THE JOY OF ABSTRACTION

An Exploration of Math, Category Theory, and Life

EUGENIA CHENG

Author of *How to Bake Pi*

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108477222
DOI: 10.1017/9781108769389

First published 2023

Printed in the United Kingdom by TJ Books Limited, Padstow Cornwall

A catalogue record for this publication is available from the British Library.

ISBN 978-1-108-47722-2 Hardback

Contents

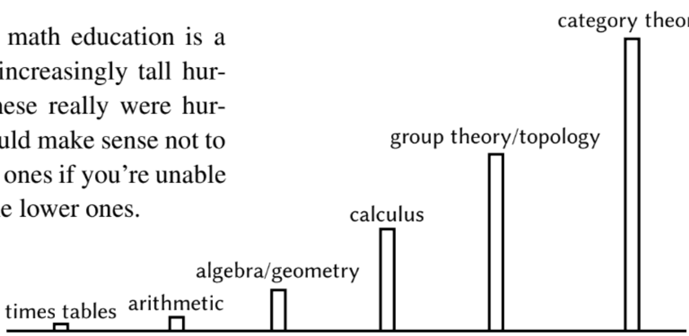
attention on what is relevant for a particular point of view and temporarily disregards the rest so that we can get to the heart of a structure or an argument. In making these connections and finding these deep structures we package up intractably complex situations into succinct units, enabling us to address yet more complicated situations and use our limited brain power to greater effect. This starts with numbers, where instead of saying "1 + 1" all the time we can call it 2, or we fit squares together and call the result a cube, and then build up to more complex mathematical structures as we'll see throughout this book.

This is what I think the power and importance of abstract mathematics are. The idea that it is relevant to the whole of life and thus illuminating for everyone may be surprising, but is demonstrated by the wide range of examples that I have found where category theory helps, despite the field being considered perhaps the "most abstract" of all mathematics. This includes examples such as privilege, sexism, racism, sexual harassment. These are not the sort of contrived real life examples involving the purchase of 17 watermelons, but are *real* real life questions, things we actually do (or should) think about in our daily lives.

If people are put off math then they are put off these ways of thinking that could really intrigue and help them. The sad part is that they are put off an entirely different kind of math usually involving algorithms, formulae, memorization and rigid rules, which is not what this abstract math is about at all. Math is misunderstood, and the first impression many people get of it is enough to put them off, forever, something that they might have been able to appreciate and benefit from if they saw it in its true light.

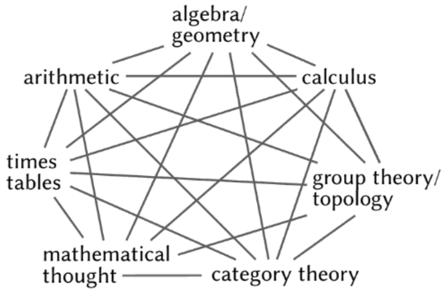
Traditional mathematics: subjects

A typical math education is a series of increasingly tall hurdles. If these really were hurdles it would make sense not to try higher ones if you're unable to clear the lower ones.



"Hurdles" model of math learning.

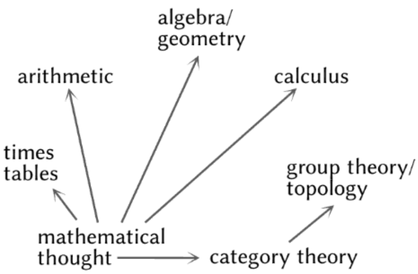
However, math is really more like an interconnected web of ideas, perhaps like this; everything is connected to everything else, and thus there are many possible routes around this web depending on what sort of brain you have.



"Interconnectedness" in math learning.

Some people do need to build up gradually through concrete examples towards abstract ideas. But not everyone is like that. For some people, the concrete examples don't make sense until they've grasped the abstract ideas or, worse, the concrete examples are so offputting that they will give up if presented with those first. When I was first introduced to single malt whisky I thought I didn't like it, but I later discovered it was because people were trying to introduce me "gently" via single malts they considered "good for beginners". It turns out I only like the extremely smoky single malts of Islay, not the sweeter, richer ones you might be expected to acclimatize with.

I am somewhat like that with math as well. My route through the web of mathematics was something like this diagram.



My progress to higher level mathematics did not use my knowledge of mathematical subjects I was taught earlier. In fact after learning category theory I went back and understood everything again and much better.

I have confirmed from several years of teaching abstract mathematics to art students that I am not the only one who prefers to use abstract ideas to illuminate concrete examples rather than the other way round. Many of these art students consider that they're bad at math because they were bad at memorizing times tables, because they're bad at mental arithmetic, and they can't solve equations. But this doesn't mean they're bad at math — it just means they're not very good at times tables, mental arithmetic and equations, an absolutely tiny part of mathematics that hardly counts as abstract at all. It turns out that they do not struggle nearly as much when we get to abstract things such as higher-dimensional spaces, subtle notions of equivalence, and category theory structures. Their blockage on mental arithmetic becomes irrelevant.

It seems to me that we are denying students entry into abstract mathematics when they struggle with non-abstract mathematics, and that this approach is counter-productive. Or perhaps some students self-select out of abstract mathematics if they did not enjoy non-abstract mathematics. This is as if we didn't let people try swimming because they are a slow runner, or if we didn't let them sing until they're good at the piano.

One aim of this book is to present abstract mathematics directly, in a way that does not depend on proficiency with other parts of mathematics. It doesn't have to matter if you didn't make it over some of those earlier hurdles.

Traditional mathematics: methods

When I studied modern languages at school there were four facets tested with different exams: reading, writing, speaking and listening. Of those, writing and speaking are "productive" where reading and listening are "receptive". For full mastery of the language all four are needed of course, but if complete fluency is beyond you it can still be rewarding to be able to do only some of these things. I later studied German for the purposes of understanding the songs of Schubert (and Brahms, Strauss, Schumann, and so on). My productive German is almost non-existent, but I can understand Romantic German poetry at a level including some nuance, and this is rewarding for me and helps me in my life as a collaborative pianist.

I think there is a notion of "productive" and "receptive" mathematics as well. Productive mathematics is about being able to answer questions, say, homework questions or exam questions, and, later on, produce original research. There is a fairly widely held view that the only way to understand math is to work through problems. There is a further view that this is the only way of doing math that is worthwhile. I would like to change that.

I view "receptive" mathematics as being about appreciating math even if you can't solve unseen problems. It's being able to follow an argument even if you wouldn't be able to build it yourself.

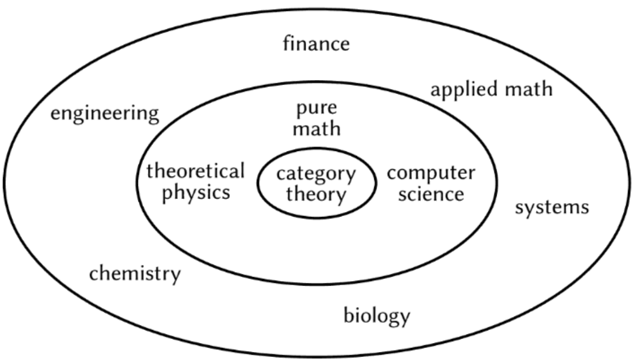
I can appreciate German poetry, restaurant food, a violin concerto, a Caravaggio, a tennis match. Imagine if appreciation were only taught by doing. I can even read a medical research paper although I can't practice medicine. The former is still valuable. In math some authors call this "mathematical tourism" with undertones of disdain. But I think tourism is fine — it would be a shame if the only options for traveling were to move somewhere to live there or else stay at home. I actually once spoke to a representative from a health insurance company who thought this was the case, and did not comprehend the concept that I might visit a different state and ask about coverage there.

One particular feature of this book is that I will not demand that the reader does any exercises in order to follow the book. It is standard in math books to exhort the reader to work through exercises, but I believe this is offputting to many non-mathematicians, as well as some mathematicians (including me). I will provide "Things to think about" from time to time, but these will really be questions to ponder rather than exercises of any sort. And one of the main purposes of those questions will be to develop our instincts for the sorts of questions that mathematicians ask. The hope is that as we progress, the reader will think of those questions spontaneously, before I have made them explicit. Thinking of "natural" next questions is one important aspect of mathematical thinking. Where working through them is beneficial to understanding what follows I will include that discussion afterwards.

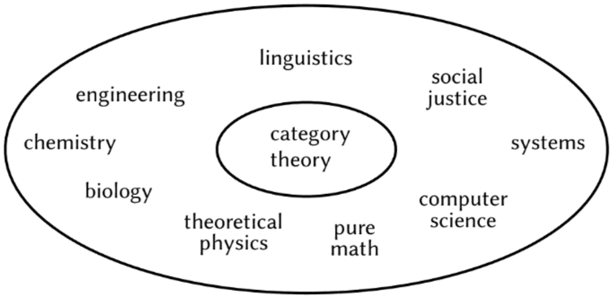
The content in this book

Category theory was introduced by Eilenberg and Mac Lane in the 1940s and has since become more or less ubiquitous in pure mathematics. In some fields it is at the level of a language, in others it is a framework, in others a tool, in others it is the foundations, in others it is what the whole structure depends on.

