Universiteit Leiden
The Netherlands

Opleiding Informatica

Probing Sequences

for Nonograms

Roos Wensveen

Supervisors:
W.A. Kosters & J.M. de Graaf

BACHELOR THESIS

Leiden Institute of Advanced Computer Science (LIACS)
www.liacs.leidenuniv.nl

July 25, 2022

Abstract

Nonograms, Japanese crosswords, are graphical logic puzzles where cells can be filled or be left blank according to the provided numbers on the side of the grid, that describe the lengths of consecutive blocks of filled cells. Solving Nonograms is done by performing two solving methods; *Simple* and *Probing*. *Simple* tries to solve a Nonogram by solely applying *H*- and *V-sweeps*, that address single rows or columns. When this approach does not succeed in fully solving the Nonogram, *Probing* is applied. Here, we try an unknown cell by filling it and applying *Simple* again, followed by leaving the cell blank and performing *Simple* once more. We can determine the usefulness of each probed cell, and by probing all cells, we get the probing sequence of a Nonogram.

This thesis analyzes the probing sequences of Nonograms. Using small Nonograms, we try to determine all sequences leading to uniquely solvable Nonogram, Nonograms with exactly two solutions and Nonograms with multiple solutions. Any relevant observations will be described as well.

Contents

1 Introduction

Nonograms, also known as Japanese crosswords, Picross and Hanjie, are graphical logic puzzles. These puzzles consist of an $m \times n$ grid, where all pixels must be colored or be left blank with respect to a provided description of each line (i.e., row or column). Traditionally, Nonograms are represented by black and white cells. However, colored Nonograms exist as well. Colored Nonograms have a description which describes the required color for each of the lines. Being invented in the 1980s, these puzzles are still frequently used in newspapers and other applications. Paint by Number is a well-known paint technique based upon the same concept. The popularity of the Nonogram rose due to Nintendo's implementation for electronic toys. Of course, the Nonograms used in these games should be solvable by humans. Solving a Nonogram manually could be done by guessing a cell, logically eliminating values of cells and overlapping all possible solutions of a small part of the Nonogram.

The goal of the player is to "solve the grid" with respect to all provided line descriptions. The line description consists of k numbers and denotes k consecutive series of filled cells. These series need to be separated by at least one blank cell. The order in which these k numbers occur, should be preserved. All these series must occur in the corresponding row or column. Once solved, one might obtain a hidden image. Solving Figure 1 results in such a hidden image.

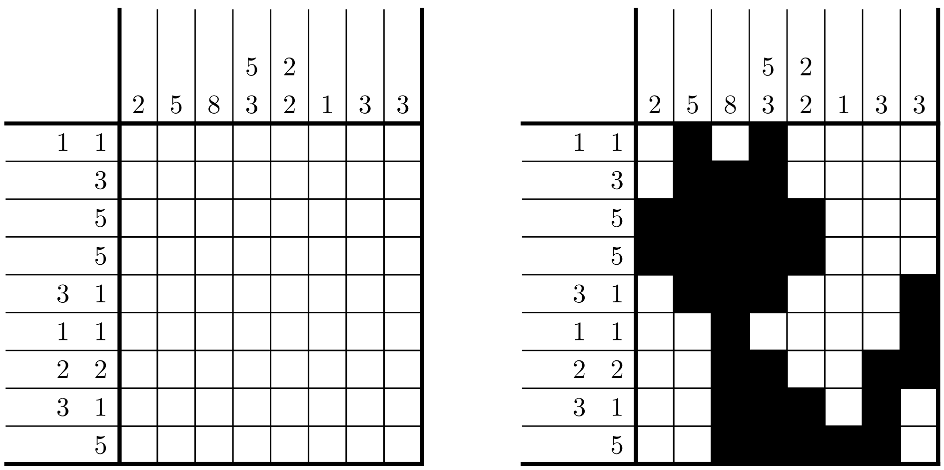


Figure 1: A Nonogram with its solution.

1.1 Solving a line

The line description "2 3 1" describes a line that should be structured as two colored cells, followed by at least one blank cell, then three colored cells followed by at least one blank cell, and finally one colored cell. Optionally, we may have blank cells at the beginning or end of the line. The number of solutions for the descriptions depends on the length of the line. When "2 3 1" is the line description with a line length equal to 8, we obtain exactly one solution, see Figure 2.

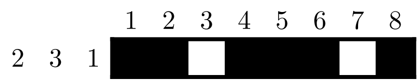


Figure 2: A line solution for a line of length 8.

Consider “2 3 1” to be the line description of a partially filled line of length 11, see Figure 3. Suppose the second cell must be left blank, which is denoted by “·”. When trying to solve this line further, we obtain multiple solutions, see Figure 4. Considering the fourth cell in each of the possible solutions, we can see that in all cases this cell has been colored. Because of this, it is certain that this cell needs to be colored. The first cell of each solution is blank, so it is certain that this cell should be blank. The third cell of all solutions is not uniformly colored as the latter solution contains a blank cell where all other solutions have a colored cell. So, as for now, it is not possible to determine whether this cell should be colored or not. All cells that have a uniformly colored cell are marked in Figure 4. By combining these four possible solutions, a general partial solution can be obtained, which can be seen in Figure 5. The cells marked with “x” cannot be colored or left blank because of the multiple solutions. When considering this line to be part of a Nonogram, this partial solution might be fully solvable in a later stage of solving.

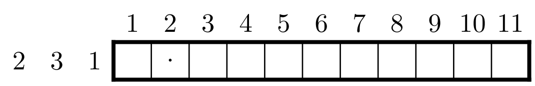


Figure 3: Partially filled line.

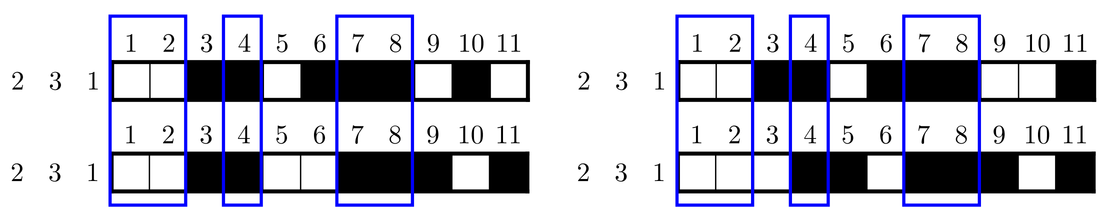


Figure 4: Multiple solutions for one line description.

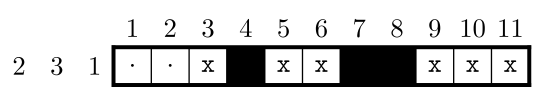


Figure 5: Partial line solution.

When considering the line description “2 3 1” once more, we have seen that it is possible to have multiple grid outcomes for the Nonogram as a whole. The Nonogram cannot perhaps be fully solved. The smallest example of this occurrence, referred to as *switching component*, can be seen in Figure 6. When a Nonogram contains a switching component, there are multiple solutions. There is no unique solution for this Nonogram.

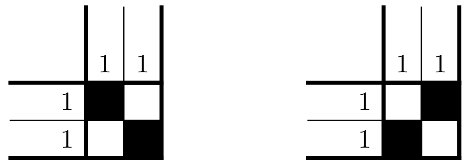


Figure 6: Switching component.

1.2 Thesis overview

First, this thesis started with an introduction of the puzzle. The basic concepts and characteristics of the Nonogram are discussed in this section. Relevant related work will then be discussed in Section 2. Then Section 3.1 will go into the definitions. These definitions will later be used as a guide to obtain a full solving approach, which will be discussed in Section 3. This section is structured to consider the different components of this solving approach. Besides the *Simple* and *Probing* solving approach, determining the difficulty of a Nonogram will be considered as well. In Section 4 the performed experiments are shown and explained. Lastly, a conclusion is given in Section 5. Here, the obtained findings and future work are discussed.

This research is a bachelor thesis for the Computer Science study program at Leiden University, supervised by Walter Kosters and Jeannette de Graaf from the Leiden Institute of Advanced Computer Science (LIACS).

2 Related Work

Several solving approaches have been created before. In *A Comparison of a Genetic Algorithm and a Depth First Search Algorithm Applied to Japanese Nonograms*, Wiggers (2004) uses a genetic algorithm (GA) and depth first search to obtain solutions [WvB04]. The DFS approach used by Wiggers generates all possible solutions of a given description. All possible solutions for each row are combined to result in all possible Nonograms. Due to the large number of possible row solutions, and thus a large number of possible Nonogram solutions, the approach is slow. The GA has a better performance. However, this faster approach does not guarantee a fully correct solution. The GA can get stuck in local optima and its population size does not work well for small Nonograms.

Yu et al. (2011) have written *An Efficient Algorithm for Solving Nonograms*, where it is proposed to use chronological backtracking (CB) to efficiently solve Nonograms [YLC11]. First, their approach uses logic rules until these can no longer be applied. After this, the method will use CB to try and find a solution. This process will be repeated until a solution has been found.

In *A Reasoning Framework for Solving Nonograms* by Batenburg and Kosters a framework for solving Nonograms has been proposed [BK08]. This framework uses logic rules to deduce new information about lines and cells of the Nonogram. Two methods of solving a line are discussed. One of these uses dynamic programming to obtain new information of a line. Recursively, the solution is extended based on the partial solutions of part of this line. The other method discussed uses discrete tomography. Here, the columns and rows are considered to be nodes, between which arcs can be drawn. Each pair of nodes connected by an arc represent which cells are filled. These approaches are then combined by using 2-SAT formulas. In *Constructing Simple Nonograms of Varying Difficulty*, Batenburg et. al consider a difficulty measure, based upon the complexity of a Nonogram and its needed actions to solve it [BHKP09]. This difficulty measure was then used to experiment with and analyze gray level images on various difficulty levels. Batenburg and Kosters have used the mentioned difficulty measure to extend their research regarding the different types of Nonograms [BK12].

