Discovery could lead to more difficult Sudoku puzzles

February 13 2010, by Lisa Zyga

A standard 9x9 Sudoku matrix. Image credit: Héctor Rodríguez.

(PhysOrg.com) -- A new analysis of number randomness in Sudoku matrices could lead to the development of more difficult and multi-dimensional Sudoku puzzles. In a recent study, mathematicians have found that the way that numbers are arranged in Sudoku puzzles is even more random than the number arrangements in randomly-generated matrices. The counter-intuitive discovery may enable researchers to develop algorithms that generate Sudoku matrices with fewer clues, making them more difficult to solve.

Mathematicians Paul Newton and Stephen DeSalvo of the University of Southern California in Los Angeles have published the results of their study in a recent issue of the Proceedings of the Royal Society A.

"I think it will help develop multi-dimensional Sudoku puzzles, and answer questions about how to give the initial [clues] in order to create a hard, but still solvable Sudoku puzzle," Newton said in an article at ABC Science.

Sudoku is a number puzzle consisting of a 9x9 grid, whose 81 boxes are filled in with the numbers 1 through 9 in a way that meets certain criteria. Each number can only appear once per row and once per column, as well as only once in each of the nine 3x3 sub-grids that make up the matrix.

In 2006, researchers (Felgenhauer and Jarvis) found that there are about 6.67×10^{21} different Sudoku matrices that satisfy these three criteria. In contrast, the total number of different 9x9 randomly generated matrices is much greater: 9^{81}. The ratio of these two numbers, or the probability of randomly generating a Sudoku matrix by randomly selecting each number in each box independently, is very small: about 3×10^{-56}. This small probability results from the constraints put on Sudoku matrices.

In their study, Newton and DeSalvo wanted to find out how exactly random a Sudoku matrix is, given these constraints. To answer this question, they generated a representative sample of about 10,000 matrices and compared them to randomly generated matrices. They were surprised to find that Sudoku matrices are actually more random than randomly-generated matrices. This result is counterintuitive since you would expect that, the more constraints on a matrix, the less random it will be.

Instead, the rules of Sudoku seem to “weed out” matrices with patterns. For example, as Newton explained, a randomly generated matrix could potentially consist of all one number, alternating numbers, or some other pattern not allowed in Sudoku. The imposed high level of number distribution in Sudoku gives it a higher level of entropy, making it more

random than random matrices.

Newton and DeSalvo predict that this greater understanding of Sudoku could lead to better Sudoku-generating algorithms that create more difficult puzzles. Currently, Sudoku puzzles require at least 17 numbers to be given in their correct boxes in order for the puzzle solver to find a unique solution. The new study could decrease that number, making it more difficult to solve the puzzles. Future algorithms might also develop more complex 3D Sudoku cubes.

"I think it will give people a lot of insight into how to produce better algorithms for constructing Sudoku matrices and it will enable ultimately the very fast learning algorithms that solve Sudoku matrices," Newton said.

Australian mathematician Marcel Jackson of Latrobe University in Melbourne, who was not affiliated with the study, added that understanding Sudoku matrices better might also be useful in coding information to minimize the effect of errors in transmission.

More information:

-- Paul K. Newton and Stephen A. DeSalvo. “The Shannon entropy of Sudoku matrices.” *Proceedings of the Royal Society A.* doi: 10.1098/rspa.2009.0522.

-- Via: ABC Science