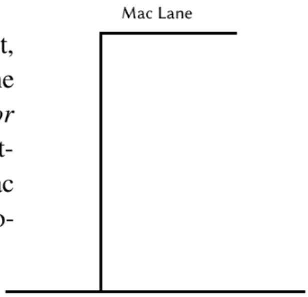
Category theory quickly found uses beyond pure math, in theoretical physics and computer science. The view of things at the end of the 20th century might be regarded like this, with the diagram showing applications moving outwards from category theory:



However, since then category theory has become increasingly pervasive, finding direct applications in a much wider range of subjects further from pure mathematics, such as ecological diversity, chemistry, systems control, engineering, air traffic control, linguistics, social justice. The picture now might be thought of as more like this:



For some time the only textbook on the subject, from which everyone had to try and learn, was the classic graduate text by Mac Lane, *Categories for the Working Mathematician* (from 1971). The situation was this: there was a huge step up to Mac Lane, which many people, even those highly motivated, failed to be able to reach.



As is the way with these things, what started as a research field had become something that graduate students (tried to) study, and eventually it trickled down into a few undergraduate courses at individual universities around the world that happened to have an expert there who wanted to teach it at that level. This spawned several much more approachable textbooks at the turn of the 21st century, notably those by Lawvere and Schanuel (1997), followed by a sort of second wave with Awodey (2006), Leinster (2014) and Riehl (2016). There was still a gap up to those books, and the gap was still insurmountable for many people who didn't have the background of an undergraduate mathematician, either in terms of experience with the formality of mathematics or background for the examples being used.[†]

In 2015 I wrote *How to Bake π*, a book about category theory for an entirely

[†] Lawvere and Schanuel include high school students in their stated target audience but I think they have in mind quite advanced ones. There are also some recent books aimed at specific types of audience, which are less in the vein of standard textbooks; see Further reading.

How to Bake π is not exactly a prerequisite but having read it will almost certainly help.

This material is developed from my teaching art students at the School of the Art Institute of Chicago. Most of the students had bad experiences of school math and many of them either can't remember any of it or have deliberately forgotten all of it as they found it so traumatic. This book seeks to be different from all of those types of experiences. It might seem long in terms of pages, but I hope you will quickly find that you can get through the pages much faster than you can for a standard math textbook. If the content here were written in a standard way it might only take 100 pages. I didn't want to make it shorter by explaining things less, so I have made it longer by explaining things more fully. I will gradually introduce formal mathematical language, and have included a glossary at the end for quick reference. I occasionally include the names of related concepts that are beyond the scope of this book, not because I think you need to know them, but in case you are interested and would like to look them up.

One obstacle to non-mathematicians trying to learn category theory is that the examples used are often taken from other parts of pure mathematics. In this book I will be sure to use examples that do not do that, including examples from daily experience such as family relationships, train journeys, freezing and thawing food, and more hard-hitting subjects such as racism and privilege. I have found that this helps non-mathematicians connect with abstract mathematics in ways that mathematical examples do not. Where I do include mathematical examples I will either introduce them fully as new concepts, or point out where they are not essential for continuing but are included for interest for those readers who have heard of them.

In particular, if you think you're bad at mental arithmetic, terrible at algebra, can't solve equations, and shudder at the thought of sketching graphs, that need not be an obstacle for you to read this book. I am not saying that you will find the book easy: abstraction is a way of thinking that takes some building of ability. We will build up through Part One of the book, and definitely take off in Part Two. It should be intellectually stretching, otherwise we wouldn't have achieved anything. But your previous experiences with math need not bar your way in here as they might previously have seemed to do. Most of all, aside from the technicalities of category theory I want to convey the joy I feel in the subject: in learning it, researching it, using it, applying it, thinking about it. More than technical prowess or a huge litany of theorems, I want to share the joy of abstraction.

PART ONE

BUILDING UP TO CATEGORIES

1

Categories: the idea

An overview of what the point of category theory is, without formality.

I like to think of category theory as the mathematics of mathematics.

I admit this phrase sounds a bit self-important, and it comes with another problem, which is the widespread misunderstandings about what mathematics actually is. This problem is multiplied (or possibly raised to the power of infinity) here by the reference of math to itself.

Another problem is that it might make it seem like you need to understand the whole of mathematics before you could possibly understand category theory. Indeed, that is not far from what the prevailing wisdom has been about studying category theory in the past: that you have to, if not understand *all* of math, at least understand a large amount of it, say up to a graduate level, before you can tackle category theory. This is why category theory has traditionally only been taught at a graduate level, and more recently sometimes to upper level undergraduates who already have a solid background in upper level pure mathematics. The received wisdom is that all the motivating examples come from other branches of pure mathematics, so you need to understand those first before you can attempt to understand category theory.

Questioning "received wisdom" is one of my favorite pastimes. I don't advocate just blindly going against it, but the trouble with received wisdom, like "common sense", is that it too often goes unquestioned.

My experience of learning and teaching category theory has been different from that received wisdom. I did first learn category theory in the traditional way, that is, only after many undergraduate courses in pure math. However, those other subjects didn't help me to understand category theory, but the other way round: category theory was much more compelling to me and I loved and understood it in its own right, whereupon it helped me to understand all those other parts of pure math that I had never really understood before.

I eventually decided to start teaching category theory directly as well, to students with essentially no background in pure mathematics. I am convinced that the ideas are interesting in their own right and that examples illustrating those ideas can be found in life, not just in pure math. That's why I'm starting this book with a chapter about those ideas.

I think we can sometimes unintentionally fall into an educational scheme of believing that we need to learn and teach math in the order in which it was developed historically, because surely that is the logical order in which ideas develop. This idea is summed up in the phrase "ontogeny recapitulates phylogeny", although that is really talking about biological development rather than learning.[†] I think this has merit at some levels. The way in which children grasp concepts of numbers probably does follow the history of how numbers developed, starting with the counting numbers 1, 2, 3, and so on, then zero, then negative numbers and fractions (maybe the other way round) and eventually irrational numbers. However, some parts of math developed because of a lack of technology, and are now somewhat redundant. It is no longer important to know how to use a slide rule. I know very few ruler and compass constructions, but this has not hindered my ability to do category theory, just like my poor skills in horse riding have not hindered my ability to drive a car. Of course, horse riding can be enjoyable, and even crucial in some walks of life, and by the same token there are some specific situations in which mental arithmetic and long division might be useful. Indeed some people simply enjoy multiplying large numbers together. However, none of those things is truly a prerequisite for meeting and benefiting from category theory.

Crucially, I think we can benefit from the ideas and techniques of category theory even outside research math and aside from direct technical applications. Mathematics is among other things a field of research, a language, and a set of specific tools for solving specific problems. But it is also a way of thinking. Category theory is a way of thinking about mathematics, thus it is a way of thinking about thinking. Thinking about how we think might sound a bit like convoluted navel-gazing, but I believe it's a good way of working out how to think better. And in a world of fake news, catchy but contentless memes, and short attention spans, I think it's rather important for those who do want to think better to find better and better ways of doing it, and share them as widely as possible rather than keeping people under a mistaken belief that you have to learn a huge quantity of pure math first.

I have gradually realized that I use the ideas and principles of category theory in all my thinking about the world, far beyond my research, and in areas that probably wouldn't be officially considered to be applications. It is these ideas and principles that I want to describe in this first chapter, before starting to delve into how category theory implements those ideas and how it trains us

[†] Also the phrase was coined by Ernst Haeckel who had some repugnant views on race and eugenics, so I'm reluctant to quote him, but technically obliged to credit him for this phrase.

is to make sure we are always aware of and specific about what context we're considering. This is relevant in all aspects of life as well. For example, the context of someone's life situation, how they grew up, what is going on for them in their personal life, and so on, has a big effect on how they behave, and what their achievements represent. The same achievement is much more impressive to me when someone has struggled against many obstructions in life, because of race, gender, gender expression, sexual orientation, poverty, family circumstance, or any number of other struggles. Sometimes this is controversially referred to as "positive discrimination" but I prefer to think of it as contextual evaluation.

1.4 Relationships

One of the crucial ways in which category theory specifies and defines context is via relationships. It takes the view that what is important in a given context is the ways in which things are related to one another, not their intrinsic characteristics. The types of relationship we consider are often key to determining what context we're in or should be in. For example, in some contexts it matters how old people are relative to one another, but in other contexts it matters what their family relationships are, or how much they earn. But if we're thinking about, say, how good different people will be at running a country, then it might not seem relevant how much money they have relative to one another. Except that in some political systems (notably the US) being very rich seems quite important in getting elected to political office.

