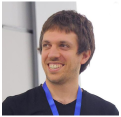
Blind Stochastic Games



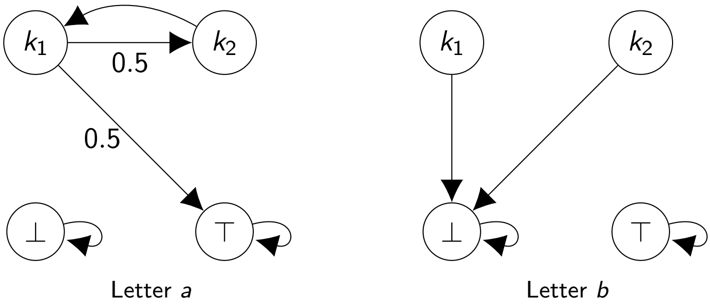
A. Asadi[1] K. Chatterjee[1] D. Lurie[2] **R. Saona**[3]


A. Shafiee[1] B. Ziliotto[2]

[1]Institute of Science and Technology Austria (ISTA)
[2]Paris Dauphine
[3]London School of Economics

University of Liverpool — November 2025

Probabilistic Finite Automata



Processing a letter defines the probabilistic transition over states.

Probabilistic Finite Automata: Language

The language of a Probabilistic Finite Automata is

$$\mathcal{L} := \left\{ w \in \Sigma^* : \mathbb{P}_{s_1}(S_{|w|} = \top) > 1/2 \right\} .$$

(In the previous example, $\mathcal{L} = aaa\Sigma^*$)

The computational problem we consider is EMPTYNESS.

$$\mathcal{L} \stackrel{?}{=} \emptyset \, .$$

Probabilistic Finite Automata: Decidability

Theorem (Madani 2003)

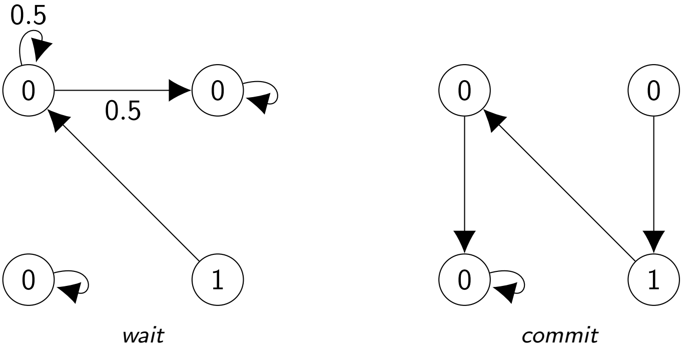
Deciding EMPTYNESS for Probabilistic Finite Automata is undecidable.

Theorem (Madani 2003)

Deciding EMPTYNESS for Probabilistic Finite Automata where every word is accepted with probability in $[0, \varepsilon] \cup [1 - \varepsilon, 1]$ *is undecidable.*

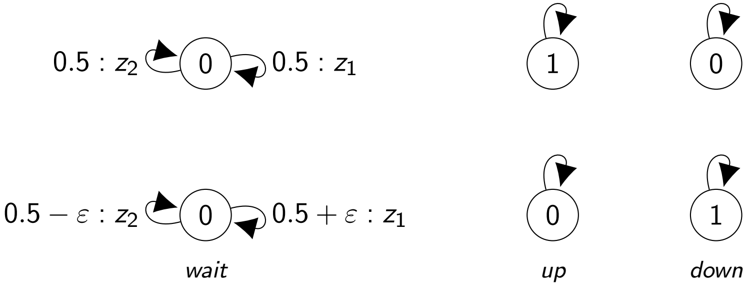
Game Theoretical view

Blind Markov Decision Process

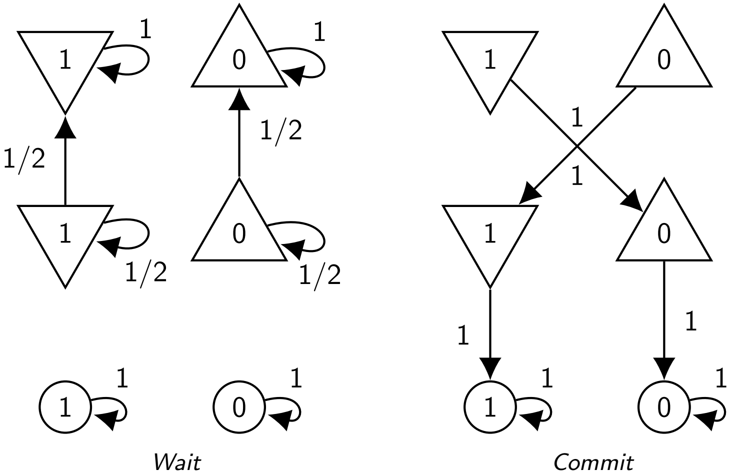


Limsup objective does not have finite-memory ε-optimal policy

Partially Observable Markov Decision Process



Discontinuity of the value of POMDPs.



