Cooperative AI

Introduction to the problem
Some key concepts

Ignacio Ojea Quintana

Plan for the lecture:

1. **Introduce and partition the scope of the problem.**
2. **Narrow our focus to 2x2 games between robots.**
3. **Review of Nash Equilibria**
4. **Classification of games between Cooperative, Mixed Interest, and Conflict.**
5. **Evolutionary Stable Equilibria definition and examples.**
6. **Replicator Dynamics, examples and extensions.**
7. **(if time) Correlated Equilibria definition and comments.**

Introduction

1. **The problem of cooperation is interdisciplinary and has been extensively studied in fields like game theory, anthropology, evolutionary biology, sociology, etc.**
   a. **We humans are a cooperative species.**
2. **Only recently cooperation has been studied by the AI community systematically:**
   a. **https://www.cooperativeai.com/ (NeurIPS 2020)**
   b. **Open Problems in Cooperative AI (Dec 2020)**
3. **First challenge: To be able to talk meaningfully about cooperation to people with different intellectual background.**

Partition Space of The Problem: Two dimensions

The *Agents* who are (or not) cooperating:

- **Humans**
- **Robots**
- **Institutions / Organizations**

The third versus first person *Perspective*:

- **The third person perspective is that of a central planner that either designs a cooperative environment (game or market) or intervenes as a third party in a given one.**
- **The first person perspective of an individual agent deciding whether to cooperate or not.**

Examples of problems across dimensions:

- **Should you cooperate in a game with other people (or robots, or institutions).**
- **How do we facilitate Human-Human cooperation (e.g. a translation algorithm).**
- **How do make sure that the AI is doing what we want it to do (Alignment and Safety).**
- **How do we develop institutions that foster Human-AI cooperation.**
- **Etc., etc., etc.**

Games and Nash Equilibria

A *normal form game* G has three components:

1. A set of players $i=1,2,3,\ldots,n$
2. A strategy or action set $S_1, S_2, \ldots, S_n$ for each player. [$S = S_1 x S_2 x \ldots x S_n$].
3. A *payoff* or *utility function* for each player *i*, $u_i{:}S \rightarrow R$.

Also, a *mixed strategy* σ_i for player *i* is a distribution over S_i. Utility functions u_i generalize to mixed strategies making use of expectations; EU_i.

A strategy profile $(\sigma_1, \sigma_2, \ldots, \sigma_n)$ is a **Nash Equilibrium** of G if for every player *i*, and every s_i in S_i:

$$EU_i(\sigma_i, \sigma_{-i}) \geq EU_i(s_i, \sigma_{-i})$$

Two Intuitions around Equilibria

The Expectational Intuition

Strategy profiles ($\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_n$) have two components for each player i:

A. A (first-order) suggestion on what they should do σ_i, and
B. A (second-order) expectation of what others will be doing σ_{-i}.

Solutions to games are strategy profiles with an equilibrium-like property:

- A and B should be compatible with maximizing expected utility.
- In other words, if **conditional on learning B players would rather not deviate from A.**

The Resilience Intuition

Strategy profiles ($\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_n$) are resilient with respect to some kind of perturbation.

In NE, the perturbation is player-wise strategy change:

Unilateral deviation is not profitable

2v2 Games

Prisoners Dilemma

	Cooperate	Defect
Cooperate	3,3	1,4
Defect	4,1	2,2

Perfect Cooperation Game

	Cooperate	Defect
Cooperate	4,4	-1,0
Defect	0,-1	-1,-1

Mixed Motive Game (Stag Hunt)

	Cooperate	Defect
Cooperate	4,4	1,3
Defect	3,1	2,2

The Importance of Social Dilemmas: They are Ubiquitous

Public Goods Game (generalization of PD)
Agents have degrees of cooperation and defections.
Each agent has a budget and then can choose to invest as much as they want from that budget into the public pool, then the pool is increased by a certain factor (>1 and less than the number of players), and equally distributed among players.
Below, a two person three degrees $2 budget and 1.5 factor example. Notice: If we were to eliminate one of the actions for both players, we get a Prisoner's Dilemma.