The approaches used by Wiggers (2004) to generate all possible Nonograms are either too slow or are not fully able to provide the correct solution of a Nonogram. The dynamic programming approach used by Batenburg and Kosters does not have these issues and serves as a solid basis. The discrete tomography described within this paper and the CB approach used by Yu et. all (2011) both use logic rules to obtain a solution.

3 Solving a Nonogram

Solving a Nonogram and trying to determine the number of possible solutions for a provided Nonogram is an NP-hard problem [UN96]. In order to solve a Nonogram, two solving techniques may be applied.

3.1 Definitions

The alphabet of a Nonogram is defined as $\Sigma = \{0, 1\}$, where "0" represents a white cell, referred to as a blank cell, and "1" a black cell, referred to as a filled cell. We expand Σ to Γ by adding "x" that represents a cell whose value is unknown. In other words, $\Gamma = \{0, 1, x\}$. A line of a Nonogram can be a row or column. For each line, a description is provided. A description d of length $k > 0$ describes an ordered series of integers > 0. This can be defined as $d = (d_1, d_2, \ldots, d_k)$. The description d_j consists of a $\sigma_j \in \Sigma$, $a_j, b_j \in \mathbb{N} \cup \{\infty\}$, denoted as $d_j = \sigma_j\{a_j, b_j\}$, where a_j and b_j represent the lower and upper bound of the number of occurrences of σ_j of description d_j, $1 \leq j \leq k$. We put $A_j = \sum_{p=1}^{j} a_p$ and $B_j = \sum_{p=1}^{j} b_p$. A description may be empty, in which case it is denoted by (0). Consider Figure 1, where for the eighth row, we have "3 1", which represents $(3, 1)$ as an ordered series. The corresponding description is

$$d = (0\{0, \infty\}, 1\{3, 3\}, 0\{1, \infty\}, 1\{1, 1\}, 0\{0, \infty\})$$

The corresponding regular expression is $0^*1^300^*1^10^*$.

An $m \times n$ Nonogram, $m > 0$ and $n > 0$, has m row descriptions and n column descriptions: $D_{\text{rows}} = \langle r_1, r_2, \ldots, r_m \rangle$ and $D_{\text{columns}} = \langle c_1, c_2, \ldots, c_n \rangle$. Row description r_i, for $i = 1, 2, \ldots, m$, describes the description of ith row of the Nonogram. Column description c_j, for $j = 1, 2, \ldots, n$, describes the description of jth column of the Nonogram. The descriptions of both the rows and columns are an ordered series. The row descriptions are denoted from top to bottom, and the column descriptions from left to right. Combining D_{rows} and D_{columns} gives all line descriptions of a provided Nonogram. Consider Figure 1 once again. This Nonogram has the descriptions

$$D_{\text{rows}} = \langle (1, 1), (3), (5), (5), (3, 1), (1, 1), (2, 2), (3, 1), (5) \rangle$$

and

$$D_{\text{columns}} = \langle (2), (5), (8), (5,3), (2,2), (1), (3), (3) \rangle$$

To solve a Nonogram, we start by considering solving lines. As seen in Section 1.1, some lines have exactly one solution, while others have multiple solutions. For lines with multiple solutions, we want to overlay all possible solutions to be able to determine whether the cells should be filled or left blank. To do so, we first discuss how to obtain an individual solution. Then we will combine those solutions to obtain a partial or full line solution. A line can be solved by the use of dynamic programming on string s with length ℓ. The solution of the line is based on solutions found for smaller fragments of this line. Based upon the findings of Batenburg and Kosters [BK08], the following formula will be used:

$$Fix(i,j) = \bigvee_{p=\max(i-b_j, A_{j-1}, L_i^{\sigma_j}(s))}^{\min(i-a_j, B_{j-1})} Fix(p, j-1)$$

Here, for each prefix of the provided line s, a solution is built up based upon earlier (partially) solved prefixes. $Fix(i,j)$ is exactly true if the first i elements of the line s can still adhere with the first j elements of the description d. This formula holds for i and j where $1 \leq i \leq \ell$ and $1 \leq j \leq k$. The summation parameter i, $A_j \leq i \leq B_j$, of $Fix(i,j)$ represents the length of the line and j describes the length of the description. Furthermore, $L_i^{\sigma_j}(s)$ represents the largest index $h \leq i$ for which $s_h \neq \sigma$ and $s_h \neq$ “x”. If this cannot occur, $L_i^{\sigma_j}(s)$ becomes 0. Now, $Fix(i,j)$ will build up until $Fix(\ell, k)$ is reached.

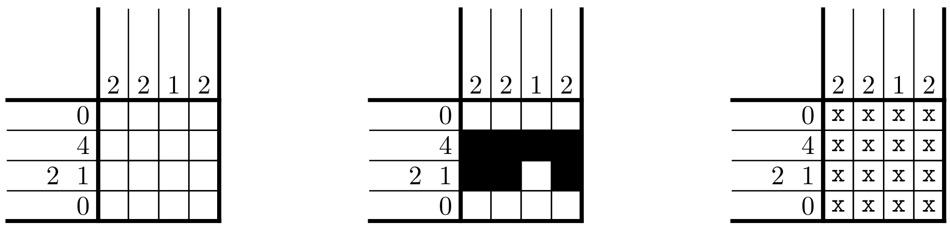
3.2 Simple

The method of solving a single line can be applied more broadly. Multiple lines (the rows or columns) are considered at once. The to be considered lines are those that contain recently altered cells or, in the first step, all lines. By only considering lines with recently altered cells, lines that will not lead to progress of solving the Nonogram are left out. Performing the line solving method on the rows, is called a *H-sweep*. Considering the columns, this will be called a *V-sweep*. The process of alternating between applying *H-* and *V-sweeps* will be repeated until a fully solved Nonogram has been found or no more progress is made, and is described in Figure 7. Figure 7 shows a *Full Sweep*. As a result, we are left with either a completely solved Nonogram or a partially solved Nonogram. When a Nonogram can be solved using only *H-* and *V-sweeps*, we consider the Nonogram to be *Simple*.

The difficulty is determined as well. The *difficulty* of a Nonogram is determined by the number of performed sweeps. To obtain a consistent difficulty, the Nonogram is first solved starting with a horizontal sweep, and later again, starting with a vertical sweep. The ultimate difficulty of a Nonogram is then the average of the two runs. Figure 8 illustrates an example of why this approach is useful. It shows that when starting with a *H-sweep*, the Nonogram can be solved after this one sweep. Therefore, its difficulty becomes 1. When starting with a *V-sweep*, no information can be found. Another sweep needs to be done to obtain the full solution. In total, the run starting with the *V-sweep* has a difficulty of 2. Thus, the final difficulty of the provided Nonogram is equal to 1.5. It should be noted that the difficulty between starting with a *H-sweep* and *V-sweep* differs by at most one.

```
switcher ← 0
while ¬solved do
    if switcher ≡ 0 (mod 2) then
        H-Sweep
    else
        V-Sweep
    end if
    switcher ← switcher + 1
end while
```

Figure 7: Full Sweep



(a) The description.

(b) Fully solved after applying *H-Sweep*.

(c) Partially solved after applying *V-Sweep*.

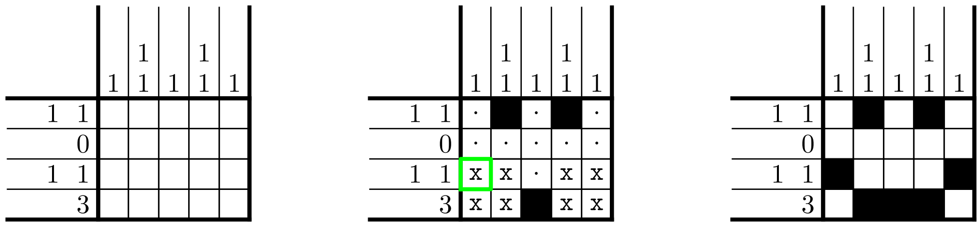
Figure 8: Differences in difficulty.

3.3 Probing

In the previous section, we considered solving Nonograms using a *Simple* approach. Some Nonograms cannot be fully solved using only this method. We consider these Nonograms to be *Non-Simple*. When *Simple* has been applied to a Nonogram and the result still contains at least one "x", the result is considered to be a partial solution. This partial solution might be further solvable. To see if this is applicable, we start using a *Probing* approach.

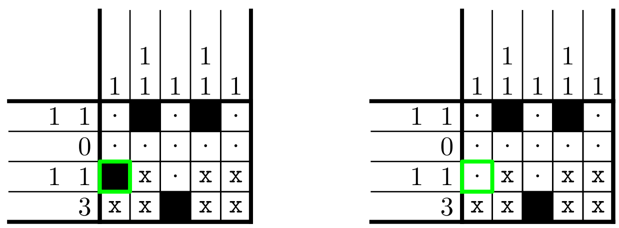
An example of a Nonogram, where probing needs to be applied, is provided in Figure 9. After applying the *Simple* approach, there are two clusters of unknown cells. When probing the marked cell in Figure 9b by filling in the cell, we get Figure 10a. Filling an unknown cell and performing *Simple* on the now new field is known as performing a *1-probe*.

After applying *Simple* to this new field, a fully solved Nonogram is obtained. There is no certainty that the fully solved Nonogram is the only solution of this Nonogram. After performing a 1-probe, the probed cell will be left blank and *Simple* is performed on this now, again, new field. Leaving the cell blank and performing *Simple* is known as a *0-probe*. Figure 10b shows a 0-probe by leaving the marked cell blank. Proceeding with this now new field, leads to a contradiction, as can be seen in Figure 11. After performing the first *H-Sweep* no contradictions occur. When performing the *V-Sweep*, the first column can be solved by filling the bottom left cell. The second column has already been partially filled due to previous *sweeps*. Since the column contains two separate ones, which matches the description, the latter cell must be left blank. The other columns are not



(a) The description. (b) After applying *Simple*. (c) After applying *Probing*.

