AI Techniques (MSc)

Multi-Agent Decision Making: Game Theory

Delft University of Technology

Dr. Nils Bulling

October 18, 2016

The next two lecture

Lecturer: Dr. Nils Bulling

Teaching Assistant: Pietro Pasotti

Course email: ai-ewi@lists.tudelft.nl

Lectures:

- Tuesday, October 12
- Tuesday, October 19

Tutorials:

- Wednesday, November 13
- Wednesday, November 20

Test:

- Wednesday, November 20 (about both lectures!)

Please note the following points

1. The **exercises** and **assignments** are very **important** for the understanding and for passing the exam!
2. I would appreciate it to hear about any flaws, typos, comments and any other feedback which helps to improve the course.

Reading Material and Copyright Information I

- The slides are quite detailed and should in general be sufficient.
- General reading :

Russel, S. and Norvig, P. (2010).
Artificial Intelligence: a Modern Approach.
Prentice Hall, 3 edition. Chapters 17.5-17.6.

- More about game theory and mechanism design:

Shoham, Y. and Leyton-Brown, K. (2009).
Multiagent Systems - Algorithmic, Game-Theoretic, and Logical Foundations.
Cambridge University Press.

Osborne, M. and Rubinstein, A. (1994).
A Course in Game Theory.
MIT Press.

Outline

Next Section

Subsection I

Overview

In this lecture we shall cover the following topics:

1. **Normal Form Games**: one shot games
2. **Extensive Form Games**: games with sequential actions
3. **Mechanism Design** and **Auctions**: designing game rules

In the lecture we will treat mechanism design on a rather informal level. Therefore, this section is split into to parts: an informal and a more formal one. For the exam, you need to understand the basic ideas.

Subsection I

Classical quantitative decision making is concerned with selecting the best option among many alternatives.

- actions (=options) of an agent: $Act = \{a_1, a_2, \ldots\}$
- set of outcomes $\mathrm{Out} = \{o_1, o_2, \ldots\}$
- utility function $\mathrm{ut} : \mathrm{Out} \to \mathbb{R}$ (defines how good an outcome is)

Can we simply select the action which maximizes ut?

In deterministic environments the agent can simply choose the action yielding the best outcome. Many problems are **non-deterministic from the agent's perspective**.

In the following we will often assume finite sets of outcomes and actions to simplify the presentation.

Non-determinism can be modelled by probabilities. In the following we focus on finite sets of outcomes.

Remark 1.1 (Probability distribution)

A **(discrete) probability distribution** *over a finite set* Out *is a function* $\mathrm{pr}: 2^{\mathrm{Out}} \to [0,1]$.

$$\mathrm{pr}(\{o_1, \ldots, o_i\}) = p\text{: likelihood that the event } \{o_1, \ldots, o_i\} \text{ occurs is } p \in [0,1].$$

As Out *is finite, we can simply assign a probability to each element* $o \in \mathrm{Out}$. *We write* $p(o)$ *for* $p(\{o\})$.

(There are more properties a probability distribution has to satisfy, cf. Definition 2.16)

Suppose $\mathrm{Out} = \{1, \ldots, 6\}$ is the outcome of a non-manipulated dice.

- Probability of "dice shows i" is $\mathrm{pr}(i) = \frac{1}{6}$.
- Probability of "an even number" is $\mathrm{pr}(\{2,4,6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$.

- Effect of an action: probability distribution.
- **probability distribution** $\mathrm{pr}(\cdot|a)$ over outcomes Out, for $a \in Act$ (conditional probability)
- A **rational decision maker** selects an action a which maximizes the agent's expected utility (**MEU**):

$$\operatorname{argmax}_{\mathbf{a} \in Actions} \sum_{\mathrm{o} \in \mathrm{Out}} \underbrace{\mathrm{ut(o)}}_{\text{Utility of } o} \cdot \underbrace{\mathrm{pr(o}|\mathbf{a})}_{\text{Probability of } o}$$

Example 1.1 (Classical decision making)

Alice (A) has ordered a book and is waiting for delivery.

- She does not remember the declared address: home or office
- Actions: "go home" (action a_h), "go to the office" (action a_o).
- Delivery address is not well readable except for Alice's name.
- Postman Bob (B) knows both possible addresses.
- Outcomes: $Out = \{HO, OO, OH, HH\}$
  - Outcome HO: Alice is at home (H) and the package is delivered to the office (O), etc.
- Bob plays a passive role and "randomly chooses" where to deliver the package.
- All this, the **structure of the setting is commonly known to all players**.

Example continued

- Likelihood of the outcomes from Alice's perspective:
  - $\mathrm{pr}(HO|a_h) = \mathrm{pr}(HH|a_h) = 0.5$;
  - $\mathrm{pr}(OO|a_o) = \mathrm{pr}(OH|a_o) = 0.5$;
  - zero for all other combinations.
- Alice's preference:
  1. receive the package.
  2. being at the office over being at home.
- **Utility function**:
  $\mathrm{ut}_A(OO) = 4$
  $\mathrm{ut}_A(HH) = 3$
  $\mathrm{ut}_A(OH) = 2$
  $\mathrm{ut}_A(HO) = 1$
- **Expected utility**:
  - for a_o is $3 = 4 \cdot 0.5 + 2 \cdot 0.5$, and
  - for a_h is $2 = 3 \cdot 0.5 + 1 \cdot 0.5$.
- **If Alice is rational she will go to the office.**

The previous example is about classical decision making as Bob did not play an active role.

In **multi-agent decision making** an agent's decision depends on and influences other agents' decisions.

Subsection I

Example 1.2 (Multi-agent decision making)

- Bob plays an active role: he **prefers to deliver package to Alice's home address**; if Alice would be at home it would be even better.
- Utility function: $\text{ut}_B(HH) = 4$, $\text{ut}_B(OH) = 3$ and $\text{ut}_B(OO) = \text{ut}_B(HO) = 0$.
- Presentation in a payoff matrix where directly assign utilities to action profiles rather than outcomes:

A\B	a_o	a_h
a_o	(4, 0)	(2, **3**)
a_h	(1, 0)	(3, **4**)

- Bob's dominant strategy: deliver package to Alice's home.
- **Bob's decision affects Alice's decision.** Why?
- **Alice is better off going home, too**. Although she prefers being in the office.

Example 1.3 (Multi-agent decision making)

We modify example 1.2

- Suppose Bob is more social and prefers an outcome at which Alice receives the package:

A\B	a_o	a_h
a_o	(4, **3**)	(2, 2)
a_h	(1, 1)	(3, **4**)

- Where should Alice an Bob go to maximize their payoff?
- Two stable points yielding outcomes OO and HH:

  No player can unilaterally deviate from his/her action to obtain a better utility. ($\rightsquigarrow$ **Nash equilibrium**)

- How to coordinate to reach one of these outcomes?

A game with more complex decisions:

Example 1.4 (Pirates and Gold—a game in extensive form)

- Five rational pirates a, b, c, d, e
- 100 gold coins
- the pirates have a rank: a is highest, e lowest
- Distribute coins such that
  - highest ranked pirate proposes a distribution
  - each pirates decide to accept or to reject the proposal, majority decides (highest ranked pirate breaks ties)
  - if rejected the proposer pirate is killed and the next highest ranked pirate makes a proposal
- Preferences of pirates (in this order and additive):
  1. stay alive
  2. maximize the number of gold coins
  3. kill other pirates

What should Pirate a propose?

Next Section

1. Introduction
2. Simple Decisions: Normal Form Games
3. Sequential Decisions: Extensive Form Games
4. A Short Introduction to Mechanism Design
5. A More Formal Introduction to Mechanism Design

Subsection I

2 Simple Decisions: Normal Form Games

Normal Form Games

Definition 2.1 (n-Person Normal Form Game)

A finite **n-person normal form game** is a tuple $\langle \text{Agt}, \mathcal{A}ct, \textbf{ut} \rangle$, where

- $\text{Agt} = \{1, \ldots, i, \ldots, n\}$ is a finite set of **players** or **agents**.
- $\mathcal{A}ct = A_1 \times \cdots \times A_i \times \cdots \times A_n$ where A_i is the finite set of actions available to player i. $\vec{a} \in \mathcal{A}ct$ is called **action profile**.
- $\textbf{ut} = \langle \textbf{ut}_1, \ldots, \textbf{ut}_i, \ldots, \textbf{ut}_n \rangle$ where $\textbf{ut}_i : \mathcal{A}ct \to \mathbb{R}$ is a utility (payoff) function for player i.

Important: all elements of the game are **commonly known** among all players.

We map action profiles directly to payoffs. (Later we will ()again) first consider outcomes and map them to payoffs.)

The higher the payoff the better. Note that the utility function is defined on action profiles.

Games can be represented graphically using an ***n*-dimensional payoff matrix**. Here is a generic picture for 2-player, 2-action games:

		Player 2	
		a_2^1	a_2^2
Player 1	a_1^1	$\mathbf{ut}_1(a_1^1, a_2^1), \mathbf{ut}_2(a_1^1, a_2^1)$	$\mathbf{ut}_1(a_1^1, a_2^1), \mathbf{ut}_2(a_1^1, a_2^2)$
	a_1^2	$\mathbf{ut}_1(a_1^2, a_2^2), \mathbf{ut}_2(a_1^2, a_2^2)$	$\mathbf{ut}_1(a_1^2, a_2^1), \mathbf{ut}_2(a_1^2, a_2^2)$

Example 2.2 (Coordination game formalized)

We would like to formalize the following decision situation: a married couple looks for evening entertainment. They prefer to go out together. This can graphically be modelled by the following game:

		Wife	
		Theatre	Restaurant
Husband	Theatre	(2,2)	(0,0)
	Restaurant	(0,0)	(1,1)

Example continued

The graphical representation of the game can be **formalized as the normal form game** $\langle \mathrm{Agt}, \mathcal{A}ct, \mathbf{ut} \rangle$ with:

- $\mathrm{Agt} = \{H, W\}$ (husband (H) and wife (W))
- $\mathcal{A}ct_1 = \mathcal{A}ct_2 = \{T, R\}$ (theatre (T) and restaurant (R))
- $\mathrm{ut}_H((T,T)) = \mathrm{ut}_W((T,T)) = 2$,
  $\mathrm{ut}_H((R,R)) = \mathrm{ut}_W((R,R)) = 1$,
  $\mathrm{ut}_H((x,y)) = \mathrm{ut}_W((x,y)) = 0$ for $x \neq y$

What we are really after are **strategies**.

Definition 2.3 (Pure strategy)

A **pure strategy** s_i for a player is a particular action that is chosen and then played.

A **pure strategy profile** s is just a sequence of pure strategies $\vec{s} = (s_1, \ldots, s_n)$, one for each agent.

In the setting of pure strategies we often use the notation s and a interchangeably.

Important: in normal form games, agents choose their action:

- independently,
- simultaneously, and
- without communication.

Definition 2.4 (Notation a_{-i}, A_{-i})

Let $\vec{a} = \langle a_1, a_2, \ldots, a_n \rangle$. We use the following notation

- $\vec{a}_{-i}$ to denote the strategy profile $\langle a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n \rangle$. the strategies of all opponents of agent i are fixed.
- For a set of strategies A, we denote by $A_{-i} = \{\vec{a}_{-i} \mid \vec{a} \in A\}$.
- For ease of notation, we also use $\mathbf{ut}_i(a_i, \vec{a}_{-i})$ to denote $\mathbf{ut}_i(\langle a_1, a_2, \ldots, a_n \rangle)$, thus

$$\langle a_1, a_2, \ldots, \boldsymbol{a_i}, \ldots, a_n \rangle \;=\; \langle \boldsymbol{a_i}, \vec{a}_{-i} \rangle.$$

  In the last vector, although the $\boldsymbol{a_i}$ is written in the first entry, we mean it to be inserted at the i'th place.

Definition 2.5 (Common Payoff Game)

A **common payoff game (team game)** is a game in which for all action profiles $\vec{a} \in A_1 \times \ldots \times A_n$ and any two agents i, j the following holds: $\mathbf{ut}_i(\vec{a}) = \mathbf{ut}_j(\vec{a})$.

In such games agents have no conflicting interests.
While a team game is on one side of the spectrum, there is another type of games which is on the opposite side:

Definition 2.6 (Strictly competitive game)

A 2-player strategic normal form game is called **strictly competitive**, if for each action profile $\vec{a} \in A_1 \times A_2$: $\mathbf{ut}_1(\vec{a}) + \mathbf{ut}_2(\vec{a}) = 0$.

Example 2.7 (Matching Pennies Game)

Two players display one side of a penny (head or tails). Player 1 wins the penny if they display the same, player 2 wins otherwise.

		Player 2	
		Head	Tails
Player 1	Head	(1, -1)	(-1, 1)
	Tails	(-1, 1)	(1, -1)

- Strictly competitive
- No stable points, ($\rightsquigarrow$ i.e. no Nash equilibria in pure strategies)

Types of Strategic Interactions

We analyse basic types of games which model different types of strategic interaction based on:

- cooperation and competition
- coordination and non-coordination

Example 2.8 (Coordination game)

Married couple looks for evening entertainment. They prefer to go out together.

		Wife	
		Theatre	Restaurant
Husband	Theatre	(2,2)	(0,0)
	Restaurant	(0,0)	(1,1)

- Game is (fully) **cooperative**
- Intuitively, no coordination is needed to obtain best outcome (which is Pareto optimal)
- Two **stable points** (Theatre,Theatre) and (Restaurant,Restaurant) $\rightsquigarrow$ Nash equilibrium
- How to decide for a stable point? It's easy in this case.

Example 2.9 (Battle of the Sexes)

Another married couple looks for evening entertainment. They also prefer to go out together, but have different views about what to do.

		Wife	
		Bach	Stravinsky
Husband	Bach	(4,3)	(2,2)
	Stravinsky	(1,1)	(3,4)

- Game has a minor competitive flavour
- Two Nash equilibria: (Bach, Bach) and (Stravinsky,Stravinsky)
- They need to **coordinate** (to agree on a NE)
- Needed for coordination: communication and agreements (some must sacrifice something)

Subsection I

We need to compare protocols. Each such protocol leads to a solution. We determine how good these solutions are according to different criteria:

Social Welfare: sum of all utilities

Pareto Optimality: A strategy profile s is Pareto-optimal, iff there is no strategy which is at least as good as the current one for all agents and strictly better for at least one agent.

