RECOGNISING ACHIEVEMENT

ADVANCED GCE

4754/01B

MATHEMATICS (MEI)

Applications of Advanced Mathematics (C4) Paper B: Comprehension

INSERT

WEDNESDAY 21 MAY 2008

Afternoon

Time: Up to 1 hour

INSTRUCTIONS TO CANDIDATES

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

This document consists of **11** printed pages and **1** blank page.

Sudoku puzzles

Introduction

Sudoku puzzles were introduced to the UK in 2004 and almost immediately became very popular. Many people found them addictive. A common form of Sudoku puzzle uses a 9×9 grid, as illustrated in Fig. 1.

- A square grid consists of small squares or *cells*. In the case of the 9×9 grid, there are 81 cells.
- The square grid is also divided into a number of smaller square grids, or *blocks*. In the case of the 9×9 grid, there are 9 blocks; each of them is a 3×3 grid and so contains 9 cells.
- When the puzzle is completed, each row, each column and each block will contain each of 9 different *symbols* exactly once. The symbols used in a 9×9 Sudoku puzzle are usually the numerals 1, 2, 3, 4, 5, 6, 7, 8 and 9 but they could just as well be letters, for example A, B, C, D, E, F, G, H and I.
- The symbols in some of the cells are already provided at the start, and these are called the *givens*.
- You have to fill in the symbols in the remaining cells. If the puzzle has been set correctly, there is only one solution; it is unique.

[Warning Do not spend time in this examination trying to complete this puzzle. You are not expected to solve it.]

								1
					3		2	
		4			9			
							9	
					4	6		
	1	7		8				
				1				7
	3		5					
9						4		

Fig. 1

In this article, the term *Sudoku* is used to describe a correctly completed grid whereas *Sudoku puzzle* is used for the initial situation where there are empty cells to be filled in. A Sudoku puzzle leads to a unique Sudoku, but the converse is not true; any Sudoku can be arrived at from many different possible starting points.

There are several variations on the basic Sudoku puzzle. Some involve grids of different shapes but, in this article, only those with square grids containing square blocks are considered. The size of the grid does not need to be 9×9; another commonly used grid is 16×16 (using the symbols 0 to 9 together with A, B, C, D, E and F), and you can even have 25×25 Sudoku puzzles. A simpler possibility is 4×4. Thus, in this article, the side length of the grid, s, is a square number and $s = b^2$, where b is the side length of each block.

Questions about Sudoku puzzles

This article investigates three questions about 9×9 Sudoku puzzles.

1. In how many different ways can you place the symbols 1 to 9 in the 81 cells of a 9×9 Sudoku grid, subject to the Sudoku rules?
2. What is the smallest number of givens that must be provided if a puzzle is to have a unique solution?
3. What is the largest number of givens that can be provided without the puzzle having a unique solution?

Approach

Two different problem-solving techniques are used in this article.

- The modelling approach is to start by working with a similar but simpler problem, see what you can learn from it and then attack the real problem; in this case the simpler problem is the Latin square puzzle.
- The other approach is to look at the same problem but on a smaller scale; you hope for one or both of two outcomes.
  - The smaller scale case will help you to find an appropriate method.
  - You can build a sequence of results from several smaller cases which will suggest a possible formula for a general result.

Latin squares

In a Latin square, the grid is not divided into smaller blocks as in a Sudoku, but each symbol must appear once in each row and in each column. An example of a 4×4 Latin square is given in Fig. 2.

1	2	3	4
3	1	4	2
2	4	1	3
4	3	2	1

Fig. 2

Notice that this Latin square could not be the solution to a Sudoku puzzle.

All Sudokus are Latin squares, but not all Latin squares are Sudokus. Consequently there are more Latin squares of any size (other than the trivial 1×1 case) than there are Sudokus, and so the number of Latin squares provides an upper bound for the number of Sudokus.

Latin squares have been known for a long time and in medieval times were used for decoration in the Islamic world. They were re-popularised in the late 1700s by Leonhard Euler who set puzzles on them that are often said to be the predecessors of Sudokus. These days, Latin squares are widely used in experimental design in statistics. In this article a puzzle which requires you to complete a Latin square is referred to as a Latin square puzzle.

