Silvia Rossi

Multiagent Interactions

Lezione n. 8

Corso di Laurea:
Informatica

Insegnamento:
Sistemi
multi-agente

Email:
silrossi@unina.it

A.A. 2014-2015

Distributed Rational Decision Making – non-cooperative Game Theory

(W: 3, MAS: 3.1, 3.2)

Game theory is the mathematical study of interaction among independent, self-interested agents.

noncooperative game theory

What does it mean to say that an agent is self-interested?

- not that they want to harm other agents
- not that they only care about things that benefit them
- that the agent has its own description of states of the world that it likes, and that its actions are motivated by this description

•Focus on decision-making where each player's decision can influence the outcomes (and hence well-being) of other players.

• Each player must consider how each other player will act in order to make its optimal choice: hence strategic considerations

• If all players have the same preferences, then game theoretic analysis is essentially redundant: there is common purpose.

• If a system has one designer, or is "owned" by a single individual, we can usually assume common purpose.

What is a Game?

•A “game” in the sense of game theory is an abstract model of a particular scenario in which self-interested players interact.

• Abstract in the sense that we only include detail relevant to the decisions that players make:

- leads to claims that game theoretic models are “toy”
- aim is to isolate issues that are central to decision making.

• Game theory origins: study of parlor games (e.g., chess)

- such games are useful for highlighting key concepts
- but the term “game” conveys something trivial :-(

Non-cooperative vs. cooperative games

- Game theory is usually sub-divided into non-cooperative and cooperative versions.

- Non-cooperative game theory is bigger and better-known:
    - it concerns settings where players must act alone.
    - Solution concepts in non-cooperative game theory relate to individual action.

- Cooperative game theory is concerned with settings where players can make binding agreements to work together, allowing for teamwork, cooperation, joint action.

Utility theory:

- quanties degree of preference across alternatives
- understand the impact of uncertainty on these preferences
- utility function: a mapping from states of the world to real numbers, indicating the agent's level of happiness with that state of the world
- Decision-theoretic rationality: take actions to maximize expected utility.

Assumptions:

- humans are rational beings
- humans always seek the best alternative in a set of possible choices
- Why assume rationality?
- narrow down the range of possibilities
- predictability

Utility Theory based on:

- rationality
- maximization of utility
    - may not be a linear function of income

It is a quantification of a person's preferences with respect to certain objects.

Utilities and Preferences

Agents are assumed to be *self-interested*: **they have preferences over how the environment is**

Assume we have just two agents: $Ag = \{i, j\}$

Assume $\Omega = \{\omega_1, \omega_2, \ldots\}$ is the set of "outcomes" that agents have preferences over

We capture preferences by *utility functions*:

