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. Given

	L	R
T	$1, -1$	$0, 0$
B	$0, 0$	$0, -1$

Suppose IGA-WoLF is played on the strategy profile $(\alpha, \beta) = (1/4, 1/4)$ with learning rate $\eta = 0.1$ and $(l_{\min}, l_{\max}) = (1, 2)$. Determine the strategy profile after one iteration.

Answer. If

$$M = \begin{array}{c} \\ \text{T} \\ \text{B} \end{array} \begin{array}{c} \text{L} \qquad \text{R} \\ \begin{pmatrix} r_{11}, c_{11} & r_{12}, c_{12} \\ r_{21}, c_{21} & r_{22}, c_{22} \end{pmatrix} \end{array}$$

then

$$\begin{cases} u \ = (r_{11} - r_{12}) - (r_{21} - r_{22}) = \ \ 1 \\ u' = (c_{11} - c_{21}) - (c_{12} - c_{22}) = -2 \end{cases}$$

so

$$\begin{aligned} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} &= \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \eta \begin{pmatrix} l_l \\ l_2 \end{pmatrix} \left(\begin{bmatrix} 0 & u \\ u' & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \begin{bmatrix} r_{12} - r_{22} \\ c_{21} - c_{22} \end{bmatrix} \right) \\ &= \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + 0.1 \begin{pmatrix} l_l \\ l_2 \end{pmatrix} \left(\begin{bmatrix} 0 & 1 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) \\ &= \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + 0.1 \begin{bmatrix} l_l(\ \ 1\,\beta + 0) \\ l_2(-2\,\alpha + 1) \end{bmatrix}. \end{aligned}$$

The dynamics is stationary when the gradient is zero, so when

$$0.1 \begin{bmatrix} l_l(\ \ 1\,\beta + 0) \\ l_2(-2\,\alpha + 1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

that is when $(\alpha, \beta) = (1/2, 0)$. This point is in the unit square, so $(\alpha^*, \beta^*) = (1/2, 0)$ is a mixed Nash equilibrium.

The eigenvalues of

$$U = \begin{bmatrix} 0 & 1 \\ -2 & 0 \end{bmatrix}$$

are imaginary, so the dynamic is centric in such a way that the strategy profiles circle around (α^*, β^*), which is the only Nash equilibrium. To determine whether the dynamics is clockwise or anti-clockwise, pick an arbitrary strategy profile $(\alpha, \beta) \neq (1/2, 0)$, let us say $(\alpha, \beta) = (0, 0)$. This profile happens to be left of the center (α^*, β^*), and the gradient at $(\alpha, \beta) = (0, 0)$ is

$$0.1 \begin{bmatrix} l_l \cdot 0 \\ l_2 \cdot 1 \end{bmatrix}$$

which regardless of l_1 and l_2 is up (both l_1 and l_2 are positive, remember). So the dynamics is clockwise.

The question was about the strategy profile $(\alpha, \beta) = (1/4, 1/4)$. This point is to the upper left of the center (α^*, β^*). With clockwise movement this means that the first player (the owner of α) moves towards the center and therefore is losing, while the second player moves away from the center and therefore is winning. Following the WoLF principle (win or learn fast), $l_1 = l_{\text{max}} = 2$, $l_2 = l_{\text{min}} = 1$, and the exact dynamics at $(\alpha, \beta) = (1/4, 1/4)$ is

$$0.1 \begin{bmatrix} l_l(\ \ 1\,\beta + 0) \\ l_2(-2\,\alpha + 1) \end{bmatrix} = 0.1 \begin{bmatrix} 2(\ \ 1\,\beta + 0) \\ 1(-2\,\alpha + 1) \end{bmatrix}.$$

It follows that the strategy profile after one iteration is

$$\begin{aligned} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{t+1} &= \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + 0.1 \begin{bmatrix} l_l(\ \ 1\,\beta + 0) \\ l_2(-2\,\alpha + 1) \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_t + 0.1 \begin{bmatrix} 2(\ \ 1\,\beta + 0) \\ 1(-2\,\alpha + 1) \end{bmatrix} \\ &= \begin{bmatrix} 1/4 \\ 1/4 \end{bmatrix} + 0.1 \begin{bmatrix} 2(\ \ 1/4 + 0) \\ 1(-1/2 + 1) \end{bmatrix} = \begin{bmatrix} 3/10 \\ 3/10 \end{bmatrix}. \end{aligned}$$

2. Suppose two agents use Bayesian learning to guide their play in *the repeated coordination game* with two actions C and D.

