MONASH University
Information Technology

Acknowledgements

This material includes content adapted from instructional resources that have been made available by the authors of our main textbook "Multiagent Systems", Yoav Shoham and Kevin Leyton-Brown, Cambridge University Press, 2009.

MONASH University
Information Technology

FIT5226
Multi-agent System & Collective Behaviour

Wk 7: Classical Game Theory (Nash)

Bernd Meyer, September 2024

Individual vs Collective Perspective

From the outside all agent's interests are equally important.

What does this mean for outcomes
that differ only on the *individual* level ?

Collective Perspective: Pareto Optimality

If an outcome o' is at least as good as another outcome o for all agents but there is at least one agent who strictly prefers o' to o, we say that o' pareto dominates o.

An outcome o^* is pareto optimal if no other outcome o pareto dominates it.

- Does every game have a pareto optimal outcome?
- Can there be multiple pareto optimal outcomes?

Pareto Optimality Example

	L	R
L	*1,1*	*0,0*
R	*0,0*	*1,1*

- Traffic Game

Pareto Optimality Example

	B	F
B	*2, 1*	*0, 0*
F	*0, 0*	*1, 2*

- Battle of the Sexes

Pareto Optimality Example

	H	T
H	*1,-1*	*-1,1*
T	*-1,1*	*1,-1*

- Matching Pennies

Pareto Optimality Example

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

Prisoners' dilemma

recall that "X is pareto optimal" does *not* mean its pareto-dominant but that it is not pareto-dominated

Pareto Optimality

- Can there be a game with no pareto-optimal outcome?

Outcomes from the individual Perspective

Now let's consider what is best for one individual player.

A strategy s for player x dominates another strategy s’ if

- for all possible choices of a strategy t for player y, the outcome of s is at least as good as that of s’
  s is better than s’ for at least one choice of t

Assume the following pay-0ff matrix:

	$t1$	$t2$	$t3$	$t4$
$s1$	-2	2	-1	1
$s2$	1	4	-2	-3
$s3$	0	2	0	1
$s4$	-1	-2	-1	5

A rational player will never play a dominated strategy

The row player A tries to maximize and therefore s1 is dominated by s3.

A will never play s1.

	$t1$	$t2$	$t3$	$t4$
~~$s1$~~	~~-2~~	~~2~~	~~-1~~	~~1~~
$s2$	1	4	-2	-3
$s3$	0	2	0	1
$s4$	-1	-2	-1	5

The column player B knows that A will never play s1.

B tries to minimize and therefore t1 is dominated by t3.

B will never play t1.

	$t1$	$t2$	$t3$	$t4$
~~$s1$~~	~~-2~~	~~2~~	~~-1~~	~~1~~
$s2$	~~1~~	4	-2	-3
$s3$	~~0~~	2	0	1
$s4$	~~-1~~	-2	-1	5

Morra

Each Player hides one or two coins.

- each player guesses how many coins the opponent has hidden.
- There is a win if and only if one player guesses correctly, the other one incorrectly.
- In case of a win an amount equal to the total number of hidden coins are paid to the winner.

Morra Payoff Matrix

<u>Strategies</u> for both players:

(1/1) Hide one coin, guess “one coin”
(1/2) Hide one coin, guess “two coins”
(2/1) Hide two coins, guess “one coin”
(2/2) Hide two coins, guess “two coins”

	(1/1)	(1/2)	(2/1)	(2/2)
(1/1)	0	2	–3	0
(1/2)	–2	0	0	3
(2/1)	3	0	0	–4
(2/2)	0	–3	4	0

As this is zero-sum, a single matrix is enough to denote the outcomes (instead of a bi-matrix notation)

We speak of the “column player” and the “row player”.

- the pay-off is given for the row player (by convention)
- the row player tries to maximize the outcome
- column player tries to minimize the outcome

Best Response

It would be straight forward to decide what to do
if we knew the actions of all others.

Let $a = \langle a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n \rangle$ be the full action profile

Let $a_{-i} = \langle a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n \rangle$

The full action profile is: $a = (a_{-i}, a_i)$

The best response to a_{-i} is

$$a_i^* \in BR(a_{-i}) \text{ iff } \forall a_i \in A_i, u_i(a_i^*, a_{-i}) \geq u_i(a_i, a_{-i})$$

Players in equilibrium

But we don’t know what the others will do, and their decisions depend on mine!

