Lecture 5: Strategic reasoning

Simon Parsons

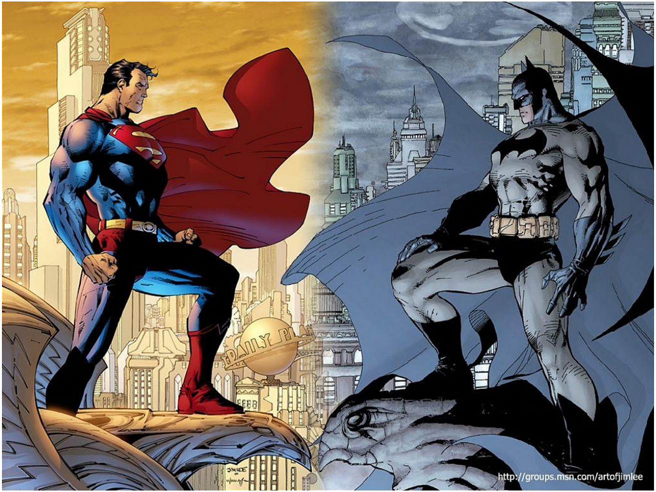
Department of Informatics
King's College London

(Version 1.2)

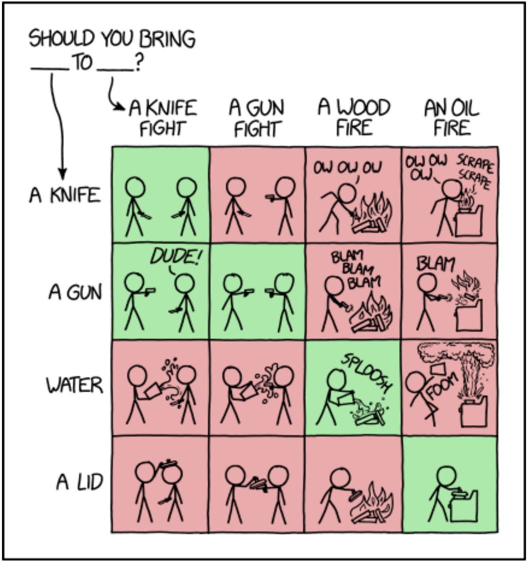
Today

- Introduction
- Probabilistic Reasoning I
- Probabilistic Reasoning II
- Sequential Decision Making
- Game Theory
- Temporal Probabilistic Reasoning
- Argumentation I
- Argumentation II
- (A peek at) Machine Learning
- AI & Ethics

Competition between agents



(groups.msn.com/artofjimlee)



(xkcd.com/1890/)

In the real world

(U.S. Army Military History Institute)

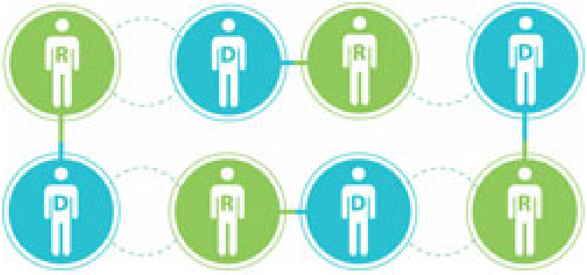
- Strategic analysis.

In the real world

(Adi Talwar/citylimits.org)

- School admissions

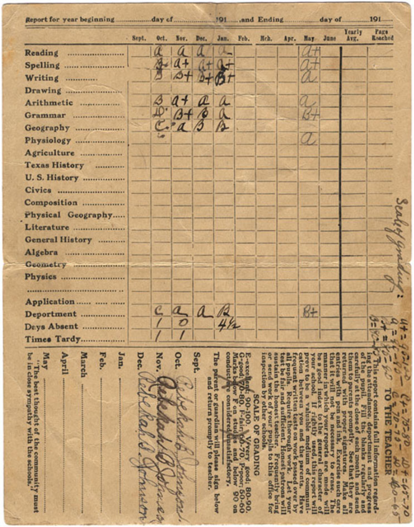
In the real world



(froedtert.com)

- Kidney exchange
- Longest chain involved 28 donors and recipients.

The grade game



(Lyndon Johnson's Report Card, US National Archives)

The grade game

- Take a piece of paper
- You are randomly paired with a partner (you do not know who!)
- You have to write X or Y on the piece of paper
- You will get a grade based on the following rules:
- If both you and your partner write X, then you both get a B
- If you write X and your partner writes Y then you get D and your partner gets A
- If you write Y and your partner writes X then you get A and your partner gets D
- If both you and your partner write Y then you both get C

  (The Grade Game, Ben Polak)

The grade game

- What would you do?