Figure 9: Progress of probing a Nonogram.



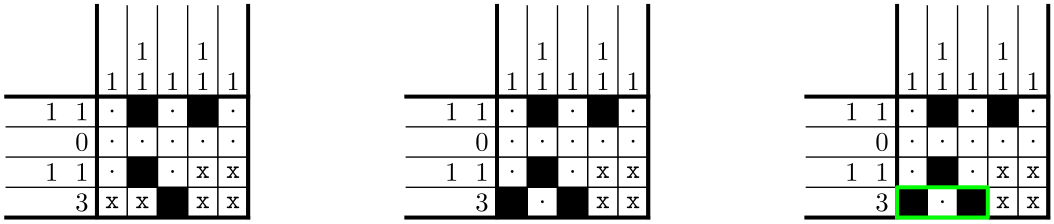
(a) A 1-probe. (b) A 0-probe.

Figure 10: A double probe.

considered, as no progression had been made in this area in the previous *sweep*. When performing the second *H-sweep*, we will only consider the last row. The last row does not match its description and thus a contradiction is obtained. Because of this, with certainty it can be said that the marked cell in Figure 10 should be filled.

As seen in Figures 9 and 10, all unknown cells are considered again for the *Probing* approach. These are marked and a 1- and 0-probe are tried for each. Performing a 1- and 0-probe on a single cell is known as a *double probe*. A *singleton probe* is a double probe that is applied to one pixel. The 1- and 0-probe are singleton probes. Probing different cells might provide different outcomes. A probe that results in a fully solved Nonogram is called a *solving probe*. A *uniquely solving probe* is a probe that results in a fully solved Nonogram where it is proven that the provided solution is the only possible solution. A *contradicting probe* is a probe that results in a contradicting result. By obtaining a contradicting probe, information about the cell is obtained with certainty, as we fill this cell with the opposite of what caused the contradiction. Figure 11 shows such a contradicting probe. A probe that does not result in a contradiction, is called a *compliant probe*. The probe may lead to a correct partial solution of the Nonogram. However, we cannot conclude with certainty that this is the only solution. With a *Set-Probing*, multiple cells are double probed at the same time.

For the Nonogram used in Figures 9, 10 and 11, probing any yet unknown cell will result in one and the same solution. These are all uniquely solving double probes.



(a) *H-Sweep.* **(b)** *V-Sweep.* **(c)** *H-Sweep.*

Figure 11: Contradicting probe.

3.4 The usefulness of a probe

Not all probes result in a solution. The gain of each probe can be classified as:

0. No progression
1. Some progression
2. Contradicting
3. Fully solving

Based upon these four categories, there are sixteen possible probes. The probe “3 2” refers to a 1-probe that results in a fully solved Nonogram and a 0-probe that results in a contradiction. The probes resulting in “2 3” or “3 2” lead to a uniquely solvable Nonogram. Probes with “3 3” indicate multiple solutions (in fact two) as both a 0-probe and 1-probe lead to a solution. The probe “2 2” states that both probes result in a contradiction. When this occurs, the Nonogram cannot have a valid solution. Probes where the 1- or 0-probe leads to 0 do not provide any information about the final state of the Nonogram. For example, having a probe that results in “0 3” means that the 0-probe leads to a fully solved Nonogram and no progression has been made by applying a 1-probe. This does not mean that this cell cannot be colored in a later stage, causing multiple solutions to be correct. So, even though the 1- or 0-probe can fully solve the Nonogram, this does not mean that this is the only solution. Furthermore, a probe that leads to no progression might be a correct filling for the cell in a later stage as well. The possible probing combinations are categorized as shown in Table 1. The category of a probe is determined by applying the Quaternary number system, using $category = 4 \cdot i_1 + i_0$, where i_1 represents the usefulness of the 1-probe and i_0 the usefulness of the 0-probe.

$0\,0^{0}$	$0\,1^{1}$	$0\,2^{2}$	$0\,3^{3}$
$1\,0^{4}$	$1\,1^{5}$	$1\,2^{6}$	$1\,3^{7}$
$2\,0^{8}$	$2\,1^{9}$	$2\,2^{10}$	$2\,3^{11}$
$3\,0^{12}$	$3\,1^{13}$	$3\,2^{14}$	$3\,3^{15}$

Table 1: The different probing combinations categorized.

Since *Probing* is applied to all unknown cells, we can determine the usefulness of each probed unknown cell. Combined, these form the *probing sequence* of a Nonogram. The probing sequence considers the rows of the Nonogram top to bottom. Per row, the unknown cells are considered from left to right.

As stated earlier, probing any cell of the example used in Figure 9 will lead to a unique solution. The corresponding probing sequence S with $|S| = 8$ is:

$$14 \;\; 11 \;\; 11 \;\; 14 \;\; 11 \;\; 14 \;\; 14 \;\; 11$$

Since this probing sequence only consists of category 11 and 14 probes, which are uniquely solving probes, double probing any cell will lead to the one possible solution.

The earlier mentioned occurrences “2 3” and “3 2” are categorized as 11 and 14, which represent a uniquely solving probe. Category 15 results in two different solutions. A probing sequence containing a category 15 and category 11 or 14 cannot occur, as this would lead to a contradicting probing sequence. The category 15 implies having exactly two solutions, whereas the category 11 and 14 imply having a unique solution. Category 10 indicates that no solution is possible. This category only occurs when the provided description causes a contradiction for both the 1- and 0-probe, which cannot happen for a correct/solvable description.

Category 0 leads with both a 1- and 0-probe to no progression. For 4×4 Nonograms, the possible line descriptions are (0), (1), (2), (3), (4), (1,1), (1,2) and (2,1). When considering a single line of such a Nonogram, performing a 1-probe on any of the unknown cells, will always lead to some new information. We assume a line where no cells have been filled or left blank. The descriptions (0), (4), (1,2) and (2,1) all lead to a fully solved line. When the description of the line is (3), applying a 1-probe on any of the cells leads to coloring at least one adjacent cell. A 1-probe for the line description (2) causes either a fully solved line, when coloring an edge cell, or a cell to be left blank, and thus some progression. As for descriptions solely consisting of ones, descriptions (1) and (1,1), coloring any cell of the line, causes the adjacent cells to be left blank. Here, again, we have some progression. When the line is partially solved, some 1-probes lead to a contradiction. The contradiction itself is new information. All possible descriptions on a 4×4 Nonogram lead to some progression using a 1-probe. Since the 1-probe leads to some progression, fully solving the line or a contradiction, category 0 cannot occur in the probing sequence of a 4×4 Nonogram.

The possible line descriptions on a 5×5 Nonogram are (0), (1), (2), (3), (4), (5), (1,1), (1,2), (1,3), (2,1), (2,2), (3,1) and (1,1,1). Here, the same reasoning as for the 4×4 Nonograms applies. Again, we assume a line where no cells have been filled or left blank. The descriptions (0), (5), (1,3), (2,2), (3,1) and (1,1,1) lead to a full solution when performing a 1-probe on any of the cells. For the description (4), a 1-probe on the edges leads to solving the line. A 1-probe on the other cells leads to coloring two other cells. Unlike on a 4×4 Nonogram, not all descriptions have some sort of progression when performing a 1-probe. When performing a 1-probe on a line of length 5 with description (3), the middle cell cannot tell anything about the other cells. Applying a 0-probe on the middle cell leads to a contradiction, and thus provides us with some information. For descriptions (1,2) and (2,1) performing a 1-probe does not give new information for all cells. For description (2,1), performing a 1-probe on the second cell of the line leads to no new information. Due to symmetry, for description (1,2) this is the case for the fourth cell. Performing a 0-probe for these cells, leads to a contradiction. For descriptions (1) and (1,1), the adjacent cells of the 1-probed cell can be left blank. For description (2), 1-probing an edge cell leads to a line solution. 1-probing the second or fourth cell leave a cell blank. 1-probing one of the edges leads to a fully solved line. Lastly, when 1-probing the middle cell, the edges can be left blank. When the line is partially solved, the probes all provide some progression, full progression or a contradiction. For all possible descriptions, there will be new information obtained. Therefore, category 0 probes cannot occur on a 5×5 Nonogram. The same reasoning applies to singly empty lines of length 6 and 7.

For probing sequences S with $|S| \geq 8$, category 0 probes may occur. An example of a single line where category 0 probes occur can be seen in Figure 12. Here, we have a single line of length 8 with description (1,2). The corresponding probing sequence is

$$4\ 4\ 4\ 0\ 0\ 4\ 1\ 4$$

Probing on the fourth or fifth cell does not provide any progression. Due to symmetry, the description (2,1) on a line of length 8 will contain category 0 probes as well.

(a) The 1-probe. (b) The 0-probe.

Figure 12: The occurrence of a category 0 probe.

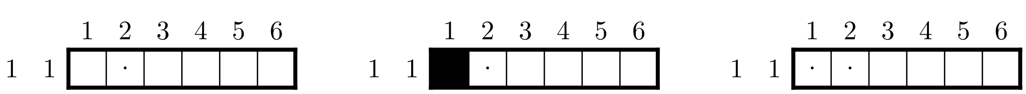
It is important to note, that it is possible to have category 0 probes on non-empty lines of length 6. So, it is possible to have category 0 probes when considering 6×6 Nonograms. Figure 13 shows an example of a non-empty line of length 6 with the corresponding probing sequence

$$0\ 4\ 4\ 4\ 4$$

This line occurs in the Nonogram shown in Figure 14, with a probing sequence equal to

$$6\ 5\ 4\ 9\ 9\ 6\ 9\ 6\ 12\ 5\ 4\ 8\ 4\ 5\ 12\ 5\ 8\ 0\ 7\ 8\ 5\ 8$$

Double probing the marked cell in Figure 14 classifies as a category 0 probe.



