# Büchi Objectives in Countable MDPs

Stefan Kiefer, Richard Mayr, Patrick Totzke

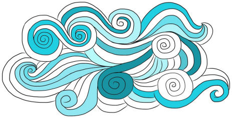
Mahsa Shirmohammadi

Gamenet workshop, April 2022, Maastricht

# Dilemma: Between a rock and a hard place



Following his departure from Circe's island home of Aeaea, Odysseus braces for the many challenges he will encounter on his journey home to his beloved Ithaca ....



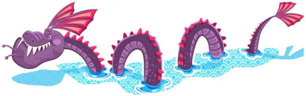
# Dilemma: Between a rock and a hard place



Scylla

strait of Messina

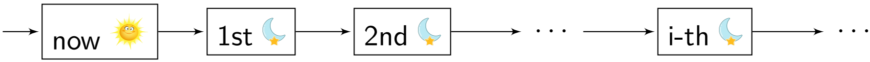
too close
inescapable!



Charybdis

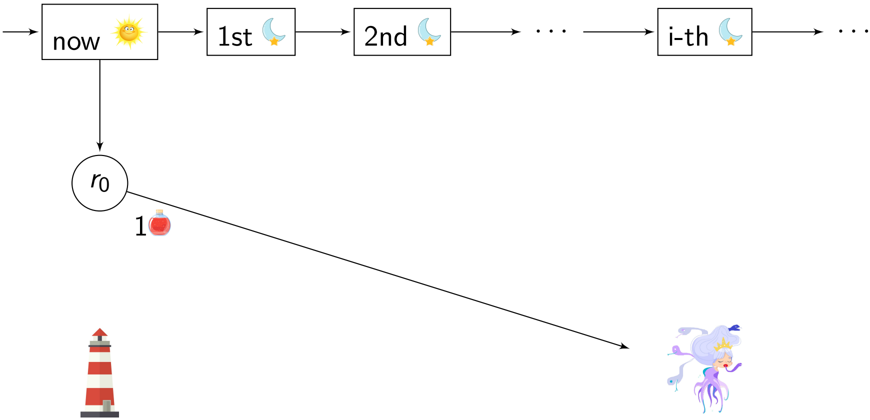
# MDP of Odysseus's dilemma





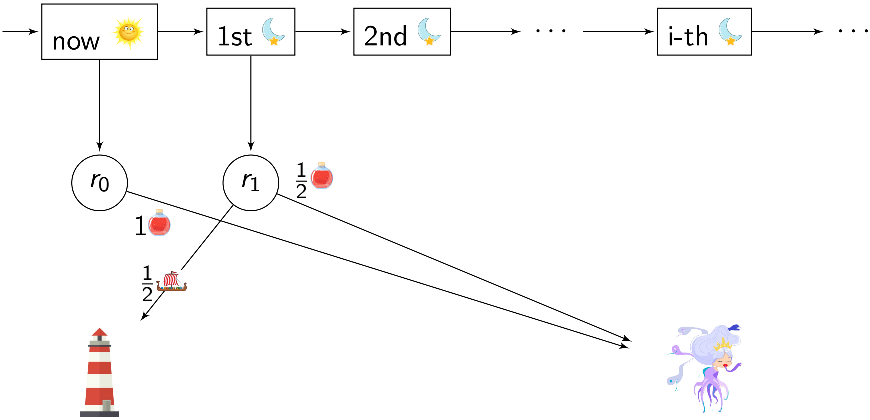

# MDP of Odysseus's dilemma





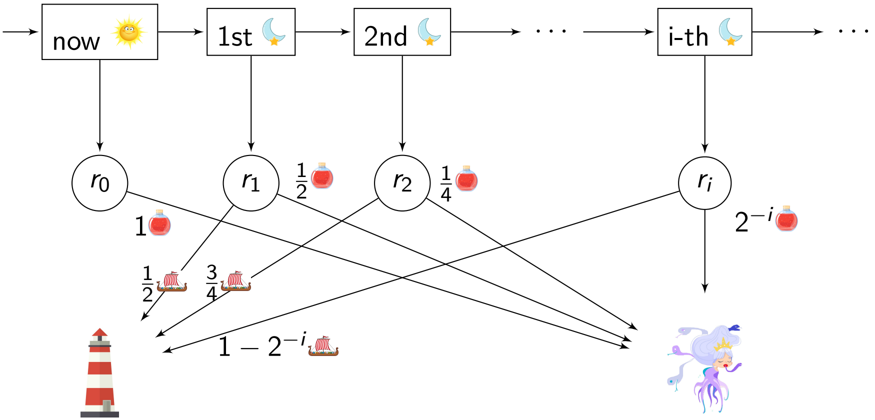
# MDP of Odysseus's dilemma





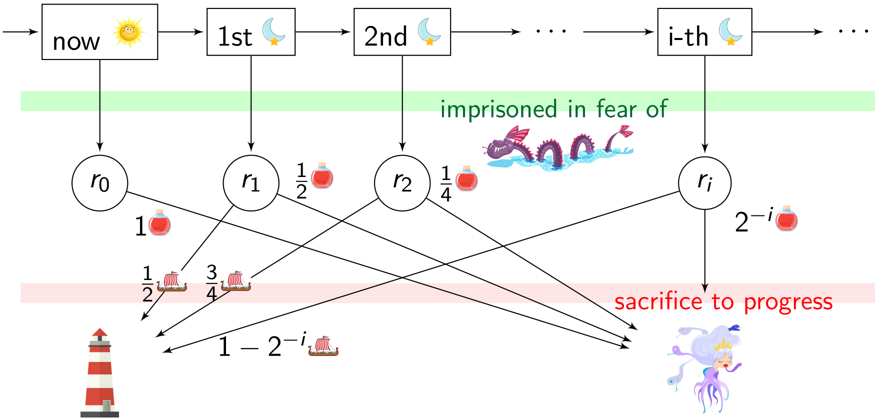
# MDP of Odysseus's dilemma



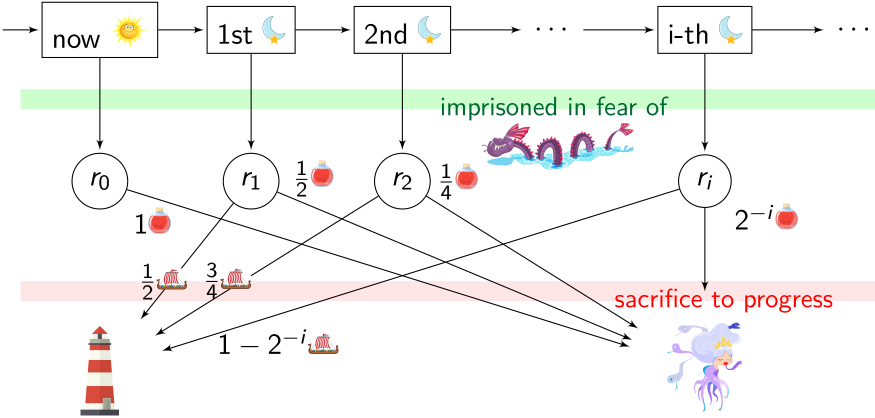


# MDP of Odysseus's dilemma





# Reachability 




The value of Reach() is 1:

for all $\epsilon > 0$,  manages to get his crew back home with $1 - \epsilon$ probability!

