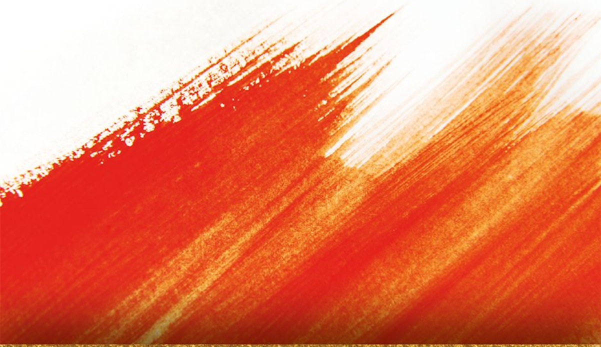
OXFORD



Modal Homotopy Type Theory

The Prospect of a New Logic for Philosophy

DAVID CORFIELD

MODAL HOMOTOPY TYPE THEORY

Modal Homotopy Type Theory

The Prospect of a New Logic for Philosophy

DAVID CORFIELD

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

First Edition published in 2020
Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2019953288

ISBN 978–0–19–885340–4

DOI: 10.1093/oso/9780198853404.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

PREFACE

The old logic put thought in fetters, while the new logic gives it wings. It has, in my opinion, introduced the same kind of advance into philosophy as Galileo introduced into physics, making it possible at last to see what kinds of problems may be capable of solution, and what kinds are beyond human powers. And where a solution appears possible, the new logic provides a method which enables us to obtain results that do not merely embody personal idiosyncrasies, but must command the assent of all who are competent to form an opinion.

(Bertrand Russell, *Logic As The Essence Of Philosophy*, 1914)

Bertrand Russell, a little over a century ago, promised a great future in philosophy for the new logic devised by Frege and Peano. The first-order predicate logic that emerged from their work, 'by an analysis of mathematics' (Russell 1914, p. 50), has certainly spread far and wide in the anglophone philosophical world. But now in the early years of the twenty-first century, a new 'new logic' has appeared, a new foundation for mathematics. It arose through a philosopher-logician, Per Martin-Löf, rethinking the composite nature of mathematical judgement and formulating his ideas in the shape of a theory of *types*. Currents from this type theory blended with others deriving from another radical reconceptualization of the foundations of mathematics, this time in the shape of *category theory*, with its emphasis on structure and transformation. The tight relationship between type theory and category had long been studied, not least in computer science departments, but the meshing of these currents was propelled to take on the form of our new logic by the insights of one of the finest mathematicians of the past few decades, Vladimir Voevodsky, who was looking to develop proof assistants capable of dealing with modern mathematics (Voevodsky 2014). This new logic allows advanced *homotopical* concepts to be directly constructed and manipulated; in particular, the mathematical concepts used in current physics. The name of this wonder-language is *homotopy type theory*. In this book I shall be arguing that *philosophy* should look to *homotopy type theory* and variants, in particular what is known as *modal* homotopy type theory, as its formal language of choice.

'Plain' homotopy type theory, familiarly known as 'HoTT', has very recently appeared (UFP 2014) as a contender to challenge set theory's long-standing claim to act as the primary foundational language for mathematics. Yet more recently, it has been proposed that to this new language there be added extra resources, in the form of what are called *modalities*, in order to express concepts such as continuity and smoothness. Taken together, I will refer to languages in this family as 'modal homotopy type theory', or 'modal HoTT' for short. Already a large body of results from quantum gauge field theory has been written up in modal HoTT (Schreiber 2013, 2014a, forthcoming). Even so, I do not intend this book primarily as a contribution to the philosophical study of mathematics, and certainly not of physics. Indeed, an exploration of what modal HoTT can bring to a philosophical treatment of physics will have to wait for another book. Regarding mathematics, since set theory is formulated as an axiomatic system built on top of a base of first-order logic, and since the major part of the debate surrounding axiom choice is only directly relevant to mathematics, then to treat set theory philosophically is inevitably to look to speak to the philosophy of mathematics. The component which is first-order logic may be separated out and its particular relevance for philosophy then studied. HoTT by contrast comes as a whole package, with its logic integrated. From the perspective of HoTT, to apply its 'logic' in philosophy is already to apply the whole calculus. Similarly, we cannot separate out naturally a modal logic from modal HoTT, so we will need to think through applications of the whole calculus to philosophy. If Russell could compare the introduction of his new logic into philosophy with that of mathematics into physics, now we can consider the prospect of the new 'new logic' being introduced simultaneously into all three disciplines.

