Cooperative AI via Decentralized Commitment Devices

Abstract

Credible commitment devices have been a popular approach for robust multi-agent coordination. However, existing commitment mechanisms face limitations like privacy, integrity, and susceptibility to mediator or user strategic behavior. It is unclear if the cooperative AI techniques we study are robust to real-world incentives and attack vectors. Fortunately, decentralized commitment devices that utilize cryptography have been deployed in the wild, and numerous studies have shown their ability to coordinate algorithmic agents, especially when agents face rational or sometimes adversarial opponents with significant economic incentives, currently in the order of several million to billions of dollars. In this paper, we illustrate potential security issues in cooperative AI via examples in the decentralization literature and, in particular, Maximal Extractable Value (MEV) [27]. We call for expanded research into decentralized commitments to advance cooperative AI capabilities for secure coordination in open environments and empirical testing frameworks to evaluate multi-agent coordination ability given real-world commitment constraints.

1 Introduction

AI coordination and safety is a necessity [72] [2] [26]. With an unrestricted empowerment of algorithmic agents, the lack of coordinated interactions between them can give rise to considerable negative externalities. For example, collusive pricing in e-commerce marketplaces or gas stations [7] can hurt consumer welfare [5] [17] [83], emergent equilibria [30] from auto-bidders in ad auctions could lead to revenue loss [51] [63] [10], high-frequency trading bots can seriously reduce the health of markets [15] [29], and sniper bots on online concert ticket platforms poses profound user experience concerns [32] [73].

However, most existing approaches to both the coordination and safety problem presume critical use of traditional social technologies like policies and regulation [40] [80] [31], which are insufficient for coordinating games with a mixture of human and algorithmic participants, especially in zero-shot scenarios [42]. For example, regulations over financial transactions [6] and ad auctions [46] have failed in coordinating or aligning the incentives of algorithmic agents. A potential way to foster secure and reliable multi-agent cooperation, filling the gaps left by traditional systems, lies in credible commitment devices [47]. Existing works on cooperative AI [23] have been focusing on enabling coordination between AIs when agents can credibly commit to running certain source codes [66] [24] and allow simulation based on those committed source codes [53] [77] [25] [28][1]. Many multi-agent reinforcement learning (MARL) works also implicitly assume that there is a credibly committed training algorithm of agents that other agents may have access to [38] (or even control over [22]) to achieve zero-shot coordination [79] [42]. Finally, studies have shown that agents learn the ability to robustly coordinate via contractual commitments across a large class of coordination games [20] [43] [62].

One popular way to implement credible commitments is via decentralized trust, namely, leveraging cryptographic schemes [75] (such as multiparty computation [57], zero-knowledge proofs [12], etc.) to give agents commitment power [39]. For example, one could use basic cryptographic signatures and commitment schemes [34] to circumvent strong negative results [1] in auction mechanisms where the auctioneer is a rational agent. Moreover, combining cryptography with distributed systems allows

[1] Here, commitment means an enforcement whose knowledge can change incentives, implementing all feasible and individually rational strategies in a game [47].

Submitted to 37th Conference on Neural Information Processing Systems (NeurIPS 2023). Do not distribute.

one to implement smart contracts [16] and use them as public broadcast channels [19] that allow a large set of credible multi-agent mediations to be implemented [18] [60]. There has been a rich literature [27] [61] [33] [55] [56] [54] [8] on the ability and security of those credible commitment devices deployed in the real world, especially when they face large-scale strategic behavior by algorithmic agents with *billion dollar* amount incentives [37] [70] [36].

In this context, the study of decentralized commitment devices provides a pragmatic framework for understanding the potential limitations of the commitment assumption in multi-agent coordination settings. Namely, *when we implement and deploy AI agents in the real world to engage in zero-shot coordination tasks, what new behaviors should we expect to emerge? Are the assumptions we make in MARL and cooperative AI realistic? What would the security properties of those commitment devices be? And with those security properties of the commitment devices in mind, will AIs still coordinate robustly?*

In this paper, we first answer those questions via a few examples that illustrate the synergy between the study of decentralized commitment devices and multi-agent security. The parallels we draw should be a stepping stone for future research agendas in evaluating robust zero-shot coordination games with AI agents. We end by calling for action on evaluating multi-agent coordination ability in the presence of a decentralized commitment device, as we believe that, aside from theory, the "crypto-economic commitments sandbox" also emerges as a *tangible platform to test multi-agent coordination with real-world incentives*.