There can also be different types of relationship between the same things in mathematics, and we might only want to focus on certain types of relationship at any given moment. It doesn't mean that the others are useless, it just means that we don't think they are relevant to the situation at hand. Or perhaps we want to study something else for now, in something a bit like a controlled experiment. Numbers themselves have various types of relationship with each other. The most obvious relationship between numbers is about size, and so we put numbers on a number line in order of size. But we could put numbers in a different diagram by indicating which numbers are divisible by others. In category theory those are two different ways of putting a category structure on the same set of numbers, by using a different type of relationship. We will go into more detail about this in Chapter 5.

The relationships used in category theory can essentially be anything, as long as they satisfy some basic principles ensuring that they can be organized in a mildly tractable way. This will guide us to the formal definition of a category. To build up to that we will look at the idea of formalism in Chapter 6, to ease into this aspect of mathematics that can sometimes be so offputting. In Chapter 7 we'll look at a particular type of relationship called equivalence relations, which satisfy many good properties making them exceedingly tractable. In fact, they satisfy too many good properties, so they are too restrictive to be broadly expressive in the way that category theory seeks.

We will see that category theory is a framework that achieves a remarkable trade-off between good behavior and expressive possibilities. If a framework demands too much good behavior then expressivity is limited, as in a totalitarian state with very strict laws. On the other hand if there are too *few* demands, then there is great potential for expressivity, but also for chaos and anarchy. Category theory achieves a productive balance between those, in the way it specifies what type of relationship it is going to study.

Part One of the book will build up to the formal definition of a category. We will then take an Interlude which will be a tour of mathematics, presenting various mathematical structures as examples of categories. The usual way of doing this is to assume that a student of category theory is already familiar with these examples and that this will help them feel comfortable with the definition of category theory. I will not do that, but will introduce those examples from scratch, taking the ideas of category theory as a starting point for introducing these mathematical topics instead. In Part Two of the book we will then look more deeply into the sorts of things we do with category theory.

1.5 Sameness

One of the main principles and aims of category theory is to have more nuanced ways of describing sameness. Sameness is a key concept in math and at a basic level this arises as equality, along with the concept of equations. Indeed, many people get the impression that math is *all* about numbers and equations. This is very far from true, especially for a category theorist. First of all, while numbers are an example of something that can be organized into a category, the whole point is to be able to study a much, much broader range of things than numbers. Secondly, category theory specifically does not deal in equations because equality is much too strong a notion of sameness in category theory.

The point is, many things that we write with an equals sign in basic math aren't really equal deep down. For example when we say $5 + 1 = 1 + 5$ we really mean that the outcomes are the same, not that the two sides of the equation are actually completely the same. Indeed, if the two sides were completely the same there would be no point writing down the equation. The whole point is that there is a sense in which the two sides are different and a sense in which the two sides are the same, and we use the sense in which they're the same to pivot between the senses in which they're different in order to make progress and build up more complex thoughts. We will go into this in Chapter 14.

Numbers and equations go together because numbers are quite straightforward concepts,† so equality is an appropriate notion of "sameness" for them. However, when we study ideas that are more complex than numbers, much more subtle notions of sameness are possible. To take a very far opposite extreme, if we are thinking about people then the notion of "equality" becomes rather complicated. When we talk about equality of people we don't mean that any two people are actually the same person (which would make no sense) but we mean something more subtle about how they should be treated, or what opportunities they deserve, or how much say they should have in our democracy. Arguments often become heated around what different people mean by "equality" for people, as there are so many possible interpretations.

Math is about trying to iron out ambiguity and have more sensible arguments. Category theory seeks to study notions of sameness that are more subtle and complex than direct equality, but still unambiguous enough to be discussed in rigorous logical arguments. Sometimes a much better question isn't to ask whether two things are equal or not, but *in what ways* they are and aren't equal, and furthermore, if we look at some way in which they're not equal, how much and in what ways do they fail to be equal? This is a level of subtlety provided by category theory which we sorely need in life too.

1.6 Characterizing things by the role they play

Category theory seeks to characterize things by the role they play in context rather than by some intrinsic characteristics. This is related to the idea of context and relationships being so important. Once we understand that objects take on very different characteristics in different contexts it becomes clearer that the whole idea of intrinsic characteristics is rather shaky.

I think this applies to people as well. I don't think I have an intrinsic personality because I behave very differently depending on what sort of situation I'm in. In some situations I'm confident and talkative, and in other situations I'm nervous and shy. Even mathematical objects do something similar, although in that case the characteristics we're thinking about aren't personality traits, but mathematical behaviors.

† Actually they're very profound, but once they're defined there's not much nuance to them.

For example, we might think the number 5 is prime "because it's only divisible by 1 and itself", but we really ought to point out that the context we're thinking of here is the whole numbers, because if we allow fractions then 5 is divisible by everything really (except 0).†

In normal life we often mix up when we're characterizing things by role and by property in the way that we use language. For example "pumpkin spice" is named after the role that this spice combination plays in classic American pumpkin pie, but it has now come to be used as a flavoring in its own right in any number of things that are not actually pumpkin pie, but it's still called pumpkin spice, which is quite confusing for non-Americans. Conversely "pound cake" is named after the fact that it's a recipe consisting of a pound each of basic cake ingredients. So it's named after an intrinsic property, and it's still called pound cake even if you change the quantity that you use. I, personally, have never made such an enormous cake.

One of the advantages of characterizing things by the role they play in context is that you can then make comparisons across different contexts, by finding things that play analogous roles in other contexts. We will talk about this when we discuss universal properties in Chapter 16. This might sound like the opposite of what I just described, as it sounds a bit like properties that are universal regardless of context, but what it actually refers to is the property of being somehow extreme or canonical within a context. This can tell us something about the objects with that property, but it can also tell us something about the context itself. If we go round looking at the highest and lowest paid employees in different companies, that tells us something about those companies, not just about the employees. It is only one piece of information (as opposed to a whole distribution of salaries across the company) but it still tells us something.

1.7 Zooming in and out

One of the powerful aspects of category theory's level of abstraction is that it enables us to zoom in and out and look at large and small scale mathematical structures in a similar light. It's like a theory that unifies the sub-atomic level with the level of galaxies. This is one of my favorite aspects of category theory.

If we study birds then we might need to make a theory of birds in order to make our study rigorous. However, that theory of birds is not itself a bird — it's one level more abstract. On the other hand if we study mathematical objects then we similarly might need a theory of them. I find it enormously

† Also this is more of a characterization than a definition.

satisfying that that theory is itself also a mathematical object, which we can then study using the same theory. Category theory is a theory of mathematics, but is itself a piece of mathematics, and so it can be used to study itself. This sounds self-referential, but what ends up happening is that although we are still in category theory we find ourselves in a slightly higher dimension of category theory. Dimensions in this case refer to levels of relationship. In basic category theory our insight begins by saying we should study relationships between objects, not just the objects themselves. But what about the relationships? If we consider those to be new mathematical objects, shouldn't we also study relationships between those? This gives us one more dimension.

Then, of course, why stop there? What about relationships between relationships between relationships? This gives us a third dimension. And really there is no logically determined place to stop, so we might keep going and end up with infinite dimensions. This is essentially where my research is, in the field of higher-dimensional category theory, and we will see a glimpse of this to finish the book. To me this is the ultimate "fixed point" of theories. If category theory is a theory of mathematics, then higher-dimensional category theory is a theory of categories. But a theory of higher-dimensional category theory is still higher-dimensional category theory.

This is not just about abstraction for the sake of it, although I do find abstraction fun in its own right. It is about subtlety. Category theory is about having more subtle ways of expressing things while still maintaining rigor, and every extra dimension gives us another layer of possible subtlety.

Subtlety and nuance are aspects of thinking that I find myself missing and longing for in daily life. So much of our discourse has become black-and-white in futile attempts to be decisive, or to grab attention, or to make devastating arguments, or to shout down the opposition. Higher-dimensional category theory trains us in balancing nuance with rigor so that we don't need to resort to black-and-white, and so that we don't *want* to either.

I think mathematics is a spectacular controlled environment in which to practice this kind of thinking. The aim is that even if the theory is not directly applicable in the rest of our lives, the thinking becomes second nature. This is how I have found category theory to help me in everyday life, surprising though it may sound.

1.8 Framework and techniques

As I have described it so far, category theory might sound like a philosophy more than anything else. But the point is that it is only *guided* by these vari-

Every academic discipline provides a way of reaching truths of some form. Each discipline is seeking a particular type of truth, and develops a method or framework for deciding what counts as true. In this era of information excess (and indeed general excess) I think understanding those methods and frameworks is far more important than knowing the truths themselves. The important thing is to know *how* to decide what should count as true — how to build good foundations on which to base our understanding. I strongly believe that this understanding of process and framework is what is most transferrable about studying any subject, especially math.

2.2 The twin disciplines of logic and abstraction

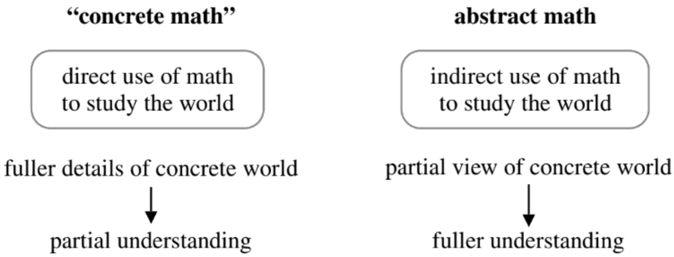
The framework of mathematics involves the twin disciplines of logic and abstraction. Math is not unique in its use of either of these things, but I regard it as being more or less defined by its use of these in combination.

