Wie viele Zahlen müssen es sein?

Untersuchung zur Lösbarkeit von Sudokus

Ganz offensichtlich macht es vielen Menschen Spaß, Sudokus zu lösen. Diese gibt es in unterschiedlichen Schwierigkeitsgraden. In der Regel sind Sudokos umso schwieriger zu lösen, je weniger Zahlen am Anfang gegeben sind. Ich habe mich nun mit der Frage beschäftigt, wie viele Zahlen mindestens gegeben sein müssen, damit ein Sudoku eindeutig lösbar ist.

1 Einleitung

1.1 Fragestellung

Zurzeit werden viele Berichte über Sudokus veröffentlicht, doch die Frage nach einer Mindestanzahl gegebener Zahlen für eine eindeutige Lösung ist noch ungeklärt. Trotz des Erstellens tausender Sudokus fand Gordon Royle kein eindeutig lösbares Sudoku mit weniger als 17 Zahlen [1]. In dieser Arbeit beschäftige ich mich damit, wie und unter welchen Voraussetzungen eine eindeutige Lösung eines Sudokus gefunden werden kann.

1.2 Spielregeln

Ein Standardsudoku besteht aus einem Gitter mit 9 Zeilen und 9 Spalten. Das Gitter selbst ist in weitere 9 Unterquadrate eingeteilt. Es tauchen Zahlen von 1 bis 9 auf. Die Vorgabe ist, dass in jeder Zeile, jeder Spalte und in jedem Quadrat die Zahlen von 1 bis 9 jeweils nur einmal auftauchen. Ein Beispiel für ein Sudokurätsel ist in Abbildung 1 dargestellt. Derartige Zahlenrätsel können streng genommen nur dann als Sudoku bezeichnet werden, wenn sie eindeutig lösbar sind, das heißt, dass es nicht mehr als eine Möglichkeit gibt, die leeren Felder eines Sudokurätsels auszufüllen.

Autorin

Ariane Papke, *1990
Wilhelmshaven

Käthe-Kollwitz-Gymnasium,
Wilhelmshaven

Eingang der Arbeit:
Mai 2007

Zur Veröffentlichung empfohlen:
Juli 2007

2 Lösungsmöglichkeiten

2.1 Quadrate absuchen

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

Abb. 1: Beispiel für ein Sudoku-Rätsel

Um ein Sudokurätsel zu lösen, kann man die 3 x 3 Quadrate absuchen. Dazu betrachtet man zwei Quadrate, die in einer Reihe (senkrecht oder waagerecht) sind und die beide zum Beispiel die Zahl 1 gegeben haben (Abbildung 2, [1]). Da diese Zahl in jeder Zeile (Spalte und Quadrat) nur einmal vorhanden sein darf, beschränkt sich die Anzahl möglicher Felder auf maximal drei. Denn wenn diese Zahl bereits in zwei benachbarten Quadraten auftaucht, bleibt nur noch eine Zeile (Spalte) innerhalb der drei Zeilen (Spalten) übrig, in denen die Zahl auftauchen kann. Da jede Zahl nur einmal in einem Quadrat auftaucht, schließen sich sechs Felder dieser Zeile aus und es sind nur noch drei Felder übrig. Die drei Felder, in die die 1 eingetragen werden kann, sind in Abb. 2 schraffiert.

Je nachdem, wie viele Zahlen vorgegeben sind, kann von hier an die Position dieser Zahl genauer festgelegt werden. Sind zum Beispiel zwei der drei Felder bereits mit anderen Zahlen belegt, ist die Zahl in das freie Feld einzutragen, denn sie muss ja genau einmal pro Zeile (Spalte) auftauchen. Gibt es mehrere Möglichkeiten, werden die den Feldern zugehörigen Spalten (Zeilen) betrachtet - denn diese Zahl darf ja auch nur einmal pro Spalte (Zeile) auftauchen (Abbildung 2, [2]). In Abbildung 2 sind die drei Quadrate auf der linken Seite zu betrachten. In der ersten und der dritten Spalte ist die 1 bereits eingetragen, es bleiben nur noch drei Felder in dem untersten Quadrat über, in die die 1 eingetragen werden kann. Da in der vorletzten Zeile der sechsten Spalte aber schon eine 1 steht, kann diese nicht noch einmal in derselben Zeile auftauchen. Das Feld in der zweiten Spalte und der vorletzten Zeile scheidet aus. Es gibt nur noch zwei Möglichkeiten für die Position der Zahl 1.

Die Suche wird auf diese Weise in Längs- und in Querrichtung so lange durchgeführt, bis keine eindeutigen Positionen der Zahlen mehr gefunden werden können.

Auch ohne gemeinsame Zahlen zweier Quadrate, die in einer Reihe liegen, kann man die zugehörige Zahl für ein Feld finden. Man sucht sich eine Zahl aus einem Quadrat aus und betrachtet die Zeilen (Spalten) der anderen beiden Quadrate, in denen diese Zahl nicht vorhanden ist. Ist eine dieser Zeilen (Spalten) in einem Quadrat schon ausgefüllt, kann diese Zahl dort nicht mehr eingetragen werden, sie muss aber in dieser Zeile auftauchen. Also gehört diese Zahl in genau diese Zeile (Spalte) des anderen Quadrates (Abb. 3, [3]). Nachdem man für diese Zahl nur noch maximal drei mögliche Felder pro Unterquadrat hat, sucht man die zugehörigen Spalten (Zeilen) ab.

