Report from Dagstuhl Seminar 24231

Stochastic Games

Nathanaël Fijalkow*[1], Jan Kretinsky*[2], Ann Nowé*[3], and Gabriel Bathie†[4]

1 **CNRS – Talence, FR.** nathanael.fijalkow@gmail.com
2 **Masaryk University – Brno, CZ.** jan.kretinsky@gmail.com
3 **Free University of Brussels, BE.** ann.nowe@vub.ac.be
4 **University of Bordeaux, FR.** gabriel.bathie@gmail.com

Abstract

The Dagstuhl Seminar on Stochastic Games brought together leading researchers and practitioners in the field to discuss recent advances, challenges, and future directions. The seminar featured a series of tutorials, invited talks, and contributed talks, which provided a comprehensive overview of the latest developments in Markov decision processes, reinforcement learning, and stochastic game theory. The seminar fostered lively discussions during open problem sessions and working groups, culminating in a collaborative exploration of open questions and potential research directions.

Seminar June 2–7, 2024 – https://www.dagstuhl.de/24231

2012 ACM Subject Classification Theory of computation → Algorithmic game theory and mechanism design; Theory of computation → Convergence and learning in games; Theory of computation → Probabilistic computation; Theory of computation → Representations of games and their complexity

Keywords and phrases Algorithmic Game Theory, Algorithms, Optimisation, Reinforcement Learning, Stochastic Games

Digital Object Identifier 10.4230/DagRep.14.6.1

1 Executive Summary

Nathanaël Fijalkow
Jan Kretinsky
Ann Nowé

The Dagstuhl Seminar on Stochastic Games brought together leading researchers and practitioners in the field to discuss recent advances, challenges, and future directions. The seminar featured a series of tutorials, invited talks, and contributed talks, which provided a comprehensive overview of the latest developments in Markov decision processes, reinforcement learning, and stochastic game theory. Key results from the seminar include novel insights into branching stochastic games, the development of new algorithms for solving concurrent and population games, and advancements in the theoretical understanding of efficient solutions for Markov decision processes. The seminar fostered lively discussions during open problem sessions and working groups, culminating in a collaborative exploration of open questions and potential research directions.

* Editor / Organizer
† Editorial Assistant / Collector

Overview of the Invited Talks and Tutorials:

- Yinyu Ye: "Progresses and Open Questions on the Markov Decision / Game Process" – Discussed recent advancements and open problems in the field of Markov decision processes and games.
- Dave Parker: "PRISM-games" – Provided a tutorial on PRISM-games, a tool for modeling and analyzing probabilistic systems.
- Kousha Etessami: "Branching MDPs, Branching Stochastic Games, and Generalizations of Newton's Method" – Explored branching Markov decision processes and games, and introduced generalizations of Newton's method for these models.
- Aaron Sidford: "Theoretical Advances in Efficiently Solving Markov Decision Processes" – Highlighted recent theoretical progress in solving Markov decision processes more efficiently.
- Sven Schewe: "Automata for Profit and Pleasure" – Discussed the applications of automata theory in both practical and theoretical contexts.

Overall, the seminar was a highly productive event, advancing the collective understanding of stochastic games and fostering future research collaborations.

2 Table of Contents

3 Overview of Talks

3.1 Solving irreducible stochastic mean-payoff games and entropy games by relative Krasnoselskii-Mann iteration

Marianne Akian (Inria & CMAP, Ecole polytechnique – Palaiseau, FR)

Joint work of Marianne Akian, Stéphane Gaubert, Ulysse Naepels, Basile Terver
Main reference Marianne Akian, Stéphane Gaubert, Ulysse Naepels, Basile Terver: "Solving Irreducible Stochastic Mean-Payoff Games and Entropy Games by Relative Krasnoselskii-Mann Iteration", in Proc. of the 48th International Symposium on Mathematical Foundations of Computer Science, MFCS 2023, August 28 to September 1, 2023, Bordeaux, France, LIPIcs, Vol. 272, pp. 10:1–10:15, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023.
URL https://doi.org/10.4230/LIPICS.MFCS.2023.10

We analyse an algorithm solving stochastic mean-payoff games, combining the ideas of relative value iteration and of Krasnoselskii–Mann damping. We derive parameterized complexity bounds for several classes of games including turn-based or concurent games satisfying irreducibility or ergodicity conditions. These bounds improve the ones of Chatterjee and Ibsen-Jensen (ICALP 2014), under the same ergodicity condition, and the one of Allamigeon, Gaubert, Katz and Skomra (ICALP 2022) in the particular case of turn-based games. We also establish parameterized complexity bounds for entropy games, a class of matrix multiplication games introduced by Asarin, Cervelle, Degorre, Dima, Horn and Kozyakin (2016). We derive all these results by methods of variational analysis, establishing contraction properties of the relative Krasnoselskii–Mann iteration with respect to Hilbert's semi-norm.