I would say that philosophy uses logic, but applies it to real questions about life experiences. Art uses abstraction, but does not primarily build on its abstractions by logic. Math uses logic and abstraction together. It uses logic to build rigorous arguments, and uses abstraction to ensure that we are working in a world where logic can be made rigorous.

This might make it sound like we can never be talking about the "real" (or rather, concrete[†]) world as we will always be working in an abstract world. While this is in some sense true, it is also reductive. Abstractions are facets of the concrete world, or views from a particular angle. While they will never give us the full explanation of the concrete world, it is still valuable to get a very full understanding of particular aspects of the concrete world. As long as we are clear that each one is only a partial view, we can then move flexibly between those different views to build up a clearer picture.

There is a subtle difference between this and the approach of studying the concrete world directly. In the direct approach we typically get only a partial understanding, because the concrete world is too messy for logic. The following diagram illustrates the difference.

[†] What is real anyway?



Thus abstract math still studies the world, just in a less direct way. Its starting point is abstraction, and the starting point for abstraction is to forget some details about a situation.

2.3 Forgetting details

Abstraction is about digging deep into a situation to find out what is at its core making it tick. Another way to think of it is about stripping away irrelevant details, or rather, stripping away details that are irrelevant to what we're thinking about now. Those details might well be relevant to something else, but we decide we don't need to think about them for the time being. Crucially, it's a careful and controlled forgetting of details, not a slapdash ignoring of details out of laziness or a desire to skew an argument in a certain direction.

If someone says "women are worse at math than men" then they are omitting crucial details and opening up ambiguities, or deliberately using data in a misleading way. This inflammatory statement has some truth *in some sense*, which is that fewer women are currently employed as math professors than men, not that there is any evidence that women are *innately* worse at math than men. It's a pedantically correct expression of the fact that women are currently doing worse in the field of mathematics than men are.[†]

Whereas if we observe that one apple together with another one apple makes two, and that one banana together with another banana makes two, and we say that one thing together with another thing makes two, then we are ignoring the detail of applehood and bananahood as that is genuinely irrelevant to the idea of one thing and another thing making two things. That is abstraction, and is how numbers come into being. Numbers are one of the first abstract concepts we come across, but we don't always think of them as being abstract. That is a good thing, as it shows that we have raised the baseline level of abstraction that we're comfortable with in our heads. This is like the fact that things can

† In *The Art of Logic* I wrote about pedantry being precision without illumination. In this case it's even worse: it's precision with active obfuscation.

seem hard at first, but later seem so easy they're second nature. It is all a sign that we've progressed.

2.4 Pros and cons

Before I go into more detail about how we perform abstraction, I want to talk about the pros and cons. It might be tempting just to talk about all the benefits of doing something, but that can be counter-productive if other people see disadvantages and think you're being dishonest or misleading. Instead, I think it's important to see the pros and cons of doing something, and weigh them up. There are rarely exclusively positives or negatives to doing something.

The advantages of abstraction, as I see them, are broadly that we unify many different situations in a sort of inclusivity of examples; this then enables us to transfer insight across different situations, and thus to gain efficiency in our thought processes by studying many things at once. The world is a complicated place and we need to simplify it in order to be able to understand it with our poor little brains. One popular way to simplify it is to ignore some of the detail, but I think that's a dangerous way of simplifying it. Another way is to make connections inside it so that our brains can deal with more of it at once. I think this is a better way. The best way overall is to become more intelligent so that the world becomes simpler relative to our brains.

So much for the advantages; I will now acknowledge some disadvantages of abstraction. One is that it does take some effort, but I do think this is about front-loading effort in order to reap rewards from it later. I think that it's an investment, and the extra effort *early* means that we can understand more things more deeply with less effort *later*.

Another disadvantage is that we lose details. However, I think this just means we shouldn't remain *exclusively* in the abstract world, but should always bear in mind that at some point some details will need to go back in. Losing the details temporarily is an important part of finding connections between situations, so again I think that this is a net positive.

Another disadvantage is that this takes us further from the normal everyday world that we're used to, which can be scary. It can be scary because it can seem like we don't have our feet on the ground any more. It can be scary because we can't touch things or feel things or see things, and we can't use a lifetime of intuition any more. However, intuition is itself a double-edged sword.† Intuition both helps and hinders us, whether in math or in life. It helps

† I've always found the metaphor of a double-edged sword a bit strange, as it doesn't seem to me that the two edges of a double-edged sword work in opposition to one another.

us in situations where we do not have enough information or time to use logic. It helps us by drawing on our experience quickly. But it is thus also limited by our experience, and if it is used instead of logic it can be dangerously misleading. For example, it's unavoidable to have a gut instinct when we meet a new person, but it's wrong to hold onto that instead of actually responding to the person as we get to know them, especially when our gut instinct is skewed by implicit bias as it (by definition) always is.

Likewise, in math it's not wrong to have intuitions about things, and indeed this is how much research gets started, by a vague idea coming from inside a mathematician's figurative gut somewhere. But the key is then to investigate it using logic and not rely too much on that intuition.

One crucial point is that the framework of building arguments by rigorous logic in math can take us much further than our intuition can. It can take us into places where we have no intuition, such as infinite-dimensional space, or worlds of numbers that have no concrete interpretation in the normal world. For example, one of the points of calculus is to understand what gives rise to interesting features in graphs, like gaps, spikes, places where the graph changes direction. This means we can seek those features even when the graph itself is much too complicated to draw, so that we can't just look for them visually.

One possible objection to this "advantage" is that you might think you'll never find yourself in a place that's so far beyond your intuition. That may well be true. But it might still be good for your brain to be stretched into those places, so that your intuition can develop. I am convinced that my years of stretching my brain in those abstract ways have enabled me to think more clearly about the world around me, and more easily make connections between situations, connections that others don't see. Often when I give my abstract explanations of arguments around social issues such as sexual harassment, sexism, racism, privilege, power imbalance and so on, people ask me how I thought of it. The answer is that a lifetime of developing my abstract mathematical brain makes these things come to me quite smoothly. It's good to train yourself to be able to do more than you think you'll need, so that the things you do need to do feel easier.

The last advantage I want to give for abstraction is that it can be fun. Fun can seem a little frivolous in trying political times, but if we only stress the utility of something it might start sounding awfully boring. I find it enormously satisfying to strip away outer layers of a situation to find its core. It appeals to my general aesthetic, which is typically that I am not so interested in superficial appearances, but care about what is going on in the heart of things, deep down, far below the surface. Abstraction in its own right really does bring me joy.

2.5 Making analogies into actual things

Abstraction comes from seeing analogies between different situations. This is a particular form of detail-forgetting, based in finding connections between different situations, rather than just arbitrarily ignoring details.

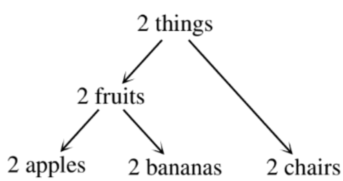
If we say that one situation "is analogous to" another situation, essentially what we are saying is that if we ignore some surface-level details in each situation then the two are really the same. In math, unlike in normal life, we don't just say that the situations are "analogous", but we make very precise what feature is the same in both situations, which is causing the analogy we want to consider. Some of what follows I have also written about in *The Art of Logic*.

If we think about two apples and two bananas, we can consider them to be analogous because they're both examples of two things. But we could also consider them to be analogous because they're both examples of two fruits. Neither of those is "right" or "wrong", neither is "better" or "worse". What we can say, however, is that "two things" is a further level of abstraction than "two fruits" because it forgets more of the details of the situation; conversely, saying "two fruits" is less abstract and leaves us closer to the actual situation.

The more abstract version takes us further away from "reality" (whatever that is), and one major upside is that it enables us to include more distant examples in our analogy. In this example it means we could include two chairs, or two monkeys, or two planets.

In a way, abstraction is like looking deeper into a situation, but it is also like taking a step back and seeing more of the big picture rather than getting lost in the details. The fact that we can find different abstractions of the same thing makes it sound ambiguous but is in fact key. I find it helpful to draw diagrams of different levels of abstraction.

Here is a diagram showing that two apples and two bananas are analogous because they're both two fruits, but also that if we go up further to the level of two things then we get to include two chairs as well.



The key in math is that we don't just say that things are *analogous*. Rather, we precisely specify our level of analogy, and then go a step further and regard that as an object in its own right, and study it. That is how we move into abstract worlds, and it is in those abstract worlds that we "do" math. In the above example that's the level of "2 things": the world of numbers.

Pinning down what is causing the analogy, rather than just saying things are analogous, is like the difference between telling someone there is a path

Things To Think About

T 2.1 What are some senses in which addition and multiplication are "the same"? What are some senses in which they are "different"?

