# Figuring It Out

NUNO CRATO

# FIGURING IT OUT

## ENTERTAINING ENCOUNTERS WITH EVERYDAY MATH

Springer

Nuno Crato
Universidade Técnica de Lisboa
Inst. Superior de Economia e Gestão
Dept. Matemática
Rua Miguel Lupi 20
1200 Lisboa
Portugal
ncrato@iseg.utl.pt

Original Portuguese edition published by Gradiva Publicações lda., Lisboa, Portugal, 2008 original title: A Matemática das Coisas

ISBN 978-3-642-04832-6 e-ISBN 978-3-642-04833-3
DOI 10.1007/978-3-642-04833-3
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2010930692



*Cover design*: eStudio Calamar S.L.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

# Preface

"When I tell people that I am a mathematician, they jokingly ask if I could help them balance their bank account. Then, when I tell them I make lots of counting mistakes, they think I must be a pretty mediocre mathematician."

That is what one mathematician friend of mine once told me, but it could as easily have come from just about any mathematician, as almost everyone in this field complains of how misunderstood the profession is. There really are a lot of people who have no idea what it is that mathematicians do.

Math, of course, is an integral part of our daily lives. The 20th century could not have been the most revolutionary one hundred years in the history of science, as indeed it was, without the extraordinary advances that took place in the field of mathematics. Computers could not have been created without binary logic, group theory and the mathematical concept of information. Telephones would not work if mathematicians hadn't developed the statistical study of signals and the algorithms to digitalize and compress data. Automated traffic lights would no doubt effect chaos, rather than order, if advances in a field of mathematics called Operations Research had not occurred.

But despite its crucial importance, mathematics is frequently viewed as an insular, even irrelevant field into which few interesting people venture, and which has little to contribute to our daily lives. Even the well-educated often demonstrate a surprising ignorance of the history of mathematics and its advancements.

I would venture that if you asked an intellectual to name two or three renowned 20th century philosophers, there would not be many who could not respond without hesitation. I would also say that most reasonably educated people could easily name two or three great contemporary composers. Many of them would also have little difficulty in identifying half a dozen modern schools of art, from cubism to minimalism. But mathematicians and fields of mathematics? Few people know who David Hilbert was or what the formalist school was, or the important part that Andrey Kolmogorov and John von Neumann played with respect to probability studies.

This book is full of stories about math, with few equations, lots of examples and many applications. Math is a fascinating science, of fundamental importance for our history and always present in our daily lives. Many things would not be possible without math: Picasso's art, online bank transactions, house numbers and A4 paper sizes, modern maps and the defeat of Hitler. Math applications appear where you would least expect them. The history of math is the history of winners.

Lisboa, Portugal Nuno Crato

# Contents

**Out of this World**

# Everyday Matters

# The Dinner Table Algorithm

If you want to invite some friends to a dinner party, but your dining table will only accommodate four people, then you might be faced with a dilemma: how do you choose three compatible dinner companions from among your five closest friends? Your buddy Art has recently broken up with his girlfriend Betty, who is now dating Charlie. Charlie and Art have managed to remain friends, but Charlie is not speaking to Dan, who won't go anywhere without Eva, who can't stand Art. So how can you choose your three dinner companions to have a pleasant, hassle-free evening? The best way, believe it or not, would be to make use of an algorithm, which is a set of rules that enable you to search systematically for an answer.

Algorithms are much loved by mathematicians as well as computer scientists. Even though some algorithms are very complex, the simplest can sometimes be the most effective. In our case we can follow a systematic process of trial and error, which may be quite an efficient algorithm, despite its apparent simplicity.

So let us start by choosing your friend A. Under the circumstances, we immediately see that you cannot possibly also invite your friend B. You could invite C, but he wouldn't come unless B was also invited. And so it goes on. It seems there is no scenario under which A could be included, which means we need to start again, this time with B, and keep going until we have found three companionable friends for the dinner party. Will that be possible in this case? Or will we have to give up, forced to admit that human relationships are more complicated than algorithms?

