6.S890: Topics in Multiagent Learning

Lecture 19

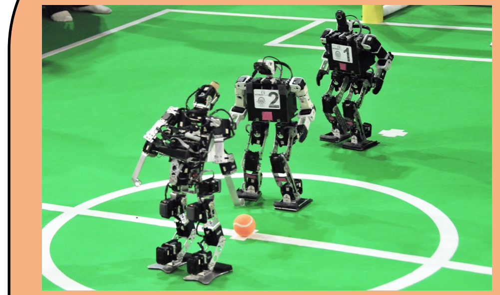
Fall 2023

Context: Increasing Interest in Multi-Agent Learning

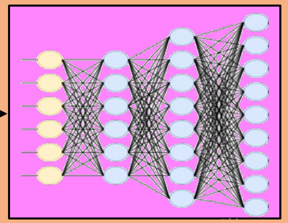


Multi-player Game-Playing:

- Superhuman Chess, Go, Poker, Gran Turismo
- Good StarCraft, Diplomacy



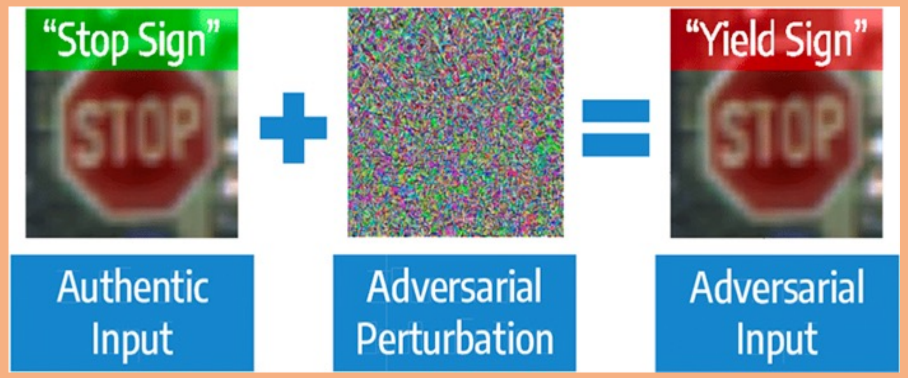
- Multi-robot interactions
- Autonomous driving
- Automated Economic policy design

$Z \sim \mathcal{N}(0, I) \longrightarrow$  $\longrightarrow$  $\sim P_{\text{target}}$

boring randomness | neural network | interesting randomness

Generative Adversarial Networks (GANs)

synthetic data generation



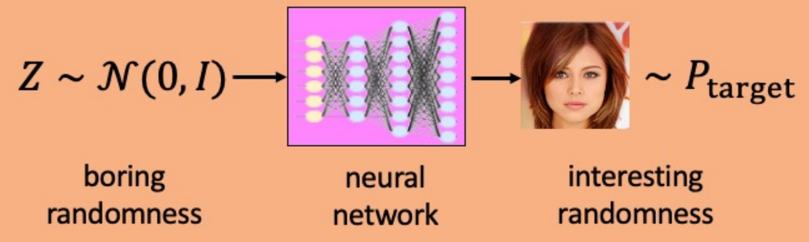
Adversarial Training

robustifying models against adversarial attacks

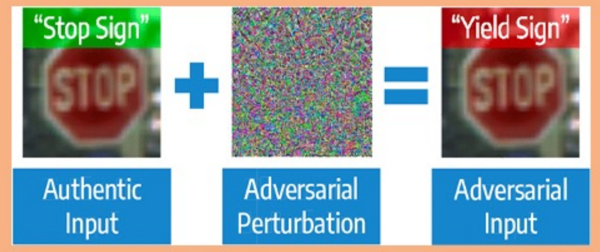
Context: Increasing Interest in Multi-Agent Learning

Multi-player Game-Playing:
- Superhuman Chess, Go, Poker, Gran Turismo
- Good StarCraft, Diplomacy

- Multi-robot interactions
- Autonomous driving
- Automated Economic policy design



Generative Adversarial Networks (GANs)
synthetic data generation



Adversarial Training
robustifying models against adversarial attacks

Important notes and caveats…

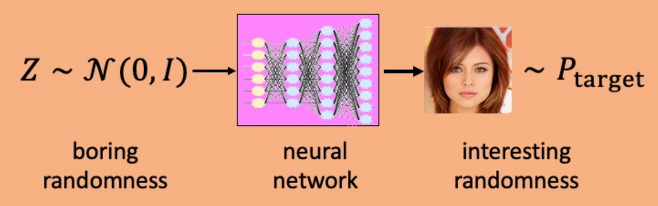
Important Caveats…

- (I) Strategic Behavior does not emerge from standard training
- (II) Naively trained models can be manipulated
- (III) Training without regard to the presence of other agents can lead to undesirable (e.g. collusive) consequences
- (IV) The optimization workhorse of Deep Learning (a.k.a. gradient descent) struggles in multi-agent settings
- (V) Finally Game Theory (namely the existence of Nash equilibrium and other types of equilibrium) breaks

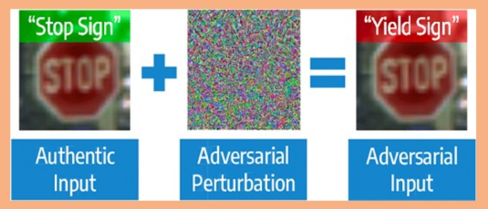
Motivating Questions

Multi-player Game-Playing:
- Superhuman GO, Poker, Gran Turismo
- Human-level Starcraft, Diplomacy