3.2 Improved bounds for strategy improvement algorithms for energy games

Dani Dorfman (Tel Aviv University, IL)

Joint work of Dani Dorfman, Haim Kaplan, Uri Zwick

Strategy improvement is a natural and well-studied family of algorithms for solving various classes of stochastic and deterministic games. We present an improved upper bound of $O(n2^n)$ on the number of iterations performed by the most natural, and most greedy, variant of the algorithm when applied to n-vertex *Energy Games*. We also obtain a similar upper bound on the expected number of iterations performed by *Random-Edge*, one of the most natural randomized variants of the algorithm.

3.3 Activating Formal Verification of Deep Reinforcement Learning Policies by Model Checking Bisimilar Latent Space Models

Florent Delgrange (Free University of Brussels, BE)

Joint work of Florent Delgrange, Ann Nowé, Guillermo A. Pérez
Main reference Florent Delgrange, Ann Nowé, Guillermo A. Pérez: "Distillation of RL Policies with Formal Guarantees via Variational Abstraction of Markov Decision Processes", in Proc. of the Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 – March 1, 2022, pp. 6497–6505, AAAI Press, 2022.
URL https://doi.org/10.1609/AAAI.V36I6.20602

Intelligent agents are computational entities that autonomously interact with an environment to achieve their design objectives. On the one hand, reinforcement learning (RL) encompasses machine learning techniques that allow agents to learn by trial and error a control policy, prescribing how to behave in the environment. Although RL is proven to converge to an optimal policy under some assumptions, the guarantees vanish with the introduction of advanced techniques, such as deep RL, to deal with high-dimensional state and action spaces. This prevents them from being widely adopted in real-world safety-critical scenarios.

On the other hand, formal methods are mathematical techniques that provide guarantees about the correctness of systems. In particular, model checking allows formally verifying the agent's behaviors in the environment. However, this typically relies on a formal description of the interaction, as well as conducting an exhaustive exploration of the state space. This poses significant challenges because the environment is seldom explicitly accessible. Even when it is, model checking suffers from the curse of dimensionality and struggles to scale to high-dimensional state and action spaces, which are common in deep RL.

We propose to tackle this challenge by leveraging the strengths of deep RL to handle realistic scenarios while integrating formal methods to provide guarantees on the agent's behaviors. Specifically, we enable formal verification of deep RL policies by learning a latent model of the environment, over which we distill the deep RL policy. The outcome is amenable for model checking and is endowed with bisimulation guarantees, which allows to lift the verification results to the original environment.

3.4 Branching MDPs, branching stochastic games, and generlizations of Newton's method

Kousha Etessami (University of Edinburgh, GB)

Over the last 20 years there has been a large body of research in theoretical computer science and verification on algorithms and complexity of analyzing and model checking infinite-state, but finitely presented, Markov chains, Markov decision processes (MDPs), and stochastic games. Many of these models add probabilistic/control/game behavior to some classic automata-theoretic models, like context-free grammars, pushdown automata, one-counter automata, etc. These models, it turns out, are also intimately related to some classic stochastic processes.

In this talk I will give a flavor of one piece of this research. I will focus on a series of results I have been involved with on algorithms and complexity of analyzing Multi-type Branching processes, Branching MDPs, and Branching stochastic games. A key aspect of these results is new algorithms and complexity bounds based on (generalizations of) Newton's method for computing the least fixed point solution for systems of monotone/probabilistic (min/max)-polynomial equations. Such equations arise, e.g., as Bellman optimality equations for Branching MDPs.

(Based on a series of joint works (2005-2020) with Alistair Stewart and Mihalis Yannakakis.)

3.5 Solving tropical polynomial systems using parametric mean-payoff games

Stéphane Gaubert (INRIA & CMAP, Ecole polytechnique – Palaiseau, FR)

Joint work of Stéphane Gaubert, Marianne Akian, Antoine Béreau
Main reference Marianne Akian, Antoine Béreau, Stéphane Gaubert: "The Tropical Nullstellensatz and Positivstellensatz for Sparse Polynomial Systems", in Proc. of the 2023 International Symposium on Symbolic and Algebraic Computation, ISSAC 2023, Tromsø, Norway, July 24-27, 2023, pp. 43–52, ACM, 2023.
URL https://doi.org/10.1145/3597066.3597089

