GSOE9210 Engineering Decisions

Victor Jauregui

vicj@cse.unsw.edu.au
www.cse.unsw.edu.au/~gs9210

Solving games

1. Modelling player behaviour
   - Solutions of zero-sum games
   - Best response
   - Repeated play; equilibria
   - Beliefs; rationalisation
   - Non strictly competitive games
   - Cooperation in games
   - Games against Nature

Outline

1. Modelling player behaviour
   - Solutions of zero-sum games
   - Best response
   - Repeated play; equilibria
   - Beliefs; rationalisation
   - Non strictly competitive games
   - Cooperation in games
   - Games against Nature

Two player zero-sum games: dominance

Consider the following zero-sum game (matrix entries are the payoffs for the row player):

	b_1	b_2	b_3	b_4
a_1	0	1	7	6
a_2	4	2	3	4
a_3	3	1	0	2
a_4	0	0	7	3

Using dominance, the solution of this game is the play (a_2, b_2).

Rational behaviour and strategic uncertainty

- In games the uncertainty for each player includes the *behaviour* of other players; *i.e.*, which strategy they'll choose
- This uncertainty can be reduced if players have *common knowledge* about the preferences and rationality of other players
- Dominance reduces *strategic uncertainty* about rational behaviour of other players (*e.g.*, rational players will never play dominated strategies)
- General principle about rational behaviour: *best response* . . .

Best response

Consider again the previous zero-sum game:

	b_1	b_2	b_3	b_4
a_1	0	1	7	6
a_2	4	2	3	4
a_3	3	1	0	2
a_4	0	0	7	3

- Play (a_2, b_2) is maximal in its column and minimal in its row
- *i.e.*, if column player plays b_2, then a_2 gives best possible outcome for row player
- Conversely, if row player plays a_2, then b_2 gives best possible outcome for column player

Best response: zero-sum games

Definition (Best response)

A player's strategy s is a *best response* to another player's strategy s' if it gives a preference maximal outcome against s'.

	b_1	b_2
a_1	${}_{*}2$	0^{*}
a_2	1^{*}	${}_{*}3$

In a zero-sum game:

- for any strategy of the column player, a best response of the row player is a strategy which maximises the column value ($*$)
- for any strategy of the row player, a best response of the column player is a strategy which minimises the row value ($*$)

Best response

	b_1	b_2	b_3	min
a_1	1^{*}	2	${}_{*}6$	1
a_2	${}_{*}7$	${}_{*}3^{*}$	3^{*}	3
a_3	3	2^{*}	5	2
max	7	3	6	

- Column player's best responses are minimal in their row
- Row player's are maximal in their column
- Against any strategy there is at least one best response; possibly more than one (*e.g.*, row 2)
- If there are multiple best responses, then they have the same payoff

Best response: *Maximin*

	b_1	b_2	b_3	min
a_1	1^{*}	2	$_{*}6$	1
a_2	$_{*}7$	$_{*}3^{*}$	3^{*}	3
a_3	3	2^{*}	5	2
max	7	3	6	

- Row player's *Maximin* strategy is best strategy against 'perfect play' by opponent
- Above, row player's *Maximin* strategy is a_2; Column player's *Maximin* strategy (*i.e.*, *miniMax* strategy) is b_2
- *Maximin* is rational play if, *e.g.*, opponent can see your move

Repeated play

	b_1	b_2	b_3
a_1	1^{*}	2	$_{*}6$
a_2	4	$_{*}3^{*}$	4
a_3	$_{*}7$	2^{*}	5

- Suppose initially row player plays a_3, hoping for best outcome; similarly column player plays b_1; play (a_3, b_1)
- Row player happy (best response)
- Column player unhappy, so switches to best response b_2; in response row player plays a_2; ...
- Play 'stabilises' at (a_2, b_2)

Equilibrium

The ‘stable’ play (a_2, b_2) has property that each of its strategies is a best response to the other.

	~~b_1~~	b_2	~~b_3~~
~~a_1~~	~~1^{*}~~	~~2~~	~~$_{*}6$~~
a_2	~~4~~	$_{*}3^{*}$	~~4~~
~~a_3~~	~~$_{*}7$~~	~~2^{*}~~	~~5~~

John F. Nash $(1928–2015^{\dagger})$

Definition (Nash equilibrium)

A play is in *equilibrium* if each of its strategies is a best response to the others.

