

GENNESARET TJUSILA¹, MATHIEU BESANÇON², MARK TURNER³,
THORSTEN KOCH⁴

How Many Clues To Give? A Bilevel Formulation For The Minimum Sudoku Clue Problem

¹  [0009-0009-7031-9993](https://orcid.org/0009-0009-7031-9993)

²  [0000-0002-6284-3033](https://orcid.org/0000-0002-6284-3033)

³  [0000-0001-7270-1496](https://orcid.org/0000-0001-7270-1496)

⁴  [0000-0002-1967-0077](https://orcid.org/0000-0002-1967-0077)

Zuse Institute Berlin
Takustr. 7
14195 Berlin
Germany


Telephone: +49 30 84185-0
Telefax: +49 30 84185-125


E-mail: bibliothek@zib.de
URL: <http://www.zib.de>


ZIB-Report (Print) ISSN 1438-0064
ZIB-Report (Internet) ISSN 2192-7782


HOW MANY CLUES TO GIVE? A BILEVEL FORMULATION FOR THE MINIMUM SUDOKU CLUE PROBLEM

A PREPRINT

 **Gennesaret Tjusila**
tjusila@campus.tu-berlin.de

 **Mathieu Besançon**[†]
besancon@zib.de

 **Mark Turner**^{**}
turner@zib.de

 **Thorsten Koch**^{* †}
koch@zib.de

May 9, 2023

ABSTRACT

It has been shown that any 9 by 9 Sudoku puzzle must contain at least 17 clues to have a unique solution. This paper investigates the more specific question: given a particular completed Sudoku grid, what is the minimum number of clues in any puzzle whose unique solution is the given grid? We call this problem the Minimum Sudoku Clue Problem (MSCP). We formulate MSCP as a binary bilevel linear program, present a class of globally valid inequalities, and provide a computational study on 50 MSCP instances of 9 by 9 Sudoku grids. Using a general bilevel solver, we solve 95% of instances to optimality, and show that the solution process benefits from the addition of a moderate amount of inequalities. Finally, we extend the proposed model to other combinatorial problems in which uniqueness of the solution is of interest.

1 Introduction

The Sudoku puzzle first appeared in the May 1979 edition of *Dell Pencil Puzzle and Word Games* [5]. Given a square integer n , the puzzle is given on a $n \times n$ grid divided into n subgrids each of size $\sqrt{n} \times \sqrt{n}$. As input, some cells are already filled with numbers between 1– n . The goal of the puzzle is to fill the rest of the cells such that each number between 1– n appears exactly once in each row, column, and subgrid. An example of a Sudoku puzzle along with its solution is given in Figure 1. For most Sudoku puzzles, uniqueness of the solution is a desirable property. We call such puzzles *valid*. It is fairly easy to construct examples of 9×9 Sudoku puzzles with 77 clues and multiple solutions (such as removing the entries marked in green in Figure 1b). One can also observe that any puzzle with at least 78 clues will always have a unique solution.

A natural question that arises is: what is the minimum number of clues that a valid puzzle can have? It is shown in [25] that the answer to this question is 17 clues. But what if the puzzle designer already has a solution grid in mind? This motivates the Minimum Sudoku Clue Problem (MSCP): what is the minimum number of clues on any valid puzzle for a given Sudoku grid?

In this paper, we make four key contributions. First, we formulate the MSCP as a binary bilevel linear program, allowing the use of generic integer bilevel methods and solvers, which to the best of our knowledge is a first in the literature. Second, we present unavoidable set inequalities, a set of globally valid inequalities, which we add at the start of the solving process to improve solver performance. Third, we provide computational results over a set of Sudoku grids to show the viability of our approach. Finally, we generalize our model to other problems which fulfill some assumption in the Fewest Clue Problem (FCP) class introduced in [6]. We note that this paper is an extension of the first author’s thesis work [31].

2 Related Work

The problem of counting the total number of $n \times n$ Sudoku grids is an open problem. For $n = 9$, it was shown in [11] that the number of Sudoku grids is around 6.671×10^{21} . A natural upper bound arises by considering that

*Zuse Institute Berlin, Department of Mathematical Optimization, Takustr. 7, 14195 Berlin

