Taming Infinity One Chunk at a Time: Concisely Represented Strategies in One-Counter MDPs

Michal Ajdarów[2] **James C. A. Main**[1]
Petr Novotný[2] Mickael Randour[1]

[1]F.R.S.-FNRS and UMONS – Université de Mons, Belgium

[2]Masaryk University, Brno, Czech Republic

ICALP 2025 – July 8, 2025

Talk overview

- We study **synthesis** and **verification** in **Markov decision processes**.
- Small controllers, i.e., **concisely represented strategies** are desirable for practical synthesis.
- In **infinite-state systems**, such representations are necessary.

Contributions

We study **one-counter Markov decision processes** (OC-MDPs).

1. We focus on **interval strategies**, a subclass of strategies that make decisions based on the current state and counter value.
2. We provide **PSPACE** algorithms for the verification and synthesis of interval strategies.

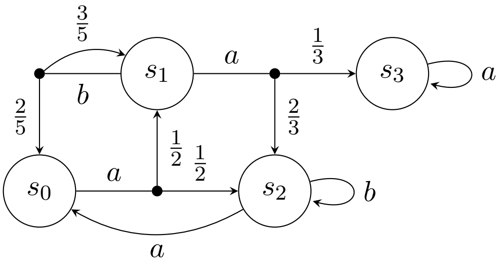
Table of contents

Table of contents

Markov decision processes

Markov decision process (MDP) $\mathcal{M}$

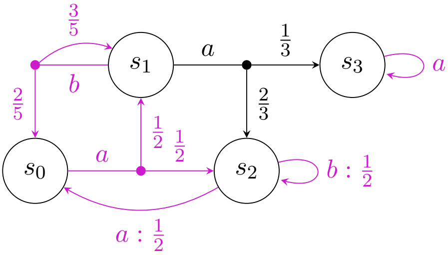
- **Finite or countable** state space S.
- **Finite** action space A.
- **Randomised** transition function $\delta\colon S \times A \to \mathcal{D}(S)$.



Plays are sequences in $(SA)^\omega$ coherent with transitions.
$\rightsquigarrow$ **Example**: $s_0 a s_1 b s_1 \dots$

Strategies and induced Markov chains

- A **strategy** is a function $\sigma\colon (SA)^*S \to \mathcal{D}(A)$.
- σ is **memoryless** if its choices depend only on the **current state**.
- We view memoryless strategies as functions $S \to \mathcal{D}(A)$.
- A memoryless strategy σ induces a **Markov chain** over S.



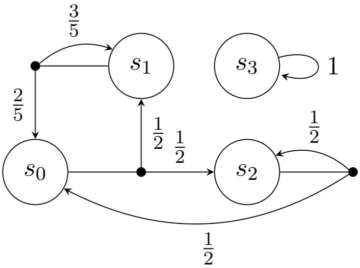


Table of contents

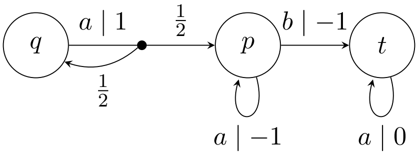
One-counter Markov decision processes

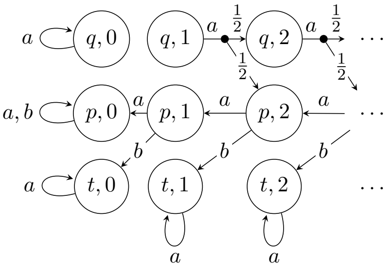
One-counter MDP (OC-MDP) $\mathcal{Q}$

- **Finite** MDP (Q, A, δ).
- **Weight function** $w \colon Q \times A \to \{-1, 0, 1\}$.

MDP $\mathcal{M}^{\leq\infty}(\mathcal{Q})$ induced by $\mathcal{Q}$

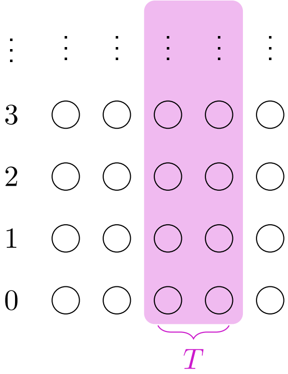
- **Countable** MDP over $S = Q \times \mathbb{N}$.
- State transitions via δ.
- Counter updates via w.



