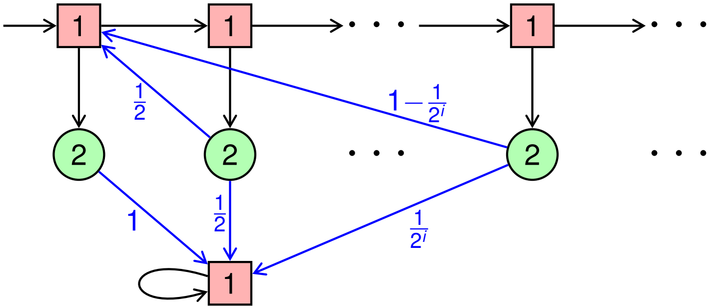
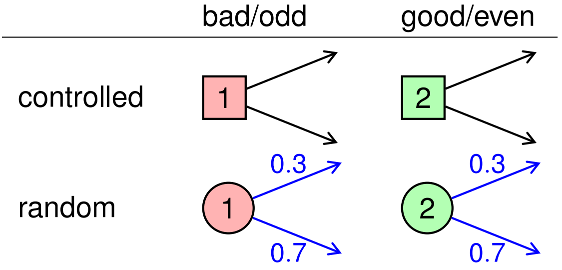
Parity Objectives in Countable MDPs

Stefan Kiefer Richard Mayr
Mahsa Shirmohammadi Dominik Wojtczak

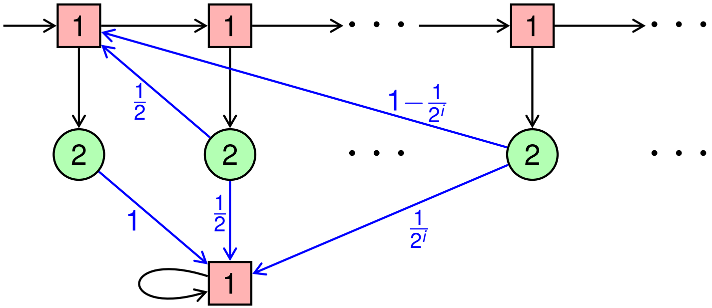
LICS 2017, Reykjavik
20 June 2017

Countable MDPs





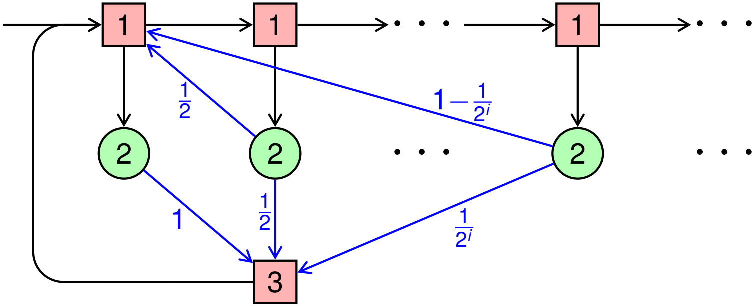
Countable MDPs



There is no almost-surely winning strategy.

$\sup_{\sigma} \Pr_{\sigma}(\text{Parity}) = 1$

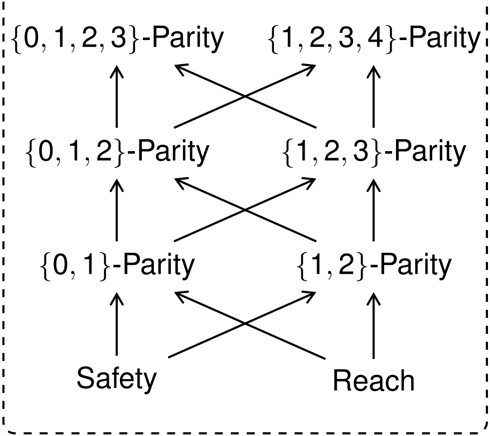
All finite-memory strategies lose almost surely.



