Univalence as a Principle of Logic

Steve Awodey

October 2016

Abstract

It is sometimes convenient or useful in mathematics to treat isomorphic structures as the same. The recently proposed Univalence Axiom for the foundations of mathematics elevates this idea to a foundational principle in the setting of Homotopy Type Theory. It states, roughly, that isomorphic structures can be identified. We explore the motivations and consequences, both mathematical and philosophical, of making such a new logical postulate.

Introduction

This talk is about the Univalence Axiom recently proposed by Voevodsky as a new principle in the foundations of mathematics. It is formulated within a new system of foundations called Homotopy Type Theory, and, roughly speaking, permits isomorphic structures to be identified. The resulting picture of the mathematical universe is rather different from that corresponding to conventional set theoretic foundations. It is not entirely alien, however, and it even has many interesting connections to some traditional philosophical issues, and that is what I hope to explain in this talk.

There are three main themes:

1. The advance in computer technology has given rise to the possibility of a new, computer-assisted, formalization of mathematics, breathing new life into the old Logicist program of Frege and Russell.

2. One of the main tools for such formalizations is MLTT, which has recently been connected to homotopy theory (of all things). But why should these remote fields be related at all? As it turns out, there are good reasons why logic and homotopy are closely connected, and they have to do with some of the oldest problems of logic and analytic philosophy.

3. In Homotopy Type Theory, the Univalence Axiom implies that isomorphic structures can be identified, in a certain, precise sense that I will explain. This sounds like the sort of thing that one hears from structuralists in the philosophy of mathematics, and

indeed, the development of mathematics within this new system does have a distinctly structural character. But the Univalence Axiom is not just a practical device for implementing structuralism within a formal system of foundations: it has a purely logical justification as well.

1 Why New Foundations for Mathematics?

There is at present a serious mismatch between the everyday practice of mathematics and the official foundations of mathematics represented by ZFC. This is evident in the fact that working mathematicians do not really pay any attention to foundations, and even violate their restrictions when it is convenient. An example is the common practice of identifying isomorphic objects with devices like "abuse of notation", which is a harmless move, but one that is outside the scope of what can actually be represented in the foundational system. The fact that this is a sort of joke, and is often accompanied by a shrug, is just evidence that the mathematician is aware of the mismatch, but doesn't really care.

This particular tension between official foundations and everyday practice really goes back to Bourbaki's "structuralist" approach, according to which the real objects of mathematics are not specific number systems but "structures", like groups, modules, spaces, etc., which can be described by operations and relations in a hierarchy of sets. The official formulation, however, was formulated in simple type theory (not axiomatic set theory) in order to be able to give a natural associated notion of "isomorphism" of structures, which was needed in order to determine which statements and properties are really relevant — namely, those that respect the isomorphisms. Thus topology is the study of spaces "up to homeomorphism", group theory of groups "up to group isomorphism", etc. — as opposed to actual sets equipped with a system of open sets, a multiplication, or whatever. In particular, e.g. the real numbers are not regarded as a specific set of sets (Dedekind cuts, say), but rather an "abstract" object, given by a particular presentation, but invariant under all isomorphisms between other possible presentations. This general methodology and point of view has now become quite pervasive in modern mathematics

Philosophers of mathematics have certainly noticed this trend, and they have developed various forms of "structuralism" based on the tools available to them: first-order logic, modal logic, etc. But they have been hampered in understanding structuralism by the absence of a comprehensive system of foundations that fully incorporates a modern, structuralist viewpoint, in the way that set theory does for non-structuralist foundations. Some use of category theory has been made, to be sure, and this is certainly a step in the right direction. But it's fair to say that no satisfactory, comprehensive, "structuralist" system of foundations has yet been formulated, either by philosophers, or by mathematicians themselves.

This situation was tolerable — at least for the mathematicians — until quite recently, when the increasing use of computational tools like computerized proof assistants and formal verification systems required a higher degree of precision and rigor in proofs than is usual in informal mathematics. [Specifically, (1) for big proofs, and (2) for hard proofs.] This has led to a reexamination of the foundations of mathematics by working mathematicians, logicians, and computer scientists involved in building such verification systems. This is not unlike previous revolutions in foundations, which were also precipitated by a need for greater practical rigor, made evident by specific mathematical advances: e.g. the introduction of ϵ-δ methods into differential calculus, and the invention of axiomatic set theory, to shore up the foundations of analysis and arithmetic.

The current advance requires foundations that are not only fully rigorous in principle, like set theory, but also natural enough to actually be used in practice, in order to formalize and verify proofs of theorems as they are produced. [Constructivity has also become an important feature in such computational proof systems, because of its close connection to computability, although the exact requirements there remain unclear.]

