Learning the Expected Core of Strictly Convex Stochastic Cooperative game

Nam Phuong Tran[1], The-Anh Ta[2], Shuqing Shi[3],

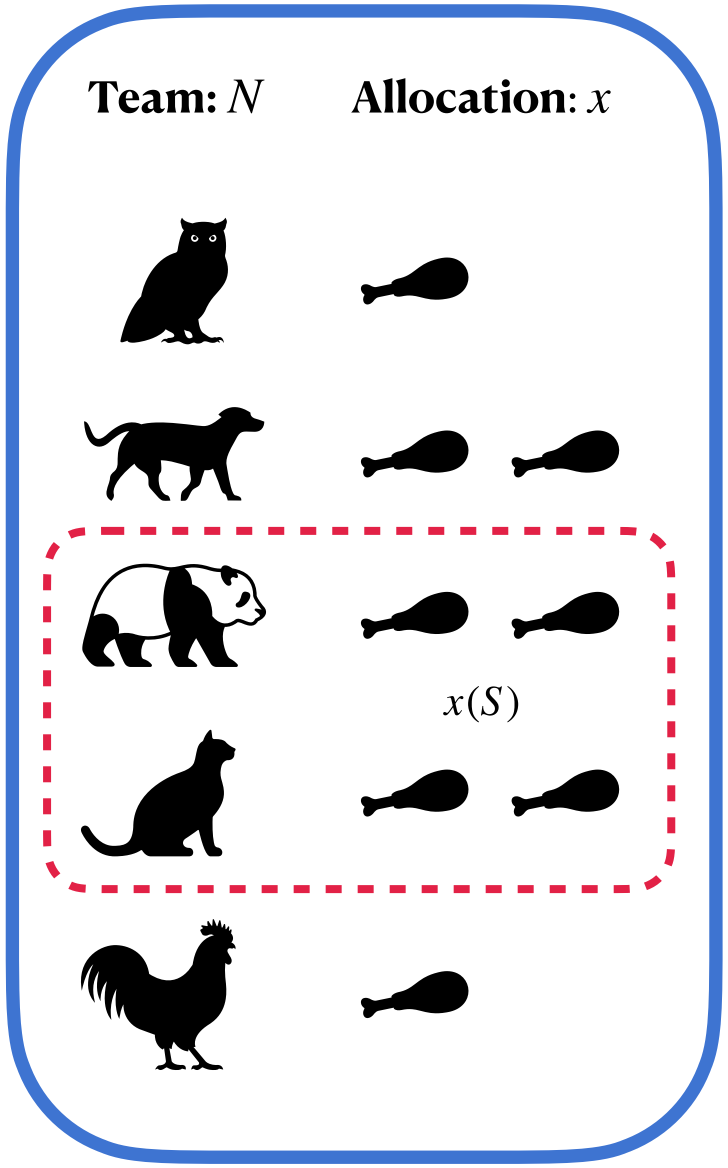
Debmalya Mandal[1], Yali Du[3], Long Tran-Thanh[1]

University of Warwick[1], Data 61 Australia[2], King's College London[3]

Reward Allocation Problem

Formulation:

- $N = \{1, \ldots, n\}$ set of players.
- Reward function $\mu : 2^N \to [0, 1]$
- Reward allocation: $x \in \mathbb{R}^n_+$,



Stable Allocation: Incentivise players to stay in N.

