

Unit 2: Modelling multiagent interactions.

Using Game Theory concepts.

Rational Agent (Game Theory)

Agents have “desires” on how the environment they want to be, and beliefs about how is currently and how it behaves. Agents can express **preferences** on these environment states.“

- Let: $Ag = \{i, j\}$
- Agents are *self-interested, each one has its own preferences* -preferences of an agent are independent of the preferences of the other agents-
- Let $\Omega = \{\omega_1, \omega_2, \dots\}$ be the set of possible “scenarios” or states of the environment.
- Each agent expresses its preferences as an *utility function*:

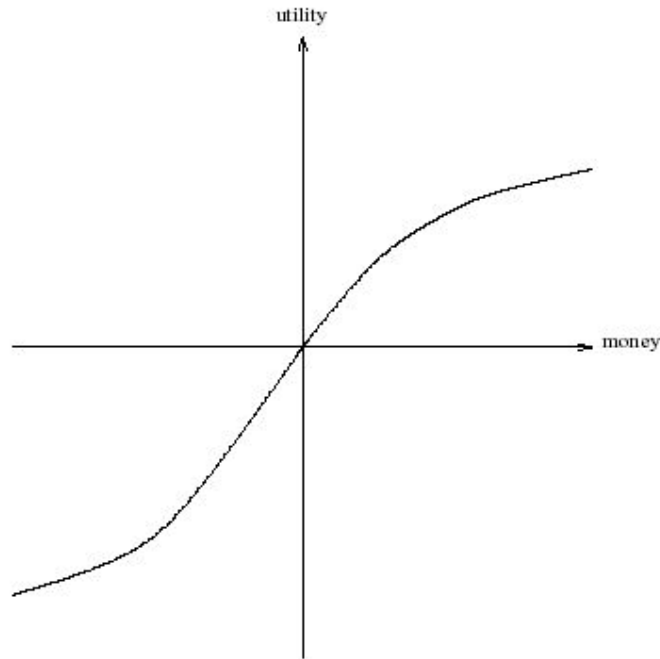
$$u_i = \Omega \rightarrow \mathbb{R}$$

← A real value is given to each scenario
- So, with utility functions scenarios can be sorted in some way:

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

What Utility stands for?

- With humans that is a hard problem. Although **utility** is not money, this is an useful analogy
- Relationship between money and utility:



Rational agent

An agent is **rational** when, choosing among their preferences, it chooses based on its **acceptable preference order**:

- An order that is complete and transitive.
- Which one is selected is done by the utility function.

Example: Make a six eggs omelette.

The agent has already beaten 5 and has 1 left, but there is a risk that it is bad.

What actions, potential scenarios and preferences exist in this example?

	Good condition (S_1)	Not good condition (S_2)
Break egg in the same plate (A_1)	6 eggs omelette (C_{11})	No omelette (C_{12})
Break egg in other plate (A_2)	Omelette with 6 eggs and dish to wash (C_{21})	Omelette with 5 eggs and dish to wash (C_{22})
Do not use egg (A_3)	5 eggs omelette (C_{31})	5 eggs omelette (C_{32})

How to sort these preferences? By using the utility values of each preference.

$$\begin{aligned}
 C_{11} &\succcurlyeq C_{21} \succcurlyeq C_{31} \approx C_{32} \succcurlyeq C_{22} \succcurlyeq C_{12} \\
 C_{11} &\succcurlyeq C_{31} \approx C_{32} \succcurlyeq C_{21} \succcurlyeq C_{12} \succcurlyeq C_{22}
 \end{aligned}$$

In which of the two states is the egg?

Expected Utility

What if the agent isn't sure what the current state of the world is? The EU (expected utility) can be calculated: EU. ("A" denotes the action selected by the agent)

$$UE(A) = \sum p(S)U[C(S, A)], \quad \text{sujeito a } \sum p(S) = 1$$

Example: *The problem of the corrupt politician (everyone knows they don't exist ;-)).*

	Evidences (S_1)	No evidences (S_2)
Deny it all (A_1)	Destitute (C_1)	Reforced (C_3)
Confess party (A_2)	Weakened (C_2)	Weakened (C_2)
Confess all (A_3)	Destitute (C_1)	Destitute (C_1)

Suppose that:

$$\begin{array}{lll} U(C_3) = 1; & U(C_2) = 0.6; & U(C_1) = 0 \text{ (these values provide a preference order)} \\ p(S_1) = 0.2; & p(S_2) = 0.8 & \end{array}$$

Expected utility

What action is the corrupt politician to take? (the one with the highest EU)

Assuming that the probability of the existence of evidence is unknown, with what probability is A2 (confess part) better than A1 (deny everything)?

