Robust Strategies for Stochastic Multi-Agent Systems

Extended Abstract

Raphaël Berthon
RWTH Aachen University
Aachen, Germany
berthon@cs.rwth-aachen.de

Joost-Pieter Katoen
RWTH Aachen University
Aachen, Germany
katoen@cs.rwth-aachen.de

Munyque Mittelmann
University of Naples Federico II
Naples, Italy
munyque.mittelmann@unina.it

Aniello Murano
University of Naples Federico II
Naples, Italy
aniello.murano@unina.it

ABSTRACT

The precise probabilities of stochastic systems are often partially unknown and may face perturbations. Finding a strategy in this setting is difficult, as it requires dealing with uncertainty on the system transitions while interacting with other agents. In this paper, we introduce the robust model checking problem for Multi-Agent Systems, in which agents play strategies that ensure the satisfaction of a specification is satisfied, even though the system probabilities are uncertain. We consider specifications in a variant of Alternating-time Temporal Logic with bounded memory.

KEYWORDS

Stochastic Multi-Agent Systems; Probabilistic Model Checking; Logics for Strategic Reasoning; Robustness

ACM Reference Format:
Raphaël Berthon, Joost-Pieter Katoen, Munyque Mittelmann, and Aniello Murano. 2025. Robust Strategies for Stochastic Multi-Agent Systems: Extended Abstract. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 3 pages.

1 INTRODUCTION

One of the first and most important developments in using formal methods for reasoning about strategies in Multi-Agent Systems (MAS) is the Alternating-time Temporal Logic (ATL) [1], which contains strategic modalities expressing cooperation and competition among agents to achieve a goal. Some aspects of MAS, such as the unpredictable behavior of agents and the occurrence of random phenomena, are uncertain. These aspects can be measured based on experiments or past observations and represented with stochastic models, such as Markov decision processes (MDP) and stochastic MAS. Probabilistic ATL (PATL) [8] extends ATL to the probabilistic setting, allowing reasoning about randomized strategic abilities of agents interacting in a system with stochastic transitions. Uncertainty in MAS may also originate from agents' partial observability of the system, but model-checking strategic abilities under imperfect information and perfect recall entails undecidability, even when restricted to deterministic MAS [10] or to a single agent as in POMDPs [14], leading to consider memoryless or bounded-memory strategies [4, 5, 14].

In many cases, the precise probabilities of the system transitions are unknown and may face perturbations. An example is model-based reinforcement learning, where agents estimate the agent-environment interaction model (e.g., a MDP [13]). Since the model is learned from their interaction with the environment, its transitions are susceptible to error. Strategizing in such a setting requires dealing with uncertainty on the system transitions while interacting with other agents, who may be cooperative or adversarial. Different approaches exist: on Markov chains, MDPs, and POMDPs, it has been proposed to consider intervals for the possible value of transition probabilities [7, 11, 14]: strategies must hold for any system whose transition set is within the interval. A way to generalize uncertainty and robustness is by considering parameters, where transition probabilities can be represented as equations over a given set of parameters [2, 12].

We introduce the *robust model checking* problem, which ensures that a temporal specification is satisfied, even though the system probabilities may suffer perturbations. By relying on PATL-like specifications, we consider *coalitional* strategies, and capture the strategic behavior of agent coalitions in probabilistic MAS with an additional uncertainty concerning the exact transition probabilities.

2 PROBABILISTIC ATL WITH BOUNDED STRATEGIES

2.1 Parametric Systems

We propose to extend stochastic multi-agent systems into *parametric systems*. In a parametric system $\mathcal{G}$, transition probabilities are replaced with equations over a set of variables X. A set of parameters resulting in all probabilities being in $[0,1]$ is *well-defined*. Once a well-defined set of parameters Val is fixed, we obtain $\mathcal{G}[Val]$, a classical stochastic MAS. We thus consider if a property holds for all valuations over X that yield a well defined set of probabilities.

We follow Definition 10.97 of [3] of strategies: a *general randomized strategy* σ with bounded recall n is a tuple $(Q, act, \Delta, start)$ where Q is a set of modes (or memory states), Δ is a randomized transition function, *act* randomly selects the next action depending on the current state, and $start$ randomly selects a starting mode for the strategy. As in [9], a strategy has finite memory n if $|Q| = n$ and

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

memoryless if $|Q| = 1$. A strategy is *observation-based* if Δ and *act* can be represented as depending on 2^{AP} and Q. This corresponds to *imperfect information*, since strategies may only depend on AP, the observed atomic propositions labeling the states, and not the states themselves. Let $Str_{a,n}$ be the set of observation-based strategies with bounded recall n for agent a and $Str_n = \cup_{a \in \mathrm{Ag}} Str_{a,n}$.