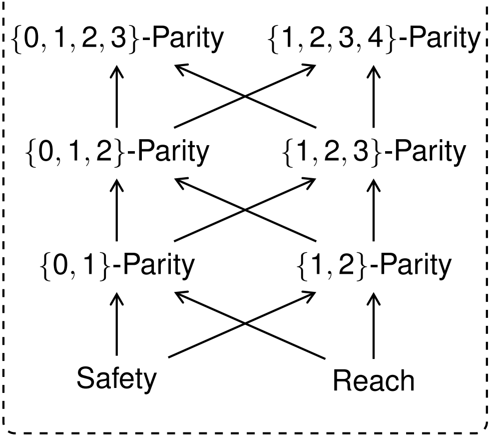
There exists an almost-surely winning strategy.

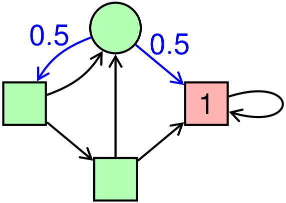
All finite-memory strategies lose almost surely.

Our Results in the Mostowski Hierarchy



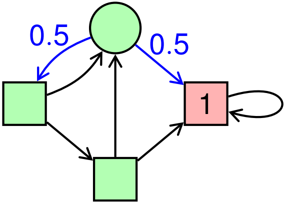
Our Results in the Mostowski Hierarchy



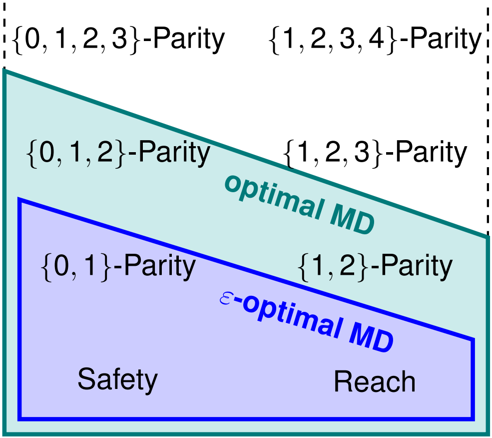


Our Results in the Mostowski Hierarchy

$\{0,1,2,3\}$-Parity	$\{1,2,3,4\}$-Parity
$\{0,1,2\}$-Parity	$\{1,2,3\}$-Parity
$\{0,1\}$-Parity	$\{1,2\}$-Parity
Safety	Reach

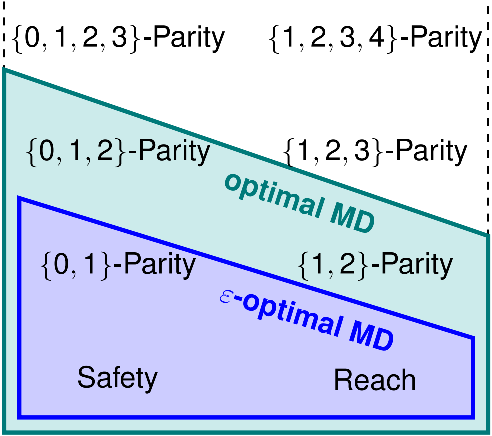


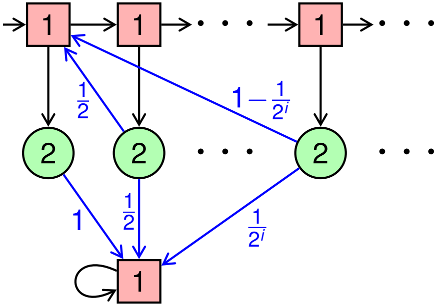
Our Results in the Mostowski Hierarchy



ε-optimal MD means: $\sup_{\sigma} \Pr_{\sigma}(\text{Parity}) = \sup_{\text{MD } \sigma} \Pr_{\sigma}(\text{Parity})$