(a) The non-empty line. (b) The 1-probe. (c) The 0-probe.

Figure 13: The occurrence of a category 0 probe.

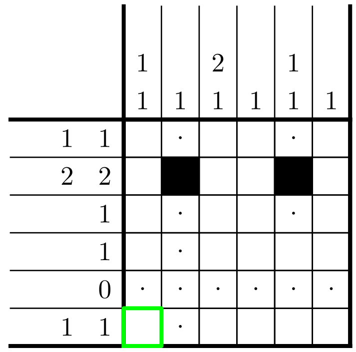


Figure 14: A 6×6 Nonogram where a category 0 probe occurs.

When considering a description d of length 1, and thus a description that consists of a single series, it can be observed that there will always be some sort of progression. Such a description on an empty line of arbitrary length ℓ, having a description for which $d = d_1 > \lfloor \frac{\ell}{2} \rfloor$, can be partially or fully, when the length is equal to ℓ, filled using *Simple*. These lines will not be empty when *Probing* is applied. When applying a 1-probe on any cell of these lines, we can color the cells between the already filled cell(s) and the probed cell, as the line description consists of only one number. The 1-probe will then lead to a fully solved line, a partially solved line or, when for example the size does not match the description, a contradiction. Applying a 0-probe allows us to leave the cells between the probed cell and the edge opposite to the filled cell(s) blank. So, 0-probing an edge cell causes the adjacent cell on the opposing side of the already filled cell(s) to be filled. For empty line descriptions, an empty line of arbitrary length, having a description for which $d = d_1 \leq \lfloor \frac{\ell}{2} \rfloor$, *Probing* needs to be applied. Applying a 1-probe on any of the cells places the filled series around the probed cell. Cells more than d_1 positions away, are left blank. So, 1-probing will give progression. A 0-probe allows us to leave the cells between the probed cell and the edge opposite to the probed cell blank. Thus, 0-probing an edge cell may not give any progression. However, 1-probing an edge will result in either some progression, full progression or a contradiction, so probing edge cells will lead to progression.

3.5 Probing sequence of a Nonogram

For this section, all 4×4 and 5×5 Nonograms have been studied. Considering "all" relates to solvable Nonograms, meaning no category 10 probes will occur in its probing sequence. Due to the large number of possible 6×6 Nonograms, not all have been considered. Out of all possible 6×6 Nonograms, $30{,}000{,}000{,}000$ have been classified. For all sizes, we will observe Nonograms that are *Non-Simple*. These observations are based on analysis of the performed experiments.

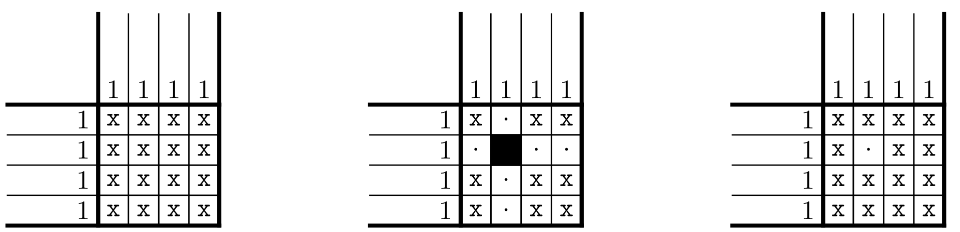
3.5.1 Nonograms of size 4 by 4

For 4×4 Nonograms, the probing sequence S of a *Non-Simple* puzzle has a minimal length of 4 and a maximal length of 16, so $4 \leq |S| \leq 16$. When a probing sequence contains a category 11 and 14 probe, then only categories 6, 8, 9 and 12 can occur. All probing sequences of uniquely solvable Nonograms contain at least one category 11 and one category 14 probe. The number of occurrences of category 11 and category 14 probes do not have to be equal. Having a category 15 probe in the probing sequence leads to exactly two distinctive solutions. The number of occurrences of category 15 does not influence the number of solutions. For the 4×4 Nonograms that have exactly two solutions, the number of occurrences of the category 15 probes falls between 2 and 16. The probing sequences containing a category 15 probe may contain a combination of probes of category 6, 8, 9 and 12. Figure 20 shows an example of a Nonogram with such a probing sequence. The corresponding probing sequence is

$$8 \;\; 9 \;\; 14 \;\; 14 \;\; 11 \;\; 14 \;\; 11 \;\; 9 \;\; 11 \;\; 6 \;\; 14$$

Having four occurrences of category 15 probes, "3 3", in the probing sequences identifies a switching component. Applying *Probing* and other solving methods will not affect the outcome, as there will be exactly two solutions. For a probing sequence consisting solely of category 15 probes, $|S| \geq 4$. The smallest occurring probing sequence that implies exactly two solutions are those who contain exactly one switching component, where $|S| = 4$. Besides the category 15 probes, probings of category 6, 8, 9 and/or 12 may occur as well.

When the probing sequence consists of all category 4 probes, the Nonogram will have multiple solutions. For example, Figure 15 describes such a case. Here, we have a Nonogram where all cells are unknown. Probing any cell results in some progress when applying a 1-probe and no progress when applying a 0-probe. An example of this can be seen in Figure 15b and 15c. Probing once will not solve this Nonogram. It does have multiple solutions, which can be obtained by probing four times total. For probing sequences of Nonograms with multiple solutions, $7 \leq |S| \leq 16$. Probing sequences of an uneven length may occur. As can be seen in Table 6, probing sequences of length 7, 9 and 11 occur. These probing sequences may contain probes of categories 4, 5, 6, 7, 8, 9, 12, 13. Categories 0, 1, 2 and 3 do not occur due to the fact that these Nonograms are too small for no progression to happen.



(a) A Nonogram after applying *Simple*.

(b) A partially solving 1-probe.

(c) A no progression 0-probe.

Figure 15: Example of a Nonogram with only category 4 probes.

When only categories 5 and 8 occur, *Simple* and *Probing* will not solve the Nonogram. An example of such a Nonogram description is shown in Figure 16. After applying *Simple*, only the last row description has given some progression, as the second and third cell can be filled. The corresponding probing sequence is

$$8 \;\; 5 \;\; 5 \;\; 8 \;\; 8 \;\; 5 \;\; 5 \;\; 8 \;\; 5 \;\; 5 \;\; 5 \;\; 5$$

Probing does not provide fully solved Nonograms, as probing these cells only leads to contradictions or partially solved Nonograms. This example could be solved using a *Set-Probing* approach.

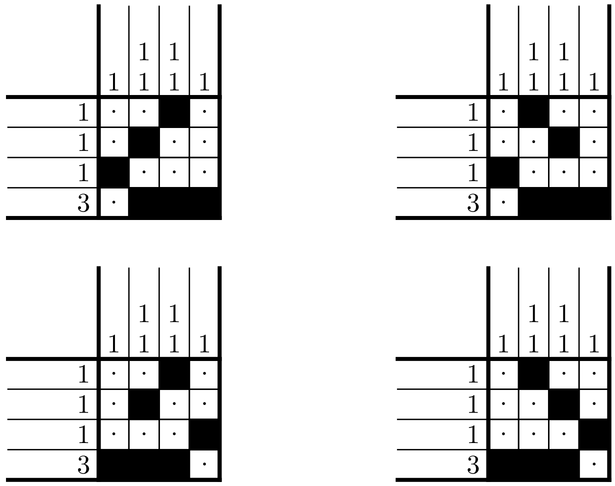


Figure 16: The solutions of a Nonogram with a probing sequence consisting solely of category 5 and 8 probes.

3.5.2 Nonograms of size 5 by 5

For 5×5 Nonograms, the length of the probing sequences S has a minimal value of 4 and a maximal value of 25, thus $4 \leq |S| \leq 25$. For Nonograms with exactly two solutions, the probing sequence has a minimal length of 4, due to the presence of a switching component. The Nonograms with multiple solutions have a probing sequence with a minimal length of 6. For all, the maximum length of the probing sequence is equal to 25.

When a probing sequence contains a category 11 and 14 probe, then categories 1, 2, 3, 4, 5, 6, 7, 8, 9, 12 and 13 can occur. Comparing these occurrences to those of the uniquely solvable 4×4 Nonograms, no clear pattern arises. Again, the number of category 11 and 14 probes do not have to be equal and both occur at least once in the probing sequence. For Nonograms with exactly two solutions, all probing sequences have at least one occurrence of a category 15 probe and at most 18 occurrences. Besides the category 15 probes, probes of categories 1, 2, 3, 4, 5, 6, 7, 8, 9, 12 and 13 may also occur.

Nonograms that have multiple solutions have $6 \leq |S| \leq 25$. These Nonograms may have a probing sequence of uneven length. Just as for the 4×4 Nonograms, having a probing sequence that consists solely out of category 5 and 8 probes results in multiple solutions which cannot simply be found using *Simple* and *Probing*. Again, probing sequences that contain solely category 4 probes, have multiple solutions. For 4×4 Nonograms with multiple solutions, category 0, 1, 2 and 3 did not occur in any of the corresponding probing sequences. Category 0 probes did not occur. For the 5×5 Nonograms, categories 1, 2 and 3 may occur. An example of such a Nonogram can be seen in Figure 17. Here, we have solely category 1 probes. Since all double probes lead to some progression at best, Nonograms such as seen in Figure 17, cannot be solved by using *Simple* and *Probing*.

	1 1	0	1 1	0	1 1
1 1		·		·	
0	·	·	·	·	·
1 1		·		·	
0	·	·	·	·	·
1 1		·		·	

Figure 17: A probing sequence for a 5×5 Nonogram consisting solely of category 1 probes.