# God of Sea

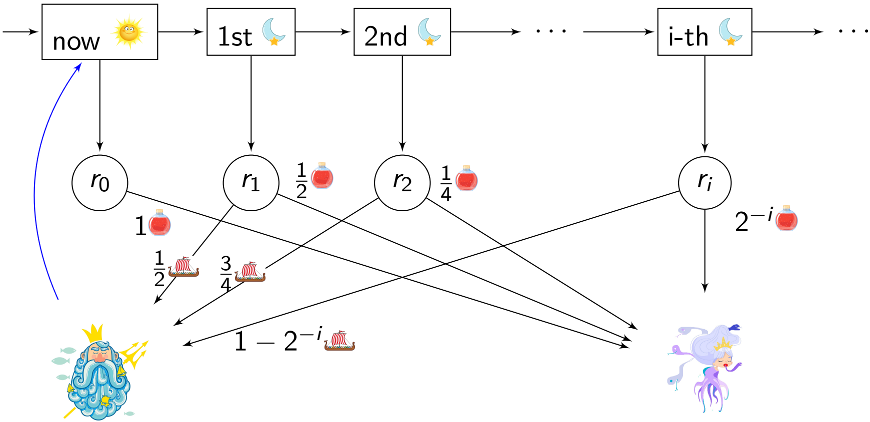


In a recent version:

after passing through Scylla meets

Poseidon  !!! Oops!

... still furious at Odysseus making his son blind.

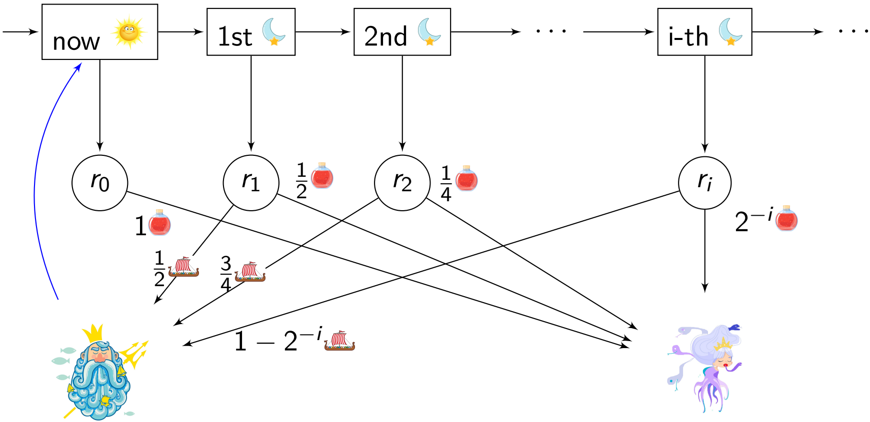



: I make you suffer!

Visit me (i.e., sacrifice to Scylla) over and over!

# Büchi



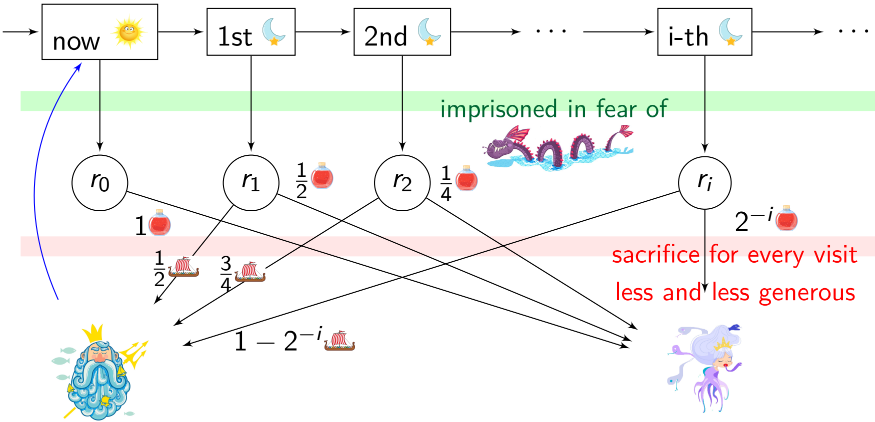



ϟ Is the value of Büchi( ) one?

for all $\epsilon > 0$, can visit $\infty$-times with probability at least $1 - \epsilon$?

# Büchi



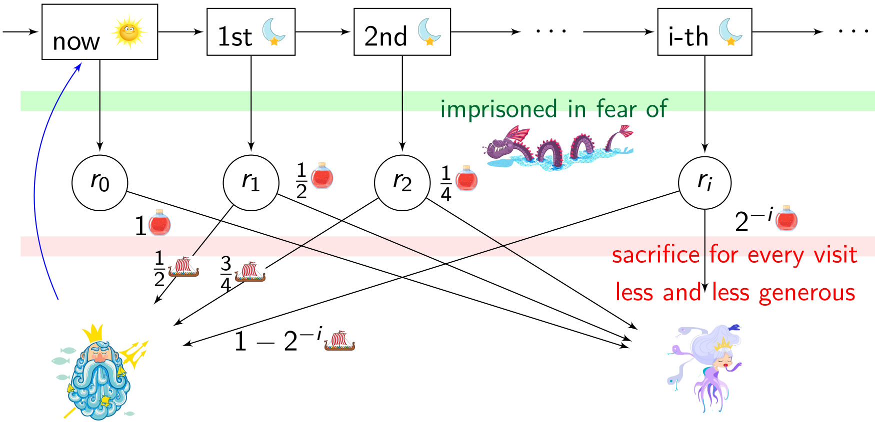



ϟ Is the value of Büchi( ) one? YES!

for all $\epsilon > 0$, can visit $\infty$-times with probability at least $1 - \epsilon$?

