

Strategy Synthesis for Partially Observable Stochastic Games with Neural Perception Mechanisms

Marta Kwiatkowska   

Department of Computer Science, University of Oxford, UK

Abstract

Strategic reasoning is essential to ensure stable multi-agent coordination in complex environments, as it enables synthesis of optimal (or near-optimal) agent strategies and equilibria that guarantee expected outcomes, even in adversarial scenarios. *Partially-observable stochastic games* (POSGs) are a natural model for real-world settings involving multiple agents, uncertainty and partial information, but lack practical algorithms for computing or approximating optimal values and strategies. Recently, progress has been made for *one-sided POSGs*, a subclass of two-agent, zero-sum POSGs where only one agent has partial information while the other agent is assumed to have full knowledge of the state, with heuristic search value iteration (HSVI) proposed for computing approximately optimal values and strategies in one-sided POSGs [1]. This model is well suited to safety-critical applications, when making worst-case assumptions about one agent; examples include the attacker in a security application, modelled, e.g., as a patrolling or pursuit-evasion game.

However, many realistic autonomous coordination scenarios involve agents perceiving *continuous environments* using *data-driven* observation functions, typically implemented as neural networks (NNs). Examples include autonomous vehicles using NNs to perform object recognition or to estimate pedestrian intention, or NN-enabled vision in an airborne pursuit-evasion scenario. Such perception mechanisms bring new challenges, notably continuous environments, which are inherently tied to NN-enabled perception because of standard training regimes. This means that naive discretisation is difficult, since decision boundaries obtained for data-driven perception are typically irregular and can be misaligned with gridding schemes for discretisation, affecting the precision of the computed strategies.

This invited paper will discuss progress with developing a model class and algorithms for one-sided POSGs with neural perception mechanisms [2, 3] that work directly with their continuous state space. Building on continuous-state POMDPs with NN perception mechanisms [4], where the key idea is that ReLU neural network classifiers induce a finite decomposition of the continuous environment into polyhedra for each classification label, a piecewise constant representation for the value, reward and perception functions is developed that forms the basis for a variant of HSVI, a point-based solution method that computes a lower and upper bound on the value function from a given belief to compute an (approximately) optimal strategy. We extend these ideas from the single-agent (POMDP) setting [4] to zero-sum POSGs. In the game setting, this involves solving a normal form game at each stage and iteration, and goes significantly beyond HSVI for finite POSGs [1].

2012 ACM Subject Classification Theory of computation → Logic and verification; Computing methodologies → Neural networks

Keywords and phrases Stochastic games, neural networks, formal verification, strategy synthesis

Digital Object Identifier 10.4230/LIPIcs.CSL.2024.5

Category Invited Talk

Related Version *Full Version*: <https://arxiv.org/abs/2310.11566>

Funding This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 834115).



© Marta Kwiatkowska;

licensed under Creative Commons License CC-BY 4.0

32nd EACSL Annual Conference on Computer Science Logic (CSL 2024).

Editors: Aniello Murano and Alexandra Silva; Article No. 5; pp. 5:1–5:2

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

References

- 1 Karel Horák, Branislav Bošanský, Vojtěch Kovařík, and Christopher Kiekintveld. Solving zero-sum one-sided partially observable stochastic games. *Artificial Intelligence*, 316:103838, 2023.
- 2 R. Yan, G. Santos, G. Norman, D. Parker, and M. Kwiatkowska. Strategy synthesis for zero-sum neuro-symbolic concurrent stochastic games. *arXiv*, 2022. Accepted to Information and Computation. [arXiv:2202.06255](https://arxiv.org/abs/2202.06255).
- 3 Rui Yan, Gabriel Santos, Gethin Norman, David Parker, and Marta Kwiatkowska. Partially observable stochastic games with neural perception mechanisms. *arXiv*, 2023. [arXiv:2310.11566](https://arxiv.org/abs/2310.11566).
- 4 Rui Yan, Gabriel Santos, Gethin Norman, David Parker, and Marta Kwiatkowska. Point-based value iteration for neuro-symbolic POMDPs. *arXiv*, 2023. [arXiv:2306.17639](https://arxiv.org/abs/2306.17639).