2 Congestion Games

Many games studied in the decentralized commitment device literature pose interesting questions about the security properties of multi-agent coordination. One such example is congestion games, a major area of interest for computational game theory and multi-agent systems.

Congestion games are a canonical example of a mixed-interest game with competitive and cooperative motives at play. Recent studies in artificial intelligence have focused on algorithms for learning equilibria and welfare effects [52] [58] [59] [50] in different subclasses of games with different assumptions on the information available to agents. Congestion games arise naturally in various contexts where sophisticated agents/bots interact. The most notable ones beyond robotics are online platforms (from online advertising [52] to Decentralized Finance (DeFi) [82] on blockchains). [54] models asset exchange on decentralized commitment devices as a routing game.

In the DeFi context, a rich evolving topology has emerged, with new nodes (assets) and new routes (liquidity pools [4], exchanges, bridges [11]) being added and a differential set of more and less sophisticated/informed agents that create and route orders (see Figure 1 in the Appendix). Specifically, traders delegate their decision-making to a program (called a "transaction") that decides on the creation and routing of orders. Complex behaviors have emerged, such as sophisticated strategies [27] for *frontrunning* or *backrunning* orders routed suboptimally [3] from an ecosystem of self-interested agents and organizations.

Strategies like those mentioned above are real-world examples of Maximal Extractable Value (MEV) [27], which refers to the externalities caused by mediators of commitment devices playing strategically. In this case, the mediator is the miner [16] or the block producer [9] responsible for ordering user transactions (commitments) and including them in the canonical chain (making them credible). The mediator can use its informational advantage from seeing other agents' strategies and parameters by simulating outcomes to identify an optimal action to insert right before (frontrunning) or right after (backrunning) to extract economic value from other agents.

Here, the multi-agent security problem comes from the fact that some agents (mediators) can simulate other agents an infinite number of times, which encourages cooperation [66] but also creates an asymmetry in payoffs. Other examples of MEV include an auctioneer in a second-price auction inserting a shill bid after observing all bidders' bids to extract value from the winner [81] [1].

In the following sections, we explore the limits of real-world commitments in both AI and blockchain. We highlight the disparities between prevailing notions of commitments and the requisites of real-world commitments and emphasize the need for robust solutions to address these limitations.

3 Limits of Real-world Commitments

When considering cooperative AI agents, there is often an implicit assumption of the existence of a trusted centralized mediator[2] responsible for orchestrating interactions among agents [77] [38]. This central mediator assumes a critical role, but recent antitrust litigation against platforms like Google's ad auction [46] serves as a reminder of the risks associated with entrusting a single all-powerful centralized entity with coordinating actions among agents despite the mediator has a long-term reputation at stake. Even when agents don't explicitly use commitments (see elaboration in Appendix B on a concrete game) or they take turns acting as mediators during contract negotiations, the risk of collusion persists. In reality, relying on identity verification to mitigate this issue is impractical, as agents and bots are more susceptible to Sybil attacks than human agents.

In real-world settings, not all applications or agents will tolerate simulation or parameter sharing (for a concrete exploit on this involving shared world models, see Appendix B). This constraint suggests we should consider the use of Privacy Enhancing Technologies (PETs) [69] to facilitate contract formation. Moreover, when formulating assumptions regarding cooperative agents, it is imperative to consider the constraints introduced by PETs, including their impact on performance, hardware requirements, and the expressiveness of contracts.

Both AI and MEV settings share common limitations. Credible commitment devices, whether permissionless or not, must offer real-world validity guarantees. To instill trust in such mechanisms, users need assurance on how commitments maintain integrity once deployed. Two approaches exist – constructing commitment devices with inherent integrity obtained via verifiable cryptographic tools such as SNARKs (Succinct Non-interactive ARguments of Knowledge) [12], or optimistic reliance on monitoring and penalizing defection [68].

Presently, real-world deployments predominantly lean toward the optimistic approach due to technical constraints. Constructing integrity guarantees by design introduces complexity and higher overhead, existing methods can now support deep neural networks of the size of MobileNet[74] or GPT2[71] with a 10-20X inference time overhead on CPU [48], which leaves a lot of room for further software and hardware speed-ups. While computationally tractable, optimistic approaches fall short in scenarios with fat-tail payoffs. There have been incidents where $20 million loss was caused by exploiting payoff asymmetry in commitments [35].