However, if no player has a unilateral incentive
to change their strategy, the situation should be stable.

We say that players use strategies in equilibrium.

Nash Equilibrium

Obviously, if **every** player plays best response to everyone else, no one has an incentive to deviate.

Thus, a situation is stable if every player is playing their best response to all other players.

$a = (a_1, \ \dots, \ a_n)$ is a **Nash Equilibrium** (*in pure strategies*) iff

$\forall i, a_i \in BR(a_{-i})$

Note: at this point, we only care bout *pure* strategies.

Nash Equilibria are Saddle Points in 2P-ZS Games

How to choose a strategy (after removing the dominated ones)??
Maximize the worst case!

A reasonable way to do this for player row player A is to calculate the minimum pay-off for each possible strategy choice si (assuming that B would react with this choice in the worst case) and to choose the si that maximizes this minimum expected pay-off.

B would do the same swapping minimization and maximization.

	$t1$	$t2$	$t3$	$t4$
$s1$	–2	2	–1	1
$s2$	1	4	–2	–3
$s3$	0	2	0	1
$s4$	–1	–2	–1	5

If A starts to play s3 (or B t3) the other player is forced to choose t3 (s3) since the expected pay-off is worse otherwise. Such a stable point is called a <u>saddle point</u>.

A saddle point is a pair of strategies such that there is no lower pay-off in the row and no higher pay-off in the column.

Nash Equilibrium Example

	L	R
L	*1,1*	*0,0*
R	*0,0*	*1,1*

- Traffic Game

Nash Equilibrium Example

	L	R
L	*1,1*	*0,0*
R	*0,0*	*1,1*

- Traffic Game

Nash Equilibrium Example

	B	F
B	*2, 1*	*0, 0*
F	*0, 0*	*1, 2*

- Battle of the Sexes

Nash Equilibrium Example

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

Prisoners' dilemma

Nash Equilibrium Example

	C	D
C	−1, −1	−4, 0
D	0, −4	−3, −3

Prisoners' dilemma

The paradox of Prisoner's dilemma: the Nash equilibrium is the only non-Pareto-optimal outcome!

Nash Equilibrium Example

	H	T
H	*1,-1*	*-1,1*
T	*-1,1*	*1,-1*

- Matching Pennies

Nash Equilibrium Example

	H	T
H	*1,-1*	*-1,1*
T	*-1,1*	*1,-1*

- Can you play a deterministic strategy in this game? (eg "always heads")
- Leave you opponent in the dark by playing randomly!

Strategy Profiles

Define a strategy s_i for agent i as any probability distribution over the actions A_i

pure strategy: only one action is played with positive probability

mixed strategy: more than one action is played with positive probability

these actions are called the <u>support</u> of the mixed strategy

Let the set of all strategies for i be S_i
Let the set of all strategy profiles be $S = S_1 \times \ldots \times S_n$

Expected Payoff (Utility)

If we adopt a mixed strategy (a set of strategy for alternating use) instead of a pure strategy (a single strategy), each individual game must chose their actions randomly (according to the distribution) in order to be unpredictable.

Let the stochastic row vector X describe the strategy for the row player (probabilities of action choices), the stochastic column vector Y the strategy for the column player.

$$1 = \sum_i x_i = \sum_i y_i; \quad p(s_i) = x_i; \quad p(t_i) = y_i$$

The expected average pay-off (per round) in an iterated game for payoff matrix $A = [a_{ij}]$ is

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_{i,j} \cdot x_i \cdot y_j = XAY$$

Mixed Nash

Best response and Nash equilibrium generalise from actions to strategies:

Best response: $s_i^* \in BR(s_{-i})$ iff $\forall s_i \in S_i, u_i(s_i^*, s_{-i}) \geq u_i(s_i, s_{-i})$

Nash: $s = \langle s_1, \ldots, s_n \rangle$$ is a Nash equilibrium iff $\forall i, s_i \in BR(s_{-i})$

EQUILIBRIUM POINTS IN N-PERSON GAMES

BY JOHN F. NASH, JR.*

PRINCETON UNIVERSITY

Communicated by S. Lefschetz, November 16, 1949

One may define a concept of an n-person game in which each player has a finite set of pure strategies and in which a definite set of payments to the n players corresponds to each n-tuple of pure strategies, one strategy being taken for each player. For mixed strategies, which are probability

Every game with a finite number of pure strategies has at least one Nash Equilibrium in (possibly) mixed strategies.