In view of the prominence accorded to logic in current anglophone philosophy, anywhere this philosophy has seen fit to employ what it takes as its standard logical tools—in the philosophy of language, in metaphysics or wherever—there is room to explore whether modal HoTT fares better. Over the course of the book, I aim to convince the reader that it does. Many philosophers throughout the twentieth century expressed their serious misgivings about the use of formal tools in philosophy, illustrated frequently by apparent failings of first-order logic to capture the nuances of natural language. I hope that I can give sufficient reason here to cast doubt on their pessimistic conclusions.

My promotion of modal HoTT is clearly, then, no minor proposal. We normally train our students in philosophy in propositional and then first-order logic. Those wishing to study metaphysics will probably also learn some modal logic. If I am right, all of this should change. New lectures and textbooks will have to be written, and a large-scale retraining exercise begin. But not only are the tools of the trade to change, the whole disciplinary landscape of philosophy will have to alter too. We have here a language which at last serves cutting-edge mathematics well, and consequently so too the theoretical physics which relies upon it. At the same time, I maintain, it makes a much better fist of capturing the structure of natural language. A radical reconfiguring of the relations of these domains to that of metaphysics is therefore to be expected.

Chapters 2 to 4 will see us build up the component parts of modal HoTT. In Chapter 2 we motivate and deploy type theory, and in particular the dependent version. Chapter 3 then explains why people have looked to represent more subtle notions of identity. Here we see in particular how the 'Ho' in HoTT is derived. Chapter 4 then introduces modalities to the type theory. This is followed up by Chap. 5, which illustrates what a particular variant of modal HoTT—the *differential cohesive* variety—can bring to the philosophy of geometry.

I begin, however, in Chap. 1 with an introductory survey of the kinds of thinking that have motivated the developments we will be covering through this book. I do so via an account of how I arrived at such a radical point of view as to wish to replace philosophy's logic. There will be allusion made to constructions that will be dealt with in closer detail in later chapters, but I hope the broadbrush picture given will help to orient readers. On the other hand, it would be quite possible for the reader to set out from Chap. 2.

The authors of the HoTT book can warn us that, because of its youth,

> This book should be regarded as a 'snapshot' of just one portion of the field, taken at the time it was written, rather than a polished exposition of a completed edifice. (UFP 2014, p. 1)

All the more, then, should such a warning apply to a book such as this. Alongside some of my own writings (Corfield 2017a, 2017b) there has been some early philosophical exploration of HoTT (see, for example, Awodey 2014, Ladyman and Presnell 2015, 2016, 2017, Tsementzis 2017, Walsh 2017), but, as I write, nothing concerning modal HoTT beyond what is included in an article of mine (Corfield 2017c). Even after a book-length treatment, I will have only scratched the surface of the philosophical relevance of modal HoTT. This is not just because I am only beginning here the task of thinking through what modal HoTT might be able to do for philosophy. It is also that the heuristic power of the HoTT programme has plenty of steam left for it to move in unforeseen ways. In view of the rapidity of developments in the field, we see further useful variants of type theory in the pipeline. Ideas here include: a *cubical* version, bringing the prospect of better computational properties; a *directed* version, which should encode one-way irreversible processes; and a *linear* version, which should be well-suited for quantum physics and stable homotopy theory. There are also *two-level* versions, which provide the means to define certain infinite structures, and a notion of 'type 2-theory' to describe higher structures. As I write these words, modal dependent type theory itself is being reformulated as 'dependent type 2-theory'. But no matter. Even if in a few years' time the consensus has shifted on the best way to frame the variety of type theories, taking efforts now to learn about some currently standard modal HoTT constructions and their applications will not be time wasted, since some form of this logical calculus is here to stay.

