Revelations: A Decidable Class of POMDPs with Omega-Regular Objectives (Extended Abstract)*

Marius Belly[1], **Nathanaël Fijalkow**[1], **Hugo Gimbert**[1], **Florian Horn**[2], **Guillermo A. Pérez**[3] and **Pierre Vandenhove**[4]
[1]CNRS, LaBRI, Université de Bordeaux, France
[2]CNRS, IRIF, Université de Paris, France
[3]University of Antwerp – Flanders Make, Antwerp, Belgium
[4]UMONS – Université de Mons, Mons, Belgium
bellymarius@gmail.com, {nathanael.fijalkow, hugo.gimbert}@u-bordeaux.fr, florian.horn@irif.fr, guillermo.perez@uantwerpen.be, pierre.vandenhove@umons.ac.be

Abstract

Partially observable Markov decision processes (POMDPs) form a prominent model for uncertainty in sequential decision making. We are interested in constructing algorithms with theoretical guarantees to determine whether the agent has a strategy ensuring a given specification with probability 1. This well-studied problem is known to be undecidable already for very simple omega-regular objectives, because of the difficulty of reasoning on uncertain events. We introduce a revelation mechanism which restricts information loss by requiring that, almost surely, the agent has eventually full information about the current state. Our main technical results are to construct exact algorithms for two classes of POMDPs called *weakly* and *strongly revealing*. Importantly, the decidable cases reduce to the analysis of a finite belief-support Markov decision process. This yields a conceptually simple and exact algorithm for a large class of POMDPs.

1 Introduction

Partially observable Markov decision processes (POMDPs) form a prominent model for uncertainty in sequential decision making. They were defined in the 1960s [Åström, 1965] for operations research and introduced in artificial intelligence by the seminal paper [Kaelbling *et al.*, 1998]. We consider POMDPs from a model-based point of view common in planning and in formal methods. Our goal is to construct exact (as opposed to approximate) algorithms that take as input a complete description of the POMDP and construct a strategy ensuring a given specification. A long line of work has established that most formulations of this problem are undecidable. For instance, even in the extreme case where the agent has no information and the goal is to reach a target state with arbitrarily high probability, complex convergence phenomena occur, implying strong undecidability results [Madani *et al.*, 2003; Gimbert and Oualhadj, 2010; Fijalkow, 2017].

In this work, we are interested in constructing *almost-sure strategies*, meaning strategies ensuring the specification with probability 1. We consider the class of omega-regular objectives (all expressible as *parity objectives*), which is a robust class including properties expressible in Linear Temporal Logic [Pnueli, 1977; Giacomo and Vardi, 2013; Bloem *et al.*, 2018]. Determining whether there exists an almost-sure strategy for the subclass of CoBüchi objectives (requiring to avoid a target from some point onwards) is undecidable [Chatterjee *et al.*, 2016; Bertrand *et al.*, 2017]. There is a vast body of work towards approximate and practical solutions: for instance, using interpolation in the belief space [Lovejoy, 1991], approximation of the value function [Hauskrecht, 2000], or Monte Carlo tree search approaches [Silver and Veness, 2010]. This is orthogonal to the current paper since we focus on exact algorithms.

Our starting point is a simple approach to construct almost-sure strategies: from the POMDP, we build a Markov decision process (MDP) whose states are *supports of the beliefs* of the POMDP. In other words, we store information about which states we can be in, but abstract away the probabilities. The *belief-support MDP* serves as a finite abstraction of the POMDP; one could expect that there exists an almost-sure strategy in the POMDP if and only if there exists one in the corresponding belief-support MDP. Unfortunately, this abstraction is neither sound nor complete; we present a simple counterexample in Figure 1.

The fundamental question we ask in this paper is whether **there are natural sufficient conditions which make the belief-support abstraction correct**. Conceptually, the failure of this abstraction is due to information loss and its accumulation over time.