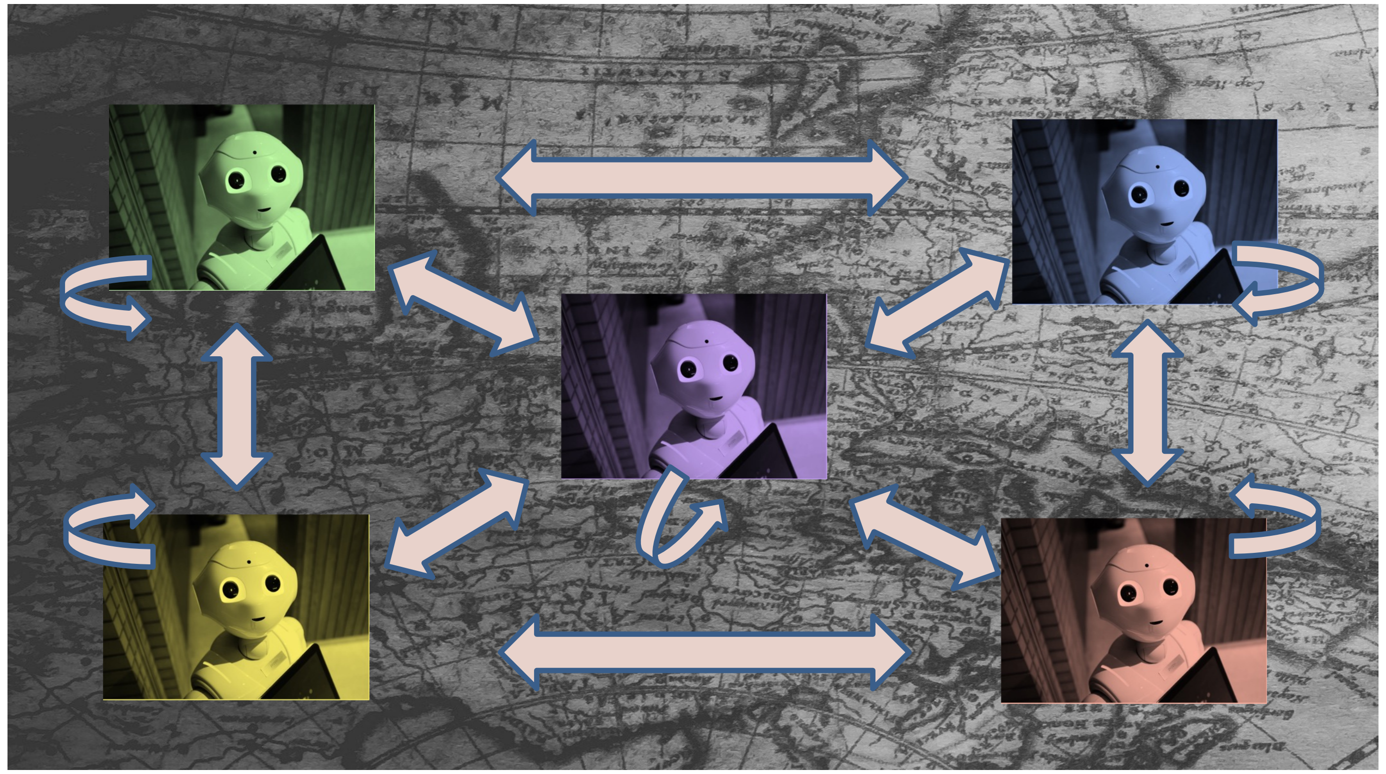
- Multi-robot interactions
- Autonomous driving
- Automated Economic policy design



Generative Adversarial Networks (GANs)
synthetic data generation



Adversarial Training
robustifying models against adversarial attacks



Practical Experience: GD vs GD (vs GD…) is cyclic or chaotic, and it is a hard engineering challenge to make it identify a good solution

What are meaningful and practically attainable optimization targets in this setting?

PARTIAL SUCCESS

Why does GD vs GD struggle even in two-player zero-sum cases?

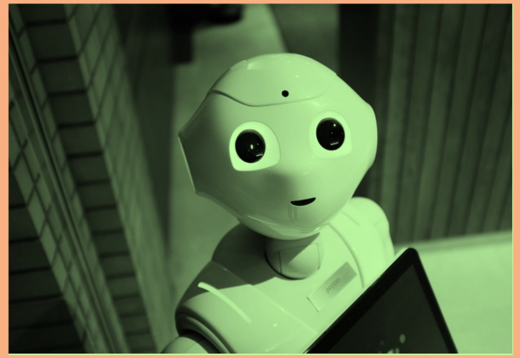
SUCCESS

Is there a generic optimization framework for Multi-Agent Deep Learning?

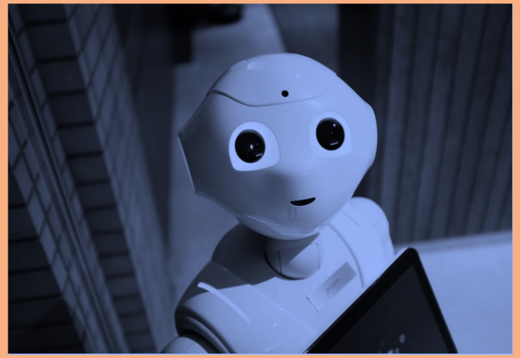
NO REAL SUCCESS YET

Local Nash Equilibrium

Setting:

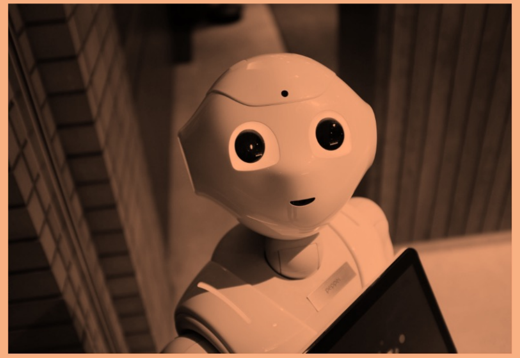


action: $x_1 \in \mathcal{X}_1 \subset \mathbb{R}^{d_1}$
goal: max $u_1(x_1, \dots, x_n)$



action: $x_2 \in \mathcal{X}_2 \subset \mathbb{R}^{d_2}$
goal: max $u_2(x_1, \dots, x_n)$

…



action: $x_n \in \mathcal{X}_n \subset \mathbb{R}^{d_n}$
goal: max $u_n(x_1, \dots, x_n)$

u_i is Lipschitz and smooth (i.e. has Lipschitz gradient) a.e.
[allow: global constraints $(x_1, x_2, \dots, x_n) \in \mathcal{S} \subseteq \times_i \mathcal{X}_i$]

Local Nash: A point $x^* = (x_1^*, \dots, x_n^*) \in \mathcal{S}$ s.t. for each player i, x_i^* is **local max** of $u_i(x_i\, ;\, x_{-i}^*)$ w.r.t. x_i

First-Order Local Nash: Take "**local max**" to mean "**1st-order local max**" i.e. max w.r.t. 1st-order Taylor appx

