Stochastic Regret Minimization in Extensive-Form Games

Gabriele Farina [1] **Christian Kroer** [2] **Tuomas Sandholm** [1 3 4 5]

[1]Computer Science Department, Carnegie Mellon University, Pittsburgh PA 15213 [2]IEOR Department, Columbia University, New York NY 10027 [3]Strategic Machine, Inc. [4]Strategy Robot, Inc. [5]Optimized Markets, Inc.. Correspondence to: Gabriele Farina <gfarina@cs.cmu.edu>, Christian Kroer <christian.kroer@columbia.edu>, Tuomas Sandholm <sandholm@cs.cmu.edu>.

Reviewer: Gyuseok Lee
Date: 10/17/2022

- Overview of paper content (0.5 points)

Typically, Extensive-Form Games (EFG) models are operationalized by computing either a Nash equilibri um of the game, or an approximate Nash equilibrium if the game is large. However, for most real-w orld games this linear program is much too large to solve, either because it does not fit in memory, or because iterations of the simplex algorithm or interior-point methods become prohibitively expensiv e due to matrix inversion. Instead, first-order methods or regret-based methods are used in practice. However, for large games, even these gradient-based methods that require traversing the entire game tree are prohibitively expensive. The original blueprint strategies had to be computed without traversin g the entire game tree, as this game tree is far too large for even a moderate amount of traversals. When full tree traversals are too expensive, stochastic methods can be used to compute approximate gradients instead. The most common stochastic method for solving large EFGs is the Monte-Carlo Cou nterfactual Regret Minimization (MCCFR) algorithm (Lanctot et al., 2009). However, beyond the MCCFR setting, stochastic methods have not been studied extensively for solving EFG. In this method, they i ntroduced a way to combine any regret-minimizing algorithm with any gradient estimator.

- Strengths (0.5 points)

The first of all, it immediately gives a significantly stronger bound on the convergence rate of MCCFR, whereby with probability 1 − p the regret grows as $O(\sqrt{T\log(1/p)})$ instead of $O(\sqrt{T/p})$ as in the ori ginal analysis—an exponentially better bound.

Second, this framework completely decouples the choice of regret minimizer and gradient estimator, th us allowing any regret minimizer to be coupled with any gradient estimator. As a result, it paves the way for many potential future investigations into stochastic methods for EFGs.

- Weaknesses (0.5 points)

When full tree traversals are too expensive, stochastic methods can be used to compute approximate gradients instead. So, I think that there is trade-off between time complexity and performance. So, th is paper should guarantee that stochastic game is better option than any other tree traversals (i.e., e xact algorithm).

- Questions or Discussion (0.5 points)

I wonder that even if the any regret minimizer and gradient estimator can be chosen, there would b e best combination. So, I think that this algorithm can develop the finding the best combination algo rithm.