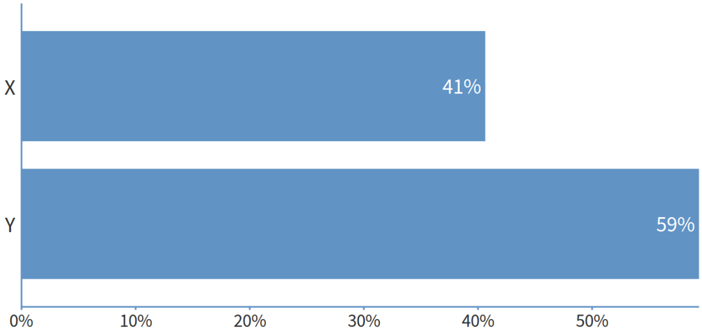
The grade game

- What would you do?
- What you get depends also on the choice of your partner.

The grade game

- What would you do?
- What you get depends also on the choice of your partner.
- This is the blueprint of strategic interaction.

The grade game



Choose the side

- Which side of the road to drive on?

(haulagetoday.com)

Choose the side

- Which side of the road to drive on?

(businessinsider.com)

Choose the side

- Which side of the road to drive on?

Any fule kno that.

(Geoffrey Williams/Ronald Searle)

Choose the side

- Which side of the road to drive on?

- Same side as everyone else

(berkshireeagle.com)

Choose the side

- How do you choose when you don't know what "everyone else" is doing.

Game theory

- Game theory is a framework for analysing interactions between a set of agents.
- Abstract specification of interactions.
- Describes each agent's preferences in terms of their utility.
    - Assume agents want to maximise utility.
- Give us a range of solution strategies with which we can make some predictions about how agents will/should interact.

Payoff Matrices

- We can characterise the “choose side” scenario in a payoff matrix

		j	
		left	right
i	left	1 1	0 0
	right	0 0	1 1

- Agent *i* is the row player
  gets the lower reward in a cell.
- Agent *j* is the column player
  gets the upper reward in a cell.

Payoff Matrices

- Actually there are two matrices here, one (call it A) that specifies the payoff to i and another B that specifies the payoff to j.
- Sometimes we'll write the payoff matrix as (A, B) in recognition of this.

Payoff Matrices

- We can characterise the grade game scenario in a payoff matrix

		j	
		Y	X
i	Y	2 2	1 4
	X	4 1	3 3

- Payoffs are the US grade points that correspond to the problem statement.
- A is 4, B is 3 etc.

Outcomes

- An outcome is what we get when we combine the actions of all the players.
- An outcome corresponds to an element of the payoff matrix

		j	
		left	right
i	up	1	0
		1	0
	down	0	1
		0	1

- We identify outcomes by the moves the players make:

$$(\textit{what } i \textit{ plays}, \textit{what } j \textit{ plays})$$

- Thus $(\textit{up}, \textit{right})$ identifies the outcome in which i plays *up* and j plays *right*

Solution Concepts

- How will a rational agent will behave in any given scenario?
- Play...
  - Dominant strategy;
  - Nash equilibrium strategy;
  - Pareto optimal strategies;
  - Strategies that maximise social welfare.

Dominant Strategies

- Given any particular strategy s (either C or D) agent i, there will be a number of possible outcomes.
- We say s_1 dominates s_2 if every outcome possible by i playing s_1 is preferred over every outcome possible by i playing s_2.
- Thus in this game:

		j: D	j: C
i	D	i: 1, j: 1	i: 1, j: 4
i	C	i: 4, j: 1	i: 4, j: 4

C dominates D for both players.

Dominant Strategies

- Two senses of “preferred”
- s_1 strongly dominates s_2 if the utility of every outcome possible by i playing s_1 is strictly greater than every outcome possible by i playing s_2.
- In other words, $u(s_1) > u(s_2)$, for all outcomes.
- s_1 weakly dominates s_2 if the utility of every outcome possible by i playing s_1 is no less than every outcome possible by i playing s_2.
- In other words, $u(s_1) \geqslant u(s_2)$, for all outcomes.

Dominant Strategies

- A rational agent will never play a dominated strategy.
- So in deciding what to do, we can delete dominated strategies.
- Unfortunately, there isn't always a unique undominated strategy.

Dominant Strategies

- Game with dominated strategies

	L	C	R
U	1 3	1 0	0 0
M	1 1	1 1	0 5
D	1 0	1 4	0 0

- Can eliminate the dominated strategies and simplify the game
- Remove R (dominated by L).

Dominant Strategies

- Game with dominated strategies

	L	C
U	3, 1	0, 1
M	1, 1	1, 1
D	0, 1	4, 1

- Can eliminate the dominated strategies and simplify the game
- Remove R (dominated by L).

Dominant Strategies

- If we are lucky, we can eliminate enough strategies so that the choice of action is obvious.

Dominant Strategies

- If we are lucky, we can eliminate enough strategies so that the choice of action is obvious.
- In general we aren't that lucky.