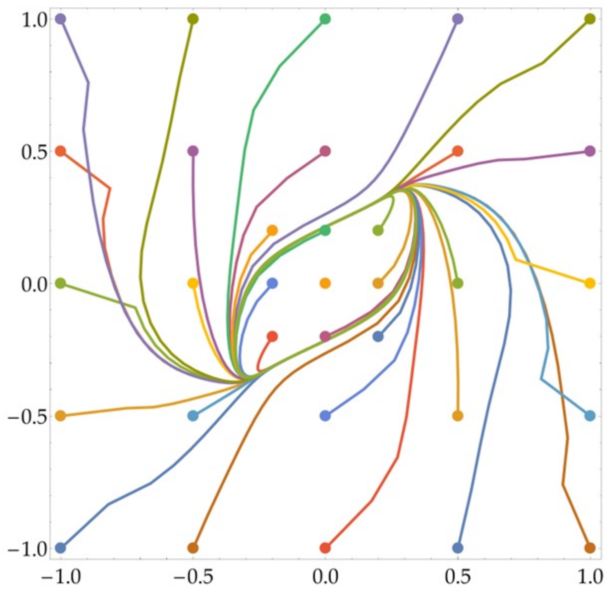
Equivalently: $\forall i$: $x_i^* = \Pi_{\mathcal{S}_i(x_{-i}^*)}\left(x_i^* + \nabla_{x_i} u_i(x_i^*\, ;\, x_{-i}^*)\right)$,
where $\mathcal{S}_i(x_{-i}^*) = \{x_i \mid (x_i\, ;\, x_{-i}^*) \in \mathcal{S}\}$, and $\Pi_{\mathcal{S}_i(x_{-i}^*)}(\cdot)$ is the Euclidean projection

Proposition: If $\mathcal{S}$ is convex and compact, a *first-order local Nash equilibrium* exists. GENERALIZES LOCAL OPT

[Daskalakis-Skoulakis-Zampetakis STOC'21]: *First-order local Nash equilibrium is intractable even for two-player zero-sum games.* EXPLAINS WHY GD vs GD FAILS BUT WORST-CASE INTRACTABILITY

Way Forward 1: Practical Local Nash Equilibrium

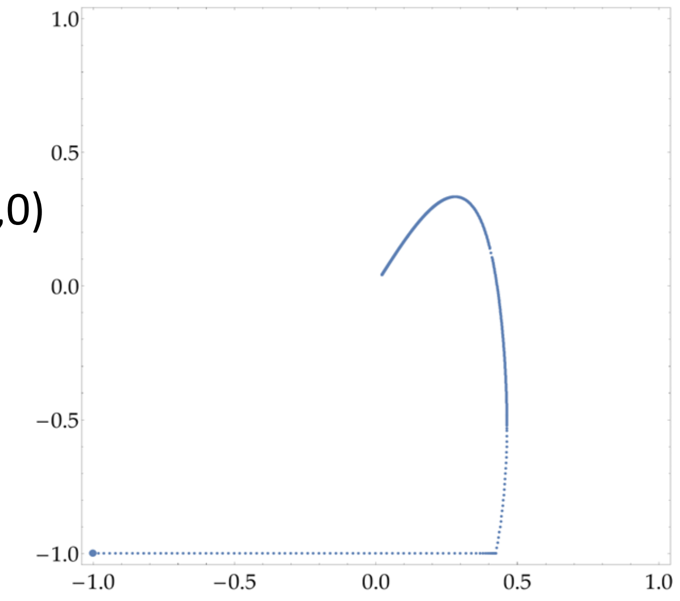
- *Practical Local Nash Equilibrium Computation*?
  - local Nash is intractable in the worst-case
  - …but can exploit connection to Brouwer fixed points to obtain 2nd-order dynamics with guaranteed (albeit necessarily not poly-time) convergence **[Daskalakis-Golowich-Skoulakis-Zampetakis COLT'23]**
  - turn it into a 1st-order method by cutting corners ?
  - identify structural properties of games under which it is efficient (beyond worst-case analysis of games)



(a) $f_1(\theta, \omega)$.

gradient descent

local Nash is at (0,0)

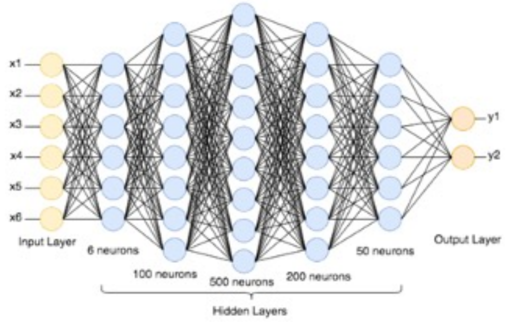


(a) $f_1(\theta, \omega)$.

our algorithm: Stay On the Ridge (or STON'R)

Way Forward 2: Consider Randomized Equilibria

- *Local* Correlated/Coarse Correlated equilibria?
  - what's a reasonable way to define it in general non-concave games?
  - …so that it is also guaranteed to exist and is tractable?
  - proposal: $||\mathbb{E}_{x^*\sim p}\left[\nabla_{x_i} u_i(x_i^*; x_{-i}^*)\right]|| \leq \varepsilon$ (formally: project to the constraint set)
  - when p has support 1 this is a local Nash eq, so this exists but is intractable
  - is there some polynomial support, so that it is tractable?
  - **[Cai-Daskalakis-Luo-Wei-Zhang'23]:** If $\mathcal{S}$ is convex and compact and the u_i's are Lipschitz and and smooth, a poly-size supported (in the dimension, in $1/\varepsilon$, in the Lipschitzness and the smoothness of the utilities) local CCE exists can be computed efficiently (using Gradient Descent) ☺

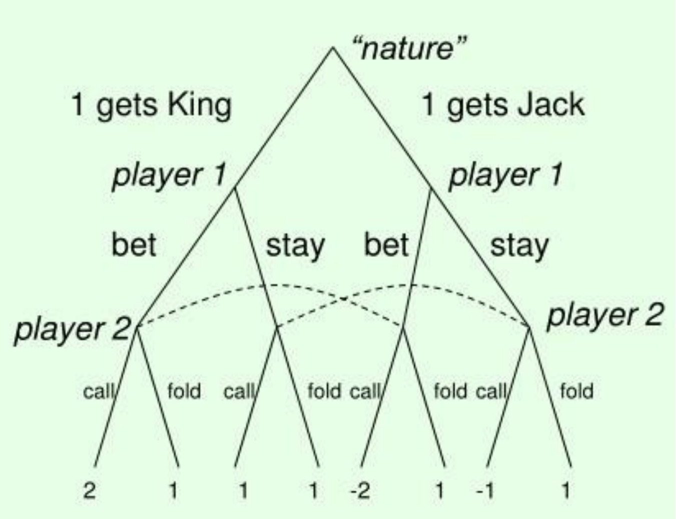


semi-agnostic

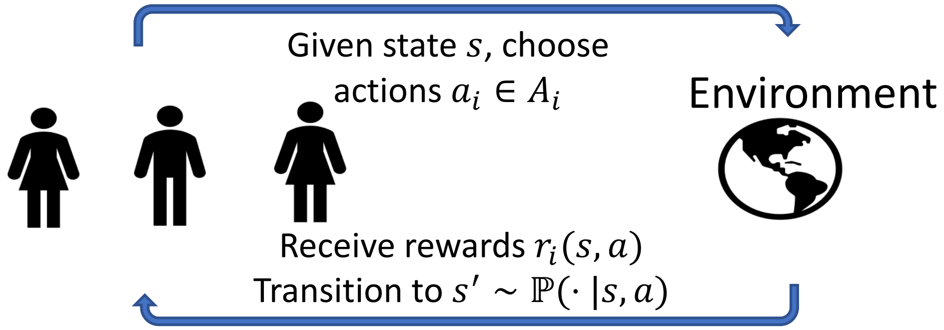
$+\ x_{t+1} \leftarrow x_t - \nabla_x \ell(x_t)\ +$