4 Security of Commitments in Cooperative AI

We underscore some security complications of commitment devices mentioned above, evident across a broad spectrum of existing cooperative AI work.

First, in situations where cooperation relies on hard-wired heuristics in the training process or the learning algorithm, such as augmenting intrinsic motivation [45] by changing reward functions or employing Centralized Training with Decentralized Execution (CTDE)[67][58], the necessity of a credible commitments to opponents' training process prior to an agent's interaction is paramount. This necessity vests significant trust in the mediator providing this proof, thereby presenting a potential vulnerability, as mediators can effectively dictate the equilibrium the agents converge to in general-sum coordination games (see Appendix B.2 for an illustrative example). This incentivizes collusion between the mediator and agents to manipulate the training process or provide non-factual proofs.

Agents using algorithms such as LOLA [38] or COLA [84], which assumes access to opponents' learning parameters, would similarly require a verifiable proof that its opponents are indeed running the reported parameters. This proof can be viewed as a commitment by either the opponent or the mediator, and it suffers from credibility problems. In reality, agents may not want to fully expose their parameters due to privacy or commercial reasons, leading them to entrust proof delegation to a mediator. However, this incentivizes the mediator to accept bribes from agents to misreport parameters as it can earn this profit risk-freely since no single agent can provide evidence of cheating [1]. Of course, those algorithms can evade the commitment credibility problem via opponent modeling using techniques like behavior cloning [78] at the cost of performance, posing an interesting meta-

[2]Here, the mediators should be explicitly interpreted as in mediated equilibrium [64] where it has commitment power to implement (coarse) correlated strategies.

coordination problem: since parameter sharing is incompatible with privacy, agents will likely resort to using entrusted mediators, and because of strategic mediator behavior, agents may resort back to using opponent modeling, giving up potential coordination gains.

In credible commitments, the problem of Maximal Extractable Value (MEV) [27] is even more apparent when agents explicitly use correlation devices to coordinate. For example, if cooperative AI approaches where agents learn payment contracts [20] or have access to signals from communication channels [21] are implemented in reality, we would likely see agents engage in MEV behavior. For example, in mediated multi-agent reinforcement learning [44], the mediator is explicitly modeled as a learning agent who can make credible commitments and takes reports from agents about their observation/reward/action pairs to design an optimal contract. It is not uncommon for users to play the meta-game here and strategically misreport inputs to the mediator or their own agents [52] [51].

Moreover, It has been demonstrated in existing MARL works that agents that coordinate using correlation devices could converge to unfair outcomes where one agent learns to propose a contract that extracts close to all of the surplus from coordination and distributes only ϵ gain the rest [20]. This is a longstanding problem in game theory where the mechanism designer, or the Stackelberg player, gets to extract the entire surplus. If those cooperative AI systems are actually deployed, we expect to see agents explicitly optimizing for being the "Stackelberg player." For example, they could disrupt the communication network via a Denial of Service (DoS) attack with their commitment such that the mediator does not see other agents' commitments or chooses to send manipulative messages [14].

Alternatively, instead of directly manipulating the mediator/correlation device, the agents may learn to do long-term optimization by learning to impact which equilibrium the policies converge to, with the resulting equilibrium formalized as *active equilibrium* [49] [50]. It has also been demonstrated that agents can learn to play the meta-game of impacting equilibrium selection in learning convergence rather than myopically optimizing for reward [85] [59]. Here, the problem is not with the credibility of the commitment but rather whether the inputs to the commitment reflect the true state of the world. Such problems are exacerbated in the face of censorship and insertion [1].

5 Call for Action

In this work, we have addressed the potential of studying decentralized commitment devices in revealing a plethora of security issues inherent in existing cooperative AI endeavors. We believe the study of those devices heralds a promising horizon for enhancing multi-agent systems' robustness and security foundations.

We also see decentralized commitment devices as a valuable addition to our arsenal of techniques to achieve more secure cooperative AI. Even though it has been demonstrated that decentralized commitment devices can mitigate such security issues of centralized mediators [64] [75] [34] [18], a significant research gap persists in their application to ameliorate multi-agent security problems within the domain of cooperative AI [26].

