Exam Performance Feedback Form

COMP30191 (was CS3191)

2006/2007

It should be pointed out that for the first time, the exam mark is not the final mark for this course unit. The final mark is calculated by applying a factor of .8 to the exam mark (taken out of 100) and adding it to the course work mark.

General remarks: 87 students sat the exam. The result was again good this year, with an average mark of 63%. In particular the first two questions, which can be practised and were similar to some of the assessed and non-assessed exercises, had a very high mark. As usual Question 3 was the hardest at all and didn't have many takers (most of whom would have been much better off picking a different one). I was particularly disappointed with the low average for Questions 4 and 5. Almost all of Question 4 was bookwork and closely related to previous exam questions, and for Question 5 parts a) and c) were straight-forward definitions and results from the notes, and Part b) is related to Assessed Exercise 4. Clearly this year's cohort could have learned something from last year's, as far as these two questions are concerned.

Question 1. All student attempted this question. All in all this question was very well answered and achieved an average mark of above 76% (the highest ever). 5 students failed to achieve a pass mark for that question, while 34 managed a mark of 18 or higher.

Reasons why marks were lost typically were:

(a) The majority of students had the right shape to their game tree—a few assumed that Alice could fold if Bob raises the stakes, while the text clearly states that she has to match his bet. Many students lost one mark because they didn't mark the information set that makes it clear that Bob cannot distinguish between the two positions where Alice has bet (holding the Queen or the King) when he has to make his move. I had made it very clear that not including this information a player has when he has to move would be penalized. A handful of people got the pay-offs wrong, giving the amount that is in the pot rather than the amount a player loses or wins.

(b) The majority of students got the strategies right. There are four for Alice:

- always bet;
- always fold;
- bet with a Queen but fold with a King;
- fold with a Queen but bet with a King.

A handful of students tried to use strategies that were only one half of a strategy ('with a Queen bet'). Bob has three strategies: He can fold, match Alice's bet, or raise the stakes. A few people got that wrong because they assumed that Bob knew Alice's card, although the text clearly states that she does not show it to him.

(c) The majority had the correct matrix for this game. The exceptions were people who had the wrong strategies, and a few who had forgotten how to calculate a matrix where there are two possible outcomes when two players choose a strategy each, although it was just the same method as for Assessed Exercise 1. Some people had Bob as the row player but gave the pay-offs for Alice, which leads to incorrect results.

(d) The correct matrix has a mixed strategy equilibrium point which can be found using (pure strategy) dominance and solving the resulting symmetric 2×2 matrix. People who had wrong matrices with pure strategy equilibrium points lost marks here because they could read theirs off in 2 seconds, whereas those with the correct matrix had to do some calculations. A mistake that was made a few times was to reduce to a matrix that has two 1s along one diagonal and 0s elsewhere; the correct matrix has a 1 and 1/2 as its non-zero entries. Some people confused the probabilities and put them into the wrong components (also see Question 2 (c)).

(e) A lot of people made mistakes calculating the value—typically because they used the more complicated method of calculating the pay-off at the equilibrium point instead of using the easier method that comes with solving a (2×2)-matrix. A few people made a hash of arguing about fairness. A surprising number wanted to play a particular role based on their interpretation of the rules of the game rather than the calculations they had just made, and a lot wouldn't have picked the equilibrium point strategy! I do wonder why they think we do all this.

Question 2. All but two students tackled this question. This question has a mistake in part b); the matrix as given cannot be reduced with dominance methods as taught on this course. While a few students apparently spotted that there was a problem none of them alerted the invigilators, so that I had no way of correcting the mistake (and giving them more time to make up for it) during the exam.

The average mark for this question was 73%, the highest ever. Three students failed to get a pass mark for it, and 15 achieved a mark of 18 or higher.

Reasons why some marks were lost typically were the following.

(a) Some students forgot to give the value, losing a mark. A few missed out one or more of the four (pure strategy) equilibrium points for some reason, even fewer tried to solve this using dominance arguments—and as a result didn't find all pure strategy ones. (Note that I did not ask for mixed strategy equilibrium points and no marks were awarded for talking about those.) Some students didn't give the actual strategy pairs but only circled some entries in the matrix. Four marks were available for the four equilibrium points, one for the correct value.

(b) Due to the mistake in the exam paper I marked this part particularly generously. Anybody who wrote that they thought there might be a mistake in the exam paper got all 5 marks. Students who tried some reductions before giving up got 4 marks. A surprising number of students were not able to properly use mixed dominance: They removed strategies although there was no parameter between 0 and 1 that solved their inequalities simultaneously. I would ordinarily give 0 marks for something like that, but in this case I was more generous. Even people who wrote nothing still got some marks. I further kept an eye on whether delays in answering this question might have held up students when they tried to solve others.