# How?

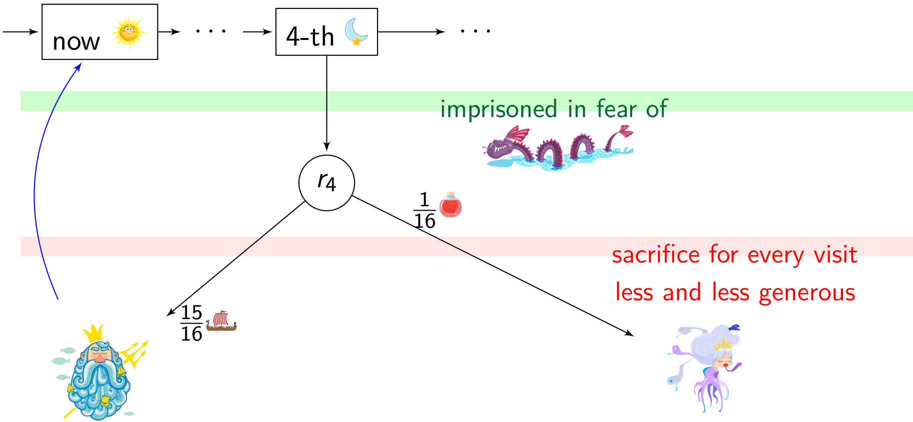





Let $\epsilon = \frac{1}{8}$, let's see how $\frac{7}{8}$ of crew visit  $\infty$-times.

# How? 1st visit



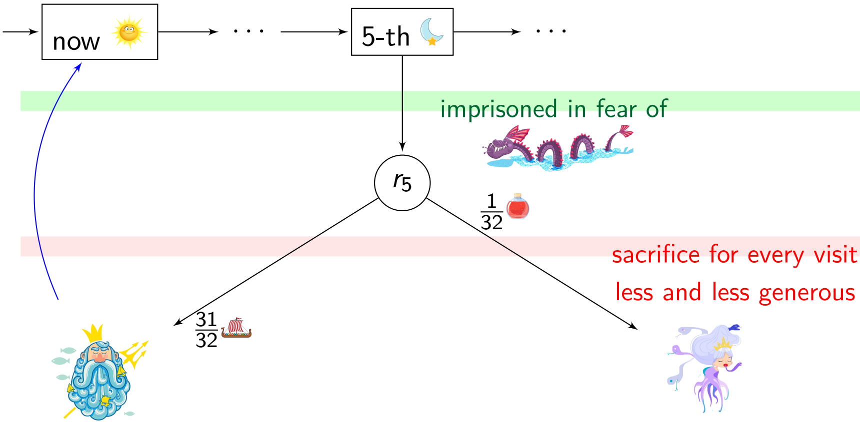



Let $\epsilon = \frac{1}{8}$, let's see how $\frac{7}{8}$ of crew visit ∞-times.

the total sacrifice so far: $\frac{1}{16}$

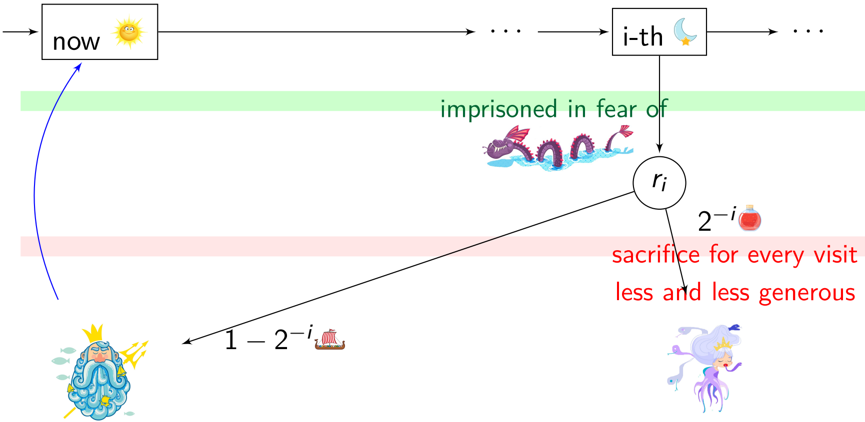
# How? 2nd visit






ϟ Let $\epsilon = \frac{1}{8}$, let's see how $\frac{7}{8}$ of crew visit $\infty$-times.

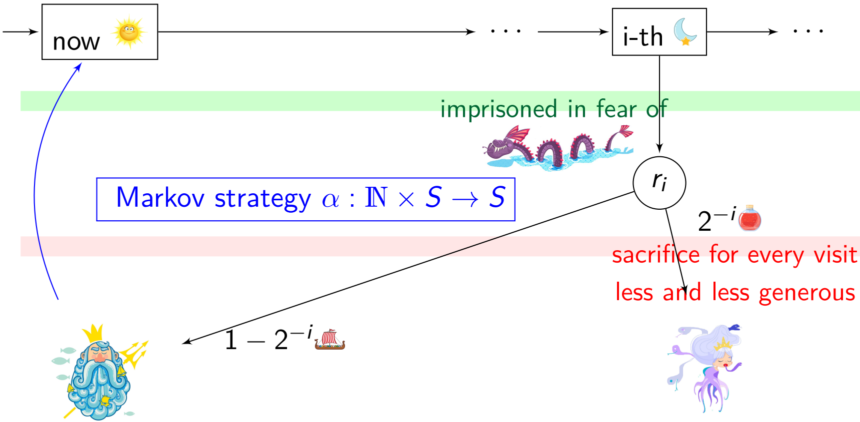
the total sacrifice so far: $\frac{1}{16} + \frac{1}{32}$




ϟ Let $\epsilon = \frac{1}{8}$, let's see how $\frac{7}{8}$ of crew visit  $\infty$-times.

the total sacrifice so far: $\frac{1}{16} + \frac{1}{32} + \cdots + \frac{1}{2^i} + \cdots = \sum_{i=4}^{\infty} \frac{1}{2^i} = \frac{1}{8}$

# How? Markov Strategy 




Let $\epsilon = \frac{1}{8}$, let's see how $\frac{7}{8}$ of crew visit ∞-times.

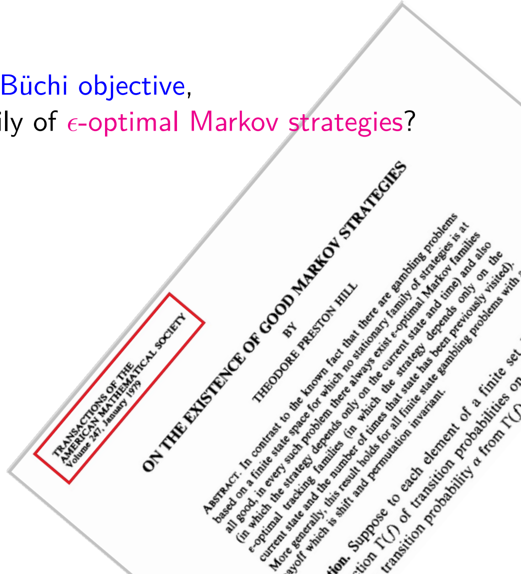
the total sacrifice so far: $\frac{1}{16} + \frac{1}{32} + \cdots + \frac{1}{2^i} + \cdots = \sum_{i=4}^{\infty} \frac{1}{2^i} = \frac{1}{8}$

# Open Problem



For countably infinite MDPs and Büchi objective,
does there always exist a family of $\epsilon$-optimal Markov strategies?