Grigoriev and Podolskii (2018) have established a tropical analogue of the effective Nullstellensatz, showing that a system of tropical polynomial equations is solvable if and only if a linearized system obtained from a truncated Macaulay matrix is solvable. They provided an upper bound of the minimal admissible truncation degree, as a function of the degrees of the tropical polynomials. We establish a tropical Nullstellensatz adapted to *sparse* tropical polynomial systems. Our approach is inspired by a construction of Canny-Emiris (1993), refined by Sturmfels (1994). This leads to an improved bound of the truncation degree, which coincides with the classical Macaulay degree in the case of $n + 1$ equations in n unknowns. We also establish a tropical Positivstellensatz, allowing one to decide the inclusion of tropical basic semialgebraic sets. This allows one to reduce decision problems for tropical semi-algebraic sets to the solution of systems of tropical linear equalities and inequalities. Finally, we shall discuss the recent development of a tropical analogue of the eigenvalue method for polynomial system solving: we show how to compute solutions of systems of tropical polynomial (in)equalities using parametric mean payoff games.

3.6 Similarities between ARRIVAL and Simple Stochastic Games

Sebastian Haslebacher (ETH Zürich, CH)

The two problems ARRIVAL and Simple Stochastic Games (SSG) share the same complexity status: Both are known to lie in NP and coNP, but not known to lie in P. Besides this, it turns out that recent algorithms for both problems use similar ideas. Concretely, I talked about similarities underlying three results in the area:

- A polynomial-time algorithm for SSG on almost acyclic graphs due to Auger, Coucheney, Strozecki [1],
- the subexponential-time algorithm for ARRIVAL due to Gärtner, Haslebacher, Hoang [3],
- and the quasi-polynomial-time algorithm for SSG on graphs of bounded treewidth due to Chatterjee, Meggendorfer, Saona, Svoboda [2].

The main question is whether the common framework of these three results can be further exploited to give better algorithms for either problem.

References

1 David Auger, Pierre Coucheney, Yann Strozecki. *Finding Optimal Strategies of Almost Acyclic Simple Stochastic Games*. Theory and Applications of Models of Computation, 8402:67–85, 2014.

2 Krishnendu Chatterjee, Tobias Meggendorfer, Raimundo Saona, Jakub Svoboda. *Faster Algorithm for Turn-based Stochastic Games with Bounded Treewidth*. In Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), pages 4590–4605, 2023.

3 Bernd Gärtner, Sebastian Haslebacher, Hung P. Hoang. *A Subexponential Algorithm for ARRIVAL*. In 48th international colloquium on automata, languages, and programming (ICALP), pages 69:1–69:14, 2021.

3.7 Open Problems in Parametric Markov Decision Processes

Sebastian Junges (Radboud University Nijmegen, NL)

Joint work of Sebastian Junges, Joost-Pieter Katoen, Guillermo A. Pérez, Tobias Winkler
Main reference Sebastian Junges, Joost-Pieter Katoen, Guillermo A. Pérez, Tobias Winkler: "The complexity of reachability in parametric Markov decision processes", J. Comput. Syst. Sci., Vol. 119, pp. 183–210, 2021.
URL https://doi.org/10.1016/J.JCSS.2021.02.006

This talk considers parametric Markov decision processes (pMDPs), an extension to Markov decision processes (MDPs) where transitions probabilities are described by polynomials over a finite set of parameters rather than precise values. Parametric MDPs have been studied in the context of verifying robust randomizing algorithms and to study decision making in partially observable environments. We first review results regarding the complexity of finding values for these parameters such that the induced MDP satisfies some maximal or minimal reachability probability constraints and then discuss three intriguing open problems.

3.8 Complexity and Representations of Controllers in Reactive Synthesis

James C. A. Main (F.R.S.-FNRS & UMONS – University of Mons, BE)

Joint work of James C. A. Main, Mickael Randour

We consider controller synthesis via game theory. Controllers are derived from strategies of players in games whenever suitable strategies can be computed. Traditionally, strategies are represented by Mealy machines, i.e., automata with outputs along their transitions. This model has been extensively studied and is particularly well-suited for automata-based game solving techniques.

The Mealy machine standard for strategy representation, in spite of its established usefulness, may not be the most relevant in certain contexts. For instance, some strategies require exponential-size Mealy machines despite admitting a small program-like representation based on counters that is thus better suited for controller implementation. In a certain sense, Mealy machines can somewhat obfuscate the structure of strategies and their complexity (e.g., with respect to implementation). These concerns have recently led to a surge of alternative models to represent strategies such as decision trees, neural networks and programs.

We motivate a multifaceted vision of strategy complexity. For instance, strategy representations provide natural measures of complexity (e.g., their size). These measures are not necessarily directly comparable to one another. We are also interested in the applicability and relevance of strategy representations, and their relationships. In addition to model-based notions of complexity, measures that are independent of the chosen representation are also relevant, such as the inclusion of randomness in decision making.