We introduce a **revelation mechanism** which restricts information loss by requiring that, almost surely, the agent has eventually about the current state. Intuitively, by forbidding information loss from accumulating for an unbounded amount of time, the revelation mechanism removes the convergence issues leading to undecidability. Practically, we conjecture that revelations are a commonly occurring phenomenon in partial observability; a canonical example is systems that reset with a small observable probability at each

*Extended abstract of the eponymous paper from the proceedings of AAAI 2025 [Belly *et al.*, 2025].

Figure 1: We consider the POMDP on the LHS: there is a single signal s, so no information is ever given about the exact state we are in (a behavior the revelation mechanisms forbid!). Yet, almost surely, we reach q_1. The *priorities* indicated on states constitute a parity condition inducing the objective "eventually never visiting q_1", which clearly cannot be ensured almost surely. We represent the *belief-support MDP* on the RHS: the two states are $\{q_0\}$ and $\{q_0, q_1\}$, and only the state $\{q_0, q_1\}$ is visited infinitely often. To assign priorities to the states of this MDP, there are two natural candidates: "maximal priority semantics" and "minimal priority semantics", meaning that we assign either the maximal or minimal priority from the states in the belief support. In this figure, we use the maximal priority semantics: the priority of $\{q_0, q_1\}$ is thus 2, so the belief-support MDP is winning. This means that the analysis of the belief-support MDP is not sound in general. By tweaking the priorities in this example, one can show that neither of these priority semantics is both sound and complete.

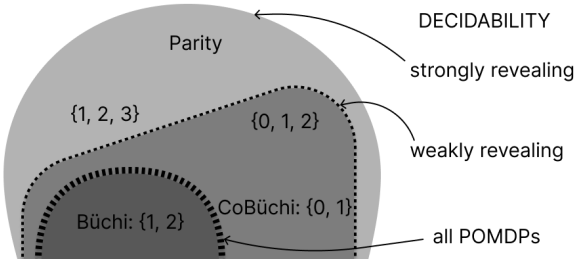

Figure 2: Summary: decidable subclasses of the *parity* objective for almost-sure strategies, depending on the revelation mechanism.

step, and therefore reset infinitely often almost surely, and where this reset is observable. We leave to future work to investigate this question further.

Our contributions. We study two properties of POMDPs based on the revelation mechanism, called *weakly* and *strongly revealing* POMDPs. We obtain (un)decidability results for both classes. The decidable cases reduce to the analysis of the finite belief-support MDP, which yields a conceptually simple and exact algorithm for a large class of POMDPs. A summary of our contributions for POMDPs is provided in Figure 2. We provide an implementation of the algorithm applied to multiple instances at https://github.com/gaperez64/pomdps-reveal.

The importance of our results can be appreciated by the following remark: instead of a subclass of POMDPs, the revelation mechanism can be seen as a new semantics for *all* POMDPs. In that sense, we obtain decidability results for an *optimistic* semantics of POMDPs which, to the best of our knowledge, has not been done before.

Related works. The closest idea to our notion of revelations that we know of is in [Berwanger and Mathew, 2017], defining a class of partial-information multi-player games with *sure* (not just almost-sure) revelations. In a quantitative setting, the idea of having actions that reveal the current state at some cost has appeared multiple times in the literature, such as in [Bertrand and Genest, 2011] for POMDPs with quantitative reachability objectives. Recently, *active-measuring POMDPs*, with a similar mechanism, have been considered in the online planning community [Bellinger *et al.*, 2021; Krale *et al.*, 2023]. Despite a different setting (online planning vs. model-checking), it carries an intuition similar to our work: precise states can be known, which helps find good strategies. Following the first version of this article, subsequent results on the model of strongly revealing POMDPs were obtained in [Asadi *et al.*, 2026].

This abstract is a shortened version of an article with the same title published in the proceedings of the AAAI 2025 conference [Belly *et al.*, 2025]. A version of the paper with the full proofs is available on arXiv [Belly *et al.*, 2024].

2 Preliminaries

A *(discrete) probability distribution* on a finite set X is a function $d\colon X \to [0,1]$ such that $\sum_{x\in X} d(x) = 1$. The set of all probability distributions on X is denoted $\mathcal{D}(X)$. The *support* $\mathsf{supp}(d)$ of a probability distribution d is the set $\{x \in X \mid d(x) > 0\}$. We let $|X|$ denote the number of elements in a set X.

