Multi-agent interaction

- Making the best decision

  We are going to look at how agents make the best decision in a multi agent domain. The key for this is revisiting the concept of Utility

  States or runs can be mapped to Utility Functions, and "success" can be mapped to outcomes (mapping 1 and 0 to a utility function).

  We can generalise success and failure further more. Being Omega the set of possible outcomes for a multi-agent interaction (game), we can match particular outcomes to real number values (the amount of utility for an agent). It's important to note a given outcome is not necessarily mapped to a particular run, since multiple runs might end up with a particular outcome. You also can't pin down what particular state is going to map to a particular outcome.

  - Example

    Example;

    $\Omega = \{\omega_1, \omega_2, \omega_3\}$: $\omega_1 = Ag_i$ wins ; $\omega_2 =$ Draw ; $\omega_3 = Ag_j$ wins

    $u_i(\omega_1) = 1, \ u_i(\omega_2) = 0, \ u_i(\omega_3) = -1$

    $u_j(\omega_1) = -1, \ u_j(\omega_2) = 0, \ u_j(\omega_3) = 1$

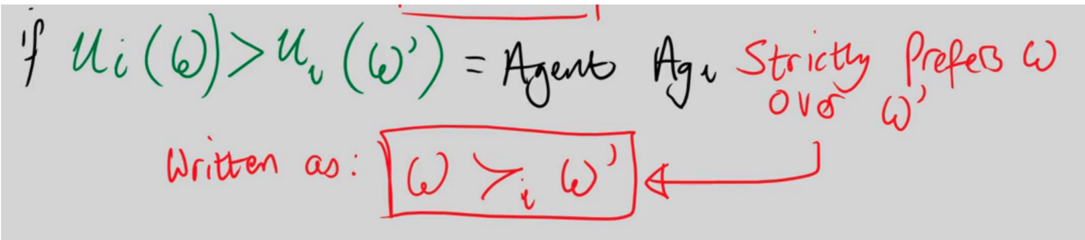
  - Preference Ordering

    If the value of a given outcome's utility is never less than the utility for another, we can say the agent prefers that outcome. Here we can say the agent prefers omega over omega'

    if $u_i(\omega) \geq u_i(\omega')$ = Agent Ag_i Prefers ω over ω'

    Written as: $\boxed{\omega \succeq_i \omega'}$

    Also, if a given outcome omega is always strictly better than another outcome omega', an agent is always going to prefer omega to omega', written as



With this, we can order an agent's preference for each of the outcomes. By doing this for all outcomes, we create a Preference Ordering over the outcomes.

- ▼ Utility

  Utility is a nice way of representing agents' preferences over different potential outcomes. It isn't exactly "money".

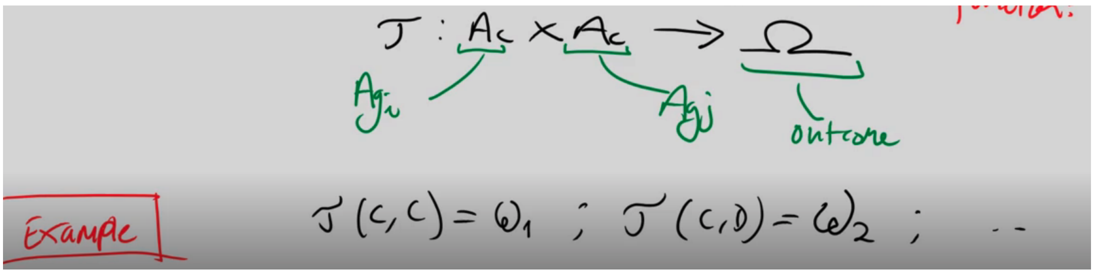
  - ▼ Example

    I have 500€ and you have 0€. We are both offered the same 100€. They will be much more useful to you (going from 0 to 100) than to me (going from 5000 to 5100). Therefore, it supposes a higher utility to you than to me, even though I'd have more money in the end.