This talk is based on joint work with Mickael Randour.

3.9 Strategy shapes for population games

Corto Mascle (University of Bordeaux, FR)

Joint work of Corto Mascle, Hugo Gimbert, Patrick Totzke

Population games are played on a non-deterministic finite automaton. Some number of tokens are placed on an initial state. At each round, one player picks a letter x, and the other answers by moving each token along an x-transition. The first player wins if she can make all tokens reach a final state, for any initial number of tokens.

In this talk we will survey the results obtained so far on this model, with a focus on the model where the second player picks transitions at random. We will sketch the latest algorithm to compute a winning strategy for the first player in that case, obtained with Hugo Gimbert and Patrick Totzke. We will also present and motivate several open problems concerning this model.

3.10 Trustworthy Reinforcement Learning, challenges and opportunities

Ann Nowé (Free University of Brussels, BE)

Main reference References are mentioned on the slides

Abstract Reinforcement Learning (RL) has long outgrown the traditional representations that guaranteed policy convergence but severely limited its application to complex domains. Modern Deep RL enables far richer and complex behaviour, yet at the cost of transparency and explainability, as well as convergence guarantees. The same has been observed in Multi-agent RL settings, where in the past it has been shown that independent learning agents, i.e. not having access to the states nor actions of the other agents, could converge to interesting solution concepts, especially. When equipped with an additional protocol, fair solutions could even be obtained. Today the state-of-the art is Centralised learning, decentralized execution, and while it can handle larger state spaces, it gives in on distributivity. After discussing some convergence proofs which hold for tabular settings, some recent developments for policy distillation were discussed and for providing formal guarantees (see also talk by Florent Delgrange).

3.11 PRISM games: Model Checking for Stochastic Games

David Parker (University of Oxford, GB)

This tutorial presents an overview of PRISM-games, which is a tool for probabilistic model checking of stochastic games. This technique provides formal modelling and analysis of multi-agent systems with probabilistic behaviour. I will describe ways to model such systems using turn-based or concurrent stochastic games, and to formally specify desired properties of the agents' strategies using probabilistic temporal logic. I will discuss the techniques that PRISM-games implements for model checking these logics, covering both zero-sum and equilibria-based properties, and present some illustrative case studies.

3.12 Synthesizing "more probabilistic" systems

Jakob Piribauer (TU Dresden, DE)

In order to provide guarantees on the behavior of a system, probabilism and non-determinism pose different challenges. Non-determinism typically necessitates a worst-case analysis, while probabilism allows for more nuanced assurances. This raises the question whether systems in which uncertainty is mainly subject to probabilism are more desirable. This talk presents an idea on how to measure the influence of non-determinism and probabilism, respectively, on the uncertainty of some quantitative aspect, e.g., the runtime, of a system. The talk aims to open a discussion on whether "more probabilistic" systems are indeed desirable and in which situations using such an objective (potentially in conjunction with further objectives) might be useful.

3.13 Automata for Profit and Pleasure

Sven Schewe (University of Liverpool, GB)

Joint work of Ernst Moritz Hahn, Mateo Perez, Sven Schewe, Fabio Somenzi, Qiyi Tang, Ashutosh Trivedi, Dominik Wojtczak, Tansholpan Zhanabekova
Main reference Ernst Moritz Hahn, Mateo Perez, Sven Schewe, Fabio Somenzi, Ashutosh Trivedi, Dominik Wojtczak: "Good-for-MDPs Automata for Probabilistic Analysis and Reinforcement Learning", in Proc. of the Tools and Algorithms for the Construction and Analysis of Systems – 26th International Conference, TACAS 2020, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2020, Dublin, Ireland, April 25-30, 2020, Proceedings, Part I, Lecture Notes in Computer Science, Vol. 12078, pp. 306–323, Springer, 2020.
URL https://doi.org/10.1007/978-3-030-45190-5_17
Main reference Sven Schewe, Qiyi Tang, Tansholpan Zhanabekova: "Deciding What Is Good-For-MDPs", in Proc. of the 34th International Conference on Concurrency Theory, CONCUR 2023, September 18-23, 2023, Antwerp, Belgium, LIPIcs, Vol. 279, pp. 35:1–35:16, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023.
URL https://doi.org/10.4230/LIPICS.CONCUR.2023.35

