TÉCNICO LISBOA

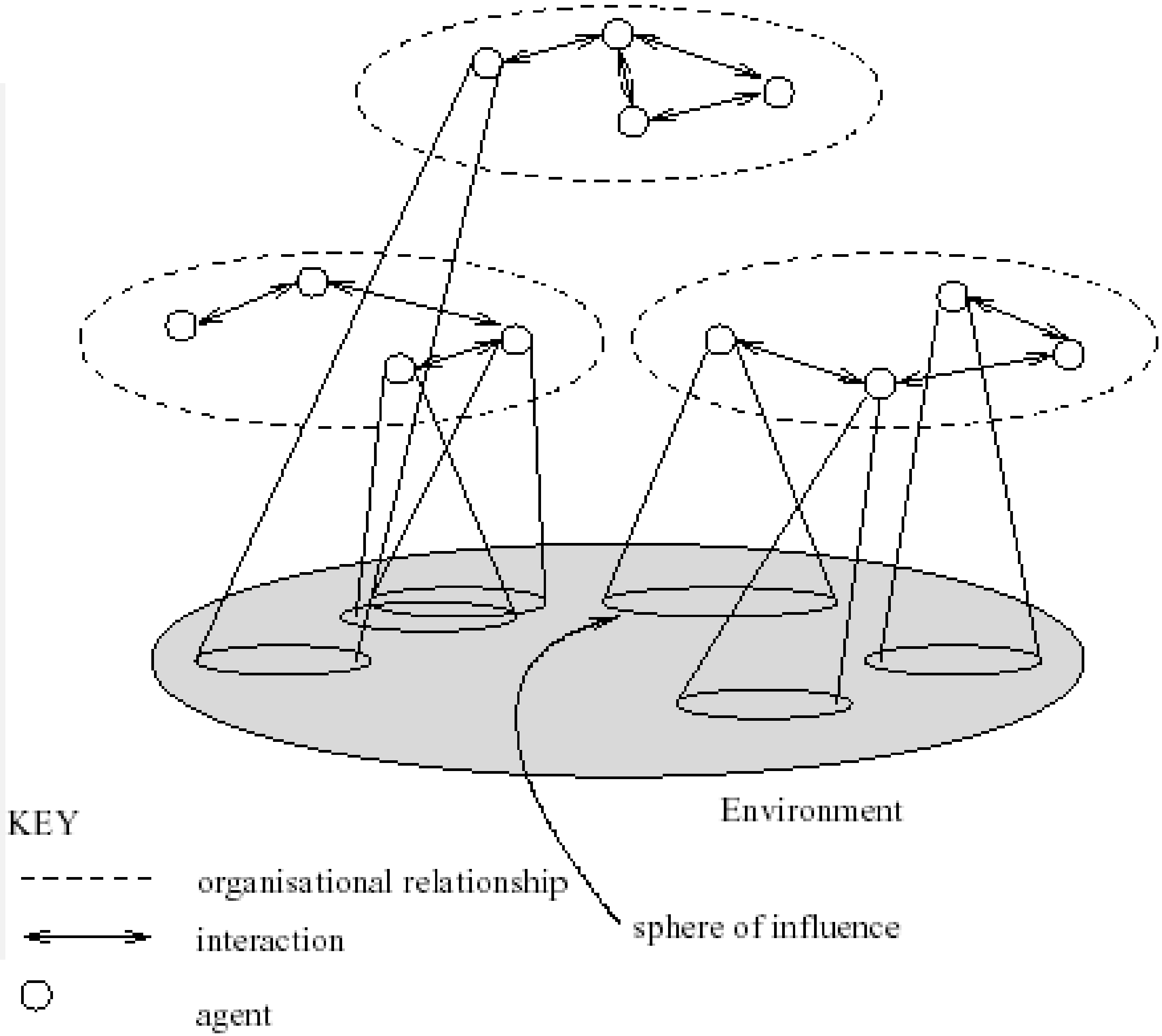
Introduction to Game Theory

Autonomous Agents and MultiAgent Systems 2015/2016

Ana Paiva

* These slides are based on the book by Prof. M. Woodridge "An Introduction to Multiagent Systems" and the online slides compiled by Professor Jeffrey S. Rosenschein. Modifications introduced by Prof. Ana Paiva or Prof. César Pimentel, are their sole responsibilit

MultiAgent Systems



MultiAgent Systems

- Multiple agents that interact on the environment
- *Spheres of influence*:
  - May *coincide*!
- *Dependencies* between agents

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
- Environment behavior given by *state transformer function*:

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

Agents thinking strategically…

- We do not act in vacuum, and agents (because they are in a society) should neither.
- Behaving in society is guided by what strategies are we feel the best when also considering the decisions of others.