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

(a) A Sudoku Puzzle

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

(b) The Solution Grid and its Unavoidable Sets

Figure 1: A Sudoku puzzle along with its solution and unavoidable sets

Sudoku grids are a subset of Latin squares with additional subgrid constraints. The enumeration of Latin squares has been extensively studied in the literature [26], which has pushed similar studies for Sudoku [1, 17]. Many of these grids are equivalent under transformations such as relabeling of digits and rotations. We call the lexicographically smallest Sudoku grid that is equivalent to a given grid under these transformations the *minlex form* of the Sudoku grid [22]. Taking these transformations into account, the number of 9×9 Sudoku grids is reduced to around 5.47×10^9 essentially different grids [29]. While our work focuses on the minimum number of clues for a given Sudoku grid, the minimum number of clues for any Sudoku grid has been shown to be 17 through a computer-assisted proof [25]. A list of nearly 50000 Sudoku puzzles with 17 clues is collected by Gordon Royle [27]. This collection is only a fraction of the possible number of Sudoku grids, heavily suggesting that most Sudoku grids do not have a 17 clue valid puzzle. Minimum bounds for the 4×4 number of clues have been derived through an algebraic process in [14] by encoding the combinatorial problem as a polynomial and analyzing its structure. Research in this direction for 9×9 grids has focused on analyzing the underlying graph structure of the Sudoku grid [4, 23] and characterizing valid Sudoku puzzles using formal logic [24].

Finding a solution to a general $n \times n$ Sudoku puzzle is ASP-complete, which implies NP-completeness of the decision problem as well as #P-completeness to count the solutions [32]. However, practical methods for solving Sudoku puzzles of size 9×9 exist in the literature [3]. There has also been recent research in making algorithms that solve Sudoku puzzles explainable for humans [2]. Given a Sudoku grid, the decision problem “is there a setting of at most k clues such that the only solution is the given grid?” is a member of the class of problem “Fewest Clue Problem” (FCP) and has been shown to be Σ_2^P -complete [6]. Mixed-integer bilevel linear programming has also been shown in [7, 20] to be Σ_2^P -complete. Therefore, transforming MSCP into a binary bilevel linear program retains the same complexity but allows for a general solving method.

To the best of our knowledge, all existing software libraries for solving MSCP are problem-specific and created by the Sudoku community, see [9] for an example. The software uses pattern-matching algorithms to quickly find so-called *unavoidable sets*, such as described in [25]. An unavoidable set is defined as a set of cells whose contents if removed will result in an invalid Sudoku puzzle. An example of such sets would be the cells marked in green, red, or blue in Figure 1b. Given a set of unavoidable sets S , we call a set of cells H a *hitting set* if for every unavoidable set in S at least one cell is contained in H . Once a large enough set of unavoidable sets has been generated, one can enumerate over all hitting sets of these unavoidable sets, starting from ones with minimal cardinality until a valid puzzle is found. Although in theory, enumerating unavoidable sets is expensive, specialized algorithms are often fast in practice owing to additional problem-specific methods, e.g. exploiting equivalence classes of Sudoku grids. We also highlight that enumeration of hitting set, in particular minimal hitting set, is an active area of research [15]. In contrast to existing software, our work uses a general mathematical optimization approach to solve MSC. We will use integer linear programming models to find unavoidable sets and generate valid inequalities to speed up the bilevel-solving process.

3 Integer Bilevel Linear Formulations of Minimum Sudoku Clue Problem

We now formulate MSCP for a Sudoku grid of size $n \times n$ where n is a square number. Let x_{ijk} be a set of binary decision variables where $i, j, k \in [n] := \{1, \dots, n\}$. The variable x_{ijk} takes value one if cell (i, j) has entry k and

zero otherwise. The variables construct an $n \times n$ Sudoku grid if they satisfy

$$\sum_{k=1}^n x_{ijk} = 1, \quad \forall i, j \in [n] \quad (G0)$$

$$\sum_{j=1}^n x_{ijk} = 1, \quad \forall i, k \in [n] \quad (G1)$$

$$\sum_{i=1}^n x_{ijk} = 1, \quad \forall j, k \in [n] \quad (G2)$$

$$\sum_{i=sp-s+1}^{sp} \sum_{j=sq-s+1}^{sq} x_{ijk} = 1, \quad \forall p, q \in [s] \text{ and } k \in [n] \quad (G3)$$

where $s := \sqrt{n}$. This is the standard Sudoku integer linear program (ILP) formulation found in the literature, see [18, 21].

