JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

Game Theory

601.464
Artificial Intelligence
TR 10.30AM—11.45AM

Topics

- Normal Form (matrix) games
    - Prisoner's Dilemma
    - Typewriter
    - Stag-Hunt
    - Rock-Paper-Scissor, and many more…
  - Elements of a Game
    - Players, actions, utilities
  - N-player Game
    - Beauty Contest
  - Mixed-Strategy Nash Equilibrium
  - Learning Algorithms

Normal Form (Matrix) Games

Prisoner's Dilemma

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

Normal Form (Matrix) Games

Prisoner's Dilemma

Action available to P2

Another action available to P2

Action available to P1

P2 col

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

P2's utility (or payoff) is 3 when in the cell (C,D), where C is P1's action and D is P2's action

P1's payoff in this cell is 0

Another action available to P1

Normal Form (Matrix) Games

Prisoner's Dilemma

P2

col

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

Best Response by P1, if P2 plays C

Normal Form (Matrix) Games

Prisoner's Dilemma

P2
col

P1 row

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

Best Response by P1, if P2 plays D

Normal Form (Matrix) Games

Prisoner's Dilemma

P2
col

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

Best Response by P2, if P1 plays C

Normal Form (Matrix) Games

Prisoner's Dilemma

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

Best Response by P2, if P1 plays D

Normal Form (Matrix) Games

Prisoner's Dilemma

$$\begin{array}{cc|c|c|}
 & & \multicolumn{2}{c}{\text{P2 col}} \\
 & & C & D \\
\hline
\text{P1 row} & C & 2, 0 & 0, 3 \\
\hline
 & D & 3, 0 & 1, 1 \\
\hline
\end{array}$$

Cell with all best responses is a Nash equilibrium

Normal Form (Matrix) Games

Prisoner's Dilemma

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

Nash equilibrium is (D, D)

Normal Form (Matrix) Games

Prisoner's Dilemma

P2
col

P1 row

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

When the game is at (D,D), if players are asked one by one, whether they would like to change their minds, the answer is no

Normal Form (Matrix) Games

Prisoner's Dilemma

P2

col

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

At a Nash equilibrium, there is no **unilateral** incentive to deviate

Normal Form (Matrix) Games

Stand/Sit

		P2 col	
		Sta	Sit
P1 row	Sta	1, 1	1, 3
	Sit	3, 1	2, 2

Nash equilibrium (Sit,Sit)

Normal Form (Matrix) Games

Technology

Nash equilibrium (A,A)

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

Nash equilibrium (B,B)

Normal Form (Matrix) Games

Hawk/Dove

P2 col

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

Nash equilibrium (H,D)

Nash equilibrium (D,H)

Normal Form (Matrix) 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

Normal Form (Matrix) 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

Well, no **pure**
Nash equilibrium

(as opposed to **mixed** strategy Nash equilibrium; more on that later…)

Normal Form (Matrix) Games

Stag Hunt

Nash equilibrium (S,S)

Nash equilibrium (H,H)

		P2 col S	P2 col H
P1 row	S	3, 3	0, 2
P1 row	H	2, 0	2, 2

Normal Form (Matrix) Games

Stag Hunt

		P2 col	
		S	H
P1 row	S	3, 3	0, 2
	H	2, 0	2, 2

Two Nash equilibria

(S,S) is **payoff** dominant because the payoff to a player is higher at (S,S), compared to the other Nash equilibrium

(H,H) is **risk** dominant. Why? Consider P1. Imagine P1 thinks it is at a Nash equilibrium, but there's been a miscue or miscommunication with P2. The drop in payoff from the Nash equilibrium (H,H) to (H,S) is lower than a drop in payoff from the Nash equilibrium of (S,S) to (S,H)

Normal Form (Matrix) 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** $\mathbf{P} = \{P1, P2, \dots\}$

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

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

- 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

P2

col

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

Nash equilibrium is (D, D)

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

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

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

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

Elements of a Game

- An action **profile** is what everyone is doing $a \in \mathbf{A}$

i's actions

everyone else's actions

- A profile can be written as $a = (a_i, a_{-i})$

-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 $a^\star = (a_i^\star, a_{-i}^\star)$

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

$$\forall i \in P$$

Prisoner's Dilemma

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

P2 col

P1 row

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

Nash equilibrium is (D, D)