For countably infinite MDPs and Büchi objective,
does there always exist a family of $\epsilon$-optimal Markov strategies?


TRANSACTIONS OF THE AMERICAN MATHEMATICAL SOCIETY
Volume 247, January 1979

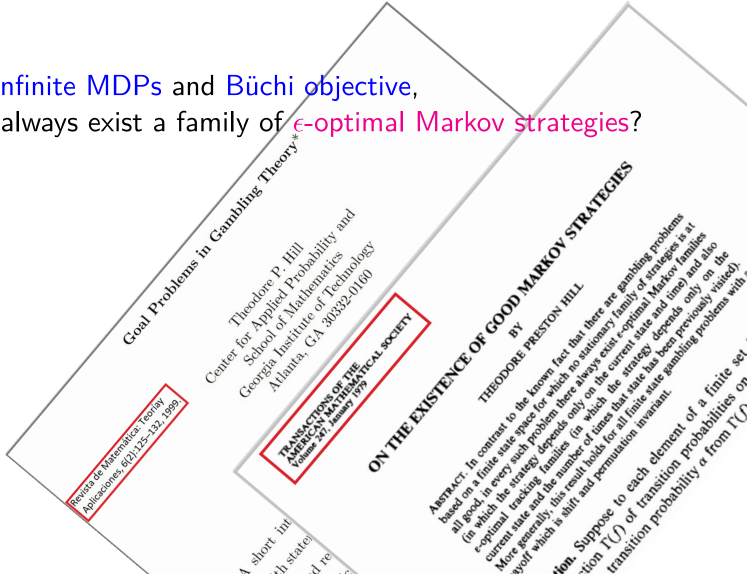
ON THE EXISTENCE OF GOOD MARKOV STRATEGIES

BY

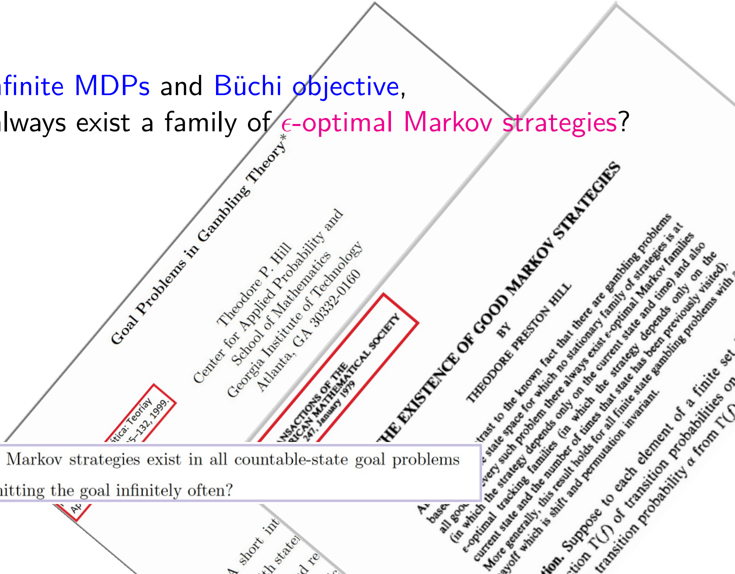
THEODORE PRESTON HILL

ABSTRACT. In contrast to the known fact that there are gambling problems based on a finite state space for which no stationary family of strategies is at all good, in every such problem there always exist ε-optimal Markov families (in which the strategy depends only on the current state and time) and also ε-optimal tracking families (in which the strategy depends only on the current state and the number of times that state has been previously visited). More generally, this result holds for all finite state gambling problems with a payoff which is shift and permutation invariant.

For countably infinite MDPs and Büchi objective,
does there always exist a family of $\epsilon$-optimal Markov strategies?

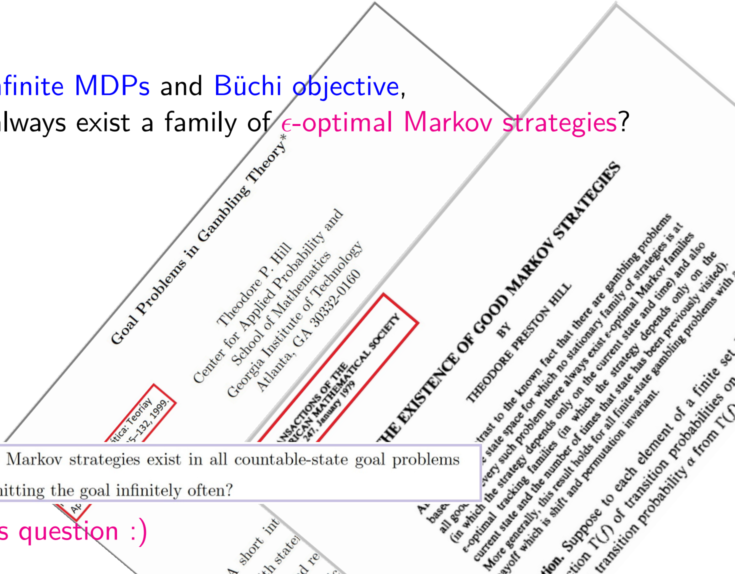


For countably infinite MDPs and Büchi objective,
does there always exist a family of $\epsilon$-optimal Markov strategies?




**Q1.** (cf. [Hi]) Do good Markov strategies exist in all countable-state goal problems with objective of hitting the goal infinitely often?

For countably infinite MDPs and Büchi objective,
does there always exist a family of $\epsilon$-optimal Markov strategies?




**Q1.** (cf. [Hi]) Do good Markov strategies exist in all countable-state goal problems with objective of hitting the goal infinitely often?

↯ we answer this question :)

# Open Problem



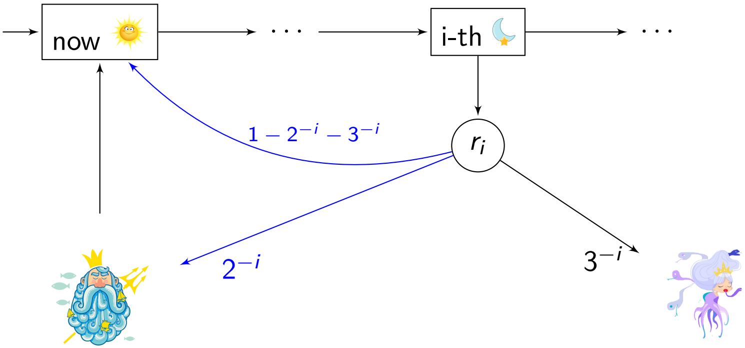
For countably infinite MDPs and Büchi objective,
does there always exist a family of $\epsilon$-optimal Markov strategies?