$+$

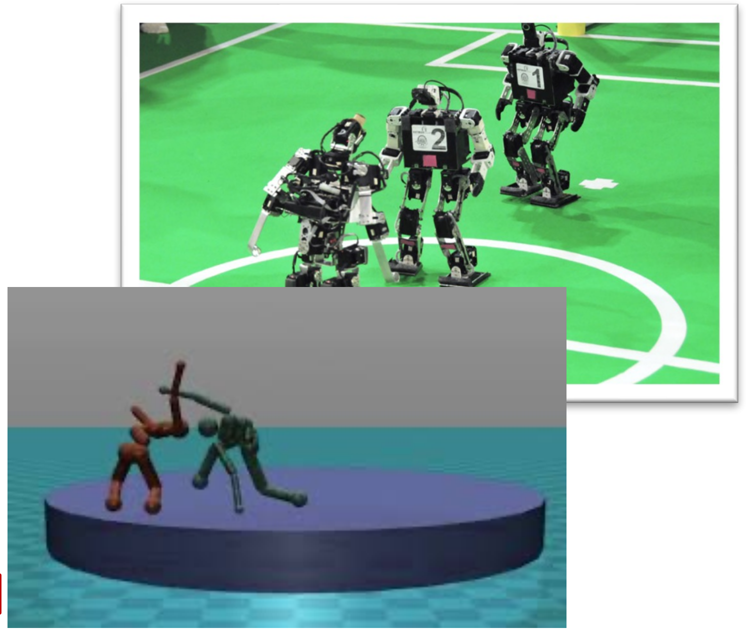
Way Forward 3: Special Structure (Lectures 9-17)



extensive form games

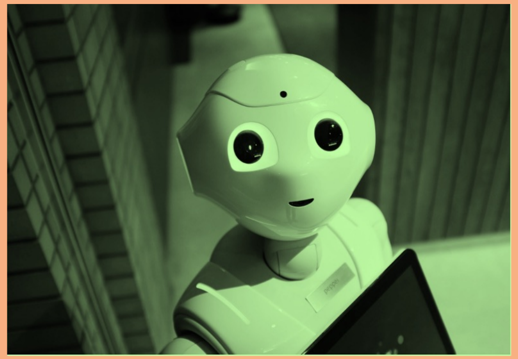


stochastic games **[Shapley'53]**
multi-agent reinforcement learning **Littman'94]**

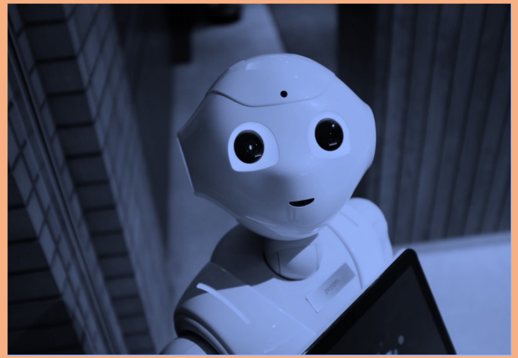


Way Forward 4: *Global* Randomized Equilibria!?!

Setting:

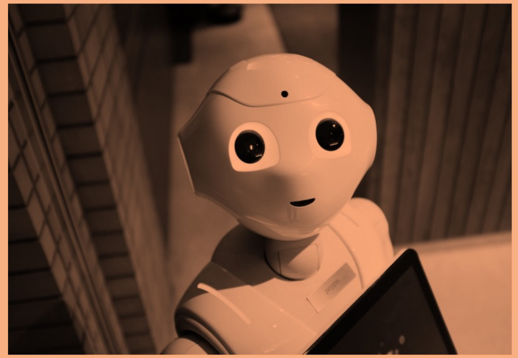


action: $x_1 \in \mathcal{X}_1 \subset \mathbb{R}^{d_1}$
goal: max $u_1(x_1, \dots, x_n)$



action: $x_2 \in \mathcal{X}_2 \subset \mathbb{R}^{d_2}$
goal: max $u_2(x_1, \dots, x_n)$

…



action: $x_n \in \mathcal{X}_n \subset \mathbb{R}^{d_n}$
goal: max $u_n(x_1, \dots, x_n)$

Nash Eq: *A collection of* $x_1^*, \dots x_n^*$ s.t. for all i, x_i: $u_i(x_i^*; x_{-i}^*) \geq u_i(x_i; x_{-i}^*)$

Randomized Nash Eq: *A collection of distributions* $p_1, \dots p_n$ s.t. for all i, x_i:

$$\mathrm{E}_{x^* \sim p_1 \times \dots \times p_n}[u_i(x_i^*; x_{-i}^*)] \geq \mathrm{E}_{x^* \sim p_1 \times \dots \times p_n}[u_i(x_i; x_{-i}^*)]$$

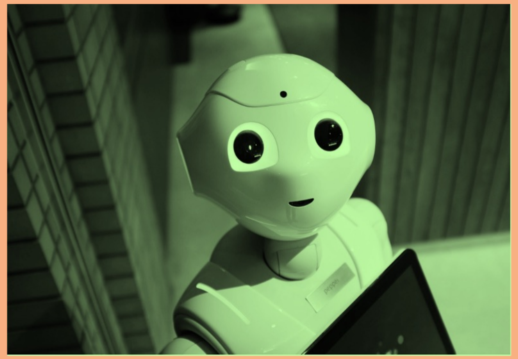
Coarse Correlated Eq: *A joint distribution of* p s.t. for all i, x_i:

$$\mathrm{E}_{x^* \sim p}[u_i(x_i^*; x_{-i}^*)] \geq \mathrm{E}_{x^* \sim p}[u_i(x_i; x_{-i}^*)]$$

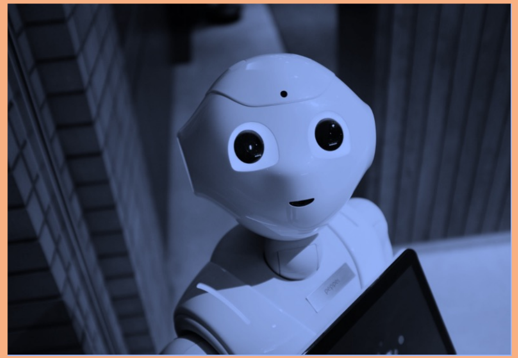
If some $u_i(x_i; x_{-i})$ is not concave in x_i for all x_{-i}, Nash equilibrium does not necessarily exist
[Glicksberg'52]: A *randomized* Nash equilibrium does exist if the $\mathcal{X}_i$'s are compact and the u_i's are continuous (and not necessarily concave), *but support could be uncountably infinite*.