> Strategy profile s is **Pareto optimal** iff
>
> there is no strategy profile $\boldsymbol{s'}$ with:
>
> (1) $\exists$ agent i : $\mathbf{ut}_i(\boldsymbol{s'}) > \mathbf{ut}_i(\boldsymbol{s})$
>
> (2) $\forall$ agents i : $\mathbf{ut}_i(\boldsymbol{s'}) \geq \mathbf{ut}_i(\boldsymbol{s})$.

Stability:

Case 1: The best strategy of an agent does not depend on the others.
Such strategies are called **dominant**.

Case 2: Strategy of an agent depends on the others.
The profile $s^*_{\mathrm{Agt}} = \langle s^*_1, s^*_2, \ldots, s^*_{|\mathrm{Agt}|} \rangle$ is called a **Nash-equilibrium**, iff

$\forall i$: s^*_i **is an optimal strategy** for agent i if all the others choose $\langle s^*_1, s^*_2, \ldots, s^*_{i-1}, s^*_{i+1}, \ldots, s^*_{|\mathrm{Agt}|} \rangle$.

Example 2.10 (Prisoners Dilemma)

Two prisoners are suspected of a crime (which they both committed). They can choose to (1) **not to confess** (to cooperate) with each other (or (2) **confess** (giving evidence that the other was involved).

		Prisoner 2: don't confess	Prisoner 2: confess
Prisoner 1	don't confess	(4,4)	(0,5)
	confess	(5,0)	(1,1)

- a higher utility means less years in prison
- incentive to cooperate $(4, 4)$, but thread of a "free-rider"
- competitive elements

Example continued

		Prisoner 2	
		don't confess	confess
Prisoner 1	don't confess	(4,4)	(0,5)
	confess	(5,0)	(1,1)

- **Social welfare (profiles):** no one confesses
- **Pareto optimal (profiles):** all are Pareto optimal, except when both confess.
- **Dominant strategy (profiles):** Both confess.
- **Nash Equilibrium (profiles):** Both confess.

What if a prisoner wants to **minimize** the worst case outcome (regardless of the behaviour of the other prisoner)? In that case the agent should confess as well.

Maxmin strategy: Risk-averse players can use this strategy to maximise their outcome of worst-case play.

Minmax strategy: Punish another player as much as possible.

Definition 2.11 (Pure Maxmin strategy)

Given a game $\langle \mathrm{Agt}, \mathcal{A}ct, \mathbf{ut} \rangle$, the **pure maxmin strategy** of player i is a **strategy** that maximises the guaranteed payoff of player i, no matter what the other players $-i$ do:

$$s_i^{\text{maxmin}} = \arg\max_{a_i \in A_i} \min_{a_{-i} \in A_{-i}} \mathbf{ut}_i(a_i, a_{-i})$$

The **maxmin value** for player i is $\max_{a_i \in A_i} \min_{a_{-i} \in A_{-i}} \mathbf{ut}_i(a_i, a_{-i})$.

The **minmax strategy** for player i is

$$s_i^{\text{minmax}} = \arg\ \min_{a_i \in A_i} \max_{a_{-i} \in A_{-i}} \mathbf{ut}_{-i}(a_i, a_{-i})$$

and its **minmax value** is $\min_{a_i \in A_i} \max_{a_{-i} \in A_{-i}} \mathbf{ut}_{-i}(a_i, a_{-i})$.

What kind of **rationality assumptions** are made about the other player? What is the strategy profile consisting only of maxmin strategies?

		Wife: Bach	Wife: Stravinsky
Husband	Bach	(4,3)	(2,2)
Husband	Stravinsky	(1,1)	(3,4)

The outcome is $(Bach, Stravinsky)$. Thus, players **do not take into consideration the rational behavior** of other players!

What would be the **best response** of the Wife if the Husband played its maxmin strategy 'Bach'?

Definition 2.12 (Best Response to a Profile)

Given a strategy profile $\vec{s}_{-i} = \langle s_1, s_2, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n \rangle$, a **best response of player i to $\vec{s}_{-i}$** is **any strategy s_i^*** $\in A_i$ such that

$$\mathbf{ut}_i(s_i^*, \vec{s}_{-i}) \geq \mathbf{ut}_i(s_i, \vec{s}_{-i})$$

for all strategies $s_i \in A_i$.

Is a best response **unique**?

Definition 2.13 (Nash Equilibrium (NE))

A strategy profile $\vec{s^*} = \langle s_1^*, s_2^*, \ldots, s_n^* \rangle$ is a **Nash equilibrium** if for all agent i,

$$s_i^* \text{ is a best response to } \vec{s^*}_{-i} = \langle s_1^*, s_2^*, \ldots, s_{i-1}^*, s_{i+1}^*, \ldots, s_n^* \rangle.$$

Example 2.14 (Battle of the Sexes)

		Wife	
		Bach	Stravinsky
Husband	Bach	(4,3)	(2,2)
	Stravinsky	(1,1)	(3,4)

- Two **Nash equilibria**: (Bach, Bach) and (Stravinsky,Stravinsky)
- They need to **coordinate** (to agree on a NE)
- Needed for coordination: communication and agreements (some player must sacrifice something)

In the case of finite, strictly competitive 2-player games there is a relation to Nash equilibria. The following famous theorem counts as the foundation of game theory:

Theorem 2.15 (von Neumann (1928))

Let G be a finite 2-person strictly competitive normal form game and let (s_1, s_2) be a Nash equilibrium in G. Then, the following hold:

1. *s_1 and s_2 is a maxmin strategy of player 1 and 2, respectively.*
2. - *The **maxmin value** for one player **is equal** to the **minmax value** of the other. The maxmin of player 1 is called value of the game.*
   - *Thus, **all Nash equilibria yield the same payoffs**.*

Existence of Nash Equilibria

- In Example 2.7 we have seen that (pure) Nash equilibria do not have to exist.
- It was shown that **all finite games with specific mathematical properties have a Nash equilibrium**.
- Later, we will consider a setting in which these properties are naturally satisfied.
- For his work, Nash was awarded the Nobel prize.

Subsection I

Mixed Strategies

A (pure) strategy is a **deterministic plan**. What about randomness?

Think about bluffing in Poker.

Mixed strategy: Probability distribution over pure strategies.

"I bluff 10% of the time."

In the following we focus on finite sets of outcomes. 2^Z denotes the powerset of a set Z.

Definition 2.16 (Probability distribution)

A **(discrete) probability distribution** over a finite set Z is a function $\mathrm{pr} : 2^Z \to [0,1]$ such that

1. $p(Y) \geq 0$ for all $Y \subseteq Z$,
2. $p(Z) = 1$, and
3. $p(X \cup Y) = p(X) + p(Y)$ for all disjoint $X, Y \subseteq Z$.

$\mathrm{pr}(X)$ defines the likelihood of event X to occur. We write $p(x)$ for $p(\{x\})$. We write $\Pi(Z)$ to refer to all probability distributions over Z.

As Z is finite, we can simply assign a probability to each element $z \in Z$. We write $p(z)$ for $p(\{z\})$.

Suppose $\mathrm{Out} = \{1, \ldots, 6\}$ is the outcome of a non-manipulated dice.

- Probability of "dice shows i" is $\mathrm{pr}(i) = \frac{1}{6}$.
- Probability of "an even number" is $\mathrm{pr}(\{2,4,6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$.

Definition 2.17 (Mixed Strategy for NF Games)

Let $\langle \text{Agt}, \mathcal{A}ct, \text{ut} \rangle$ be normal form game.
The set of **mixed strategies for player** i is the set $S_i = \Pi(A_i)$.
The set of mixed strategy profiles is $S = S_1 \times \ldots \times S_n$.
This is also called the **strategy space** of the game.

For a set of actions $A_i = \{a_1, \ldots, a_k\}$ we represent a mixed strategy $s_i \in \Pi(A_i)$ as a tuple:

$$(\mathbf{a_1} : \mathbf{x_1}, \ldots, \mathbf{a_k} : \mathbf{x_k})$$

meaning that $s_i(a_i) = x_i$, e.g. $(a_1 : 0.5, a_2 : 0.25, a_3 : 0.25)$.

Note, that a pure strategy is a degenerated mixed strategy which assigns all probability to a single pure strategy/action.

Definition 2.18 (Support)

The **support** of a mixed strategy is the set of actions that are assigned non-zero probabilities.

What is the payoff of a mixed strategies? We have to take into account the probability with which an action is chosen.

Definition 2.19 (Expected Utility for player i)

The **expected utility** for player i of the mixed strategy profile $(s_1, \dots, s_n)$ is defined as

$$\mathbf{ut}_i^{expected}(s_1, \dots, s_n) = \sum_{\vec{a} \in \mathrm{Act}} \underbrace{\mathbf{ut}_i(\vec{a})}_{\text{utility of } \vec{a} \text{ for } i} \underbrace{\prod_{j=1}^{n} s_j(a_j)}_{\text{probability that } \vec{a} \text{ occurs}} .$$

In the following we also write $\mathbf{ut}_i(s)$ for $\mathbf{ut}_i^{expected}(s)$ if clear from context.

Mixed versions of maxmin strategies are defined as before.

Definition 2.20 (Maxmin strategy)

Given a game $\langle \mathrm{Agt}, \mathcal{A}ct, \mathbf{ut} \rangle$, the **(mixed) maxmin strategy** of player i is a **strategy** that maximises the guaranteed payoff of player i, no matter what the other players $-i$ do:

$$\arg\max_{s_i \in S_i} \min_{s_{-i} \in S_{-i}} \mathbf{ut}_i(s_i, s_{-i})$$

The **maxmin value** for player i is $\max_{s_i \in S_i} \min_{s_{-i} \in S_{-i}} \mathbf{ut}_i(s_i, s_{-i})$.

The **minmax strategy** for player i is

$$\arg\ \min_{s_i \in S_i} \max_{s_{-i} \in S_{-i}} \mathbf{ut}_{-i}(s_i, s_{-i})$$

and its **minmax value** is $\min_{s_i \in S_i} \max_{s_{-i} \in S_{-i}} \mathbf{ut}_{-i}(s_i, s_{-i})$.

Example 2.21 (Two-finger morra)

- Two players: odd player O, even player E
- Actions: each can show 1 finger, or 2 fingers
- Payoffs: Suppose f fingers are show in total.
  - If f is even, then E gets f EUR from O.
  - If f is odd, then O gets f EUR from E.

		O	
		one	two
E	one	(2,-2)	(-3,3)
	two	(-3,3)	(4,-4)

How should rational players play? (Fingers are shown **simultaneously**). The whole game is commonly known to all players.

Example 2.22 (Maxmin of Morra Game)

First, we compute the **maxmin value** of the Morra game from Example 2.21.

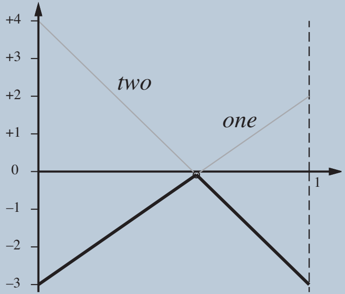
		O	
		one	two
E	one	(2,-2)	(-3,3)
	two	(-3,3)	(4,-4)

Suppose E plays $(one : x, two : 1 - x)$, against the action one and two of the opponent, respectively (the numbers refer to columns as O's strategy is fixed): $\rightsquigarrow$ Blackboard

- $\mathrm{ut}_E((one : x, two : 1 - x), \mathbf{one}) = 2x - 3(1 - x)$
- $\mathrm{ut}_E((one : x, two : 1 - x), \mathbf{two}) = -3x + 4(1 - x)$

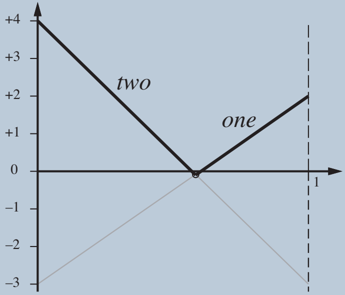
How can E maximize its worst-case payoff? Both expressions must give the same utility (otherwise the opponent would go for the worse action): $2x - 3(1 - x) = -3x + 4(1 - x)$ iff $x = \frac{7}{12}$.

Example continued



- against **one**: $\mathrm{ut}_E((one : x, two : 1-x), one) = 5x - 3$
- against **two**: $\mathrm{ut}_E((one : x, two : 1-x), two) = -7x + 4$
- $5x - 3 = -7x + 4$ iff $x = \frac{7}{12}$
- Player E's maxmin strategy is $(one : \frac{7}{12}, two : \frac{5}{12})$.
- Player E's maxmin value is: $-\frac{1}{12}$

Example continued



- Similarly, we obtain that player O's maxmin strategy is $(one : \frac{7}{12}, two : \frac{5}{12})$.
- But, player O's maxmin value is: $\frac{1}{12}$

The strategy profile $s = ((one : \frac{7}{12}, two : \frac{5}{12}), (one : \frac{7}{12}, two : \frac{5}{12}))$ is the **unique Nash equilibrium.** Note that $ut_E(s) = -\frac{1}{12}$ and $ut_O(s) = \frac{1}{12}$. **Thus, it is better to be player O.**

Example continued

An alternative way to compute the maxmin value of E is to directly compute the maxmin value:

$$\begin{aligned} s_E & \quad \operatorname{argmax}_{s_E} \min_{s_O} \quad ut_E((one : x, two : 1-x), (one : y, two : 1-y)) \\ = & \quad \operatorname{argmax}_{s_E} \min_{s_O} \quad \underbrace{2xy - 3(1-x)y - 3x(1-y) + 4(1-x)(1-y)}_{=f(x,y)} \end{aligned}$$

- Player O tries to minimize the expression. The slope in y is given by the first derivative of f to y:

$$\frac{df(x,y)}{dy} = 12x - 7$$

- E is best off, when f does not slope at all, i.e. $12x - 7 = 0$ which implies $x = \frac{7}{12}$.

Pictures are taken from [Russel and Norvig, 2010].

Example 2.23 (Nash equilibria in Battle of Sexes)

		Wife	
		Theatre	Restaurant
Husband	Theatre	(4,3)	(2,2)
	Restaurant	(1,1)	(3,4)

$\rightsquigarrow$ Blackboard

How to compute a **mixed NE**? What can be say about the support? Suppose $((T : \mathbf{x}, R : \mathbf{1} - \mathbf{x}), (T : \mathbf{y}, R : \mathbf{1} - \mathbf{y}))$ is a mixed Nash equilibrium of the Battle of Sexes game with $0 < x < 1$.