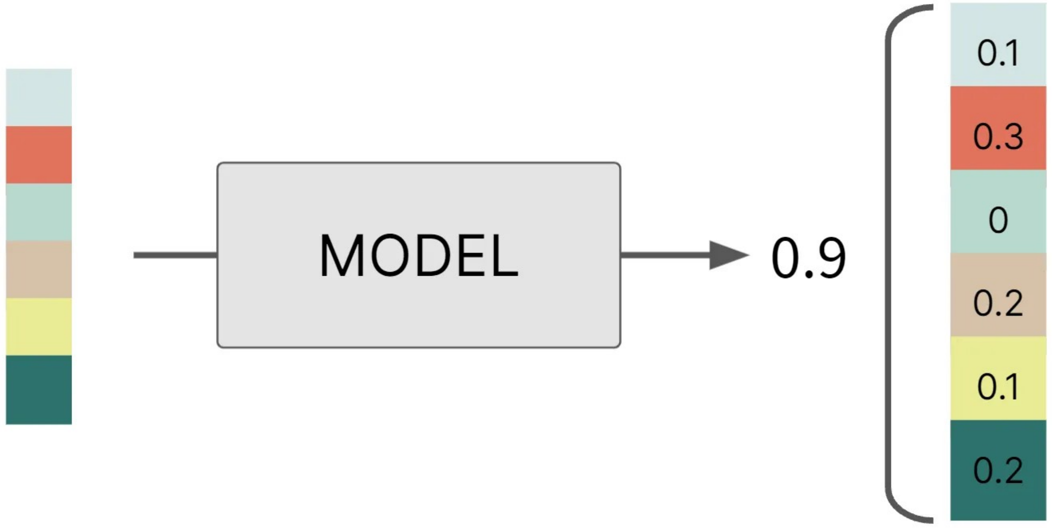
- $x(S) \geq \mu(S), \forall S \subseteq N$

Core: *Set of all stable allocations.*

Motivation

Credit assignment in CTDE

Credit assignment in model explainability



Research Questions

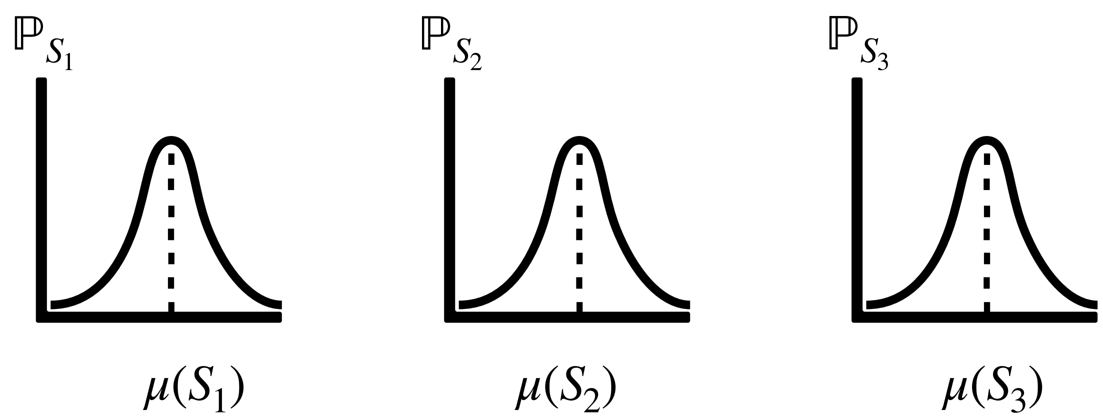
Classic Deterministic Setting:

- Compute core: require perfect knowledge of reward function μ.

Stochastic setting:

- The reward follows some unknown distributions.
- Learn the stable allocation through interacting with the environment.
- **Question**: How many samples (interactions) needed to output a *"stable" allocation*, with high probability.

Expected Core of Stochastic Game



Definition of Stochastic Game

- $G = (N, \mathbb{P}),\ \mathbb{P} = \{\mathbb{P}_S \text{ is a distribution on } [0,1], S \subset N\}$.
- Expected reward: $\mu(S) = \mathbb{E}_{r \sim \mathbb{P}_S}[r]$.
- $\mathbb{P}$ is unknown to the players. Need to learn through data.
- **Expected Core (E-Core)**: Set of all stable allocation in expectation
    - $x(S) \geq \mu(S), \forall S \subseteq N$

Interaction Protocol

There is a principal: game proceed in rounds.

Round t:

- Principal queries a coalition $S_t \subset N$
- Environment return $r_t \sim \mathbb{P}_{S_t}$.

Question : Are there some classes of the game where principal can learn expected stable allocation with probability at least $(1 - \delta)$ by poly$(n, \log(\delta^{-1}))$.

