Incentive Compatible Active Learning

Federico Echenique (Caltech)
Siddharth Prasad (CMU)

ITCS 2020

Motivation: model of active learning with incentives for experimental design in economics

- Economic experiments: learner seeks to elicit parameters governing agent's preferences
- Experiments always incentivized: payoff to agent depends on questions/answers in the experiment.

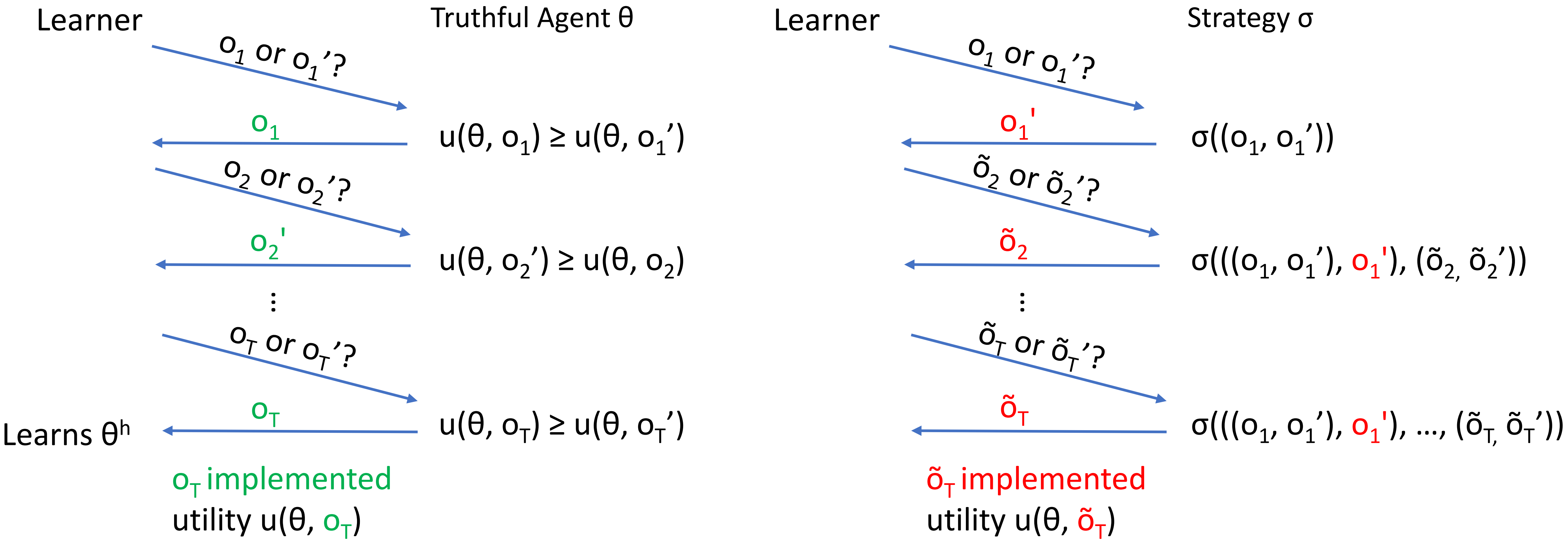
- Passive vs. Active learning:
  - Passive: labeled points (x_t, $y_t \in \{0, 1\}$) produced by unknown distribution. How many samples needed to learn?
  - Active: learner chooses points x_t on which to reveal label. How many labels required to learn?
    - Known as *membership queries* model in the active learning literature

- Large and growing body of work on both *economic applications of PAC learning* [Beigman and Vohra, EC '06; Zadimoghaddam and Roth, WINE '12; Balcan, Daniely, Mehta, Urner, and Vazirani, WINE '14; Basu and Echenique, EC '18; Chase and Prasad, ITCS '19] and *learning with incentive issues* [Abernethy, Chen, Ho, and Waggoner, EC '15; Hardt, Megiddo, Papadimitriou, and Wooters, ITCS '16; Liu and Chen, EC '17; Chen, Podimata, Procaccia, and Shah, EC '18].