Let G be a Sudoku grid given as an $n \times n$ matrix with entries in $[n]$. The leader problem of our binary bilevel linear program will act as a ‘‘puzzle setter’’, and determine which entries of the Sudoku grid are given as clues. The follower problem will act as an ‘‘adversary’’ that tries to find a solution different from the given Sudoku grid. Concretely, our model is as follows

$$\begin{aligned} \min_{x, y, z} \quad & \sum_{i=1}^n \sum_{j=1}^n y_{ij} \\ \text{s.t.} \quad & z = 1 \\ & y_{ij} \in \{0, 1\}, \quad \forall i, j \in [n] \\ & (x, z) \in S(y) \end{aligned} \quad (V1)$$

where $S(y)$ is the set of optimum solutions to the y -parameterized follower problem

$$\begin{aligned} \min_{x, z} \quad & z \\ \text{s.t.} \quad & (G0) - (G3) \\ & x_{ijG_{ij}} \geq y_{ij}, \quad \forall i, j \in [n] \quad (F1) \\ & \sum_{i=1}^n \sum_{j=1}^n x_{ijG_{ij}} - z \leq n^2 - 1 \quad (N1) \\ & x_{ijk}, z \in \{0, 1\}, \quad \forall i, j, k \in [n]. \end{aligned}$$

The leader decision variable, y_{ij} , determines whether the entry of a cell (i, j) is given to the follower problem as a clue. The objective function of the leader problem is the number of clues given. Constraint (F1) requires the follower problem to adhere to these given clues. Constraint (N1) requires that the Sudoku grid defined by the set of decision variables x_{ijk} with $i, j, k \in [n]$ is different from G . This constraint can be relaxed by setting z to one and taking a penalty. The intuition of the leader constraint (V1) is as follows: The objective of the follower is to minimize this penalty. If the puzzle determined by the leader problem has multiple solutions, the follower can find a feasible solution with a penalty of zero. However, if this is not possible, then the puzzle determined by the leader is a valid puzzle and the only option the follower has is to take the penalty.

Finally, we highlight that the high-point relaxation is always trivially achieved by setting $z = 1$, $y_{ij} = 0$ for all $i, j \in [n]$ and x to be another Sudoku grid not equal to G , by permuting digits for instance. This weakness of the relaxation suggests the hardness of the bilevel problem.

4 Strengthening The Bilevel Formulation Through Valid Inequalities

Consider the Sudoku grid given in Figure 1b. We can swap the 3’s and 8’s in the green marked cells to get a new Sudoku grid G' that has the same entries except for the cells marked in green. Thus, any valid puzzle P must have at least one clue in one of the green-marked cells. Similarly, we observe that it is possible to change the entries of

cells marked in blue or red. Thus, there must also be at least one clue in the cells marked red and one clue in the cells marked blue. We call a set of cells U an *unavoidable set* for a Sudoku grid G if there exists a Sudoku grid $G' \neq G$ that differs from G only on cells in U . We call an unavoidable set *minimally unavoidable* if it contains no subset that is again unavoidable. In what follows, we represent Sudoku grids as $n \times n$ matrices with entries from $[n]$ and Sudoku puzzles as $n \times n$ matrices with entries from $[n] \cup \{0\}$ where 0 marks an empty cell.

Proposition 1. *Let G be a grid and P a puzzle such that G is a solution of P . Then P is a valid puzzle, if and only if, for every minimally unavoidable set U of G there exists a cell $(i, j) \in U$ that is given as a clue, i.e., $P_{ij} \neq 0$.*

Proof. To show sufficiency suppose that there exists a minimally unavoidable set U such that $P_{ij} = 0$ for all cells $(i, j) \in U$. By definition there exists a Sudoku grid $G' \neq G$ which differs from G only in the entries of cells that are in U . Since $P_{ij} = 0$ for all cells $(i, j) \in U$ then G' is also a solution of P . Thus P is not a valid puzzle.

To show necessity, suppose that P is not a valid puzzle and there exists a Sudoku grid G' which is a solution of P and $G' \neq G$. We define

$$U := \{(i, j) \in \{1, \dots, n\}^2 \mid G'_{ij} \neq G_{ij}\}$$

as the set of cells whose entry in G is different from its entry in G' . By construction, U is an unavoidable set and $P_{ij} = 0$ for all $(i, j) \in U$, as otherwise, their entries would be identical. If U is minimally unavoidable then we are done, otherwise, a subset of U is again unavoidable. Since U is finite, we can iterate the process until we end up with a minimally unavoidable set. \square

Corollary 2. *Let G be a Sudoku grid and U be an arbitrary minimal unavoidable set. Then, the inequality*

$$\sum_{(i,j) \in U} y_{ij} \geq 1 \tag{U}$$

*is a globally valid inequality for the leader of our bilevel program. We call this inequality the **unavoidable set inequality** corresponding to U*