$a^\star$ $\quad$ (C, C) $\quad$ (C, D) $\quad$ (D, C) $\quad$ (D, D) $\quad$ a'

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

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

N-Player Game

- Beauty Contest
  - N players — players
  - Pick a number between 0 and 100 — action
  - Winner is the closest to half the average — utility

Collin:	10		100	0
Brandon:	0		100	0
Maya:	23		100	0
David:	73		100	0
Andrew:	100		100	0
Bruce:	88		100	0
Abigail:	26		100	0
	Avg: ~45.7		Avg: 100	Avg: 0
	Half: ~22.9		Half: 50	Half: 0

Any player (other than Maya) would want to change their minds, if asked unilaterally (i.e., one by one)

Not a Nash equilibrium

Asked unliterally, any player would be better off by choosing a lower number

Not a Nash equilibrium

Asked unliterally, no player is better off by changing their choice

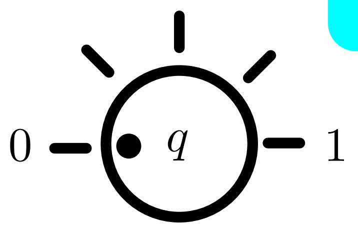
No incentive for unilateral deviation

Nash equilibrium

Mixed-Strategy

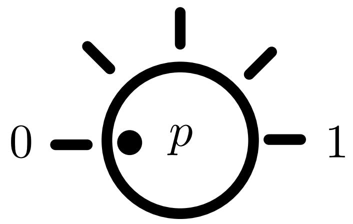
P2

P2 mixes its actions with probability q



P1 mixes its actions with probability p

P1

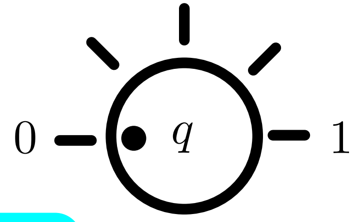


		q L	$1-q$ R
p	U	4, 1	0, 0
$1-p$	D	0, 0	1, 4

P2 picks R with probability 1-q

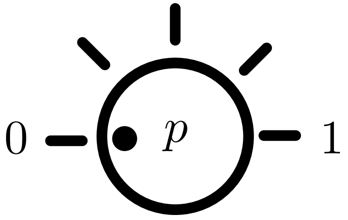
Mixed-Strategy

P2



Pure Nash equilibrium (U,L)

P1

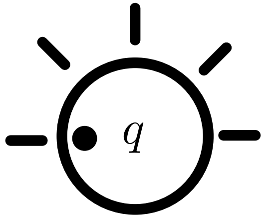


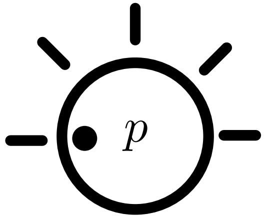
		q L	$1-q$ R
p	U	4, 1	0, 0
$1-p$	D	0, 0	1, 4

Pure Nash equilibrium (D,R)

Mixed-Strategy

P2

0  1

P1 0  1

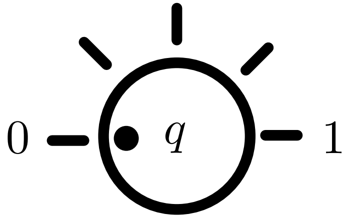
		q	$1-q$
		L	R
p	U	4, 1	0, 0
$1-p$	D	0, 0	1, 4

P2 chooses L over R if

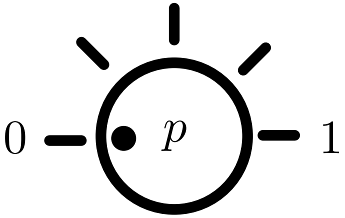
$$(1 \cdot p) + (0 \cdot (1-p)) > (0 \cdot p) + (4 \cdot (1-p))$$

Mixed-Strategy

P2



P1



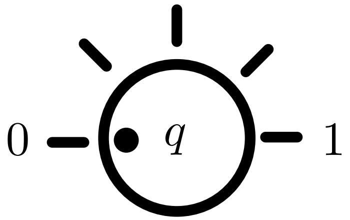
		q L	$1-q$ R
p	U	4, 1	0, 0
$1-p$	D	0, 0	1, 4

