Spreadsheets in Education (eJSiE)

Volume 4 | Issue 1 Article 3

6-19-2010

Spreadsheet-Based Sudoku as a Tool for Teaching Logical Deduction

Clarence C. Y. Kwan
McMaster University, kwanc@mcmaster.ca

Spreadsheet-Based Sudoku as a Tool for Teaching Logical Deduction

Abstract

Drawing on the contributions of two recent studies by Sugden (2007) and Luterbach, Rodriguez, and Milling (2010) in Spreadsheets in Education, this study illustrates that Excel-based Sudoku is an effective tool for teaching logical deduction. This study also illustrates some essential Sudoku techniques, all based on logical deduction. Relevant information from the Sudoku grid is extracted by using Excel features, thus making it easier for students to explore opportunities to apply suitable Sudoku techniques throughout the solution process.

Keywords

Sudoku, logical deduction, spreadsheet

Spreadsheet-Based Sudoku as a Tool for Teaching Logical Deduction

Clarence C.Y. Kwan
DeGroote School of Business
McMaster University
Hamilton, Ontario L8S 4M4
Canada
E-mail: kwanc@mcmaster.ca

June 2010

Abstract

Drawing on the contributions of two recent studies by Sugden (2007) and Luterbach, Rodriguez, and Milling (2010) in *Spreadsheets in Education*, this study illustrates that Excel-based Sudoku is an effective tool for teaching logical deduction. This study also illustrates some essential Sudoku techniques, all based on logical deduction. Relevant information from the Sudoku grid is extracted by using Excel features, thus making it easier for students to explore opportunities to apply suitable Sudoku techniques throughout the solution process.

Submitted March 2010, revised and accepted June 2010.

Key words: Sudoku, logical deduction, spreadsheet.

1 Introduction

Sudoku is a popular number-placement game. The standard version of Sudoku is played on a 9×9 grid, which is divided into nine 3×3 sub-grids, also known as boxes, in the same manner that a matrix is partitioned into block matrices. A new puzzle starts with a grid where some cells are already filled with integers from 1 to 9. The task is to fill all vacant cells, also with such integers, satisfying the requirement that all nine numbers in each row, in each column, and in each box be distinct. The popularity of Sudoku can be attributed to both the simplicity of its rules for beginners and the sophistication of some of its advanced techniques to sustain the interest of experienced players.

Besides the many daily Sudoku puzzles that appear in major newspapers worldwide, there are also numerous websites that allow the game at different levels to be played on the internet. While there are various websites that offer advice on Sudoku strategies, there is also an online encyclopedia, called Sudopedia, with voluntary contributions from

Sudoku experts and enthusiasts.[1] Beginners and experienced players alike now have easy access to all aspects of the game. Gradually, the language of the game is becoming more standardized. Though still confusing to outsiders and beginners, terms such as *X-wing*, *XY-wing*, *Swordfish*, *Jellyfish*, and many others are already part of the standard Sudoku vocabulary.

As Delahaye (2006) reports in an article in *Scientific American*, although solving a Sudoku puzzle requires not even arithmetic, the game actually raises many intriguing mathematical issues. By posing some basic and open questions, Taalman (2007) reiterates that Sudoku deserves to be taken seriously. Indeed, Sudoku has already attracted the attention of many educators and researchers, including those outside mathematical fields. Collectively, their work has contributed to the advancement of mathematics and science education. Here are some recent examples:

Several chemists report the use of Sudoku puzzles as teaching tools to help students learn organic functional groups, i.e., groups of atoms within molecules that are involved in chemical reactions characteristic of those molecules; see, for example, Crute and Myers (2007), Perez and Lamoureux (2007), and Welsh (2007). Crook (2009) presents, with the aid of concepts in set theory, a pencil-and-paper approach to solve Sudoku puzzles. Lorch and Lorch (2008) show how to incorporate methods for counting Sudoku puzzles into an introductory abstract algebra course.

Further, Dahl (2009) associates Sudoku with a class of permutation matrices. Herzberg and Murty (2007) relate Sudoku to chromatic polynomials — polynomials studied in algebraic graph theory — by treating each Sudoku puzzle as an exercise in proper coloring of the vertices of a graph. Sander (2009), in contrast, considers Sudoku from the perspective of spectral graph theory. Bailey, Cameron, and Connelly (2008) connect Sudoku to various mathematical and statistical topics. Lee, Goodwin, and Johnson-Laird (2008) present a theory on how individuals solve Sudoku puzzles from a psychological perspective and provide collaborative experimental results.

With electronic spreadsheets such as Microsoft Excel^{TM} becoming more familiar to students in recent years, Sudoku puzzles on spreadsheets are becoming more suitable as an educational tool. Specifically, Weiss and Rasmussen (2007) show how Sudoku puzzles can be solved with Premium Solver^{TM}, which is an upgraded version of Excel Solver^{TM} to accommodate a large number of constraints on the placement of numbers in individual vacant cells. As the task requires coding in Visual Basic for Applications (VBA), students can use Sudoku puzzles for programming exercises. In contrast, Sugden (2007) sets up an Excel worksheet with cell formulas alone to help students solve Sudoku puzzles and uses it as a tool for teaching set theory. More recently, to compare student performance in paper-based and spreadsheet-based approaches to complete Sudoku puzzles in Grades 4 and 5 mathematics classes, Luterbach, Rodriguez, and Milling (2010, hereafter LRM) use Excel's data validation to help students prevent Sudoku rule violations.

Drawing on the contributions of the Sugden and LRM studies, this study is intended to make Sudoku more effective for classroom use. It shows how Excel-based Sudoku can be used as a tool for teaching logical deduction in mathematics classes. In core mathemat-

[1] See <http://www.sudopedia.org/wiki/Main_Page>.

ics courses at the secondary level, such as algebra, geometry, and trigonometry, students are taught various mathematical proofs and are assigned related exercises for practice. Although Sudoku is not directly connected to any specific topics in these courses, except for being an excellent example for teaching set theory, the logical deduction skills that students can acquire from solving Sudoku puzzles are indeed valuable.

To benefit more from Sudoku, however, students have to learn some essential techniques of the game. For ease of exposition of the materials that appear later in this study, such techniques, though generally accessible from internet sources such as those linked to the Sudopedia website, are also covered here. Such techniques, once acquired, will help students solve many Sudoku puzzles that are more difficult than those considered in the Sugden and LRM studies, thus improving their logical deduction skills in the process. Unfortunately, to cover such techniques in classroom settings, especially for paper-based Sudoku, is very time-consuming. Even for classrooms with internet access, to solve Sudoku puzzles online in class, for the purpose of illustrating Sudoku techniques to students, is not an ideal option.[2]

We now identify the specific features in the Excel files accompanying the Sugden and LRM articles that are relevant for purposes of this study. Notably, by using Sugden's Excel worksheet to solve a Sudoku puzzle, players can use the automatically generated clues on the grid — which include highlighted numbers for all default entries and vacant cells with exactly two candidates — to make progress toward the final solution of the game. In easy Sudoku puzzles such as those in the Sugden and LRM studies, enumeration of the candidates for the corresponding vacant cells will lead to successive revelations of default entries and will allow each puzzle to be solved without much effort. For more difficult Sudoku puzzles, however, enumeration of the candidates alone may not lead to their final solutions. Once no more default entries are revealed, to extract relevant information from the Sudoku grid for the game to proceed can be a challenge for inexperienced players.