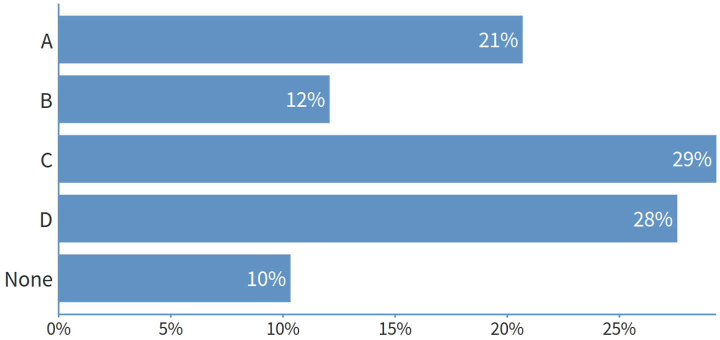
Dominant Strategies

- Consider this scenario:

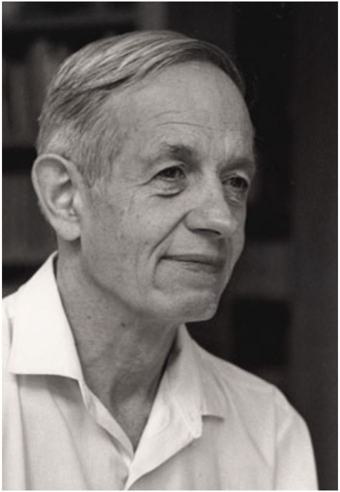
		j	
		C	D
i	A	1 2	4 3
	B	2 3	3 2

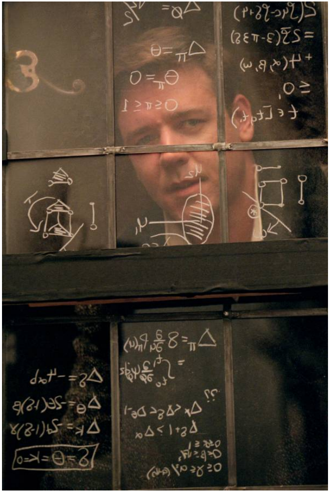
- Are there any dominated strategies?

Are there any dominated strategies?



Nash Equilibrium





John Forbes Nash.

(princeton.edu, Universal Pictures/DreamWorks)

Nash Equilibrium

- In general, we will say that two strategies s_1 and s_2 are in Nash equilibrium (NE) if:
  1. under the assumption that agent i plays s_1, agent j can do no better than play s_2; and
  2. under the assumption that agent j plays s_2, agent i can do no better than play s_1.
- Neither agent has any incentive to deviate from a NE.
- Eh?

KING'S College LONDON

Nash Equilibrium

- Let's consider the payoff matrix for the grade game:

		j	
		Y	X
i	Y	2 2	1 4
	X	4 1	3 3

- Here the Nash equilibrium is (Y, Y).
- If i assumes that j is playing Y, then i's best response is to play Y.
- Similarly for j.

Nash Equilibrium

- If two strategies are best responses to each other, then they are in Nash equilibrium.

Nash Equilibrium

- In a game like this you can find the NE by cycling through the outcomes, asking if either agent can improve its payoff by switching its strategy.

		j	
		Y	X
	Y	2	1
i		2	4
	X	4	3
		1	3

- Thus, for example, (X, Y) is not an NE because i can switch its payoff from 1 to 2 by switching from X to Y.

Nash Equilibrium

- More formally:

  A pair of strategies (i^*, j^*) is a Nash equilibrium solution to the game (A, B) if:

$$\begin{aligned} \forall i, a_{i*,j*} &\geqslant a_{i,j*} \\ \forall j, b_{i*,j*} &\geqslant b_{i*,j} \end{aligned}$$

- That is, (i^*, j^*) is a Nash equilibrium if:
  - If j plays j^*, then i^* gives the best outcome for i.
  - If i plays i^*, then j^* gives the best outcome for i.

Nash Equilibrium

- Unfortunately:
    1. Not every interaction scenario has a pure strategy NE.
    2. Some interaction scenarios have more than one NE.

Nash Equilibrium

- This game has two pure strategy NEs, (C, C) and (D, D):

		j	
		D	C
	D	5	1
i		3	2
	C	0	3
		2	3

- In both cases, a single agent can't unilaterally improve its payoff.

Nash Equilibrium

- This game has no pure strategy NE:

		j	
		D	C
	D	2	1
i		1	2
	C	0	1
		2	1

- For every outcome, one of the agents will improve its utility by switching its strategy.
- We can find a form of NE in such games, but we need to go beyond pure strategies.