▷ Is it all about reducing the risk of facing dangerous monsters?



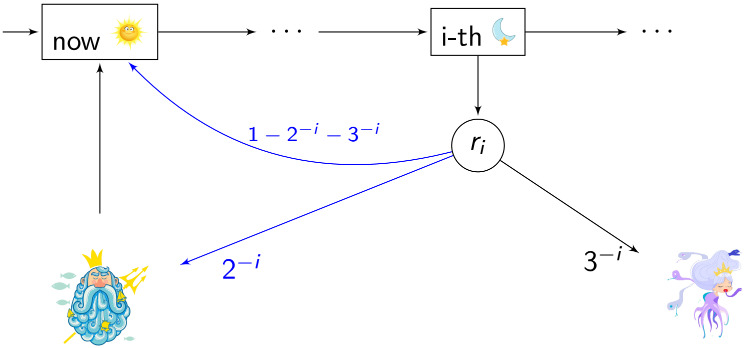
# 2-nd challenge





# 2-nd challenge

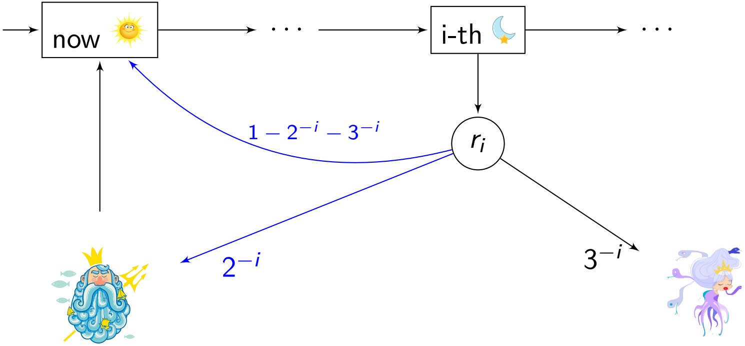





▷ The Markov strategy that, after i-th visit to ☀, picks $r_{i+1}$ attains 0!

# 2-nd challenge

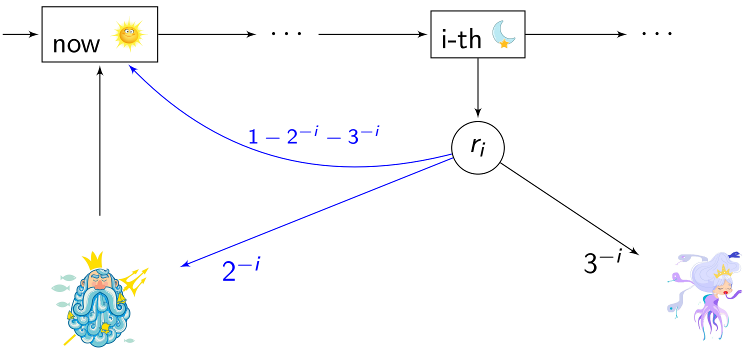





▷ The Markov strategy that, after i-th visit to , picks $r_{i+1}$ attains 0!

the expected number of visits to Poseidon is at most 1

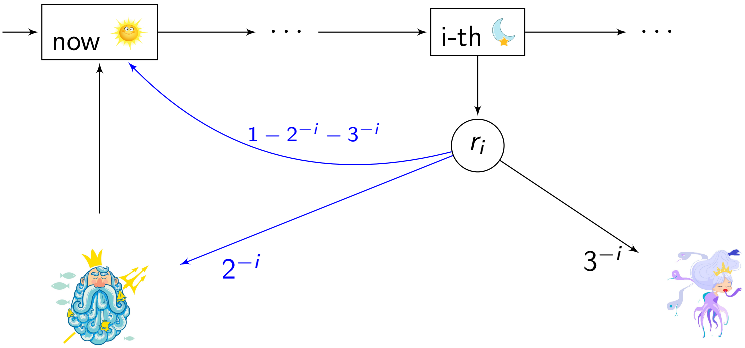
$< \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 1$




- ▷ A strategy that picks each $r_i$ for $2^i$ times achieves Büchi positively!

# 2-nd challenge



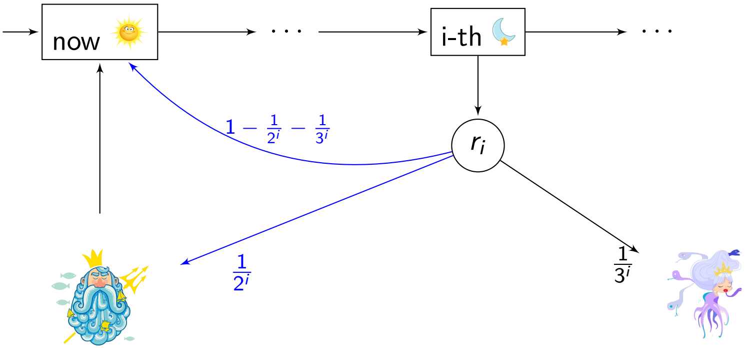



▷ A strategy that picks each $r_i$ for $2^i$ times achieves Büchi positively!

Bound the total sacrifice by $1 - c$ (technical).

The probability of revisit ☀ after each visit $\geq c$

# 2-nd challenge




▷ A strategy that picks each $r_i$ for $2^i$ times achieves Büchi positively!

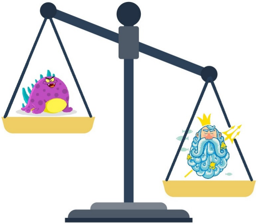
What is the probability to not visiting Poseidon after $i$-th phase (for large $i$)

$$\approx \prod_{k=i}^{\infty} c(1 - \tfrac{1}{2^k})^{2^k} = 0$$

(since $\sum_{k=i}^{\infty} 2^k \log(c(1 - \frac{1}{2^k}))$ is non-convergent)

For countably infinite MDPs and Büchi objective,
does there always exist a family of $\epsilon$-optimal Markov strategies?

- ▷ it is not all about reducing the risk of facing dangerous monsters
- ▷ but rather about a good compromise between progress and loss

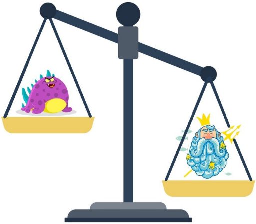


# Counter-example

For countably infinite MDPs and Büchi objective,
does there always exist a family of $\epsilon$-optimal Markov strategies?

NOOOOoOO!

- ▷ it is not all about reducing the risk of facing dangerous monsters
- ▷ but rather about a good compromise between progress and loss



# Counter-example



For countably infinite MDPs and Büchi objective,
does there always exist a family of $\epsilon$-optimal Markov strategies?