We advocate for intensified exploration into the nexus between decentralized commitment devices and cooperative AI. A critical need exists for empirical demonstrations of AI's coordination capacity across an expansive array of general-sum games utilizing decentralized commitments alongside a comprehensive evaluation of the real-world constraints inherent in these commitments. This call to action echoes the imperative to bolster AI's capability to do few-shot coordination in environments with commitments and robustly defend against security attacks on those commitments.

It is also pivotal to traverse beyond theoretical discussions, propelling the implementation of extant cooperative AI research atop a decentralized commitment device framework. This strategic move will invariably augment the exposure to both adversarial and real-world incentives, offering a more holistic and realistic testing ground for cooperative AI, transcending the confines of an insulated environment.

The fulfillment of these objectives will significantly advance our understanding and capabilities in cooperative AI, underscoring its potential to flourish securely and efficiently in real-world applications, thus contributing profoundly to the broader AI research landscape. As we advance in the field of multi-agent systems, it is essential to recognize the role that decentralized commitments can play.

References

[1] Akbarpour, M., Li, S.: Credible auctions: A trilemma. Econometrica **88**(2), 425–467 (2020)

[2] Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., Mané, D.: Concrete problems in ai safety. arXiv preprint arXiv:1606.06565 (2016)

[3] Angeris, G., Evans, A., Chitra, T., Boyd, S.: Optimal routing for constant function market makers. In: Proceedings of the 23rd ACM Conference on Economics and Computation. pp. 115–128 (2022)

[4] Angeris, G., Kao, H.T., Chiang, R., Noyes, C., Chitra, T.: An analysis of uniswap markets (2021)

[5] of Justice Department of Antitrust, U.D.: Former e-commerce executive charged with price fixing in the antitrust division's first online marketplace prosecution (September 20th 2023), `https://www.justice.gov/opa/pr/former-e-commerce-executive-charged-price-fixing-antitrust-divisions-first-online-marketplace`

[6] of Appeals, U.S.C.: Newton v. merrill, lynch, pierce, fenner (1998), `https://casetext.com/case/newton-v-merrill-lynch-pierce-fenner-3`

[7] Assad, S., Clark, R., Ershov, D., Xu, L.: Algorithmic pricing and competition: Empirical evidence from the german retail gasoline market (2020)

[8] Babel, K., Javaheripi, M., Ji, Y., Kelkar, M., Koushanfar, F., Juels, A.: Lanturn: Measuring economic security of smart contracts through adaptive learning. Cryptology ePrint Archive (2023)

[9] Bahrani, M., Garimidi, P., Roughgarden, T.: Transaction fee mechanism design with active block producers. arXiv preprint arXiv:2307.01686 (2023)

[10] Banchio, M., Skrzypacz, A.: Artificial intelligence and auction design. In: Proceedings of the 23rd ACM Conference on Economics and Computation. pp. 30–31 (2022)

[11] Belchior, R., Vasconcelos, A., Guerreiro, S., Correia, M.: A survey on blockchain interoperability: Past, present, and future trends. ACM Computing Surveys (CSUR) **54**(8), 1–41 (2021)

[12] Bitansky, N., Canetti, R., Chiesa, A., Goldwasser, S., Lin, H., Rubinstein, A., Tromer, E.: The hunting of the snark (2014), `https://eprint.iacr.org/2014/580`, citation Key: cryptoeprint:2014/580 tex.howpublished: Cryptology ePrint Archive, Paper 2014/580

[13] Blum, A., Hajiaghayi, M., Ligett, K., Roth, A.: Regret minimization and the price of total anarchy. In: Proceedings of the fortieth annual ACM symposium on Theory of computing. pp. 373–382 (2008)

[14] Blumenkamp, J., Prorok, A.: The emergence of adversarial communication in multi-agent reinforcement learning. In: Conference on Robot Learning. pp. 1394–1414. PMLR (2021)

[15] Budish, E., Cramton, P., Shim, J.: The high-frequency trading arms race: Frequent batch auctions as a market design response. The Quarterly Journal of Economics **130**(4), 1547–1621 (2015)

[16] Buterin, V., et al.: A next-generation smart contract and decentralized application platform. white paper **3**(37), 2–1 (2014)

[17] Chen, L., Mislove, A., Wilson, C.: An empirical analysis of algorithmic pricing on amazon marketplace. In: Proceedings of the 25th international conference on World Wide Web. pp. 1339–1349 (2016)

