Does Homotopy Type Theory Provide a Foundation for Mathematics?

Friday 24th October, 2014

Abstract

Homotopy Type Theory (HoTT) is a putative new foundation for mathematics grounded in constructive intensional type theory, that offers an alternative to the foundations provided by ZFC set theory and category theory. This paper explains and motivates an account of how to define, justify and think about HoTT in a way that is self-contained, and argues that it can serve as an autonomous foundation for mathematics.

We first consider various questions that a foundation for mathematics might be expected to answer, and find that the standard formulation of HoTT as presented in the "HoTT Book" does not answer many of them. More importantly, the way HoTT is developed in the HoTT Book suggests that it is not a candidate *autonomous* foundation since it explicitly depends upon other fields of mathematics, in particular homotopy theory.

We give an alternative presentation of HoTT that does not depend upon sophisticated ideas from other parts of mathematics, and in particular makes no reference to homotopy theory (but is compatible with the homotopy interpretation). Our elaboration of HoTT is based on a new interpretation of types as *mathematical concepts*, which accords with the intensional nature of the type theory.

Contents

1 Introduction

Homotopy Type Theory (HoTT) is first and foremost a research programme within mathematics that connects algebraic topology with logic, computer science, and category theory. Its name derives from the way it integrates homotopy theory (which concerns spaces, points and paths) and formal type theory (as pioneered by Russell, Church, and Gödel, and developed in computer science) by interpreting types as spaces and terms of them as points in those spaces. Hence, the extant text on the theory [1], which we call 'the HoTT Book', involves homotopy theory throughout.

The authors of the HoTT Book are also concerned to develop and promote HoTT for working mathematicians, and to establish it as a foundation for mathematics – the subtitle of the HoTT Book is 'Univalent Foundations of Mathematics'.[1] Here the authors mean 'foundation' in the sense of a framework or language for mathematical practice. They make a strong case that HoTT can indeed serve as a foundation in this sense, demonstrating how to characterise mathematical structures such as natural numbers, real numbers, and groups in the language of HoTT, and how to use it to formalise proofs in homotopy theory.

However, philosophers often mean something stronger by 'foundation for mathematics' – they require a foundation to provide not just a language but also a conceptual and epistemological basis for mathematics, and moreover one that can be formulated without relying upon any other existing foundation.[2] If a system is to serve as a foundation for mathematics in this stronger sense it must be possible to present it in a way that does not make reference to sophisticated mathematical

---

[1] Indeed, much of the interest in HoTT is due to the fact that it may be regarded as a 'programming language for mathematics', and it is formulated in a way that facilitates automated computer proof checking [1, p. 10].

[2] This is the notion of an autonomous foundation similar to that of Linnebo and Pettigrew [2] that we define below.

domains such as homotopy theory.

This paper explains and motivates an account of how to define, motivate and think about HoTT in a way that is self-contained, and argues that so construed it can serve as an autonomous foundation for mathematics. (Of course, this paper does not present all the technical details of HoTT – for this, see [1] or [3] – but it does introduce its essential features.) It is important to note that we do not alter the framework of HoTT, but rather reconstruct and interpret the formal theory given in the HoTT Book, often answering questions that the authors of it did not address rather than answering them differently. This way of thinking about HoTT is compatible with the homotopy interpretation but does not presuppose it.

In order to judge whether HoTT provides a foundation for mathematics it is essential to be clear about what is meant by 'foundation'. There are many related notions in the folklore of the field, but not everyone in the field agrees about what is and is not required of a foundation of mathematics. It is precisely because this is not made make explicit that some disputes about the foundations of mathematics involve the protagonists talking past each other. In the next section we therefore list a number of questions of different kinds, pertaining to semantics, metaphysics, and epistemology, which a foundation for mathematics might be required to answer.[3]

Many foundational programmes give answers to some questions that make other questions particularly difficult. For example, if the proposed ontology of mathematics consists of abstract entities with no causal connection to the physical world then this raises the question of why we should believe in such entities, and how we can come to have access to them at all (the 'Access Problem' [4, pp. 30-33]). Ideally a foundational programme should answer philosophical questions about mathematics without generating new, more difficult questions, and its answers should hang together with each other, with mathematical practice, and with our broader understanding of the world. For example, its epistemological claims should follow naturally from the ontology without positing new mysterious ways for the mind to gain access to knowledge of non-physical things. The interpretation of HoTT we propose in this paper is intended to meet these desiderata, give answers to the questions of Section 2, avoid the Access Problem, and fit well with mathematical practice.

Our primary aim is not to advocate HoTT as a foundation for mathematics.[4] Nor is

---

[3] We make no great claims for the originality of our characterisation of foundation, but we are not aware of any similar complete characterisation in the literature. The most detailed treatment that we have found is the one due to Linnebo and Pettigrew [2], but ours differs very significantly from theirs. We give further references to extant discussions of foundations in footnote 8.

[4] Although we do believe that it has a number of advantages over the usual foundation

it to argue that mathematics must have a foundation that answers all the questions set out below.[5] Rather, our intention is to show that HoTT is a viable candidate autonomous foundation for mathematics over and above being a framework for mathematical practice, and to draw the attention of philosophers of mathematics to this new research programme that, as we argue below, offers an interestingly different way of thinking about mathematics.

Section 2 characterises a foundation for mathematics. Section 3 introduces HoTT. Section 4 argues that the existing presentation of HoTT is not autonomous. Section 5 explains our interpretation of HoTT. Section 6 outlines our novel autonomous justification of the way identity is handled in HoTT (summarising the argument of our companion paper [21]). Section 7 shows how our interpretation of HoTT answers the questions of Section 2. Section 8 replies to some possible objections and Section 9 briefly concludes by raising the issue of whether HoTT under our interpretation ought to be adopted as a foundation for mathematics.

2   What is a foundation for mathematics?

Any discussion of the foundations of mathematics is complicated by the fact that advocates of different foundational programmes often have different ideas about what it is to provide a foundation, and what is taken to be required is not always made explicit.[6] As mentioned above, one sense of foundation is that of a unifying language and conceptual framework, such as that of ZFC set theory. Another sense goes beyond this by adding *particular* definitions for mathematical entities in terms of that language (such as the point-set definition of topological space in terms of sets).[7]

On the other hand, as also pointed out above, philosophers often seek a stronger kind of foundation for mathematics. Beyond merely giving a language for mathematics, a foundation in this sense involves providing a grounding for mathematics

---

provided by ZFC set theory, some of which are described in Section 9.1.

  [5] Some authors deny that any foundation for mathematics is necessary at all (see e.g. [5]). However, notice that some 'anti-foundationalists' who deny that we need and/or can have foundations of a particular kind may be happy with foundations in a weaker sense. For example, Awodey [6, 7] rejects a particular conception of foundationalism, but nonetheless he does think we need to give answers to some of the questions we set out below. Likewise Shapiro [8] rejects 'foundationalism' while doing work in foundations more generally.

  [6] See [9] for recent reviews of the main foundational programmes.

  [7] The distinction between these two positions essentially comes down to a question of how much of the groundwork of mathematics should be counted as part of the foundation rather than as mathematical work built upon the foundational language. No philosophical issues relevant to the present discussion hang on this, and so we set aside this issue.

in pre-mathematical ideas, and answering semantic, metaphysical, epistemological, and/or methodological questions about mathematics.[8] In what follows we will use the word 'foundation' to denote a foundation in this stronger sense unless otherwise indicated, reserving the work 'framework' for the weaker sense of foundation as language and conceptual framework. This is a vague distinction but the notion of a framework is intended distinguish the foundational work of many mathematicians that is not concerned with the details of epistemology and metaphysics, from most foundational work in the philosophy of mathematics.

Even those who are interested in foundations in the strong sense do not agree about their required features. There are therefore no necessary and sufficient conditions for a system to constitute a foundation. Rather, in the next subsection we offer a characterisation of five components that a foundation might be expected to have, and articulate a series of questions that can be asked for each component. In Section 2.2 we introduce an important criterion, namely that a foundation be *autonomous*, in the sense that it is not built upon sophisticated extant mathematics. In the rest of the paper we use the analysis of this section to explore the foundational status of HoTT.

2.1 A characterisation of a Foundation for Mathematics

There are five interrelated components to a foundation for mathematics, and each generates a series of questions which may or may not have answers in the case of a given putative foundation for mathematics. Note that we do not claim that a putative foundation must be able to answer all these questions, nor must it definitively settle all the issues considered by philosophers of mathematics. However, to provide a foundation in the stronger sense a system must at least say *something* in relation to questions of semantics, metaphysics, and epistemology (even if only to say that no interpretation, metaphysics and/or epistemology is to be given).

(1) A single **framework** in which to cast, and perhaps also develop, some or all of existing mathematics. This framework involves a mathematical language and theory that may be studied in its own right as well.

Questions: How are complex higher-level elements constructed or composed out of more basic ones? Is it built on formal logic or not?[9] What role, if any, do axioms play?

---

[8] See [10] for one explicit account of this sense of a foundation for mathematics. For a discussion of different varieties of foundations see [8, Chapter 2]. The classic discussion of the foundations of mathematics is Part I of [4].

