Taking Sudoku Seriously

Laura Taalman, James Madison University

You've seen them played in coffee shops, on planes, and maybe even in the back of the room during class. These days it seems that everyone is filling in gerechte designs of order 9 with square subregions. But is it math?

The rules of the game

For those of you who spent the last two years in seclusion, here are the basics of the game known as Sudoku. A ***Sudoku board*** is a 9×9 matrix of integers with the property that in every row, in every column, and in every one of nine 3×3 "blocks" (see Figure 1), each of the integers from 1 to 9 appears exactly once. A ***Sudoku puzzle*** is a partially filled-in Sudoku board that can be completed in exactly one way. Notice that we only call something a "puzzle" if it has a unique solution (all well-made Sudoku puzzles have this property), and also that any given Sudoku board has many possible puzzles. Figure 2 shows two different Sudoku puzzles that have the same solution. The "game", of course, is to extend a given Sudoku puzzle to its unique Sudoku board.

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

Figure 1: *An example of a Sudoku board.*

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

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

Figure 2: *Two Sudoku puzzles with the same solution, namely the board in Figure 1.*

In both puzzles shown in Figure 2, the set of cells containing the clues has 180 degree rotational symmetry (just as the black squares in a crossword puzzle traditionally have). By convention, most quality Sudoku puzzles that are published today have this property – although mathematically, we will not assume that puzzles are symmetric unless we explicitly say so. Notice that Sudoku boards are just Latin squares with an additional "block" condition.

Counting boards

There are many questions one could ask about Sudoku, but perhaps the most basic one is:

Question 1. *How many Sudoku boards are there?*

In 2005, Felgenhauer and Jarvis used a computer algorithm that counted certain equivalence classes of Sudoku boards to conclude that there are $6,670,903,752,021,072,936,960$ different Sudoku boards. A proof that does not use computers is not yet known. (Here is your chance to make your mark!) To get a feel for how Felgenhauer and Jarvis counted the 9×9 Sudoku boards, let's run through a similar argument for 4×4 boards.

A ***Shidoku board*** is a 4×4 matrix of integers with the property that in every row, every column, and every 2×2 "block", each of the integers from 1 to 4 appears exactly once. A ***Shidoku puzzle*** is a partially filled-in Shidoku board with a unique solution; see Figure 3.

			2
	1		
		4	
3			

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

Figure 3: *A Shidoku puzzle and its solution board.*

Theorem 1. *There are 288 different Shidoku boards.*

There are many ways that we could prove this fact, but we will follow a procedure that is similar to what Felgenhauer and Jarvis did in the 9×9 case.

Proof. First, some terminology: Any Shidoku board whose first row, first column, and first block are "ordered" as shown in Figure 4 will be called an ***ordered Shidoku board***. We will argue that every Shidoku board is in some sense equivalent to an ordered Shidoku board, and then count the possible ordered boards.

Given any Shidoku board, we can permute the choice of symbols 1, 2, 3, and 4 so that the first 2×2 block in the board is ordered as in Figure 4.

1	2	3	4
3	4		
2			
4			

Figure 4: *An ordering on the entries of the first row, column, and block of a Shidoku board.*

Note there are 4! such permutations. Now by swapping the last two columns (if necessary) we can finish ordering the first row, and by swapping the last two rows (if necessary) we can finish ordering the first column, so as to obtain an ordered Shidoku board. Note that the row and column swaps represent $2 \cdot 2 = 4$ choices. Therefore, every Shidoku board can be turned into an ordered Shidoku board by permutations and symmetries, and every ordered Shidoku board represents $4! \cdot 4 = 96$ different Shidoku boards.

To count the number of different Shidoku boards we need only count the number of *ordered* Shidoku boards, and then multiply by 96. It is a simple matter to count the ordered Shidoku boards: Simply start with the ordered entries in Figure 4 and then "play" Shidoku, keeping track of any choices you make along the way. This leads to the three boards shown in Figure 5. Therefore there are $3 \cdot 96 = 288$ different Shidoku boards. ■

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

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

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

Figure 5: *The three ordered Shidoku boards.*

One way that the 9×9 argument for answering Question 1 is harder than the proof above is that we must consider more than one possible "ordering" of the first row, column, and block. Call an ***ordered Sudoku board*** one whose first block is as in Figure 6, and whose first row and column both end with two 3-digit sequences, each with increasing digits, in lexographic order. For example, in the ordering shown in Figure 6, the first row ends with 458 and 679, and the first column ends with 269 and 358. To count all the possible Sudoku boards, one would have to first count all of these types of orderings, and then count the number of boards that complete each of these orderings. This is in fact the sort of thing that Felgenhauer and Jarvis did to reduce the search space of their counting algorithms.

