JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

Game Theory Contd.

Musad Haque
18-Oct-2023

Agenda

- ~~2x2 Games~~
- ~~Nash Equilibrium~~
- ~~Elements of a Game~~
- N-Player Games
- Pageant Game
- Learning Algorithms

2x2 Games

Prisoner's Dilemma

		P2 C	P2 D
P1	C	2,2	0,3
P1	D	3,0	1,1

2x2 Games

Prisoner's Dilemma

		P2	
		C	D
P1	C	2,2	0,3
	D	3,0	1,1

Nash equilibrium (D,D)

No unilateral incentive to deviate

2x2 Games

Technology Game

		P2 A	P2 B
P1	A	4,4	0,0
	B	0,0	3,3

2x2 Games

Hawk/Dove Game

		P2	
		H	D
P1	H	-1,-1	4,0
	D	0,4	1,1

2x2 Games

Stag-Hunt Game

P2

P1		L	S
	L	3,3	0,2
	S	2,0	2,2

2x2 Games

Rock, Paper, Scissors

P2 col

P1 row		R	P	S
	R	0, 0	-1, 1	1, -1
	P	1, -1	0, 0	-1, 1
	S	-1, 1	1, -1	0, 0

No
Nash equilibrium

2x2 Games

a,b	c,d
e,f	g,h

Where do these numbers come from?

Descriptive agenda: They come from the social sciences

Prescriptive agenda: We choose them as designers of multiagent systems

Elements of a Game

- A set of **players** $\mathbf{P} = \{P1, P2, \ldots\}$
- Each player has a set of **actions**, $\mathbf{A}_i \quad i \in \mathbf{P}$

which results in a joint action set $\mathbf{A} = \mathbf{A}_{P1} \times \mathbf{A}_{P2} \times \ldots$

- A **utility** for each player (defined over the joint action set) $u_i : \mathbf{A} \to \mathbb{R} \qquad i \in \mathbf{P}$

Prisoner's Dilemma

$\mathbf{P}$={row,col}

P2 col

P1 row

	C	D
C	2, 0	0, 3
D	3, 0	1, 1

Nash equilibrium is (D, D)

$$\mathbf{A}_{row} = \{C, D\}$$
$$\mathbf{A}_{col} = \{C, D\}$$

$$\mathbf{A} = \mathbf{A}_{row} \times \mathbf{A}_{col}$$
$$= \{(C,C),(C,D),(D,C),(D,D)\}$$

$$u_{row}(C,C) = 2 \qquad u_{col}(C,C) = 2$$
$$u_{row}(C,D) = 0 \qquad u_{col}(C,D) = 3$$
$$u_{row}(D,C) = 3 \qquad u_{col}(D,C) = 0$$
$$u_{row}(D,D) = 1 \qquad u_{col}(D,D) = 1$$

Elements of a Game

- An action **profile** is what everyone is doing $a \in \mathbf{A}$
- A profile can be written as $a = (a_i, a_{-i})$

i's actions

everyone else's actions

-i is Game Theory notation used to indicate "not i," and could signify more than one such player

- Action profile is a **Nash equilibrium** if $a^\star = (a_i^\star, a_{-i}^\star)$

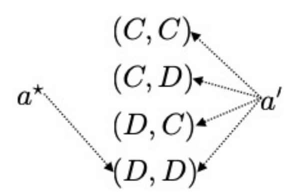
$$u_i(a_i^\star, a_{-i}^\star) \geq u_i(a_i', a_{-i}^\star) \quad \forall a_i' \in \mathbf{A}_i$$
$$\forall i \in P$$

Prisoner's Dilemma

$a = (C, C)$
or $a = (C, D)$
or $a = (D, C)$
or $a = (D, D)$

		P2 col C	P2 col D
P1 row	C	2, 0	0, 3
	D	3, 0	1, 1

Nash equilibrium is (D, D)



$$u_{row}(D, D) \geq u_{row}(\cdot, D)$$
$$u_{col}(D, D) \geq u_{col}(D, \cdot)$$

N-Player Game: Pageant Game

- Everyone on this call (**players**)
- Pick a number between 0 and 100 (**actions**)
- Winner: closest to half the average (**utilities**)

- Values…
- Average = ?
- Half = ?
- Winner = ?
- Nash:

N-Player Game: Pageant Game

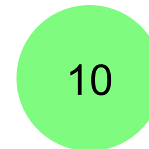
- Everyone on this call (**players**)
- Pick a number between 0 and 100 (**actions**)
- Winner: closest to half the average (**utilities**)