Way Forward 4: *Global* Randomized Equilibria!?!

Setting:

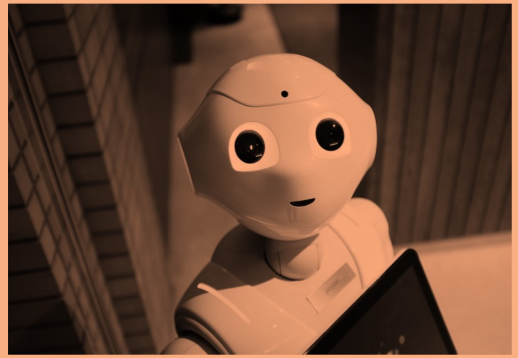


action: $x_1 \in \mathcal{X}_1 \subset \mathbb{R}^{d_1}$
goal: max $u_1(x_1, \dots, x_n)$



action: $x_2 \in \mathcal{X}_2 \subset \mathbb{R}^{d_2}$
goal: max $u_2(x_1, \dots, x_n)$

…



action: $x_n \in \mathcal{X}_n \subset \mathbb{R}^{d_n}$
goal: max $u_n(x_1, \dots, x_n)$

Nash Eq: *A collection of* $x_1^*, \dots x_n^*$ s.t. for all i, x_i: $u_i(x_i^*; x_{-i}^*) \geq u_i(x_i; x_{-i}^*)$

Randomized Nash Eq: *A collection of distributions* $p_1, \dots p_n$ s.t. for all i, x_i:

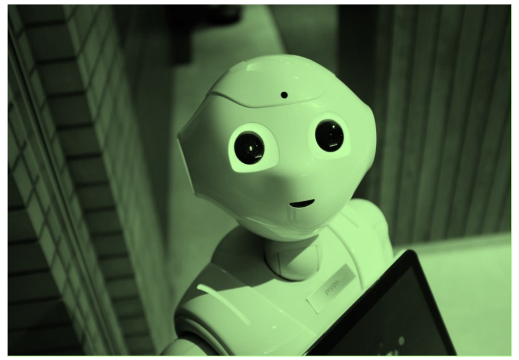
$$\mathrm{E}_{x^* \sim p_1 \times \dots \times p_n}[u_i(x_i^*; x_{-i}^*)] \geq \mathrm{E}_{x^* \sim p_1 \times \dots \times p_n}[u_i(x_i; x_{-i}^*)]$$

Coarse Correlated Eq: *A joint distribution of* p s.t. for all i, x_i:

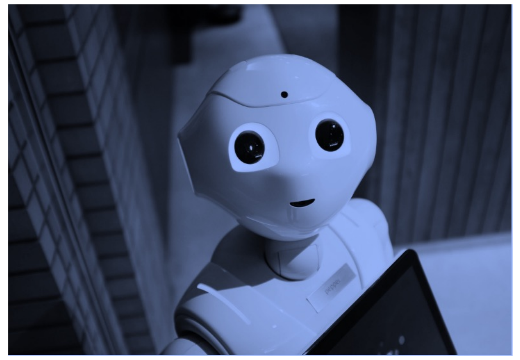
$$\mathrm{E}_{x^* \sim p}[u_i(x_i^*; x_{-i}^*)] \geq \mathrm{E}_{x^* \sim p}[u_i(x_i; x_{-i}^*)]$$

If some $u_i(x_i; x_{-i})$ is not concave in x_i for all x_{-i}, Nash equilibrium does not necessarily exist
If the $\mathcal{X}_i$'s are non-compact, even randomized Nash/correlated eq do not necessarily exist

Infinite/Non-Parametric Games

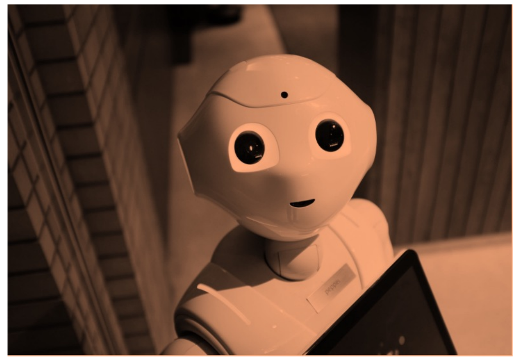


action: $x_1 \in \mathcal{X}_1$
goal: max $u_1(x_1, \ldots, x_n)$



action: $x_2 \in \mathcal{X}_2$
goal: max $u_2(x_1, \ldots, x_n)$

…



action: $x_n \in \mathcal{X}_n$
goal: max $u_n(x_1, \ldots, x_n)$

- Action sets $\mathcal{X}_i$: high-dimensional or infinite-dimensional/non-parametric
- Utilities u_i : arbitrary functions $u_i: \times_i \mathcal{X}_i \to \mathbb{R}$
- Questions I want to ask:

Under what conditions do there exist **global** Nash/Correlated/Coarse Correlated Equilibria?

Are there simple methods converging to equilibria in a finite number of steps?

- For Q1: I hope that the answer depends on some complexity measure of the u_i's that I can identify
- For Q2: by "simple" I want that each step can be executed efficiently

"Guess the larger number" Game

Player 2 (max player)

Player 1 (min player)

	1	2	3	4	...
1	1	1	1	1	...
2	-1	1	1	1	...
3	-1	-1	1	1	...
4	-1	-1	-1	1	...
...	...	...	...	...	...

A two-player zero-sum game where:

- $\mathcal{X}_1 = \mathcal{X}_2 = \mathbb{N}$
- $u_1(x_1, x_2) = -u_2(x_1, x_2) = 1_{x_1 \geq x_2} - 1_{x_1 < x_2}$
- (so table shows utility of Player 2)

Fact: "Guess the larger number" game has no Nash equilibrium (not even a very coarse approximate one).