The underlying pure strategies must give the same expected utility against the opponent's mixed strategy: "Theater against $(T : \mathbf{y}, R : \mathbf{1} - \mathbf{y})$" = "Restaurant against $(T : \mathbf{y}, R : \mathbf{1} - \mathbf{y})$"

$$4y + 2(1 - y) \;=\; y + 3(1 - y) \quad \Rightarrow \mathbf{y} = \frac{1}{4}$$

Analogously for player 2: $3x + 1 - x = 2x + 4(1 - x) \quad \Rightarrow \mathbf{x} = \frac{3}{4}$

Example continued

Hence, a mixed Nash equilibrium is: $((T : \frac{3}{4}, R : \frac{1}{4}), (T : \frac{1}{4}, R : \frac{3}{4}))$.

All mixed Nash equilibria of the game are:

- $((T : \frac{3}{4}, R : \frac{1}{4}), (T : \frac{1}{4}, R : \frac{3}{4}))$
- $((T : 1, R : 0), (T : 1, R : 0))$
- $((T : 0, R : 1), (T : 0, R : 1))$

		Wife	
		Theatre	Restaurant
Husband	Theatre	(4,3)	(2,2)
	Restaurant	(1,1)	(3,4)

Note the difference to the computation of maxmin strategies!

Let $G = \langle \text{Agt}, \mathcal{A}ct, \mathbf{ut} \rangle$ be a finite normal form game. The **mixed extension** of G is the game

$$\Pi(G) = \langle \text{Agt}, \mathbf{S}, \mathbf{ut}_i^{expected} \rangle$$

where $S = S_1 \times \dots S_n$ and $S_i = \Pi(\mathcal{A}_i)$.
It is just a special normal form game. Which properties does it enjoy?

Theorem 2.24 (of Nash (1950))

Every finite normal form game ***has a Nash equilibrium*** *in mixed strategies.*

Nash's theorem can be proved by the fixed point theorems of Kakutani or Brouwer.

Exercise 2.1

Each of 100 people $A = \{1, \ldots, 100\}$ *announces a natural number from* 1 *to* 100*. A prize of 10 EUR is split equally between all the people whose number is closest to* $\frac{2}{3}$ *of the average number of the numbers chosen by all players.*
Which is the pure unique Nash equilibrium? Are there any other mixed Nash equilibria?

Subsection I

2 Simple Decisions: Normal Form Games

What can you say about the strategies of the column player B?

A\B	a_o	a_h
a_o	(4, 0)	(2, 2)
a_h	(1, 1)	(3, 4)

Definition 2.25 (Dominated action: Weakly, Strictly)

An **action a_i is strictly dominated** for an agent i, if there exists a mixed strategy $s'_i \in S_i$ that strictly dominates it, i.e. that

$$\mathbf{ut}_i(\langle s'_i, \vec{a}_{-i}\rangle) > \mathbf{ut}_i(\langle a_i, \vec{a}_{-i}\rangle)$$

for all action profiles $\vec{a}_{-i} = \langle a_1, \ldots a_{i-1}, a_{i+1} \ldots a_n\rangle \in A_{-i}$. We say that a action a_i is **weakly dominated** for an agent i, if in the above inequality we have $\geq$ instead of $>$ and the inequality is **strict for at least one** of the other $\vec{a}_{-i}$.

Lemma 2.26 (Iterated elimination of dominated strategies)

- *Given a 2-player normal form game. All* **strictly dominated** *actions can be* **eliminated** *without changing the Nash equilibria.*
- *This results in a* **finite series of reduced games**.
- *The final result does not depend on the order of the eliminations.*

Note: the last lemma is not true for **weakly** dominated strategies. There, the order **does** matter.

Note that we eliminate only **pure** strategies, i.e. actions. Such a strategy might be **dominated** by a **mixed** strategy.

Exercise 2.2

- *What are the pure Nash equilibria?*
- *What are the mixed Nash equilibria?*
- *What is the reduced game?*

$\rightsquigarrow$ *Blackboard*

	L	**C**	**R**
U	$\langle 3, 2\rangle$	$\langle 2, 1\rangle$	$\langle 3, 1\rangle$
M	$\langle 1, 1\rangle$	$\langle 1, 1\rangle$	$\langle 2, 2\rangle$
D	$\langle 0, 1\rangle$	$\langle 4, 2\rangle$	$\langle 0, 1\rangle$

1. *Eliminate row* **M**: *dominated by* $(U : 1, M : 0, D : 0)$
2. *Eliminate column* **R**: *dominated by* $(L : 0.5, M : 0, C : 0.5)$

This leads to

	L	**C**
U	$\langle 3, 2\rangle$	$\langle 2, 1\rangle$
D	$\langle 0, 1\rangle$	$\langle 4, 2\rangle$

- Pure Nash equilibria: (U, L) and (D, C)
- Mixed Nash equilibrium: $((U : \frac{1}{2}, M : 0, D : \frac{1}{2}), (L : \frac{2}{5}, C : \frac{3}{5}, R : 0))$

Compute the mixed Nash equilibria of the Battle of Sexes (Coordination) game. $\rightsquigarrow$ Exercise

Next Section

In normal form games each player selects a strategy at the beginning of the game once and for all.

Is this how decisions are made in more complex/real world situations?

Subsection I

Extensive Form Games

We have previously introduced normal form games (Definition 2.1 on Slide 22). This notion does not allow to deal with sequential decision making.

Extensive form (tree form) games

Unlike games in normal form, those in **extensive form** do not assume that all moves between players are made simultaneously. This leads to a **tree form**, and allows to introduce **strategies**, that take into account the **history** of the game.

We distinguish between **perfect** and **imperfect** information games. While the former assume that the players have **complete** knowledge about the current state of the game, the latter do not: a player might **not know** exactly which node it is in.

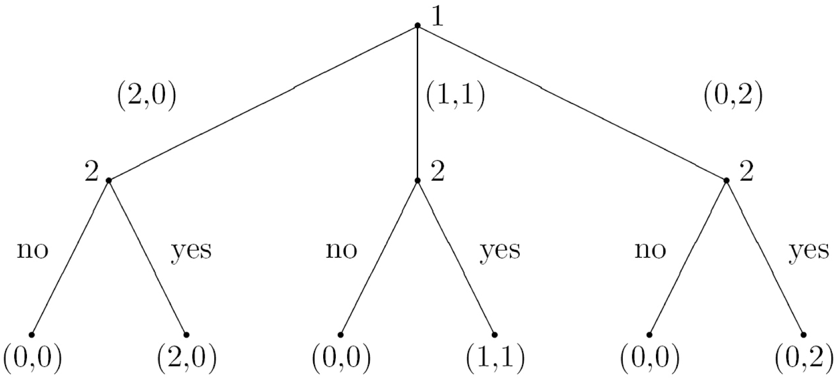


Figure 3.1: The Sharing game.

The following definition defines extensive form games as a tree. 1 and 2 are the players. Tuples and non leaf nodes, and "no" and "yes" indicate actions. Leaf nodes are labelled by utility values.

Definition 3.1 (Perfect Extensive Form Games)

A **finite perfect information game in extensive form** is a tuple $\Gamma = \langle \text{Agt}, \mathcal{A}ct, H, Z, \alpha, \rho, \sigma, \text{ut}_1, \dots, \text{ut}_n \rangle$ where

- Agt is a set of n players, $\mathcal{A}ct$ is a set of actions
- H is a set of non-terminal nodes, Z a set of terminal nodes, $H \cap Z = \emptyset$, $H \cup Z$ forms a tree (see below) with a **root node**,
- $\alpha : H \to 2^{\mathcal{A}ct}$ assigns to each node a set of actions,
- $\rho : H \to \text{Agt}$ assigns to each non-terminal node a player who chooses an action at that node,
- $\sigma : H \times A \to H \cup Z$ assigns to each (node,action) a successor node such that $\sigma(h_1, a_1) = \sigma(h_2, a_2)$ implies $h_1 \neq h_2$ and $a_1 \neq a_2$),
- $\text{ut}_i : Z \to \mathbb{R}$ is the utility functions of player i.

We often identify a node with the sequence of actions that leads to it from the root note. The sequence is unique and called **history**.

We use the notions node and history interchangeably.
Such games can be visualised as trees. Let us consider the "Sharing Game".

Example 3.2 (Sharing Game)

The game consists of two rounds. In the first, player 1 offers a certain share

1. 2 for player 1, 0 for player 2, $\rightsquigarrow$ action $(2,0)$
2. 1 for player 1, 1 for player 2, $\rightsquigarrow$ action $(1,1)$
3. 0 for player 1, 2 for player 2. $\rightsquigarrow$ action $(0,2)$

Player 2 can accept, or refuse. In the latter case, nobody gets anything.

Example continued

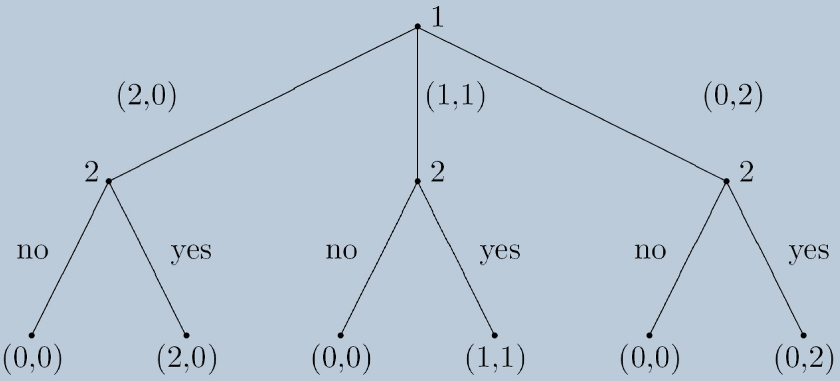


Figure 3.2: The Sharing game.

Which offer should player 1 make? A **strategy** defines what to do in each situation. How many different strategies are there for player 2?

Example continued

Formally, we can define it as the extensive form game

$$\Gamma = \langle \mathrm{Agt}, \mathcal{A}\mathrm{ct}, H, Z, \alpha, \rho, \sigma, \mathrm{ut}_1, \ldots, \mathrm{ut}_n \rangle$$

with

- $\mathrm{Agt} = \{1, 2\}$
- $\mathcal{A}\mathrm{ct} = \{(2,0), (1,1), (0,2), no, yes\}$
- $H = \{\epsilon\} \cup \{(2,0), (1,1), (0,2)\}$ (we use ϵ to denote the root node/ empty history)
- $Z = \{(u, v) \mid u \in \{(2,0), (1,1), (0,2)\}, v \in \{y, n\}\}$
- $\alpha(\epsilon) = \{(2,0), (1,1), (0,2)\}$, $\alpha(x) = \{no, yes\}$ for $x \in \{(2,0), (1,1), (0,2)\}$
- $\rho(\epsilon) = 1$ and $\rho(x) = 2$ for $x \neq \epsilon$.
- $\sigma(\epsilon, x) = x$, and for $x \neq \epsilon$ we define $\sigma(x, no) = (x, n)$, $\sigma(x, yes) = (x, y)$

Example continued

It remains to define the utility functions of the two players:

- $ut_1((x, n)) = ut_2((x, n)) = 0$ for $x \in Z$
- $ut_1(((2, 0), y)) = 2$
- $ut_1(((1, 1), y)) = 1$
- $ut_1(((0, 2), y)) = 0$
- $ut_2(((2, 0), y)) = 0$
- $ut_2(((1, 1), y)) = 1$
- $ut_2(((0, 2), y)) = 2$

Of course, there are many other ways to formalize the game as an extensive form game. For example, we could simplify the notation if we would also use *no* and *yes* in the definition of histories rather than *n* and *y*, respectively.

Strategies in extensive form games

A strategy for a player in an extensive form game defines an action for every node at which it is the player's turn.

Definition 3.3 (Strategies in Extensive Form Games)

Let $\Gamma = \langle \mathrm{Agt}, \mathcal{A}ct, H, Z, \alpha, \rho, \sigma, \mathrm{ut}_1, \ldots, \mathrm{ut}_n \rangle$ be a finite perfect information game in extensive form. A **(pure) strategy** for player i in Γ is **function**

$$\mathbf{s_i} : \{\mathbf{h} \in \mathbf{H} \mid \underbrace{\rho(h) = i}_{i\text{'s turn}}\} \to \bigcup_{\mathbf{h} \in \mathbf{H}} \alpha(\mathbf{h}) \quad \text{such that} \quad \mathbf{s_i}(\mathbf{h}) \in \alpha(\mathbf{h})$$

that assigns a legal action to each node owned by i. The set of all pure strategies of i is denoted by S_i. We define $S = \times_{i=1}^{n} S_i$.

We represent a pure strategy of player i as elements of $\Pi_{h \in H, \rho(h)=i} \alpha(h)$. It is represented as a **vector** $\langle a_1, \dots, a_r \rangle$ or simply $a_1 \dots a_r$ where $a_1, \dots, a_r$ are i's choices at the respective histories.

Some important nodes:

- When denoting a strategy for a player by $a_1 a_2 a_3 \dots$ we assume an implicit ordering of the histories at which it is the players turn:
    - we start from **top-to-bottom**
    - and **left-to-right**
- A strategy is a conditional plan. It assigns choices to each history, also those not reachable given a strategy profile.
- There are exponentially many strategies in the number of decision nodes for an agent.

Given a **pure strategy profile** $s = (s_1, \ldots, s_n)$, containing a strategy for each player, the **outcome of s in game Γ** is the terminal node which results from all players following their strategy in s.

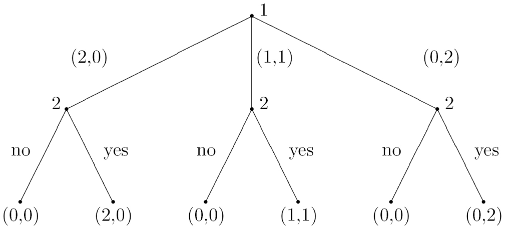
Definition 3.4 (Outcome)

The **outcome $O_\Gamma(S)$ of a pure strategy profile s in game Γ** is defined as follow:

$\mathbf{O_\Gamma(s) = h}$ where $h \in Z$ is the terminal node in Γ which is reached if all players follows their strategy in s from the root of Γ.

We omit Γ if clear from context.