- ▼ Outcomes and payoff matrices

  - ▼ Actions to outcomes

    We have to gents: i and j. Both simultaneously choose an action: c or d (cooperate or defect). The combinations of actions they choose will result in a particular outcome from omega. They can't see the action chosen by the other agent until they themselves have chosen (like in rock-paper-scissors). We also now say an action results in an outcome, similar to a state. We can map a combination of actions to a particular outcome.



We can also map outcomes to utility

From Vid. 6.1, we can map outcomes to utility:

eg: $u_i : \Omega \to \mathbb{R}$

$u_i(\omega_1) = 1$

So we can abuse the notation in order to list all the action combinations

From Vid. 6.1, we can map outcomes to utility:

eg: $u_i : \Omega \to \mathbb{R}$

$u_i(\omega_1) = 1$

Previously, all combinations of actions gave distinct outcomes

So...

We can "abuse" the notation in the following manner: $\tau(C,C) = \omega_1 \rightsquigarrow u_i(C,C) = 1$

Likewise, we can abuse this notation to show agent preferences related to outcomes and actions

eg: $(C,C) \succeq_i (C,D) \succ_i (D,C) \succeq_i (D,D)$

- Preferences beget strategies

  Consider the following utility functions

  $u_i(D,D)=4,\ u_i(D,C)=4,\ u_i(C,D)=1,\ u_i(C,C)=1$

  $u_j(D,D)=4,\ u_j(D,C)=1,\ u_j(C,D)=4,\ u_j(C,C)=1$

  If we map those utility functions into our preferences

$$(D,D) \succeq_i (D,C) \succ_i (C,D) \succeq_i (C,C)$$
$$(D,D) \succeq_j (C,D) \succ_j (D,C) \succeq_j (C,C)$$

We can see both agents prefer to defect, no matter the other one's moves. This is called a dominant strategy.

We can map those utility functions to a Payoff Matrix

	i defects	i cooperates
j defects	4, 4	1, 4
j cooperates	4, 1	1, 1

Each cell in the matrix corresponds to a possible outcome. We can see the outcomes of utility functions in the third column from before correspond to the top right cell in the matrix (1 and 4).

- ▼ Strategies and Equilibria
  - ▼ Dominant Strategy

    A strategy c is dominant for a player i if, no matter what strategy agent j does, agent i is going to do at least as well playing c than he would have done playing anything else.

    It's a way of making the decision of what to do a lot easier, because it guarantees the best outcome.

    Consider the following utility functions

    $$u_i(D,D)=4,\ u_i(D,C)=4,\ u_i(C,D)=1,\ u_i(C,C)=1$$
    $$u_j(D,D)=4,\ u_j(D,C)=1,\ u_j(C,D)=4,\ u_j(C,C)=1$$

    Let's say agent j plays action d. We want to know what is the dominant strategy for agent i. Agent i can do no better than playing d.

Let's say agent j plays c. Agent i gets 1 payoff when plays c and 4 payoff when plays d.