It is important to note that naming conventions have not been definitively settled yet. Some people have looked to distinguish *Univalent Foundations* from HoTT, but for the duration of this book, I shall be working on the understanding that HoTT is a variety of intensional Martin-Löf dependent type theory with higher inductive types and satisfying the univalence axiom. I shall be referring to all of these ingredients in the course of the book as we need them.[1]

Portions of the contents of two chapters of this book have appeared elsewhere. Chapter 3 is a reworking of a rather clumsily named paper of mine, 'Expressing "The Structure Of" in Homotopy Type Theory' (Corfield 2017a). In the setting of a book I now have the space to include considerably more detail on important ingredients of HoTT. Similarly, Chap. 5 provides an opportunity to recast my article 'Reviving the Philosophy of Geometry' (Corfield 2017c) in the context of a more elaborate discussion of the current situation in geometry, and with the background prepared by the preceding chapters, I have the opportunity to enter into much more detail about the geometric modalities involved.

[1] This is in line with the 'HoTT Book' (UFP 2014), which provides the most detailed account of the system in a reasonably accessible way. See also Shulman (2017) for an excellent introduction.

The pleasure I have taken in my research path over many years has been greatly enhanced by interacting with some of the people responsible for the construction of modal HoTT. In particular, over the past ten years I have been a steering committee member of the wonderful *n*Lab wiki and a participant in its discussion forum, the *n*Forum. The generosity of the contributors there, including two of the pivotal figures in the development of modal HoTT, Mike Shulman and Urs Schreiber, never ceases to astound me. It is a genuine privilege to witness the foundations of mathematics and physics being developed before your own eyes. Many of the ideas contained in this book have come from them, or been developed in discussions with them. For much more exposition about many of the concepts and broader outlooks treated in this book, the *n*Lab is an invaluable resource to go alongside more conventional publications, such as the HoTT book (UFP 2014). I will provide additional pointers to these and other resources in the Further Reading section towards the end of the book. The *n*Lab was spun off from a blog which Urs, Mike, I and others run, the *n*-Category Café. Its founding mission when Urs and I set it up with John Baez was to explore what higher category theory might mean for mathematics, physics and philosophy. We always operated under the firmly held tenet that careful explanation could bring clarity even to the most advanced concepts. I have endeavoured to adhere to this ideal. New contributors/editors to the *n*Lab are always welcome. Visit the *n*Forum for advice about how to get started.

In addition to my *n*Lab coworkers, I should like to thank various members of audiences which have heard me: in particular, Jon Williamson, Colin McLarty, Zhaohui Luo and my doctoral student, Gavin Thomson.

I dedicate this book with much love to my mother and to the memory of my father.

CONTENTS

1

A Path to a New Logic

1.1 First Encounters

It is striking, although I imagine not a terribly uncommon experience, when you come to realize through the course of a life of intellectual enquiry that you haven't moved so very far from where you first began. Nearly thirty years ago I came to philosophy from mathematics, full of zeal to discover what should be said about the richness of the latter's concepts. I could see that beautiful thematic ideas were to be found manifesting themselves across the historic, yet frequently breached, boundaries regimenting its branches, and I pored over the writings of the philosophical master of these *supra*-mathematical ideas, Albert Lautman. I was convinced that a formal language I had just begun to study, *category theory*, was much better suited than alternatives to speak to these ubiquitous concepts, and that Lautman could be seen as anticipating its invention.[1] This seemed to be the case even with mathematical logic, as Lambek and Scott (1986) had recently shown. *Adjoint functors*, those mainstays of category theory, appeared to be just as prevalent in logic as in the rest of mathematics, something William Lawvere had demonstrated a couple of decades earlier (Lawvere 1969). In fact, I agreed wholeheartedly with Lautman that logic should be seen as of a piece with mathematics, and thought to look to category theory's treatment of logic to explain this to the philosophical community I was hoping to join.

Consequently, my Masters thesis addressed the relationship between two styles of semantics for intuitionistic logic, one couched in the traditional philosophical language of judgement and warrant, the other in spatial terms. Intuitionism, in this sense, has a long history stretching back to the sometimes mystical writings of the Dutch mathematician, L E J Brouwer, in the first decades of the twentieth century. Supposedly against his wishes, a logic had been extracted from these thoughts by Andrey Kolmogorov and Arend Heyting. Later, we could read philosophers, such as Michael Dummett, who argued that intuitionistic logic makes good sense of the distinctions in meaning Brouwer had unearthed, which were typically collapsed in classical hands.

[1] See the French mathematician Jean Dieudonné's comments to this effect in the preface to Lautman (2006), and also Corfield (2010).