Instead of $\mathrm{ut}_i(O(s))$ we also write $\mathrm{ut}_i(s)$.

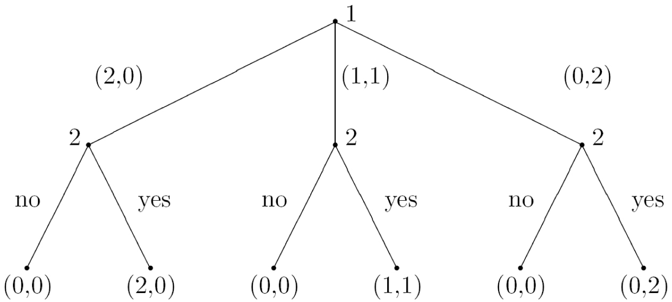


In the sharing game, a pure strategy for player 2 is $\langle \text{no}, \text{yes}, \text{no} \rangle$. A (better) one is $\langle \text{no}, \text{yes}, \text{yes} \rangle$.

Best response, Nash Equlibrium, et.c

Note that the definitions of **best response** and **Nash equilibrium** and other concepts carry over (literally) to games in extensive form.

Note that in the following we are talking only about **pure strategy profiles**.



What are the NE's in the sharing game?

- $\langle (2,0), \text{yyy} \rangle$ $\quad \langle (2,0), \text{nnn} \rangle$ $\quad \langle (2,0), \text{nny} \rangle$
- $\langle (2,0), \text{ynn} \rangle$ $\quad \langle (2,0), \text{yny} \rangle$ $\quad \langle (2,0), \text{yyn} \rangle$
- $\langle (1,1), \text{nyn} \rangle$ $\quad \langle (1,1), \text{nyy} \rangle$ $\quad \langle (0,2), \text{nny} \rangle$

The action profile $\langle (2,0), \text{nyy} \rangle$ is not a NE.

We claim that only $\langle (2,0), \text{yyy} \rangle$ and $\langle (1,1), \text{nyy} \rangle$ makes sense. Why? The others contain **non-credible threats**.

Transforming extensive form games

Lemma 3.5 (Extensive form $\hookrightarrow$ Normal form)

Each game Γ in **perfect information extensive form** *can be transformed to a game $NF(\Gamma)$ in* **normal form** *such that the pure strategy spaces correspond. It is called the* **reduced normal form of Γ.**

Proof.

- A strategy profile determines a unique path from the root $\emptyset$ of the game to one of the terminal nodes (and hence also a single profile of payoffs): $O(s)$.
- Therefore one can construct the corresponding normal form game $NF(\Gamma)$ by enumerating all strategy profiles and filling the payoff matrix with the resulting payoffs.

$\square$

Sharing Game in normal form

1\ 2	nnn	nny	nyn	nyy	ynn	yny	yyn	yyy
$(2,0)$	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(2, 0)	(2, 0)	(2, 0)	(2, 0)
$(1,1)$	(0, 0)	(0, 0)	(1, 1)	(1, 1)	(0, 0)	(0, 0)	(1, 1)	(1, 1)
$(0,2)$	(0, 0)	(0, 2)	(0, 0)	(0, 2)	(0, 0)	(0, 2)	(0, 0)	(0, 2)

Do Nash equilibria in the reduced normal form game always exist?

We cannot transform any normal form game to an equivalent extensive form as the following theorem shows:

Theorem 3.6 (Zermelo, 1913)

For each perfect information game in extensive form **there exists** *a pure strategy Nash equilibrium.*

The theorem follow from Theorem 3.9.

We do not need mixed strategies for perfect information extensive games to ensure the existence of a Nash equilibrium.

Example 3.7 (Unintended Nash equilibria)

Consider the following game in extensive form. The Nash equilibria are: (A, R) and (B, L).

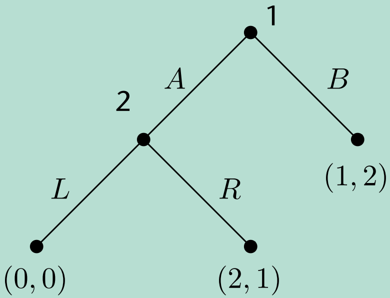


Figure 3.3: Unintended Equilibrium [Osborne and Rubinstein, 1994].

What do you think about the following Nash equilibrium (B, L)?

Would player two actually play L if its decision node were reached?

It is only a Nash equilibrium because of the incredible treat playing of L.

This leads to the notion of **subgame perfect Nash equilibria**.

Definition 3.8 (Subgame Perfect NE (SPE))

Let Γ be a perfect information game in extensive form.

Subgame: A **subgame of Γ rooted at node h** is the restriction of Γ to the descendants of h. The game is denoted Γ_h.

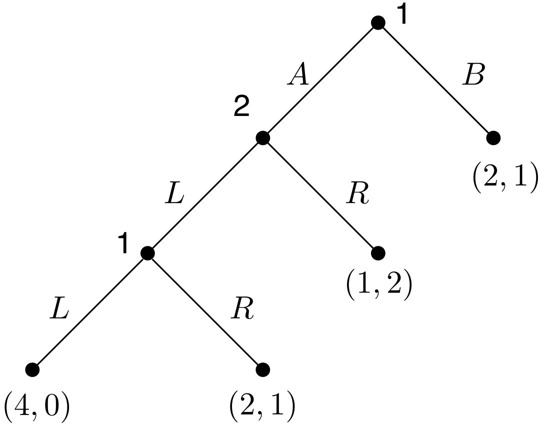
SPE: The **subgame perfect Nash equilibria (SPE)** of a perfect information game Γ in extensive form are those Nash equilibria of Γ, that are also Nash equilibria for all subgames Γ' of Γ.

What are the SPE's in the Sharing game (Example 3.2)?

Theorem 3.9 (Existence of SPE)

For each **finite perfect information** *game in extensive form* **there exists a SPE**.

The proof is by induction on the length of histories. It is constructive.



What is the SPE?

The (unique) SPE is:
$(\underbrace{BL}_{s_1}, \underbrace{R}_{s_2})$

Proof.

By **backwards induction** we construct a subgame perfect Nash equilibrium s. Let $l(\Gamma)$ denote the length of the longest history in Γ.

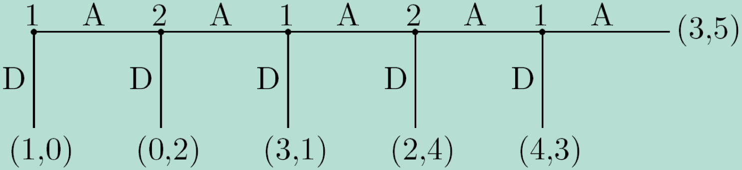
1. Let h be a terminal node and h' be the previous node. Moreover, let it be player i's move in h'. Then, we define $s_i(h')$ to be the action which maximizes i's utility. We proceed like this for all such histories h'.
2. Suppose now that s is a subgame perfect Nash equilibrium for all nodes h with $l(\Gamma_h) \leq k$. Consider a history h' of length $k+1$.
3. As before the player whose move it is in h' chooses an action which maximizes its payoff assuming that all other players follow s. We proceed like this for all histories h'.
4. The constructed strategy s is a subgame perfect Nash equilibrium.

□

This is similar to the multi-player Minimax algorithm.

Example 3.10 (Centipede Game)

This is a two person game which illustrates that even the notion of SPE can be critical.



Picture: [Osborne and Rubinstein, 1994]

Which strategy would you perform?

- There is a unique SPE: All players always choose D.
- This is rational, but humans often do not behave like that.
- Experiments show, that humans start with going across and do a down only towards the end of the game.

Let's return to the initial example:

Example 3.11 (Pirates and Gold)

- Five rational pirates a, b, c, d, e
- 100 gold coins
- the pirates have a rank: a is highest, e lowest
- Distribute coins such that
  - highest ranked pirate proposes a distribution
  - each pirates decide to accept or to reject the proposal, majority decides (highest ranked pirate breaks ties)
  - if rejected the proposer pirate is killed and the next highest ranked pirate makes a proposal
- Preferences of pirates (in this order and additive):
  1. stay alive
  2. maximize the number of gold coins
  3. kill other pirates

What should Pirate a propose? $\rightsquigarrow$ Exercise

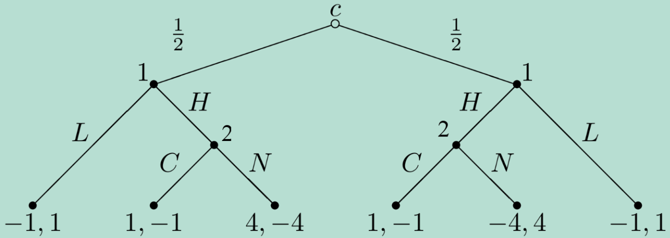
Subsection I

Chance Moves

Often, it is helpful to have player called **nature** or **chance**. Nature chooses an action randomly according to a given probability distribution over actions.

- We often denote the player "**chance**" by c.
- Given $h \in H$ with $\rho(h) = c$, there is a probability distribution $\beta(h)$ on the set of actions $\alpha(h)$.
- That is, action $a \in \alpha(h)$ is played with probability $\beta(h)(a)$.
- For $h \neq h'$, each two $\beta(h)$, $\beta(h')$ are independent distributions.
- The outcome is now an **expected outcome** $O(s)$ yielding and **expected utility** $\mathrm{ut}_i^{\textbf{expected}}(s)$
- $O(s)$ defines a probability distribution over terminal nodes.
- As before, we simply write $\mathrm{ut}_i(s)$ for $\mathrm{ut}_i^{\text{expected}}(s)$.

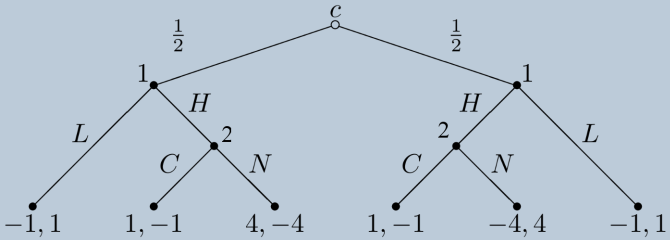
Example 3.12 (Extensive form game with chance node)



In the following table we list the **expected utility** of the associated normal form game.

		Player 2			
		CC	NN	CN	NC
Player 1	LH	$0,0$	$-\frac{5}{2},\frac{5}{2}$	$-\frac{5}{2},\frac{5}{2}$	$0,0$
	LL	$-1,1$	$-1,1$	$-1,1$	$-1,1$
	HH	$1,-1$	$0,0$	$-\frac{3}{2},\frac{3}{2}$	$\frac{5}{2},-\frac{5}{2}$
	HL	$0,0$	$\frac{3}{2},-\frac{3}{2}$	$0,0$	$\frac{3}{2},-\frac{3}{2}$

Example continued



What is the subgame perfect equilibrium of the game?

- Player 2 plays *CN*.
- Then, player 1 plays, *HL*.
- The **unique SPE** is (*HL*, *CN*).
- The payoff for both players is 0.
- Note that this reasoning was not affected by the chance node.
- Because the action selection of a player is independent.

Subsection I

3 Sequential Decisions: Extensive Form Games

- Perfect Information
- Extensive Form Games with Chance Moves
- Extensive Form Games with Imperfect Information
- Repeated Game
- Final Thoughts

Imperfect Information

- In an extensive game with perfect information, the player does know all previous moves (and also the payoffs that result).

- In an extensive game with **imperfect** information, a player might not be completely informed about the past history.
- An agent may not have enough memory to recall all past events.

The Idea:

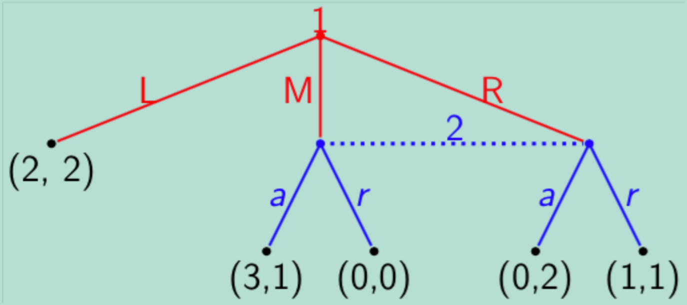
- An **extensive game** is nothing else than a **tree**. Thus each node is unique and carries with it the path from the root (the history that lead to it).

- In order to model that a player does not perfectly know the past events, we introduce an **equivalence relation** on the nodes: **equivalent** nodes cannot be **distinguished** by a player.

- All nodes in one equivalence class must be assigned the same actions: otherwise the player could distinguish them.

Recall: an equivalence relation is a relation which is reflexive, transitive, and symmetric.

Example 3.13 (Incomplete information extensive form game)



The dotted line defines player s's **information set**. It can not distinguish the two histories M and R.

A **partition of a set X** is a set of subsets $\{X_1, \ldots X_r\}$ where $X_i \subseteq X$:

1. $\bigcup_{j=1}^{r} X_j = X$ and
2. $X_i \cap X_j = \emptyset$ for $i \neq j$.

For example in abstract terms, $\{\{1,2\},\{3\}\}$ is a partition of the set $\{1,2,3\}$.
Note that a partition can be seen as the classes of an equivalence relation.

Definition 3.14 (Information partition and set)

Let $H_i \subseteq H$ be the set of histories after which it is the turn of player $i \in \mathrm{Agt}$, we define **an information partition of agent i of H_i** as a partition $I_i = \{I_{i1}, \ldots, I_{ir}\}$ of H_i. Its elements I_{ij} are called **information sets**.

Each $I_{ij} \in I_i$ encodes a set of histories that are indistinguishable for agent i, i.e. all $h \in I_{ij}$ look the same for i.

Definition 3.15 (Extensive Games, Imperfect Inf.)