The feature of Sugden's Excel worksheet — that it automatically lists all candidates for each vacant cell — should appeal to experienced Sudoku players who consider the initial enumeration of candidates a chore. Such a feature should also make Sugden's Excel worksheet an ideal template for teachers to set up Sudoku puzzles for classroom use. Teachers can access Sudoku puzzles at different levels of difficulty from public sources. With the template at hand, they no longer have to perform the tedious task of enumerating the individual candidates themselves. Instead, they can directly use information from the template to check the listing accuracy of the initial candidates for the vacant

[2] Freely accessible Sudoku puzzles from the internet typically either disallow players to enumerate their own candidates for any vacant cells or give them the option to see, but not to revise manually, the internally generated candidate lists. As will soon be clear, having the option to revise the candidate lists is important in the process of solving a seemingly difficult Sudoku puzzle by logical deduction. Web Sudoku, <http://www.websudoku.com/>, is a rare exception; this free website does allow players to enter up to five candidates for each vacant cell. However, depending on the puzzle being solved, some vacant cells may have more than five candidates, typically during early stages of the game. Incomplete candidate lists may cause inconvenience or even confusion for inexperienced players. Further, most free Sudoku websites, including the one mentioned above, do not allow any partially solved puzzles to be saved and resumed at a later date.

cells that students deduce. To sustain the interest of experienced students, they can also let these students solve challenging Sudoku puzzles where the initial candidates for some or all vacant cells have already been identified.

The feature of data validation in LRM's Excel worksheet should appeal to teachers who use Sudoku as an educational tool, as mistakes in entering numbers to vacant cells of Sudoku puzzles can cause frustration of inexperienced players of the game and interfere with the learning process. However, the way an Excel-based Sudoku puzzle is set up in LRM does not allow students to keep track of the candidates for each vacant cell. Thus, the Excel worksheet in LRM is not readily suitable for classroom use in teaching logical deduction, especially when more difficult Sudoku puzzles are involved.

This study extends the Sugden and LRM approaches as follows: Specifically, it extends the LRM formula for data validation of each cell to allow also the entry of a multiple-digit integer, where each digit represents a candidate. For example, if the number 247 is entered to a vacant cell, this means that the cell is still considered to be vacant and that 2, 4, and 7 are all candidates for the cell. Such flexibility in data entry allows Excel-based Sudoku to retain a nice feature of the traditional paper-based version of the game, which allows easy changes to any pencil markings on the grid as the game progresses. In addition, when solving a Sudoku puzzle on Excel, students can extract some specific information from the grid to allow the game to proceed. Exactly what information from a Sudoku grid is generally relevant will soon be clear in this study.

What makes Sudoku unusual is the fact that solving Sudoku puzzles requires no prior mathematical knowledge, apart from knowing how to enumerate distinct numbers or items in general. Thus, the game can be introduced to students at almost any grade level. To give teachers the flexibility in choosing a suitable Excel-based Sudoku puzzle for classroom use, this study provides two versions of the game. If students are experienced enough not to require help to prevent Sudoku rule violations, then the version without data validation is already adequate. For less experienced students, however, the version with data validation is obviously better.

Excel-based Sudoku also has the advantage of being able to update immediately the results of all computations involved and to display any selected information on the computer monitor. If the classroom has an electronic overhead projector, such information can be displayed on a large screen as well. With the aid of technology, it becomes easier for students to participate in, and contribute to, the process of shortening the candidate lists for various vacant cells on the Sudoku grid. Further, it is relatively simple to choose any part of the Excel worksheet for display, to show the solution process, to highlight any specific cells with color, and, if necessary, to retrace previous steps in response to questions from students. These convenient Excel features all have contributions to make Excel-based Sudoku an educational tool that can enhance the classroom experience of students.

This article is organized as follows: Section 2 illustrates several essential Sudoku techniques. The section is not intended to duplicate many known Sudoku techniques. Rather, it is intended to help students recognize, more readily, typical situations that require logical deduction for the game to proceed.[3] Section 3 presents an Excel-based example to illustrate the solution process for a Sudoku puzzle that requires some of the essential techniques as described in Section 2. Section 4 provides some concluding remarks.

Separately, the Appendix of this article states, illustrates, and proves, also by logical deduction, two interesting Sudoku properties. The first property pertains to the patterns of number placements on some Sudoku sub-grids. Although the property tends to contribute only marginally to the solution process and thus is not covered in the main text, its proof is a good example to illustrate logical deduction.

The second property pertains to the issue of uniqueness of solutions to Sudoku puzzles. The property has an interesting implication; it leads to an advanced Sudoku technique. Although situations requiring this specific technique are seldom encountered by casual Sudoku players and thus are not considered in the main text, the technique itself is actually a good classroom example to illustrate logical deduction in the context of Sudoku.

2 Elimination of redundant candidates by logical deduction

In the language of set theory, the candidates — which include redundant candidates — for each vacant cell on a standard 9×9 Sudoku grid are elements of the complement of the union of the three sets of integers already in the row, the column, and the box that contain the vacant cell in question. A redundant candidate is an integer that can be ruled out because it must eventually be placed in a connected cell. Here, two cells are connected if they share a common row, a common column, or a common box. The idea of solving a given Sudoku puzzle is to reduce the number of redundant candidates for each vacant cell, by logical deduction, until all vacant cells are filled with correct numbers.

For ease of exposition below, we follow a commonly used convention to label the individual cells, rows, columns, and boxes of the Sudoku grid. The numbering of rows and columns is analogous to that for matrices; row and column numbers are, in ascending orders, from top to bottom and from left to right, respectively. The cell that is in row i and column j is labeled as $RiCj$, for $i, j = 1, 2, \ldots, 9$. The numbering of the nine 3×3

[3]Although this study — more specifically, Section 2 — can serve as an introduction to Sudoku, the materials covered may still be too technical for readers who are unfamiliar with the game. Such readers will benefit more from this study once they have gained some hands-on experience from solving simple Sudoku puzzles such as those in the Sugden and LRM studies.

boxes is as follows:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$									
$R2$		Box 1			Box 2			Box 3	
$R3$									
$R4$									
$R5$		Box 4			Box 5			Box 6	
$R6$									
$R7$									
$R8$		Box 7			Box 8			Box 9	
$R9$									

As solving each Sudoku puzzle without guessing is a process of logically eliminating redundant candidates, we provide below nine examples, with each illustrating a potentially applicable task. Students can learn from these examples how logical deduction works in different situations. From a pedagogic perspective, it is useful to give students some exercises to gain hands-on experience with various Sudoku techniques. Such experience will help them recognize, without any hints from others, situations to which some specific techniques are applicable. We can assess how well a student has acquired such techniques by knowing at which stage of the solution process the student cannot proceed any further.