Equilibrium: belief interpretation

	b_1	b_2	b_3
a_1	1^{*}	2	6
a_2	$_{*}4$	$_{*}3^{*}$	4
a_3	$_{*}7$	2^{*}	5

- If row player believes column player will play b_2, then row player cannot improve outcome, and vice versa
- More generally, if each player believes the other will play according to their equilibrium strategy, then neither can improve their outcome by deviating from their equilibrium strategy

Equilibrium: existence and uniqueness

- Not all games have an equilibrium . . . in pure strategies

	b_1	b_2
a_1	${}_*2$	0^*
a_2	1^*	${}_*3$

- Some games may have multiple equilibria:

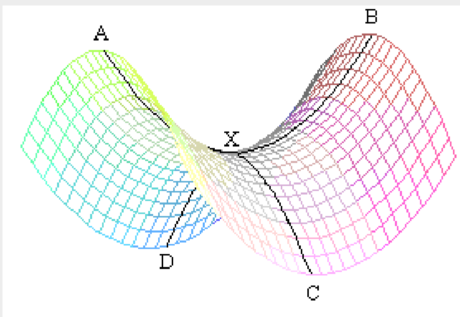
	b_1	b_2	b_3	b_4
a_1	4	${}_*2^*$	5	${}_*2^*$
a_2	2	1	-1	-2
a_3	3	${}_*2^*$	4	${}_*2^*$
a_4	-1	0	6	1

Zero-sum games: saddle points

Definition (Saddle point)

An entry in a zero-sum game is called a *saddle point* iff it is minimal in its row and maximal in its column.

	b_1	b_2	b_3
a_1	1^*	3	4
a_2	${}_*7$	${}_*5^*$	6^*
a_3	3^*	4	${}_*8$



Theorem (Minimax)

In a zero sum game, saddle points represent equilibria.

Zero sum games: solutions

Theorem

If a zero sum game has an equilibrium, then it corresponds to the players playing Maximin strategies.

	b_1	b_2	b_3	min
a_1	1	3	4	1
a_2	7	${}_{*}5^{*}$	6	5
a_3	3	4	8	3
max	7	5	8	

Because the matrix entries are the payoffs for the row player, the column player's *Maximin* strategy translates to a *miniMax* strategy.

Zero-sum games: equilibrium

	b_1	b_2	b_3	min
a_1	1	3	4	1
a_2	7	${}_{*}5^{*}$	6	5
a_3	3	4	8	3
max	7	5	8	

Theorem (Unique value)

All equilibria in a zero sum game yield the same payoffs. This payoff is said to be the value of the game.

- The value of the game above is 5
- Equilibria in zero-sum games are *Maximin* strategies (*miniMax* for column player)

Zero-sum games: finding saddle points

	b_1	b_2	b_3	min
a_1	1	3	4	1
a_2	7	5*	6	5
a_3	3	4	8	3
max	7	5	8	

- Saddle points are *Maximin* strategies
- To find them:
  - Use *Maximin* to evaluate each of the players' strategies (*i.e.*, *miniMax* for column player)
  - If the *Maximin* values agree for any play (*e.g.*, 5 above), then that is a saddle point of the game

Behaviour and beliefs

- A game matrix represents possible outcomes, but says nothing about the players' *behaviour*; *i.e.*, which strategies the players should play
- Dominance and best response are principles about rational *behaviour*
- An agent's behaviour should depend on its *beliefs* about the other players' behaviour (including likelihoods)
- In order to better explain behaviour we must formulate an agent's beliefs

Rational behaviour: rationalisation

Rational behaviour principle: best response

A rational player should not play an strategy which is not a best responses to any of its opponent's strategies.

Definition (Rationalisable strategies)

A strategy is *rationalisable* for a player if it is a best response to some rational strategy of the other players.

- Only rationalisable strategies should be considered by players; *i.e.*, non-rationalisable strategies can be eliminated
- A dominated strategy is never rationalisable*

Theorem

A rationalisable play will survive elimination by iterated dominance.