Proof: Suppose (P, Q) is a pair of distributions over $\mathbb{N}$.

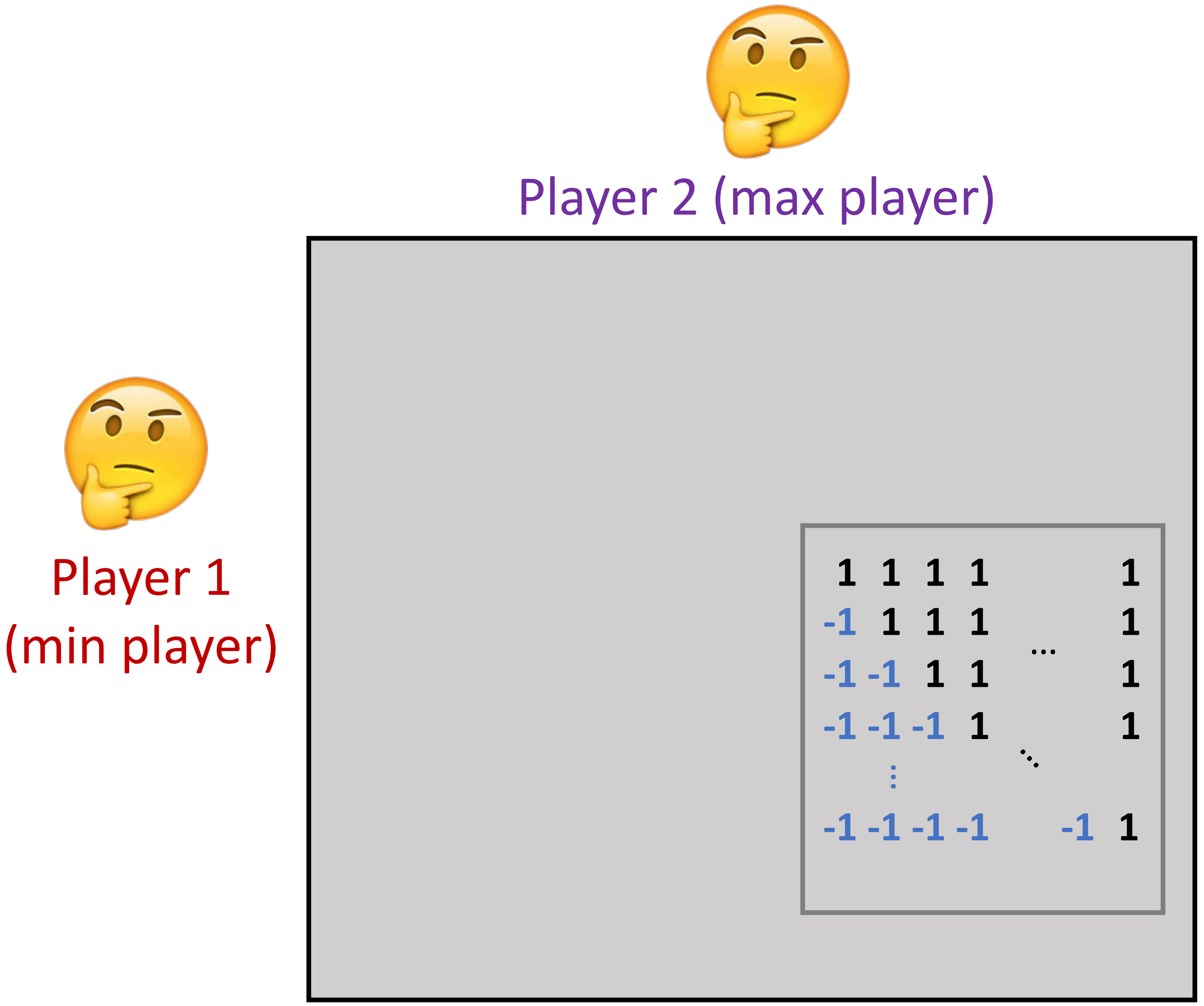
- Suppose WLOG that Player 2 has expected utility ≥ 0 under (P, Q).
- Can find $x \in \mathbb{N}$ such that x is greater than 0.999 fraction of the mass of Q.
- If min-player deviates to x her utility is > 0.99.

So "Guess the larger number game" is an obstacle to the existence of Nash equilibrium.

What if we exclude "Guess the larger number"?

- Surprising fact: "Guess the larger number" game is the only obstacle to the existence of Nash equilibrium in $\{-1,1\}$-valued two-player zero-sum games!

Theorem [Hanneke-Livni-Moran'21]: If an (infinite) $\{-1,1\}$-valued two-player zero-sum game has no subgame which is "Guess the larger number," then it has an ϵ-approximate Nash equilibrium for all $\epsilon > 0$.