Modal Homotopy Type Theory: The Prospect of a New Logic for Philosophy. David Corfield, Oxford University Press (2020). © David Corfield. DOI: 10.1093/oso/9780198853404.001.0001

In the meantime, however, a second style of semantics for this logic had emerged mid century from the work of McKinsey and Tarski, in a form which resembled somewhat the use of Venn diagrams to represent sets and their intersections. To illustrate something of the difference between these two kinds of semantics, consider the connective *implies* in propositional logic. On the one hand, there is the *proof-theoretic* interpretation whereby a proof of $A \to B$ is a means to transform any warrant for the truth of A into a warrant for the truth of B. On the other hand, propositions in the *topological* semantics are interpreted spatially as interiors of domains, technically as *open subsets* of a topological space, denoted $A \mapsto (A)$. Then $A \to B$ is interpreted as the largest open set whose intersection with (A) is contained in (B). Dummett (1977) tried to assure us that the proof-theoretic interpretation was the proper one, and that the topological one was something of a fluke. While most readers will surely agree that the proof-theoretic account presents a more readily graspable meaning, at the time it struck me as implausible that the discovery of topological models for intuitionistic logic wasn't pointing us in some important direction, an indication again that logic ought to be considered as subsumed within mathematics. The tie between intuitionism and topology goes right back to the days of Brouwer, he himself being one of the founding fathers of topology from 1909, just a year after his first published criticism of the classical logical principle of excluded middle. It seemed to me at the time, then, that category theory could provide a resolution to the opposition between these twin semantics with its invention through the 1960s of *toposes*, categories which combine the logical and the spatial, and that this needed philosophical attention.[2]

During research for this work, I remember reading Göran Sundholm's resolution of the puzzle of the 'Donkey sentence' (Sundholm 1986) using a constructive type theory, a form of intuitionistic language, which resonated with Dummett's ideas. The original sentence is:

- If a farmer owns a donkey, then he beats it,

but Sundholm's point can be treated in terms of the simpler sentence

- If John owns a donkey, then he beats it.[3]

The problem here is that we expect there to be a compositional account of the meaning of this sentence as given by the sentence structure. At first glance, it appears that there is an existential quantifier involved in the antecedent of a conditional proposition, as signalled by the indefinite article. However, a beginner's attempt to use one is ill-formed, the final x being unbound:

- $\exists x(Donkey(x) \& Owns(John, x)) \to Beats(John, x)$.

[2] Very recently in Shulman (forthcoming) we see an answer provided by modal homotopy type theory, as I explain in Chap. 5.

[3] I have heard the complaint that these are not the kinds of sentences ever likely to be uttered. Of course, it's easy to find examples with just the same structure which are more plausibly spoken: 'If a customer has purchased a faulty radio from my store, then she will receive a full refund for it.'

On the other hand, if the scope of $\exists$ is extended to the end of the sentence by a shift of parenthesis, then the new sentence means that there is some entity for which, if it is the case that it is a donkey owned by John, it is beaten by him. But then this is made true by a sheep owned by John or a donkey owned by Jane, whether beaten by John or not. Clearly, this is not the intended meaning.

The alternative in standard first-order logic is to rephrase the sentence as something like:

- Any donkey that John owns is beaten by him,

and then to render it formally as

- $\forall x(Donkey(x) \& Owns(John, x) \rightarrow Beats(John, x))$.

But can it really be the case that we comprehend such sentences by first performing this kind of radical transformation? Compositional accounts surely have the advantage of plausibility as concerns language acquisition and comprehension.

Sundholm showed how we could have our cake and eat it, a compositional and faithful account of the sentence using the resources of dependent type theory. Propositions in type theory are *types* whose elements are warrants for their truth. The idea, then, is to construct a type such that an element plays the role of justifying the Donkey sentence. Such an element would be of the kind that, when presented with a donkey together with a warrant that John owns it, would deliver a proof that he beats it. The resources of Martin-Löf's type theory allow precisely this. Its dependent sum (or sometimes *pair*) construction, $\sum$, provides a way to construct *paired* types, used here to form the type *Donkey owned by John*, $\sum x : Donkey(Owns(John, x))$. An element of this type has two components: a donkey and a proof of ownership by John. We can *project* out from any such pair to its components, using p for the first and q for the second.