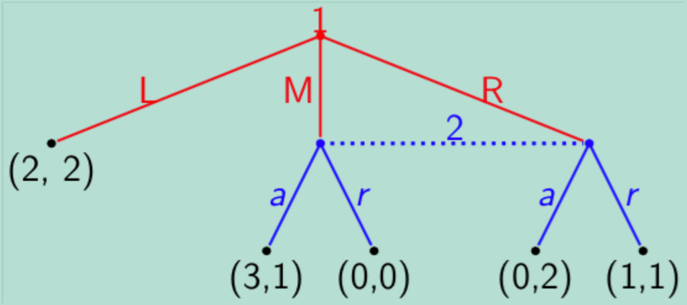
A **finite imperfect information game in extensive form** is a tuple $G = \langle \mathrm{Agt}, \mathcal{A}ct, H, Z, \alpha, \rho, \sigma, \mathrm{ut}_1, \dots, \mathrm{ut}_n, \mathbf{I_1}, \dots, \mathbf{I_n} \rangle$ where

- $\langle \mathrm{Agt}, \mathcal{A}ct, H, Z, \alpha, \rho, \sigma, \mathrm{ut}_1, \dots, \mathrm{ut}_n \rangle$ is a perfect information game in the sense of Definition 3.1 on Slide 69,
- $I_i = \{I_{i1}, \dots, I_{ik}\}$ are **information partitions of** $\{h \in H : \ \rho(h) = i\}$ such that

$$h, h' \in I_{ij} \in I_i \textbf{ implies } \alpha(h) = \alpha(h').$$

- Thus, we also write $\alpha(I_{ij})$ to refer to $\alpha(h)$ for any $h \in I_{ij}$.

Example 3.16 (Incomplete information extensive form game)



What are the pure strategies in this game?

Mappings from information sets to actions, that is, the same action must be assigned to indistinguishable histories.

How to define a pure strategy in an incomplete information extensive form game?

Definition 3.17 (Pure strategy)

Given an imperfect information game in extensive form, a **pure strategy for player *i*** is a function from information sets of the player to its actions:

$$\mathbf{s_i} : \mathbf{I_i} \to \bigcup_{\mathbf{h} \in \mathbf{H}} \alpha(\mathbf{h}) \quad \text{such that} \quad \mathbf{s_i}(\mathbf{h}) \in \alpha(\mathbf{h})$$

Again, we can represent this by a vector $\langle a_1, \ldots, a_k \rangle$ with $a_j \in \alpha(I_{ij})$ where $I_i = \{I_{i1}, \ldots I_{ik}\}$ is the information partition of player i.

Note that an extensive form game with perfect information is simply an imperfect information extensive form game in which each information set is a singleton.

Can we model prisoner's dilemma as an extensive game with imperfect information?

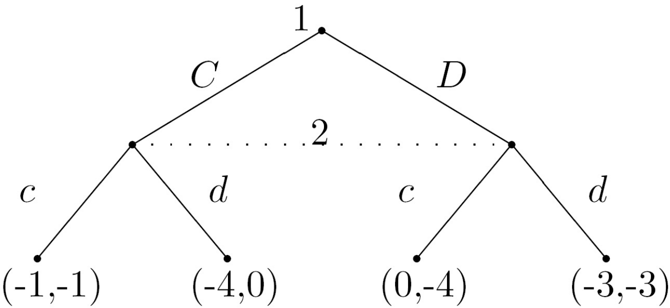
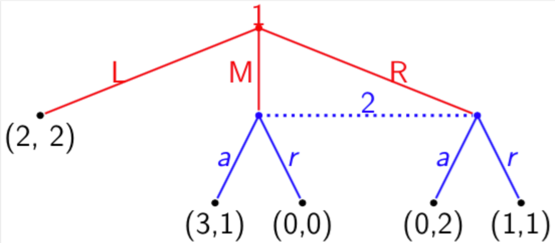


Figure 3.4: Prisoner's dilemma in extensive form [Osborne and Rubinstein, 1994].

There is a pure strategy Nash equilibrium.

But we could have chosen to switch player 1 with player 2.



Example 3.18

What are the pure Nash equilibria of the game shown in Figure 3.16?

As in the perfect information case, we can define the **reduced normal form game:**

		Player 2	
		a	r
Player 1	L	(2, 2)	(2, 2)
	M	(3, 1)	(0, 0)
	R	(0, 2)	(1, 1)

Nash equilibria are (L, r) and (M, a). But 2 does not know whether M or R has been played. Here a is better in both situations.

Mixed vs. Behavioral Strategies

Randomization over strategies is also defined for extensive form games. There are two prominent approaches to introduce probabilistic decision making:

- **Mixed strategies:** they define probability distributions over the players' pure strategies.
- **Behavioral strategies:** they define independent probability distributions over actions; that is, in each information set the agents randomize over their actions independently of previous choices.

It can be shown that over specific classes of extensive form games both approaches are outcome equivalent.

Definition 3.19 (Mixed strategies)

Let $G = \langle \mathrm{Agt}, \mathcal{A}ct, H, Z, \alpha, \rho, \sigma, \mathrm{ut}_1, \ldots, \mathrm{ut}_n, I_1, \ldots, I_n \rangle$ be an imperfect information game in extensive form. A **mixed** strategy σ_i of player i is a probability distribution **over i's pure strategies** S_i. The set of all mixed strategies is denoted Σ_i.

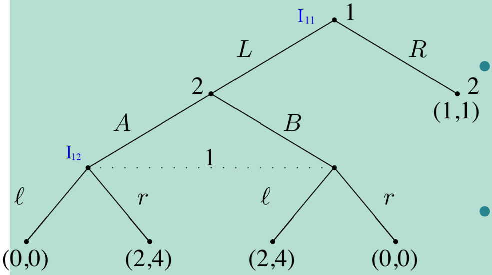
A pure strategy $s_i \in S_i$ is **consistent** with a history h if h can result if i plays according to s_i.
For $\sigma_i \in \Sigma_i$ we write $\sigma_i(h)$ as the **sum of all probabilities** $\sigma_i(s_i)$ where s_i is consistent with h. (Example⇝ Blackboard)

Definition 3.20 (Outcome of mixed strategy)

The **outcome** $O(\sigma)$ **of a mixed strategy profile** $\sigma = (\sigma_1, \ldots, \sigma_n)$ is a probability distribution over terminal nodes $h \in Z$ defined as follows:

$$O_\sigma(h) = \prod_{i=1}^{n} \sigma_i(h)$$

Example 3.21 (Mixed strategies)



- pure strategies of 1: Ll, Lr, Rl, Rr
- **mixed strategy of 1**:
  - $\sigma_1(Ll) = 0.5$
  - $\sigma_1(Lr) = 0.1$
  - $\sigma_1(Rl) = \sigma_1(Rr) = 0.2$
- **mixed strategy of 2**: $\sigma_2(A) = 0$, $\sigma_2(B) = 1$

Let $\sigma = (\sigma_1, \sigma_2)$. **Outcomes**:

- $O_\sigma(LAl) = 0 \cdot 0.5 = 0$
- $O_\sigma(R) = 0.2 + 0.2 = 0.4$
- $O_\sigma(LBl) = 1 \cdot 0.5 = 0.5$
- $O_\sigma(LBr) = 0.1.$

Definition 3.22 (Behavioral strategy)

Let $G = \langle \text{Agt}, Act, H, Z, \alpha, \rho, \sigma, \text{ut}_1, \dots, \text{ut}_n, I_1, \dots, I_n \rangle$ be an imperfect information game in extensive form. A **behavioral** strategy of player i β_i assigns to each information set $I_{i_j} \in I_i$ a probability distributions $\beta(I_{i_j})$ over the set of actions $\alpha(I_{i_j})$.

We define by $\beta(h)(a)$ the probability $\beta(I_{i_j})(a)$ for the action a for player i if $h \in I_{i_j}$. The set of all mixed strategies is denoted B_i.

Definition 3.23 (Outcome of behavioral strategy)

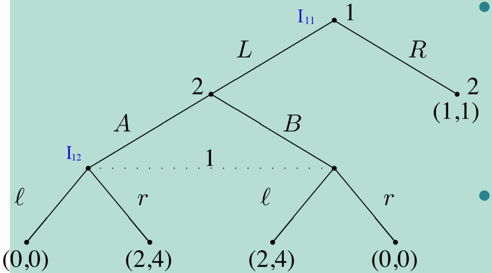
The **outcome $O(\beta)$ of a behavioral strategy profile** $\beta = (\beta_1, \dots, \beta_n)$ is a probability distribution over terminal nodes $h \in Z$ where $(n_1, \dots, n_k)$ the path leading to the node defined as follows:

$$O_\beta(h) = \prod_{i=1}^{k-1} \beta_{\rho(n_i)}(n_i)((n_i, n_{i+1}))$$

where (n_i, n_{i+1}) denotes the unique action leading from n_i to n_{i+1}.

A mixed and behavioural strategy are **outcome equivalent** if for all pure strategies of the opponents, they induce the same probabilities over terminal states.

Example 3.24 (Behavioral strategies)



- behavioral strategy of 1:
  $\beta_1(I_{11})(L) = 0.6$
  $\beta_1(I_{11})(R) = 0.4$
  $\beta_1(I_{12})(l) = \frac{5}{6}$
  $\beta_1(I_{12})(r) = \frac{1}{6}$
- behavioral strategy of 2:
  $\beta_2(L)(A) = 0,\ \beta_2(L)(B) = 1$

Let $\beta = (\beta_1, \beta_2)$. **Outcomes**:

- $O_\beta(LAl) = 0 \cdot 0.6 \cdot \frac{5}{6} = 0$
- $O_\beta(R) = 0.4$
- $O_\beta(LBr) = 0.6 \cdot 1 \cdot \frac{1}{6} = 0.1$
- $O_\beta(LBl) = 0.6 \cdot 1 \cdot \frac{5}{6} = 0.5$

Note that σ_1 from Example 3.21 gives the same outcome.

A mixed and behavioural strategy are **outcome equivalent** for a player if for all pure strategies of the opponents, they induce the same probabilities over terminal states.

Definition 3.25 (Expected utility)

We define the **expected utility** of a mixed or behavioural strategy f as

$$\mathrm{ut}_{\mathbf{i}}^{\mathrm{expected}}(\mathbf{f}) = \sum_{\mathbf{h} \in \mathbf{Z}} \mathrm{ut}_{\mathbf{i}}(\mathbf{h}) \cdot \mathbf{O_f}(\mathbf{h})$$

Again, we often simply write $\mathrm{ut}_i(f)$ for $\mathrm{ut}_i^{\mathrm{expected}}(f)$.

- The main difference is that for **behavioral** strategies, at each node the probability distribution is started freshly; they are independent. Even if a player ends up in the same partition, she can choose independently of her previous choice.
- Whereas for **mixed** strategies, this choice is **not** independent: there is **just one single** distribution that relates the possible choices.

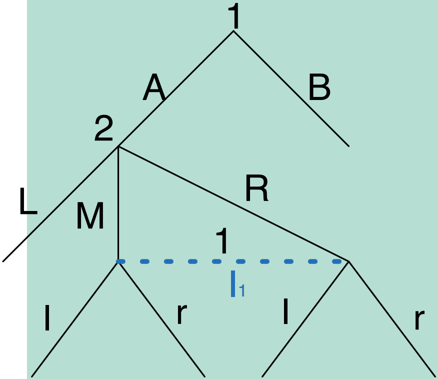
Remark 3.1

There are mixed strategies for which there are no behavioral strategies with the same outcome and vice versa. They also yields different notions of Nash equilibrium.

Exercise 3.1

Give an extensive form game such that there is a mixed-strategy which has no outcome equivalent behavioral strategy.

Example 3.26 (Mixed vs. Behavioral Strategy)



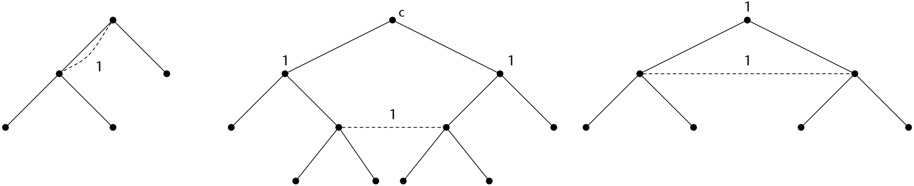
Mixed Strategy of player 1:

- $\langle B, r\rangle$ with probability 0.4
- $\langle B, l\rangle$ with probability 0.1
- $\langle A, l\rangle$ with probability 0.5

Outcome equivalent behavioural strategy:

- $\beta(\emptyset)(A) = \beta(\emptyset)(B) = \mathbf{0.5}$
- In the information set $I_{11} = \{(A, M), (A, R)\} \in I_1$: $\beta(I_1)(l) = \mathbf{1}$.

Intuitively speaking, a game of **perfect recall** is one, in which players do not forget past events. That is, in each history in the same information set, the player must have made the **same experiences**. Which games have perfect recall?



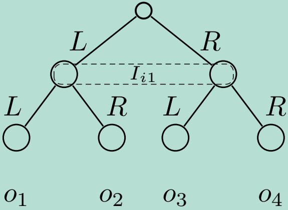
Picture: [Osborne and Rubinstein, 1994]
None of them: **(1)** agent forgets that an action was performed. **(2)** forgets about distinguishable subtrees **(3)** it forgets its own action.

Theorem 3.27 (Behavioral = Mixed (Kuhn, 1953))

Let Γ be a **game of perfect recall**. *Then for any mixed strategy of agent i there is an outcome equivalent behavioral one of agent i.*

Example 3.28 (Mixed, no behavioural strategy)

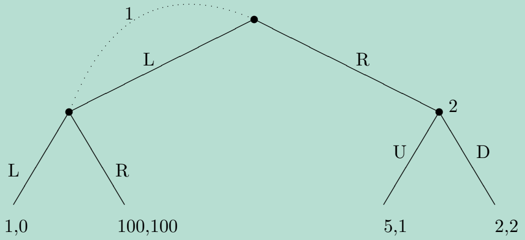
We consider a one player game. At the start node, the player can choose L or R resulting in two nodes which can not be distinguished by the player. Again, L or R can be played and result in the four outcomes o_1, o_2, o_3, o_4.



- What is the outcome of the mixed strategy $\langle \mathbf{LL} : \frac{1}{2}, \mathbf{RR} : \frac{1}{2} \rangle$?
- The outcome is $\frac{1}{2}o_1 + \frac{1}{2}o_4$
- No behavioral strategy results in this outcome. $\rightsquigarrow$ Exercise
- Therefore mixed strategies are not necessarily behavioral.

Example 3.29 (Game of imperfect recall: behavioural, no mixed)

We consider the following game



For **mixed** strategies, $\langle R, D \rangle$ is the unique NE. But in the context of **behavioral** strategies, the following behavioral strategy is a better response of player 1 to D: $(L : \frac{98}{198}, R : \frac{100}{198})$. Work out the details $\rightsquigarrow$ Exercise

Subsection I

So, far a game was played once. What if we **repeatedly** play the same game? How often should a game be repeated?