How the utilities of the Cs are decided?
How do they relate to each other?
Cardinally or ordinally



Problem

Von Neumann-Morgenstern Utility

Evaluate the intermediate consequences (Cs) based on the best and worst available consequences acting in a lottery.

Lotteries are defined (consequences tied with probabilities).

$$L = (p_1 C_1, p_2 C_2, \dots, p_n C_n), \text{ sujeto a } \sum p_i = 1$$

The player's indifference is sought between opting for an intermediate consequence and the probability of playing a lottery in which only extreme consequences can come out.

For example, with the problem of the corrupt politician:

C2 is not indifferent to (0.4C3, 0.6C1) (it prefers C2 to that lottery),

C2 is indifferent to (0.6C3, 0.4C1)

The best and worst consequence utilities are normalized to 1 and 0.

$$U(C_3) = 1$$

$$U(C_1) = 0$$

The utilities of intermediate consequences are calculated using the expected utility of the lottery that makes them indifferent:

$$U(C_2) = UE(L(0.6C_3, 0.4C_1)) = 0.6*1 + 0.4*0 = 0.6$$

Von Neumann-Morgenstern utility

The Von Neumann-Morgenstern utility function determines which Lottery is the most suitable for each possible action: choosing the lottery associated with the action that is most likely to result in the best possible consequence.

Example: Given two actions A1 and A2 with, respectively, lotteries L1 and L2. Suppose C1 be the best consequence, C4 the worst consequence, and it is known that the intermediate options show the indifferences indicated below. Which action will choose the agent: A1 or A2?

$$C_2 \approx (0.7C_1, 0.3C_4)$$

$$C_3 \approx (0.4C_1, 0.6C_4)$$

$$L_1 = (0.4C_1, 0.1C_2, 0.2C_3, 0.3C_4)$$

$$L_2 = (0.2C_1, 0.3C_2, 0.3C_3, 0.2C_4)$$

$$\begin{aligned} L_1 &= (0.4C_1, 0.1[(0.7C_1, 0.3C_4)], 0.2[(0.4C_1, 0.6C_4)], 0.3C_4) = \\ &= (0.4C_1, 0.07C_1, 0.03C_4, 0.08C_1, 0.12C_4, 0.3C_4) = (0.55C_1, 0.45C_4) \end{aligned}$$

Exercise: Calculate L2 according to the previous pattern. What action does the agent choose?

Attitude towards risk

Risk is a key to measure the intensity of the agent's preferences.

Is the agent indifferent between:

- 1) play a lottery (obtain the expected utility of the lottery: $UE(L) = \sum pU(C_i)$); either
- 2) obtain the expected value of the lottery $E(L) = \sum pC_i$?

There are three possibilities:

- **Risk neutral agent**. It is indifferent to himself.
- **Risk prone agent**. He prefers to play the lottery.
- **Risk averse agent**. Prefers the expected value of the lottery.

Attitude towards risk

Suppose a **linear** utility function: $U(x) = 4x + 2$.

The agent is offered the lottery $L = [4/5 \text{ 1 euro}, 1/5 \text{ 16 euro}]$

The expected value of the lottery L is $E(L) = 4/5 \cdot 1 + 1/5 \cdot 16 = 4$ euros

$$U(E(L)) = U(4) = 4 \cdot 4 + 2 = 18$$

$$UE(L) = 4/5 \cdot U(1) + 1/5 \cdot U(16) = 4/5 \cdot 6 + 1/5 \cdot 66 = (66+24)/5 = 90/5 = 18$$

$$U(E(L)) = UE(L) \text{ The agent is } \textbf{risk neutral}$$

Assume a **nonlinear** utility function: $U(x) = 4 \sqrt{x}$

Then,

$$U(E(L)) = U(4/5 \cdot 1 + 1/5 \cdot 16) = U(4) = 4 \sqrt{4} = 8 \text{ euros}$$