What could be greater fun than toying around with formal structures? One particularly beautiful structure to play with are automata over infinite words, and there is really no need to give any supporting argument for the *pleasure* part in the title. But how about profit? When using ω-regular languages as target languages for practical applications like Markov chain model checking, MDP model checking and reinforcement learning, reactive synthesis, or as a target for an infinite word played out in a two player game, the classic approach has been to first produce a deterministic automaton D that recognises that language. This deterministic automaton is quite useful: we can essentially play on the syntactic product of the structure and use the acceptance mechanism it inherits from the automaton as target. This is beautiful and moves all the heavy lifting to the required automata transformations. But when we want even more profit in addition to the pleasure, the question arises whether deterministic automata are the best we can do. They are clearly good enough: determinism is as restrictive as it gets, and easily guarantees that one can just work on the product. But what we really want is the reverse: we want an automaton, so that we can work on the product, and determinism is just maximally restrictive, and therefore good enough for everything. At Dagstuhl, all will know that we can lift quite a few restrictions and instead turn to the gains we can make when we focus on the real needs of being able to work on the product. For Markov chains, this could be unambiguous automata, for MDPs this could be good-for-MDP automata, and for synthesis and games, this could be good-for-games automata. We will shed a light to a few nooks and corners of the vast room available open questions and answers, with a bias towards MDPs analysis in general and reinforcement learning in particular.

3.14 Theoretical Advances in Efficiently Solving Markov Decision Processes

Aaron Sidford (Stanford University, US)

Joint work of Yujia Jin, Ishani Karmarkar, Aaron Sidford, Jiayi Wang
Main reference Yujia Jin, Ishani Karmarkar, Aaron Sidford, Jiayi Wang: "Truncated Variance Reduced Value Iteration", CoRR, Vol. abs/2405.12952, 2024.
URL https://doi.org/10.48550/ARXIV.2405.12952

Markov Decision Processes (MDPs) are a fundamental mathematical model for reasoning about uncertainty and have a foundational role in the theory of reinforcement learning. Over the past decade, there have been substantial advances in the design and analysis of algorithms for provably computing approximately optimal policies for MDPs in a variety of settings. In this talk, I will survey these advances touching upon optimization tools of potential broader utility. Additionally, I will discuss recent joint work with Ishani Karmarkar, Jiayi Wang, and Yujia Jin on solving MDPS (arXiv:2405.12952).

3.15 Solving concurrent games in PSPACE

Patrick Totzke (University of Liverpool, GB)

Joint work of Sougata Bose, Rasmus Ibsen-Jensen, Patrick Totzke
Main reference Sougata Bose, Rasmus Ibsen-Jensen, Patrick Totzke: "Bounded-Memory Strategies in Partial-Information Games", in Proc. of the 39th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2024, Tallinn, Estonia, July 8-11, 2024, pp. 17:1–17:14, ACM, 2024.
URL https://doi.org/10.1145/3661814.3662096

Very recently, progress has been made on new nondeterministic upper bounds for solving concurrent Mean-Payoff and related games, resulting in approximating algorithms at level 2 of the polynomial hierarchy, FNP[NP], for approximating values and equilibria. In this talk I give a brief overview of the technique, which is based on Frederiksen and Miltersen's (ISAAC 2013) work that uses floating point representations to approximate doubly-exponentially small values. We have extended this to mean-payoff objectives under partial info, and for any fixed number of players and shown that this implies improved upper bounds for various well-known types of games. Directions for discussions:

- Is there hope to get it down to plain FNP?
- Ultimately, I would like to use SAT/SMT solvers to implement these procedures. I would welcome any advice or discussions about implementation details: Which libraries etc to use to deal with (doubly exponentially small) probabilities?
- I would be interested in hearing suggestions about more applications, theoretical and practical.

References

1 Søren Kristoffer Stiil Frederiksen and Peter Bro Miltersen. *Approximating the Value of a Concurrent Reachability Game in the Polynomial Time Hierarchy* In International Symposium on Algorithms and Computation (ISAAC 2013). 457–467.

3.16 Decidability of Omega-Regular Objectives for POMDPs with Revelations

Pierre Vandenhove (University of Bordeaux, FR)

Joint work of Pierre Vandenhove, Fijalkow Nathanaël, Hugo Gimbert, Guillermo A. Pérez

We consider partially observable Markov decision processes (POMDPs) with parity objectives. We study the qualitative problem of deciding whether there exists an almost surely winning strategy. Such a problem is undecidable in general, already for coBüchi objectives.

We introduce two decidable properties requiring that, almost surely and infinitely often, the exact state can be deduced (which is called a revelation). Assuming the first property, we show that coBüchi is decidable, while parity with three priorities is still undecidable. Assuming the second property, we show that parity is decidable. Technically, the decidable cases all reduce to the analysis of the finite belief support MDP. We also consider partially observable zero-sum games and show that coBüchi is still undecidable under the revealing properties.