Example 1: The following is a simple case where correct entries to two vacant cells can be identified directly:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$		6	2						
$R2$		3		2		1		4	
$R3$							1		8

Here, as integer 1 is already present in rows 2 and 3, and in boxes 2 and 3, its default position in box 1 is $R1C1$. Likewise, the default position for integer 2 in box 3 is $R3C8$. Given these direct results, as shown in the following, there is no need to enumerate all candidates for $R1C1$ and $R3C8$:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	1	6	2						
$R2$		3		2		1		4	
$R3$							1	2	8

The same idea also works for a set of vertically stacked boxes. The method here to identify default entries is best performed at the start of each Sudoku game, prior to enumerating all candidates for the remaining vacant cells on the grid. However, as the game progresses, similar opportunities may also arise for the method to be used.

Example 2: The following example illustrates how some candidates can be eliminated:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	9	347	2	8	47	1	345	356	34

Here, the four single-digit entries are given or solved numbers for row 1, and the five multiple-digit entries provide the candidate lists for the individual vacant cells, with each digit representing a candidate. The numbers 3, 4, and 7 must eventually be placed in $R1C2$, $R1C5$, and $R1C9$, where the candidates are digits of 347, 47, and 34, respectively. Thus, we can safely remove candidates 3 and 4 from $R1C7$; this will give us a 5 for that cell. Now, with both 3 and 5 ruled out, cell $R1C8$ becomes a 6, as shown in the following:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	9	347	2	8	47	1	5	6	34

Notice that the idea of locked cells as illustrated in the example also applies to entries in a common column or in a common box. The method tends to be used frequently for solving Sudoku puzzles.

Example 3: The following is a similar but slightly more complicated example:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	1245	24579	2578	6	14	39	139	38	1389

Here, the candidates 2, 5, and 7 for row 1 appear in two or three of the cells $R1C1$, $R1C2$, and $R1C3$ and nowhere else in the row. As 2, 5, and 7 will eventually be placed in these three cells, we can eliminate all others candidates there. The result of the step is as follows:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	25	257	257	6	14	39	139	38	1389

Now, as the candidacy of 4 for row 1 is confined to $R1C5$, the number must be placed there. The end result, therefore, is as follows:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	25	257	257	6	4	39	139	38	1389

Notice that, as the numbers 2, 5, and 7 will eventually fill the first three rows of box 1, none of them are eligible as candidates for the remaining vacant cells of the box (although the rest of the box is not shown in the example).

Example 4: The following 3×9 sub-grid can be used to illustrate the removal of a candidate for a vacant cell:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	67	3	2	5678	68	58	9	1	4
$R2$	4	68	168	9	168	2	3	5	7
$R3$	9	5	17	347	13	14	2	8	6

As row 1 must have an 8 and it can only be placed in one of $R1C4$, $R1C5$, and $R1C6$, its candidacy for $R2C5$ must be ruled out. Alternatively, as box 1 must have an 8 and it can only be placed in $R2C2$ or $R2C3$, the remaining vacant cell in row 2 cannot have the same number.

The relevant information for ruling out the candidacy of 8 for $R2C5$ above can be presented alternatively as follows:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$				8	8	8			
$R2$		8	8		8				
$R3$									

Each number on this sub-grid is not a given or solved number of the Sudoku puzzle being considered; rather, it is only a candidate for the corresponding vacant cell. Clearly, with the potential locations of 8 in box 1 and in row 1 locked, its candidacy for $R2C5$ can be ruled out.

The ideas as illustrated in the above four examples are generally enough for solving many Sudoku puzzles, including those broadly classified as "hard" or "evil."[4] However, to help students develop further skills in logical thinking, it is also useful to show them some solution methods that are beyond the basic types, once they have become more experienced with the game. As many advanced Sudoku techniques are based on considerations of individual integers or pairs of integers as candidates for various vacant cells, we illustrate some of them in the examples below.

Example 5: Like the alternative approach in Example 4, the following grid shows all potential locations of a specific number, which is a 5 here:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	5	5				5			
$R2$	5	5			5				
$R3$									
$R4$				5		5			5
$R5$				5	5	5			5
$R6$									
$R7$		(5)						(5)	
$R8$				5		5			
$R9$		(5)						(5)	

The vacant cells $R7C2$, $R7C8$, $R9C2$, and $R9C8$, where the integer 5 is shown in parentheses, are of special interest. These four cells can be viewed as the corners of a rectangle. As each of rows 7 and 9 has two potential locations for a 5, the same integer must eventually be placed at opposite corners of the rectangle. Regardless of whether $R7C2$ or

[4] See, for example, Web Sudoku <http://www.websudoku.com/> for billions of Sudoku puzzles in four categories, which includes also the "easy" and "medium" types.

$R9C2$ ends up having a 5, the candidacy of the integer for $R1C2$ and $R2C2$ can be ruled out.[5]

Example 6: The following grid, which shows the initial locations of the integer 6 as a candidate for various vacant cells, is an extension of the same ideas in the above example:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$		6				6	6		
$R2$	6	6	(6)		(6)				
$R3$		6					6		
$R4$	6					6			
$R5$			(6)			6		(6)	6
$R6$	6			6					6
$R7$		6	(6)		(6)			(6)	
$R8$	6								6
$R9$	6			6		6	6		6

Here, the integer that is a candidate for each vacant cell at the intersections of any of rows 2, 5, and 7 and any of columns 3, 5, and 8 is shown in parentheses. The nine corners of the group of rectangles formed by intersecting the three rows and the three columns contain seven vacant cells for which a 6 can potentially be placed.

As each of the three columns must have a 6, the integer, which appears three times in each column on the above grid, must eventually be placed in each of the three rows. Thus, the candidacy of the integer for all vacant cells in the same rows not covered by the corners of the group of rectangles can safely be ruled out. They include $R2C1$, $R2C2$, $R5C6$, $R5C9$, and $R7C2$.[6]

For box 1, as a 6 must be placed in its column 2, the candidacy of the integer for $R2C3$ can be ruled out. This in turn leaves $R2C5$ as the default position for 6 in row 2 and leads to the removal of 6 as a candidate for $R1C6$ and for $R7C5$. Now, with $R1C2$, $R1C7$, $R3C2$, and $R3C7$ being the four corners of the rectangle in a situation analogous to that in example 5, the candidacy of 6 for $R9C7$ can also be ruled out. The end result,

[5] In the language of Sudoku, the technique here is called *X-wing*.

[6] The technique here is called *Swordfish* in the language of Sudoku; it is an extension of *X-wing* by considering an extra row and an extra column to place a common candidate. A further extension, which considers cases of four rows and four columns, is called *Jellyfish*.

with $R2C5$ solved, is as follows:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$		6					6		
$R2$					(6)				
$R3$		6					6		
$R4$	6					6			
$R5$			(6)					(6)	
$R6$	6			6					6
$R7$			(6)					(6)	
$R8$	6								6
$R9$	6			6		6			6

Example 7: The following is another example, which shows a different way to utilize the information on the candidacy of a specific integer — a 9 here — for various vacant cells on a 6×9 sub-grid:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	9			9	9				
$R2$	9	(9)		9	9			9	9
$R3$	9	9		9				9	
$R4$	9			9	9				9
$R5$	9	9			9			9	9
$R6$		9						9	

Of special interest here is the 9 in $R2C2$, which is shown in parentheses.

