Robotics and Autonomous Systems

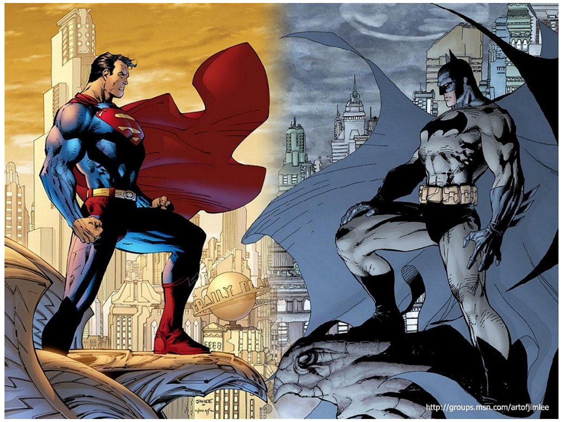
Lecture 27: More on self-interest

Richard Williams

Department of Computer Science
University of Liverpool

Today

- In the last lecture we started to look at competition between agents



- Today we look more into this.

Competition between agents


- Situation is more like this.

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.

Congestion Game

- Agents using TCP to communicate.
  - If packets collide, should back-off.
- Works if everyone does this.
- But what if agents could choose a defective implementation that doesn't back-off?
  - In a collision, their message would get sent quicker.
- But what if everyone did this?
  - Outcome depends on what other agents do.

Congestion Game

- Capture this as:

		i	
		defect	correct
j	defect	-3 (top right), -3 (bottom left)	-4 (top right), 0 (bottom left)
	correct	0 (top right), -4 (bottom left)	-1 (top right), -1 (bottom left)

- Agent *i* is the column player.
- Agent *j* is the row player.

Congestion Game

- Two obvious questions we can ask in this scenario:
  - What should an individual agent do?
  - How does the game get played — how do both agents together act?
- Game theory offers some ideas about how to answer these questions.

Congestion Game

- What should an individual agent do?
  - Depends on what the other agent does.
- How does the game get played — how do both agents together act?
  - Equilibrium.

Congestion Game

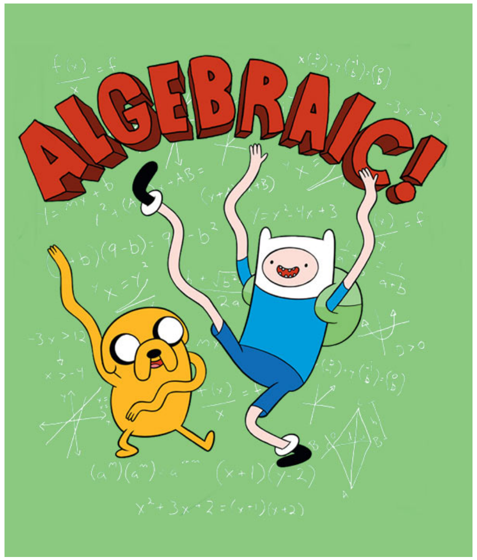
- As with all good games, the congestion game captures some underlying truths about the world at an abstract level:

- (Though you might want to alter the payoffs somewhat.)

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

Normal form games



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.
- In the congestion game, the set of pure strategies is:

  $\{correct, defect\}$
- 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:

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

Common payoff games

- In between is the El Farol bar problem:

- If everyone goes to the bar it is no fun, but if only some people go then everyone who goes has a good time.
  Should you go or not?

Constant sum games

- Matching pennies

	heads	tails
heads	-1 1	1 -1
tails	1 -1	-1 1

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

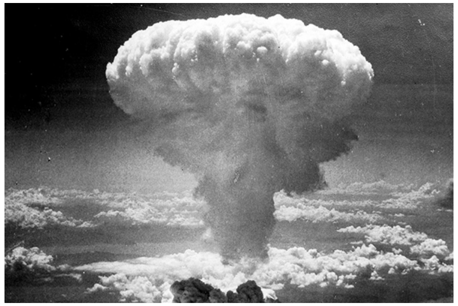
Zero-sum games

- 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.
- Zero sum implies strictly competitive.



