TÉCNICO LISBOA

Autonomous Agents and Multiagent Systems

MSc in Computer Science and Engineering

First Exam (2022-2023)

Instructions

- You have 120 minutes to complete the exam.
- Make sure that your exam has a total of 13 pages. Also, check if there are no missing sheets, then write your full name and student number on this page (and your student number on all pages).
- The exam has 14 questions, with a maximum score of 20 points. The questions have different levels of difficulty. The point value of each question is provided next to the question number.
- *If you get stuck in a question, move on.* You should start with the more straightforward questions to secure those points before moving on to the more complex questions.
- *No interaction with the faculty is allowed during the exam.* If you are unclear about a question, clearly indicate the unclear part and answer the question to the best of your ability.
- Please provide your answer in the space below each question. If you make a mess, clearly indicate your answer.
- This exam is a closed-book assessment, whereby students are NOT allowed to bring books or other reference material into the examination room. You may bring only ONE A4 page of handwritten notes, in your OWN handwriting. Typed notes or a copy of someone else's notes are not allowed.
- You may use a calculator, but any other type of electronic or communication equipment is not allowed.
- **Good luck!**

1 Agent architectures

Question 1. (**1 pts.**)
What is the difference between pro-active and reactive behavior? Explain each behavior.

Write your answer here:

Solution 1.
While the reactive behavior only responds to what it can observe, the pro-active behavior is driven by a goal and plans to achieve that goal.
An agent with reactive behavior maintains an ongoing interaction with its environment and responds to changes that occur in it (in time for the response to be useful).
Pro-active behavior is the same as a goal-oriented behavior, whereby an agent is generating goals and attempting to achieve these goals.

Question 2. (**1 pts.**)
Why does the reactive agent architecture have a short-term view and is unable to learn? Explain.

Write your answer here:

Solution 2.
The short-term view means that the decisions are only based on local information (observations). The reactive agent does not plan the decision it must make in order to achieve a goal.
The reactive agent's rules do not evolve and thus is unable to learn.

2 Normal-form games

Question 3. (1 pts.)
Define a normal-form game and explain its essential components.

Write your answer here:

Solution 3.
A normal-form game is a mathematical representation of a strategic interaction between two or more agents. It consists of a set of agents, a set of actions available to each agent, and a payoff function that defines the payoffs or outcomes associated with each combination of actions.

Question 4. (1 pts.)
Given a payoff matrix for a 2-agent normal-form game, how does one determine if it has a strictly dominaned action? Give an example.

Write your answer here:

Solution 4.
A strictly dominant action occurs when an action yields the highest payoff regardless of the actions chosen by the other agent.

	Cooperate	Defect
Cooperate	3, 3	0, 5
Defect	5, 0	1, 1

For instance, the action *Defect* for agent 1 is a strictly dominant action because agent 1 always receives the highest payoff with this action regardless of the actions chosen by agent 2.

Question 5. (1 pts.)
Discuss the concept of Nash equilibrium in normal-form games and provide an example.

Write your answer here:

Solution 5.
Nash equilibrium is a solution concept in game theory where no agent has an incentive to unilaterally deviate from their chosen action, given the actions of the other agents. In other words, it is a joint action where no agent can improve their payoff by changing their action alone.
An example is the Prisoner's Dilemma, where (*Defect*, *Defect*) is a Nash equilibrium because no agent can improve their payoff by changing their action alone.

	Cooperate	Defect
Cooperate	3, 3	0, 5
Defect	5, 0	1, 1

Question 6. (**1 pts.**)

Define the concept of mixed strategy Nash equilibrium in normal-form games. How does it differ from pure strategy Nash equilibrium? Does every normal-form game have a Nash equilibrium?

Write your answer here:

Solution 6.

In a mixed strategy equilibrium, agents choose their actions probabilistically, rather than deterministically. It represents a situation where agents randomize their actions to achieve the best possible expected payoff.
This differs from pure strategy Nash equilibrium, where agents choose a specific action with certainty and the joint action is a Nash equilibrium (as defined in the previous answer).
Every normal-form game with a finite number of actions for each agent, has at least one mixed strategy Nash equilibrium.