	Pool 2	Pool 1	Pool 0
Pool 2	3, 3	2.25,3.25	1.5,3.5
Pool 1	3.25,2.25	2.5,2.5	1.75,2.75
Pool 0	3.5,1.5	2.75,1.75	2,2

What mixed strategy NE?

Zero Sum / Conflict (Matching Pennies)

	Cooperate	Defect
Cooperate	1,-1	-1,1
Defect	-1,1	1,-1

John Nash proved the existence of equilibria, provided that we allow for mixed strategies.

First notice that in Matching Pennies there are no pure strategy NE. Given any pair of pure strategies, one of the players would change their action.

Nevertheless, there is a mixed strategy in NE in which both players evenly mix their actions, and they each get an expected payoff of zero.

How to Compute (mixed) NE of a Mixed Motive Game

	Column	p	(1-p)
Row		Cooperate	Defect
q	Cooperate	4,4	1,3
(1-q)	Defect	3,1	2,2

Suppose we have a NE in which Row cooperates with probability *q*, and defects with probability (*1-q*); and Column cooperates with probability *p* and defects with probability *(1-p)*. Furthermore, ***assume*** *q≠0 ,1* and *p≠0,1*.

- Important: If *q* is in a NE, Row will select *q* in a way that makes Column indifferent between playing Cooperate and Defect. If Column is not indifferent, then they could improve by taking a deterministic action (Cooperate or Defect), against our assumption that it is a NE with *p≠0 ,1*.

$EU_C(\text{Cooperate}) = 4*q + 1*(1-q) = 3q+1 = EU_C(\text{Defect}) = 3*q + 2*(1-q)=q+2 \rightarrow q=1/2$
$EU_R(\text{Cooperate}) = 4*p + 1*(1-p) = 3p+1 = EU_R(\text{Defect}) = 3*p + 2*(1-p)=p+2 \rightarrow p=1/2$

2v2 Games Classification

Social Dilemma:

1. There is an outcome that is pareto (and/or welfare) dominant over all of the NE in the game.

Common Interest Game:

1. There is an outcome that is pareto (and/or welfare) dominant over all of the possible outcomes.
2. All of the NE of the game secure the dominant outcome.

Mixed Motives Game:

1. There are pareto (and/or welfare) outcomes dominant over all of the outcomes in the game.
2. Some of the NE in the game secore that optimal payoff, but others do not.

Zero Sum Game

1. The game is zero sum (hence no pareto outcome).

2v2 Games

Prisoners Dilemma

	Cooperate	Defect
Cooperate	3,3	1,4
Defect	4,1	2,2

Perfect Cooperation Game

	Cooperate	Defect
Cooperate	4,4	-1,0
Defect	0,-1	-1,-1

Mixed Motive Game (Stag Hunt)

	Cooperate	Defect
Cooperate	4,4	1,3
Defect	3,1	2,2

Zero Sum / Conflict (Matching Pennies)

	Cooperate	Defect
Cooperate	1,-1	-1,1
Defect	-1,1	1,-1

Equilibria

A strategy profile $(\sigma_1, \ldots, \sigma_n)$ is a **Nash Equilibrium** of G if for every player *i*, and every s_i in S_i:

$$EU_i(\sigma_{i,} \sigma_{-i}) \geq EU_i(s_{i,} \sigma_{-i})$$

A strategy profile $(\sigma_1, \ldots, \sigma_n)$ is an **Evolutionary Stable Nash Equilibrium** of G if for every player *i*, and every s_i in S_i :and s_{-i} in S_{-i}:

- $EU_i(\sigma_{i,} \sigma_{-i}) > EU_i(s_{i,} \sigma_{-i})$, or
- $EU_i(\sigma_{i,} \sigma_{-i}) = EU_i(s_{i,} \sigma_{-i})$ and