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

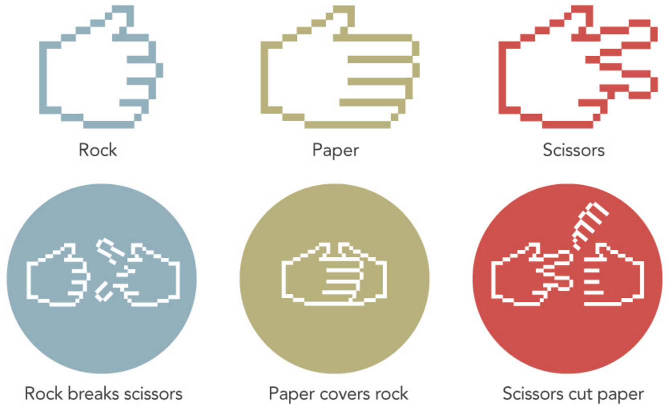
Zero-sum games

- Rock, paper, scissors:

is another constant/zero sum game.
- Game in two senses.

The rules

- Rules for “rock, paper, scissors”.



Rock, paper, scissors

- How would you play?

i

j		rock	paper	scissors
	rock	0 0	1 -1	-1 1
j	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.
- More formally, for a game with two actions a_1 and a_2, i picks a vector of probabilities:

$$x = (x_1, x_2)$$

where

$$\sum_k x_k = 1$$

and

$$x_k \geqslant 0$$

- i then picks action a_1 with probability x_1 and a_2 with probability a_2.

Mixed strategy

- To determine the mixed strategy, *i* needs then to compute the best values of x_1 and x_2.
- These will be the values which give *i* the highest payoff given the options that *j* can choose and the joint payoffs that result.
- In the next lecture we will get into the computation of expected values, which is one way to analyse this.
- But for now we will look at a simple graphical method
  - Only works for very simple cases.

Mixed strategy

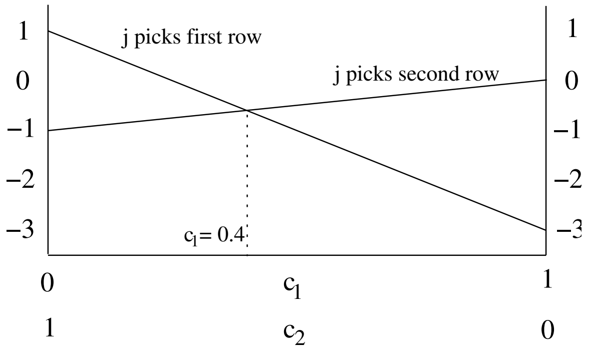
- Let's consider the payoff matrix:

j \ *i*	a_1	a_2
a_1	-3 3	1 -1
a_2	0 0	-1 1

and compute mixed strategies to be used by the players.

Mixed strategy

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



- *i* picks the probability of a_1 so that it is indifferent to what *j* picks.

Mixed strategy

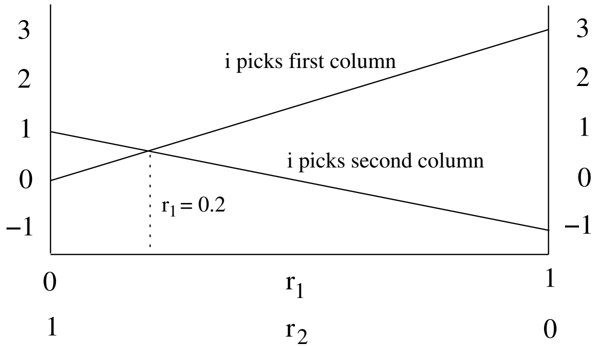
- *j* can analyse the problem in terms of a probability vector

$$y = (y_1, y_2)$$

and come up with a similar picture.

Mixed strategy

- *j*'s analysis would be something like this:



Mixed strategy

- This analysis will help *i* and *j* choose a mixed strategy for the specific case in which the payoffs to the two agents are equal and opposite for each outcome.