[18] Chitra, T., Ferreira, M.V., Kulkarni, K.: Credible, optimal auctions via blockchains. arXiv preprint arXiv:2301.12532 (2023)

[19] Choudhuri, A.R., Green, M., Jain, A., Kaptchuk, G., Miers, I.: Fairness in an unfair world: Fair multiparty computation from public bulletin boards. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. pp. 719–728 (2017)

[20] Christoffersen, P.J., Haupt, A.A., Hadfield-Menell, D.: Get it in writing: Formal contracts mitigate social dilemmas in multi-agent rl. arXiv preprint arXiv:2208.10469 (2022)

[21] Cigler, L., Faltings, B.: Decentralized anti-coordination through multi-agent learning. Journal of Artificial Intelligence Research **47**, 441–473 (2013)

[22] Claus, C., Boutilier, C.: The dynamics of reinforcement learning in cooperative multiagent systems. AAAI/IAAI **1998**(746-752), 2 (1998)

[23] Conitzer, V., Oesterheld, C.: Foundations of cooperative ai. AAAI-23 Senior Member Blue Sky Ideas track (2022)

[24] Critch, A.: A parametric, resource-bounded generalization of löb's theorem, and a robust cooperation criterion for open-source game theory. The Journal of Symbolic Logic **84**(4), 1368–1381 (2019)

[25] Critch, A., Dennis, M., Russell, S.: Cooperative and uncooperative institution designs: Surprises and problems in open-source game theory. arXiv preprint arXiv:2208.07006 (2022)

[26] Dafoe, A., Hughes, E., Bachrach, Y., Collins, T., McKee, K.R., Leibo, J.Z., Larson, K., Graepel, T.: Open problems in cooperative ai. arXiv preprint arXiv:2012.08630 (2020)

[27] Daian, P., Goldfeder, S., Kell, T., Li, Y., Zhao, X., Bentov, I., Breidenbach, L., Juels, A.: Flash boys 2.0: Frontrunning in decentralized exchanges, miner extractable value, and consensus instability. In: 2020 IEEE Symposium on Security and Privacy (SP). pp. 910–927. IEEE (2020)

[28] DiGiovanni, A., Clifton, J.: Commitment games with conditional information disclosure. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 37, pp. 5616–5623 (2023)

[29] Easley, D., De Prado, M.M.L., O'Hara, M.: The microstructure of the "flash crash": flow toxicity, liquidity crashes, and the probability of informed trading. The Journal of Portfolio Management **37**(2), 118–128 (2011)

[30] Edelman, B., Ostrovsky, M.: Strategic bidder behavior in sponsored search auctions. Decision support systems **43**(1), 192–198 (2007)

[31] Edwards, L.: Regulating ai in europe: four problems and four solutions. Retrieved March **15**, 2022 (2022)

[32] Elbeshbishi, S.: 'industrial-scale ticket scalping.' senators grill ticketmaster over taylor swift concert fiasco (September 20th 2023), https://web.archive.org/web/20230124233707/https://www.usatoday.com/story/news/politics/2023/01/24/senate-judiciary-taylor-swift-ticketmaster/11091086002/

[33] Ferreira, M.V., Parkes, D.C.: Credible decentralized exchange design via verifiable sequencing rules. arXiv preprint arXiv:2209.15569 (2022)

[34] Ferreira, M.V., Weinberg, S.M.: Credible, truthful, and two-round (optimal) auctions via cryptographic commitments. In: Proceedings of the 21st ACM Conference on Economics and Computation. pp. 683–712 (2020)

[35] Flashbots: Post mortem: April 3rd, 2023 mev-boost relay incident and related timing issue (2023), https://collective.flashbots.net/t/post-mortem-april-3rd-2023-mev-boost-relay-incident-and-related-timing-issue/1540/1

[36] Flashbots: Flashbots transparency dashboard (September 20th 2023), https://transparency.flashbots.net/

[37] Flashbots: Mev-explorev1 (September 20th 2023), https://explore.flashbots.net/

[38] Foerster, J.N., Chen, R.Y., Al-Shedivat, M., Whiteson, S., Abbeel, P., Mordatch, I.: Learning with opponent-learning awareness. arXiv preprint arXiv:1709.04326 (2017)

[39] Griffith, V.: Ethereum is game-changing technology, literally. (2019), https://medium.com/@virgilgr/ethereum-is-game-changing-technology-literally-d67e01a01cf8