2.1 POMDPs

A *partially observable Markov decision process* (POMDP) is a tuple $\mathcal{P} = (Q, \mathsf{Act}, \mathsf{Sig}, \delta, q_0)$ such that Q is a finite set of *states*, Act is a finite set of *actions*, Sig is a finite set of *signals*, $\delta\colon Q \times \mathsf{Act} \to \mathcal{D}(\mathsf{Sig} \times Q)$ is the *transition function*, and $q_0 \in Q$ is an *initial state*.

A *Markov decision process* (MDP) is a tuple $\mathcal{M} = (Q, \mathsf{Act}, \delta, q_0)$ where $\delta\colon Q \times \mathsf{Act} \to \mathcal{D}(Q)$. Formally, an MDP $\mathcal{M} = (Q, \mathsf{Act}, \delta, q_0)$ can be seen as a POMDP $\mathcal{P} = (Q, \mathsf{Act}, \mathsf{Sig}, \delta, q_0)$ such that $\mathsf{Sig} = \{s_q \mid q \in Q\}$ and for all $q, q', q'' \in Q$ and $a \in \mathsf{Act}$, $\delta(q,a)(s_{q''}, q') > 0$ if and only if $q' = q''$. In practice, it means that the last signal always uniquely determines the current state. MDPs have "complete observation", whereas POMDPs have "partial observation". For a POMDP $\mathcal{P} = (Q, \mathsf{Act}, \mathsf{Sig}, \delta, q_0)$, we define the *underlying MDP of* $\mathcal{P}$ to be the MDP $(Q, \mathsf{Act}, \delta', q_0)$ with $\delta'(q,a)(q') = \sum_{s\in\mathsf{Sig}} \delta(q,a)(s,q')$.

Remark 1. *The observable information in POMDPs is here provided through* signals *along transitions. This contrasts with state-based* observations *that partition the state space, which are also frequently used to model POMDPs. Both models are polynomially equivalent. Using signals here makes the definition of* strongly revealing *(Definition 2) more natural, which is why we opted for this convention.*

We refer to the full version for the classical definitions of plays, strategies, probability measures, and objectives.

2.2 Beliefs and Belief Supports

Let $\mathcal{P} = (Q, \mathsf{Act}, \mathsf{Sig}, \delta, q_0)$ be a POMDP. A *belief* $\mathfrak{b} \in \mathcal{D}(Q)$ is a probability distribution on Q. A *belief support* $b \in 2^Q \setminus \{\emptyset\}$ is the support of a belief. For brevity, we write $2^Q_\emptyset$ for $2^Q \setminus \{\emptyset\}$. At every step, beliefs and belief supports can be updated when playing an action and observing a signal. We show how to do so for belief supports: we define a function $\mathcal{B}\colon 2^Q_\emptyset \times \mathsf{Act} \times \mathsf{Sig} \to 2^Q_\emptyset$ that updates the belief support. For $b \in 2^Q_\emptyset$, $a \in \mathsf{Act}$, $s \in \mathsf{Sig}$, we define $\mathcal{B}(b, a, s) = \{q' \in Q \mid \exists q \in b, \delta(q, a)(s, q') > 0\}$. We extend this function in a natural way to a function $\mathcal{B}^*\colon 2^Q_\emptyset \times (\mathsf{Act} \cdot \mathsf{Sig})^* \to 2^Q_\emptyset$.

Beliefs carry more information than belief supports, as they contain the exact probability of being in a particular state, while belief supports only contain the qualitative information of the possible current states. Observe that when the belief support is a singleton (i.e., $b = \{q\}$ for some $q \in Q$), knowing the precise belief does not yield more information than knowing the belief support, as all the probability mass is in one of the states. Our "revealing" restrictions on POMDPs defined later will exploit this fact.

