USING MULTI AGENT RE-ENFORCEMENT LEARNING & GAME THEORY TO COMBAT ASYMMETRICAL TRANSACTIONS ONLINE: A LITERATURE REVIEW

APRIL 2021

GROUP 10

ABDULAZIZ ALGAIN (201560413)

SHUBHAM SINGH (201538011)

JONE CHONG (201533109)

USMAN SHOUKAT (201537600)

Abstract

Marketplaces have moved from offline to online, however, online marketplaces have been designed in a way that mimics offline marketplaces, and this has brought a lot of the problems that exist in the real world to the virtual world. We believe that information asymmetry can be resolved using technology, particularly (MARL) Multi-Agent Reinforcement Learning and the proper transfer of information between buyers and sellers. However, part of our hypothesis is that users should be incentivized to share local information with the other party to benefit everyone long term at least to cooperate. We attempted to examine the extant literature and try to combine different mathematical and scientific concepts and techniques to try to see what could work as a potential solution. What we have found that is the extant literature has a number of very promising case studies about using MARL to combat resistance to knowledge transfer, which leads to information asymmetry and ultimately market failure in marketplaces, yet there are not specific examples that consider all the necessary factors that make or break a marketplace. Conclusions: in contrast to other papers which focus on MARL in other areas (ecommerce, robotics and traffic management) our paper believes that MARL can be used to improve online marketplaces, however, it has to be used in conjunction with Game Theory as well as other techniques such as Q-learning to get the best results.

Introduction

Society has clearly benefited from the advancement of technology especially in the field of (AI) Artificial intelligence where complexity prevents agents from making intelligent decisions fast. The rise of AI has helped accelerate adoption of the technology in both ecommerce and marketplaces.

However, we are yet to see MARL used to its full potential in online marketplaces. Similar to agents in the physical world such as humans on roads, or a traffic system, where agents or drivers in this case benefit from MARL, also marketplaces would likely benefit from MARL to increase efficiency and improve satisfaction for both buyers and sellers.

The problem of MARL has been investigated by prior works, but some of these efforts could be making assumptions that could prevent application of MARL in general multiagent problems, such as online marketplaces. Online Marketplaces are a unique example in the sense that they require domain expertise, such in the area psychology, motivation and the art of selling, before it can be applied well. However, we believe that if the interactions are modelled mathematically in the design phase of the marketplace, then it would lessen the need for domain expertise and it would help us break free from the shortcomings that come with a fixed way of looking at problems that come with highly specific expertise.

Our approach, which uses math and science (Game Theory, Reinforcement learning) to understand human interactions in online marketplaces and reinforce beneficial behaviour. We will focus on addressing the peer-to-peer relationship in both the cooperative vs competitive transactions in a multi-agent system that utilizes reinforcement learning. Important to say, that these roles the peers play are not fixed; these agents learn to assume the role of seller and/or buyer at the appropriate moments, and subsequently requesting and providing access to information in order to improve performance and learning to sell or buy more effectively. Novel methods show promise that agents not only can learn to cooperate significantly faster, but also learn to coordinate in tasks where existing methods fail.

Game Theory

Background

Game theory is the strategic interaction of two or more players exposed to a competitive environment, with the aim of maximising their respective payoffs. Game theory is a successor of the zero-sum game. It is utilised in situations where competitors are utility rational and fully aware of the set rules and the consequences of their resultant actions. [1]

Game theory has been the centre of attention of many fields, including economics, political science, biology, et cetera. Its presence in computer science is increasing at a very rapid rate over the recent years. Many areas of computer science, especially artificial intelligence (AI), theory, e-commerce, and networking, are using game theory to solve various problems, which are being occasionally featured in leading journals and conferences. One of the reasons for advancements in this area is a technology push: computer scientists get characterised by the same mindset as mathematics and scientific game theory.

Ongoing Work

Game theory is becoming increasingly popular in logical designing. Computer Science uses game semantics for interactive computations, and most of the theoretical concepts of multi-agent systems and reinforcement learning are based on game theory. Game theory is proving very helpful in predicting the prices and outcome of releases of different products in competitive settings. [2]

Most of the areas where computer science and game theory are collaborating today are compact game representations, the complexity and algorithms for computing solutions, algorithmic aspects of mechanism design, game-theoretic analysis inspired by specific applications, multi-agent learning, logics of knowledge and belief, and other logical aspects of games.

