MULTIAGENT DECISION MAKING

Introduction

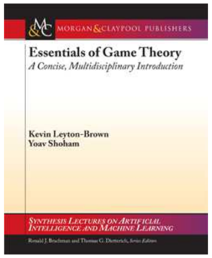
- Last time we looked at how an agent can make sequential decisions
  - Where outcome/utility depends on a sequence of decisions/actions.
- All under the control of one agent.
  - Static world
- What if we have more than one agent?
  - Utility depends on what all the agents do.
- This is the domain of *game theory*

What is 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.
- Game theory is *not* about being selfish.

Book?

- Textbook doesn't say much about game theory.
- Instead:



ISBN 978-159-829-5931

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 -3	-4 0
	correct	0 -4	-1 -1

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

- 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 qestions.

- 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.

- 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

The Prisoner's Dilemma

(Okay, so if you know the movie, this isn't a 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.

- Payoff matrix for prisoner's dilemma:

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

- What should each agent do?

- Payoff matrix for prisoner's dilemma:

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

- Well?

- In fact:

		i: defect	i: coop
j	defect	d	b
		d	c
	coop	c	a
		b	a

- any game with $c > a > d > b$ is a prisoner's dilemma.

Common payoff games

- Coordination game

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

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

- The misanthropes' (un)coordination game:

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

- 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

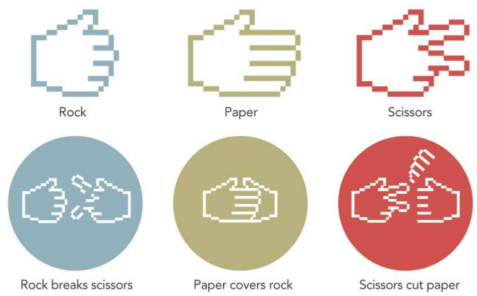
	heads	tails
heads	-1	1
	1	-1
tails	1	-1
	-1	1

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

- Rock, paper, scissors:

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

- Rules for “rock, paper, scissors”.



- As a normal form game:

	rock	paper	scissors
rock	0 0	1 -1	-1 1
paper	-1 1	0 0	1 -1
scissors	1 -1	-1 1	0 0

- This game is “zero-sum” since the utilities sum to zero.

- Zero sum games are pure competition:
  - If one player wins, the other loses.
- Few real situations other than games in the recreational sense are really zero sum.

General sum games

- 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.

Strategies

- We analyze games in terms of *strategies*, that is what they 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 Prisoner's Dilemma (PD), the set of pure strategies is:

  cooperate defect
- We need more than just pure strategies in many cases.

- *Mixed strategies* are probability distributions over pure strategies.
- In the PD, one mixed strategy is:

  pick cooperate with probability 0.7, pick defect with probability 0.3
- Clearly there are many possible pure strategies.
  - Strategy set S_i is set of all probability distributions over A_i. (Actions may vary between agents).
- For the PD, a strategy set of an agent is all possible probability distributions over cooperate and defect.
- A *profile* is a selection of strategies, one for each agent.
- A set of *mixed strategy profiles* is $S_1 \times \ldots \times S_n$.

- The *support* for a mixed strategy s_i is the set of pure strategies $\{a_i | s(a_i) > 0\}$
  - All the actions that might be picked by a particular mixed strategy.
- The payoff of a mixed strategy is the expected utility of the strategy:

$$u_i(s) = \sum_{a \in A} u_i(a) \prod_{j=1}^{n} s_j(a_j)$$

Solution concepts

- For an agent acting alone we can compute the *optimal strategy*
  - maximises the expected utility.
- In a multiagent setting this isn't very meaningful.
- Best strategy depends on what others are doing.
- *Solution concepts* identify sets of outcomes (subsets of the whole) that are interesting in some way.
- External view — Pareto optimality.
- Internal view — Nash equilibrium.

Pareto optimality

- In multiagent settings it is hard to define "best solution".
  - Can't easily handle tradeoffs between agents' utilities.
- Which is best outcome in the battle of the sexes?

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

- If your children were playing, which outcome would you like to see?