2.3 The Belief-Support MDP

For a POMDP $\mathcal{P} = (Q, \mathsf{Act}, \mathsf{Sig}, \delta, q_0)$, the *belief-support MDP of* $\mathcal{P}$ is the MDP $\mathcal{P}_\mathcal{B} = (2^Q_\emptyset, \mathsf{Act}, \delta_\mathcal{B}, \{q_0\})$ where for $b, b' \in 2^Q_\emptyset$ and $a \in \mathsf{Act}$, $\delta_\mathcal{B}(b, a)(b') > 0$ if and only if there is $s \in \mathsf{Sig}$ such that $\mathcal{B}(b, a, s) = b'$. We assume the distribution to be uniform over successors with positive probability.

For several simple objectives, the POMDP and its belief-support MDP behave in a similar way. For example, sets of belief supports that can be reached with a positive probability are the same in the POMDP and its belief-support MDP; if a set of belief supports is reachable almost surely in the POMDP, it is also the case in the belief-support MDP.

3 Weakly Revealing POMDPs

We define here our first *revealing* property for POMDPs, which requires that, infinitely often and almost surely, the current state can be deduced by looking at the previous observable history. Formally, we write $B^\mathcal{P}_{\mathsf{sing}} = \{\{q\} \mid q \in Q\}$ for the set of singleton belief supports of a POMDP $\mathcal{P} = (Q, \mathsf{Act}, \mathsf{Sig}, \delta, q_0)$. An observable history $h \in (\mathsf{Act} \cdot \mathsf{Sig})^*$ such that $\mathcal{B}^*(\{q_0\}, h) \in B^\mathcal{P}_{\mathsf{sing}}$ is called a *revelation*.

Definition 1 (Weakly revealing)**.** *A POMDP $\mathcal{P}$ is* weakly revealing *if, for all strategies $\sigma \in \Sigma(\mathcal{P})$, we have $\mathbb{P}^\mathcal{P}_\sigma\left[\mathsf{Büchi}(B^\mathcal{P}_{\mathsf{sing}})\right] = 1$; i.e., for all strategies, infinitely many revelations occur almost surely.*

In particular, POMDPs that "reset" infinitely often, and whose reset can be observed with a dedicated signal, are weakly revealing.

The weakly revealing property is decidable.

Theorem 1. *Deciding whether a POMDP is weakly revealing is EXPTIME-complete.*

3.1 Soundness of the Belief-Support MDP

In this section, we show that, for *weakly revealing* POMDPs, the existence of an almost-sure strategy in the belief-support MDP (with an appropriate priority function) implies the existence of an almost-sure strategy in the POMDP.

For the priority function of the belief-support MDP, we consider the "maximal priority" semantics. Formally, let $\mathcal{P} = (Q, \mathsf{Act}, \mathsf{Sig}, \delta, q_0)$ be a POMDP, and $\mathcal{P}_\mathcal{B}$ be its belief-support MDP. Let $p\colon Q \to \{0, \dots, n\}$ be a priority function on $\mathcal{P}$, inducing the objective $\mathsf{Parity}(p)$. We extend this function to the belief-support MDP: for $b \in 2^Q_\emptyset$, we define $p_\mathcal{B}(b) = \max\{p(q) \mid q \in b\}$.

Without any assumption, the belief-support MDP may be unsound, already for Büchi objectives; there may be an almost-sure strategy in the belief-support MDP, but not in the POMDP. An example illustrating this was given in Figure 1. Surprisingly, it is sound for CoBüchi objectives without further assumption. Using "$\max$" (and not "$\min$") turns out to be the right choice in our setting. Intuitively, under the right revealing assumptions and the right strategies, if a belief support is visited infinitely often, then all its states will be visited infinitely often, so the maximal priority of the belief support is the one that matters given the parity objective.

Under the weakly revealing semantics, almost-sure strategies of the belief-support MDP carry over to the POMDP for all parity objectives. In other words, the analysis of the belief-support MDP is sound. We recall that pure memoryless strategies suffice to reach the optimal value for parity objectives in MDPs [Chatterjee and Henzinger, 2012].

