CrossMark

Identity and intensionality in Univalent Foundations and philosophy

Staffan Angere[1] ⬤

Abstract The Univalent Foundations project constitutes what is arguably the most serious challenge to set-theoretic foundations of mathematics since intuitionism. Like intuitionism, it differs both in its philosophical motivations and its mathematical-logical apparatus. In this paper we will focus on one such difference: Univalent Foundations' reliance on an intensional rather than extensional logic, through its use of *intensional Martin-Löf type theory*. To this, UF adds what may be regarded as certain extensionality principles, although it is not immediately clear how these principles are to be interpreted philosophically. In fact, this framework gives an interesting example of a kind of border case between intensional and extensional mathematics. Our main purpose will be the philosophical investigation of this system, and the relation of the concepts of intensionality it satisfies to more traditional philosophical or logical concepts such as those of Carnap and Quine.

1 The Univalent Foundations project

The Univalent Foundations (UF) project is an approach to the foundations of mathematics currently under rapid development. It entails a major change both in philosophical motivation and form, ultimately taking its main inspiration from geometry rather than grammar or traditional logic. Michael Harris—himself working in the rather different field of number theory, and not being a fan of "Foundations" for math-

✉ Staffan Angere
  staffan.angere@fil.lu.se

1 Department of Philosophy, University of Lund, Lund, Sweden

ematics, despite his otherwise interest in mathematical–philosophical questions—still makes bold claims about its philosophical relevance:

> It's impossible to overstate the consequences for philosophy, especially the philosophy of Mathematics, if Voevodsky's proposed new Foundations were adopted. By replacing the principle of identity by a more flexible account modeled on space, the new approach poses a clear challenge, on which I cannot elaborate here, to the philosophy underlying "identity politics"; it also undermines the case for analytic philosophy to seek guidance in the metaphysics of set theory, as in W.V.O Quine's slogan "to be is to be the value of a variable." (Harris 2015, p. 219)

This paper constitutes a first attempt at a philosophical analysis of some of the concepts underlying this framework, and of their possible uses in philosophy. As such, it is to be seen as an orienting preliminary work, outlining a number of problems and approaches rather than an attempt at coming to any firm conclusions. It is my hope, however, that the potential importance Univalent Foundations, and the comparably small amount of work on its philosophy which has so far been done, still warrant its publication.

An example of the difference in approach between the usual foundations and UF is that the latter is based on type theory—more precisely Martin-Löf's intensional intuitionistic type theory (Martin-Löf 1998)—rather than predicate logic and set theory. Martin-Löf (M-L) type theory is a very versatile framework, capable of unifying logic, computation and mathematics. UF relies on a certain toplogical interpretation of this type theory, which is known as *Homotopy Type Theory* (HoTT).

While predicate logic makes a distinction between terms and formulae, with the former combined by functions and the latter by connectives and quantifiers, M-L type theory's main distinction is between terms and *types*, which on the traditional reading are a version of sets. The connection to assertion is made through *judgments*, which are either of the form

$$\Gamma \vdash a \,:\, A$$

which means that a is of the type A, given a *context* Γ, which is a sequence of assignments of types to terms, or

$$\Gamma \vdash a \equiv b \,:\, X$$

which means that a and b are definitionally equal objects of type X in the context Γ.[1] As we will see later, however, the meaning of this "definitionally" is not entirely trivial to spell out, so we will usually employ the alternative term "judgmentally equal".

---

[1] This is not strictly true, as HoTT also contains a judgment type expressing that a sequence of judgments is a valid context, i.e. can appear to the left of the turnstile. M-L type theory in its standard form also makes a difference between identity judgments about objects and identity judgments about types, but as HoTT assumes universes containing all lower-level types, the latter are expressible through the kind of identity judgment described here as well.

Unlike sentences in FOL, judgments cannot be combined in any way. They appear as premises or conclusions in inference rules.

Type theories usually contain many rules for introducing new types given others. Of particular significance are the *function types* $A \rightarrow B$, whose elements are the functions from A to B. A function, being an element of a type, is thus also a term. Another rule, of particular importance to HoTT, is one allowing the introduction of *universe* types. This means that we assume the existence of a sequence of primitive types $\mathcal{U}_0, \mathcal{U}_1, \mathcal{U}_2, \ldots$, such that $\mathcal{U}_0$ is the type of all the "usual" types, $\mathcal{U}_1$ all those of $\mathcal{U}_0$ as well as those (such as $\mathcal{U}_0$ itself) which are too big to be in $\mathcal{U}_0$, etc. We will usually follow custom in not writing out the indices for these, but instead rely on typical ambiguity to assign them on a case-by-case basis.

These are properties that are common to several contemporary type theories. The main addition made by M-L type theory is that of *dependent* types and functions. A dependent type (or *type family*) P on the type X is an assignment of a type $P(x)$ to each element $x : X$, i.e. a function $P : X \rightarrow \mathcal{U}$. A *dependent function type*, written as

$$\prod_{x : X} P(x)$$

is the type of all functions from the type X to type $P(x)$, for any $x : X$, or in other words the functions that assign some element of the dependent type $P(x)$ to each element of x. The elements of such a type are referred to as *dependent functions*. To illustrate, we have drawn one such function $f : \prod_{x : X} P(x)$ in Fig. 1.

Fig. 1 A dependent function $f : \prod_{x : X} P(x)$
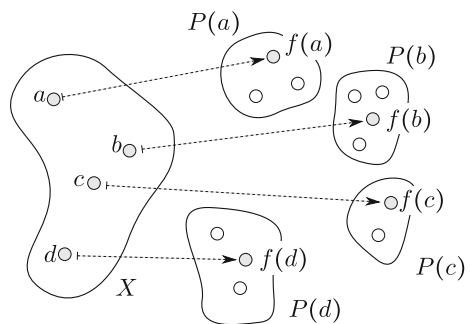
In terms of set theory, a dependent type on a set X would be a function P from X to the class V of all sets, i.e. an assignment of a set to each element of C.[2] A dependent function would then be a set-theoretic function $f : X \to V$ such that

$$f(x) \in P(x)$$

for all $x \in X$, i.e. a choice function for the set $P[X]$. A dependent function type $f : \prod_{x \,:\, X} P(x)$ is non-empty iff the type family P it is defined on assigns a non-empty type to every element of X.

Another example of a dependent type is the *dependent pair type*

$$\sum_{x \,:\, X} P(x)$$

In set theory, a dependent pair corresponds to an ordered pair $\langle x, y \rangle$ where $x \in X$ and $y \in P(x)$, and the dependent pair type to a disjoint union. In the logical (proof-theoretical) interpretation we will describe later in this section, dependent pairs are also used to model the existential quantifier.

In many type theories, it is common to have a specific type Ω of *truth values* (or propositions, as they are sometimes called), with at least two elements $\top$ and $\bot$ representing *truth* and *falsity*. A predicate P on the elements of the type X is then interpretable as a function $f_P : X \to \Omega$ such that $f_P(x) = \top$ iff $P(x)$ holds, for $x \in X$. M-L type theory, in contrast, does not assume such a type, but instead relies on the Curry–Howard correspondence, or *formulae-as-types* interpretation (Howard 1980), according to which *any* type can be interpreted as a proposition. One way of doing this is to interpret a proposition X as the type of *proofs that X*. Like in the Brouwer–Heyting–Kolmogorov interpretation of intuitionistic logic, we can then say that a proposition is true iff it has some proof, i.e. is non-empty as a type.[3]

To illustrate, it is useful to consider another common type-forming operator, namely the *product type $A \times B$* of A and B, which is definable in terms of the dependent function type. In the set-theoretic interpretation, such a type is exactly the Cartesian product of the sets A and B. But just as such a product equals $\emptyset$ iff A or B are empty, we also have that $A \times B = \mathbf{0}$, where $\mathbf{0}$ is the empty type, iff either A or B are empty. As an empty type is a false proposition, we can therefore see that $A \times B$, when interpreted this way,

---

[2] There is a case to be made for introducing and describing HoTT without giving translations to set-theoretic vocabulary, since HoTT can work as a foundation of mathematics independent of, say, ZFC. However, since the main purpose of these first sections are to introduce HoTT to philosophers of mathematics and logic, and set theory is what many such philosophers—at least in the analytic school—are best accustomed to thinking in, allowing ourselves the use of such a crutch may help. We just have to remember that the set theoretical constructions we present are pedagogical tools rather than parts of any conceptual underpinnings. Since set theory is not presupposed, we are usually not going to be very careful about e.g. cardinality restrictions that might be important if one wants to give formal models of HoTT in extensions of ZFC.

[3] In HoTT, it is sometimes assumed that only certain types of sufficiently simple structure are propositions, in order to make sure that the laws governing them correspond exactly to the ones of intuitionistic logic. Such propositions are referred to as *mere* propositions (Univalent Foundations Program 2013, pp. 111–113). In this paper, we will not make this distinction, and thus simply use 'proposition' as a synonym for 'type'.

is true in exactly the same circumstances as the conjunction of the propositions A and B.[4]

The Curry–Howard correspondence can also be extended in order to deal with correlates of open formulae rather than sentences. A *property P*, which can be instantiated by the elements of the type X, can be interpreted as a dependent type $P : X \to \mathcal{U}$ which assigns the type $P(x)$ of all proofs that x has property P to each element $x : X$. Again, if we take a proposition to be true iff it has a proof, $P(x)$ will be a non-empty type iff x has the property P. Since a dependent function type $\prod_{x : X} P(x)$ is non-empty iff $P(x)$ is non-empty for each $x : X$, it follows that $\prod_{x : X} P(x)$ works as the proposition *for all x of type X, $P(x)$*.

There is one consequence of this approach which is particularly noteworthy: while, in classical logic, $P(x)$ is simply true or false of an object x, and in intuitionistic logic, $P(x)$ is more generally assigned an element in a Heyting algebra, in intensional M-L type theory we always need a *specific* element of $P(x)$ to express $P(x)$'s truth, and thus more information. Although intuitionistic logic is also sometimes characterised as requiring specific proofs, e.g. in taking $\varphi \vee \psi$ to be true only if a proof exists either for φ or for ψ, it does not keep track of the identities of such proofs except insofar as they count as proofs of the same sentences. Intensional M-L type theory, in contrast, provides the possibility of distinguishing between individual proofs even of the same sentences. A type theory of this kind is called *proof relevant* (Univalent Foundations Program 2013, p. 8).

2 The varieties of identity

It seems to be a common assumption in analytic (but of course not continental) philosophy that identity is a fairly simple concept. A forerunner here is Frege, who in the afterword to the second volume of the *Grundgesetze*—the one in which he reacts to Russell's paradox—argues against taking classes to be improper objects since this would require them to have their own notion of identity. Identity, he says, "is a relation given in so determinate a way that it is inconceivable that different kinds of it could occur" (2013, p. 254). More recently, we find David Lewis expressing quite similar feelings:

> Identity is utterly simple and unproblematic. Everything is identical to itself; nothing is ever identical to anything except itself. There is never any problem about what makes something identical to itself; nothing can ever fail to be. (1986, pp. 192–193)

Making a distinction with the notion of *reference*, McGinn rather matter-of-factly writes in his book *Logical Properties* that "one could not write a good book entitled *The Varieties of Identity*" (2000, p. 1).

---

[4] There is a "deeper" structural explanation of this which does not depend on a set-theoretic interpretation. Both the product of types in a category of types and functions between them, and the conjunction of sentences in a language $\mathcal{L}$ seen as a category with an single arrow $p \to q$ whenever q follows from p, are instances of the same category-theoretical construction, namely that of a (category-theoretical) product of objects.

But such a view can only be defended if one has already decided that the concept of identity of, say, classical predicate logic, is the correct one, in which case the simplicity and lack of problems that Lewis alludes to naturally spring from the fact that a decision has already been made, even if unconsciously. As soon as we leave classical logic's safe confines, the conceptual landscape opes up. In the classical presentation of his type theory Martin-Löf (1984) distinguishes four forms of identity[5]:

(*i*) The *definitional equality* $a \equiv_{df.} b$. This is "the equivalence relation generated by abbreviatory definitions, changes of bound variables and the principle of substituting equals for equals" (Martin-Löf 1984, p. 31). It is a purely syntactic notion, and is therefore not strictly the same as the identity judgment $\vdash a \equiv b : X$, which—to start with—also makes reference to a specific type. In Martin-Löf (1984), it is the only form of identity that the author holds to be intensional.

(*ii*) The *equality of elements* $a \equiv b : X$, where X is a type of which a and b are elements. This is perhaps the version most like classical identity in that it allows replacing equals for equals arbitrarily, subject only to type restrictions. This is what we have referred to as the judgmental equality.

(*iii*) The *equality of types* $X \equiv Y$, which consists in them having judgmentally equal elements.

(*iv*) The *identity type family* $\mathsf{Id}_C(x, y)$ through which every pair of elements a, b of C is assigned a specific type $\mathsf{Id}_C(a, b)$. Using the Curry–Howard correspondence, the elements of this type are interpretable as *proofs of the proposition $a = b$*, and this form of identity is therefore also referred to as *propositional*. Furthermore, if $a : C$, there is a postulated element $\mathsf{refl}_C(a)$ of $\mathsf{Id}_C(a, a)$ called a's *reflexivity proof*, which corresponds to a canonical proof that $a = a$; e.g. the one obtained by a single application of equality introduction.

In *extensional* M-L type theory the existence of an element of $\mathsf{Id}_C(a, b)$ is postulated to imply that $a \equiv b : C$; this is sometimes referred to as the *equality reflection rule*. If added as an extra postulate to non-extensional M-L type theory the equality reflection rule entails that if $p : \mathsf{Id}_C(a, b)$, then $p \equiv \mathsf{refl}_C(a) : \mathsf{Id}_C(a, a)$, so the reflexivity proof of a (or, equivalently, b) is the only proof of the propositional equality of a and b that we have (Univalent Foundations Program 2013, p. 104). We therefore say that extensional type theory has a *uniqueness of identities*.

In the *intensional* version [as presented in e.g. Martin-Löf (1998)] the non-emptiness of $\mathsf{Id}_C(a, b)$ is strictly weaker than $a \equiv b \in C$, and instead motivates a type of induction principle. Very informally, this principle allows us to transfer the structure of one type which is a value of a dependent function $\prod_{x : X} P(x)$ for an argument a, to the type which is a value of any argument b which is propositionally equal to a. Formally it is rather complex, but for our present cases it suffices to consider a consequence of it called the *indiscernibility of identicals* (Univalent Foundations Program 2013, pp. 44–45):[6]

---

[5] We will follow standard mathematical practice in treating the words 'equal' and 'identical' as interchangeable. Differences in which of these words we use will therefore merely be due to stylistic factors.

