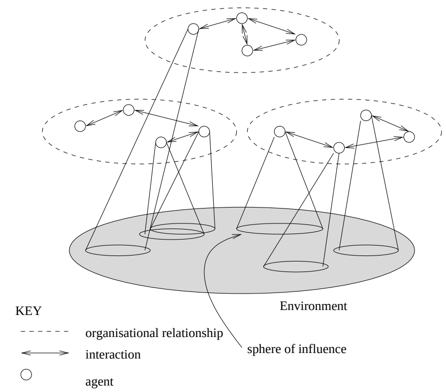
LECTURE 7: MULTIAGENT INTERACTIONS

An Introduction to Multiagent Systems

CIS 716.5, Spring 2010

What are Multiagent Systems?



Thus a multiagent system contains a number of agents . . .

- . . . which interact through communication . . .
- . . . are able to act in an environment . . .
- . . . have different “spheres of influence” (which may coincide). . .
- . . . will be linked by other (organisational) relationships.

Utilities and Preferences

- Assume we have just two agents: $Ag = \{i, j\}$.
- Agents are assumed to be *self-interested*: they *have preferences over how the environment is*.
- Assume $\Omega = \{\omega_1, \omega_2, \ldots\}$ is the set of “outcomes” that agents have preferences over.
- We capture preferences by *utility functions*:

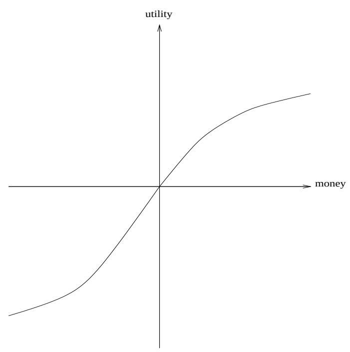
$$u_i : \Omega \rightarrow I\!R$$
$$u_j : \Omega \rightarrow I\!R$$

- Utility functions lead to *preference orderings* over outcomes:

$$\omega \succeq_i \omega' \quad \text{means} \quad u_i(\omega) \geq u_i(\omega')$$
$$\omega \succ_i \omega' \quad \text{means} \quad u_i(\omega) > u_i(\omega')$$

What is Utility?

- Typical relationship between utility & money:



- Utility is *not* money.

Multiagent Encounters

- We need a model of the environment in which these agents will act. . .
  - agents simultaneously choose an action to perform, and as a result of the actions they select, an outcome in Ω will result;
  - the *actual* outcome depends on the *combination* of actions;
  - assume each agent has just two possible actions that it can perform C (“cooperate”) and “D” (“defect”).
- Environment behaviour given by *state transformer function*:

$$\tau : \underbrace{Ac}_{\text{agent } i\text{'s action}} \times \underbrace{Ac}_{\text{agent } j\text{'s action}} \rightarrow \Omega$$

- Here is a state transformer function:

$$\tau(D,D) = \omega_1 \quad \tau(D,C) = \omega_2 \quad \tau(C,D) = \omega_3 \quad \tau(C,C) = \omega_4$$

(This environment is sensitive to actions of both agents.)

- Here is another:

$$\tau(D,D) = \omega_1 \quad \tau(D,C) = \omega_1 \quad \tau(C,D) = \omega_1 \quad \tau(C,C) = \omega_1$$

(Neither agent has any influence in this environment.)

- And here is another:

$$\tau(D,D) = \omega_1 \quad \tau(D,C) = \omega_2 \quad \tau(C,D) = \omega_1 \quad \tau(C,C) = \omega_2$$

(This environment is controlled by j.)