Investigation 1: How many 9×9 Sudokus are there?

The first investigation involves the number of ways that it is possible to allocate the symbols 1, 2, 3, …, 9 to the 81 cells on a Sudoku grid.

The first approach to the problem is to look at Latin squares of different sizes.

Trivially, there is one possible 1×1 Latin square.

There are two possible 2×2 Latin squares and these are shown in Fig. 3.

1	2
2	1

2	1
1	2

Fig. 3

What about 3×3 Latin squares? Start by thinking about the top row.

- Choose the top left symbol first. It can be 1 or 2 or 3 so there are 3 possibilities.
- Now choose the top middle symbol; whatever the top left symbol, there are now 2 remaining symbols to choose from for this cell.
- There is only 1 possibility for the last symbol in the row, in the top right cell.
- So the number of ways of filling the cells in the top row is $3 \times 2 \times 1 = 3! = 6$.

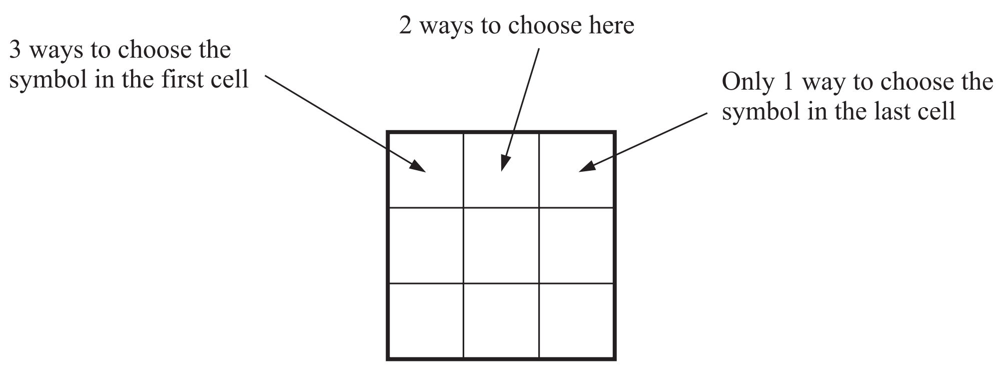

Fig. 4

Now think about the left-hand column.

- The symbol in the top cell has already been chosen.
- There remain 2 possible symbols for the middle left cell.
- There is only 1 possibility for the symbol at the bottom of the column.
- So for any given top row, the number of ways of filling the rest of the left-hand column is $2 \times 1 = 2! = 2$.

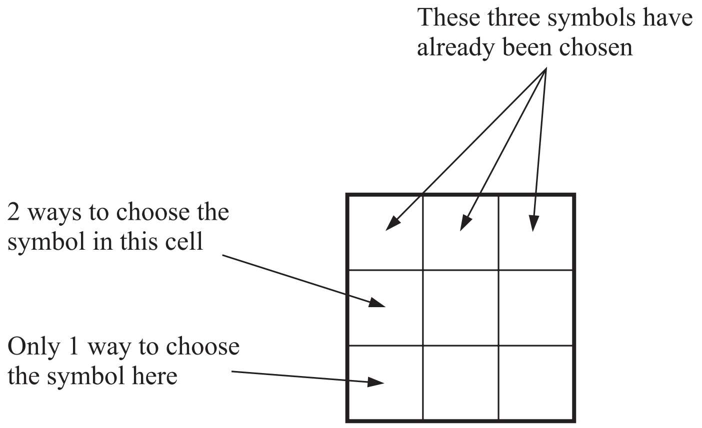

Fig. 5

In each case there is only one way to fill in the four remaining cells, as shown in the example below where the top row is (1 2 3) and the left-hand column is (1 3 2).