N. Crato, *Figuring It Out*, DOI 10.1007/978-3-642-04833-3_1, 

This type of problem is known as a satisfiability problem. Mathematicians call them SAT problems, which keeps things simpler. The dinner party mentioned above is an example of a "2-SAT" problem, as each restriction contains two variables ("A or B", "A and C", etc.). The problem would become more complicated if Art, Charlie and Dan were inseparable, i.e., if we had to take three variables in each restriction into account ("A and B and C" or "A and C or D").

Such problems are known as "3-SAT" problems. And it is also possible to imagine restrictions of a more general type, which give rise to "*k*-SAT" problems.

Although this example may seem trivial, a similar approach can be applied to many basic tasks, such as drawing up timetables in large schools, organizing conferences, or planning flight schedules for airlines. It is the basis of a new branch of mathematics called computational complexity, which aims to study and classify problems in terms of their inherent difficulty. When such SAT problems could only be solved manually, one after the other, it was difficult to study many of their characteristics. However, from the time it became possible to employ computers to solve them, attempts have been made to study the complexity of the processes used to solve them, i.e., the algorithms, and to evaluate the time that it takes a computer to solve them.

In 1959, Richard Karp was still a 24-year-old mathematician who had just earned a Ph.D. from Harvard and begun to work at the IBM research laboratory at Yorktown Heights, NY. At the time, computers were in their infancy, but the invention of transistors made it possible to incorporate more and more elaborately designed circuits. Karp's task at IBM was to find an automatic process for designing circuits with as few transistors as possible. Written as a computer program, the algorithm he wrote was limited to checking out all the possible circuits and calculating their costs. Later, in 1985, when he was presented with the prestigious Turing Award given by the Association for Computing Machinery, Karp recalled that although this approach seemed simple, it contained a basic problem: "The number of circuits that the program had to comb through grew at a furious rate as the number of input variables increased, and, as a consequence, we could never progress beyond the solution of toy problems."[1] Karp, who spent 10 years more at IBM before becoming a professor at Caltech, had identified a phenomenon that came to occupy the attention of hundreds of researchers and to generate thousands of studies: the problems might well be simple and the technique might be easy to apply, but they could rapidly grow to become impossible to solve, even when using the most powerful computers. Mathematicians, logicians and computer scientists spent many years subsequently trying to devise more efficient algorithms, but always arrived at the same result: there are problems that can be resolved simply and that have a complexity that increases in a controlled fashion, and there are problems that quickly become impossible to solve because their complexity increases exponentially with the number of variables and restrictions.

At present, a distinction is made between the "type P" problems, in which the complexity increases in polynomial time with the rise in the number of variables, and the "non-P" type problems, in which this does not happen. In particular, there is a class of non-P problems that are all reducible to each other and whose solution can be checked in polynomial time. These are the so-called "NP-complete" problems (nondeterministic polynomial). Even though solutions for these problems can be checked efficiently, to find such solutions there are known algorithms that increase dramatically in computing time (more than polynomially) as their dimension grows. These problems thus become impractical when the number of variables increases. It is still not known if type NP-complete problems are amenable to a type P approach. This question was also posed by Karp in 1985 during his Turing Award speech, but even today remains a major unsolved issue in computer science. Specialists assume that these are two different and irreducible types of problems, but they have not been able to prove this yet.

Our dinner table dilemma, which is a 2-SAT problem, is of type P. Even if we had to select thirty persons from a group of 50 instead of having to choose three of our five friends, a computer program could find

[1] From Karp's 1985 Turing Award lecture "Combinatorics, complexity, and randomness" (in http://awards.acm.org)

a solution rapidly or indicate that there is no possible solution, which would be equally important to know.

And if we were, say, holding an event at the UN and had to select 300 persons from a list of 500 possible guests, this would indeed keep the computer busy for a little longer, but we would still have an answer in a reasonable amount of time.

Strangely enough, though, we enter another world entirely when we move on to a 3-SAT problem by inserting restrictions such as "either not including Art and Betty or Charlie". We then cross the line dividing type P problems, for which we will eventually find a solution, from NP-complete problems, when having a few dozen friends is enough to make it impossible for any computer in the world to organize our dinner table in time.

# Cutting the Christmas Cake

When a small cake has to be cut in two pieces to be shared by two people, and the person who cuts the cake is also the person who chooses which half to take, then there is no guarantee that one of the two people will not be disadvantaged. The best way to avoid any complaints about the division of the cake is for one person to cut the cake and the other to choose which half to take. This way, it is in the first person's interest to divide the cake as fairly as possible, as otherwise he or she might very well end up with the smaller piece. It is a wise solution, requiring that two persons, basically motivated by egotism, cooperate with one another in such a way that neither is deprived of a fair share.