Preference elicitation setup

- Type space Θ, equipped with metric d, Θ bounded wrt d.
- Outcome space O.
- Agent of type $\theta \in \Theta$ has utility $u(\theta, o)$ for outcome $o \in O$.
- θ induces preference relation $\succcurlyeq$ where $o \succcurlyeq o'$ iff $u(\theta, o) \geq u(\theta, o')$.

Learner executes learning algorithm to learn agent's true type $\theta \in \Theta$.



- **Learning:** $d(\theta, \theta^h) \leq \varepsilon$ w.p. $\geq 1-\delta$. Min number of rounds $q(\varepsilon, \delta)$ achieving this: *learning complexity*
- **Incentive compatibility:** For all types θ, strategies σ, $u(\theta, o_T(\text{truthful})) \geq u(\theta, o_T(\sigma)) - \tau$ w.p. $\geq 1-\nu$. Min number of rounds $T(\tau, \nu)$ achieving this: *IC complexity*.
- **$(\varepsilon, \delta, \tau, \nu)$-incentive compatible learning algorithm:** achieves (ε, δ)-learning and (τ, ν)-incentive compatibility. $\text{Max}(q(\varepsilon, \delta), T(\tau, \nu))$ called *IC learning complexity.*

Discretizing the type space:

Suppose $s : \Theta \to O$ is such that $u(\theta, s(\theta')) > u(\theta, s(\theta''))$ iff $d(\theta, \theta') < d(\theta, \theta'')$ for all $\theta, \theta', \theta''$ (s is *effective*).

Example. *Expected utility preferences over $O = R^n$.*

- Agent's type is a probability $\alpha = (\alpha_1,\ldots, \alpha_n)$ [belief over n uncertain "states of the world"]
- Learner offers outcomes $x \in R^n$ where x_i is reward if state i is realized
- $u(\alpha, x) = E_{i\leftarrow\alpha}(x_i) = \alpha.x$

Spherical scoring rule: $s(\alpha) = \frac{\alpha}{\|\alpha\|}$ satisfies $E_\alpha(s(\beta)) > E_\alpha(s(\gamma))$ iff $dist(\alpha, \beta) < dist(\alpha, \gamma)$ (renormalized L_2 dist)

IC search over type space: Let $\{\theta_1,\ldots,\theta_N\}$ be an ε-cover of Θ wrt d.

Iterate over cover to find most preferred outcome $s(\theta_{best})$ using choices "$s(\theta_i)$ or $s(\theta_j)$?".

Agent of type θ will answer according to the θ_{best} closest to θ.

Deterministic (ε, 0)-learning, (0, 0)-incentive compatibility

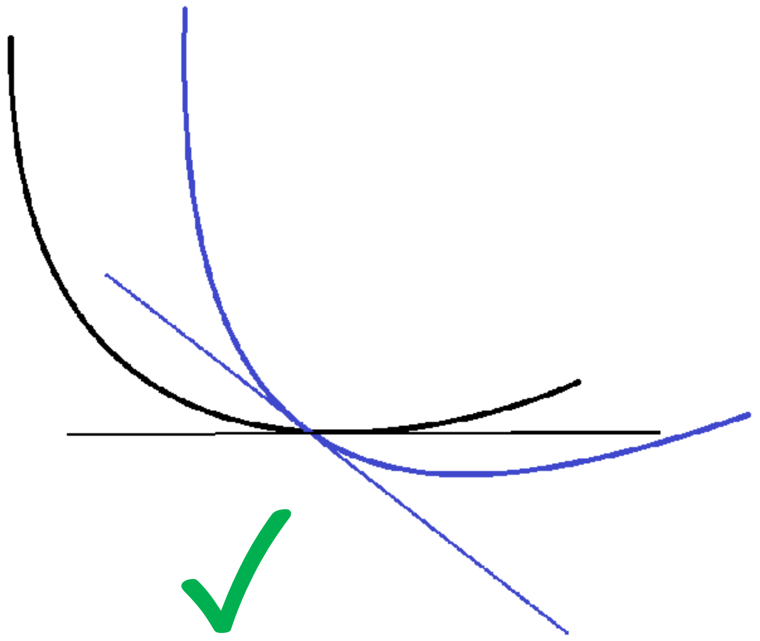
Inefficient: N might be large, e.g. $N_\varepsilon(\Delta_n) = O((n/\varepsilon)^n)$ wrt tvd

