Workshop on Game Theory and Machine Learning

LSE, Th-Fr 19-20 October 2023

Talk titles and abstracts

Ted Turocy and Rahul Savani, University of East Anglia, University of Liverpool, and The Alan Turing Institute

Title: Automated Analysis of Strategic Interactions: The Roadmap for Gambit 17 and Beyond

Gambit, http://www.gambit-project.org/, is the main software package for doing computational game theory. This talk will be about a new project, funded by the Alan Turing Institute, that will see Gambit significantly improved and extended over the next three years. We want to update and extend Gambit's support for equilibrium computation, which was Gambit's traditional focus, as well as integrating it with modern techniques for solving huge games, such as Multi-agent Reinforcement Learning. The talk is designed to encourage discussion and input about the direction of the project, and how Gambit can best serve the needs of the computational game theory community.

Galit Ashkenazi-Golan, Sahar Jahani, Katerina Papadaki, Edward Plumb, Bernhard von Stengel, Department of Mathematics, LSE

Title: What we learned from EquiLearn

Abstract: EquiLearn, funded by the CIVICA consortium from November 2022 to October 2023, is a research project to study machine learning by multiple agents in an interactive environment. We consider a model of competition between firms of a duopoly with demand inertia, decades ago invented by Selten and studied with behavioral experiments. The learning system is modeled in two stages in a "double oracle" approach. The first stage uses reinforcement learning in the dynamic pricing game, in order to find a good pricing strategy. The second stage models a population of strategies which are in equilibrium and define the learning environment for the first stage. A newly learned strategy, if successful, is added to the population, for which a new equilibrium is computed. Reinforcement learning at the first stage, which we tried with many approaches, is a computational bottleneck; finding an equilibrium at the second stage is decidedly not. We discuss these results and the planned next steps, which include: applying the framework to simpler games with similar characteristics of being able to cooperate with, exploit, or fight an opponent; comparing it with other pricing games where ML agents learn how to collude; considering games with more

than two players; and studying how individual gain is compatible with evolutionary success.

Yongzhao Wang, University of Liverpool and The Alan Turing Institute

Title: An Introduction to Empirical Game-Theoretic Analysis

Abstract: The methodology of empirical game-theoretic analysis (EGTA) provides a broad toolbox of techniques for game reasoning with models based on simulation data. As many multiagent systems of interest are not easily expressed or tackled analytically, EGTA offers an alternative approach whereby a space of strategies is examined through simulation, combined with game model induction and inference. The number of strategies that can be explicitly incorporated in game models is significantly limited by computational constraints, hence the selection of strategies to include is pivotally important. For accurate analysis results, we require that the included strategies are high-performing and cover the key strategic issues. The challenge of efficiently assembling an effective portfolio of strategies for EGTA is called the strategy exploration problem. This talk introduces modern developments and applications of EGTA, with a focus on addressing the strategy exploration problem

Konstantinos Varsos, University of Liverpool and The Alan Turing Institute

Title: Forward Looking Best-Response Multiplicative Weights Update Methods for Bilinear Zero-Sum Games

Abstract: We focus on learning algorithms for the classical problem of finding Nash equilibria in bilinear zero-sum games. The proposed method, which is a variant of Optimistic Mirror Descent, uses a large learning rate for the intermediate gradient step which essentially leads to computing (approximate) best response strategies against the profile of the previous iteration. Although counter-intuitive at first sight due to the irrationally large, for an iterative algorithm, intermediate learning step, we prove that the method guarantees last-iterate convergence to an equilibrium. Particularly, we show that the algorithm reaches first an epsilon-approximate Nash equilibrium, until the method becomes a contracting map, and converges to the exact equilibrium. Furthermore, we provide experimental comparisons with the optimistic variant of the multiplicative weights update method and show the significant practical potential of our algorithm since it offers substantial gains in terms of accelerated convergence. Finally, we present recent research directions in learning dynamics.

Clemens Possnig, University of Waterloo, Canada

Title: Strategic Communication and Algorithmic Advice

Abstract: We study the strategic interaction between a machine-learning-based algorithmic recommendation system and a rational, best-responding receiver when the two have slightly different objectives. Our running example is a revenue maximizing platform (e.g. Airbnb) that offers price recommendations to a profit maximizer who faces positive marginal costs. We show a robust pattern where the algorithm learns to play a cut-off strategy pooling messages when its private information suggests actions in the direction of its preference bias, while sending mostly separate signals otherwise. We characterize this pattern in terms of primitives and show that it improves on welfare against the best cheap-talk equilibrium of Crawford-Sobel.

Dominik Karos, University of Bielefeld

Title: From prejudice to racial profiling and back

Abstract: A designer conducts random searches to detect criminals, and may condition the search probability on individuals' appearance. She updates her belief about the distribution of criminals across appearances using her search results, but incorrectly takes her sample distribution for the population distribution. In equilibrium she employs optimal search probabilities given her belief, and her belief is consistent with her findings. We provide sufficient conditions for the existence of an equilibrium and show that she will be discriminating against an appearance if and only if she overestimates the probability of this appearance being criminal. Notably, the ranking of two appearances' being criminal may be reversed in equilibrium.