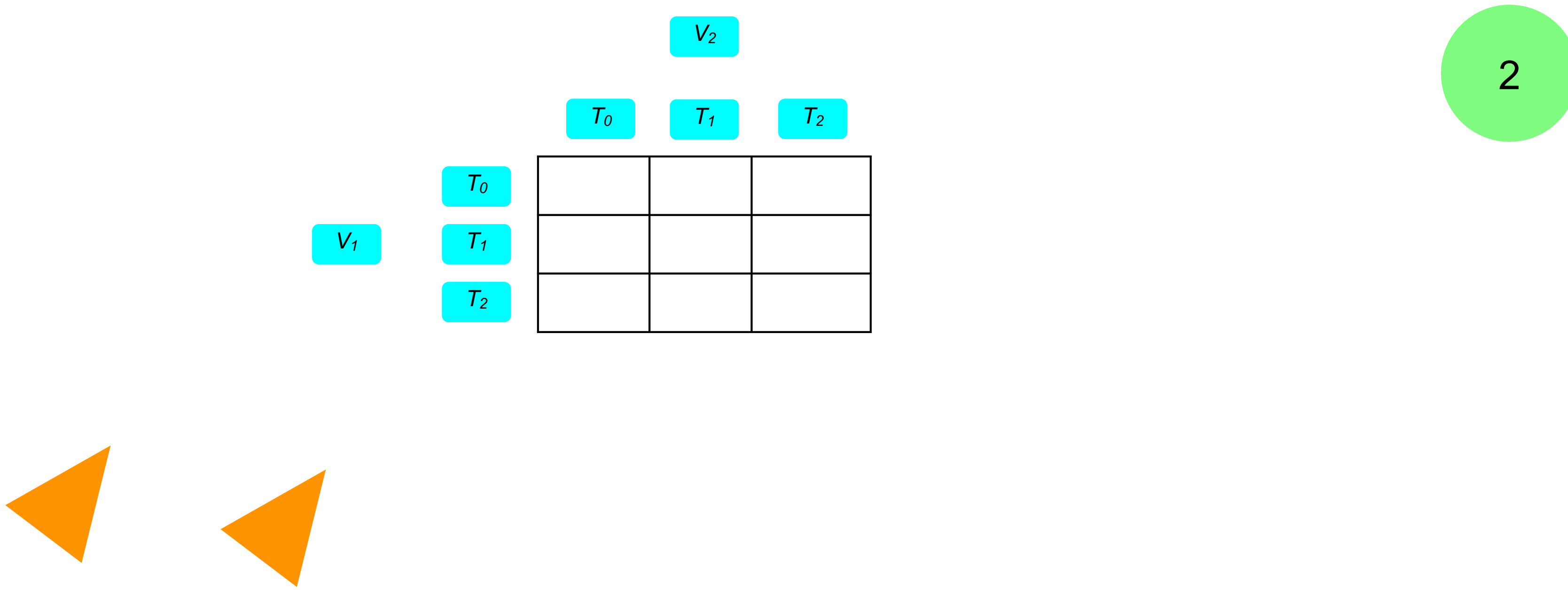
- 77, 50, 49, 45, 37, 100; sum = 358
- Average = 59.67
- Half = 29.83
- Winner = 37
- Nash: (0 for everyone)

Design Considerations

1. Global Objective	Mission, system-level plan, multiagent designer
2. Individual Utility Design	*Align* with global so that that we have a potential game (Nash is optimal)
3. Learning Algorithms	Negotiation mechanism for agents with convergence to agreeable assignments (Nash)

Example





Potential Games

Global objective

$$U_i(a_i', a_{-i}) - U_i(a_i'', a_{-i}) = \phi(a_i', a_{-i}) - \phi(a_i'', a_{-i})$$

Individual utility

Improvement in individual utility is equal to improvement in the "potential function"