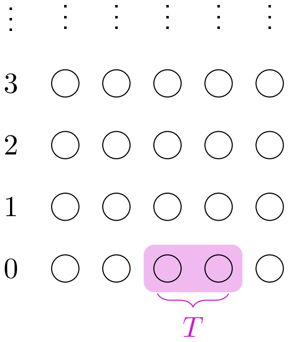


Objectives

- Let $T \subseteq Q$ **be a target**.
- We study variants of **reachability objectives**.



State reachability $\mathsf{Reach}(T)$



Selective termination $\mathsf{Term}(T)$

Interval strategies (1/2)

We study two classes of **memoryless strategies** of $\mathcal{M}^{\leq\infty}(\mathcal{Q})$.

Open-ended interval strategies (OEIS)

σ is an OEIS if $\exists\, k_0 \in \mathbb{N}$ s.t. $\forall\, q \in Q$ and $\forall\, k \geq k_0$, $\sigma(q,k) = \sigma(q,k_0)$.

$\mathbb{N}_0$	1	2	$\cdots$	$k_0 - 1$	k_0	$k_0 + 1$	$\cdots$
Q	σ_1	σ_2	$\cdots$	σ_{k_0-1}	σ_{k_0}	σ_{k_0}	$\cdots$

constant

Group counter values in intervals

Inter.	I_1	I_2	$\cdots$	$I_d = [\![k_0, \infty]\!]$
Q	τ_1	τ_2	$\cdots$	$\tau_d = \sigma_{k_0}$

$\rightsquigarrow$ **Finite partition** of $\mathbb{N}_0$ into **intervals**

Interval strategies (2/2)

We study two classes of **memoryless strategies** of $\mathcal{M}^{\leq\infty}(\mathcal{Q})$.

Cyclic interval strategies (CIS)

σ is a CIS if $\exists\,\rho \in \mathbb{N}_0$ s.t. $\forall\, q \in Q$ and $\forall\, k \in \mathbb{N}$, $\sigma(q, k) = \sigma(q, k + \rho)$.

$\mathbb{N}_0$	1	2	$\cdots$	ρ	$\rho+1$	$\rho+2$	$\cdots$
Q	σ_1	σ_2	$\cdots$	σ_ρ	σ_1	σ_2	$\cdots$

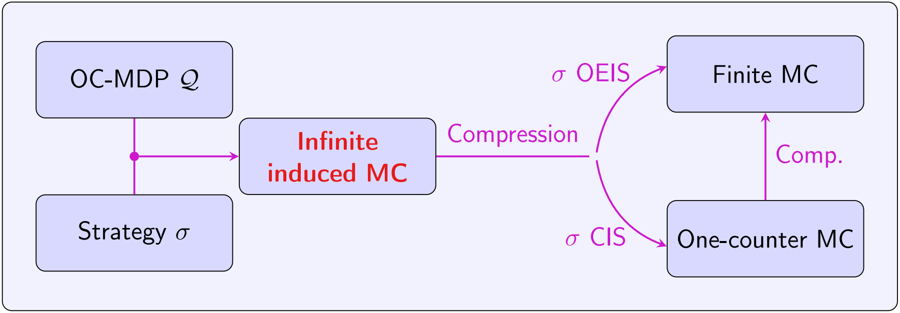
$\rightarrow$ CISs can be represented via an **interval partition** of $[\![1, \rho]\!]$.

Table of contents

Interval strategy verification problem

Interval strategy verification problem

Decide whether $\mathbb{P}^{\sigma}_{\mathcal{M}^{\leq\infty}(\mathcal{Q}),s_{\text{init}}}(\Omega) \geq \theta$ given an **interval strategy** σ, **an objective** $\Omega \in \{\text{Reach}(T), \text{Term}(T)\}$, a **threshold** $\theta \in \mathbb{Q} \cap [0,1]$ and an **initial configuration** $s_{\text{init}} \in Q \times \mathbb{N}$.

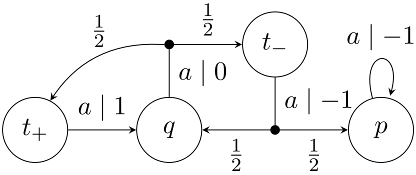


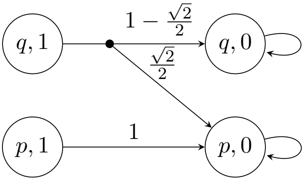
Compressing for an unbounded interval

Compressing for $I = [\![k_0, \infty]\!]$

- We only keep (q, k_0) for all states $q \in Q$.
- The successors of (q, k_0) are of the form $(p, k_0 - 1)$.