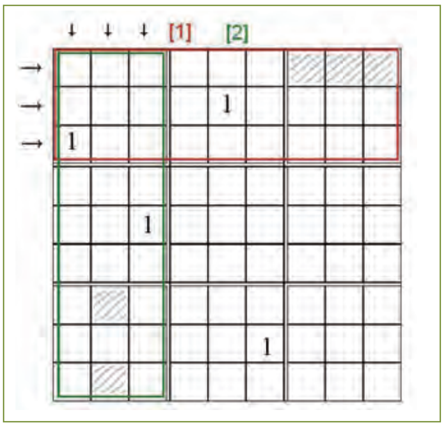

Abb. 2: Gemeinsame Zahlen und mögliche Felder

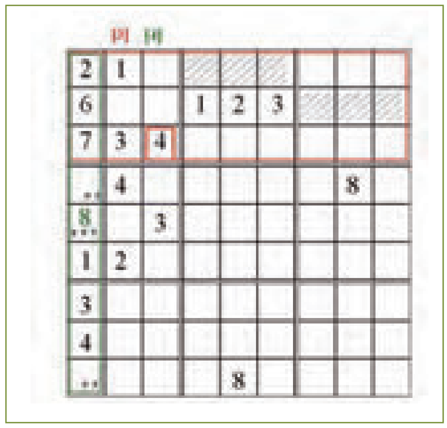

Abb. 3: Ausschlussverfahren

2.2 Einzelne Felder absuchen

Man kann sich auch ein Feld heraussuchen, bei dem viele Felder in der gleichen Zeile, Spalte und dem gleichen Quadrat bereits ausgefüllt sind. Für dieses Feld notiert man sich dann die möglichen Zahlen (in der Abbildung klein eingetragen). Passt eine Zahl nur in ein Feld einer Zeile, Spalte, eines Quadrats, so muss die Zahl in das ausgesuchte Feld gehören (Abbildung 3, [4]).

3 Mathematische Formulierung

Die bisher beschriebenen Strategien sind zwar gut geeignet fürs Rätseln, lassen sich aber so nicht in einem Programm realisieren, das Sudokus löst. Im Folgenden wird eine Methode zur Umsetzung eines Sudokus in ein Programm näher untersucht [2].

3.1 Definition mit Hilfe der Schnittmenge

Innerhalb eines Sudokurätsels tauchen die Zahlen von 1 bis 9 auf, also ist die Menge M dieser Zahlen $M = \{1,\ldots,9\}$. Nun werden noch die Mengen der Zahlen in den Zeilen (Z), Spalten (S) und Quadraten (Q) definiert: Z_z ist die Menge aller Zahlen in jeder Zeile (Abbildung 4, [5]). Analog hierzu ist S_s in Bezug auf die Spalten (Abbildung 4, [6]) definiert. Im Weiteren

