Last Time:

- [ ] collaboration / assistance / coordination

Lecture 11

HRI, FALL '25

Andrea Bajcsy

This Time:

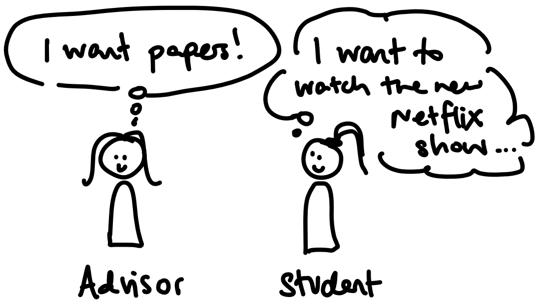
- [ ] game theory

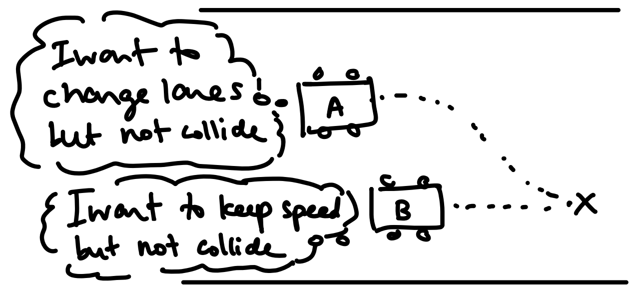
Resources:

- Tamer Başar and Geert Jan Olsder. "Dynamic noncooperative game theory." SIAM, 1998
- Rufus Isaacs. "Differential Games I: Introduction." RAND Corporation, 1954
- David Fridovich-Keil. "Smooth Game Theory". 2024.

Why game theory?

A mathematical language for describing "coupled" decision-making problems. ex.





⇒ preferences conflict! so each actor will choose strategy which in some way accounts for the other

⇒ not full of conflict! e.x. both cars want to cooperate on the collision part...

Time is also key in games. For example, in chess, each player knows they will get to move in the future.

Dynamic games of this type are particularly interesting in AI / Robotics / HRI.

Terminology

- agents ("multi-agent RL") (or players ("game theory")) – who is playing the game.
- policies (or strategies) – how the players choose their decisions

One may categorize games across several key axes

- finite / infinite – game is "finite" if players have only finitely many actions to pick from (e.g. $A = \{\leftarrow, \rightarrow, \uparrow, \downarrow\}$) "Infinite" if the actions available form a continuum (e.g. $A \subseteq \mathbb{R}^n$)

- static / dynamic / differential –
  - ↳ "static" if played @ a single instance in time (e.g. 1 round of rock paper scissors)

↳ "dynamic" if play continues over a period of time (ex. chess)

↳ "differential" if it is played in continuous time

(ex. you are optimizing player A & B's trajectories which are solutions to ODEs : $\dot{x}_A = f_A(x_A, a_A)$ & $\dot{x}_B = f_B(x_B, a_B)$)

- <u>zero / general sum</u> –
  - ↳ "zero-sum" are games in which players' objectives add to zero! (i.e. your win is my loss). These model perfectly adversarial problems
  - ↳ "general sum" are games with arbitrary player objectives
- <u>unconstrained / constrained</u> – some games have additional constraints on player's actions
- <u>pure / mixed strategies</u> –
  - ↳ "pure" strategies are deterministic.
  - ↳ "mixed" strategy is stochastic
    → in rock - paper - scissors, mixed strategies are optimal!

The most relevant type of game we will see in this class is:

"infinite, dynamic, general-sum, constrained" games w/ "pure strategies"

HOW DO WE FORMULATE GAMES ?

Let's start with simple static, finite, pure strategy games

<u>Ex.</u> Prisoners' Dillemma

2 prisoners suspected of crime. Prisoners <u>are</u> guilty, but police need confession b/c they don't have enough evidence. Police tell each prisoner they have 2 options: {confess (C), quiet (Q)}

Since each player has these options, there are 4 outcomes:

CC: Both confess → both given 2 yr. sentence

QQ: Both quiet → both given 1 yr. sentence

CQ/QC: The one that confesses gets 0 jail time but the other gets 3 yr. sentence.