- Though we can't say which outcome is best, we can say that some outcomes are better than others.
- s *Pareto dominates* s' if for all i, $u_i(s) \geq u_i(s')$ and there is some j such that $u_j(s) > u_j(s')$.
- Defines a partial order over strategies.
- s is *Pareto optimal* if there is no s' such that s' Pareto dominates s.
- "Pareto optimal" is also described as "strictly Pareto efficient".

- Which outcome(s) is/are Pareto dominant in the battle of the sexes?

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

- How about for the Prisoner's Dilemma:

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

- Pareto optimality is a rather weak concept.

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

Nash equilibrium

- Pareto dominance doesn't tell an individual agent how to play the game.
- Nash equilibrium is more useful from this perspective.
- If I know how you will play the game, I can maximise. I choose my *best response*.
- i's best response to the strategy profile s_{-i} is the mixed strategy $s_i^* \in S$ such that $u_i(s_i^*, s_{-i}) \geq u_i(s_i, s_{-i})$ for all s_i.

- Best response is not a solution concept since we don't, in general, know what other agents will do.
- But we build the idea of *Nash equilibrium* (NE) on top of it.
- A strategy profile $s = (s_1, \ldots, s_n)$ is a Nash equilibrium if, for all agents i, s_i is a best response to s_{-i}.
- NE is stable, since no agent can do better by switching strategy while everyone else sticks.
- Every game (within reason) has a (mixed strategy) Nash equilibrium.

- What is the Nash equilibrium for the Prisoner's Dilemma:

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

- What about for the battle of the sexes?

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

Dominated strategies

- Let s_i and s_i' are strategies of i. S_{-i} is the set of strategy profiles of the other players.
- s_i *strictly dominates* s_i' if $u(s_i, s_{-i}) > u(s_i', s_{-i})$ for all $s_{-i} \in S_{-i}$.
- s_i *weakly dominates* s_i' if $u(s_i, s_{-i}) \geq u(s_i', s_{-i})$ for all $s_{-i} \in S_{-i}$ and $u(s_i, s_{-i}) > u(s_i', s_{-i})$ for at least one s_{-i}
- s_i *very weakly dominates* s_i' if $u(s_i, s_{-i}) \geq u(s_i', s_{-i})$ for all $s_{-i} \in S_{-i}$
- A *dominant* strategy is one that dominates all others.
- A strategy profile in which every s_i is dominant for i is a Nash equilibrium
  - equilibrium in dominant strategies

Aside — Prisoner's Dilemma redux

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

- "defect" is a dominant strategy.

Dominated strategies II

- Game with dominated strategies

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

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

Dominated strategies III

- Game with dominated strategies

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

- M is now dominated by the mixed strategy that picks U and L with equal probability.
- It was not dominated before we removed R.

Dominated strategies IV

- Final game

	L	C
U	1 3	1 0
L	1 0	1 4

- Removing dominated strategies will not remove any Nash equilibria.
- If we only use strict dominance, the order of elimination doesn't matter.

Computing equilibria

- If we can eliminate enough dominated strategies, we can easily determine the NE.
- Not always possible.
- When no dominant strategy, computationally difficulty to establish the NE.
  - NP hard
  - Basically a search through a huge space of possible mixed strategies
- How else can we decide what strategy to use?

- We can do it by learning.
- Agent *A* observes agent *B*'s choices, and formulates the best response to them.
- If *B* i splaying a mixed strategy, this will correctly identify the best response.
  - However, may be susceptible to manipulation.
- What if both agents are learning?
- Turns out that under some conditions, agents can work their way to the NE.

Evolutionarily Stable Strategies

- Consider a large population of agents playing a two player game.
  - Equilibrium strategy
- Is the equilibrium strategy stable against some fraction of the population switching to a different strategy.
- A mixed strategy s is an *evolutionarily stable strategy* if for all other strategies s':
  - $u(s, s) > u(s', s)$; or
  - $u(s, s) = u(s', s)$ and $u(s, s') = u(s', s')$

- Hawk/Dove game

	hawk	dove
hawk	-2, -2	6, 0
dove	0, 6	3, 3

- Unique symmetric Nash equilibrium, $(3/5, 2/5)$.
- Also the unique ESS.
- But, for example, $(dove, dove)$ is not an ESS, though it is Pareto optimal.