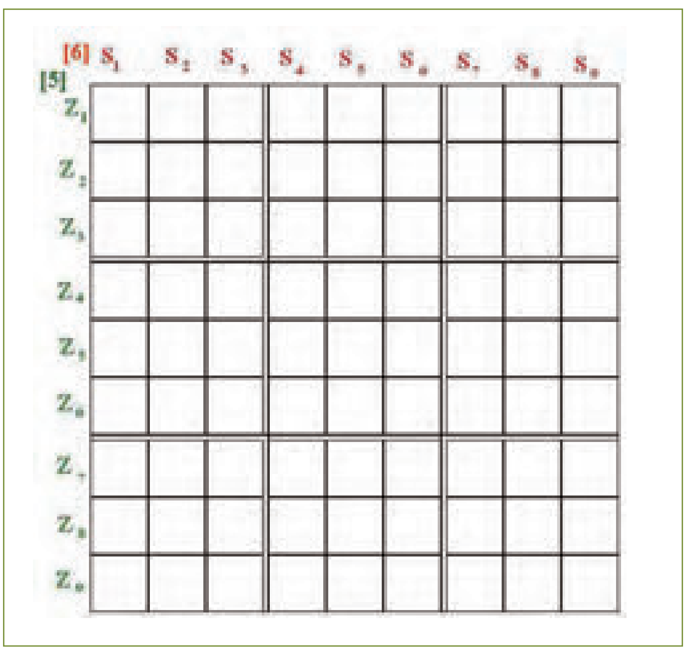

Abbildung 4: Mengen im Sudoku

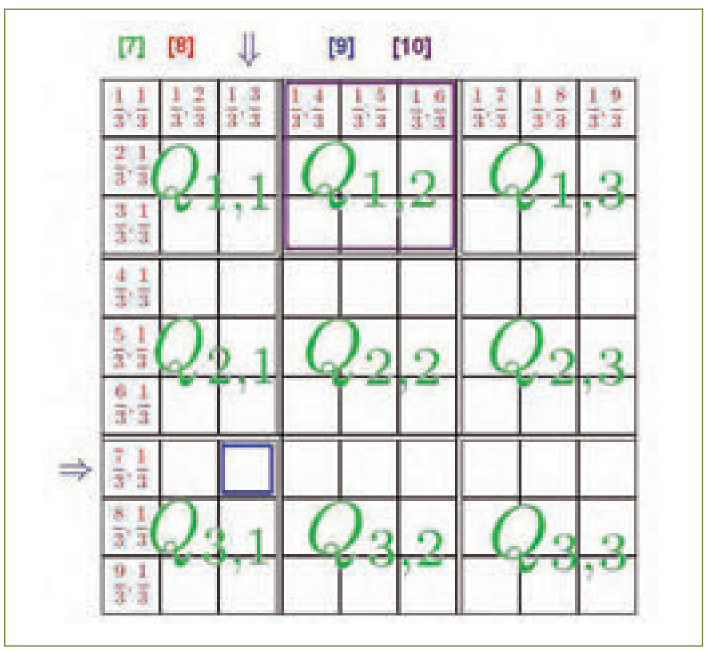

Abbildung 5: Einteilung in Unterquadrate

wird der Index ‚z' für die Zeile, ‚s' für die Spalte verwendet. Ein bestimmtes Feld des Sudokus wird mit $F_{z,s}$ bezeichnet.

Es gibt 9 Unterquadrate eines Sudokurätsels, deren Zahlenmengen mit $Q_{1,1},...,Q_{3,3}$ bezeichnet werden (Abbildung 5, [7]). Also sind je drei Felder eine Einheit, es wird in Dreierschritten vorgegangen, da die Seiten der Quadrate genau drei Felder lang sind. Dabei werden der erste Wert, der Wert der Zeilen, und der zweite Wert, der Wert der Spalten, durch drei geteilt. Allgemein bezeichnet man die Menge der Zahlen der Unterquadrate als $Q_{\lceil \frac{z}{3} \rceil, \lceil \frac{s}{3} \rceil}$ (Abbildung 5, [8]).

Dabei wurden Aufrundungsklammern verwendet, die folgende Bedeutung haben: alle Terme, die sich innerhalb dieser Klammern befinden, werden, sofern nicht ganzzahlig, aufgerundet. Zum Beispiel ist $\lceil 2{,}1 \rceil = 3$ und $\lceil -1{,}9 \rceil = -1$.

Diese Klammern sind notwendig, da es zum Beispiel kein Quadrat gibt, dessen Seitenlänge $\frac{8}{3}$ Einheiten beträgt. Wäre also $\frac{z}{3} = \frac{8}{3}$, dann wäre die Zahlenmenge in $Q_{3,\lceil \frac{s}{3} \rceil}$ gemeint.

Der Quotient muss aufgerundet werden, da die Zahlen von 1 bis 9 durch drei geteilt werden. Dabei muss das Ergebnis jeweils 1, 2 oder 3 sein. Sobald eine Zahl größer als ein Vielfaches von drei ist, befindet sich die zugehörige Zeile (Spalte) bereits in der nächsten Einheit.

So gehört die Zahl in dem Feld der Zeile 7, Spalte 3 zu der Menge der Zahlen des Unterquadrates $Q_{3,1}$ (Abbildung 5, [9]), $Q_{1,2}$ beinhaltet die Menge der Zahlen des zweiten Quadrates von links in der obersten Reihe (Abbildung 5, [10]).

Dann ergibt sich für die Menge $N_{z,s}$ der möglichen Zahlen eines Feldes $F_{z,s}$ die Gleichung

$$N_{z,s} = (M \setminus Z_z) \cap (M \setminus S_s) \cap (M \setminus Q_{\lceil \frac{z}{3} \rceil, \lceil \frac{s}{3} \rceil})$$

Erklärung: $(M \setminus Z_z)$ ist die Menge aller Zahlen, die in M, *aber nicht* in Z vorkommen, also die Differenz von M und Z. Da in M alle Zahlen von 1 bis 9 vorkommen, in Z aber nur gegebene Zahlen der Zeilen, sind in $(M \setminus Z_z)$ alle Zahlen, die nicht in Z vorkommen. Das Gleiche gilt in Bezug auf die Spalten S. Der letzte Teil hat dieselbe Funktion, gilt aber für die Unterquadrate Q.

Entscheidend ist jetzt die Schnittmenge ($\cap$). Werden zwei Mengen „geschnitten", so ist das Ergebnis dieses Schnitts die Menge aller Zahlen, die in beiden Mengen auftauchen. Werden die Mengen Z, S und Q geschnitten, so schneidet man die Menge der Zahlen, die nicht in Z vorkommen mit der Menge der Zahlen, die nicht in S und nicht in Q vorkommen. Das Ergebnis ist dann eine Menge von Zahlen, die weder in S noch in Z noch in Q vorhanden sind. Dies ist die Menge N aller Zahlen, die für das Feld $F_{z,s}$ in Frage kommen.

Beispiel: Für das Feld $F_{5,5}$ vom Sudoku in Abbildung 1 gilt $M = \{1,...,9\}$, $Z_5 = \{1,4,6,9\}$, und $S_5 = \{3,5,6,7\}$. Dann ist

$$N_{5,5} = (M \setminus Z_5) \cap (M \setminus S_5) \cap (M \setminus Q_{2,2})$$

$$= \{2,3,5,7,8\} \cap \{1,2,4,8,9\} \cap \{1,2,3,8,9\}$$

$$= \{2,8\} \cap \{1,2,3,8,9\} = \{2,8\}.$$ Es kommen also nur die 2 und die 8 in Frage.

