	47	64
	zconf	(10000)	10004	20005	TO	TO	TO*	TO*	4	4
		(100000)	100004	200005	TO	TO	TO$^+$	TO$^+$	255	263
conditional	brp	(256, 2)	10757	13827	–	–	–	–	< 1	1
		(256, 5)	20744	27651	–	–	–	–	1	3
	crowds	(15, 5)	592060	1754860					5	50
		(20, 5)	2061951	7374951					14	174

Result of State Elimination





State Elimination

?

Rational function over the parameters

Numerically instable

large

Used by many

Result of State Elimination

41 States - 138 Transitions - 2 Parameters

The numerator has 408 terms,
The denominator is the product of 48 linear polynomials

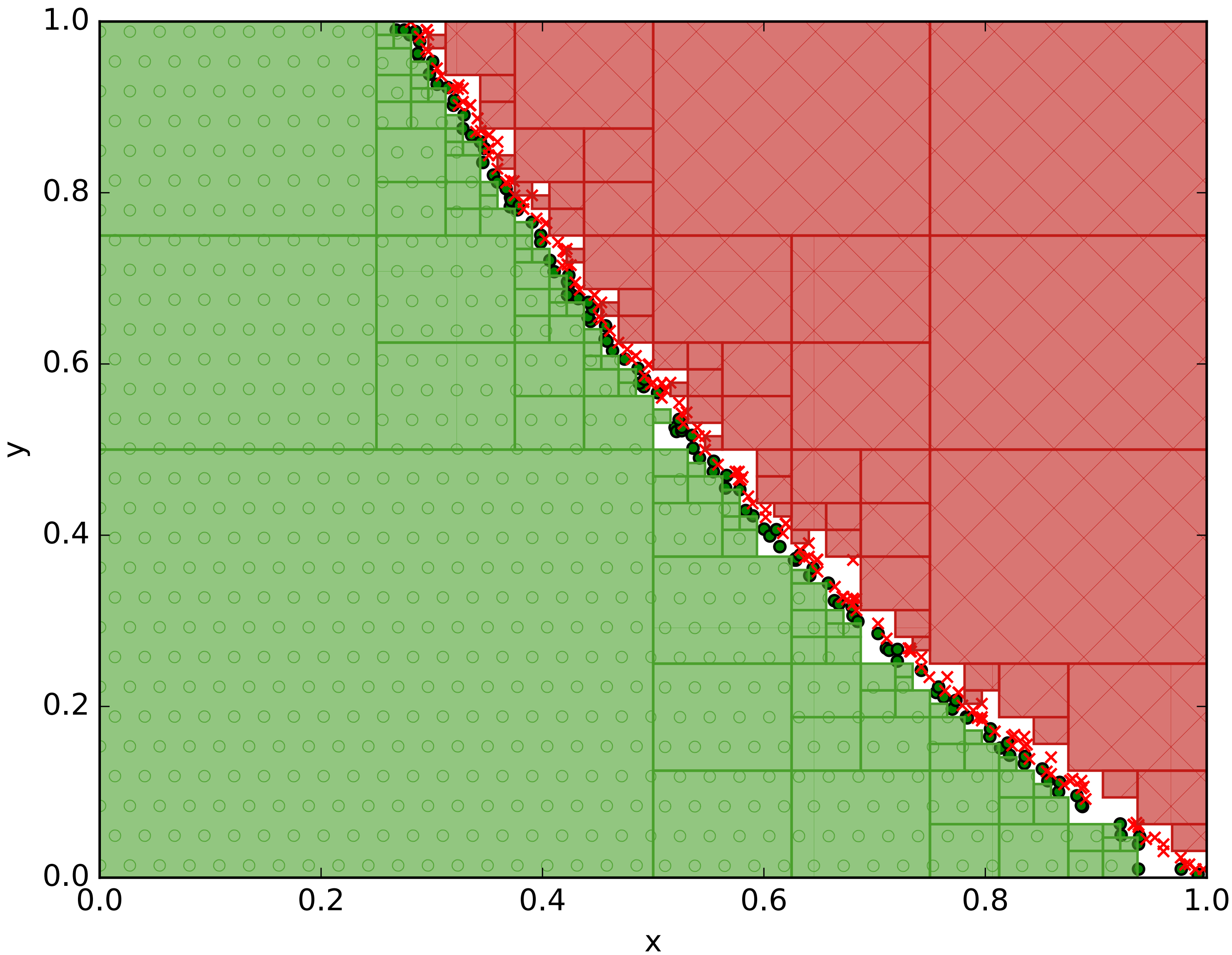
5 seconds

What is the Output of Parameter Synthesis?

So far, three options have been considered in the literature

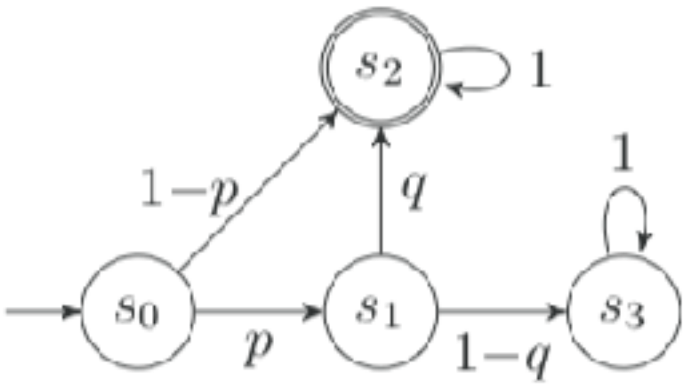
- Option A:
  A generalisation of the output of non-parametric Markov Chain model checking
- Option B:
  A concise description of parameter values that yield satisfactory performance
- Option C:
  One parameter valuation that yields satisfactory performance

