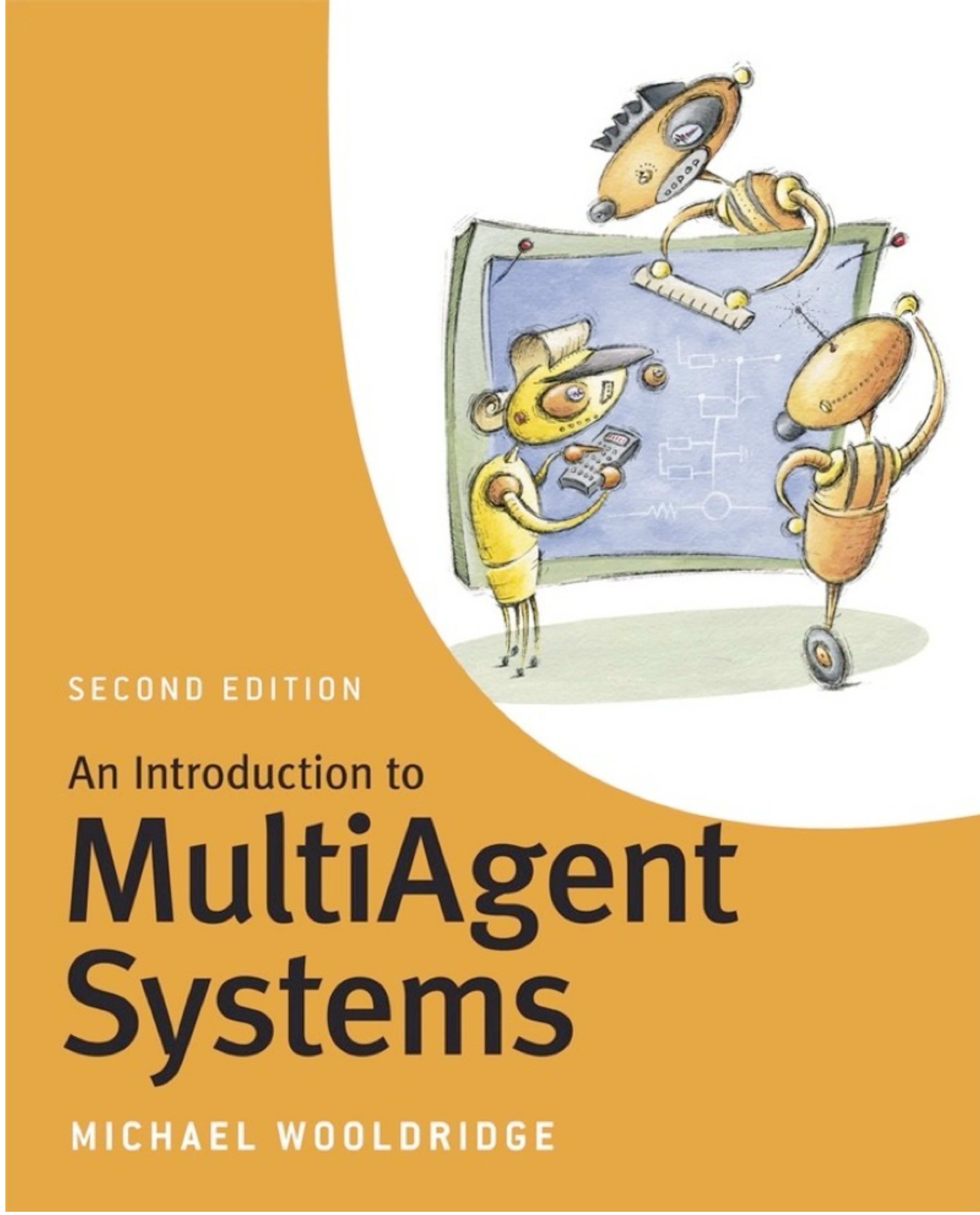
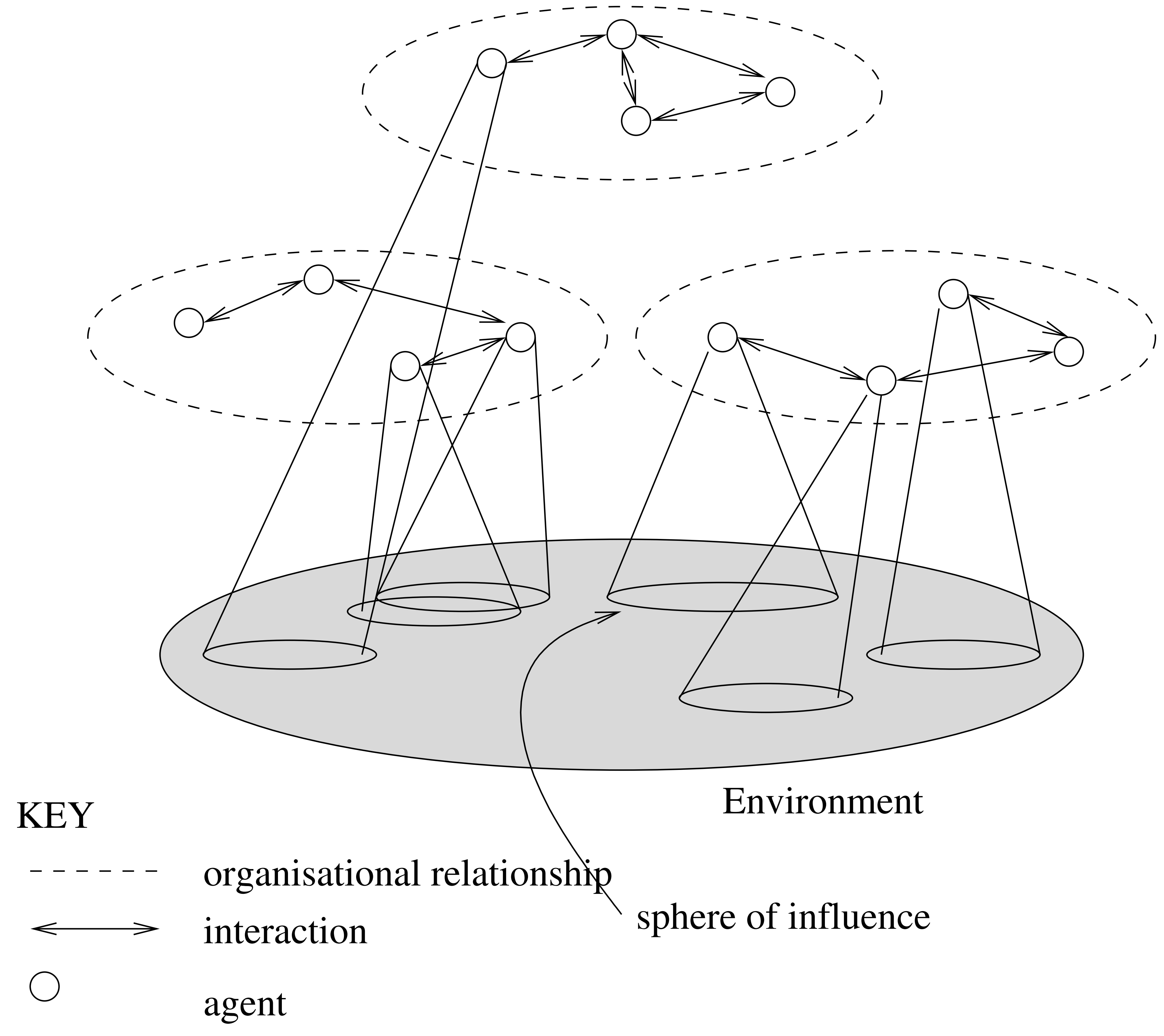
COMP310
Multi-Agent Systems
Chapter 11 - Multi-Agent Interactions