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

Figure 6: *One of many orderings on the first row, column, and block of a Sudoku board.*

It is worth pointing out that in our proof that there are 288 Shidoku puzzles, we did not use all the possible Shidoku "symmetries", but we used exactly what we needed to produce classes of Shidoku puzzles that were the same size, which was particularly convenient for our counting argument. If we take into account all the different ways that one Shidoku puzzle can be transformed into another, then there are actually only *two* classes of Shidoku puzzles, not three. Here is an exercise for you: Show that by taking the transpose and then permuting the number labels $1, 2, 3, 4$ we can change the third ordered Shidoku board in Figure 5 into the second. This means that there are only two "essentially different" Shidoku boards. In the 9×9 case, Jarvis and Russell have shown (again, a computer was involved at some point) that when all Sudoku symmetries are taken into account, there are only $5,472,730,538$ "essentially different" Sudoku boards.

Minimal Sudoku puzzles

Perhaps the second most basic question to ask about Sudoku puzzles is the following:

Question 2. *What is the minimum number of clues that a Sudoku puzzle can have?*

Keep in mind that we only consider something a "puzzle" if it has a unique solution, so Question 2 is asking how few initial clues can completely determine an entire Sudoku board. The answer to this question is not known.

Gordon Royle, who is the author of the book *Algebraic Graph Theory* in the Springer series "Graduate Texts in Mathematics", maintains a collection of over $36,000$ different Sudoku puzzles with 17 clues. Nobody has yet found any examples of Sudoku puzzles with 16 clues. This strongly suggests that the answer to our question is 17 – but of course, a proof this does not make. (Here is another open problem you could work on!) If we require that the Sudoku puzzle clues have 180 rotational symmetry, then the least known number of clues that determines a unique solution increases to 18. Figure 7 is an example of such a puzzle. It is not known whether or not 18 is in fact the minimum.

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

Figure 7: *A symmetric Sudoku puzzle with only 18 clues.*

In the 4×4 case this question is much more approachable, especially considering that there are only two types of essentially different Shidoku boards (represented by the first two boards in Figure 5, which we'll call ***type-1*** and ***type-2*** boards).

Theorem 2. *The minimum number of clues that a Shidoku puzzle can have is 4.*

Proof. The Shidoku puzzle that we saw in Figure 3 has only 4 clues, so it suffices to prove that no Shidoku puzzle can have less than 4 clues.

Consider the board in Figure 8 (note this is the type-1 Shidoku board). We will call a collection of cells on a given Shidoku board an ***unavoidable set*** if *every* puzzle for that board must have at least one clue in that set. For example, the four yellow cells in Figure 8 make up an unavoidable set: Even if a puzzle for this board included all 12 non-yellow cells on the board as clues, there would still not be enough information to determine what numbers should be placed in the yellow cells. Therefore *every* puzzle whose solution is the type-1 Shidoku board must contain at least one clue in the yellow set.

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

Figure 8: *Four disjoint unavoidable sets on a type-1 Shidoku board.*

Since the type-1 board in Figure 8 has four disjoint unavoidable sets, any puzzle whose solution is that board must contain at least four clues. A similar argument for type-2 boards applies. Therefore, the minimum number of clues that any Shidoku puzzle can have is 4. ■

Extending the basic questions

A natural followup to Question 1 is to ask:

Question 3. *How many Sudoku puzzles are there?*

Remember that each Sudoku board correesponds to many possible puzzles. For example, the board itself is a (very stupid) puzzle. Also, taking as clues any 80 of the 81 entries of a board produces an (admittedly easy) puzzle. In fact, every one of these 80-clue puzzles is well-defined, i.e. has a unique solution. Similarly, all 79-clue puzzles and all 78-clue puzzles are well-defined. Not so for 77 clues, however. Here is an exercise for you: Find 77 clues that do *not* determine a unique Sudoku board.

Question 3 might be more meaningful if we considered counting only the ***irreducible*** Sudoku puzzles, that is, the puzzles for which every clue is necessary, in the sense that removing any one clue would result in a set of clues with non-unique solution. In the 4×4 Shidoku case, this is an accessible problem. For example, it is not hard to show that the total number of 6-clue irreducible puzzles whose solution is any of the three Shidoku boards in Figure 5 is 16. Each of these 16 puzzles represents 96 puzzles (see the proof of Theorem 1), and thus there are a total of $16 \cdot 96 = 1536$ irreducible 6-clue Shidoku puzzles.