Parameter Space Partitioning



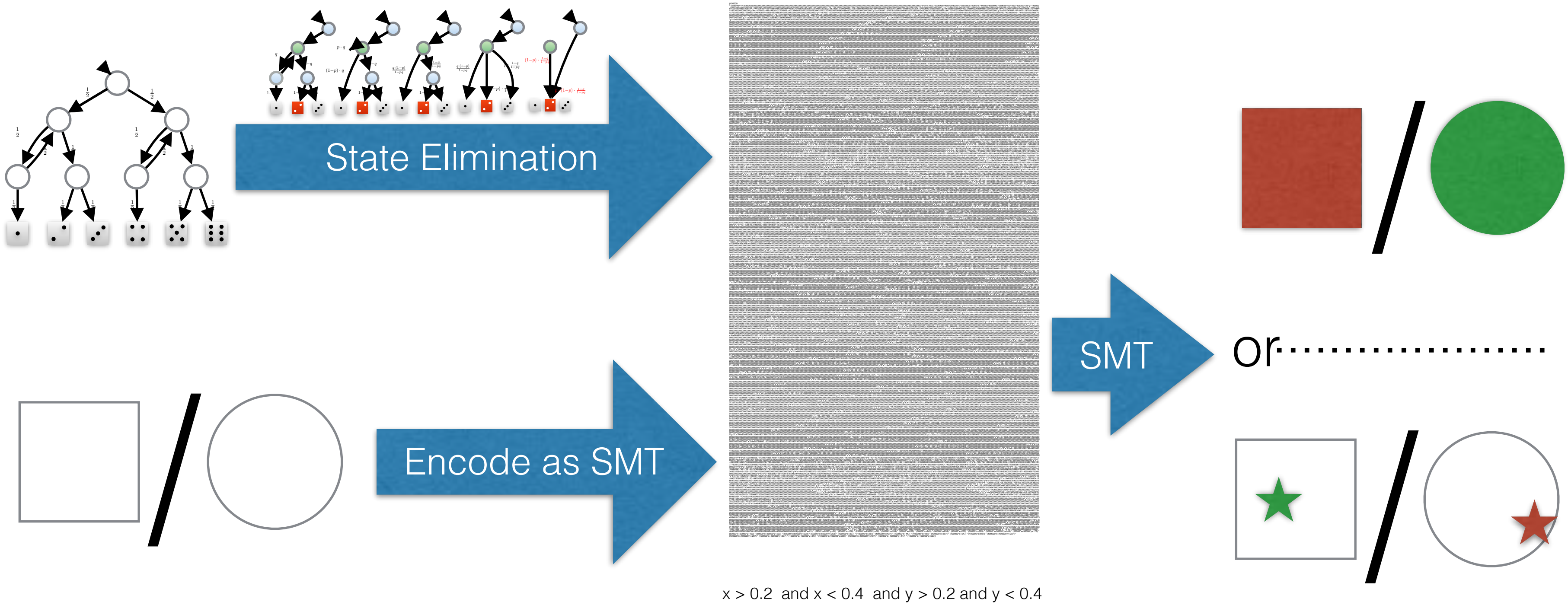
SMT-Encoding

Encoding transition system, satisfiability (region)



