Project title	Finding the Minimum Number of Sudoku Clues Through Information Theory
Owner	Félix Balado Pumarino
Supervisor	Félix Balado Pumarino

Description

A Sudoku grid is a 9x9 square grid such that every row, column, and 3x3 square subgrid contain a permutation of the numbers 1 to 9.

In a Sudoku puzzle, a few squares out of the 81 forming a chosen Sudoku grid are given as clues to the player, who has to guess the content of the remaining ones. In 2013, McGuire et al found through a mathematically-informed empirical search that 17 is the minimum number of clues for a Sudoku puzzle to be uniquely solvable. However there is no formula yet that tells us why this should be so.

Core

Firstly, the student will become acquainted with Sudoku and its most relevant literature. A Sudoku grid can be seen as a 9-ary code where only certain codewords are allowed. Thus a Sudoku puzzle can be seen as putting a Sudoku grid (codeword) through a 9-ary erasure communications channel. The task of the player is to _decode_ the original codeword (i.e. the original grid) from the noisy output of the channel. Since it is known how

many Sudoku grids exist (i.e., the number of Sudoku grids has been enumerated), the rate of the code is also known.

For this reason, information theory should be able to allow us to investigate the minimum number of clues in an analytical way. With the help of the advisor, the student will study the problem from the point of view of standard information theory, where the uses of the channel go asymptotically to infinity.

Advanced

To refine the results above it will be necessary to resort to zero-error capacity in the erasure channel, over a finite number of channel uses (i.e., 81 uses).

Dataset

Catalogue of known Sudoku grids

Requirements

Strong interest in information theory, and not afraid of a challenge.