Alternating-controller blind stochastic game with two actions with limit value but no undiscounted value

Computational Problems

- **Synthesis of policies**
  Compute an ε-optimal strategy
- **Qualitative Reachability**
  Is the reachability value 1?
- **Value approximation**
  Approximate the value
- **Property checking**
  Is my Blind Stochastic Game particularly easy to solve?

Contributions

Synthesis of policies

The value of POMDPs exists, but it is undecidable to approximate.

Theorem (MOR 2021)

Every POMDP with liminf average objective has **finite-memory** *ε-optimal strategies.*

Corollary (MOR 2021)

Every Blind MDP with liminf average objective has **finite-recall** *ε-optimal strategies.*

Qualitative Reachability

Is the reachability value of my POMDP 1?
In general, this problem is also undecidable.
Restricting to finite-memory policies does not it easier.

Theorem (UAI 2025)

Deciding if the reachability value of a POMDP is 1 *with constant-memory polices is NP-complete.*

Theorem (AAAI 2026)

If the state is revealed with positive probability in each step, then deciding if the reachability value is 1 *is in EXPTIME.*

Value approximation

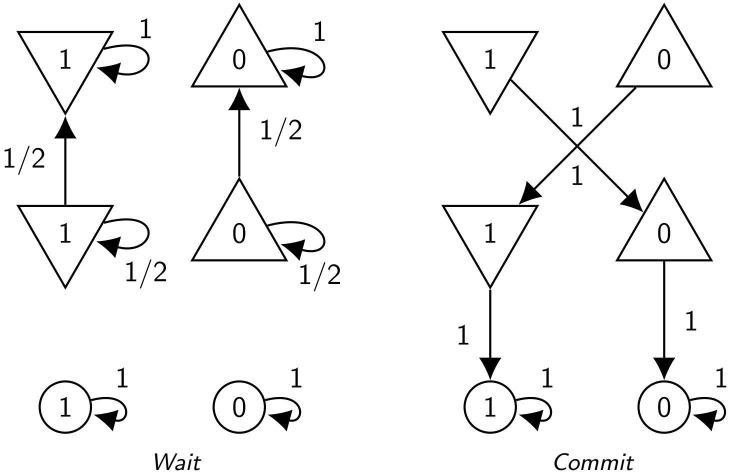
Value does not exist in general blind stochastic games.

We define a subclass where the (undiscounted) value

- exists
- is robust upon perturbations
- can be approximated
- can not be computed exactly

Ergodic
Blind Stochastic Games

Blind Stochastic Games



Alternating-controller blind stochastic game with two actions with limit value but no undiscounted value

Difficulty: Absorbing states

Difficulty:
Absorbing states can **accumulate arbitrarily small contributions**. So, the player(s) behaviour depends on nonapproximable effects because, in the limit value, they are infinitely patient.

Definitions

Blind Stochastic Games

A Blind Stochastic Game is a tuple $\Gamma = (\mathcal{K}, \mathcal{I}, \mathcal{J}, \delta, r, s_1)$ where

- $\mathcal{K}$ is a finite set of **states**.
- $\mathcal{I}$ and $\mathcal{J}$ are finite sets of **actions** for each player.
- $\delta \colon \mathcal{K} \times \mathcal{I} \times \mathcal{J} \to \Delta(\mathcal{K})$ is a probabilistic **transition** function.
- $r \colon \mathcal{K} \to \mathbb{R}$ is a **reward** function.
- $k_1 \in \mathcal{K}$ is an **initial state**.

Model. Players know the game Γ.
They play simultaneously and observe each others actions.
Therefore, **they have the same belief** over the current state.

Limit Value

Denote σ and τ general strategies for the players.
For $\lambda \in (0,1)$, the λ-objective of the players is to optimize

$$\gamma_\lambda(\sigma, \tau) := \mathbb{E}_{k_1}^{\sigma,\tau} \left(\lambda \sum_{t=1}^{\infty} (1-\lambda)^{t-1} \, r(K_t) \right) .$$

The discounted value is defined as

$$\mathrm{val}_\lambda := \min_\sigma \max_\tau \gamma_\lambda(\sigma, \tau) = \max_\tau \min_\sigma \gamma_\lambda(\sigma, \tau) .$$

The (limit) value is defined as

$$\mathrm{val} := \lim_{\lambda \to 0^+} \mathrm{val}_\lambda \, .$$

Previous results

Mertens' Conjecture

Conjecture (1987, International Congress of Mathematics)

In every (zero-sum) stochastic game, the (limit) value exists.