G: $\{-1,1\}$-valued two-player zero-sum game

Threshold dimension of G: size of largest threshold sub-matrix

[Hanneke-Livni-Moran'21]: **Tr(G)** finite $\Rightarrow$ Minimax Eq exists

Claim: **Tr(G)** finite $\Leftrightarrow$ **Littlestone dimension** of G finite*

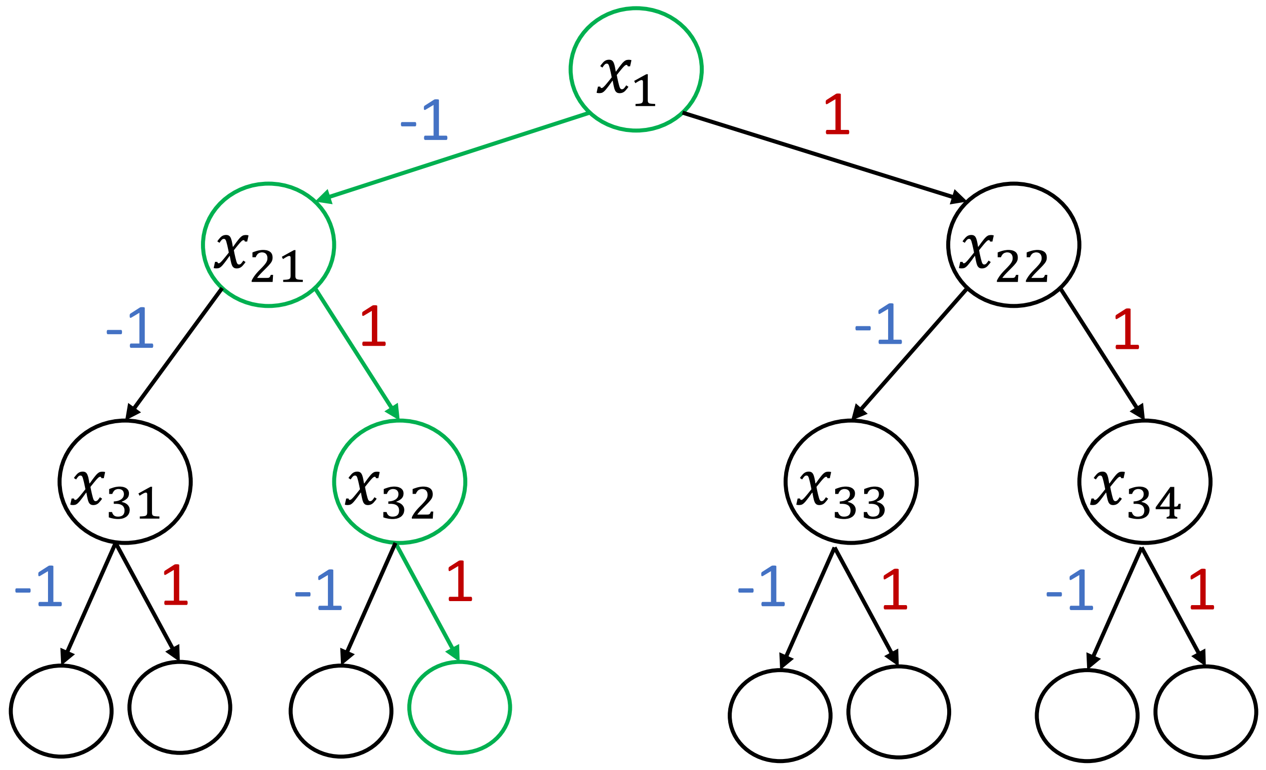
*: define Littlestone dimension of G in next slide

[Parenthesis: Littlestone dimension of a Concept Class

- H: binary classifiers over feature set $\mathcal{X}$
- TL;DR:
  - $\mathbf{Ldim}(H)$: characterizes whether and how well (in terms of regret) classifiers can be online learned from a sequence of adversarial data $(x_t, b_t) \in \mathcal{X} \times \{\pm 1\}$
  - [Analogously to how $\mathbf{VC}(H)$ dimension characterizes learnability of H given a batch of i.i.d. data]
- Detailed description:
  - Consider online learning setting where for $t = 1, \ldots, T$:
    - learner chooses distribution p_t over $h_t \in H$
    - adversary chooses $(x_t, b_t) \in \mathcal{X} \times \{\pm 1\}$ (with knowledge of learner's distribution)
    - learner samples $h_t \sim p_t$ and experiences loss $\ell(h_t(x_t), b_t)) = \frac{1 - h_t(x_t) \cdot b_t}{2}$ (i.e. 1 if prediction is wrong ow 0)
  - Learner's goal: minimize expected regret $\sum_t \ell(h_t(x_t), b_t)) - \min_h \sum_t \ell(h(x_t), b_t))$
  - Clearly can get expected regret $O\left(\sqrt{T \cdot \log |H|}\right)$ (by doing MWU over H)
  - But what if H is infinite?
  - **[Rakhlin-Sridharan-Tewari'15, Hanneke-Livni-Moran'21]:** can get expected regret $\tilde{O}\left(\sqrt{T \cdot \mathbf{Ldim}(H)}\right)$
    - $\mathbf{Ldim}(H)$ may be finite even when H is infinite; also $\mathbf{Ldim}(H) \leq \log |H|$ always

Littlestone dimension: formal definition

- H: binary classifiers over feature set $\mathcal{X}$
- Detailed definition of **Ldim**(H) considers trees, whose internal vertices are labeled by $\mathcal{X}$ and edges by +1 or -1



Defn: For a binary tree with all internal nodes labeled by elements of $\mathcal{X}$, edges labeled by $\{-1,1\}$:

- It is **shattered** by H if for each leaf ℓ there is some $h_\ell \in H$ which labels all nodes on the root-to-leaf path for ℓ according to the labels on the edges.
- E.g., for the green leaf:

  need $h_\ell(x_1) = -1, h_\ell(x_{21}) = 1, h_\ell(x_{32}) = 1.$

Defn: Littlestone dimension of hypothesis class H, denoted **Ldim**(H), is largest d so that there exists tree of depth d shattered by H.

Littlestone dimension of a Game

Littlestone dimension of a Concept Class

- H: binary classifiers over feature set $\mathcal{X}$
- TL;DR:
  - $\mathbf{Ldim}(H)$: characterizes whether and how well (in terms of regret) classifiers can be online learned from a sequence of adversarial data $(x_t, b_t) \in \mathcal{X} \times \{\pm 1\}$
  - [Analogously to how $\mathbf{VC}(H)$ dimension characterizes learnability of H given a batch of i.i.d. data]

  - **Claim:** can get expected regret $\tilde{O}\left(\sqrt{T \cdot \mathbf{Ldim}(H)}\right)$ (which may be finite even when *H* is infinite!)

