Agent-Based Systems

Michael Rovatsos
mrovatso@inf.ed.ac.uk

Lecture 8 – Multiagent Interactions

Where are we?

Last time . . .

- Coordination: managing interactions effectively
- Different methods for coordination
- Partial global planning: achieving a global view through information exchange
- Joint intentions: extending the BDI paradigm to include joint intentions, collective commitments and conventions
- Mutual modelling: taking the role of the other to predict their actions
- Norms and social laws: coordination through offline/emergent constraints on agent behaviour
- Multiagent planning and synchronisation, plan merging

Today . . .

- **Multiagent Interactions**

Multiagent interactions

- We have looked at agent communication, but not described how it is used in actual agent interactions
- In itself, communication does not have much effect on the agents
- Now, we are going to look at interactions in which agents *affect* each other through their actions
- Assume agents to have “spheres of influence” that they control in the environment
- Also, we assume that the welfare (goal achievement, utility) of each agent at least partially depends on the actions of others
- This part of the lecture will deal with what agents should *do* in the presence of other agents (which also do stuff)

Preferences and utilities

- We first need an abstract model of interactions
- Assume $O = \{o_1, \ldots o_n\}$ a set of possible outcomes (e.g. possible "runs" of the system until final states are reached)
- A **preference ordering** $\succ_i \subseteq O \times O$ for agent i is a total, antisymmetric, transitive relation on O, i.e.
  - $o \succ_i o' \Rightarrow o' \not\succ_i o$
  - $o \succ_i o' \wedge o' \succ o'' \Rightarrow o \succ_i o''$
  - $\forall o, o' \in O$ either $o \succ_i o'$ or $o' \succ_i o$
- Such an ordering can be used to express strict preferences of an agent over O (write $\succeq_i$ if also reflexive, i.e. $o \succeq_i o$)

Preferences and utilities

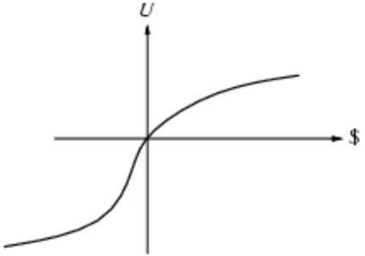
- Preferences are often expressed through a **utility function** $u_i : O \rightarrow \mathbb{R}$:

$$u_i(o) > u_i(o') \Leftrightarrow o \succ o', \quad u_i(o) \geq u_i(o') \Leftrightarrow o \succeq o'$$

- Utilities make representing preferences easier because the ordering follows naturally if we use real numbers
- Often, people falsely associate utility directly with money!
- Intuitively, the utility of money depends on how much money one already has
- Therefore, utility does not increase proportionally with monetary wealth

Preferences and utilities

- The utility of money:



- Empirical evidence suggests utility of money is often very close to logarithm function for humans
- This shows that utility function depends on agent's risk aversion attitude (value of additional utility depending on current "wealth")

Multiagent encounters

- Applying the above to a multiagent setting, we need to consider several agents' actions and the outcomes they lead to
- For now, restrict ourselves to two players and identical sets of actions
- Abstract architecture: state transformer function becomes

$$\tau : Ac \times Ac \rightarrow O$$

  where *Ac* are the actions of each of the two agents
- Outcome depends on other's actions!
- For pairs $(a_1, a_2), (a'_1, a'_2) \in Ac \times Ac$ we can write

$$(a_1, a_2) \succeq (a'_1, a'_2) \text{ iff } \tau(a_1, a_2) \succeq \tau(a'_1, a'_2)$$

  (similarly for $\succ$ and utilities $u_{1/2}(\tau(a_1, a_2)))$
- We consider agents to be rational if they prefer actions that lead to preferred outcomes

Example: 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.
- **Payoff matrix** for this game:

1 \ 2	C	D
C	(3,3)	(0,5)
D	(5,0)	(1,1)

Game theory

- Mathematical study of interaction problems of this sort
- Basic model: agents perform simultaneous actions (potentially over several stages), the actual outcome depends on the combination of action chosen by all agents
- **Normal-form games**: final result reached in single step (in contrast to **extensive-form games**)
  - Agents $\{1, \ldots, n\}$, S_i=set of (pure) **strategies** for agent i, $S = \times_{i=1}^{n} S_i$ space of **joint strategies**
  - Utility functions $u_i : S \to \mathbb{R}$ map joint strategies to utilities
  - A probability distribution $\sigma_i : S_i \to [0, 1]$ is called a **mixed strategy** of agent i (can be extended to joint strategies)
- Game theory is concerned with the study of this kind of games (in particular developing solution concepts for games)