Beliefs and behaviour

- Beliefs about the other players' play can be represented by a mixture of the other players' pure strategies
- Player A assigns to player B's strategy b_j a 'proportion' p_j if A's belief in the 'degree of likelihood' that B will play b_j is p_j
- Recall that utilities encode preferences in the presence of uncertainty (risk)

Best response to beliefs: zero-sum games

- Suppose player A believes that player B is twice as likely to play b_2 as b_1; *i.e.*, B will play b_1 with probability $\frac{1}{3}$ and b_2 with probability $\frac{2}{3}$
- Let $\beta \sim (\frac{1}{3}, \frac{2}{3})$ represent A's 'belief' about B's behaviour

B

A	b_1	b_2
a_1	2	0
a_2	1	3

- For belief β calculate the *Bayes* values of A's strategies:

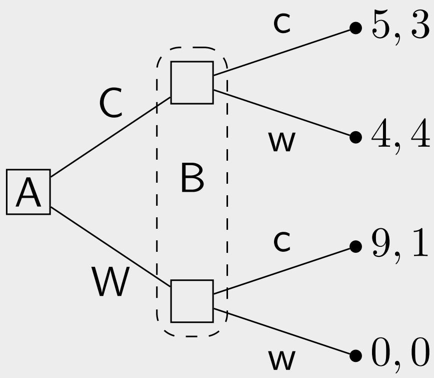
$$V_B^{\beta}(a_1) = \tfrac{1}{3}(2) + \tfrac{2}{3}(0) = \tfrac{2}{3}$$
$$V_B^{\beta}(a_2) = \tfrac{1}{3}(1) + \tfrac{2}{3}(3) = \tfrac{7}{3}$$

- Therefore, A's best response given belief β about B is a_2.

Rationalisation of behaviour and belief

- Any strategy that is not a best response for any belief β about the other players' will not be played; *i.e.*, it should receive degree of belief (*i.e.*, probability) 0
- In general, a strategy is *rationalisable* iff it is *Bayes* for some belief β (not just for some pure strategy)
- Compare *rationalisability* and *admissibility*
- In a zero-sum game, a player's rationalisable strategies must be on the player's 'admissibility frontier'

Non zero-sum games: best response



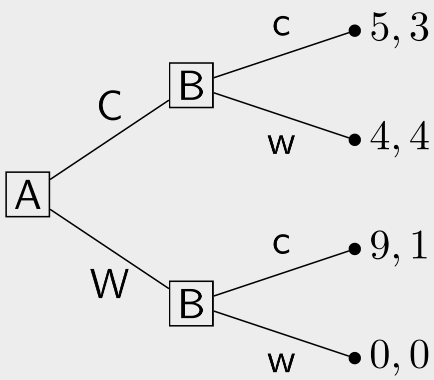
B

A		c	w
	C	5, 3	4, 4
	W	9, 1	0, 0

- If Alice were to wait, then Bob's best counter-move would be to climb
- Conversely, if Bob were to climb, then Alice's best counter-move would be to wait below

Solving games

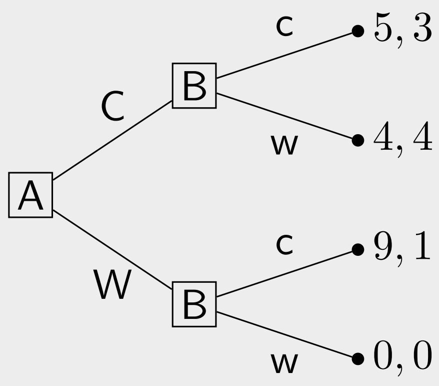
- What if Alice moves first?

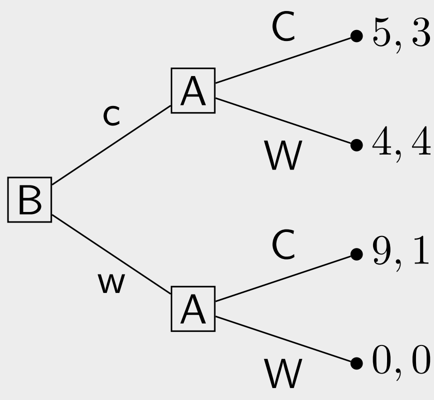


Exercises

- What is Bob's best response to Alice waiting? To Alice Climbing?
- Are there any equilibrium pairs/points? If so, which are they?

Equilibrium and solutions





Exercise

For the problems above, find all the equilibrium plays.

- In games that aren't strictly competitive, determining which equilibrium points are solutions is less clear, because opportunities for co-operation should be considered
- Other considerations include: group benefit (Pareto optimality), initial tendencies (equilibrium), *etc*.