Why Game Theory?

- Game theory studies what happens when self interested agents interact.

- A game is “ a competitive activity…in which players contend with each other according to a set of rules”- Dictionary

Game Theory

Def: A formal way to analyze interaction among a group of rational agents who behave strategically.

Some Assumptions

2 Agents: i and j
Self Interested: Each with its preferences and desires
Both *must choose* an action *simultaneously* (or without knowing the action of the other)
The combination of the actions determines an outcome

The Grade Game (O Jogo das notas)

Here is how grades may be assigned for the Agents course Test.

- If you put α and your pair puts β, then you will get grade 19, and your pair grade 10.
- if both you and your pair put α, then you both will get grade 13.
- if you put β and your pair puts α, then you will get grade 10, and your pair grade 19.
- if both you and your pair put β, then you will both get grade 16."

Elements in a Game

- Players: that form the group and are the decision makers.
- Interactions: the actions of one player affects the actions of another player.
- Strategy: an individual player accounts for his independence in deciding what action to take, that is its strategy.

.


Lets play!

Strategies and Outcomes

- The possible choices, α or β that an agent (or a player) can make are called \`strategies'.
- The results obtained (in this case the grades (for example 13,19) are \`outcomes'.

Outcome Matrix

Pair

Me

α	
β	

Outcome Matrix

Pair

Me

	α	β
α		
β		

Outcome Matrix

Pair

Me

	α	β
α	13,13	
β		

Outcome Matrix

Pair

Me

	α	β
α	13,13	
β	10,19	

Outcome Matrix

Pair

Me

	α	β
α	13,13	19,10
β	10,19	

Outcome Matrix

Pair

Me		α	β
	α	13,13	19,10
	β	10,19	16,16

What strategy should an intelligent agent follow?

- Depends on what is “important” for the agent…..
- -> UTILITY Function

Game theory and Utility Theory

- The dominant approach to modeling an agent's interest is *utility theory*.
- A utility function is a mapping from states of the world to real numbers- and these numbers should be interpreted as the agent's level of happiness in the given states.

Def: Normal form game

- A *finite n-person normal form* game is a tuple (N,A,u), where:
  - N is the finite set of players, indexed by i;
  - A= A1 x A2 x … x An where An is a finite set of actions available to player i (strategies).
  - u = (u1,…un) where ui: A -> R is a real-valued utility (payoff) function for player i.

Payoff Matrices

- We can characterize then a *payoff matrix*:

j

i		defect	coop
	defect	1,1	1,4
	coop	4,1	4,4

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

And for the grades Game?

Possible payoff for the grades game

“Evil gits”: the agent only cares for his/her own grade
- get 13 (payoff 0)
- get 10 (payoff -1)
- get 16 (payoff 1)
- get 19 (payoff 3)

Pair

Me

	α	β
α	0,0	3, -1
β	-1,3	1,1

Possible payoff for the grades game

"Nice agents": the agent gets happy if not only him but the other gets a good grade

- get 13 (payoff 1)
- get 10 (payoff 0)
- get 16 (payoff 4)
- get 19 (payoff 3)

Pair

Me

	α	β
α	1,1	3, 0
β	0,3	4,4

The Chocolate's Dilemma

- Two students are told that:
  - if one defects and the other cooperates, the defector will get 4 chocolates, and the other will get zero.
  - if both defect, then each will take one chocolate
  - If both collaborate, then each will get 3 chocolates

Let’s Play!

TÉCNICO LISBOA

The Chocolate's Dilemma

Payoffs

- Get 4 Chocolates (payoff 4)
- Get 3 Chocolates (payoff 2)
- Get 1 Chocolate (payoff 1)
- Don't get any Chocolate (payoff 0)

Pair

Me

	Cooperate	Defect
Cooperate	3,3	0,4
Defect	4,0	1,1

What type of Game?

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 (Defect) and the other does not (Cooperate), 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

Prisoner’s Dilemma: summary

– If one confesses and the other does not, the confessor (defector) is *freed* and the other is jailed for *3 years*

– If both confess: jail for *2 years*

– If neither confesses (cooperate): jail for *1 year*

The prisoner’s Dilemma

- What to do?