Then the dependent product construction, $\prod$, provides a way to construct *function* types. An element of this type will send an element of the input type to one in the output type. The complete type for the proposition is then

- $\prod z : (\sum x : Donkey(Owns(John, x)) Beats(John, p(z)))$.

An element of this type takes any donkey owned by John to a proof that John beats that donkey. The 'it' of the Donkey sentence is represented by $p(z)$, the donkey component of the pair coupling the animal to a proof of its ownership by John. The product ($\prod$) at the beginning also confirms the thought that there's a form of hypothetical occurring, a 'whenever', and the sum ($\sum$) is a form of existential quantification.

All those years ago, I was not duly impressed with this contribution to the philosophy of language. What I cared about then was the promotion of category theory. Sundholm was relying on a formalism known after its inventor as *Martin-Löf dependent type theory*, and this had a semantics in what were called *locally cartesian closed categories*, the toposes mentioned above being a special case of these. What I might then have pursued was the ever-tighter embrace into which computer scientists, constructive type theorists and category theorists were entering. They were converging under the guidance of a vision dubbed by the computer scientist, Robert Harper, *computational trinitarianism*. Tongue in cheek, he glosses this position as follows:

> The central dogma of computational trinitarianism holds that Logic, Languages, and Categories are but three manifestations of one divine notion of computation. There is no preferred route to enlightenment: each aspect provides insights that comprise the experience of computation in our lives. (Harper 2011)

The serious point is that these three corners allow a literal 'triangulation' of the value of a new development put forward in any one of them. Any proposed construction in one field had better make good sense in the other two.

Constructive logic	Programming languages

Category theory

Taking category theory as the representative of mathematics, he continues:

> Imagine a world in which logic, programming, and mathematics are unified, in which every proof corresponds to a program, every program to a mapping, every mapping to a proof! Imagine a world in which *the code is the math*, in which there is no separation between the reasoning and the execution, no difference between the language of mathematics and the language of computing. Trinitarianism is the *central organizing principle* of a theory of computation that integrates, unifies, and enriches the language of logic, programming, and mathematics. It provides a *framework for discovery*, as well as analysis, of computational phenomena. An innovation in one aspect must have implications for the other; a good idea is a good idea, in whatever form it may arise. If an idea does not make good sense logically, categorially, and typically . . . then it cannot be a manifestation of the divine. (Harper 2011)

There is a crucial lesson here—one that philosophers in particular have not always been the first to grasp. It is remarkably easy to convince yourself that you are doing good work when devising a formal system to capture some topic or other. Proper triangulation, where achievable, is enormously reassuring, then. If several (at least partially) independent sources agree that some construction or other looks right according to its own view of the matter, then there is a much greater chance that it will stand the test of time. It was with this lesson in mind that Sundholm had explained:

> In this manner, then, the type-theoretic abstractions suffice to solve the problem of the pronominal back-reference in [the Donkey sentence]. It should be noted that there is nothing *ad hoc* about the treatment, since all the notions used have been introduced for mathematical reasons in complete independence of the problem posed by [the Donkey sentence]. (Sundholm 1986, p. 503)

Indeed, the dependent sum and product constructions used there had already suggested themselves to Per Martin-Löf. As we shall see later, these constructions are also extremely natural from a category-theoretic perspective, namely, as *adjoints*.

Now, each of the corners of Harper's 'Trinity' comes with a very intricate history. It is important, then, to consider overlaps and differences of emphasis between them to understand the degree of their independence. Over many years, Martin-Löf has given us evolving versions of his constructive type theory. He achieved this by reaching back to Kant, Brentano, Frege and Husserl, that is, to traditions which take great care to articulate the judgemental aspects of logic, and then also to the school of Intuitionism in mathematics, founded by Brouwer and Heyting. It is not uncommon to hear people describe this latter movement, involving the denial of excluded middle and double negation elimination, as a failed revolution (see, for example, van Fraassen 2002, p. 239n8) which had been conclusively defeated in the world of mathematics at least as far back as the 1920s. Notably, one of the acknowledged world leaders of mathematics could say in that decade:

> Taking the principle of excluded middle from the mathematician would be the same, say, as proscribing the telescope to the astronomer or to the boxer the use of his fists. To prohibit existence statements and the principle of excluded middle is tantamount to relinquishing the science of mathematics altogether. (Hilbert 1927, p. 476)

However, constructivism refused to go away, and indeed computer science has very much taken it to heart. Any history of this convergence would dwell on the contribution of Andrei Markov, son of the identically named mathematician of Markov chain fame, whose brand of constructivism rested on the computational principles of recursion theory.