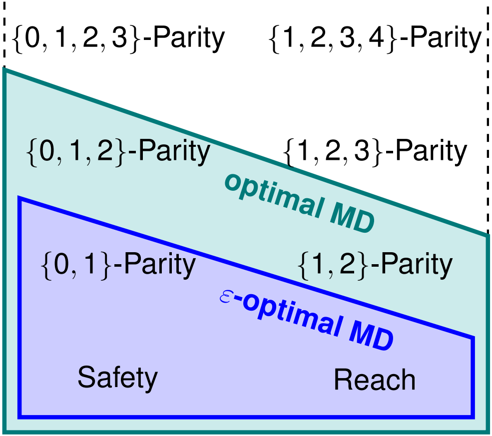
Our Results in the Mostowski Hierarchy

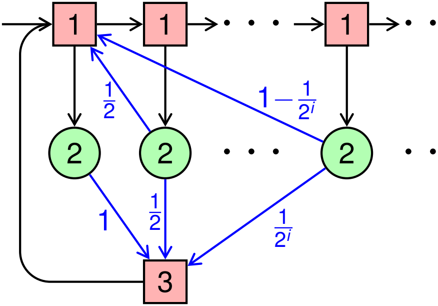




ε-optimal MD means: $\sup_{\sigma} \Pr_{\sigma}(\text{Parity}) = \sup_{\text{MD } \sigma} \Pr_{\sigma}(\text{Parity})$

Our Results in the Mostowski Hierarchy

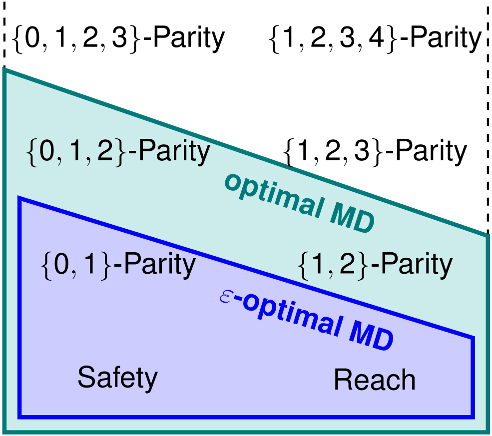


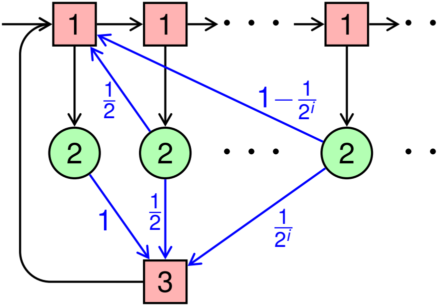


ε-optimal MD means: $\sup_{\sigma} \Pr_{\sigma}(\text{Parity}) = \sup_{\text{MD } \sigma} \Pr_{\sigma}(\text{Parity})$

optimal MD means: if $\exists$ optimal σ, then $\exists$ optimal σ that is MD

Our Results in the Mostowski Hierarchy





ε-optimal MD means: $\sup_{\sigma} \Pr_{\sigma}(\text{Parity}) = \sup_{\text{MD } \sigma} \Pr_{\sigma}(\text{Parity})$

optimal MD means: if $\exists$ optimal σ, then $\exists$ optimal σ that is MD

Dichotomy between MD and infinite memory; contrast to finite MDPs

Optimal MD-Strategies

Theorem

Consider a countable-state MDP with $\{0,1,2\}$-parity objective.

If there exists an optimal strategy,
then there exists an optimal strategy that is MD.

"Optimal strategies for $\{0,1,2\}$-parity may be chosen MD."