[40] Hacker, P., Engel, A., Mauer, M.: Regulating chatgpt and other large generative ai models. In: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. pp. 1112–1123 (2023)

[41] Harris, K., Anagnostides, I., Farina, G., Khodak, M., Wu, Z.S., Sandholm, T.: Meta-learning in games. arXiv preprint arXiv:2209.14110 (2022)

[42] Hu, H., Lerer, A., Peysakhovich, A., Foerster, J.: "other-play" for zero-shot coordination. In: International Conference on Machine Learning. pp. 4399–4410. PMLR (2020)

[43] Hughes, E., Anthony, T.W., Eccles, T., Leibo, J.Z., Balduzzi, D., Bachrach, Y.: Learning to resolve alliance dilemmas in many-player zero-sum games. arXiv preprint arXiv:2003.00799 (2020)

[44] Ivanov, D., Zisman, I., Chernyshev, K.: Mediated multi-agent reinforcement learning. arXiv preprint arXiv:2306.08419 (2023)

[45] Jaques, N., Lazaridou, A., Hughes, E., Gulcehre, C., Ortega, P., Strouse, D., Leibo, J.Z., De Freitas, N.: Social influence as intrinsic motivation for multi-agent deep reinforcement learning. In: International conference on machine learning. pp. 3040–3049. PMLR (2019)

[46] of Public Affairs at U.S. Department of Justice, O.: Justice department sues google for monopolizing digital advertising technologies. (2023), https://www.justice.gov/opa/pr/justice-department-sues-google-monopolizing-digital-advertising-technologies

[47] Kalai, A.T., Kalai, E., Lehrer, E., Samet, D.: A commitment folk theorem. Games and Economic Behavior **69**(1), 127–137 (2010)

[48] Kang, D., Hashimoto, T., Stoica, I., Sun, Y.: Scaling up trustless dnn inference with zero-knowledge proofs (arXiv:2210.08674) (Oct 2022), http://arxiv.org/abs/2210.08674

[49] Kim, D.K., Riemer, M., Liu, M., Foerster, J., Everett, M., Sun, C., Tesauro, G., How, J.P.: Influencing long-term behavior in multiagent reinforcement learning. Advances in Neural Information Processing Systems **35**, 18808–18821 (2022)

[50] Kim, D.K., Riemer, M., Liu, M., Foerster, J.N., Tesauro, G., How, J.P.: Game-theoretical perspectives on active equilibria: A preferred solution concept over nash equilibria. arXiv preprint arXiv:2210.16175 (2022)

[51] Kolumbus, Y., Nisan, N.: Auctions between regret-minimizing agents. In: Proceedings of the ACM Web Conference 2022. pp. 100–111 (2022)

[52] Kolumbus, Y., Nisan, N.: How and why to manipulate your own agent: On the incentives of users of learning agents. Advances in Neural Information Processing Systems **35**, 28080–28094 (2022)

[53] Kovarik, V., Oesterheld, C., Conitzer, V.: Game theory with simulation of other players. arXiv preprint arXiv:2305.11261 (2023)

[54] Kulkarni, K., Diamandis, T., Chitra, T.: Towards a theory of maximal extractable value i: Constant function market makers. arXiv preprint arXiv:2207.11835 (2022)

[55] Landis, D., Schwartzbach, N.I.: Side contract commitment attacks on blockchains. arXiv preprint arXiv:2301.08523 (2023)

[56] Landis, D., Schwartzbach, N.I.: Stackelberg attacks on auctions and blockchain transaction fee mechanisms. arXiv preprint arXiv:2305.02178 (2023)

[57] Lindell, Y.: Secure multiparty computation. Communications of The Acm **64**(1), 86–96 (Dec 2020). https://doi.org/10.1145/3387108

[58] Lowe, R., Wu, Y.I., Tamar, A., Harb, J., Pieter Abbeel, O., Mordatch, I.: Multi-agent actor-critic for mixed cooperative-competitive environments. Advances in neural information processing systems **30** (2017)

[59] Lu, C., Willi, T., De Witt, C.A.S., Foerster, J.: Model-free opponent shaping. In: International Conference on Machine Learning. pp. 14398–14411. PMLR (2022)

[60] Matsushima, H., Noda, S.: Mechanism design with blockchain enforcement. Available at SSRN 3554512 (2020)

