OCR
RECOGNISING ACHIEVEMENT

ADVANCED GCE

MATHEMATICS (MEI) 4754B

Applications of Advanced Mathematics (C4) Paper B: Comprehension
INSERT

Monday 1 June 2009
Morning

Duration: Up to 1 hour

4754B

INSTRUCTIONS TO CANDIDATES

- This insert contains the text for use with the questions.

INFORMATION FOR CANDIDATES

- This document consists of **8** pages. Any blank pages are indicated.

© OCR 2009 [T/102/2653] OCR is an exempt Charity
RP–8J18

The Prisoner's Dilemma

Introduction

During the 1st World War a curious sort of truce often occurred between the two sides. Between fierce battles there were long periods when nothing much happened; the soldiers were living in trenches quite close to the enemy and, without any conversations taking place, understandings often arose that they would not shoot at each other. A British officer visiting the front line was horrified by what he saw.

I was astonished to observe German soldiers walking about within rifle range behind their own line. Our men appeared to take no notice. I privately made up my mind to do away with that sort of thing … ; such things should not be allowed. These people evidently did not know there was a war on. Both sides apparently believed in the policy of 'live and let live'.

This did not mean that the soldiers on the two sides had become friends. They did not know each other and they certainly would shoot to kill when the next major battle occurred, but in the meantime co-operation was a better policy.

How could such behaviour arise spontaneously? This article looks at a mathematical model that describes it and allows it to be simulated.

What happened across the trenches of the 1st World War is just one example of a general situation in which opposing groups, between whom there is no mutual trust or friendship, nonetheless find it beneficial to co-operate. Other examples include an arms race between two countries and commercial competition between companies. Another situation arises when each of two suspects is offered a lighter prison sentence in exchange for giving information against the other; this has given rise to the general name 'The Prisoner's Dilemma' for all such situations.

Modelling the situation

Imagine the situation in the 1st World War. On any occasion the soldiers on each side had two options.

- They could 'co-operate' with the other side by not shooting. (Option C)
- They could 'defect' by breaking the agreement and shooting. (Option D)

So between the two sides, there were four possibilities as shown in Table 1.

	Side 2 co-operates	Side 2 defects
Side 1 co-operates	C C	C D
Side 1 defects	D C	D D

Table 1

The 'benefits' and 'costs' to each side of these different situations may be described as follows.

- In the situation C C, the two sides co-operate and they both benefit; no one gets shot (and so killed or injured).
- In the situation D D, both sides defect and shoot at each other. There is a cost to both sides because some of their soldiers get shot.

PMT

- In the situation D C, Side 1 unexpectedly defects by breaking the agreement and shooting some of the enemy soldiers. This is of short-term benefit to Side 1 by advancing the war effort; in contrast, Side 2 pays the cost of losing some soldiers.

- The fourth situation C D is the mirror image of D C. In this case there is a cost to Side 1 because some of their soldiers are shot, and there is a short-term benefit to Side 2.

There are thus four possible levels of benefit that may be modelled as follows.

Both sides **c**o-operate (C C). Each benefits by c units.

Both sides **d**efect (D D). Each benefits by d units, where d is negative.

One side co-operates and the other defects (D C or C D):

> the defecting side benefits by b units from **b**reaking the agreement and the co-operating side, for whom this is the **w**orst possible outcome, benefits by w units, where w is negative.

The situation being modelled means that

$$b > c > d > w.$$

Turning the model into a game

This model has been the subject of extensive study, using the technique of turning it into a game. There are two players and, at each turn, they declare C (co-operate) or D (defect) at the same time as each other. In this article the various benefits are set as the following values, although other values are commonly used.

When both players co-operate (C C), each scores 1 point: $c = 1$.

When both players defect (D D), each scores −1 point: $d = -1$.

When one player co-operates and the other defects (D C or C D), the player who co-operates scores −2 points: $w = -2$. The defecting player benefits by 3 points: $b = 3$.

This raises two questions.

- What is a good strategy for the game?
- What, if anything, does the game tell us about human behaviour?