[9] Some foundational programmes presuppose a background formal logic as with, for example, axiomatic set theory, or with logicism which reduces mathematics to logic. Others incorporate a

(2) A **semantics** in the sense of an account of the basic concepts of the framework and how the theoretical terms of (1) are to be understood, how concepts are related, and an account of how the rules that are used to manipulate the concepts are to be understood.

Questions: What are the basic concepts? Are statements expressed in the language to be understood as potentially having truth values, and if so is the logic bivalent? How are the rules governing the terms to be understood? Is the theory extensional or intensional?

(3) A **metaphysics** that spells out the ontological status of any entities posited in (2).

Questions: Does the metaphysics posit any objects at all?[10] Is the ontology (if any) to be understood as mind-dependent or mind-independent? What is the relationship between mathematical reality (if any) and physical reality?

(4) An **epistemology** in the sense of an account of how we are able to know the truths (if any) of mathematics, given (2) and (3) (which may also include an account of the applicability of mathematics) and a justification of the axioms and rules of the framework.

Questions: Given the answers to the above questions, what account, if any, do we give of mathematical knowledge? In particular, if basic entities are posited, how do we know about them?[11] Given the role of proof in mathematical practice, what is the relationship between mathematical knowledge and proof? If there are axioms, what is their epistemological status – for example, are they taken to be known, or are they taken to be merely hypothetical statements that form the antecedent of conditionals? What is the relationship between mathematical knowledge and knowledge of physical reality?

(5) A **methodology** for mathematical practice based on some or all of the above.

Questions: How is the foundation to be used in practice? In particular, how is it to be applied in the physical sciences?

---

logic informally by way of the rules that govern the mathematical theory, or advocate a particular logic for metaphysical and/or epistemological reasons, as with constructive logic and Intuitionism. See [8] for a detailed investigation of the relationship between mathematics and logic.

[10] We may take a 'nominalist' position, denying (or at least not asserting) that there are any mathematical objects [11], in which case our answers to subsequent questions must be compatible with this.

[11] For example, do we take the indispensability argument to justify a belief in mathematical objects? See [12].

2.2 Autonomy

The answers given to the above questions may in some presentations be expressed in terms of some existing sophisticated mathematics. For example, we might give an account of category theory that assumes a background of ZFC set theory and defines a category to be a set of objects and a set of morphisms. However, if this were the only way to present the system then it would be hard to justify calling it a 'foundation' at all, since the concepts that were actually taken to be (conceptually, logically, or ontologically) fundamental would be those of the background mathematical system.

We call a presentation of a putative foundation *autonomous* iff the answers given to the above questions do not depend upon sophisticated ideas from other parts of mathematics.[12] An autonomous foundation (i.e. one with an autonomous presentation) must therefore be grounded in pre-mathematical ideas, and all of its rules must be justified by appeal to pre-mathematical principles.[13] (Of course, it may also have other non-autonomous presentations which may have been the original motivation for devising the system, and may even be more convenient or fruitful ways to think about it. Autonomy requires only that there is *some* way of arriving at the system via pre-mathematical considerations.)

As we noted above, mathematicians are often not at all concerned with the epistemology and metaphysics of mathematics, and so there is no reason to demand that a foundation in the weaker sense of *framework* for mathematics be autonomous. It is therefore entirely appropriate that these issues are not considered in the HoTT Book, and that the presentation in the HoTT Book weaves the homotopy theory and the type theory together from the beginning.

Hence, while in what follows we argue that the presentation of HoTT given by the authors of the HoTT Book is not autonomous, this should not be taken as criticism of their project in general or of the HoTT Book in particular, since their aim in writing the book was to communicate their work to other mathematicians and computer scientists, and not to provide a foundation for mathematics in the philosophers' sense outlined above. It is because we think that HoTT is so promising as a framework for mathematics that we also think it is worth exploring the extent to which it is a foundation more broadly. Our work should therefore be understood as complementary to the HoTT Book.

---

[12] We take this terminology from Linnebo and Pettigrew [2], who further divide the notion into 'logical autonomy', 'conceptual autonomy', and 'justificatory autonomy'.

[13] As John Mayberry puts it: "the primitive concepts of mathematics are those in terms of which all other mathematical concepts are ultimately defined, but which themselves are grasped directly, if grasped at all, without the mediation of a definition" [10, p. 18].

3   The basic concepts and features of HoTT

In this section we summarise the basic language of HoTT and consider some of the motivation for its rules.[14]

HoTT is built on the constructive intensional type theory of Martin-Löf [13]. The basic framework of HoTT is a formal system consisting of *types* and *tokens*, in which each token belongs to exactly one type.[15] We write 'a : A' to denote that token a is of type A.

3.1   The basic rules

We begin with functions, which are defined as in lambda calculus [14] by means of substitution into expressions.[16] A function of type $A \to B$ is something that, when given a token of type A, returns a token of type B. This is defined by an expression containing zero, one, or more instances of a variable, such that when all instances of the variable are replaced by an expression denoting a token of A (making sure to avoid re-using variable names in a way that changes the meaning of the expression) the resulting expression denotes a token of B.

Having defined this basic notion we can use it in stating the other definitions, thereby avoiding having to talk in terms of expressions denoting tokens and types.

A few basic rules for constructing new types from old ones are provided. To define a type we require the following components:

- A *type former* giving the type itself;

- One or more *token constructors*: functions whose output is of the type being defined;

---

[14] This brief survey just introduces the main ideas of HoTT. For a much more detailed exposition see [1] or [3].

[15] Terminology varies in the literature – in the HoTT Book the words 'term', 'object', 'element', and 'point' are used interchangeably for what we are calling a 'token' of a type. However, each of these words carries a connotation that we wish to avoid: 'term' suggests something syntactic, 'object' begs the question about the semantics of the interpretation, 'element' suggests ideas from set theory, and 'point' suggests a spatial interpretation. We therefore prefer the word 'token', which is (we hope) metaphysically neutral, and which we intend to be understood as a term of art independently of other usage.

[16] While lambda calculus is a sophisticated domain of study, the basic notion of substitution in an expression is a simple pre-mathematical one, familiar to anyone who, for example, is able to use pronouns in natural language. We therefore do not consider this to be an obstruction to the autonomous status of HoTT.

- An *elimination rule*: a general method for constructing functions whose input type is the type being defined.

We may also provide *computation rules* specifying how the token constructors and elimination rule interact.

The basic operations of the language of HoTT are as follows:

- As noted above, given any two types A and B we can form the *function type* A → B. Tokens of this type are functions that take a token of A as input and return a token of B as output.

- Given any two types A and B we can form the *product* type, A × B. A token of A × B is a pair (a, b), where a : A and b : B. The elimination rule is given by the obvious projection functions A × B → A and A × B → B.

- Given any two types A and B we can form the *coproduct* type A + B. Tokens of A + B are formed by the functions inl : A → A + B and inr : B → A + B. The elimination rule says that if we have functions g_l : A → C and g_r : B → C (for some type C) then we can form a function g : A + B → C that examines its input and applies either g_l or g_r to it as appropriate.

- We can generalise the function type to lift the restriction that the output of a function is always of a particular fixed type regardless of what token of the input type it is given. Given a type A and a 'family of types indexed by A' (i.e. a means of specifying, for each a : A, a type P(a)), we can define the *dependent function type* which we write as ⟨x : A⟩ → P(x) or $\prod_{x:A} P(x)$. Tokens of this type, called 'dependent functions', are functions that return a token of type P(a) when given a : A as input.

- Similarly, we can generalise product types to lift the restriction that the second component of a pair is always of a particular fixed type regardless of what the first component is. Given a type A and a family of types P(a) indexed by A as above, we can define the *dependent pair type* which we write as ⟨x : A⟩ × P(x) or $\sum_{x:A} P(x)$. Tokens of this type, called 'dependent pairs', are pairs (x, q), where x : A and q : P(x).

- There is a Unit type 1 which has a single token former producing the single token ∗ : 1. To define a function f : 1 → C (for any given output type C) we need a single token c : C to serve as the value of f(∗).

- There is a Zero type 0 which has no token formers. For any given output type C there is a function $!_C$: 0 → C.

10

The above system is not all of HoTT, since it does not provide anything that corresponds to identity or equality. We call the system sketched above HoTT$^-$ to distinguish it from full HoTT. The rules for identity are introduced in Section 3.3 and the elimination rule for identity is of particular importance in what follows.

3.2   Propositions as Types

The above type definitions can be motivated and justified via an interpretation of types as *propositions*. It was noted by Curry and Howard (and also by Brouwer, Heyting, and Kolmogorov) that there is a correspondence between type theory and natural deduction. Operations in type theory that involve the construction of an output token from some input tokens (computations) correspond to inferences in constructive logic. For example, if we have a function f of type $A \to B$ and a token x of type A then applying f to x gives, by definition, a token of type B. This is formally parallel to the rule of *modus ponens*: if proposition $A \Rightarrow B$ is true, and proposition A is true, then it follows that proposition B is also true. The Curry-Howard correspondence (or 'equivalence' or 'isomorphism') extends this to other logical operations.