3.5.3 Nonograms of size 6 by 6

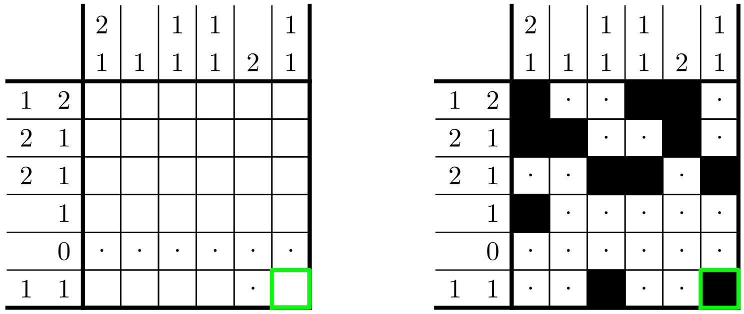
Of all 6×6 Nonograms, $30,000,000,000$ have been classified. These have been generated starting with an empty Nonogram followed by all Nonograms with one filled cell, expanding the number of filled cells. It is to note that the following observations are not based upon all 6×6 Nonograms, but just a part of these: in fact, 44% of all 6×6 Nonograms have been examined.

For 6×6 *Non-Simple* Nonograms, the length of the probing sequences S has a minimal value of 4 and a maximal value of 36, thus $4 \leq |S| \leq 36$. Just like for probing sequences of 4×4 and 5×5 Nonograms, uniquely solvable probing sequences of 6×6 Nonograms have a minimal length of 8. Again, because of the switching component, the probing sequence of Nonograms with exactly two solutions have a minimal length of 4. Just like probing sequences of 5×5 puzzles, Nonograms with multiple solutions have probing sequences with a minimal length of 6.

Just as for the 5×5 Nonograms, for probing sequences that contain a category 11 and 14 probe, the categories 1, 2, 3, 4, 5, 6, 7, 8, 9, 12 and 13 may occur. Again, the numbers of category 11 and 14 probes do not have to be equal and both categories occur at least once in the probing sequence. There are uniquely solvable 6×6 Nonograms where probing sequences contain a category 0 probe. An example of such a Nonogram with a probing sequence contain a category 0 probe, can be seen in Figure 18, where the category 0 probe cell is marked green. The corresponding probing sequence is:

$$14\ \ 8\ \ 11\ \ 14\ \ 14\ \ 11\ \ 14\ \ 14\ \ 11\ \ 11\ \ 14\ \ 11\ \ 11\ \ 11\ \ 14\ \ 14\ \ 11\ \ 14\ \ 14\ \ 8\ \ 8\ \ 8\ \ 11\ \ 8\ \ 7\ \ 8\ \ 4\ \ 4\ \ 0$$

For Nonograms with exactly two solutions, all probing sequences have at least one occurrence of a category 15 probe. Probes of categories 1, 2, 3, 4, 5, 6, 7, 8, 9, 12 and 13 may also occur. Nonograms that have multiple solutions have $6 \leq |S| \leq 36$. Again, these Nonograms may have a probing sequence of uneven length. Just as for the 4×4 and 5×5 Nonograms, probes of categories 1, 2, 3, 4, 5, 6, 7, 8, 9, 12 and 13 may occur. Different from 4×4 and 5×5 Nonograms, there are probing sequences where category 0 probes occur, as we have seen in Figure 14. Again, probing sequences that contain solely category 4 probes, have multiple solutions.



(a) Partially solved Nonogram.

(b) Fully solved Nonogram after *Probing*.

Figure 18: A uniquely solvable Nonogram with a probing sequence containing a category 0 probe.

3.6 Difficulty

The difficulty of a Nonogram partly determines how hard a Nonogram is to solve. The difficulty of the Nonograms can mostly be classified as *Simple* or *Difficult*, see [BK12]. The difficulty of simple Nonograms is based on the number of switches between performing *H-* and *V-Sweeps*, which is classified to have a *1-star* difficulty, as can be seen in Table 2.

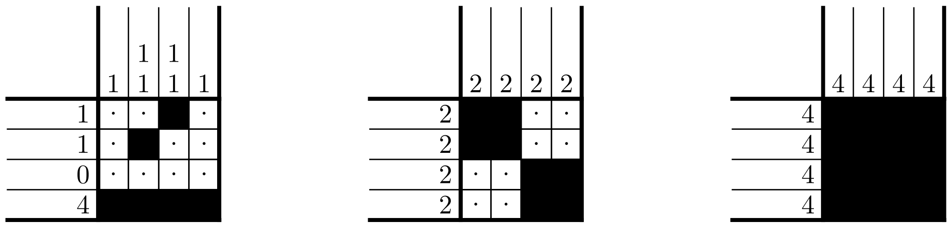
Table 2 shows an overview of the difficulty categories. When a Nonogram is not solvable using only *Simple*, we apply *Probing*. The switching of strategy and performing a more costly approach causes the difficulty of the Nonogram to increase. *Probing* may lead to uniquely solvable Nonograms, Nonograms with exactly two solutions or Nonograms with multiple solutions (> 2). When *Probing* leads to one of those cases, *Probing* once is enough to determine its difficulty. Nonograms that need to be solved by applying both *Simple* and *Probing* are classified to have a *2-star* difficulty. Its difficulty may be determined by the number of unknown cells on which we need to apply *Probing*. *Probing* one cell might provide more progression than the other, so usefulness of probes should also be taken into account. Nonograms for which only partial solutions can be obtained may be unsolvable, not unique or have multiple solutions. While solving most of the uniquely solvable Nonograms using solely *Simple* and *Probing*, a number of Nonograms need a *Set-Probing* approach to be solved. This approach would be more costly than applying *Simple* and *Probing*, and thus would have a higher difficulty. When *Simple* and *Probing* leaves a Nonogram unsolved, then the difficulty is categorized as a *3-star* difficulty, see Table 2.

Difficulty	**Solved by**	**Specific difficulty**
$\star$	Simple	Number of switches between performing *H-* and *V-Sweeps*
$\star\star$	Simple & Probing	Number of unknown cells & usefulness of the probes
$\star\star\star$	Simple & Probing, Unsolved	Number of unknown cells & usefulness of the probes & *Set-Probing*

Table 2: Difficulty classification.

4 Experiments

Most Nonograms can be mirrored and rotated. When a Nonogram can be mirrored and rotated, there are eight similar Nonograms in total. Figure 19a shows an example of a Nonogram where the rotations and mirroring result in 8 different Nonograms. However, this does not hold for all Nonograms. Not all Nonograms differ from their rotations or mirroring, see Figure 19b. Here, we see a Nonogram where rotating and mirroring lead to the same two possible Nonograms. It is also possible to have only one possible Nonogram, as can be seen in Figure 19c.



(a) 8 similar Nonograms in total.

(b) 2 similar Nonograms in total.

(c) Only one possible Nonogram.

Figure 19: Rotations and mirroring Nonograms.

The following experiments have been performed on all 4×4, 5×5 and 6×6 Nonograms. For a 4×4 Nonogram, we have a total number of 16 cells. Each of these cells can be filled or be left blank, meaning we have two options for each cell. In total, there are $2^{16} = 65,536$ possible Nonograms of size 4×4. For a 5×5 Nonogram, there are $2^{25} = 33,554,432$ possible Nonograms and for 6×6 Nonograms this comes down to $2^{36} = 68,719,476,736$ possible Nonograms. These Nonograms have been classified according to their probing sequences. The classification can be seen in Table 3. As can be seen in Table 3, $30,000,000,000$ out of $68,719,476,736$ 6×6 Nonograms have been classified. The provided percentages are relative to the classified $30,000,000,000$. The ratio of *Simple* and *Probing* is shown in Table 3. Here, *Probing* is split into three categories: probing that leads to a uniquely solved Nonogram, probing that leads to exactly two solutions and the remaining Nonograms. The remaining category contains Nonograms with multiple solutions and unsolved uniquely solvable Nonograms. The uniquely solvable Nonograms can be distinguished by checking whether their description is unique.

4.1 Experiments on 4 by 4 Nonograms

Tables 4, 5 and 6 describe the probing sequences S on all 4×4 *Non-Simple* Nonograms. These tables describe the probing sequence after applying *Probing* once. *Simple* did not fully solve these Nonograms. The Nonograms that were fully solved using *Simple* are thus not shown in these tables. The probing sequences in Tables 4, 5 and 6 are sorted in lexicographic order. The actual probing sequences may vary in order.

For 4×4 Nonograms, a uniquely solvable Nonogram contains at least three category 11 probes and four category 14 probes, see Table 4. When considering the possible probing sequences for uniquely solvable Nonograms, it might seem that only categories 6, 8, 9 and 12 occur, besides the

Size	Simple	Probing Unique	Probing Exactly Two	Probing Rest	Total
4×4	51,234	1,128	10,100	3,074	65,536
	(78.18%)	(1.72%)	(15.41%)	(4.69%)	
5×5	24,976,511	330,602	5,526,788	2,720,531	33,554,432
	(74.44%)	(0.99%)	(16.47%)	(8.11%)	
6×6	21,590,009,602	507,898,588	3.208.556.453	4,693,535,357	30,000,000,000 of
	(71,97%)	(1,69%)	(10,70%)	(15,65%)	68,719,476,736

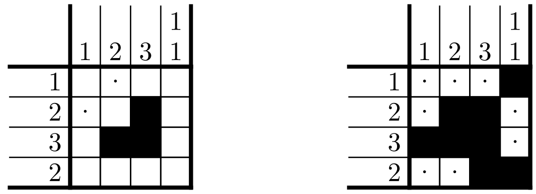
Table 3: Probing classification for different sizes.