Nash Equilibrium

Nash equilibrium gives the best outcome in a competitive environment, meaning that none of the players can increase their payoffs after reaching the Nash Equilibrium. For an algorithm to achieve convergence to Nash equilibria, the modelling information of the game is required, and assumption is taken that players can observe other players' actions. [3]

The ongoing work in this area is to look for the feasibility of network growth of two or more players. The following open questions are yet to be answered. Considering that the networks will be built by selfish and rational players, who of which would decide to build relations? Will the resulting network be functional? How would the influence be spread through the network? See [Fabrikant et al. 2003; Kempe et al. 2003] to look for more on this emerging area of computer science.

Apart from that, the computation of sample Nash equilibrium is very complex, and this area is of much focus for computer scientists, especially the theory community. [4]

Multi-Agent Systems

Background

Multi-Agent Systems (MAS) has been a long-standing body of research in the field of Artificial Intelligence. MAS is composed of several (semi) autonomous distributed entities, known as agents, which interact in a shared environment to achieve common or conflicting goals [5]. They are not necessarily homogenous and could be a software abstraction, robotic systems and even human beings as individual agents. The earliest development in the field of MAS was in 1980, even before the technical term was firmly established, was Contract Net Protocol [6]. It developed a communication mechanism for the managers and contractors to allow resource allocation, bidding and faster decision making, similar to how Uber offers its drivers ride contracts which they can accept or reject. MAS research advanced in the late 1980s when game theory started being incorporated into MAS to study the behaviour of agents in complex environments. [7] talks about automated negotiations, which enabled networked agents to strategically and rapidly negotiate without human intervention.

Recent Advancements

Apart from academic research, Multi-Agent Systems have found some niche areas of industry application. Video games have both competitive and cooperative agents in a complex environment and thus make a good platform for MAS research [8]. In autonomous vehicles, computer vision and machine learning are used to assess the surrounding of a vehicle and this information is fed into agent planning systems to negotiate among other autonomous vehicles in a road traffic, thus relying on MAS research [9]. Another important aspect of MAS research is robotic logistics and planning. Path-finding algorithms, task allocation, and scheduling are part of the coordination system that form an active research topic in MAS [10]. MAS are widely advocated for use in mobile networks, smart grids, transportation, manufacturing and disaster relief. E-commerce, trading and marketplace transactions performed through software agents is an emerging reality, it was envisioned in the late 1990. Literature suggests the use of electronic agents has seen a growing interest in electronic agents that can assist firms and customers with their electronic commercial transactions. Depending on the application, the software agents can play different roles in online transactions, for instance, buying and selling agents, auction manager, shop bots, pricing bots, et. cetera. In [11], the author investigates optimal pricing strategies of a selling agent that is randomly matched with several heterogeneous buying agents whose reservation prices are initially unknown. They carry out multi-agent system simulations of this dynamic pricing decision problem, and we discuss some properties of the price dynamics one can observe on such marketplaces. However, at present, the design of these intelligent agents remains an issue. Some potential theoretical challenges and practical concerns with agent mediated marketplace such as rationality, optimality, liability, interaction standards, policies, risk assessment with the agent-based systems are needed to be looked at in detail [12].

Challenges

The scope of research in MAS has expanded over time. While game theory is nevertheless continuously researched with MAS, new concepts such as Reinforcement Learning have recently been studied in association with the field [13]. Reinforcement learning has enabled AI researchers to develop systems that learn and adapt from past observations and make better decisions, making the agents 'Intelligent'. To help understand the current scope of MAS, compared to fairly recent topics like deep learning and reinforcement learning, it has not received the same attention from the researchers and industry. Because of their data-driven approach, different machine learning frameworks can be evaluated based on a given dataset and well-defined metrics. Due to the multi-disciplinary nature of MAS, where they are designed for use on different tasks, it becomes difficult to evaluate the success of the MAS techniques [14]. It is important to develop standardised benchmarks to rigorously compare and better evaluate the success of the MAS to understand their potential better and increase confidence in research and industry.

Multi-Agent Reinforcement Learning

Background

A reinforcement learning agent is modelled to interact with the environment and make sequential decisions. The environment is commonly expressed as a Markov decision process (MDP), a discrete-time stochastic control process with the formal definition of the tuple (S, A, P, R, γ). S denotes the State, A denotes action, P denotes the probability that action A in State S will lead to another state, R denotes the immediate/ expected reward received after the transition from one state S to another, and γ is the discount factor that balances immediate and future rewards. [15]

In multi-agent reinforcement learning, the environment involves more than one agent making sequential decisions. The states and rewards of the system are influenced by the collective agents, with each agent optimising its own rewards (Figure 1).