Per Martin-Löf's *meaning interpretation* of propositions [15] makes this more than a merely formal parallel. In his system every proposition may be represented by a collection of things that stand as 'evidence' or 'proofs' of that proposition, and logical reasoning is then simply the manipulation of pieces of evidence according to rules. The Curry-Howard correspondence is understood as saying that to each proposition there corresponds a type, the tokens of which are proofs (or as we say, 'certificates') of that proposition.[17] Thus a proposition is true just in case there is a token of the corresponding type (the type is said to be 'inhabited'). This leads to the Brouwer-Heyting-Kolmogorov (BHK) interpretation [16, p. 161], in which a certificate to a compound proposition can be understood in terms of certificates to its constituent propositions. The basic rules for constructing new types from old ones given above then correspond to basic logical operations such as conjunction and disjunction, as spelled out below:

- The *implication* relation $A \Rightarrow B$ between propositions is represented by the *function type* $A \to B$. The application of a function to an input corresponds

---

[17] Again, terminology varies in the literature. Martin-Löf [15] uses the word 'proofs', which we avoid for reasons explained below. In the HoTT Book the words 'witness' and sometimes 'evidence' are used. We prefer the word 'certificate' over the other alternatives, since this makes the right analogy with everyday life: whereas evidence *indicates* the truth of a proposition, and a witness *observes* the truth, a certificate is something whose existence makes a proposition true – consider a legal document such as a driving licence or a marriage certificate.

to *modus ponens*, as described above.

- The *conjunction* of two propositions is represented by the *product* of the corresponding types, `A` × `B`. The token construction rule corresponds to the rule of &-introduction: from a certificate to `A` and a certificate to `B` we can produce a certificate to the conjunction. Similarly, the elimination rules correspond to &-elimination.

- The *disjunction* of two propositions is represented by the *coproduct* of the corresponding types, `A` + `B`. The token construction rules correspond to the ∨-introduction rule of constructive logic: to produce a certificate of $A \vee B$ we must have either a certificate of A or a certificate of B. The elimination rule corresponds to ∨-elimination: given certificates to $A \Rightarrow C$ and $B \Rightarrow C$ then we can produce a certificate to $(A \vee B) \Rightarrow C$.

- The Zero type corresponds to a contradictory proposition, hence no certificate to this proposition can be (directly) produced.[18] The elimination rule corresponds to the Law of Explosion: from a contradictory proposition any consequence follows.

- Given a proposition A with corresponding type `A`, the negation of A corresponds to the type `A` → `0`. A token of `A` → `0` is therefore a function that, if it were given a token of `A`, would produce a token of `0` which would certify the truth of a contradiction.

- A universally quantified proposition corresponds to the dependent function type $\prod_{\text{x:A}} \text{P(x)}$. A token of this type is a function that, when given a token `x` of type `A`, returns a token of type `P(x)`. The family of types indexed by `A` therefore corresponds to a *predicate* on `A`, and a dependent function of the above type is a certificate to the fact that all tokens of type `A` satisfy that predicate.

- An existentially quantified proposition corresponds to the dependent pair type $\sum_{\text{x:A}} \text{P(x)}$. A token of this type is a pair whose first component is a token `x` of `A` and whose second component is a certificate for the fact that this `x` satisfies the predicate. This is precisely what is required in constructive logic for a certificate to an existentially quantified proposition.

Thus the propositions as types interpretation in conjunction with the Curry-Howard correspondence and the BHK interpretation leads straightforwardly to

---

[18] Consistency of the system then consists in the claim that no token of `0` can be produced by any means.

the rules of HoTT$^-$. Moreover, this interpretation can be motivated and explained without any recourse to existing mathematics such as set theory.[19] What has been said so far can therefore be part of an autonomous foundation for mathematics.

The logic that is encoded into the rules of type formation and token construction is *constructive*: a proof is a construction process for building up a certificate to the conclusion. Since these rules give no way to construct a certificate for $A \vee \neg A$ or $\neg\neg A \Rightarrow A$ for arbitrary propositions A, the Law of Excluded Middle (LEM) and the law of Double Negation Elimination (DNE) do not hold as laws of logic. (However, wherever these rules are needed in a proof they can be assumed as premises and so classical reasoning can be recovered as required.)

Applying these type formation rules we can build up many complex types corresponding to complex mathematical propositions. A proof of such a proposition from some given premises consists of a sequence of applications of token constructors and elimination rules, beginning with the given certificates of the premises and ending with a certificate of the proposition.

The above rules and definitions provide the basic vocabulary of the language of HoTT (aside from identity, to be introduced in the next section). However, we can freely introduce new types by providing definitions: in particular, alongside types corresponding to propositions we can also introduce types that are understood more naturally as *kinds of mathematical entities*, such as the natural numbers $\mathbb{N}$. Again, we define such a type by specifying the type former (which gives the name of the type, perhaps depending upon some inputs), the token constructors (functions that return tokens of the type when given suitable inputs) and elimination rules (ways of defining functions from that type to any other).

This way of handling mathematical objects and mathematical propositions gives a natural way of mixing them together – for example, to pick out particular objects satisfying certain properties, or to assert that an object satisfies a particular property as part of its definition.

3.3　Identity

The above definitions are not sufficient to provide a language for mathematics, since they give us no way to assert that two things are equal, and so for example even simple truths of number theory cannot be stated.

The statement that two things are equal is a proposition, and so we must introduce a new type definition for this kind of proposition. Thus for each type `C` and any

---

[19] We must, of course, interpret 'domains of quantification' correctly: not as sets whose 'elements' are members, but as types whose 'elements' are tokens.

two tokens a : C and b : C we must have the type $\text{Id}_\text{C}(\text{a}, \text{b})$ that corresponds to the proposition that a and b are equal. A token of that type, which is called an 'identification' of a and b, is a certificate to the proposition.

The only identifications that can be produced with no further assumptions are the *trivial self-identifications*. Thus for any token a : C there is a token $\text{refl}_\text{a} : \text{Id}_\text{C}(\text{a}, \text{a})$ (where 'refl' is short for 'reflexivity of identity'). The function that maps each token a to its trivial self-identification refl_a is the token constructor for the identity type. However, although this is the only kind of identification that can be freely constructed, we cannot prove that these are the only identifications that exist. That is, without additional premises we cannot construct non-trivial identifications, but we also cannot prove that none exist.[20]

To complete the definition of the identity type we must also give its elimination rule, which is called 'path induction'. Just as mathematical induction on $\mathbb{N}$ allows us to prove that some condition holds of all natural numbers without considering each natural number individually, likewise path induction allows us to prove that some condition holds of all identifications in a particular identity type. In an induction on $\mathbb{N}$ we first prove that the condition holds of 0 (the base case), then show that if it holds of some n then it also holds of $n + 1$ (the inductive step). A proof using path induction is even simpler than this, since only a base case needs to be proved – the inductive step holds as a theorem for all properties that we might consider. That is, the principle of path induction says that, for any property P of identifications, if we can prove that P holds of all trivial self-identifications refl_a then it holds of all identifications.

Path induction is a very powerful tool since in many cases it makes proofs about identifications very simple. However, its interpretation and justification are not so straightforward. One natural way to interpret it is as a statement that all identifications are trivial self-identifications. However, as noted above this is not something that can be proved in HoTT, and moreover this is not the intended interpretation, since it eliminates all the structure of 'higher identity types' that are of particular interest in HoTT [1, pp. 66-67].

Identity types are of fundamental importance in HoTT. In particular, firstly, identity types and especially 'higher identity types' are fundamental to the homotopy interpretation, which provides a more tractable approach to problems previously treated in n-category theory. Second, the 'Univalence axiom' due to Voevodsky is about identity types and it was only when this axiom was added to the theory that its true power and significance became clear. This is illustrated by the fact the the subtitle of [1] is 'Univalent Foundations of Mathematics'.

---

[20] This is a common situation in constructive logic: consider, for example, the non-zero infinitesimals in Synthetic Differential Geometry [17].

However, the propositions as types interpretation set out in the previous section does not extend to give an account of identifications between tokens. Thinking of tokens as certificates to propositions gives us no reason to adopt path induction. Instead, the HoTT Book explains and justifies the principle via the homotopy interpretation of HoTT [1, p. 67], which we sketch below.

3.4   The homotopy interpretation

One of the major innovations of Homotopy Type Theory is the alternative interpretation of types and tokens it provides using ideas from homotopy theory, which was found by Awodey and Warren [18]. In this section we give a very brief survey of homotopy theory and the main points of this interpretation.[21]

Homotopies can be thought of as continuous distortions between functions, or between the images of functions. Facts about homotopy theory are therefore only given 'up to continuous distortions', and only facts that are preserved by all such distortions are well-defined. Homotopy is usually presented by starting with topological spaces. Given two such spaces X and Y, we say that continuous maps $f, g : X \to Y$ are *homotopic*, written '$f \sim g$', just if there is a continuous map $h : [0,1] \times X \to Y$ with $h(0,x) = f(x)$ and $h(1,x) = g(x)$ for all $x \in X$. Such a map is a *homotopy* between f and g. For example, any two curves between the same pair of points in the Euclidean plane are homotopic to one another, because they can be continuously deformed into one another. However, in a space with a hole in it (such as an annulus) there can be paths between two points that are not homotopic, since a path going one way around the hole cannot be continuously deformed into a path going the other way around the hole.[22]

Two spaces X and Y are *homotopy equivalent* if there are maps $f : X \to Y$ and $f' : Y \to X$ such that $f' \circ f \sim id_X$ and $f \circ f' \sim id_Y$. This is an equivalence relation between topological spaces, so we can define the equivalence class $[X]$ of all topological spaces homotopy equivalent to X, called the *homotopy type* of X. Homotopy theory does not distinguish between spaces that are homotopy equivalent, and thus homotopy types, rather than the topological spaces themselves, are the basic objects of study in homotopy theory.