Proposition 1 (Soundness)**.** *Let $\mathcal{P} = (Q, \mathsf{Act}, \mathsf{Sig}, \delta, q_0)$ be a weakly revealing POMDP with priority function p, and let $\mathcal{P}_\mathcal{B}$ be its belief-support MDP with priority function $p_\mathcal{B}$. Assume there is an almost-sure strategy $\sigma_\mathcal{B}$ for $\mathsf{Parity}(p_\mathcal{B})$ in $\mathcal{P}_\mathcal{B}$; by [Chatterjee and Henzinger, 2012], we may assume $\sigma_\mathcal{B}$ to be pure and memoryless. Then, there is an almost-sure strategy for $\mathsf{Parity}(p)$ in $\mathcal{P}$.*

3.2 Decidability for Priorities 0, 1, and 2

We show that the existence of an almost-sure strategy in a weakly revealing POMDP implies the existence of an almost-sure strategy in its belief-support MDP when priorities are in $\{0, 1, 2\}$. This provides a converse to Proposition 1 when priorities are restricted to $\{0, 1, 2\}$. We emphasize that parity objectives with priorities $\{0, 1, 2\}$ encompass both Büchi and CoBüchi objectives. This result is false without the weakly revealing assumption; see the simple POMDP in Figure 1.

Proposition 2 (Completeness)**.** *Let $\mathcal{P} = (Q, \mathsf{Act}, \mathsf{Sig}, \delta, q_0)$ be a weakly revealing POMDP with priority function p with values in $\{0, 1, 2\}$. Let $\mathcal{P}_\mathcal{B}$ be its belief-support MDP with priority function $p_\mathcal{B}$. If there is an almost-sure strategy for $\mathsf{Parity}(p)$ in $\mathcal{P}$, then there is an almost-sure strategy for $\mathsf{Parity}(p_\mathcal{B})$ in $\mathcal{P}_\mathcal{B}$.*

Together with an EXPTIME-hardness proof, this yields the following result.

Theorem 2. *The existence of an almost-sure strategy for parity objectives with priorities in $\{0, 1, 2\}$ in weakly revealing POMDPs is EXPTIME-complete.*

3.3 Undecidability for Priorities 1, 2, and 3

The previous section suggests that analyzing the belief-support MDP is a sound and complete approach for weakly revealing POMDPs with parity objectives with priorities in $\{0, 1, 2\}$. One may wonder whether it is complete for any priority function, i.e., if the existence of an almost-sure strategy in a weakly revealing POMDP implies the existence of an almost-sure strategy in its belief-support MDP. Unfortunately, this fails to hold in general, already for priority functions taking values in $\{1, 2, 3\}$. Our decidability result is therefore optimal w.r.t. the priorities used.

Theorem 3 (Undecidability)**.** *The existence of an almost-sure strategy in weakly revealing POMDPs with a parity objective with priorities in $\{1, 2, 3\}$ is undecidable.*

4 Strongly Revealing POMDPs

In this section, we introduce *strongly revealing POMDPs*, a stronger property entailing that infinitely many revelations occur in a POMDP almost surely. We show that the existence of almost-sure strategies is decidable for strongly revealing POMDPs with arbitrary parity objectives.

We define a notion of *revealing signals*: for an action a and a state q' of a POMDP $\mathcal{P} = (Q, \mathsf{Act}, \mathsf{Sig}, \delta, q_0)$, we define $\mathsf{Revealing}(a, q') = \{s \in \mathsf{Sig} \mid \forall r, r' \in Q, r' \neq q' \implies \delta(r, a)(s, r') = 0\}$ to be the set of signals that indicate surely that the next state is q' after playing a. For convenience, we define $\mathsf{Succ}(q, a) = \{q' \in Q \mid \exists s \in \mathsf{Sig}, \delta(q, a)(s, q') > 0\}$ and $\mathsf{Succ}(q, a, s) = \{q' \in Q \mid \delta(q, a)(s, q') > 0\}$.