   Compute Agent 1's probability distribution over Agent 2's response rules after history CC, CC, DC, CD, if Agent 1 gives equal priors to the following response rules:

   all-C, all-D, TFT, unforgiving (a.k.a. "grim trigger"), mix $C = 1/2$, mix $C = 2/3$, fictitious play.

Answer. $P(h|s)$:

	all-C	all-D	TFT	GT	C=1/2	C=2/3	FP
CC	1	0	1	1	1/2	2/3	1/2
CC	1	0	1	1	1/4	4/9	1/2
DC	1	0	1	1	1/8	8/27	1/2
CD	0	0	1	1	1/16	8/81	0

The first entry of FP is $1/2$ because at the first round FP has nothing to go on and must randomise. (This is a direct consequence of the definition of fictitious play: randomise when the counts are tied.)

$P(h|s)P(s)$ (divide everything by 7):

	all-C	all-D	TFT	GT	C=1/2	C=2/3	FP
CC	1/7	0	1/7	1/7	1/14	2/21	1/14
CC	1/7	0	1/7	1/7	1/28	4/63	1/14
DC	1/7	0	1/7	1/7	1/56	8/189	1/14
CD	0	0	1/7	1/7	1/112	8/567	0

$P(s|h)$ (normalise last row, whether that is the last row from the first table or the last row of the second table):

	all-C	all-D	TFT	GT	C=1/2	C=2/3	FP
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
DC	0	0	0.463	0.463	0.0289	0.0457	0

Multiple-choice questions

1. With mixed strategies (α, β) for Row and Col, respectively, give the gradient in IGA of the game with the following payoff matrix

$$\begin{pmatrix} -3, 2 & 3, 1 \\ -3, 0 & -3, -2 \end{pmatrix}$$

   (a) $(-6\alpha + 6, -\beta + 2)$.
   √ $(-6\beta + 6, -\alpha + 2)$.
   (c) $(-\alpha + 2, -6\beta + 6)$.
   (d) $(-\beta + 2, -6\alpha + 6)$.

Explanation. See slide “Gradient of expected payoff”.

2. IGA doesn’t converge in the following case.

   (a) Real eigenvalue, stationary point within $[0, 1]^2$.
   (b) Real eigenvalue, stationary point outside $[0, 1]^2$.
   √ Imaginary eigenvalue, stationary point within $[0, 1]^2$.
   (d) Imaginary eigenvalue, stationary point outside $[0, 1]^2$.

Explanation. See slide “Convergence of IGA (Singh et al., 2000)”.

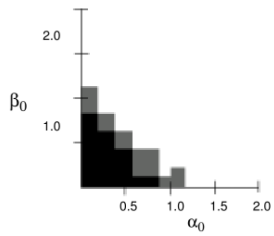
3. Which statements are true?

   i) Satisficing play stabilises if, at some round, all players are satisfied.
   ii) Satisficing play stabilises only if, at some round, all players are satisfied.

   √ *i)* and *ii)*.
   (b) *i)*.
   (c) *ii)*.
   (d) None.

Explanation. If and only if: if none of the players has an incentive to change, then they indeed won’t change and they won’t have an incentive to change in the next round for the same reason. As a result all players are locked into their current action profile forever. See slide “Self-play: possible dynamics”.

4. From: “Satisficing and Learning Cooperation in the Prisoner’s Dilemma”, Stimpson *et al.*, 2001.



Initial aspiration of player A on x-axis; Initial aspiration of player B on y-axis. White: convergence to *(C, C)*; black: convergence to *(D, D)*; grey: periodic or chaotic behaviour.

The a-symmetry is caused by different

   (a) Payoffs (σ, δ).
   (b) Persistence rates.
   √ Starting actions.
   (d) Aspiration levels.

> **Explanation.** See source article (or slides).

5. Joss-10% plays Tit-for-tat and defects 10% of the time at random moments. Give

$$\Pr\{\ h \mid \text{Agent 2 plays Joss-10\%}\ \}$$

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

(a) 0.009
(b) 0.081
(c) 0.729
√ Another answer.

