Büchi Objectives in Countable MDPs

Stefan Kiefer, Richard Mayr, Patrick Totzke

Mahsa Shirmohammadi

Gamenet workshop, April 2022, Maastricht
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

now → 1st → 2nd → ⋯ → i-th → ⋯
MDP of Odysseus’s dilemma
MDP of Odysseus's dilemma
MDP of Odysseus’s dilemma

\[
\begin{align*}
&\text{now} & \rightarrow & \text{1st} & \rightarrow & \text{2nd} & \rightarrow & \cdots & \rightarrow & \text{i-th} & \rightarrow & \cdots \\
& r_0 & \rightarrow & r_1 & \rightarrow & r_2 & \rightarrow & \cdots & \rightarrow & r_i & \rightarrow & \cdots \\
& 1 & \rightarrow & \frac{1}{2} & \rightarrow & \frac{1}{4} & \rightarrow & \cdots & \rightarrow & 2^{-i} & \rightarrow & \cdots \\
& 1 - 2^{-i} & \rightarrow & \frac{1}{2} & \rightarrow & \frac{3}{4} & \rightarrow & \cdots & \rightarrow & \text{1 - 2^{-i}} & \rightarrow & \cdots \\
& \text{lighthouse} & \rightarrow & \text{ship} & \rightarrow & \text{ship} & \rightarrow & \cdots & \rightarrow & \text{ship} & \rightarrow & \cdots \\
& \text{Odysseus} & \rightarrow & \text{Odysseus} & \rightarrow & \text{Odysseus} & \rightarrow & \cdots & \rightarrow & \text{Odysseus} & \rightarrow & \cdots \\
\end{align*}
\]
MDP of Odysseus’s dilemma

now → 1st → 2nd → ⋯ → i-th → ⋯

- $r_0$ to $r_1$: $\frac{1}{2}$
- $r_1$ to $r_2$: $\frac{1}{4}$
- $r_i$ to $r_i$: $2^{-i}$
- $r_0$ to $r_i$: $\frac{1}{2}$
- $r_1$ to $r_i$: $\frac{3}{4}$

imprisoned in fear of sacrifice to progress
The value of \( \text{Reach}(x) \) is 1:

for all \( \epsilon > 0 \), manages to get his crew back home with \( 1 - \epsilon \) probability!
In a recent version:

after passing through Scylla meets Poseidon

!!! Oops!

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

```
now → 1st → 2nd → ... → i-th → ...

r0

r1 \[\frac{1}{2}\]

r2 \[\frac{1}{4}\]

ri \[2^{-i}\]

\[\frac{1}{2}\] \[\frac{3}{4}\]

1 - 2^{-i}

: I make you suffer!

Visit me (i.e., sacrifice to Scylla) over and over!
```
Is the value of Büchi(Ω) one?

for all $\epsilon > 0$, can visit $\infty$-times with probability at least $1 - \epsilon$?
Is the value of Büchi(🧙) one? **YES!**

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

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

\[

c_{1} \rightarrow \ldots \rightarrow c_{4} \rightarrow \ldots
\]

\[
	ext{imprisoned in fear of}
\]

\[
	ext{sacrifice for every visit}
\]

\[
\text{less and less generous}
\]

\[
\text{Let } \epsilon = \frac{1}{8}, \text{ let’s see how } \frac{7}{8} \text{ of crew visit } \infty \text{-times.}
\]

\[
\text{the total sacrifice so far: } \frac{1}{16}
\]

\[
f_{4}
\]

\[
\frac{15}{16}
\]

\[
\frac{1}{16}
\]
How? 2nd visit

now \rightarrow \cdots \rightarrow 5\text{-th} \rightarrow \cdots

imprisoned in fear of

sacrifice for every visit
less and less generous

\[ \frac{1}{32} \]

\[ \frac{31}{32} \]

\[ \frac{1}{16} + \frac{1}{32} \]

Let \( \epsilon = \frac{1}{8} \), let’s see how \( \frac{7}{8} \) of crew visit \( \star \) \( \infty \)-times.

the total sacrifice so far: \( \frac{1}{16} + \frac{1}{32} \)
Let $\epsilon = \frac{1}{8}$, let’s see how $\frac{7}{8}$ of crew visit $\odot$ $\infty$-times. The total sacrifice so far: $\frac{1}{16} + \frac{1}{32} + \cdots + \frac{1}{2^i} + \cdots = \sum_{i=4}^{\infty} \frac{1}{2^i} = \frac{1}{8}$
Let $\epsilon = \frac{1}{8}$, let’s see how $\frac{7}{8}$ of the crew visit $\diamondsuit \infty$-times.