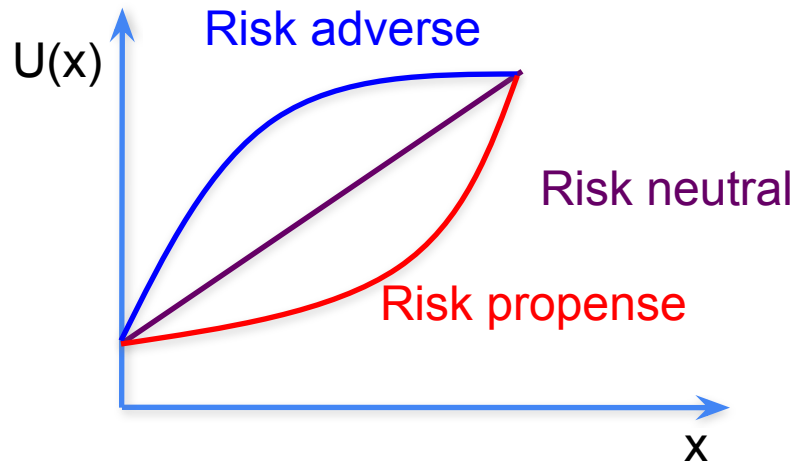
$$UE(L) = 4/5 \cdot U(1) + 1/5 \cdot U(16) = (4/5) \cdot 4 \sqrt{1} + (1/5) \cdot 4 \sqrt{16} = (16+16)/5 = 6.4 \text{ euros}$$

\Rightarrow

$$U(E(L)) > UE(L) \text{ The agent is } \textbf{risk averse}$$

What will an agent (neutral, prone, or adverse) choose if $p = 0.3$ to get 10 euros and $p = 0.7$ to get 2?

Attitude towards risk



But, do we, humans, act like this, with this type of calculations in our interactions?

Numerous experiments show that people deviate from the results provided by utility theory.

The paradox of Allais

	S (p=0.01)	T (p=0.1)	U (p=0.89)
L	500.000	500,000	500.000
M	0	2.500.000	500.000
L'	500.000	500.000	0
M'	0	2.500.000	0

What do you choose choose L or M? What do you choose L' or M'? ... **Choose now!**

Many people tend to choose L (between L and M) and M' (between L' and M')

Normalize utilities such that $U(2.500.000) = 1$; $U(500.000) = z$; $U(0) = 0$ where $0 < z < 1$

What must be hold to choose L over M? And M' over L'?

Calculate $UE(L)$, $UE(M)$, $UE(L')$ y $UE(M')$...

Is it consistent to choose L and M'?

Strategic **Multi-Agent** Encounters (Game Theory)

We need a model of the environment in which agents act:

- The agents simultaneously choose an action to execute and as a result of its execution, a scenario of Ω will be obtained.
- Such a current scenario depends on the combination of actions
- There are utility functions of VN-M for each combination of strategies

Behavior of the environment will be given by the state transformation function:

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

Strategic multi-agent meetings

- Example of the state transformation 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 the actions of both agents)

- Other example:

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

(No agent influences the environment)

- Another more example:

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

- *(Agent j controls the environment)*

Rational action

- Suppose that both agents can influence the scenario and that their utility functions are:

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

- Abusing the notation a bit...

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

- Agent i's preferences are: $C, C \succeq_i C, D \succ_i D, C \succeq_i D, D$

- “C” is the rational choice for i.

(Because i prefers all scenarios that appear by running C over all scenarios that appear by running D.)

Payoff matrices

- The previous scenario can be characterized by payoff matrices (normal form):

		i	
		Defect	Coop
j	Defect	1,1	1,4
	Coop	4,1	4,4

- Agent i *column*
- Agent j *row*

Payoff matrices (more examples)

		J2		
		l	m	r
J1	U	4,3	5,1	6,2
	M	2,1	8,4	3,6
	D	3,0	9,6	3,8

How can we analyze this kind of payoff matrices to know which actions will be taken?
Are there any “general” properties that can be drawn from them?