3 Applications of a Nash equilibrium: Cournot Model

Question 7. (**1 pts.**)
Define the Cournot model and explain its assumptions.

Write your answer here:

Solution 7.
The Cournot model is an economic model of imperfect competition that analyzes the behavior of firms in an oligopolistic market. The model assumes that firms produce identical products and make output decisions simultaneously. The key assumptions of the Cournot model include: (a) firms choose their output levels independently, (b) firms have complete information about market demand and their competitors' costs, (c) firms aim to maximize their profits, and (d) there are a small number of firms in the market.

Question 8. (**2 pts.**)
Consider a market with two firms that produce homogeneous products. Firm A has a constant marginal cost of \$20 per unit, while firm B has a constant marginal cost of \$30 per unit. The market demand function is given by $P = 200 - 2Q$, where Q is the total quantity and P is the price. Calculate the Cournot equilibrium output for each firm and the price.

Write your answer here:

Solution 8.
To find the Cournot equilibrium output and price, we can use the following steps:

- Firm A's payoff is $\pi_A = Pq_A - c_A = (200 - 2(q_A + q_B))q_A - 20q_A = -2q_A^2 - 2q_Bq_A + 180q_A$
- Firm B's payoff is $\pi_B = Pq_B - c_B = (200 - 2(q_A + q_B))q_B - 30q_B = -2q_B^2 - 2q_Aq_B + 170q_B$

- Firm A's first order condition: $\frac{d}{dq_A}\pi_A = 0$

  $-4q_A - 2q_B + 180 = 0$

  $q_A = \frac{90-q_B}{2}$

- Firm B's first order condition: $\frac{d}{dq_B}\pi_B = 0$

  $-4q_B - 2q_A + 170 = 0$

  $q_B = \frac{85-q_A}{2}$

- We now need to solve the following pair of equations:

  $q_A = \frac{90-q_B}{2}$

  $q_B = \frac{85-q_A}{2}$

- Hence, the Nash equilibrium is:

  $q_A = \frac{95}{3}$

  $q_B = \frac{80}{3}$

- And the price is $P = \frac{250}{3}$

4 Extensive-form games

Question 9. (**1 pts.**)
Define an extensive-form game and explain its key components.

Write your answer here:

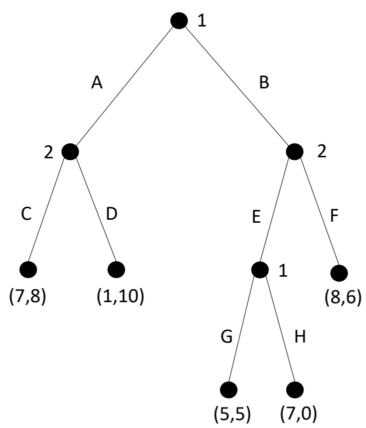
Solution 9.
An extensive-form game is a mathematical representation of a strategic interaction among multiple decision-makers over time. It consists of agents, actions available to each agent at each decision point, a game tree representing the sequence of actions and outcomes, and payoffs associated with the outcomes.

Question 10. (**1 pts.**)
Explain the concept of perfect information in extensive-form games. How does it differ from imperfect information?

Write your answer here:

Solution 10.
Perfect information in extensive-form games refers to a situation where all agents have complete knowledge about the actions taken by other agents at each decision point. Imperfect information, on the other hand, occurs when agents have limited or incomplete knowledge about the actions or strategies of other agents, often due to information asymmetry or private information.

Question 11. (**3 pts.**)
Consider the following extensive-form game:



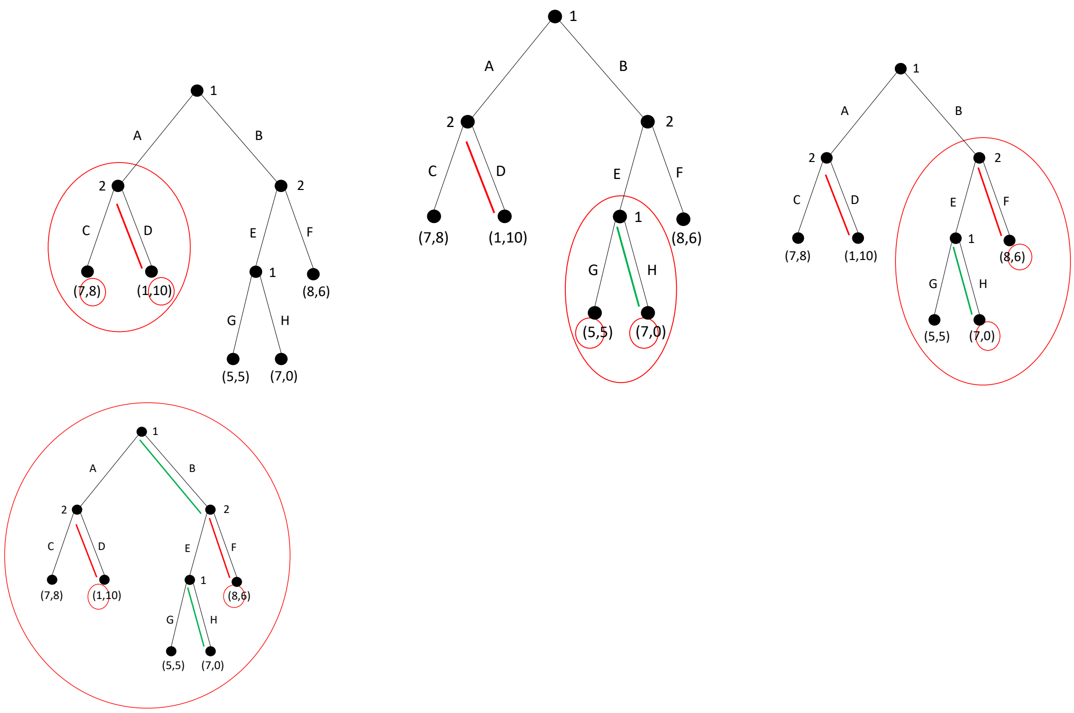
Describe the concept of backward induction and its application in solving extensive-form games. Solve the extensive-form game above using backward induction.

Write your answer here:

Solution 11.

Backward induction is a procedure for extensive-form games to compute the subgame-perfect equilibrium. The agents identify the equilibria in the "bottom-most" subgame trees, consider that these equilibria will be played, and back up and consider increasingly larger trees.

Applying backward induction:



Hence, the subgame-perfect equilibrium is {(B,H),(D,F)}

5 Bayesian games

Question 12. (**1 pts.**)
Define a Bayesian game and explain the key components involved.

Write your answer here:

Solution 12.
A Bayesian game is a strategic interaction among agents where each agent has incomplete information about the game being played. It extends the concept of a standard game to include an agent's private information, which affects their decision-making process. The key components of a Bayesian game include the type space (possible types each agent can have), action space (possible actions available to each agent), payoff functions (specifying the utility of each agent based on their type and the actions taken by all agents), and a prior distribution over the types.

Question 13. (**3 pts.**)
Consider the two-agent game below, whereby the payoffs depend on β:

for $\beta = 0$:

	X	Y
X	0, 4	4, 2
Y	2, 3	6, 2

for $\beta = 1$:

	X	Y
X	1, 3	2, 4
Y	3, 2	1, 3

where $P(\beta = 0) = p$ and $P(\beta = 1) = (1 - p)$, and $p \in [0, 1]$ is an unknown parameter. Only agent 2 knows whether $\beta = 0$ or $\beta = 1$.

First, formalize the game above as a Bayesian game. Second, find the interval of values of p such that agent 1 plays action Y under the Bayesian Nash equilibrium.

Write your answer here:

Solution 13.