Dominance and Best Response Strategies

- Two simple and very common criteria for rational decision making in games
- Strategy $s \in S_i$ is said to **dominate** $s' \in S_i$ iff

$$\forall s_{-i} \in S_{-i} \quad u_i(s, s_{-i}) \geq u_i(s', s_{-i})$$

  ($s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$, same abbrev. used for S)
- Dominated strategies can be safely deleted from the set of strategies, a rational agent will never play them
- Some games are solvable in **dominant strategy equilibrium**, i.e. all agents have a single (pure/mixed) strategy that dominates all other strategies

Dominance and Best Response Strategies

- Strategy $s \in S_i$ is a **best response** to strategies $s_{-i} \in S_{-i}$ iff

$$\forall s' \in S_i, s' \neq s \quad u_i(s, s_{-i}) \geq u_i(s', s_{-i})$$

- Weaker notion, only considers optimal reaction to a *specific* behaviour of other agents
- Unlike dominant strategies, best-response strategies (trivially) always exist
- Strict versions of the above relations require that "$>$" holds' for at least one s'
- Replace s_i/s_{-i} above by σ_i/σ_{-i} and you can extend the definitions for dominant/best-response strategies to mixed strategies

Nash Equilibrium

- Nash (1951) defined the most famous equilibrium concept for normal-form games
- A joint strategy $s \in S$ is said to be in (pure-strategy) **Nash equilibrium** (NE), iff

$$\forall i \in \{1, \ldots n\} \forall s_i' \in S_i \quad u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i})$$

- Intuitively, this means that no agent has an incentive to deviate from this strategy combination
- Very appealing notion, because it can be shown that a (mixed-strategy) NE always exists
- But also some problems:
  - Not always unique, how to agree on one of them?
  - Proof of existence does not provide method to actually find it
  - Many games do not have pure-strategy NE

Example

The Prisoner's Dilemma: Nash equilibrium is not Pareto efficient (or: no one will dare to cooperate although mutual cooperation is preferred over mutual defection)

2 / 1	C	D
C	(3,3)	(0,5)
D	(5,0)	(1,1)

General conditions on utilities: $DC \succ CC \succ DD \succ CD$ (from first player's point of view) and $u(CC) > \frac{u(DC)+u(CD)}{2}$

Example

The Coordination Game: No temptation to defect, but two equilibria (hard to know which one will be chosen by other party)

2 / 1	A	B
A	(1,1)	(-1,-1)
B	(-1,-1)	(1,1)

The Evolution of cooperation?

- In **zero-sum/constant-sum** games one agent loses what the other wins (e.g. Chess) ➡ no potential for cooperation
- Typical **non-zero sum game**: there is a potential for cooperation but how should it emerge among self-interested agents?
- This situation occurs in many real life cases:
  - Nuclear arms race
  - Tragedy of the commons
  - “Free rider” problems
- Axelrod’s tournament (1984): a very interesting study of such interaction situations
- Iterated Prisoner’s Dilemma was played among many different strategies (how to play against different opponents?)

The evolution of cooperation?

- In single-shot PD, defection is the rational solution
- In (infinitely) iterated case, cooperation is the rational choice in the PD
- But not if game has a fixed, known length (“backward induction” problem)
- TIT FOR TAT strategy performed best against a variety of strategies (this does not mean it is the best strategy, though!)
- Axelrod’s conclusions from this:
  - don’t be envious, don’t be the first to defect, reciprocate defection and cooperation (don’t hold grudges), don’t be too clever

Critique

While game-theoretic/decision-theoretic approaches are currently very popular, there is also some criticism:

- How far can we get in terms of cooperation while assuming purely self-interested agents?
  - Good for economic interactions but how about other social processes?
  - In a sense, these approaches assume “worst case” of possible agent behaviour and disregard higher (more fragile) levels of cooperation
- Although mathematically rigorous,
  - . . . the proofs only work under simplifying assumptions
  - . . . often don’t consider irrational behaviour
  - . . . can only deal with a “utilitised” world
- Relationship to goal-directed, rational reasoning (e.g. BDI) and to deductive reasoning complex and not entirely clear

Summary

- Discussed simple, abstract models of multiagent encounters
- Utilities, preferences and outcomes
- Game-theoretic models and solution concepts
- Examples: Prisoner's Dilemma, Coordination Game
- Axelrod's tournament its conclusions and critique
- Next time: **Social Choice**