Revelations: A Decidable Class of POMDPs with Omega-Regular Objectives

Pierre Vandenhove

Joint work with Marius Belly, Nathanaël Fijalkow, Hugo Gimbert, Florian Horn, Guillermo A. Pérez

LaBRI, Université de Bordeaux (*now at Université de Mons*)

Journées annuelles du GT Vérif 2024 – November 20, 2024

LABORATOIRE BORDELAIS DE RECHERCHE EN INFORMATIQUE LaBRI

UMONS Université de Mons

Outline

Partially observable Markov decision processes (POMDPs):

- stochasticity,
- nondeterminism,
- **uncertainty** about the actual state.

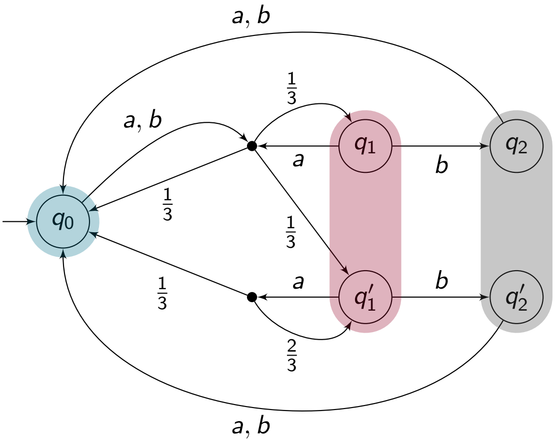
Goal

Strategy synthesis for **parity objectives** ($\rightsquigarrow$ ω-regular objectives).
Undecidable in general; **decidable subclasses**?

Means

Two subclasses with probabilistic guarantees about sometimes **knowing the actual state**.
Natural algorithm that applies to this class.

Partially observable MDPs



States Q, **initial state** q_0, **actions** Act, **observations** Obs.
Strategies are functions $(\mathrm{Act} \times \mathrm{Obs})^* \to \mathcal{D}(\mathrm{Act})$.

Objective

- Function $p \colon Q \to \{0, \ldots, d\}$ assigning **priorities** to **states**.
- **Parity objective**: the **maximal** priority seen infinitely often is **even**.
- Common subclasses:
  - **Büchi**: $p \colon Q \to \{1, 2\}$: something good (2) occurs infinitely often,
  - **coBüchi**: $p \colon Q \to \{0, 1\}$: something bad (1) occurs finitely often.
- **Almost-sure** strategies; "qualitative".

Decidability in POMDPs[1,2]

- Almost-sure **reachability**, **safety**, and **Büchi** are **EXPTIME-complete**.
- Almost-sure **coBüchi** (and therefore **parity**) are **undecidable**.

Undecidability already for **probabilistic automata** ($|\mathrm{Obs}| = 1$).
Quantitative problems (e.g., value-1 problem) are undecidable for reachability objectives.[3]

---

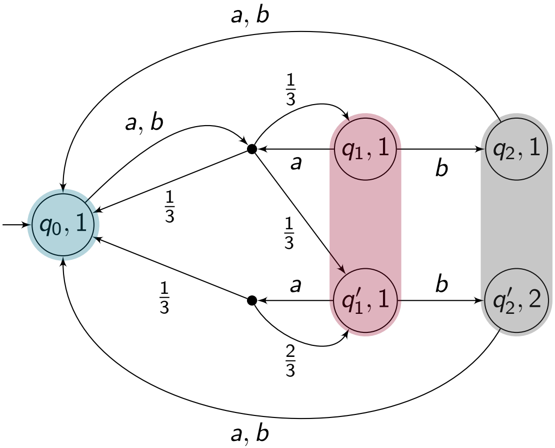
[1] Baier, Größer, and Bertrand, "Probabilistic ω-automata", 2012.

[2] Chatterjee, Chmelik, and Tracol, "What is decidable about partially observable Markov decision processes with ω-regular objectives", 2016.

[3] Gimbert and Oualhadj, "Probabilistic Automata on Finite Words: Decidable and Undecidable Problems", 2010.