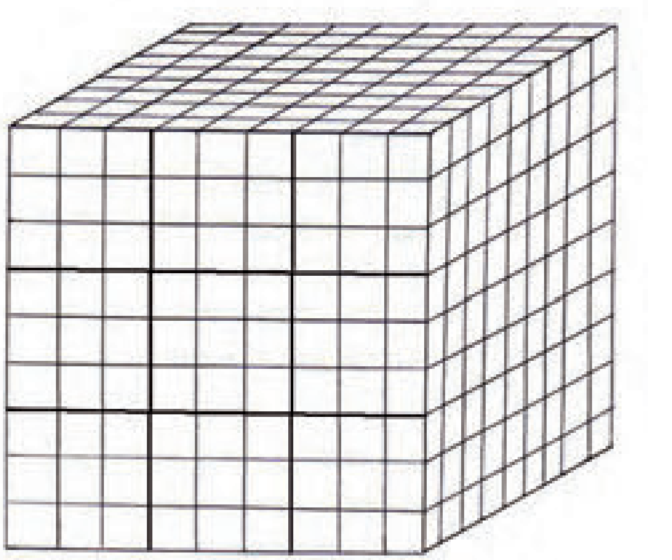

Abbildung 6: Aufbau eines Sudokus, um es mit einem Programm zu lösen

Diese mathematische Definition zur Lösung eines Sudokurätsels kann zur Umsetzung in ein Programm hilfreich sein.

4 Programm zur Lösung

4.1 Grundidee

Ein Lösungsprogramm für Sudokurätsel könnte wie folgt aussehen [3]:

Man stellt sich ein Sudokugitter vor. Es ist zweidimensional. Dahinter befinden sich dann für jedes Feld jeweils die Zahlen von 1 bis 9. Man stellt sich vor, dass die Zahlen mit Hilfe von Würfeln dargestellt werden, die die jeweilige Zahl repräsentieren (Abbildung 6). So hat zum Beispiel der sechste Würfel hinter dem Sudokugitter die Zahl 6. Wird die 6 eingetragen, so erhält der sechste Würfel hinter dem jeweiligen Feld den Wert 1, alle restlichen Würfel hinter dem Feld erhalten den Wert 0 (Abbildung 7).

Also befinden sich in einem leeren Gitter hinter jedem Feld 9 kleine Würfel mit den Zahlen von 1 bis 9. Es entsteht ein dreidimensionaler Körper, ein Würfel mit der Länge 9 (9 Spalten), Höhe 9 (9 Zeilen) und der Breite 9 (9 Zahlen).

Dem Programm muss nun noch beigebracht werden, dass nur einem der neun Würfel hinter einem Feld der Wert 1 zugeteilt werden darf, also nur eine Zahl in ein Feld gehört. Weiterhin sollen nicht zwei Würfel der gleichen Zahl in einer Zeile, Spalte und in einem Quadrat den Wert 1 haben.

Dass ein Würfel nur einmal auftauchen kann, wird im Binärcode angegeben. Nur ein Würfel passt in ein Feld, dieser wird „angeschaltet", indem der Computer ihm die binäre Zahl 1 zuschreibt. Die Null wird den Würfeln zugeschrieben, die mit einer Zahl nummeriert sind, die nicht in das Feld passt. Die Summe aller Würfel in einem Feld ist also 1.

In jeder Zeile, Spalte und jedem Quadrat dürfen nicht mehrere Würfel mit der gleichen Zahl sein, die Summe der Würfel einer Zahl in einer Zeile, Spalte, einem Quadrat ist damit 1.

Das bedeutet: Die Würfel, die sich hinter dem Feld befinden, können in dieses Feld gesetzt werden. Liegt nun aber ein Sudokurätsel vor, so sind bereits Zahlen eingetragen, die wiederum andere Zahlen in den umliegenden Feldern ausschließen.

Um die spezielle Struktur der Unbekannten (Zuordnung der binären Werte) auszunutzen und hinterher ganzzahlige Lösungen für die

