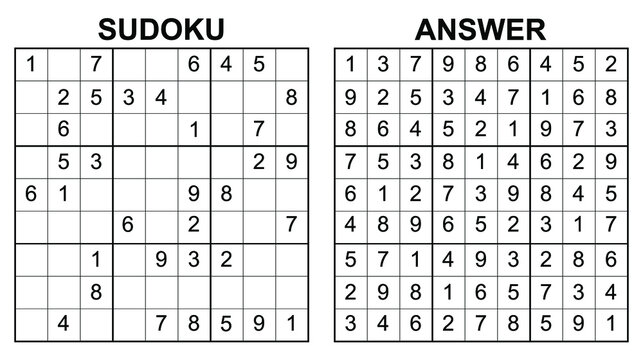
Sudoku Solver

Martin Rode, Marvin Üürike

Introduction

Sudoku is a popular number puzzle game that originated in Japan. The objective of the game is to fill a 9x9 grid with digits from 1 to 9 in such a way that each column, each row, and each of the nine 3x3 subgrids (also known as "boxes") contains all the digits from 1 to 9 without repetition. The puzzle starts with some cells pre-filled with numbers, and the player must fill in the remaining cells following these rules.

In this project, we aim to develop an AI Sudoku solver using the Depth-First Search (DFS) algorithm combined with various logical constraints. The goal is to enhance the solver's efficiency and accuracy in providing solutions to any valid Sudoku puzzle and to compare the effectiveness of different logical constraints in speeding up the solving process.



Strategies

Basic:

In this strategy, the solver inserts numbers from 1 to 9 into cells row by row, trying each possibility until a solution is found or backtracking occurs if a conflict is encountered. It's a straightforward approach where the solver systematically explores possible solutions without any optimization. This method is simple but may lead to inefficient exploration of the search space.

Original Legal Values:

Before solving the puzzle, the solver calculates the legal values for each cell based on the initial puzzle configuration and stores them in a cache. During the solving process, the solver only considers these pre-computed legal values for each cell. This reduces the search space compared to the basic approach but does not adapt to changes made during solving.

Updating Legal Values with Cell Prioritization:

This strategy involves dynamically updating the cache of legal values for each cell after every entry. Additionally, cells with the fewest legal values are prioritized for exploration. By continuously updating the cache of legal values and strategically prioritizing cells, this method can significantly improve solving efficiency, especially for hard sudokus.

Hidden Singles with Cell prioritization:

Cells with the fewest legal values are prioritized for exploration. After inserting a value into a cell, the solver updates the cache of legal values for other cells affected by this entry. Additionally, it checks each block to see if the inserted value remains the only legal option for any cell within the block. If so, the excessive legal options for that cell are removed from the cache, accelerating subsequent iterations. Hidden singles represent efficient opportunities for progression in solving the puzzle, providing definitive solutions without the need for further exploration or backtracking. This mechanism can be thought of as a deep update of the cache, optimizing the solving process.

Priority Value search:

This strategy prioritizes trying the legal values of a cell based on their occurrence frequency in the legal values of other cells within the same row, column, or block. Starting with the least occurring values increases the likelihood of making progress quickly. By prioritizing values with lower occurrence frequencies, the solver strategically explores fewer common possibilities first, potentially leading to faster solutions.

Results

To test the implementation of strategies 300 ‘hard’ sudokus were downloaded online and solved. For each algorithm and each sudoku the number of calls and total time were computed.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Name** | **Avg calls** | **Min calls** | **Max calls** | **Solved** | **Failed** | **Call limit** |
| Basic | 10669 | 54 | 99660 | 231 | 69 | 100000 |
| Original Legal Values | 43705 | 52 | 490894 | 250 | 50 | 500000 |
| Updating Legal Values  with Cell Prioritization | 3179 | 48 | 156323 | 300 | 0 | 500000 |
| Hidden Singles  with Cell prioritization | 1704 | 48 | 56694 | 300 | 0 | 500000 |
| Priority Value search | 747 | 48 | 32040 | 300 | 0 | 500000 |

It should be noted that in the current scenario, the number of function calls was restricted to 100,000 for the basic strategy and 500,000 for the others. It is observed that with the Basic and Original Legal Values strategies, multiple Sudoku puzzles remained unsolved within these limits. Consequently, this limitation could skew the average number of calls downward compared to scenarios where no call limit was imposed. However, the introduction of smarter strategies resulted in a significant reduction in the number of calls required to solve 100% of the Sudoku puzzles. Particularly noteworthy is the unexpected success of the Priority Value Search strategy—a technique not typically intuitive for human Sudoku solving.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Name** | **Avg time** | **Min time** | **Max time** | **Solved** | **Failed** | **Call limit** |
| Basic | 2.26690 | 0.00197 | 35.74368 | 231 | 69 | 100000 |
| Original Legal Values | 0.15200 | 0.00041 | 1.82349 | 250 | 50 | 500000 |
| Updating Legal Values  with Cell Prioritization | 0.02300 | 0.00045 | 1.10812 | 300 | 0 | 500000 |
| Hidden Singles  with Cell prioritization | 0.02805 | 0.00083 | 1.78878 | 300 | 0 | 500000 |
| Priority Value search | 0.02698 | 0.0016 | 1.1052 | 300 | 0 | 500000 |

A significant improvement in solving time is evident when comparing the simple Original Legal Values strategy to the Basic strategy. This is expected, as the former involves far fewer checks to verify if value criteria are met. Another tenfold improvement is observed with the Updating the Cache of Legal Values strategy. However, the strategies that required the fewest number of function calls, such as Priority Search and Hidden Singles, took slightly longer than the straightforward cache update approach. This increased time could be attributed to the additional processing required to identify hidden singles or count the frequencies of possibilities, which slows down the algorithm.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Name** | **Avg Calls** | **Avg time** | **Solved** | **Failed** | **Call limit** |
| Basic | 10669 | 2.26690 | 231 | 69 | 100000 |
| Original Legal Values | 43705 | 0.15200 | 250 | 50 | 500000 |
| Updating Legal Values  with Cell Prioritization | 3179 | 0.02300 | 300 | 0 | 500000 |
| Hidden Singles  with Cell prioritization | 2276 | 0.03764 | 300 | 0 | 500000 |
| Priority Value search | 2166 | 0.04585 | 300 | 0 | 500000 |

In summary, our exploration of various Sudoku solving strategies has shown significant improvements in solving efficiency. Simple approaches like Original Legal Values reduced solving time substantially, while more advanced techniques like Updating the Cache of Legal Values achieved even greater efficiency gains. Although some strategies incurred slightly longer solving times due to additional processing, their effectiveness in reducing function calls underscores the importance of exploring diverse solving approaches.