We can easily see why there should be a connection by considering an informal proof of the simple result that in a room with some people inside it, there must be a youngest person (being perhaps joint youngest):

- Assume not, so that we have a nonempty room where for everyone present there is a younger person.
- Then (without loss of generality) select someone in the room.
- By assumption, that person has someone younger than themselves. Choose one such person. Then again, that person has someone younger than them, and so on.
- In this way, we can create a sequence of people in the room whose ages grow ever younger of any length we like. Since there are only finitely many people present, any sufficiently long such sequence must have at least one person appear more than once. But then, by the transitivity of the 'younger than' relation, that someone is younger than themselves. Contradiction.
- So it can't possibly be that no one is the youngest. Therefore, someone must be the youngest.

This is a valid classical proof, but one which is needlessly non-constructive. Compare this with a method which actually allows you to lay your hands on the youngest:

- Place the room occupants in any order.
- Take the first person and compare their age with the next person.
- If older, switch to the new person. Otherwise, continue with the original person. Proceed to the third person and do likewise.
- Repeat until you reach the end of the line.
- The resulting person is youngest.

Of course, there are other ways to perform this task. I could arrange for the occupants to meet in tennis-tournament fashion, the younger of any pairing proceeding to the next round. Taking ages to be measured in a number of years, each of these are algorithms which effectively find a lowest natural number in a set (or perhaps multiset) of natural numbers.

From the constructive perspective, the needlessly non-constructive proof is still a proof of something, it's just not a proof of 'for a nonempty, finite (multi)set of natural numbers, there is a smallest member'. Instead, it is a proof of absurdity on assuming that in such a (multi)set no element is smallest. 'Proof of absurdity of ... ' is the form that constructive negation takes. The negation of a proposition, $\neg P$, is defined as $P \to \perp$, a proof of which when presented with a proof of P yields a proof of absurdity. To argue by *reductio* as here, we would need the rule that $\neg\neg P$ implies P. Famously, in constructive logic we do not have this. We will see through the book that the constructive element within homotopy type theory (HoTT) will require us to be much more sensitive than is commonly the case in our use of negation.

Here in this case of a finite collection, our proof is evidently needlessly non-constructive, but there are cases involving infinitely large domains where I can avoid a *reductio* argument by invoking constructively valid principles. For instance, a proof that all natural numbers may be expressed as products of prime numbers can deploy the so-called *principle of strong induction* rather than the typical classical strategy of showing the absurdity of assuming there to be a least natural number that is not so expressible. On the other hand, there are cases where no valid constructive principle can step into the breach, such as when looking to establish that any real number is either less than zero, equal to zero or more than zero.

Clarifying the constructive-computer science connection, the so-called *Curry–Howard correspondence* associates constructive proofs of propositions to computer programs carrying out corresponding tasks. A proof is like an algorithm. A proof of, say, *A implies B* corresponds to a program which transforms an input of type A to an output of type B, in line with the interpretation by intuitionistic logic.

So if constructivism in logic and programming theory is integrally related, what, then, of the third corner, category theory? Well, one of its originators, Saunders Mac Lane, while a student in Göttingen in 1933, had an important brush with the kind of philosophy that informed Martin-Löf, including phenomenology through Oskar Becker (see McLarty 2007). Becker was someone whose philosophy inspired him to work out a formalism for intuitionistic logic, so there may be direct philosophical connections worth exploring. On the other hand, it is hard to overestimate the influence of William Lawvere on the category-theoretic outlook. The title of one of his papers, 'Taking Categories Seriously' (Lawvere 1986), beautifully encapsulates his outlook. To think category theory could provide a foundational language for mathematics (and physics) and to push through with this goal was extremely courageous, and is very far from being sufficiently recognized, even today.