This well-known anecdote is applicable to many situations in our day-to-day lives, and not only ones involving cakes. However, the problem becomes more difficult when the cake has to be divided among more than two persons. How would you divide a cake among three people, for instance? Two cut and one chooses? Couldn't two of them conspire to deprive the third of a fair share? And what if many more people wanted a slice of the action? What if a cake had to be shared by twenty equally sweet-toothed persons?

That is not a trivial problem, and mathematicians are beginning to develop algorithms for equal shares. These algorithms can be applied in very diverse areas, ranging from personal matters like the sharing of an inheritance to affairs of state such as establishing international borders.

The "one cuts, the other chooses" algorithm can be directly applied to some situations in which more than two people are involved. If four people want a slice of the cake, for example, the algorithm is applied in

N. Crato, *Figuring It Out*, DOI 10.1007/978-3-642-04833-3_2,

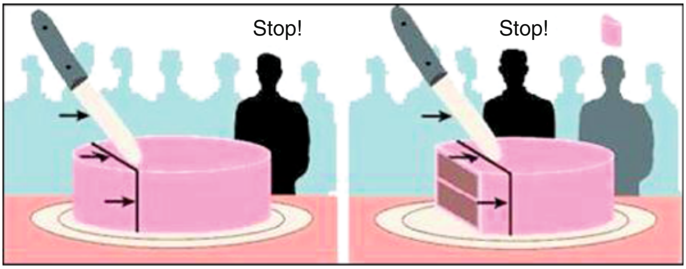
two steps. We start by assigning the four people to two groups, each containing two members. One of the groups then cuts the cake in two, and the other group chooses its half. In the second step each group divides its half of the cake, using the established procedure of one person cutting and the other choosing.

It is easy to see that this repetitive method can also work well with eight participants or indeed with any number that is a power of two. It is not so simple to find a solution when three persons want to share the cake. But if you think carefully about it, you will find a solution to this problem. Can you suggest a way?

However, mathematicians don't like methods that work only in special cases; they prefer to devise algorithms that can be more widely applied. The ideal would be to find methods that could be applied to any number of persons. One of these methods, first proposed by the Polish mathematicians Stefan Banach (1892–1945) and Bronislaw Knaster (1893–1980), resolves the problem utilizing what has been dubbed the "moving knife procedure". It is easier to explain if we take a loaf cake as an example.

The persons who want a slice of the cake gather round it while one of them begins to slide the knife along the cake. The knife keeps moving until one of the participants says "Stop!". At this precise moment the knife stops moving and a slice is cut from the cake and handed to the person who said "Stop". This person then has a slice that he or she considers to be at least a fair share of the cake – if he or she had thought that the knife had not yet traveled far enough to provide a fair share, then he or she would have remained silent. Now the others also had the chance to say "Stop", but they did not do so. So presumably they did not consider that the slice of cake offered was larger than a fair share – otherwise they would have claimed this slice.

After being given a slice, the first participant leaves the game while the knife continues to travel along the cake until one of the remaining participants says "Stop!" and is given the corresponding slice. This process is repeated until only two participants remain in the game. At this stage, the first person to speak receives the slice that is cut and the other receives the remainder of the cake.




The moving knife method can be used to divide a homogeneous cake into equal portions for an arbitrary number of persons. One person moves a knife along the cake until one of the participants says "Stop!" and claims the slice of cake that is cut at that point. The procedure is continued until another participant claims a slice, and so on until the cake has been divided into slices for each person

The interesting thing about this method is that, even considering the fallibility of each of the participants in assessing the right moment to say "Stop", none of them can claim that he or she has been disadvantaged. If any person has not in fact received their fair share, then it is their own fault, as he or she did not speak up at the right time.