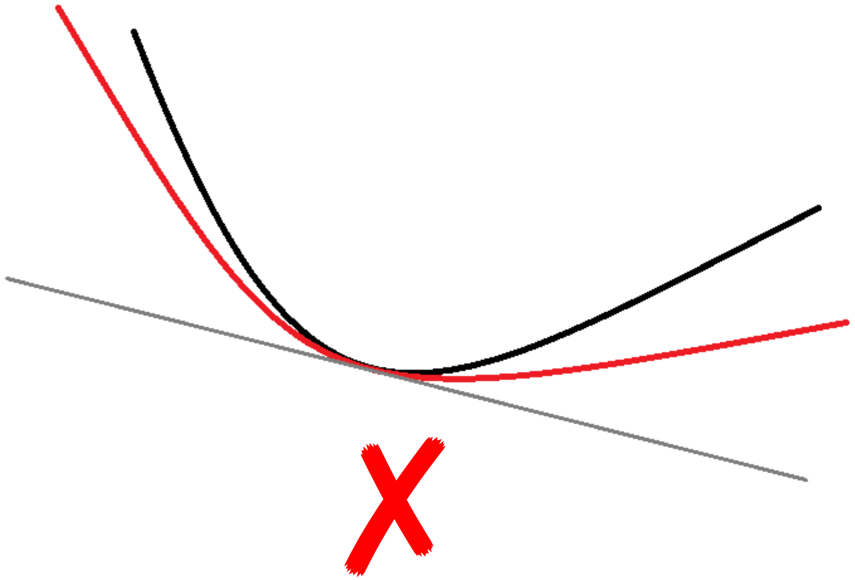
Hyperplane Uniqueness: a sufficient condition for (potentially inefficient) IC learnability

- Outcome space R^n. Agent type induces preference relation $\succcurlyeq$ over R^n.
- Upper contour set: $C(x) = \{y : y \succcurlyeq x\}$

Theorem. If all upper contour sets $C(x)$ are closed and convex, any supporting hyperplane of $C(x)$ is unique, and *hyperplane uniqueness* holds, then there is a metric d on Θ wrt which Θ is $(\varepsilon, 0, 0, 0)$-IC learnable.

Hyperplane uniqueness: For every $x \in R^n$ and types $\succcurlyeq_1$, $\succcurlyeq_2$ the supporting hyperplanes of the upper contour sets $C_1(x)$ and $C_2(x)$ are distinct.





$O(n^{3/2} \log n \max\{\log(n/\varepsilon), \log(1/\tau)\})$ IC learning algorithm for Expected Utility preferences

