facebook Artificial Intelligence Research

ROBUST MULTI-AGENT COUNTERFACTUAL PREDICTION

Alex Peysakhovich
Adam Lerer
Christian Kroer

PROBLEM

Environment: We observe some existing multi-agent system (e.g. an ad auction, a traffic system, a school assignment mechanism).

Question: What would happen if we changed the rules?

Issue: Real world agents are strategic, if you change the rules, they will change their behavior.

SETUP

Formalization

We consider the standard one-shot Bayesian game setup. There are N players which each have a type $\theta_i \in \Theta$ drawn from an unknown distribution $\mathcal{F}$. This type is assumed to represent their preferences and private information. For example, in the case of auctions this type describes the valuations of each player for each object.

Definition 1. *A game $\mathcal{G}$ has a set of actions for each player $\mathcal{A}_i$ with generic element a_i. After each player chooses their action, the players receive utilities given by $u_i^{\mathcal{G}}(a_1, \dots, a_N, \theta_i)$.*

Definition 2. *An Bayesian Nash equilibrium (BNE) is a strategy profile σ^* such that for each player i, all possible types θ_i for that player which have positive probability under $\mathcal{F}$, and any other strategy σ_i' we have*

$$\mathbb{E}_{\mathcal{F}}\left[u_i^{\mathcal{G}}(\sigma_i^*(\theta_i), \sigma_{-i}^*(\theta_{-i}), \theta_i)\right] \geq \mathbb{E}_{\mathcal{F}}\left[u_i^{\mathcal{G}}(\sigma_i'(\theta_i), \sigma_{-i}^*(\theta_{-i}), \theta_i)\right].$$

Question 1. *Suppose we have a dataset $\mathcal{D}$ of actions played in $\mathcal{G}$. What can we say about what would happen if we changed the underlying game to $\mathcal{G}'$?*

Standard Assumptions
(Inverse Reinforcement Learning or Structural Modeling)

Assumption 1 (Equilibrium)**.** *Data is drawn from a BNE of $\mathcal{G}$ and play in $\mathcal{G}'$ will form a BNE.*

Assumption 2 (Identification)**.** *For any possible distribution of types $\mathcal{F}$ and associated BNE σ^* there does not exist another distribution of types $\mathcal{F}'$ and BNE σ'^* that induces the same distribution of actions.*

Assumption 3 (Uniqueness in $\mathcal{G}'$)**.** *Given $\mathcal{F}$ there is a unique BNE in $\mathcal{G}'$.*

Assumption 4 (Specification)**.** *$\mathcal{G}$ and $\mathcal{G}'$ include the correct specifications of individuals' reward functions.*

RMAC as Relaxation

Revelation Game

$$\text{Regret}_j^{\mathcal{G}}(\hat{\theta}_j, \mathcal{D}_{-j}) = \max_{a_j} \mathbb{E}\left[u_j^{\mathcal{G}}(a_j, \hat{\theta}_j, \mathcal{D}_{-j})\right] - \mathbb{E}\left[u_j^{\mathcal{G}}(d_j, \hat{\theta}_j, \mathcal{D}_{-j})\right].$$

$$\text{Regret}_j^{\mathcal{G}'}(\hat{a}_j, \hat{\theta}_j, \hat{a}_{-j}) = \max_{a_j} \mathbb{E}\left[u_j^{\mathcal{G}'}(a_j, \hat{\theta}_j, \hat{a}_{-j})\right] - \mathbb{E}\left[u_j^{\mathcal{G}'}(\hat{a}_j, \hat{\theta}_j, \hat{a}_{-j})\right].$$

$$\mathcal{L}_j^{rev}(\hat{\theta}_j, \hat{a}_j, \hat{a}_{-j}, \mathcal{D}) = \max\{\text{Regret}_j^{\mathcal{G}}(d_j, \hat{\theta}_j, \mathcal{D}), \text{Regret}_j^{\mathcal{G}'}(\hat{a}_j, \hat{\theta}_j, \hat{a}_{-j})\}.$$

Theorem 1. *If assumptions 1-3 are satisfied then the revelation game has a unique BNE where each agent reveals their true type and counterfactual action.*

Theorem 1. *Let $(\mathcal{G}, \mathcal{G}')$ be the real game/counterfactual game, let $(\mathcal{G}_m, \mathcal{G}_m')$ be misspecified versions of these two games with same type/action spaces but*