Example

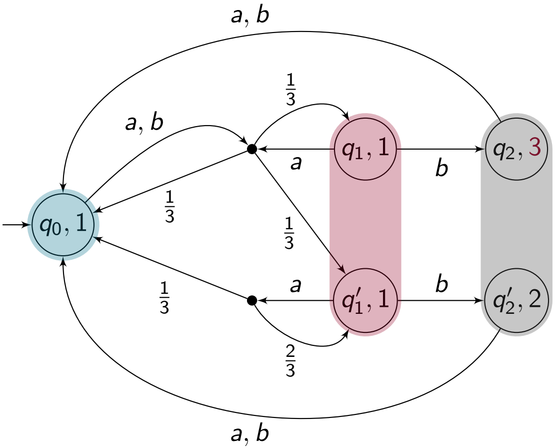
Added priorities 1, 2 to the previous POMDP.



Almost-sure strategy? Yes! Move to q_2/q_2' infinitely often.

Example

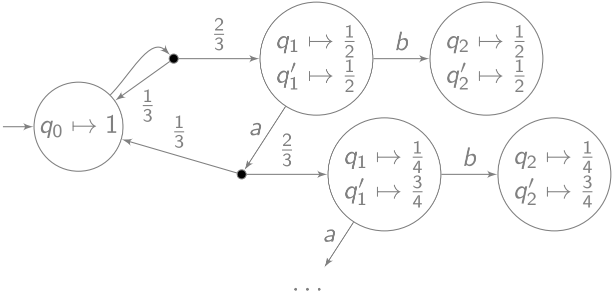
Added priorities 1, 2, 3 to the previous POMDP. **Changed the priority of q_2 to 3.**



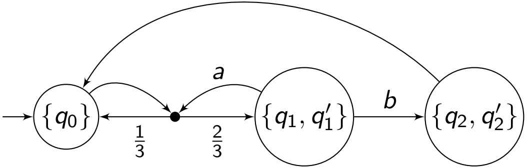
Almost-sure strategy? Yes! Move to q_2/q_2' when *increasingly high probability* to be in q_1'.

Belief (support) MDP

POMDPs induce **infinite belief MDPs**:



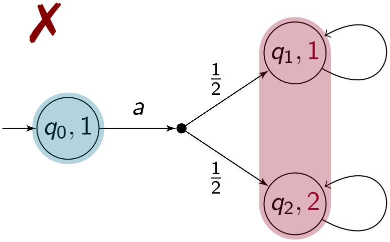
Finite: only keep belief **supports**:



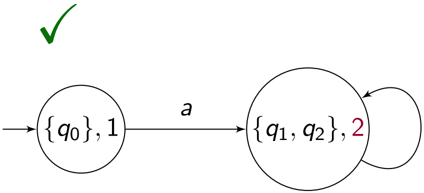
When does the analysis of the belief **support** MDP suffice?

Non-soundness of the belief support MDP

No almost-sure strategy in the POMDP, but OK in the belief support MDP.



POMDP

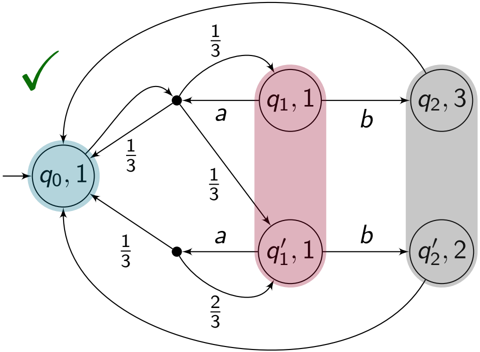


Belief support MDP

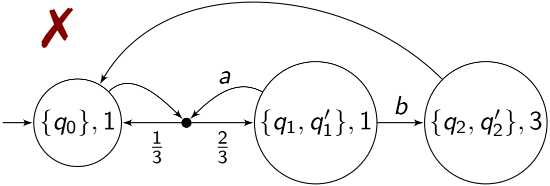
(Technical detail: how to lift the priority function? Take the **max**.)

Incompleteness of the belief support MDP

Almost-sure strategy in the POMDP, **not** in the belief support MDP.