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

$u_j = \Omega \rightarrow \mathrm{R}$

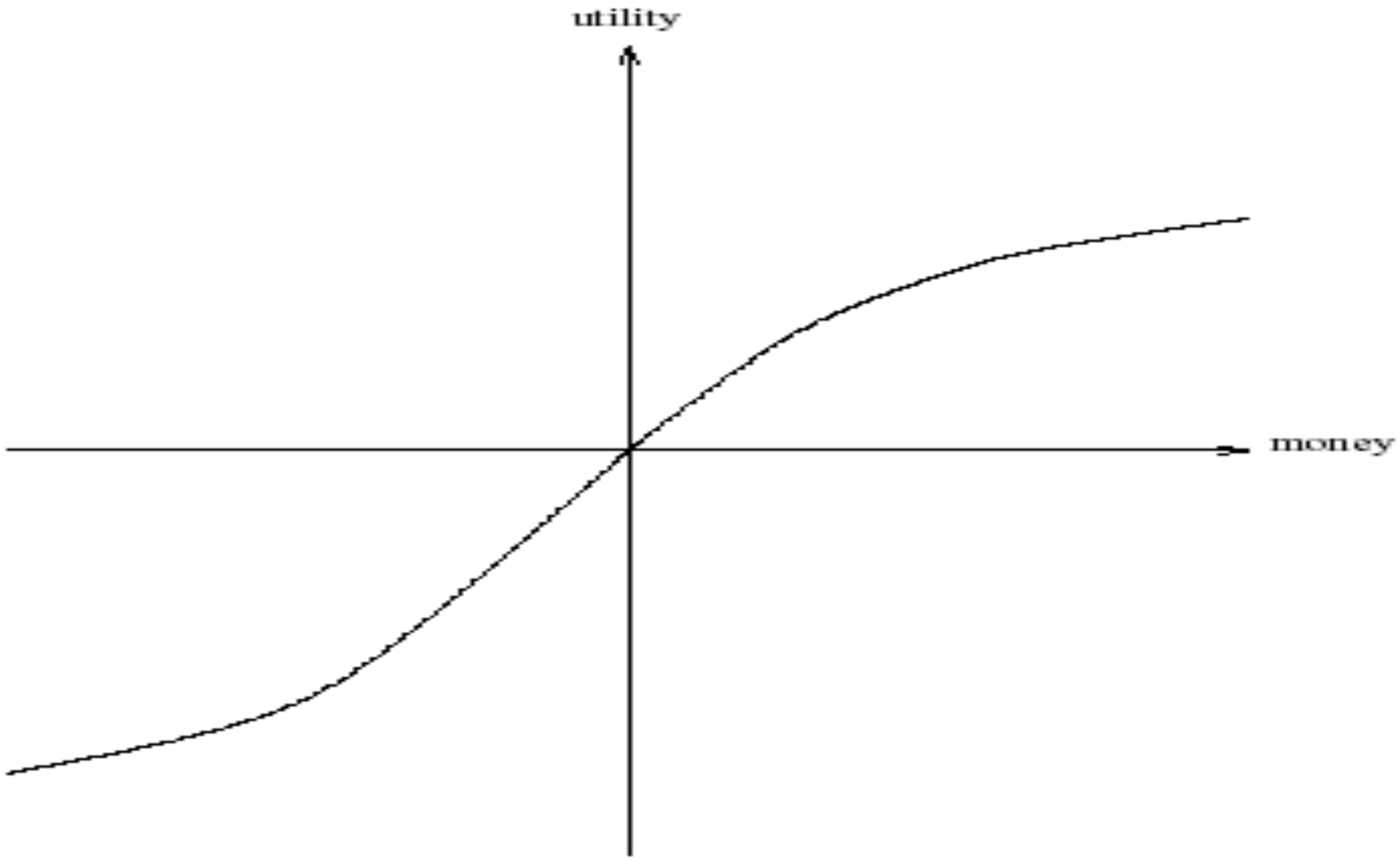
Utility functions lead to *preference orderings* over outcomes:

$\omega \geq_i \omega'$ means $u_i(\omega)$ >= $u_i(\omega')$

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

Utility is *not* money (but it is a useful analogy)

Typical relationship between utility & money:



Il soggetto deve essere in grado di:

Determinare l’insieme di scelta (le azioni);

Una relazione che lega le azioni alle conseguenze;

Ordinare tutte le conseguenze possibili;

Selezionare l’azione migliore.

1. Scelta in condizioni di certezza: ad ogni azione e’ associata una ed una sola conseguenza.

Nell’ambito del processo di scelta razionale questo problema diventa banale una volta che il decisore abbia definito l’insieme delle scelte ed ordinato tutte le possibili conseguenze.

2. Scelta in condizioni di incertezza: ad ogni azione sono associate piu’ conseguenze, in base ad una distribuzione di probabilita’ data.

Se la probabilita’ e’ oggettiva ->> SCELTA IN CONDIZIONI DI RISCHIO

Se la probabilita’ e’ soggettiva ->>SCELTA IN CONDIZIONI DI INCERTEZZA

PROBLEMA. Prendere una decisione in cui le conseguenze sono incerte e tale incertezza è quantificabile in modo non ambiguo.

L'incertezza dipende dalla presenza di più di uno stato di natura.

Si assume che le probabilità con cui i vari stati si verificano sia nota.