- Application to zero sum games 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

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

Nash equilibrium

- Last time 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:

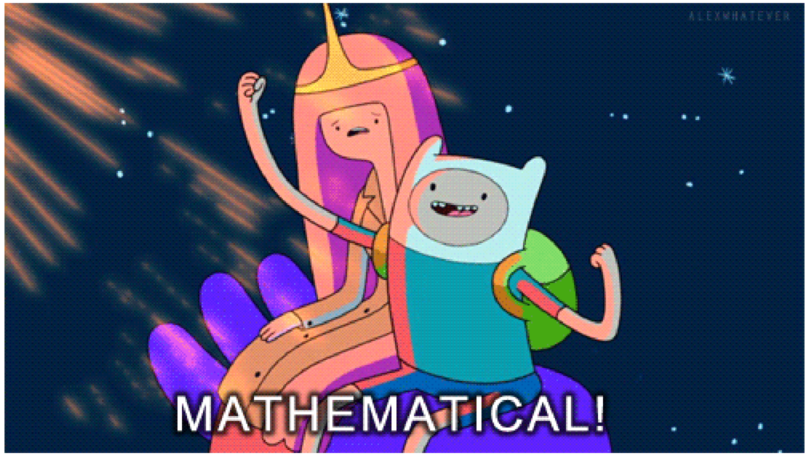
		i	
		D	C
j	D	2 1	1 2
	C	0 2	1 1

- 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



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.

The Prisoner's Dilemma

The Prisoner's Dilemma

Two men 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:

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

- 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.

What Did You Do?

- The Prisoner's Dilemma is the same game as the "grade game".
  Just has a different back story.
- When you played that,
  18 of you chose "defect".
  6 of you chose "cooperate".

Solution Concepts

- Payoff matrix for prisoner's dilemma:

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

- There is no dominant strategy (in the game theory sense).
- (D, D) is the only Nash equilibrium.
- All outcomes except (D, D) are Pareto optimal.
- (C, C) maximises social welfare.

The Paradox

- 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.
- 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
- 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.
  (Hurrah!)
- 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 $n - 1$, you have an incentive to defect, to gain that extra bit of payoff.
  But this makes round $n - 2$ the last "real", 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.

Agh!

- That seems to suggest that you should never cooperate.
- 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 New York, people don't behave so badly.
  Or, maybe more accurately, they don't behave badly all the time.
- 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.
- But is always co-operating the best approach?

Axelrod's Tournament

- Suppose you play iterated prisoner's dilemma (IPD) against a range of opponents.
- What approach should you choose, so as to maximise your overall payoff?
- Is it better to defect, and hope to find suckers to rip-off?
- Or is it better to cooperate, and try to find other friendly folk to cooperate with?

Axelrod's Tournament

- Robert Axelrod (1984) investigated this problem.
- He ran a computer tournament for programs playing the iterated prisoner's dilemma.
- Axelrod hosted the tournament and various researchers sent in approaches for playing the game.

Example Strategies in Axelrod's Tournament

- ALLD:
  "Always defect" — the hawk strategy;
- TIT-FOR-TAT:
  1. On round $u = 0$, cooperate.
  2. On round $u > 0$, do what your opponent did on round $u - 1$.
- TESTER:
  On 1st round, defect. If the opponent retaliated, then play TIT-FOR-TAT. Otherwise intersperse cooperation & defection.
- JOSS:
  As TIT-FOR-TAT, except periodically defect.

Axelrod's Tournament

- Surprisingly TIT-FOR-TAT for won.
- But don't read too much into this.
  - Turns out that TIT-FOR-TWO-TATS would have done better.
- In scenarios like the IPD tournament, the best approach depends heavily on what the full set of approaches is.
- TIT-FOR-TAT did well because there were other players it could co-operate with.
  - In scenarios with different strategy mixes it would not win.
- Suggests that there is some value in cooperating, at least some of the time.

Recipes for Success

Axelrod suggests the following rules for succeeding in his tournament:

- Don't be envious:
  Don't play as if it were zero sum!
- Be nice:
  Start by cooperating, and reciprocate cooperation.
- Retaliate appropriately:
  Always punish defection immediately, but use "measured" force — don't overdo it.
- Don't hold grudges:
  Always reciprocate cooperation immediately.

Summary

- Have looked a bit further at game theory and 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.