Dr Terry R. Payne
Department of Computer Science
UNIVERSITY OF LIVERPOOL



What are Multi-Agent Systems?

- A multiagent system contains a number of agents that:
  - interact through communication;
  - are able to act in an environment;
  - have different “spheres of influence” (which may coincide); and
  - will be linked by other (organisational) relationships.



- We will look at how agents decide how to interact in ***competitive*** situations.

Utilities and Preferences

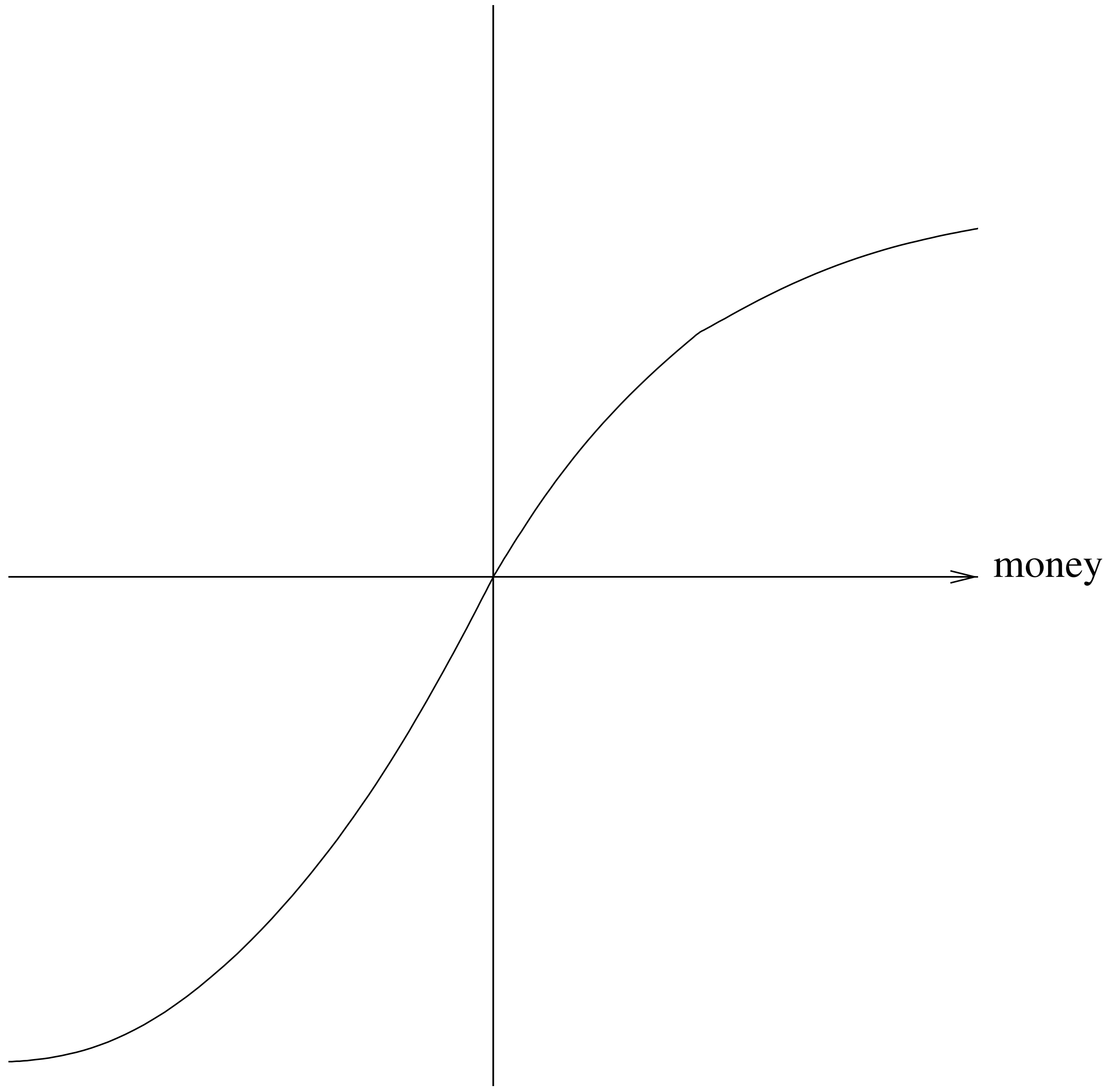
- Our Assumptions:
  - Assume we have just two agents: $Ag = \{i, j\}$
  - Agents are assumed to be ***self-interested*** i.e. they have preferences over how the environment is.
  - Assume $\Omega = \{\omega_1, \omega_2, \ldots\}$ is the set of "outcomes" that agents have preferences over.
- We capture preferences by ***utility functions***, represented as real numbers ($\mathbb{R}$):

$$u_i : \Omega \rightarrow \mathbb{R}$$
$$u_j : \Omega \rightarrow \mathbb{R}$$

- Utility functions lead to ***preference orderings*** over outcomes, e.g.:

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

- where ω and ω' are both possible outcomes of Ω



Utility is not money. Just a way to encode preferences.

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:
  - i.e. $Ac = \{C,D\}$, where
    - C ("***cooperate***") and
    - D ("***defect***")

- Environment behaviour given by ***state transformer function*** τ
  - (introduced in Chapter 2): $\tau : \underbrace{Ac}_{\text{agent } i\text{'s action}} \times \underbrace{Ac}_{\text{agent } j\text{'s action}} \to \Omega$

Multiagent Encounters

- Here is a state transformer function $\tau(i,j)$
  - This environment is sensitive to actions of both agents.

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

- With this state transformer, neither agent has any influence in this environment.

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

- With this one, the environment is controlled by j

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

Rational Action

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

$$u_i(\omega_1)=1 \quad u_i(\omega_2)=1 \quad u_i(\omega_3)=4 \quad u_i(\omega_4)=4$$

$$u_j(\omega_1)=1 \quad u_j(\omega_2)=4 \quad u_j(\omega_3)=1 \quad u_j(\omega_4)=4$$

  - With a bit of abuse of notation:

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

  - Then agent i's preferences are $(C, C) \succcurlyeq_i (C, D) \succ_i (D, C) \succcurlyeq_i (D, D)$
  - In this case, what should i do?