- If a mixed strategy s is an evolutionarily stable strategy, then it is a Nash equilibrium.
- Any ESS is a best response to itself, and is therefore an NE.
- The reverse does not hold — only strict Nash equilibria are ESS.
- In a two-player game, given a mixed strategy s, if (s, s) is strict Nash equilibrium, then s is an evolutionarily stable strategy.
- Interesting because we can *learn* ESS and hence NE.

- There are several sets of *replicator dynamics* which explain how agents should change their strategy based on how strategies payoff in the current mix.
- If agents adjust their strategy this way, the population will converge to NE.
- Can look at this as either:
  - the behavior of a large population; or
  - a computational process for establishing the mixed strategy NE.
- A population of 100 agents where the ESS has 10 agents using strategy s_1, 20 using s_2 and 70 using s_3 tells us that $\langle(s_1, 0.1), (s_2, 0.2), (s_3, 0.7)\rangle$ is an NE.

Iterated Games

- All we did so far assumes that agents only play the game with each other once.
  - One-shot games
- Even when learning, the assumption is that agents are trying to establish how an opponent plays the one-shot game.
  - Either playing an anonymous crowd of opponents; or
  - Observing an agent who is doing this.
- Different behavior when you know that you will play the same agent multiple times.
- Shadow of the future.

Iterated Prisoner's dilemma

- 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 infinititely repeated prisoner's dilemma.*
- 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.

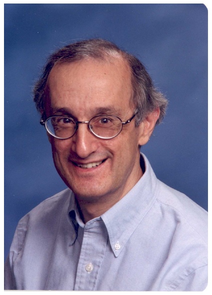
  This is the *backwards induction* problem.
- Playing the prisoner's dilemma with a fixed, finite, pre-determined, commonly known number of rounds, defection is the best strategy.

- 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?



- Robert Axelrod (1984) investigated this problem, with 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.

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.

- 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.

Recipes for Success in Axelrod's Tournament

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.

Sequential Games

- In normal-form games we assume moves are simultaneous.
- Another area of game theory studies *sequential* games.
  - Players take it in turns
- We don't have time to look at this.
- Can always map the sequence of moves into a strategy, and consider this to be a very big normal form game.

Markov Games

- Can think of Markov games as a hybrid of MDPs and games.
- In MDPs, choices about action determine the next state an agent is in.
  - Actions can vary with state
- In games, joint choices of actions determine payoffs.
- In Markov games, joint choices of action determine the next state and agent is in.
  - The game that defines outcomes can vary with state.
- Markov games are very much at the cutting edge of research.

Bayesian Games

- Everything we have done so far assumes agents know what game they are playing.
- Assume that:
  - Number of players
  - Set of actions
  - Payoffs

  are common knowledge across all players.
- Now look at games of *incomplete information* or *Bayesian* games.
- Represent the lack of knowledge with a probability distribution over a set of games
  - Agents' beliefs about which game they are playing.

- All these games have the same number of players and strategy space.
  - Not a very restrictive assumption.
  - Pad games if necessary with dominated strategies.
- Agents' beliefs are posteriors, based on a common prior conditioned on private signals.
  - Start the same, experience differs.

- Bayesian game over some familiar games

MP

2, 0	0, 2
0, 2	2, 0

$p = 0.3$

PD

2, 2	0, 3
3, 0	1, 1

$p = 0.1$

Coord

2, 2	0, 0
0, 0	1, 1

$p = 0.2$

BoS

2, 1	0, 0
0, 0	1, 2

$p = 0.4$

- Row player can only distinguish between (MP, PD) and $(Coord, Bos)$.

- To decide what to do, the row player has to look at his possible choices and use the probability distribution over the possible outcomes to determine his expected payoff.
- Does the same to establish expected payoffs of the other agent.
- Then apply the usual kinds of analysis, suitably complicated by the expected value calculations.
- Mixed strategies add another layer of expectation.

Summary

- Today we looked at some game theory.
- In artificial intelligence, game theory is used to help analyze interactions between agents.
- However, it is (has been) an area of growing interest and importance in computer science as a whole.
- Lots of work on figuring out incentives to make systems work as we want them to.
  - Mechanism design.
- Lots more we could say about multiagent decision-making also:
  - Voting
  - Social choice theory