3.17 General-sum stochastic games: Turn-based v.s. simultaneous play

Muthukumar Vidya (Georgia Institute of Technology – Atlanta, US)

Joint work of Muthukumar Vidya, Aaron Sidford
Main reference Yujia Jin, Vidya Muthukumar, Aaron Sidford: "The Complexity of Infinite-Horizon General-Sum Stochastic Games", in Proc. of the 14th Innovations in Theoretical Computer Science Conference, ITCS 2023, January 10-13, 2023, MIT, Cambridge, Massachusetts, USA, LIPIcs, Vol. 251, pp. 76:1–76:20, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023.
URL https://doi.org/10.4230/LIPICS.ITCS.2023.76

We study the complexity of computing stationary Nash equilibrium (NE) in n-player infinite-horizon general-sum stochastic games. We focus on the problem of computing NE in such stochastic games when each player is restricted to choosing a stationary policy and rewards are discounted. First, we prove that computing such NE is in PPAD (in addition to clearly being PPAD-hard). Second, we consider turn-based specializations of such games where at each state there is at most a single player that can take actions and show that these (seemingly simpler) games remain PPAD-hard. Third, we show that under further structural assumptions on the rewards computing NE in such turn-based games is possible in polynomial time. Towards achieving these results we establish structural facts about stochastic games of broader utility, including monotonicity of utilities under single-state single-action changes and reductions to settings where each player controls a single state.

3.18 An Overview of Stochastic Games Benchmarks

Maximilian Weininger (IST Austria – Klosterneuburg, AT)

Evaluating, validating and comparing algorithms for solving stochastic games requires a set of benchmarks. This set should be "realistic", easy-to-use, and structurally diverse. I provide an overview of the state-of-the-art benchmark sets, pointing out their shortcomings and sketching how we can improve them.

My insights on this topic are based on my continued work on algorithms for stochastic games, in particular [1, 2, 3, 4, 5], and the discussions we had about the MDP benchmark set at the previous Dagstuhl Seminar 24134.

References

1 Roman Andriushchenko, Alexander Bork, Carlos E. Budde, Milan Ceska, Kush Grover, Ernst Moritz Hahn, Arnd Hartmanns, Bryant Israelsen, Nils Jansen, Joshua Jeppson, Sebastian Junges, Maximilian A. Köhl, Bettina Könighofer, Jan Kretínský, Tobias Meggendorfer, David Parker, Stefan Pranger, Tim Quatmann, Enno Ruijters, Landon Taylor, Matthias Volk, Maximilian Weininger, and Zhen Zhang. Tools at the frontiers of quantitative verification. *CoRR*, abs/2405.13583, 2024.

2 Muqsit Azeem, Alexandros Evangelidis, Jan Kretínský, Alexander Slivinskiy, and Maximilian Weininger. Optimistic and topological value iteration for simple stochastic games. In *ATVA*, volume 13505 of *Lecture Notes in Computer Science*, pages 285–302. Springer, 2022.

3 Arnd Hartmanns, Sebastian Junges, Tim Quatmann, and Maximilian Weininger. A practitioner's guide to MDP model checking algorithms. In *TACAS (1)*, volume 13993 of *Lecture Notes in Computer Science*, pages 469–488. Springer, 2023.

4 Jan Kretínský, Emanuel Ramneantu, Alexander Slivinskiy, and Maximilian Weininger. Comparison of algorithms for simple stochastic games. *Inf. Comput.*, 289(Part):104885, 2022.

5 Tobias Meggendorfer and Maximilian Weininger. Playing games with your PET: extending the partial exploration tool to stochastic games. *CoRR*, abs/2405.03885, 2024.

4 Working groups

4.1 Auction-Based Scheduling

Guy Avni (University of Haifa, IL)

Joint work of Guy Avni, Kaushik Mallik, Suman Sadhukhan
Main reference Guy Avni, Kaushik Mallik, Suman Sadhukhan: "Auction-Based Scheduling", in Proc. of the Tools and Algorithms for the Construction and Analysis of Systems, pp. 153–172, Springer Nature Switzerland, 2024.
URL https://doi.org/10.1007/978-3-031-57256-2_8

We discussed the details of the auction-based scheduling framework and its required background from bidding games. We discussed possible relations with history deterministic automata. We discussed extensions of the framework in other domains, specifically in MDPs, which we agreed might be an interesting future direction for study.

4.2 Solving Markov Decision Processes by adding a discount

Nathanaël Fijalkow (CNRS – Talence, FR)

In this working group we discussed how we could use the solutions for solving discounted Markov Decision Processes to the undiscounted case. This was an excuse to revisit the existing methods for the discounted case. Some interesting directions were discussed.