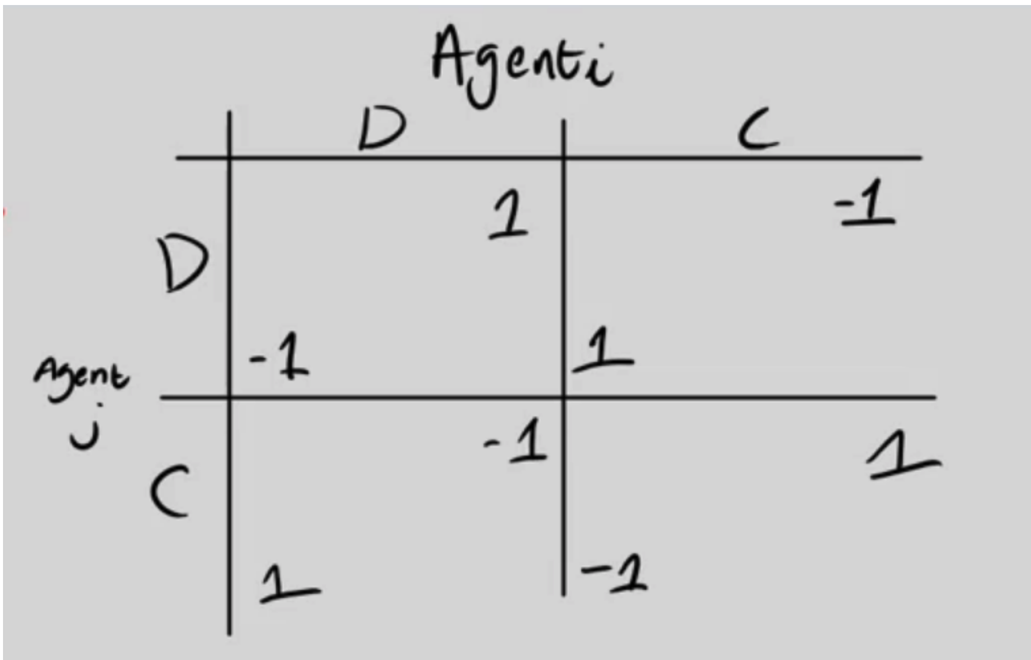
When a particular strategy is the best response to all the opposition's strategies (in this case, playing d), it is considered a Dominant Strategy.

▼ Nash Equilibrium

Two agents (i and j) playing strategies Si and Sj are in nash equilibrium if, when Agi plays Si, Agj does no better than playing Sj, and when Agj plays Sj, Agi does no better than playing Si. In other words, Si is the best response to Sj and Sj is the best response to Si. (For example, two cars driving both on their left side of the road will keep doing so)

▼ Mixed Strategy Nash Equilibrium

Consider the following Payoff Matrix



In the outcome (d, d) agent j regrets choice, and in (c, d) agent i regrets choice. This means (d, d) is not in NE, since agent j will regret playing d. On the other hand, (c, d), agent i will regret their choice and want to play d, getting a better payoff. Therefore, there is no pure strategy NE here.

A strategy is not just an action ("play this action"), so we can introduce some uncertainty or randomness into agent strategies.

This can be useful if, even after fixing our strategy, it's payoff still depends on what the other agent does. In those cases, we might want to fix it with a random component (for example having a 0.5 probability of playing either c or d). This is, in itself, a NE. This is because, if i picks c or d randomly, j cannot do better by also picking randomly c or d.

Every game in which every player has a finite set of possible strategies has a NE in mixed strategies.

- Pareto Optimality

  Pareto Optimality is a property of an outcome in relation to all other outcomes (embodying some kind of fairness).