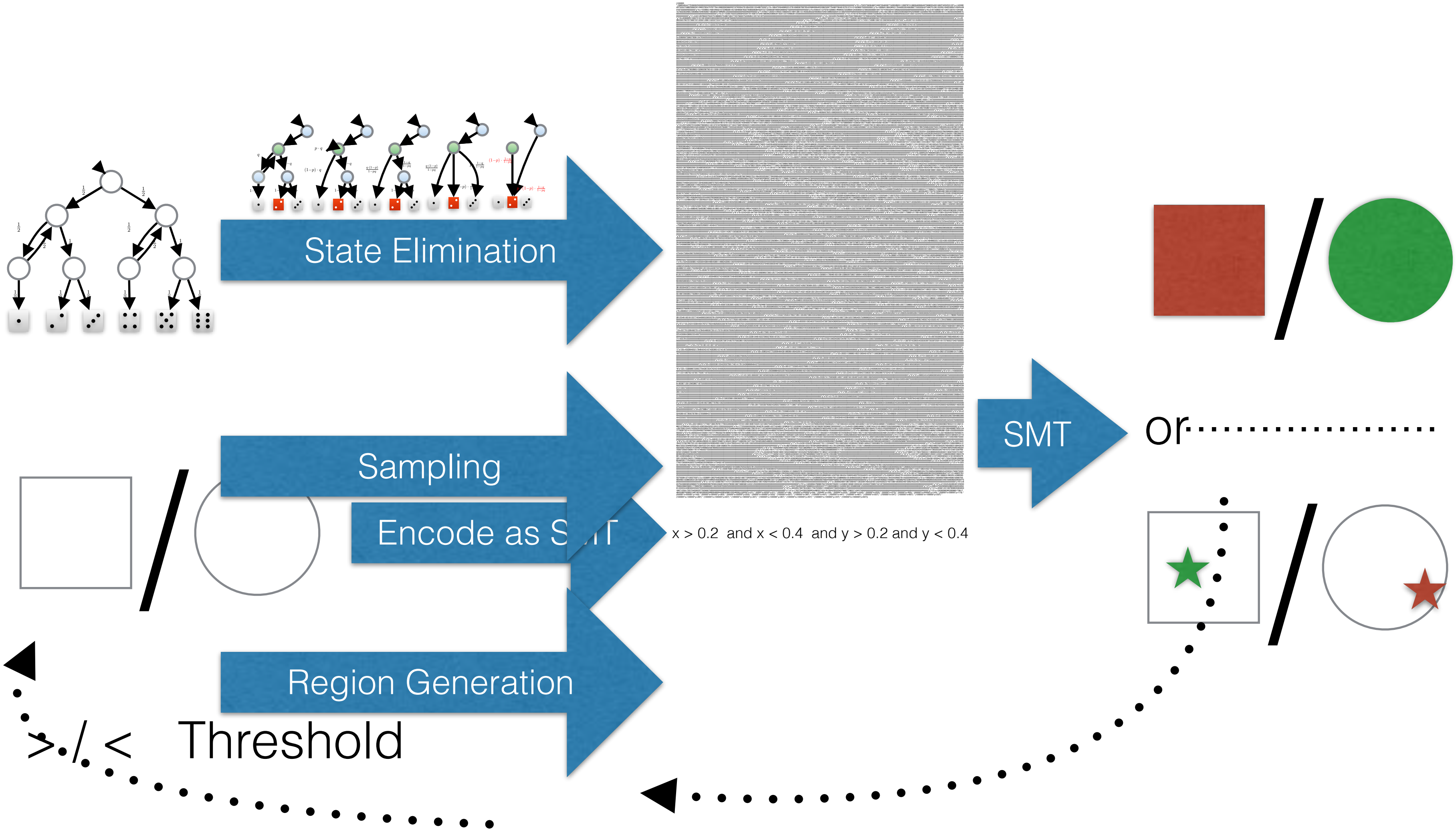
Checking Regions

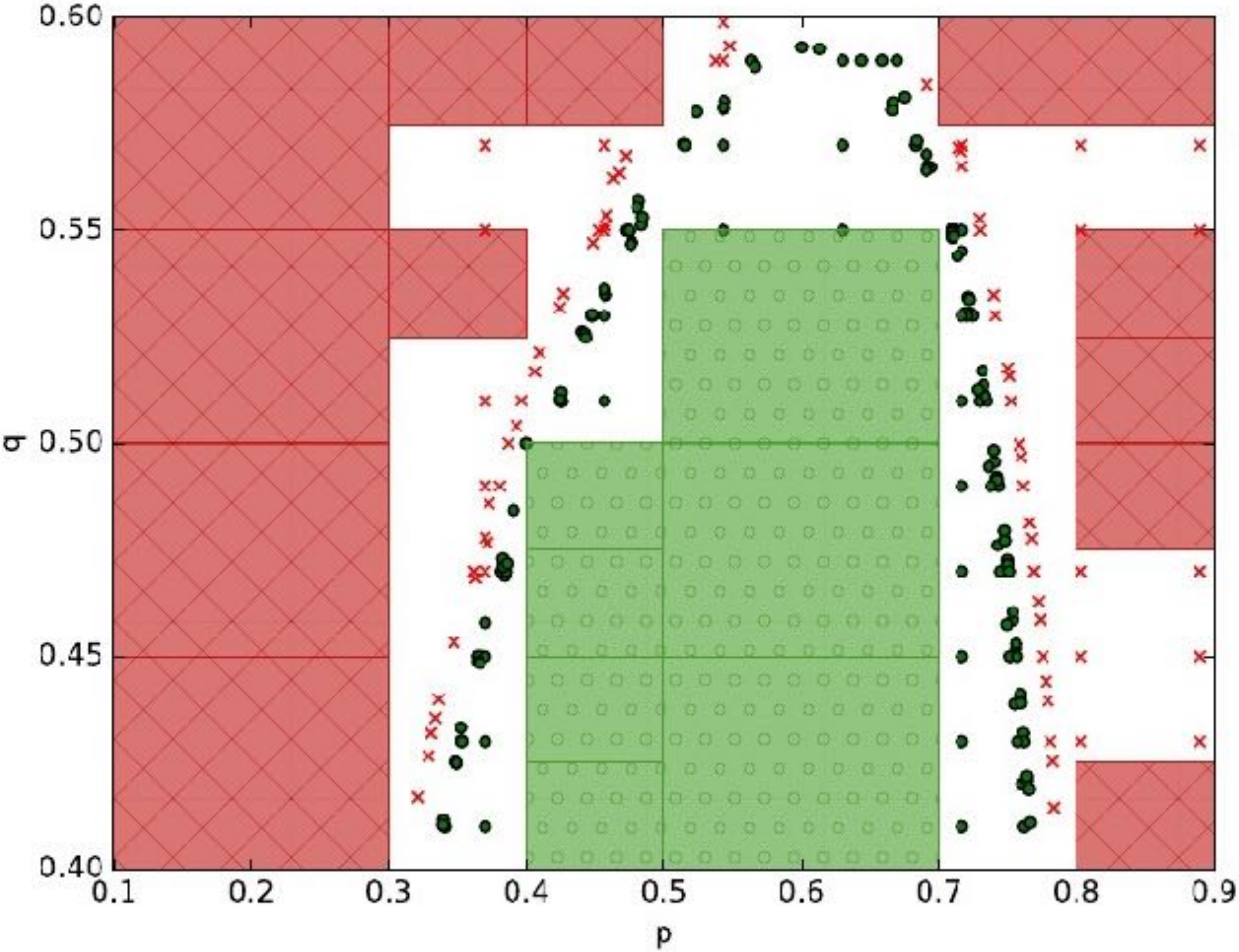
[CAV 2015]