The game can be formalized as a Bayesian game as follows:

- $N = \{\text{Agent}_1, \text{Agent}_2\}$ is the set of agents;
- $A_1 = A_2 = \{X, Y\}$ is the set of actions for each agent with $A = A_1 \times A_2$;
- $\theta_1 = \{t_1\}$ and $\theta_2 = \{0, 1\}$ are the set of types for each agent with $\theta = \theta_1 \times \theta_2$;
- $P(\theta_1 = t_1, \theta_2 = 0) = p$ and $P(\theta_1 = t_1, \theta_2 = 1) = (1 - p)$ is a prior over types;
- $u_1(a_1, a_2, \theta_1, \theta_2)$ and $u_2(a_1, a_2, \theta_1, \theta_2)$ are given by the payoff matrices above and $u = (u_1, u_2)$.

Considering agent 2:

For $\theta_2 = 0$, X strictly dominates Y. Hence, $a_2^*(\theta_2 = 0) = X$.

For $\theta_2 = 1$, Y strictly dominates X. Hence, $a_2^*(\theta_2 = 1) = Y$.

Considering agent 1:

The expected utility for action X is given by:

$$\begin{aligned}
EU_1(X|\theta_1) &= P(\theta_2 = 0|\theta_1 = t_1) \cdot u_1(X, a_2^*(\theta_2 = 0), t_1, 0) + P(\theta_2 = 1|\theta_1 = t_1) \cdot u_1(X, a_2^*(\theta_2 = 1), t_1, 1) \\
&= P(\theta_2 = 0|\theta_1 = t_1) \cdot u_1(X, X, t_1, 0) + P(\theta_2 = 1|\theta_1 = t_1) \cdot u_1(X, Y, t_1, 1) \\
&= p \cdot 0 + (1 - p) \cdot 2
\end{aligned}$$

The expected utility for action Y is given by:

$$\begin{aligned}
EU_1(Y|\theta_1) &= P(\theta_2 = 0|\theta_1 = t_1) \cdot u_1(Y, a_2^*(\theta_2 = 0), t_1, 0) + P(\theta_2 = 1|\theta_1 = t_1) \cdot u_1(Y, a_2^*(\theta_2 = 1), t_1, 1) \\
&= P(\theta_2 = 0|\theta_1 = t_1) \cdot u_1(Y, X, t_1, 0) + P(\theta_2 = 1|\theta_1 = t_1) \cdot u_1(Y, Y, t_1, 1) \\
&= p \cdot 2 + (1 - p) \cdot 1
\end{aligned}$$

For agent 1 to select action Y under the Bayesian Nash equilibrium we need to have that $EU_1(Y|\theta_1) > EU_1(X|\theta_1)$:

$$\begin{aligned}
EU_1(Y|\theta_1) &> EU_1(X|\theta_1) \\
2p + 1 \cdot (1 - p) &> 2 \cdot (1 - p) \\
p &> 1/3
\end{aligned}$$

Thus, agent 1 selects $a_1^*(t_1) = Y$ in the Bayesian Nash equilibrium $a^* = (a_1^*, a_2^*)$ whenever $p > 1/3$.

6 Repeated games

The following stage game of the Prisoner's Dilemma is played repeatedly:

	C	D
C	10, 10	0, 15
D	15, 0	4, 4

Question 14. (**2 pts.**)
Could we sustain cooperation if the stage game above is played with an infinite number of steps? Compute the solution and give an interpretation of the final result.

Write your answer here:

Solution 14.
We can use future discounted rewards in order to obtain the solution.

If agents 1 and 2 always cooperate:
$10 + \beta 10 + \beta^2 10 + \beta^3 10 + ... = \frac{10}{1-\beta}$

If agent 1 defects and agent 2 cooperates in the first time step, then agent 2 will defect in all the other time steps (trigger strategy):
$15 + \beta 4 + \beta^2 4 + \beta^3 4 + ... = 15 + \beta \frac{4}{1-\beta}$

The difference between these two strategies is:
$-5 + \beta 6 + \beta^2 6 + \beta^3 6 + ... = -5 + \beta \frac{6}{1-\beta}$

To sustain cooperation:
$-5 + \beta \frac{6}{1-\beta} \geq 0$
$\beta \frac{6}{1-\beta} \geq 5$
$6\beta \geq 5 - 5\beta$
$\beta \geq \frac{5}{11}$

Interpretation: if we want to sustain cooperation, the agent needs to care about tomorrow at least $\frac{5}{11}$ more than he cares about today!