To be sure, there are already modern systems of proof verification based on set theory (Mizar) and HOL (HOL-light). But these rely on the usual formalization of mathematics in such conventional systems of foundations, and this often involves unnatural coding and inconvenient forms of reasoning.

The Coq proof system, by contrast, is based on a modern system of type theory, which turns out to allow for formalizations that are closer to everyday mathematical reasoning. It has recently been used e.g. to formalize several significant mathematical results — the Feit-Thompson Odd Order theorem, for instance, which is a crucial step in the celebrated classification of finite simple groups. For reasons that I will explain next, this particular system also admits a new interpretation that brings it even closer to everyday mathematical reasoning, and closes the gap between formal foundations and structuralist practice.

2 Homotopy Type Theory

The system of intensional Martin-Löf type theory that underlies Coq was originally developed as a logical basis for constructive mathematics. It is motivated by strict principles that give it an internal coherence and unity not shared by systems based on a more practical desire to "capture as much math as possible". This distinctly constructive character also makes it well-suited to use in proof assistants and computerized mathematics. But its peculiar *intensionality* has until recently been a source of considerable difficulty in using it to formalize classical mathematics.

In a bit more detail, the equality relation in particular was not well-understood, even by the experts. In MLTT, one can have a proof p that $x = y$, say,

$$p : x = y$$

and one can also have another such proof,

$$q : x = y.$$

And now the question can arise whether $p = q$!

$$? : p = q$$

And if that's not bad enough, one may again have different proofs

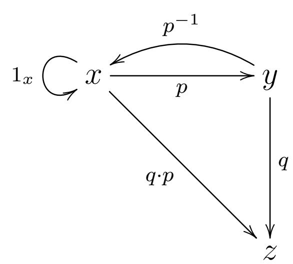
$$\alpha, \beta : p = q\,,$$

and then ask whether $\alpha = \beta$, and so on.

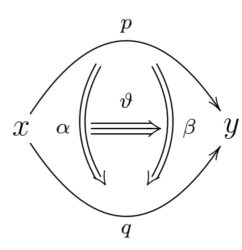
You see, in this system (as is often the case in constructive mathematics), proofs themselves are first-class mathematical objects, and they can enter into further propositions and constructions — a point of view that is somewhat unusual in classical mathematics. This can be very useful constructively, however, as when from a constructive proof that for all x there is a y such that blah, one can extract a function taking the x's to the y's — this is the sort of thing that constructive mathematics is good at. But the proliferation of proofs of identities, and especially the "higher" proofs of identities of identity proofs, etc., were until very recently regarded as simply junk in the system.

The first indication that these proofs of identities are actually significant was the observation (by Hofmann and Streicher in 1998) that the identity proofs in this system endow each type with the structure of a groupoid — in much the same way that the identity relation in first-order logic determines an equivalence relation on the set of terms of a first-order theory. (Equivalence relations are special groupoids.)

A groupoid is a generalized group, with the multiplication being only a partial operation — or equivalently, a category in which every arrow has an inverse.



In fact, the types in ITT are actually endowed with a structure more like what is called an ∞-groupoid — a concept also occurring in higher category theory and topology. An infinity groupoid is like a groupoid, but with higher dimensions of arrows. I will give a bit more detail about this below. The higher identity proofs, it turns out, are much like the higher arrows in such an ∞-groupoid.



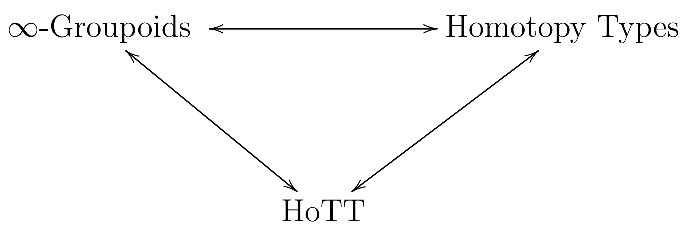
Now, one of the last contributions of the great mathematician Alexander Grothendieck — who was responsible for some of the most fundamental advances in 20th-century mathematics

— was the idea that ∞-groupoids can be used to represent homotopy types of spaces. This is now known as the *Homotopy Hypothesis*.

$$\infty\text{-Groupoids} \longleftrightarrow \text{Homotopy Types}$$

A homotopy type is an equivalence class of spaces under the relation of homotopy equivalence — roughly, one space can be continuousy deformed into the other without tearing, like the old example of the coffee cup and the donut. These have the same homotopy type, but the hollow soccer ball and the solid bowling ball have different types, both from each other and from the donut.