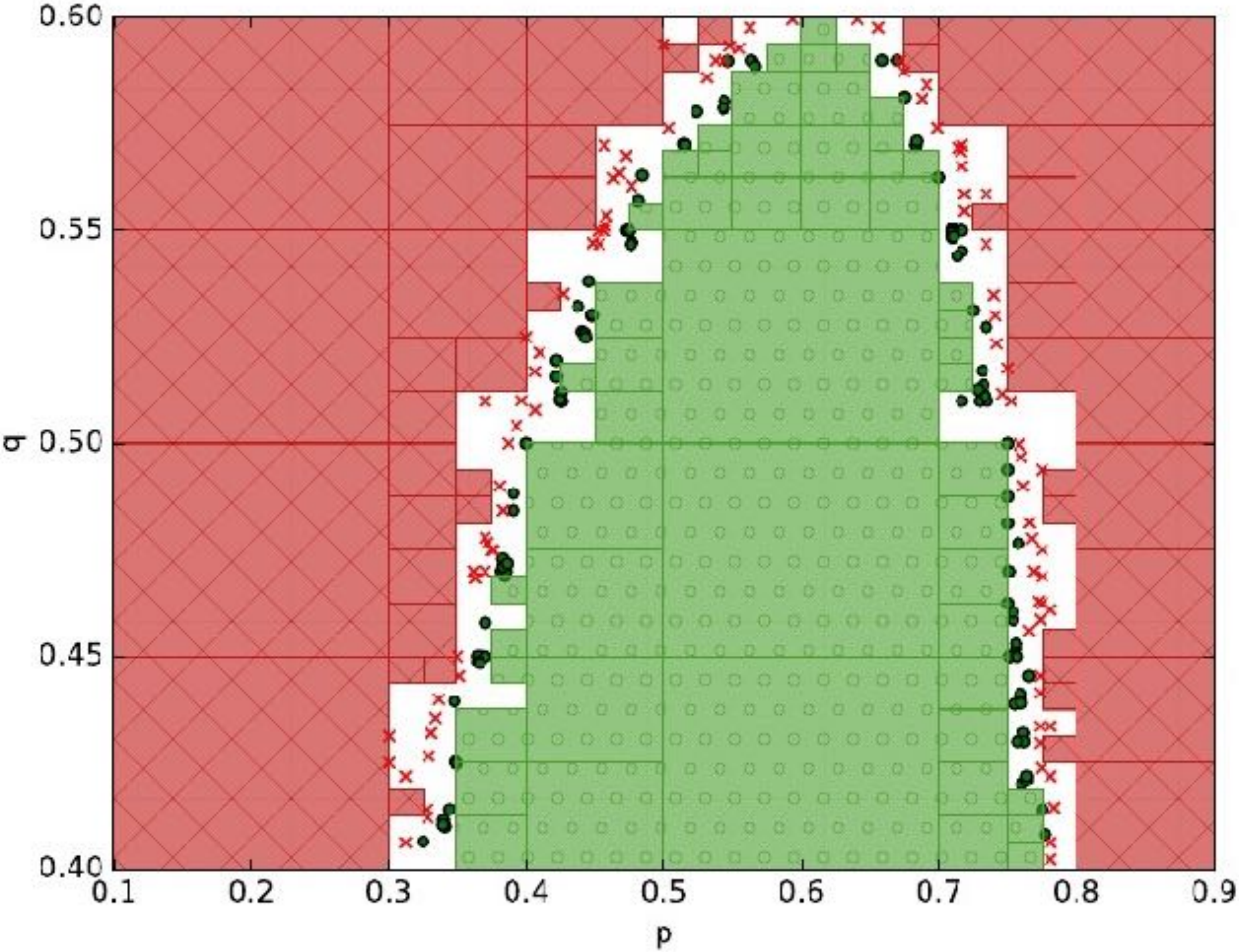
> / < Threshold

Prophesy

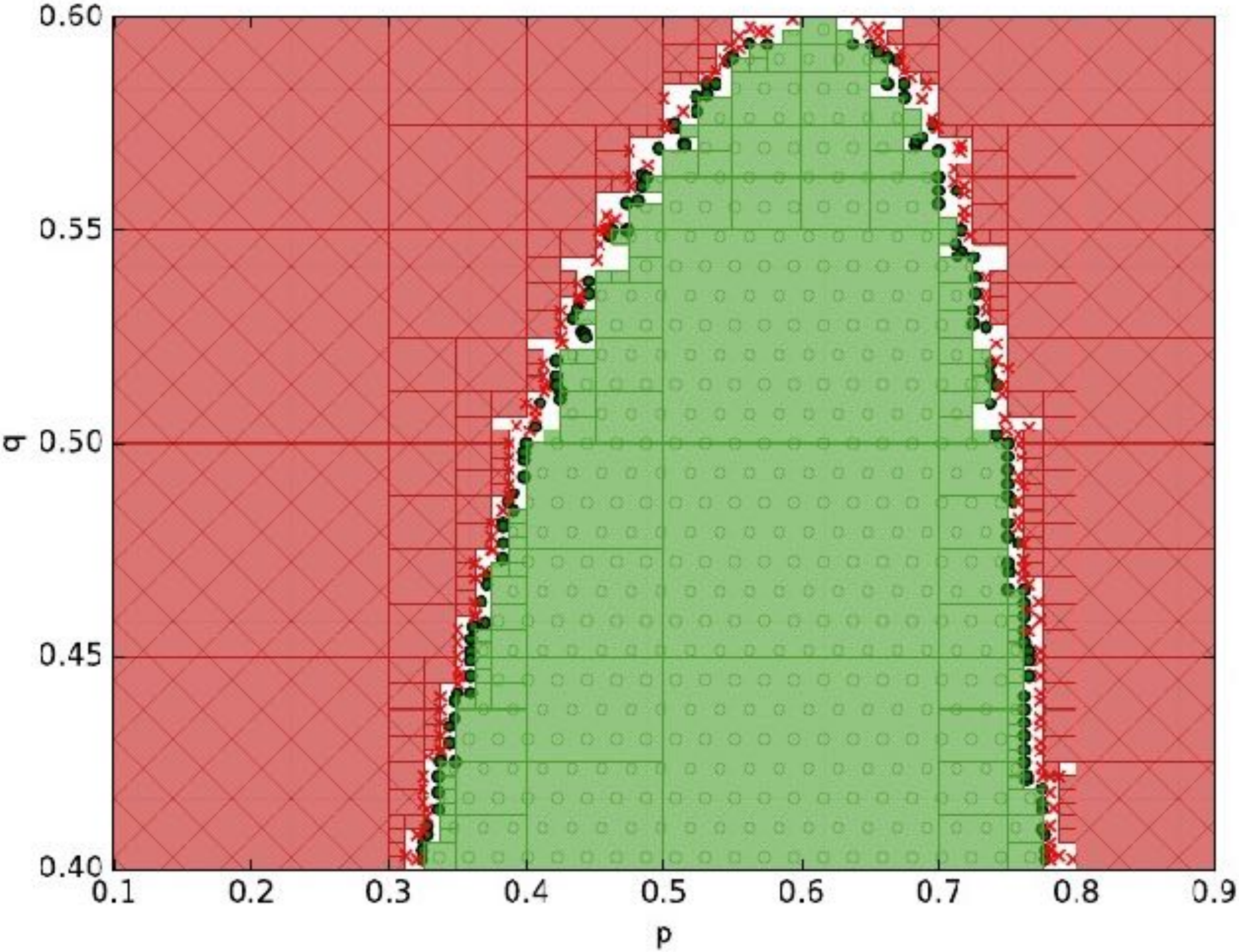

State Elimination
SMT
or
Sampling
Encode as S
x > 0.2 and x < 0.4 and y > 0.2 and y < 0.4
Region Generation
>/< Threshold



0.60
0.55
0.50
0.45
0.40
q
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
p



0.60
0.55
0.50
0.45
0.40
q
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
p


0.60
0.55
0.50
0.45
0.40
q
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
p

Complexity

Does there exist a parameter valuation
such that the induced MC induces probability larger than 1/2?

Facts

- Definitely in the existential theory of the reals (ETR), see earlier.

Thus, in PSPACE.

- Via rational function: number of variables == number of parameters

Thus, as ETR is polynomial for fixed variables,
PTIME for fixed number of parameters (graph-preserving)

- With two objectives: NP-hard (graph-preserving)
- NP-hard (**non-graph preserving**)

Also known to be SQRS-hard (graph-preserving)

Checking Regions w SMT

#4 parameters seems out of reach

Obtaining Rational Function → function is typically huge

Hard to develop heuristics

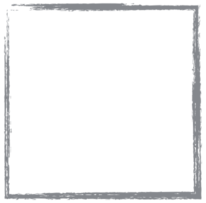
SMT performance unpredictable

Solver does too much work

Rational function is exact everywhere

pMC is a 1.5 player game

pMC is a set of *instantiated* DTMCs



For all points u in the region:

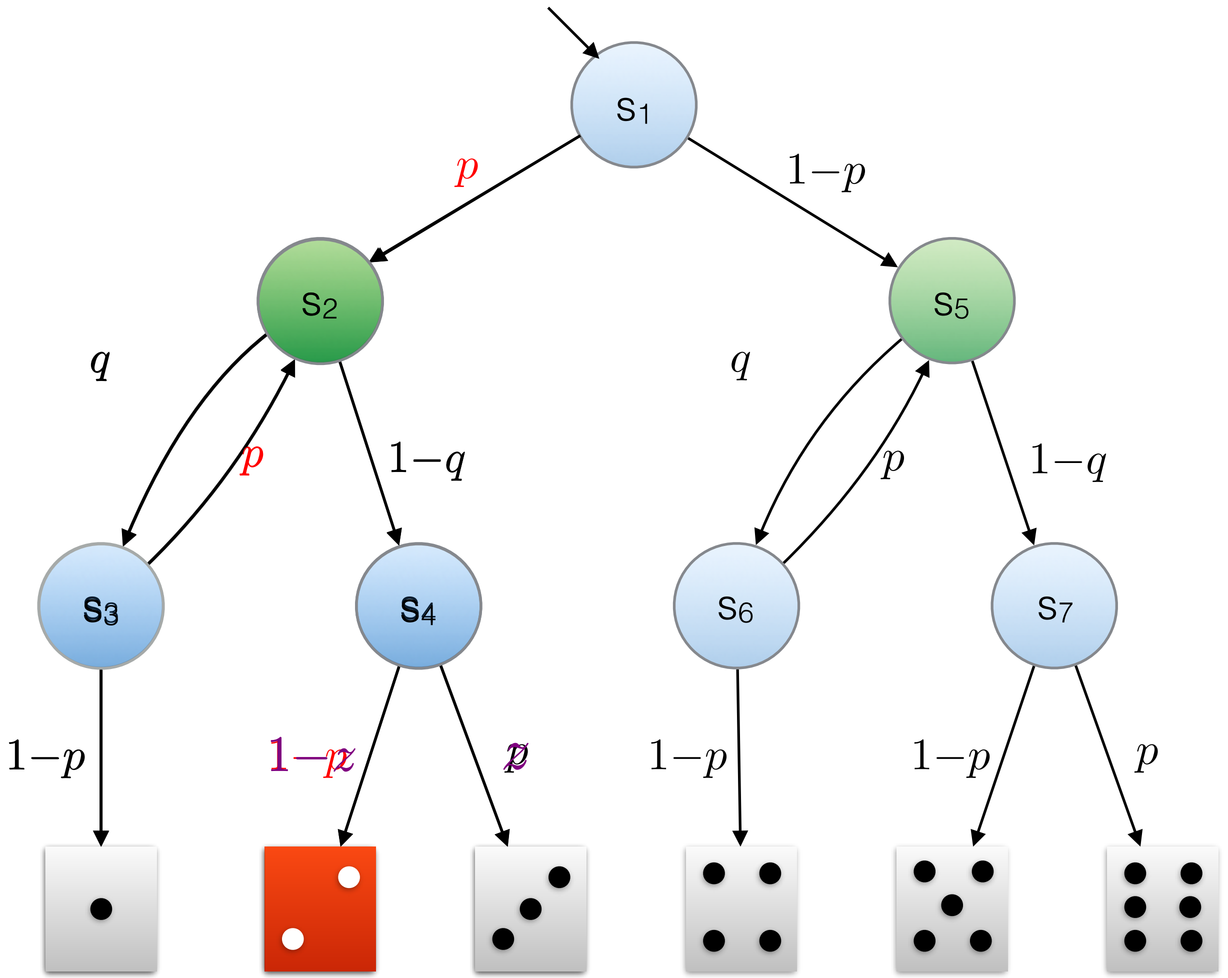
Probability in DTMC(u) < Threshold?