- i prefers all outcomes that arise through C over all outcomes that arise through D.
  - Thus C is the ***rational choice*** for i.

Payoff Matrices

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

In this case, i **cooperates** and gains a **utility of 4**; whereas j **defects** and gains a **utility of only 1**.

		i	
		defect	coop
j	defect	1 (upper), 1 (lower)	4 (upper), 1 (lower)
	coop	1 (upper), 4 (lower)	4 (upper), 4 (lower)

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

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) that agent i can play, 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 the game opposite, C dominates D for both players.

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

Dominant Strategies

- A rational agent will never play a ***dominated strategy***.
  - i.e, a strategy that is dominated (and thus inferior) by another
- So in deciding what to do, we can delete dominated strategies.
  - Unfortunately, there isn't always a unique un-dominated strategy.

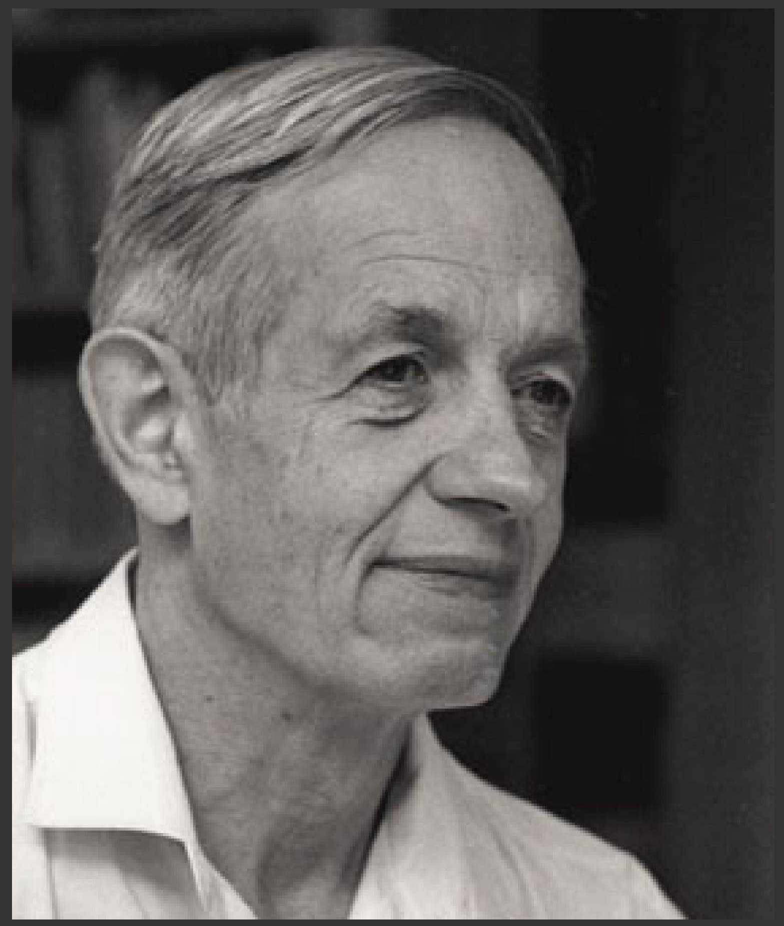
		i	
		defect	coop
j	defect	1 1	4 1
	coop	1 4	4 4