Definition 2. *POMDP $\mathcal{P} = (Q, \mathsf{Act}, \mathsf{Sig}, \delta, q_0)$ is* strongly revealing *if any transition between two states for a given action in the underlying MDP of $\mathcal{P}$ can also happen with a revealing signal. Formally, $\mathcal{P}$ is strongly revealing if for all $q, q' \in Q$ and $a \in \mathsf{Act}$, if $q' \in \mathsf{Succ}(q, a)$, then there is $s \in \mathsf{Revealing}(a, q')$ such that $q' \in \mathsf{Succ}(q, a, s)$.*

Under this definition, the set of belief supports $B^{\mathcal{P}}_{\mathsf{sing}}$ is visited infinitely often almost surely from the initial state for any given strategy, so a strongly revealing POMDP is in particular weakly revealing. The strongly revealing property can be decided in polynomial time in the size of a POMDP by simply analyzing every transition.

Example 1. *We give an example of a strongly revealing POMDP inspired by the* tiger *of [Cassandra* et al.*, 1994], depicted in Figure 3. In this environment, an agent has to open the left or the right door, with action a_{L} or a_{R}, respectively. One of them has a (deadly) tiger behind it. Fortunately, the agent can choose to wait and listen (action $a_?$) to inform its decision. Listening results in a signal that is biased towards the true state, i.e., the signal can be s_{L} or s_{R} and the former is more likely if the tiger really is on the left, and vice versa.*

We present our version of the tiger environment in which listening guarantees one will eventually discern behind which door there is a tiger. This is achieved by adding new revealing signals $s_{\mathsf{L}!}$ and $s_{\mathsf{R}!}$ which, importantly, can only be obtained when the tiger is on the left or on the right, respectively. To keep things interesting, these signals can only be obtained with a small probability (yet, their presence already ensures that the POMDP is strongly revealing*). We also add signals for death ($s_\bot$) and victory ($s_\top$), which are missing from the original tiger environment.*

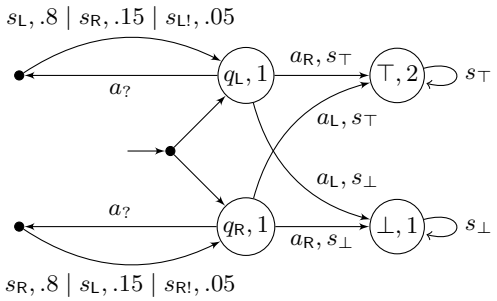

Figure 3: Strongly revealing *tiger* (Example 1).

The addition of the signals $s_{\mathsf{L}!}$ and $s_{\mathsf{R}!}$ makes the objective easier to satisfy and therefore changes the semantics of the POMDP; indeed, it is possible to gather more information than in the original tiger environment. However, revealing the deadlock states $\top$ and $\bot$ through signals $s_\top$ and $s_\bot$ is necessary to make the POMDP strongly revealing but does not make the objective easier to satisfy, as no strategy can exit these states anyway.

4.1 Decidability of Parity with Strong Revelations

The soundness of the analysis of the belief-support MDP for strongly revealing POMDPs follows from Proposition 1; it remains to show completeness.

Proposition 3. *Let $\mathcal{P} = (Q, \mathsf{Act}, \mathsf{Sig}, \delta, q_0)$ be a strongly revealing POMDP with a parity objective with priority function p, and $\mathcal{P}_{\mathcal{B}}$ be its belief-support MDP with priority function $p_{\mathcal{B}}$. If there is an almost-sure strategy for $\mathsf{Parity}(p)$ in $\mathcal{P}$, then there is an almost-sure strategy for $\mathsf{Parity}(p_{\mathcal{B}})$ in $\mathcal{P}_{\mathcal{B}}$.*

We obtain as before the decidability of the problem by reducing to the analysis of the belief-support MDP. The proof is similar to that of Theorem 2, simply replacing the use of Proposition 2 by Proposition 3. Together with a proof of EXPTIME-hardness, this yields the following result.