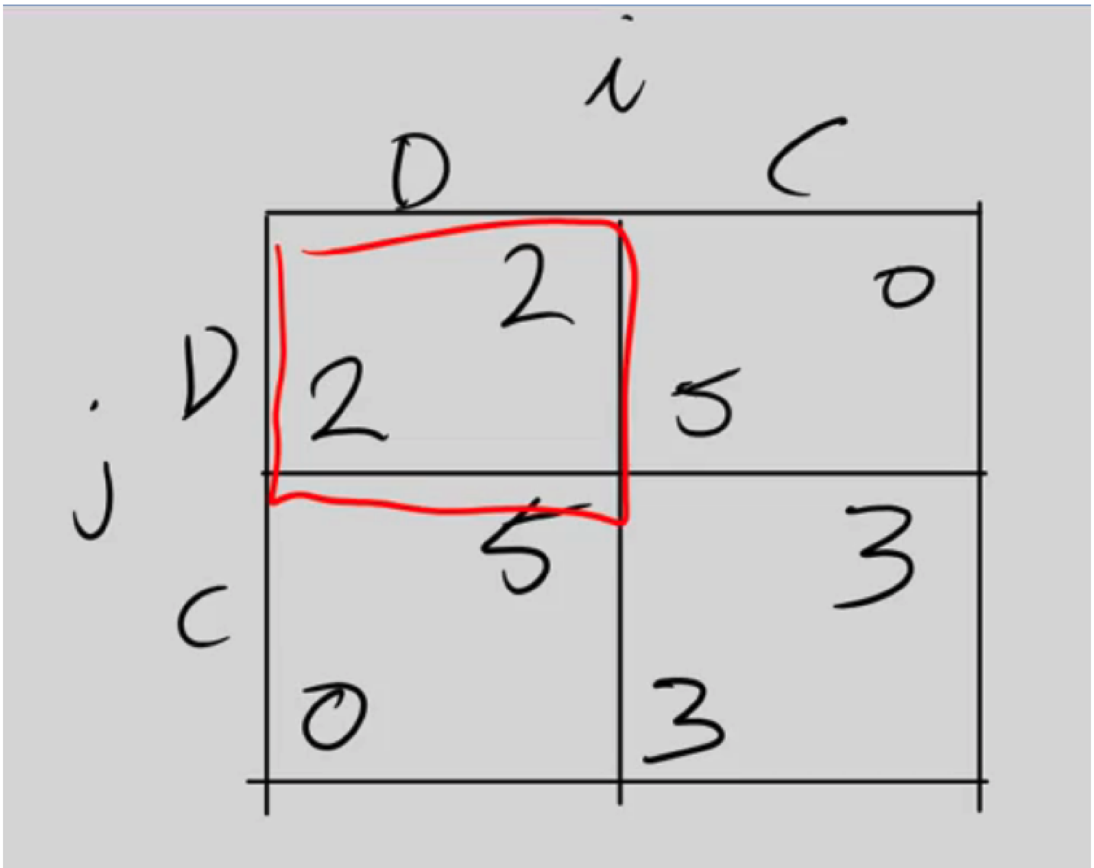
  A Pareto Optimal Outcome is one where no agent can get a better payoff without making at least one agent worse off.

  A non-Pareto Optimal Outcome is one where an agent can obtain a better payoff without putting any other agent in a disadvantage.

  This way, no agent will object moving to a Pareto Optimal outcome.

  - Is an outcome Pareto Optimal?

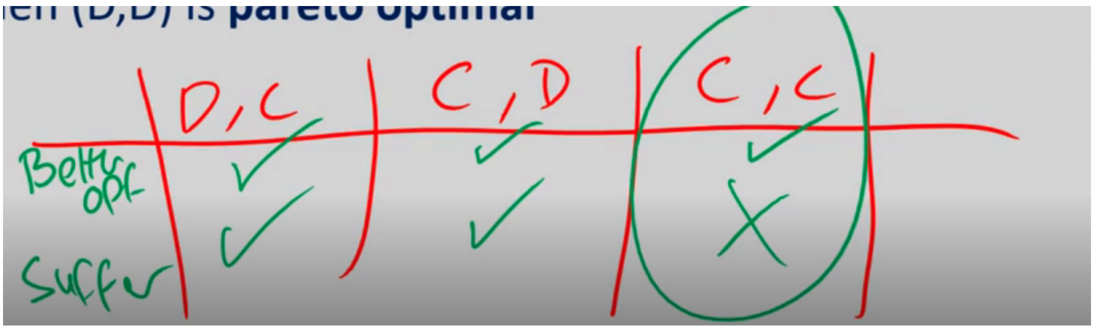
    As an example we are going to check if (d, d) is Pareto Optimal by looking for a counter example. We do this by checking every other outcome on the Payoff Matrix.



We need to check wether or not any other agent will be better off (we want "yes") or will suffer (we want "no").

If the above fails for all other outcomes, then (d, d) is Pareto Optimal.

We know (d, d) is not Pareto Optimal because both agents could move to (c, c) and be better off.