Convert this into table of outcomes

Prisoner 1's sentence:

P1 \ P2	C	Q
C	2	0
Q	3	1

$$M_1 = \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix}$$

Prisoner 2's sentence:

P1 \ P2	C	Q
C	2	3
Q	0	1

$$M_2 = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}$$

Imagine P1 decides to confess (C) and P2 decides to stay quiet (Q)

We can encode these decisions (or "actions") in vector form:

$$x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \begin{matrix} \{0,1\} \text{ if confess} \\ \{0,1\} \text{ if quiet} \end{matrix} \qquad x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

x_i this is decision of i^{th} player

The game from any $i \in \{1, 2\}^{th}$ players' perspective is optimizing:

$$J_i(x_1, x_2) = x_1^T M_i x_2$$

POV of player i

plug in each decision

evaluate "cost" with that player's cost matrix

Each player is choosing $x_1, x_2 \in \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ action space $= \mathcal{A}$

Can we compute the best strategy?

If P1 does not know what P2 will do, how can we obtain minimal cost?

Well, P1 can play a <u>security strategy</u>: → idea: minimize P1's cost even if P2 plays adversarially.

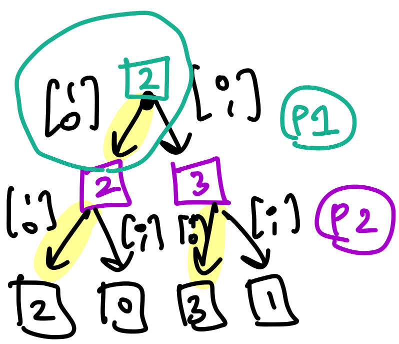
↓ BUT, not zero-sum b/c one players gain isn't others equal loss (i.e. total payoff = 0)

P1: $$\min_{x_1} \left(\max_{x_2} x_1^T M_1 x_2 \right)$$ ← from P1's POV.

Q what order of play is encoded above? i.e. what information does each player have access to when choosing strategy?

A Read from left to right. Player 1 first selects a strategy x_1, then Player 2 chooses response x_2 which maximizes $x_1^T M_1 x_2$ given knowledge of x_1!

ex. (cont.)



$$\min_{x_1} \left(\max_{x_2} x_1^T M_1 x_2 \right)$$ ← from P1's POV.

↓ if we did $\arg\max_{x_1} (\ldots\ldots) = \begin{bmatrix}1\\0\end{bmatrix}$

So to actually <u>solve games</u> we need to understand that the solution — called an <u>equilibrium</u> — can broadly take two forms for general-sum games:

- <u>Nash Equilibrium</u> (NE) of a 2 player game is a pair of strategies (x_1^*, x_2^*) s.t.

$$J_i(x_i^*, x_{\neg i}^*) \leq J_i(x_i, x_{\neg i}^*) \quad \forall x_i \in X_i$$

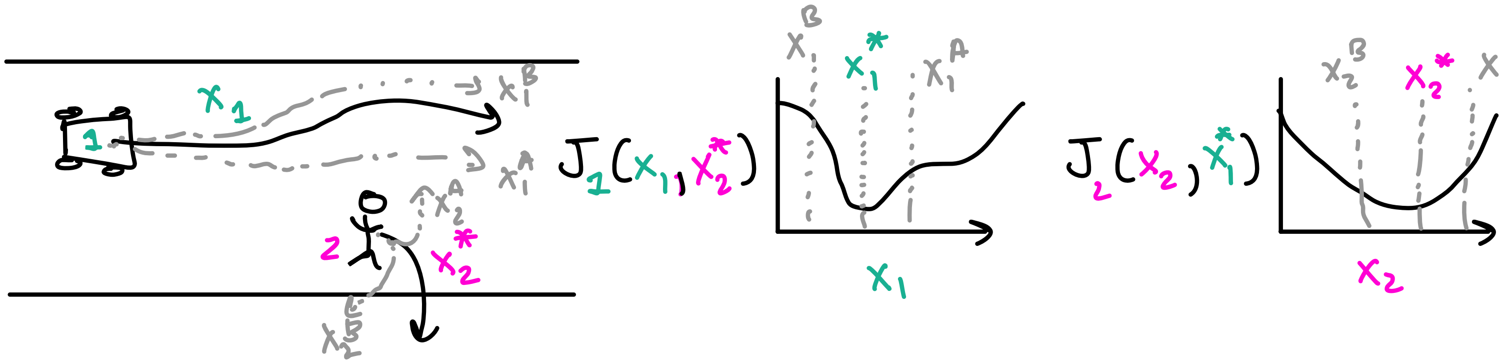
Intuition: no player has unilateral incentive to deviate!