[6] Although the name "indiscernibility of identicals" has been used with various meanings in philosophy, we will from now on reserve it for this type theoretic principle. For a reasonably intuitive defense of the full equality elimination rule, see Martin-Löf (1975, p. 81).

$$\frac{\Gamma, A \,:\, \mathcal{U}, \, x \,:\, A, \, y \,:\, A, \, p \,:\, \mathsf{Id}_A(x, y) \vdash C \,:\, A \to \mathcal{U}}{\Gamma \vdash \mathsf{ind}_=(x, y, p) \,:\, C(x) \to C(y)}$$

Here, $\mathsf{ind}_=$ is a specific function that assigns each identity between objects x, y of type A a function from $C(x)$ to $C(y)$. Remembering the interpretation of type families on A as properties of A's elements (i.e. $C(x)$ contains proofs that x has property C) we can see that this rule guarantees that if we have some proof that $C(x)$ holds as well as a proof $p \,:\, \mathsf{Id}_A(x, y)$, then we can construct a proof that $C(y)$ as well by applying $\mathsf{ind}_=$. Expressed in more traditional terms: if x has a property and x is identical to y, then y has that property as well.

HoTT assumes only the second and the fourth of Martin-Löf's forms of identity as primitive, although arguably the second form also plays parts of the roles of the first and the third.

Since the existence of universes means that types and not only terms can be assigned types, Martin-Löf's equality of types $X \equiv Y$ is representable as $X \equiv Y \,:\, \mathcal{U}$. As HoTT also assumes inference rules that entail that $a \,:\, B$ follows from $a \,:\, A$ and $A \equiv B \,:\, \mathcal{U}$, the fact that A and B have judgmentally equal elements is still a consequence of their judgmental equality.

The question of *propositional* equality between types, i.e. their identity type families, is significantly more interesting. Any well-defined type X should specify rules for constructing an element of $\mathsf{Id}_X(a, b)$. Under the Curry–Howard correspondence, this means that the introduction of X as a type also should make it clear what a proof of the propositional identity of elements of X. This is easy in an extensional type theory: since identities are unique, identity proofs are always identical to the reflexivity proof, so any identity proof is really a reflexivity proof, even if it may not look like one. In intensional type theory this approach is not possible, and the question arises as to what equality between its types really means. As we shall see in Sect. 4, it is here that univalent foundations makes one of its most significant contributions.

3 Some basic concepts of homotopy theory

In order to describe HoTT's approach to equality it is necessary to go through a few mathematical preliminaries. We will try to keep these to a minimum in order to simplify as much as possible; for example, the admittedly crucial notion of *fibration* will hardly be discussed at all. We assume at the outset that the reader is familiar with general point-set topology, i.e. concepts such as topological spaces, continuous functions, homeomorphisms etc.

Several semantics exist for type theory, such as ones based on category theory or theories of computation. HoTT distinguishes itself by instead supplying a *geometric* (or more precisely, topological) semantics. The main notion used is *homotopy*. This
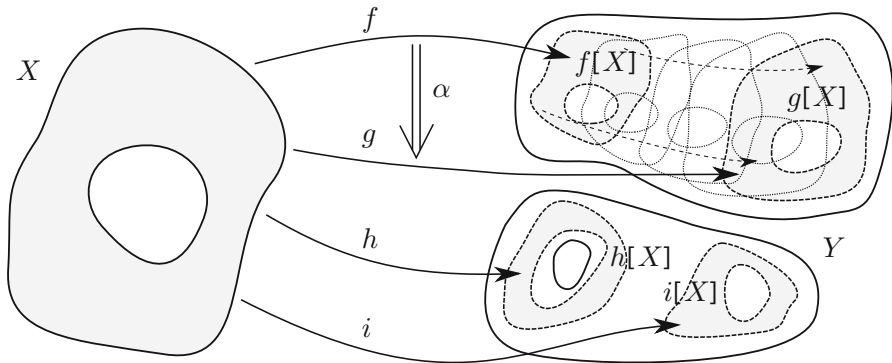
Fig. 2 Four continuous functions and a homotopy

is a central concept in algebraic topology, which the study the relationships between topological and algebraic structures—in particular groups and groupoids.[7]

While the idea of one space being continuously transformed into another is central to all topology, homotopy theory adds consideration of how such transformations may themselves continuously transform into one another. This makes it a kind of "higher-order" version of topology. The central concept is that of a homotopy between two continuous functions $f, g : X \to Y$, where X and Y are topological spaces. Formally, a homotopy α from f to g is represented as a function $\alpha : I \times X \to Y$, where I is the unit interval [0, 1] with its usual (interval) topology, such that the equations

$$\alpha(0, x) = f(x)$$
$$\alpha(1, x) = g(x)$$

hold for all $x \in X$. For an example, consider the four continuous mappings f, g, h, i of the space X into the two-component space Y depicted in Fig. 2.

Here α is one of many homotopies from f to g. On the other hand, there are no homotopies from g to h, because a necessary condition for a homotopy $g \to h$ to exist is that every point $g(x)$ is path-connected to its corresponding point $h(x)$, and no such paths can exist since $g[X]$ and $h[X]$ are in different components. That this condition is not *sufficient* can be seen by considering h and i: although we can draw a path from every point $h(x)$ to its correlate $i(x)$, we cannot do so in a *continuous* way. Graphically, the fact that h maps X around a hole in Y means that $h[X]$ cannot be transformed into $i[X]$ without tearing it apart, i.e. introducing discontinuities. This illustrates the fact that the existence or non-existence of homotopies is a global phenomenon, which is generally not possible to decide just by looking at individual points of a space one by one.

---

[7] A *groupoid* is a category in which all arrows are invertible. It can also be seen as a form of group where not all elements can be multiplied with one another. See Brown (2006) for an insightful exposition of groupoids as applied to homotopy theory.
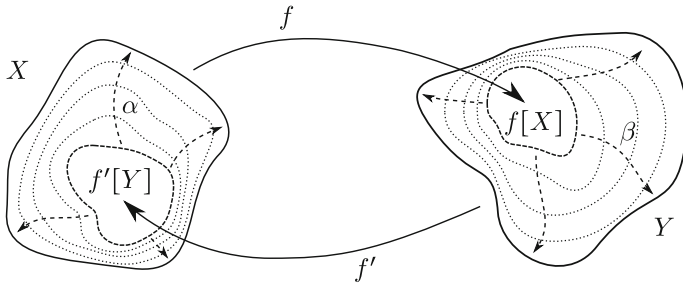
Fig. 3 A homotopy equivalence

We write $f \sim g$ and call f and g *homotopic* if there is some homotopy from f to g. It is easy to see that whenever there is a homotopy $\alpha : f \to g$, we can get a homotopy $\alpha' : g \to f$ going the other way by setting

$$\alpha'(t, x) = \alpha(1 - t, x)$$

which simply reverses the direction. Likewise, given two homotopies $\alpha : f \to g$ and $\beta : g \to h$, we can define a composite $\beta \circ \alpha : f \to h$ as

$$(\beta \circ \alpha)(t, x) = \begin{cases} \alpha(2t, x) & \text{if } t < {}^{1}\!/_{2} \\ \beta(2t - 1, x) & \text{if } t \geqslant {}^{1}\!/_{2} \end{cases}$$

Since the constant function $1_f : f \to f$ defined as $f(t, x) = x$ is also a homotopy, it follows that $\sim$ is an equivalence relation. *Homotopy theory* uses this equivalence to disregard the difference between homotopic functions and thereby obtain more powerful methods for proving theorems about spaces. While general topology does not differentiate between spaces that are stable under continuous transformation, homotopy theory does not differentiate between *functions* that are continuously deformable into one another either.

Formally, let **hTop** be the category that has all toplogical spaces as objects and all homotopy classes of continuous functions between them. Thus **hTop** has the same objects as the category **Top** of topological spaces and continuous functions, but the arrows of **hTop** are equivalence classes of those of **Top**. As is well known, the isomorphisms of **Top** are precisely the homeomorphic functions. In **hTop**, the isomorphisms (i.e. those morphisms that have an inverse) come out as equivalence classes of so-called *homotopy equivalences*.

A continuous function $f : X \to Y$ is a homotopy equivalence iff it has a *homotopy inverse*: a function f' such that $f' \circ f \sim 1_X$ and $f \circ f' \sim 1_Y$, where 1_X and 1_Y are the identity functions on X and Y, respectively.[8] See Fig. 3 for an example.

---

[8] It is often useful to define the homotopy equivalence more explicitly as a quadruple f, f', α, β where $\alpha : f' \circ f \to 1_X$ and $\beta : 1_Y \to f \circ f'$. Since every function that has a homotopy inverse usually has more than one, and since the existence of a homotopy between functions usually implies the existence of many others, this second form gives more information than the first one. Whether this extra information is relevant or not depends on the context.

Fig. 4 A path homotopy
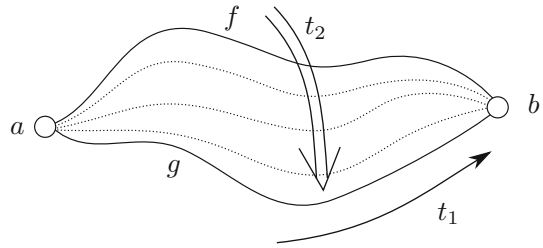
Any homeomorphism is a homotopy equivalence, but the converse does not hold: a disc is homotopy equivalent to a point, but since they have different cardinalities, they cannot be homeomorphic. Furthermore, the n-cube is homotopy equivalent to the $n + 1$-cube, but they are not homeomorphic because they have different topological dimensions. The class of all spaces homotopy equivalent to a space X is called its *homotopy type*.[9] Spaces that are homotopy equivalent to a point are called *contractible*.

From the fact that homotopies are themselves continuous functions it follows that their construction is iterable so that there can be homotopies between homotopies. For example, a homotopy γ between the homotopies α and β above is a continuous function $\gamma : I^2 \times X \to Y$ such that

$$\gamma(0, t, x) = \alpha(t, x)$$
$$\gamma(1, t, x) = \beta(t, x)$$

where $t \in I$ and $x \in X$. A *path* in X with endpoints $a, b \in X$ is a continuous function $f : I \to X$ for which $f(0) = a$ and $f(1) = b$, or equivalently, a homotopy between two functions $f, g : 1 \to X$, where 1 is a one-point topological space. *Path homotopies* are homotopies between paths, homotopies between such homotopies, etc. See Fig. 4, where f, g are functions of t_1 and the homotopy a function of t_1 and t_2.

Since a path is also a homotopy between functions $f, g : 1 \to X$, we can compose them like we compose all homotopies. While this does not give a composition that is associative (we generally have that $h \circ (g \circ f) \neq (h \circ g) \circ f$), replacing the paths from a to b by their homotopy classes restores associativity, and also ensures that all path-classes have inverses. This means that the homotopy classes of paths between the points of a topological space constitute a groupoid with the points as objects and the classes of paths as arrows. If we consider *loops*, i.e. paths $h : a \to a$ whose endpoints and starting points coincide, their equivalence classes make up a group. For a path connected space (one in which there is some path between each pair of points), every choice of point a—or *basepoint*—leads to an isomorphic group, which is called the space's *fundamental group*.

---

[9] We have followed the usual mathematical practice in defining a homotopy type as a class of spaces. However, this is not the only way to do it: a different method, which may be more in line with univalent foundations as e.g. Voevodsky (2010) envisions them is to define homotopy types directly via Kan complexes. For a philosophical discussion on the metaphysical and epistemological status of homotopy types see Marquis (2013).

Fig. 5 A non-trivial path
$f : a \to a$
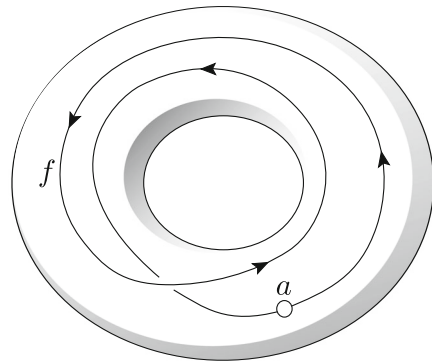
Some spaces, such as the disc and the point, are *trivial* in the sense that their fundamental group is simply the one-element group. An example of a space with a *non-trivial* group of path classes is the solid torus of Fig. 5. Here, there is a specific homotopy class of paths for each number of times the path goes around the central hole. No two of these are identical since the hole in the middle makes it impossible to contract any of the loops to a point. Furthermore, the direction that we go around in matters. In algebraic topology this is expressed by saying that the fundamental group of the solid torus is $\langle \mathbb{Z}, + \rangle$. In other words, the homotopy class of any path in the solid torus is determined by a whole number, and composing such paths gives a number which is the sum of the numbers of the paths composed.

As we noted, the construction of homotopies—and therefore also path homotopies—is iterable. This entails that a topological space does not only have groups of paths, but also groups of higher-dimensional objects. The most important of these are the spheres of various dimensions, although we will not here go into how such spheres are composed with one another. Each dimension n of such spheres gives rise to a group called the *nth homotopy group*. These groups are used in order to model iterated identities in HoTT, i.e. propositional identities that hold between identity proofs.

Apart from homotopies HoTT also relies on another important topological concept, namely that of *fibration*. Intuitively, a fibration is a continuous assignment of a topological space called a *fibre* to each point of another space called the *base space*, which satisfies a so-called *lifting property* that allows topological structures (such as paths) on the base space to be mapped to the space consisting of all the fibres. The conditions on which such assignments count as fibrations ensure that all fibres assigned to the same path-connected component of the base space are homotopy equivalent in a continuous way. Usually, a fibration is expressed as a continuous function $\pi : E \to B$, where B is the base space, E is the union of all the fibres and is called the *entire* space, and π is called the *projection* of E onto B. The fibres are definable as the inverse images of points of B under π, i.e. the fibre on point x is the space $\pi^{-1}[x]$.