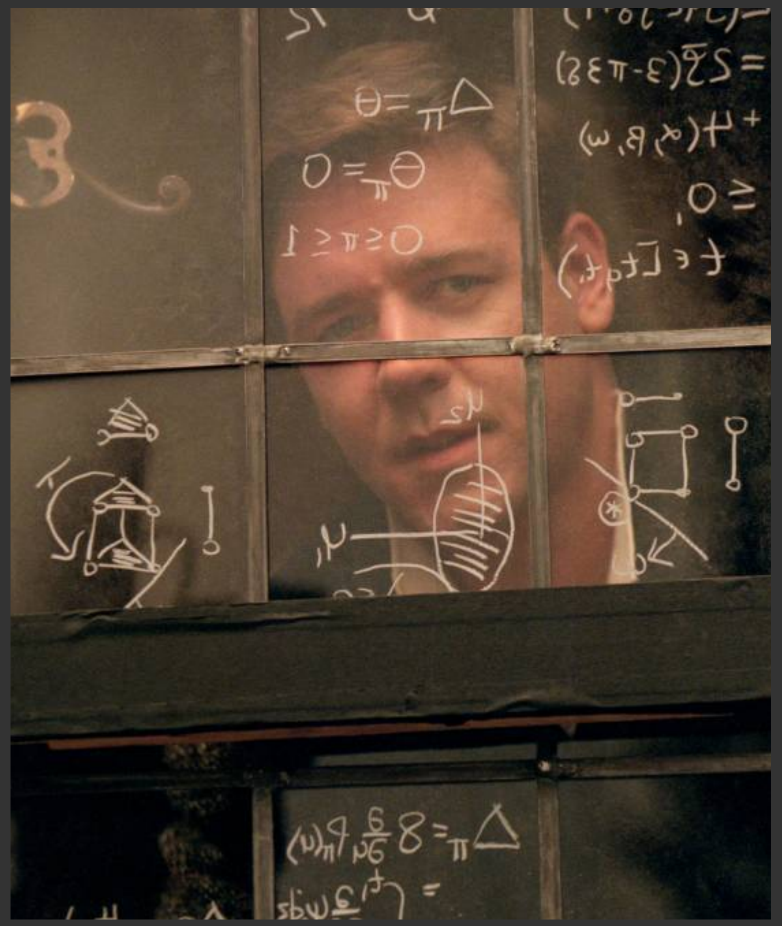
Nash Equilibrium

- In general, we will say that two strategies s_1 and s_2 are in Nash equilibrium (NE) if:
  - under the assumption that agent *i* plays s_1, agent j can do no better than play s_2;
    - I.e. *if **I** drive on the left side of the road, **you** can do no better than also driving on the left!*
  - under the assumption that agent j plays s_2, agent i can do no better than play s_1.
    - I.e. *if **you** drive on the left side of the road, **I** can do no better than also driving on the left!*
- Neither agent has any incentive to deviate from a ***Nash Equilibrium (NE)***.

John Forbes Nash

(Nobel Laureate in Economics)





Portrayed by Russel Crowe in the film "A Beautiful Mind"

Nash Equilibrium

- Consider the payoff matrix opposite:
  - Here the ***Nash equilibrium (NE)*** is (D, D).
  - In a game like this you can find the NE by cycling through the outcomes, asking if either agent can improve its payoff by switching its strategy.

		i	
		defect	coop
j	defect	5 3	1 2
	coop	0 2	0 1

- Thus, for example, (C, D) is not a NE because i can switch its payoff from 1 to 5 by switching from C to D.

		i	
		defect	coop
j	defect	5 3	1 2
	coop	0 2	0 1

Nash Equilibrium

- More formally:
  - A strategy (i^*, j^*) is a ***pure strategy Nash Equilibrium*** solution to the game (A, B) if:

$$\forall i,\ a_{i*,j*} \geq a_{i,j*}$$
$$\forall j,\ b_{i*,j*} \geq b_{i*,j}$$

- Unfortunately:
  - Not every interaction scenario has a ***pure strategy*** Nash Equilibrium (NE).
  - Some interaction scenarios have ***more than one pure strategy*** Nash Equilibrium (NE).

Nash Equilibrium

- The game opposite (upper) has two pure strategy NEs, *(C, C)* and *(D, D)*
  - In both cases, a single agent can't unilaterally improve its payoff.

		i	
		defect	coop
j	defect	5 3	1 2
	coop	0 2	3 3

- In the game opposite game (lower) has no pure strategy NE
  - For every outcome, one of the agents will improve its utility by switching its strategy.
  - We can find a form of NE in such games, but we need to go beyond pure strategies.

		i	
		defect	coop
j	defect	2 1	1 2
	coop	0 2	1 1

Mixed Strategy Nash equilibrium

- Matching Pennies
  - Players i and j simultaneously choose the face of a coin, either "heads" or "tails".
  - If they show the same face, then i wins, while if they show different faces, then j wins.

		i	
		heads	tails
j	heads	1 -1	-1 1
	tails	-1 1	1 -1

- NO pair of strategies forms a pure strategy NE:
  - whatever pair of strategies is chosen, somebody will wish they had done something else.

- The solution is to allow ***mixed strategies***:
  - play "heads" with probability 0.5
  - play "tails" with probability 0.5.

- This is a Mixed Nash Equilibrium strategy.

Mixed Strategy Nash equilibrium

- Consider the Game Rock/Paper/Scissors
  - Paper covers rock
  - Scissors cut paper
  - Rock blunts scissors
- This has the following payoff matrix

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

- What should you do?
  - ***Choose a strategy at random!***

Mixed Strategies

- A mixed strategy has the form
  - play α_1 with probability p_1
  - play α_2 with probability p_2
  - ...
  - play α_k with probability p_k.
  - such that $p_1+p_2+\cdots+p_k=1$.
- Nash proved that:
  - ***every finite game has a Nash equilibrium in mixed strategies.***

Nash's Theorem

*Nash proved that **every finite game has a Nash equilibrium in mixed strategies**. (Unlike the case for pure strategies.)*

So this result overcomes the lack of solutions; but there still may be more than one Nash equilibrium. . .