P2 chooses L over R if $p > \frac{4}{5}$

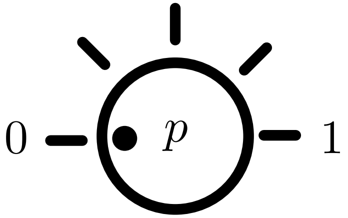
P2 chooses R over L if $(1 \cdot p) + (0 \cdot (1-p)) < (0 \cdot p) + (4 \cdot (1-p))$

Mixed-Strategy

P2



P1



		q L	$1-q$ R
p	U	4, 1	0, 0
$1-p$	D	0, 0	1, 4

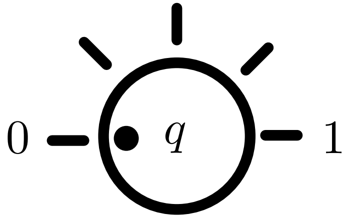
P2 chooses L over R if $p > \frac{4}{5}$

P2 chooses R over L if $p < \frac{4}{5}$

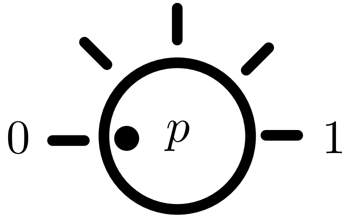
P2 indifferent between L and R $(1 \cdot p) + (0 \cdot (1-p)) = (0 \cdot p) + (4 \cdot (1-p))$

Mixed-Strategy

P2



P1



		q L	$1-q$ R
p	U	4, 1	0, 0
$1-p$	D	0, 0	1, 4

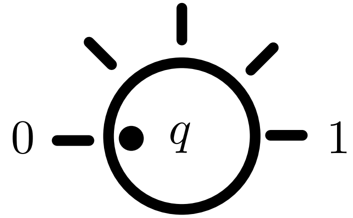
P2 chooses L over R if $p > \frac{4}{5}$

P2 chooses R over L if $p < \frac{4}{5}$

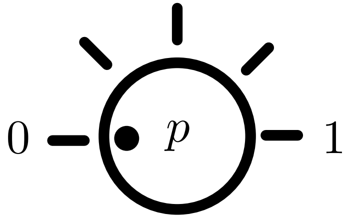
P2 indifferent between L and R $p = \frac{4}{5}$

Mixed-Strategy

P2



P1



		q L	$1-q$ R
p	U	4, 1	0, 0
$1-p$	D	0, 0	1, 4

P1 chooses U over D if

$$(4 \cdot q) + (0 \cdot (1 - q)) > (0 \cdot q) + (1 \cdot (1 - q))$$

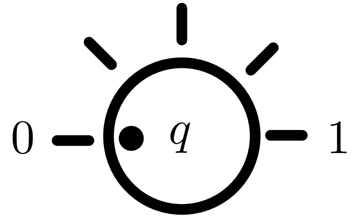
P2 chooses L over R if $p > \frac{4}{5}$

P2 chooses R over L if $p < \frac{4}{5}$

P2 indifferent between L and R $p = \frac{4}{5}$

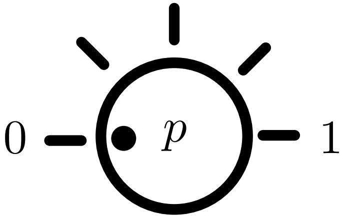
Mixed-Strategy

P2



		q	$1-q$
		L	R
p	U	4, 1	0, 0
$1-p$	D	0, 0	1, 4

P1



P1 chooses U over D if $q > \frac{1}{5}$

P1 chooses D over U if $(4 \cdot q) + (0 \cdot (1-q)) < (0 \cdot q) + (1 \cdot (1-q))$

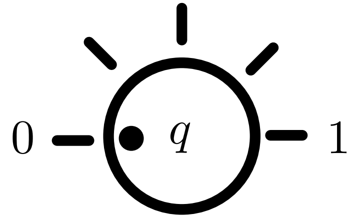
P2 chooses L over R if $p > \frac{4}{5}$

P2 chooses R over L if $p < \frac{4}{5}$

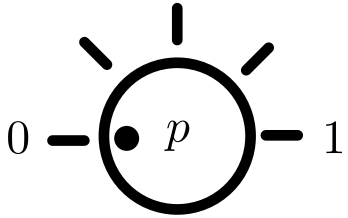
P2 indifferent between L and R $p = \frac{4}{5}$