Utility Design

$$U_g(a) = \sum_{\tau_j \in \tau} U_{\tau_j}(a)$$

$$U_i(a) = \sum_{\tau_j \in \tau} U_{\tau_j}(a)$$

$$U_i(a) = \sum_{\tau_j \in \mathbf{A}_i} U_{\tau_j}(a)$$

$$U_i(a) = \frac{U_{\tau_j}(a)}{n_{\tau_j}(a)}$$

$$U_i(a) = U_{\tau_j}(a_i, a_{-i}) - U_{\tau_j}(\tau_0, a_{-i})$$

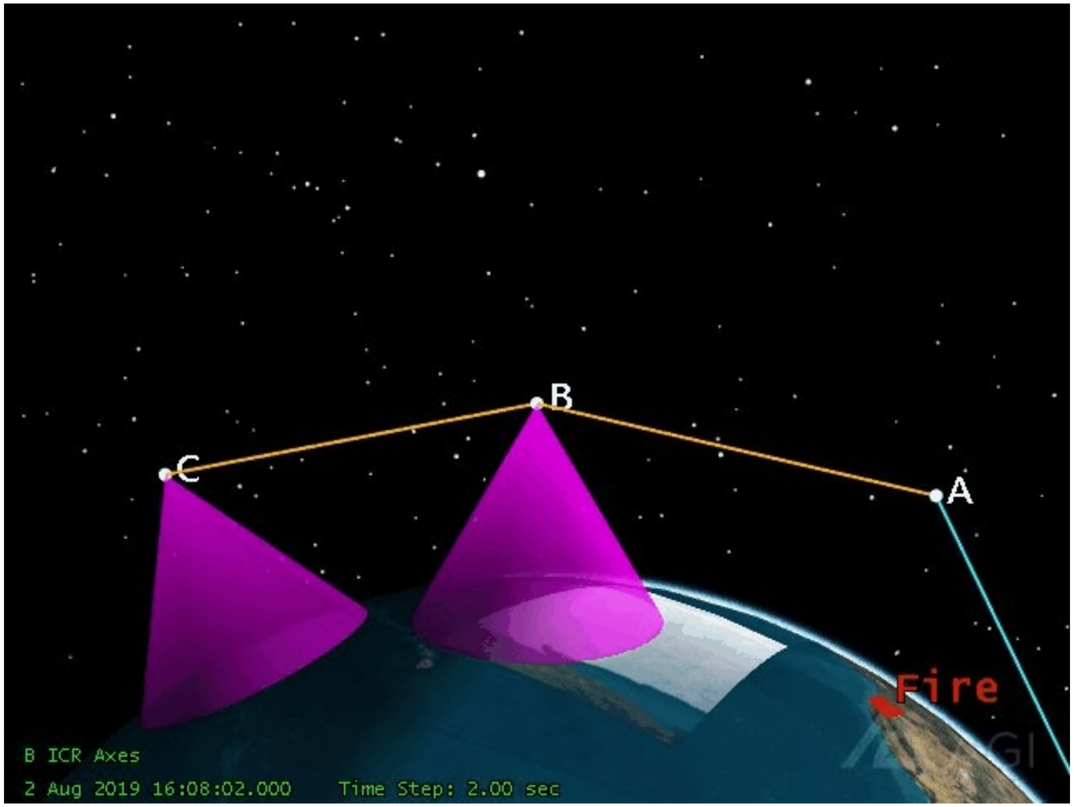
Utility	Description	Pros	Cons	Comments
Identical Interest (IIU)	Set individual utility to global objective	Potential Game with global objective as potential function	Each agent needs everyone else's assignment	Information burden on agents; learnability issue
Range-Restricted (RRU)	IIU with range restriction	Potential Game with global objective as potential function	Learnability issue	Global information burden alleviated
Equally Shared (ESU)	Distribute value equally	Information burden alleviated	Still a potential game, but not for the objective we care about	Similar to so called Congestion Games
Wonderful Life (WLU)	Benefit with me minus benefit without me	Potential Game with global objective as potential function	Optimal solution agreeable, but so are suboptimal ones	Marginal contribution structure

Learning Algorithms

Algorithm	Description	Pros	Cons	Comments
Fictitious Play (FP)	Maintain empirical frequency of other players	Converges to NE for Potential Games	Computationally expensive; assumes opponent using stationary dist	Optimize over product of histograms
Regret Monitoring (RM)	Propose an action based on regret of not proposing it in the past	Converges to NE (Potential)	Infinite memory (info still lingers)	Calculate regret by replaying from the beginning
Regret Monitoring with Fading Memory (RMFM)	Same as above, but discount the past	Converges to NE (Potential)	Infinite memory (info still lingers)	Converges for *Ordinal* Potential (broader class)
Spatial Adaptive Play (SAP)	Select an action that maximizes based on opponent's actions from yesterday	Converges to NE (Potential)	Token-based; slower than FP and RM	Low compute burden; optimal with high prob; parallel update
Selective Spatial Adaptive Play (SSAP)	Same as above, but over a selected set	Converges to NE (Potential)	Token-based; faster than SAP	Optimal with high prob

Backup Material

On-orbit Tipping and Cueing



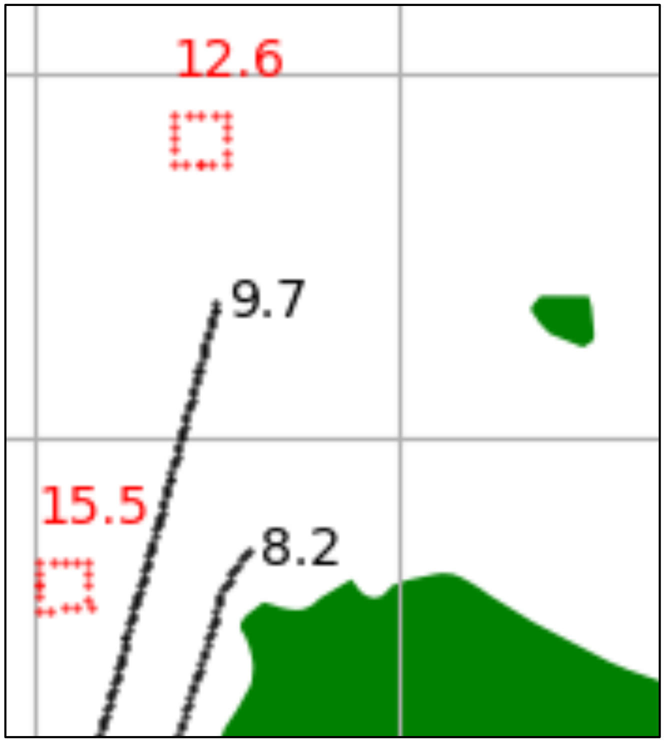
Game Design

1. Global Objective — Total events under custody

2. Individual Utility Design — Wonder Life Utility

3. Learning Algorithms — Spatial Adaptive Play

Pattern of Life



JOHNS HOPKINS

APPLIED PHYSICS LABORATORY