No-regret Learning Dynamics for Sequential Correlated Equilibria

Extended Abstract

Hugh Zhang
Harvard University
Cambridge, Massachusetts, United States
hughzhang@seas.harvard.edu

ABSTRACT

While no-regret learning procedures that converge to correlated equilibria have long been known to exist for normal form games, their analogue in sequential games remains less clear. We propose the *sequential correlated equilibrium*, a solution concept that extends the correlated equilibrium to sequential games while also guaranteeing sequential rationality even for mediator recommendations off the path of play. Additionally, we show that any internal regret minimization procedure designed for normal-form games can be efficiently extended to sequential games and use this to design no-regret learning dynamics that converge to the set of sequential correlated equilibria.

KEYWORDS

internal regret; counterfactual regret minimization; sequential games; correlated equilibria; sequential rationality

ACM Reference Format:
Hugh Zhang. 2023. No-regret Learning Dynamics for Sequential Correlated Equilibria: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

Algorithmic game theory aims to discover simple, efficient algorithms that converge to equilibria. One popular family of such algorithms is no-regret learning dynamics, whereupon players play a game repeatedly and choose their future actions using a simple learning procedure over their past results. A celebrated result in computational game theory is that no-regret learning dynamics lead to a Nash equilibrium in two-player zero-sum games if each player chooses their future strategies with probabilities proportional to their regret for not choosing that strategy in the past [2]. More recently, this result was extended for correlated equilibria in general normal-form games [3, 11, 12].

This paper asks whether we can achieve similar results in sequential games. One recent successful approach in this direction has been counterfactual regret minimization [20], which converges to a Nash equilibrium in two-player zero-sum sequential games of perfect recall. Counterfactual regret minimization uses a tree decomposition approach to break up the original game into a number of "minigames," each of which can be efficiently optimized independently with an external regret minimization algorithm. Counterfactual regret minimization has led to significant successes in computational game theory, such as building agents which have surpassed humans in several popular forms of poker [4–6, 16].

These recent successes in two-player zero-sum games have led to increased interest in learning in general games with more than two-players and with non zero-sum utility functions. Since computing a Nash equilibrium is known to be computationally hard in such cases [8, 9, 18], focus has been turned to alternative solution concepts such as the correlated equilibrium [1]. Directly extending counterfactual regret minimization to converge to the set of correlated equilibria in general sequential games runs into the hurdle of having to satisfy an exponentially large number of incentive constraints. Several extensions of the correlated equilibrium, such as the extensive-form correlated equilibrium [19] or the autonomous correlated equilibrium [10] have been proposed and are known to be tractably computable in sequential games [14]. However, these equilibria do not typically guarantee any form of *sequential rationality* [15], in that they may require players to play suboptimally after departing from the equilibrium path.This paper has two main contributions. First, we propose a novel solution concept called the *sequential correlated equilibrium*. Second, we describe S-CFR, a no-regret learning procedure that converges to the set of sequential

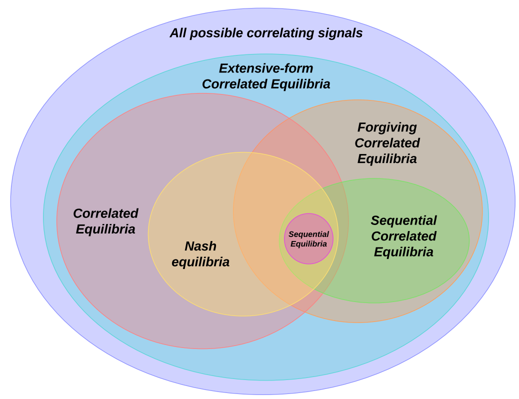

Figure 1: No method is currently known to efficiently compute correlated equilibria (and thus by extension, Nash and sequential equilibria) in general sequential games. The first major contribution of this paper is the sequential correlated equilibrium, a solution concept that includes the sequential rationality guarantees of the sequential equilibrium while still being computable in polynomial time.

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

Table 1: A comparison of the computability and sequential rationality properties of various equilibria

Equilibrium	*Complexity in Normal-form Games*	*Complexity in Sequential Games*	*Guarantees Sequentially Rationality?*
Nash [17]	PPAD [8, 9]	PPAD [8, 9]	No
Correlated [1]	Polynomial [13]	Unknown	No
Extensive-form Correlated [19]	Polynomial [19]	Polynomial [19]	No
Sequential [15]	PPAD [8, 9]	PPAD [8, 9]	Yes
Sequential Correlated (ours)	Polynomial	Polynomial	Yes

correlated equilibria for any finite sequential game of perfect recall. Alongside [7], which was completed concurrently and independently from our work, this paper provides the first extension of no internal regret learning dynamics to sequential games.