A very simple (or literally, *trivial*) example of a fibration is the space $I \times I$, i.e. the 2-dimensional square with side 1, with π taken to be the orthogoal projection onto the first coordinate. As a fibration it can be seen as an assignment of a copy of I to each
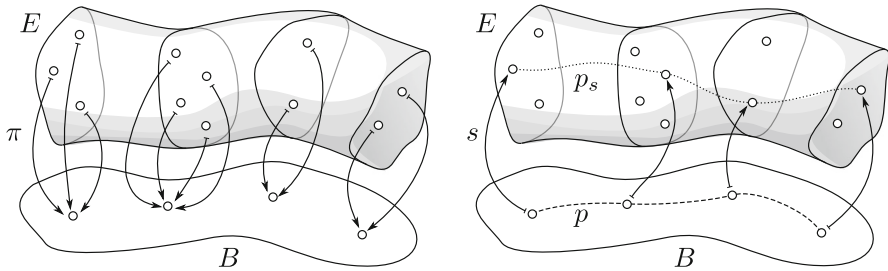
Fig. 6 A fibration $\pi : E \to B$ and a section s with a lifted path p

point in I. Interesting fibrations, however, tend to have both base spaces and fibres with more complex topological structure.

Since a fibration $\pi : E \to B$ associates a space with each point of B we can define functions $s : B \to E$ such that $s(x) \in \pi^{-1}[x]$, i.e. assignments of elements of a point's fibre to that point. Such a function is known as a *section* of the fibration. The lifting property of fibrations ensures that the application of a section s to a path $p : a \to b$ in B results in a path $p_s : s(a) \to s(b)$, and furthermore that this path is unique up to homotopy. Figure 6 gives an example.

A section in our trivial example $\pi : I \times I \to I$ is simply a continuous graph in the usual sense, defined on and taking values in the interval $[0, 1]$. Therefore a path in the x-coordinate I is a function $p : I \to I$, and such a path naturally gives rise to a path taking values in the section.

Fibrations are used to model dependent types in HoTT. This is of special importance for us, since the propositional identity type family $\mathsf{Id}_A(x, y)$ of a type A is an example of such a dependent type, assigning a type to each pair of elements of A. Unfortunately, time and space restrictions mean that we will not be able to delve any deeper into their definition. In what follows, we will attempt to go around these limitations by mainly focusing on identity types of specific pairs of points, rather than the more general type families.[10]

4 Univalence and the semantics of homotopy type theory

Homotopies are, as we mentioned, central to the semantics of HoTT.[11] While the domain of a model of classical first-order logic is a set, terms refer to elements of that set, and relations and functions have subsets of Cartesian powers of the domain as their extension, the semantics of HoTT is thoroughly topological. Intuitively and informally, we may to a first approximation assume the following [cf. Univalent Foundations Program (2013), p. 11]:

---

[10] A classic treatment of the concepts of this section from a point-set topological perspective is given in Hatcher (2006).

[11] More or less everything in this section is based on the author's present understanding of the Univalent Foundations Project's book Univalent Foundations Project's book (2013). Any possible misunderstandings are purely due to him.

(*i*) A type is a topological space, and a term of a type is a point in that space. Thus $a : X$ is true iff a is a point in X, and $a \equiv b : X$ is true iff a and b are the same point of X.

(*ii*) A function $f : A \to B$ is a continuous function from A to B; thus all functions that can be described in HoTT are continuous.

(*iii*) A dependent type $F : A \to \mathcal{U}$ is a fibration. The space corresponding to the type $F(x)$ is the fibre on x.

(*iv*) A dependent function $f : \prod_{x : X} F(x)$ is a section of the fibration F.

(*v*) The elements of an identity type $\mathsf{Id}_X(a, b)$ are the paths from a to b in the space X.

(*vi*) Identities between identities are homotopies between these paths, and the constant path on a is the reflexivity proof $\mathsf{refl}_X(a)$.

(*vii*) When the *univalence* axiom, one of the main innovations of UF, is added, identities between types become homotopy equivalences between their corresponding spaces.

Since we will avoid formal, rigorous derivations in what follows, we will mainly use informal characterisations like this one of the semantics of HoTT as a basis for our arguments. This gives some advantages in terms of keeping the discussion visual—which is one of the things that HoTT, much more than set theory, makes possible—but of course has the disadvantage that what we will not really create proofs proper. Still, I believe that for our philosophical purposes, an approach like this not only acceptable but also in some ways preferable.

When we need to go beyond geometric intuition (e.g. to actually prove things, or when the spaces involved become too hard to visualize) there are several directions open. To begin with, HoTT is being developed in formal proof assistants such as Coq and Agda, which means that the highest possible degree of rigour is available when needed. For more semantically oriented approaches, one can follow Awodey and Warren (2009) and employ Quillen's (1967) notion of *model category* in order to equip the types with a well-defined topological structure. Furthermore, HoTT also has a second, category-theoretical semantics, according to which types are modeled by higher groupoids.[12] Because of its greater intuitive content we have chosen to focus on the topological semantics here, even if the category-theoretical version often is useful as a complement.

Since a space's homotopies between paths, its homotopies between homotopies, etc. can have complex group structures of their own, identities holding between elements of other identity types can have such structure as well. Temporarily suppressing the types of Id and refl and relying on context to sort them out, we call the type

$$\mathsf{Id}(\mathsf{refl}^{n-1}(a), \mathsf{refl}^{n-1}(a))$$

---

[12] While the general concept of *higher category* is still in development, that of a *higher groupoid* and its slight generalization, the $(\infty, 1)$-category, is starting to crystallise somewhat. For one of the more well-known approaches, which is also rather topological in spirit, see Lurie (2009).

the *n-iterated identity type* on the point a, where the exponent $n - 1$ denotes repeated application.[13] Thus, as expected, the usual identity type is the 1-iterated one.

As all types—including the identity types—are spaces, even the types of identities holding between an element and itself can be complex, so the structure of an n-iterated identity type can be highly nontrivial. In extensional type theory, the uniqueness of identities removes all such complexity. In intensional type theory, the *univalence axiom* helps manage it by allowing the treatment of some of the elements of these types as the same, much as going to homotopy classes of paths in homotopy theory allows the treatment of a multitude of paths as identical. In both cases, simplicity is increased without losing the more important features of the structures in question.

With these remarks as motivation, we will now begin to gradually work our way towards stating the univalence axiom. Let f and g be elements of the dependent function type $\prod_{a \,:\, A} P(a)$, i.e. assignments of a point of the space $P(a)$ to each element $a \,:\, A$. To capture type-theoretic homotopy we first need to define the following intermediate type:

$$f \sim g \equiv_{df.} \prod_{a \,:\, A} \mathsf{Id}_{P(a)}(f(a), g(a))$$

Read out, this says that $f \sim g$ is the dependent function type consisting of assignments of an identity proof between $f(a)$ and $g(a)$ to each element $a \,:\, A$. An element of it can thus be seen as a continuous selection of paths between the values taken by f and g. We may also say that $f \sim g$ is the type of proofs that f and g, interpreted as continuous functions, are homotopic, i.e. that there is a homotopy between them.

Using this type, the type of proofs that a function f is a homotopy equivalence can be defined as

$$\mathsf{isequiv}\, f \equiv_{df.} \left(\sum_{g:B \to A} f \circ g \sim 1_B \right) \times \left(\sum_{h:B \to A} h \circ f \sim 1_A \right)$$

This definition makes $\mathsf{isequiv}\, f$ the type of pairs of proofs that there are functions $g, h \,:\, B \to A$ such that $g \circ f$ is homotopic to the identity function on B, and $f \circ h$ to the identity function on A. We let the type $A \simeq B \,:\, \mathcal{U}$ be the type that contains all homotopy equivalences $f : A \to B$.

Relying on this sequence of concepts, it is finally possible to state the univalence axiom. In fact, it turns out to be remarkably simple; it just allows us to postulate the existence of a homotopy equivalence

$$(A \simeq B) \simeq \mathsf{Id}_{\mathcal{U}}(A, B)$$

for any types A, B. Informally, it says that the type of homotopy equivalences between A and B is itself homotopy equivalent to the type of identities between A and B.

---

[13] *Repeated* may perhaps not be the correct word in this case since we strictly speaking have a separate function refl_X for a different type X for each application. The n-iterated identity type is referred to as the *n-fold iterated loop space* in Univalent Foundations Program (2013, p. 70).

In terms of the Curry–Howard correspondence, it implies that given a proof that $f : A \to B$ is a homotopy equivalence, we can always construct a proof that there is an identity between A and B.

As expected, the main motivation for univalence comes from the homotopy semantics, according to which it says that we are always allowed to treat homotopy equivalent spaces as equal. In a way, this accords with the structuralist attempts to identify isomorphic structures, i.e. to take isomorphism as the identity condition for structures (Awodey 2014). And in fact such an identification, at least for algebraic structures, follows from the univalence axiom (Univalent Foundations Program 2013, pp. 96–99). It is important to note, however, that the possibility of both the identification of isomorphic structures and of homotopy equivalent spaces requires the multiplicity of non-equal identities supplied by intensional type theory; it is impossible to do so using classical extensional equality, as can be seen by considering e.g. the automorphism group of any object, which requires the existence of non-trivial isomorphisms to be meaningful. Or, for that matter, the fact that sets are isomorphic when they have the same cardinality, but making any two sets with different members identical would result in contradiction.

The Curry–Howard correspondence can also be used to illustrate the univalence axiom. We first note that the existence of functions $f : P \to Q$ and $g : Q \to P$ is sufficient to guarantee that P is true iff Q is, but this does not say anything about the nature of the proofs involved in this equivalence, and neither does it say anything about the identity conditions of the propositions themselves. Since univalence allows us to construct a proof that P and Q are the same from proof of $P \simeq Q$, it bears directly on the question of when propositions are to count as identical. What we need to do in order to see what it says to is to get a grasp of the type $P \simeq Q$ when P and Q are interpreted as propositions.

Following traditional usage, we say that a type theory is *propositionally extensional* if we can construct a proof of $P = Q$ from proofs of $P \to Q$ and $Q \to P$. Univalence does not give propositional extensionality in this sense, since the mere existence of functions between two types is insufficient to draw the conclusion that these types are identical. The exception is when we limit ourselves to types P that have at most one member, and such types are precisely the ones that are referred to as mere propositions in HoTT (Univalent Foundations Program 2013, p. 118). What follows from univalence is a principle we will call *weak* propositional extensionality, in distinction to the traditional kind, which we will refer to as *strong*. To describe it, we first have to say something more about the propositions-as-types interpretation.

We said that a proposition P is to be interpreted as the type of its proofs. But in the particular case of conditionals, a proof of $P \to Q$ is interpretable both as a *conditional proof* of Q from the premise that P, and as a *way to transform* any proof of P to a proof of Q. The first interpretation indicates that conditional proofs are the kind of things that can be *concatenated* by placing one after the other, as long as the premises of the second are proved as part of the first. The second lets us *apply* a conditional proof $f : P \to Q$ to a proof $p : P$ to get an unconditional proof $f(p) : Q$. These interpretations are not in conflict: if we take $f : P \to Q$ to be a proof of Q given a premise P, then concatenation of proofs $f \circ p$ also gives the method of

application. Thus, when both composition and application are possible, we should have that $f(p) = f \circ p$.

Every proposition P has an associated *identity proof* $1_P : P \to P$, which *qua* conditional proof may be seen the obvious derivation of P from P obtained by merely reiterating P. Applying this to $\simeq$ means that, intuitively, a proof of $P \simeq Q$ can be obtained by giving proofs $f : P \to Q$ and $g : Q \to P$, such that $g \circ f$—the concatenation of f and g—is homotopic to the identity proof of P, and $f \circ g$ to that of Q.

What does it mean to say that one conditional proof is homotopic to another? Looking at the definition of $\sim$, we see that a proof of $g \circ f \sim 1_P$ ought to correspond to a (non-empty) collection of proofs that $g(f(p)) = 1_P(p) = p$ for all proofs $p : P$, or perhaps to a uniform method of constructing a proof of $g(f(p)) = p$ given any $p : P$. Such a collection or method cannot, of course, exist unless it actually is the case that $g(f(p)) = p$ for all $p : P$, so this requires that $g(f(p))$ and p are indeed the same proof. Since this must hold for $f(g(q))$ and q as well, we see that $P \simeq Q$ has the consequence that the type of proofs of P must be isomorphic (in the category-theoretic sense of being invertibly mappable) to the type of proofs of Q. The question of when $P \simeq Q$ holds thus ultimately rests on when proofs are to count as identical, which means that univalence gives a way of producing identity conditions of propositions from those of proofs. We refer to the validity of the inference from the isomorphism of P with Q to $P = Q$, when P and Q are propositions, as *weak propositional extensionality*.

Whether weak propositional extensionality holds or not depends on which identity conditions for proofs and propositions we adopt. For instance, strong propositional extensionality implies weak, since the existence of an isomorphism between P and Q requires the existence of functions $P \to Q$ and $Q \to P$. But weaker conditions suffice. For any proposition P, let the *lemma family L_P* be a type family $L_P : P \to \mathcal{U}$ that maps any proof $p : P$ to the type of all theorems (in the mathematician's sense, rather than merely the logician's) that are proved and used as lemmas in p. We will consider the assumption that $p = q \to L(p) = L(q)$, which means that if two proofs use and rely on different lemmas, they are not equal; this arguably conforms to at least one common interpretation of what it means to be the same proof in ordinary mathematics. It also follows from the type theoretic indiscernibility of identicals if L_P is a well-defined type family.

Let $P \simeq Q$ and take both P and Q to be sufficiently important to count as "theorems". If one exists, let $p : P$ be a proof of P such that $Q : L(p)$ does not hold, or alternatively take $q : Q$ such that not $P : L(q)$. Consider two cases:

(*i*) No such p or q exists, in which case it is impossible to prove either of P or Q without proving the other. If $P \neq Q$, this means that P and Q are essentially unprovable, since no one of them can be proved first. Disregarding this possibility, we must have that $P = Q$.

(*ii*) Either such a p or such a q exists. Without loss of generality, assume that the first of these is the case. Assume that f, g are conditional proofs $f : P \to Q$ and $g : Q \to P$ such that $g(f(p)) = p$. Then $Q : L(g(f(p)))$, since $f(p)$ is a proof of Q and $g(f(p))$ is obtained by concatenating $f(p)$ with g. So $L(g(f(p))) \neq L(p)$, and from the assumption that equal proofs have the same

lemmas, that $g(f(p)) \neq p$, which yields a contradiction. Thus no such f, g can exist, so $P \simeq Q$ cannot hold.

Thus weak propositional extensionality follows from some quite reasonable assumptions about proof identity. Furthermore, univalence can, informally, be seen as a proof relevant version of weak propositional extensionality.