▼ Social Welfare

Social Welfare is the sum of the payoff generated by all the agents in a game (suitable if you're the owner). It disregards individual agents, but it's useful if you have an entire group of agents working towards a problem, since you won't care about a given agent having a small amount of utility and rather be concerned about all the agents having a good total payoff. It is expressed as

$$Sw = \sum_{i \in Ag} U_i(w)$$

We can maximise SW by taking the max SW of the values of all the outcomes within a game

▼ Prisoners' Dilemma

We have 2 prisoners charged with a crime who can't communicate with each other. they are told that, if one of them confesses to the crime and the other doesn't, the confessor will be freed and the other one jailed for 3 years; if both confess, both will be jailed for 2 years; and if none confesses, both will be jailed for 1 year.

We can think of this as a game to solve it.

	Agent *i* defects	Agent *i* cooperates
Agent *j* defects	2 2	0 5
Agent *j* cooperates	5 0	3 3

$(D, C) \succ_i (C, C) \succ_i (D, D) \succ_i (C, D)$

$(C, D) \succ_j (C, C) \succ_j (D, D) \succ_j (D, C)$

If you were one of the prisoners, your dominant strategy (DS) would be to defect, since it's the best response to all actions by the other agent. This will give them both a payoff of 2.

There is a single NE, of the outcome (d, d).

However, this is not the best players can do. If they both switched from (d, d) to (c, c), they could get a payoff of 3. But going from our DS, if we assume the other player will defect, that is also the best thing to do, therefor being always the most rational thing to do.

The only output that is not Pareto Optimal is (d, d), and (c, c) maximises SW.

The fact that with the rational action utility is wasted and agents can do better by cooperating, even though the rational thing is to defect, is why this is referred to as a dilemma.

In order to protect ourselves from this (called the "suckers' payoff"), we can have Program Equilibria, saying "I'll collaborate if you will", a series of "if" statements in which we compare strategies of different agents in order to get us to the most preferable outcome.

- ▼ Other games
  - ▼ Stag Hunt

    Two hunters can choose between cooperating and capturing a stag, or working alone and capturing a rabbit. If one hunter tries to capture the stag alone, they will fail. Hunters are always going to prefer getting a stag rather than a rabbit.

    The problem here is in trusting the other agent to also cooperate, or going for a rabbit alone.

    The difference between this and the Prisoners' Dilemma is that the mutual cooperation is the preferred outcome, rather than you defecting while your opponent cooperates.

    It's Payoff Matrix is

	Agent *i* defects	Agent *i* cooperates
Agent *j* defects	1 1	0 2
Agent *j* cooperates	2 0	3 3

    $(C,C) \succeq_i (D,C) \succ_i (D,D) \succeq_i (C,D)$

    $(C,C) \succ_j (C,D) \succ_j (D,D) \succ_j (D,C)$

    - Two NE: (D,D) and (C,C)
    - Max SW: (C,C)
    - Pareto Optimal: (C,C)

    Of course we have Nash Equilibrium in (d, d) and (c, c), and the max SW is (c, c), being this last one also the Pareto Optimal outcome.
  - ▼ Chicken

Two drivers drive towards the end of a cliff, having to turn away at some point in order not to die. The winner is the last one to turn away. It's payoff matrix is

	Agent *i* defects	Agent *i* cooperates
Agent *j* defects	0 0	1 3
Agent *j* cooperates	3 1	2 2

$(D,C) \succ_i (C,C) \succ_i (C,D) \succ_i (D,D)$

$(C,D) \succ_j (C,C) \succ_j (D,C) \succ_j (D,D)$

An agent will play the game depending on how brave or foolish it believes it's opponent to be. If it thinks it's opponent is braver, it's better to turn away, since the other one will keep on driving; whereas if the agent believes that it's opponent is less brave than him, it should keep driving.

The problem arises when both agents believe the other is less brave, since both agents will stay in the car, hence both will die.

There are two NE: (d, c) and (c, d)

The ones with max SW and the ones that are Pareto Optimal are the ones where both drivers stay alive (c, d), (d, c) and (c, c)