- **Finitely repeated** games: often, do not not behave intuitively, but can be treated with backwards induction.
- **Infinitely repeated** games: often, more intuitive, but require more technical details.

Example 3.30 (Finitely iterated prisoner dilemma)

After each round, the players know the action the other player has played. So, a strategy can take past plays into account.

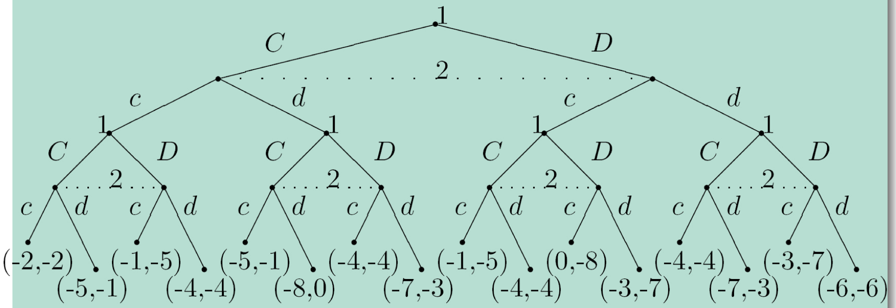


Figure 3.5: Iterated Prisoners Dilemma.

What is a SPE? Defecting at the last step makes sense.

Example continued

How are the utilities computed? It is the **(finite) sum** of each outcome of the stage game. Player i receives the

$$\textbf{accumulated reward}: \sum_{k=1}^{2} r_i^{(k)}$$

Lemma 3.31 (NE for iterated prisoners dilemma)

In the finitely repeated version of the prisoners dilemma, the **only subgame perfect Nash equilibrium** *is the strategy in which both players always defect.*

What if the game has an **infinite horizon**? We cannot assign a payoff to terminal nodes—there are none.

How to define the overall reward of a player?

Definition 3.32 (Average and discounted reward)

Let $r_i^{(1)}, r_i^{(2)}, \ldots$ denote an infinite sequence of rewards for player i. We define the **average reward** of player i as

$$\limsup_{k \to \infty} \frac{1}{k} \sum_{j=1}^{k} r_i^{(j)}.$$

We define the **future δ-discounted reward** of player i as

$$\sum_{j=1}^{\infty} \delta^{\mathbf{j}} r_i^{(j)}$$

where δ is a discount factor, $0 \leq \delta \leq 1$.

Two interpretation of the discounted reward:

- Players care more about the near future.
- Players care the same about present and future, but consider it possible that the repeated game ends after stage j with probability $1 - \delta$.

The **strategy space** is huge. **Finite automata** can be used to represent a subclass of these strategies. For example, a simple strategy where a player always defects can be represented with a single state only.

Finite Bargaining

Example 3.33 (Finite Bargaining with discount)

Finite bargining model with two agents want to split a dollar.

- An **offer** is a tuple $(x, 1-x)$ where $x \in [0, 1]$.
- Suppose $\boldsymbol{\delta} = 0.9$ is a discount factor (i.e. in every round the value v of the dollar is decreased to $v\delta$).
- Suppose 1 is the last player to offer (in round n, n even).
- Then the outcome depends on the number of rounds. In round $n+1$ they get nothing.

This is nothing else than an extensive form game. What is the unique subgame perfect Nash equilibrium in the depicted setting?

Example continued

Round	1's share	2's share	Total value	Offerer
⋮	⋮	⋮	⋮	⋮
$n-3$	0.819	0.181	0.9^{n-4}	*2*
$n-2$	0.91	0.09	0.9^{n-3}	*1*
$n-1$	0.9	0.1	0.9^{n-2}	*2*
n	1	0	0.9^{n-1}	*1*

- 1 can offer $(1,0)$ in round n $\rightsquigarrow$ utility $(\delta^{n-1},0)$
- 2 can offer $(\delta, 1-\delta)$ in round $n-1$ $\rightsquigarrow$ utility $(\delta\delta^{n-2},(1-\delta)\delta^{n-2})$
- 1 can offer $(\underbrace{1-(1-\delta)\delta}_{1-\delta+\delta^2},(1-\delta)\delta)$ in round $n-2$.
- …1 can offer $(x,1-x)$ in round 0 with
  $x = 1-\delta+\delta^2-\ldots+\delta^n = \sum_{i=0}^{n}(-1)^n\delta^n = \frac{\mathbf{1-\delta}}{\mathbf{1-\delta^2}}$
- The subgame perfect Nash equilibrium yields the outcome $(\frac{1-\delta}{1-\delta^2}, 1-\frac{1-\delta}{1-\delta^2})$. An agreement is reached in the first round.

Example continued

Derivation: We show that $\sum_{i=0}^{n}(-1)^n\delta^n = \frac{1-\delta}{1-\delta^2}$:

$$
\begin{aligned}
\sum_{i=0}^{n}(-1)^n\delta^i &= \sum_{i=0}^{n/2}\delta^{2i} - \sum_{i=1}^{n/2}\delta^{2i-1} \\
&= \sum_{i=0}^{n/2}(\delta^2)^i - \delta\sum_{i=0}^{n/2}(\delta^2)^{i-1} \\
&= \frac{1-(\delta^2)^{n/2+1}}{1-\delta^2} - \delta\frac{1-(\delta^2)^{n/2-1}}{1-\delta^2} \\
&= \frac{1-\delta}{1-\delta^2}
\end{aligned}
$$

Above we use the property of the geometric series (partial sum of the first n terms): $\sum_{i=0}^{n} q^i = \frac{1-q^{n+1}}{1-q}$ for $0 < q < 1$.

Infinite Bargaining

Bargaining is another example of a repeated game.

Example 3.34 (Infinite bargaining with discounts)

- Two players, 1 and 2, bargain about how to split goods worth initially $w_0 = 1$ EUR.
- After each round without agreement, the worth of the goods reduces by discount rates δ_1 (for player a_1) and δ_2 (for player a_2), $0 < \delta_1, \delta_2 < 1$.
- After t rounds, the goods are worth only $\langle \delta_1^t, \delta_2^t \rangle$. Subsequently, 1 (if t is even) or 2 (if t is odd) makes an offer to split the goods in proportions $\langle x, 1 - x \rangle$, and the other player accepts or rejects it.
- If the offer is accepted, then 1 gets $x\delta_1^t$, and 2 gets $(1 - x)\delta_2^t$; otherwise the game continues.

Theorem 3.35 (Unique solution for infinite bargaining [Rubinstein, 1982])

Consider the two player infinite bargaining setting with discount rates δ_1 and δ_2, $0 < \delta_1, \delta_2 < 1$. We define the following two strategies for 1 and 2, respectively:

s_1: *offers* $(\kappa_2, 1-\kappa_2)$ *and agrees to all* $(y, 1-y)$ *iff* $y \geq 1-\kappa_1$;

s_2: *offers* $(1-\kappa_1, \kappa_1)$ *and agrees to all* $(y, 1-y)$ *iff* $1-y \geq \kappa_2$;

where $\kappa_i = \frac{1-\boldsymbol{\delta}_i}{1-\boldsymbol{\delta}_1\boldsymbol{\delta}_2}$ *for* $i = 1, 2$. *Then,*

1. (s_1, s_2) *is the unique subgame perfect Nash equilibrium.*
2. **Agent *1* gets** κ_1.
3. *Agent 2 gets the rest,* $1-\kappa_1$.
4. *Agreement is reached in the first round.*

Proof.

First, we show that (s_1, s_2) is indeed a subgame perfect Nash equilibrium. (Hint: show that in no single step an agent can deviate from the strategy to improve its payoff (**one deviation property**)) $\rightsquigarrow$ Exercise The following gives a intuition of the proof of the unique SPNE:

Round	1's share	2's share	offerer
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$t-2$	$1-\delta_2(1-\delta_1 M_1)$		1
$t-1$		$1-\delta_1 M_1$	2
t	M_1		1
$\vdots$	$\vdots$	$\vdots$	$\vdots$

- Let M_1 and m_1 be the **maximal and minimal undiscounted share** that agent *1* can possibly get in any subgame perfect Nash equilibrium when she is offering, respectively.

Proof.

- If 2 offers x in $t-1$: 1 will reject if $x \le \delta_1 m_1$ and accept if $x \ge \delta_1 M_1$.
- Thus, if 1 offers y in $t-2$: 2 will reject if $y \le \delta_2(1-\delta_1 m_1)$ and accept if $y \ge \delta_2(1-\delta_1 M_1)$.
- The latter shows that 1 gets at most $1-\delta_2(1-\delta_1 m_1)$, thus

$$m_1 \ge 1-\delta_2(1-\delta_1 m_1) \quad \Rightarrow \quad m_1 \ge \frac{1-\delta_2}{1-\delta_1\delta_2}$$

- The latter shows that 1 can keep at least $1-\delta_2(1-\delta_1 M_1)$, thus

$$M_1 \le 1-\delta_2(1-\delta_1 M_1) \quad \Rightarrow \quad M_1 \le \frac{1-\delta_2}{1-\delta_1\delta_2}$$

- As $m_1 \le M_1$ by definition, it follows that $m_1 = M_1 = \frac{1-\delta_2}{1-\delta_1\delta_2}$.

The argument for player 2 is done analogously. The agreement is reached in the first round. $\square$

Subsection I

Some further models

Game theory offers tools to analyse the behavior of players. Other types of games are:

- Bayesian games: uncertainty about characteristics of other player
- Bounded rationality: e.g. due to bounded resources

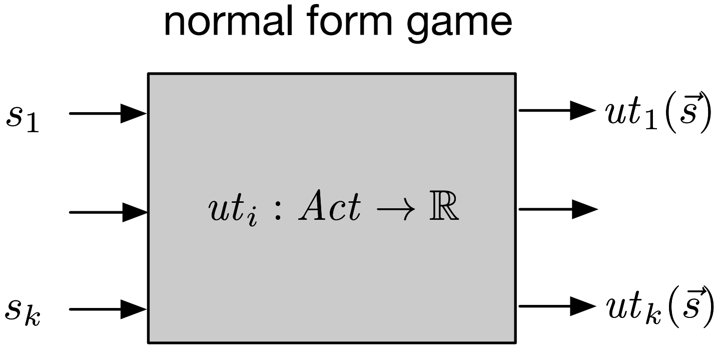
Game theory has many applications in multi-agent systems and economics:

- auction design
- prediction of markets
- protocol/interaction analysis
- security games
- negotiation
- cooperative game theory analyses the outcomes of cooperative teams and the stability of coalitions.
- etc.

The aim of this section was to give a you a brief introduction to game theory and decision making in multi-agent systems.

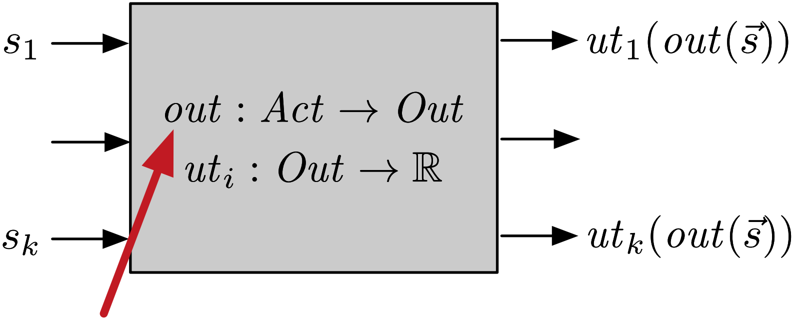
Next Section

Normal Form Game



Outcomes

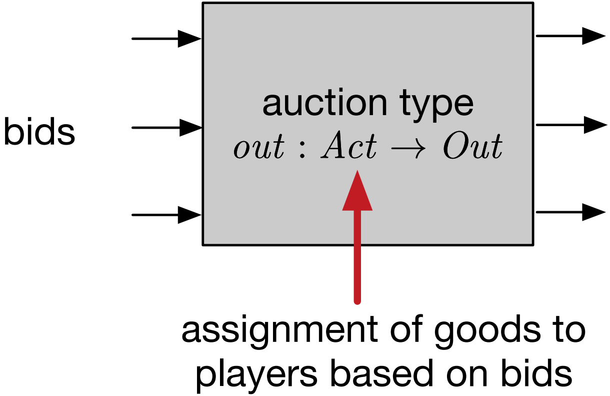
normal form game (alternative formulation)



defines the rules of the game

Mechanism design is about **designing the rules of the game**

Example of a Mechanism: Auctions



Players submit their bids and the auctioning rules decide on the allocation of the goods.

Auctions are special types of mechanisms:

- players propose a bid (action).
- the auction mechanism defines who wins the auction.
- a social choice function can be used to study properties of the auction, compare e.g. quasi-linear mechanisms and revenue maximization.

Auctions have been shown useful to allocate goods in multi-agent systems.

Types of auctions:

first-price open cry: (English, japanese auction), as usual.

first-price sealed bid: bidding without knowing the other bids.

Dutch auction: (**descending** auction) the seller lowers the price until it is taken (flower market).

second-price sealed bid: (**Vickrey** auction) Highest bidder wins, but the price is the second highest bid! (similar to eBay)

Three different auction settings:

private value (IPV): Value depends only on the bidder (cake).

common value (CV): Value depends only on other bidders (treasury bills).

correlated/interdependent value: Partly on own's value, partly on others.

Remark 4.1

Vickrey auctions *are, e.g., be used to*

- *allocate computation resources in operating systems,*
- *allocate bandwidth in computer networks,*
- *control building heating.*

Why can they be better suited than, e.g. English auctions or Dutch auctions? Less communication is needed.

Vickrey auction protocol

1. Auctioneer presents item.
2. Each bidder propose a bid, sealed in an envelope, for the item.
3. Outcome: bidder with highest bid wins.
4. Winner has to pay the amount of the second highest bid.

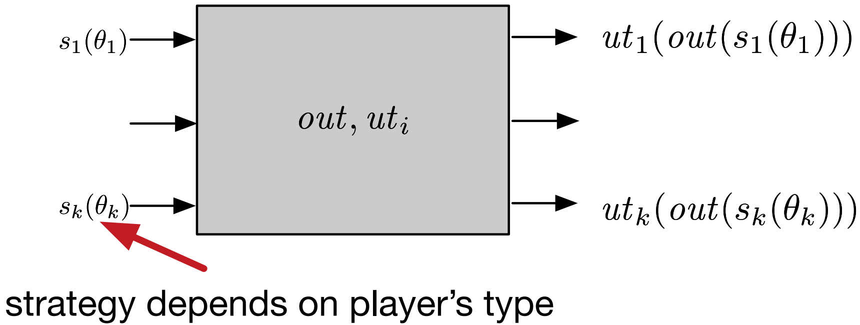
This is a **mechanism**: Given the bids of the bidders, the auction mechanism determines an outcome.