Addition and multiplication are both *binary operations*: they take two inputs and produce one answer at the end. In the first instance they are binary operations on numbers, but as math progresses through different levels of abstraction we find ways of defining things like addition and multiplication on other types of mathematical object as well.

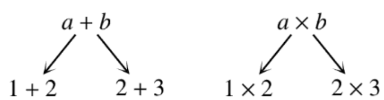
As binary operations, they have some features in common, including that the order in which we put the numbers doesn't matter (which is *commutativity*) and nor does how we parenthesize (which is *associativity*). However, addition and multiplication behave differently in various ways. For example, addition can always be "undone" by subtraction, but multiplication can only sometimes be "undone" by division: multiplication by 0 can't be undone by division. This is sometimes thought of as "you can't divide by 0" but we'll come up with some better abstract accounts of this later.

2.7 Abstraction journey through levels of math

As mathematics progresses, aspects of it become more and more abstract. There is a sort of progression where we move through levels of abstraction gradually, through the following steps:

1. see an analogy between some different things,
2. specify what we are regarding as causing the analogy,
3. regard that thing as a new, more abstract, concept in its own right,
4. become comfortable with those new abstract concepts and not really think of them as being that abstract any more,
5. see an analogy between some of those new concepts,
6. iterate...

One of the advantages of taking abstract concepts seriously as new objects is that we can then build on them in this way. Here's an example of an initial process of abstraction in basic math.



This is the infamous process of "turning numbers into letters", which is the stage of math many people tell me is where they hit their limit.

(In fact there was a level below that, where we went from objects like apples and bananas to numbers in the first place.) Why have the numbers turned into letters? It's so that we can see things that are true abstractly, regardless of what exact numbers we're using. For example:

$$\begin{aligned} 1+2 &= 2+1 \\ 1+3 &= 3+1 \\ 2+3 &= 3+2 \\ 5+7 &= 7+5 \\ &\vdots \end{aligned}$$

Something analogous is going on in all of these situations, and it would be impossible for us to list all the combinations of numbers for which this is true as there are infinitely many of them. We could describe this in words as "if we add two numbers together it doesn't matter what order we put them", but this is a bit long-winded.

The concise abstract way of saying it is: given any numbers a and b,

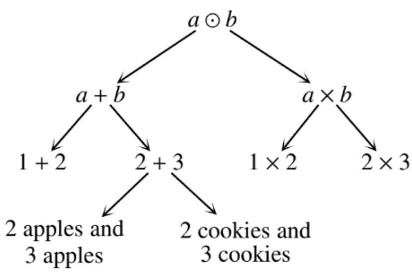
$$a + b = b + a.$$

We have "turned the numbers into letters" so that we can make a statement about tons of different numbers at once, and make precise what pattern it is that is causing the analogy that we see. Not only is this more concise and thus quicker to write down (and mathematicians are very lazy about writing down long-winded things repeatedly), but the abstract formulation can help us to go a step further and pin down similarities with other situations.

But there's a level more abstract as well, in the direction we were going at the end of the previous section. If we think about similarities between addition and multiplication we see that they have some things in common. For a start, they are both processes that take two numbers (at a basic level) and use them to produce an answer. The processes also have some properties that we noticed, such as commutativity and associativity.

When we call them a "process that takes two numbers and produces an answer" that is a further level of abstraction. It's an analogy between addition and multiplication.

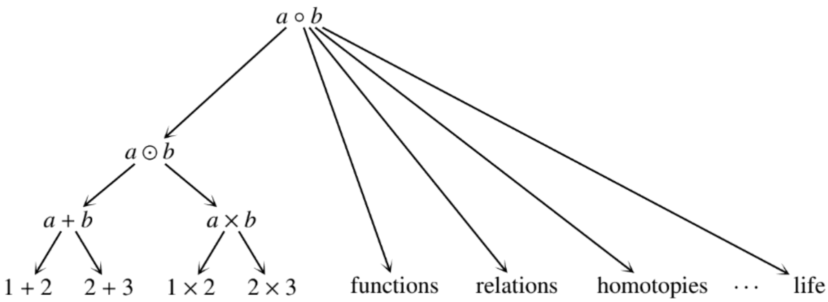
Here is a diagram showing that new level, with the symbol $\odot$ representing a binary operation that could be $+$, $\times$ or something else. There is a journey of abstraction up through levels of this diagram that is a bit like the journey through math education.



At the very bottom level we have the sorts of things you might do in pre-school or kindergarten where you play around with familiar objects and get nudged in the direction of thinking about numbers. At the next level up we have arithmetic as done in elementary school, perhaps, and then we move into algebra as done a little bit further on at school. The top level here, with the abstract binary operation, is the kind of thing we study in "abstract algebra" if we take some higher level math at university. Binary operations are studied in group theory, for example, and this is one of the topics we'll come back to. Incidentally I always find the term "abstract algebra" quite strange, because all algebra is abstract and, as I've described, what we even consider to count as "abstract" changes as we get more used to more abstraction.

There is indeed a further level of abstraction, which one might call "very abstract algebra". At this level we can think about more subtle ways of combining two things, where instead of taking just any two things and producing an answer, we can only take two things that fit together like in a jigsaw puzzle. For example, we can think about train journeys, where we can take one train journey and then another, to make a longer train journey — but this only makes sense if the second journey starts where the first one ends, so that you can actually change train there. This means you can't combine *any old* train journeys to make longer ones, but only those that meet up suitably where one ends and the other begins.

This is the sort of way we'll be combining things in category theory. Binary operations are still an example of this, but as with all our higher levels of abstraction, we will now be able to include many more examples of things that are more subtle than binary operations. This includes almost every branch of math, as they almost all (or maybe even all) involve some form of this way of combining things. Here is a diagram showing that, including the names of some of the mathematical topics we'll be exploring later in this book.



I've included "life" in the examples here, to emphasize that the higher level of abstraction may seem further away from normal life, but at the same time, the higher level is what enables us to unify a much wider range of examples, including examples from normal life that are not usually included in abstract mathematics. I think this is like swinging from a rope — if you hang the rope from a higher pivot then you can swing much further, provided you have a long rope. It's also like shining a light from above — if you raise it higher then the light will become less bright but you will illuminate a wider area. The result is perhaps more of an overview and an idea of context, rather than a close-up of the details. However, you won't lose the details forever as long as you retain the ability to move the light up and down. Furthermore, if you can find a way to make the light itself brighter, then you can see more detail and more context at the same time. I think this is an important aspect of becoming more intelligent, and that abstract mathematics can help us with that.

"More abstract" doesn't necessarily mean less relevant and it doesn't necessarily mean harder either. Too often we assume that things get harder and harder as we move up through those levels of abstraction, and thus that we shouldn't try to move up until we've mastered the previous level. However I think this is one of the things that can hold some people back or keep them excluded from mathematics. Actually the higher levels might be easier for some people, either because, like me, they enjoy abstraction and find it more satisfying, or because it encompasses more examples and those might be more motivating than the examples included at the lower levels. If you're stuck at the level of $a + b$ and the only examples involve adding numbers together then you may well feel that the whole thing is tedious and not much help to you either. After all, some things are fun, some things are useful, and some things are both, but the things that are neither are really the pits.

I think we should stop using the lower levels of abstraction as a prerequisite for the higher ones. If at the higher levels you get to deal with examples from life that you care about more than numbers, perhaps examples involving people and humanity, then it could be a whole lot more motivating. Plus, if you enjoy making connections, seeing through superficiality, and shining light, then those higher levels are not just useful but also fun.

3

Patterns

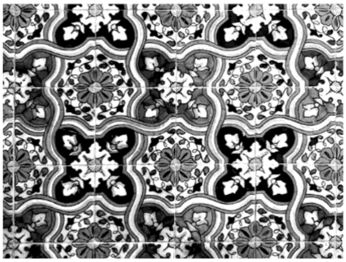
We can make something abstract, then find patterns, then ask if those patterns are caused by some abstract structure. This chapter still has little formality.

3.1 Mathematics as pattern spotting

Patterns are aesthetically pleasing, but they are also about efficiency. They are a way to use a small amount of information to generate a larger amount of information, or a small amount of brain power to understand a large amount of information.

Humans have used patterns across history and across most, if not all, cultures. Patterns are used for designs on fabric, on floors, on walls.

One carefully designed tile can be used to generate a large and intricate pattern covering a very large surface, such as in this tiled wall at the Presidential Palace of Panama. If you look closely you can see that the pattern is actually made from just one tile, rotated and placed at different orientations.†



Using one tile like this requires much less "information" than drawing an entire mural from scratch. But aside from this sort of efficiency, patterns can be very satisfying. They give our brains something to latch onto, so that they don't get too overwhelmed. It's like when a chorus or refrain comes back in between each verse of a song.

Mathematics often involves spotting patterns. This can help us understand what is going on in general, so that we can use less of our brain power to understand more things.

† It might be hard to see in black-and-white, but you can see a color version at this link: eugeniacheng.com/tile/.