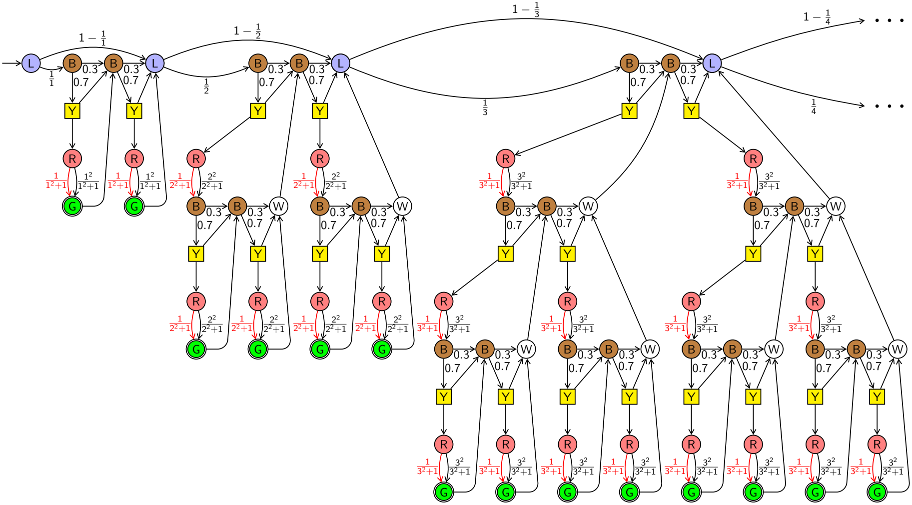
NOOOOoOO!

$\triangleright$ We build an acyclic MDP where $\epsilon$-optimal strategies cannot be Markov.

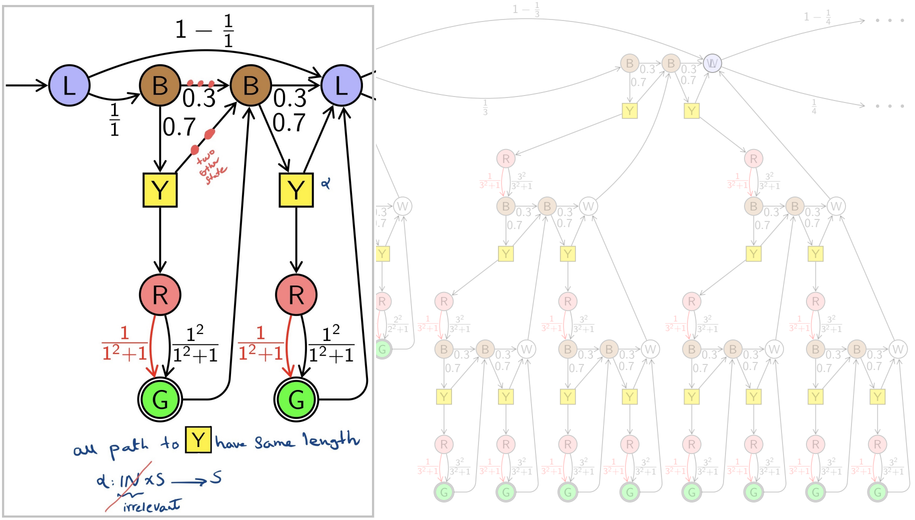
Markov strategy $\alpha : \mathbb{N} \times S \to S$

# Counter-example





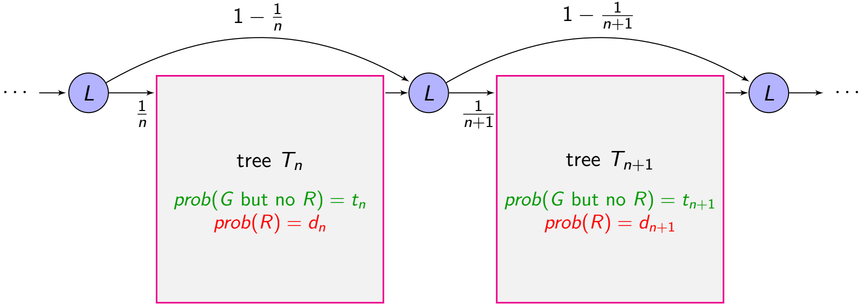
# Counter-example



# Markov not useful



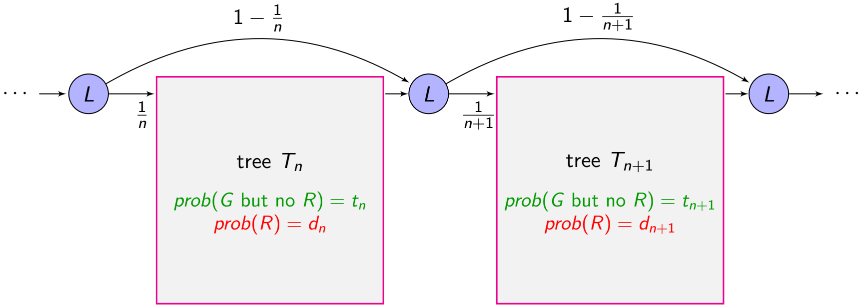
Claim. For Büchi(G) and no R-edge, all Markov strategies attain only 0!



# Markov not useful



Claim. For Büchi(G) and no R-edge, all Markov strategies attain only 0!

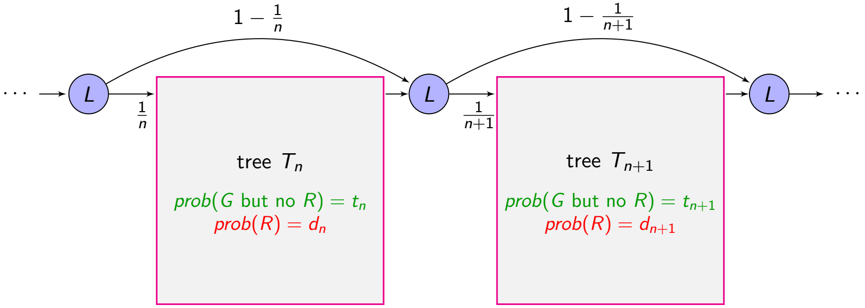



Expected number of visits to $G$ is $\sum \frac{1}{n} t_n$

$\sum \frac{1}{n} t_n$ must be divergent!

# Markov not useful

Claim. For Büchi(G) and no R-edge, all Markov strategies attain only 0!