This method seems to be perfect, but it fails to take some interesting aspects into consideration. It works well with a homogeneous cake, but would it work with a cake that has various ingredients that are distributed irregularly, like a Christmas cake? Would it be possible to devise an algorithm that guarantees that each person ends up with an equal quantity of glacé cherries, almonds, sultanas and dough? An answer to this question is provided by a theorem the Polish mathematician Hugo Steinhaus (1887–1972) proved in the 1940's and that came to be known by the curious name of the "ham sandwich theorem". Let us take a three-dimensional object with three components such as a sandwich consisting of bread, butter and ham – it does not matter if these components are distributed equally or not, are concentrated in different areas or are spread uniformly. What this theorem proves is that there is always a plane that divides the object in two halves in such a way that each half contains an equal quantity of the three components. In other words, even if the ham or the butter are distributed unequally, there is always a way to cut the sandwich into two completely equal halves.

In the case of a two-dimensional object an equal division works only with two components. Let us suppose that salt and pepper are spread on a table, for example. Steinhaus's theorem proves that there is always a straight line that divides the surface of the table into two sections containing equal quantities of salt and pepper. If there were three ingredients, let us say salt, pepper and sugar, it is easy to imagine a concentration of the substances in three different places so that it would be impossible to draw a straight line that would divide them equally. Generally the theorem states that for *n* dimensions there is always a hyperplane that simultaneously divides *n* components equally. As it seems that we live in a three-dimensional world, and as the Christmas cake has many more ingredients than just three, we have just learned that no knife exists that can cut slices of Christmas cake containing equal quantities of all the ingredients.

# Oranges and Computers

For more than 2000 years mathematics has been making progress by means of rigorous proofs, based on explicit assumptions and logical arguments. The arguments should be faultless. But how can their validity be checked? This has always been the subject of debate and has never been completely resolved. The issue was rekindled at the end of the 20th century, when some prestigious mathematical journals accepted proofs completed with the help of computers. Should these proofs be accepted as legitimate? Should they even be considered mathematical proofs?

One of these disputes involved a well-known and easily understood problem: what is the best way to stack spheres? Is it the way that supermarkets sometimes stack oranges, in little pyramids structured in layers, with each orange sitting in the space between those on the layer below? This system seems more efficient than piling one orange exactly on top of another, for instance. But aren't there other more efficient ways to stack them?

Legend has it that this particular mathematical problem originated in a question that the English explorer Sir Walter Raleigh (1552–1618) posed to the scientist Thomas Harriot (1560–1621). Raleigh was interested in finding a procedure for rapidly estimating the quantity of his munitions. For this purpose, he wanted to be able to calculate the number of cannonballs in each pile simply by inspecting it, without having to count them. Harriot was able to provide him with a correct and simple answer for square pyramidal piles: if each side of the bottom layer of the pile has $k$ cannonballs, then the stack consists of $k(1+k)(1+2k)/6$ cannonballs. So, for instance, if the bottom layer of a square pyramidal

N. Crato, *Figuring It Out*, DOI 10.1007/978-3-642-04833-3_3, 

pile has four balls on each side, then the pile has a total of 30 balls. You could check this yourself by stacking thirty oranges of your own.

Harriot studied various ways of stacking balls. Years later, he brought up the problem in a discussion with the German astronomer Johannes Kepler (1571–1630), who posed an even more interesting question: what is the most efficient way of packing spheres?

Kepler conjectured that the best way would be to put balls in parallel layers, with each layer disposed along a hexagonal grid. Balls on layers below and above should be inserted on the spaces formed by the balls on the other layers. Kepler concluded that there was no better solution than this one but he was unable to prove it mathematically. Centuries passed, and the problem became known as the sphere-packing problem. The astronomer's supposition became known as the Kepler conjecture. It was always admitted that the supposition was true, but nobody ever succeeded in proving it with absolute certainty.

Then in 1998, Thomas C. Hales, a mathematics professor at the University of Michigan, surprised the scientific community by providing a proof. After this, Kepler's conjecture seemed to have ceased being a simple hypothesis and to have become a perfectly proven theorem. However, there was a problem with all this. Just one minor problem. . . the proof had been derived with the help of a computer.