POMDP



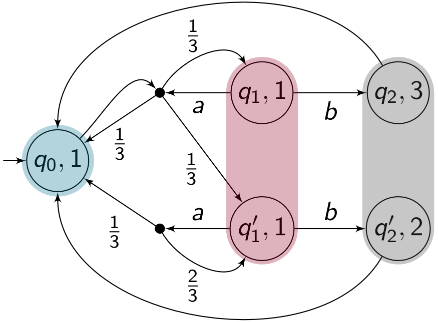
Belief support MDP

First revealing property

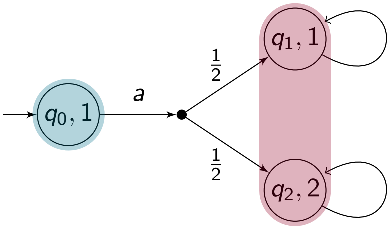
First revealing property

Property 1

A POMDP is **weakly revealing** if for all strategies,
almost surely, the **current state can be known** infinitely often.



Weakly revealing (q_0 is visited infinitely often)



Not weakly revealing

First revealing property

Property 1

A POMDP is **weakly revealing** if for all strategies,
almost surely, the **current state can be known** infinitely often.

When a revealing history happens, as much information in the finite belief **support** MDP as in the infinite belief MDP.

$$\{q_0\} \approx q_0 \mapsto 1$$

Includes POMDPs that *reset* to the initial state with probability 1.

Weakly revealing POMDPs

"Weakly revealing" is a semantic property:

Deciding the property

Deciding whether a POMDP is **weakly revealing** is EXPTIME-hard and in 2-EXPTIME (**update**: actually EXPTIME-complete — work in progress).

Let $\mathcal{P}$ be a **weakly revealing** POMDP with a parity objective.

Soundness for **parity**

Almost-sure winning strategy in the **belief support MDP** of $\mathcal{P} \implies$ also in **POMDP** $\mathcal{P}$.

Completeness for **priorities** $\{0, 1, 2\}$

Almost-sure winning strategy in **POMDP** $\mathcal{P} \implies$ also in the **belief support MDP** of $\mathcal{P}$.

Analysing the belief support MDP is **sound** and **complete** for parity $\{0, 1, 2\}$.

Decidability of weakly revealing POMDPs

Decidability

Almost-sure **parity** $\{0,1,2\}$ for **weakly revealing** POMDPs is EXPTIME-complete.

Algorithm: solve the **belief support MDP** $\rightsquigarrow$ in EXPTIME.
EXPTIME-hardness: already for coBüchi; reduction from almost-sure safety in POMDPs.

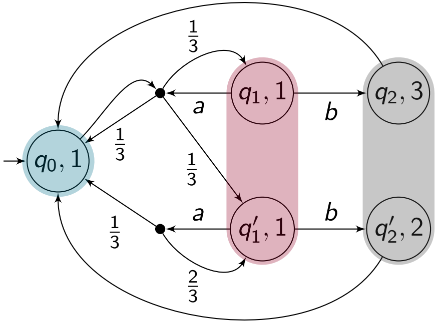
Compared to general POMDPs:
$\rightsquigarrow$ makes **coBüchi decidable**,
$\rightsquigarrow$ gives a (conceptually) **simpler algorithm** for Büchi (state space is 2^Q, instead of $Q \times 2^Q$ in general[4]).

Pure exponential strategies ($2^Q \to \mathsf{Act}$) suffice; this bound is tight.

[4] Baier, Größer, and Bertrand, "Probabilistic ω-automata", 2012.

Full parity still not decidable

Belief support MDP is "incomplete" for this **weakly revealing** POMDP with priorities 1, 2, 3:



Undecidability

Almost-sure **parity** $\{1, 2, 3\}$ is **undecidable** for **weakly revealing** POMDPs.

Reduction from the value-1 problem for probabilistic automata.[5]

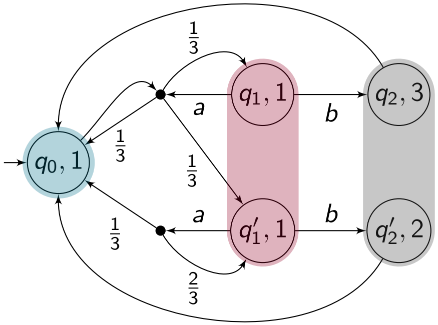
[5] Gimbert and Oualhadj, "Probabilistic Automata on Finite Words: Decidable and Undecidable Problems", 2010.