Mixed-Strategy

P2



P1



		q L	$1-q$ R
p	U	4, 1	0, 0
$1-p$	D	0, 0	1, 4

P2 chooses L over R if $p > \frac{4}{5}$

P2 chooses R over L if $p < \frac{4}{5}$

P2 indifferent between L and R $p = \frac{4}{5}$

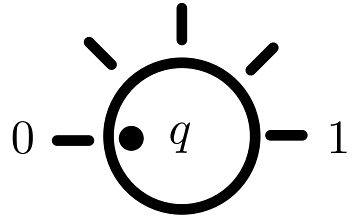
P1 chooses U over D if $q > \frac{1}{5}$

P1 chooses D over U if $q < \frac{1}{5}$

P1 indifferent between U and D $(4 \cdot q) + (0 \cdot (1-q)) = (0 \cdot q) + (1 \cdot (1-q))$

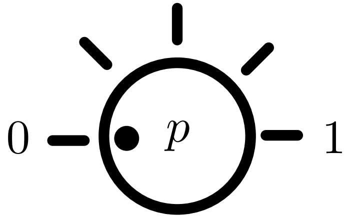
Mixed-Strategy

P2



		q	$1-q$
		L	R
p	U	4, 1	0, 0
$1-p$	D	0, 0	1, 4

P1



P1 chooses U over D if $q > \frac{1}{5}$

P1 chooses D over U if $q < \frac{1}{5}$

P1 indifferent between U and D $q = \frac{1}{5}$

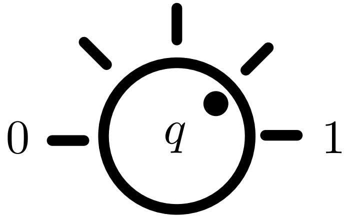
P2 chooses L over R if $p > \frac{4}{5}$

P2 chooses R over L if $p < \frac{4}{5}$

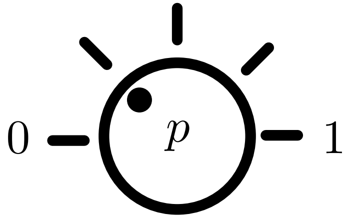
P2 indifferent between L and R $p = \frac{4}{5}$

Mixed-Strategy

P2



P1



		q L	$1-q$ R
p	U	4, 1	0, 0
$1-p$	D	0, 0	1, 4

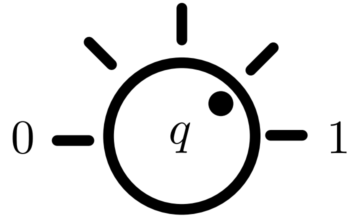
$$(p = \frac{4}{5}, q = \frac{1}{5})$$

Mixed-strategy Nash equilibrium

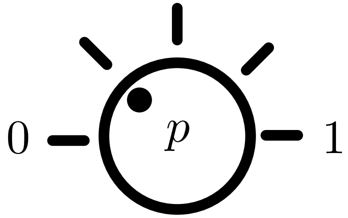
No unilateral incentive to deviate from this setting