5 Varieties of extensionality

We have now introduced some of the fundamental ideas of HoTT, and it is time to delve into matters pertaining more directly to questions of intensionality and extensionality. An *extensional type* in HoTT is one in which all of the n-iterated identity types are trivial, or in the topological interpretation, one whose identity types are all contractible spaces. *Extensional type theory* is type theory in which one may infer a judgmental equality $a \equiv b : X$ from the existence of an identity proof $p : \mathsf{Id}_X(a, b)$. It is certainly not obvious what this has to do with traditional notions of extensionality and intensionality used in logic and philosophy. In this section and the next we will review some of these, and trace out connections between them and to concepts in HoTT. In doing so, we are attempting a kind of conceptual analysis, although one that consists in tracing logical fibres through a Wittgensteinian family resemblence, rather than a search for essences or sufficient and necessary conditions.

Figure 7 provides a road map of the notions we will discuss. Concepts pertaining to HoTT are in grey boxes, and arrows represent logical consequence, although in some
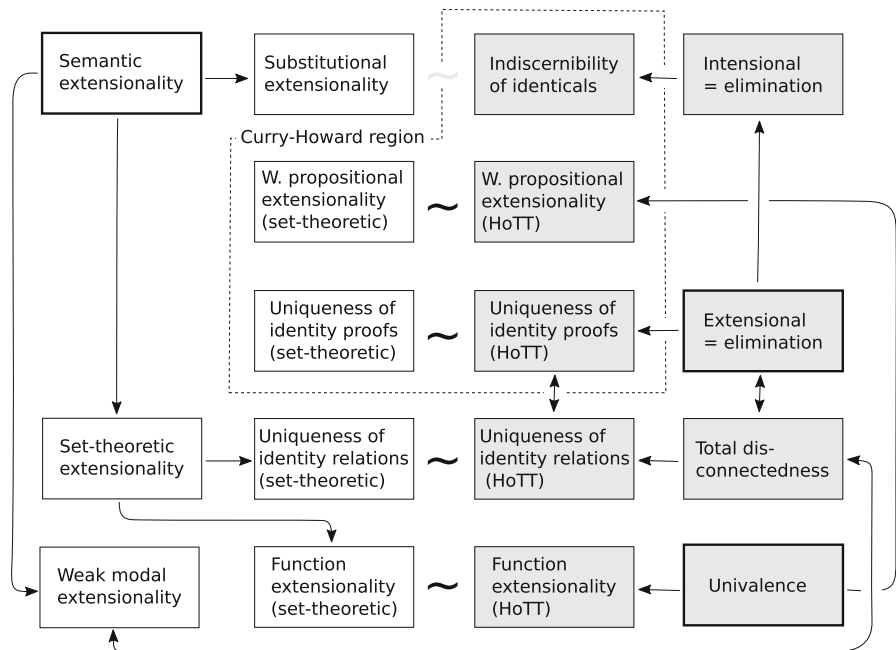

Fig. 7 Some concepts of extensionality

cases only given specific interpretations of the terms involved. A '∼' in the diagram stands for an analogy; a more careful investigation would preferably flesh these out in terms of definite structural relationships.

The two main concepts of extensionality relevant to HoTT—*univalence* and *extensional = elimination* (i.e. the equality reflection rule)—appear on the far right; of these the first is considered an optional (but often crucial) postulate, and the second would make HoTT trivial if it held universally, although it can be applied to individual types. The present section will concern itself with the boxes *semantic extensionality*, *substitutional extensionality*, and *set-theoretic extensionality*, which are all concepts that come directly out of traditional logic. The next section is about the boxes immediately to the right of set-theoretic extensionality: weak propositional extensionality, uniqueness of identity proofs, uniqueness of identity relations, and function extensionality, presented set theoretically. These are notions that are of importance for us mainly because of their structural similarities to corresponding concepts in HoTT, although we will see that in two of the cases, the logical ties to set-theoretic extensionality turn out to be weak.

Section 7 starts out from the concepts of intensionality in HoTT, and traces their relationships with semantic extensionality, especially via the *indiscernibility of identicals*, as it was introduced in Sect. 2. Here we also discuss the other forms of extensionality in the grey boxes, and their connections to the more traditional forms we have discussed in Sects. 5 and 6. We will see that the fact that HoTT essentially has two fundamental and non-equivalent notions of extensionality—uniqueness of identities and univalence—gives rise to complications when they are to be related to classical concepts, partly because it simply cannot be the case that both of them correspond to what philosophers or classical logicians usually mean by "extensionality".

The result of Sects. 5–7 will be largely negative in the sense that few concrete connections will be found between extensionality in classical logic and extensionality in HoTT. Section 8 attempts to alleviate this by providing a different link, going via modal logic and the topological semantics of HoTT.

Semantic extensionality, which we will begin with, may be seen as fundamental, but also merely schematic. We will call a context (e.g. a sentence, theory or language) *semantically extensional* iff the extension of any complex expression in it is a function of the extensions of its component expressions. This does not rule out that expressions *have* non-extensional features, but only that such features do not affect extensions. Allowing such *inessential* non-extensional features seems necessary if one is to have a workable concept, since e.g. terms in first-order logic certainly *have* non-extensional properties, such as syntactical ones, even if these do not matter in determining complex expressions' extensions.

Given an appropriate binary coextension relation coex between expressions, we may write the semantic extensionality of the expression ψ as the axiom schema

$$\xi \; \mathsf{coex} \; \xi' \to \psi \; \mathsf{coex} \; \psi[\xi'/\xi]$$

where ξ, ξ' and $\psi[\xi'/\xi]$ all have to be well-formed expressions. It is clear that unless we say something about what makes expressions coextensional, this tells us very little.

Various forms of assumptions about expressions' coextensionality give rise to various other concepts of extensionality, two of which we will discuss here.

The first, which will be called *substitutional extensionality*, can be summarised as the requirement that the truth value of a sentence φ does not change when the replacement of a term s with a term t is made, given that $s = t$ is true. In first-order logic it can be expressed simply as the axiom schema

$$(\varphi \wedge s = t) \rightarrow \varphi[t/s]$$

Substitutional extensionality is closely related to Quine's notion of a *referentially transparent context*, in which both the substitution of coreferential terms and the rule of existential instantiation are truth preserving, as in first-order logic (Quine 1961b). It can be obatained from semantic extensionality by laying down the definitions

$$\varphi \; \mathsf{coex} \; \psi \equiv_{df.} \varphi \leftrightarrow \psi$$
$$s \; \mathsf{coex} \; t \equiv_{df.} s = t$$

for formulae φ, ψ and terms s, t.

The second form of extensionality we will discuss here is the set theoretical[14] one, embodied in the *axiom of extensionality*

$$\forall x \; (x \in c \leftrightarrow x \in d) \leftrightarrow c = d$$

Indeed, not assuming this axiom is often taken to be sufficient for making mathematics intensional, as in e.g. the classic Feferman (1985). How is this set theoretic extensionality related to the substitutional or semantic versions?

It is trivial that substitutional extensionality does not imply set extensionality, since every instance of substitutional extensionality is a theorem of FOL with identity, but the set-theoretic axiom of extensionality is not. To investigate whether the converse holds, it is naturally crucial to not assume FOL to already have identity. Luckily, set theory can be developed without primitive identity, and the axiom of extensionality taken as a *definition* of the identity relation, given that what we will call the *weak axiom of set extensionality*

$$\forall x \; (x \in c \leftrightarrow x \in d) \rightarrow \forall y \; (c \in y \leftrightarrow d \in y)$$

is imposed. This is needed to guarantee that sets with the same members are themselves members of the same sets, and is a theorem of the usual extensionality axiom in first-order logic *with* identity.

In a first-order language *in which the only predicate is* $\in$, the weak axiom of set extensionality does imply substitutional extensionality, as can be proved through induction on the complexity of formulae. If we have other primitive predicates as well,

---

[14] Since we will focus on extensionality, which is roughly the same in more or less any extensional set theory in existence, we will usually not need to pick which one we mean when we talk about 'set theory'. In the few cases where it matters, substitute 'ZFC'.

$c = d$ does not guarantee that c and d are substitutable *salva veritate*. Thus weak set extensionality neither implies nor is implied by substitutional extensionality.

However, set extensionality can be made to follow from semantic extensionality by making a specific definition of the relation **coex** for terms that refer to sets. Assuming the same definition of **coex** for formulae as above, and restricting that definition for terms to ones *not* referring to sets, we define

$$c \text{ coex } d \equiv_{df.} \forall x \ (x \in c \leftrightarrow x \in d)$$

whenever c or d are terms for sets. Given this definition, it can be shown that the weak axiom of set extensionality follows in FOL without identity, and the set theoretic axiom of extensionality follows in a language which has identity:

Proposition 1 *If c coex $d \leftrightarrow \forall x \ (x \in c \leftrightarrow x \in d)$ then semantic extensionality entails weak set extensionality in a language without identity, and set extensionality in a language with identity.*

Proof From semantic extensionality it follows that c and d having the same members is sufficient for full substitutivity *salva veritate*. Applying this to the theorem $\forall y \ (c \in y \leftrightarrow c \in y)$ gives weak set extensionality, and applying it to $c = c$ gives set extensionality. $\qquad\square$

We have thus found that although set extensionality and substitutional extensionality are logically independent, they both follow from semantic extensionality given appropriate definitions of what it means for various forms of expressions to be coextensional. This is why we referred to the latter as *schematic*: it can be made to fit several other concepts, depending on how **coex** is defined.

6 Alternate forms of extensionality in set theory

This section will construct—using set-theoretic components—a number of concepts of extensionality that do not usually make an appearance in set theory, but which have a important roles to play in intuitionistic type theory. The concepts in question are function extensionality, weak propositional extensionality, and two forms of uniqueness of identity. Both one of the forms of uniqueness of identity and weak propositional extensionality will furthermore require using a version of the Curry–Howard correspondence adapted for set theory, which will be presented later in this section. However, we begin with function extensionality, i.e. the implication

$$\forall x \ (f(x) = g(x)) \rightarrow f = g$$

This corresponds to defining

$$f \text{ coex } g \equiv_{df.} \forall x \ (f(x) = g(x))$$

for function terms f, g. To interpret this set theoretically, it is easiest to take the function application notation $f(x)$ to be contextually defined by a schema such as

$$s = f(t) \equiv_{df.} \langle t, s \rangle \in f$$

for arbitrary terms s, t. Function extensionality is then equivalent to

$$\forall x \forall y \; (\langle x, y \rangle \in f \leftrightarrow \langle x, y \rangle \in g) \rightarrow f = g$$

whenever f and g are sets of ordered pairs satisfying the requirements of a total function. But this condition is guaranteed to hold—for any encoding $\langle \cdot , \cdot \rangle$ of ordered pairs as sets—by set extensionality.

For the other direction, i.e. from function extensionality to set extensionality, we need that every set must also be a function. This does not hold given the above interpretation of functions as sets of ordered pairs. To be sure, just as it is possible to interpret functions as sets, it is possible to interpret any set f as a function, e.g. by distinguishing a certain 2-element set Ω (say $\Omega = \{0, 1\}$) and then taking f to be the set of arguments for which f takes the value 1 rather than 0. More precisely, we would define the function notation $f(x)$ through the equivalence

$$s = f(t) \equiv_{df.} (1 = s \wedge t \in f) \vee (0 = s \wedge t \notin f)$$

We can then to derive set extensionality from function extensionality as follows:

$$\forall x \; (f(x) = g(x)) \rightarrow f = g \Leftrightarrow$$
$$\forall x \; ((1 = f(x) \wedge x \in g) \vee (0 = f(x) \wedge x \notin g)) \rightarrow f = g \Leftrightarrow$$
$$\forall x \; ((((1 = 1 \wedge x \in f) \vee (1 = 0 \wedge x \notin f)) \wedge x \in g)$$
$$\vee (((0 = 1 \wedge x \in f) \vee (1 = 1 \wedge x \notin f)) \wedge x \notin g)) \rightarrow f = g \Leftrightarrow$$
$$\forall x \; ((x \in f \wedge x \in g) \vee (x \notin f \wedge x \notin g)) \rightarrow f = g \Leftrightarrow$$
$$\forall x \; (x \in f \leftrightarrow x \in g) \rightarrow f = g$$

However, such an interpretation is incompatible with the previous one, so at most we can see function extensionality as following from set extensionality, or set extensionality as following from function extensionality, but we cannot have both. As the interpretation of functions as sets of ordered pairs is by far the more common one in set theory, we take the former option.

The second form of extensionality we will consider here is one that is clearly type theoretical: uniqueness of identities. Indeed, its strongest form, the equality reflection rule

$$\frac{\Gamma \vdash p \; : \; \mathsf{Id}_X(a, b)}{\Gamma \vdash a \equiv b \; : \; X}$$

is what distinguishes extensional Martin-Löf type theory from intensional. In terms of the homotopy semantics, a type theory that validates this rule is one in which every identity type is a one-point space.

In HoTT, it is also useful to discuss weaker versions which correspond to the condition that identity types have to be *contractible*, i.e. homotopy equivalent to one-point spaces, but not necessarily one-point spaces themselves. In type theoretical terms, this means that for each $p \;:\; \mathsf{Id}_X(a, a)$, there is a higher identity $p' \;:\; \mathsf{Id}_{\mathsf{Id}_X(a,a)}(p, \mathsf{refl}_X(a))$, or differently put, that all propositional identities are themselves *propositionally*, but not necessarily judgmentally, equal.

To translate this to set theory we will consider two interpretations of the identity type $\mathsf{Id}_X(a, b)$: as a type of *relations* or *relation instances* holding between a and b, and as a type of *proofs of a proposition* $a = b$. We refer to these interpretations as *uniqueness of identity relations* and *uniqueness of identity proofs*, and will discuss the first of these directly, and the second, since it relies crucially on interpreting types as propositions, after we have introduced a version of the Curry–Howard correspondence for set theory. Although they come out equivalent in HoTT, the same does not hold in set theory.

From a philosophical viewpoint, there is in itself nothing strange with more than one instance of a relation holding between different entities. It is possible for someone to be *related to* me through different lineages, such as when pairs of siblings marry one another so that, say, my sister's wife is also my wife's sister. Objects may be *similar* in more than one respect. And mathematical structures are often *embeddable* in one another in more than one way. So the idea that the identity relation—so long as we take it to be a relation at all in any substantial sense of the word—could be treated as non-unique should not be seen as obviously incoherent. Different interpretations of relations' identity conditions may give us different answers.