Theorem 4. *The existence of an almost-sure strategy for parity in strongly revealing POMDPs is EXPTIME-complete.*

4.2 Undecidability of Strongly Revealing Games

We briefly consider here whether the revealing semantics helps in *zero-sum two-player partial-information games*. In general, such games with CoBüchi objectives are undecidable (they encompass POMDPs) while Büchi games are decidable for almost-sure strategies [Bertrand *et al.*, 2017; Chatterjee *et al.*, 2013]. We show a negative result: the existence of an almost-sure strategy in two-player CoBüchi games with partial information is undecidable, even when restricted to games satisfying a natural extension of the strongly revealing property.

Informally, a game is strongly revealing if for any possible transition between two states with a pair of actions, there is a positive probability of a signal that reveals the second state. This definition generalizes naturally the one for POMDPs.

Theorem 5. *The existence of an almost-sure strategy for Player 1 in strongly revealing CoBüchi games is undecidable.*

Acknowledgments

This work was partially supported by the SAIF project, funded by the "France 2030" government investment plan managed by the French National Research Agency, under the reference ANR-23-PEIA-0006. Pierre Vandenhove was funded by ANR project G4S (ANR-21-CE48-0010-01) and by the Belgian Fonds de la Recherche Scientifique – FNRS through the CDR grant "INSTAP" no. J.0228.26. This work was sparked by discussions with Guillaume Vigeral and Bruno Ziliotto, following a talk on a related model [Vigeral and Ziliotto, 2022].

References

[Asadi *et al.*, 2026] Ali Asadi, Krishnendu Chatterjee, David Lurie, and Raimundo Saona. Revealing POMDPs: Qualitative and quantitative analysis for parity objectives. In Sven Koenig, Chad Jenkins, and Matthew E. Taylor, editors, *Proceedings of the 40th AAAI Conference on Artificial Intelligence, AAAI 2026, Singapore, January 20-27, 2026*, pages 36146–36154. AAAI Press, 2026.

[Åström, 1965] Karl Johan Åström. Optimal Control of Markov Processes with Incomplete State Information I. *Journal of Mathematical Analysis and Applications*, 10:174–205, 1965.

[Bellinger *et al.*, 2021] Colin Bellinger, Rory Coles, Mark Crowley, and Isaac Tamblyn. Active measure reinforcement learning for observation cost minimization. In Luiza Antonie and Pooya Moradian Zadeh, editors, *Canadian Conference on Artificial Intelligence*. Canadian Artificial Intelligence Association, 2021.

[Belly *et al.*, 2024] Marius Belly, Nathanaël Fijalkow, Hugo Gimbert, Florian Horn, Guillermo A. Pérez, and Pierre Vandenhove. Revelations: A decidable class of POMDPs with omega-regular objectives. *CoRR*, abs/2412.12063, 2024.

[Belly *et al.*, 2025] Marius Belly, Nathanaël Fijalkow, Hugo Gimbert, Florian Horn, Guillermo A. Pérez, and Pierre Vandenhove. Revelations: A decidable class of POMDPs with omega-regular objectives. In Toby Walsh, Julie Shah, and Zico Kolter, editors, *Proceedings of the 39th Annual AAAI Conference on Artificial Intelligence, February 25 – March 4, 2025, Philadelphia, PA, USA*, pages 26454–26462, Washington, DC, USA, 2025. AAAI Press.

[Bertrand and Genest, 2011] Nathalie Bertrand and Blaise Genest. Minimal disclosure in partially observable Markov decision processes. In Supratik Chakraborty and Amit Kumar, editors, *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2011, December 12–14, 2011, Mumbai, India*, volume 13 of *LIPIcs*, pages 411–422. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2011.

[Bertrand *et al.*, 2017] Nathalie Bertrand, Blaise Genest, and Hugo Gimbert. Qualitative determinacy and decidability of stochastic games with signals. *Journal of the ACM*, 64(5):33:1–33:48, 2017.

[Berwanger and Mathew, 2017] Dietmar Berwanger and Anup Basil Mathew. Infinite games with finite knowledge gaps. *Information and Computation*, 254:217–237, 2017.

[Bloem *et al.*, 2018] Roderick Bloem, Krishnendu Chatterjee, and Barbara Jobstmann. Graph games and reactive synthesis. In Edmund M. Clarke, Thomas A. Henzinger, Helmut Veith, and Roderick Bloem, editors, *Handbook of Model Checking*, pages 921–962. Springer, 2018.