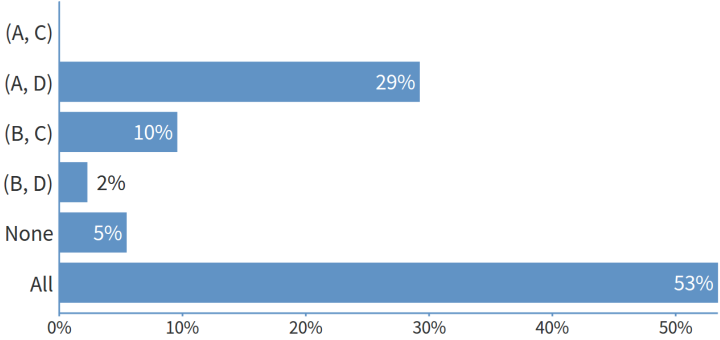
Nash equilibria?

- Consider this scenario (again):

		j	
		C	D
i	A	1 2	4 3
	B	2 3	3 2

- Are there any Nash equilibria?

Are there any Nash Equilibria?



Pareto Optimality

- An outcome is said to be Pareto optimal (or Pareto efficient) if there is no other outcome that makes one agent better off without making another agent worse off.
- If an outcome is Pareto optimal, then at least one agent will be reluctant to move away from it (because this agent will be worse off).
- If an outcome ω is not Pareto optimal, then there is another outcome ω' that makes everyone as happy, if not happier, than ω.

Pareto Optimality

- Can argue as follows:
- "Reasonable" agents would agree to move to ω' from ω if ω is not Pareto optimal and ω' is.
- Even if a given agent doesn't directly benefit from ω', others can benefit without it suffering.

(frugalentrepreneur.com)

Pareto Optimality

- This game has one Pareto efficient outcome, (D, D).

		j	
		D	C
	D	5	1
i		3	2
	C	0	0
		2	1

- There is no solution in which either agent does better.

Pareto Optimality

- This next game has two Pareto efficient outcomes, (C, D) and (D, C).

		j	
		D	C
	D	1	4
i		3	1
	C	1	1
		4	1

- Note that Pareto efficiency doesn't necessarily mean *fair*.
- Just that you can't move away and make one agent better off without making the other worse off.

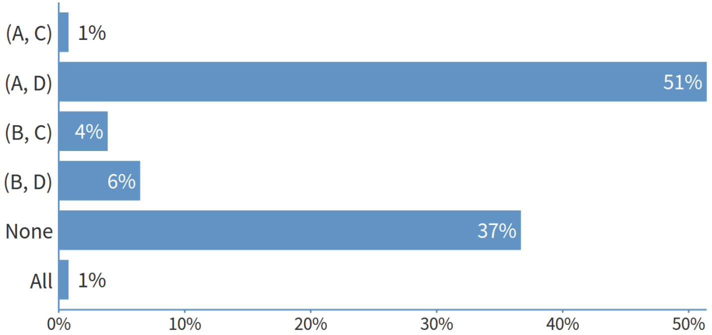
Pareto optimal?

- Consider this scenario (again):

		j	
		C	D
i	A	i: 2, j: 1	i: 3, j: 4
	B	i: 3, j: 2	i: 2, j: 3

- Are there any Pareto optimal outcomes?

Pareto optimal outcomes



Pareto Optimality

- Pareto optimality is a rather weak concept.

(coolfunpedia.blogspot.co.uk)

- What is the Pareto optimal way to divide a pile of money between A and B?

Social Welfare

- The social welfare of an outcome ω is the sum of the utilities that each agent gets from ω:

$$\sum_{i \in Ag} u_i(\omega)$$

- Think of it as the "total amount of money in the system".
- As a solution concept, may be appropriate when the whole system (all agents) has a single owner (then overall benefit of the system is important, not individuals).

Social Welfare

- As a solution concept it doesn't consider the benefits to individuals.

(telegraph.co.uk)

- A very skewed outcome can maximise social welfare.

- In both these games, (C, C) maximises social welfare.

		j	
		D	C
i	D	2 2	1 1
	C	3 3	4 4

		j	
		D	C
i	D	2 2	1 1
	C	3 3	9 0

Normal form games

(Pendleton Ward/Cartoon Network)

Normal form games

- An n-person, finite, normal form game is a tuple (N, A, u), where
  - N is a finite set of players.
  - $A = A_1 \times \ldots \times A_n$ where A_i is a finite set of actions available to i.
    Each $a = (a_1, \ldots, a_n) \in A$ is an action profile.
  - $u = (u_1, \ldots, u_n)$ where $u_i : A \mapsto \Re$ is a real-valued utility function for i.
- Naturally represented by an n-dimensional matrix

Strategies

- We analyze games in terms of strategies, that is what agents decide to do.
  - Combined with what the other agent(s) do(es) this jointly determines the payoff.
- An agent's strategy set is its set of available choices.
- Can just be the set of actions — pure strategies.
- We need more than just pure strategies in many cases.
  - Will discuss this later

Payoff matrix

- Here is the payoff matrix from the "choose which side" (of the road) game:

		j	
		left	right
i	left	1	0
		1	0
	right	0	1
		0	1

- We can classify games by the form of the payoff matrix.