Length	Probing sequence	Occurrences
8	11 11 11 11 14 14 14 14	348
10	11 11 11 11 11 14 14 14 14 14	8
11	6 8 9 9 11 11 11 14 14 14 14	8
12	8 11 11 11 11 11 11 14 14 14 14 14	36
12	11 11 11 11 11 11 14 14 14 14 14 14	48
16	6 6 6 8 9 9 9 9 9 11 11 11 14 14 14 14	32
16	6 8 9 9 11 11 11 11 11 11 14 14 14 14 14 14	128
16	6 9 9 11 11 11 11 11 11 11 12 14 14 14 14 14	32
16	8 8 8 8 11 11 11 11 11 11 14 14 14 14 14 14	128
16	8 11 11 11 11 11 11 11 11 14 14 14 14 14 14 14	288
16	11 11 11 11 11 11 11 11 14 14 14 14 14 14 14 14	72
		1,128

Table 4: Probing sequences of all uniquely solvable *Non-Simple* 4×4 Nonograms after *Probing*.

distinguishable categories 11 and 14. However, this pattern disappears when considering larger Nonograms. Besides the required categories 11 and 14, 5×5 Nonograms may contain probes of category 1, 2, 3, 4, 5, 6, 7, 8, 9, 12 and/or 13. For uniquely solvable 5×5 Nonograms, the probing sequences have at least one category 11 probe and at least one category 14 probe. Larger Nonograms may thus contain fewer category 11 and category 14 probes. Something similar holds for Nonograms with multiple or exactly two solutions, as the patterns disappear when considering larger Nonograms. Nonograms with exactly two solutions have probing sequences that require at least one category 15 probe and do not contain any probe of category 0, 10, 11 or 14.

Table 4 shows that there are uniquely solvable *Non-Simple* Nonograms with probing sequences of length 11. This sequence occurs eight times in total. An example of such a Nonogram can be seen in Figure 20. Figure 20a shows the partially solved Nonogram for which Figure 20b is the unique solution. The remaining seven uniquely solvable *Non-Simple* Nonograms with a probing sequence of length 11 are rotations and mirrorings of the provided example. When looking at the number of occurrences of each probing sequence in Tables 4, 5 and 6, we see that most of these sequences occur a multiple of eight times. As described earlier, Nonograms can be mirrored and rotated in such a way that there are eight similar, yet distinctively different, Nonograms in total.



(a) Partially solved Nonogram after applying *Simple*.

(b) Fully solved Nonogram after *Probing*.

Figure 20: A uniquely solvable Nonogram with a probing sequence of uneven length.

Length	Probing sequence	Occurrences
4	15 15 15 15	8,296
6	15 15 15 15 15 15	120
8	15 15 15 15 15 15 15 15	696
10	6 6 9 9 15 15 15 15 15 15	8
10	8 8 15 15 15 15 15 15 15 15	48
10	15 15 15 15 15 15 15 15 15 15	8
11	8 15 15 15 15 15 15 15 15 15 15	8
12	6 6 6 6 9 9 9 9 15 15 15 15	16
12	6 6 6 8 9 9 9 9 15 15 15 15	32
12	6 6 6 9 9 9 15 15 15 15 15 15	8
12	6 8 9 9 15 15 15 15 15 15 15 15	48
12	15 15 15 15 15 15 15 15 15 15 15 15	72
16	6 8 9 9 15 15 15 15 15 15 15 15 15 15 15 15	32
16	6 6 6 6 6 8 9 9 9 9 9 9 15 15 15 15	128
16	6 6 6 6 6 9 9 9 9 9 9 9 12 15 15 15	32
16	6 6 6 6 8 8 8 8 9 9 9 9 15 15 15 15	96
16	6 6 6 6 8 8 8 9 9 9 9 9 12 15 15 15	64
16	6 6 6 6 8 8 9 9 9 9 9 9 12 12 15 15	128
16	6 6 6 8 8 8 9 9 9 9 12 15 15 15 15 15	32
16	6 6 8 8 9 9 9 9 12 12 15 15 15 15 15 15	32
16	6 6 8 9 9 9 15 15 15 15 15 15 15 15 15 15	32
16	8 8 8 8 15 15 15 15 15 15 15 15 15 15 15 15	152
16	15 15 15 15 15 15 15 15 15 15 15 15 15 15 15 15	12
		10,100

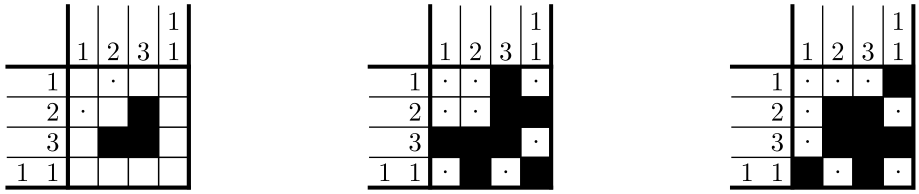
Table 5: Probing sequences with two solutions for all 4×4 Nonograms.

In Table 5, a probing sequence S containing sixteen category 15 probes can be observed. This probing sequence has a maximal length for the 4×4 Nonogram. Probing any cell results in two solutions. Since the descriptions of probing sequences that contain a category 15 lead to exactly two solutions, one description leads to two different Nonograms. The twelve Nonograms represented in Table 5 by the probing sequence of length 16 that consists solely of category 15 probes, have in total six different descriptions, which are:

- $D_{\text{rows}} = \langle(2),(2),(2),(2)\rangle$ and $D_{\text{columns}} = \langle(2),(2),(2),(2)\rangle$
- $D_{\text{rows}} = \langle(2),(1,1),(1,1),(2)\rangle$ and $D_{\text{columns}} = \langle(1,1),(2),(2),(1,1)\rangle$
- $D_{\text{rows}} = \langle(2),(2),(2),(2)\rangle$ and $D_{\text{columns}} = \langle(1,1),(1,1),(1,1),(1,1)\rangle$
- $D_{\text{rows}} = \langle(1,1),(2),(2),(1,1)\rangle$ and $D_{\text{columns}} = \langle(2),(1,1),(1,1),(2)\rangle$
- $D_{\text{rows}} = \langle(1,1),(1,1),(1,1),(1,1)\rangle$ and $D_{\text{columns}} = \langle(2),(2),(2),(2)\rangle$

Length	**Probing sequence**	**Occurrences**
7	7 7 12 13 13 13 13	276
8	5 5 5 5 12 12 12 12	208
8	7 7 7 7 13 13 13 13	72
9	4 4 4 4 4 4 4 4 4	96
10	5 5 7 7 12 12 13 13 13 13	96
10	6 7 7 9 9 12 12 13 13 13	288
10	7 7 7 7 8 8 13 13 13 13	288
10	7 7 7 12 13 13 13 13 13 13	16
11	4 7 7 7 7 13 13 13 13 13 13	16
11	5 5 5 5 7 12 12 12 12 13 13	20
12	4 4 4 4 5 5 5 5 12 12 12 12	288
12	5 5 5 5 5 5 5 5 8 8 8 8	36
12	6 6 7 7 9 9 9 12 13 13 13 13	48
12	6 7 7 7 7 8 9 9 13 13 13 13	96
12	6 7 7 7 9 9 12 13 13 13 13 13	48
12	7 7 7 7 12 13 13 13 13 13 13 13	16
16	4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	24
16	4 4 4 4 4 4 5 5 5 5 7 12 12 12 13 13	144
16	4 4 4 4 5 5 5 5 5 5 5 5 5 5 5 5	32
16	4 4 4 5 5 5 5 5 7 8 8 12 12 12 12 13	192
16	4 4 6 6 7 7 9 9 9 9 12 12 13 13 13 13	64
16	4 5 5 5 5 5 5 5 8 8 12 12 12 12 12 12	96
16	4 5 5 5 5 6 7 7 8 8 9 9 12 13 13 13	128
16	4 6 6 6 7 9 9 9 9 9 9 12 12 12 13 13	48
16	4 6 6 7 7 7 7 9 9 9 13 13 13 13 13 13	64
16	4 6 7 7 7 7 8 8 9 9 12 13 13 13 13 13	128
16	5 5 5 5 5 5 7 7 8 8 12 12 13 13 13 13	64
16	5 5 5 5 6 6 7 9 9 9 12 12 12 12 13 13	80
16	6 6 6 7 7 7 9 9 9 9 12 13 13 13 13 13	48
16	6 6 6 7 7 8 8 8 9 9 9 12 13 13 13 13	48
16	7 7 7 7 7 7 7 7 13 13 13 13 13 13 13 13	6
		3,074

Table 6: Probing sequences with multiple solutions for all 4×4 Nonograms.

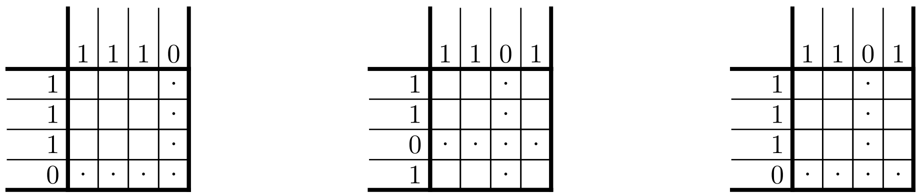


(a) Partially solved Nonogram.

(b) A possible solution for the provided description.

(c) A possible solution for the provided description.

Figure 21: Another probing sequence of uneven length for a Nonogram with exactly two solutions.



(a) A Nonogram that can be rotated into four different Nonograms.

(b) A Nonogram that can be rotated into four different Nonograms.

(c) A Nonogram that can be rotated and mirrored into eight different Nonograms.

Figure 22: Three Nonograms that represent all 4×4 Nonograms with a probing sequence of length 9 leading to multiple solutions.