It turns out that 6 clues is the maximum number of clues that an irreducible Shidoku puzzle can have. This can be shown easily by computer enumeration, and it is not too hard to show directly that this maximum can be at most 8, but it would of course be preferable to have a direct argument that 6 is the maximum. We can ask the same question in the context of 9×9 puzzles:

Question 4. *What is the maximum number of independent clues that a Sudoku puzzle can have?*

The most independent clues that I have ever seen in a Sudoku puzzle is 33. Is 33 the maximum? Or is the maximum number much larger than that? Here again you have an opportunity to make your mark in this new field of research.

We can also extend our questions by extending our notion of Sudoku. For example, we could add additional constraints to our Sudoku boards and then ask the same questions of that smaller class of boards. For example, consider the subset of Sudoku boards with no repeated entries in any of the marked diagonals shown in Figure 9. What is the

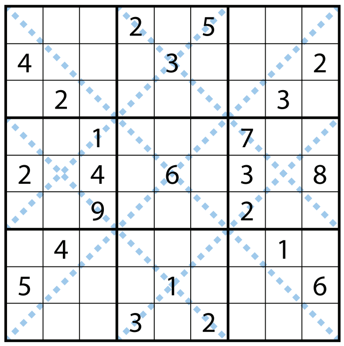

Figure 9: *Snowflake: Each row, column, and block must contain 1–9 exactly once, with no repeated entries in any marked diagonal.*

minimum number of clues that such a "Snowflake" puzzle can have? What is the maximum number of independent clues? How many Snowflake boards exist? How many Snowflake puzzles exist? Everything we asked about regular Sudoku boards can now be asked about Snowflake boards. And if we make up another special case or type of Sudoku boards, we can ask the same questions again! Just for fun, Figure 10 shows another Sudoku variation for you to think about. (Of course I give you a puzzle and not a board, so you can also have the fun of *playing* the puzzle if you like.)

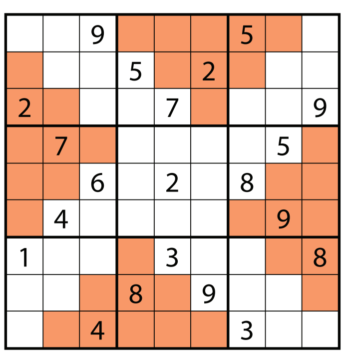

Figure 10: *Pyramids: Each row, column, block, and pyramid-region must contain 1–9 exactly once.*

More Open Questions

There are of course many other questions we can ask about Sudoku puzzles. For example, you may have heard of Sudoku solving techniques with funny names like "Swordfish" and "X-Wings". A natural question to ask is this: Given a certain set of solving techniques, what types of Sudoku puzzles can be solved by those techniques? A possibly easier question than this would be to ask how many Sudoku puzzles can be solved using only "scanning", without any fancy logical arguments or solving techniques at all.

Here is another source of questions: Since Sudoku puzzles are at their hearts 9×9 ***Latin squares*** (meaning that each row and column has no repeated entries) with an additional "block" condition, we can possibly extend some results about Latin squares to results about Sudoku puzzles. For example, you can make an alternate argument for the fact that there are 288 Shidoku boards by using the known fact that there are 576 Latin squares of order 4, and then showing that half of them fail the extra block condition. In the 9×9 case it is *not* true that the number of Sudoku boards is half of the number of Latin squares of order 9, but nonetheless it might prove a fruitful avenue of study to consider the set of all 9×9 Latin squares and then count the ones that fail to be Sudoku boards. Anything one can ask about Latin squares can also be asked about Sudoku puzzles; for example, it can be shown that the maximum number of "mutually orthogonal" (I'll let you look that one up on your own) Sudoku boards is 6.

In a similar vein, we can think of Sudoku puzzles as ***gerechte designs***, which is just a fancy way of saying that they are Latin squares with an extra set of regions (namely, the 3×3 blocks) each containing 1–9 exactly once. This opens up the possibility of generalizing Sudoku by considering regions that are different than the usual 3×3 blocks. For example, the "Jigsaw" puzzle in Figure 11 is a generalized Sudoku puzzle that is a different type of gerechte design.

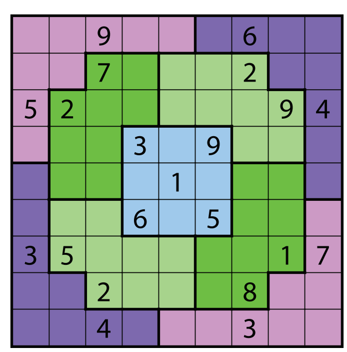

Figure 11: *Jigsaw: Each row, column, and jigsaw-region must contain 1–9 exactly once.*

Now things that are known about gerechte designs can be brought to bear on Sudoku questions.