2 EQUILIBRIA IN SEQUENTIAL GAMES

For space reasons, we relegate a detailed discussion of the notation and no regret learning procedure to the full paper and merely describe the proposed solution concepts in the extended abstract.

2.1 Forgiving Correlated Equilibria

Before formally defining sequential correlated equilibria, we will define an additional solution concept which will be useful to describe the sequential correlated equilibrium and the learning dynamics that converge to it. A *forgiving correlated equilibrium* is similar to an extensive-form correlated equilibrium except that the mediator will "forgive" player errors. That is, in a forgiving correlated equilibrium, mediator recommendations continue to be given to each player *even after they deviate from previous recommendations.* Furthermore, these recommendations will continue to be incentive compatible. Thus, it will never be in a player's interest to deviate from a mediator recommendation, even if that player has already done so at a previous point in the game. Formally, we can define the forgiving correlated equilibrium as follows:

Definition 2.1. Let h be a probability distribution over strategy profiles. A signal h forms a *forgiving correlated equilibrium* if it satisfies:

$$\max_{I \in \mathcal{I}} \max_{\vec{a} \in S(I)} \max_{d \in \mathcal{D}_{P(I)}} \sum_{s \in \Sigma} h(s) O(I, \vec{a}, s)(u(I, d, s) - u(I, s)) \leq 0.$$

2.2 Sequential Correlated Equilibrium

Inspired by the sequential equilibrium [15], which addresses this issue for Nash equilibria in sequential games, we propose the *sequential correlated equilibrium.* The sequential correlated equilibrium is a computationally-tractable extension of the correlated equilibrium solution concept to sequential games that includes guarantees of sequential rationality at all infosets in the game, not merely those that are on the path of play. Say a probability distribution h has full support if it places positive probability on every pure strategy profile $s \in \Sigma$. Define a belief β that corresponds to a signal h with full support as

$$\beta^h(s \mid I, \vec{a}) = \frac{h(s) O(I, \vec{a}, s)}{\sum_{s \in \Sigma} h(s) O(I, \vec{a}, s)}.$$

These are the Bayes consistent beliefs over the strategy profile describing the mediator's recommendations, conditional on arriving at an infoset I and receiving a signal history $\vec{a}$.

Definition 2.2. Let $(h_n)_n$ be a sequence of signals, all of which have full support, that converge to some signal h^* (not necessarily full support). Let β^n be the corresponding Bayes updated beliefs to the signal h_n as defined above and let $\beta^* = \lim_{n \to \infty} \beta^n$. $(h^*, \beta^*, (h_n)_n)$ forms a *sequential correlated equilibrium* if it satisfies

$$\max_{I \in \mathcal{I}} \max_{\vec{a} \in S(I)} \max_{d \in \mathcal{D}_{P(I)}} \sum_{s \in \Sigma} \beta^h(s \mid I, \vec{a})(u(I, d, s) - u(I, s)) \leq 0.$$

Figure 1 illustrates the relationship between the forgiving correlated equilibrium, the extensive-form correlated equilibrium, the canonical correlated equilibrium, the Nash equilibrium, and the sequential forgiving correlated equilibrium described in Section **??**.

Sequential correlated equilibria are the limit points of a sequence of forgiving correlated equilibria with full support (see Section **??**), along with their corresponding beliefs. This observation will form the backbone of our construction of no-regret learning dynamics that converge to sequential correlated equilibria.

3 RELATED WORK

3.1 Equilibria

We extensively discuss the relationship between sequential correlated equilibria and other equilibria in Section 3 of our full paper and give only a brief summary here. The Nash equilibrium is the most well-known solution concept in game theory, but is computationally hard to compute [8, 9]. Its leading competitor, the correlated equilibrium, is computationally feasible in normal-form games [13] but not in sequential games. The extensive-form correlated equilibria is a tractable extension of the correlated equilibrium to sequential games [14, 19], but does not come with guarantees of sequential rationality. The sequential equilibrium [15] guarantees sequential rationality, but as it is a refinement of the Nash equilibrium, it is also computationally hard to compute.