Common payoff games

- "Choose which side" game

	left	right
left	1, 1	0, 0
right	0, 0	1, 1

  Also called the *coordination game*
- Any game with $u_i(a) = u_j(a)$ for all $a \in A_i \times A_j$ is a common payoff game.

Common payoff games

- The misanthropes' (un)coordination game:

	left	right
left	0, 0	1, 1
right	1, 1	0, 0

Here we try to avoid each other.

Misanthrope

http://www.thebadchemicals.com

Constant sum games

- Matching pennies

	heads	tails
heads	-1 (col), 1 (row)	1 (col), -1 (row)
tails	1 (col), -1 (row)	-1 (col), 1 (row)

- Any game with $u_i(a) + u_j(a) = c$ for all $a \in A_i \times A_j$ is a constant sum game.

Zero-sum games

- A particular category of constant sum games are zero-sum games.
- Where utilities sum to zero:

$$u_1(a_i) + u_j(\omega) = 0 \quad \text{for all } a \in A_i \times A_j$$

Zero-sum games

- Where preferences of agents are diametrically opposed we have strictly competitive scenarios.

(Library of Congress)

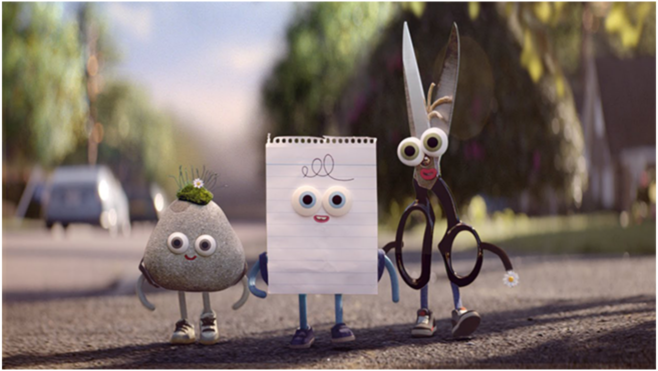
- Zero sum implies strictly competitive.

Zero-sum games

- Zero-sum encounters in real life are very rare . . . but people tend to act in many scenarios as if they were zero-sum.
- Most encounters have some room in the set of outcomes for agents to find (somewhat) mutually beneficial outcomes.

Zero-sum games

- Rock, paper, scissors:



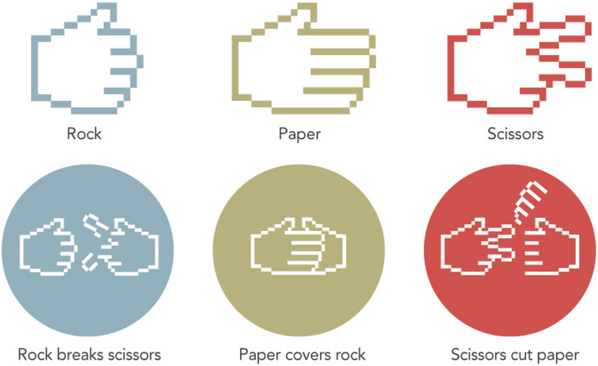
(Google/Droga5)

is another constant/zero sum game.

- Game in two senses.

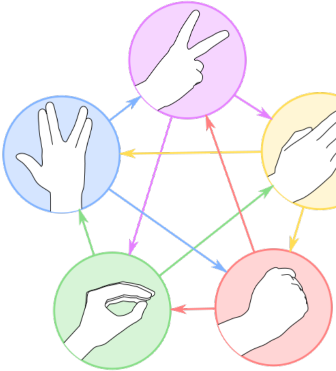
The rules

- Rules for “rock, paper, scissors”.



(eyemotive.com)

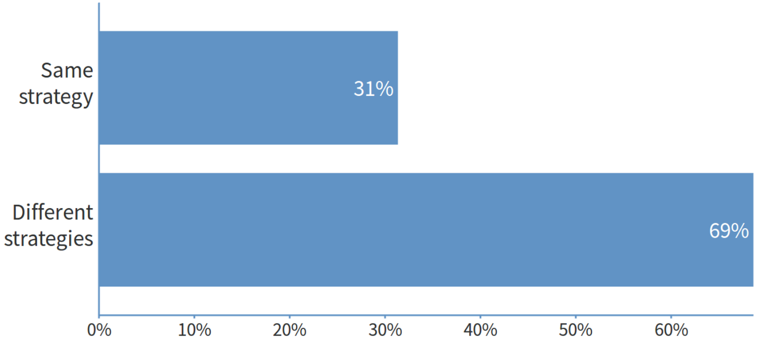
Rock, paper scissors, lizard, Spock



(DMacks/Nojhan/Wikipedia)

http://www.youtube.com/watch?v=x5Q6-wMx-K8

Rock, paper scissors



- How would you play?