3.2 Patterns as analogies

Patterns are really analogies between different situations, which is what category theory is going to be all about. At a visual level, we could talk about "stripes", for example, and understand that we mean alternating lines in different colors. We might not know the specific details – how wide the lines are, what the colors are, what direction they're pointing – but there is something analogous going on between all different situations involving stripes, and the abstract concept behind it is the concept of a "stripe". Patterns for clothes are also analogies, this time between different items of clothing. A dress pattern is like an abstract version of the dress.

In our previous examples of numbers we were looking at analogies between numbers. With the repeating patterns of final digits on multiples of numbers it was analogies between multiples. The numbers 10, 20, 30, 40, 50, and so on are all analogous via the fact that they consist of two digits: some number n followed by 0.

With the 12-hour clock the pattern on the table was an analogy between the different rows: in each row, the numbers go up one by one to the right (and the number 1 counts as "one up" from the number 12). The only difference between the rows is then what number they start at.

The analogy between clocks with different numbers of hours is then a sort of meta-pattern — an analogy between different patterns. The thing the different tables have in common is the "shifting diagonal" pattern. We have, in a sense, gone up another level of abstraction to spot a pattern among patterns.[†]

Things To Think About

T 3.3 What patterns can you think of in other parts of life? In what sense could they be thought of as analogies between different situations? You could think about patterns in music, social behavior, politics, history, virus spread, language (in terms of vocabulary and grammar)...

3.3 Patterns as signs of structure

Spotting patterns in math is often a starting point for developing a theory. We take the pattern as a sign of some sort of abstract structure, and we investigate what abstract structure is causing that pattern.

[†] At this point in writing I went into lockdown for the COVID-19 crisis and finished the draft without leaving the house again. I feel the need to mark it here.

Things To Think About

T 3.4 Here is an addition table for the numbers 1 to 4. Can you find a line of symmetry on this table, that is, a line where we could fold the grid in half and the two sides would match up. Why is that line of symmetry there? What about in a multiplication table for the same numbers?

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

Here's the addition table for the numbers 1 to 4 with a line of symmetry marked. I have also highlighted an example of a pair of numbers that correspond to each other according to this line of reflection. The one on the lower left is the entry for $3 + 1$ whereas the one on the upper right is the entry for $1 + 3$.

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

We can *see* that these entries are the same, but the *reason* they are the same is that $3+1 = 1+3$, which we might know as the commutativity of addition. If you check any other pair of numbers that correspond under the symmetry, you will find that they are all pairs of the form $a + b$ and $b + a$. The entries on the diagonal where the line is actually drawn are all examples of $x+x$ so switching the order doesn't change the entry. That is a form of symmetry in itself: in the expression $a + b = b + a$, if a and b are both x then the left and right become the same. We say the equation is symmetric in a and b.

An analogous phenomenon happens in the multiplication table, with the line of symmetry now being a visual sign of commutativity of multiplication. Often when we spot visual patterns we ask ourselves what abstract or algebraic structure is giving rise to that visual pattern.

Things To Think About

T 3.5 Here is a grid of the numbers 0 to 99. We have already talked about the patterns for multiples of 2, 5, and 10. If we mark in all the multiples of 3, what pattern arises and why? What about multiples of 9?

0	1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	28	29
30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49
50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69
70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89
90	91	92	93	94	95	96	97	98	99

Here is a picture of the multiples of 3 on the number grid. When we only listed their last digits it was less obvious how much of a pattern there was, because it seemed a bit random: 0, 3, 6, 9, 2, 5, 8, 1, 4, 7. However, when we draw them on this square it's visually quite striking that the multiples of 3 go in diagonal stripes, a bit like the diagonal stripe pattern we saw on the addition table above.

0	1	2	**3**	4	5	**6**	7	8	**9**
10	11	**12**	13	14	**15**	16	17	**18**	19
20	**21**	22	23	**24**	25	26	**27**	28	29
30	31	32	**33**	34	35	**36**	37	38	**39**
40	41	**42**	43	44	**45**	46	47	**48**	49
50	**51**	52	53	**54**	55	56	**57**	58	59
60	61	62	**63**	64	65	**66**	67	68	**69**
70	71	**72**	73	74	**75**	76	77	**78**	79
80	**81**	82	83	**84**	85	86	**87**	88	89
90	91	92	**93**	94	95	**96**	97	98	**99**

This is because the last multiple of 3 under 10 is one less than it, and so when we move down a row in this table, the pattern shifts one to the left. If the last multiple were 2 less than it, then the pattern would shift two to the left and be less striking. In this picture of multiples of 9 we see something similar happening, just with fewer stripes.

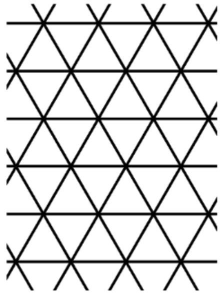
This pattern is essentially where we get that cute trick for the 9 times table where you hold down one finger at a time and read the number off the remaining fingers. So for 4×9 you can hold up all 10 fingers, then put down the fourth one from the left, and read off 3 from the left of that and 6 from the right to get 36. Tricks like that can be a way to bypass understanding or a way to deepen understanding.

0	1	2	3	4	5	6	7	8	**9**
10	11	12	13	14	15	16	17	**18**	19
20	21	22	23	24	25	26	**27**	28	29
30	31	32	33	34	35	**36**	37	38	39
40	41	42	43	44	**45**	46	47	48	49
50	51	52	53	**54**	55	56	57	58	59
60	61	62	**63**	64	65	66	67	68	69
70	71	**72**	73	74	75	76	77	78	79
80	**81**	82	83	84	85	86	87	88	89
90	91	92	93	94	95	96	97	98	**99**

Things To Think About

T 3.6 How does that trick generalize for other multiples? We will have to change what base we're working in.

A general principle of pattern spotting is that *visual* patterns might be easy to spot in simple examples, but the *abstract* structure might be easier to reason with, use or even verify, in more complex situations. It would be harder to draw the table for multiples of a much larger number, and if the patterns in the table were less obvious then it would be harder to see them. In situations of higher dimensions it is then even harder.



It is fairly easy to see all sorts of patterns of different sized triangles in this grid. However if this were a 3-dimensional space filled with triangular pyramids it would be rather harder to see, and in higher-dimensional space we can't even fit it into our physical world. But those are very helpful structures in many fields of research; it just requires more abstract ways of expressing them.

3.4 Abstract structure as a type of pattern

If category theory is the mathematics of mathematics, then categorical structure is about patterns in patterns. It's about seeing the same pattern in different places, and about making analogies between patterns.

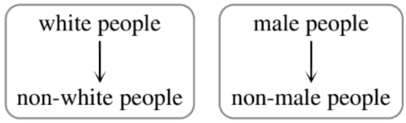
Here's an example. We might talk about a "mother–daughter" relationship abstractly, as opposed to thinking about a specific mother's relationship with their daughter. Now we could think about a relationship between someone's grandmother and mother. This is another type of "mother–daughter" relationship; it's just a particular type where another generation also does exist. The difference this makes to our considerations depends on the context.

Here's a tiny family tree for that structure. In a family tree we're not taking into account any context other than parent–child relationships. In the diagram there is no difference in the abstract structure depicted between the grandmother and her daughter, or the mother and her daughter.

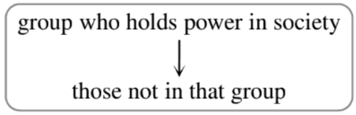
However, if we're writing a book about sociology, or the psychology of family units, or about motherhood, then we might well want to think about how the relationship between a mother and her daughter changes when the daughter has her own daughter in turn. However, the abstract similarity between the grandmother–mother and the mother–daughter relationships is still an aspect of what frames this question.

Another example is if we think about power dynamics between different groups of people in society. White people hold structural power over non-white people, and male people hold structural power over non-male people. This is about overall structures, not individuals — it doesn't mean that every white person holds power over every individual non-white person, or that every male person holds power over each individual non-male person. It's about the way the structures of society are set up. In any case even if you dispute this fact we can still depict the abstract structure that I am describing, because we can describe abstract structure as a separate issue from the question of how the abstract structure manifests itself in "real life".

We could depict these power structures like this, emphasizing the analogy that I am claiming exists between the two structures.



We could emphasize it further by going one level more abstract to this, which immediately unifies many situations.



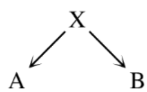
This now includes all sorts of other examples such as straight people over non-straight people, cisgender[†] people over trans people, rich people over non-rich people, educated people over non-educated people, employed over unemployed, people with homes over people experiencing homelessness, and so on.

3.5 Abstraction helps us see patterns

Finding the abstract structure in situations and expressing it in some way, often as a diagram to make it less abstract, can help us see the patterns and relationships between situations. It can help clear our mind of clutter and distraction and emotions. Distraction and emotions aren't bad per se, they can just get in the way of us seeing the actual structure of an argument rather than the window-dressing. Sleight of hand and flattering clothing can be fun, but if we're at the doctor's getting diagnosed for something it would be much more productive to show the whole truth and not be afraid of getting naked.