Second revealing property

Second revealing property

Property 2

A POMDP is **strongly revealing** if for every transition $q \xrightarrow{a} q'$, there is a non-zero probability of **observing** q'.

- Syntactic property.
- Strongly revealing $\implies$ weakly revealing.
- 

Not strongly revealing: $q_1 \xrightarrow{a} q'_1$ is a possible transition, but nothing can reveal q'_1 with certainty.

Strongly revealing: results

Completeness for parity

Almost-sure winning strategy in **strongly revealing POMDP** $\mathcal{P} \implies$ also in the **belief support MDP** of $\mathcal{P}$.

Soundness for full parity follows already from weakly revealing POMDPs.

Theorem

Almost-sure **parity** for **strongly revealing** POMDPs is EXPTIME-complete.

Already EXPTIME-hard for coBüchi.

Optimistic semantic

Another way to see the strongly revealing property:

Optimistic semantic

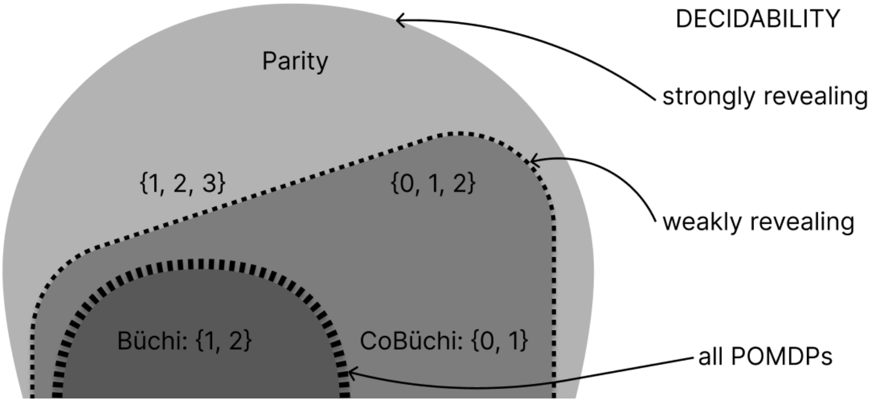
From a POMDP $\mathcal{P}$, one can define a related **strongly revealing** POMDP $\mathcal{P}_{\text{opt}}$ by adding a small probability of a "revelation" along all transitions.

Proposition

If there is **no** almost-sure strategy in $\mathcal{P}_{\text{opt}}$, then this is **also the case in $\mathcal{P}$**.

Finer approach than looking at the underlying MDP (which assumes that *all* states are revealed).

Summary for POMDPs



Decidable subclasses for *parity* POMDPs depending on the **revelation** mechanism.

Decidability frontier when we move to **games**: **games with partial observation** are still **undecidable** for coBüchi under **strong revelations**.

Related works

A few works with similar approaches:

- Models with "**sure** revelations" (not just almost sure).[6]
  $\rightsquigarrow$ Even **games** are decidable!
- We study strategies $2^Q \to \mathsf{Act}$ and give conditions for their sufficiency. Similar studies exist for (less general) "**memoryless**" strategies $\mathsf{Obs} \to \mathsf{Act}$.[7]
- *Active-measuring POMDPs*: a cost may be paid to acquire additional information about the next state.[8]

[6] Berwanger and Mathew, "Infinite games with finite knowledge gaps", 2017.

[7] Vlassis, Littman, and Barber, "On the Computational Complexity of Stochastic Controller Optimization in POMDPs", 2012.

[8] Bellinger et al., "Active Measure Reinforcement Learning for Observation Cost Minimization", 2021; Krale, Simão, and Jansen, "Act-Then-Measure: Reinforcement Learning for Partially Observable Environments with Active Measuring", 2023.

Final comments

Implementation of the algorithms available at
`https://github.com/gaperez64/pomdps-reveal`.

Open problems:

- Larger class where the **belief support MDP** is sound and complete?
- Larger **decidable classes** for coBüchi/parity?
- **More general models** that the revealing mechanisms make decidable?

Thanks!