In the homotopy interpretation of the basic language of HoTT we interpret types as homotopy types or 'spaces'. It is then natural to interpret tokens of a type

---

[21] We use the familiar language of set theory and classical logic for this presentation, rather than casting everything into the language of HoTT itself.

[22] These ideas have important applications in physics. For example, the celebrated Aharonov-Bohm effect depends upon the fact that the space in which the electrons move is punctured by a 'hole', namely the solenoid around which the electrons pass. [19, 20]

as 'points' in a space. However, due to the 'flexible' nature of homotopy types described above we must be careful about this. The points of topological space have what we might call absolute identity, being elements of the underlying set. But a homotopy equivalence will in general map a given point $x \in X$ to some other $x' \in X$, and so when we work with homotopy types the absolute identity of the points is lost. Rather, we must say that a token belonging to a type is interpreted as a function from a one-point space into the space.

Given two points a and b in a space X, a *path* between them is a function $\gamma : [0,1] \to X$ with $\gamma(0) = a$ and $\gamma(1) = b$. However, given any such path, X can be smoothly distorted by *retracting* the path along its length toward a. Thus a space containing two distinct points and a path between them is homotopic to a space in which both points coincide (and the path is just a constant path at this point). We may therefore interpret a path between points as an *identification* of those points. Thus the identity type $\mathtt{Id_X(a,b)}$ corresponds to the *path space* of paths from a to b. This also gives a straightforward justification for the principle of path induction: since any path is homotopic to a constant path (which corresponds to a trivial self-identification), any property (that respects homotopy) that holds of all trivial self-identifications must hold of all identifications.

This concludes our brief presentation of the standard formulation of HoTT.[23] In the next section we begin to consider its foundational status.

4 Answers to Some Foundational Questions from HoTT

The presentation of HoTT given in the previous section allows us to answer many of the questions posed in Section 2.

Framework The theory makes use of the language of HoTT, sketched above. Complex expressions are constructed systematically and precisely from basic ones by application of the rules.

HoTT is not built upon a background logic in the way that axiomatic ZFC is. Rather, constructive logic is incorporated into the mathematical theory, and propositional and predicate logic can be expressed within it. The basic language of HoTT does not involve an axiomatic theory, and the role that axioms play in other foundations is played instead by the rules of type formation. However, additional assumptions can be made, in the form of tokens

---

[23] For more details see [1, Chapter 1].

of types that are assumed as premises or axioms without being constructed from other premises.[24] One assumption that may be treated as axiomatic is the existence of the single token $*$ of the Unit type $\mathbf{1}$, as this is posited as part of the definition of that type rather than constructed from other resources.

Semantics The interpretation of types and tokens given in the HoTT Book is the homotopy interpretation summarised above, in which types are spaces (as understood in homotopy theory), tokens are points of spaces, and identifications are paths in spaces between the points that they identify.

Statements in the language are true if the corresponding type is inhabited, but since the Law of Excluded Middle is not adopted we do not have bivalence.

The rules of HoTT⁻ are justified by analogy with the introduction and elimination rules of propositional and predicate logic. The rules can be understood as telling us how to construct a token of one type given tokens of other types, and therefore tell us when some type's being inhabited follows from some other types' being inhabited.

The theory is intensional: types are distinguished from one another by their definitions (i.e. by their constructions) rather than by their extensions. Of course, since each token belongs to exactly one type there is no sense in which two inhabited types could have the same extension. But uninhabited types – for example, 'even divisor of 9' and 'even divisor of 11' – are treated as distinct. Moreover, any two types that might be understood as containing (representatives of) the same mathematical entities under different presentations – for example, 'positive integer less than 3' and 'integer exponent n for which $a^n + b^n = c^n$ has a solution in the positive integers' – are nonetheless considered to be distinct types.[25] Thus any distinction that can be made gives rise to a distinction between types. The justification given in the HoTT Book for using a theory that is intensional in this way is the fact that it enables the homotopy interpretation [1, pp. 71, 128].

Metaphysics Nothing is said about the metaphysics. In particular, it is not clear whether the tokens of types literally are mathematical objects (for example, natural numbers) or whether they represent them in some sense. We are free

---

[24] An important example of this is the Univalence Axiom [1, Section 2.10], which is central to the research programme developing Homotopy Type Theory. This is beyond the scope of the present paper, but identity types and Univalence are also of great philosophical significance, particularly in relation to mathematical structuralism. See [7].

[25] The addition of the Univalence axiom [1, Section 2.10] weakens these distinctions, since it gives identifications between types that are 'equivalent' in a suitable sense. But the basic language of the theory considered here does not include Univalence.

to apply any metaphysical interpretation we prefer. Nothing is said about applicability in the physical sciences.

Epistemology Since no metaphysics is given, we are likewise free to bring any epistemological account of the basic entities that we prefer. However, the constructive logic does have epistemological implications, since it gives an account of how mathematical knowledge can be had *given* any particular initial assumptions.

Methodology HoTT can be used informally, as is done in much of the HoTT Book. But such informal reasoning can be translated into a computer programming language. It can be used directly to construct formal proofs. The idea that the foundations for mathematics should be applicable to mathematical practice and to computational methods in mathematics, as is arguably not the case with set theory, is a major motivation for HoTT.[26] In HoTT the practicality of computer-assisted reasoning and proof-verification narrows the gap between everyday practice and idealised formal reasoning. Also arguably computability requires or at least is facilitated by a constructive framework.

4.1 Autonomy of this presentation of HoTT?

The standard presentation of HoTT is not autonomous because throughout the 'HoTT Book' the terminology and concepts of homotopy form a central part of the presentation of HoTT. In the book no distinction is made between the type theory per se and the homotopy interpretation of it.

As the authors themselves say, "Homotopy Type Theory (HoTT) interprets type theory from a homotopical perspective" [1, p. 3]. They go on to explain that "the key new idea of the homotopy interpretation" involves interpreting the logical notion of identity in terms of the existence of a path [1, p. 5]. Spaces, points, and paths are given centre stage when explaining major new ideas, such as the Univalence axiom. The fundamental principle of path induction is justified using homotopical reasoning (as sketched in Section 3.4 above). As mentioned above, the justification for using an intensional theory is the fact that it enables the homotopy interpretation.

Of course the fact that the homotopy interpretation is used throughout the presentation does not settle its exact role in the theory itself. Since much of the innovative use of HoTT is derived from the homotopy interpretation it is completely natural to want the reader become familiar with that interpretation *ab*

---

[26]IMathematical theories are not typically rendered into set-theoretic terms by casting them in first- or second-order predicate logic in accordance with the axioms of ZFC.

initio, so that they can use the theory effectively. It does not follow that the fundamentals of the theory inherently depend upon homotopy theory, or that an alternative presentation cannot be given.

In what follows we give an alternative account of the aspects of the standard interpretation of HoTT that depend upon homotopy theory or some other sophisticated mathematics in order to produce an interpretation of HoTT that can be an autonomous foundation for mathematics.

The issues that must be addressed are:

1. How can we understand types without the homotopy interpretation?

2. How are the rules of HoTT to be justified on the basis of that understanding? In particular, how is path induction to be justified without reference to the homotopy interpretation?

5 The interpretation of tokens and types

In the standard presentation of HoTT types are to be interpreted either as propositions or as spaces (or more specifically as homotopy types). As we pointed out in Section 3.3, the interpretation of types as propositions does not give a justification of path induction or other features of identity types. The homotopy interpretation does give a justification of path induction, and pervades the way of thinking about identity presented in the HoTT Book, but clearly depends upon the sophisticated machinery of homotopy theory and is therefore incompatible with the autonomy of the foundation.

One response to this problem is simply to drop the interpretation of types from our account of HoTT altogether, adopt a 'null semantics', and thus take an extreme *formalist* approach.[27] That is, we could say that the foundation provided by HoTT consists only of a formal language with rules for the manipulation of symbols, but that the symbols should not be interpreted as denoting (or having any other connection to) 'mathematical objects' of any kind. This is clearly compatible with the foundation's being autonomous, and furthermore eliminates any need to explain or justify any metaphysical assumptions, since none are made. We will not rehearse the objections to formalism here, but we do show that a richer foundational account can be given.

We might alternatively try to preserve as much of the standard presentation as possible by giving an account of spaces that contains the features needed to support

---

[27] Formalism in general comes in different forms, see Detlefsen's essay in [9].

this interpretation, but which is grounded in pre-mathematical intuitions rather than homotopy theory. That is, by inspection of our intuitive concept of what a space is, we might hope to recover those features of homotopy theory that are used in the homotopy interpretation of HoTT without needing to set up all the mathematical machinery that is standardly used to define it. However, it is not at all clear that such an argument – that the required features of homotopy theory are all present in our intuitive notion of space – can be defended, and we do not pursue this approach further in the present paper.