4.3 Finding Tarski Fixed Points

Sebastian Haslebacher (ETH Zürich, CH)

The complexity of solving Condon's and (approximately solving) Shapley's stochastic games are longstanding open questions. Etessami et al. [1] proved that both problems reduce to the problem of finding a fixed point of a monotone function on an integer grid. Concretely, such a function maps $\{1,\ldots,N\}^d$ into itself and is monotone with respect to the coordinate-wise partial order $\leq_c$. A well-known theorem of Tarski guarantees that these conditions guarantee the existence of a fixed point. Hence, the computational problem of finding it is often called TARSKI.

The best algorithms to date solve TARSKI using at most $O(\log^{\lceil \frac{2d}{3} \rceil} N)$ queries [2], or $O(\log^{\lceil \frac{d+1}{2} \rceil} N)$ queries [3] if d is considered to be a constant, respectively. All of these algorithms are time efficient, i.e. the next query can be determined in time polynomial in d and $\log N$. Conversely, the best query lower bound is $\Omega(\log^2(N)$ [1]. This means that there is a big gap in terms of query complexity, and closing this gap would be a major achievement.

Recently, Chen et al. [4] gave a query-efficient algorithm for the problem of finding an approximate fixed point of a contraction map (with respect to the infinity norm). Unfortunately, their algorithm is not time-efficient (if it were, this would have tremendous implications e.g. for stochastic games). But naturally, the question arises whether similar techniques could be used to obtain better query upper bounds for Tarski.

The discussion in the working session revolved around these recent results. Concretely, the above mentioned literature was discussed, and we tried to understand whether the techniques from [4] might indeed be useful for Tarski as well.

References

1 Kousha Etessami, Christos Papadimitriou, Aviad Rubinstein, and Mihalis Yannakakis. *Tarski's Theorem, Supermodular Games, and the Complexity of Equilibria*. In 11th Innovations in Theoretical Computer Science Conference (ITCS), pages 18:1–18:19, 2020.
2 John Fearnley, Dömötör Pálvölgyi, and Rahul Savani. *A Faster Algorithm for Finding Tarski Fixed Points*. ACM Transactions on Algorithms, 18(3):1–23, July 2022.
3 Xi Chen and Yuhao Li. *Improved Upper Bounds for Finding Tarski Fixed Points*. In Proceedings of the 23rd ACM Conference on Economics and Computation, EC '22, pages 1108–1118, July 2022.
4 Xi Chen, Yuhao Li, and Mihalis Yannakakis. *Computing a Fixed Point of Contraction Maps in Polynomial Queries*. In Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, pages 1364–1373, June 2024.

4.4 Discussion on Parametric Markov Chains and Partially Observable MDPs

Sebastian Junges (Radboud University Nijmegen, NL)

Main reference Sebastian Junges, Nils Jansen, Ralf Wimmer, Tim Quatmann, Leonore Winterer, Joost-Pieter Katoen, Bernd Becker: "Finite-State Controllers of POMDPs using Parameter Synthesis", in Proc. of the Thirty-Fourth Conference on Uncertainty in Artificial Intelligence, UAI 2018, Monterey, California, USA, August 6-10, 2018, pp. 519–529, AUAI Press, 2018.
URL https://auai.org/uai2018/proceedings/papers/195.pdf

This blackboard session studied the connection between decision making in Partially Observable Markov Decision Processes (POMDPs) and parametric Markov chains (pMCs), which were discussed at length in an earlier talk. After reviewing results from the main reference given below a discussion revolved mostly about the possibilities to extend these results in different directions, e.g., by partitioning the set of parameters in controllable and uncontrollable parameters, by connecting Partially Observable Stochastic Games (or Concurrent Stochastic Games) with pMDPs. This lead to a discussion of the important features which make parameter synthesis in MDPs hard.

4.5 Measures of benchmarks difficulty and strategy structure

Jan Kretinsky (Masaryk University – Brno, CZ)

Main reference Jan Kretínský, Emanuel Ramneantu, Alexander Slivinskiy, Maximilian Weininger: "Comparison of algorithms for simple stochastic games", Inf. Comput., Vol. 289(Part), p. 104885, 2022.
URL https://doi.org/10.1016/J.IC.2022.104885

This break-out session was concerned with two questions. Firstly, how to measure the difficulty and structure of benchmarks so that we can pinpoint what makes real ones easy (or hard) and what to focus one in order to design practical algorithms [1]. Second, how to describe the fine structure of strategies and utilize it in order to represent them (i) explainably and (ii) so that they are easier to compute [2].