  $EU_i(\sigma_{i,} s_{-i}) > EU_i(s_{i,} s_{-i})$

Where they come apart

Sample Game with NE that are not ESS

	Cooperate	Defect
Cooperate	2,2	1,2
Defect	2,1	2,2

Here (C,C) and (D,D) are NE.
However (D,D) is ESS while (C,C) is not.
This is because:

- $EU_R(C,C) = EU_R(D,C)$, but
- $EU_R(C,D) \not> EU_R(D,D)$

Replicator Dynamics

Evolutionary Dynamics:

1. Initial State: We assume the environment can only hold a fixed population of agents, some of them with the cooperator gene, and some with the defector gene. Assume the same proportion of both types.
2. Dynamics:
   a. Each agent is randomly paired with another agent (so that pairing is proportional to the current population), they play the 2v2 game and each is given a fitness reward.
   b. Compute the average fitness across all agents, and for each type of agent (C or D).
   c. Update the population of C and D in proportion to how distant they were with respect to the mean fitness.
3. Run the process until it stabilizes.

Pseudo Code (Git):

Init:

xA0, xB0 = initial proportions of cooperators and defectors

Compute fitness:

fA1 = $\sum_j$(prob of interacting with j)*EU_C(C,j)
fB1 = $\sum_i$(prob of interacting with j)*EU_D(i,D)
ft = xA1 * fA1 + (1 - xA1) * fB1

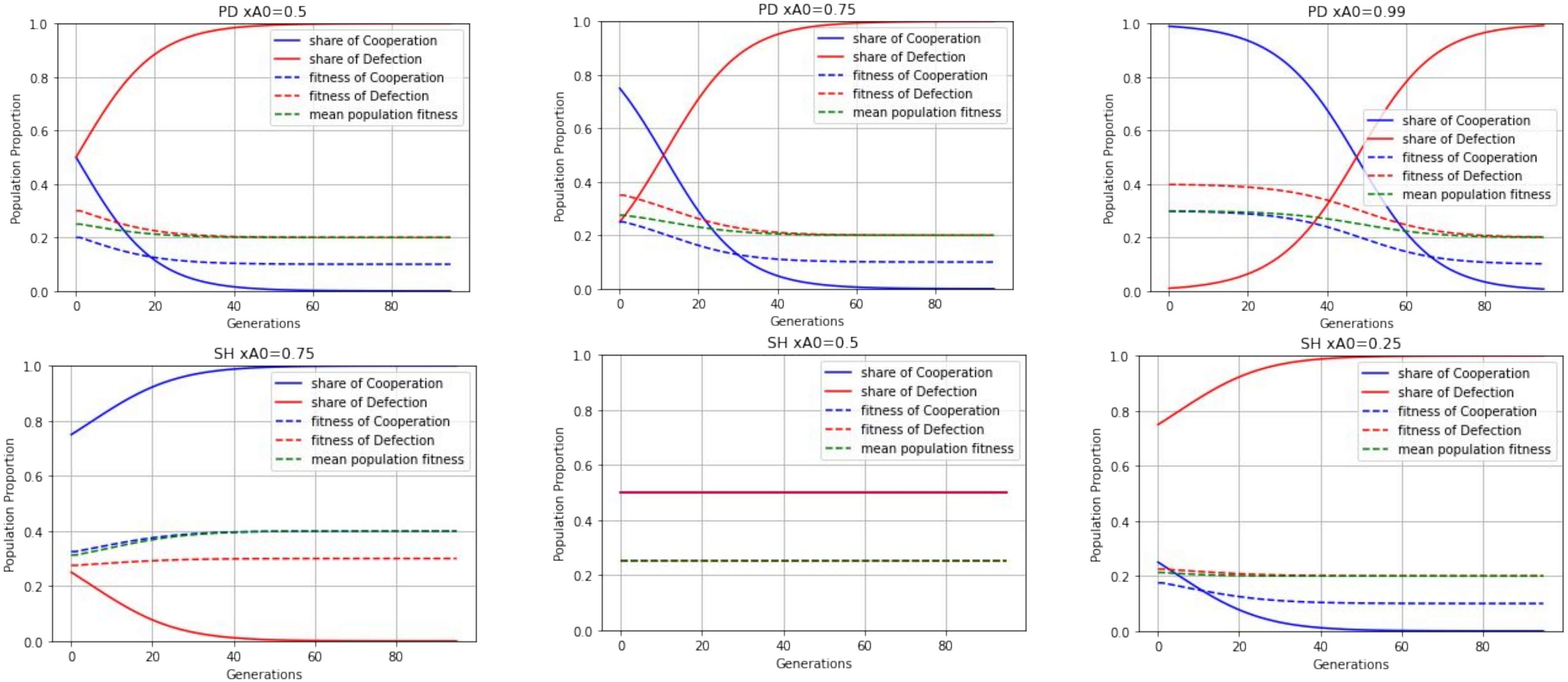
Dynamic Equation:

xA1 = xA0 + (xA0 * (fA0- ft)) * dt)

xB1 =xB0 + (xB0 * (fB0 - ft)) * dt)

At each time step, populations increase or decrease in proportion to their distance to the mean fitness.

Some Simulations: PD and SH



Differential Equations and Basins of Attraction

State of the system:

s = a vector with the population proportion for each type

Fitness:

fA(s) = $\sum_j$(prob of interacting with j)*EU_C(C,j) = $\sum_j s_j$*EU_C(C,j)
fB(s) = $\sum_i$(prob of interacting with j)*EU_C(i,D) = $\sum_i s_i$*EU_C(i,D)

Average Fitness:

F = s_A*fA + s_B*fB

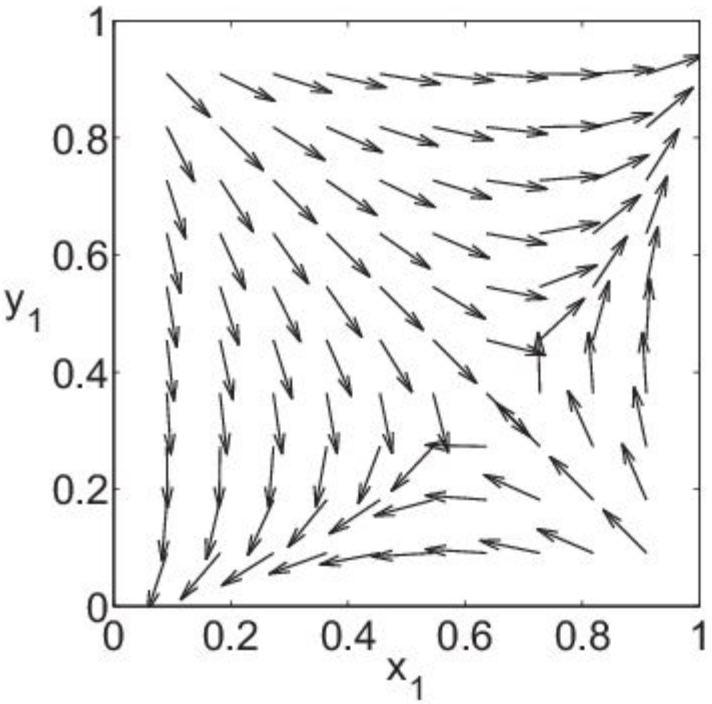
Dynamic Equations:

dA/dt = xA(fA(s)-F))
dB/dt = xB(fB(s)-F))

dA = (xA(fA(s)-F)))*dt
dB = (xB(fB(s)-F)))*dt

Solve:

dA = 0
dB = 0



Uses of the Replicator Dynamics

You can apply the replicator dynamics to a more complex set of agents:

- Agents with conditional strategies on type encountered.
- Agents with non-neutral risk attitudes.
- More complex and iterated games.