Payoffs
-Get free (payoff 4)
-One year in jail (payoff 3)
- Two years in jail (payoff 2)
- Three years in jail (payoff 1)

The prisoner's Dilemma

Payoffs

- Get free (payoff 4)
- One year in jail (payoff 3)
- Two years in jail (payoff 2)
- Three years in jail (payoff 1)

Pair

Me

	Cooperate	Defect
Cooperate	3,3	1, 4
Defect	4,1	2,2

TÉCNICO LISBOA

The Prisoner's Dilemma

- The *individual rational* action is *to 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!

Concept: “Dominant Strategies”

- Given any particular strategy of 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
- A rational agent will never play a dominated strategy
- So in deciding what to do, our agents can *delete dominated strategies*
- Unfortunately, there isn’t always a unique undominated strategy

Dominant Strategies in the Grades Game

My strategy α strictly dominates my strategy β if my payoff from α is strictly higher than that from β regardless of others' choices.

Pair

Me		$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$
	α	0,0	3, -1
	β	-1,3	1,1

Dominant Strategies in the Grades Game

My strategy α strictly dominates my strategy β if my payoff from α is strictly higher than that from β regardless of others' choices.

Pair

Me		α	β
	α	0,0	3, -1
	β	-1,3	1,1

Dominant Strategies in the prisoner's Dilemma

Payoffs

- Get free (payoff 4)
- One year in jail (payoff 3)
- Two years in jail (payoff 2)
- Three years in jail (payoff 1)

Pair

Me

	Cooperate	Defect
Cooperate	3,3	1, 4
Defect	4,1	2,2

Dominant Strategies

- To decide the best strategy the agent should first eliminate the dominated strategies.

..then… it should analyze the payoffs putting itself on the shoes of the other agents…

Nash Equilibria

- s_1 and s_2 are in *Nash equilibrium* iff:
  - Assuming *i* plays s_1, *j* *can do no better* than play s_2
  - Assuming *j* plays s_2, *i* *can do no better* than play s_1
- *Neither agent has any incentive to deviate from a Nash equilibrium!*
- Example: Choosing the side of the road when driving in UK or USA

Example: Nash Equilibrium

P2

P1

	left	centre	right
up	0,4	4,0	5,3
mid	4,0	0,4	5,3
down	3,5	3,5	6,6

Question: is there a Nash Equilibrium?

Example: Nash Equilibrium

P2

P1

	left	centre	right
up	0,4	4,0	5,3
mid	4,0	0,4	5,3
down	3,5	3,5	6,6

P1 best response:
- P2 left: P1 mid
- P2 centre: P1 up
- P2 right: P1 down

P2 best response:
- P1 up: P2 left
- P1 mid: P2 centre
- P1 down: P2 right

Nash Equilibrium: down, right

Nash Equilibrium

	i defects	*i* cooperates
j **defects**	**1,1**	**1,4**
j **cooperates**	**4,1**	$\boxed{\textbf{4,4}}$

- Problems:
  - *Not every scenario has* a pure strategy Nash equilibrium
  - *Some scenarios have more than one* pure strategy Nash equilibrium

Pareto Efficiency

- An outcome is *Pareto efficient* if no other outcome improves a player's utility without making someone else worse off.
- *Not really a solution!*
  - Example: Dividing a chocolate cake

Pareto Efficiency

	i defects	*i* cooperates
j defects	2,2	4,0
j cooperates	0,4	2,2

- *All outcomes are Pareto efficient!*

Back to the Prisoner's Dilemma

- Defecting is the *dominant strategy*!
- (D,D) is a *Nash equilibrium*!
- It is the "rational" choice!
- But *(C,C) is better than (D,D)*!

Cooperation in The Prisoner’s Dilemma

- 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 (there are good ones!)
  - The other prisoner is my twin!
  - The shadow of the future… (what happens if we play again?)

But what happens if we play again?

The iterative chocolate game..

Two teams… several rounds…

The Iterated Prisoner's Dilemma

- One answer: *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 infinitively repeated prisoner's dilemma* (Hurrah!)

Yet….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 seems to be the best strategy

Axelrod’s Tournament

- Suppose you play iterated prisoner’s dilemma against a *range* of opponents… What strategy should you choose, so as to maximize your overall payoff?
- Axelrod (1984) investigated this problem, with a computer tournament for programs playing the prisoner’s dilemma

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 and defection.
- JOSS:
  - As TIT-FOR-TAT, except periodically defect

TÉCNICO LISBOA

What strategies were the best?

Recipes for Success in Axelrod's Tournament

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

Split or Steal

- Not exactly the P.D.! (but similar)
- But players are communicating…

- http://www.youtube.com/watch?v=p3Uos2fzIJ0

- http://www.youtube.com/watch?v=S0qjK3TWZE8

Further Discussion

(and an interview with John Nash)