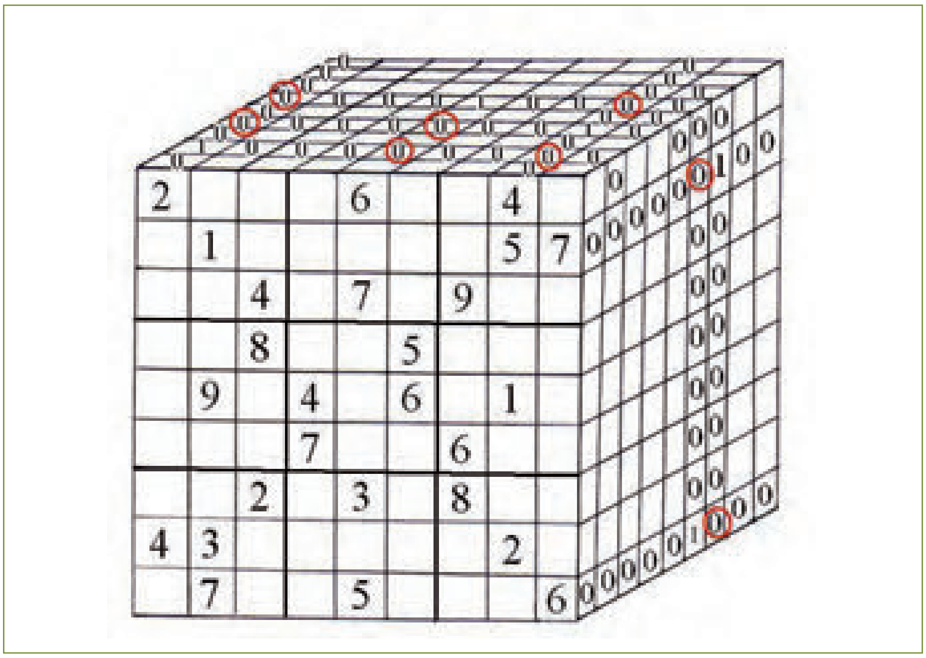

Abbildung 7: Darstellung eines Sudokus, um es zu lösen – Überschneidungen (rot umrandet)

Zahlen im Feld zu erhalten, werden so genannte Ganzzahlige-Programme verwendet. Eine Programmiersprache, die die Ganzzahligkeit ausnutzt, ist ZIMPL.

Man benötigt weiterhin ein Lösungsprogramm (SCIP) und eine Datei, in der man die gegebenen Zahlen bereits definiert. Das Programm hat die Endung *.zpl.

Der Nachteil dieses Lösungsweges ist, dass das Programm immer eine Lösung findet, sofern kein Widerspruch in den gegebenen Zahlen besteht. Damit ist nicht jedes Sudoku, zu dem das Programm eine Lösung findet, eindeutig. Jedoch zeigt das Programm an, ob das vorliegende Sudoku eindeutig ist, indem es die Anzahl der insgesamt eliminierten Unbekannten anzeigt. Ist diese 729 (Anzahl der Würfel und damit Unbekannten in einem leeren Sudoku), so ist das Sudoku eindeutig lösbar.

4.2 Lösung eines Sudokurätsels

Aus den oben genannten Bedingungen, die sich für die Würfel hinter dem Sudokurätsel ergeben, erstellt das Programm mehrere Gleichungen. Die Unbekannten sind die kleinen Würfel, die in das Rätsel eingesetzt werden können.

Das Problem wird von SCIP 0.9 gelöst [7]. Es gibt $9^3 = 729$ Unbekannte, da der große Würfel genau so viele kleine Würfel besitzt. Je nach Anzahl der bereits gegebenen Zahlen gibt es unterschiedlich viele „Constraints“ (Nebenbedingungen). Die Nebenbedingungen entsprechen der Summe aus den Gleichungen und den gegebenen Zahlen.

Die gegebenen Zahlen werden in einer *.dat-Datei tabellarisch aufgelistet. Dabei sind in den ersten zwei Spalten die Zeile und die Spalte der einzutragenden Zahl aufgelistet und in der dritten Spalte befindet sich die Zahl, die in dieses Feld eingesetzt wird.

Im Folgenden wird das Beispielrätsel aus Abbildung 1 gelöst. Es hat 350 Nebenbedingungen, da 324 Gleichungen (wie diese Zahl zu Stande kommt ist in „Herleitung mit Hilfe des Programms“ erläutert) erstellt werden und 26 Zahlen gegeben sind. Nach Optimierung des Problems werden alle Zahlen angezeigt, die eingetragen werden sollen, auch diejenigen, die bereits in der *.dat-Datei angegeben wurden. Dies geschieht in Form von „x#Zeile#Spalte#Zahl“, zum Beispiel gilt für die Zahl, die in das obere linke Feld eingetragen werden soll x#1#1#2. Die Reihenfolge von Zeile und Spalte hängt von der Eingabe in der *.dat-Datei ab; da das Sudoku durch die Vertauschung nur gespiegelt wurde, ergeben sich keine Unterschiede.

					x		a	
		x		a				
	a							x
x			a					
a				x				
						a	x	
	x				a			
			x					a
		a						

Abbildung 8: Beispiel für eine Anordnung von 17 Zahlen

5 Mindestanzahl gegebener Zahlen für eine eindeutige Lösung

5.1 Lösbarkeit von Gleichungssystemen

Gaußsches Eliminationsverfahren

Dieses Verfahren wird verwendet, um nach und nach Unbekannte aus den Gleichungen eines Gleichungssystems zu eliminieren und am Ende den genauen Wert für eine Unbekannte zu erhalten. Ist dies erreicht worden, wird dieser Wert in eine andere Gleichung eingesetzt. Daraus lassen sich nach und nach die Werte der Unbekannten eines Gleichungssystems bestimmen.

Auf die genaue Vorgehensweise wird an dieser Stelle nicht weiter eingegangen [8]. Unter Anwendung dieses Verfahrens lassen sich Aussagen über die Lösungsmenge des Gleichungssystems treffen.

Eine eindeutige Lösungsmenge des Gleichungssystems kann erst dann ermittelt werden, wenn mindestens so viele Gleichungen wie Unbekannte vorliegen. Diese Überlegung wird für die folgende Herleitung verwendet.