Set theory has its own interpretation of what a relation is: a set of ordered pairs. The *identity relation* on a set A is the diagonal $\Delta A = \{\langle x, x \rangle \mid x \in A\}$. An *instance* of such a relation, for an element $a \in A$, can then fairly reasonably be interpreted as a pair $\langle a, a \rangle$. Thus the question of whether identities are unique in set theory comes down to whether such pairs are unique. Insofar as we take them to be set-theoretic constructions (e.g. as defined by Wiener, Hausdorff or Kuratowski) this uniqueness follows from the set-theoretic axiom of extensionality, although the converse does not hold unless we somehow manage to encode ordered pairs in such a way that *all* sets are also ordered pairs. We thus take uniqueness of identity relations to be a consequence of set extensionality, but set extensionality to not be a consequence of uniqueness of identity relations.

The next two extensionality concepts—uniqueness of identity proofs and weak propositional extensionality—both depend on the Curry–Howard interpretation of propositions as types. But since types are the type-theoretic analogue of sets, what we need is a version of the Curry–Howard correspondence for set theory. In fact, such a correspondence can be set up, since it does not depend critically on anything specific to type theories rather than set theory. The guiding idea is to go back to the correspondence's origins in the Brouwer–Heyting–Kolmogorov semantics of intuitionistic logic, which are easily adapted to classical logic.[15]

---

[15] We should come clear at once about the fact that neither the usefulness nor the beauty that the Curry–Howard correspondence has in type theory will be features of the set theoretical version; it will rather resemble something like a translation of *The Waste Land* into first-order logic. This naturally invites the

To simplify, we will only consider a very simple language $\mathcal{L}$ without variables or quantifiers, but with individual constants c_1, c_2, ..., possibly a set of primitive predicates P_1, P_2, ..., an identity predicate $=$, and the classical connectives $\bot$, $\neg$, $\wedge$, $\vee$, $\rightarrow$. This gives us a kind of middle ground between sentential and predicate logic, which is easy to work with but still allows us to discuss identity.[16] A treatment of full first-order logic in the style we will do here would require use of dependent sets, and although this can be done, it requires far more elaborate set-theoretic constructions than we will need to use for our limited language. In particular, dependent functions can be treated with an elegance in type theory that seems to be unavailable in set theory.

As metalanguage, we use ZFC. Just as the Curry–Howard correspondence allows us to interpret a type as a proposition in type theory, we want to be able to interpret *sets* as propositions, and in particular non-empty sets as true propositions. In the following, we will use s, t, u, t', t_1, t_2, etc. as metavariables ranging over terms (in the present case, individual constants) of $\mathcal{L}$. Let a *CH model* $\mathfrak{M}$ of $\mathcal{L}$ be

- (i) an assignment of a set $\mathfrak{M}(P(t_1, \ldots, t_n))$ to each atomic sentence $P(t_1, \ldots, t_n)$, together with
- (ii) a bijective function $\mathsf{refl}_{\mathfrak{M}}$ from the terms of $\mathcal{L}$ to a set $\mathsf{CRefl}_{\mathfrak{M}}$, which we call the set of *canonical reflexivity proofs*, and
- (iii) functions

$$\mathsf{symm}_{\mathfrak{M}} : \mathfrak{M}(s = t) \rightarrow \mathfrak{M}(t = s)$$
$$\mathsf{trans}_{\mathfrak{M}} : \mathfrak{M}(s = t) \times \mathfrak{M}(t = u) \rightarrow \mathfrak{M}(s = u)$$
$$\mathsf{idelim}_{\mathfrak{M}} : \mathfrak{M}(t_k = u) \times \mathfrak{M}(P(t_1, \ldots, t_k, \ldots, t_n)) \rightarrow \mathfrak{M}(P(t_1, \ldots, u, \ldots, t_n))$$

for all the terms of $\mathcal{L}$. These functions will be needed to guarantee that identity performs as expected.

A *valuation* based on $\mathfrak{M}$ is a function $[\![\cdot]\!]_{\mathfrak{M}}$ that assigns sets to sentences of $\mathcal{L}$ such that

$$\mathsf{refl}_{\mathfrak{M}}(t) \in [\![t = t]\!]_{\mathfrak{M}}$$

and the following recursive clauses hold:

$$[\![\bot]\!]_{\mathfrak{M}} = \emptyset$$
$$[\![P(t_1, \ldots, t_n)]\!]_{\mathfrak{M}} = \mathfrak{M}(P(t_1, \ldots, t_n))$$
$$[\![\varphi \wedge \psi]\!]_{\mathfrak{M}} = [\![\varphi]\!]_{\mathfrak{M}} \times [\![\psi]\!]_{\mathfrak{M}}$$
$$[\![\varphi \vee \psi]\!]_{\mathfrak{M}} = [\![\varphi]\!]_{\mathfrak{M}} \sqcup [\![\psi]\!]_{\mathfrak{M}}$$

Footnote 15 continued

question of why we want to do it at all, and if such translations can tell us anything about extensionality in type theory *or* set theory. The answer to this is that *some* translation will always be necessary, since the concepts we are studying seem to differ so radically from each other. One should however make sure that we do not forget that everything we say is relative to the translations we have made.

[16] This is, essentially, the kind of language described in Barker-Plummer et al. (2011, chs. 1–8) as *propositional logic*.

$$[\![\varphi \rightarrow \psi]\!]_{\mathfrak{M}} = ([\![\psi]\!]_{\mathfrak{M}})^{[\![\varphi]\!]_{\mathfrak{M}}}$$

$$[\![\neg\varphi]\!]_{\mathfrak{M}} = [\![\varphi \rightarrow \perp]\!]_{\mathfrak{M}}$$

The '⊔' in the clause for disjunction stands for disjoint union; a regular union could also have been used, but we have gone with the disjoint version in order to stay closer to the type theoretic treatment. As in the Curry–Howard correspondence or the Brouwer–Heyting–Kolmogorov semantics, we refer to the set $[\![\varphi]\!]_{\mathfrak{M}}$ as the set of *proofs* of the sentence φ. Since the proof sets of implications have functions as members, we can *apply* a proof $q \in [\![\varphi \rightarrow \psi]\!]_{\mathfrak{M}}$ to a proof $p \in [\![\varphi]\!]_{\mathfrak{M}}$ to obtain a proof $r \in [\![\psi]\!]_{\mathfrak{M}}$, and we can *compose* a proof $p \in [\![P \rightarrow Q]\!]_{\mathfrak{M}}$ with a proof $q \in [\![Q \rightarrow R]\!]_{\mathfrak{M}}$ to obtain a proof $r \in [\![P \rightarrow R]\!]_{\mathfrak{M}}$; we write both these operations as $r = q \circ p$. We say that φ is *true* in the model $\mathfrak{M}$ iff $[\![\varphi]\!]_{\mathfrak{M}} \neq \emptyset$. A sentence is *logically true* iff it is true in all models.

Let $\mathcal{L}^{+}$ be the first-order language that extends $\mathcal{L}$ by adding first-order variables and quantifiers, and define logical truth in $\mathcal{L}^{+}$ in the usual Tarskian way. Then we have the following:

Theorem 1 *φ is a logical truth of $\mathcal{L}$ iff φ is a logical truth of $\mathcal{L}^{+}$, for all sentences $\varphi \in \mathcal{L}$.*

Proof It is sufficient to show that any first-order model $\mathfrak{M}^{+}$ of a sentence $\varphi \in \mathcal{L}$ may be assigned a CH model $\mathfrak{M}$ of the same sentence φ, and vice versa. The left-to-right direction can be done by assigning atomic sentences true in $\mathfrak{M}^{+}$ arbitrary non-empty sets, and the right-to-left direction by making a first-order model from equivalence classes of $\mathcal{L}$'s individual constants, much as in the Henkin construction commonly used to prove first-order completeness. The proof then proceeds by checking that the recursive clauses given above for the connectives give the same truth conditions as the usual ones for classical logic, e.g. that $[\![\varphi]\!]_{\mathfrak{M}} \times [\![\psi]\!]_{\mathfrak{M}}$ is non-empty iff both $[\![\varphi]\!]_{\mathfrak{M}}$ and $[\![\psi]\!]_{\mathfrak{M}}$ are, etc. This follows from the standard set-theoretic definitions of the operators in question.   □

We have thus set up something similar to the Curry–Howard correspondence for a fragment of first-order logic: sets can be taken as the semantic values of sentences in such a way that true sentences have non-empty sets as values, just as true propositions under the Curry–Howard correspondence correspond to non-empty types. We will use this semantics to transfer the concepts of weak propositional extensionality and uniqueness of identity proofs from HoTT to set theory, in order to see how they relate to the set-theoretic axiom of extensionality.

The question of whether identity proofs are unique translates directly to whether the sets $[\![s = t]\!]_{\mathfrak{M}}$ can have more than one element. But neither this nor its negation follows from purely set-theoretical assumptions like the axiom of extensionality: if we want it to hold, we have to add it as an extra condition in our definition of CH model. As such, the decision is part of the semantics rather than the metalanguage that semantics is implemented in.

This is actually not very different from how the situation is in type theory. The Curry–Howard correspondence in HoTT is also best seen as a *semantics* for many-sorted first-order logic, which allows discussion of truth or falsity to be carried out

in terms of discussion about types.[17] What obscures this fact somewhat is that the logic (the discourse about truth and falsity) and the type theory (the discourse about types/sets/properties) share the same syntax. Nevertheless, it is clear that uniqueness of identity proofs is just as independent as an assumption in type theory as it is in the semantics in terms of CH models we have sketched here. For example, in Univalent Foundations Program (2013, p. 118), what is called "traditional logical notation", which is an alternative propositions-as-types interpretation for HoTT, is defined using a truncation operator $\| \cdot \|$ that discards all higher identity structure of types. As presented, this notation does not mention how to interpret first-order identity, but it would be easy to extend it to do so by making the definition

$$x = y \ =_{df.} \ \| \mathsf{Id}_A(x, y) \|$$

where A is the type of x and y. Under this modified semantics it is reasonable to hold that identity *proofs* would be unique in HoTT, while identity *relations* could still be non-unique, since these deal with the types directly, and not merely through the propositions-as-types interpretation.

Let us move on to weak propositional extensionality in set theory. The traditional (strong) version, according to which propositions are equal when interderivable, is actually *incompatible* with the semantics outlined here: although all logical truths are interderivable, they will generally, by the recursive clauses we have given here, have wildly different sets of proofs. More generally, the strong version of propositional extensionality is incompatible with proof-relevant mathematics (or logic), which both our current semantics, and that of HoTT, adhere to.

The case with the homotopy-theoretic version is more subtle. In Sect. 4 we described weak propositional extensionality as allowing us to construct an element of $\mathsf{Id}_\mathcal{U}(P, Q)$ from elements of

$$\mathsf{Id}_{P \to P}(q \circ p, \mathsf{refl}_\mathcal{U}(P \to P))$$

and

$$\mathsf{Id}_{Q \to Q}(p \circ q, \mathsf{refl}_\mathcal{U}(Q \to Q))$$

where $p : P \to Q$ and $q : Q \to P$. To translate this directly to our set-theoretic Curry–Howard semantics, we would need to introduce reflexive or canonical proofs of sentences of the form $P \to P$, much as we did for those of the form $t = t$ when defining CH models. While this can be done, it does however turn out that much of what we want to say can be shown in another way, which does not require us to expand the semantics.

For any model $\mathfrak{M}$ of $\mathcal{L}$, let $\mathcal{L}|_\mathfrak{M}$ be the fragment of $\mathfrak{L}$ that does not contain $\bot$ and only contains those atomic sentences $P(t_1, \ldots, t_n)$ or identities $s = t$ not assigned

---

[17] A similar point is made in Howard's own exposition (1980) as he criticises his earlier use of the phrase 'formulae-as-types notion of construction' as indicating that formulae *are* types, when what he meant is that formulae are *names* of types, i.e. refer to types. But this is the same as saying that the 'formulae-as-types notion of construction' is actually a semantics.

empty sets by $\mathfrak{M}$. Let a *logical* model $\mathfrak{M}$ be one in which no two atomic sentences or identities are assigned the same set of proofs. It is clear that, at least on the intended interpretation, logical models are the only ones that are reasonable, since models which are not logical are ones in which two logically independent sentences have exactly the same proofs.

Proposition 2 *Let $\mathfrak{M}$ be a logical model of $\mathcal{L}$. Then $[\![\cdot]\!]_{\mathfrak{M}}$ is injective on the fragment $\mathcal{L}|_{\mathfrak{M}}$.*

Proof By induction on the truth clauses. □

The upshot of this is that, insofar as there are some atomic sentences that are true, there will always be a significant part of $\mathcal{L}$ wherein two different sentences *never* express the same proposition, so no non-trivial form of weak propositional extensionality can hold at all. In particular, the existence of functions

$$p : [\![\varphi]\!]_{\mathfrak{M}} \to [\![\psi]\!]_{\mathfrak{M}}$$
$$q : [\![\psi]\!]_{\mathfrak{M}} \to [\![\varphi]\!]_{\mathfrak{M}}$$

cannot imply that $[\![\varphi]\!]_{\mathfrak{M}} = [\![\psi]\!]_{\mathfrak{M}}$, unless we have that $\varphi = \psi$ or that $[\![\varphi]\!]_{\mathfrak{M}} = [\![\psi]\!]_{\mathfrak{M}} = \emptyset$.

What makes weak propositional extensionality possible in HoTT is the presence of a weaker form of identity in the metalanguage: HoTT's weak propositional extensionality does not say that P and Q are judgmentally equal, but only propositionally so. A strict form of equality, like the one of classical logic, or judgmental equality, will always make it impossible for propositions to be equal as long as we use something like the proof-relevant Curry–Howard semantics given here. The very point of having the clause for $\varphi \wedge \psi$ be $[\![\varphi]\!]_{\mathfrak{M}} \times [\![\psi]\!]_{\mathfrak{M}}$ rather than, say, $[\![\varphi]\!]_{\mathfrak{M}} \cap [\![\psi]\!]_{\mathfrak{M}}$ (which would give the same truth conditions) is that the former does not discard any information about the proofs we have used for φ and ψ, while the latter typically does. It is only the presence of an equality relation like the propositional one in intensional type theory, whose elimination rule does not discard information either, but only shows how to translate it to other types, that makes it possible to equate propositions with one another even if they do not have strictly (judgmentally) the same proofs.

In summary, we have found that set extensionality implies function extensionality as well as the set-theoretic uniqueness of identity relations, but that neither of these entail set extensionality. Using the set theoretic Curry–Howard semantics, we then showed that not only is set extensionality independent of uniqueness of identity proofs, but set theory, at least with extensionality, is incompatible with weak propositional extensionality. This is a consequence of proof relevance together with the absence of weaker forms of identity in classical logic.