Finding Nash

Finding Nash (Example)

	B	F
B	*2, 1*	*0, 0*
F	*0, 0*	*1, 2*

Computing Nash equilibria is easy when you can guess the support

For BoS, let's look for an equilibrium where all actions are part of the support

Finding Nash (Example)

	B	F
B	*2, 1*	*0, 0*
F	*0, 0*	*1, 2*
	p	$1\text{-}p$

$$u_1(B) = u_1(F)$$

$$2p + 0(1-p) = 0p + 1(1-p)$$

$$p = \frac{1}{3}$$

Let player 2 play B with p, F with 1 – p.

Player 1 is expected to best-respond with a mixed strategy,
so Player 2 must make them indifferent to choosing between B and F (why?)

Finding Nash (Example)

	B	F
B	*2, 1*	*0, 0*
F	*0, 0*	*1, 2*

Conversely, Player 1 must randomize to make Player 2 indifferent.

Why is player 1 willing to randomize?

Finding Nash (Example)

	B	F	
B	*2, 1*	*0, 0*	q
F	*0, 0*	*1, 2*	$1\text{-}q$

$$u_2(B) = u_2(F)$$

$$q + 0(1 - q) = 0q + 2(1 - q)$$

$$q = \frac{2}{3}$$

Thus the mixed strategy pair

$\left(\frac{2}{3}, \frac{1}{3}\right)$ for Player 1 (row) and $\left(\frac{1}{3}, \frac{2}{3}\right)$ for Player 2 (column)

is a Nash equilibrium *in mixed strategies*

Finding mixed Nash in general

- If the support of the mixed Nash Equilibria is known, the problem reduces to a system of linear equations.
- The set of supports for a finite game is a combinatorial space (size 2^n for *n* actions)
- Finding mixed in general Nash Equilibria is a combinatorial problem.

Solving 2-player 0-sum Games as Linear Programs

The task of player B (column) is to

$$\text{minimize} \quad \max_i \sum_{j=1}^{n} a_{i,j} \cdot y_j$$

$$\text{subject to} \quad 1 = \sum_{j=1}^{n} y_j \wedge \forall j : 0 \le y_j \le 1$$

This is equivalent to the LP

$$\text{minimize} \quad z'$$

$$\text{subject to} \quad \begin{cases} \forall i : z' - \sum_{j=1}^{n} a_{i,j} \cdot y_j \ge 0 \\ 1 = \sum_{j=1}^{n} y_j \wedge \forall j : 0 \le y_j \le 1 \end{cases}$$

To prove this witness that for at least one i the first constraint in the LP must be tight (in order for z to be optimal). Since no constraint can be violated this must be the most restrictive constraint. Therefore

$$z = \max_i \sum_{j=1}^{n} a_{i,j} \cdot y_j$$

The Dual Role of Player A

With exactly the same (dual) reasoning (swapping minimization for maximization), Player A (the row player) arrives at the conclusion that

Player A's task is to

$$\text{maximize} \quad \min_j \sum_{i=1}^{m} a_{i,j} \cdot x_i$$

$$\text{subject to} \quad 1 = \sum_{i=1}^{m} x_i \wedge \forall i : 0 \leq x_i \leq 1$$

which is equivalent to the LP

$$\text{maximize} \quad z$$

$$\text{subject to} \quad \begin{cases} \forall j : z - \sum_{i=1}^{m} a_{i,j} \cdot x_i \leq 0 \\ 1 = \sum_{i=1}^{m} x_i \wedge \forall i : 0 \leq x_i \leq 1 \end{cases}$$

The MiniMax Theorem

For every m x n matrix A there is a stochastic row vector x-opt of length m and a stochastic column vector y-opt of length n such that

$$max_X \; XAY_{opt} = min_Y \; X_{opt}AY$$

The original proof for that in the exclusive context of game theory is quite long and involved. After reduction to LP the proof has been much simplified by recasting it into the Duality Theory of linear programming.