Impossibility result of learning Expected Core

In general games: may not be possible (**for any finite samples**).

Theorem 1: Suppose the E-Core is not full-dimensional, then with finite samples, no algorithm can output a expected stable allocation with probability at least 0.8.

$\varsigma-$**Strictly convex game (in expectation):**

- Expected reward function is strictly supermodular:

$$\mu(S \cup \{i\}) - \mu(S) \geq \mu(C \cup \{i\}) - \mu(C) + \varsigma, \quad \forall C \subset S$$

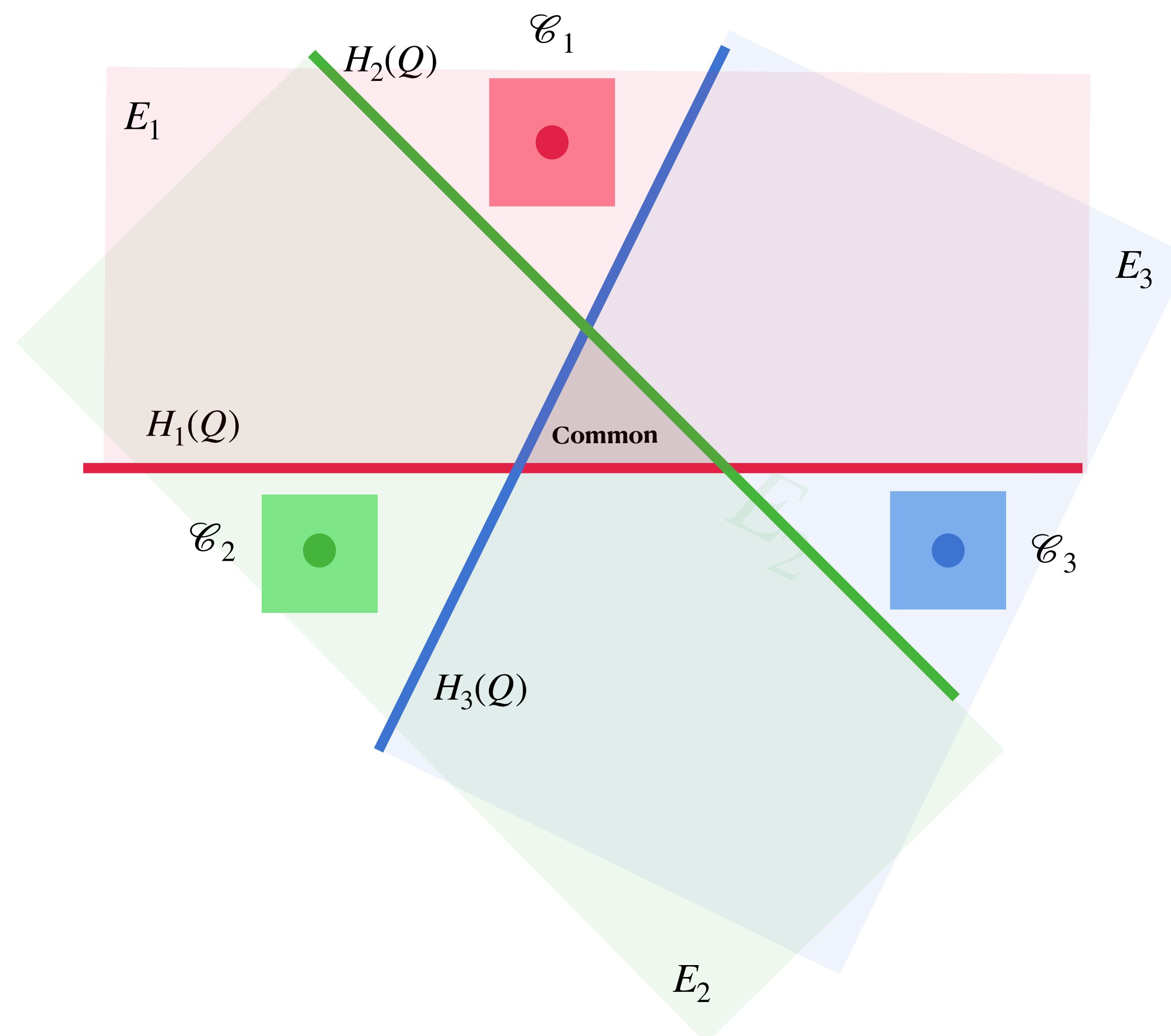
Strict convexity guarantees that the core is full-dimensional, $(n-1)$.