Rock, paper, scissors

		j		
		rock	paper	scissors
	rock	0 0	1 -1	-1 1
i	paper	-1 1	0 0	1 -1
	scissors	1 -1	-1 1	0 0

Mixed strategy

- Chances are you would play a mixed strategy.
- You would:
  - sometimes play rock,
  - sometimes play paper; and
  - sometimes play scissors.
- A fixed/pure strategy is easy for an adaptive player to beat.

Mixed strategy

- A mixed strategy is just a probability distribution across a set of pure strategies.
- So, for a game where agent i has two actions a_1 and a_2, a mixed strategy for i is a probability distirbution:

$$MS_i = \{P(a_1), P(a_2)\}$$

- Given this mixed strategy, when i comes to play, they pick action a_1 with probability $P(a_1)$ and a_2 with probability $P(a_2)$.

Mixed strategy

- To determine the mixed strategy, *i* needs then to compute the best values of $P(a_1)$ and $P(a_2)$.
- These will be the values which give *i* the highest expected payoff given the options that *j* can choose and the joint payoffs that result.
- We could write down the expected payoffs of different mixed strategies and pick the one that optimises expected payoff.
- There is also a simple graphical method which works for very simple cases.
- Will look at this method next.

Mixed strategy

- Let's consider the payoff matrix:

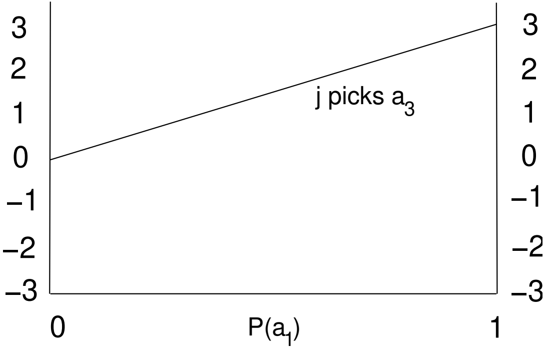
		j	
		a_3	a_4
i	a_1	3, -3	-1, 1
	a_2	0, 0	1, -1

- We want to compute mixed strategies to be used by the players.
- That means decide $P(a_1)$ and $P(a_2)$ etc.

Mixed strategy

- i's analysis of this game would be something like this.

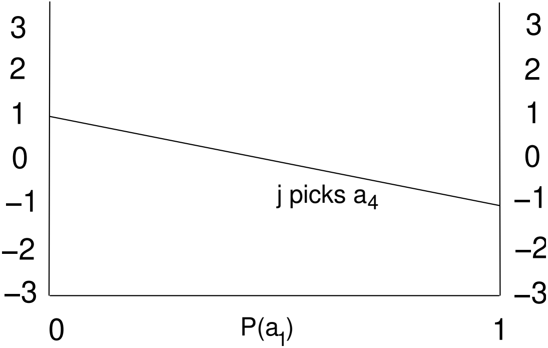
Mixed strategy



- If j picks a_3, i's payoff will be 3 or 0 depending on whether i picks a_1 or a_2.
- The expected payoff therefore varies along the line, as $P(a_1)$ varies from 0 to 1.

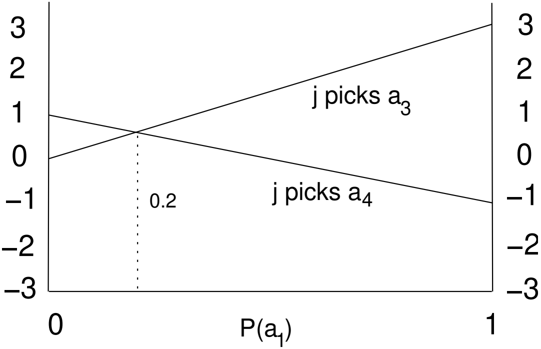
KING'S College LONDON

Mixed strategy



- If j picks a_4, i's payoff will be -1 or 1 depending on whether i picks a_1 or a_2.
- The expected payoff therefore varies along the line, as $P(a_1)$ varies from 0 to 1.

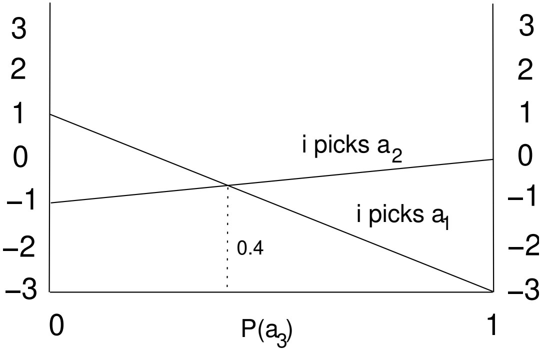
Mixed strategy



- Where the lines intersect, *i* has the same expected payoff whatever *j* does.
- This is a rational choice of mixed strategy.