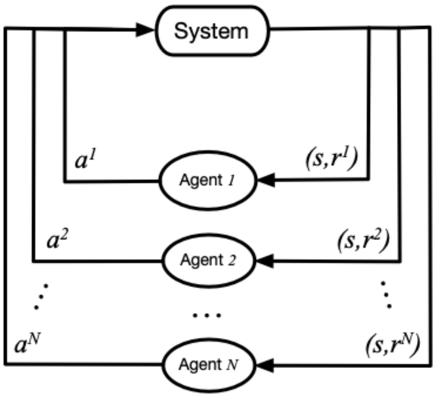

Figure 1 shows the Markov Game Process [15].

Past and recent research efforts of Multi-Agent Reinforcement Learning in online marketplaces

Multi-agent Reinforcement Learning can be used to observe strategic marketplace activity such as the economic decisions of buyer and seller agents in a simulated environment. In [16], seller and buyer agents were modelled with sellers having price adaptations, and it was examined that the pricing strategy of the seller agents converged to the market price of their respective products.

In [17], model-based reinforcement learning is utilised for finding bilateral negotiation optimisation between buyer and seller. A mediator agent was used to make the negotiation process more efficient by determining a benchmark of utility between buyer and seller.

[18], examines single and multi-agent Q-learning, where two seller agents had to adapt and take alternate turns to price their products. The seller activity of transactions and rewards were studied, and it was observed that the seller agents focus on instant reward maximisation in comparison to long-term rewards.

A buyer agent-centric approach of reinforcement learning was studied by [19], where the incorporation of a reputation system for selling agents was factored in. Buyer agents learned to choose sellers according to reputation, optimising the expected product values. Seller agents adapted to the behaviour of buyer agents by price and quality adjustment. The study ultimately demonstrated that implementing reinforcement learning methods with reputation mechanisms improved the performance of buyer agents and increased satisfaction for both buyer and seller agents.

Challenges

Regarding multi-agent reinforcement learning, [20] states the challenges of incorporating a more significant number of agents into the MDP model, along with the preliminary study of joint states and actions of the agents. Real-world applications require these complex elements to be characterised. Furthermore, investigation of agent coordination within global and partial state space information would be required. In the study done by [17], they clarified the challenge of the agents having inadequate utilisation of historical information and reasoning. Also, the agents' inability to adapt to dynamic parameters (such as time and price values in the MDP model), as opposed to pre-determined parameters, show that optimisation would be required for real-world application and marketplace integration.

Conclusion

Deep Reinforcement learning for sequential decision-making problems has progressed greatly in the last few years especially in the fields of autonomous vehicles, robotics and game playing as well [15], [21]. What is unique about these areas is that they all involve multiple players or decision makers, which sets the stage for MARL (multi agent reinforcement learning). The challenge for our team has been that while there are many examples of MARL in the extant literature it was difficult to find enough examples that fit our problem statement which is to use MARL for marketplaces, more specifically to combat informational asymmetry in the transactions that take place in these marketplaces.

Future Work

Our plan or solution, after we are done with the literature review, is to do more primary research and to design a marketplace and using these principles to build a marketplace but utilizing the characteristics of MAS to test out solutions to the aforementioned problems in the context of a marketplace.

Bibliography

[1] R. B. Myerson, ‘Game theory: Analysis of conflict’, Games and Economic Behavior, vol. 3, no. 3, pp. 387–391, 1991.

[2] Y. Shoham and K. Leyton-Brown, Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations. Cambridge: Cambridge University Press, 2008.

[3] P. Frihauf, M. Krstic, and T. Basar, ‘Nash Equilibrium Seeking in Noncooperative Games’, IEEE Transactions on Automatic Control, vol. 57, no. 5, pp. 1192–1207, May 2012, doi: 10.1109/TAC.2011.2173412.

[4] Y. Shoham, ‘Computer science and game theory’, Commun. ACM, vol. 51, no. 8, pp. 74–79, Aug. 2008, doi: 10.1145/1378704.1378721.