5.2 Herleitung mit Hilfe des Programms

Mit den Überlegungen aus dem Kapitel “Programm zur Lösung” umfasst das Sudokurätsel nun im Ganzen $729=9^3$ Unbekannte. In einem ersten Schritt wird nun überlegt, wie viele Gleichungen sich aus den Vorgaben des Sudokus ergeben:

Ein Sudokurätsel umfasst 81 Felder, hinter denen sich je 9 kleine “Würfel” (Unbekannte) befinden, die die Werte 0 oder 1 annehmen können. Jeder Würfel gibt an, ob die jeweilige Zahl im Feld steht (Wert 1) oder nicht (Wert 0).

Die Summe der Unbekannten eines Feldes, also der Summe der 9 Würfel hinter einem Feld, ist gleich der binären Zahl 1. Das heißt, dass nur eine der Zahlen von 1 bis 9 in das Feld eingesetzt wird. Für jede Würfelreihe, also für jedes Feld, wird eine Gleichung erstellt.

Um also die Bedingung zu erfüllen, dass in jedes Feld nur eine Zahl gehört, werden 81 Gleichungen aufgestellt, pro Feld eine Gleichung (man betrachtet die Frontseite des Würfels).

Eine weitere Vorschrift zum Lösen von Sudokus ist, dass keine Zahl zweimal oder mehr in einer Zeile steht. Dazu wird pro Zeile eine Gleichung aufgestellt. Ein Rätsel hat 9 Zeilen und es müssen 9 verschiedene Zahlen eingesetzt werden. Damit werden pro Zahl 9 Gleichungen erstellt, da sie in allen 9 Zeilen je einmal vorkommen muss. Insgesamt entstehen also $9^2 = 81$ Gleichungen (man dreht den Würfel um 90° um die y-Achse).

Ebenso soll beachtet werden, dass jede Zahl pro Spalte nur einmal auftaucht. Analog zur Vorgehensweise bei den Zeilen werden 9 Gleichungen für die Spalten, also erneut 9 Gleichungen pro Unbekannte, benötigt. Durch diese Vorgabe entstehen ebenfalls 81 Gleichungen (man dreht den Würfel um 90° um die x-Achse).

Zuletzt darf in jedem Unterquadrat eine Zahl nur einmal vorkommen.

In einem Unterquadrat gibt es neun Felder, in die neun Unbekannte eingesetzt werden sollen. Da auch hier keine Zahl doppelt auftauchen darf, ist die Summe der Unbekannten, die einen festen und gleichen Wert haben, nämlich 1. Durch diese Vorgabe entsteht pro Unbekannte eine Gleichung, also liefert diese Bedingung 9 Gleichungen pro Unterquadrat. Da es 9 Unterquadrate gibt, entstehen 81 Gleichungen.

Insgesamt ergeben sich $4 \cdot 9^2 = 324$ Gleichungen.

Dieses Gleichungssystem kann nur dann eindeutig lösbar sein, wenn es mindestens so viele Gleichungen wie Unbekannte gibt.

Um herauszufinden, wie viele Zahlen dazu eingesetzt werden müssen, überlegt man sich, wie viele Unbekannte durch Einsetzen einer Zahl wegfallen. Wird eine Zahl eingesetzt, so entfallen 8 kleine Würfel, die sich hinter dem Feld befinden und einen anderen Wert als diese Zahl haben und ein Würfel, also eine Unbekannte, deren Wert eingesetzt wird. Hier entfallen also 9 Unbekannte. Des Weiteren darf die Zahl nicht mehr in derselben Zeile oder Spalte auftauchen. In dieser Zeile und Spalte können je 8 Unbekannte nicht mehr eingesetzt werden, damit entfallen weitere 16 Unbekannte.

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

Abbildung 9: Ungleiche Verteilung

In dem Unterquadrat dieser Zahl können nochmals 4 Unbekannte entfallen, denn vier von den verbleibenden acht Feldern in dem Unterquadrat liegen in derselben Zeile oder Spalte der eingesetzten Zahl. Also fallen durch Einsetzen einer Zahl in das Ausgangssudoku bis zu 29 Unbekannte weg.

Um nun aber genauer betrachten zu können, wie viele Unbekannte durch Einsetzen einer Zahl tatsächlich entfallen können, wird ein Sudoku konstruiert, bei dem es zu möglichst wenig Überschneidungen kommt, das heißt, dass möglichst viele Unbekannte durch Einsetzen einer Zahl entfallen. Diese Konstruktion eines Sudokus wäre also ideal, sodass eine Mindestanzahl gegebener Zahlen unter der Voraussetzung für eine eindeutige Lösung bestimmt werden kann.

Ein Sudoku mit 17 gegebenen Zahlen könnte wie in Abbildung 8 dargestellt aussehen. Im Folgenden werden die entfallenden Unbekannten für dieses Beispiel untersucht:

Zuerst werden 9 Zahlen eingesetzt, die weder eine Zeile, noch eine Spalte noch ein Unterquadrat gemeinsam haben (diese werden mit "a" in Abbildung 8 bezeichnet). Unter der Voraussetzung, dass dies 9 verschiedene Zahlen sind, entfallen damit höchstens $9 \cdot 29 = 261$ Unbekannte, es verbleiben minimal 468 Unbekannte, wenn 9 verschiedene Zahlen eingesetzt werden.

