MATHEMATICS OF SUDOKU

Vira Babenko
University of Utah
April 16th 2014
(based on “Taking Sudoku Seriously” by Laura Taalman)
Mathematics, Magic, & Mystery

Math Awareness Month ★ April 2014

Go to mathaware.org to find
30 days of videos and articles on mathematical magic tricks, mysteries, puzzles, illusions, and more!
The rules of the game:

A Sudoku board is a 9x9 matrix of integers with the property that in every row, in every column, and in every one of nine 3x3 "blocks", each of the integers from 1 to 9 appears exactly once.

An example of a Sudoku board:

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

Photo from Wikipedia
A Sudoku puzzle is a partially filled-in Sudoku board that can be completed in exactly one way.

Notice:
1. we only call something a “puzzle” if it has a unique solution
2. that any given Sudoku board has many possible puzzles.
Two different Sudoku puzzles that have the same solution

The goal is to extend a given Sudoku puzzle to its unique Sudoku board.
History

- In the XVIII century Leonhard Euler invented the game «Carré latin» («Latin square»).

In combinatorics and in experimental design, a Latin square is an $n \times n$ array filled with $n$ different symbols, each occurring exactly once in each row and exactly once in each column.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
</tbody>
</table>
History

- Number puzzles appeared in newspapers in the late 19th century (November 19, 1892) (French puzzle setters began experimenting with removing numbers from magic squares)
- On July 6, 1895, a Paris-based daily, refined the puzzle so that it was almost a modern Sudoku. It simplified the 9×9 magic square puzzle so that each row, column and broken diagonals contained only the numbers 1–9, but did not mark the sub-squares.

From La France newspaper, July 6, 1895
History

- These weekly puzzles were a feature of French newspapers for about a decade but disappeared about the time of World War I.

- It came back in the 70s of the last century in North America were invented special numerical crossword puzzles. Thus, the U.S. Sudoku first appeared in 1979 in the journal «Dell Puzzle Magazine». Then it was called «Number Place».

- Sudoku gained real popularity in the 1980-1990's, when the Japanese magazine «Nikoli» began to regularly publish on their pages this puzzle (since 1986).

- Today Sudoku - a mandatory component of many newspapers.
There are many questions one could ask about Sudoku.

- **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.*
To get a feel for how Felgenhauer and Jarvis counted the 9x9 Sudoku boards, let’s run through a similar argument for 4x4 boards.

- A **Shidoku** board is a 4x4 matrix of integers with the property that in every row, every column, and every 2x2 “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.
A Shidoku puzzle and its solution board.
Theorem 1.
There are 288 different Shidoku boards.

Proof.
Terminology: Any Shidoku board whose first row, first column, and first block are “ordered” as shown in Figure will be called an ordered Shidoku board.

```
1 2 3 4
3 4   
2   
4
```
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 2x2 block in the board is ordered as in Figure.

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.

- Every Shidoku board can be turned into an ordered Shidoku board by permutations and symmetries.

- 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 simple to count the ordered Shidoku boards:
- start with the ordered entries in Figure and then “play” Shidoku, keeping track of any choices you make along the way.
This leads to the three boards shown here:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Therefore there are $3^9 = 288$ different Shidoku boards.
One reason why the 9x9 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, and whose first row and column both end with two 3-digit sequences, each with increasing digits, in lexicographic order.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>8</th>
<th>6</th>
<th>7</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 what Felgenhauer and Jarvis did to reduce the search space of their counting algorithms.
Sometimes question about “essentially different” combinations is more important. In our proof that there are 288 Shidoku puzzles, we did not use all the possible Shidoku “symmetries”.

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.

In the 9x9 case, Jarvis and Russell have shown 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.
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. But for a long time it was not known whether there is a uniquely solvable sudoku with 16 tips.

Distributed computing project Sudoku@vtaiwan on BOINC platform was looking for such Sudoku and in January 2012, there was evidence that uniquely solvable sudoku with 16 tips does not exist.
In the 4x4 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, 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 before has only 4 clues, so it suffices to prove that no Shidoku puzzle can have less than 4 clues.

```
  1 2 3
  4
```
Consider the board in next Figure (note: this is the type-1 Shidoku board).

![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 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.
Since the type-1 board 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 follow-up to Question 1 is to ask:

- **Question 3.**
  How many Sudoku puzzles are there?
Remember that each Sudoku board corresponds 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 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.
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 nonunique solution.

In the 4x4 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 is 16. Each of these 16 puzzles represents 96 puzzles, 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 9x9 puzzles:

- **Question 4.**
  What is the maximum number of independent clues that a Sudoku puzzle can have?
The most independent clues that I saw mentioned was 33.

Is 33 the maximum? Or is the maximum number much larger than that?

It’s not known yet.
We can also extend our questions by extending our notion of Sudoku.

For example, we could add further constraints to our Sudoku boards and then ask the same questions of that smaller class of boards.
We can consider the subset of Sudoku boards with no repeated entries in any of the marked diagonals shown in Figure.
- What is the 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!
DUIDOKU

**How to Play:**
You make the first move. Make a legal Shidoku move: a number from 1 to 4 so that you don't put the same number in a single row, column, or cluster. Certain moves will block other moves, so try to be the last one to make a legal move!
GAME TIME

We need two teams
References:

- “Taking Sudoku Seriously” by Laura Taalman (www.maa.org/mathhorizons)
- Most of the puzzles in this talk were created by Laura Taalman and Philip Riley (otherwise known as Brainfreeze Puzzles)
- www.brainfreezepuzzles.com
- Photos are taking from wikipedia and hd-pictures.ru
THANK YOU

Pies time! ;)