We now show that 9 cannot be a candidate for $R2C2$. To do so, let us consider row 6, where there are only two potential locations for 9. If it is placed in $R6C2$, no other cells in column 2 can have the same number. If it is placed in $R6C8$ instead, then, with its candidacy for $R2C8$ and $R3C8$ ruled out, its default location in box 3 is $R2C9$. Accordingly, no other cells in row 2 can have the same number. Thus, in either case, $R2C2$, which belongs to row 2 and column 2, cannot have 9 as a candidate.

Example 8: A similar idea to rule out the candidacy of a number for a vacant cell can sometimes work well in situations where two specific numbers are the only candidates for several vacant cells. The following 6×9 sub-grid shows an example where the integers 1 and 3 are such candidates:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	(13)				1				3
$R2$									
$R3$			[13]		1		3		3
$R4$		13	13	[13]					
$R5$	[13]						(13)		
$R6$				(13)					[13]

Here, the entries in parentheses or brackets are those where 1 and 3 are the only candidates. The remaining entries are those where there are other candidates besides the digits indicated. If one of the two integers 1 and 3 is placed in any cell indicated by (13), the same integer must be placed in all other cells indicated by (13) and the remaining integer must be placed in all cells indicated by [13]. Therefore, regardless of the eventual placements of the two integers in these differently labeled cells, the candidacy of 3 for $R1C9$ and $R3C7$ can be ruled out, as the union of the row and the column containing each of the two cells will always have a 3.[7]

Example 9: The following is another situation where some vacant cells on a 3×9 sub-grid have exactly two candidates as shown:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$					57			27	15
$R2$	15			17		35			
$R3$			24					27	15

Here, each of $R1C5$, $R2C1$, and $R2C4$ must have one of 1, 5, and 7, regardless of the exact placements of the three integers there. As $R2C6$ shares the same row with $R2C1$ and $R2C4$ and shares the same box with $R1C5$, it cannot have 5 as a candidate. Thus, by default, 3 is the integer that $R2C6$ can have. Notice that the same idea also works if $R2C6$ has more than two candidates. If so, the candidacy of 5 there can be ruled out; however, the cell remains unsolved.[8]

3 An example of Excel-based Sudoku

As illustrated repeatedly in various examples in Section 2, a candidate for a vacant cell is ruled out if it can be established that the same candidate must eventually be placed in one of the connected vacant cells. However, the applicable techniques for the task do vary, depending on the individual situations involved. Therefore, a challenge for inexperienced Sudoku players is to recognize which technique to attempt, when overwhelmed by a vast amount of information on the 9×9 grid of a partially solved Sudoku puzzle.

With the aid of various Excel features, the locations of specific integers or pairs of integers as candidates for various vacant cells can be extracted and displayed separately. Before proceeding, we must emphasize that the role of Excel here is not to solve any Sudoku puzzles for students, nor to trivialize the solution process. The use of Excel features is intended to make it easier for students to recognize situations for which some specific Sudoku techniques known to them are potentially applicable. It is still the task of students to solve each given puzzle.

For Sudoku puzzles encountered by general players of the game, the techniques as described in the first seven examples in Section 2 are usually adequate. To apply these

[7] In the language of Sudoku, the technique here is called *Remote Pairs*.

[8] In the language of Sudoku, this technique to rule out the candidacy of an integer for a vacant cell by considering three vacant cells connected to it and among themselves, where three integers must eventually be placed, is called *XY-wing*.

techniques to situations in Examples 4-7 requires only the locations of each of the nine integers as candidates for various vacant cells. To apply more advanced techniques, including those for situations in Examples 8 and 9, requires two additional grids, with one showing all candidates for each vacant cell and the other showing only cases where each cell in question has exactly two candidates. In both cases, the given and solved cells need not be displayed.

Technically, once the candidates for each vacant cell are correctly identified, the given and solved cells are irrelevant in the solution process. The reason is that the candidates for each vacant cell are elements of the complement of the union of the sets of integers in the same row, in the same column, and in the same box containing the vacant cell in question. There will not be Sudoku rule violations between any selected candidates and the given and filled numbers on the grid. If there are any violations at all, they must be among the selected candidates themselves. Thus, an inspection of the separate Sudoku grids containing the candidates alone is often sufficient to find ways to rule out redundant candidates in the solution process.

An example of Excel-based Sudoku is provided below. The corresponding Excel file has been set up to accommodate Sudoku puzzles of different levels of difficulty. The puzzle for the illustration has been obtained randomly from Web Sudoku <http://www.websudoku.com/> on the internet, under the "Evil Puzzle" category. The name of the category notwithstanding, it turns out that the puzzle requires only some of the techniques as illustrated in the first seven examples in Section 2. In fact, there is no need to examine the Excel worksheet for any opportunities that the locations of vacant cells with exactly two candidates may provide. However, students new to Sudoku will most likely find the required techniques for solving the puzzle very difficult to comprehend. Thus, from a pedagogic perspective, it is better to use the Excel worksheet to illustrate some easier puzzles before attempting any puzzles similar to the one here.

3.1 Features of the Excel file

A representative part of the Excel worksheet, showing the start of the solution process, is provided below. The corresponding technical details are also shown.[9]

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1		Enter only integers here:																			
2	Game	5679	15679	2	8	367	6	169	13569	4		Givens			2	8					4
3		679	4679	3	247	1	5	269	269	8					3		1	5			8
4		5678	145678	15678	9	23467	46	126	12356	123567						9					
5		2567	3	9	12457	2457	14	8	12456	1256				3	9				8		
6		2578	12578	1578	6	24578	3	1249	12459	1259						6		3			
7		2568	12568	4	125	258	189	3	7	12569					4				3	7	
8		356789	56789	5678	145	4568	2	1469	134689	1369								2			
9		1	2568	568	3	9	468	7	2468	26			1			3	9		7		
10		4	2689	68	1	68	7	5	123689	12369			4					7	5		
11																					
12	Text	5679	15679	2	8	367	6	169	13569	4											
13		679	4679	3	247	1	5	269	269	8											
14		5678	145678	15678	9	23467	46	126	12356	123567											
15		2567	3	9	12457	2457	14	8	12456	1256											
16		2578	12578	1578	6	24578	3	1249	12459	1259											
17		2568	12568	4	125	258	189	3	7	12569											
18		356789	56789	5678	145	4568	2	1469	134689	1369											
19		1	2568	568	3	9	468	7	2468	26											
20		4	2689	68	1	68	7	5	123689	12369											
21																					
22	Pairs											All	5679	15679			367		169	13569	
23													679	4679		247			269	269	
24							46						5678	145678	15678		23467	46	126	12356	123567
25							14						2567			12457	2457	14		12456	1256
26													2578	12578	1578		24578		1249	12459	1259
27													2568	12568		125	258	189			12569
28													356789	56789	5678	145	4568		1469	134689	1369
29										26				2568	568			468		2468	26
30				68		68								2689	68		68			123689	12369
31																					
32	One		1					1	1												
33																					
34			1	1				1	1	1											
35					1		1		1	1											
36			1	1				1	1	1											
37			1		1		1			1											
38					1			1	1	1											
39																					
40									1	1											

[9] Readers who wish to bypass the technical descriptions in this subsection can do so without affecting the readability of the remainder of the article.