We give a method to generate the set of unavoidable sets \mathcal{U} . Let $m \in \mathbb{N}$ with $m \geq 1$. Consider the m -parameterized integer linear program,

$$\begin{aligned} \min_x \quad & 0 \\ \text{s.t.} \quad & (G0) - (G3) \\ & \sum_{i=1}^n \sum_{j=1}^n x_{ijG_{ij}} = n^2 - m \\ & x_{ijk} \in \{0, 1\}, \quad \forall i, j, k \in [n], \end{aligned} \tag{D1}$$

The integer linear program gives us a Sudoku grid G' which differs from G in exactly m entries. We get that

$$U := \{(i, j) \in \{1, \dots, n\}^2 \mid G_{ij} \neq G'_{ij}\}.$$

is an unavoidable set of G by construction. We start with $m = 1$ and repeatedly solve the ILP, adding in each iteration the no-good cut constraint

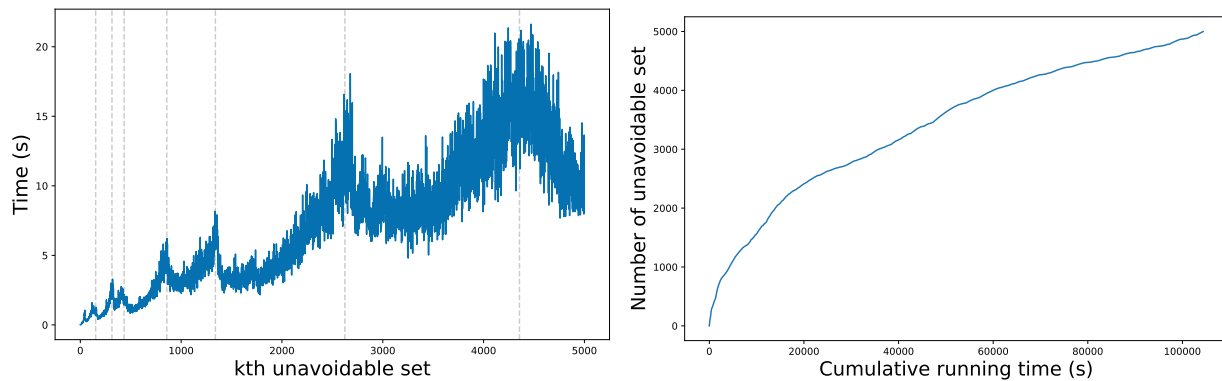
$$\sum_{(i,j) \in U} x_{ijG_{ij}} \geq 1 \tag{N2}$$

which bars the ILP from returning any G' , whose associated unavoidable set is a superset of U . The ILP will thus return a different unavoidable set of size m in each iteration. Once all unavoidable sets of size m have been generated, we move on to $m + 1$. Note that we could have equivalently formulated this as a minimization problem.

Proposition 3. *For the procedure described above it applies*

- (i) *At each iteration, the resulting unavoidable set will always be a minimally unavoidable set*
- (ii) *Repeating the procedure eventually yields all minimal unavoidable sets*

Proof. To show (i), let \bar{U} be an unavoidable set that is not minimal and $U \subset \bar{U}$ a minimal unavoidable set with $m := |\bar{U}|$. When generating all unavoidable sets of size m , a no-good cut for U will also be added to the formulation. Thus, any G' which generates \bar{U} will be infeasible because $U \subset \bar{U}$. We get (ii) by construction. \square



(a) Geometric mean of time needed to generate the k th unavoidable set (b) The number of unavoidable set generated as a function of time

Figure 2: Experiment Results For Cut Generation

5 Computational Results

In this section, we investigate the performance of our models for solving MSCP over 50 instances of 9×9 Sudoku grids. All of our computations run on a single thread of an Intel Xeon E5-2630V4 2.2 GHz. A wall-clock time limit of 4 days and a memory limit of 16 GB was used for each run. The algorithm to generate unavoidable set cuts uses Gurobi 9.5.1 [16] as an ILP solver, and we use the bilevel solver from [12] to solve the main model, where the authors granted us a license upon request. The solver uses CPLEX 12.7 [19] to solve linear programming relaxations. The code used for this section along with the computational results can be found in <https://github.com/gtjasila/minimum-sudoku>.

The 50 instances are split into two groups of 25. The first group is randomly selected from a list of Sudoku puzzles with 17 clues [28]. The second group is randomly selected from the list of Sudoku puzzles with a difficulty rating of more than 11 (the maximum difficulty rating being 12) maintained by the new Sudoku players forum². The known puzzles for all instances in this second group contain more than 20 clues each. To get a diverse instance set, we also ensure that we select Sudokus with different minlex forms [22]. To convert the Sudoku grids to minlex form, we use the code from [8].