Dominant strategies

- It is said that s_1 dominates s_2 if every possible scenario obtained by i when playing s_1 is preferred over every possible scenario obtained by i when playing s_2 .
- A **rational agent** will never play a dominated strategy
- Therefore, a rational agent must eliminate the dominated strategies
- Unfortunately, not always they exist.

		J2	
		l	r
J1	U	-1,3	2,1
	D	0,2	3,4

What rational choice will J1 and J2 make?

Dominant strategies

Iteratively remove dominated strategies:

		J2		
		l	m	r
J1	U	4,3	5,1	6,2
	M	2,1	8,4	3,6
	D	3,0	9,6	3,8

		J2	
		l	r
J1	U	4,3	6,2
	M	2,1	3,6
	D	3,0	3,8

		J2	
		l	r
J1	U	4,3	6,2
	D	3,0	3,8

		J2	
		l	r
J1	U	4,3	6,2

		J2
		l
J1	U	4,3

Dominant strategies and rationality

		J2	
		l	r
J1	U	8,10	-100,9
	D	7,6	6,5

(U,l) is the solution with payments (8,10)

What if we are not sure of the rationality of J2?

Suppose J1 believes that J2 is 98% rational. What would be the best strategy for J1?

$$UE_{J_1}(U) = 0.98 \cdot 8 + 0.02 \cdot (-100) = 5.84$$

$$UE_{J_1}(D) = 0.98 \cdot 7 + 0.02 \cdot 6 = 6.98$$

The process of eliminating dominant strategies does not always lead to an adequate solution (especially if we play with people who do not work 100% with the standard mathematical criterion of rational agent)

Nash equilibrium

- In general, we will say that two strategies s_1 and s_2 are in Nash equilibrium if:

Under the assumption that agent i plays s_1 , agent j can do no better than play s_2

Y

Under the assumption that agent j plays s_2 , agent i can do no better than play s_1 .



John Nash.

Neither agent has an incentive to deviate from the Nash equilibrium.

Nevertheless:

Not every interaction situation has a pure Nash equilibrium.

Some interaction situations have more than one Nash equilibrium.

It can be computationally expensive, in the worst case, to calculate the Nash equilibrium (NP-complete)

Nash equilibrium

Calculate the Nash equilibria of the following games in Normal Form.

		J2	
		s_1	s_2
J1	s_1	1,1	1,1
	s_2	2,-1	-10,-2
	s_3	-1,-2	0,-1

		J2	
		s_1	s_2
J1	s_1	1,1	0,4
	s_2	-1,3	3,-5

Mixed Nash Equilibrium

Behavior of insects according to the role of Hunter or Prey

		Prey	
		Active	Passive
Hunter	Active	2,-7	6,-8
	Passive	3,-6	-1,0

There are no pure equilibrium strategies. If Prey is Active, Hunter is Passive => Passive Hunter, Prey is Passive => Passive Prey, Active Hunter => Active Hunter, Active Prey...

What probability for the prey makes the hunter indifferent to the choice of any of his strategies?

Let p be the probability that the prey is Active ($1-p$ that it is Passive)

$$2p + 6(1-p) = 3p - (1-p) \Rightarrow p = 7/8$$

What probability for the hunter makes the prey indifferent to the choice of any of his strategies? Let q be the probability that the hunter is Active. **Solve**.

Mixed Nash Equilibrium

The sex battle

		One (J1)	
		PrettyWoman	Rocky
Another (J2)	PrettyWoman	3,1	0,0
	Rocky	0,0	1,3

They both want to go to the movies together, but prefer different movies.

Solve. Compute all Nash equilibria in this game.

Competitive and Zero-Sum Interactions

- If the preferences of the agents are totally opposite, then we have strictly competitive situations.
- Zero-Sum Encounters are those in which the sum of the utilities is zero:
 $u_i(\omega) + u_j(\omega) = 0$ for all $\omega \in \Omega$
Therefore it implies a strictly competitive situation (J1 wins with pairs, J2 wins with odds)

		J2	
		Pairs	Odds
J1	Pairs	1,-1	-1,1
	Odds	-1,1	1,-1

Competitive and Zero-Sum Interactions

		J2		
		Rock	Paper	Scissors
J1	Rock	0,0	-1,1	1,-1
	Paper	1,-1	0,0	-1,1
	Scissors	-1,1	1,-1	0,0

Zero-sum encounters are rare in real life...although people tend to act sometimes as if we were in one of them.