[5] Wooldridge, M, An Introduction to Multiagent Systems. Wiley. 2012

[6] Smith, “The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver,” in IEEE Transactions on Computers, vol. C-29, no. 12, pp. 1104-1113, Dec. 1980, doi: 10.1109/TC.1980.1675516.

[7] Beam, C., & Segev, A, Automated Negotiations: A Survey of the State of the Art. Wirtsch., 39, 263-268. 1997.

[8] Vinyals, Oriol & Babuschkin, Igor & Czarnecki, Wojciech & Mathieu, Michaël & Dudzik, Andrew & Chung, Junyoung & Choi, David & Powell, Richard & Ewalds, Timo & Georgiev, Petko & Oh, Junhyuk & Horgan, Dan & Kroiss, Manuel & Danihelka, Ivo & Huang, Aja & Sifre, Laurent & Cai, Trevor & Agapiou, John & Jaderberg, Max & Silver, David. 2019. Grandmaster level in StarCraft II using multi-agent reinforcement learning. Nature. 575. 10.1038/s41586-019-1724-z.

[9] Campos-Rodriguez, R., Gonzalez-Jimenez, L., Cervantes-Alvarez, F., Amezcua-Garcia, F., & Fernandez-Garcia, M, Multiagent Systems in Automotive Applications. Multi-Agent Systems, 2017, 1. https://doi.org/10.5772/intechopen.69687

[10] Li, Jun-tao. Design Optimisation of Amazon Robotics. Automation, Control and Intelligent Systems. 4. 48. 2016. 10.11648/j.acis.20160402.17.

[11] Leloup, B., Pricing with local interactions on agent-based electronic marketplaces. Electronic Commerce Research and Applications, 2(2), 187–198. 2003. https://doi.org/10.1016/s1567-4223(03)00023-1

[12] Guttman, A. Moukas, P. Maes, Agent-mediated commerce: a survey, Knowledge Engineering Review 13 (2) the range of issues over which an agreement can be 147–159. 1998.

[13] Zhang, K., Yang, Z., Başar, T., Multi-Agent Reinforcement Learning: A Selective Overview of Theories and Algorithms. 2019.

[14] Aylett, R., Brazier, F., Jennings, N., Luck, M., Nwana, H., & Preist, C. Agent Systems and Applications. The Knowledge Engineering Review, 13(3), 303-308. 1998. doi:10.1017/S0269888998003038

[15] K. Zhang, Z. Yang, and T. Başar, ‘Multi-Agent Reinforcement Learning: A Selective Overview of Theories and Algorithms’, arXiv:1911.10635 [cs, stat], Nov. 2019, Accessed: Apr. 16, 2021. [Online]. Available: http://arxiv.org/abs/1911.10635.

[16] Kutschinski, Uthmann, and Polani, ‘Learning competitive pricing strategies by multi-agent reinforcement learning’, Journal of Economic Dynamics and Control, vol. 27, no. 11–12, pp. 2207–2218, Sep. 2003, doi: 10.1016/S0165-1889(02)00122-7.

[17] L. Chen, H. Dong, and Y. Zhou, ‘A reinforcement learning optimised negotiation method based on mediator agent’, Expert Systems with Applications, vol. 41, no. 16, pp. 7630–7640, Nov. 2014, doi: 10.1016/j.eswa.2014.06.003.

[18] G. Tesauro and J. O. Kephart, ‘Pricing in Agent Economies Using Multi-Agent Q-Learning’, Autonomous Agents and Multi-Agent Systems, vol. 5, no. 3, pp. 289–304, Sep. 2002, doi: 10.1023/A:1015504423309.

[19] T. Tran and R. Cohen, ‘A Reputation–Oriented Reinforcement Learning Strategy for Agents in Electronic Marketplaces’, Computational Intelligence, vol. 18, no. 4, pp. 550–565, 2002, doi: https://doi.org/10.1111/1467-8640.t01-1-00203.

[20] W. Zemzem and M. Tagina, ‘Cooperative Multi-Agent Systems Using Distributed Reinforcement Learning Techniques’, Procedia Computer Science, vol. 126, pp. 517–526, Jan. 2018, doi: 10.1016/j.procs.2018.07.286.

[21] Zhang, K., Sun, T., Tao, Y., Genc, S., Mallya, S., & Başar, T. (2020). Robust Multi-Agent Reinforcement Learning with Model Uncertainty. NeurIPS.