Rational Action

- Suppose we have the case where *both* agents can influence the outcome, and they have utility functions as follows:

$$\begin{array}{llll} u_i(\omega_1) = 1 & u_i(\omega_2) = 1 & u_i(\omega_3) = 4 & u_i(\omega_4) = 4 \\ u_j(\omega_1) = 1 & u_j(\omega_2) = 4 & u_j(\omega_3) = 1 & u_j(\omega_4) = 4 \end{array}$$

- With a bit of abuse of notation:

$$\begin{array}{llll} u_i(D,D) = 1 & u_i(D,C) = 1 & u_i(C,D) = 4 & u_i(C,C) = 4 \\ u_j(D,D) = 1 & u_j(D,C) = 4 & u_j(C,D) = 1 & u_j(C,C) = 4 \end{array}$$

- Then agent i's preferences are:

$$C,C \succeq_i C,D \quad \succ_i \quad D,C \succeq_i D,D$$

- In this case, what should i do?

- i prefers all outcomes that arise through C over all outcomes that arise through D.

$$C,C \succeq_i C,D \quad \succ_i \quad D,C \succeq_i D,D$$

- Thus C is the *rational choice* for i.

Payoff Matrices

- We can characterise the previous scenario in a *payoff matrix*

		i	
		defect	coop
j	defect	1 1	4 1
	coop	1 4	4 4

- Agent i is the *column player* and gets the upper reward in a cell.
- Agent j is the *row player* and gets the lower reward in a cell.
- 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 game as (A, B) in recognition of this.

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

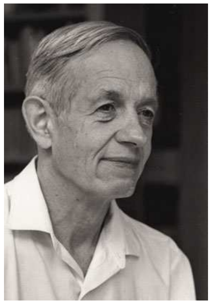
- Given any particular strategy s (either C or D) agent i, there will be a number of possible outcomes.
- We say s_1 *dominates* s_2 if every outcome possible by i playing s_1 is preferred over every outcome possible by i playing s_2.
- Thus in this game:

		i	
		defect	coop
j	defect	1 1	4 1
	coop	1 4	4 4

C dominates D for both players.

- 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.

Nash Equilibrium



John Forbes Nash.

- 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.*

- More formally:

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

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

- Unfortunately:
  1. *Not every interaction scenario has a* pure strategy *NE.*
  2. *Some interaction scenarios have more than one NE.*
- So we need more than just the notion of pure strategies.

Mixed-strategy Nash Equilibrium

- Consider the game "rock, paper, scissors".

- Play a few rounds.

- How did you play?

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

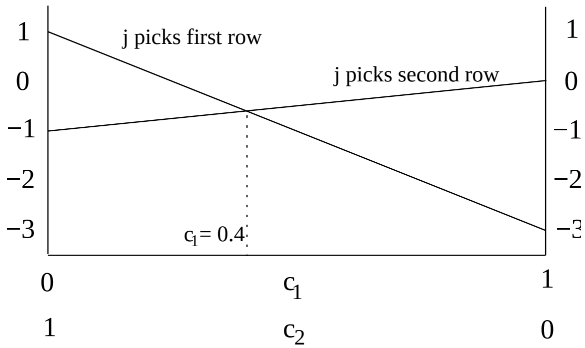
- Chances are you used a *mixed* strategy. You sometimes played rock, sometimes paper and sometimes scissors.

- A mixed strategy is just a probability distribution across a set of pure strategies.
- More formally, i picks a vector of probabilities:
$$x = (x_1, x_2)$$
where
$$\sum_u x_u = 1$$
and
$$x_u \geq 0$$
- i then picks strategy u with probability x_u.
- To determine the 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 expected payoff for its mixed strategy.

- Let's consider the payoff matrix:

		i	
		defect	coop
j	defect	-3 3	1 -1
	coop	0 0	- 1 1

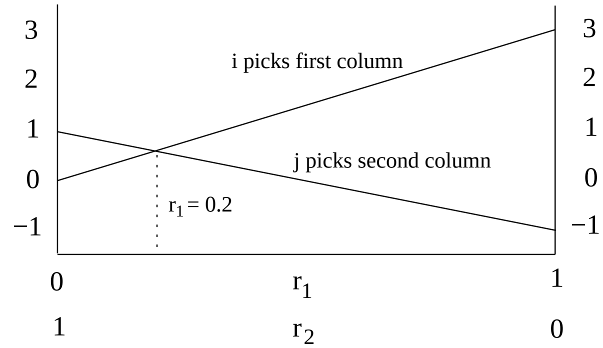
- *i*'s analysis would be something like this:



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

$$y = (y_1, y_2)$$

and come up with a similar picture:



- Now, let's consider the payoff's the players will expect.
- With i having mixed strategy (x_1, x_2) and j having (y_1, y_2), the value of the game to j will be:

$$\begin{aligned} V &= 3x_1y_1 + 0(1 - x_1)y_1 \\ &\quad - x_1(1 - y_1) + (1 - x_1)(1 - y_1) \\ &= 5x_1y_1 - y_1 - 2x_1 + 1 \end{aligned}$$

- Now, let's assume that j uses $y_1^* = 0.4$ as calculated above. Then:

$$\begin{aligned} V &= 5(0.4x_1) - 0.4 - 2(x_1) + 1 \\ &= 0.6 \end{aligned}$$

- Similarly, if i picks $x_1^* = 0.2$ then:

$$V = 0.6$$

- The neat thing is that the expected value for one player does not depend upon the strategy of the other player.
- Thus a mixed strategy is one which works whatever the other player does.
- Thus it doesn't matter if you know what the other players' strategy is, or if they know what your strategy is.

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

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

- In other words, x^* gives a higher *expected* value to P1 than any other strategy when P2 plays y^*.
- Similarly, y^* gives a higher *expected* value to P2 than any other strategy when P1 plays x^*.

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.
  "Reasonable" agents would agree to move to ω' in this case. (Even if I don't directly benefit from ω', you can benefit without me suffering.)

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).

Competitive and Zero-Sum Interactions

- Where preferences of agents are diametrically opposed we have *strictly competitive* scenarios.
- Zero-sum encounters are those where utilities sum to zero:

$$u_i(\omega) + u_j(\omega) = 0 \quad \text{for all } \omega \in \Omega.$$

- 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.

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.

- Payoff matrix for prisoner's dilemma:

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

- Top left: If both defect, then both get punishment for mutual defection.
- Top right: If i cooperates and j defects, i gets sucker's payoff of 1, while j gets 4.
- Bottom left: If j cooperates and i defects, j gets sucker's payoff of 1, while i gets 4.
- Bottom right: Reward for mutual cooperation.

What Should You Do?

- The *individual 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!

Solution Concepts

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

- 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 ("why don't I keep mine. . . ")
  - free rider systems — public transport;
  - in the UK — television licenses.
- The prisoner's dilemma is *ubiquitous*.
- Can we recover cooperation?

Arguments for Recovering Cooperation

- Conclusions that some have drawn from this analysis:
  - the game theory notion of rational action is wrong!
  - somehow the dilemma is being formulated wrongly
- Arguments to recover cooperation:
  - We are not all Machiavelli!
  - The other prisoner is my twin!
  - The shadow of the future. . .

We are not all Machiavelli

- "We aren't all that hard-boiled, and besides, people really do act altruistically."
- Well maybe.
- Lots of "altruism" is something else:
  - Either there is some delayed reciprocity; or
  - There are mechanisms to punish defection.
- There is a reason why the IRS audits people's taxes after all :-)
- Altruism may be something that makes us feel good, so we are prepared to pay for it.

  "It's ok to be a sucker when you can afford it"

  (Doc, Cannery Row).

The Other Prisoner is My Twin

- Argue that both prisoner's will think alike and decide that it is best to cooperate.
- If they are twins, they must think along the same lines, right?
- Well, if this is the case, we aren't really playing the Prisoner's Dilemma.
- Possibly more to the point is that if you know the other person is going to cooperate, you are *still* better off defecting.

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.
- *Cooperation is the rational choice in the infinititely repeated prisoner's dilemma.*

  (Hurrah!)
- But what if there are a finite number of repetitions?

