2021-Farina-1

Simple Uncoupled No-Regret Learning Dynamics for Extensive-Form Correlated Equilibrium

---

- **Date**: 05-21-2021
- **Link**
- **Authors:**
  - Farina, Gabriele
  - Celli, Andrea
  - Marchesi, Alberto
  - Gatti, Nicola
- **Cites:**
  - 1974-Aumann
  - 2008-Gordon
  - 2008-Huang
  - 2008-von-Stengel
  - 2017-Moravcik
  - 2018-Brown-0
- **Cited by:**
- **Keywords:** #extensive-form-game #extensive-form-correlated-equilibrium
- **Collections:**
- **Status:** #in-progress

---

0 - Abstract

- Study of normal-form equilibria:
  - Simple uncoupled no-regret learning dynamics that converge to correlated equilibria in normal-form games has long been a known result.
  - An equilibrium is achieved when all players seek to minimize their internal regret.

- Study of extensive-form equilibria:
  - Key differences from normal-form games:
    - Sequential and simultaneous moves.
    - Imperfect information.

  - Challenges in computing extensive-form equilibrium (EFCE):
    - EFCE has more complex constraints.
    - EFCE correlation device must account for changes in the agent's beliefs as they make observations during game play.

- Paper contributions:
  - Introduce the first uncoupled no-regret dynamics that converge to the set of EFCEs in n-player general sum EFGs with perfect recall.
  - Can be computed in polynomial time wrt game tree size.
  - Solution converges to a $O(1/\sqrt{T})$-approximate EFCE after T game repetitions.

1 - Introduction

- Correlated vs Uncorrelated equilibrium:
  - *Non-correlated equilibrium* - assumes interaction among players is *decentralized*.
    - Each player determines their strategy independently from the other players.
    - Strategy space is the product of independent strategies.
    - *Nash equilibrium* (NE) - an element of the un-correlated strategy space.
    - Notable results in this area, heads-up no limit poker:
      - 2018-Brown-0
      - 2017-Moravcik.

  - *Correlated equilibrium* (CE) - assumes interaction among players is centralized, commonly modeled by an external *mediator* that coordinates strategies amongst the players.
    - Correlated strategy is a probability distribution over joint action profiles.
    - CE is achieved when no player wants a better strategy from the mediator.
    - Introduced by 1974-Aumann.

  - Strengths of correlated equilibrium:
    - Many real-world interactions require cooperation
      (i.e. optimizing for general-sum utilities).
    - CE can be computed in polynomial time.
    - *Uncoupled learning* - players learn by observing their own payoffs and the strategies of the other players. No knowledge of the other player's payoff functions is required.

  - Weaknesses of non-correlated equilibrium:

- Prone to equilibrium selection issues.
    - How can players select an equilibrium without inter-communication?
- Computing an NE is computationally intractable.
- NEs do not produce strong social welfare, compared to CEs.
- Uncoupled learning dynamics for NE is not well understood.
    - Currently, only defined for the two-player zero-sum case.

- CE and extensive-form games:
  - *Extensive-form correlated equilibrium* (EFCE) - extension of CE to extensive-form games.
    - See 2008-von-Stengel.

  - Extending CE to sequential interactions:
    - The mediator gives a recommended action to each player at a decision point before a sequential interaction occurs.
      - The actions given by the mediator must form a CE strategy.
    - The players are free to deviate from the recommended action.
    - If a player does choose to deviate, then they lose all future recommendations.

  - **Main challenge:** the mediator must account for changes in the agents' beliefs as they make observations through out the game.

- Previous approaches to computing EFCEs for general-sum games:
  - EFCEs can be solved in polynomial wrt to the game tree size
    (without using uncoupled learning dynamics):
    - Shown using multiple methods:
      - Using a variation of the *Ellipsoid Against Hope* algorithm. 2008-Huang.
      - Sampling-based algorithm based on MCMC. 2009-Dudik.
    - These methods are too slow to be used in practice.

- *Paper contribution* - faster EFCE computation using uncoupled learning dynamics:
  - *Phi-regret minimization* - regret defined wrt a given set of linear transformations on the decision set.

  - *Trigger agent* - the trigger agent may deviate from recommended behavior.
    - Introduced by 2008-Gordon.

  - *Trigger regret* - internal regret formulation for the *trigger agent*, defined by the *canonical trigger deviation functions*.
    - This is a modification of *phi-regret minimization*.

  - *Canonical trigger deviation functions* - the set of linear transformations that allows for the trigger agent to behave within the definition of EFCE.

  - **Core result:** If each player plays according to any uncoupled no-regret learning dynamics that minimizes trigger regret, then the resulting empirical frequency of play approaches the set of EFCEs.

  - Provide an algorithm that minimizes trigger regret, converging to a solution in $O(1/\sqrt{T})$ time, where T game repetitions.