Interacciones Competitivas y de Suma-Cero

No es sólo suma = 0, si no también suma = k, como en el caso del problema del penalty.

¿Cómo calcular un equilibrio en un juego de suma cero?

Aplicando el Teorema *Minimax*.

$$S_1 = \operatorname{argmax}_{s' \in S_1} \min_{s_2 \in S_2} u_1(s', s_2)$$

El jugador 1 ha de escoger la estrategia que maximice su utilidad sabiendo que el jugador 2 va a tratar de minimizarla (y de esta forma maximiza la utilidad del jugador 2)

		Portero	
		Izq	Der
Lanzador	Izq	0.6, 0.4	0.8, 0.2
	Der	0.9, 0.1	0.7, 0.3

$$S_2 = \operatorname{argmin}_{s' \in S_2} \max_{s_1 \in S_1} u_1(s_1, s')$$

El jugador 2 ha de escoger la estrategia que minimice la utilidad del jugador 1 (y de esta forma máxima la suya) sabiendo que el jugador 1 va a tratar de maximizarla.

$$\min_{s_2 \in S_2} u_1(s', s_2) = [s'(I)s_2(I) * 0.6 + s'(I)s_2(D) * 0.8 + s'(D)s_2(I) * 0.9 + s'(D)s_2(D) * 0.7]$$

$$\min_{s_2 \in S_2} u_1(s', s_2) = [s'(I)s_2(I) * 0.6 + s'(I)(1 - s_2(I)) * 0.8 + (1 - s'(I))s_2(I) * 0.9 + (1 - s'(I))(1 - s_2(I)) * 0.7]$$

Interacciones Competitivas y de Suma-Cero

Primera derivada con respecto a $s_2(I) = 0$:

$$0.2 - s'(I) * 0.4 = 0$$

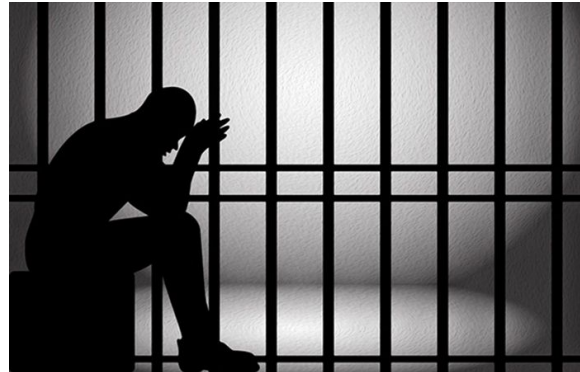
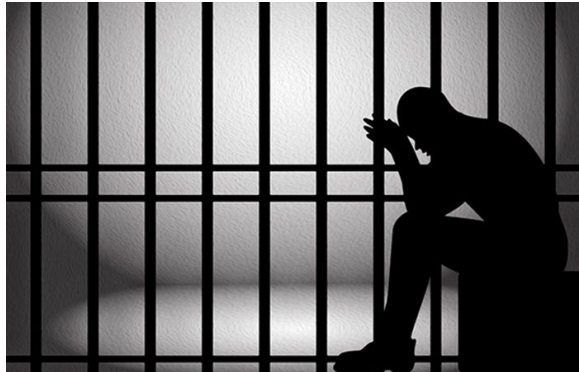
Luego $s'(I) = \frac{1}{2}$ y, por lo tanto, $s'(D) = \frac{1}{2}$ para el jugador lanzador.

Resolver. Calcular $s'(I)$ para el jugador portero
(Pista: el resultado debe ser $s'(I) = \frac{1}{4}$ y $s'(D) = \frac{3}{4}$)

A estos resultados también se puede llegar sin utilizar el teorema maxmin mediante el cálculo general de estrategias en equilibrio de los juegos de dos jugadores.

The Prisoner's Dilemma

- Two people are accused of a crime and are locked in individual cells with no possibility of communication between them.
- The following deal is offered:
 - If you confess and the other does not confess, the confessor will go free (for helping) and the other will be imprisoned for three years
 - If both confess, they will each be sentenced to two years.
 - Both defendants know that if neither confesses they will be jailed for one year (for minor crimes)



The Prisoner's Dilemma

payoff matrix of the
prisoner's dilemma:

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

It is considered that the higher the utility, the better the agent does. For this reason, the value of spending 3 years in prison has been modified to utility 1 and that of going free to utility 4...