In the next subsection we will argue why tokens should not be interpreted as mathematical objects, before introducing our interpretation of types and tokens as concepts in Section 5.2.

5.1   Tokens as mathematical objects?

A natural way to interpret tokens and types is to say that tokens correspond to mathematical objects and types correspond to kinds of mathematical objects, so that `a : A` says that object `a` is of kind `A`. However, there are problems with this interpretation.

First, it commits us to a strong kind of Platonism, since we can only adopt this interpretation if we believe that there are mathematical objects; we would then have to give an account of the metaphysics of these objects, and our epistemic access to them.

Second, this interpretation does not naturally accord with the intensional treatment of types. In particular, while a description of an object may omit some of its properties and leave it partially specified, any actual object must possess all the relevant properties, whether they are specified in its description or not. Thus, for example, any actual triangle must be either equilateral, isosceles, or scalene; an abstract triangle that is merely triangular and not, more specifically, also equilateral, isosceles, or scalene is therefore problematic.

Third, whereas complex types are constructed from simpler types by application of the type formers, there is no corresponding sense in which complex kinds are composed from simpler kinds. The definition or description of a kind is of course constructed by a logical assembly of simpler notions, but the kind itself is not.

Fourth, this interpretation does not explain why we take each token to belong to exactly one type. We would normally say that a given mathematical object can be of multiple kinds, just as a physical object can be of multiple kinds – for example, a particular dog is of the kinds *dog*, *mammal*, and *pet*, and a particular number may be of the kinds *natural number*, *odd number*, and *prime number*. The restriction

to a single type for each token is, on this interpretation, both unexplained and unnatural.

Finally, this interpretation arguably gets the relationship between tokens and types backwards. When we classify physical objects into kinds, we take the objects themselves to be primary and the kinds to be secondary, dependent upon the objects. However, in the type theory sketched above the order is the other way around: token constructors can only be specified once the type itself has been defined. To adopt this interpretation we would therefore need to explain this reversal.

Hence, while this approach is initially appealing, it does not accord well with the features of the type theory to be interpreted.

5.2 Tokens and types as concepts

Our alternative interpretation originates in the observation that, we can have mathematical *concepts* whether or not there are any mathematical objects. For example, we have the idea of what it is to be a natural number, whether or not we believe that there are any such things.

We could therefore say that, rather than picking out mathematical objects, tokens correspond to specific mathematical concepts; and rather than denoting kinds of mathematical object, types correspond to general mathematical concepts. Thus `a : A` would say that the specific concept `a` falls under, or is an instance of, the general concept `A`.

It is clear that concepts are pre-mathematical. That is, without having learned any mathematics someone can understand the idea of a 'concept', and at no point does the development of this understanding require ideas from mathematics. Thus this interpretation is suitable for an autonomous foundation for mathematics.

To develop this interpretation we must say a few things about concepts. Of course we do not here give a complete account of concepts, but we highlight some features that are relevant to the present discussion.

1. The existence of concepts does not depend upon the existence of specific objects that they may represent. So, for example, existence of the concept of 'unicorn' does not require that there be any unicorns in the world, and having the concept does not commit us to believing that any exist. Indeed, it is necessary to have the concept 'unicorn' even to be able to frame the denial that such things actually exist.

2. Concepts can be of the concrete or abstract, and can be specific or general. For example, 'mammal' is a concrete general concept, while 'justice' is an abstract specific concept. We take it that for any specific object or general kind there is a corresponding concept. Thus a foundation that takes concepts as primitive entities is compatible with Platonism about mathematics.[28]

3. While concepts only have effects via mental activity, they may or may not depend for their existence on mental activity. That is, the concept 'triangle' may exist even while no-one is presently thinking about triangles, and moreover the concept 'elliptic curve' may have existed even before anyone first conceived of elliptic curves – that is, we might say that the concept did not spring into existence when the definition was first stated. Thus a foundation that takes concepts as primitive entities is not thereby committed to saying that mathematics is subjective or intersubjective rather than being concerned with an objective mind-independent subject matter.

4. Concepts are intensional. That is, they correspond (roughly) to descriptions rather than to extensional collections. Hence (to use two famous examples), 'the morning star' and 'the evening star' are two distinct specific intensions, although they have the same extension; and 'human' and 'featherless biped' are distinct general intensions although they have the same extension.[29] We can have empty concepts, even *necessarily* empty concepts.

5. Concepts can be composed and refined to form more complex concepts.

6. Being intensional, concepts can be partially specified. So, for example, whereas any specific triangle must be equilateral, isosceles, or scalene, the concept 'triangle' is simply 'plane figure with 3 sides' and no relationship between its side-lengths is required. Concepts can therefore be just as precisely specified as required, in a way that objects cannot.

7. Specific concepts can stand in relation to general concepts just as specific objects stand in relation to kinds; for example, courage is a virtue.

Let's consider again the problems that faced the interpretation of tokens and types as objects and kinds. Since the existence of concepts does not depend upon the existence of entities, we are not committed to mathematical Platonism. The intensional treatment of types fits perfectly with an interpretation of types as concepts,

---

[28] It is an open question whether the posited objects would then be redundant in mathematical practice, and we do not address this here.

[29] Moreover, concepts are *hyperintensional*, since even intensions that are necessarily coextensional, such as 'equiangular triangle' and 'equilinear triangle', are distinct.

since concepts are themselves intensional. Complex concepts may be composed from simpler ones, just as complex descriptions can be composed.

However, the interpretation of tokens and types as specific and general concepts respectively does not answer the remaining two issues, namely why each token belongs to exactly one type, and the dependence relationship between tokens and types. Our account must therefore be slightly refined to address these points.

Since concepts are intensional they can be specified in ways that emphasise some particular feature or aspect. We therefore interpret a token not as the concept of some specific mathematical thing, but rather as the concept of a specific thing *qua* instance of a general concept. This explains why each token belongs to exactly one type, and gets the order of dependence right: we must have the general concept first before we can have the concept of some specific thing *qua* instance of that general concept.

Finally, how does this interpretation accord with the treatment of identity in HoTT? What does it mean to identify two specific concepts (*qua* instance of some general concept)?

Consider, for example, two triangles at two different locations in the Euclidean plane. Since they are at different locations they are clearly distinct – two triangles in the Euclidean plane are identical just if they precisely the same three vertices. However, if we abstract away their locations the criterion of identity becomes *congruence* (i.e. having the same angles and side-lengths). If we further abstract away considerations of size then the criterion of identity becomes *similarity* (i.e. having the same angles). Thus two triangles in the Euclidean plane may be numerically distinct whilst being identical *qua* instance of some general concept (whereby some of their specific features are abstracted away).

By allowing us naturally to abstract away some features and retain others the interpretation as concepts fits well with the treatment of identity.[30] Moreover, this gives an illustration as to how a pair of tokens may be identified in multiple distinct ways. For example, two triangles can be shown to be similar by giving a correspondence between their angles, or alternatively by showing that their side-lengths are the same.

This brief survey shows that the interpretation of types and tokens as concepts accords well with the intensional nature of the type theory, and how how we think about types, terms and their identity. Note that this interpretation is an extension,

---

[30] Note, however, that this is not necessarily the only way we might have non-trivial identifications. We might, for example, simply stipulate that there are two numerically distinct instances of some general concept that are identified (i.e. making their distinctness merely a 'brute fact') without providing some other type from which they are abstracted.

rather than a replacement, of the propositions as types view.

6 Justifying the elimination rule for identity

We separate identity from the rest of the basic language of HoTT (which we call HoTT$^-$) because the justification of the rules of HoTT$^-$ that follows from the BHK interpretation, as described above, does not extend to the rules for identity types in HoTT.

The type former and token constructor are straightforward. The type former for identity types says that for any tokens `a`, `b` of the same type we can form the proposition `a = b` asserting that they are equal.[31] However, tokens of different types cannot be equated. The token constructor for identity types allows us to say that any token of any type is identical to itself, i.e. that identity is reflexive. It is clear both that this is an essential defining property of identity (so we are justified in having this token constructor), and that we should not in general be able to construct identifications between arbitrary pairs of tokens (so no further token constructors are required).[32]

However, the elimination rule for the identity type, namely the principle of path induction explained in Section 3.3, is not so easily justified. Recall that it says that if a property holds of all trivial self-identifications then it holds of all identifications. The justification given in the HoTT Book (and summarised in Section 3.4) depends upon the interpretation of identifications as paths in (homotopy) spaces – since any path is homotopic to a constant path (since it can be continuously retracted), properties of constant paths are shared by all paths. However, as we argue above, such a justification cannot form part of an autonomous foundation for mathematics, for which an alternative account must be found.

In a companion paper [21] we investigate this problem, and set out a justification for path induction that depends only on HoTT$^-$ under our interpretation (which we have argued above is suitable for an autonomous foundation) and two other principles that can be justified independently of the homotopy interpretation and do not depend upon any other advanced mathematics. Without going into all the technical details, we sketch the argument in the remainder of this section.