2 Main Ideas

[ATVA 2016]

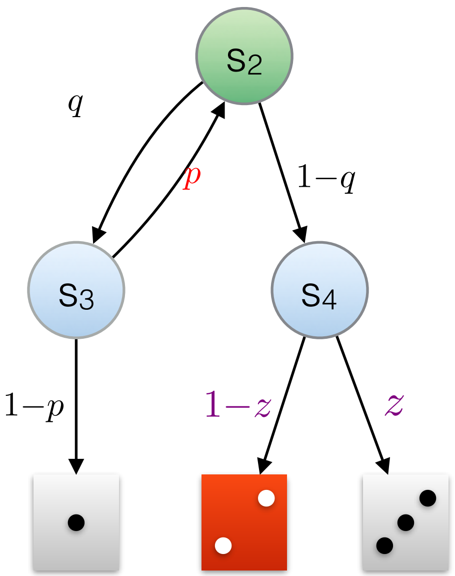
1	2

Relaxation



Relaxed Markov Chain

[ATVA 2016]

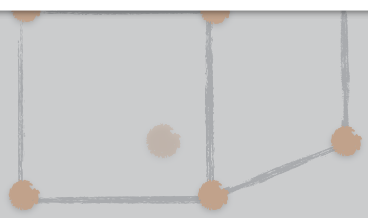


P(◊ )= $(1-z) \cdot \frac{1-q}{1-pq}$

...st at 1 state;

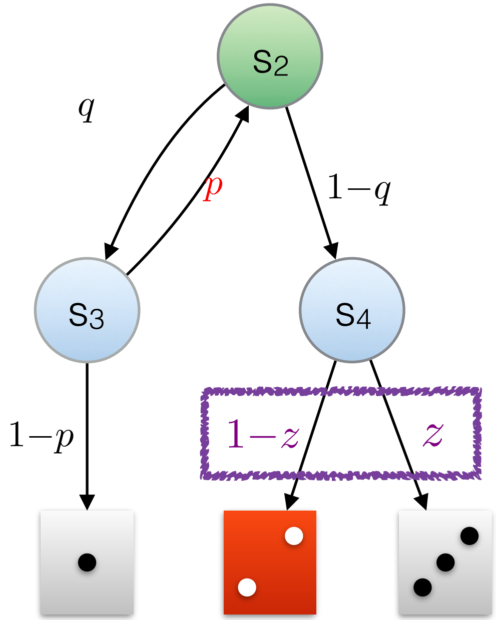
...at most degree 1

...t extremal values

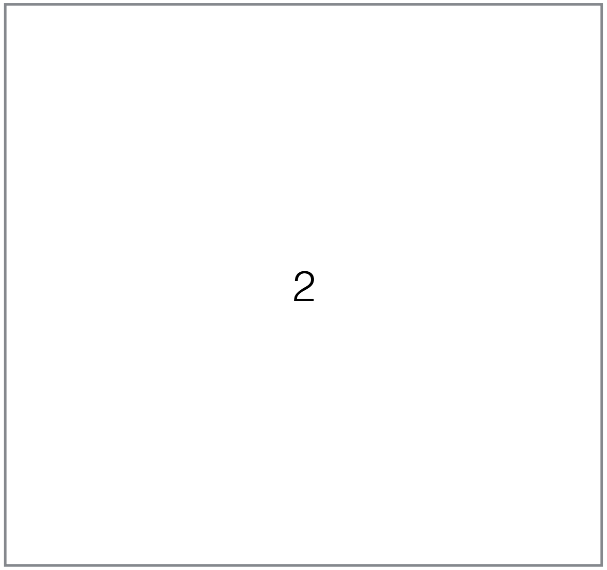


2 Main Ideas

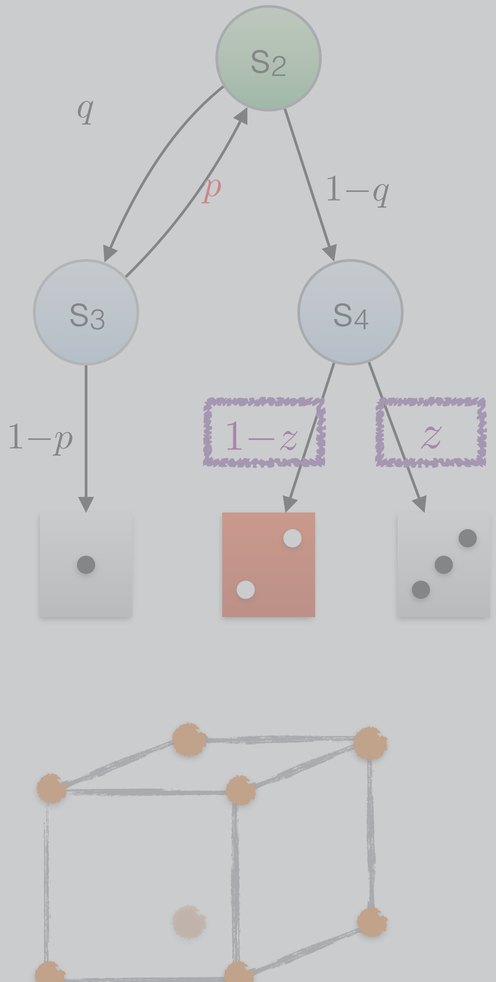
[ATVA 2016]

