Stable-Predictive Optimistic Counterfactual Regret Minimization

Gabriele Farina[1] Christian Kroer[2]
Noam Brown[1] Tuomas Sandholm[1,3]

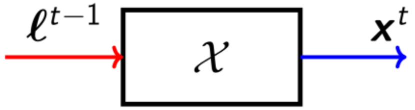
[1] Computer Science Department, Carnegie Mellon University
[2] IEOR Department, Columbia University
[3] Strategic Machine, Inc.; Strategy Robot, Inc.; Optimized Markets, Inc.

Recent Interest in Extensive-Form Games (EFGs)

- EFGs are games played on a game tree
  - Can capture both **sequential** and **simultaneous** moves
  - Can capture **private information**
- **Application**: recent breakthroughs show that it is possible to compute approximate Nash equilibria in large poker games:
  - Heads-Up Limit Texas Hold'Em [Bowling, Burch, Johanson and Tammelin, Science 2015]
  - Heads-Up No-Limit Texas Hold'Em
    - The game has 10^{161} decision points (before abstraction)!
    - Finally reached **superhuman** level (after 20 years of effort) [Brown and Sandholm, Science 2017]

Counterfactual Regret Minimization (CFR)

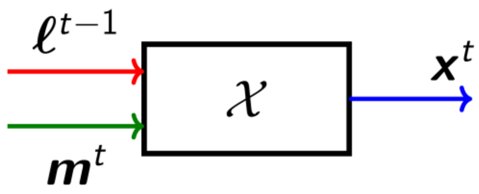
- Defines a class of **regret minimizers**



- Specifically designed for EFGs: regret is **minimized locally** at each decision point in the game
  - By taking into account the **combinatorial structure** of the game tree, it enables **game-specific techniques**, such as pruning subtrees, and warm starting different parts of the tree separately
- Convergence rate $\Theta\left(T^{-1/2}\right)$
- **Practical state of the art for approximating Nash equilibrium in EFGs for 10+ years** (when used in conjunction with alternation and other techniques)

Optimistic (aka *Predictive*) Regret Minimization

- Recent development in online learning
- Idea: inform device with **prediction of next loss**



  - **Accurate** prediction $\Longrightarrow$ **small** regret
  - Several optimistic/predictive regret minimizers are known in the literature, notably Optimistic Follow-the-Regularized-Leader (OFTRL)
  - **Enables convergence rate** of $\Theta(T^{-1})$ to Nash equilibrium in matrix games

- **Natural idea: can we combine CFR's idea of *local* regret minimization with the improved convergence rate of predictive regret minimization?**

Our Contributions

- We present the **first CFR variant which breaks the $\boldsymbol{\Theta(T^{-1/2})}$ convergence rate to Nash equilibrium**, where T is the number of iterations. Our algorithm converges to a Nash equilibrium at the improved rate $O(T^{-3/4})$
- Our algorithm is based on the notion of "**stable-predictive" regret minimizers**, which are a particular type of predictive regret minimizers that we introduce
- Our algorithm **operates locally at each decision point**. We show how different local regret minimizers should be set up differently at different parts of the game tree
  - Main idea: the stability parameter of the different regret minimizers drops exponentially fast with the depth of the decision point
  - Any stable-predictive regret minimizer (such as OFTRL) can be used as long as it respects the requirements on the stability parameter

Poster: **Pacific Ballroom #152** 06:30 - 09:00 pm