Expected number of visits to $G$ is $\sum \frac{1}{n} t_n$

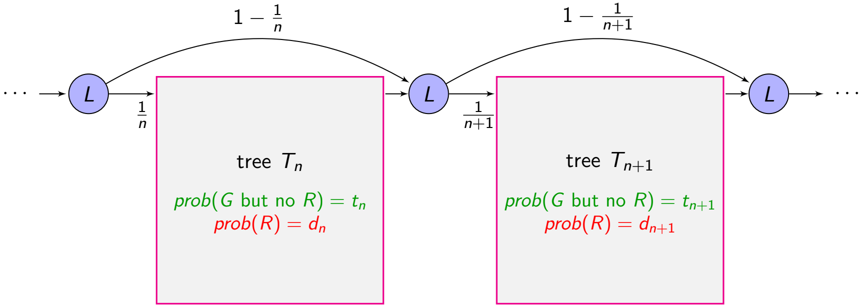
$\sum \frac{1}{n} t_n$ must be divergent!

The probability of R is $\leq \sum \frac{1}{n} d_n$

$\sum \frac{1}{n} d_n$ must be convergent

# Markov not useful

Claim. For Büchi(G) and no R-edge, all Markov strategies attain only 0!




Expected number of visits to $G$ is $\sum \frac{1}{n} t_n$

$\sum \frac{1}{n} t_n$ must be divergent!

The probability of R is $\leq \sum \frac{1}{n} d_n$

$\sum \frac{1}{n} d_n$ must be convergent

By a careful analysis we shows that $d_n \geq 0.008 t_n$ (difficult).

# Counter-example



For countably infinite MDPs and Büchi objective,
does there always exist a family of $\epsilon$-optimal Markov strategies?

NOOOOoOO!

▷ We showed an acyclic MDPs that $\epsilon$-optimal strategies cannot be Markov; however, the value of Büchi(G) is 1 (technical).

# Strategy complexity



For countably infinite MDPs and Büchi objective,
does there always exist a family of $\epsilon$-optimal Markov strategies?

NOOOOoOO!

Theorem. For Büchi, there are always $\epsilon$-optimal 1-bit Markov strategies.

$$\alpha : \mathbb{N} \times S \times \{0,1\} \to S \qquad \text{(necessary and sufficient)}$$

# Use 1-bit of memory? 

- For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies

1-bit strategy $\alpha : S \times \{$  ,  $\} \to S$

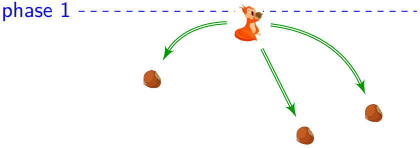
# Use 1-bit of memory?



▷ For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies

1-bit strategy $\alpha : S \times \{$, $\} \to S$

Fix $\epsilon > 0$.



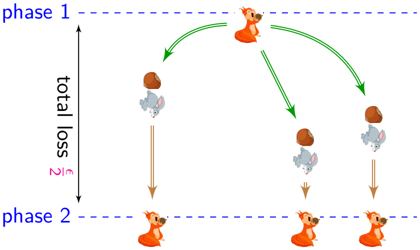

 follows $\frac{\epsilon}{4}$-optimal Reach()

# Use 1-bit of memory? 

▷ For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies

1-bit strategy $\alpha : S \times \{\ ,\ \} \to S$

Fix $\epsilon > 0$.




follows $\frac{\epsilon}{4}$-optimal Reach( )

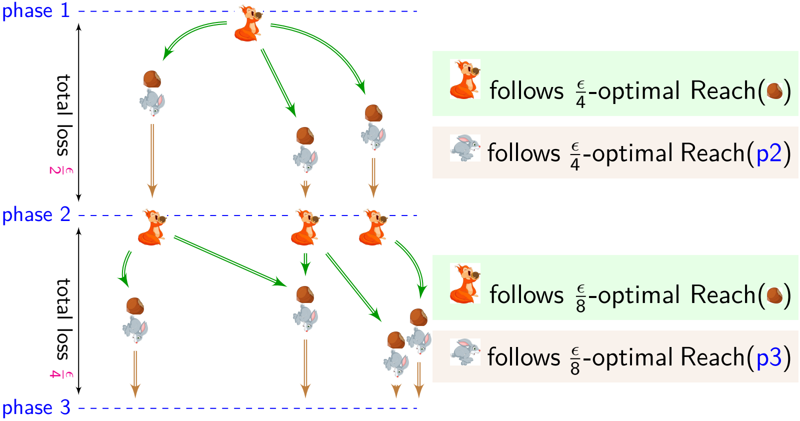
follows $\frac{\epsilon}{4}$-optimal Reach(p2)

# Use 1-bit of memory?

▷ For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies

1-bit strategy $\alpha : S \times \{\ldots\} \to S$

Fix $\epsilon > 0$.

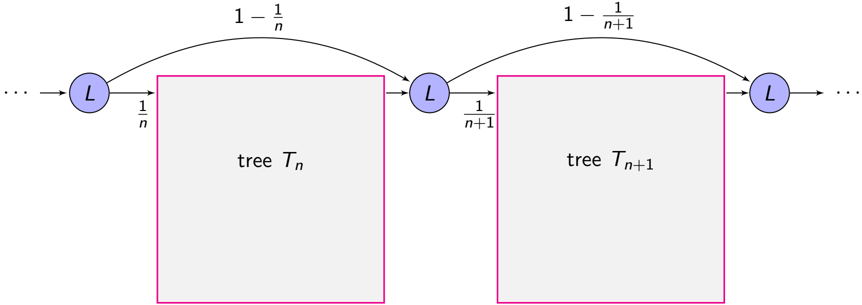


# Use 1-bit of memory?



▷ For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies

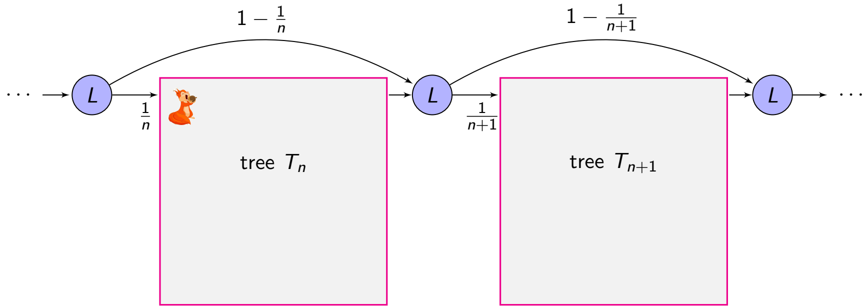
1-bit strategy $\alpha : S \times \{$  ,  $\} \to S$



# Use 1-bit of memory?

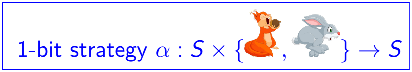
▷ For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies

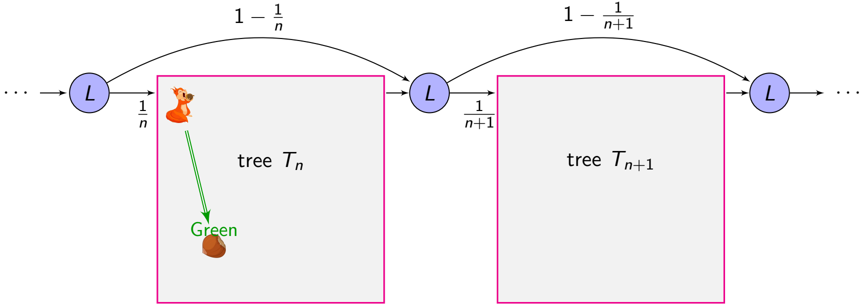
1-bit strategy $\alpha : S \times \{$  ,  $\} \to S$



# Use 1-bit of memory?

▷ For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies

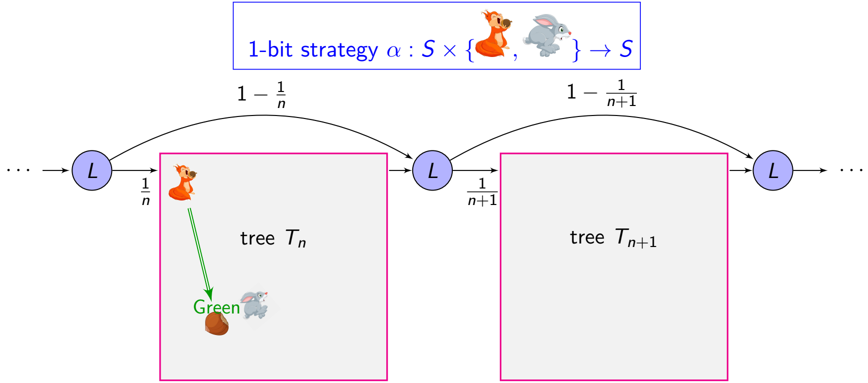




# Use 1-bit of memory?



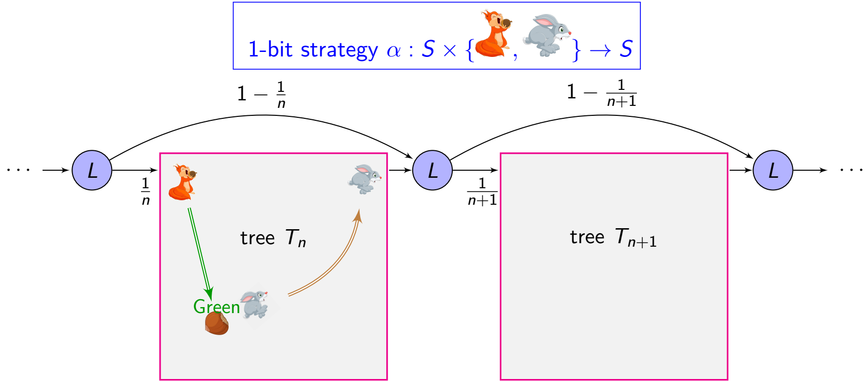
- ▷ For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies



# Use 1-bit of memory?



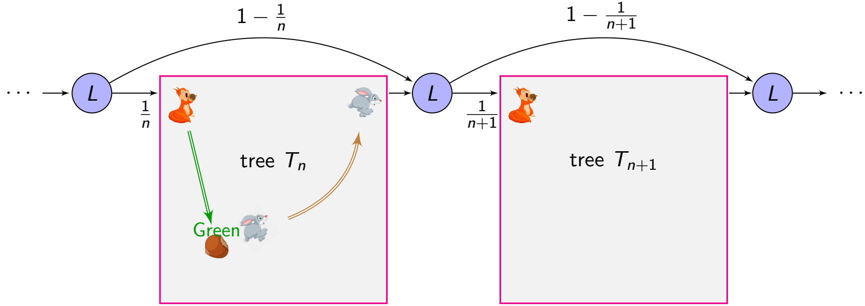
▷ For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies



# Use 1-bit of memory? 

▷ For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies

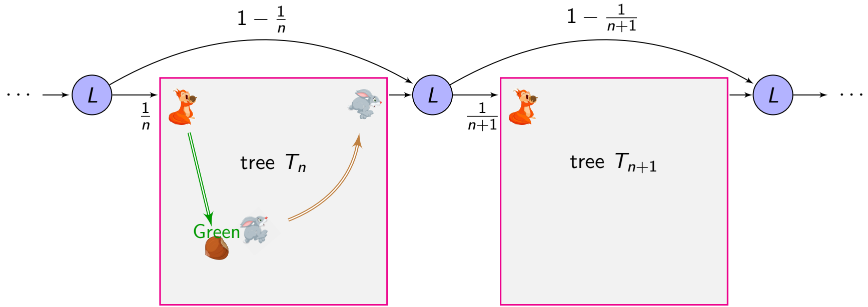
1-bit strategy $\alpha : S \times \{$ 🐿, 🐇 $\} \to S$



# Use 1-bit of memory?

▷ For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies

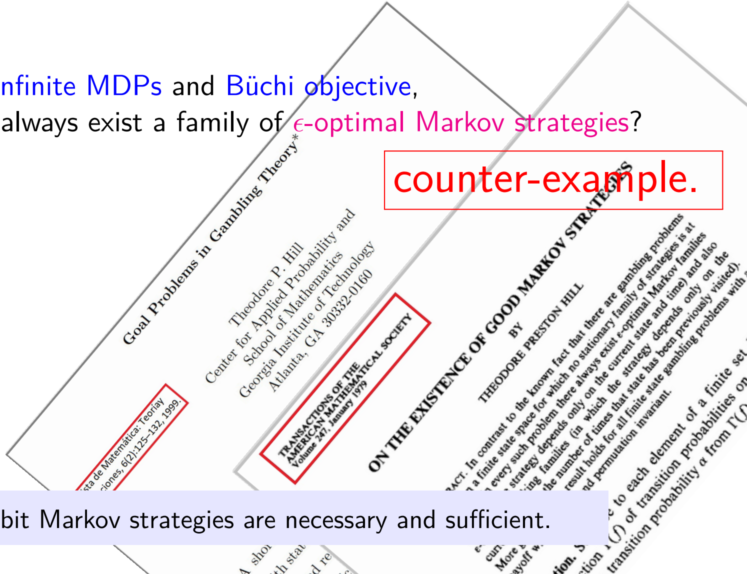
1-bit strategy $\alpha : S \times \{ \; , \; \} \to S$




For all $\epsilon > 0$, a starving-squirrel-and-panic-rabbit strategy achieves $1 - \epsilon$.

# Summary: Strategy complexity *

For countably infinite MDPs and Büchi objective,
does there always exist a family of $\epsilon$-optimal Markov strategies?




counter-example.

**Theorem.** 1-bit Markov strategies are necessary and sufficient.