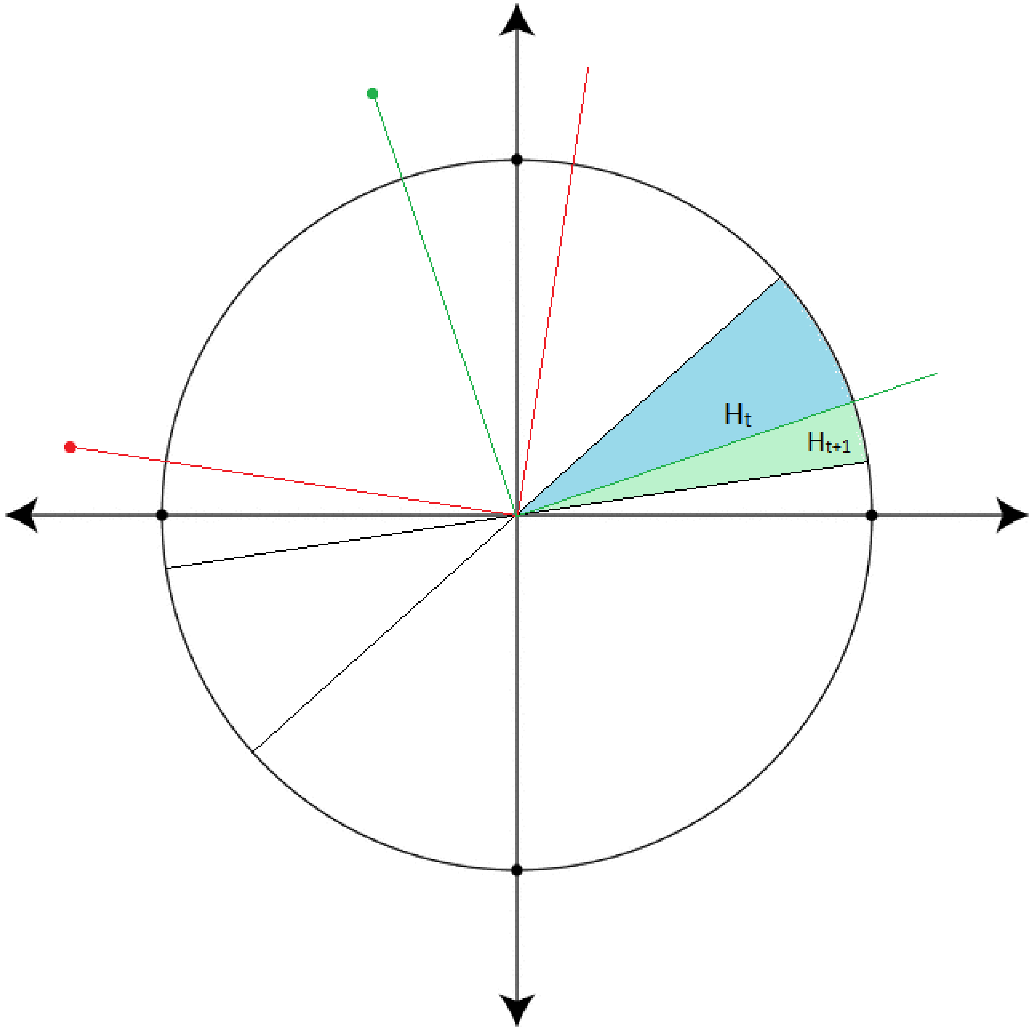
- Agent's type is a probability $\alpha = (\alpha_1, \ldots, \alpha_n)$ [belief over n uncertain "states of the world"]
- Learner asks agent to make choices (x_t, y_t), $x_t, y_t \in R^n$ encoding rewards for realization of each state.
- $u(\alpha, x) = E_\alpha(x) = \alpha.x$

Disagreement based active learning of linear separators (from [Dasgupta2011]):

1. $H_1 = S^{n-1}$
2. For t = 1,2,…:
    1. Receive unlabeled point x_t
    2. If there are $w, w' \in H_t$ such that $\text{sign}(w.x_t) \neq \text{sign}(w'.x_t)$, get 0/1 label for x_t. $H_{t+1} = \{w \in H_t : \text{sign}(w.x_t) = \text{label}(x_t)\}$
    3. Else $H_{t+1} = H_t$

Requests $O(\theta\ VC(H) \log(1/\varepsilon))$ labels

θ = "disagreement coefficient"

