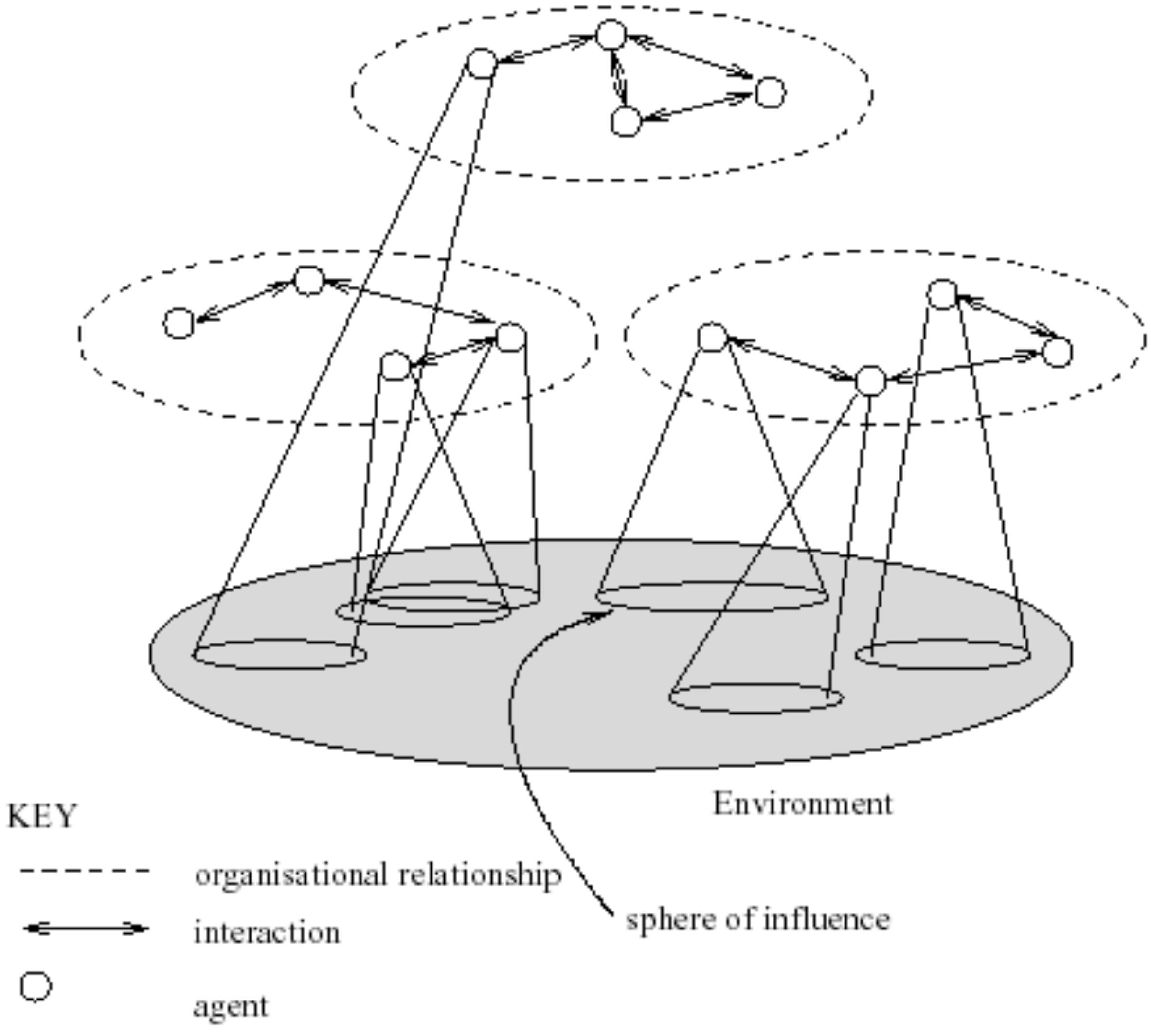
MULTIAGENT INTERACTIONS

What are *Multiagent* Systems?



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 (organizational) 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*:

$$u_i = \Omega \rightarrow \text{ú}$$
$$u_j = \Omega \rightarrow \text{ú}$$

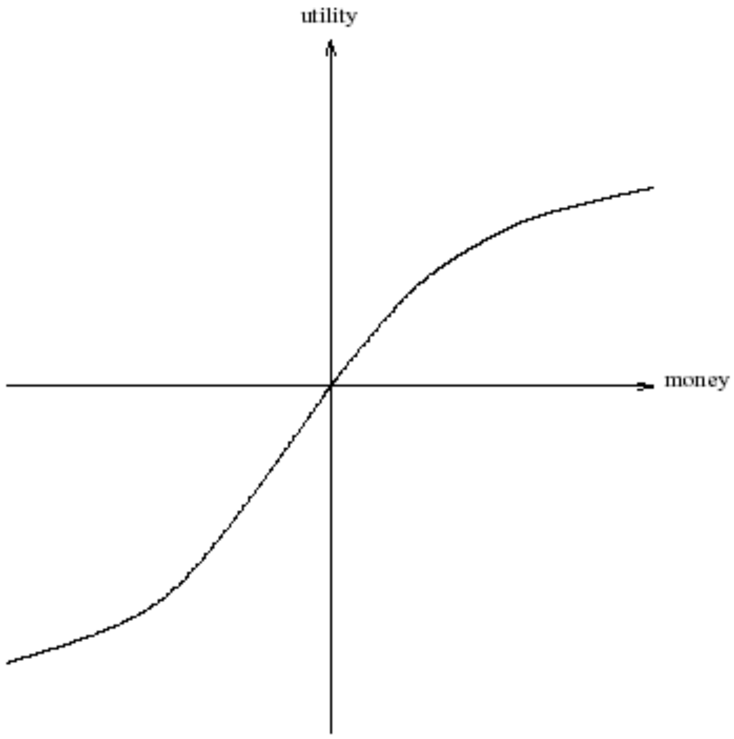
- Utility functions lead to *preference orderings* over outcomes:

ω š$_i$ ω' means $u_i(\omega)$ \$ $u_i(\omega')$

ω ™$_i$ ω' means $u_i(\omega) > u_i(\omega')$

What is Utility?

- Utility is *not* money (but it is a useful analogy)
- Typical relationship between utility & 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 behavior given by *state transformer function*:

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

Multiagent Encounters

- 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:

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

- 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:

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

- "*C*" is the *rational choice* for i.
(Because i prefers all outcomes that arise through *C* over all outcomes that arise through *D*.)

Payoff Matrices

- We can characterize 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*
- Agent j is the *row player*

Dominant Strategies

- Given any particular strategy (either *C* or *D*) 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, we can *delete dominated strategies*
- Unfortunately, there isn't always a unique undominated strategy

Nash Equilibrium

- In general, we will say that two strategies s_1 *and* s_2 are in Nash equilibrium 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 Nash equilibrium*
- Unfortunately:
  1. *Not every interaction scenario has a Nash equilibrium*
  2. *Some interaction scenarios have more than one Nash equilibrium*

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:

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

- 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

The Prisoner's Dilemma

- Payoff matrix for prisoner's dilemma:

		i	
		defect	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

The Prisoner's Dilemma

- 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!

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!
  - The other prisoner is my twin!
  - The shadow of the future…

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 infinititely repeated prisoner's dilemma* (Hurrah!)

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

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

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

Game of Chicken

- Consider another type of encounter — the *game of chicken*:

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

(Think of James Dean in *Rebel without a Cause*: 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.)

- Strategies (c,d) and (d,c) are in Nash equilibrium

Other Symmetric 2 x 2 Games

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