Here, the puzzle has been manually entered into the block M2:U10 of an Excel 2007 worksheet. The numbers in M2:U10 have been pasted to Sugden's Excel file to generate a complete set of candidates for each vacant cell. Subsequently, the values generated there have been pasted to the cells B2:J10 above. To allow all candidates for individual vacant cells on the Sudoku grid to be displayed, we have opted for "shrink to fit" in Excel's cell formatting.[10] The Excel file here has been saved as an Excel 97-2003 workbook.

Alternatively, the solution process can start with the puzzle itself. With the values in M2:U10 directly pasted to B2:J10, the candidates for all vacant cells are to be enumerated manually. Another option is to delete the values of some cells that Sugden's Excel template provides. In such a case, the starting point of the solution process is to identify the candidates for these vacant cells. Regardless of the starting point of the solution process, if the feature of data validation is present in the 81 cells in B2:J10, it is important to ensure that only values be pasted there. Pasting the contents of other cells there, other than their values, will cause the loss of the feature of data validation.

We now use a representative cell among the 81 cells in B2:J10 to illustrate how data validation can be set up. For example, after selecting the cell B2, which represents $R1C1$ of the Sudoku grid, we select "Data Validation" in the "Data Tools" group on the "Data" tab of the menu bar. We then select "Custom" from the menu of "Allow" and paste the following formula in the formula box provided:

```
=AND(INT(CELL("contents"))=CELL("contents"),OR(CELL("contents")>9,AND(CELL
("contents")>0,CELL("contents")<10,COUNTIF($B2:$J2,CELL("contents"))=1,COUNTIF
(B$2:B$10,CELL("contents"))=1,COUNTIF($B$2:$D$4,CELL("contents"))=1)))
```

This formula, which extends the idea of LRM, allows a positive integer to be entered in the cell B2. To be acceptable, the integer must either be greater than 9 or satisfy the condition that it has not already appeared in the filled cells among B2:B10, B2:J2, or B2:D4. The use of the COUNTIF function is similar to that in LRM; it ensures the uniqueness of each single-digit entry by ruling out the presence of the same integer in any connected cells.[11] The use of absolute rows or columns (as denoted by $ here) is for convenience when editing various similar formulas for the 81 individual cells requiring data validation.

The block B12:J20 contains the text version of the numbers in B2:J10. Specifically, the cell B12, which uses the formula "=TEXT(B2,"#")", are copied to the block B12:J20. This conversion makes it easier for each specific candidate to be extracted from the set of candidates for each vacant cell, in the same manner as the various displays in Examples 4-7 in Section 2. The block B32:J40, where the locations of the integer 1 as a candidate for various vacant cells on the Sudoku grid are shown, is only an example. The rest, including $2, 3, \ldots, 9$, though not explicitly shown above, are in B42:J120 of the Excel worksheet.

[10] This option is available in Excel's "Format Cells: Alignment" dialog box.

[11] See the Appendix of the LRM study for a description of a representative validation formula there. The formula above extends that in LRM by using also the OR function to accommodate multiple-digit entries.

The cell B32 of the worksheet, which represents $R1C1$ of the Sudoku grid, uses the formula "=IF(B2<10,"",IF(COUNTIF(B12,"*1*")>0,1,""))"; the formula is pasted to the block B32:J40. For the cell B42 of the worksheet not shown above, which also represents $R1C1$ of the Sudoku grid for the display of the candidacy of the integer 2 there, the formula is revised to "=IF(B2<10,"",IF(COUNTIF(B12,"*2*")>0,2,""))" and pasted to the block B42:J50. The formulas for capturing the candidacy of the remaining integers are analogous.

To illustrate how each cell formula in B32:J40 works, consider C32, for example. The cell represents $R1C2$ on the Sudoku grid, where the five candidates are 1, 5, 6, 7, and 9. The text version of this multiple-digit number, 15679, has a 1, and its presence is recognized by "*1*" in the formula "=IF(C2<10,"",IF(COUNTIF(C12,"*1*")>0,1,""))" where the wildcard (*) captures all other candidates on the list. As the criterion for C12 to have a 1 is satisfied, the function COUNTIF provides a positive count and a 1 is displayed in C32. In the case of F32, the corresponding formula is "=IF(F2<10,"",IF(COUNTIF(F12,"*1*")>0,1,""))". The candidates for $R1C5$, captured by F12, are 3, 6, and 7, and thus there is no 1 on the list. Accordingly, a blank is resulted for F32.

The blocks B22:J30 and M22:U30 show all occurrences of pairs of candidates and all candidates, respectively. The formulas for B22 and M22, which are each pasted to the remainder of the corresponding block, are "=IF(AND(B2>0,B2<10),"",IF(AND(B2>9,B2<100),B2,""))" and "=IF(B2<10,"",B2)", respectively. For these two blocks, numbers in B2:J10 can be used directly.

3.2 The solution process

For ease of exposition, we now return to the same way Sudoku cells are labeled in Section 2. The block B2:J10 in the Excel worksheet is duplicated as follows:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	5679	15679	2	8	367	6	169	13569	4
$R2$	679	4679	3	247	1	5	269	269	8
$R3$	5678	145678	15678	9	23467	46	126	12356	123567
$R4$	2567	3	9	12457	2457	14	8	12456	1256
$R5$	2578	12578	1578	6	24578	3	1249	12459	1259
$R6$	2568	12568	4	125	258	189	3	7	12569
$R7$	356789	56789	5678	145	4568	2	1469	134689	1369
$R8$	1	2568	568	3	9	468	7	2468	26
$R9$	4	2689	68	1	68	7	5	123689	12369

Here, each of $R1C6$ and $R9C4$ has a single candidate. In addition, following the same idea as illustrated in Example 1 in Section 2, we are able to identify the default location of 3 in box 7, as well as the default location of 7 in box 3. With these four vacant cells solved, what follows is a series of obvious steps to remove some redundant candidates and to reach the solved numbers for some of them. The following is at the stage where some less obvious forms of logical deduction are required for the process to continue:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	579	1579	2	8	37	6	19	1359	4
$R2$	679	4679	3	27	1	5	269	269	8
$R3$	568	1568	156	9	23	4	126	12356	7
$R4$	2567	3	9	2457	2457	1	8	2456	256
$R5$	2578	12578	157	6	24578	3	1249	12459	1259
$R6$	2568	12568	4	25	258	9	3	7	1256
$R7$	3	5679	567	45	45	2	1469	14689	169
$R8$	1	256	56	3	9	8	7	246	26
$R9$	4	29	8	1	6	7	5	239	239

At this stage of the solution process, the locations of 2, 3, 5, and 9 as candidates for various vacant cells on the Sudoku grid are as follows, with the entries at relevant locations shown in parentheses (in the article only):

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$									
$R2$				2			2	2	
$R3$					2		2	2	
$R4$	2			2	2			2	2
$R5$	2	2			2		2	2	2
$R6$	2	2		2	2				2
$R7$									
$R8$		(2)						2	2
$R9$		(2)						2	2

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$					(3)			(3)	
$R2$									
$R3$					(3)			(3)	
⋮									
$R7$									
$R8$									
$R9$								3	3

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	5	5						(5)	
$R2$									
$R3$	5	5	5					(5)	
$R4$	5			5	5			5	5
$R5$	5	5	5		5			5	5
$R6$	5	5		5	5				5
$R7$		5	5	5	5				
$R8$		5	5						
$R9$									

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	9	9					9	9	
$R2$	9	9					9	9	
$R3$									
$R4$									
$R5$							9	9	9
$R6$									
$R7$		(9)					9	9	9
$R8$									
$R9$		(9)						9	9

What is common about the pairs of vertically positioned vacant cells on the four grids above, all with the corresponding specific candidates shown in parentheses, is that each pair reveals the only feasible locations for a specific candidate in a given box. Thus, for the same ideas as shown in Example 5 in Section 2, we can rule out the candidacy of 2 for $R5C2$ and $R6C2$, 3 for $R9C8$, 5 for $R4C8$ and $R5C8$, and 9 for $R1C2$ and $R2C2$.