Example: σ choosing a **in all states** for the interval $\mathbb{N}_0 = [\![1, \infty]\!]$.





- Transition probabilities can be **irrational**.
- They are the **least solution** of a quadratic **system of equations**.[1]

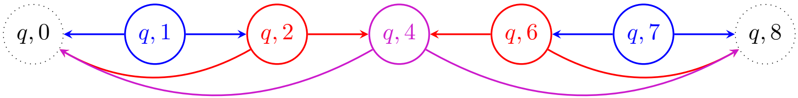
[1] Kucera et al., "Model Checking Probabilistic Pushdown Automata", LMCS 2006.

Compressing bounded intervals

Why should we compress a bounded interval I?

- The bounds of I are encoded in **binary**.
- Thus $Q \times I$ is of **exponential size**.
- We do **not** analyse the compressed Markov chain **on the fly**.

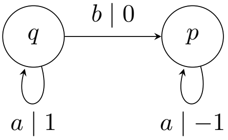
Main idea: using counter jumps that are **powers of two**.

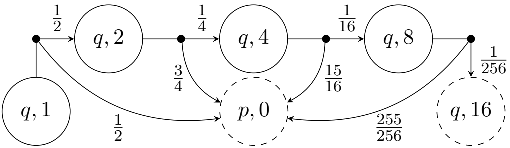


- The construction requires that $|I| = 2^\beta - 1$ for some $\beta \in \mathbb{N}_0$.
- We retain at most $2\beta - 1$ counter values.

Compressing bounded intervals
Transition probabilities

Example: σ playing **uniformly at random** for the interval $[\![1, 15]\!]$.





- Transition probabilities can require **exponential-size** representations.
- They are the **least solution** of a quadratic **system of equations**.

Verification via compressed Markov chains

A compressed Markov chain $\mathcal{C}_{\mathcal{I}}^{\sigma}$

- has a **polynomial-size** state space and
- transition probabilities given by **polynomial-size systems**, and
- **preserves termination probabilities**.

Verification via the theory of the reals

We have formulae (in the signature $\{0, 1, +, -, \cdot, \leq\}$)

- $\Phi_{\delta}^{\mathcal{I}}(\mathbf{x}, \mathbf{z}^{\sigma})$ for **transition probabilities** of $\mathcal{C}_{\mathcal{I}}^{\sigma}$;
- $\Phi_{\Omega}^{\mathcal{I}}(\mathbf{x}, \mathbf{y})$ for **termination probabilities** from configurations of $\mathcal{C}_{\mathcal{I}}^{\sigma}$.

We can solve the verification problem by checking if

$$\mathbb{R} \models \forall \mathbf{x} \forall \mathbf{y} \left(\Phi_{\delta}^{\mathcal{I}}(\mathbf{x}, \mathbf{z}^{\sigma}) \wedge \Phi_{\Omega}^{\mathcal{I}}(\mathbf{x}, \mathbf{y})\right) \implies y_{s_{\mathsf{init}}} \geq \theta.$$

Complexity of verification

Unbounded counter		Bounded counter
OEIS	**CIS**	**OEIS**
co-ETR	co-ETR	$\mathsf{P}^{\mathsf{PosSLP}}$
Square-root-sum-hard [EWY10][2]		Square-root-sum-hard

[2] Etessami et al., "Quasi-Birth-Death Processes, Tree-Like QBDs, Probabilistic 1-Counter Automata, and Pushdown Systems", Perform. Evaluation 2010.

Table of contents

Interval strategy synthesis

We consider **two variants** for the **interval strategy synthesis problem.**

Fixed-structure OEIS synthesis problem

Given a **finite interval partition $\mathcal{I}$ of $\mathbb{N}$**, does there exists an OEIS σ **based on $\mathcal{I}$** such that $\mathbb{P}^{\sigma}_{s_{\mathsf{init}}}(\mathsf{Term}(T)) \geq \theta$.

Parameterised OEIS synthesis problem

Given **parameters d and $n \in \mathbb{N}_0$**, does there exists an interval partition $\mathcal{I}$ of $\mathbb{N}$ and an OEIS σ such that

1. $|\mathcal{I}| \leq d$;
2. all bounded $I \in \mathcal{I}$ satisfy $|I| \leq n$;
3. σ is **based on $\mathcal{I}$** and
4. $\mathbb{P}^{\sigma}_{s_{\mathsf{init}}}(\mathsf{Term}(T)) \geq \theta$.