One of my favorite examples is the way I have been drawing diagrams of analogies and levels of abstraction. This way of specifying their structure has helped me pin down much more clearly where disagreements around analogies are coming from. I described it in *The Art of Logic* in terms of disagreements largely taking two possible forms.

It starts by someone invoking an analogy in this form, between two ideas A and B, but crucially without specifying what abstract level X they're referring to.



[†] Cisgender people are those whose gender identity matches the one they were assigned at birth.

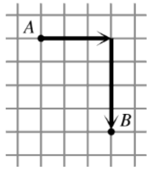
dry. I think a more accurate view is that it's a living tree, whose base is fixed but it can still sway and its branches and roots still grow. An impatient explorer might say "Trees are boring. They don't move." And yes, perhaps compared with lions and elephants they don't move. But if you take time to stare at them long enough they might become fascinating.

4.1 Distance

We're going to look at a world in which some familiar things behave differently from usual. Although actually it's a very common world in "real life", it just behaves differently from some common mathematical worlds that might be regarded as "fixed", which is not really *its* fault.

We're going to think about being in a city with streets laid out on a grid — so probably not a European city. Of course, even American cities aren't on a *perfect* grid — there are usually some diagonal streets somewhere, but we're going to imagine a perfectly regular grid with evenly spaced parallel roads.

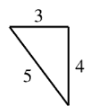
Now, when traveling from A to B we can't take the diagonal because there isn't a street there. Instead, the distance we'd actually have to go along the streets will be something like in this diagram. We have to go 3 blocks east and 4 blocks south making a total of 7 blocks.



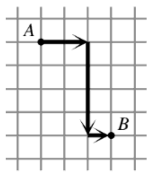
Calculating the distance "as the crow flies", that is, the direct distance through the air in a straight line regardless of obstacles, is rather academic as we can't usually make use of that route (unless we're pinpointing a location by sound or something, as in the film "Taken 2").

Things To Think About

T 4.1 We can use Pythagoras to calculate the distance as the crow flies which in this case will be 5 blocks. Is the road distance always more than the crow distance?

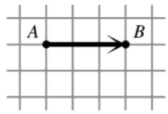


Is the road distance between two places always the same even if we turn at different points, for example as in this diagram?

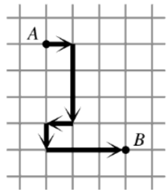


This type of distance is sometimes called "taxi-cab" distance (as if we all travel around in taxis all the time).

The taxi distance can be the same as the crow distance if *A* and *B* are on the same street, but the taxi distance can never be *smaller* than the crow distance.



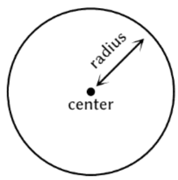
Also it doesn't matter where you turn as long as you don't go back on yourself, because wherever you turn you still have to cover the same number of blocks east and the same number of blocks south, making the same total. By contrast the path shown on the right covers more blocks because it does go back on itself.



It might seem like I'm trying to say something complicated, but it really is just like counting the blocks you actually walk when you walk around a city laid out on a grid. If your brain is flashing "Math class!" warning lights then you might be trying to read too much into this. No hypotenuse is involved.

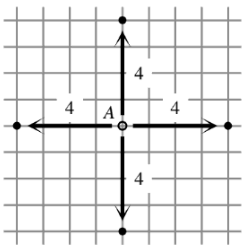
We can now think about what a circle looks like in this world. You might think a circle is just a familiar shape, but there is a *reason* that shape looks like that, and in math we are interested in reasons, or ways we can characterize things precisely.

How could you describe a circle to someone down the phone? One way is to say that you pick a center and a distance and draw all the points that are this distance away from that center. That chosen distance is called the *radius*.

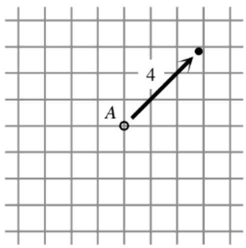


But in the taxi world we can't do distance along diagonals, so what will "all the points the same distance from a chosen center" look like? Let's try a circle of radius 4 blocks.

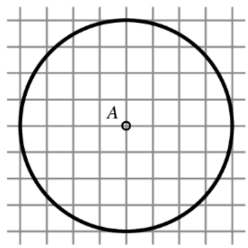
Here are the most obvious points that are a distance of 4 blocks from *A*.



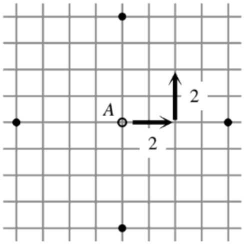
Here is something we can't do.



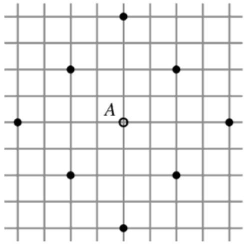
So we definitely can't get all these points — most aren't on the grid at all.



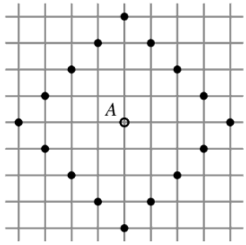
However we can go 4 blocks with a turn in the middle.



If we do this in all possible directions we get these points.

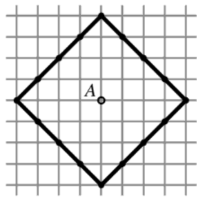


If we also do 3 blocks and 1 block to make 4 we get all these points.



The last picture is what a circle looks like in this taxi world. After filling in the points for going “2 blocks and 2 blocks” you might have seen the pattern to help you realize you could also do 3 and 1, which is good.

However, you might also have been tempted to join the dots like this picture. You are welcome to do so on paper but it won’t mean anything in the taxi world because those lines are not lines we can travel on. They include points we can’t get to in the taxi world — the taxi world circle really doesn’t include those lines.



So a circle in this taxi world is just a collection of disconnected dots in a diamond shape. How does that make you feel? Do you feel uncomfortable, as if this somehow violates natural geometry? Or do you feel tickled that a circle can look so funny? Both reactions are valid. The main thing is to appreciate that even some of the most basic things we think we know are only true in a particular context, and things can look very different in another context. It is important to be

a) clear what context we’re considering at any given moment (which we usually aren’t in basic math lessons), and
b) open to shifting context and finding different things that can happen.

Technicalities

What we have done here is find a different scheme for measuring distance between points in space. In fact there are many different possible ways of measuring distance and these are called *metrics*. Not every scheme for measuring distance will be a reasonable one, and in order to study this and any abstract concept rigorously we decide on criteria for what should count as reasonable. A metric space is then a set of points endowed with a metric. The idea of a metric space is to focus our attention on not just the points we're thinking about, but the *type of distance* we're thinking about.

The usual way of measuring distance "as the crow flies" is called the Euclidean metric. The taxi-cab way really is called the taxi-cab metric or the rectilinear metric. More formally it is called the L_1 metric and is the first in an infinite series of L_n metrics. The next one, L_2, is in fact the Euclidean metric.

4.2 Worlds of numbers

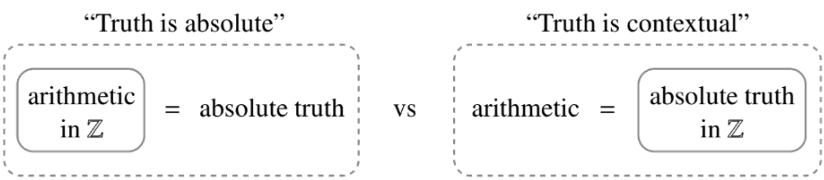
Many people say to me "Well, one plus one just *does* equal two." I reply, "In some worlds it's zero."

It's true that $1 + 1 = 2$ in ordinary numbers. But that's because it's how "ordinary numbers" are defined. Most people who think $1 + 1$ just *does* equal 2 are not considering that this is only true in some contexts and not others, as they're so used to one particular context. This is a bit like people who've never visited another country, and don't realize that people drive on the other side of the road in other places. Some people who haven't traveled don't understand that some ways of doing things are highly cultural and possibly arbitrary. For example:

- "it's math not maths" (not in the UK);
- "steering wheels are on the right of a car" (in the UK).

Some people can't imagine not having a car and others can't imagine having one.

We have seen that distance is contextual and thus "circles" are also dependent on context. We will now see a way that the behavior of numbers is also dependent on context. So all the arithmetic we are forced to learn in school is contextual, not fixed; it's not an absolute truth of the universe, unless we take its context as part of that truth. The context is the integers, that is, all the whole numbers: positive, negative and zero. The set of integers is often written as $\mathbb{Z}$. Here is a diagram showing those different points of view.



Arithmetic might seem like absolute truth if you think the integers are the only possible context. But in fact most people *do* know other contexts, they just don't come to mind when thinking about arithmetic. For example if you dump a pile of sand onto a pile of sand you still just get one pile of sand; it's just bigger.

Things To Think About

T 4.2 What contexts can you think of in which 1 + 1 is something other than 2? Can you think of other contexts in which it's 1? What about 0, or 3 or more? What about other ways in which arithmetic sometimes works differently?

Here are some places where arithmetic works differently.