Sample Complexity

Theorem 2: For ς-strictly convex game, with probability at least, after

$$T \cong n^{15}\varsigma^{-4}\log(\delta^{-1}),$$

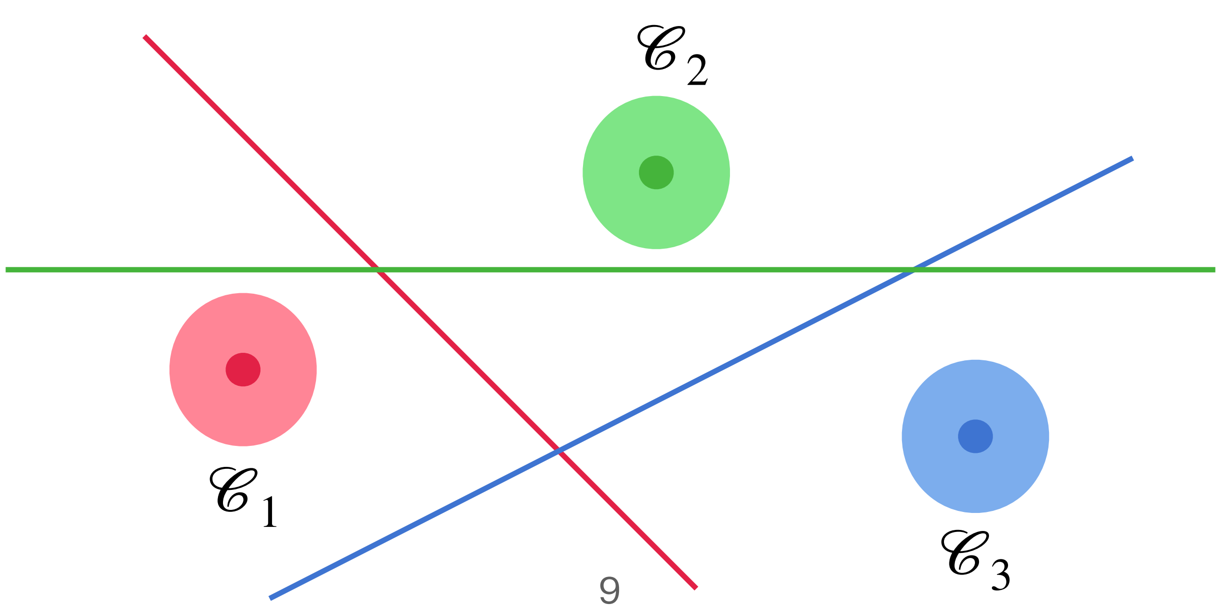
the algorithm stops and returns a stable allocation.



Separation Hyperplane for multiple convex set

Sufficient Stopping Condition: Using Separation hyperplanes

Theorem 3(Separation Hyperplane for multiple convex set): Assume that $\{\mathscr{C}_p\}_{p\in[n]}$ is are mutually disjoint compact and convex subsets in (n-1)-Euclidean space. Suppose that there does not exist a (n-2)-dimensional hyperplane that intersects with the interior of confidence sets. Then for each $p \in [n]$, there exist a hyperplane separating $\mathscr{C}_p$ from $\bigcup_{q\neq p} \mathscr{C}_q$.



Conclusion

- We propose an algorithm that can learn a stable allocation with polynomial number of n for a class of strictly convex game.
- Analysis involves proving new result from convex geometry, e.g., *extension of hyperplane separation theorem.*