

# Sudoku : A constraint satisfaction problem

Aditya Joshi, Barada p Acharya, Bhushan Pagariya, Vinay Bagade  
University of California, Irvine

## Abstract

In this report, we describe Sudoku as a constraint satisfaction problem. CSP is viewed as consisting of four major components namely the variable heuristic, value heuristic, domain reduction and early detection, this report explores the efficiency of various methods used in each of the above components and displays statistical results for the same.

## A. Introduction

Sudoku is a puzzle where the player is given an 9X9 grid and some of the cells in the grid are assigned a value the player has to fill up the grid in such a way that in a row, column and a box of 3X3 numbers appear with no repetition. Each cell can take values in  $\{1,2,...9\}$ . The literal meaning of Sudoku in Japanese is ‘single number’ and the puzzle is among the most commonly featured puzzles in newspapers and magazines.

## B. Problem Description

Variables:  $X_{ij}$  is a variable corresponding to the cell  $M_{ij}$  of the matrix, for  $1 \leq i, j \leq 9$

Domains: Domains of each  $X_{ij}$  corresponds to  $\{1,2,3,...9\}$

Constraints: alldiff constraints between elements of the same row, column and a square matrix of 3X3.

Note: The square matrix of 3X3 is not any 3X3 subset but one which has its first cell on a row  $k$  where  $k \% 3 = 1$ .

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

Figure1: Example problem and solution

## C. Input

The input is a sequence of 81 integers in the range 0 to 9 given in row major format. 0 corresponds to the empty cell.

**Data Structure:** The input is stored in an adjacency list with every row corresponding to a cell. The adjacency of list will always contain the allowed values corresponding to that cell. If the solution exists and is found every variable consists of one value is the list.

## D. Our Solver

The three important steps to solve a constraint satisfaction problem(CSP) are:

1. Selection of variables
2. Assigning value to the variables
3. Constraint Propagation and backtracking

### 1. Selection of variable

While solving any CSP, one of the critical thing is the order in which variables are selected. We are using Most Restricted Variable (MRV) and Maximum Degree (MD) variable selection heuristics in our project.

#### **Most Restricted Variable (MRV)**

MRV heuristic selects variable having least possible legal values in its domain. This heuristic is also known as ‘fail-first’ heuristic because it picks variable that is most likely to cause failure soon, thereby pruning the search tree. If some variable X has no legal value left, the MRV will select X and failure will be detected immediately.

#### **Maximum Degree (MD)**

If there are two variables having least possible legal values in their domain. So in this scenario we are Maximum Degree heuristic as a tiebreaker. MD heuristic will select a variable having largest number of constraints on other unassigned variables. MD heuristic attempts to reduce branching factor on future choices by selecting variable having maximum branching factor in early iteration.

### 2. Assigning value to the variables

Once the variable is selected, next important thing is to assign a value to that variable. We are using Least Constraining Value (LCS) heuristic and Probabilistic heuristic(PR) to decide order in which values are selected.

#### **Least Constraining Value (LCS)**

LCS heuristic selects the value which rules out the fewest choices from the neighbouring unassigned variable. In general, LCS heuristic is trying to leave the maximum flexibility for subsequent variable assignment.

### **Probabilistic Heuristic (PR)**

Probabilistic heuristic selects a value which is most likely to be a valid assignment. We are using following formula for computing probability of each value and then we are choosing a value having highest probability.

Probability of a value K to be a valid assignment =  $((9 - X)/9)$

where X : the number of occurrences of value K in assigned variables

## **3. Constraint propagation and Backtracking**

### **Naive Backtracking (BT)**

A number (1 to 9) is randomly selected and before assigning it to a variable, we do a constraint satisfaction check, then we check recursively that this assignment will lead to a solution or not. If the assignment won't lead to a solution, then we try next number for the variable. And if none of the number (1 to 9) leads to the solution, we return false. The problem with this approach is that it takes a lot of time. We can improve on this naive approach by doing some sort of preprocessing and constraint propagation. Some of the approaches are discussed below.

### **Forward Checking (FC)**

The Forward checking algorithm is the most elementary form of the constraint propagation. It eliminates in advance, the possibilities that do not match the constraints from the domains of unassigned variables connected to the variable whose values is being assigned. For example, if we assign the digit '1' to cell c1, we eliminate '1' from all variables in the corresponding row, column and box.

### **Arc Consistency (AC)**

The arc consistency is a kind of preprocessing step before the actual backtracking. What we mean by this is that if two variables connected by an arc in constraint graph are arc consistent then each of admissible values of one variable is consistent with some admissible value of the other variable.