7 Intensionality in HoTT

Martin-Löf (2015) has explained that the use of the word "intensional" for a non-propositional form of identity was influenced by a footnote in Gödel's *Dialectica* paper:

> Identität zwischen Funktionen ist als intensionale oder Definitionsgleichheit zu verstehen.[18]

A related use of "intensional" as a synonym for "definitional" appears in e.g. Tait (1967), who Martin-Löf also refers to in his work. In conversation, Martin-Löf has said that he currently does not approve of such a use, since he believes that intensional equality should be equality of *meaning*, while definitional equality is a purely syntactic concept (Martin-Löf 2015). We have so far followed the earlier—Gödelian— tradition here in not specifically requiring a connection between intensionality and meaning, but simply focused on non-extensionality. The members of the Univalent Foundations program seem to interpret HoTT's intensionality in a rather similar way:

> Intensional type theory is so named because its judgmental equality, $x \equiv y$, is a very intensional equality: it says essentially that x and y "have the same definition", after we expand the defining equations of functions. By contrast, the propositional equality type $x = y$ is more extensional, even in the axiom-free intensional type theory [...] we can prove by induction that $n + m = m + n$ for all $m, n\ :\ \mathbb{N}$, but we cannot say that $n + m \equiv m + n$ for all $m, n\ :\ \mathbb{N}$, since the definition of addition treats its arguments asymmetrically. (Univalent Foundations Program 2013, p. 102)

The notation '$x = y$' used here is an abbreviation for the type $\mathsf{Id}_A(x, y)$, where $x, y\ :\ A$. It is understood that a *type theory* (or perhaps a logic) is intensional if it makes a difference between extensional and intensional identity, and that having the same definition is a relation that is strictly stronger than mere coextensionality.[19] The appearance of quotation marks around 'have the same definition' in the quote, however, indicates that this is an informal characterisation using an analogy.

From a λ-calculus point of view, there is a formal version of it which says that $a \equiv b\ :\ X$ iff $a, b\ :\ X$ and a and b reduce to the same term by applying β-reduction, i.e. application of a function to its arguments (Univalent Foundations Program 2013, p. 422). The existence of a propositional equality $p\ :\ \mathsf{Id}(a, b)$ does not ensure that a and b can be rewritten as the same term in this way. This is an explanation in terms of syntax, which does not mention HoTT's semantics. As we have already said, $a \equiv b : X$

---

[18] "Identity between functions is to be understood as intensional or definitional equality" (Godel 1958, p. 283).

[19] It is curious that in their book on Martin-Löf type theory, Nordström, Petersson & Smith refer to an identity type that is equivalent to judgmental equality as *extensional*, and to the one that appears as the propositional identity in HoTT as *intensional* (Nordstrom et al. 1990, pp. 57–62). This seems to be the exact opposite of the position taken by the members of the Univalent Foundations program, as well as being in conflict with their earlier occasional use of 'intensional' as a synonym for 'definitional'.

holds iff a and b refer to precisely the same point of the space X, and $p : \mathsf{Id}_X(a, b)$ holds if p is a path from a to b. This would indicate that a term's *intension* in HoTT can be seen as a point of a space X, and its extension as a path-connected component of X. Similarly, a type's intension would be a space and its extension a homotopy type. Both of these choices conform to the Fregean principle that intension (or meaning) is to serve as a way to present extension, so that the same extension may be presented in many different ways. For terms, these ways are represented by points, and for types, by specific topological spaces.

Collecting these suggestions in a table gives us the following:

	Intension	Extension
Term	Point	Path-connected component
Type	Topological space	Homotopy type

Of course, not *any* space can be assigned to any type. For example, identity types will have to be assigned path spaces, dependent types fibrations, etc. In any case, our interpretation would seem to validate the Fregean priciple that intension determines extension.[20] Somewhat surprisingly, what it does *not* show is that HoTT is semantically non-extensional in our sense. Since all functions are continuous, it follows that if s is path-connected to t, then $f(s)$ will be path-connected to $f(t)$, so function application preserves coextensionality. Likewise, the type forming operations we have, such as making product spaces or fibrations, also satisfy the requirement that applying them to homotopy equivalent spaces yield homotopy equivalent results. Thus intensions *do not* matter when the extensions of terms and types are to be determined, which is exactly what semantic extensionality says should make these things extensional.

To find intensionality we have to go further and consider not only terms and types, but also judgments. There are two kinds of these that we have discussed: the typing judgment $a : A$ and the equality judgment $a \equiv b : X$. Following the practice of Sect. 5, we will take judgments to be coextensional when they are both true or both false, with truth determined by the semantics we outlined in the beginning of Sect. 4. That is, $a : A$ is true iff a is a point in the space A, and $a \equiv b : X$ is true iff a and b are the same point, and this point is in the space X.

Beginning with the typing judgment, we can quickly see that replacement of coextensional terms will not change a judgment's truth value: if $a : A$, then we must also have $a' : A$ for any point a' which is path-connected to a, since a prerequisite for there to be a path between two points is that they are in the same space. On the other hand, a space A being homotopy equivalent to a space B is not sufficient for them to

---

[20] This may depend on how we interpret 'point' in the table: if the same point can be a member of two or more topological spaces, then a single point may be in two different path-connected components in these spaces. What we should say then is that *given* a term's type (i.e. its topological space), its intension determines its extension. But there are similar caveats applicable to Frege's original principle that *Sinn* determines *Bedeutung*: we cannot find out what a term's reference is in our world merely from knowing its meaning; we need to know which world is ours too.

have the same points, so we may very well have $a : A$ and $p : \mathsf{Id}_{\mathcal{U}}(A, B)$, but not $a : B$.

Equality judgments also violate replacement *salva veritate* of coextensional types. But they do not permit replacement of coextensional terms either, since the truth of $a \equiv b : X$ under the semantics we used depends on the exact identity of the points a and b refer to, rather than merely whether there is a path between them. So, if our proposed interpretation of coextensionality of judgments is correct, we have found a sense in which HoTT is indeed *semantically* intensional.

There is a possible source of confusion here, however. What we have done in taking a judgment's extension to be a truth value is essentially to pass over to the substitutional version of extensionality, and shown that HoTT does not satisfy this. But, as we said in Sect. 2, HoTT satisfies something rather close to substitutional extensionality: the indiscernibility of identicals. As we described it, this consequence of the rules of propositional identity in intensional type theory lets us construct an element of a dependent type $C(y)$ given an element of $C(x)$ and an identity $p : \mathsf{Id}_A(x, y)$. Since the existence of an element of $C(y)$ is interpreted as y having the property C, this would mean that $p : \mathsf{Id}_A(x, y)$ is sufficient for x and y to have the same properties. But how can things have the exact same properties while still giving rise to failures of substitutivity?

The answer is that a *property*, interpreted according to the Curry–Howard correspondence, is not the same thing as a *type* of HoTT. Indeed, they are not even part of the same language, strictly speaking. Recall our insistence in the last section that the Curry–Howard correspondence should really be seen as a semantics. Then *property* and *proposition* are concepts of the object language, while *type* is of the metalanguage. To be more precise, the object language, although it uses symbols from the metalanguage, is just a many-sorted first order logic, and does not use the judgmental form of equality or equalities between properties (Univalent Foundations Program 2013, pp. 4144, 118). This is why it satisfies substitution extensionality, while the full type theory of the metalanguage does not, and it is also why the $\sim$ denoting analogy between substitutional extensionality and the indiscernibility of identicals in Fig. 7 is greyed out: although they may be glossed similarly, they are very different concepts that apply to different systems. A similar phenomenon happens in Tarskian first-order semantics when we do not take identity to be logical, so that it can denote any congruence relation on the model. In such a case, we will still have full substitutivity of identicals in the object language, but can have failure of substitutivity in the metalanguage, i.e. we can have that $\mathfrak{M} \vDash s = t$ with s and t denoting different elements of $\mathfrak{M}$'s domain.

Having examined the relationship between HoTT and semantic and substitutional extensionality, we will now briefly consider the other analogies in Fig. 7. Weak propositional extensionality, as we noted in the last section, is incompatible with the Curry–Howard semantics as implemented in set theory. This fact, which depends on classical logic's lack of weaker forms of identity, also affects possible interpretations of the univalence axiom itself in set theory. To show this, we make a simple translation of this axiom: we say that a set theory is *univalent* iff, for any two sets a and b such that there are functions $f : a \to b$ and $g : b \to a$ such that $g \circ f = \Delta a$

and $f \circ g = \Delta b$, where Δa and Δb are the identity functions on a and b, we have that $a = b$.

Proposition 3 *Any univalent extension of ZF is inconsistent.*[21]

Proof For the necessary functions f and g to exist, it is clearly sufficient that g should be the inverse of f. But the existence of such functions is guaranteed merely by a and b having the same cardinality, so a univalent set theory must identify all sets of the same cardinality. Any extension of ZF, however, validates e.g. $\{\emptyset\} \neq \{\{\emptyset\}\}$.   □

Function extensionality offers a very different circumstance: it is compatible with both set theory and HoTT. In set theory, it follows from set extensionality, and in HoTT it follows from univalence (Univalent Foundations Program 2013, pp. 144–146), and in both cases, the converse does not hold. It thus supplies an example of where Univalence in HoTT actually plays the same role as extensionality does in set theory, and highlights how complex the relationships we are investigating really are.

Uniqueness of identities also points to this: we have split this concept into two, but in HoTT it is not obvious that this is a meaningful distinction to make. Uniqueness of identity *proofs* and uniqueness of identity *relations* in HoTT are just different interpretations of the same condition, at least insofar as we do not modify the Curry–Howard correspondence to interpret identities in terms of truncated types, as was indicated in the previous section. On the other hand, in set theory, one of them follows from set extensionality, while the other does not.

In fact, set extensionality itself provides a curious case. In any type theory, the common membership of even a single entity is sufficient to make types identical, while set extensionality requires common membership of *all* elements. In fact, even the *propositional* identity of a and b is enough to make any type they are in judgmentally equal. Thus intensional type theory, even without any additional axioms such as univalence, is in one way far more extensional than set theory—at least when 'extensional' is interpreted in the sense of the set theoretic axiom of extensionality.

Summing up, we have seen that the relationships between traditional forms of extensionality and those of HoTT are far from straightforward. The most promising connection we have found may be between semantic extensionality and the intension-extension characterisation of the HoTT semantics we sketched in this section, although this connection shows up only when we extend the semantics to also cover judgments. The intensionality of HoTT *as a system* comes from not having the extensional rule of = elimination, which would make propositional and judgmental identities logically equivalent. However, semantic extensionality is a rather thin concept, and it would be useful to see if we can work out this connection in more detail. This will be purpose of the next section.

---

[21] Interestingly, as an anonymous referee pointed out, *categorical* set theory is compatible with univalence in the sense we use it here. In fact, a topos with an added axiom making it skeletal is univalent in our sense.

8 Extensionality in modal logic

One characterisation of the divide between intensional and extensional sentences was made by Carnap (1956), primarily as an explication of ideas of Frege. The fundamental principle is, much as for Quine, who as we noted championed a version of substitutional extensionality, that φ is an extensional context if no replacement of coreferential terms can change the truth value of a sentence they appear in. An intensional context is one where this does not need to hold, but where substitutivity *salva veritate* among *intensionally* equivalent terms still holds. For Carnap, such an intensional equivalence is given by the identity of the *individual concepts* that are associated with the equivalent terms.

What sets this notion apart from Quine's is its very close relation to modality: in Carnap's logic, intensional equivalence is the same as *necessary* equivalence, so that we have that $\Box(\varphi \leftrightarrow \psi)$ is true iff φ and ψ are intensionally equivalent, i.e. have the same intension. In the specific case of identity, this entails that s and t are intensionally equal iff $\Box s = t$ holds. We will call a modal logic *weakly extensional* iff it validates the inference rule $s = t \dashv\vdash \Box s = t$, i.e. if it makes no difference between necessary and contingent equality. That this is not the same as extensionality *tout court* can be seen from the existence of the Kripkean theory of equality as always being necessary without having all other modal distinctions (e.g. between sentences) collapse. However, since weak extensionality will be the concept that corresponds best to the one in HoTT, we will from here on use the word 'extensional' simply as an abbreviation for 'weakly extensional'. This usage also fits far better with our connection to Carnap, for whom a modal logic in which identity between terms was always necessary would lose much of its point.

The connection through modal logic is promising not least because modal logic and topology have been known, since at least McKinsey and Tarski (1944), to go very well together. We want to use this fact to build up a sequence of logical relationships

$$\textit{Classical logic} \sim \textit{Modal logic} \sim \textit{Topological Semantics} \sim \textit{HoTT}$$

that allow us to compare notions—such as extensionality—in classical or modal logic directly to those in HoTT. Our goal in this section is to show that *the same* condition on the topological semantics makes both the classical-style modal logic we will describe here and HoTT extensional.

For our purposes it is easiest to limit ourselves to a *monadic* first-order modal logic $\mathcal{L}$, which we also take to have an existence predicate E.[22] Let a model $\mathfrak{M}$ of $\mathcal{L}$ be a triple $\langle \mathcal{O}, \mathsf{ref}, \Omega \rangle$ where

(*i*) $\mathcal{O}$ is a nonempty T_0 topological space[23] which we will call the *ontology*,

---

[22] Extensions to full first-order logic can be obtained by complicating our theory to map non-unary predicates to product spaces. Extensions to second-order modal logics would also be fairly easy to achieve. We will not go into details of how to carry out these here.

[23] Recall that a space is T_0 (alt. *Kolmogorov*) iff any two points in it are topologically distinguishable, i.e. if for any points a, b there is some open set X such that either $a \in X$ and $b \notin X$ or $a \notin X$ and $b \in X$. Given that we will interpret open sets as properties and points as referents of terms, this condition therefore comes out as a version of the identity of indiscernibles.

(ii) ref is a function from the individual constants of $\mathcal{L}$ to points in $\mathcal{O}$,

(iii) Ω is the set of all functions $\omega : Pred \rightarrow \wp\mathcal{O}$, where $Pred$ is the set of predicates in $\mathcal{L}$, such that for each $P \in Pred$ we have that

(a) $\omega(P)$ is open in $\mathcal{O}$ and

(b) $\omega(P) \subseteq \omega(\mathsf{E})$.

We call the elements of Ω *worlds* since they correspond to distributions of properties–including existence—among possible entities. The open subsets of Ω can be interpreted as *possible* properties, e.g. ones whose potential extensions are well-defined enough. The conditions on topological spaces then say that if P can have any number of extensions, then the union of these is also a possible extension, and if P can have two extensions X and Y, then $X \cap Y$ is also a possible extension. The function ref is known as the *reference* function. Note that it is not world relative, so individual constants are rigid. We will see that there are still models that are not weakly extensional.