Hales had explicitly resolved many of the steps that were required to prove the hypothesis, but he had left others to be tested automatically using software specially written for this purpose. He claimed that combining the results from the computer with his own work would unquestionably prove the theorem. This was not the first time that a proof had been made with the assistance of a computer. In 1976, Wolfgang Haken and Kenneth Appel, from the University of Illinois, had also used a computer to attain another of the great goals of mathematics – the proof of the four colors theorem, which posits that four colors are sufficient to color a flat map in such a way that no two adjacent regions have the same color. And in 1996 Larry Wos and William McCune, of the Argonne Laboratory in the USA, used logical software to provide proof of another famous supposition, the "Robbins conjecture", a deep statement in mathematical logic.

As soon as Hales announced his achievement, the *Annals of Mathematics*, a prestigious scientific journal, offered to publish his work, but as is usual in academic circles, only after it had been peer-reviewed, that is reviewed by fellow experts. It then took years of work before a panel of 12 experts declared that they had been defeated by the enormity of the task. They confirmed that they were 99% certain that the proof was valid, but they could not succeed in independently verifying all the steps the computer had performed. The editor of the journal regretfully wrote to Hales that while the experts had approached their task with unprecedented vigor, they had become completely exhausted before being able to complete the verification.

The editors of *Annals of Mathematics* did eventually decide to accept the work performed by Hales, though they would only publish those parts that had been verified via explicit logical reasoning, as is normal in the field of mathematics. The computational parts of Hales' proof were published in another, more specialized journal, *Discrete and Computational Geometry*. The provision of computer-generated proof has thus been implicitly admitted into the realm of pure mathematics, but it continues to be regarded with suspicion. Will this ever change?

# When Two and Two Don't Make Four

Two and two always makes four. But the *four* can result from the sum "two plus two" or from the sum "one plus three". It would seem impossible to differentiate between the two fours. However, this problem has a tremendous practical importance for statistics.

In 1919, two American political scientists, William Ogburn and Inez Goltra, published a study on the voting behavior of Oregon women who had recently registered to vote for the first time. The two investigators only knew the total number of votes cast in the election, but had no information on voting patterns according to gender. "Even though the method of voting makes it impossible to count women's votes" they wrote, "one wonders if there is not some indirect method of solving the problem".[1] They decided to estimate the correlation between the number of votes cast in each district with the number of women who had voted in that district. In this way, in the districts with more women, they could attribute the departures from the mean to the higher number of women voters. Still, as the investigators themselves conceded, their method was fallible, as there could have been another explanation: men could have changed their voting habits in those districts that had a greater number of women.

The problem of reconstructing individual behavior from aggregate data came to be known as the *ecological inference problem* (as ecology is the science that is concerned with the relationships between the

[1] W. F. Ogburn and I. Goltra, *Political Science Quarterly* **34**, 413–433, 1919.

N. Crato, *Figuring It Out*, DOI 10.1007/978-3-642-04833-3_4, 

elements and their environment), but very few basic steps were taken to solve it.

Thirty years later an American sociologist named William Robinson published a study that decisively influenced the future methodology of the social sciences. Essentially, Robinson showed that the existing methods at that time did not permit the reconstruction of partial data from aggregate data, and he afterwards coined the expression “ecological fallacy” to describe the faulty inferences that could be drawn as a result. Robinson’s study cast doubt on several strands of sociological investigation. Geopolitical studies, which were flourishing in France, Germany and the USA, practically ground to a halt when the validity of the methods then used was questioned.

However, the ecological inference problem is still a pressing question in applied statistics. The questions posed by the studies are too important for scientists to simply accept that no solution exists. The prime example that is usually cited is the attempt to understand the political and electoral success of the Nazi party in the early 1930s. In this case it is necessary to differentiate between the groups and classes that supported Hitler’s rise to power. The sociologists have based such studies on the data for each electoral district, for which only aggregate data is available. They have no other option.

Another prime example of the importance of ecological inference is taken from epidemiology. The total number of persons affected during an outbreak of disease is often known, but the specific areas of the population that are most affected are often much less evident. The data are aggregated in the hospitals, but in less developed countries it is always very difficult to process them so that the zones where the epidemic is spreading most rapidly can be pinpointed quickly. An efficient method for comparing aggregate data with the existing parceled information (for instance, in some better-organized health centers) could be used to detect the origin of the epidemic and to help save many human lives.

Yet another example comes from marketing. The success or failure of an advertising campaign in attracting new customers can usually be measured, as can the age and income distribution of the target population. Nevertheless, it may be too costly to carry out the research that