2.2 The Logic $\mathsf{PATL_b}$

We introduce the Probabilistic Alternating-Time Temporal Logic with bounded strategies ($\mathsf{PATL_b}$) defined as follows:

Definition 1. The syntax of $\mathsf{PATL_b}$ is defined by the grammar

$$\varphi ::= p \mid \varphi \vee \varphi \mid \neg\varphi \mid \langle\langle C \rangle\rangle_k^{\bowtie d}(\mathbf{X}\varphi) \mid \langle\langle C \rangle\rangle_k^{\bowtie d}(\varphi \mathbf{U} \varphi) \mid \langle\langle C \rangle\rangle_k^{\bowtie d}(\varphi \mathbf{R} \varphi)$$

where $p \in \mathrm{AP}$, $k \in \mathbb{N}$, $C \subseteq \mathrm{Ag}$, d is a rational constant in $[0, 1]$, and $\bowtie \in \{\leq, <, >, \geq\}$. Most of these operators are classical, except for $\langle\langle C \rangle\rangle_k^{\bowtie d}\varphi$, that asserts that there exists an observation-based strategy with complexity at most k for the coalition C to collaboratively enforce φ with a probability in relation $\bowtie$ with constant d.

Given a coalition strategy $\boldsymbol{\sigma_C} \in \prod_{a \in C} Str_a^{\rho}$, the set of possible outcomes of $\boldsymbol{\sigma_C}$ from a path π (i.e., a finite sequence of states) to be the set $out_C(\boldsymbol{\sigma_C}, \pi) = \{out((\boldsymbol{\sigma_C}, \boldsymbol{\sigma_{-C}}), \pi) : \boldsymbol{\sigma_{-C}} \in \prod_{a \in \mathrm{Ag}_{-C}} Str_a\}$ of probability measures that the players in C enforce when they follow the strategy $\boldsymbol{\sigma_C}$, namely, for each $a \in \mathrm{Ag}$, player a follows strategy σ_a in $\boldsymbol{\sigma_C}$. We use $\mu_{\pi}^{\boldsymbol{\sigma_C}}$ to range over the measures in $out_C(\boldsymbol{\sigma_C}, \pi)$.

Definition 2. $\mathsf{PATL_b}$ formulas are interpreted in a stochastic system $\mathcal{G}$ and a path π. Most of the semantics is classical, except for the coalition operator, defined as follows[1]:

$$\mathcal{G}, \pi \models \langle\langle C \rangle\rangle_k^{\bowtie d}\varphi \quad \text{iff } \exists \boldsymbol{\sigma_C} \in \prod_{a \in C} Str_{a,k} \text{ s.t. } \forall \mu_{\pi_0}^{\boldsymbol{\sigma_C}} \in out_C(\boldsymbol{\sigma_C}, \pi_0),$$
$$\mu_{\pi_0}^{\boldsymbol{\sigma_C}}(\{\pi' : \mathcal{G}, \pi' \models \varphi\}) \bowtie d$$

3 THE ROBUST $\mathsf{PATL_b}$ MODEL CHECKING PROBLEM

We introduce the model checking problem for $\mathsf{PATL_b}$. Three cases are of interest: (i) In the most general case, we have an arbitrary *parametric system.* (ii) When every transition may be perturbed with at most a fixed ε, we have *interval perturbations* on the system. (iii) When only a few critical components have an uncertain behavior, we can assume the number of perturbed transitions is *fixed.*

Definition 3. For a parametric system $\mathcal{G}$, state $s \in \mathrm{St}$, and formula φ in $\mathsf{PATL_b}$, the parametric model checking problem for $\mathsf{PATL_b}$ consists in deciding if for all well-defined valuations, $\mathcal{G}[Val], s \models \varphi$.

Example 1. *Let us consider a system $\mathcal{G}_{river}$ with two companies sharing the usage of a river. At every step, each company has two available actions: discharge wastewater directly into the river (action d) or treat it before discharging it into the river (action t). Atomic propositions state whether the river's water quality has reached low (proposition low) or high levels (proposition high). The system is shown in Figure 1. The propositions low and high are true only in state q_{low} and q_{high}, resp. In state q_{normal}, no proposition is true, representing that the water quality is normal. If both companies discharge the wastewater, the water quality is guaranteed to decrease (from high to normal, and normal to low). Similarly, if both companies treat the water, the quality will increase (from low to normal, and normal to high). When only one company treats the water while the other discharges the wastewater, the effect has a degree of uncertainty.*

When in q_{normal}, the probability that the water quality increases when only one company treats the water may be slightly higher than planned, and so the probability to go to q_{high} is $0.5 + x$. As a consequence, the probability to go to q_{low} under this situation is $0.5 - x$, to compensate. At the same time, the probability to stay in q_{high} with only one company treating water may be slightly lower than expected. but follows the same trend, hence it is $0.75 - x$. Still, we can check that, as long as x stays within 0 and 0.25, a company alone can always make sure to have at least probability 0.375 of having a good water quality within two time steps by treating the water.