Die Zahlen, die dann eingesetzt werden (also ab der zehnten Zahl), werden mit x gekennzeichnet. Es soll bestimmt werden, wie viele Zahlen noch eingesetzt werden müssen, damit maximal 324 Unbekannte verbleiben, um eine eindeutige Lösbarkeit des Sudokus zu erreichen.

Die Zahlen, die zusätzlich zu den bisher 9 eingetragenen Zahlen eingesetzt werden, haben jeweils genau drei Felder in derselben Zeile, Spalte oder dem Unterquadrat gemein. Entweder hat eine zusätzlich eingetragene Zahl je eine Zahl (a) in derselben Zeile, in derselben Spalte und in demselben Unterquadrat, oder eine Zahl steht in demselben Unterquadrat und in derselben Zeile und eine weitere Zahl steht in der gleichen Spalte.

Pro gemeinsamem Feld entfallen durch Einsetzen einer solchen Zahl zwei Unbekannte weniger, da dem a-ten Würfel des einen Feldes bereits zuvor der Wert 0 zugeschrieben wurde und der Unbekannten des anderen Feldes die 0 zugeteilt wurde, da diese Zahl nur einmal pro Zeile, Spalte und Unterquadrat auftauchen darf, es kommt zu Überschneidungen wie in Abbildung 7 (rot umrandet) dargestellt. Damit entfallen durch Einsetzen einer Zahl x maximal 29-6=23 Unbekannte.

Insgesamt verbleiben dann, wenn 16 Zahlen eingetragen werden, unter der Voraussetzung, dass alle bisher eingetragenen Zahlen a und x jeweils voneinander verschieden sind (was in einem 9 x 9 Sudoku nicht realisierbar ist), Unbekannte mit 308 Gleichungen. Die Anzahl Gleichungen hat sich um die Anzahl der eingesetzten Zahlen verringert, denn den Unbekannten hinter den Feldern wurden bereits Werte zugeordnet und die Gleichungen sind damit wahr. In diesem Fall verblieben 324-16=308 Gleichungen.

Nach diesen Überlegungen gäbe es Sudokus mit 16 gegebenen Zahlen, aber nur, wenn diese eingetragenen 16 Zahlen komplett verschieden wären. Tauchen Zahlen mehr als einmal auf, so entfallen weniger Unbekannte, denn den Unbekannten, die in derselben Zeile, Spalte oder dem Unterquadrat liegen, würde zweimal der Wert 0 zugeordnet werden. Je öfter eine Zahl auftaucht, umso weniger Unbekannte entfallen je weiterer Zahl desselben Wertes.

Ist eine Zahl n zweimal in einem Sudoku vorhanden, so gibt es mindestens ein Feld, deren n-te Unbekannte zweimal den Wert 0 erhält. Also entfallen durch doppelte Zahlen mindestens zwei Unbekannte weniger. Es entfallen genau zwei Unbekannte weniger, wenn sich die Zahlen in Unterquadraten, die nicht nebeneinander liegen, befinden, denn dann haben sie genau zwei gemeinsame Felder. Liegen die beiden Unterquadrate, in denen die gleichen Zahlen eingetragen sind, nebeneinander, so weisen beide Zahlen drei Feldern je eines Unterquadrates zweimal denselben Wert zu (in Abbildung 2 schreiben die Einsen je drei (ersten) Feldern der zwei oberen linken Unterquadrate den Wert 0 zu), also entfallen sechs Unbekannte weniger. Pro doppelter Zahl entfallen zwischen zwei und sechs Unbekannte weniger.

In einem Sudoku müssen mindestens 8 voneinander verschiedene Zahlen gegeben sein, um eine eindeutige Lösbarkeit zu erhalten, da die Zahlen sonst untereinander ausgetauscht werden könnten. Würde eine Zahl neunmal eingetragen werden, so verblieben bei 16 gegebenen Zahlen nur 7 verschiedene weitere Zahlen, es gäbe also, wie gefordert, 8 voneinander verschiedene Zahlen. Da aber dann sehr viele Überschneidungen existieren, entfallen durch jede weitere Zahl

1								3
			1		3			
		3				1		
	1			3				
				1			3	
	3						1	
		1				3		
			3		1			
3								1

Abbildung 10: Geringe Zahlenvielfalt

(Zahl taucht dreimal, viermal,..., neunmal auf) nur noch sehr wenige Unbekannte (analoge Überlegungen ergeben, dass bei einer dreimal eingetragenen Zahl zwischen sechs und achtzehn Unbekannte weniger entfallen). Damit muss es mehr als eine Zahl geben, die mindestens doppelt auftaucht.

Taucht eine Zahl doppelt bzw. dreimal auf (weiterhin vorausgesetzt, dass alle anderen Zahlen voneinander verschieden sind), so verbleiben bereits mindestens $729\text{-}9 \cdot 29\text{-}7 \cdot 23+2=309$ beziehungsweise. $729\text{-}9 \cdot 29\text{-}7 \cdot 23+6=313$ Unbekannte mit 16 eingesetzten Zahlen und 308 Gleichungen, das Gleichungssystem besitzt keine eindeutige Lösungsmenge mehr.