Is it a good mechanism? Which properties does it have? (will be made more precise later.)

- bidder who bids most gets the item
- maximizes revenue of the auctioneer (under some conditions)
- truthful bidding is dominant strategy (under some conditions)

How do players decide on their strategy?

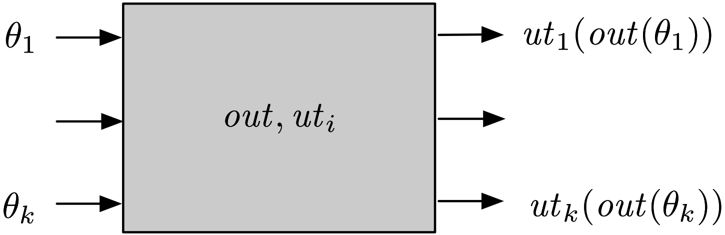
Type of a player θ_i captures the **private information** of the player.



Strategy proof

Auction: a type could model the **value** of a good to an agent.

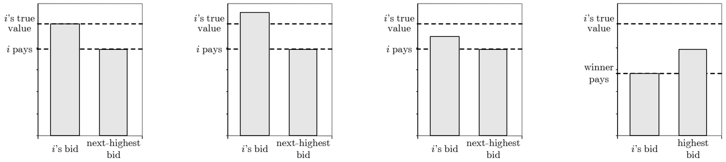
- Example: $s(\theta_{100EUR})$: I bid 80 EUR, and the good has a value of 100 EUR for the agent.
- **Direct mechanism**: the action set is identical to the agent's types:



- A direct mechanism is **truthful** or **strategy proof** if for each player announcing its true type is a dominant strategy.

Theorem 4.1 (Private-value Vickrey auctions)

The **dominant strategy** *of a bidder in a private-value Vickrey auction* **is to bid the true valuation**.

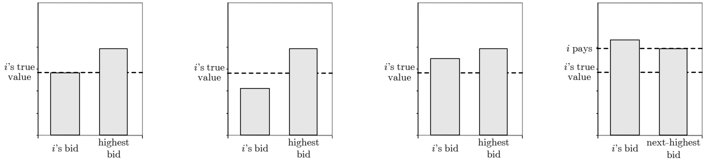


(a) Bidding honestly, i has the highest bid.

(b) i bids higher and still wins.

(c) i bids lower and still wins.

(d) i bids even lower and loses.



(e) Bidding honestly, i does not have the highest bid.

(f) i bids lower and still loses.

(g) i bids higher and still loses.

(h) i bids even higher and wins.

Picture: [Shoham and Leyton-Brown, 2009]

Remark 4.2 (Some final remarks)

- *Vickrey auctions* **maximize the expected revenue** *of the auctioneer.*
- *For* **non risk-neutral agents***, there can be a difference in the* **expected revenue** *for second price versus first price auctions.*
- *For* **interdependent items** *(e.g. two delivery tasks which provide synergies), the global optimal solution may not be reached by independently truthful bidding.*
- *Why are second price auctions not that popular among humans?*
  1. *Lying auctioneer*
  2. *When the results are published, subcontractors know the true valuations and what the winner saved. So they might want to share the profit.*

Vickrey-Clarke-Groves Mechanisms

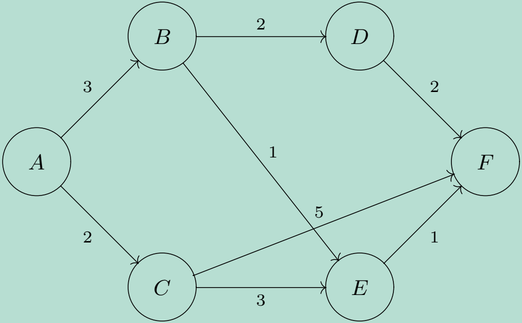
- Idee: the mechanism can **pay or charge** agents money
- a social desireable outcome can be achieved in such a way that the social benefit is distributed among the agents
- all **externalities** are made explicit and are internalised

Vickrey-Clarke-Groves mechanism works as follows:

1. mechanism asks each agent for its type θ_i^* (private information)
2. mechanism computes the **optimal outcome** c^* based on types
3. mechanism computes for each i:
   - $W_{-i} =$ sum of payoffs wrt. the best solution c' without i
   - $B_{-i} =$ sum of payoffs apart from i wrt. the best solution c^* (with i)
4. Agent i has to pay $\mathrm{pay}_i(\theta^*) = W_{-i} - B_{-i}$ as a tax.
5. Agent i gets utility $v_i(c^*) - \mathrm{pay}_i(\theta^*)$ where v_i denotes the value of outcome c^* for player i.

- So, if other agents are better off without i's contribution, i has to pay.
- An agent pays the **social costs** of its participation.
- Payments can be negative, meaning that agents get money.

Example 4.2 (VCG-Mechanism: transportation network)



Picture: [Shoham and Leyton-Brown, 2009]

Each transition between XY belongs to an agent XY.

Transitions are labelled with costs; thus, the optimal outcome will minimize costs. Define a mechanism to find the shortest path $ABEF$ in the model.

Example continued

- The mechanism computes the shortest path based on the announced types θ^*. The true types gives: $c = (A, B, E, F)$.
- Suppose all agents announce the true costs, then the payment $\text{pay}_i(\theta) = p_i$ is obtained as follows:

  $i \in \{AC, CE, CF, BD, DF\}$: pay 0 as they are not part of the shortest path.

  $i = AB$: is charged $\text{pay}_i(\theta) = (-6) - (-2) = -4$, i.e. is paid 4.

  $i = BE$: is charged $\text{pay}_i(\theta) = (-6) - (-4) = -2$, i.e. is paid 2.

  $i = EF$: is charged $\text{pay}_i(\theta) = (-7) - (-4) = -3$, i.e. is paid 3.
- Note that EF is paid more than BE, although they have the same costs, i.e. EF is more important than BE (has more market power).

Example continued

- So, given these payments the agents' utility computes as follows:

$$\begin{aligned}
\text{Utility of } AB &= -3+4=1\\
\text{Utility of } BE &= -1+2=1\\
\text{Utility of } EF &= -1+3=2\\
\text{Utility of others} &= 0+0=0
\end{aligned}$$

Example continued

What happens if, e.g., AB lies about its costs and announces 3.5 rather than 3?

- Then, $\text{pay}_{AB}(\theta^*) = (-6) - (-2) = -4$, i.e. is paid 4. Thus, it does not affect its paid money as its payoff is independent of its own announcement.

Now, suppose AB is even more greedy and announces 4.5.

- Then, the optimal solution would be $ACEF$ and hence, AB is paid 0. Its utility is 0 as it is not part of the shortest path.
- Clearly, a payment of 4 is better than 0. The agent is better off announcing the true costs.

In the example, what can we say about the payment made by the mechanism and the utility received (transportation costs)? It is $5 - 9 = -4$. The mechanism **pays more than it gets**.

Remark 4.3

Some properties of VCG mechanisms:

- **strategy proof**: *announcing the true type is a dominant strategy*
- **efficient**: *the socially optimal solution is computed by the mechanism*
- **not budget balanced**: *the mechanism may spend more money than it gets*
- **not individual rational**: *agents may be better off not taking part in the mechanism*

Next Section

Subsection I

5 A More Formal Introduction to Mechanism Design

Mechanism Design

Quasi-Linear Mechanisms

A utility function is defined as $\mathrm{ut}_i : \mathrm{Out} \to \mathbb{R}$. Two different functions encode different characteristics of a player:

- $\mathrm{ut}_i(car) = 12$
- $\mathrm{ut}'_i(car) = 40$

The latter function values a car more than the former. Rather than having different utility functions we use **types**.

Definition 5.1 (Type and utility functions)

A **type of player i** contains all private information about player i. It is denoted by θ_i. The set of all possible types of player i is denoted by Θ_i, and write $\Theta = \times_{i \in \mathrm{Agt}} \Theta_i$ for the set of type vectors.
A **(type-based) utility function** is defined as $\mathrm{ut}_i : \mathrm{Out} \times \Theta_i \to \mathbb{R}$.

Now we can model the example above by: $\mathrm{ut}_i(car, like-bikes) = 12$ and $\mathrm{ut}_i(car, like-cars) = 40$.
Here, a utility function is independent of the other agents' types. This can be generalized.

Game Form and Mechanisms

Definition 5.2 (Implementation Setting and Mechanism)

A finite **implementation setting** is a tuple $\langle \mathrm{Agt}, \mathcal{A}ct, \mathrm{Out}, \mathrm{out}, \Theta, \mathbf{ut}\rangle$, where

- $\mathrm{Agt} = \{1, \ldots, i, \ldots, n\}$ and $\mathcal{A}ct = A_1 \times \cdots \times A_i \times \cdots \times A_n$ are defined as before.
- Θ_i is the set of **types** of player i, $\Theta = \times_{i \in \mathrm{Agt}} \Theta_i$.
- Out is the set of **outcomes**.
- The **outcome function** $\mathrm{out} : \mathcal{A}ct \to O$ assigns to each action profile an outcome.
- $\mathbf{ut} = \langle \mathbf{ut}_1, \ldots, \mathbf{ut}_i, \ldots, \mathbf{ut}_n\rangle$ where $\mathbf{ut}_i : \mathrm{Out} \times \Theta_i \to \mathbb{R}$ is a (type-based) **utility function** for player i.

We call $M = \langle \mathrm{Agt}, \mathcal{A}ct, \mathrm{Out}, \mathrm{out}\rangle$ a **mechanism** or **game form**.

- For a fixed type vector $\theta \in \Theta$ the implementation setting $\langle \text{Agt}, Act, \text{Out}, \text{out}, \Theta, \textbf{ut} \rangle$ is simply a **normal form game** $\langle \text{Agt}, Act, \text{Out}, \text{out}, \textbf{ut}(\cdot, \theta) \rangle$
- Often, a Bayesian setting is considered. In that case a common-prior probability distribution p (commonly known to all players) on the type vectors Θ is included in the game.
- Several other extensions are possible.

Example 5.3 (Public School: Mechanism)

- Suppose the a public school with costs c should be built.
- **Outcomes**: *yes* (build the school) and *no* (don't build it)
- Each of the n players has to contribute $\frac{c}{n}$ EUR.
- **Actions**: $Act_i \subseteq \mathbb{R}^+$ announces how much the player values the school.
- **Outcome function** for $s_i \in Act_i$, $i \in \text{Agt}$:

$$\text{out}(s_1, \ldots, s_n) = \begin{cases} \text{yes} & (\sum_{i=1}^{n} s_i) - c \geq 0 \\ \text{no} & \text{otherwise} \end{cases}$$

  The school is built if the public benefit is greater equal the costs.
- The mechanism is given by $M = (\{1, \ldots, n\}, Act, \text{Out}, \Theta, \text{out})$

A social choice function f assigns to $\theta \in \Theta$ the **socially desirable outcome**.

Definition 5.4 (Social Choice Function)

Let Out be a set of outcomes. A **social choice function** is defined as $f : \Theta \to \text{Out}$.

Example 5.5 (Vickrey auction)

We return to our example of a Vickrey auction. Given a type vector θ the social choice function f specifies that:

$f(\theta)$: bidder who values the good most gets it (maximizing the revenue of the auctioneer) (for the price of the second highest bid).

How to link utility functions to the actual bids?

Example 5.6 (Public School: Social Choice Function)

- Players i values of the school $\theta_i \in \mathbb{R}^+$ (private information)
- Utility function (if project is built: benefit minus partial costs):

$$\mathrm{ut}_i(o, \theta_i) = \begin{cases} \theta_i - \frac{c}{n} & \text{o=yes} \\ 0 & \text{otherwise.} \end{cases}$$

The utility is the benefit the school has for the player, if it is built.

- Social choice function:

$$f(\theta_1, \dots, \theta_n) = \begin{cases} \text{yes} & \sum_{i=1}^{n} \mathrm{ut}_i(yes, \theta_i) = (\sum_{i=1}^{n} \theta_i) - c \geq 0 \\ \text{no} & \text{otherwise} \end{cases}$$

So, the school should be built if the social benefit is at least as high as the costs.

Given a social choice function f, we would like to construct a mechanism such that the **outcome of the rational behaviors** agree with the social choice function.

Definition 5.7 (Implementation)

Let M be a mechanism and f be a social choice function. M **implements f in dominant strategies**, if for all type vectors $\theta \in \Theta$,

- there is a dominant action profile $a = (a_1, \ldots, a_n)$ in the normal form game $(M, \mathrm{ut}(\cdot, \theta))$ and
- for all such dominant strategy profiles $a = (a_1, \ldots, a_n)$ we have that $M(a) = f(\theta)$.

When we talk about equilibria in the following we refer to dominant strategy profiles.

Remark 5.1 (Implementation in NE)

Implementation can also be defined with respect to Nash equilibria and other solution concepts.

Example 5.8 (Public School: Implementation)

- We are given the game $M = (\{1,\ldots,n\}, Act, \mathrm{Out}, \mathrm{out}, \mathrm{ut}(\cdot,\theta))$ defined as before with

$$\mathrm{ut}_i(o,\theta_i) = \begin{cases} \theta_i - \frac{c}{n} & \text{o=yes} \\ 0 & \text{otherwise.} \end{cases}$$

  and the actions by $Act_i = \mathbb{R}^+$, specifying the announced value of i for the public school.
- Social choice function:
  $f(\theta) = \text{yes iff } \sum_{i=1}^n \mathrm{ut}_i(yes,\theta_i) = \sum_{i=1}^n \theta_i - c \geq 0$
- What is a dominant strategy of i? $s_i^* = 0$ is dominant if $\theta_i < \frac{c}{n}$.
- Does the mechanism implement f in dominant strategies?
- No! Agents with $\theta_i < \frac{c}{n}$ have an incentive to under-report $s_i = 0$, and players with $\theta_i > \frac{c}{n}$ to over-report $s_i = c$. **They lie**!

We want to study truthfulness more formally.

Definition 5.9 (Direct mechanism)

A mechanism is called **direct** if the players' action sets correspond to a set of types, $\mathcal{A}ct_i = \Theta_i$.