Moreover, we can use the concept of gerechte designs to extend our concept of Sudoku to include even more exotic puzzles. For example, the "Wrap Up" puzzle in Figure 12 is an example of a so-called "multiple" gerechte design (which just means that there are rows, columns, blocks, and yet *another* set of 9 regions to contend with).

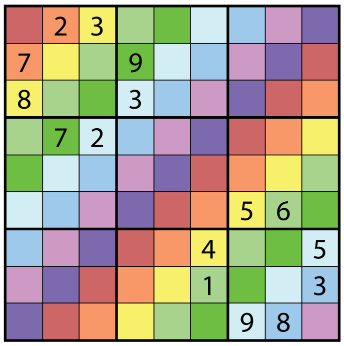

Figure 12: *Wrap Up: Each row, column, block, and color-region must contain 1–9 exactly once. Note that the color-regions wrap around the square in "torus" fashion.*

These generalizations lead to even more questions: Which types of regions can be "extra" regions? What percentage of Sudoku boards satisfy a given extra region condition? What is the maximum number of extra regions that a Sudoku board can have?

Of course, we can go beyond just adding extra regions; we could also add some *adding*! For example, the puzzle in Figure 13 has regions that add to particular amounts (and you must figure out the amounts). It is actually relatively difficult to come up with a partition of a Sudoku board into connected regions such that each region adds to one of five different possible sums. Does every Sudoku board have such a partition? Multiple such partitions? What are the restrictions on the values of the sums?

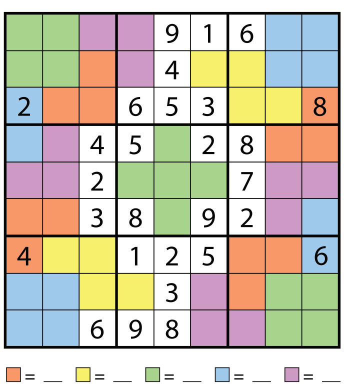

Figure 13: *Mystery Sums: Each row, column, and block must contain 1–9 exactly once, with no repeated entries in any connected color-region. Color-regions of the same color add to the same sum.*

If you don't like sums, how about *orders*? For example, consider the puzzle in Figure 14. Of course, this leads to even more questions: How many Sudoku boards can be totally filled up with "worms" all of size 3 or greater? What is the maximum number of ways that a Sudoku board can be "wormed"?

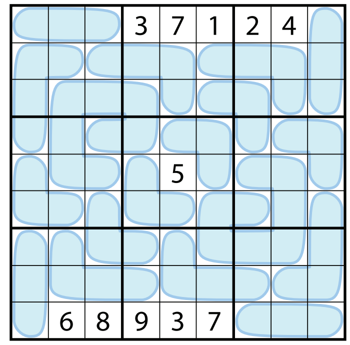

Figure 14: *Worms: Each row, column, and block must contain 1–9 exactly once, and the entries in each worm either increase or decrease monotonically (although not necessary sequentially; e.g. 2, 3, 6, 8 is allowable).*

Every Sudoku variation you can think of will come with its own set of interesting open questions. Mathematicians young and old have studied Sudoku puzzles and their variations from diverse mathematical perspectives that include graph coloring, chromatic polynomials, Gröbner bases, the "rook problem", magic squares, permutation group theory, and Diophantine equations. There is math to be had here. So get working!

For Further Reading

All of the puzzles in this article were created by Laura Taalman and Philip Riley (otherwise known as *Brainfreeze Puzzles*). For more puzzles, see:

▷ `www.brainfreezepuzzles.com`

▷ *Color Sudoku*, Riley and Taalman, Sterling Publishing 2007.

Here are some current mathematical research articles that relate to Sudoku puzzles:

▷ Felgenhauer and Jarvis, *Enumerating possible Sudoku grids.*

▷ Jarvis and Russell, *There are 5472730538 essen-*

tially different Sudoku grids... and the Sudoku symmetry group.

▷ Bailey, Cameron, and Connelly, *Sudoku, gerechte designs, resolutions, affine space, spreads, reguli, and Hamming codes.*

▷ Gupta, *Some results on Su Doku.*

▷ Eppstein, *Nonrepetitive Paths and Cycles in Graphs with Application to Sudoku.*

▷ Van Hoeve, *The Alldifferent Constraint: a Systematic Overview.*

▷ Seta and Yato, *Complexity and Completeness of Finding Another Solution and Its Application to Puzzles.*