Backwards Induction

- But... 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...

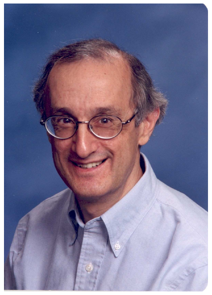
  But this makes round $n - 2$ the last "real", and so you have an incentive to defect there, too.

  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.

Axelrod's Tournament

- Suppose you play iterated prisoner's dilemma against a *range* of opponents . . .
  What strategy 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 prisoner's dilemma.
- Axelrod hosted the tournament and various researchers sent in strategies 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 :-)

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.

Stag hunt

The stag hunt scenario can be described by:

> You and a friend decide it would be a great joke to show up on the last day of school with some ridiculous haircut. Egged on by your clique, you both *swear* you'll get the haircut.
>
> A night of indecision follows. As you anticipate your parents' and teachers reactions [. . .] you start wondering if your friend is really going to go through with the plan.

> Not that you don't want the plan to succeed: the best possible outcome would be for both of you to get the haircut.
>
> The trouble is, it would be awful to be the *only* one to show up with the haircut. That would be the worst possible outcome.
>
> You're not above enjoying your friend's embarrassment. If you *didn't* get the haircut, but the friend did, and looked like a real jerk, that would be almost as good as if you both got the haircut.

- A payoff matrix like:

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

  describes this scenario.
- The difference from the prisoner's dilemma is that now it is better if you both co-operate than if you defect while the other co-operates.

- There are two Nash equilibrium solutions:
  - Both co-operate
  - Both defect
- The same scenario occurs in mutinies and in the orginal hunting scenario that gave it its name (as well as the hunting practices of some animals).

Game of Chicken

The game of chicken gets its name from a rather silly, macho "game" that was supposedly popular amongst juvenile delinquents in 1950s America; the game was immortalised by James Dean in the 1950s film *Rebel without a Cause*. The purpose of the game is to establish who is bravest of the two players.

The game is played by both players driving their cars at high speed towards a cliff. The idea is that the least brave of the two (the "chicken") will be the first to drop out of the game by jumping out of the speeding car. The winner is the one who lasts longest in the car. Of course, if *neither* player jumps out of the car, then both cars fly off the cliff, taking their foolish passengers to a fiery death on the rocks that undoubtedly lie at the foot of the cliff.

- Chicken has a payoff matrix like:

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

(Swerving = coop, driving straight = defect.)

- Difference to prisoner's dilemma:

Mutual defection is most feared outcome.

(Whereas sucker's payoff is most feared in prisoner's dilemma.)

- There is no dominant strategy (in our sense).
- Strategy pairs (C, D) and (D, C) are Nash equilibria.
- All outcomes except (D, D) are Pareto optimal.
- All outcomes except (D, D) maximise social welfare.

Other Symmetric 2 x 2 Games

- Given the 4 possible outcomes of (symmetric) cooperate/defect games, there are 24 possible orderings on outcomes.
  - $CC \succ_i CD \succ_i DC \succ_i DD$
    Cooperation dominates.
  - $DC \succ_i DD \succ_i CC \succ_i CD$
    Deadlock. You will always do best by defecting.
  - $DC \succ_i CC \succ_i DD \succ_i CD$
    Prisoner's dilemma.
  - $DC \succ_i CC \succ_i CD \succ_i DD$
    Chicken.
  - $CC \succ_i DC \succ_i DD \succ_i CD$
    Stag hunt.

Summary

- This lecture has looked at agent interactions, and one approach to characterising them.
- The approach we have looked at here is that of *game theory*, a powerful tool for analysing interactions.
- We looked at solution concepts of Nash equilibrium and Pareto optimality.
- We then looked at the classic Prisoner's Dilemma, and how the game can be analysed using game theory.
- We also looked at the iterated Prisoner's Dilemma, and other canonical 2×2 games.