---

[31] Being able to *form* this proposition for any two tokens is not the same as being able to *prove* the proposition for any two tokens, of course. Forming the proposition amounts to forming the type, but proving amounts to showing the type is inhabited.

[32] We might think that, since symmetry and transitivity are also essential characteristics of identity, they should also be built into to the definition via new token constructors. However, they don't need to be imposed separately because they can be derived from the more basic principle we introduce below as we show in [21].

One obvious application of identity types is to express facts about the uniqueness of tokens of other types. For example, intuitively there should be exactly one token of the Unit type, namely the token $* : \mathbf{1}$ given by the token constructor for this type. Since HoTT uses constructive logic we cannot assert the *non-existence* of tokens. However, we can assert that all tokens of the Unit type are identical to the token $* : \mathbf{1}$, thereby saying that there is exactly one token of $\mathbf{1}$ *up to identity* in $\mathbf{1}$. In other words, we cannot rule out the existence of tokens of $\mathbf{1}$ other than $*$, but we can say that any that exist are equal to $*$ in $\mathbf{1}$.

More generally, for each type we have a uniqueness principle characterising the tokens of that type up to suitable identifications. Often the relevant uniqueness principle says that every token of the type is equal to the output of a token constructor, but this is not the case for every type.

In particular, the uniqueness principle for an identity type cannot say that every identification $\mathtt{p} : \mathtt{Id}_{\mathtt{C}}(\mathtt{a}, \mathtt{b})$ is equal to the output of a token constructor, since the constructors only produce identifications of the form $\mathtt{refl}_{\mathtt{x}} : \mathtt{Id}_{\mathtt{C}}(\mathtt{x}, \mathtt{x})$, and we cannot identify tokens of different types.

To characterise the uniqueness principle for identity types we therefore need to define a type that contains representatives of both arbitrary identifications such as $\mathtt{p} : \mathtt{Id}_{\mathtt{C}}(\mathtt{a}, \mathtt{b})$ and trivial self-identifications such as $\mathtt{refl}_{\mathtt{x}} : \mathtt{Id}_{\mathtt{C}}(\mathtt{x}, \mathtt{x})$.

For this purpose we define, for each $\mathtt{a} : \mathtt{C}$, the 'based identity type' at $\mathtt{a}$

$$\mathtt{E}_{\mathtt{a}} \; :\equiv \; \sum_{\mathtt{x}:\mathtt{C}} \mathtt{Id}_{\mathtt{C}}(\mathtt{a}, \mathtt{x})$$

The tokens of this type are pairs $(\mathtt{b}, \mathtt{p})$ consisting of a token $\mathtt{b} : \mathtt{C}$ and an identification $\mathtt{p}$ between that token and the given token $\mathtt{a} : \mathtt{C}$. In particular, one token of this type is $(\mathtt{a}, \mathtt{refl}_{\mathtt{a}})$. Thus for any identification $\mathtt{q} : \mathtt{Id}_{\mathtt{C}}(\mathtt{x}, \mathtt{y})$ we can find a counterpart $(\mathtt{x}, \mathtt{q})$ to it in the based identity type $\mathtt{E}_{\mathtt{x}}$, alongside a counterpart to the trivial self-identification $\mathtt{refl}_{\mathtt{x}}$.[33]

Naively, we would expect that the only token of $\mathtt{C}$ that is identical with $\mathtt{a}$ is $\mathtt{a}$ itself, and moreover that it should be self-identical in just one way. Thus naively we would expect $\mathtt{E}_{\mathtt{a}}$ to have only a single token. However, we have seen above that we cannot rule out the existence of non-trivial identifications, and thus tokens of $\mathtt{E}_{\mathtt{a}}$ other than $(\mathtt{a}, \mathtt{refl}_{\mathtt{a}})$. As with the Unit type, the role of the uniqueness

---

[33] We could instead define the 'free identity type' whose tokens are *triples* $(\mathtt{a}, \mathtt{b}, \mathtt{p})$ consisting of two tokens of $\mathtt{C}$ and an identification between them. This type also contains a counterpart to every identification, alongside counterparts to trivial self-identifications. However, the justification of the elimination rule for identity types is more easily expressed in terms of the based identity type defined above, so we restrict attention to this, and thus directly justify what is known as 'based path induction' from which path induction also follows as shown in subsection 1.12.2 of [1].

principle is to say that if any such tokens exist then they are equal to $(\mathtt{a}, \mathtt{refl_a})$ in $\mathtt{E_a}$. More precisely, for any token $(\mathtt{b}, \mathtt{p}) : \mathtt{E_a}$ there is an identification between it and $(\mathtt{a}, \mathtt{refl_a})$. We express this formally by saying that the following type

$$\prod_{(\mathtt{b},\mathtt{p}):\mathtt{E_a}} \mathtt{Id_{E_a}}((\mathtt{a}, \mathtt{refl_a}), (\mathtt{b}, \mathtt{p}))$$

is inhabited. This is the uniqueness principle for identity types.

The elimination rule for identity type now follows as an application of the principle of *substitution salva veritate* – that identical elements of a proposition can be substituted without changing the truth the proposition.[34] That is, if we can prove that some property holds of all trivial self-identifications – or, more precisely, all pairs $(\mathtt{a}, \mathtt{refl_a})$ for any $\mathtt{a} : \mathtt{C}$ – then it must likewise hold for any pair $(\mathtt{b}, \mathtt{p})$, since the uniqueness principle says that this is identical to $(\mathtt{a}, \mathtt{refl_a})$.

Thus by applying the uniqueness principle for identity types and the principle of *substitution salva veritate*, each of which is justified without reference to sophisticated existing mathematics, we derive the elimination rule for identity types.[35]

7   The foundations of HoTT without homotopy

In Section 4 we give answers to the questions of Section 2 for the presentation of HoTT given in the HoTT Book. In this section we answer these questions for the new interpretation of HoTT that we explained in Section 5.2.

7.1   Framework

The framework of the theory is as described in Section 3.1: the language of HoTT, incorporating constructive logic, and assuming no axioms (except for the existence of the single token of the Unit type).

7.2   Semantics

Types and tokens are interpreted as *concepts* (rather than spaces, as in the homotopy interpretation). In particular, a type is interpreted as a general mathematical

---

[34] One might object that this principle is in conflict with intensional treatment of types. In Section 8.4 we explain why this does not pose a problem.

[35] For a detailed version of this argument, along with proofs that other properties of identity such as symmetry and transitivity follow from *substitution salva veritate*, see [21].

concept, while a token of a given type is interpreted as a more specific mathematical concept *qua* instance of the general concept. This accords with the fact that each token belongs to exactly one type. Since 'concept' is a pre-mathematical notion, this interpretation is admissible as part of an autonomous foundation for mathematics.

Expressions in the language are the names of types and tokens. Those naming types correspond to propositions. A proposition is 'true' just if the corresponding type is inhabited (i.e. there is a token of that type, which we call a 'certificate' to the proposition). There is no way in the language of HoTT to express the absence or non-existence of a token. The negation of a proposition P is represented by the type $P \to 0$, where P is the type corresponding to proposition P and 0 is a type that by definition has no token constructors (corresponding to a contradiction). The logic of HoTT is not bivalent, since the inability to construct a token of P does not guarantee that a token of $P \to 0$ can be constructed, and vice versa.

The rules governing the formation of types are understood as ways of composing concepts to form more complex concepts, or as ways of combining propositions to form more complex propositions. They follow from the Curry-Howard correspondence between logical operations and operations on types. However, we depart slightly from the standard presentation of the Curry-Howard correspondence, in that the tokens of types are not to be thought of as 'proofs' of the corresponding propositions but rather as certificates to their truth. A proof of a proposition is the construction of a certificate to that proposition by a sequence of applications of the token construction rules. Two different such processes can result in construction of the same token, and so proofs and tokens are not in one-to-one correspondence.

When we work formally in HoTT we construct expressions in the language according to the formal rules. These expressions are taken to be the names of tokens and types of the theory. The rules are chosen such that if a construction process begins with non-contradictory expressions that all name tokens (i.e. none of the expressions are 'empty names') then the result will also name a token (i.e. the rules preserve non-emptiness of names).

7.3   Metaphysics

Since we interpret tokens and types as concepts, the only metaphysical commitment required is to the existence of concepts. That human thought involves concepts is an uncontroversial position, and our interpretation does not require that concepts have any greater metaphysical status than is commonly attributed to them [22].

Just as the existence of a concept such as 'unicorn' does not require the existence of actual unicorns, likewise our interpretation of tokens and types as mathematical concepts does not require the existence of mathematical objects. However, it is compatible with such beliefs. Thus a Platonist can take the concept, say, 'equilateral triangle' to be the concept corresponding to the abstract equilateral triangle (after filling in some account of how we come to know about these abstract objects in a way that lets us form the corresponding concepts).