Proven in many special cases of stochastic games.

Limit Value: Existence

Theorem (2002, Rosenberg & Solan & Vieille, Annals of Statistics)

Every blind 1-player stochastic game has a (limit) value.

Limit Value: Nonexistence

Theorem (2016, Bruno Ziliotto, Annals of Probability)

There exists a blind stochastic game where the (limit) value does not exist.

Limit Value: Undecidability

Theorem (Madani 2003)

Deciding EMPTYNESS for Probabilistic Finite Automata where every word is accepted with probability in $[0, \varepsilon] \cup [1 - \varepsilon, 1]$ *is undecidable.*

Theorem (2003, Madani & Hanks & Condon, Artificial Intelligence)

The problem of recognizing blind MDPs with value almost 1 *is undecidable.*

Ergodic transitions

Ergodicity: Forgetting where you come from

In Markov Chains, an ergodic transition probability P satisfies

$$\lim_{n \to \infty} P^n = \mathbb{1} \mu^\top .$$

Equivalently, for all $p \in \Delta(\mathcal{K})$, we have that

$$p^\top \lim_{n \to \infty} P^n = \mu^\top .$$

In particular, for all $k, \widetilde{k} \in \mathcal{K}$, for all $k' \in \mathcal{K}$

$$\lim_{n \to \infty} \left| P^n_{k,k'} - P^n_{\widetilde{k},k'} \right| = 0 .$$

Coefficient of Ergodicity

Definition (Coefficient of Ergodicity)

Given a matrix $P \in \mathbb{R}^{\mathcal{K} \times \mathcal{K}}$, define

$$\operatorname{erg}(P) := \max_{k, \widetilde{k} \in \mathcal{K}} \sum_{k' \in \mathcal{K}} \left| P_{k,k'} - P_{\widetilde{k},k'} \right| .$$

Note that

- $\operatorname{erg}(PQ) \leq \operatorname{erg}(P) \ \operatorname{erg}(Q)$.
- $\operatorname{erg}(P) = 0$ if and only if $P = \mathbb{1}\mu^\top$.

Ergodic Blind Stochastic Games

Definition (Ergodic blind stochastic game)

For all action pairs $(i, j) \in \mathcal{I} \times \mathcal{J}$,

$$\operatorname{erg}\left(P(i,j)\right) < 1\,.$$

Lemma

Consider an ergodic blind stochastic game. For all $\varepsilon > 0$*, there exists an integer* n_ε *such that,*
for all $n \geq n_\varepsilon$ *and tuples of action pairs* $(i_1, j_1), \ldots, (i_n, j_n)$*,*

$$\operatorname{erg}\left(P(i_1,j_1) \cdots P(i_n,j_n)\right) \leq \varepsilon\,.$$

Intuitively, the current belief is approximated by considering only the last n_ε actions:
no need to remember your initial distribution!

Our Contributions

Limit Value: Existence

Theorem (MOR 2026)

Every ergodic blind stochastic game has a limit value.

Proof sketch.

- Construct a finite stochastic game based on n_ε steps at a time.
- Belief dynamics remain close between the original and approximated model.
- Finite-stage payoff remain close between the models.

□

Limit Value: Approximability

Theorem (MOR 2026)

Approxmating the limit value of an ergodic blind stochastic game can be done in 2-EXPSPACE.

Proof sketch.

- The previous construction requires 2-EXP states.
- Approximating the limit value can be done by solving a sentence of the first order theory of the reals, which is PSPACE on the input.

□

Limit Value: Undecidability

Theorem (MOR 2026)

The problem of recognizing lower and upper bounds of the limit value of ergodic blind MDPs is undecidable.

Proof sketch.

- Consider an arbitrary blind MDP.
- Add a positive transition to a new state and a restart action.
- These modifications do not change the limit value, because the controller is infinitely patient.
- Remarkably, the transitions are now ergodic!

□

Summary of Contributions

Blind Class	**Existence**	**Approximation**	**Exact**
SGs	No	–	–
Ergodic SGs	**Yes**	**2-EXPSPACE**	**Undecidable**
MDPs	Yes	Undecidable	Undecidable
Ergodic MDPs	Yes	**2-EXPSAPCE**	**Undecidable**

Summary of results

Thank you!