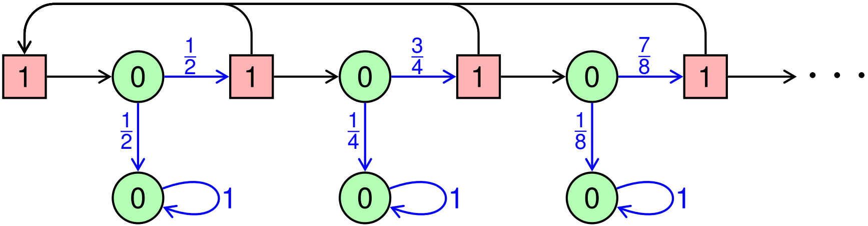
Theorem

Almost-surely winning strategies for co-Büchi may be chosen MD.

Suppose there is an almost-surely winning strategy σ.

Focus on states used by σ. They all have an a.s. winning strategy.

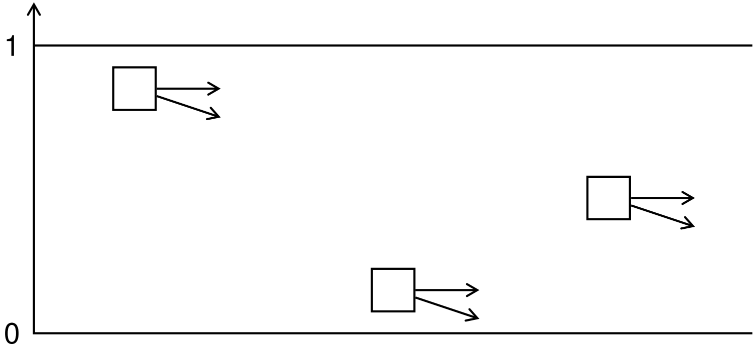
Set a more ambitious goal: Safety (= never see 1 again)



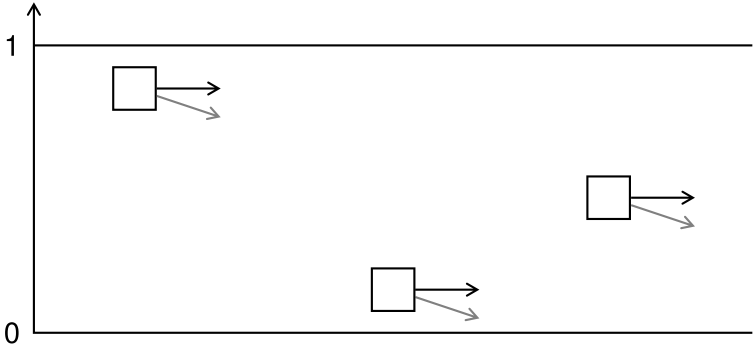
Always playing for safety is too greedy.

An Optimal MD-Strategy for Co-Büchi

$\max_{\sigma} \Pr_{\sigma}$ (never see [1] or (1) again)



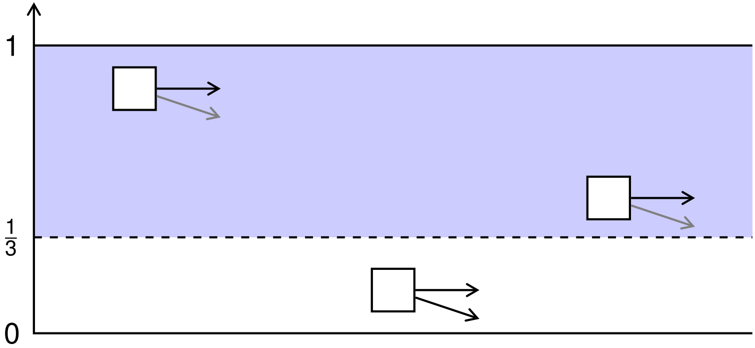
$\max_\sigma \Pr_\sigma$(never see 1 or 1 again)