Philosophically, Lawvere's points of nineteenth-century reference differ from Martin-Löf's, to include Hegel, Grassmann and Cantor. Very important historical work needs to be done to think through intellectual connections between his sources. Even so, it is quite possible to indicate quickly something of the commonality of category theory with constructivism in terms of certain structural observations. Let's consider again double negation, in particular the contrast between its introduction and its elimination rules. Constructive logicians are happy with the inference from the truth of P to the truth of $\neg\neg P$, but not the other way around, $\neg\neg P$ to P. Rewriting these rules in constructive guise points us to the essential difference: P *true* $\vdash (P \to \bot) \to \bot$ *true* is an example of a very common mathematical construction. When a function, $f : A \to B$, is applied to an element, $a : A$, it results in an element $f(a) : B$. Thus, dually, any $a : A$ provides a means to turn any such function, f, into an element of B. There is a natural *application* pairing, $app : A \times B^A \to B$, which may be transformed (*curried*, to use the jargon) to an associated mapping, $A \to B^{B^A}$. Here an element a of A is sent to the map 'evaluate at a'.

Considering this phenomenon in the framework of propositions, and choosing B as absurdity, we arrive at double negation introduction. This tallies with how the constructive logician explains things: if I have a proof of P, then consider the situation where I also have a proof of $\neg P$, or, in other words, a way of converting a proof of P into a proof of absurdity. In that case, I would have a proof of absurdity. So a proof of P is a means to transform a proof of $\neg P$ to absurdity, or, in other words, a proof of $\neg\neg P$.

The questions for the category theorist to ask, then, are: In what kinds of setting are these constructions possible? Where do we have for any two objects an object of functions between them? Where do we even have the ability to form a pair of items of different kinds, here a function and an argument? Answering these kinds of questions has been at the heart of categorical logic, and Lawvere has forged the way in extracting such common, deep principles. Even if people have heard of the structural set theory due to him known as the Elementary Theory of the Category of Sets (ETCS), or of the attempt to take a category of categories as foundational or even of his *elementary toposes*, one should not ignore Lawvere's success more generally in extracting the common essence of apparently diverse settings. In his paper 'Metric Spaces, Generalized Logic and Closed Categories' (Lawvere 1973), he speaks of a 'generalized logic' which makes apparent the commonality between ordinary categories and a kind of metric space, a space equipped with a distance. Categories collect together objects of a certain kind and relate them by transformations of the appropriate kind. Then for any two objects, X, Y, in a category, there is a 'hom-set', $Hom(X, Y)$, that is, the set of arrows or morphisms representing appropriate transformations between the objects. One of the fundamental rules of a category is the capacity to compose such arrows in the situation where the tail of one matches the head of another. So we require a composition map: $Hom(A,B) \times Hom(B,C) \to Hom(A,C)$. Now, Lawvere noted that this strongly resembles the triangle inequality, $d(x,y) + d(y,z) \geq d(x,z)$, which holds in metric spaces – it is always at least as far to go two sides around a triangle as along the third edge. He then showed that metric spaces can be seen as a variant of the former, where instead of the relationship between two objects being captured in a *set* of arrows, now the relationship between two points of space is determined by a *distance*, an element in the extended non-negative real numbers. So categories other than the category of sets may play the role of providing values for $Hom(A, B)$, and to do so they have to be equipped with additional structure. In the case here, a category which is enriched in the *monoidal* partially ordered set $([0, \infty], \geq)$ is a kind of metric space.

General enriched category theory allows a range of such monoidal categories to provide values for Hom objects, including the basic choice of the category of truth values. This is a category with only two objects, $\top$ and $\bot$. As required by a category, both objects have their identity arrow, representing the implications from *True* to *True* and *False* to *False*, and there is one further arrow from $\bot$ to $\top$, corresponding to the inference *False* to *True*. Evidently, we will not have an arrow in the other direction. In addition, we need a monoidal structure, here simply acting to multiply truth values by the ordinary propositional connective of conjunction. Now a category enriched in truth values amounts to a partially ordered set, a common mathematical structure.[4]

Returning to our search for an environment in which the pairing and unpairing of objects and the formation of function spaces are represented, category theory has determined the necessary features as that of being *cartesian* and being *closed*, respectively. These are very frequently encountered properties of categories. For a category to be cartesian it needs to possess a (finite) product structure, so in particular for any pair of objects, A and B, a product, which is an object denoted $A \times B$, with projections to A and B, such that any object with a pair of maps to A and B factors uniquely through the product. Being closed means that for any pair of objects, B and C, there is an object C^B, such that $Hom(A \times B, C) \cong Hom(A, C^B)$. Combining these two features gives us the concept of a cartesian closed category. These are just the ingredients for the kind of setting where $A \mapsto B^{B^A}$ becomes a natural map.[5]