> **Explanation.** Let $s_2 =$ Joss-10% for Player 2.
>
> $$\begin{aligned}\Pr\{h|s_2\} &= \Pr\{(C,C)|s_2\} \times \Pr\{(D,C)|s,(C,C)\} \\ &\qquad \times \Pr\{(C,D)|s,(C,C),(D,C)\}\ \ldots \\ &= 0.9 \times 0.9 \times 1 \times 0.9 \times 0.1 \\ &= 0.0729.\end{aligned}$$

6. In what respect differs the subjective distribution of play from the true distribution of play in Bayesian learning?

√ Reply rules are replaced by forecasting rules.
(b) Reply rules are replaced by predictive learning rules.
(c) Forecasting rules are replaced by predictive learning rules.
(d) Forecasting rules are replaced by behavioural strategies.

7. In Peyton Young's model of hypothesis testing, players may be thrown out of equilibrium occasionally.

(a) This is impossible.
√ This is a Type I error (to drop an otherwise correct model of opponent's play).
(c) This is a Type II error (to stick to an incorrect model of opponent's play).
(d) This is because at least one player maintains an incorrect model of opponent's play.

> **Explanation.** Type I error: all players maintain a correct model opponent's play while, at the same time, at least one player misrepresents opponent's play and changes strategy.
>
> To maintain an incorrect model of opponent's play due to bad sampling is a Type II error.

8. In a certain instance (you may read: implementation) of Peyton Young's generalised framework of hypothesis testing, the revised models are mappings from bounded histories to the product of mixed strategies, with probabilities rounded to two decimals.

What can be said about this particular instance?

(a) Models are of bounded recall, model revision is flexible.
(b) Models are of bounded recall, model revision is not flexible.
(c) Models are not of bounded recall, can't say whether model revision is flexible.
√ Can't say whether models are of bounded recall, and model revision is not flexible.

> **Explanation.** In order to be able to say whether models are of bounded recall we also need to know whether response functions themselves are of bounded recall.
>
> Model revision is not flexible because the distribution of revised models does not lie dense in Δ^M_{-i}. (For the same reason as to why every finite set, in particular $K = \{0, 0.01, 0.02, 0.03, \ldots, 0.99, 1.00\}$, does not lie dense in $[0,1]$.)

9. There exist uncoupled, 1-recall strategies that guarantee a.s. convergence of play to pure NE of the stage game (in all games where such equilibria exist).

(a) True, this can be shown through marginalisation.
(b) True, this can be shown by creating a Markov process.
(c) False, unless the stage game has real eigenvalues.
√ False, unless the stage game is generic.

> **Explanation.** First two entries in table “Overview of learning NE in MAL (column ‘Result’)” on slides.

10. There exist uncoupled 2-recall response rules that guarantee almost sure convergence of play to pure NE of the stage game in every game where such equilibria exist.

(a) True, this can be shown through marginalisation.
√ True, this can be shown by creating a Markov process.
(c) False, unless the stage game has real eigenvalues.
(d) False, unless the stage game is generic.

> **Explanation.** Third entry in table “Overview of learning NE in MAL (column ‘Result’)” on slides.

11. In comparing MAL algorithms empirically, why is it injudicious (i.e., not wise) to use the grand table as an end point of analysis?

(a) Stronger algorithms dominate weaker algorithms.
√ Weaker algorithms may blur the relationships among stronger algorithms.
(c) Entries possess a margin of error due to variance.
(d) Entries that correspond to probabilistic algorithms are realisations rather than expectations.

> **Explanation.** Evolutionary approaches, knock-out tournaments, and rank elimination are all are designed to get rid of the weaker algorithms as competitors (“partners”) to the stronger algorithms.

12. To consider the grand table as the payoff matrix of a 2-player game in normal form has the following benefit.

(a) NE correspond to fully mixed limit points of the replicator dynamic.
√ NE correspond to fixed points (not necessarily limit points) of the replicator dynamic.
(c) Strict NE correspond to fixed points of the replicator dynamic.
(d) Strict NE correspond to fully mixed limit points of the replicator dynamic.

> **Explanation.** (1) Every fully mixed limit point of the replicator dynamic is a Nash equilibrium. (2) Every Nash-equilibrium is a fixed point of the replicator dynamic. The converse of (1) and (2) are not true because species that are absent can never be introduced through the replicator.

Document last modified on Friday 28^{th} June, 2019 at 10:11.