Non strictly competitive games

Example (The Prisoner's Dilemma)

Alice and Bob are suspects in a joint crime. The police doesn't yet have enough evidence to convict both/either, so it is trying to get either to implicate the other. The police inspector offers each separately a reduced sentence if they defect (D) by implicating their accomplice.
If both suspects defect they will get a moderate sentence each (2 years). A suspect who defects will get immunity, and the other will get the full sentence (3 years). If neither defects—*i.e.*, they both cooperate (C) with each other—both will be charged for only a minor offence (1 year).

	d	c
D	$1, 1$	$3, 0$
C	$0, 3$	$2, 2$

The payoff is the *reduction* in the player's sentence: $3 - s$, where $s \in \{0, 1, 2, 3\}$ is the length of the sentence.

Cooperation in games

A \ B	d	c
D	$1,1$	$3,0$
C	$0,3$	$\boxed{2,2}$

- Individual rationalisation (dominance) suggests that they should both defect (Dd); however mutual cooperation (Cc) is better for both
- In games that aren't strictly competitive cooperation may be possible
- What's best for individuals (individual rationalisation) may not be best for the group, and vice versa
- Here play Cc gives each player a better payoff than the individually rationalisable play Dd

The Prisoner's Dilemma

Definition (Pareto optimality)

An outcome is *Pareto optimal* iff there is no other outcome which is at least as good or better for all the agents.

Pareto principle

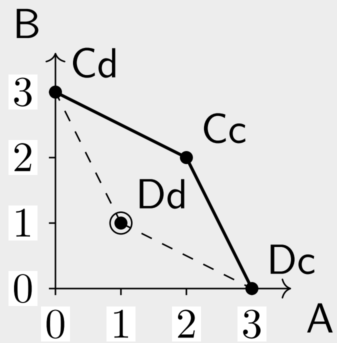
Pareto optimal outcomes are optimal for the group.

Consider the two-player *play diagram* on the right, where:

v_1 is the payoff to Prisoner 1

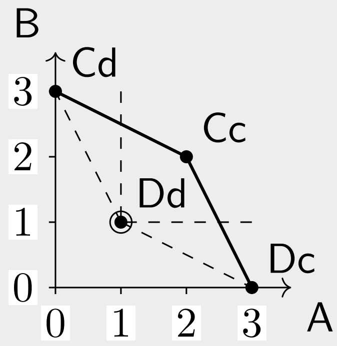
v_2 is the payoff to Prisoner 2

Pareto optimal outcomes represented by points on solid line



The Prisoner's Dilemma

A \ B	d	c
D	1, 1	3, 0
C	0, 3	2, 2

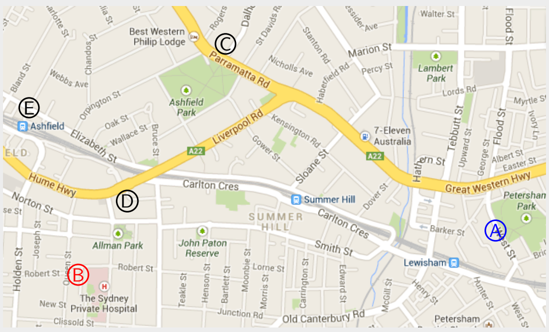


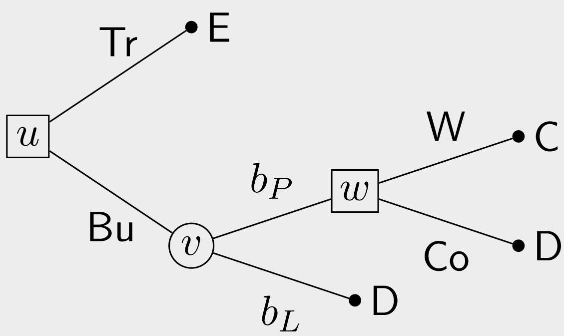
- The equilibrium is Dd (circled)
- The Pareto optimal outcomes are: Cc, Cd, Dc
- Play Cc, which is Pareto optimal, is better than Dd for both players

Conclusion

In two-player non strictly competitive games, what's best for the individual may not be best for the group; *i.e.*, *cooperation* preferable.

'Nature' as a player





- Single agent decisions can be regarded as games against a neutral player called 'Nature', or 'Chance', who has no preferences
- Game in which some of the players' preferences are unknown are said to have *incomplete informtation*—as opposed to imperfect information, in which information sets may have multiple nodes
- In extensive form, Nature's moves take place at chance nodes, and its moves correspond to chance events

Summary

- Best response strategies
- Equilibrium in games
- Rationalisation
- Group preference and Pareto optimality; cooperation
- Single agent decisions are ‘games against nature’