3. Scelta in condizioni di interazione strategica: ad ogni azione sono associate piu’ conseguenze, ma ora cio’ dipende dalle scelte effettuate da altri soggetti razionali.

Alice has three options:

- going to the club (c),
- going to a movie (m),
- or watching a video at home (h).

If she is on her own, Alice has a utility:

- u(c) = 100,
- u(m) = 50,
- u(h) = 50.

Bob is Alice’s nemesis;

- If Alice runs into Bob at the movies, she can try to ignore him and only suffers a disutility of 40;
- If she sees him at the club he will pester her endlessly, yielding her a disutility of 90.
- Bob prefers the club: he is there 60% of the time, spending the rest of his time at the movie theater.

Carol is Alice’s friend.

- Carol increases Alice’s utility for either activity by a factor of 1.5 (after taking into account the possible disutility of running into Bob).
- Carol can be found at the club 25% of the time, and the movie theater 75% of the time.

List Alice’s utility for each possible state of the world…

	$B = c$	$B = m$
$C = c$	15	150
$C = m$	10	100

$A = c$

	$B = c$	$B = m$
$C = c$	50	10
$C = m$	75	15

$A = m$

Alice's expected utility for c:

0.25(0.6*15 + 0.4*150) + 0.75(0.6*10 + 0.4*100) = 51.75

Alice's expected utility for m:

0.25(0.6*50 + 0.4*10) + 0.75(0.6*75 + 0.4*15) = 46.75

Alice's expected utility for h: 50.

Alice prefers to go to the club (though Bob is often there and Carol rarely is), and prefers staying home to going to the movies (though Bob is usually not at the movies and Carol almost always is).

Why would anyone argue with the idea that an agent's preferences could be described using a utility function as we just did?

why should a single-dimensional function be enough to explain preferences over an arbitrarily complicated set of alternatives?

Why should an agent's response to uncertainty be captured purely by the expected value of his utility function?

If o1 and o2 are outcomes

$o_1 \succeq o_2$ means o1 is at least as desirable as o2.

- read this as "the agent weakly prefers o1 to o2"

$o_1 \sim o_2$ means $o_1 \succeq o_2$ and $o_2 \succeq o_1$.

- read this as "the agent is indifferent between o1 and o2."

$o_1 \succ o_2$ means $o_1 \succeq o_2$ and $o_2 \not\succeq o_1$.

- read this as "the agent strictly prefers o1 to o2"

An agent may not know the outcomes of his actions, but may instead only have a probability distribution over the outcomes.

Definition (lottery)

A lottery is a probability distribution over outcomes. It is written $[p_1:o_1;\ p_2:o_2;\ \ldots;\ p_k:o_k]$

where the o_i are outcomes and $p_i > 0$ such that

$$\sum_i p_i = 1$$

The lottery species that outcome o_i occurs with probability p_i.

We will consider lotteries to be outcomes.

Definition (Completeness)

A preference relationship must be defined between every pair of outcomes:

$$\forall o_1 \forall o_2 \; o_1 \succeq o_2 \text{ or } o_2 \succeq o_1$$

Denition (Transitivity)

Preferences must be transitive:

$$\text{if } o_1 \succeq o_2 \text{ and } o_2 \succeq o_3 \text{ then } o_1 \succeq o_3$$

This makes good sense: otherwise

$$o_1 \succeq o_2 \text{ and } o_2 \succeq o_3 \text{ and } o_3 \succ o_1$$

An agent should be prepared to pay some amount to swap between an outcome they prefer less and an outcome they prefer more

Intransitive preferences mean we can construct a “money pump"!

Denition (Monotonicity)

An agent prefers a larger chance of getting a better outcome to a smaller chance:

If $o_1 \succ o_2$ and $p > q$ then

$$[p : o_1, 1 - p : o_2] \succ [q : o_1, 1 - q : o_2]$$

Let $P_l(o_i)$ denote the probability that outcome o_i is selected by lottery l.