Generally, arc consistency is implemented using a queue where all the arcs of a constraint graph are first pushed into the queue and then the algorithm pops an arc from the queue and prunes the domain of the corresponding connected variables till they are consistent with each other and arcs corresponding to all the variables connected to the pruned variable are inserted back into the queue. This process is repeated while the queue is not empty.

What we implemented here in our project is not exactly arc consistency, but a flavour of the same. Here, in the project we recursively prune the domains of the unassigned variables which are connected to a variable whose domain has been reduced to one. The function is recursively called on any variables whose domain was reduced to one because of our previous assignment.

### Early Detection (ED)

Another algorithm we have implemented for domain reduction during constraint propagation is Early detection. Here, after assigning the value to the variable we check for the corresponding row, column and box if the number of unassigned variables is more than the number of unique values they can take if that is the case then we stop immediately.

## E. Observation

A sample collection of sudoku based upon difficulty level (like easy, medium and hard) is collected from the internet. Several combinations of the above-defined algorithm were applied upon problem set and the results have been analyzed.

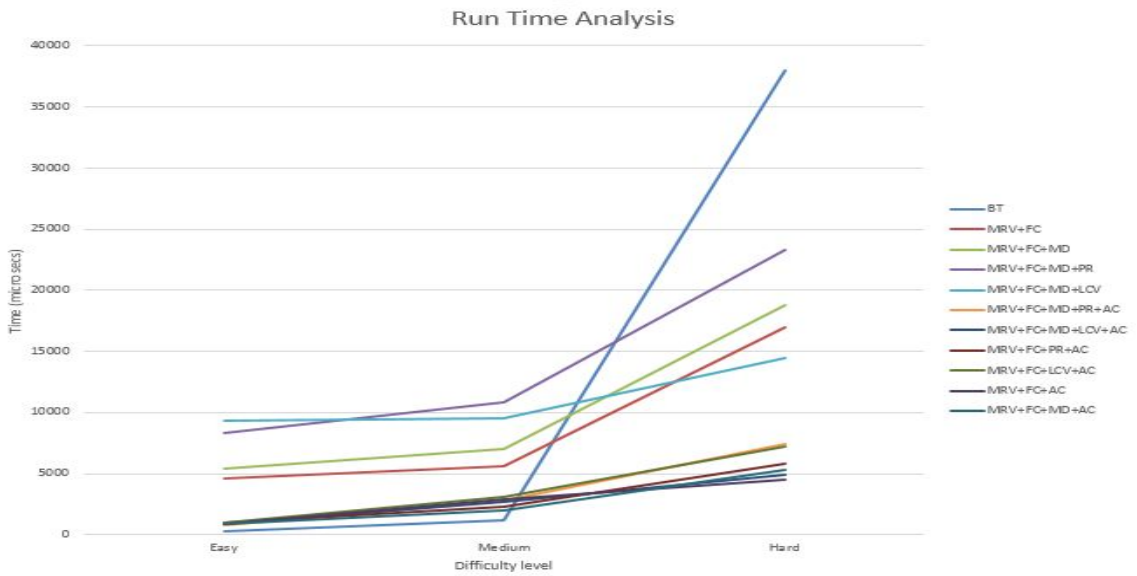


Figure 2: Run Time Analysis plot between algorithms BT, MRV+FC, MRV+FC+MD, MRV+FC+MD+PR, MRV+FC+MD+LCV, MRV+FC+MD+PR+AC, MRV+FC+MD+LCV+AC, MRV+FC+PR+AC, MRV+FC+LCV+AC, MRV+FC+AC, MRV+FC+MD+AC.

From the above graph, we can assert that run-time of an algorithm like brute-force backtracking search is exponentially increasing with difficulty level of the problem, However, CSP algorithms are maintaining a steady increase in run-time values. Among various combinations of CSP, we can see that MRV+FC+AC, MRV+FC+PR+AC, MRV+FC+LCV+AC algorithms are generating better results than other. So we will select these algorithms for our further study in the project.

## F. Experiment Results

The experiments for all problems were run on a laptop with 2.4GHz, Core i5 processor and 4 GB of memory under Window 7.