Mixed strategy

- *j* can do the same kind of analysis:



Mixed strategy

- This analysis will help i and j choose a mixed strategy in zero-sum games.

(Archives of the Institute of Advanced Study, Princeton)

- This approach is due to von Neumann.

General sum games

- Battle of the Outmoded Gender Stereotypes
  - aka Battle of the Sexes

	this	that
this	1 2	0 0
that	0 0	2 1

(Time-Life/Getty)

- Game contains elements of cooperation and competition.
- The interplay between these is what makes general sum games interesting.

Negotiation

- Interplay between cooperation and competition leads to negotiation
- See, for example, the work of Sarit Kraus.

(law-train.eu)

Nash equilibrium

- Earlier we introduced the notion of Nash equilibrium as a solution concept for general sum games.
- (We didn't describe it in exactly those terms.)
- Looked at pure strategy Nash equilibrium.
- Issue was that not every game has a pure strategy Nash equilibrium.

Nash equilibrium

- For example:

		j	
		D	C
i	D	2 1	1 2
	C	0 2	1 1

- Has no pure strategy NE.

Nash equilibrium

- The notion of Nash equilibrium extends to mixed strategies.
- And every game has at least one mixed strategy Nash equilibrium.

Nash equilibrium

- For a game with payoff matrices A (to i) and B (to j), a mixed strategy (x^*, y^*) is a Nash equilibrium solution if:

$$\begin{aligned} \forall x, x^* A y^{*T} &\geqslant x A y^{*T} \\ \forall y, x^* B y^{*T} &\geqslant x^* B y^T \end{aligned}$$

- In other words, x^* gives a higher expected value to i than any other strategy when j plays y^*.
- Similarly, y^* gives a higher expected value to j than any other strategy when i plays x^*.

Nash equilibrium

- Unfortunately, this doesn't solve the problem of which Nash equilibrium you should play.

(Wolf Films/NBC)

The Prisoner's Dilemma

Two suspects are collectively charged with a crime and held in separate cells, with no way of meeting or communicating.
They are told that:

- *if one confesses and the other does not, the confessor will be freed, and the other will be jailed for three years;*
- *if both confess, then each will be jailed for two years.*

Both prisoners know that if neither confesses, then they will each be jailed for one year.

The Prisoner's Dilemma

- Payoff matrix for prisoner's dilemma:

		j defect	j coop
i	defect	i: 2, j: 2	i: 4, j: 1
i	coop	i: 1, j: 4	i: 3, j: 3

  "defect" = confess
  Numbers are payoffs to players, not years in jail.

- What should each agent do?

What Should You Do?

- The individually rational action is defect.
  This guarantees a payoff of no worse than 2, whereas cooperating guarantees a payoff of at most 1.
- So defection is the best response to all possible strategies: both agents defect, and get payoff = 2.
- But intuition says this is not the best outcome:
  Surely they should both cooperate and each get payoff of 3!
- This is why the PD game is interesting — the analysis seems to give us a paradoxical answer.

Solution Concepts

- Payoff matrix for prisoner's dilemma:

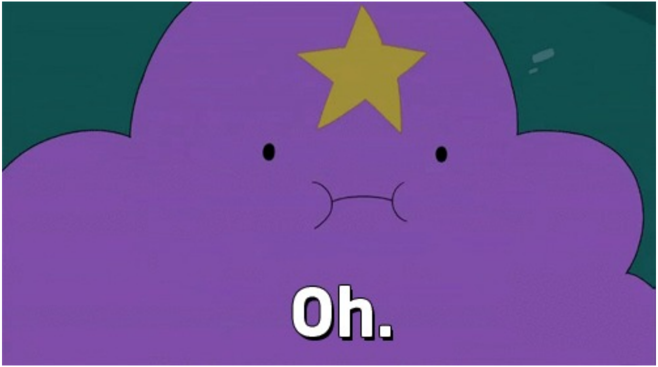
		j	
		defect	coop
i	defect	2 2	1 4
	coop	4 1	3 3

- (D, D) is the only Nash equilibrium.
- All outcomes except (D, D) are Pareto optimal.
- (C, C) maximises social welfare.

The Paradox

- The fact that the Nash Equilibrium is not the co-operative solution is problematic.
- This apparent paradox is the fundamental problem of multi-agent interactions.
- It appears to imply that cooperation will not occur in societies of self-interested agents.

Oh!