First, we evaluate the performance of the unavoidable sets generating procedure. For each of our 9×9 instances, we generate 5000 minimal unavoidable sets. In all instances, we generate all minimal unavoidable sets of size 16 or less. We observed no unavoidable sets of size 5 and 7, which leads us to conjecture that none exist for any instance. In 39 out of 50 instances, we generated all minimal unavoidable sets of size less than or equal to 17. We plot the geometric mean of the time needed to generate the k th unavoidable set over all 50 of our instances in Figure 2a. We see that generally, the time needed to generate an unavoidable set increases as k gets larger. An interesting feature of the figure is the periodic peaks. Looking deeper into the result of individual instances, we see that as we try to enumerate all minimal unavoidable sets of size $n \in \mathbb{N}$, the time increases in each iteration. This is expected as in each iteration there are fewer and fewer minimal unavoidable sets of size n available, and thus, they become increasingly hard to find. To visualize this effect, we computed the average number of unavoidable sets less than or equal to n for $n = 11, \dots, 17$ (for the instance in which not all unavoidable sets of size 17 have been found, we assume the number of unavoidable set of size less than 17 to be 5000) and drew them as vertical lines in Figure 2a. One can think of these lines as the average point where an instance switches from searching for unavoidable sets of size n to size $n + 1$. The leftmost line represents $n = 11$.

Figure 2a does not catch how extreme these peaks can be. To see this effect, we provide the frequency distribution table of the generation time of unavoidable sets in Table 1. Though the majority of the minimally unavoidable sets (94.10%) can be generated in less than 1 minute, some minimal unavoidable sets are very hard to find with the longest taking nearly 3 hours to find.

Lastly, it is important to remember that we are not obliged to generate all unavoidable sets since their sole function is to help reduce the feasible region of our bilevel program and improve performance. For this reason, we find it

²<http://forum.enjoysudoku.com/the-hardest-Sudokus-new-thread-t6539-600.html#p277835>

generation time [s]	# of unavoidable sets
≤ 1	30763
1 – 10	136009
10 – 30	42389
30 – 60	26092
60 – 300	13809
300 – 600	542
600 – 1800	298
1800 – 3600	62
3600 – 7200	31
≥ 7200	5

Table 1: Frequency distribution table of the time needed to generate unavoidable set

helpful to plot the average number of cuts generated as a function of time. To do this, for each instance I and each $n \in [5000]$, we calculate the cumulative time our model takes to generate n unavoidable sets of instance I . We then take the geometric mean of the cumulative time for each n over all the instances and plot the result as a function of n . The resulting plot is shown in Figure 2b. The figure reiterates that generating unavoidable sets is quicker in the beginning and shows how it becomes more difficult over time. It takes less than 20000 seconds to generate the first 2000 unavoidable sets and nearly 40000 seconds to generate the next 2000.

We will now test the effect of unavoidable set inequalities on our model by varying the number of inequalities that are used. For our initial analysis, we do not take into account the time needed to generate the unavoidable sets. We decide to test 500, 1000, 3000, and 5000 unavoidable set inequalities, where we use the first n inequalities generated by our unavoidable set generating algorithm. A summary of the optimization results is shown in Table 2. 45 out of the 50 instances of size 9×9 solved to optimality in at least one solver setting. Interestingly, all instances in the 17 clues puzzle group solve to optimality in at least one solver setting and they generally solve faster than the instance group with no 17 clue puzzle, see Figure 4.

# of unavoidable set	# of instances	
	optimal	time limit
500	37	13
1000	43	7
3000	36	14
5000	29	21

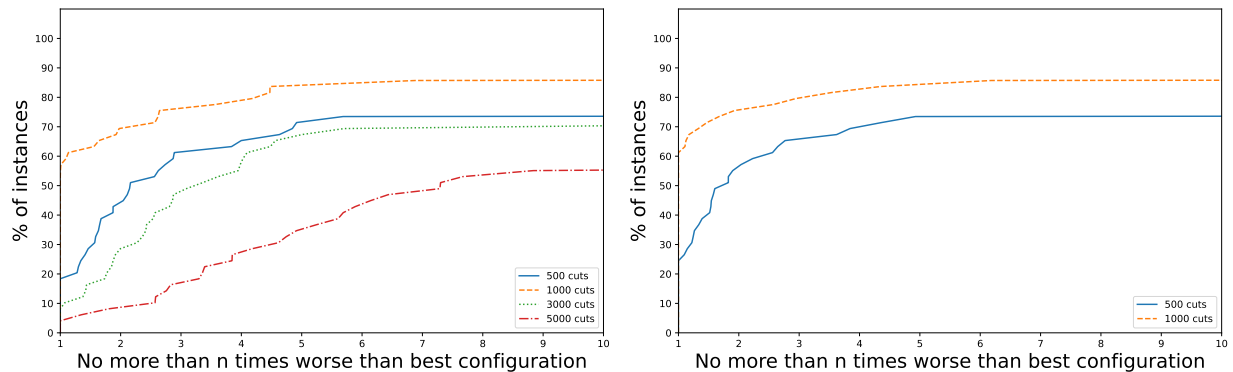
Table 2: Summary of end result for 9×9 standard bilevel model with different number of unavoidable set cuts

