Name: ..... Student card number: .....

Hand in this sheet only.

Rules

- ID required.
- You are not allowed to leave the exam room during the first 30 minutes.
- Scratch paper is handed out. You cannot use your own. It is possible to request additional scratch paper from the invigilator.

  The use of markers is not permitted.
- If you want to go to the toilet, raise your finger to warn a security guard. He or she will give you permission to go and walk with you to the toilet. Toilet visits are not permitted during the first and last half hour of the exam. You may only visit the toilet once.

  It is forbidden to take a telephone or similar electronic devices to the toilet.
- After you have left the examination room, you are not allowed to stay in the corridors / hall immediately outside due to noise. You follow the instructions of the invigilator.

Instructions

- There are open questions and multiple-choice questions.
- Every multiple-choice question has exactly one correct answer. In some cases, other answers may be “almost correct ” or “partly correct”. In such cases the best answer applies.

  Answer in the appropriate boxes by placing a cross. If you make a mistake, scratch the cross and put a cross in another box.

  Each correctly answered multiple-choice item yields one point.
- Answers to open questions are entered in the boxes (open rectangles)

  First draft your answer. Then fill the box.

  Each correctly answered open item yields two points, unless indicated otherwise.
- Because there are different versions of the exam, the order of the multiple-choice questions does not always correspond with the order of the material as discussed in the lectures.
- It is possible to request a new answer sheet as well as additional scratch paper from the invigilator. Our stock of answer sheets is finite, first come first serve.

  Good luck!

Multiple-choice answers

	A	*B*	*C*	*D*
1.				
2.				
3.				
4.				

	A	*B*	*C*	*D*
5.				
6.				
7.				
8.				

	A	*B*	*C*	*D*
9.				
10.				
11.				
12.				

Open questions—first draft your answer, then fill the box

1. Row is a 10% noise (plays a random action 10% of the time at random moments) fictitious player in a normal form game with payoffs

	L	C	R
T	$0,2$	$1,0$	$4,0$
M	$1,1$	$0,1$	$5,2$
B	$3,1$	$2,3$	$0,1$

Determine the strategy of Row at round eight if the history of play is

$$h = TC, BC, TL, BL, TR, TR, TL.$$

Answer. Projected mixed strategy of Col after seven rounds: L: 3/7; C: 2/7; R: 2/7.

Expected payoffs

$$\begin{aligned}
\mathrm{E}[T] &= \frac{3}{7}0 + \frac{2}{7}1 + \frac{2}{7}4 = \frac{10}{7},\\
\mathrm{E}[M] &= \frac{3}{7}1 + \frac{2}{7}0 + \frac{2}{7}5 = \frac{13}{7},\\
\mathrm{E}[B] &= \frac{3}{7}3 + \frac{2}{7}2 + \frac{2}{7}0 = \frac{13}{7}.
\end{aligned}$$

If Row would be a pure fictitious player, it would thus playe either M or B, each with probability $1/2$.

However, Row isn't a pure fictitious player, instead it randomizes 10% at the time. If F represents the event that Row does not randomize and plays according to fictitious play, then

$$\begin{aligned}
\Pr\{T|h\} &= \Pr\{T|F,h\}\Pr\{F|h\} + \Pr\{T|\overline{F},h\}\Pr\{\overline{F}|h\} = 0\frac{9}{10} + \frac{1}{3}\frac{1}{10} = \frac{1}{30}\ (= \frac{2}{60}).\\
\Pr\{M|h\} &= \Pr\{M|F,h\}\Pr\{F|h\} + \Pr\{M|\overline{F},h\}\Pr\{\overline{F}|h\} = \frac{1}{2}\frac{9}{10} + \frac{1}{3}\frac{1}{10} = \frac{29}{60}.
\end{aligned}$$

Similarly,

$$\Pr\{B|h\} = \cdots = \frac{29}{60}.$$

(It may be verified that the probabilities add up to 1.)

2. Two satisficing agents play the prisoner's dilemma with begin state C, initial aspiration 5, and persistence rate $1/2$. Determine the action profiles of the first five rounds.

Answer. Given the standard payoffs in the prisoner's dilemma

	C	D
C	$3,3$	$0,5$
D	$5,0$	$1,1$

we have

round	actions	payoffs	aspirations
0			$(5,5)$
1	(C,C)	$(3,3)$	$(4,4)$
2	(D,D)	$(1,1)$	$(2.5,2.5)$
3	(C,C)	$(3,3)$	$(2.75,2.75)$

From then on the payoffs are larger than the aspirations so that players won't change their actions, from which in turn it follows that play has converged. Answer: (C,C), (D,D), (C,C), (C,C), (C,C).

Multiple-choice questions

1. Describe the MAL algorithm named Bully.

   (a) Punish the opponent as much as possible by playing opponent's max-min.

   (b) Punish the opponent as much as possible by playing opponent's min-max.

   √ Play one action throughout, in the hope that the opponent is a follower.

   (d) Play your part of a NE, in the hope that the opponent is a follower.