For example, if $l = [0.3{:}o_1;\ 0.7{:}[0.8{:}o_2;\ 0.2{:}o_1]]$ then $P_l(o_1) = 0.44$ and $P_l(o_3) = 0$.

Definition (Decomposability ("no fun in gambling"))

$$\text{If } \forall o_i \in O,\ P_{\ell_1}(o_i) = P_{\ell_2}(o_i) \text{ then } \ell_1 \sim \ell_2$$

Definition (Substitutability)

If $o_1 \sim o_2$ then for all sequences of one or more outcomes $o_3, \ldots, o_k$ and sets of probabilities $p, p_3, \ldots, p_k$ for which $p + \sum_{i=3}^{k} p_i = 1$,
$[p : o_1, p_3 : o_3, \ldots, p_k : o_k] \sim [p : o_2, p_3 : o_3, \ldots, p_k : o_k]$.

Definition (Continuity)

Suppose $o_1 \succ o_2$ and $o_2 \succ o_3$, then there exists a $p \in [0, 1]$ such that $o_2 \sim [p : o_1, 1 - p : o_3]$.

Theorem (von Neumann and Morgenstern, 1944)

If an agent's preference relation satisfies the axioms Completeness, Transitivity, Decomposability, Substitutability, Monotonicity and Continuity then **there exists a function**

u : O -> [0; 1] with the properties that:

1. $u(o_1) \geq u(o_2)$ iff the agent prefers o_1 to o_2; and
2. $u([p_1 : o_1, \ldots, p_k : o_k]) = \sum_{i=1}^{k} p_i u(o_i)$.

Proof idea:

define the utility of the best outcome u(o) = 1 and of the worst u(o) = 0. Now define the utility of each other outcome o as the p for which .

$$o \sim [p : \overline{o}; (1-p) : \underline{o}]$$

Se le preferenze di un individuo soddisfano un insieme di assiomi allora le scelte di quell’individuo massimizzano l’utilità attesa e quindi sono razionali

Giochi strategici

… what if Bob hates Alice and wants to avoid her too,
…while Carol is indifferent to seeing Alice and has a crush on Bob?

In this case, we might want to revisit our previous assumption that Bob and Carol will act randomly without caring about what the other two agents do.

To study such settings, we turn to game theory

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*:

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

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*.)

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:

$$C, C \succeq_i C, D \qquad \succ_i \qquad D, C \succeq_i D, D$$

If you were agent i in this scenario, what would you choose to do – cooperate or defect?

Payoff Matrices

We can characterize the previous scenario in a *payoff matrix*:

Agent i is the *column player*

Agent j is the *row player*

No strategic thinking

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

Normal-form game

A (finite, n-person) normal-form game is a tuple (N,A, u), where:

- *N is a finite set of n players, indexed by i;*
- *$A = A1 \times \cdot \cdot \cdot \times An$, where Ai is a finite set of actions action available to player i.*
  - *Each vector $a = (a1, . . . , an) \in A$ is called an action profile;*
- *$u = (u1, . . . , un)$ where $ui : A \rightarrow R$ is a real-valued utility (or payoff)* function *for player i.*

A common-payoff game is a game in which for all action profiles $a \in A_1 \times \cdots \times A_n$ and any pair of agents i, j, it is the case that $u_i(a) = u_j(a)$

	a	b
a	1,1	0,0
b	0,0	1,1

pure coordination games or team games

no conflicting interests

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 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

Head and Tail

If pennies match then 2 pays 1,

if they differ then 1 pays 2

no equilibrium!

Game is strict competitive

	H	T
H	1,-1	-1,1
T	-1,1	1,-1

games can include elements of both coordination and competition

$N = \{1,2\}$

$A_1 = \{B,S\}$

$A_2 = \{B,S\}$

u_1, u_2 in figura

B: Bach

S: Strawinsky

Battle of the Sexes

	B	S
B	2,1	0,0
S	0,0	1,2

How will a rational agent will behave in any given scenario? Play. . .

- Strategies that maximise social welfare;
- Pareto optimal strategies;
- Nash equilibrium strategy;
- Dominant strategy.