We plot the resulting performance profile [10] in Figure 3a. We observe that adding too few or too many inequalities results in slower optimization times. Nearly 60% of the instances solve fastest on models that use 1000 unavoidable set inequalities, followed by slightly under 20% of instances that solve fastest on models that use 500 unavoidable set inequalities. This claim is also supported when we see that we solve to optimality in most instances when we are using 1000 unavoidable sets inequalities. By looking deeper into node-level data as presented in Table 3, we see that too few inequalities result in a huge increase of nodes processed to prove optimality, while too many inequalities result in a huge decrease in node throughput. The best choice is therefore likely to be in the middle.

# of unavoidable set	node count	time per node [s]	total runtime [s]
500	2503382	0.028	70312
1000	2097150	0.031	64941
3000	1126612	0.082	92931
5000	887962	0.139	123472

Table 3: Geometric average of node count, time per node, and runtime of different settings

Finally, we take into account the time needed to generate the inequalities. Note that we only need to compare models with 500 and 1000 inequalities since 1000 inequalities models outperform the 3000 and 5000 inequalities model cut



(a) Performance profile of different number of unavoidable set on 9×9 instances (b) Performance profile of different number of unavoidable set on 9×9 instances with augmented time

Figure 3: Experiment Results For Solving 9 by 9 instances

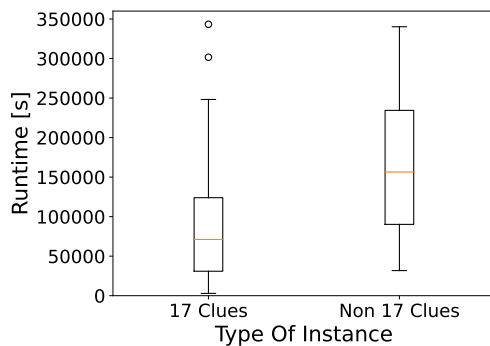


Figure 4: Comparison Of Solving Time For 17 Clue Instances and Non 17 Clue Instances

models. To compare if it is worth generating the 500 extra unavoidable sets, we compute the time needed to solve the bilevel instance plus the time needed to generate the unavoidable sets and plot the performance profile for instances with 500 and 1000 inequalities. The calculation is done by using the data from the generating unavoidable set experiments. The resulting plot is shown in 3b. We see that even accounting for the unavoidable set cut generation time, using 1000 inequalities is still superior to using 500 inequalities.

For our experiments, we also obtained preliminary results with the MiBS solver [30]. Even on easy instances however, we quickly observed that MiBS required much more time than the solver from [13]. We believe that this is in large part due to the inability of MiBS to find a primal solution to our problem.

6 Generalization of the Model to other Fewest Clue Problems

A desire for unique solutions is not only relevant to Sudoku, with other example problems being Slither Link and Cross Sum [32]. This motivates the definition of the ‘‘Fewest Clue Problem’’ (FCP) class in [6]. In this section, we show how our model can also be adapted for FCP problems of other puzzles which have a linear binary formulation.

We restate the definition of the Fewest Clue Problem as in [6]. Let A be a problem in NP. We denote with R_A the set of instance-certificate pairs where the certificates are binary strings of length l . For a given instance I of A , we call a string $c \in \{0, 1, \perp\}$ a *clue* if there exists a certificate c^* such that $(I, c^*) \in R_A$ and $c_i = c_i^*$ for all indices $i \in [k]$ where $c_i \neq \perp$. The symbol \perp can be interpreted as a missing or non-specified entry. We call c^* a *satisfying solution* to clue c . The *size* of a clue is the number of non \perp characters.

We define FCP A to be the decision problem: given an instance I , a certificate c^* and an integer k , does there exist a clue c of size at most k for which the unique satisfying solution is c^* ? We note that our definition is a slight variant of that proposed in [6].