(c) Most students knew how to solve a (2×2)-matrix. In a few cases marks were lost because

- Player 1 and Player 2 were confused;
- students made mistakes in solving the equations—a few of them obtained non-sensical results (a probability of 0, or a negative one) and still did not go back to check their work;
- a very few students inserted the value they got from solving the equations into the wrong position in the tuple describing the equilibrium point strategy;
- a few students don't seem to have caught on to the fact that the easiest way of calculating the value of the game is to take the value calculated by solving the equations and inserting it into either of the equations—the resulting number is the value of the game.

(d) The given matrix can be reduced to a (3×3)-matrix using dominance arguments. Many students performed an *impossible* reduction because from the given equilibrium points they concluded that they had to get down to (2×2). A few people did not realize that in order to show that the given equilibrium point is correct they were supposed to use Proposition 2.4 in the way shown on page 45 of the notes, and asked for in Exercise 14 (a). One or two managed to produce a proper proof without that.

Question 3. Only 6 students decided to answer this question, which was the hardest in the exam (and which had been announced as such a number of times). One of these didn't make a serious attempt and received 0 marks, three could write something sensible but couldn't give proofs and got 7 or 8 marks, and two made decent attempts. The average mark was 43%.

In the serious attempts of solving the question the mistakes were

(a) realizing that an inductive argument was required, but being unable to properly prove the induction step;

(b) being unable to connect the values calculated by the minimax algorithm with the stated objective;

(c) not realizing there had to be infinitely many equilibrium points, or being vague about the reasons why.

Question 4. This question was attempted by 21 students. It had an average mark of 48%. Four students failed to get a passing mark for the question, and nobody got a mark of 18 or higher.

Many of the answers were somewhat confused and contained details that were wrong. Reasons why marks were typically lost are the following:

(a) Not fully describing how the alpha-beta algorithm works, not presenting an example step-by-step, making mistakes regarding how the bounds are updated, having given values for nodes other than the roots, forgetting to answer why it is faster than the minimax algorithm, or making wrong statements about that. There were ten marks to be had for this question, and many answers didn't give enough detail to be awarded most of the marks.

(b) The one thing that future sitters of this exam should take to heart is that they should answer a question like this as if the reader didn't really know much about game playing programs. Very few students attempted to describe how a game playing program works, and often were confused about details. In particular, the alpha-beta search algorithm searches the tree *to a given depth and only applies the evaluation function to nodes at that depth.* It then calculates an estimated value for each position reachable by the next move by treating these values as if they were the real things. Many answers contained details that the question didn't ask for, and for which no marks were available.

(c) This was a somewhat speculative question. I expected to be told that many available programs can beat the average player and, that even the world's best player (at the time) has been beaten by a program (although one fine-tuned to do well against him. I further expected some comments regarding the different approaches humans and computers take: A computer will search a very large part of the game tree, and the reason it may do well against humans is that it is using a brute force approach. On the other hand, good human players will be able to determine only a handful of possible moves to look at.

Question 5. Sixty students attempted this question, and 24 of them did not manage to get a passing mark (that is, they had a mark below 8). Many of these same students had achieved quite high marks in Questions 1 and/or 2. Only one managed a mark of 18 or better. I was disappointed with the general quality of the answers to this question. The marks were not low because people ran out of time (although some answers clearly had been written in haste) but because many did not give me the information I had asked for.

Marks were typically lost for the following reasons.

(a) This question was worth five marks, three for the definition and two for the two example strategies. Many students failed to give a proper definition of collective stability (Definition 12), or merely gave the part that says that a strategy cannot be invaded without saying what that means. For three marks more than just one sentence was required. Also, many students stated that TitForTat is collectively stable, but failed to add that this is only the case if w is large enough.

(b) This is a question similar to Assessed Exercise 4, but only a few students managed to give a proper proof. Many were confused about what exactly was required here, quite a few only did a proof for a one round PD game (rather than the indefinitely repeated version as asked), and a few didn't use general pay-offs, filling in their own numbers.

(c) All I wanted here was that collectively stability implies territorial stability (Proposition 6.8). I got much irrelevant information on how the two systems differ, and much of that was even wrong.

(d) A lot of people apparently didn't properly read this question. I asked them to pick a simulation *from the literature*. Any of the scenarios described in Section 6.5 would have been fine. Only a few people wrote about these, and nobody really answered my question regarding the choice of strategies which were represented in the system. Some people had a decent go at using Axelrod's territorial system as an example, but mostly what I got here was information irrelevant to the question as asked.