4 ACKNOWLEDGEMENTS

I would like to thank my undergraduate thesis advisor Gabriel Carroll for his guidance and supervision for my undergraduate thesis, which led to this paper. I'd also like to thank my graduate advisor David Parkes for his help and advice in preparing this submission.

REFERENCES

[1] Robert J Aumann. 1974. Subjectivity and correlation in randomized strategies. *Journal of mathematical Economics* 1, 1 (1974), 67–96. Publisher: Elsevier.

[2] David Blackwell. 1956. An analog of the minimax theorem for vector payoffs. *Pacific J. Math.* 6, 1 (1956), 1–8. Publisher: Pacific Journal of Mathematics.

[3] Avrim Blum and Yishay Mansour. 2007. From external to internal regret. *Journal of Machine Learning Research* 8, Jun (2007), 1307–1324.

[4] Michael Bowling, Neil Burch, Michael Johanson, and Oskari Tammelin. 2015. Heads-up limit hold'em poker is solved. *Science* 347, 6218 (2015), 145–149. Publisher: American Association for the Advancement of Science.

[5] Noam Brown and Tuomas Sandholm. 2018. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science* 359, 6374 (2018), 418–424. Publisher: American Association for the Advancement of Science.

[6] Noam Brown and Tuomas Sandholm. 2019. Superhuman AI for multiplayer poker. *Science* 365, 6456 (2019), 885–890. Publisher: American Association for the Advancement of Science.

[7] Andrea Celli, Alberto Marchesi, Gabriele Farina, and Nicola Gatti. 2020. No-Regret Learning Dynamics for Extensive-Form Correlated Equilibrium. *Advances in Neural Information Processing Systems* 33 (2020), 7722–7732. https://proceedings.neurips.cc/paper/2020/hash/5763abe87ed1938799203fb6e8650025-Abstract.html

[8] Xi Chen, Xiaotie Deng, and Shang-Hua Teng. 2009. Settling the complexity of computing two-player Nash equilibria. *Journal of the ACM (JACM)* 56, 3 (2009), 1–57. Publisher: ACM New York, NY, USA.

[9] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. 2009. The Complexity of Computing a Nash Equilibrium. *SIAM J. Comput.* 39, 1 (Jan. 2009), 195–259. https://doi.org/10.1137/070699652 Publisher: Society for Industrial and Applied Mathematics.

[10] Francoise Forges. 1986. Correlated equilibria in repeated games with lack of information on one side: a model with verifiable types. *International Journal of Game Theory* 15, 2 (1986), 65–82. Publisher: Springer.

[11] Dean P Foster and Rakesh V Vohra. 1997. Calibrated learning and correlated equilibrium. *Games and Economic Behavior* 21, 1-2 (1997), 40.

[12] Sergiu Hart and Andreu Mas-Colell. 2000. A simple adaptive procedure leading to correlated equilibrium. *Econometrica* 68, 5 (2000), 1127–1150. Publisher: Wiley Online Library.

[13] Sergiu Hart and David Schmeidler. 1989. Existence of Correlated Equilibria. *Mathematics of Operations Research* 14, 1 (1989), 18–25. http://www.jstor.org/stable/3689835 Publisher: INFORMS.

[14] Wan Huang and Bernhard von Stengel. 2008. Computing an extensive-form correlated equilibrium in polynomial time. In *International Workshop on Internet and Network Economics*. Springer, 506–513.

[15] David M Kreps and Robert Wilson. 1982. Sequential equilibria. *Econometrica* 50, 4 (1982), 863–894.

[16] Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisý, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. 2017. DeepStack: Expert-Level Artificial Intelligence in No-Limit Poker. *Science* 356, 6337 (May 2017), 508–513. https://doi.org/10.1126/science.aam6960 arXiv: 1701.01724.

[17] John Nash. 1951. Non-cooperative games. *Annals of mathematics* 54, 2 (1951), 286–295.

[18] Aviad Rubinstein. 2019. *Hardness of Approximation Between P and NP.* Morgan & Claypool.

[19] Bernhard von Stengel and Francoise Forges. 2008. Extensive-form correlated equilibrium: Definition and computational complexity. *Mathematics of Operations Research* 33, 4 (2008), 1002–1022. Publisher: INFORMS.

[20] Martin Zinkevich, Michael Johanson, Michael Bowling, and Carmelo Piccione. 2008. Regret minimization in games with incomplete information. In *Advances in neural information processing systems*. 1729–1736.