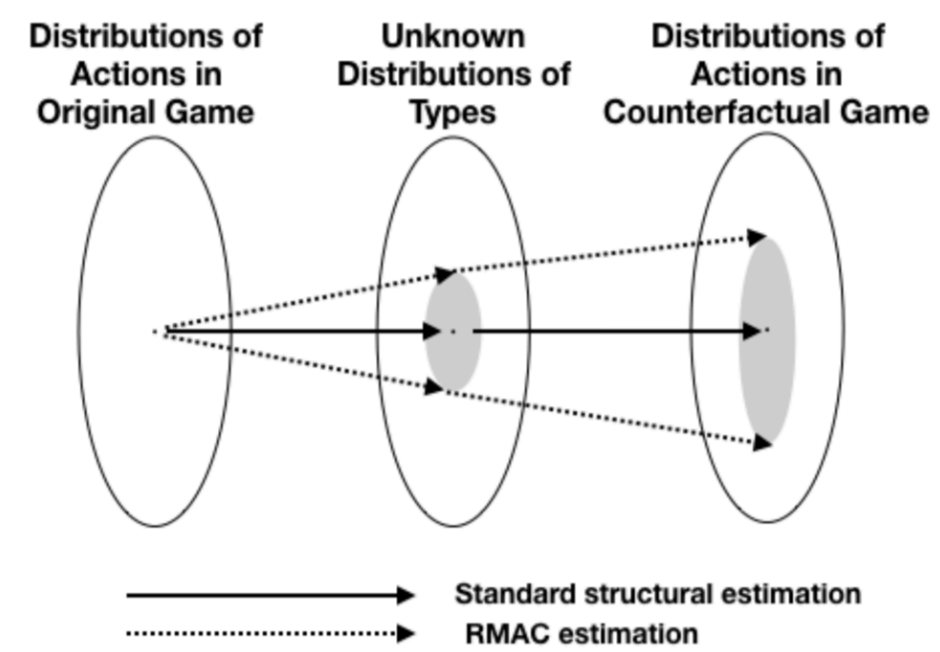
$$||u_m - u||_\infty \leq \frac{\epsilon}{2}$$

and

$$||u_m' - u'||_\infty \leq \frac{\epsilon}{2}.$$

Let $\mathcal{D}$ be some data. If $r^* = (\hat{a}, \hat{\theta})$ is a BNE of the real revelation game corresponding to $(\mathcal{G}, \mathcal{G}', \mathcal{D})$ then r^* is an ϵ-BNE of the misspecified revelation game corresponding to $(\mathcal{G}_m, \mathcal{G}_m', \mathcal{D})$

RMAC BOUNDS



SOLVING FOR RMAC BOUNDS

Negative Result

Theorem 2. *It is NP-hard to compute the robust counterfactual estimate even if each data-point j has only a single feasible type, and there are only two data points. It is also NP-hard even if there is no objective function, a finite number of feasible types, and $\mathcal{G}'$ has only two players.*

Solving RMAC with First Order Methods

Algorithm 1 Revelation Fictitious Play

Input: $\epsilon, \mathcal{D}, V, \mathcal{G}, \mathcal{G}'$, if pessimistic then $\alpha = -1$, if optimistic then $\alpha = 1$
Randomly initialize $\hat{\theta}_i^0, \hat{a}_i^0$
for $t = 0, \dots$ while not converged **do**
  Let $\bar{a}_{-i}^t$ be the historical distribution of $\hat{a}_{-i}^{t'}$ for $t' \in \{0, \dots, t\}$
  Let $\bar{\sigma}_{-i}^t$ be the (mixed) strategy profile implied by the historical distribution of $(\hat{\theta}_{-i}^{t'}, \hat{a}_{-i}^{t'})$
  Let the set of low-regret revelation game actions be

$$\hat{\mathcal{C}}_i^t = \{(\hat{\theta}_i, \hat{a}_i) \in \Omega \times \mathcal{A} \mid \mathcal{L}_i^{rev}(\hat{\theta}_i, \hat{a}_i, \bar{a}_{-i}^t, \mathcal{D}) \leq \epsilon\}$$

  Breaking ties randomly, update guesses for each datapoint

$$(\hat{\theta}_i^{t+1}, \hat{a}_i^{t+1}) = \text{argmax}_{\hat{\theta}_i, \hat{a}_i \in \hat{\mathcal{C}}_i^t} \left[\alpha V(\hat{\theta}_i, \hat{a}_i, \bar{\sigma}_{-i}^t)\right].$$

Definition 4. *RFP* converges *to a mixed strategy σ^* if* $\lim_{t \to \infty} \bar{\sigma}^t = \sigma^*$.