Even without invoking mathematical objects to be the 'targets' of mathematical concepts, one could still maintain that concepts have a mind-independent status, i.e. that the concept 'triangle' continues to exist even while no-one is thinking about triangles, and that the concept 'elliptic curve' did not come into existence at the moment someone first gave the definition. However, this is not a necessary part of the interpretation, and we could instead take concepts to be mind-dependent, with corresponding implications for the status of mathematics itself.[36]

7.4   Epistemology

Since the metaphysics of this interpretation does not include mathematical objects, there is no issue of explaining how we come to know about abstract objects. We take it that there is no corresponding problem with regard to explaining how we have epistemic access to the concepts we form. Since complex concepts are formed step-by-step from more basic ones, our access to any given concept can be traced backward via the given definition to whatever elementary concepts are assumed given.

Similarly, since the truth of a proposition corresponds to the existence of a certificate, and a proof is a step-by-step construction of a certificate to the conclusion from certificates to the premises, a straightforward account of the status of mathematical knowledge follows. Any claim to mathematical knowledge can be demonstrated by showing a specific way of constructing a certificate to the corresponding propositon. This is true even of proofs of negated propositions, since the classical technique of Indirect Proof (proving P by deriving contradiction from the negation of P) is disallowed – since HoTT uses constructive logic, proofs always result in positive evidence for their conclusions.

Where a proof depends upon premises it of course provides only conditional knowledge of its conclusion. In this case the assumed premises – and moreover the

---

[36] Since, as noted in Section 1, our aim in this paper is to defend the claim that HoTT is a candidate autonomous foundation for mathematics *at all*, we will not explore further the question of what version of this foundation is most satisfactory, or what further assumptions are required to support any particular version.

particular certificates to those premises that are assumed – are made completely explicit.

As detailed above, the constructive logic guarantees that if the input expressions are non-empty then the output expressions are non-empty as well. We don't know in general that arbitrary types are inhabited, but we assume (without justification) that there is at least one mathematical entity, namely the token $*$ of the Unit type 1. There are no other axioms of the basic theory.[37]

The basic rules of the fragment of HoTT that we called HoTT$^-$ are justified as the instantiations of the elementary logical operations such as conjunction and disjunction, all of which can be understood and motivated pre-mathematically. The specific form of the rules comes from the interpretation of logical reasoning as the manipulation of certificates to propositions, which leads (via the BHK interpretation) to the constructive logic that is used in HoTT.

The type former and token constructor for the identity type are easily justified, as outlined in Section 6. In Section 6 we sketch how to derive path induction, the elimination rule for identity types, from pre-mathematical notions. (A more detailed discussion of this is given in our companion paper [21].)

7.5 Methodology

The methodology for the use of HoTT in mathematical reasoning is unchanged from the presentation in the HoTT Book. While the language of HoTT is completely formal, it can also be used as part of informal reasoning (just as set-theoretic reasoning can be used in arguments that are not rendered into completely formal manipulations in first order logic).

In general we proceed by formulating types corresponding to the kinds of mathematical entities under discussion, then from these form types expressing the proposition to be proved and the premises to be assumed. We introduce names for the tokens of the premises we assume given, and then seek to produce a token of the conclusion (or to argue informally that such a token can be constructed).

Note that this methodology does not require that every type involved be completely formalised down to the tiniest detail. We can just as well introduce 'placeholder' types, so long as whatever particular properties we require of them for a given proof can either be derived or are explicitly assumed. So we can still reason about

---

[37] The 'Univalent Foundations' research program introduces the powerful 'univalence axiom', which allows types that are 'equivalent' (in a formally defined sense) to be equated. However, we consider this to be outside the basic language of HoTT, and so beyond the scope of the present paper.

entities whose existence cannot be constructively derived (such as non-principal ultrafilters) by assuming their existence (and any relevant properties) as premises.

8 Possible objections to this account

In this section we address a number of objections that may be raised to the interpretation we set out above.

8.1 What are concepts?

A foundation for mathematics is supposed to clear up the mystery of what mathematical entities are and how we come to know about them, given that (according to many interpretations) they are abstract entities with no causal powers. While the current interpretation does not invoke abstract objects, and so avoids this problem, it may be objected that we have simple traded one mystery for another: namely the question of what concepts are, and what their metaphysical status is.

However, although we do invoke concepts in our interpretation of tokens and types, the features of concepts that we rely upon in Section 5.2 are only those straightforward features that follow from our intuitive understanding of concepts. Thus we do not depend upon any advanced or intricate theory of concepts, and therefore do not need to give a comprehensive detailed account of concepts. Furthermore, concepts are arguably needed for any account of thought more generally and are not an additional requirement for mathematics in particular.

Since concepts play such a central role in this proposed account of HoTT as a foundation for mathematics, some features of the foundation provided by HoTT will depend upon how we think about concepts. For example, as mentioned in Section 7.3, the mind-dependent or mind-independent status of mathematics will depend, on this interpretation, on whether concepts themselves are mind-dependent or mind-independent. Similarly, one might be concerned to explain how different mathematicians appear to be talking about the same domain when (on some accounts) each has his or her own independent concepts.

We acknowledge this dependence but do not take it to be an objection to the account of HoTT as a foundation. Rather, it means that there are a number of interpretational varieties of the foundation available, which append different accounts of concepts and thus give different views on the status of mathematical entities. Moreover, interpreting mathematical talk as talk about concepts integrates problems in philosophy of mathematics into this wider domain. So, for

example, questions about the inter-subjectivity of mathematics are subsumed into questions about the inter-subjectivity of concepts more generally.[38]

8.2 Isn't this just Brouwerian Intuitionism?

By claiming that mathematical language talks of concepts not mathematical objects, and that mathematical reasoning is the manipulation of concepts in the mind using constructive aka 'intuitionistic' logic, have we not just re-invented Brouwerian Intuitionism?

Brouwer's intuitionism positively rejected consideration of mathematical objects, whereas our account is entirely compatible with a Platonist position (since the existence of objects that can be treated by the mathematician entails the existence of concepts under which they can be understood).

Brouwer's position puts the activity of the human mind at the centre, whereas our interpretation is compatible with an understanding of concepts that grants them a mind-independent status. It is compatible with mathematical reasoning being carried out by an automated process in a computer without any human being involved (after the initial setup of the computer program itself, of course). So, for example, a sufficiently large and complex calculation in a computer might make use of types that correspond to concepts that have never been conceived by a human reasoner, and may (for all we know) be too complex for any human mind to conceive.[39]

Brouwer held that "mathematics is an essentially languageless activity of the mind having its origin in the perception of a move of time [which] may described as the falling apart of a life moment into two distinct things, one of which gives way to the other, but is retained by memory" [24, pp. 4-5]. No such hypothesis plays any role in our interpretation of HoTT as a foundation for mathematics, and indeed we give no account of how mathematical concepts originate. Moreover, while we leave it open that one could take concepts to be necessarily pre-linguistic, any concepts that are involved in mathematical activity must, on our account, be given a precise and rigorous definition in the formal language of HoTT.

Finally, Brouwer had truth and falsity of propositions be a temporal matter, with propositions only gaining a truth value when they are first proved or disproved

---

[38] Of course, the answers we give in the case of mathematics might still be specific to that domain. For example, even without giving a full account of how concepts can be shared between subjects we might still be able to give an account of how the concepts of mathematics, being rigorously defined in a formal language, can be shared.

[39] Although this last point would not be relevant if we took the reasoner in the Intuitionism to be the idealised 'creating subject' [23].

and having no truth value at all before then. Our interpretation makes no such claim, and allows that propositions may have truth values – or rather, that types may have tokens – that are simply unknown to us.

8.3 Duplicated objects

Mathematicians standardly consider, for example, the number 2, as a single entity that appears in multiple different domains – the natural numbers, the reals, the complex numbers, and in subsets such as the even numbers, the prime numbers, and the even primes. However, in the formalism of HoTT this cannot be the case, since each token belongs to exactly one type. This appears to require us needlessly to multiply entities, producing 'clones' or 'counterparts' that inhabit different types. This, it might be argued, is inelegant and in conflict with our intuitions.

In response we recall that the tokens of the theory are interpreted not as mathematical objects, but as specific concepts *qua* instances of general concepts. Even if we think that 'the number 2' is a singular entity, it is clear that the *conception* of 2 as a rational number is distinct from the *conception* of 2 as a natural number. This is a natural distinction, not one artificially imposed. Moreover, we have no concept of 'the number 2' except as a member of some number system – we do not conceive of the number 2, separately conceive of the natural numbers, and then determine that the former is a member of the latter.

Rather, it is these concepts – '2 qua natural number', '2 qua rational number', and so on – to which we can plausibly claim to have direct access. We have no concept of 'the number 2 simpliciter', and any putative access to 'the number 2 itself' (as a mathematical object) is notoriously shrouded in mystery.

8.4 Intensionality and *substitution salva veritate*

As mentioned above, our justification of path induction (the elimination rule for identity types) depends upon the principle of *substitution salva veritate*. However, the failure of this principle is a hallmark of 'opaque contexts', namely ones in which co-referring expressions *cannot* be substituted while preserving the truth of a sentence. Such contexts are generated by intensionality. Since the treatment of types in HoTT is intensional, surely *substitution salva veritate* is routinely violated, and so cannot be a defining principle of identity?