[61] Mazorra, B., Della Penna, N.: The cost of sybils, credible commitments, and false-name proof mechanisms. arXiv preprint arXiv:2301.12813 (2023)

[62] McAleer, S., Lanier, J., Dennis, M., Baldi, P., Fox, R.: Improving social welfare while preserving autonomy via a pareto mediator. arXiv preprint arXiv:2106.03927 (2021)

[63] Mehta, A., Perlroth, A.: Auctions without commitment in the auto-bidding world. arXiv preprint arXiv:2301.07312 (2023)

[64] Monderer, D., Tennenholtz, M.: Strong mediated equilibrium. Artificial Intelligence **173**(1), 180–195 (2009)

[65] Moulin, H., Vial, J.P.: Strategically zero-sum games: the class of games whose completely mixed equilibria cannot be improved upon. International Journal of Game Theory **7**, 201–221 (1978)

[66] Oesterheld, C., Treutlein, J., Grosse, R., Conitzer, V., Foerster, J.: Similarity-based cooperation. arXiv preprint arXiv:2211.14468 (2022)

[67] Oliehoek, F.A., Amato, C., et al.: A concise introduction to decentralized POMDPs, vol. 1. Springer (2016)

[68] Ostrom, E.: Governing the commons: The evolution of institutions for collective action. Cambridge university press (1990)

[69] PPTTT, P.P.T.T.T.: UN handbook on privacy-preserving computation techniques (2019), https://unstats.un.org/bigdata/task-teams/privacy/index.cshtml, citation Key: big2019handbook

[70] Qin, K., Zhou, L., Livshits, B., Gervais, A.: Attacking the defi ecosystem with flash loans for fun and profit. In: International conference on financial cryptography and data security. pp. 3–32. Springer (2021)

[71] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I., et al.: Language models are unsupervised multitask learners. OpenAI blog **1**(8) (2019)

[72] Russell, S.: Human compatible: Artificial intelligence and the problem of control. Penguin (2019)

[73] Salcedo, A.: Angry taylor swift fans push lawmakers to take on ticketmaster (September 20th 2023), https://www.washingtonpost.com/business/2023/06/11/taylor-swift-ticketmaster-ticket-meltdown-bills/

[74] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4510–4520 (2018)

[75] Shi, E., Chung, H., Wu, K.: What can cryptography do for decentralized mechanism design. arXiv preprint arXiv:2209.14462 (2022)

[76] Skyrms, B.: Quasi-conventions. Synthese **201**(3), 99 (2023)

[77] Tennenholtz, M.: Program equilibrium. Games and Economic Behavior **49**(2), 363–373 (2004)

[78] Torabi, F., Warnell, G., Stone, P.: Behavioral cloning from observation. arXiv preprint arXiv:1805.01954 (2018)

[79] Treutlein, J., Dennis, M., Oesterheld, C., Foerster, J.: A new formalism, method and open issues for zero-shot coordination. In: International Conference on Machine Learning. pp. 10413–10423. PMLR (2021)

[80] Union, E.: Regulatory framework proposal on artificial intelligence (September 20th 2023), https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

[81] Vickrey, W.: Counterspeculation, auctions, and competitive sealed tenders. The Journal of finance **16**(1), 8–37 (1961)

[82] Werner, S., Perez, D., Gudgeon, L., Klages-Mundt, A., Harz, D., Knottenbelt, W.: Sok: Decentralized finance (defi). In: Proceedings of the 4th ACM Conference on Advances in Financial Technologies. pp. 30–46 (2022)

[83] Wieting, M., Sapi, G.: Algorithms in the marketplace: An empirical analysis of automated pricing in e-commerce. Available at SSRN 3945137 (2021)

[84] Willi, T., Letcher, A.H., Treutlein, J., Foerster, J.: Cola: consistent learning with opponent-learning awareness. In: International Conference on Machine Learning. pp. 23804–23831. PMLR (2022)

[85] Xie, A., Losey, D., Tolsma, R., Finn, C., Sadigh, D.: Learning latent representations to influence multi-agent interaction. In: Conference on robot learning. pp. 575–588. PMLR (2021)

A DeFi network and routing

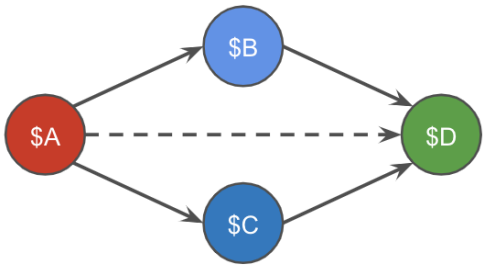

Figure 1: Example DeFi network with 4 assets and routes between them. Full lines are smart contract exchanges and dotted line is a (potential) direct swap.