Telling the time. 2 hours later than 11 o'clock isn't 13 o'clock unless you're using a 24-hour clock. On a 12-hour clock it's 1 o'clock, that is $11 + 2 = 1$. On a 24-hour clock 2 hours later than 23 o'clock is not 25 o'clock, it's 1 o'clock, that is $23+2 = 1$. (While we don't usually say "23 o'clock" out loud in English, it does happen in French.)

I'm not not hungry. Particular kinds of children find it amusing to say things like "I'm not not hungry" to mean "I am hungry". If we count the instances of "not" we get $1 + 1 = 0$.

Rotations. If you rotate on the spot by one quarter-turn four times in the same direction you get back to where you started, as if you had done zero quarter-turns. So if we count the quarter-turns, $1 + 1 + 1 + 1 = 0$. We could generalize this to any n by rotating n times by $\frac{1}{n}$ of a turn each time.

Mixing paint. If you add one color paint to another you do not get two colors, you get one color.† Likewise a pile of sand or drop of water. So $1 + 1 = 1$.

Pairs. If one pair of tennis players meets up with another pair for an afternoon of tennis, there are 6 potential pairs of tennis players among them, if everyone is happy to partner with any of the others.‡

The first three of these situations all have something in common and we are now going to express exactly what that analogy might be.

† This example was brought to my attention by my art students at SAIC.

‡ This example was also brought to my attention by my art students at SAIC, though in a less child-friendly formulation.

5

Relationships

The idea of studying things via their relationships with other things. Revisiting some of the concepts we've already met, and reframing them as types of relationship, to start getting used to the idea of relationships as something quite general.

In the last chapter we saw that objects take on very different qualities in different contexts. Now we're going to see that different contexts can be provided by looking at different types of relationship. For example, if we relate people by age we get a different context from if we relate them by wealth or power.

One example we investigated was taxi-cab distance. Distance can be viewed as a relationship between locations, and the taxi-cab relationship gives us a different context from the "crow distance" relationship. There are also other possibilities — we could take one-way streets into account, or we could use walking distance, which might be different, since cars and pedestrians often have access to different routes.

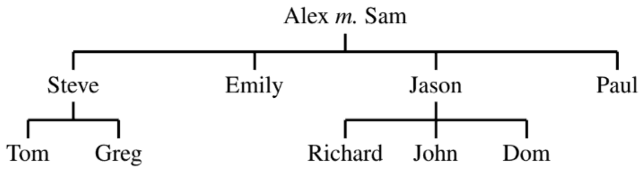
In the case of the n-hour clocks we saw a different type of relationship between numbers, in which, for example, $1+1$ can equal 0. This equation is really a relationship between the numbers 1 and 0. We do not have this relationship in the ordinary numbers, where we only have $1 - 1 = 0$ (and $-1 + 1 = 0$). So the existence of the relationship $1 + 1 = 0$ tells us we are in the context of the 2-hour clock, technically called the integers modulo 2.

In the case of the zero world, everything is related by being the same. Of course, everything isn't *actually* the same, but is *considered* to be the same in that world. This is an important distinction that we will keep coming back to. When we focus on context, we are looking at how things appear in *that* context. Things can appear the same in one context but not another, just like when I take my glasses off and everyone looks more or less the same to me.

In this chapter we will develop a way of dealing with relationships that heads towards the way in which category theory deals with them. We will also start drawing diagrams in the way they are drawn in category theory.

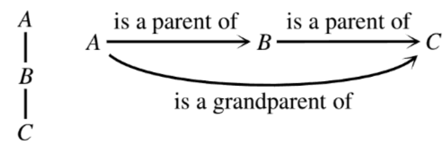
5.1 Family relationships

We sometimes depict family relationships in a family tree like this:

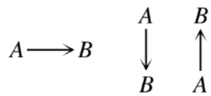


There are only three types of relationship directly depicted here: marriage, parent–child, and siblings. In fact we could view "parent–child" and "sibling" as part of the same depiction, in which case we are depicting only two types of relationship. In any case, we don't need to depict grandparents directly, because we can deduce those relationships from two consecutive parent–child relationships.

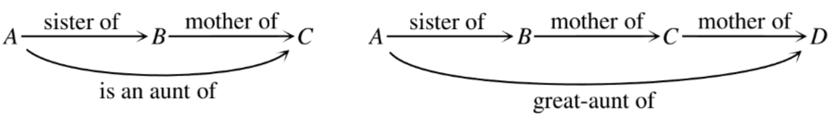
Here are two ways to depict this. The first one is a little more rigid as it depends on positions on the page.



If we represent a relationship using arrows rather than physical positioning on the page, we can draw things any of these ways up (and others) without affecting the relationship we're expressing:



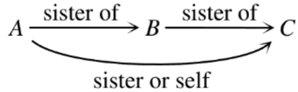
That is the point of the arrowhead. Different choices might help us visually, so the flexibility is beneficial (as flexibility typically is). The arrows also encourage us to "travel" along them to deduce other relationships such as:

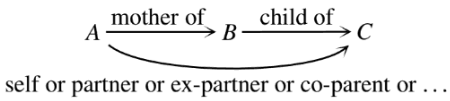


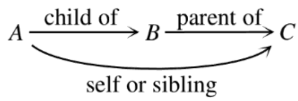
Things To Think About

T 5.1 Can you think of any situations where we can travel along two arrows and the resulting relationship could be several different things?

Note that we all have a relationship with ourself: $A \xrightarrow{\text{self}} A$ for any person A. For this among other reasons things might be ambiguous as in these examples.







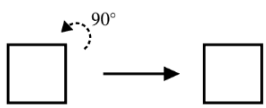
We are going to see that this way of depicting and compiling relationships is remarkably fruitful. It is general and flexible enough to be usable in a vast range of situations, and illuminating enough to have become a widespread technique in modern math and central to category theory. However we will see that we do need to impose some conditions to make sure we don't have ambiguities.

5.2 Symmetry

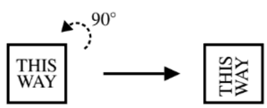
Some things might not initially seem like a type of relationship, but can be viewed like that by shifting our point of view slightly. We are going to see that categories are built from a very general type of relationship, so if we can view something as a relationship we have a chance of being able to study it with category theory. This is often how we find new examples of existing mathematical structures — it's not exactly that the example is new, but we look at it in a new way so that we see the sense in which it is an example. It's a bit like the fact that if we consider traffic like a fluid then we can understand its flow better using the math of fluid dynamics, leading to effective (and perhaps counter-intuitive) methods for easing congestion.

Symmetry is something that we might think of as a property, but we can alternatively think of it as a relationship between an object and itself. For example a square has four types of rotational symmetry: rotation by 0°, 90°, 180° or 270°. (It doesn't matter which direction we pick as long as we're consistent.)

The symmetry can be seen as this property of a square: if you rotate it by any of these angles it goes back to looking like itself.

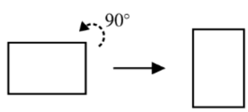


You can't tell the difference unless I put something on it.



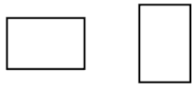
Now, the *fact* that we can do this is a property. It's a property that a rectangle, for example, doesn't possess.

A rectangle in general looks different after we rotate it 90° even without anything written on it.

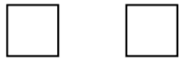


In abstract math we are moving away from facts and moving towards processes. The *process* of turning a shape around is a relationship.

In the case of the rectangle it's a relationship between these two pictures.

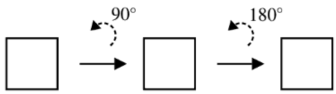


In the case of the square it's a relationship between these two pictures: the square and itself.

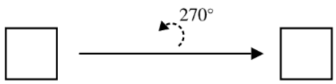


Note that now we're thinking of symmetry as a process or relationship we can ask what happens if we do one process and then another.

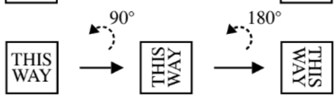
For example if we do these two processes



the end result is the same as doing this all in one go.



We can see this by checking that with the words written on it looks like this:



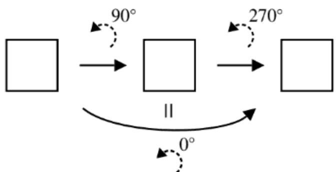
Things To Think About

T 5.2 What happens if we do 90° and then 270°? Remember the end result should be one of our rotations: 0°, 90°, 180°, 270°.

If we rotate by 90° and then 270° that's the same as rotating by 360°, but this isn't in our list of rotations because it's "the same" as doing 0°.

Here "the same" means the *result* of rotating by 360° is the same as the *result* of rotating by 0° as shown here.

We could put all that information in this single diagram, which has the added benefit of looking just like the diagrams we drew for family relations previously. (The symbol looking like two short vertical lines is a rotated equals sign.)



We could depict all these relationships in a table like our previous addition and multiplication tables, to help us keep track of what is going on. So far we have these relationships.

	0	90	180	270
0				
90				
180		270		
270		0		