(Pendleton Ward/Cartoon Network)

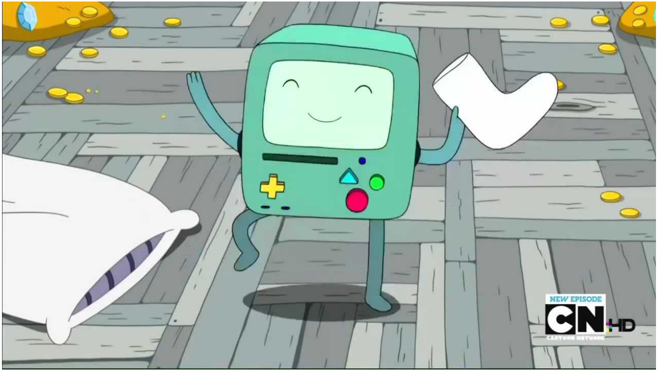
The Paradox

- Real world examples:
  - nuclear arms reduction/proliferation
  - free rider systems — public transport, file sharing;
  - in the UK — television licenses.
  - climate change — to reduce or not reduce emissions
  - doping in sport
  - resource depletion
  - Golden Balls
- The prisoner's dilemma is ubiquitous.
- Can we recover cooperation?

The Shadow of the Future

- *Play the game more than once*.
  If you know you will be meeting your opponent again, then the incentive to defect appears to evaporate.
  - If you defect, you can be punished (compared to the co-operation reward.)
  - If you get suckered, then what you lose can be amortised over the rest of the iterations, making it a small loss.
- Cooperation is (provably) the rational choice in the infinitely repeated prisoner's dilemma.

Yay!



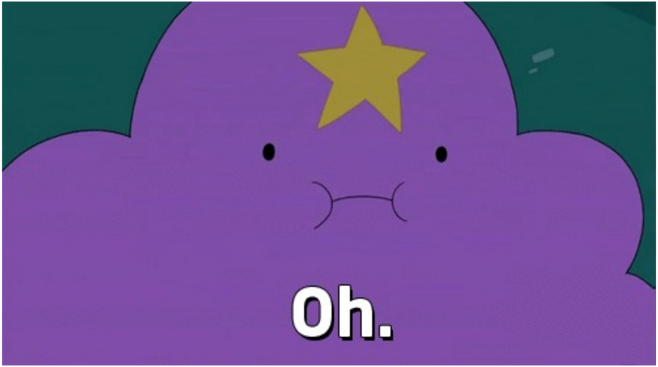
(Pendleton Ward/Cartoon Network)

- But what if there are a finite number of repetitions?

Backwards Induction

- Suppose you both know that you will play the game exactly n times.
  On round n, you have an incentive to defect, to gain that extra bit of payoff.
  But this makes round $n - 1$ the last "real" round, and so you have an incentive to defect there, too.
  This is the backwards induction problem.
- Playing the prisoner's dilemma with a fixed, finite, pre-determined, commonly known number of rounds, defection is the best strategy.
- That seems to suggest that you should never cooperate.

Oh!



(Pendleton Ward/Cartoon Network)

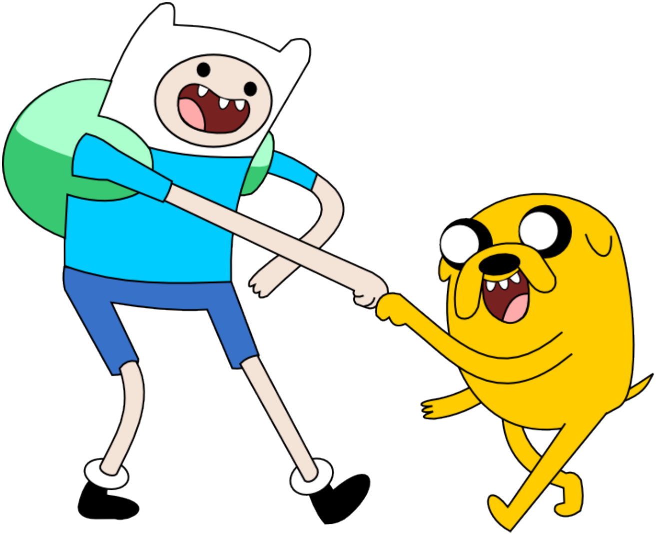
But

- So how does cooperation arise? Why does it make sense?
- After all, there does seem to be such a thing as society, and even in a big city like London, people don't behave so badly.
- Or, maybe more accurately, they don't behave badly all the time.

But

- Turns out that...
- As long as you have some probability of repeating the interaction, co-operation can have a better expected payoff.
- As long as there are enough co-operative folk out there, you can come out ahead by co-operating.

Mathematical!



Summary

- Have looked at strategic reasoning in the presence of other agents.
- Covered some ideas from game theory and discussed what it can do for us.
- Lots more we haven't covered...
- Game theory helps us to get a handle on some of the aspects of cooperation between self-interested agents.
- Rarely any definitive answers.
- Given human interactions, that should not surprise us.