Andrea Celli, Bocconi University

Title: No-Regret Learning in Bilateral Trade via Global Budget Balance

Abstract: Bilateral trade revolves around the challenge of facilitating transactions between two strategic agents — a seller and a buyer — both of whom have a private valuations for the item. We study the online version of the problem, in which at each time step a new seller and buyer arrive. The learner's task is to set a price for each agent, without any knowledge about their valuations. The sequence of sellers and buyers is chosen by an oblivious adversary. In this setting, known negative results rule out the possibility of designing algorithms with sublinear regret when the learner has to guarantee budget balance for each iteration. We introduce the notion of global budget balance, which requires the agent to be budget balanced only over the entire time horizon. By requiring global budget balance, we provide the first no-regret algorithms for bilateral trade with adversarial inputs under various feedback models (full feedback and one-bit feedback). We complement these results with a nearly-matching lower bound.

Maryam Kamgarpour, EPFL Lausanne

Title: Learning Equilibria with Payoff-Based Information

Abstract: Nash equilibria are sought-after outcomes in a non-cooperative multi-agent system as no agent finds it profitable to unilaterally deviate from her Nash equilibrium strategy. How can players learn equilibrium strategies based only on the payoff information, namely, evaluations of their objective functions? This question arises in applications ranging from auctions and routing to power markets, where closed-form expression of a player's objective function is a priori unavailable. I will discuss our approach to address this problem in static repeated games leveraging monotonicity of the game. For non-monotone games with continuous actions, I will discuss our approach to designing an efficient no-regret algorithm.

Cesare Carissimo, ETH Zurich

Title: Cooperative Outcomes of Q-Learning in Congestion Games

Abstract: Traffic on roads, packets on the Internet, and electricity on power grids share a structure abstracted in congestion games, where self-interested behaviour can lead to socially sub-optimal results. However, through simulation experiments we find that Q-learning algorithms show a potential for seemingly collaborative outcomes. These results closely resemble those of collusion in pricing games. In this talk we will fist discuss examples of congestion games where this behaviour arises and then introduce a model of recommendation which can amplify these effects. Our simulation code will be made available as an open source package.

Bei Peng, University of Liverpool

Title: Overcoming Relative Overgeneralisation for Cooperative Multi-Agent Reinforcement Learning

Abstract: Many real-world problems involve multiple agents acting and interacting in a shared environment to achieve some common goal, which can be naturally modelled as cooperative multi-agent systems. Multi-agent reinforcement learning (MARL) can be used to learn optimal decision making in many of these complex uncertain and dynamic multi-agent systems. In this talk, I first introduce the MARL paradigm and overview some of the key challenges in developing MARL algorithms that can efficiently learn decentralised policies for a group of agents. I then focus on discussing one of our recent works that addresses the multi-agent pathology of relative overgeneralisation, which can prevent agents from solving cooperative tasks requiring significant coordination. In this work, we propose a new MARL approach called Universal Value Exploration

(UneVEn) to better overcome relative overgeneralisation. To learn a target task exhibiting relative overgeneralisation, UneVEn learns a set of related tasks simultaneously and uses the policies of already solved related tasks to improve the joint exploration process of all agents in the target task. Our empirical results show that UneVEn can solve challenging cooperative tasks where other state-of-the-art MARL methods fail.

Stephanos Leonardos, King's College London

Title: Exploration-Exploitation in Multi-Agent Learning

Abstract: Exploration-exploitation is a practical tool in multi-agent learning; however, its effects in the performance of learning algorithms are far from understood. In this talk, we study smooth Q-learning (SQL), a continuous-time variant of the standard Q-learning algorithm. We show (1) that SQL has bounded regret in arbitrary games for a cost model that explicitly captures the trade-off between utility maximization and exploration costs and (2) that it always converges to the set of quantal-response equilibria (QRE), the standard solution concept for games under bounded rationality, in weighted potential (i.e., coordination) games with heterogeneous learning agents. Regarding the effects of exploration in collective system performance, we show that, over time, the system undergoes phase transitions where the number and stability of equilibria can change dramatically given an infinitesimal change to the exploration hyperparameter. We then turn to competitive settings, modelled as weighted zero-sum polymatrix games, and show that SQL always converges to the unique QRE when all agents use positive exploration rates. Fast convergence of SQL to QRE applies regardless of the number of agents and without any need for hyperparameter fine-tuning. As showcased by our experiments, these results provide an algorithmic approach to the currently open problem of equilibrium selection in competitive multi-agent settings.

Maria Polukarov, King's College London

Title: Reaching Collective Decisions

Abstract: Collective decision-making (with voting protocols being its central tool) is a key component of many applications, ranging from large-scale political elections to small-scale decisions such as scheduling a meeting or allocating a local budget. Based on the fact, that voters may have incentives to strategically misreport their preferences, much of the literature in computational social choice focuses on evaluating voting mechanisms by their resistance to strategic behaviours and uses computational complexity as a barrier to them. We take another natural approach and analyse voting scenarios from a game-theoretic perspective, viewing strategic parties as players and

examining possible stable outcomes of their interaction (i.e., equilibria). In particular, we are interested in the outcomes that can result from natural iterative processes, such as best-response dynamics or its restricted variants. Convergence of such procedures is a highly desirable property of the game, since, from a system-wide perspective, it implies that a system has a deterministic stable state that can be reached by the players without any centralised control and/or communication. We derive conditions on how the iterative process should (or not) be restricted to achieve convergence and, on the other hand, offer a framework where such restricted voting dynamics are justified as boundedly rational decisions under uncertainty.