The numbers 4 and 5 are locked in $R7C4$ and $R7C5$, and thus they cannot be candidates for any other cells in row 7. Likewise, in box 3, as 3 and 5 must eventually be placed in $R1C8$ and $R2C8$, the candidacy of all other integers for the two cells can be ruled out. The locked integers 2 and 9 in $R9C2$ and $R9C8$ also lead to the determination of $R9C9$. These three situations are similar to those illustrated in Examples 3 and 4 in Section 2.

As the process continues, the candidacy of 4 in $R7C7$ is ruled out, and the default location for 4 in column 7 becomes $R5C7$. After some corresponding changes to remove more redundant candidates, the grid appears as follows:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	579	157	2	8	37	6	19	35	4
$R2$	679	467	3	27	1	5	269	269	8
$R3$	568	1568	156	9	23	4	126	35	7
$R4$	2567	3	9	2457	2457	1	8	26	256
$R5$	2578	1578	157	6	2578	3	4	129	1259
$R6$	2568	1568	4	25	258	9	3	7	1256
$R7$	3	679	67	45	45	2	169	1689	169
$R8$	1	256	56	3	9	8	7	246	26
$R9$	4	29	8	1	6	7	5	29	3

The locked integers 2, 6, and 9 in $R2C8$, $R4C8$, and $R9C8$ in column 8 allow $R5C8$, $R7C8$, and $R8C8$ to be determined. These simple steps, in turn, lead to the determination of $R6C2$, the revelation of locked integers 2, 5, and 6 in $R4C8$, $R4C9$, and $R6C9$ of box 6, and subsequently the determination of $R5C9$.

Following the same idea as illustrated in Example 1 in Section 2, we can solve $R2C2$, $R3C3$, and $R1C7$ as well. Now, with 3, 5, and 7 being locked integers in $R1C2$, $R1C5$, and $R1C8$ of row 1, $R1C1$ can also be solved. After some obvious steps, we have the following:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	9	57	2	8	37	6	1	35	4
$R2$	67	4	3	27	1	5	269	269	8
$R3$	568	568	1	9	23	4	26	35	7
$R4$	2567	3	9	2457	2457	1	8	26	256
$R5$	2578	578	57	6	2578	3	4	1	9
$R6$	2568	1	4	25	258	9	3	7	256
$R7$	3	679	67	45	45	2	69	8	1
$R8$	1	256	56	3	9	8	7	4	26
$R9$	4	29	8	1	6	7	5	29	3

At this stage, the candidacy of the integer 6 for the remaining vacant cells is shown in the following:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$									
$R2$	6						6	6	
$R3$	6	6					6		
$R4$	(6)							6	6
$R5$									
$R6$	(6)								6
$R7$		6	6				6		
$R8$		6	6						6
$R9$									

As a 6 (shown in parentheses here) must eventually be placed in $R4C1$ or $R6C1$, we can rule out the candidacy of 6 for $R2C1$ and $R3C1$. With $R3C2$ being its default position in box 1, its candidacy for $R3C7$, $R7C2$, and $R8C2$ can be ruled out as well. The remaining steps are straightforward. The following is the final solution:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	9	5	2	8	7	6	1	3	4
$R2$	7	4	3	2	1	5	6	9	8
$R3$	8	6	1	9	3	4	2	5	7
$R4$	2	3	9	7	4	1	8	6	5
$R5$	5	8	7	6	2	3	4	1	9
$R6$	6	1	4	4	8	9	3	7	2
$R7$	3	7	6	4	5	2	9	8	1
$R8$	1	2	5	3	9	8	7	4	6
$R9$	4	9	8	1	6	7	5	2	3

4 Concluding remarks

The educational role of the number-placement game Sudoku has been confirmed by various academic studies in the recent literature. Notwithstanding the general perception that Sudoku does not require mathematics, the game has actually raised many intriguing mathematical issues. In response to the growing attention to Sudoku by educators, this study is intended to use Sudoku as a tool to help students develop valuable skills in logical deduction. It draws on the contributions of two recent articles in *Spreadsheets in Education*, by Sugden (2007) and Luterbach, Rodriguez, and Milling (2010). Specifically, by using Excel-based Sudoku puzzles, it makes the game more suitable for use in classroom settings.

The appeal of Sudoku can be attributed by both its exceptionally simple rules for beginners and various sophisticated techniques to sustain the interest of experienced players. As solving a Sudoku puzzle only requires a player's ability to enumerate numbers or objects, the game can be introduced to students at any grade level. From a pedagogic perspective, illustrative examples such as those in Section 2 of this study are suitable for students at higher grades. For students at lower grades, however, it is better to use, or at least start with, simpler examples. Regardless of the pedagogic approach for introducing the game to the classroom, it is the idea that no guessing be allowed that ensures students pay special attention to logical reasoning in order to narrow down the candidate list for each vacant cell in the process of solving a given Sudoku puzzle.

Should guessing be always disallowed? To put the question in a proper context, suppose that a student is unable to proceed any further after exhausting all solution methods that are known to the student based on logical deduction. In such a case, should guessing be permissible for the game to proceed? This is a controversial issue among Sudoku players. The proponents of guessing as a technique of the last resort argue that, as a wrong choice between the two final candidates for a vacant cell will eventually lead to violations of the Sudoku rules somewhere on the grid, guessing is not any different from proof by contradiction. As an example, the algorithm presented by Crook (2009) to solve Sudoku puzzles — which requires only the essential Sudoku techniques like those in the first four examples in Section 2 of this study — explicitly allows guessing as the last resort. In contrast, the opponents argue that guessing is fundamentally different from proof by contradiction, as the player is totally unaware of the eventual outcome for placing an unjustified number on the grid. They also argue that guessing, if allowed, would trivialize the elegant game of Sudoku.

The pursuit of elegance in how a Sudoku puzzle is solved does require a considerable time commitment from the player in order to acquire and practise advanced techniques. However, if the placements of individual numbers on the Sudoku grid are always based on logical deduction, solving a challenging puzzle will give the player considerable satisfaction and pride. Indeed, it is the challenge to solve highly difficult puzzles that sustains the interest of many Sudoku enthusiasts.