You can slightly modify the dynamics:

- The replicator-mutator dynamics introduces some perturbations in the population, so that an epsilon of the total agents adopt a random strategy at each time step.
- The replicator-imitator dynamics allows individual agents to imitate the best performers.

You can make the interactions more complex:

- Replicator Dynamics can be done on networks representing topological spaces.

The central research question associated with these techniques is:

How can we explain the evolution of cooperation?

Famously, Robert Axelrod did tournaments involving people submitting strategies to an iterated Prisoners Dilemma.

Discovery: TIT-FOR-TAT did pretty well.

Summary and Final Remarks

Today we covered:

1. Big overview of the subject of Cooperative AI.
2. Review of Nash Equilibria
3. Classification of games 2v2 games.
4. Evolutionary Stable Equilibria definition and examples.
5. Replicator Dynamics, examples and extensions
6. (if time) Correlated Equilibria definition and examples.

Looking ahead on Cooperative AI as a research field.

Pros: Lots of interesting avenues

- Trust
- Communication and signaling
- Game theoretic equilibria
- Commitments
- Mechanism design
- The evolution of social behavior
- Nudging

Cons: Old problem

- There is a lot of literature on it.
- There is a lot of *diverse* literature on it, from different fields.
- The concept might be too plastic.

Correlated Equilibria

A strategy profile ($\sigma_1, \ldots, \sigma_n$) is a **Nash Equilibrium** of G if for every player i, and every s_i in S_i:

$$EU_i(\sigma_i, \sigma_{-i}) \geq EU_i(s_i, \sigma_{-i})$$

Scenario:

- Six friends with diverse preferences, each friend with a distinct preferred restaurant.
- Second best reward if if they go together to another restaurant,
- Third reward if they go to their favorite restaurant but not together

Solution: Use a (shared) fair die as a correlating mechanism, and publicly commit to go to the restaurant associated the outcome of the die.

- After observing the outcome of the die, and assuming others will fulfill their commitments, they each expect with certainty all others to go to its associated restaurant, and on that expectation they would rather comply with their own commitment. This is a correlated equilibrium.

Correlated Equilibria

<u>Correlating Device (S,{Hi},p):</u>

- ***S*** set of states or possible outcomes
- ***{H_i}*** a partition of S for each agent i
- ***p*** a common prior on S for all of the agents

<u>Correlated Strategy: fi</u>

- $f_i: S \rightarrow S_i$
- f_i is measurable with respect to ***{H_i}*** (i.e. it assigns the same action to elements of the same partition component)

A profile of correlated strategies $(f_1,\ldots,f_n)$ is a Correlated Equilibrium iff for every i and correlated strategy g_i:

$$\Sigma_{s \text{ in } S}\, EU_i(f_i(s),f_{-i}(s))*p(s) \geq \Sigma_{s \text{ in } S}\, EU_i(g_i(s),f_{-i}(s))*p(s)$$

Good properties of Correlated Equilibria

- Under the assumption of common priors, if all players are Bayes rational (i.e. they maximize expected utility conditionalizing on the information), the resulting strategy tuple would be a correlated equilibrium (Aumann's 1987) (yes, the converse is also true).

- They are computationally simpler than NE. For example, for many scenarios you can find them in polynomial time (Aumann 1987, Papadimitriou & Roughgarden 2008).

- Evolutionary Stable Strategies, as well as stability under some modeling of rational deliberation, are all correlated equilibria (but the converse does not hold).

- There are well studied relations between \textit{conventions} and correlated equilibria (Vandeerschraaf 1995, Skyrms 2014). Relatedly, there are connections with the notion of *salience*.

- Social Norms can also be conceptualized using correlated equilibria (Morsky & Akcay 2019).

Thanks!