Example 5.10 (Public School: Direct Mechanism)

In the previous example, we simply take:

- $Act_i = \Theta_i$ (the players announce their possible types)
- Outcome function:
$$\text{out}(\theta) = \begin{cases} \text{yes} & \sum_{i=1}^{n} \mathsf{ut}_\mathbf{i}(yes, \theta_i) = (\sum_{i=1}^{n} \theta_i) - c \geq 0 \\ \text{no} & \text{otherwise} \end{cases}$$
- Types are private information: players can lie.

Challenge: Define mechanisms in which players cannot benefit from lying.

We consider the setting of Example 5.5.

- M: (direct) auction mechanism
- f: bidder who values item most should get it

How to ensure that each bidder declares its **true** type?

Definition 5.11 (Truthful)

In a normal form game $(M, \mathrm{ut}(\cdot, \theta))$ with a direct mechanism M, θ_i is interpreted as the **true type of** i. We denote the **announced type** of a player i by $\theta_{\mathbf{i}}^* \in \Theta_i = Act_i$

A direct mechanism is **truthful** if for each agent i a dominant strategy θ_i^* in $(M, \mathrm{ut}(\cdot, \theta))$ is to announce its true type $\theta_i^* = \theta_i$.

Example 5.12 (Public School: Non Truthful)

We have seen that the direct mechanism from Example 5.10 is neither truthful nor does it implement f in dominant strategies.

Theorem 5.13 (Revelation principle)

If there is a mechanism M which implements f in dominant strategies then there is a direct and truthful mechanism which does so too.

Subsection I

The definition of a Mechanism is quite general. We define a special class of mechanisms.

Definition 5.14 (Quasi-linear Mechanism)

A **quasi-linear mechanism** is a mechanism $\langle \mathrm{Agt}, \mathcal{A}\mathrm{ct}, \mathrm{Out}, \Theta, \mathrm{out} \rangle$ with:

- $\mathrm{Out} = \mathrm{Ch} \times \mathbb{R}^n$ (outcome of choice and monetary component) where Ch is a set of **choices**, and n the number of players.
- $\mathrm{out}(a) = (\mathrm{ch}(a), \mathrm{pay}(a))$ where
  - $\mathrm{ch} : Act \to \mathrm{Ch}$ is the **choice function**
    Intuition: defines the choice resulting from the action tuple
  - $\mathrm{pay} : Act \to \mathbb{R}^n$ is the **payment function**.
    We write $\mathrm{pay}_i(a) = p_i$ to refer to p_i in $\mathrm{pay}(a) = (p_1, \ldots, p_n)$.
    Intuition: amount agents are charged (or paid by) the mechanism.

Other notions like **direct** and **truthful** are defined analogously.

Definition 5.15 (Quasi-linear utility function)

A **quasi-linear utility function** is defined as:

$$\mathrm{ut}_i((c,p),\theta_i) = \mathrm{v}_i(c,\theta_i) - f_i(p_i)$$

where $p = (p_1, \ldots, p_n)$ and

- $\mathrm{v}_i : \mathrm{Ch} \times \Theta_i \to \mathbb{R}^+$ is player i's **valuation function** (of choices),
- $f_i : \mathbb{R} \to \mathbb{R}$ is a strictly monotonically increasing function, modelling the player's risk attitude.

In a quasi-linear utility function:

- The players' valuation depends on the selected choice only.
- The mechanism can charge or pay the player some amount.
- We often write $\mathrm{v}_i(c)$ for $\mathrm{v}_i(c,\theta)$ when the type is clear from context. (Note that the utility function is based on the player's true type)

Remark 5.2 (Risk attitude)

Function f_i allows to model a player's **risk attitude**, *i.e. how it values a payoff p_i.*

A millionaire mY value 100 EUR less than a homeless person.

We distinguish three types:

- *risk averse*
- *risk neutral* $(f(x) = x)$
- *risk seeking* (e.g. $f(x) = e^{0.1x}$)

We assume that f_i **is commonly known**. *Thus, it is sufficient to specify the valuation function v_i. Therefore, we also write (M, v) for (M, ut). In the following we also assume that all players are risk neutral, setting $f_i(x) = x$.*

Note that quasi-linear mechanisms and utility functions allow to define social choice functions based on payoffs. For example:

- A mechanism M is **revenue maximizing** if all equilibria a in $(M, \mathrm{ut}(\cdot, \theta))$ maximize

$$\sum_{i \in \mathrm{Agt}} \mathrm{pay}_i(a).$$

- A mechanism M is **efficient** if

  all equilibria $\mathbf{a}$ in $(M, \mathrm{ut}(\cdot, \theta))$ maximize $\sum_{i \in \mathrm{Agt}} \mathrm{v}_i(\mathrm{ch}(\mathbf{a}), \theta_i)$.

  (Note that efficiency is defined wrt. the agents' true type.)

- A mechanism M is **budget balanced** if all equilibria a in $(M, \mathrm{ut}(\cdot, \theta))$ it holds that $\sum_{i \in \mathrm{Agt}} \mathrm{pay}_i(a) = 0$.

We consider again direct mechanisms, i.e, the agents' actions are their sets of types. θ^* represents the announced type.

Definition 5.16 (Groves Mechanism)

A **Groves mechanism** is a direct, quasi-linear mechanism with the following outcome and payment function:

$$\begin{aligned}
\mathsf{ch}(\theta^*) &= \operatorname{argmax}_{c \in \mathsf{Ch}} \sum_{i \in \mathsf{Agt}} \mathsf{v}_i(c, \theta^*_i) \\
\mathsf{pay}_i(\theta^*) &= h_i(\theta^*{}_{-i}) - \sum_{i \neq j \in \mathsf{Agt}} \mathsf{v}_j(\mathsf{ch}(\theta^*), \theta^*_j)
\end{aligned}$$

for a quasi linear utility function ut with valuation function v where $h_i : \Theta_{-i} \to \mathbb{R}$ is an arbitrary function. Note that the functions are defined wrt. the announced type θ^*.

Some notes about the mechanism:

- externalities are internalized
- agent i has to pay $h_i(\theta^*{}_{-i})$ (payment depending on other players types only) and is paid $\sum_{i \neq j \in \mathrm{Agt}} \mathrm{v}_j(\mathrm{ch}(\theta^*), \theta^*_j)$ (other players utility given the **current choice**)
- the choice depends on the valuation of all agents
- utility of a player depends on valuation and payments

Note that an agents is better off when other agents are better off. So, it is interested in maximizing the other agents' utility as well.

Theorem 5.17

Any Groves mechanism is **truthful** *and* **efficient**.

In the setting of quasi-linear utility functions, Groves mechanisms are the only quasi-linear mechanisms that are efficient (Theorem of Green-Laffont).

Proof.

Agent i wants to maximize $\max_{\theta_i^*} \mathrm{ut}_i((\mathrm{ch}(\theta^*), \mathrm{pay}(\theta^*)), \theta_i)$. This is the same as maximizing

$$\max_{\theta_i^*} \left(\mathrm{v}_i(\mathrm{ch}(\theta^*), \theta_i) - \underbrace{h_i(\theta^*_{-i})}_{\text{irrelevant for maximum}} + \sum_{i \neq j} \mathrm{v}_j(\mathrm{ch}(\theta^*), \theta_j^*) \right)$$

which is equivalent to

$$\max_{c \in \mathrm{Ch}} \left(\mathbf{v_i(c, \theta_i)} + \sum_{i \neq j} \mathrm{v}_j(c, \theta_j^*) \right)$$

The mechanisms is built to maximize $\max_{c \in \mathrm{Ch}} \left(\mathbf{v_i(c, \theta_i^*)} + \sum_{i \neq j} \mathrm{v}_j(c, \theta_j^*) \right)$. So, the agent's dominant strategy is to announce its true type $\theta_i = \theta_i^*$ and let the mechanism to maximize the outcome accordingly. □

Definition 5.18 (Vickrey-Clarke-Groves (VCG) Mechanism)

A **Vickrey-Clarke-Groves (VCG) mechanism** is a Groves mechanism with

$$\begin{aligned} \mathrm{ch}(\theta^*) &= \mathrm{argmax}_{c\in \mathrm{Ch}} \sum_{i\in \mathrm{Agt}} \mathrm{v}_i(c,\theta_i^*) \\ \mathrm{pay}_i(\theta^*) &= h_i(\theta^*{}_{-i}) - \sum_{i\neq j\in \mathrm{Agt}} \mathrm{v}_j(\mathrm{ch}(\theta^*),\theta_j^*) \end{aligned}$$

where

$$h_i(\theta^*{}_{-i}) = \sum_{i\neq j\in \mathrm{Agt}} \mathrm{v}_j(\mathrm{ch}(\theta^*_{-i}),\theta_j^*)$$ is the **Clarke tax**

and $\mathrm{ch}(\theta^*_{-i})$ defines the choice that would have been made by the mechanism if i had not participated.

Remark 5.3

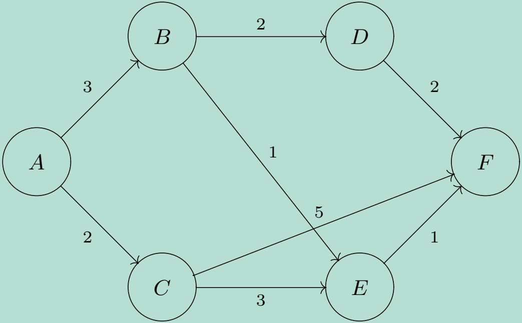
We are a bit sloppy regarding $h_i(\theta^*_{-i})$ and assume that the choice $\mathrm{ch}(\theta^*_{-i})$ *is contained in* Ch.

- Suppose players announce their true type:

$$\mathrm{pay}_i(\theta) = \underbrace{\sum_{i \neq j \in \mathrm{Agt}} \mathrm{v}_j(\mathrm{ch}(\theta_{-i}), \theta_j)}_{\text{charge of } i\text{: utility of other agents if agent } i \text{ would not have participated}} - \underbrace{\sum_{i \neq j \in \mathrm{Agt}} \mathrm{v}_j(\mathrm{ch}(\theta), \theta_j)}_{\text{paid to } i\text{: utility of all other agents wrt. current outcome}}$$

- If an agent does not change the outcome, it has to pay 0.
- Player costs: best outcome of the others (without the player) minus current outcome.
- A player pays the **societal costs** of its participation.

Example 5.19 (VCG-Mechanism: transportation network)



Picture: [Shoham and Leyton-Brown, 2009]

Each transition between XY belongs to an agent XY.

Transitions are labelled with costs; thus, ch will minimize costs.
Define a mechanism to find the shortest path $ABEF$ in the model.

How to ensure that agents announce their true costs?

Example continued

- The choice function $\text{ch}(\theta^*) = c$ returns the shortest path based on the announced type θ^*. The true types gives: $c = (A, B, E, F)$.
- Suppose all agents announce the true costs, then the payment $\text{pay}_i(\theta) = p_i$ is obtained as follows:

  $i \in \{AC, CE, CF, BD, DF\}$: pay 0 as $\text{ch}(\theta_{-i}) = \text{ch}(\theta)$.

  $i = AB$: is charged $\text{pay}_i(\theta) = (-6) - (-2) = -4$, i.e. is paid 4.

  $i = BE$: is charged $\text{pay}_i(\theta) = (-6) - (-4) = -2$, i.e. is paid 2.

  $i = EF$: is charged $\text{pay}_i(\theta) = (-7) - (-4) = -3$, i.e. is paid 3.
- Note that EF is paid more than BE, although they have the same costs, i.e. EF is more important than BE (has more market power).

Example continued

- So, given these payments the agents' utility computes as follows:

$$\begin{aligned}
\mathrm{ut}_{AB}((c,p),\theta_{AB}) &= -3+4=1\\
\mathrm{ut}_{BE}((c,p),\theta_{BE}) &= -1+2=1\\
\mathrm{ut}_{EF}((c,p),\theta_{EF}) &= -1+3=2\\
\mathrm{ut}_{x}((c,p),\theta_{x}) &= 0+0=0 \quad \text{for all other } x
\end{aligned}$$

Example continued

What happens if, e.g., AB lies about its costs and announces 3.5 rather than 3?

- Then, $\text{pay}_{AB}(\theta^*) = (-6) - (-2) = -4$, i.e. is paid 4. Thus, it does not affect its paid money as its payoff is independent of its own announcement.

Now, suppose AB is even more greedy and announces 4.5.

- Then, the optimal solution would be $ACEF$ and hence, AB is paid 0. Its utility is 0 as it is not part of the shortest path.
- Clearly, a payment of 4 is better than 0. The agent is better off announcing the true costs.

In the example, what can we say about the payment made by the mechanism and the utility received (transportation costs)? It is $5 - 9 = -4$. The mechanism **pays more than it gets**. That is, the mechanism is **not budget balanced**.

Exercise 5.1 (Public school setting)

Formulate the public school setting (strating from Example 5.3) in the quasi-linear setting and show that the VCG-mechanism is truthful and implements the social choice function in dominant strategies.

Some drawbacks of VCG mechanism:

- As seen in the previous example, it is not **budget balanced**.
- Moreover, it is not **individual rational**: players may be better off not taking part at all.
- With appropriate restrictions weaker versions hold: **ex post individual rationality** and **weak budget balance**.
- Other drawbacks (which also hold in other mechanisms) include:
  - Agents must fully disclose private information.
  - Lying can be beneficial for groups of agents.
  - VCG-mechanism is not frugal: paid costs may not be "optimal".
  - Removing agents could increase revenue.
  - Mechanism is computationally intractable. The computation of the maximum in ch(θ) can be intractable (**NP**-hard).

Conclusion

- Agents take their decisions autonomously.
- Mechanisms need to take this into consideration.
- Mechanisms, in particular auctions, can e.g. be used to coordinate behavior of agents.

References I

Osborne, M. and Rubinstein, A. (1994).
A Course in Game Theory.
MIT Press.

Rubinstein, A. (1982).
Perfect equilibrium in a bargaining model.
Econometrica, 50(1):97–109.

Russel, S. and Norvig, P. (2010).
Artificial Intelligence: a Modern Approach.
Prentice Hall, 3 edition.

Shoham, Y. and Leyton-Brown, K. (2009).
Multiagent Systems - Algorithmic, Game-Theoretic, and Logical Foundations.
Cambridge University Press.