The issue of whether guessing is a legitimate Sudoku technique is best left for Sudoku experts to decide. For us, as educators, what matters is that students can acquire valuable skills in logical deduction from Sudoku to complement what they can learn in the core mathematics curriculum. We can pre-screen each Sudoku puzzle to ensure that the Sudoku techniques acquired by students in the classroom are sufficient to allow the solution to be reached without guessing. It is hoped that this study can generate further interest in Sudoku among educators. With collective effort, the effectiveness of Sudoku as a teaching tool can be greatly enhanced.

Finally, although this study does not wish to enter the contentious debate on guessing, it must acknowledge that an Excel-based Sudoku puzzle accommodates guessing much more easily than does the paper version. When solving Sudoku puzzles on Excel, backtracking — the process to retract all subsequent steps caused by an incorrect entry to a vacant cell — which requires only some repeated strokes on the “undo” button of Excel, is straightforward. In addition, as the Excel worksheet containing a partially solved Sudoku puzzle can easily be duplicated, a copy of the worksheet made just before placing an unjustified number in a vacant cell can be used to check whether it eventually leads to any Sudoku rule violations.

Acknowledgements

The author wishes to thank the anonymous reviewers for helpful comments and suggestions.

References

[1] Bailey, R.A., Cameron, P.J., and Connelly, R. (2008) Sudoku, Gerechte Designs, Resolutions, Affine Space, Spreads, Reguli, and Hamming Codes. *American Mathematical Monthly*, 115(5): 383-404.

[2] Crook, J.F. (2009) A Pencil-and-Paper Algorithm for Solving Sudoku Puzzles. *Notices of the American Mathematical Society*, 56(4): 460-468.

[3] Crute, T.D., and Myers, S.A. (2007) Sudoku Puzzles as Chemistry Learning Tools. *Journal of Chemical Education*, 84(4): 612-613.

[4] Dahl, G. (2009) Permutation Matrices Related to Sudoku. *Linear Algebra and its Applications*, 430: 2457-2463.

[5] Delahaye, J.-P. (2006) The Science behind Sudoku. *Scientific American*, June: 80-87.

[6] Herzberg, A.M., and Murty, M.R. (2007), Sudoku Squares and Chromatic Polynomials. *Notices of the American Mathematical Society*, 54(6): 708-717.

[7] Lee, N.Y.L., Goodwin, P.G., and Johnson-Laird, P.N. (2008) The Psychological Puzzle of Sudoku. *Thinking & Reasoning*, 14(4): 342-364.

[8] Lorch, C., and Lorch, J. (2008) Enumerating Small Sudoku Puzzles in a First Abstract Algebra Course. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 18(2): 149-157.

[9] Luterbach, K.J., Rodriguez, D., and Milling, S. (2010) A Comparison of Two Methods for Comparing Sudoku Puzzles. *Spreadsheets in Education*, 3(3): Article 1.

[10] Perez, A.L., and Lamoureux, G. (2007) Sudoku Puzzles for First-Year Organic Chemistry Students. *Journal of Chemical Education*, 84(4): 614.

[11] Sander, T. (2009) Sudoku Graphs Are Integral. *The Electronic Journal of Combinatorics*. 16: #N25.

[12] Sugden, S.J. (2007) The Spreadsheet as a Tool for Teaching Set Theory: Part 1 – an Excel Lesson Plan to Help Solve Sudokus. *Spreadsheets in Education*, 2(3): Article 5.

[13] Taalman, L. (2007) Taking Sudoku Seriously. *Math Horizons*, September: 5-9.

[14] Weiss, H.J., and Rasmussen, R.A. (2007) Advanced Lessons on the Craft of Optimization Modeling Based on Modeling Sudoku in Excel. *INFORMS Transactions on Education*, 7(2): 228-237.

[15] Welsh, M.J. (2007) Chemistry of Art and Color Sudoku Puzzles. *Journal of Chemical Education*, 84(4): 610-611.

Appendix

The Appendix presents two interesting Sudoku properties that are suitable for classroom discussions. Depending on the grade levels at which Sudoku is utilized as an educational tool, the two properties can be presented either formally or informally.

Property 1: *On any 3×9 sub-grid of a solved Sudoku puzzle consisting of three adjacent 9-cell rows and three adjacent 3×3 boxes, the same two numbers in each 3-cell row of a box always appear in a different 3-cell row of each of the remaining boxes. Further, if the same three numbers in a 3-cell row of a box appear again in a different 3-cell row of any of the remaining boxes, then all three numbers in each 3-cell row of any box must appear again in the same manner in the remaining boxes.*

Notice that the property also holds for a 9×3 sub-grid consisting of three adjacent boxes in a column. In such a case, we simply substitute, respectively, the words “row” and “rows” with “column” and “columns” in the description of the property. The following is an example of the first part of this property:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	2	4	5	8	3	6	1	9	7
$R2$	8	1	6	7	9	2	5	4	3
$R3$	9	7	3	5	1	4	8	2	6

Here, the same two numbers 4 and 5 appear in row 1 of box 1, in row 3 of box 2, and in row 2 of box 3. Analogous statements can be made about the two numbers 6 and 8 and about the two numbers 7 and 9. The following is an example of the second part of this property:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	1	2	4	6	8	7	3	5	9
$R2$	3	5	9	4	1	2	6	7	8
$R3$	7	6	8	9	5	3	2	4	1

Here, the same three numbers 3, 5, and 9 appear in row 2 of box 1, in row 3 of box 2, and in row 1 of box 3. Analogous statements can be made about the three numbers 1, 2, and 4, as well as three numbers 6, 7, and 8.

To prove this property without any loss of generality, let us place the three numbers 1, 2, and 3 in row 1 of box 1. The exact placement of the three numbers is unimportant here. There are four scenarios in the placement of the three numbers 1, 2, and 3 in box 2. Row 2 of box 2 can have three, two, one, or none of the three numbers, with row 3 of the same box taking up the rest, if any. As we can interchange rows 2 and 3 or boxes 2 and 3 without affecting the validity of the reasoning, we only have to consider the scenarios where either three or two of the three numbers 1, 2, and 3 appear in row 2 of box 2.

If all of 1, 2, and 3 appear in row 2 of box 2, then, by default, they must also appear in row 3 of box 3. Subsequently, if three other numbers, such as 4, 5, and 6, are to be placed in row 2 of box 1, they must also appear in row 1 of box 3 and then in row 3 of box 2. Finally, the remaining vacant cells in each box must be filled with 7, 8, and 9.

The 3×9 sub-grid will look like the following, with the exact placement of the numbers in each row unspecified:

	$C1$ to $C3$	$C4$ to $C6$	$C7$ to $C9$
$R1$	123	789	456
$R2$	456	123	789
$R3$	789	456	123

Here, each entry actually represents three numbers that are placed in each 3-cell row of the corresponding box. In this scenario, the second part of the property is confirmed. Notice that, in this scenario, the same two numbers in each 3-cell row of a box, as indicated in the first part of the property, can be any two of the three numbers there.