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.***

This game has one Pareto efficient outcome: (D, D)

		i	
		defect	coop
j	defect	5 3	1 2
	coop	0 2	0 1

There is no solution in which either agent does better

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).
  - It doesn't consider the benefits to individuals.
  - A very skewed outcome can maximise social welfare.

In both these games, *(C, C)* maximises social welfare

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

		i: defect	i: coop
j	defect	i: 2, j: 2	i: 1, j: 1
j	coop	i: 3, j: 3	i: 7, j: 0

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:

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

  - Zero sum encounters are bad news: for me to get +ve utility you have to get negative utility! ***The best outcome for me is the worst for you***!
  - ***Zero sum encounters in real life are very rare*** . . . but people tend to act in many scenarios as if they were zero sum.

- Most games have some room in the set of outcomes for agents to find (somewhat) mutually beneficial outcomes.

The Prisoner's Dilemma

- Payoff matrix for prisoner's dilemma:

		i defect	i coop
j	defect	2 2	1 4
	coop	4 1	3 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 (i.e. neither confess).

As neither men want to admit to being guilty, **cooperation** means **not** confessing!

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 (C,D) or (D,C), the confessor will be freed, and the other will be jailed for three years;*
- *if both confess (D,D), then each will be jailed for two years.*

Both prisoners know that if neither confesses (C,C), then they will each be jailed for one year.

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!
- This is why the Prisoners Dilemma game is interesting
  - The analysis seems to give us a ***paradoxical*** answer.

Solution Concepts

- The dominant strategy here is to defect.
- (D, D) is the only Nash equilibrium.
- All outcomes **except** (D, D) are Pareto optimal.
- (C, C) maximises social welfare.

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

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, file sharing;
  - in the UK — television licenses.
  - in the US — funding for NPR/PBS.
- The prisoner's dilemma is ubiquitous.
  - Can we recover cooperation?

Solution Concepts

- The dominant strategy here is to defect.
- *(D, D)* is the only Nash equilibrium.
- All outcomes **except** *(D, D)* are Pareto optimal.
- *(C, C)* maximises social welfare.

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

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!
  - Program equilibria and mediators
  - The shadow of the future. . .

We are not all Machiavelli

- Lots of “altruism” is something else:
  - Either there is some delayed reciprocity; or
  - There are mechanisms to punish defection.
- There is a reason why HMRC (or the IRS in the US) audits people’s taxes :-)
- Altruism may be something that makes us feel good
  - This is why we are prepared to pay for it.

“... We aren’t all that hard-boiled, and besides, people really do act altruistically ...”