Puzzle	Brute-force Backtracking		MRV+FC		MRV+FC+AC		MRV+FC+Prob+AC		MRV+FC+LCV+AC	
	Runtime ( $\mu$ s)	No. of backtracks	Runtime ( $\mu$ s)	No. of backtracks	Runtime ( $\mu$ s)	No. of backtracks	Runtime ( $\mu$ s)	No. of backtracks	Runtime ( $\mu$ s)	No. of backtracks
Easy	780	278	2322	20	775	4	1076	7	593	4
Medium1	3973	1924	4322	66	2346	14	2428	15	2595	19
Medium2	1593	789	2351	12	957	2	937	2	944	2
Hard1	1822	864	2664	0	1907	8	1593	8	1350	6
Hard2	3915	1626	1986	5	1509	4	1381	5	1301	4
Evil1	17479	9737	14105	340	3392	19	2375	14	12547	137
Evil2	83359	76285	11988	287	1776	5	5037	54	1316	4
Evil3	1237	1311	1986	0	4223	27	1778	23	1519	6

Table 1: Comparison of Brute-force backtracking, MRV+FC, MRV+FC+AC, MRV+FC+Prob+AC and MRV+FC+LCV+AC algorithms.

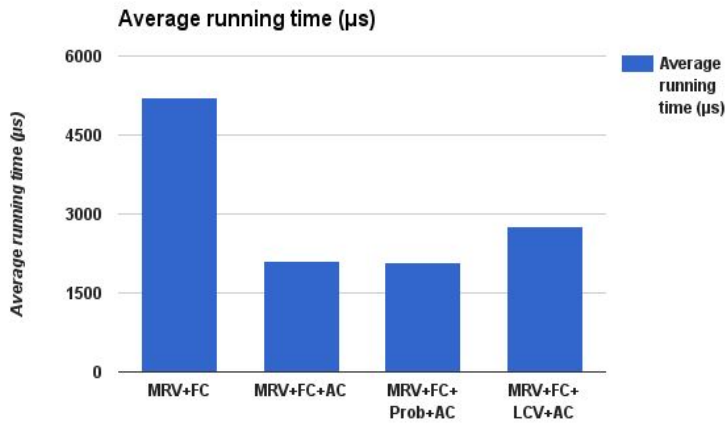


Figure 3(a): Runtime analysis of MRV+FC, MRV+FC+AC, MRV+FC+Prob+AC and MRV+FC+LCV+AC algorithms

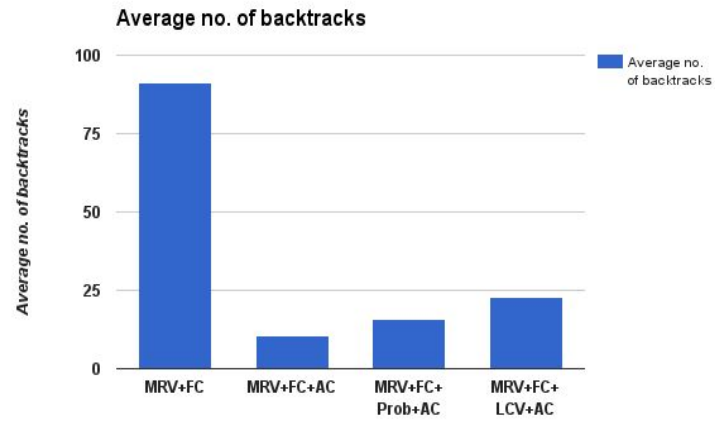


Figure 3b: Backtrack comparison of MRV+FC, MRV+FC+AC, MRV+FC+Prob+AC and MRV+FC+LCV+AC algorithms

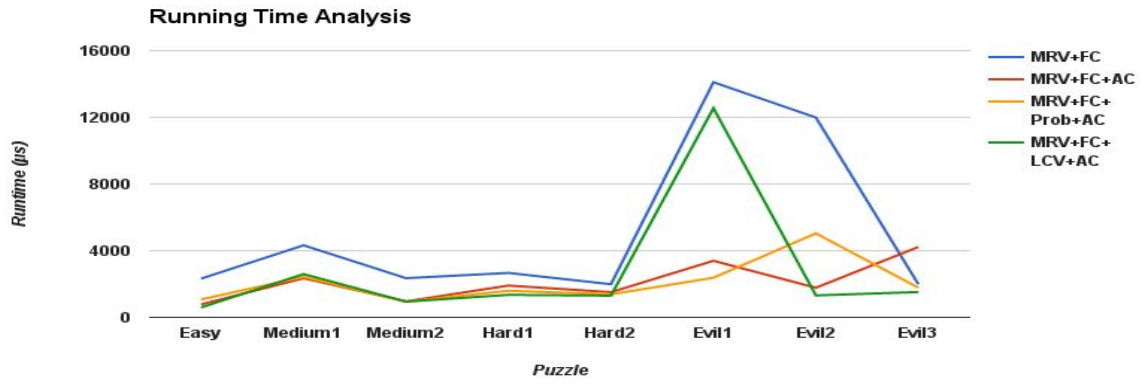


Figure 4: Runtime analysis of MRV+FC, MRV+FC+AC, MRV+FC+Prob+AC and MRV+FC+LCV+AC algorithms for different problems