0. Playing the safest action everywhere is not ok.

An Optimal MD-Strategy for Co-Büchi

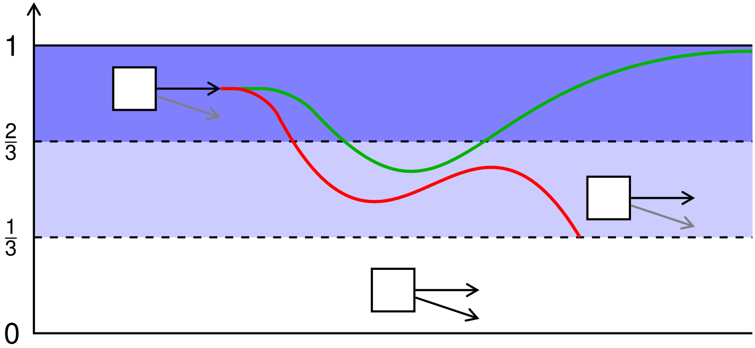
$\max_\sigma \Pr_\sigma$(never see [1] or (1) again)



0. Playing the safest action everywhere is not ok.
1. Fixing the safest action in the blue region is ok.

An Optimal MD-Strategy for Co-Büchi

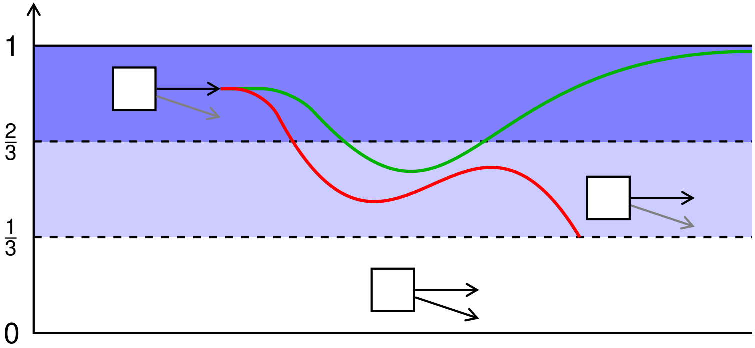
$\max_\sigma \Pr_\sigma$(never see 1 or 1 again)



0. Playing the safest action everywhere is not ok.
1. Fixing the safest action in the blue region is ok.

An Optimal MD-Strategy for Co-Büchi

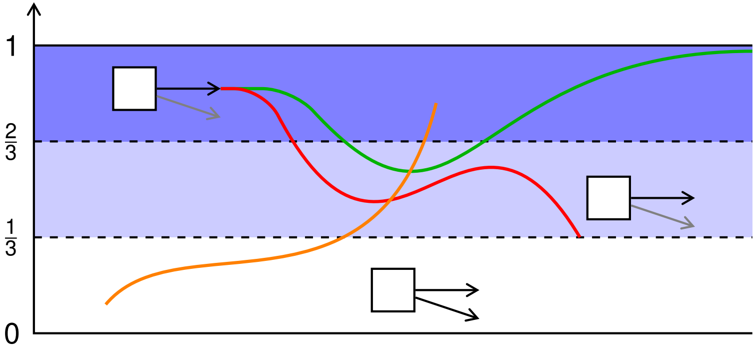
$\max_\sigma \Pr_\sigma$(never see 1 or 1 again)



0. Playing the safest action everywhere is not ok.
1. Fixing the safest action in the blue region is ok.
2. Once we are in dark blue: with prob $\geq \frac{1}{2}$ we stay in blue.

An Optimal MD-Strategy for Co-Büchi

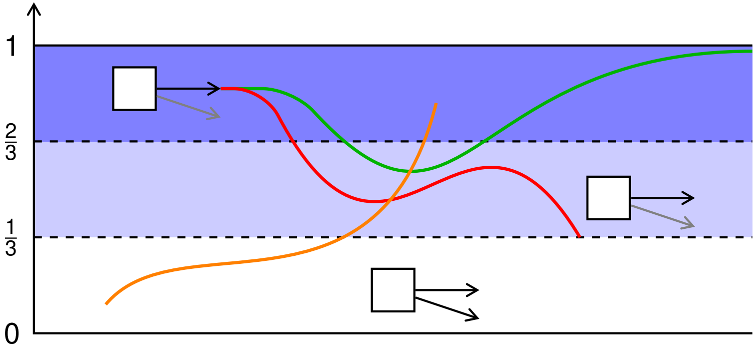
$\max_\sigma \Pr_\sigma$(never see 1 or 1 again)