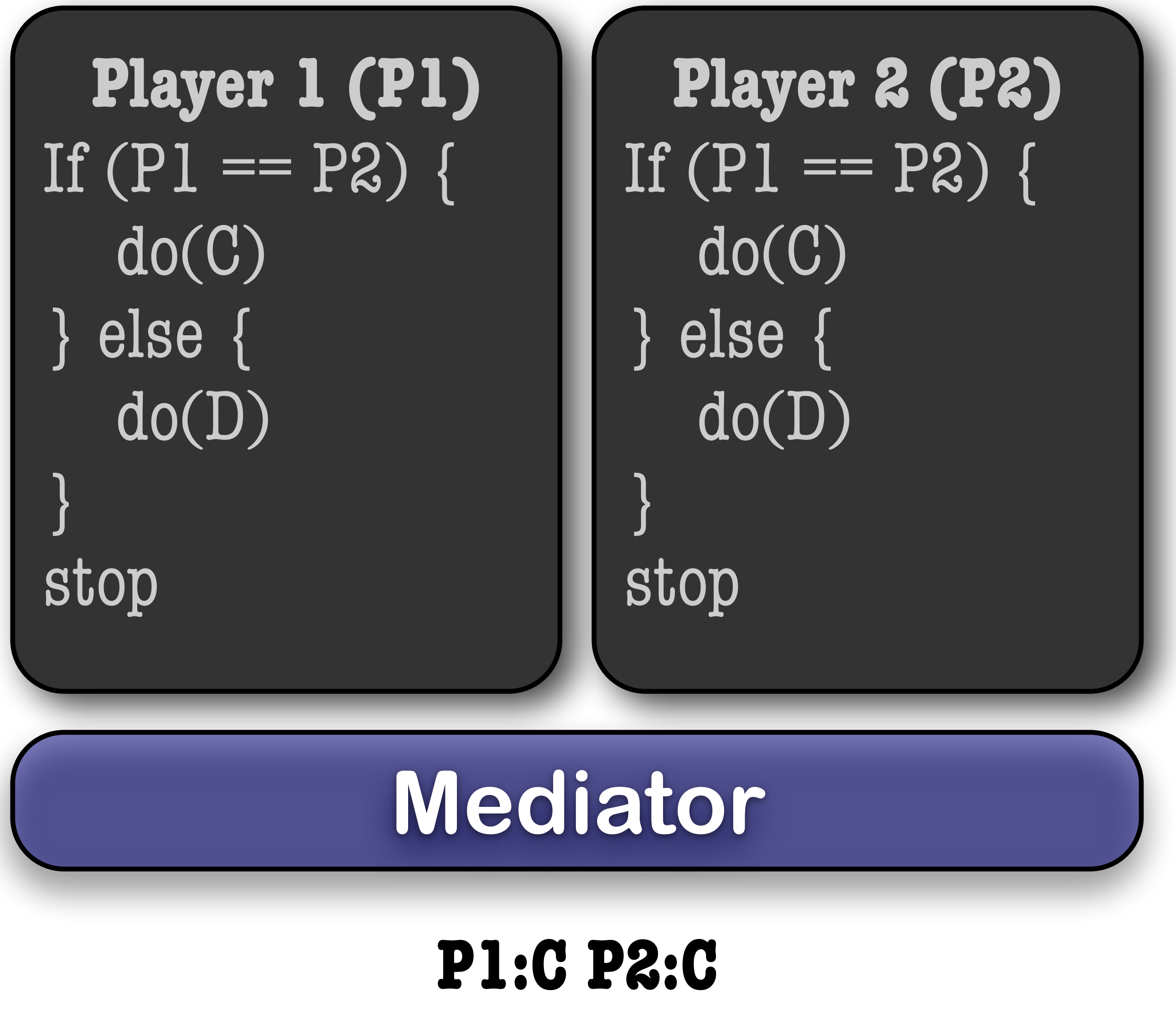
The Other Prisoner is My Twin

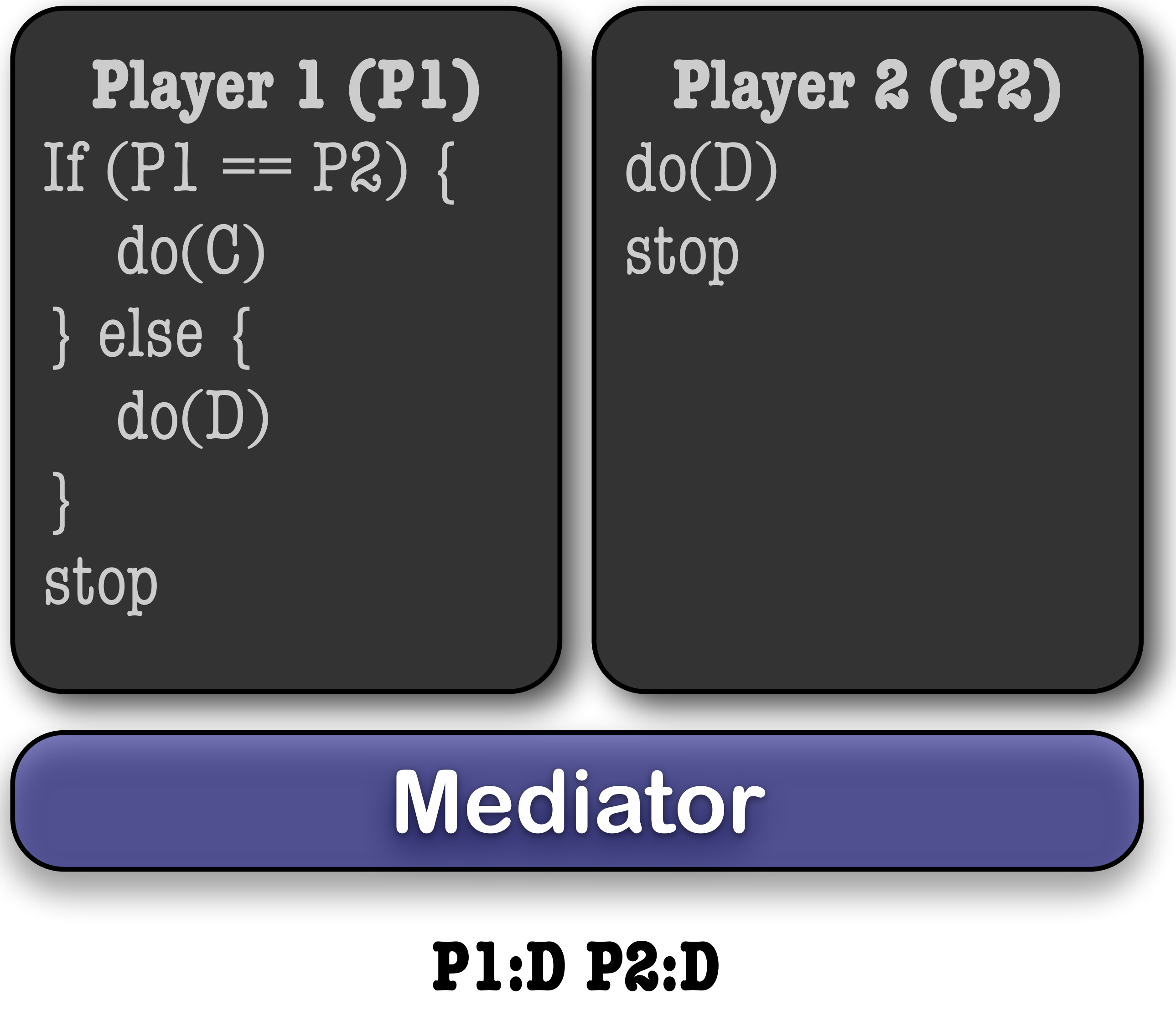
- Argue that both prisoner's will think alike and decide that it is best to cooperate.
  - If they are twins, they must think along the same lines, right?
  - (Or they have some agreement that they won't talk.)
- Well, if this is the case, we aren't really playing the Prisoner's Dilemma!
- Possibly more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