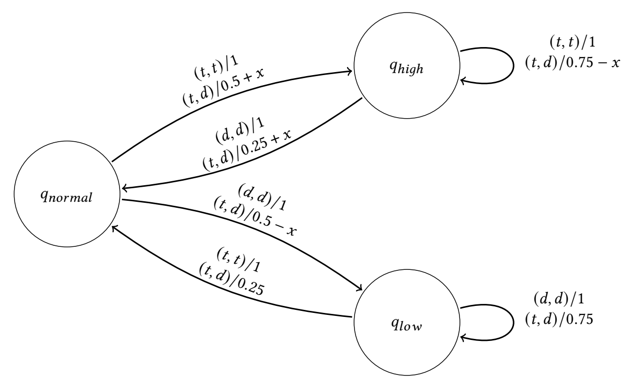

Figure 1: The parametric system $\mathcal{G}_{river}^{p}$ where some transitions may change together depending on $x \in [0, 0.25]$.

4 CONCLUSION

This paper introduces the intricate problem of verifying the robustness of strategies for agents operating within stochastic multi-agent systems (MAS) that are prone to perturbations or variations. We introduced $\mathsf{PATL_b}$, a logic tailored for reasoning about observation-based bounded memory strategies, which we model as automata. In our context, perturbations are parameterized, and we explored two distinct cases: one where the number of parameters is fixed, and another where the perturbations can assume any value within a bounded interval.

The robust model-checking problem guarantees that strategies are resilient to various types of perturbations in models, which is crucial in applications where exact probabilities of random events are imprecise. Considering bounded memory allows agents to retain relevant information while avoiding the undecidability issues associated with the combination of perfect recall and imperfect information in ATL-based formalisms [10].

ACKNOWLEDGMENTS

This research has been supported by the EU H2020 Marie Skłodowska-Curie project with grant agreement No 101105549, PRIN project RIPER (No. 20203FFYLK), PNRR MUR project PE0000013-FAIR, and DFG Project POMPOM (KA 1462/6-1).

[1] See [6] for the complete definition.

REFERENCES

[1] Rajeev Alur, Thomas A. Henzinger, and Orna Kupferman. 2002. Alternating-time temporal logic. *J. ACM* 49, 5 (2002), 672–713. https://doi.org/10.1145/585265.585270

[2] Christel Baier, Christian Hensel, Lisa Hutschenreiter, Sebastian Junges, Joost-Pieter Katoen, and Joachim Klein. 2020. Parametric Markov chains: PCTL complexity and fraction-free Gaussian elimination. *Information and Computation* 272 (2020), 104504.

[3] Christel Baier and Joost-Pieter Katoen. 2008. *Principles of model checking*. MIT Press.

[4] Francesco Belardinelli, Wojciech Jamroga, Munyque Mittelmann, and Aniello Murano. 2023. Strategic Abilities of Forgetful Agents in Stochastic Environments. In *KR*. IJCAI Organization, 726–731.

[5] Francesco Belardinelli, Wojtek Jamroga, Munyque Mittelmann, and Aniello Murano. 2024. Verification of Stochastic Multi-Agent Systems with Forgetful Strategies. In *AAMAS*. Int. Foundation for Autonomous Agents and Multiagent Systems / ACM, Richland, SC, 160–169.

[6] Raphaël Berthon, Joost-Pieter Katoen, Munyque Mittelmann, and Aniello Murano. 2024. Natural Strategic Ability in Stochastic Multi-Agent Systems. In *AAAI*. AAAI Press, 17308–17316.

[7] Taolue Chen, Tingting Han, and Marta Z. Kwiatkowska. 2013. On the complexity of model checking interval-valued discrete time Markov chains. *Inf. Process. Lett.* 113, 7 (2013), 210–216. https://doi.org/10.1016/J.IPL.2013.01.004

[8] Taolue Chen and Jian Lu. 2007. Probabilistic alternating-time temporal logic and model checking algorithm. In *Proc. of FSKD*. IEEE Computer Society, 35–39.

[9] Julien Cristau, Claire David, and Florian Horn. 2010. How do we remember the past in randomised strategies?. In *GANDALF (EPTCS, Vol. 25)*. 30–39.

[10] Catalin Dima and Ferucio Laurentiu Tiplea. 2011. Model-checking ATL under Imperfect Information and Perfect Recall Semantics is Undecidable. *CoRR* abs/1102.4225 (2011).

[11] Ernst Moritz Hahn, Vahid Hashemi, Holger Hermanns, Morteza Lahijanian, and Andrea Turrini. 2019. Interval Markov decision processes with multiple objectives: from robust strategies to Pareto curves. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 29, 4 (2019), 1–31.

[12] Sebastian Junges, Joost-Pieter Katoen, Guillermo A. Pérez, and Tobias Winkler. 2021. The complexity of reachability in parametric Markov decision processes. *J. Comput. Syst. Sci.* 119 (2021), 183–210. https://doi.org/10.1016/J.JCSS.2021.02.006

[13] Thomas M Moerland, Joost Broekens, Aske Plaat, Catholijn M Jonker, et al. 2023. Model-based reinforcement learning: A survey. *Foundations and Trends® in Machine Learning* 16, 1 (2023), 1–118.

[14] Marnix Suilen, Nils Jansen, Murat Cubuktepe, and Ufuk Topcu. 2020. Robust Policy Synthesis for Uncertain POMDPs via Convex Optimization. In *IJCAI*. ijcai.org, 4113–4120.