For example, in sentences describing the contents of a person's thoughts or beliefs, if that person is unaware of the fact that two names or descriptions refer to the

same entity then substituting one for the other may change the truth of a sentence. For example, the sentence 'S is thinking about the inventor of bifocal glasses' can be true, while the sentence 'S is thinking about the first US Postmaster General' is arguably false, if S does not know that both of these descriptions pick out Benjamin Franklin.

The standard diagnosis of this phenomenon is that in some contexts names and descriptions are implicitly associated with a 'mode of presentation'. When a sentence omits this mode of presentation (as is usually the case) it can be rendered ambiguous, and filling in different modes of presentation can change the truth value. Thus in the case above S is thinking about Benjamin Franklin under the mode of presentation *inventor of bifocal glasses* and not under the mode of presentation *first US Postmaster General*. Thus S's thinking about Franklin can render the first sentence true and the second sentence false, even though the object of S's thought, Benjamin Franklin, is the extension of both descriptions.

Arguably, eliminating ambiguity about intensions solves this problem, and this is the case according to our interpretation of HoTT. The intension of a token is the type that it belongs to – for example, the intension of '2 qua rational number' is *rational number*. Since each token belongs to exactly one type there can be no ambiguity about the intension under which it is considered, and so the intensional treatment of types does not generate an opaque context, and *substitution salva veritate* is not violated.

9   Conclusion

It is shown above that, suitably interpreted, HoTT can be an autonomous foundation for mathematics in the strong sense that involves providing an account of the semantics, metaphysics and epistemology as well as a framework. While the framework in question is that presented in the HoTT Book [1, Chapter 1], we give a new interpretation of the tokens and types of the theory and an account of the metaphysical commitments and the epistemology that goes with it.

Our account of the semantics of type theory via concepts is metaphysically relatively conservative. No special ontology of abstract objects is required for mathematics. If concepts are regarded as abstract objects then either we need them anyway to account for non-mathematical thought, or some reductive or eliminative account of them can be given, in which case it can be applied to mathematical concepts as well. We have no such account in mind, but do not rule one out, and we do not address the ontology of concepts further here.

As far as epistemology goes, the present view requires that concepts can be accessed

by thought, and again there is no special mystery about this in the mathematical case. When it comes to knowledge of mathematical truths, where these are construed as conceptual facts, any philosophy of mathematics that respects mathematical practice must take proof to be the primary means by which mathematical knowledge is established. In HoTT any informal proof can always be translated into a completely formal proof and the logic of the latter is constructive. Thus knowledge of the truth of theorems is reduced to knowledge of the existence of certificates to those theorems, and a proof is a complete and explicit specification of a method to produce such a certificate by the application of simple rules.

Note that none of this rules out an ontology of abstract objects. However, we do not need to posit the latter in order to do mathematics. Our interpretation is compatible with stronger positions that have such an ontology of mathematics, but it does not give any additional support to them. Similarly, the epistemological position proposed here provides a basis for mathematical knowledge via constructive proofs, but does not rule out that it may be had in other ways, such as via proofs in classical logic (or even by some faculty of intellectual perception). For example, a particular proof may assume the applicability of the Law of Excluded Middle or some other additional premise that is not a consequence of the axioms of the theory; such a proof is then conditional on these assumptions, which must be made explicit. Unlike some strong intuitionist positions, the epistemology proposed here does not rule out the use of the Law of Excluded Middle, it simply requires that such uses be explicitly flagged up as assumptions that must either be justified or assumed as premises of the proof. In short, the metaphysical and epistemological positions proposed here do not restrict the ontological or mathematical assertions that can be made but they do limit the assertions for which we can claim justification without further premises.

9.1   Advantages of this foundation

HoTT understood as above or as in the HoTT book has many advantages as a foundation for mathematics. Type theories help us avoid paradox and errors, and the constructive framework of HoTT facilitates automatic proof verification. The intensionality of the theory allows for the finest of distinctions to be made. In terms of flexibility and the naturalness with which structures can be represented it is like category theory rather than set theory, but HoTT also unifies logic with mathematics via the Curry-Howard correspondence. Predicativity is optional as is function extensionality. As Awodey [6] argues, HoTT is also a promising framework for structuralism in the philosophy of mathematics.

In mathematical practice every well-formulated definition is the careful expression

of some concept, usually formed by the logical combination of simpler concepts. We have direct access to concepts, as well as to pictures and symbols, in thinking about mathematics. Mathematical concepts are often taken to be about mathematical objects, but (even setting aside problems of the nature of these abstract objects and our access to them) the relationship between concepts and objects is not straightforward. It can be a matter of great mathematical effort to determine that a given mathematical concept denotes a particular object, and it is often a substantial mathematical achievement to show that two concepts are extensionally equal, and in a particular case this may be suspected and hypothesised for a long time without being proved.[40]

Hence, it is plausible that concepts and the definitions that directly correspond to them should be taken as the primary entities of mathematics, while the objects that they (appear to) denote should be taken as at best secondary and at worse redundant. It follows that our foundational language for mathematics should be understood as picking out concepts rather than objects. This implies that the language should be *intensional*, maintaining a distinction between conceptually distinct descriptions, even when they have the same extension. This is in contrast to the extensional foundations given by ZFC set theory, for example, which takes the objects (i.e. sets) as primary and therefore collapses together co-extensional definitions.

Our concepts interpretation of HoTT motivates the intensionality of the theory, coheres with its constructive nature, explains how identity types are thought of and used, and is independent of but compatible with the homotopy interpretation. Whether or not these advantages are sufficient to show that HoTT so construed or otherwise should be adopted as a foundation for mathematics is a matter for future work.

10   Acknowledgments

---

[40] Consider, for example, the concepts 'positive natural number less than 3' and 'natural number n for which the Fermat equation $x^n + y^n = z^n$ has at least one solution in the positive integers'.

References

[1] The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics.* http://homotopytypetheory.org/book, Institute for Advanced Study, 2013.

[2] Øystein Linnebo and Richard Pettigrew. Category Theory as an Autonomous Foundation. *Philosophia Mathematica*, 19(3):227–254, 2011.

[3] Authors. A Primer on Homotopy Type Theory. 2014.

[4] Paul Benacerraf and Hilary Putnam, editors. *Philosophy of mathematics: selected readings.* Cambridge University Press, 2nd edition, 1983.

[5] Hilary Putnam. Mathematics without foundations. In Paul Benacerraf and Hilary Putnam, editors, *Philosophy of mathematics: selected readings.* Cambridge University Press, 2nd edition, 1983.

[6] Steve Awodey. Structure in mathematics and logic: a categorical perspective. *Philosophia Mathematica*, 4(3):209–237, 1996.

[7] Steve Awodey. Structuralism, invariance, and univalence. *Philosophia Mathematica*, 22(1):1–11, 2014.

[8] Stewart Shapiro. *Foundations without Foundationalism: A case for second-order logic.* Oxford University Press, 1991.

[9] Stewart Shapiro, editor. *The Oxford Handbook Handbook of Philosophy of Mathematics and Logic.* Oxford University Press, 2005.

[10] John P. Mayberry. What is Required of a Foundation for Mathematics? *Philosophia Mathematica*, 2(1):16–35, 1994.

[11] Hartry H. Field. *Science Without Numbers: A Defence of Nominalism.* Princeton University Press, 1980.

[12] M. Colyvan. *The Indispensability of Mathematics.* Oxford University Press, 2001.

[13] Per Martin-Löf. An intuitionistic theory of types: predicative part. In H. E. Rose and J. C. Shepherdson, editors, *Logic Colloquium (1973)*, pages 73–118. North-Holland, 1974.

[14] A. Church. An unsolvable problem of elementary number theory. *American Journal of Mathematics*, 58(2):345–363, 1936.

[15] Per Martin-Löf. On the Meanings of the Logical Constants and the Justifications of the Logical Laws. *Nordic Journal of Philosophical Logic*, 1(1):11–60, 1996.

[16] A. S. Troelstra. History of constructivism in the 20th century. In *Set Theory, Arithmetic, and Foundations of Mathematics*. Cambridge University Press, 2011.

[17] John L. Bell. *A Primer of Infinitesimal Analysis*. Cambridge University Press, 2008.

[18] Steve Awodey and Michael Warren. Homotopy theoretic models of identity types. *Mathematical Proceedings of the Cambridge Philosophical Society*, 146(1), 2009.

[19] Y. Aharonov and D. Bohm. Significance of electromagnetic potentials in the quantum theory. *Phys. Rev.*, 115:485–491, Aug 1959.

[20] Richard Healey. Nonlocality and the Aharonov-Bohm Effect. *Philosophy of Science*, 64(1):18–41, 1997.

[21] Authors. Identity in Homotopy Type Theory, Part I: The Justification of Path Induction. 2014.

[22] Daniel Isaacson. Mathematical intuition and objectivity. In Alexander George, editor, *Mathematics and Mind*, pages 118–140. Oxford University Press, 1994.

[23] L. E. J. Brouwer. Essentially negative properties. *Indagationes Mathematicae*, 10:322–323, 1948.

[24] L. E. J. Brouwer. *Brouwer's Cambridge lectures on Intuitionism*. Cambridge University Press, 1981.