- ❑ Top Left: If they both give themselves away (don't cooperate with each other) they both get punished
- ❑ Top Right: If j snitches (doesn't cooperate) and i doesn't snitch (cooperates), j gets the snitch convicted (utility 4), while i goes to prison for 3 years (utility 1).
- What is the rational individual action for each agent? Discussion ...

The Prisoner's Dilemma

Individual rational action is defective

This guarantees a payoff no worse than 2, while cooperating guarantees a payoff of at most 1.

Then default is the best answer for all possible strategies: both agents default, and they get a payoff = 2

But intuition says that this is not the best scenario:

If they both cooperated they would get a payoff = 3!

The Prisoner's Dilemma

This apparent paradox is the so-called fundamental problem of multiagent interactions.

It seems to imply that cooperation will not occur in societies of self-interested agents.

Some real examples:

- Reduction of nuclear arsenals (“and if I agree, but I don't destroy them...”)

- Uncontrolled payment systems — public transport;

The prisoner's dilemma is very common.

Why then do we cooperate?

Arguments to Cooperate

- We could think that in the previous analysis:
 - The notion of rational action in game theory is wrong
 - The dilemma has some error in its statement that I can't find...

Arguments to cooperate:

- ☐ We are not so Machiavellian!
- ☐ Blind trust in the other prisoner (the other prisoner is my brother, ...)
- ☐ And if I find him when he goes out from prison...?

The Iterated Prisoner's Dilemma

Let's try to play the game more than once

...

It is no longer so clear that snitching is good, in fact:

The rational choice in the infinitely iterated prisoner's dilemma is to cooperate!

Induction from the end

- But...what if we both know that we will play the game exactly n times?
In round $n - 1$, you have the incentive to default, to get more in the last round...
This makes round $n - 2$ the last “real” one, and so defect would be interesting as well.
....

Playing the prisoner's dilemma with a known, finite, fixed, predetermined number of rounds again shows that "Don't Cooperate" (default) is the best strategy.

Axelrod's Tournament

- ❑ Suppose we play the prisoner's dilemma repeatedly with **several opponents...**
- ❑ What strategy should we choose, to maximize the total profit?
- ❑ Axelrod (1984) investigated this problem through a computerized tournament of programs that played the prisoner's dilemma.



Robert Axelrod. The Evolution of Cooperation.

Some Axelrod Tournament Strategies

- ALLD:
 - “Always defect” —shark strategy;
- TIT-FOR-TAT:
 1. In the first round, cooperate
 2. In round $u > 1$, do what your opponent did to you in round $u - 1$
- TESTER:
 - On the first round, I default. If the opponent misses you then play TIT-FOR-TAT. Otherwise intersperse cooperation and non-cooperation.
- JOSS:
 - As TIT-FOR-TAT, except sporadically defect

Recipes for Success in the Axelrod Tournament

- Axelrod suggests the following rules to be successful in your tournament:
- Do not be jealous:
Don't play like it's a zero sum game
- Be polite:
Start cooperating, and return cooperation
- Measured revenge:
Always punish “non-cooperation” immediately, but do not over-punish
- Be quick to return the good:
Cooperate with whoever cooperates with you immediately

The Chicken game

- Another kind of encounter

the Chicken game

		i	
		Defect	Coop
j	Defect	1,1	4,2
	Coop	2,4	3,3

Two cars driving off a cliff, who will brake first?

In this case defect mutually is the worst scenario.

The strategies (C,D) and (D,C) are in Nash equilibrium

Another symmetric 2×2 games

- With 2 agents and 2 actions there are 24 possible games
 - $CC \succ_i CD \succ_i DC \succ_i DD$
Cooperation dominates
 - $DC \succ_i DD \succ_i CC \succ_i CD$
Deadlock. The best is always defect
 - $DC \succ_i CC \succ_i DD \succ_i CD$
Prisoner's dilemma
 - $DC \succ_i CC \succ_i CD \succ_i DD$
Chicken (Who will brake before reaching the precipice?)
 - $CC \succ_i DC \succ_i DD \succ_i CD$
joke for two