The total sacrifice so far:

$$\frac{1}{16} + \frac{1}{32} + \cdots + \frac{1}{2^i} + \cdots = \sum_{i=4}^{\infty} \frac{1}{2^i} = \frac{1}{8}$$
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?
For countably infinite MDPs and Büchi objective, does there always exist a family of $\epsilon$-optimal Markov strategies?
Open Problem

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?
Open Problem

For countably infinite MDPs and Büchi objective, does there always exist a family of $\varepsilon$-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 :)
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

\[ r_i = 1 - 2^{-i} - 3^{-i} \]
The Markov strategy that, after i-th visit to \(
\star
\), picks \( r_{i+1} \) attains 0!
The Markov strategy that, after i-th visit to \( \star \), 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!
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 $\star$ 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 $i$)

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

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

For countably infinite MDPs and Büchi objective, does there always exist a family of $\varepsilon$-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
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

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

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

Markov strategy $\alpha : \mathbb{N} \times S \rightarrow S$
Counter-example
**Claim.** For Büchi($G$) and no $R$-edge, all Markov strategies attain only 0!

\[
\text{tree } T_n \\
\text{prob}(G \text{ but no } R) = t_n \\
\text{prob}(R) = d_n
\]

\[
\text{tree } T_{n+1} \\
\text{prob}(G \text{ but no } R) = t_{n+1} \\
\text{prob}(R) = d_{n+1}
\]
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!
Claim. For Büchi(G) and no R-edge, all Markov strategies attain only 0!

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

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

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

$\sum \frac{1}{n} d_n$ must be convergent
Claim. For Büchi(G) and no R-edge, all Markov strategies attain only 0!

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

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

By a careful analysis we shows that \( d_n \geq 0.008 t_n \) (difficult).
For countably infinite MDPs and Büchi objective, does there always exist a family of $\epsilon$-optimal Markov strategies?

We showed an acyclic MDPs that $\epsilon$-optimal strategies cannot be Markov; however, the value of $\text{Büchi}(G)$ is 1 (technical).
For countably infinite MDPs and Büchi objective, does there always exist a family of $\epsilon$-optimal Markov strategies?

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

$$\alpha : \mathbb{N} \times S \times \{0, 1\} \rightarrow S \quad \text{(necessary and sufficient)}$$
For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies.

1-bit strategy $\alpha : S \times \{\text{fox}, \text{rabbit}\} \rightarrow S$
For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies.

1-bit strategy $\alpha : S \times \{\bullet, \star\} \rightarrow S$

Fix $\epsilon > 0$. Phase 1 follows $\frac{\epsilon}{4}$-optimal Reach($\bullet$).
For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies.

Fix $\epsilon > 0$.

1-bit strategy $\alpha : S \times \{1, 2\} \rightarrow S$

- phase 1
- phase 2

Total loss $\leq \frac{\epsilon}{4}$

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

follows $\frac{\epsilon}{4}$-optimal Reach(p2)
For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies.

Fix $\epsilon > 0$.

1-bit strategy $\alpha : S \times \{\text{Fox}, \text{Rabbit}\} \rightarrow S$

- Phase 1 follows $\frac{\epsilon}{4}$-optimal Reach(\text{Fox})
- Phase 2 follows $\frac{\epsilon}{4}$-optimal Reach(\text{Rabbit})
- Phase 3 follows $\frac{\epsilon}{8}$-optimal Reach(\text{Fox})
- Phase 3 follows $\frac{\epsilon}{8}$-optimal Reach(\text{Rabbit})
For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies.

1-bit strategy $\alpha : S \times \{\, \text{L}, \text{R} \,\} \to S$
For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies.

1-bit strategy $\alpha : S \times \{\text{fox}, \text{rabbit}\} \rightarrow S$

- $1 - \frac{1}{n}$
- $1 - \frac{1}{n+1}$

Tree $T_n$

Tree $T_{n+1}$
For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies.

1-bit strategy $\alpha : S \times \{\text{Red}, \text{Green}\} \rightarrow S$
For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies

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

\[
\begin{align*}
1 - \frac{1}{n} & \quad \text{to} \quad 1 - \frac{1}{n+1} \\
\frac{1}{n} & \quad \text{to} \quad \frac{1}{n+1}
\end{align*}
\]
For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies. A 1-bit strategy $\alpha : S \times \{\text{Red}, \text{Blue}\} \rightarrow S$ is used. The transition probabilities are $1 - \frac{1}{n}$ and $1 - \frac{1}{n+1}$. The diagram illustrates the process with trees $T_n$ and $T_{n+1}$, and the label 'Green' on the edge leading to $T_n$. The transitions are shown from left to right with probabilities $\frac{1}{n}$ and $\frac{1}{n+1}$. The diagram is a visual representation of the strategy and the transition probabilities.
For Büchi and acyclic MDPs, there are always \( \epsilon \)-optimal 1-bit strategies.

1-bit strategy \( \alpha : S \times \{0, 1\} \rightarrow S \)

\[
1 - \frac{1}{n} \quad \text{tree } T_n
\]

\[
1 - \frac{1}{n+1} \quad \text{tree } T_{n+1}
\]
For Büchi and acyclic MDPs, there are always $\epsilon$-optimal 1-bit strategies.

A 1-bit strategy $\alpha : S \times \{\text{left}, \text{right}\} \to S$ is defined.

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