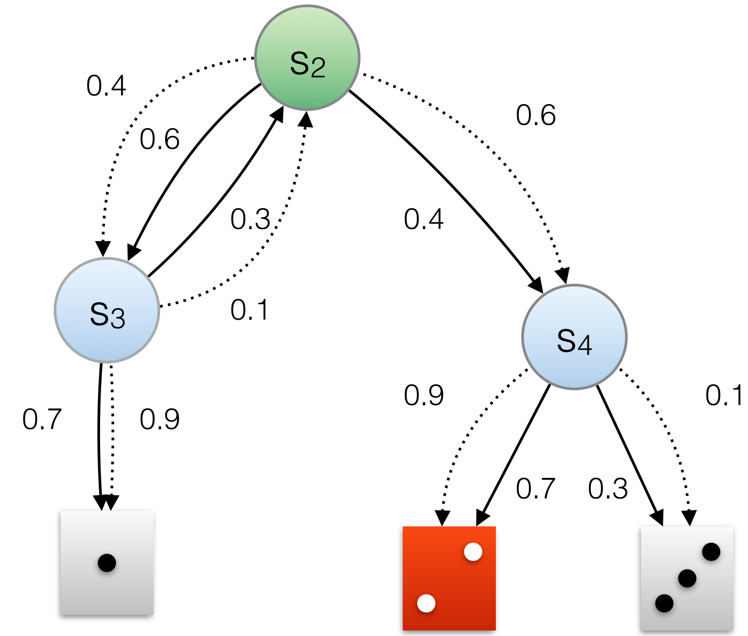


Relaxation



Lifting

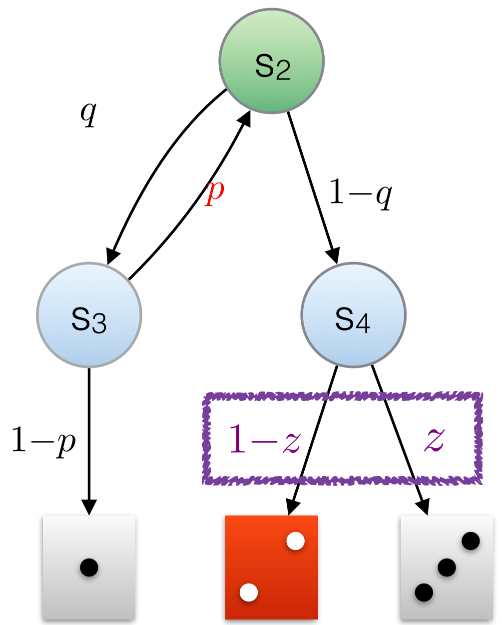




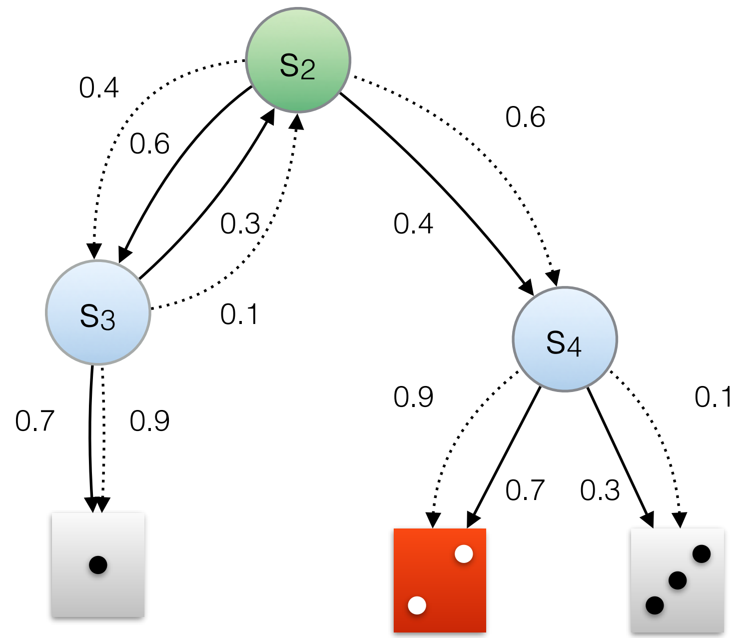
We obtain an MDP

2 Main Ideas

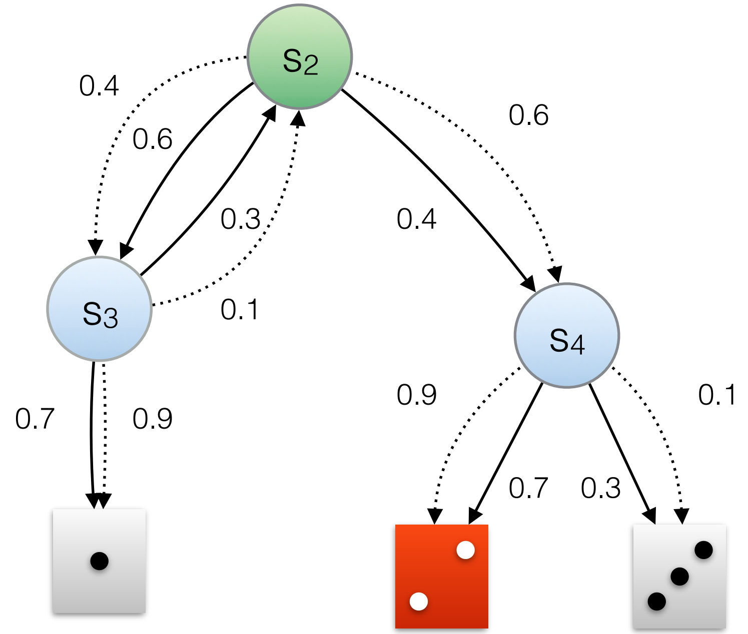
[ATVA 2016]