Playing a single round

Imagine that you are one of the two players. Start with the case when there is only a single round of the game. It is possible to apply simple logic to this situation. Remember that both players declare at the same moment.

The other player is going to declare either C or D.

Take first the case when the other player declares C.

> If you declare C, you score 1 point.
>
> If you declare D, you score 3 points.

So you are better to declare D.

© OCR 2009

PMT

Now take the case when the other player declares D.

If you declare C, you score –2 points.

If you declare D, you score –1 point.

So again you are better off to declare D.

So, whatever the other player declares, your better option is D.

However, the other player can be expected to apply the same logic and so also to declare D. So the outcome is predictable as being D D, worth –1 point to each player. This seems paradoxical when this is not the best possible outcome for either player, since C C would give both players a score of 1 point.

This result does, however, make sense in terms of human behaviour if neither party expects to meet the other again. Soldiers might shoot to kill in a one-off war-time encounter, and in a single commercial transaction both parties might seek to make as much money out of the other as possible.

Co-operation becomes more likely when the two parties expect to have a long-term relationship.

A large number of rounds

Now suppose that you are playing the Prisoner's Dilemma game with a very large number of rounds, with no end in sight. What is a good strategy ?

Random choice

One possible strategy is to choose C and D at random, both with probability $\frac{1}{2}$. Suppose that both players do this independently. Then on any move there are 4 possible outcomes: C C, C D, D C and D D. The scores for these are shown in Table 2.

Outcome	C	C	C	D	D	C	D	D
Scores	1	1	–2	3	3	–2	–1	–1

Table 2

These four outcomes are all equally likely so each has a probability of occurring on any move of $\frac{1}{4}$; on average each will occur once every 4 rounds. The average score for each player works out to be 0.25 points per round.

Constant choice

Another simple strategy is to make the same choice, either C or D, on every round. The problem is that your opponent will soon realise what you are doing and will seek to exploit it. Table 3 shows a case when you always co-operate. At some point, in this example on the third round, your opponent will defect and will continue to do so once it is clear that you will continue to co-operate. The game settles down with you scoring –2 points every round and your opponent 3 points.

© OCR 2009

Round	You	Opponent	Your score	Opponent's score
1	C	C	1	1
2	C	C	1	1
3	C	D	−2	3
4	C	D	−2	3
5	C	D	−2	3
6	C	D	−2	3
7	C	D	−2	3
8	C	D	−2	3
9	C	D	−2	3
10	C	D	−2	3
…	…	…	…	…

Table 3

Clearly the longer the game goes on the closer your average score approaches −2 points per round and that of your opponent approaches 3.

Table 4 shows a possible game when you always defect.

Round	You	Opponent	Your score	Opponent's score
1	D	C	3	−2
2	D	C	3	−2
3	D	D	−1	−1
4	D	D	−1	−1
5	D	C	3	−2
6	D	D	−1	−1
7	D	D	−1	−1
8	D	D	−1	−1
9	D	D	−1	−1
10	D	D	−1	−1
…	…	…	…	…

Table 4

After some attempts at co-operation your opponent realises that the best strategy playing against you is also to defect. Once the game settles down you both score −1 point at each round.

Both of these two fixed-choice strategies result in your obtaining a lower average score per round than you would have done by choosing at random. They illustrate the fact that your opponent will try to learn from your play and then to benefit from it. Even though you do not talk to your opponent, communication is still taking place through your actions, as happened between the trenches of the 1st World War.

© OCR 2009

PMT

Tit-for-tat

An alternative strategy to adopt is 'Tit-for-tat'. In this, you always do the same thing as your opponent did last time. So the first few rounds of a game might be as in Table 5.

Round	You	Opponent	Your score	Opponent's score
1	C	C	1	1
2	C	C	1	1
3	C	C	1	1
4	C	D	−2	3
5	D	D	−1	−1
6	D	C	3	−2
7	C	C	1	1
8	C	D	−2	3
9	D	C	3	−2
10	C	C	1	1
...	...	...	...	...