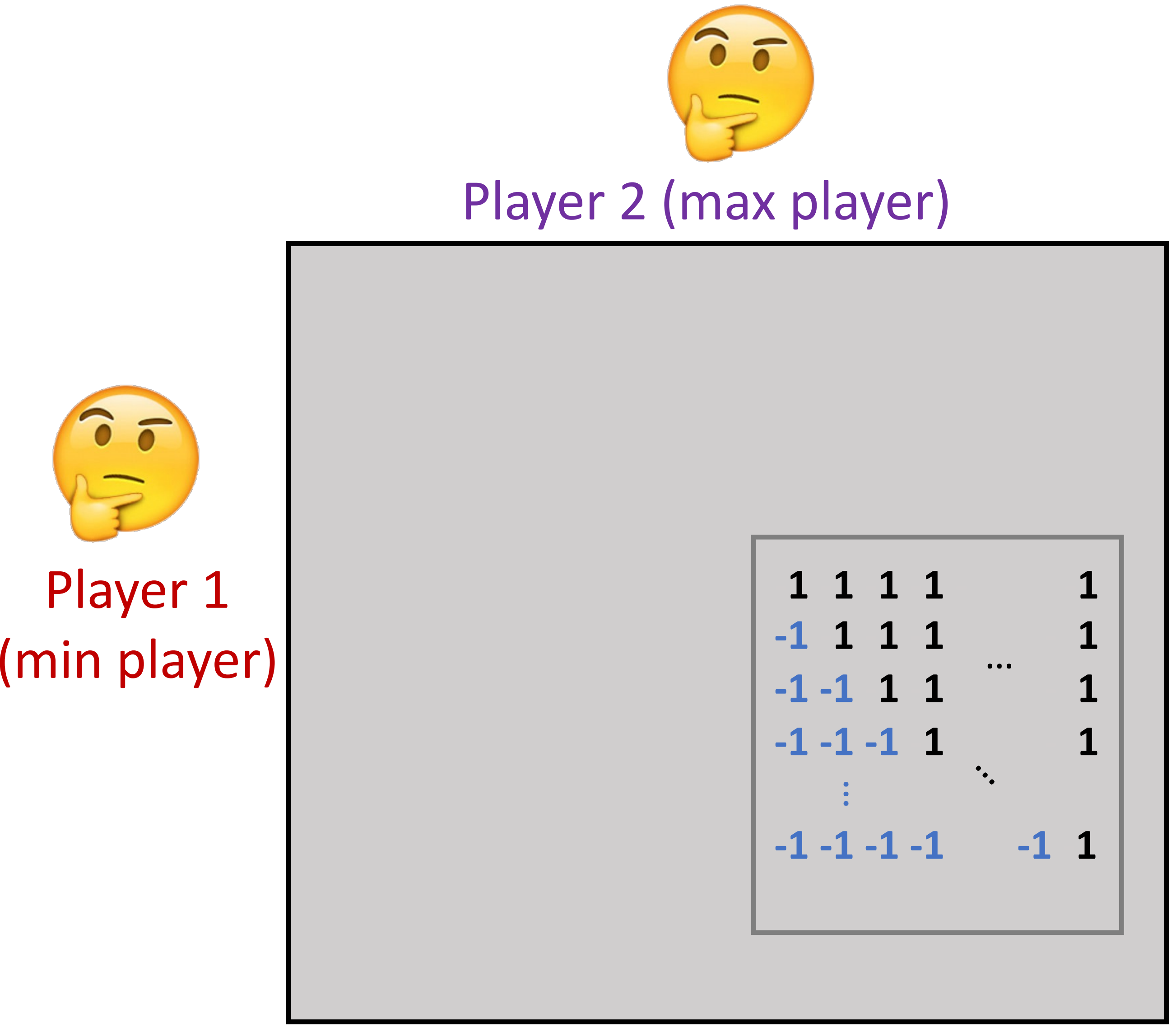
Littlestone dimension of a Game

- G: a multiplayer $\{\pm 1\}$-valued game with utilities $u_i: \mathcal{X}_1 \times \cdots \times \mathcal{X}_n \to \{\pm 1\}$
- For each player, consider the function class $H_i \coloneqq \{u_i(x_i, \cdot) \mid x_i \in \mathcal{X}_i\}$
  - H_i contains binary classifiers mapping each x_{-i} to ± 1
- **Littlestone dimension** of G is $\max_i \{\mathbf{Ldim}(H_i)\}$

What if we exclude "Guess the larger number"?

- Surprising fact: "Guess the larger number" game is the only obstacle to the existence of Nash equilibrium in $\{-1,1\}$-valued two-player zero-sum games!

Theorem [Hanneke-Livni-Moran'21]: If an (infinite) $\{-1,1\}$-valued two-player zero-sum game has no subgame which is "Guess the larger number," then it has an ϵ-approximate Nash equilibrium for all $\epsilon > 0$.



G: {−1,1}-valued two-player zero-sum game

Threshold dimension of G: size of largest threshold sub-matrix

[Hanneke-Livni-Moran'21]: **Tr(G)** finite $\Rightarrow$ Minimax Eq exists

Claim: **Tr(G)** finite $\Leftrightarrow$ **Littlestone dimension** of G finite

Littlestone dimension of G: $\max\{\mathbf{Ldim}(H_1), \mathbf{Ldim}(H_2)\}$

where $H_1 :=$ {rows of G viewed as binary classifiers over $\mathcal{X}_2$}
$H_2 :=$ {columns of G viewed as binary classifiers of $\mathcal{X}_1$}

$\mathbf{Ldim}(H)$: characterizes online learnability of H (from stream of examples)
(analogous to $\boldsymbol{VC}(H)$ which characterizes batch learning)

Suggests: perhaps equilibria can be found through learning…

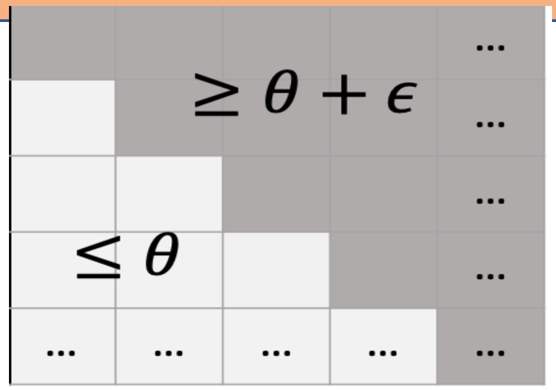
How about real-valued games?

- Surprising fact: "Guess the larger number" game is the only obstacle to the existence of Nash equilibrium in $\{-1,1\}$-valued two-player zero-sum games!

[Hanneke-Livni-Moran'21]: If an (infinite) $\{-1,1\}$-valued two-player zero-sum game has no subgame which is "Guess the larger number" (a.k.a. has finite **Tr**(G) $\Leftrightarrow$ finite **Lit**(G)) then it has an ϵ-approximate Nash eq for all $\epsilon > 0$.

[Daskalakis-Golowich'21] (Real-valued generalization of the above; informal):
If an (infinite) real-valued two-player zero-sum game has no subgame which is **ϵ-close to** some "scaling" of "Guess the larger number," then it has $O(\epsilon)$-approximate Nash equilibrium.

Formal result: requires finiteness of **ϵ-Fat Threshold** or **ϵ-sequential fat shattering dimension** (which are respectively generalizations of threshold dimension and Littlestone dimension to real-valued functions).

- **Def:** **ϵ-FatTr**(G) is the largest subgame satisfying 

for some θ.

- **Def:** ϵ**-seqFat**$(G) = \max_i \epsilon$**–seqFat**(H_i) where $H_i := \{u_i(x_i,\cdot) \mid x_i \in \mathcal{X}_i\}$

[Rakhlin-Sridharan-Tewari'15

- TL;DR: **ϵ–seqFat**(H) characterizes online learnability of concept class H; achievable regret: $O(\epsilon \cdot T) + \tilde{O}\left(\sqrt{T \cdot \epsilon\text{–seqFat}(H)}\right)$

Next Time: Equilibrium Learning?