Program Equilibria

- The strategy you really want to play in the prisoner's dilemma is: ***I'll cooperate if he will***
  - Program equilibria provide one way of enabling this.

- Each agent submits a ***program strategy*** to a ***mediator*** which ***jointly executes*** the strategies.
  - Crucially, strategies can be ***conditioned on the strategies of the others***.
  - The best response to this program:
    - ***submit the same program***, giving an outcome of (C, C)!





The Iterated Prisoner's Dilemma (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.
    - 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 infinitely 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?

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.

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

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

- Surprisingly TIT-FOR-TAT for won.
  - But don't read too much into this :-)
- In scenarios like the Iterated Prisoner's Dilemma (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.

JOSS

As TIT-FOR-TAT, except periodically defect.

ALL-D

"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.

Recipes for Success in Axelrod's 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

- A group of hunters goes stag hunting.
  - If ***they all stay focussed on the stag***, they will catch it and all have a lot of food.
  - If ***some of them head off to catch rabbits***, the stag will escape.

- In this case the rabbit hunters will have some small amount of food and the (remaining) stag hunters will go hungry.
  - What should each hunter do?

> *"...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..."*
>
> *Mike Wooldridge*

Stag Hunt

Stag Hunt Payoff Matrix

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

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.

- Two Nash equilibrium solutions *(C, C)* and *(D, D)*.
  - If you know I'll co-operate, the best you can do is to co-operate as well.
  - If you know I'll defect, then that is the best you can do as well.
  - Social welfare is maximised by *(C, C)*.
  - The only Pareto efficient outcome is *(C, C)*.
- As usual with Nash equilibrium, theory gives us no real help in deciding what the other party will do.
  - Hence the worrying about the haircut.
- The same scenario occurs in mutinies and strikes.
  - We would all be better off if our hated captain is deposed, but if some of us give in, we will all be hanged.

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 the following payoff matrix:

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

jump = coop
stay in car = 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
  - Strategy pairs *(C, D)* and *(D, C)* are Nash equilibria.
    - If I think you will stay in the car, I should jump out.
    - If I think you will jump out of the car, I should stay in.
  - All outcomes except *(D, D)* are Pareto optimal.
  - All outcomes except *(D, D)* maximise social welfare.

Other Symmetric 2x2 Games

- Given the 4 possible outcomes of (symmetric) cooperate/defect games, there are 24 possible orderings on outcomes.
  - The textbook lists them all, but here we give the combinations
- These are more abstract descriptions of the games than the payoff matrices we considered.
- All payoff matrices consistent with these preferences orders are instances of the games.

First, cases with dominant solutions:

Cooperation dominates

$CC >_i CD >_i DC >_i DD$

$CC >_i CD >_i DD >_i DC$

Deadlock (You will always do best by defecting)

$DC >_i DD >_i CC >_i CD$

$DC >_i DD >_i CD >_i DD$

Games that we looked at in detail:

Prisoner's dilemma.

$DC >_i CC >_i DD >_i CD$

Chicken

$DC >_i CC >_i CD >_i DD$

Stag Hunt

$CC >_i DC >_i DD >_i CD$

Summary

- This chapter 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.

Class Reading (Chapter 11):

"The Evolution of Cooperation", R. Axelrod. Basic Books, New York, 1984.

This is a book, but is quite easy going, and most of the important content is conveyed in the first few chapters. Appart from anything else, it is beautifully written, and it is easy to see why so many readers are enthusiastic about it. However, read it in conjunction with Binmore's critiques, cited in the course text.