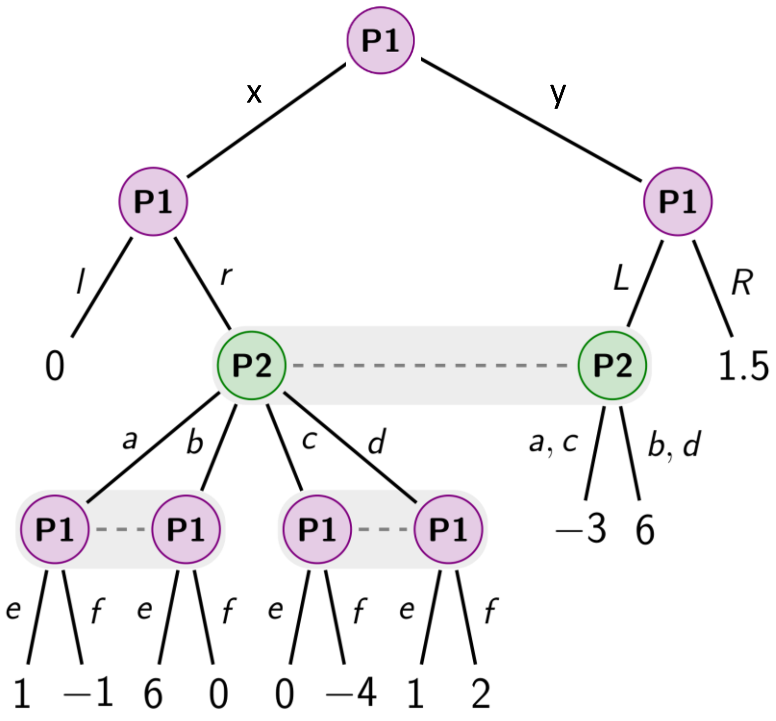
Efficient Regret Minimization Algorithm for Extensive-Form Correlated Equilibrium

Gabriele Farina[1] Chun Kai Ling[1] Fei Fang[2]
Tuomas Sandholm[1,3,4,5]

[1] Computer Science Department, Carnegie Mellon University
[2] Institute for Software Research, Carnegie Mellon University
[3] Strategic Machine, Inc.
[4] Strategy Robot, Inc.
[5] Optimized Markets, Inc.

Extensive-Form Games



- Can capture sequential and simultaneous moves
- Private information
- Each information set contains a set of “indistinguishable” tree nodes
- We assume perfect recall: no player forgets what the player knew earlier

Extensive-Form Correlated Equilibrium (EFCE)

- Introduced by von Stengel and Forges in 2008
- Correlation device selects a recommended strategy for each player before the game starts
  - The correlated **distribution** of strategies is known in advance to all players
- Recommendations are revealed incrementally, move by move, as the players progress in the game tree
  - A recommended move is only revealed to the acting player when the player reaches the decision point for which the recommendation is relevant
  - Players are free to not follow the recommendation, at the cost of future recommendations

Extensive-Form Correlated Equilibrium (EFCE)

- An optimal (e.g., social-welfare-maximizing) mediator that is provably incentive-compatible can be constructed in polynomial time in two-player general-sum games with no chance moves [von Stengel and Forges, 2008]
  - Players can be induced to play strategies with significantly higher social welfare than Nash equilibrium…
  - …despite the fact that each player is free to not follow the recommendations
  - Added benefit: players get told what to do---they do not need to come up with their own optimal strategy as in Nash equilibrium

Computing EFCEs

- Original formulation [von Stengel and Forges, 2008] is based on linear programming
  - Does not scale beyond toy problems
  - Prohibitive amount of memory (>500GB for a game with 1M sequences per player)
- Another paper of ours in NeurIPS-19 ("Correlation in Extensive-Form Games: Saddle-Point Formulation and Benchmarks") formulates the problem as a bilinear saddle point problem and proposes a method based on projected subgradient descent
  - Transforms problem into a zero-sum game between a mediator and deviator, the latter of which is finding the worst possible deviation by the players for the given correlation plan given by the mediator
  - Scales better than an LP, but still faces issues with large games. The main hurdle is the projection onto the set of feasible EFCEs

Regret minimization has become a standard module in leading approaches for finding Nash equilibrium in very large, zero-sum extensive form games
[Bowling et al. Science 2015; Moravcik et al. Science 2017; Brown and Sandholm, Science 2017&2019]

Q: Can regret minimization be used to compute optimal EFCEs in two-player games without chance moves?

A: Yes. We give the first efficient regret minimization algorithm that operates on the set of correlation plans

- Significantly more complicated than the Nash equilibrium case
  - The constraints that define the set of correlation plans lack the clean, hierarchical structure of sequential strategies
  - The constraints form cycles!