[Cassandra *et al.*, 1994] Anthony R. Cassandra, Leslie Pack Kaelbling, and Michael L. Littman. Acting optimally in partially observable stochastic domains. In Barbara Hayes-Roth and Richard E. Korf, editors, *Proceedings of the 12th National Conference on Artificial Intelligence, Seattle, WA, USA, July 31 – August 4, 1994, Volume 2*, pages 1023–1028. AAAI Press / The MIT Press, 1994.

[Chatterjee and Henzinger, 2012] Krishnendu Chatterjee and Thomas A. Henzinger. A survey of stochastic ω-regular games. *Journal of Computer and System Sciences*, 78(2):394–413, 2012.

[Chatterjee *et al.*, 2013] Krishnendu Chatterjee, Laurent Doyen, and Thomas A. Henzinger. A survey of partial-observation stochastic parity games. *Formal Methods in System Design*, 43(2):268–284, 2013.

[Chatterjee *et al.*, 2016] Krishnendu Chatterjee, Martin Chmelik, and Mathieu Tracol. What is decidable about partially observable Markov decision processes with ω-regular objectives. *Journal of Computer and System Sciences*, 82(5):878–911, 2016.

[Fijalkow, 2017] Nathanaël Fijalkow. Undecidability results for probabilistic automata. *ACM SIGLOG News*, 4(4):10–17, 2017.

[Giacomo and Vardi, 2013] Giuseppe De Giacomo and Moshe Y. Vardi. Linear temporal logic and linear dynamic logic on finite traces. In Francesca Rossi, editor, *International Joint Conference on Artificial Intelligence*, IJCAI'13, pages 854–860. IJCAI/AAAI, 2013.

[Gimbert and Oualhadj, 2010] Hugo Gimbert and Youssouf Oualhadj. Probabilistic automata on finite words: Decidable and undecidable problems. In Samson Abramsky, Cyril Gavoille, Claude Kirchner, Friedhelm Meyer auf der Heide, and Paul G. Spirakis, editors, *International Colloquium on Automata, Languages and Programming, ICALP*, volume 6199 of *Lecture Notes in Computer Science*, pages 527–538. Springer, 2010.

[Hauskrecht, 2000] Milos Hauskrecht. Value-function approximations for partially observable Markov decision processes. *Journal of Artificial Intelligence Research*, 13:33–94, 2000.

[Kaelbling *et al.*, 1998] Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1):99–134, 1998.

[Krale *et al.*, 2023] Merlijn Krale, Thiago D. Simão, and Nils Jansen. Act-then-measure: Reinforcement learning

for partially observable environments with active measuring. In Sven Koenig, Roni Stern, and Mauro Vallati, editors, *Proceedings of the Thirty-Third International Conference on Automated Planning and Scheduling, Prague, Czech Republic, July 8–13, 2023*, pages 212–220. AAAI Press, 2023.

[Lovejoy, 1991] William S. Lovejoy. Computationally feasible bounds for partially observed Markov decision processes. *Operations Research*, 39(1):162–175, 1991.

[Madani *et al.*, 2003] Omid Madani, Steve Hanks, and Anne Condon. On the undecidability of probabilistic planning and related stochastic optimization problems. *Artificial Intelligence*, 147(1-2):5–34, 2003.

[Pnueli, 1977] Amir Pnueli. The temporal logic of programs. In *Symposium on Foundations of Computer Science*, SFCS'77, 1977.

[Silver and Veness, 2010] David Silver and Joel Veness. Monte-Carlo planning in large POMDPs. In *Conference on Neural Information Processing Systems, NIPS*, pages 2164–2172. Curran Associates, Inc., 2010.

[Vigeral and Ziliotto, 2022] Guillaume Vigeral and Bruno Ziliotto. Zero-sum stochastic games with intermittent observation of the state. Communication at the "Current Trends in Graph and Stochastic Games" workshop (GAMENET'22), 2022.