Damit ist kein eindeutig lösbares Sudoku mit 16 gegebenen Zahlen realisierbar. Sind 17 Zahlen gegeben, so verbleiben mindestens $729\text{-}9 \cdot 29\text{-}8 \cdot 23=284$ Unbekannte mit 307 Gleichungen. Tauchen dabei 7 Zahlen doppelt und eine Zahl dreimal auf, so liegen noch weniger Unbekannte als Gleichungen vor, nämlich $284+2 \cdot 7+6=306$ Unbekannte.

Für die Konstruktion bedeutet dies, dass viele eindeutig lösbare Sudokus mit 17 Zahlen je zwei Zahlen in acht Unterquadraten haben und sich eine Zahl alleine in einem Unterquadrat befindet. Von dem Beispiel abweichende Konstruktionen sind ebenfalls möglich. Sind in sieben Unterquadraten je zwei Zahlen, in einem genau eine und in einem genau drei Zahlen eingetragen, so gleicht sich die Zahl entfallender Unbekannter aus. In dem einen Unterquadrat kommt es daher zu weniger Zahlen in derselben Zeile, Spalte, demselben Unterquadrat und in dem Unterquadrat mit drei Zahlen finden mehrere Überschneidungen statt. Somit kann kein Sudoku mit 16 gegebenen Zahlen existieren, es ist aber möglich, ein Sudoku mit mindestens 17 gegebenen Zahlen zu finden, dass eindeutig lösbar ist.

6 Kritische Betrachtung

Fraglich ist, ob 17 Zahlen immer reichen, damit ein Sudoku eindeutig lösbar ist. In den Beispielen auf den Abbildungen 9 und 10 sind sogar 18 Zahlen vorgegeben, trotzdem kann keine eindeutige Lösung gefunden werden.

Wenn es viele Zeilen (Spalten, Quadrate) gibt, die wenig bis gar nicht ausgefüllt sind, kann keine eindeutige Lösung gefunden werden (wären zum Beispiel nur zwei Spalten ausgefüllt, wie in Abbildung 9 gezeigt). Eine gleichmäßige Verteilung erweist sich bei Sudokus als vorteilhaft. Dadurch, dass ganze Spalten ausgefüllt sind, kommt es zu vielen Überschneidungen. Im Mittel sollten nach den vorangegangenen Untersuchungen zwei Zahlen pro Unterquadrat auftauchen, damit eine eindeutige Lösung existiert.

Ein weiteres Beispiel: Wären nur Einsen und Dreien vorgegeben wie in Abbildung 10, könnte keine eindeutige Lösung gefunden werden, es sollten möglichst viele unterschiedliche Zahlen gegeben sein. Insbesondere für die hergeleitete Zahl 17 als Mindestanzahl gegebener Zahlen gibt es Einschränkungen. Es gibt vergleichsweise wenig Sudokus mit 17 gegebenen Zahlen. Dies liegt daran, dass die Anordnung der Zahlen auch die Anzahl der entfallenden Unbekannten bestimmt.

7 Zusammenfassung und Ausblick

Ein Lösungsprogramm erstellt aus den Sudokuregeln Gleichungen und eliminiert schrittweise Unbekannte. Daraus lässt sich ableiten, dass mindestens 17 gegeben sein müssen, damit das Gleichungssystem und damit auch das Sudoku eindeutig lösbar ist.

In weiteren Untersuchungen werde ich mich näher mit der Verteilung der Zahlen und der Anzahl verschiedener Zahlen in einem Sudoku beschäftigen, da diese die Mindestanzahl gegebener Zahlen beeinflussen. Es wird untersucht, wie viele Überschneidungen sich durch Einsetzen einzelner Zahlen ergeben, um die Mindestanzahl mit Hilfe der Mathematik weiter zu erfassen.

Literatur

[1] http://www.spiegel.de/wissenschaft/mensch/0,1518,488244,00.html

[2] http://de.wikipedia.org/wiki/Sudoku#Algorithmisch

[3] Zum Aufbau eines Lösungsprogramms (Volker Kaibel, Thorsten Koch): http://www.zib.de/Publications/Reports/ZR-06-28.pdf

[4] Blum, Wolfgang: Neun Ziffern gegen einen. In: Süddeutsche Zeitung Wissen. 2006, Nr. 12, S. 42-51

[5] Vorgehensweise zur Ermittlung aller möglichen Sudokurätsel (Bertram Felgenhauer, Frazer Jarvis): http://www.afjarvis.staff.shef.ac.uk/sudoku/sudoku.pdf

[6] ZIMPL, Homepage des Konrad-Zuse-Zentrum für Informationen Berlin (ZIB): http://www.zib.de/zimpl

[7] SCIP (Lösungsprogramm): http://scip.zib.de/

[8] Nipp, Kaspar; Stoffer, Daniel: Lineare Algebra: Eine Einführung für Ingenieure unter besonderer Berücksichtigung numerischer Aspekte. Zürich: Verlag der Fachvereine Zürich, 1992

Dank

Für die Unterstützung und Geduld meines Projektbetreuers Ulf Neubacher möchte ich mich an dieser Stelle bedanken. Ebenso gebührt Thorsten Koch und Armin Fügenschuh mein Dank, da sie mich bei der Erstellung eines Lösungsprogramms unterstützt haben.