Ingredient 1: Scaled Extension

- Powerful operation for constructing certain structured sets, including strategy spaces. We use it to construct the space of EFCEs
- Idea: extend $\mathcal{X}$ with a scaled version of $\mathcal{Y}$

$$\mathcal{X} \stackrel{h}{\triangleleft} \mathcal{Y} := \{(\boldsymbol{x}, \boldsymbol{y}) : \boldsymbol{x} \in \mathcal{X},\ \boldsymbol{y} \in h(\boldsymbol{x})\mathcal{Y}\}.$$

- Scaled extension preserves convexity and compactness of $\mathcal{X}$ and $\mathcal{Y}$

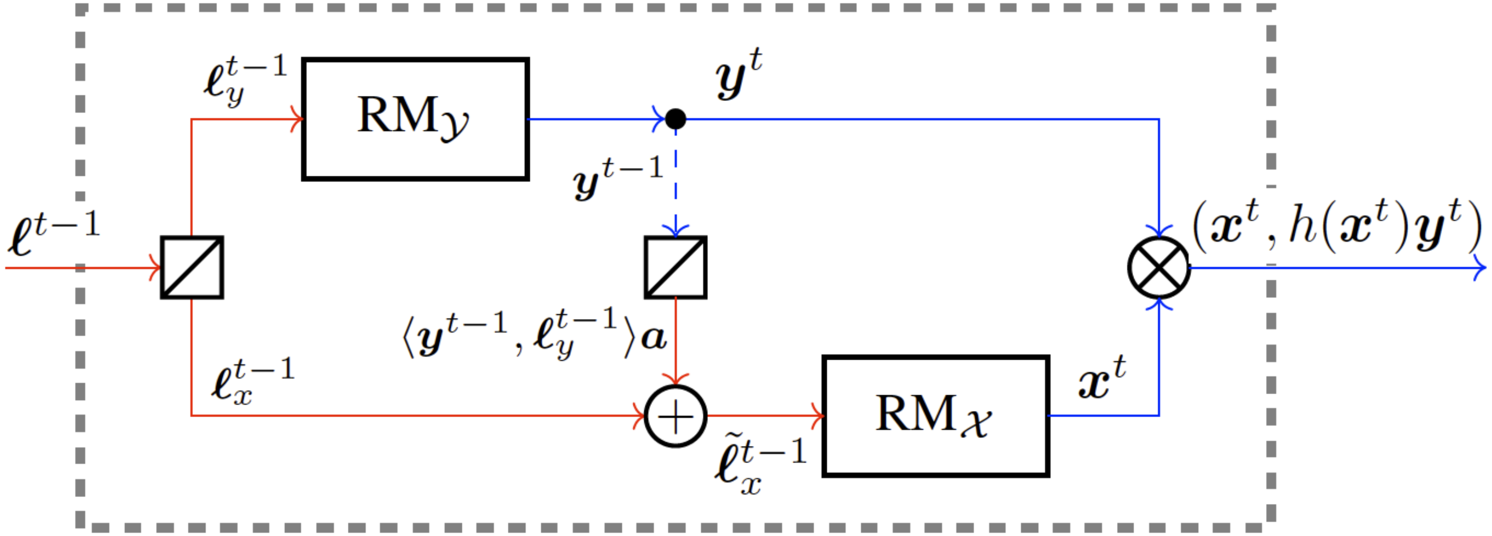
Ingredient 2: Correlation plans as composition of scaled extensions

- Some of the constraints that define the space of correlation plans are redundant and can be safely eliminated
- We propose an algorithm which can safely identify which of these constraints are redundant and removes them
- The remaining constraints form a tree
- The set generated by the remaining constraints can be equivalently generated by composing several scaled extension operations

Ingredient 3: Regret Circuits

[Farina, Kroer, Sandholm ICML'19]

- General methodology for constructing regret minimizers obtained from convexity-preserving operations
  - Given regret minimizers for convex sets $\mathcal{X}$ and $\mathcal{Y}$, can we compose them and construct a regret minimizer for, say, the convex hull/Cartesian product/intersection of $\mathcal{X}$ and $\mathcal{Y}$?
- In this NeurIPS-19 paper we construct a regret circuit for the scaled extension operation



Summary of main contributions

- We introduce **scaled extension**, a novel convexity-preserving operation between sets
- For games with no chance: space of correlation plans may be constructed top down using a series of scaled extension operators
- We show that an **efficient regret minimizer** for the scaled extension of two sets can be constructed starting from any regret minimizer for each individual set
  - Regret circuit approach as in Farina, Kroer, Sandholm [ICML'19]
- Therefore: optimal EFCEs in two-player games without chance can be computed using regret minimization
  - Much faster than subgradient descent
  - Does not need projections: it is guaranteed to always produce feasible iterates