Let a *variable assignment s* on $\mathfrak{M}$ be a function from variables of $\mathcal{L}$ to points in $\mathcal{O}$, and let a *valuation v* be the union of a variable assignment and the function ref. As usual, we write $v[c/x]$ for the valuation which is like v except for a replacement of the value at x with c.

The truth conditions are mostly straightforward:

$$\mathfrak{M}, \omega \vDash_v P(t) \Leftrightarrow v(t) \in \omega(P)$$
$$\mathfrak{M}, \omega \vDash_v \neg\varphi \Leftrightarrow \mathfrak{M}, \omega \nvDash_v \varphi$$
$$\mathfrak{M}, \omega \vDash_v \varphi \wedge \psi \Leftrightarrow \mathfrak{M}, \omega \vDash_v \varphi \text{ and } \mathfrak{M}, \omega \vDash_v \psi$$
$$\mathfrak{M}, \omega \vDash_v \varphi \vee \psi \Leftrightarrow \mathfrak{M}, \omega \vDash_v \varphi \text{ or } \mathfrak{M}, \omega \vDash_v \psi$$
$$\mathfrak{M}, \omega \vDash_v \varphi \rightarrow \psi \Leftrightarrow \mathfrak{M}, \omega \nvDash_v \varphi \text{ or } \mathfrak{M}, \omega \vDash_v \psi$$
$$\mathfrak{M}, \omega \vDash_v \forall x\varphi \Leftrightarrow \mathfrak{M}, \omega \vDash_{v[c/x]} \varphi \text{ for all } c \in \omega(\mathsf{E})$$
$$\mathfrak{M}, \omega \vDash_v \exists x\varphi \Leftrightarrow \mathfrak{M}, \omega \vDash_{v[c/x]} \varphi \text{ for some } c \in \omega(\mathsf{E})$$
$$\mathfrak{M}, \omega \vDash_v \Box\varphi \Leftrightarrow \mathfrak{M}, \omega' \vDash_v \varphi \text{ for all } \omega' \in \Omega$$
$$\mathfrak{M}, \omega \vDash_v \Diamond\varphi \Leftrightarrow \mathfrak{M}, \omega' \vDash_v \varphi \text{ for some } \omega' \in \Omega$$

From these is follows that atomic formulae with non-denoting terms are always false. We have followed Carnap in taking $\Box$ to express the global modality, i.e. truth in every world. Free variables range over all possibilities, but the quantifiers only range over that which exists.

To model identity we will need to introduce a new concept. Let the *entity space* $\mathcal{E}_\mathcal{O}(\omega)$ of a world ω in the ontology $\mathcal{O}$ be the coarsest topological space that results from adding all sets $\omega(P)$ to $\mathcal{O}$'s *closed* sets. In this space, every extension of a predicate is not only open, but clopen (i.e. both closed and open). This, in turn, entails that $\mathcal{E}_\mathcal{O}(\omega)$ will be partitioned into maximally connected subspaces, each of which has elements that have the same distribution of properties, as illustrated in Fig. 8.

Intuitively, one can think of the ontology $\mathcal{O}$ of a model as a vaguely defined "slum of possibilities" in Quine's sense (Quine 1961a), without firm identity conditions on
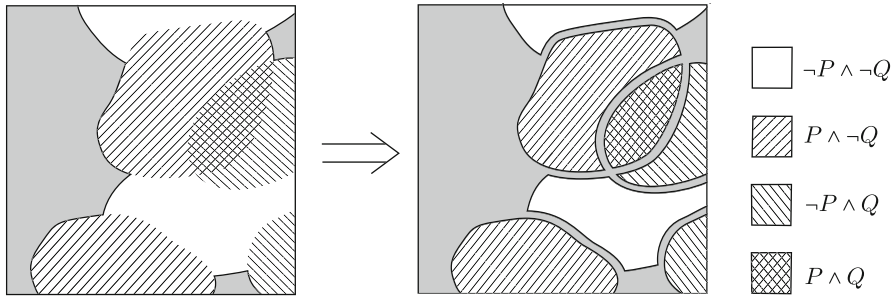
Fig. 8 Going to entity space

its constituents. A world ω is a specification of which chunks of possibility instantiate which properties, including existence, and ω's entity space is the existing part of possibility cut up into well-individuated entities, where the individuation is done based on the properties assigned by ω.

Write $a \sim_\omega b$ when there is a path from a to b in the space $\mathcal{E}_\mathcal{O}(\omega)$, and let the truth condition for identity formulae be

$$\mathfrak{M}, \omega \vDash_v s = t \text{ iff } v(s) \sim_\omega v(t)$$

From the fact that the clopen sets in a topological space are the unions of its (path-)connected components, together with our definition of $\mathcal{E}_\mathcal{O}(\omega)$, we get that if $a \sim_\omega b$, then $a \in \omega(P)$ iff $b \in \omega(P)$ for all $a, b \in \mathcal{O}$ and $P \in Pred$. By the truth condition we have imposed it follows that the substitutivity of identicals is valid:

$$\mathfrak{M}, \omega \vDash_v (s = t \wedge \varphi(s)) \to \varphi(t)$$

As is easily checked, we do not have weak extensionality. But while identity between terms thus does not guarantee strict identity between the terms' referents, we can show that *necessary* identity does:

Proposition 4 $\mathfrak{M}, \omega \vDash_v \Box s = t$ *iff* $v(s) = v(t)$.

Proof The right-to-left direction follows from the existence of the constant path $r : v(s) \to v(t)$ together with the fact that we have taken singular terms to be rigid. For the converse, assume that $v(s) \neq v(t)$. Because of the T_0 condition there must then be some open set S such that $v(s) \in S$ and $v(t) \notin S$, or vice versa. Since all open sets are possible properties, there is therefore some world ω' where $S = \omega'(\mathsf{E})$, and because $\mathcal{E}_\mathcal{O}(\omega)$ therefore must have S not only open but also closed, there can be no path between $v(s)$ and $v(t)$ in $\mathcal{E}_\mathcal{O}(\omega)$, so they cannot be equal. $\Box$

Corollary 1 $\mathfrak{M}, \omega \vDash_v \Box \forall x \, (P(x) \leftrightarrow Q(x))$ *iff* $\omega(P) = \omega(Q)$.

Corollary 2 $\mathfrak{M}, \omega \vDash_v \Box s = t$ *iff* $\mathfrak{M}, \omega \vDash_v \Box(\varphi(s) \leftrightarrow \varphi(t))$.

Call a model $\mathfrak{M}$ *extensional* iff $\mathfrak{M} \vDash_v s = t \to \Box s = t$, and *disconnected* iff its ontology is a totally path disconnected space, i.e. one in which the only path connected components are singletons. With these preliminary remarks in place, and the following lemma, we can state the main result of this section:

Lemma 1 *A model $\mathfrak{M}$ is extensional iff it is disconnected.*

Proof Assume that $\mathcal{O}$ is totally disconnected. Then, whenever we have that $v(s)$ is path connected to $v(t)$, we also have that $v(s) = v(t)$, which by Proposition 4 entails that $\mathfrak{M}, \omega \vDash_v s = t \to \Box s = t$. Conversely, assume that there are elements $a, b \in \mathcal{O}$ such that $a \neq b$ but $a \sim_\omega b$ for some world ω. Let v be a valuation such that $v(s) = a$ and $v(t) = b$. Then it follows, again by Proposition 4, that although we by assumption have $\mathfrak{M}, \omega \vDash_v s = t$, we cannot have $\mathfrak{M}, \omega \vDash_v \Box s = t$. $\Box$

Theorem 2 *Let A be a type in HoTT modeled by the ontology $\mathcal{O}$ of the model $\mathfrak{M}$. Then the following are equivalent:*

- *A has judgmentally unique identities.*
- *$\mathfrak{M}$ is extensional.*

Proof From the semantics of the propositional identity types we have that $p : \mathsf{Id}_A(a, b)$ iff p is a path $a \to b$. But if A is totally path disconnected there can be only one such path, namely the constant path. Conversely, assume that p is a path from a to b, and q is a path from b to a. If $a = b$ and p is not the constant path, there must be some $t \in [0, 1]$ such that $p(t) \neq a$, and thus paths $p_1 : a \to p(t)$ and $p_2 : p(t) \to a$, so it suffices to consider the case $a \neq b$. But if $a \neq b$ there will always be at least a countable infinity of other paths from a to b, e.g. $p_1 = p \circ q \circ p$, $p_2 = p \circ q \circ p \circ q \circ p$, etc. Thus $\mathsf{Id}_A(a, b)$ must have more than one element. The result follows by applying Proposition 1. $\Box$

We have thus finally managed to set up a connection between a version of HoTT's concept of extensionality and a more traditional interpretation thereof. That our conception is a version of Carnap's intension-extension distinction can be seen by, like Carnap, defining *having the same intension* as the truth of $\Box(\varphi \leftrightarrow \psi)$ for sentences, $\Box s = t$ for terms, and $\Box \forall x \, (P(x) \leftrightarrow Q(x))$ for predicates. It follows that the intension of a term correlates uniquely to a point, like in the HoTT interpretation of the previous section. The ontology itself, being a topological space X, naturally corresponds to a type in HoTT, and the extension of a predicate P, as a subset of X, can easily be interpreted as a specification of which fibres of a dependent type $P' : X \to \mathcal{U}$ are non-empty, i.e. those objects to which the property P' applies.[24]

To be sure, we have only described a very simple toy model—in fact a version of the setoid interpretation of types (Hofmann 1995, pp. 159–185) dressed up in topological clothing—and it is quite possible that there is a limit to how far it can be taken.

---

[24] It should perhaps be noted that the topological semantics given here is quite different from the one of Awodey and Kishida (2008), which is a more direct generalisation of McKinsey and Tarski (1944), and which, in distinction to our semantics, validates the Kripkean inference $s = t \vdash \Box s = t$. We also note that HoTT furthermore contains its own notion of *modality* (Univalent Foundations Program 2013, pp. 245–249), but this concept is again rather different from the one presented here.

Consequently, what we can say so far is merely that our modal logic and HoTT are similar by having similar semantics (i.e. ones in terms of topological spaces and paths), rather than that we can see them to be equivalent in some deeper sense. And in fact, there are important differences between HoTT and the logic we have discussed here which makes such deeper equivalence unlikely. In particular, HoTT is far more expressive: the logic we outlined, like all other classical logics, is not proof relevant. The concept of homotopy, which is arguably one of the most central ones in HoTT (as can be seen from e.g. the name!) does not appear in our semantics at all, so we have not been able to see what univalence would mean for us. While it would be possible, and perhaps also worthwhile, to extend the semantics sketched here in order to strengthen the similarities between classical modal logic and HoTT, we will stop here for now.

9 Discussion

If nothing else, the last sections should illustrate that the question of HoTT's intensionality is highly complex. While we have supplied a collection of logical models for making comparisons between different conceptions of intensionality and extensionality, there seems to be fairly little in common between these. As was announced at the beginning, these are at most some preliminary investigations. With that in mind, we will still attempt to say something about philosophical questions that may have arisen.

As one of the primary inventions of HoTT is its interpretation of identity, one of these philosophical questions would be if that concept is still *identity* in a sufficiently similar sense to an informal notion to be useful for philosophy. Quine famously argued that *all* differences in logical systems were down to different meanings assigned to the systems' sentences and logical constants (Quine 1986). Thus, for example intuitionistic logic is about informal provability rather than truth or falsity, i.e. '*p*' in intuitionistic logic means '*p* is provable' while '*p*' in classical logic means '*p* is true'. The intuitionistic conjunction $p \lor q$ means '*p* or *q* are provable', while the classical means '*p* or *q* is true'. If the classical-logical meaning were the one we commonly use and mean by 'identity', this would entail that the concept of identity used in HoTT is simply not *real* identity, but something else.

A related question arises from the indiscernibility of identicals as expressed in HoTT. Under the interpretation of properties as dependent types, this does indeed say that identical objects have the same properties. Insofar as properties *are* dependent types, such a characterisation is definitely closer to Leibniz's original version of the principle than the contemporary one about predicates, formulae and satisfaction. But are propositions *really* types, and are properties really type families?

The present Quinean argument against interpreting propositional equality as identity, explicitly spelled out, would be that unless a relation guarantees substitutivity *salva veritate*, it is not identity. Such arguments have previously been used e.g. to say that identity over time is not true identity (in support of 4-dimensionalism), and that identity across worlds is not true identity (in support of counterpart semantics).

We should note that for judgmental equality, full substitutivity holds by the inference rules assumed for this judgment type (Univalent Foundations Program 2013, p. 427). It is propositional identity that poses problems, and the Quinean argument would be

in support of the thesis that $p \; : \; \mathsf{Id}_X(a, b)$ does not say, nor even entail, that a and b are identical.

Now, the existence of a propositional identity $p \; : \; \mathsf{Id}_X(a, b)$ does not show that the structures defined in terms of a are strictly the same as those defined in terms of b, but rather that we can *transport*, or perhaps *translate*, structures defined in terms of a to ones defined in terms of b in an appropriate way.[25] An inference such as

$$\frac{P \; : \; X \to \mathcal{U}, \; a \; : \; X \vdash c \; : \; P(a) \qquad P \; : \; X \to \mathcal{U}, \; a \; : \; X, \; b \; : \; X \vdash p \; : \; \mathsf{Id}_X(a, b)}{P \; : \; X \to \mathcal{U}, \; b \; : \; X \vdash c \; : \; P(b)}$$

cannot be valid since $P(a)$ may be a different type than $P(b)$, which means that they can have no elements in common. This does not mean that they cannot be propositionally equal, and indeed the univalence axiom guarantees that they are. We may refer to the kind of substitutivity that allows the replacement of any term with another equal to it as *syntactic* since it does not rely on any specific interpretation of types but rather means that the terms on both sides can be rewritten, using syntactic rules, so that they come out as literally the same.

Since judgmental equality allows full syntactic substitutivity, it may at first seem like this is the concept that has the most in common with equality in classical logic. But judgmental equality also lacks many features that the '=' predicate has. Foremost among these is its inability to be embedded in terms or types, which makes it far less versatile, and rather indicates that although it may look like classical identity, it does not do the job that this concept does. Propositional identity, in contrast, can do more of the things we use classical identity for and also, according to the interpretation given in Sect. 7, signifies sameness of extension. Insofar as we take the classical identity $a = b$ to also mean that a and b have the same extension, we therefore have an argument for the primacy of propositional identity when comparing HoTT to classical logic.