0. Playing the safest action everywhere is not ok.
1. Fixing the safest action in the blue region is ok.
2. Once we are in dark blue: with prob $\geq \frac{1}{2}$ we stay in blue.

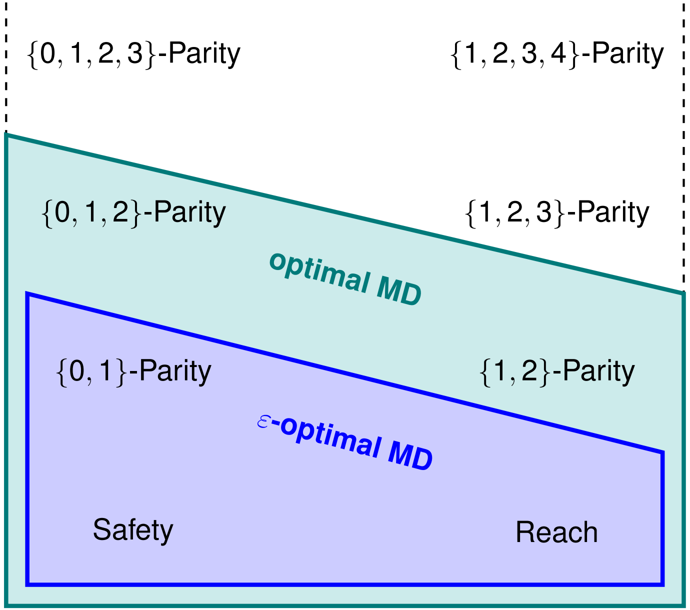
An Optimal MD-Strategy for Co-Büchi

$\max_\sigma \Pr_\sigma$(never see 1 or 1 again)

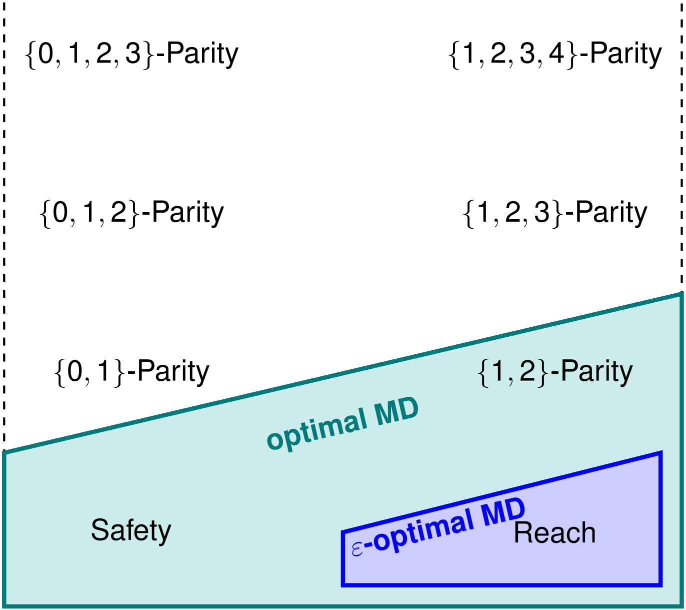


0. Playing the safest action everywhere is not ok.
1. Fixing the safest action in the blue region is ok.
2. Once we are in dark blue: with prob $\geq \frac{1}{2}$ we stay in blue.
3. The a.s. winning strategy for 1. gets us in dark blue a.s.