- $D_{\text{rows}} = \langle(1,1),(1,1),(1,1),(1,1)\rangle$ and $D_{\text{columns}} = \langle(1,1),(1,1),(1,1),(1,1)\rangle$

Just like the uniquely solvable *Non-Simple* 4×4 Nonograms, Table 5 shows Nonograms with probing sequences where $|S| = 11$. There are in total eight Nonograms with this uneven length. Two of these are shown in Figure 21b and 21c, where Figure 21a shows the partially solved Nonogram after applying *Simple* with its descriptions. There are in total four different descriptions, as each description leads to exactly two possible solutions. The remaining Nonograms, for which $|S| = 11$ and exactly two solutions exist, are rotations of the provided example.

Different from the uniquely solvable *Non-Simple* Nonograms and Nonograms that have exactly two solutions, the remaining Nonograms, with probing sequences of uneven length, cannot simply be rotated or mirrored into other Nonograms. Table 6 shows that there are numerous Nonograms with $|S| = 7$, $|S| = 9$ and $|S| = 11$. A probing sequence of length 9 only occurs when the probing sequence consists of solely category 4 probes, which happens 98 times. Figure 22 shows three Nonograms that represent all possible Nonograms with a probing sequence of length 9 that consist solely of category 4 probes. The Nonograms in Figure 22a and 22b each represent four Nonograms in total, due to rotations. Each of these rotations has six possible solutions. Figure 22c represents eight Nonograms with each six solutions. In total, $4 \times 6 + 4 \times 4 + 8 \times 6 = 96$ solutions with a probing sequence of length 9 solely consist of category 4 probes, which is supported by the results in Table 6.

Table 6 shows all remaining Nonograms. Due to the manner of classification, it is possible that this category contains some uniquely solvable Nonograms. When considering 4×4 Nonograms, we only have Nonograms with more than two solutions within this category. Multiple Nonograms lead to the same description, meaning that the description leads to multiple solutions as well. All 3,074 Nonograms of this category have more than two solutions, and thus no uniquely solvable Nonograms are being found by applying *Simple* and *Probing*.

4.2 Experiments on 5 by 5 Nonograms

The probing sequences of uniquely solvable Nonograms have a minimal length of 8. As can be seen in Table 3, there are 330,602 uniquely solvable 5×5 Nonograms that can be found after applying *Probing*. These 330,602 Nonograms lead to 3,665 different probing sequences. We will consider *pure* probing sequences, which are probing sequences that consist solely of the characterizing categories. Within this category of uniquely solvable Nonograms, pure probing sequences consist solely of category 11 and category 14 probes. In total, 13 out of the 3,665 different probing sequences are pure probing sequences, see Table 7. These pure probing sequences make up 35.8% of all probing sequences within this category. Probing any unknown cell of these Nonograms solves the Nonogram completely. Thus, these pure probing sequences indicate that all double probes are uniquely solving double probes. The other Nonograms of this category contain more varying probing sequences. An example is given in Figure 23, with a probing sequence of

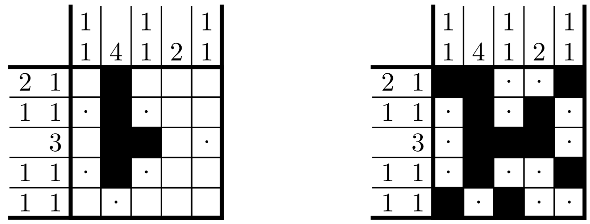
6 9 11 14 14 11 7 13 11 14 2 6 8 11

Length	**Probing sequence**	**Occurrences**
8	11 11 11 11 14 14 14 14	71,824
10	11 11 11 11 11 14 14 14 14 14	8,080
11	11 11 11 11 11 11 14 14 14 14 14	48
12	11 11 11 11 11 11 14 14 14 14 14 14	18,640
13	11 11 11 11 11 11 11 14 14 14 14 14 14	968
14	11 11 11 11 11 11 11 14 14 14 14 14 14 14	4,716
15	11 11 11 11 11 11 11 11 14 14 14 14 14 14 14	4,716
16	11 11 11 11 11 11 11 11 11 14 14 14 14 14 14 14	16
16	11 11 11 11 11 11 11 11 14 14 14 14 14 14 14 14	8,334
17	11 11 11 11 11 11 11 11 11 14 14 14 14 14 14 14 14	672
18	11 11 11 11 11 11 11 11 11 11 14 14 14 14 14 14 14 14	152
19	11 11 11 11 11 11 11 11 11 11 14 14 14 14 14 14 14 14 14	72
20	11 11 11 11 11 11 11 11 11 11 11 14 14 14 14 14 14 14 14 14	24
		118,262

Table 7: Pure probing sequences for uniquely solvable 5×5 Nonograms.

Considering the pure probing sequences for the category with exactly two solutions, we will consider probing sequences consisting solely of category 15 probes to be pure. Table 8 shows the pure probing sequences for this category. This category contains, in total, 3,583 different probing sequences. The 16 probing sequences of Table 8 make up 91.0% of all probing sequences within this category. An example of a Nonogram with the probing sequence

15 15 15 15



(a) A partially solved Nonogram after applying *Simple* and *Probing*.

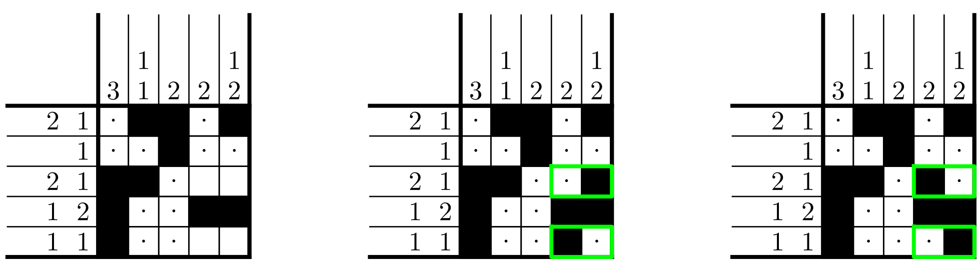
(b) The uniquely solvable Nonogram.

Figure 23: A uniquely solvable 5×5 Nonogram.

which can also be seen in Table 8, is shown in Figure 24. The probing sequences indicate a switching component, which is marked in Figures 24b and 24c.

Length	**Probing sequence**	**Occurrences**
4	15 15 15 15	4,363,030
6	15 15 15 15 15 15	189,940
8	15 15 15 15 15 15 15 15	407,528
10	15 15 15 15 15 15 15 15 15 15	13,328
12	15 15 15 15 15 15 15 15 15 15 15 15	43,340
14	15 15 15 15 15 15 15 15 15 15 15 15 15 15	808
16	15 15 15 15 15 15 15 15 15 15 15 15 15 15 15 15	10,678
		5,028,652

Table 8: Pure probing sequences for category 15 5×5 Nonograms.



(a) A partially solved Nonogram after applying *Simple* and *Probing*.

(b) A corresponding Nonogram.

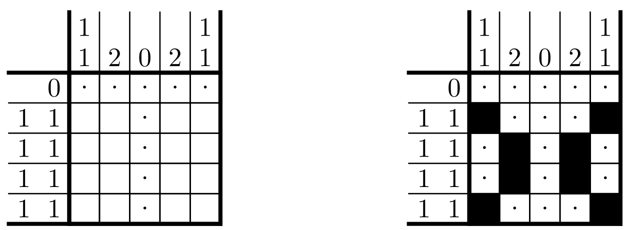
(c) Another corresponding Nonogram.

Figure 24: A 5×5 Nonogram with exactly two solutions.

For 5×5 Nonograms, the remaining category does contain uniquely solvable Nonograms. Uniquely solvable Nonograms within this remaining category can not be found solely using *Simple* and *Probing*. The 2,720,531 Nonograms of the remaining category make up 15,586 different probing sequences. There are 2,462 uniquely solvable Nonograms in the remaining category, which make up

282 different probing sequences. Out of these 282 different probing sequences, there are 40 probing sequences that represent both uniquely solvable Nonograms and Nonograms with two or more solutions. The other probing sequences will only result in a uniquely solvable Nonogram, classified in the remaining category. Table 9 shows the 40 probing sequences of the 2,462 uniquely solvable Nonograms which have not been classified as such. The occurrence of each probing sequence for the uniquely solvable Nonograms are shown against the total number of occurrences of each probing sequence within the remaining category. In total, there are 340 uniquely solvable Nonograms which an ambiguous probing sequence. The total number of Nonograms that have an ambiguous probing sequence is 2,380. An example of a uniquely solvable Nonogram categorized in the remaining category is shown in Figure 25 with a probing sequence of

$$6\ 9\ 6\ 9\ 9\ 6\ 9\ 6\ 9\ 6\ 9\ 6\ 6\ 9\ 6\ 9$$



(a) A partially solved Nonogram after applying *Simple* and *Probing*.

(b) The uniquely solvable Nonogram.

Figure 25: A uniquely solvable Nonogram out of the remaining category.