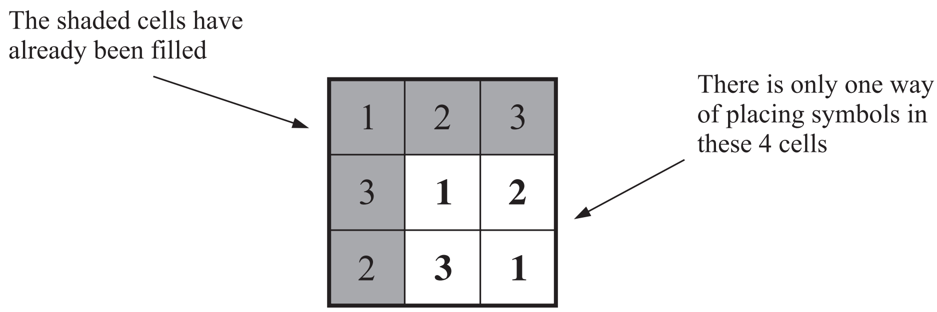

Fig. 6

So the number of 3×3 Latin squares is given by

	Top row		Left-hand column		Rest		
	6	$\times$	2	$\times$	1	=	12.
i.e.	3!	$\times$	2!	$\times$	1		

This process is illustrated in Figs. 4, 5 and 6.

The next step is to investigate the 4×4 Latin square, using the same approach.

- The number of ways of filling the cells in the top row is $4 \times 3 \times 2 \times 1 = 4! = 24$.
- The top cell in the left-hand column has already been filled, so there are just 3 cells remaining; there are $3 \times 2 \times 1 = 3! = 6$ ways of filling these cells.
- That leaves 9 cells, as shown in the example in Fig. 7. There is nothing special about the symbols in the top row and left-hand column; they are just a typical selection.

1	4	2	3
3			
2			
4			

These 9 cells have still to be filled in

Fig. 7

The next step is to find the number of ways of filling the remaining 9 cells. The answer turns out to be 4.

So the total number of 4×4 Latin squares is given by

Top row		Left-hand column		Rest		
$4!$	$\times$	$3!$	$\times$	4	$=$	576.

A similar procedure can be used to find the number of 5×5 Latin squares. Considering the top row and the left-hand column gives $5! \times 4!$ but counting the rest is distinctly messy. There are actually 56 ways, so the number of 5×5 Latin squares is $5! \times 4! \times 56 = 161\,280$.

The hope that finding the numbers of Latin squares of sides 1, 2, 3, 4, … would suggest a possible formula now looks rather forlorn.

Table 8 gives the numbers of Latin squares for sides of up to 11. The sequence is remarkable because the numbers get so large so quickly. At the time of writing this article, no general formula is known and no-one has worked out the exact number for 12 or beyond.

Side	Number of Latin squares
1	1
2	2
3	12
4	576
5	161 280
6	812 851 200
7	61 479 419 904 000
8	108 776 032 459 082 956 800
9	5 524 751 496 156 892 842 531 225 600
10	9 982 437 658 213 039 871 725 064 756 920 320 000
11	776 966 836 171 770 144 107 444 346 734 230 682 311 065 600 000

Table 8

So what does this tell you about the numbers of possible Sudokus?

The side, s, of a Sudoku grid must be a square number, so $s = 1$ or 4 or 9 or 16 or … .

Apart from the trivial case when $s = 1$, the number of possible Sudokus is always less than the number of Latin squares for the same value of s. So the number of possible 4×4 Sudokus is less than 576 and the number of 9×9 Sudokus is less than the 28-digit number in Table 8.

Given the difficulty of finding a formula for the number of Latin squares, and the fact that the Sudoku grid is by its nature more complicated, it would be surprising if there were an easy general formula for the number of Sudokus of side s.

It is not, however, difficult to find the number of 4×4 Sudokus, using a similar method to that for Latin squares. Fig. 9 shows one of the $4! = 24$ ways of completing the top row.

Because of Sudoku rules the symbol in this cell must be either 3 or 4

1	2	3	4

Fig. 9

It turns out that there are 16 different ways of filling in the remaining cells while keeping to the Sudoku rules. One of these ways is shown in Fig. 10.

1	2	3	4
3	4	2	1
2	1	4	3
4	3	1	2

Fig. 10

So there are 384 possible 4×4 Sudokus. Notice that this is indeed fewer than the 576 Latin squares of the same size.