Mixed-Strategy

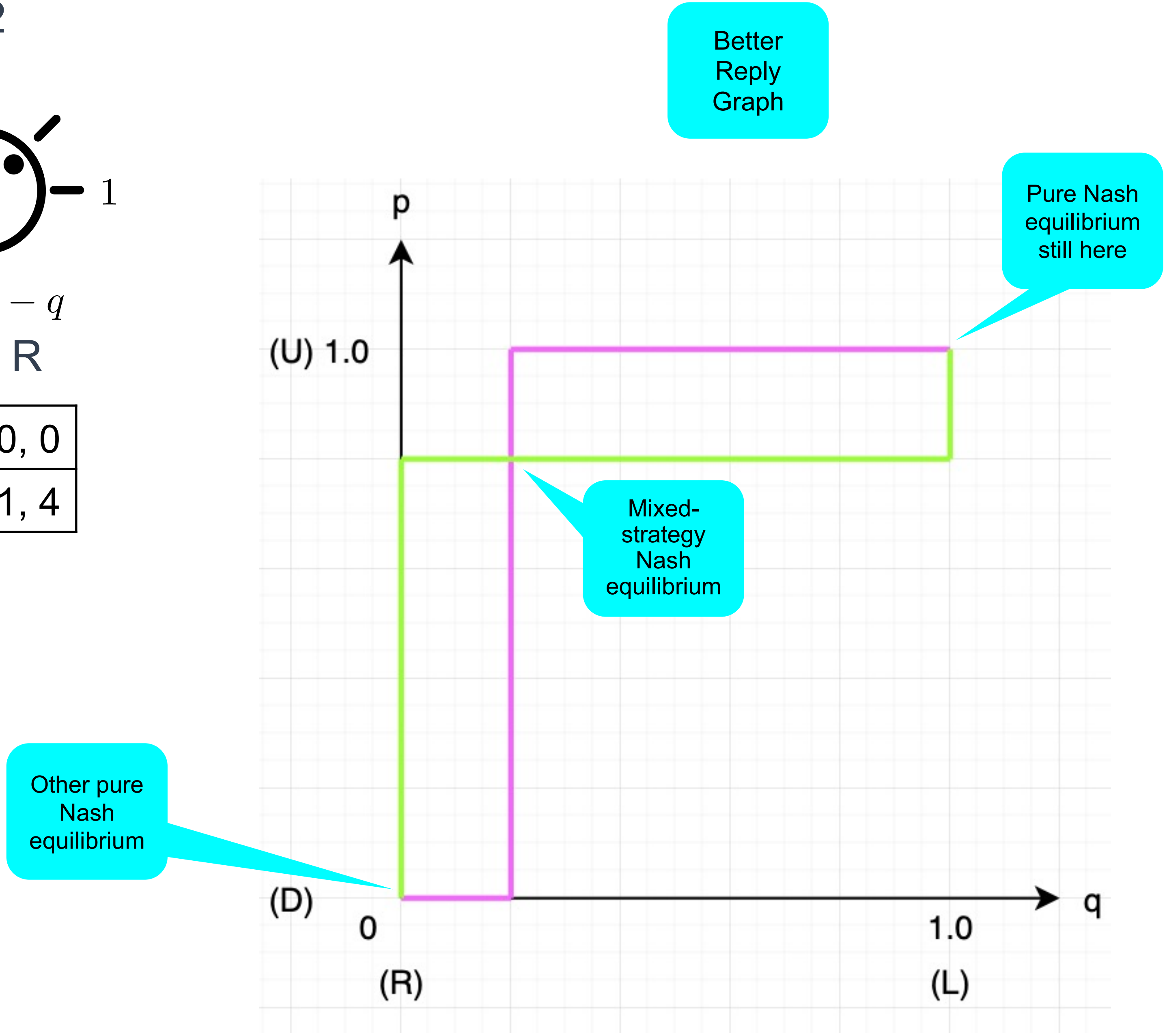
P2



P1



		q	$1-q$
		L	R
p	U	4, 1	0, 0
$1-p$	D	0, 0	1, 4



Learning Algorithms

Looks like a utility function because it takes a joint action set like (C,C) and outputs a real number

The difference is, this function is not specific to any player, so there's no "i" subscript.

- If there exists a *potential function* $\Phi : \mathbf{A} \rightarrow \mathbb{R}$

such that

Difference in potential function

$$u_i(a_i', a_{-i}^{\cdot}) - u_i(a_i'', a_{-i}^{\cdot}) = \Phi(a_i', a_{-i}^{\cdot}) - \Phi(a_i'', a_{-i}^{\cdot}) \qquad \forall i \in \mathbf{P}$$

Difference in utility for player i between choosing two different actions

then we have a *potential game*

Learning Algorithms

- If we have a potential game, then applying one of the well-known **learning algorithms** (often referred to as **negotiation mechanisms**) converges the actions of all the players to a Nash equilibrium

- Examples of learning algorithms:
  - *Fictitious Play*
  - *Spatial Adaptive Play*
  - *Regret Monitoring*
  - *Best Response with Inertia*

- Best Response with Inertia

$$a_i(t) = \begin{cases} BR(a_{-i}(t-1)) & \text{with high porbability} \\ a_i(t-1) & \text{otherwise} \end{cases}$$

$$\forall i \in \mathbf{P}$$

On a given day t, Player i, with high probability, performs a **best response** to the other player's action from yesterday (t-1)

Sometimes Player i is lazy, and just does what it did yesterday (hence, "inertia" in the name)

Day 1's actions are picked at random

JOHNS HOPKINS

APPLIED PHYSICS LABORATORY