So to actually solve games we need to understand that the solution — called an equilibrium — can broadly take two forms for general-sum games:

- Nash Equilibrium (NE) of a 2 player game is a pair of strategies (x_1^*, x_2^*) s.t.

$$\underbrace{J_i(x_i^*, x_{\neg i}^*)}_{\text{for player } i\text{, the cost of } x_i^* \text{ (when holding other player constant) is LOWER or equal}} \leq J_i(x_i, x_{\neg i}^*) \quad \forall x_i \in X_i$$

... to any alternative $x_i \in X_i$.



Intuition: no player has a unilateral incentive to deviate!

ex (continued, Prisoners' dilemma)

$$M_1 = \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix} \qquad M_2 = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}$$

The pure strategies $x_1^* = x_2^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ form the unique Nash equilibrium of the game ⇒ BOTH prisoners confess!

⊛ NOTE: in this case, it happens that x_1^* and x_2^* are also security strategies!

- <u>Stackelberg Equilibrium</u> (SE) which can be seen as the general-sum extension of security strategies.

⇒ Stackelberg equilibria occur when one player must commit to a strategy that the other views before deciding its strategy. The player who pre-commits is called the "leader" and the other is "follower"

(P1, <u>leader</u>) $$x_1^* = \arg\min_{x_1} J_1(x_1, x_2^*(x_1))$$

(P2, <u>follower</u>) $$\text{s.t. } x_2^*(x_1) = \arg\min_{x_2} J_2(x_1, x_2)$$

! <u>NOTE</u>: Nash ≠ Stackelberg, in general!

<u>How can we algorithmically solve games?</u>

$\boxed{\text{ALGO 0}}$ ITERATED BEST RESPONSE (IBR)

↳ this will help us compute Nash Equilibrium

<u>INPUT</u>: initial strategies (x_1, x_2) → $\{x_i\}_{i=1}^{N}$

while not converged do:

  for $i = 1, 2 \dots N$

    $$x_i \leftarrow x_i^*(x_{\neg i}) = \underbrace{\arg\min_{x_i} J_i(x_i, x_{\neg i})}_{\text{using equilibrium conditions as update rule}} \quad // \; P_i \text{ best response}$$

player i — all other players

return converged (x_1^*, x_2^*)

⊕ easy to implement w/ standard optimization tools

⊕ if it converges, then it finds NE

⊖ no convergence guarantees (cycling!!)
⊖ slow convergence ⇒ many expensive optimization steps in the inner loop.

can we do better?

→ outside the scope of this class, but there are modern solvers (e.g. PATH*) which can find solutions well + fast!

ex. in Peters et. al. "Contingency Games", R-AL 2024. for 5 player, 25 t-step, 3,208 decision variables

⇒ get solution in $\boxed{35\ ms}$

→ other algos: Monte Carlo Tree Search (MCTS)

iterative Linear Quadratic Games (Fridovich-Keil, 2019)

Why is game theory helpful in HRI?

1) joint prediction + planning!

→ recall how before we would call traj. forecasting model to give us $P(x_H^{0:T} \mid x_H^0)$ and then robot would use predictions to plan ⇒ but this doesn't account for influence of R's actions on H!

2) some settings, games can well-describe interaction phenomena (e.g. economics).