B Where does the Commitment Happen?

In MARL and Cooperative Game Theory more generally, agents are sometimes said to act as though they are committed, but the commitment lives outside the model. For example, Coarse Correlated Equilibrium [65] is an equilibrium refinement popular in MARL. One appealing feature is that a commonly used algorithm, external regret minimization, converges quickly to a CCE. Projected Gradient Descent in a MARL context results in a CCE as well.[41] It is a highly tractable equilibrium concept, but it has some unintuitive features related to commitment.[3] We will discuss it in some depth as an illustrative example of how difficult commitment can be to achieve in practice.

CCE requires a signal that is commonly observed by the agents, and it further requires that, within the game, agents cannot condition on their own behavior. This latter condition presumes strong commitment: if a player is about to choose x based on the commonly observed signal, a CCE assumes that the player is not allowed to instead choose an option y even if they are certain y will give higher payoffs. To observe the force of this assumption, and consider plausible implementations, we will illustrate with an example.

B.1 CCE in a Stop Light Game

Consider the Stop Light Game (adapted from [76], with a signal set $S = \{FS, SF, SC, CS\}$, and a probability distribution function $\pi : S \to [0, 1]$. If the probabilities for each signal are $\pi(FS) = \pi(SF) = \frac{1}{3}$ and $\pi(SC) = \pi(CS) = \frac{1}{6}$, each player obeying the signals–e.g. choosing (Fast, Stop) conditional on signal FS–is a CCE.

	Fast	Caution	Stop
Fast	(0,0)	(3,1)	(7,2)
Caution	(1,3)	(2.1,2.1)	(6,2)
Stop	(2,7)	(2,6)	(4,4)

Table 1: Payoff table for stop light game

However, note that whenever Column (Row) receives the signal CS (SC) they strictly prefer to choose Fast instead of Caution. Doing so gives them a certain payoff increase of 1. If players could unilaterally defect to their preferred choice, such Coarse Correlated Equilibria would be anti-correlated.

Players who could bribe the mediator in some way can also improve their outcome, even without formally defecting. If Column bribes the mediator to signal FS, Column receives a guaranteed payoff of 7 which is higher than their CCE payoff (of 4.3 in expectation.) In principle, since the signal is

[3]This was noted by the original authors who observe that their proposal is "not, strictly speaking, non-cooperative", which is to say that it leaves some strategic elements of the game unmodeled.

ostensibly random and thus an observed $\hat{\pi}(FS) = 1$ is not ruled out, the mediator could profit about 2 per draw.

B.2 Methods of Commitment

One way to improve the incentive compatibility of a CCE in the Stop Light game would be to credibly commit to destroy some amount greater than 1 if the non-recommended strategy is chosen. This is effectively the "decentralized commitment device" method.

Another is to try and restrict the choices of the agents in some other way. Among machines, one commitment would be to not use an algorithm that penalizes losses at each decision node, such as swap regret minimization.[4] But this would seem to just beg the question of why swap regret minimization is not chosen if it provides better payoffs in the actual game.

A further possibility is conditional commitment by mutual inspection, as in Tennenholtz's Program Equilibria [77]. Two cars at an intersection could potentially settle on $(Caution, Stop)|S = (CS)$ by inspecting the others' program and conditioning on it being the same as well as the signal CS. But a player that somehow learned to misunderstand CS as, say, FS does strictly better. And so too perhaps would a car owner who was able to modify their car to misinterpret CS as FS.

Maybe such errant behavior could be ruled out if it required a changing of the program itself[5]. But the implementation of the game itself offers further opportunities for a type of "non-cooperation". For instance, implementing a naive "Program Equilibrium" program in the game Matching Pennies would mean that the column program exploits the matching behavior of the row program. Some commitment to the game form itself appears to be needed here.

[4] Swap regret minimization is of the form "every time I chose action i, I should have chosen action j instead", and thus would choose $Disobey|S = CS$ in our reduced stop light game.[13]

[5] Though this may not be the case–presumably you would need to be assured of the same inputs, or have the ability to check state up in until the moment of the execution of the game.