So now combining the idea of enrichment by truth values with the structure provided by being cartesian closed, we arrive at a category that behaves like a collection of propositions where arrows correspond to entailment. Consider a category of propositions with an arrow from P to Q whenever Q follows from P. Then the product of P and Q is the conjunction $P\&Q$, and exponential objects are of the form $P \to Q$, hypothetical propositions. The cartesian closedness property of the category here amounts to $P \vdash Q \to R$ if and only if $P\&Q \vdash R$. Then, since we have the *modus ponens* entailment $P\&(P \to Q) \vdash Q$, so it follows that we have an entailment $P \vdash (P \to Q) \to Q$. In particular, when Q is $\bot$, that is, falsity, then there is a natural map from P to $\neg\neg P$.

To produce a map in the opposite direction, $\neg\neg P$ to P, as classical logic requires, will force us to add a much more specific structure to our categorical setting. There is a broad literature here which works with a cluster of relevant concepts involving *involution*, dagger categories, duals in vector spaces and Chu spaces, right up to star-autonomous categories, structures which have been made popular by categorical approaches to quantum mechanics (Abramsky and Coecke 2008). For our purposes, note that instead of the generic Q that we

[4] Strictly speaking, one might say a *preorder*, where it is possible for $x \leq y \leq x$ without x and y being identical.
[5] We will revisit these structures in greater detail in Chap. 2.

find in the entailment $P \vdash (P \to Q) \to Q$, to provide an entailment in the other direction will involve a specific choice of object, known as a dualizing object (Corfield 2017d).

What I have sketched above is just the very beginning of a highly developed relationship between category theory and logic, both proof-theoretic and model-theoretic aspects, which has fed into theoretical computer science in a vast number of ways. We have a range of conditions which categorical settings may possess in accordance with different logical principles. A forensic scrutiny of the different pieces of a logical framework allows for subtle variants, including the typed lambda-calculus corresponding to cartesian closed categories, and versions of linear logic corresponding to varieties of monoidal category. You want conjunction, and there must be products. Weakening and deletion, and the category should be cartesian. A deduction theorem, and it should be cartesian closed. Meanwhile, reasoning in a higher-order constructive logic is possible in toposes. Finally, the subject of this book, HoTT, has been constructed with a view to providing the logical calculus for an $(\infty, 1)$-topos, a kind of higher category that has been devised by mathematicians in recent years. As a computationally salient language, proof assistants based on HoTT are flourishing. A new chapter for the *trinitarian* thesis is being written. I should now explain the need for higher categories.

1.2 Next Steps

Let me resume my own path which led me next to *groupoids* and then *higher-dimensional algebra*, vital ingredients for the HoTT outlook. As we shall see in Chap. 3, groupoids mark the third stage in a hierarchy which sets out with *propositions* and then *sets*. Perhaps the simplest way to think of them is as a common generalization of *equivalence relations* and of *groups*. If an equivalence relation divides a collection of entities into subcollections of mutually equivalent entities, a group can be thought of as collecting the ways in which a single entity is self-identical. So, we may divide, say, a population of people into equivalence classes of those individuals who have the same age. Between any two people the relation 'having the same age' is a yes–no matter. In particular, any person has the same age as herself. On the other hand, to think of groups in a similar light, consider a single object which is multiply related to itself. For instance, take a symmetric butterfly which looks identical to its mirror image, or take the twenty-four rotations of a cube which leave it invariant. The collections of such symmetries form groups.

In diagrammatic form, an equivalence relation corresponds to a set of elements arranged in separate clusters, where each pair in the same cluster is joined by a linking edge, including a reflexive loop at each point. In our same-aged people, the lines mark the equality of age. The corresponding diagram for a group of symmetries is a single point with a collection of loops corresponding to the group elements or symmetries. For the cube, then, we have a single point with twenty-four loops (along with a way to compose the corresponding transformations). There were several properly mathematical motivations for doing so, but even without these one might think to combine the concept of an equivalence relation with that of a group. The resulting concept is that of a groupoid. With this concept in hand, we have a way of representing a collection of entities, any pair of which may be inequivalent