Relaxation



Lifting

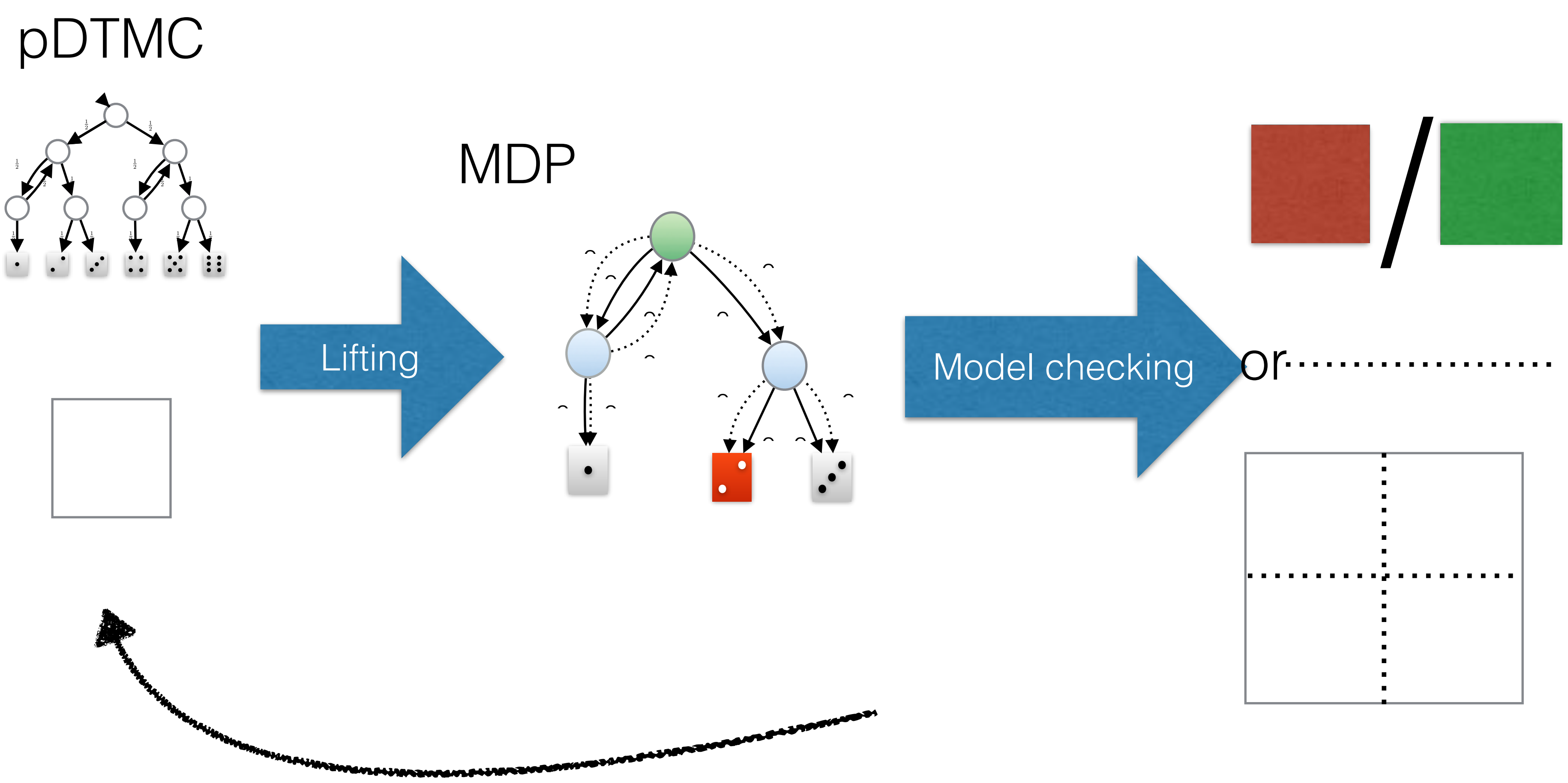


Lifting

Lifting (Relaxation (***D***)) = Lifting (***D***)

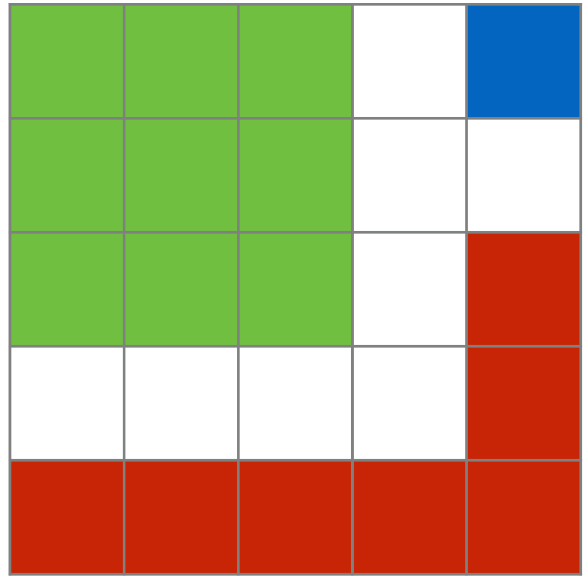
Checking Regions

[ATVA 2016]



Parameter Lifting Performance

[ATVA 2016]



		instance	φ	#pars	#states	#trans	#par trans	ℓ	safe	unsafe	neither	unkn
pMC	brp	(256,5)	E	2	20 744	27 651	13814	51	14.9%	79.2%	5.8%	0.2%
		(256,5)	P	4	20 744	27 651	13814	71	7.5%	51.0%	40.6%	0.8%
	crowds	(10,5)	P	2	104 512	246 082	51 480	44	54.4%	41.1%	4.2%	0.3%
	nand	(10,5)	P	2	35 112	52 647	25 370	21	21.4%	68.5%	6.9%	3.2%
pMDP	brp	(256,5)	P	2	40 721	55 143	27 800	153	6.6%	90.4%	3.0%	0.0%
	consensus	(4,2)	P	4	22 656	75 232	29 376	357	2.6%	87.0%	10.4%	0.0%
	sav	(6,2,2)	P	4	379	1 127	552	2	44.0%	15.4%	35.4%	5.3%
	zeroconf	(2)	P	2	88 858	203 550	80 088	186	16.6%	77.3%	5.6%	0.5%

Parameter Lifting Performance

Covering 95%

SMT TO (1h) everywhere

	benchmark	instance	φ	#pars	#states	#trans	PLA #regions	PLA direct	PLA bisim	PRISM best
pMC	brp	(256,5)	$\mathbb{P}$	2	19 720	26 627	37	6	14	TO
		(4096,5)	$\mathbb{P}$	2	315 400	425 987	13	**233**	TO	TO
		(256,5)	$\mathbb{E}$	2	20 744	27 651	195	**8**	15	TO
		(4096,5)	$\mathbb{E}$	2	331 784	442 371	195	502	**417**	TO
		(16,5)	$\mathbb{E}$	4	1 304	1 731	1 251 220	2 764	**1 597**	TO
		(32,5)	$\mathbb{E}$	4	2 600	3 459	1 031 893	TO	**2 722**	TO
		(256,5)	$\mathbb{E}$	4	20 744	27 651		TO	TO	TO
	crowds	(10,5)	$\mathbb{P}$	2	104 512	246 082	123	17	**6**	2038
		(15,7)	$\mathbb{P}$	2	8 364 409	25 108 729	116	1 880	**518**	TO
		(20,7)	$\mathbb{P}$	2	45 421 597	164 432 797	119	TO	**2 935**	TO
	nand	(10,5)	$\mathbb{P}$	2	35 112	52 647	469	**22**	30	TO[2]
		(25,5)	$\mathbb{P}$	2	865 592	1 347 047	360	**735**	2 061	TO
pMDP	brp	(256,5)	$\mathbb{P}$	2	40 721	55 143	37	**85**	3 359	TO
		(4096,5)	$\mathbb{P}$	2	647 441	876 903	13	**3 424**	TO	TO
	consensus	(2,2)	$\mathbb{P}$	2	272	492	119	**<1**	<1	31[4]
		(2,32)	$\mathbb{P}$	2	4 112	7 692	108	**113**	141	TO[4]
		(4,2)	$\mathbb{P}$	4	22 656	75 232	6 125	**1 866**	2 022	TO[4]
		(4,4)	$\mathbb{P}$	4	43 136	144 352	–	TO	TO	TO[4]
	sav	(6,2,2)	$\mathbb{P}$	2	379	1 127	162	**<1**	<1	TO[4]
		(100,10,10)	$\mathbb{P}$	2	1 307 395	6 474 535	37	**1 612**	TO	TO
		(6,2,2)	$\mathbb{P}$	4	379	1 127	621 175	944	**917**	TO[4]
		(10,3,3)	$\mathbb{P}$	4	1 850	6 561		TO	TO	TO[4]
	zeroconf	(2)	$\mathbb{P}$	2	88 858	203 550	186	**86**	1 295	TO
		(5)	$\mathbb{P}$	2	494 930	1 133 781	403	**2 400**	TO	TO

What is the Output of Parameter Synthesis?