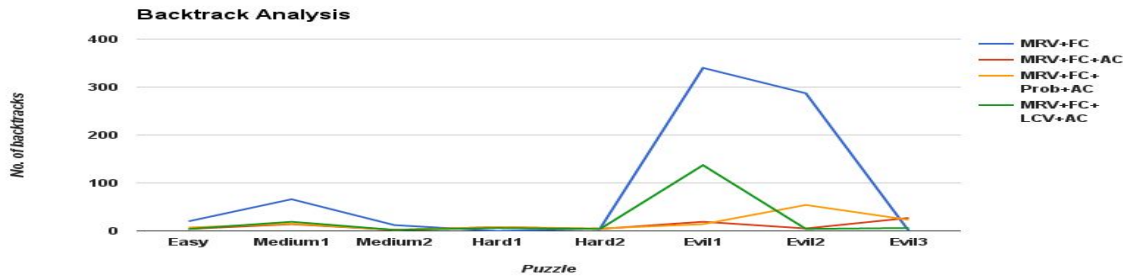


Figure 5: Backtrack analysis of MRV+FC, MRV+FC+AC, MRV+FC+Prob+AC and MRV+FC+LCV+AC algorithms for different problems.

## G. Conclusion

Based on the run-time analysis, we can infer that algorithm using CSP takes less run-time than brute-force Backtracking approach. These gains are seen due to preprocessing of data for domain reduction and early detection failure scenario. We can see that CSP using arc-consistency generally gives better solution CSP without arc-consistency. Our algorithm provides an efficient way to solve the sudoku puzzle by formulating the puzzle as a Constraint Satisfaction Problem and applying concepts such as arc-consistency, forward checking, constraint propagation, early detection with using various value and variable heuristic.

## References

- [1] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach (3rd Edition)*
- [2] Wikipedia page on Sudoku. <https://en.wikipedia.org/wiki/Sudoku>
- [3] Brute-force backtracking algorithm. <http://www.geeksforgeeks.org/backtracking-set-7-sudoku>

## Results for benchmark test-cases

### Puzzle 1

Easy

0640038512000009600058000000200410900360501400102600300  
00004200058000009642100580

MRV+FC	MRV+FC+AC:	MRV+FC+Prob+AC	MRV+FC+LCV+AC
964723851 283415967 175896324 527341698 836957142 419268735 391584276 758632419 642179583	964723851 283415967 175896324 527341698 836957142 419268735 391584276 758632419 642179583	964723851 283415967 175896324 527341698 836957142 419268735 391584276 758632419 642179583	964723851 283415967 175896324 527341698 836957142 419268735 391584276 758632419 642179583
**Time :- 2322 **Backtracks :- 20	**Time :- 775 **Backtrack:- 4	**Time :- 1076 **Backtracks :- 7	**Time :- 593 **Backtracks :- 4

### Puzzle 2

Medium1

7205000060951600000007009042109000030090004006000040913  
02006000000053620900001045

MRV+FC	MRV+FC+AC:	MRV+FC+Prob+AC	MRV+FC+LCV+AC
723549816 495168237 861732954 214985763 539617482 678324591 352496178 147853629 986271345	723549816 495168237 861732954 214985763 539617482 678324591 352496178 147853629 986271345	723549816 495168237 861732954 214985763 539617482 678324591 352496178 147853629 986271345	723549816 495168237 861732954 214985763 539617482 678324591 352496178 147853629 986271345
**Time :- 4322 **Backtracks :- 66	**Time :- 2346 **Backtrack:- 14	**Time :- 2428 **Backtracks :- 15	**Time :- 2595 **Backtracks :- 19

### Puzzle 3

Medium2

0008000420060500000001420060800040350506730103205000704  
00761000000020400190005000