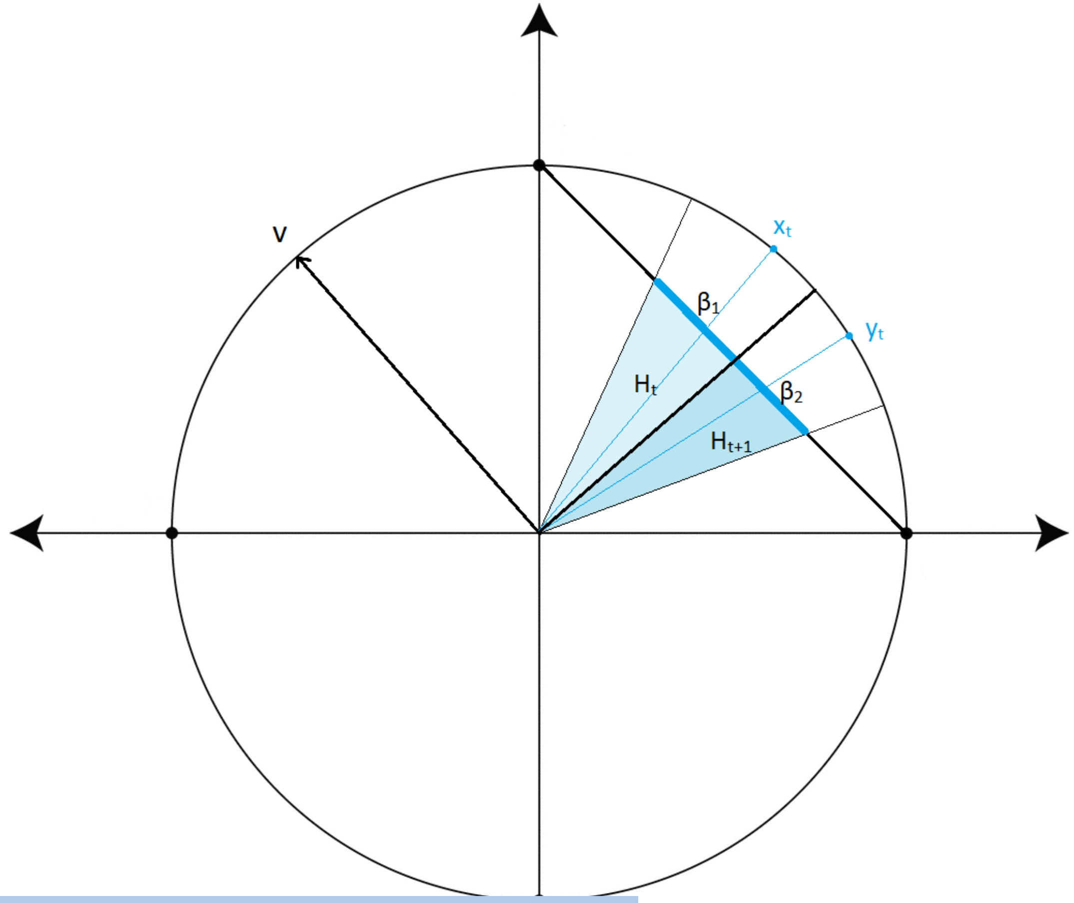


$O(n^{3/2} \log n \max\{\log(n/\varepsilon), \log(1/\tau)\})$ IC learning algorithm for Expected Utility preferences

- Agent's type is a probability $\alpha = (\alpha_1,\ldots, \alpha_n)$ [belief over n uncertain "states of the world"]
- Learner asks agent to make choices (x_t, y_t), $x_t, y_t \in R^n$ encoding rewards for realization of each state.
- $u(\alpha, x) = E_\alpha(x) = \alpha.x$

Use spherical scoring rule to ensure incentive compatibility

1. $H_1 = \Delta_n$
2. For $t = 1,2,\ldots,T$:
    1. Choose v uniformly at random from S^{n-1} (if no disagreement, resample)
    2. Find β_1, β_2 s.t. $s(\beta_1) - s(\beta_2)$ is a scalar multiple of v. Ask agent to choose between $x_t = s(\beta_1)$ and $y_t = s(\beta_2)$.
3. Pay agent based on preference from (x_T, y_T).



$O(n^{3/2} \log n \log(n/\varepsilon))$ rounds: any best responding agent reports within ε-tvd of true type (whp)

Question: is there a combinatorial characterization of IC learning? Like VC dimension for PAC learning, Littlestone dimension for online learning, private learning (conjectured), etc.