> **Explanation.** Bully is not there to punish the opponent, but to produce a highest payoff as possible in the hope that the opponent is a follower. In effect, Bully achieves this by determining which action yields the highest reward if the opponent is a follower. This requires determining the action that yields a highest reward when the opponent maximizes on that action.

2. For a subgame-perfect NE in a repeated game, the payoff profiles must be

   (a) Admissible and countable.

   (b) Admissible and enforceable.

   √ Feasible and enforceable.

   (d) Feasible and admissible.

> **Explanation.** Cf. slides.

3. (Reinforcement learning.) What does it mean for an action to be *dominant* in a normal form game?

   (a) For some history of play, the actions's expected payoff is higher than any other action.

   (b) For all histories of play, the action's expected payoff is higher than any other action.

   (c) For all histories of play, there is some margin, such that the action's expected payoff is at least this margin higher than any other action's expected payoff.

   √ For some margin, for all histories of play, it's expected payoff is at least this margin higher than any other action's expected payoff.

> **Explanation.** Reinforcement learning, slide "Unboundedness of propensities, and convergence", Theorem 1.

4. Regret matching uses

   (a) Regrets.

   (b) Negative parts of regrets.

   √ Positive parts of regrets.

   (d) Mean square error of regrets.

5. Suppose two fictitious players play matching pennies.

   (a) The empirical frequencies do not converge.

   (b) The empirical frequencies converge, but not necessarily to a NE.

   (c) If the empirical frequencies converge, they converge to a NE.

   √ The empirical frequencies converge to a NE.

Explanation. Slide "Convergent empirical distribution of strategies".

6. Which of the following statements is true?

   i) Every evolutionarily stable state is a Nash equilibrium.

   ii) Every evolutionarily stable state is a neutrally stable state.

   √ Both.

   (b) Only *i)*.

   (c) Only *ii)*.

   (d) None.

7. Of a certain game, the stationary point of the gradient dynamic lies outside the unit square. Which of the following statements is true?

   i) There is a mixed NE.

   ii) IGA converges.

   (a) Both.

   (b) Only *i)*.

   √ Only *ii)*.

   (d) None.

Explanation. A mixed NE is (the only) stationary point of the gradient dynamic. IGA always converges, except when the mixed NE lies inside the unit square and the dynamics is concentric.

8. Satisficing play.

   i) Initial aspirations do matter (for the outcome, that is).

   ii) Play converges to Pareto-optimal action profiles.

   (a) Both.

   √ Only *i)*.

   (c) Only *ii)*.

   (d) None.

Explanation. The truth of *i)* was shown through diagrams of Stimpson and Goodrich. In fact this was one of the main points of their paper. The falsity of *ii)* can be understoopd as follows. With small persistence rates, aspiration profiles may overshoot Pareto-optimal action profiles and may get stuck at sub-optimal Pareto profiles.

9. Eatherly cooperates in the first round, and mirrors the (projected) mixed strategy of the opponent in subsequent rounds. Give

$$\Pr\{\ h \mid \text{Agent 2 plays Eatherly}\% \ \}$$

   if the history of play is CC, DC, CD, CC, CD.

   (a) 0.009

   √ 0.083

   (c) 0.729

   (d) Another answer.

Explanation. Let $s_2 =$ Eatherly for Player 2.

$$\begin{aligned}\Pr\{h|s_2\} &= \Pr\{(C,C)|s_2\} \times \Pr\{(D,C)|s,(C,C)\} \\ &\quad \times \Pr\{(C,D)|s,(C,C),(D,C)\} \ \ldots \\ &= 1.0 \times 1.0 \times 0.5 \times 0.67 \times 0.25 \\ &= 0.083.\end{aligned}$$

10. Consider a game in normal form where players enjoy normalised payoffs and engage in naive hypothesis testing with soft-max responses. For every $\epsilon > 0$, if the responses are sufficiently rational, the test tolerances are sufficiently fine, and the amounts of data collected are comparable and sufficiently large, then the responses constitute [choose the strongest]:

    (a) a Nash equilibrium a.s.

    (b) a Nash equilibrium at least $1 - \epsilon$ of the time.

    (c) an ϵ-equilibrium a.s.

    $\surd$ an ϵ-equilibrium at least $1 - \epsilon$ of the time a.s.

11. Uncoupled learning

    $\surd$ can observe actions but not the payoffs of the opponent(s).

    (b) can observe actions but not the strategy of the opponent(s).

    (c) can observe payoffs, but not the actions of the opponent(s).

    (d) can observe neither actions nor payoffs of the opponent(s).

12. Rank elimination in MAL tournaments:

    (a) Sort by performance, eliminate the worst.

    $\surd$ Sort by performance, eliminate the worst, and subtract all payoffs earned against the eliminated performer from the revenues of all survivors.

    (c) If the difference between the cumulative returns of the two worst performers becomes too large, the worst performer is removed.

    (d) If the global ranking has not changed within a certain number of tournaments, the worst performer is removed.

Happy holidays!

Scratch paper.