We use the following notion of local optimality (analogously defined for optimistic V):

Definition 5. *A mixed ϵ-BNE σ^* of the revelation game is* locally V-optimal *if*

$$V(\sigma^*) \leq V(\theta_j, a_j, \sigma_{-j}^*)$$

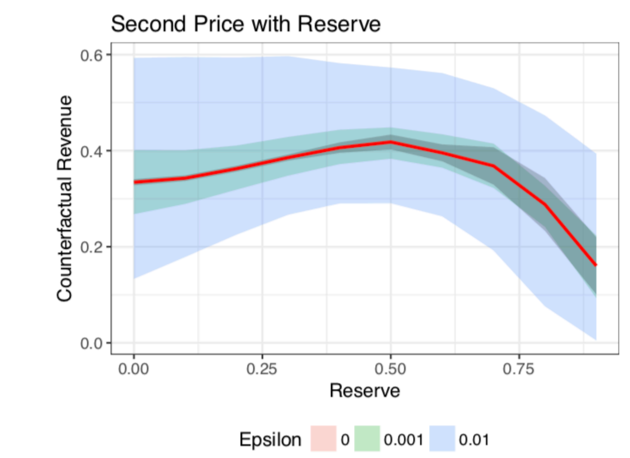
for any data-player j and unilateral deviation (θ_j, a_j) where[5]

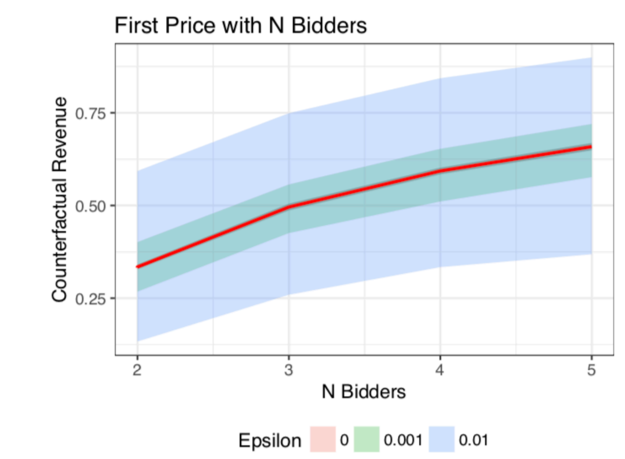
$$\mathbb{E}_{(\theta_{-j}, a_{-j}) \sim \sigma_{-j}^*}[\mathcal{L}_j^{rev}(\theta_j, a_j, a_{-j}, \mathcal{D})] < \epsilon.$$

Theorem 3. *If RFP converges to σ^* then σ^* is a locally V-optimal ϵ-BNE of the revelation game.*

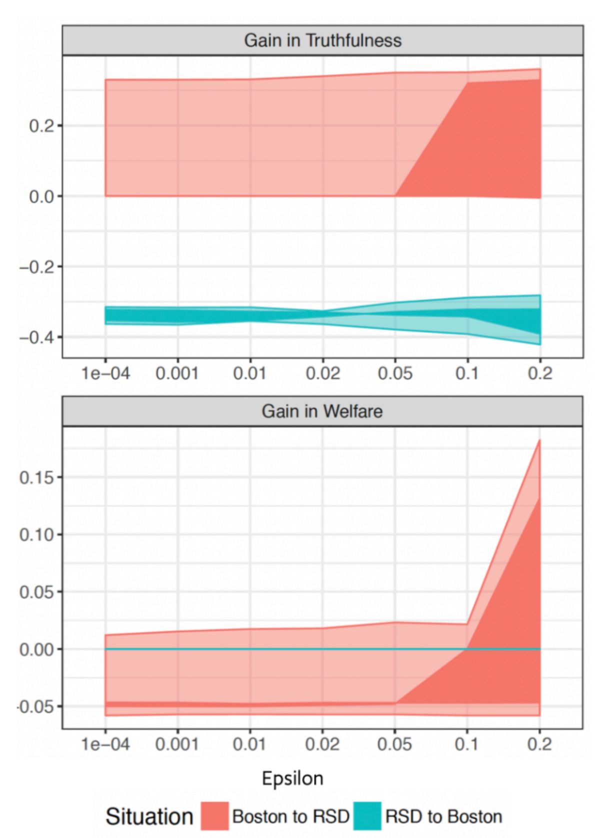
EXPERIMENTS

Auctions





School Choice



Social Choice