The probing sequence

$$6\ 6\ 6\ 6\ 6\ 7\ 8\ 8\ 8\ 8\ 8\ 9\ 9\ 9\ 9\ 9\ 9\ 12\ 13\ 13$$

occurs 24 times within the remaining category and covers 8 different uniquely solvable Nonograms. The descriptions of these 8 different Nonograms are:

- $D_{\text{rows}} = \langle(2),(3),(1,1),(1),(2)\rangle$ and $D_{\text{columns}} = \langle((2),(2,1),(2),(2),(1)\rangle$
- $D_{\text{rows}} = \langle(1),(2),(2),(1,2),(2)\rangle$ and $D_{\text{columns}} = \langle((2),(1),(1,1),(3),(2)\rangle$
- $D_{\text{rows}} = \langle(2),(2,1),(2),(2),(1)\rangle$ and $D_{\text{columns}} = \langle((2),(3),(1,1),(1),(2)\rangle$
- $D_{\text{rows}} = \langle(2),(1,2),(2),(2),(1)\rangle$ and $D_{\text{columns}} = \langle((2),(1),(1,1),(3),(2)\rangle$
- $D_{\text{rows}} = \langle(2),(1),(1,1),(3),(2)\rangle$ and $D_{\text{columns}} = \langle((1),(2),(2),(1,2),(2)\rangle$
- $D_{\text{rows}} = \langle(2),(1),(1,1),(3),(2)\rangle$ and $D_{\text{columns}} = \langle((2),(1,2),(2),(2),(1)\rangle$
- $D_{\text{rows}} = \langle(2),(3),(1,1),(1),(2)\rangle$ and $D_{\text{columns}} = \langle((1),(2),(2),(2,1),(2)\rangle$
- $D_{\text{rows}} = \langle(1),(2),(2),(2,1),(2)\rangle$ and $D_{\text{columns}} = \langle((2),(3),(1,1),(1),(2)\rangle$

Length	Probing sequence	Occurrences
20	4 4 6 6 6 6 7 8 8 8 8 9 9 9 9 9 9 12 13 13	8/24
20	6 6 6 6 6 7 8 8 8 8 8 8 9 9 9 9 9 12 12 13	8/184
20	6 6 6 6 6 7 8 8 8 8 8 9 9 9 9 9 9 12 13 13	8/24
20	4 4 6 6 7 7 7 8 8 8 8 8 8 9 9 12 13 13 13 13	8/32
21	6 6 6 6 6 7 8 8 8 8 8 8 9 9 9 9 9 9 12 12 13	8/24
21	4 6 6 6 6 7 7 8 8 8 8 9 9 9 9 9 12 12 13 13 13	8/32
22	4 4 5 5 6 6 6 6 8 8 8 8 8 8 9 9 9 9 9 9 12 12	8/104
22	5 5 6 6 6 6 6 6 8 8 8 8 8 9 9 9 9 9 9 9 12 12	8/24
22	4 5 5 6 6 6 6 8 8 8 8 8 8 8 8 9 9 9 9 9 12 12	8/168
22	4 4 4 4 4 6 6 7 8 8 8 8 8 8 8 9 9 9 9 12 13 13	8/32
22	4 4 5 5 6 6 6 7 8 8 8 8 8 8 8 9 9 9 9 12 13 13	8/24
22	4 4 6 6 6 6 7 8 8 8 8 8 8 8 8 9 9 9 9 12 13 13	8/24
22	4 4 6 6 6 6 7 8 8 8 8 8 8 8 9 9 9 9 9 12 13 13	8/40
22	4 6 6 6 6 7 8 8 8 8 8 8 8 8 9 9 9 9 9 12 12 13	8/200
22	4 6 6 6 6 7 8 8 8 8 8 8 8 9 9 9 9 9 9 12 12 13	8/24
22	4 4 4 4 5 5 5 5 6 6 7 8 8 8 8 9 9 9 12 12 12 13	8/24
22	4 4 6 6 6 7 7 7 8 8 8 8 8 8 9 9 9 9 12 13 13 13	8/32
22	4 4 6 6 6 7 7 8 8 8 8 8 8 8 9 9 9 9 12 13 13 13	8/32
23	4 4 6 6 6 6 6 6 8 8 8 8 8 8 8 9 9 9 9 9 9 12 12	16/80
23	4 4 4 4 4 4 4 4 5 5 5 5 6 6 8 8 8 8 8 9 9 12 12	8/208
23	4 4 4 4 6 6 6 6 7 8 8 8 8 8 8 9 9 9 9 9 12 12 13	8/32
23	4 4 4 4 4 6 6 6 7 8 8 8 8 8 8 8 9 9 9 12 12 12 13	8/32
24	4 4 5 5 5 5 6 6 6 6 8 8 8 8 8 8 8 9 9 9 9 9 12 12	8/24
24	4 4 4 4 6 6 6 7 8 8 8 8 8 8 8 8 8 9 9 9 9 12 12 13	8/24
24	4 4 4 5 5 5 5 6 6 8 8 8 8 8 8 8 8 8 9 9 9 12 12 12	8/32
24	4 4 6 6 6 6 7 8 8 8 8 8 8 8 8 8 8 9 9 9 9 12 12 13	8/56
24	4 4 6 6 6 6 7 8 8 8 8 8 8 8 8 8 9 9 9 9 9 12 12 13	8/56
24	4 6 6 6 6 6 6 7 8 8 8 8 8 8 8 9 9 9 9 9 9 12 12 13	8/104
24	4 6 6 6 6 6 7 8 8 8 8 8 8 8 8 8 9 9 9 9 9 12 12 13	8/40
24	4 4 4 6 6 6 6 7 8 8 8 8 8 8 8 8 9 9 9 9 12 12 12 13	8/32
24	4 4 6 6 6 6 7 7 8 8 8 8 8 8 8 9 9 9 9 9 12 12 13 13	8/56
25	4 4 4 4 4 4 4 4 4 4 4 4 5 5 5 5 5 5 8 8 8 8 8 8 12	16/144
25	4 4 4 4 4 4 4 4 4 4 6 6 6 8 8 8 8 8 8 8 9 9 9 9 12	8/24
25	4 4 4 4 6 6 6 6 6 8 8 8 8 8 8 8 8 9 9 9 9 9 9 12 12	16/32
25	4 4 4 4 4 4 4 4 4 4 4 4 4 5 5 5 5 8 8 8 8 8 8 12 12	4/36
25	4 4 4 4 4 4 4 4 4 4 4 4 4 5 5 6 8 8 8 8 8 8 9 12 12	8/64
25	4 4 4 4 4 4 4 4 4 4 5 5 5 5 5 5 8 8 8 8 8 8 8 12 12	8/152
25	4 4 4 4 4 4 4 4 4 6 6 6 8 8 8 8 8 8 8 9 9 9 9 12 12	8/56
25	4 4 4 4 4 4 4 4 5 5 5 5 6 8 8 8 8 8 8 8 8 9 12 12 12	8/24
25	4 4 4 4 4 4 6 6 6 6 8 8 8 8 8 8 9 9 9 9 9 12 12 12 12	8/24
		340/ 2,380

Table 9: Probing sequences of 5×5 Nonograms of the remaining category which are uniquely solvable.

The other occurrences of this probing sequences result in 16 different descriptions which lead to more than one solution.

Nonograms of size $n \times n$, with $n \leq 4$, are too small for *Simple* and *Probing* not to solve the provided description. However, considering larger Nonograms, there are uniquely solvable Nonograms which are not going to be classified correspondingly with the current solvers. By applying a *Set-Probing* approach, these Nonograms might be classified as uniquely solvable rather than having two or more solutions.

5 Conclusions and Further Research

Japanese puzzles, Nonograms, have a description which is used to try to find a solution. To solve Nonograms *Simple* and *Probing*, as described in Section 3, have been used. *Simple* tries to solve Nonograms by only considering *H-* and *V-sweeps*. The *H-* and *V-sweeps* try to solve complete lines. To solve a line, dynamic programming is used as a basis. Ultimately, if this method does not lead to a uniquely solved Nonogram, *Probing* is applied, where cells are tried as a filled cell followed by leaving the cell blank. The result of the probed cells is stored in the probing sequence of the Nonogram. There are four classifications regarding the usefulness of a probe: no progression, some progression, a contradiction and fully solving. The usefulness of each double probe is then determined by applying the Quaternary number system. There are sixteen categories, 0 to 15, regarding the usefulness of the double probe.

Some of the defined probe categories indicate a classification. A category 10 probe indicates that neither a 0-probe nor a 1-probe leads to any progression. Category 15 probes indicate that there are exactly two solutions. Probing sequences that consist solely of category 4 probes seem to indicate multiple solutions. Category 11 and 14 probes result in a uniquely solvable Nonogram. By the use of the stated categories, the Nonograms can be divided into different categories. We differentiate Nonograms that can be solved by solely using *Simple* and those where *Probing* is needed. Those where *Probing* is applied are then divided into those that have a unique solution found after probing, those that have exactly two solutions and a remaining category where most have more than two solutions. Apart from the earlier mentioned categories, no clear patterns occur when trying to determine the number of possible solutions with respect to a provided description.

As briefly mentioned in Section 4, there might be some uniquely solvable Nonograms in the remaining category of *Probing*. For future work, these could be extracted from this category. By applying an alternative *Probing* approach using multiple double probes at the same time, known as a *Set-Probing* approach, these uniquely solvable Nonograms might be found. The *Set-Probing* approach may be analyzed in a similar manner as described in Section 3.4 and Section 4. By doing so, we will get another difficulty classification, due to its extended approach. The experiments could be extended by considering larger $n \times n$ Nonograms to see how the probing sequences behave.

References

[BHKP09] K. Joost Batenburg, Sjoerd Henstra, Walter A. Kosters, and Willem Jan Palenstijn. Constructing simple Nonograms of varying difficulty. *Pure Mathematics and Applications*, 20:1–15, 2009.

[BK08] K. Joost Batenburg and Walter A. Kosters. A reasoning framework for solving Nonograms. In *Proceedings International Workshop on Combinatorial Image Analysis*, pages 372–383. Springer, 2008.

[BK12] K. Joost Batenburg and Walter A. Kosters. On the difficulty of Nonograms. *ICGA Journal*, 35(4):195–205, 2012.

[UN96] Nobuhisa Ueda and Tadaaki Nagao. NP-completeness results for Nonogram via parsimonious reductions. *Technical Report TR96-0008*, 1996.

[WvB04] Wouter Wiggers and Willem van Bergen. A comparison of a genetic algorithm and a depth first search algorithm applied to Japanese Nonograms. In *Proceedings Twente Student Conference on IT*, 2004.

[YLC11] Chiung-Hsueh Yu, Hui-Lung Lee, and Ling-Hwei Chen. An efficient algorithm for solving Nonograms. *Applied Intelligence*, 35(1):18–31, 2011.