The complexity of interval strategy synthesis

- **Pure strategies**: **guess** a strategy and **verify** it.
- **Randomised strategies**: **quantify existentially** on strategy variables of the verification formula.

	Unbounded counter		Bounded counter
	OEIS	**CIS**	**OEIS**
Pure	$\mathsf{NP}^{\mathsf{ETR}}$	$\mathsf{NP}^{\mathsf{ETR}}$	$\mathsf{NP}^{\mathsf{PosSLP}}$
Random	PSPACE	PSPACE	$\mathsf{NP}^{\mathsf{ETR}}$
	Square-root-sum-hard and NP-hard		

References I

[EWY10] Kousha Etessami, Dominik Wojtczak, and Mihalis Yannakakis. "Quasi-Birth-Death Processes, Tree-Like QBDs, Probabilistic 1-Counter Automata, and Pushdown Systems". In: *Perform. Evaluation* 67.9 (2010), pp. 837–857. DOI: `10.1016/J.PEVA.2009.12.009`. URL: `https://doi.org/10.1016/j.peva.2009.12.009`.

[KEM06] Antonín Kucera, Javier Esparza, and Richard Mayr. "Model Checking Probabilistic Pushdown Automata". In: *Log. Methods Comput. Sci.* 2.1 (2006). DOI: `10.2168/LMCS-2(1:2)2006`. URL: `https://doi.org/10.2168/LMCS-2(1:2)2006`.

[PB24] Jakob Piribauer and Christel Baier. "Positivity-hardness results on Markov decision processes". In: *TheoretiCS* 3 (2024). DOI: `10.46298/THEORETICS.24.9`. URL: `https://doi.org/10.46298/theoretics.24.9`.

References II

[Tiw92] Prasoon Tiwari. "A problem that is easier to solve on the unit-cost algebraic RAM". In: *J. Complex.* 8.4 (1992), pp. 393–397. DOI: 10.1016/0885-064X(92)90003-T. URL: https://doi.org/10.1016/0885-064X(92)90003-T.

Synthesis in one-counter MDPs

- A sequence $(x_n)_{n\in\mathbb{N}}$ is a **linear recurrence sequence** if there exist coefficients $\alpha_1, \dots, \alpha_\ell \in \mathbb{R}$ such that for all $n \geq \ell$,

$$x_n = \alpha_1 \cdot x_{n-1} + \cdots + \alpha_\ell \cdot x_{n-\ell}.$$

- The **positivity problem** asks whether all terms of a **linear recurrence sequence** are non-negative.

Decidability status of optimal synthesis in OC-MDPs [PB24][3]

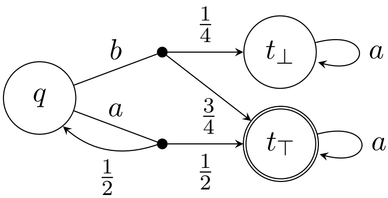
The **positivity problem** is reducible in **polynomial time** to the problem of deciding whether there exists **a strategy** σ such that $\mathbb{P}^{\sigma}_{s_{\mathsf{init}}}(\mathsf{Term}(Q)) > \theta$.

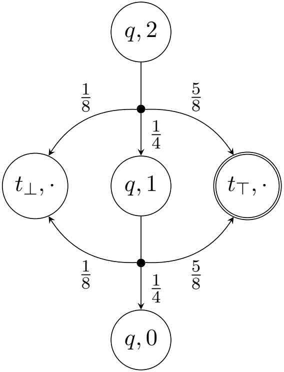
[3] Piribauer and Baier, "Positivity-hardness results on Markov decision processes", TheoretiCS 2024

Randomisation provides more power

We can ensure **higher thresholds** with **randomised strategies**.

Example: reachability with **no memory** with **two steps**.





Square-root-sum problem (see [Tiw92])[4]

Given $x_1, \ldots, x_n, y \in \mathbb{N}$, decide whether

$$\sum_{i=1}^{n} \sqrt{x_i} \geq y.$$

Can be solved in **polynomial time** with **unit-cost arithmetic** [Tiw92].

[4] Tiwari, "A problem that is easier to solve on the unit-cost algebraic RAM", J. Complex. 1992.