Xopt and *Yopt* are therefore the strategies that the two players should adopt under the above reasoning. The common pay-off of these strategies is referred to as the “value” *v* of the game.

A can make a guaranteed win of $v and B can protect themselves from losses greater than $v.

It follows that the value of a fair game must be v=0.

A symmetric game (-*aji* = *aij*) is always fair.

Unexpected Conclusion

Corollary (Gloating):

In a game of the above type your opponent can gain nothing from you announcing your strategy if it is optimal.

You can announce your strategy without damaging your prospects.

Homework: Contemplate the <u>rational</u> value of bluffing.

- The following are NP-complete problems. Given a bi-matrix game, does it have…
    - at least two Nash Equilibria?
    - A Nash Equilibrium in which player 1 has at least a given amount of payoff
    - A Nash Equilibrium with support size at least a given number.
    - A Nash Equilibrium whose support does not contain strategy s?

"Algorithmic Game Theory", Nissan et al (2007).

The Price of Anarchy - or: Is Nash good?

The "Price of Anarchy" is the payoff lost due to the lack of central coordination.

For example, in the prisoners' dilemma (below),
central coordination (communication) would enable the prisoners to achieve -2 in total

due to a lack of coordination they will only obtain -6 at Nash.

This "price of anarchy" is usually defined as the quotient of these quantities, here 2/6

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

Correlated Equilibrium

If there is intelligent life on other planets, in a majority of them, they would have discovered correlated equilibrium before Nash equilibrium.

– Roger Myerson

Correlated Equilibrium (Example 1: BoS)

	B	F
B	*2, 1*	*0, 0*
F	*0, 0*	*1, 2*

- The ideal fair outcome is to play B/F equally often $\left(\frac{1}{2}, \frac{1}{2}\right)$
- There is no way to achieve this. Players may and will miscoordinate!

Correlated Equilibrium (Example 2: Traffic)

	Go	Stop
Go	*-10, -10*	*1, 0*
Stop	*0, 1*	*-1, -1*

- Shouldn't we have a traffic light?
- The traffic light could also be random!!!

- a random variable with a commonly known distribution
- and a private signal to each player about the outcome is sufficient!

- signal doesn't determine the outcome or others' signals;
- it correlates them

Correlated Equilibrium (Def)

Given an n-agent game $G = (N, A, u)$, a correlated equilibrium is a tuple (v, π, σ), where v is a tuple of random variables $v = (v_1, \ldots, v_n)$ with respective domains $D = (D_1, \ldots, D_n)$, π is a joint distribution over v, $\sigma = (\sigma_1, \ldots, \sigma_n)$ is a vector of mappings $\sigma_i : D_i \mapsto A_i$, and for each agent i and every mapping $\sigma_i' : D_i \mapsto A_i$

$$\sum_{d \in D} \pi(d) u_i \left(\sigma_1(d_1), \ldots, \sigma_i(d_i), \ldots, \sigma_n(d_n)\right)$$

$$\geq \sum_{d \in D} \pi(d) u_i \left(\sigma_1(d_1), \ldots, \sigma_i'(d_i), \ldots, \sigma_n(d_n)\right)$$

Existence

For every Nash equilibrium there exists a corresponding correlated equilibrium

Proof as an exercise

Not every correlated equilibrium is equivalent to a Nash equilibrium.

Thus correlated equilibrium is a weaker notion than Nash.

Take home lessons

- The core difficulty with finding the optimal strategy is that desirable actions depend on the actions of other players. We thus have circular dependencies.
- If a player knew the actions of all other players (and the payoff structure), they could simply choose a “best response.”
- A stable outcome would be if every player plays the best response to all other players.
- This is called the Nash equilibrium. It constitutes a situation where no player has an incentive to deviate unilaterally from the combination of strategies being played.
- Alternative equilibrium definition (e.g. correlated equilibrium) exist
- Nash equilibria can be defined for pure strategies (single action) or mixed strategies (probability distributions over actions)
- Determining Nash equilibria is generally computationally hard.
- For 2-player zero-sum games they can be computed efficiently as linear programs.
- This is based on the MiniMax theorem, which can be understood as a corollary of the Duality Theorem of linear programming (even though historically it isn’t).