The method used involved both logic and counting. This mixture of logic and counting has been applied successfully to the 9×9 Sudoku by Felgenhauer and Jarvis of the University of Sheffield, but with the difference that, following clever programming, the counting was done by computer. The final outcome was the 22-digit number 6 670 903 752 021 072 936 960.

Investigation 2 **What is the smallest number of givens that must be provided if a puzzle is to have a unique solution?**

A completed 4×4 Latin square contains the symbols 1, 2, 3 and 4. Look at Fig. 11 as a starting grid for a Latin square puzzle. There are 9 givens and 7 blank cells to be filled in. Only three of the symbols, in this case 1, 2 and 3, appear but you can still find the symbols for the remaining cells easily enough, and there is only one way of doing so.

2	1	3	
1	3		2
3		2	
	2		

Fig. 11

However, if only two of the four symbols, say 1 and 2, appear at the start, there must be more than one way of filling in the grid because the other two symbols, in this case 3 and 4, are interchangeable. The same is true for a 4×4 Sudoku puzzle.

So it is certain that for there to be a unique solution to a 4×4 Latin square or Sudoku puzzle, the givens must include at least three distinct symbols. It follows that the number of givens must be at least three.

However, a systematic search shows that there is no unique solution for either a Latin square puzzle or a Sudoku puzzle if there are just three givens.

The next case to try is four givens. Fig. 12 shows a Latin square puzzle with four givens and a unique solution.

1		3	
	4		
		1	

Fig. 12

Fig. 13 shows a Sudoku puzzle, also with four givens and a unique solution.

1		3	
	4	2	

Fig. 13

Interestingly, neither of these is a starting grid for the other type of puzzle.

The situation for the 9×9 Sudoku puzzle is much more complicated because of the increased number of possibilities but the required method is still systematic search, albeit with a well programmed computer. So far the smallest possible number of givens that anyone has found is 17, and many examples are now known of such puzzles, including that shown in Fig. 1. At the time of writing this article no-one has found a Sudoku puzzle with just 16 givens.

Investigation 3 **What is the largest number of givens that can be provided without a 9×9 Sudoku puzzle having a unique solution?**

Look at the 4×4 Latin square puzzle in Fig. 14. The givens cover twelve out of the sixteen cells. However, even with this number of givens, the remaining four cells cannot be filled in uniquely. There are two possible ways to complete the puzzle. This is because the four missing entries form an *embedded* Latin square.

4	2	3	1
		2	4
		4	2
2	4	1	3

Fig. 14

Fig. 15 shows the two possible arrangements for the symbols in the four remaining cells. In both arrangements the four symbols form their own 2×2 Latin square.

1	3
3	1

3	1
1	3

Fig. 15

The four symbols in a 2×2 embedded Latin square do not need to be next to each other but they do need to lie in the four corners of a rectangle. There are three other embedded Latin squares in Fig. 14; one of them is illustrated in Fig. 16.

4	2	3	1
		2	4
		4	2
2	4	1	3

Fig. 16

The situation would be the same if those twelve givens were provided for the equivalent 4×4 Sudoku puzzle.

If a Latin square puzzle has an embedded Latin square, it will be impossible to solve unless at least one of the symbols in the embedded square is a given. This is also sometimes true for a Sudoku puzzle, as in Fig. 16.

So the largest number of givens that can be provided without a 4×4 Latin square puzzle having a unique solution is at least twelve. In fact it is possible to prove that it is exactly twelve.

In the case of the 9×9 Sudoku puzzle shown in Fig. 17, there are 77 givens and just four cells remain to be filled. However those four form an embedded Latin square and the puzzle does not have a unique solution.

So the largest number of givens that can fail to give a unique solution for a 9×9 Sudoku puzzle is at least 77, and it can be proved that it is exactly 77.

7	1	8	2	5	9	3	4	6
4			3	8	7	1	9	5
5	3	9	4	1	6	8	2	7
1	4	7	6	9	3	2	5	8
8			5	4	1	9	7	3
3	9	5	8	7	2	6	1	4
2	5	4	9	3	8	7	6	1
9	7	3	1	6	5	4	8	2
6	8	1	7	2	4	5	3	9

Fig. 17

BLANK PAGE