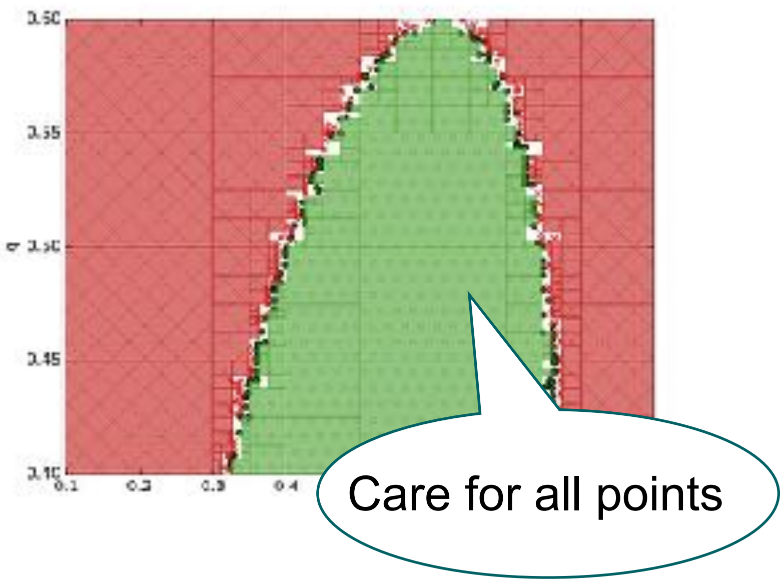
So far, three options have been considered in the literature

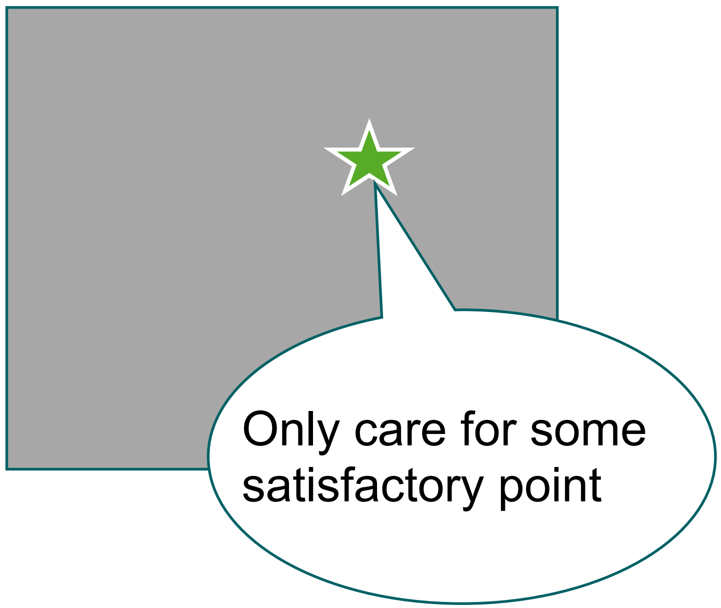
- Option A:
  A generalisation of the output of non-parametric Markov Chain model checking
- Option B:
  A concise description of parameter values that yield satisfactory performance
- Option C:
  One parameter valuation that yields satisfactory performance

Finding satisfactory parameter valuations

A more practical approach

- Methods so far are limited to ~10 parameters.
- Applications as e.g. originating from control under partial observability require thousands of parameters.



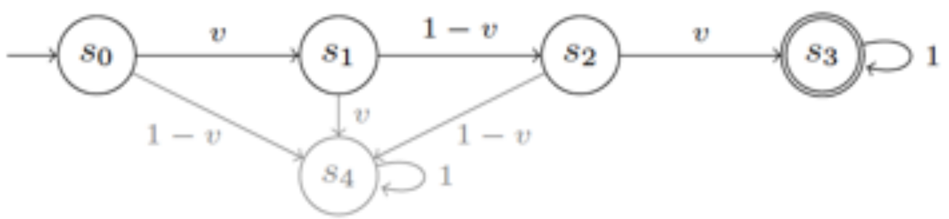


Curse of dimensionality

In more than 100 dimensions, sampling becomes a challenge.

- Sampling via model checking instantiated Markov chains.
- Select sampling points via, e.g., particle swarm optimisation: Fitness via model checking result.
- Neglects much of the underlying structure.

Use (non)linear optimisation?
Convex optimisation problems use structure more



pMCs typically are not convex NLPs

Solution: Linearise concave part, boost with model checking

Empirical results

[ATVA 18]

Problem			Info			PSO			SMT	CCP		
Set	Inst	Spec	States	Trans.	Par.	tmin	tmax	**tavg**	t	t	solv	iter
Brp	16,2	$P_{\leq 0.1}$	98	194	**2**	**0**	0	**0**	40	0	30%	3
Brp	512,5	$P_{<0.1}$	6146	**12290**	**2**	**24**	36	**28**	TO	33	**24%**	3
Crowds	10,5	$P_{<0.1}$	42	82	**2**	4	5	5	8	**4**	2%	4
Nand	5,10	$P_{<0.05}$	10492	20982	**2**	21	51	28	TO	**22**	21%	2
Zeroconf	10000	$E_{\leq 10010}$	10003	20004	**2**	2	4	**3**	TO	57	81%	3
GridA	4	$P_{>0.84}$	1026	2098	**72**	11	11	**11**	TO	22	81%	11
GridB	8,5	$P_{\geq 0.84}$	8653	17369	**700**	409	440	427	TO	**213**	84%	8
GridB	10,6	$P_{\geq 0.84}$	16941	33958	**1290**	533	567	553	TO	**426**	84%	7
GridC	6	$E_{\leq 4.8}$	1665	305	**168**	261	274	267	TO	**169**	90%	23
Maze	5	$E_{<14}$	1303	2658	**590**	213	230	219	TO	**67**	89%	8
Maze	5	$E_{<6}$	1303	2658	**590**			TO	TO	**422**	85%	97
Maze	7	$E_{\leq 6}$	2580	5233	**1176**	–	–	TO	TO	**740**	90%	60
Netw	5,2	$E_{<11.5}$	21746	63158	**2420**	312	523	359	TO	**207**	39%	3
Netw	5,2	$E_{\leq 10.5}$	21746	63158	**2420**	–	–	TO	TO	**210**	38%	4
Netw	4,3	$E_{<11.5}$	38055	97335	**4545**	–	–	TO	TO	MO	-	-
Repud	8,5	$P_{>0.1}$	1487	3002	**360**	16	22	18	TO	**4**	36%	2
Repud	8,5	$P_{\leq 0.05}$	1487	3002	**360**	273	324	293	TO	**14**	72%	4
Repud	16,2	$P_{<0.01}$	790	1606	**96**	–	–	TO	TO	**15**	78%	9
Repud	16,2	$P_{\geq 0.062}$	790	1606	**96**	–	–	TO	TO	TO	-	-