MRV+FC	MRV+FC+AC:	MRV+FC+Prob+AC	MRV+FC+LCV+AC
519836742	519836742	519836742	519836742
246957183	246957183	246957183	246957183
873142596	873142596	873142596	873142596
687214935	687214935	687214935	687214935
954673218	954673218	954673218	954673218
321598674	321598674	321598674	321598674
435761829	435761829	435761829	435761829
768329451	768329451	768329451	768329451
192485367	192485367	192485367	192485367
**Time :- 2351	**Time :- 957	**Time :- 937	**Time :- 944
**Backtracks :- 12	**Backtrack:- 2	**Backtracks :- 2	**Backtracks :- 2

### Puzzle 4

Hard1

0400500000100800200600001041003009000076492000020010083  
05000090020090070000060050

MRV+FC	MRV+FC+AC:	MRV+FC+Prob+AC	MRV+FC+LCV+AC
248156739	248156739	248156739	248156739
713984526	713984526	713984526	713984526
569732184	569732184	569732184	569732184
156328947	156328947	156328947	156328947
837649215	837649215	837649215	837649215
492571368	492571368	492571368	492571368
385417692	385417692	385417692	385417692
621895473	621895473	621895473	621895473
974263851	974263851	974263851	974263851
**Time :- 2664	**Time :- 1907	**Time :- 1593	**Time :- 1350
**Backtracks :- 0	**Backtrack:- 8	**Backtracks :- 8	**Backtracks :- 6



## Puzzle 5

Hard2

0950100060700390000000070530180000004000700010000005808  
40700000000260040500080720

MRV+FC	MRV+FC+AC:	MRV+FC+Prob+AC	MRV+FC+LCV+AC
395418276 276539814 184627953 718956432 453872691 629143587 842791365 937265148 561384729	395418276 276539814 184627953 718956432 453872691 629143587 842791365 937265148 561384729	395418276 276539814 184627953 718956432 453872691 629143587 842791365 937265148 561384729	395418276 276539814 184627953 718956432 453872691 629143587 842791365 937265148 561384729
**Time :- 1986 **Backtracks :- 5	**Time :- 1509 **Backtrack:- 4	**Time :- 1381 **Backtracks :- 5	**Time :- 1301 **Backtracks :- 4

## Puzzle 6

Evil1

0098002000400010001007200904000000603000100050700000040  
50037006000100050004005900

MRV+FC	MRV+FC+AC:	MRV+FC+Prob+AC	MRV+FC+LCV+AC
769853241 243691578 185724693 421578369 396412785 578369124 852937416 937146852 614285937	769853241 243691578 185724693 421578369 396412785 578369124 852937416 937146852 614285937	769853241 243691578 185724693 421578369 396412785 578369124 852937416 937146852 614285937	769853241 243691578 185724693 421578369 396412785 578369124 852937416 937146852 614285937
**Time :- 14105 **Backtracks :- 340	**Time :- 3392 **Backtrack:- 19	**Time :- 2375 **Backtracks :- 14	**Time :- 12547 **Backtracks :- 137

## Puzzle 7

Evil2

8000000029305000700000015000000400600069154000900300000  
08600000040007018600000007

MRV+FC:	MRV+FC+AC:	MRV+FC+Prob+AC	MRV+FC+LCV+AC
865473192 931562874 427891536 312748965 786915423 594236781 178624359 249357618 653189247	865473192 931562874 427891536 312748965 786915423 594236781 178624359 249357618 653189247	865473192 931562874 427891536 312748965 786915423 594236781 178624359 249357618 653189247	865473192 931562874 427891536 312748965 786915423 594236781 178624359 249357618 653189247
**Time :- 11988 **Backtracks :- 287	**Time :- 1776 **Backtrack:- 5	**Time :- 5037 **Backtracks :- 54	**Time :- 1316 **Backtracks :- 4

## Puzzle 8

Evil3

4207000000000000016010005080000650077000200091004300002  
030008069000000000000004072

MRV+FC:	MRV+FC+AC:	MRV+FC+Prob+AC	MRV+FC+LCV+AC
425718693 398546721 671293548 832965417 754821369 169437285 243179856 987652134 516384972	425718693 398546721 671293548 832965417 754821369 169437285 243179856 987652134 516384972	425718693 398546721 671293548 832965417 754821369 169437285 243179856 987652134 516384972	425718693 398546721 671293548 832965417 754821369 169437285 243179856 987652134 516384972
**Time :- 1986 **Backtracks :- 0	**Time :- 4223 **Backtrack:- 27	**Time :- 1778 **Backtracks :- 23	**Time :- 1519 **Backtracks :- 6