In the remaining scenario where only 1 and 2 are to be placed in row 2 of box 2, we have the following:

	$C1$ to $C3$	$C4$ to $C6$	$C7$ to $C9$
$R1$	123		
$R2$		12	3
$R3$		3	12

The two numbers in the remaining cells of row 2 of box 3, such as 4 and 5, must appear in row 3 of box 1 and in row 1 of box 2, as shown in the following:

	$C1$ to $C3$	$C4$ to $C6$	$C7$ to $C9$
$R1$	123	45	
$R2$		12	345
$R3$	45	3	12

Subsequently, the number in the remaining cell in row 2 of box 2, such as 6, must also be in the remaining cell in row 3 of box 1 and in row 1 of box 3, as shown in the following:

	$C1$ to $C3$	$C4$ to $C6$	$C7$ to $C9$
$R1$	123	45	6
$R2$		126	345
$R3$	456	3	12

If a number such as 7 appears in remaining cell of row 1 of box 2, it must also appear in the remaining cell in row 3 of box 3 and in row 2 of box 1, as shown in the following:

	$C1$ to $C3$	$C4$ to $C6$	$C7$ to $C9$
$R1$	123	457	6
$R2$	7	126	345
$R3$	456	3	127

With the remaining cells in each box filled with 8 and 9, we have the following:

	$C1$ to $C3$	$C4$ to $C6$	$C7$ to $C9$
$R1$	123	457	689
$R2$	789	126	345
$R3$	456	389	127

As each pair of numbers among $(1, 2)$, $(4, 5)$, and $(8, 9)$ always appears in the same row of each box, the first part of Property 1 is confirmed. Notice that, in the above proof, the numbers $1, 2, 3, \ldots, 9$ that are placed successively on the 3×9 sub-grid can also be any permutations of these nine numbers. The 3×9 sub-grid can also cover boxes 4 to 6 or boxes 7 to 9. The proof is analogous if the property pertains to any three vertically stacked boxes.

Property 2: *If each of the vacant cells of a partially solved Sudoku puzzle has exactly two candidates, none of which can be ruled out by logical deduction, then the puzzle has either no solution or non-unique solutions.*

This property pertains to an issue that has attracted the attention of Sudoku experts and enthusiasts. Although each Sudoku puzzle as provided in daily newspapers or posted on websites always has a unique solution, this uniqueness, however, is not a universal feature of the Sudoku game. The following 3×9 sub-grid of an almost completed Sudoku board is an illustration of the issue:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	8	1	6	4	9	3	7	2	5
$R2$	7	9	3	5	2	18	4	18	6
$R3$	4	5	2	7	6	18	9	18	3

Here, like the examples in Section 2, each single-digit entry is a given or solved number, and each double-digit entry captures the two candidates for the corresponding vacant cell. The rest of the Sudoku board, which is not shown here, has been solved. It is clear that this Sudoku puzzle does not have a unique solution, as there are two ways to place candidates 1 and 8 in the four vacant cells.

The following is another example:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	59	59	4	8	7	2	3	1	6
$R2$	8	2	6	3	4	1	57	9	57
$R3$	1	3	7	5	6	9	8	2	4
$R4$	2	6	5	1	8	4	9	7	3
$R5$	7	1	9	2	5	3	6	4	8
$R6$	4	8	3	6	9	7	1	5	2
$R7$	3	59	1	7	2	6	4	8	59
$R8$	59	47	2	49	3	8	57	6	1
$R9$	6	47	8	49	1	5	2	3	79

Here, the number of vacant cells cannot be reduced any further by logical deduction. What is peculiar here is that each of the remaining vacant cells has exactly two candidates.

To see this feature more clearly, let us filter out the given and solved cells, as the following sub-grid shows:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	59	59							
$R2$							57		57
$R3$									
⋮									
$R7$		59							59
$R8$	59	47		49			57		
$R9$		47		49					79

There are two solutions to this puzzle; they can be reached by arbitrarily selecting one of the two candidates for any vacant cell. For example, if we place a 5 in $R1C1$ to start the process to eliminate redundant candidates, a solution will be reached; if we place a 9 there instead, an alternative solution will be reached. The example illustrates that, if each vacant cell of a Sudoku puzzle has exactly two candidates, its solution can be non-unique.

This property can easily be proven once we recognize the fact that both candidates for each vacant cell are already consistent with the given and solved numbers on the Sudoku grid. As neither candidates can be ruled out by logical deduction, there will not be any violations of the Sudoku rules between each arbitrarily chosen candidate and the given and solved numbers on the Sudoku grid. An arbitrary choice between the two candidates for any vacant cell will lead to a series of default entries to the remaining vacant cells. If there are no violations of the Sudoku rules, both solutions are acceptable. Otherwise, the Sudoku puzzle in question does not have a solution.

Why is this property interesting? As Sudoku puzzles are intended to be solved by logical deduction without guessing, the situation where each of the remaining vacant cells has exactly two candidates does create a dilemma; in such a situation, guessing is required for the game to proceed, although the end results, if any, will follow immediately. As explained below, in similar situations where all but one vacant cells have exactly two candidates for each cell, this property actually allows us to rule out candidates in the exceptional cell. As such situations are seldom encountered by casual players of the game, it is useful for teachers to show students with specific examples how logical deduction can be utilized to ensure the attainment of the solutions without guessing.

The following is such an example:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R1$	8	9	2	4	6	3	5	1	7
$R2$	6	3	7	5	8	1	9	4	2
$R3$	4	1	5	7	9	2	6	3	8
$R4$	5	248	1	38	34	9	47	27	6
$R5$	9	6	48	2	7	48	3	5	1
$R6$	7	24	3	1	5	6	8	29	49
$R7$	2	48	9	6	1	5	47	78	3
$R8$	3	7	48	9	2	48	1	6	5
$R9$	1	5	6	38	34	7	2	89	49

To see the placements of all vacant cells more clearly, let us filter out the given and solved cells to produce the following sub-grid:

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$	$C7$	$C8$	$C9$
$R4$		248		38	34		47	27	
$R5$			48			48			
$R6$		24						29	49
$R7$		48					47	78	
$R8$			48			48			
$R9$				38	34			89	49

Here, among the 19 vacant cells, $R4C2$ is the only cell not having exactly two candidates; it has three instead. If we remove the 4 from the candidate list 2, 4, and 8 in $R4C2$, then each of the three vacant cells in column 2 and in box 4 will have a distinct pair of candidates among 2, 4, and 8, and each of the five vacant cells in row 4 will have a distinct pair of candidates among 2, 3, 4, 7, and 8. As none of these candidates can be ruled out by logical deduction, the situation will be the same as what is described in Property 2.

Notice that having 2 and 8 as the candidates for $R4C2$ implies that one of them will eventually be placed there. However, as the placement of either of them there will lead to a non-unique solution or no solution, they cannot be acceptable if a unique solution is sought. Thus, the 4 on the candidate list 2, 4, and 8 must be the default choice for $R4C2$. With this cell solved, the remaining steps are trivial.[12]

[12] As 4 is the correct choice for $R4C2$ in the example, choosing either 2 or 8 instead will lead to violations of Sudoku rules among some of 19 cells on the grid. In the language of some Sudoku experts and enthusiasts, the situation as described in Property 2 is called a *Bivalue Universal Grave (BUG)*, and thus the technique used in the example is sometimes called *BUG Removal.*