We make the assumption that there exists an l -dimensional polytope \mathcal{Q} such that c is a valid certificate if and only if $c \in \mathcal{Q}$ and is binary. The FCP A can be written as a bilevel optimization problem as follows:

$$\begin{aligned} \min_{x,y,z} \quad & \sum_{i=1}^l y_i \\ \text{s.t.} \quad & z = 1 \\ & y_i \in \{0, 1\}, \quad \forall i \in [l] \\ & (x, z) \in S(y) \end{aligned}$$

where $S(y)$ is the set of optimum solutions to the y -parameterized follower problem:

$$\begin{aligned} \min_{x,z} \quad & z \\ \text{s.t.} \quad & x \in \mathcal{Q} \\ & x_i \geq y_i, \quad \forall i \in [l], c_i^* = 1 \\ & \sum_{i \in [l], c_i^* = 1} x_i + \sum_{i \in [l], c_i^* = 0} (1 - x_i) - z \leq l - 1 \\ & x_i, z \in \{0, 1\}, \quad \forall i \in [l]. \end{aligned} \tag{NG}$$

The leader program determines which indices are given in the clue, while the follower tries to find an alternative solution respecting the clue. Constraint (NG) is a no-good constraint prohibiting the assignment $x = c^*$ if $z = 0$. It is trivially fulfilled if $z = 1$, it is a generalization of the equivalent constraint of the Sudoku-specific model presented in Section 3.

7 Conclusion and Outlook

In this paper, we have shown that the Minimum Sudoku Clue problem can be formulated and solved as a binary bilevel linear programming problem. By introducing unavoidable-set inequalities, we showed that the formulation can be tightened, and that solver performance can be improved. Our models are able to compute a provable optimal solution to the Minimum Sudoku Clue problem in 95% of instances. Despite these performance results, the inherent complexity of the Minimum Sudoku Clue problem and the more general Fewest Clue problem complicates scaling to larger instances. Unlike specialized ad hoc enumeration techniques developed in the Sudoku literature [9] however, our approach naturally benefits from the continued improved performance of mixed-integer programming solvers.

We see three main avenues of future research for the Minimum Sudoku Clue problem. First, we can use faster unavoidable set finding algorithms such as the one proposed by [25]. Second, we can develop formulations that exploit the symmetries of Sudoku grids. Third, we can develop a branch-and-cut approach leveraging unavoidable set inequalities to separate non-feasible solutions throughout the branch-and-bound process instead of initially applying a large number of inequalities.

Acknowledgments

The work for this article has been conducted in the Research Campus MODAL funded by the German Federal Ministry of Education and Research (BMBF) (fund numbers 05M14ZAM, 05M20ZBM). The described research activities are funded by the Federal Ministry for Economic Affairs and Energy within the project UNSEEN (ID: 03EI1004-C). We thank Markus Sinnl and coauthors for providing us a license to their bilevel solver, Fakultät II at the Technische Universität Berlin for allowing us to use their HPC facility, and Kai Hoppmann-Baum for initial advice on the integer formulation.

References

- [1] D. BEREND, *On the number of Sudoku squares*, Discrete Mathematics, 341 (2018), pp. 3241–3248.
- [2] J. BÜRGER, *Sudoku Assistant – an AI assistant combining machine learning and reasoning – Leuven.AI Stories – ai.kuleuven.be*. <https://ai.kuleuven.be/stories/post/2023-02-08-sudoku/>, 2023. [Accessed 24-Apr-2023].
- [3] L. C. COELHO AND G. LAPORTE, *A comparison of several enumerative algorithms for Sudoku*, Journal of the Operational Research Society, 65 (2014), pp. 1602–1610.