References

1 Jan Kretinsky, Emanuel Ramneantu, Alexander Slivinskiy, Maximilian Weininger. Comparison of Algorithms for Simple Stochastic Games. In GandALF, pp. 131-148. EPTCS, 2020.
2 Pranav Ashok, Mathias Jackermeier, Jan Kretinsky, Christoph Weinhuber, Maximilian Weininger, Mayank Yadav. dtControl 2.0: Explainable Strategy Representation via Decision Tree Learning Steered by Experts. In TACAS (2), pp. 326-345. Springer, 2021.

4.6 Speed of victory in population games

Corto Mascle (University of Bordeaux, FR)

Joint work of Corto Mascle, Nathanaël Fijalkow

In 2017, Bertrand, Dewaskar, Genest and Gimbert introduced a population games, inspired by questions from biology on the control of yeast populations [1]. A population game is described by an NFA with one initial and one final state. Some number of tokens are placed on an initial state of an NFA. Two players then play alternately: Controller chooses a letter a, then Environment moves each token along an a-labelled transitions chosen at random. Controller wins if all tokens end up on the same final state eventually. The question is then: does Controller have a winning strategy against any number of tokens?

A randomized version of those games was later presented by Colcombet, Fijalkow and Ohlmann [2]: this is the model studied in this working group. In this framework, the Environment picks the next transition of each token uniformly at random. The new question is: Can Controller win with probability 1 against any number of tokens? That first paper established decidablity of this problem. However, the expected time needed by Controller to win may in general be exponential in the number of tokens.

This working group was dedicated to the search for decidable characterizations of games where Controller can win in polynomial and polylogarithmic time. We made progress in several directions, Most notably, we built a clear plan to characterize the polylogarithmic case, leaving reducing the initial problem to a couple of lemmas that we believe we can prove with a little more work.

References

1 Nathalie Bertrand and Miheer Dewaskar and Blaise Genest and Hugo Gimbert. *Controlling a population*. CONCUR 2017., https://doi.org/10.4230/LIPIcs.CONCUR.2017.12, 2017
2 Thomas Colcombet and Nathanaël Fijalkow and Pierre Ohlmann. *Controlling a random population*. Log. Methods Comput. Sci., https://doi.org/10.46298/lmcs-17(4:12)2021, 2021

Participants

- Marianne Akian
Inria & CMAP, Ecole polytechnique – Palaiseau, FR
- Xavier Allamigeon
INRIA & Ecole Polytechnique – Palaiseau, FR
- Guy Avni
University of Haifa, IL
- Christel Baier
TU Dresden, DE
- Gabriel Bathie
University of Bordeaux, FR
- Eline Bovy
Radboud University Nijmegen, NL
- Antonio Casares
University of Bordeaux, FR
- Florent Delgrange
Free University of Brussels, BE
- Dani Dorfman
Tel Aviv University, IL
- Kousha Etessami
University of Edinburgh, GB
- Nathanaël Fijalkow
CNRS – Talence, FR
- Stéphane Gaubert
INRIA & CMAP, Ecole polytechnique – Palaiseau, FR
- Kristoffer Arnsfelt Hansen
Aarhus University, DK
- Sebastian Haslebacher
ETH Zürich, CH
- Sebastian Junges
Radboud University Nijmegen, NL
- Marcin Jurdzinski
University of Warwick – Coventry, GB
- Jan Kretinsky
Masaryk University – Brno, CZ
- Antonin Kucera
Masaryk University – Brno, CZ
- Karoliina Lehtinen
Aix-Marseille University, FR
- Georg Loho
FU Berlin, DE & University of Twente – Enschede, NL
- James C. A. Main
F.R.S.-FNRS & UMONS – University of Mons, BE
- Corto Mascle
University of Bordeaux, FR
- Théo Matricon
University of Bordeaux, FR
- Petr Novotný
Masaryk University – Brno, CZ
- Ann Nowé
Free University of Brussels, BE
- David Parker
University of Oxford, GB
- Guillermo A. Pérez
University of Antwerp, BE
- Jakob Piribauer
TU Dresden, DE
- Roxana Radulescu
VU – Brussels, BE
- Sven Schewe
University of Liverpool, GB
- Guruprerana Shabadi
IMDEA Software Institute – Madrid, ES
- Aaron Sidford
Stanford University, US
- Mateusz Skomra
LAAS – Toulouse, FR
- Yann Strozecki
University of Versailles, FR
- Marnix Suilen
Radboud University Nijmegen, NL
- K. S. Thejaswini
IST Austria – Klosterneuburg, AT
- Patrick Totzke
University of Liverpool, GB
- Pierre Vandenhove
University of Bordeaux, FR
- Muthukumar Vidya
Georgia Institute of Technology – Atlanta, US
- Maximilian Weininger
IST Austria – Klosterneuburg, AT
- Yinyu Ye
Stanford University, US & Shanghai JiaoTong University, CN