Furthermore, although propositional identity does not give full syntactic sub-stitutability on the syntactic level, the Curry–Howard interpretation implies that it embodies another version of Leibniz's law, as shown in the fact that it satisfies the indiscernibility of identicals. Unlike syntactic substitutability, this *semantic* substitu-tivity does not guarantee that *symbols* can be replaced, but only that the truth values of propositions under the Curry–Howard correspondence are the same. It thus relies on this specific interpretation of a type's reference or extension.

I think complications such as these should be taken as an indication that the question of which is the "real" identity in HoTT may be pointless. Informal language is vague and mutable, and to draw conclusions about logical concepts from it is hazardous, to say the least. Especially when it comes to a locution such as *is the same as*, the possible meanings seem almost inexhaustible, and there are good reasons to hold that

---

[25] The same distinction can be framed as one of *invariance* vs. *covariance*. A feature is *invariant* under a transformation if it is the same before and after the transformation is applied. It is *covariant* if, roughly, applying the transformation gives a *corresponding* feature in the transformation's image. If I take a perfect photograph of another photograph, the image on the first photograph will not be strictly the same as the second, but it will picture the same thing, and it will *look* the same in every way, so the pictures *covary* under the transformation *taking a perfect photograph*, but they are not *invariant*. See also Rodin (2013, pp. 253–255).

the ones embodied in classical predicate logic are some of the lesser used ones. This holds even in mathematics, for whose purpose that logic was created: a mathematical uniqueness proof, for example, only very rarely shows that something is unique in the sense of the '=' of predicate logic.

Luckily, finding the "true" concept of identity is not what we were after here. What is important is to what degree the concept can be used for the job it has been introduced to do. Sometimes parts of this job will involve replacing informal concepts in certain contexts, but sometimes it consists in purely technical applications: one may compare with the concept of 'energy' in physics, which has many uses in the subject's theories, but very few connections to pretheoretical notions with the same name.[26]

The primary subjects to which HoTT is to be applied are mathematics and logic, and here there can be few doubts that its conceptions of identity, property and proposition *can* do the work we want them to since they all reduce to traditional notions when we add axioms such uniqueness of identities or the excluded third. Arguing that they are *more* useful than traditional notions requires showing how they can be used to solve problems or explain relationships that the traditional notions cannot. So far, it seems quite likely that this is indeed the case. But even if this should not hold, nothing forces us to have *one* notion of identity: formal systems are not themselves weapons of imperialism, although some of their advocates may occasionally try to use them as such.

So much for mathematics and logic. For philosophers, it would be particularly interesting to see how the homotopy-theoretical concepts of identity, property and proposition can be useful for our subject as well; according to the pragmatic conception of formal systems advocated here this will indeed be the primary motivation for their acceptance. Although I will not try to take up Harris's suggestion, in the quote we started with, that the philosophy behind "identity politics" will be challenged by HoTT, I still want to end this paper by sketching a few ideas regarding possible applications HoTT to important philosophical problems.

Because of intensionality's strong connection to identity, we have placed a special focus on this concept. Classical-logical identity, under the standard Tarskian semantics, is a rather trivial affair; in first-order logic it is arguably useful for little more than counting. In contrast, the propositional identities of HoTT allow one to model a lot of structural relationships. For example, it is possible that one could use them to give a theory of family resemblances in Wittgenstein's sense. This would be done by letting a path $a \rightarrow b$ represent *exact similarity in a respect* between a and b, and the composition of paths $f : a \rightarrow b$ and $g : b \rightarrow c$ a family resemblance. For example, f might model sameness of eye colour, and g sameness of hair colour.

This could, in itself, easily lead to trivialities since such chains usually can be set up between any two entities. But identity in HoTT is not a yes or no question, but a higher-order structure. That some family resemblance holds between any two entities

---

[26] The connection to Carnap's notion of explication should be obvious (Carnap 1950, ch. 1). In our case, we have a pretheoretical concept (identity) or type of phrase ('a is the same as b') which is given different explications: one in terms of classical logic, and one in terms of the propositional identity of HoTT. We should not expect one of these to be more "correct" than the other, and neither should we expect one of them to be more useful than the other in *every* circumstance.

does not imply that identity is trivial, since specific identities are what count. In terms of homotopy theory, the fact that two spaces are both path-connected does not imply that they are even homotopically the same: path-connectedness takes into account only the one-dimensional structure of a space. Similarly, higher-order identities between identities may be used in order to represent exact similarities between similarities, which could result in a way to describe properties in a purely structural manner.

This connection to structuralism points rather naturally to another possible application. What if one were to use HoTT to build a version of Carnap's *Aufbau* (2003)? One of Carnap's main goals was a structuralist one, namely to show how scientific knowledge can be represented purely in terms of logical structure, without having to assume prior interpretations of autopsychological predicates. As he knew, what has become known as the Newman objection makes this impossible in classical logic, since versions of this language without interpreted predicates—even in higher-order logic—can only distinguish between cardinalities. HoTT, on the other hand, can be used to represent topological spaces in all their (homotopy-invariant) complexity. At least the homotopy invariant topological structure of the world should therefore be possible to capture using only the logical machinery of HoTT.[27]

HoTT also has an additional advantage as a language for projects such as the *Aufbau*: its amenability to computer implementation. Indeed, one of the motivating factors in the design of UF has been to make possible a fully formalised, completely rigorous approach to mathematics. To this purpose most of the mathematics developed in HoTT has been written in proof assistants such as *Coq* and *Agda*. This allows philosophers to build on these results and thereby gain access to an ever growing supply of mathematical structures. Developing an *Aufbau*-style project using HoTT, using libraries of such and further results, would not only ensure a greater level of exactness than was available to Carnap, but also aid in making the project possible to carry out as a collaborative effort. Computer-implemented proofs and definitions can be efficiently managed via version control systems, much like large programming projects. Such systems can be used to keep every participant's assumptions explicit, and to make it far clearer which parts of different philosophers' works are compatible, which build upon one another, and which are contradictory.

Moving on to HoTT's representation of properties as type families, we may note that such a representation could be useful for philosophical notions of properties as well. Since a type family $P : A \to \mathcal{U}$, under the topological interpretation, is a fibration which assigns a (possibly empty) topological space to each point of a type A, these spaces can be seen as *determinables*, e.g. mass, temperature, velocity etc. A dependent function $f : \prod_{x:A} P(x)$ can then be interpreted as a specification of the precise values (e.g. $2\,\text{kg}$, $280\,\text{K}$, $3\mathbf{e}_x + 2\mathbf{e}_y + 1\mathbf{e}_z$ m/s) these determinables take.

---

[27] As an example of where concepts from HoTT may help, consider Carnap's characterisation of the visual sense as the unique five-dimensional sense (Carnap 2003, p. 183). This is not really an optimal description, since merely assuming a five-dimensional space does not impose any limitation on which points have to be defined. For example, Carnap's conception is compatible with a visual field where some points have no colour and others have several. A more fitting representation of the visual sense would be as a fibration, where the base space represents the visual field, while the fibres contain the possible colours each point in this field can have. In type-theoretical terms, the contents of the visual sense would be a dependent function. Topologically, they would be a section.

This connects nicely to how properties of space-time points are handled in general relativity and in gauge theories such as the standard model, which may be useful if one wishes to develop metaphysical frameworks that are applicable to contemporary physics as well.

As a final example, also from metaphysics, the Curry–Howard interpretation of propositions can be connected to truthmaker theory by taking each element of a proposition to be a truthmaker for it. This is quite natural in view of the strong similarities between truthmaker theory and the Brouwer–Heyting–Kolmogorov interpretation of intuitionistic logic, of which the Curry–Howard correspondence can be seen as a successor. But the topological structure of HoTT gives us more than that: it allows us to discuss relationships among truthmakers as well, thus potentially enhancing truthmaker theory's use as a tool for inferring ontological structure from our theories.[28]

For applications such as these it may also be appropriate to consider alternatives to univalence. Univalence corresponds to a structuralist principle of the form

$$homotopy\ equivalence \Rightarrow identity$$

but many other versions of the antecedent side of this conditional imaginable. *Isomorphism*, for example, is not the same as homotopy equivalence except in certain special categories such as $\mathcal{T}$. A more direct *isomorphism* $\Rightarrow$ *identity* principle would supposedly correspond to allowing the construction of a canonical identity proof $i\ :\ \mathsf{Id}_\mathcal{U}(X, Y)$ from *judgmental* identities $g \circ f \equiv_{df.} 1_X\ :\ \mathcal{U}$ and $f \circ g \equiv_{df.} 1_Y\ :\ \mathcal{U}$. Because it has a stronger antecedent, such a principle would be weaker than univalence.

Another possibility would be to insert more structure into the topological semantics in order to increase its expressiveness. Adding differential information to make types correspond to differential manifolds rather than bare topological spaces would make a physically correct treatment of time and space possible. In such a framework, the appropriate antecedent of the structuralist principle would be diffeomorphism. Since any differentiable manifold is a topological space, and any differential map is continuous, such spaces ought to still validate the inference rules of HoTT, with the exception of univalence. The interesting question concerns which further logical operations or principles might be possible to introduce based on the additional structure available.

All of these applications are admittedly very speculative. We do not know if any of them works until someone has attempted to work out their details, although I hope that they illustrate some of the possible uses homotopy type theory and univalent foundations, and their approach to identity, may have for philosophy. But the existence of this system also helps remind us of something even more important: the set-theoretic and classical-logical foundations that held such sway over twentieth century philosophy, and which continue to shape so much of our thinking, are only one possibility among many. Extensionality, intensionality, identity, proposition, and property are all

---

[28] For a discussion on the use of truthmaker theory for the purpose of inferring ontology from logical structure, see Angere (2015).

concepts for which other, no less rigorous or coherent, frameworks are available for our use.[29]

References

Angere, S. (2015). The logical structure of truthmaking. *Journal of Philosophical Logic*, *44*(4), 351–374.

Awodey, S. (2014). Structuralism, invariance, and univalence. *Philosophia Mathematica*, *22*(1), 1–11.

Awodey, S., & Kishida, K. (2008). Topology and modality: The topological interpretation of first-order modal logic. *Review of Symbolic Logic*, *1*, 146–166.

Awodey, S., & Warren, M. (2009). Homotopy theoretic models of identity types. *Mathematical Proceedings of the Cambridge Philosophical Society*, *146*(1), 45–55.

Barker-Plummer, D., Barwise, J., & Etchemendy, J. (2011). *Language, proof, and logic* (2nd ed.). Chicago: University of Chicago Press.

Brown, R. (2006). *Topology and groupoids*. Charleston: BookSurge Publishing.

Carnap, R. (1937). The logical syntax of language. Kegan Paul.

Carnap, R. (1950). *Logical foundations of probability*. Chicago: University of Chicago Press.

Carnap, R. (1956). *Meaning and necessity* (2nd ed.). Chicago: University of Chicago Press.

Carnap, R. (2003). *The logical structure of the world and pseudoproblems in philosophy*. Open Court Publishing.

Feferman, S. (1985). Intensionality in mathematics. *Journal of Philosophical Logic*, *14*(1), 41–55.

Frege, G. (2013). Basic Laws of Arithmetic (P. Ebert, M. Rossberg, C. Wright, & A. Philip, Trans.). Oxford: Oxford University Press.

Gödel, K. (1958). Über eine bisher noch nicht benutzte erweiterung des finiten standpunktes. *Dialectica*, pp. 280–287.

Harris, M. (2015). *Mathematics without apologies*. Princeton: Princeton University Press.

Hatcher, A. (2006). *Algebraic topology*. Cambridge: Cambridge University Press.

Hofmann, M. (1995). *Extensional concepts in intensional type theory*. PhD thesis, University of Edinburgh.

Howard, W. A. (1980). The formulae-as-types notion of construction. In H. B. Curry (Ed.), *Essays on combinatory logic, lambda calculus and formalism* (pp. 479–490). New York: Academic Press.

Lewis, D. (1986). *On the plurality of worlds*. New York: Wiley.

Lurie, J. (2009). *Higher topos theory*. Princeton: Princeton University Press.

Marquis, J.-P. (2013). Mathematical forms and forms of mathematics: Leaving the shores of extensional mathematics. *Synthese*, *190*(12), 2141–2164.

Martin-Löf, P. (1975). An intuitionistic theory of types: Predicative part. In *Logic Colloquium '73 (Bristol 73). Studies in logic and the foundations of mathematics*. North-Holland: Amsterdam, pp 73–118.

Martin-Löf, P. (1984). *Intuitionistic type theory*. Berkeley: Bibliopolis.

Martin-Löf, P. (1998). An intuitionistic theory of types. In G. Sambin & J. M. Smith (Eds.), *Twenty-five years of constructive type theory*. Oxford: Oxford University Press.

Martin-Löf, P. (2015). Personal communication, Stockholm, December 2015.

May, J. P. (1999). A concise course in algebraic topology. University of Chicago Press.

McGinn, C. (2000). *Logical properties*. Oxford: Oxford University Press.

McKinsey, J. C. C., & Tarski, A. (1944). The algebra of topology. *Annals of Mathematics*, *45*, 141–191.

Nordström, B., Petersson, K., & Smith, J. M. (1990). *Programming in Martin-Löf's type theory*. Oxford: Oxford University Press.

Quillen, D. G. (1967). *Homotopical algebra*. Berlin: Springer.

Quine, W. V. (1951). Two dogmas of empiricism. The Philosophical Review, 20–43.

Quine, W. V. (1961a). Reference and modality. In *From a logical point of view* (2nd ed.). Harper & Row Publishers, pp. 139–157.

Quine, W. V. (1961b). On what there is. In *From a logical point of view* (2nd ed.). Harper & Row Publishers. pp. 1–19.

Quine, W. (1986). *Philosophy of logic* (2nd ed.). Cambridge: Harvard University Press.

Rodin, A. (2013). *Axiomatic method and category theory*. Berlin: Springer.

Tait, W. (1967). Intensional interpretations of functionals of finite type. *Journal of Symbolic Logic*, *32*(2), 198–212.

The Univalent Foundations Program. (2013). *Homotopy type theory: Univalent foundations of mathematics*. http://homotopytypetheory.org/book. Institute for Advanced Study.

Voevodsky, V. (2010). Univalent foundations project. NSF grant application. http://www.math.ias.edu/~vladimir/Site3/Univalent_Foundations_files/univalent_foundations_project.pdf.