When MD Suffices For Finitely Branching MDPs



When MD Suffices For Infinitely Branching MDPs



Context of the Paper

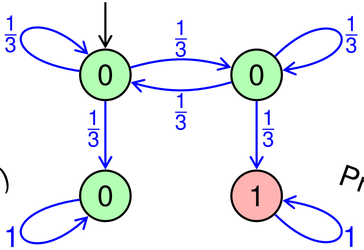
Our work: countable MDPs
Other work: mostly finite MDPs

Our work: maximizing the probability of Parity objectives
Other work: maximizing expected (discounted) total/average reward/cost

Our work: general countable MDPs
Other work: countable MDPs arising from specific models:

- recursive MDPs
- nondeterministic probabilistic lossy channel systems
- VASS-induced MDPs
- one-counter MDPs
- controlled queueing systems
- controlled multitype branching processes
- ...

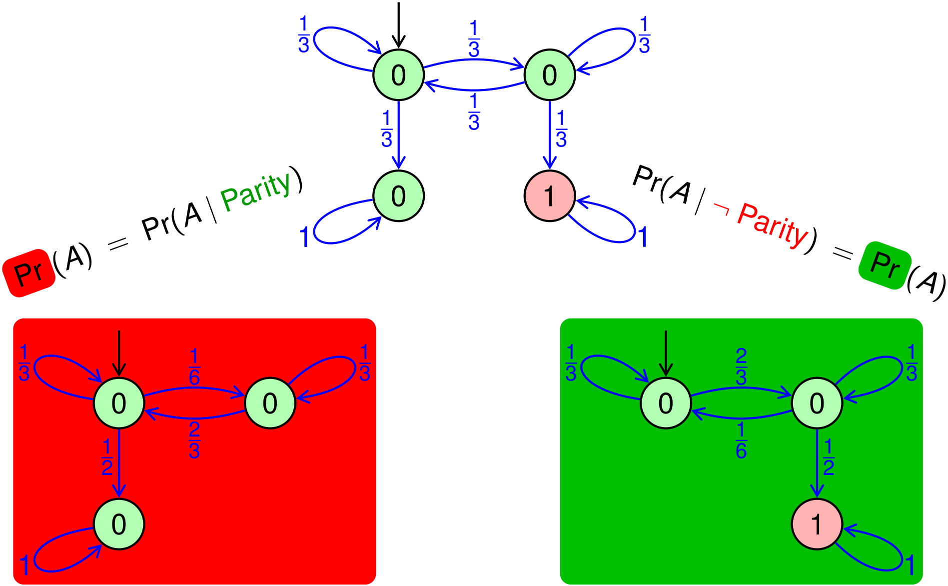
Conditioning a Markov Chain



$\mathrm{Pr}(A) = \mathrm{Pr}(A \mid \text{Parity})$

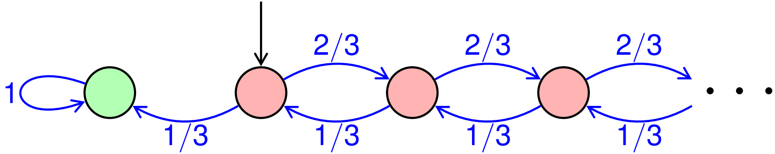
$\mathrm{Pr}(A \mid \neg\,\text{Parity}) = \mathrm{Pr}(A)$

Conditioning a Markov Chain



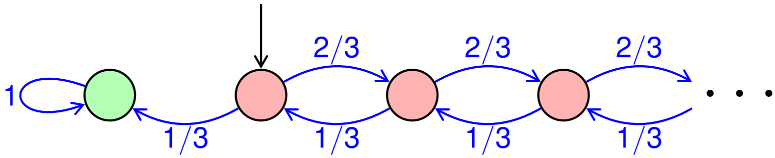
Infinite Markov chains are very different from finite ones.

Gambler's ruin:



Infinite Markov chains are very different from finite ones.

Gambler's ruin:



Dependence on exact probabilities