Table 5

In this example, although you both start off co-operating scoring 1 point per round, in Round 4 your opponent defects and so scores 3 points on that round to your −2. However, you respond immediately by defecting on Round 5 and you continue to defect until the round after your opponent next co-operates; thus your opponent co-operates on Round 6 and you co-operate again on Round 7.

Over Rounds 4, 5 and 6 your opponent scores $3 + (-1) + (-2) = 0$ points and this is less than the 3 points that would have resulted from co-operating for those three rounds.

In Round 8 your opponent tests you by defecting again and you respond immediately by defecting in Round 9. By Round 10 you are back to mutual co-operation but this short-term defection has cost your opponent 1 point.

Once your Tit-for-tat strategy is clear, the only sensible thing for your opponent to do is to co-operate at every round, leading to a long-term average score of 1 point per round for both of you.

Thus Tit-for-tat is a strategy which allows long-term co-operation to evolve. It was used strictly in the 1st World War. Both sides knew that if they fired at the other, there would be instant retaliation.

The example in Table 5 illustrates an important feature of the scoring system for the Prisoner's Dilemma game. In Rounds 8 and 9, your opponent scored $3 + (-2) = 1$ point; since this was less than the $2 \times 1 = 2$ points available for continued co-operation, defection did not pay. The benefits b, c, d and w were assigned the values 3, 1, −1 and −2 respectively, but these are not the only possible values that obey the inequality on line 47, $b > c > d > w$. If, for example, b were given the value 10 and the other values remained the same, then a short-term defection would be a profitable thing to do.

So for long-term co-operation to be the best option, the following further condition must be fulfilled.

$$b + w < 2c.$$

© OCR 2009

PMT

This inequality may be written as $b < 2c - w$ and may be interpreted as saying that long-term co-operation is only possible if the benefit from defection, b, is not too great.

Competitions for the best strategy

Because the Prisoner's Dilemma can be used to model a great variety of situations, it has attracted a large amount of academic interest; many research papers have been written about it. In 1984 Robert Axelrod published *The Evolution of Co-operation*; it has now become a classic book on the subject. In this he reported on a large computer-based competition to determine the best strategy. There were 62 entries, each in the form of a computer program, from a wide variety of sources. Each of them played all the others over 200 rounds, and the whole exercise was carried out 5 times. The winning strategy was Tit-for-tat.

Axelrod analysed the most successful strategies and found they all had certain characteristics, which he described in the following terms.

- They were all *nice*; that is they would not defect before the opponent did.
- They would always *retaliate* when an opponent defected.
- They were *forgiving*, returning to co-operation once the opponent ceased to defect.
- They were *non-envious*, seeking to maximise their own benefit rather than to reduce that of their opponents.

If this sounds rather idealistic, it is not. It is a statement that a selfish individual acting entirely out of self-interest will nonetheless behave in ways that are generally thought to be good.

In conclusion

This article introduces the Prisoner's Dilemma game, but it only scratches the surface. The game itself can be refined in various ways but there is much more that can be learnt from it even in its simplest form.

It is said that one of the challenges for mathematics is to develop techniques for studying and predicting human behaviour. The Prisoner's Dilemma provides one example of how using a suitable game may allow this to be done.

PMT

OCR
RECOGNISING ACHIEVEMENT

Copyright Information

OCR is committed to seeking permission to reproduce all third-party content that it uses in its assessment materials. OCR has attempted to identify and contact all copyright holders whose work is used in this paper. To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced in the OCR Copyright Acknowledgements Booklet. This is produced for each series of examinations, is given to all schools that receive assessment material and is freely available to download from our public website (www.ocr.org.uk) after the live examination series.

If OCR has unwittingly failed to correctly acknowledge or clear any third-party content in this assessment material, OCR will be happy to correct its mistake at the earliest possible opportunity.

For queries or further information please contact the Copyright Team, First Floor, 9 Hills Road, Cambridge CB2 1PB.

OCR is part of the Cambridge Assessment Group; Cambridge Assessment is the brand name of University of Cambridge Local Examinations Syndicate (UCLES), which is itself a department of the University of Cambridge.

© OCR 2009