Exam Performance Feedback

COMP34120—Questions 1 and 2

2011/2012

The following only concerns Questions 1 and 2 of this exam.

It should be pointed out that the exam mark is not the final mark for this course unit. The final mark is calculated by applying a factor of .4 to the exam mark (taken out of 100), multiplying each of the coursework marks with .3, and adding the three together.General remarks: 44 students sat the exam.

Question 1. Again I disappointed with the many serious mistakes found in many of the answers for this question. Overall, performance was better than in the previous year. The question was worth twenty marks, and of the 37 students attempting it eleven received a mark of seven or lower (that is a failing grade) and while six managed a first class mark. The average mark was 9.5. Students who had marks of five or lower wrote down very little that was both relevant and true.

a) I thought this was easy, but a surprising number of students did not follow the simple instructions of the game given. In particular, the first thing that happens is that Alice draws a card. Since she is allowed to see her card this does make a difference for her strategies! Apart from not following the structure of the game typical mistakes were forgetting probabilities for the chance move, not drawing information sets correctly (Bob does not know which card Alice has, and that is the only point of uncertainty), and giving incorrect pay=offs. The game tree consists of a maximum of three moves and has thirteen nodes.

   Alice has four strategies, but many students did not describe the correct ones. Her strategies are:

   - Bet on either card.
   - Fold on either card.
   - Bet on King, fold on Queen.
   - Fold on King, bet on Queen.

   I was happy to accept any definition, either verbally or as a subtree of the game tree. Since Bob does not know which card Alice has he has three strategies: Fold, accept the bet, double the bet.

   Many students did not describe the strategies at all but tried to argue somehow how the two players should play based on pay-offs they might reach. The notion of 'strategy' is mathematically defined in the introductory chapter of my notes, which seems to have escaped a number of students.

b) There are two pure strategy equilibrium points in the game, including the choice of strategy that leads to the pay-off of $(5, 5)$. This is the best outcome either player can hope for so that is the choice of strategy they will both play in practice. A few students missed the second equilibrium point, losing a mark, but by and large this was well answered.

c) Most students only addressed the issue of playing such a game in normal form, and then they described the same calculation in a number of different ways. The model answer reserved quite a few marks for discussing the extensive case which most students missed out on.

d) I probably shouldn't have specified a particular form of game here, but in practice that doesn't seem to have made a difference. Many students could write down either a game tree or a matrix that satisfied the requirement (if in normal form it must have either two identical entries in one row that are maximal in their respective columns, or two identical entries in a column that are minimal in their respective rows). This cannot be achieved with four different pay-offs since in a two-person zero-sum game all equilibrium points must lead to the same pay-off.

Question 2. This question was attempted by 36 students. The average mark was 10.1 out of 20. Eight students got a failing mark and four managed a first class answer. This was a bit different from last year by specifically asking about what students have taken away from the Semester 1 project.

a) Almost all students correctly stated that Kalah is a 2-person zero-sum game of perfect information without chance (although quite a few forgot one of these properties), which means that Theorem 1.1 applies. A solution is given by an equilibrium point, and we know that it must be the case that either one of the two players has a winning strategy or that they can both force a draw. (It is currently not known which of these applies for Kalah with seven or more wells, but the solution is still known to exist—something I made a big song and dance about in the relevant lecture.) I got the impression that those who listened well in the revision session had an advantage here. This part was certainly better answered than last year.

b) There were a lot of descriptions that weren't very meaningful—most students were able to list components, but the description of how it all fit together to produce a move were often missing, vague, or did not make sense (I'm sure the evaluation function is not applied to each position encountered, for example—or at least I hope so). Ironically a number of students claimed components as the best feature in part c) which they hadn't even mentioned in part b)!

c)–e) Giving a best/worst feature plus 'one extra idea' is, of course, easy. Each of these questions was worth four marks, and I expected arguments to be given for why the feature mentioned was as important as claimed. Very little evidence was provided, and very few students even referred to their program's performance in the tournament. For the 'best feature' I also wanted to know why students thought it was as good (not just as important) as they claimed, and again very little evidence was given. Similarly for the worst feature. Again, for the 'single additional idea' I was looking for arguments as to

- why it would have made a big difference to the performance of the program in question (and it should be a big one),
- why the supposed idea would have been feasible,
- why nothing else would have worked so well.

Often students gave one fairly general issue ('Our program would have been better if it had had an opening database') but then struggled to give more than very generic reasons for why this would be the one important change.

I was surprised with how little some students wrote for the four marks, and how little reference was made to any testing the students had performed themselves. (I wasn't looking for detail, just the big picture. 'We know that our heuristic function was excellent because we did x testing, and because we could beat other programs which looked further ahead', or words to that effect.)