- [4] J. COOPER AND A. KIRKPATRICK, *Critical sets for Sudoku and general graph colorings*, Discrete Mathematics, 315-316 (2014), pp. 112–119.
- [5] J.-P. DELAHAYE, *The science behind Sudoku*, Scientific American, 294 (2006), pp. 80–7.
- [6] E. D. DEMAINE, F. MA, A. SCHVARTZMAN, E. WAINGARTEN, AND S. AARONSON, *The fewest clues problem*, Theoretical Computer Science, 748 (2018), pp. 28–39.
- [7] S. T. DENEGRE, *Interdiction and Discrete Bilevel Linear Programming*, PhD thesis, Lehigh University, 2011.
- [8] M. DEVERIN, *Minlex Form: Min and Max Lists / Chaining*. <http://forum.enjoysudoku.com/minlex-form-min-and-max-lists-chaining-t30325-15.html>. [Accessed 24-Apr-2023].
- [9] M. DOBRICHEV, *GridChecker, an exhaustive puzzle enumerator*. <http://forum.enjoysudoku.com/gridchecker-an-exhaustive-puzzle-enumerator-t30071.html>. [Accessed 26-Apr-2023].
- [10] E. D. DOLAN AND J. J. MORÉ, *Benchmarking optimization software with performance profiles*, Mathematical Programming, 91 (2002), pp. 201–213.
- [11] B. FELGENHAUER AND F. JARVIS, *Mathematics of Sudoku I*, Mathematical Spectrum, 39 (2006), pp. 15–22.
- [12] M. FISCHETTI, I. LJUBIĆ, M. MONACI, AND M. SINNL, *A new general-purpose algorithm for mixed-integer bilevel linear programs*, Operations Research, 65 (2017), pp. 1615–1637.
- [13] ———, *On the use of intersection cuts for bilevel optimization*, Mathematical Programming, 172 (2017), pp. 77–103.
- [14] A. FORROW AND J. R. SCHMITT, *Approaching the minimum number of clues Sudoku problem via the polynomial method*, 2013.
- [15] A. GAINER-DEWAR AND P. VERA-LICONA, *The minimal hitting set generation problem: algorithms and computation*, SIAM Journal on Discrete Mathematics, 31 (2017), pp. 63–100.
- [16] GUROBI OPTIMIZATION, LLC, *Gurobi Optimizer Reference Manual*, 2022.
- [17] A. M. HERZBERG AND M. R. MURTY, *Sudoku squares and chromatic polynomials*, Notices of the AMS, 54 (2007), pp. 708–717.
- [18] T. HÜRLIMANN, *Puzzles and Games: A Mathematical Modeling Approach*, Department of Informatics, University of Fribourg, Fribourg, Switzerland, 4th ed., 2016.
- [19] IBM, *IBM ILOG CPLEX 12.8 User's Manual*, IBM ILOG CPLEX Division, Incline Village, NV, 2017.
- [20] R. G. JEROSLOW, *The polynomial hierarchy and a simple model for competitive analysis*, Mathematical programming, 32 (1985), pp. 146–164.
- [21] T. KOCH, *Rapid Mathematical Programming or How to Solve Sudoku Puzzles in a Few Seconds*, in Operations Research Proceedings 2005, H.-D. Haasis, H. Kopfer, and J. Schönberger, eds., Berlin, Heidelberg, 2006, Springer Berlin Heidelberg, pp. 21–26.
- [22] C. LASS, *Minimal number of clues for sudokus*, Open Computer Science, 2 (2012), pp. 143–151.
- [23] G.-C. LAU, J. M. JEYASEELI, W.-C. SHIU, AND S. ARUMUGAM, *Sudoku number of graphs*, arXiv preprint arXiv:2206.08106, (2022).
- [24] D. MAŠULOVIĆ, *Deducibility in Sudoku*, arXiv preprint arXiv:2212.01053, (2022).
- [25] G. MCGUIRE, B. TUGEMANN, AND G. CIVARIO, *There is no 16-clue Sudoku: Solving the Sudoku minimum number of clues problem via hitting set enumeration*, Experimental Mathematics, 23 (2014), pp. 190–217.
- [26] B. D. MCKAY AND I. M. WANLESS, *On the number of Latin squares*, Annals of combinatorics, 9 (2005), pp. 335–344.
- [27] G. ROYLE, *Minimum Sudoku*. <https://web.archive.org/web/20120730100322/http://mapleta.maths.uwa.edu.au/~gordon/sudokumin.php>. [Accessed 23-Apr-2023].
- [28] ———, *Minimum Sudoku*. <https://web.archive.org/web/20060220092603/http://www.csse.uwa.edu.au/~gordon/sudokumin.php>.
- [29] E. RUSSELL AND F. JARVIS, *Mathematics of Sudoku II*, Mathematical Spectrum, 39 (2006), pp. 54–58.
- [30] S. TAHERNEJAD, T. K. RALPHS, AND S. T. DENEGRE, *A branch-and-cut algorithm for mixed integer bilevel linear optimization problems and its implementation*, Mathematical Programming Computation, 12 (2020), pp. 529–568.
- [31] G. K. TJUSILA, *A Bilevel Model for the Minimum Sudoku Clue Problem*, Bachelor's Thesis, Technische Universität Berlin, 2022. available on [https://github.com/gtjusila/minimum-sudoku/blob/main/Bachelorarbeit`Gennesaret`Tjusila`407687.pdf](https://github.com/gtjusila/minimum-sudoku/blob/main/Bachelorarbeit%20Gennesaret%20Tjusila%20407687.pdf).
- [32] T. YATO AND T. SETA, *Complexity and completeness of finding another solution and its application to puzzles*, IEICE TRANSACTIONS on Fundamentals of Electronics, 86 (2003), pp. 1052–1060.