Now when the same concept arises independently in three different branches of mathematics — here the notion of ∞-groupoid in logic, algebra, and topology — well, that is worthy of note. The basic idea of homotopy type theory was to use this connection between identity proofs, ∞-groupoids, and homotopy, to relate type theory to higher category theory and homotopy. Once you understand the connection between type theory and groupoids, and that between groupoids and homotopy, there is then a fairly obvious way of relating type theory directly to homotopy. And in fact, it turns out that the homotopical interpretation "closes the triangle" in a particularly neat way, because it also extends the already well-known topological interpretation of the (simply-typed) λ-calculus.



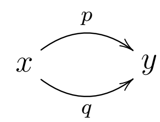
Without going into too much detail, and for those of you familiar with the λ-calculus, the types in that theory can be represented as spaces, and the terms as continuous mappings between them, with the basic constructions of product $A \times B$ and function type $A \to B$ being modelled by the corresponding constructions of spaces. Under that interpretation, continuity becomes a mathematical model of computibility, since computability was long ago shown to be equivalent to λ-definability in the work of Kleene, Rosser, and Church. The λ-calculus itself was originally formulated by Church, but its main idea goes all the way back to Frege.

Now MLTT is an extension of λ-calculus — if a very broad one — and the homotopy interpretation of MLTT agrees with this topological interpretation on the subsystem consisting of λ-calculus. The main new idea of the extension is that of interpreting an identity $p : x = y$ as a "continuous path" from the point x to the point y, in the space interpreting the type of x and y.

$$p : x = y \qquad\qquad x \xrightarrow{p} y$$

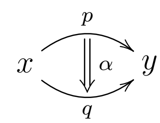
Another such identity $q : x = y$ is then just another path from x to y,

$$q : x = y$$



And a higher identity $\alpha : p = q$ is then a continuous deformation of the path p into the path q, i.e. a homotopy.

$$\alpha : p = q$$



From this it follows almost immediately that a *dependent* type (one of the essential extensions of λ-calculus in MLTT) must be interpreted by what the topologist calls a "fibration", and that two functions are equal just if they are homotopic.

This homotopical interpretation of intensional MLTT (originally due to Awodey and Warren) then suggests that the important notion of "identity" for *types* is not really equality, or even isomorphism, but rather *homotopy equivalence*,

$$A \simeq B$$

the concept developed by topologists for studying spaces in the setting of homotopy theory. And indeed, celebrated homotopy theorist Vladimir Voevodsky came up with exactly this idea, after having independently arrived at his own version of the homotopy interpretation while learning to use the Coq proof assistant. It was in that setting that he formulated the Univalence Axiom, which implies that two types that are homotopy equivalent can be identified,

$$A \simeq B \to A = B$$

More precisely, it says that the type of all equivalences between two types is itself equivalent to that of all their identifications:

$$(A \simeq B) \simeq (A = B) \tag{UA}$$

Remember that in type theory, proofs and propositions are first class objects, so the "propositions" $A \simeq B$ and $A = B$ are themselves types that can figure into further constructions and propositions, such as UA itself. Recalling that in homotopy type theory the identifications $p : A = B$ are understood as paths, UA then says that, in the space $\mathcal{U}$ of all types, the continuous paths between any two types are the equivalences between them.

The UA thus specifies a previously underdetermined notion in type theory, namely the identity relation on the universe of all types. I hasten to add that it should be thought of as *expanding* the notion of equality to become that of equivalence, rather than collapsing that of equivalence to "literal" identity. That is to say, UA does *not* say that all equivalent types are strictly the same. Exactly how this is possible formally is something I will return to later — it has to do with the intensionality of the type theory.

Informally, this move is much the same as the "structural" methodology that we have already described as being used elsewhere in modern mathematics since Burbaki: we can have different "presentations" of the same structure by different structured objects. We then work with the presentations in a way that is invariant under the relevant notion of equivalence, in order to reason indirectly about the structures they present. In homotopy type theory, one can work with types and terms up to equivalence, without formally identifying them. This has the advantage that one can still make use of the presentations, which may have better computational and combinatorial properties than would the abstract objects resulting from actually "making the identification".

Let us consider the example of intensional versus extensional type theory. The extensional theory has an apparently "stronger" notion of equality, because it permits one to simply substitute equals for equals in all contexts. In the intensional system, by contrast, one can have $a = b$ and a statement $\Phi(a)$ and yet not have $\Phi(b)$.

$$\frac{a = b \qquad \Phi(a)}{\Phi(b)}\ ?$$

So there are some contexts Φ that are "intensional", in the sense that they do not strictly respect the system's notion of equality. (A similar thing happens in quantified modal logic, which is therefore also said to be intensional.)

This seemingly strange behavior of equality in intensional type theory is however quite useful, because it permits the system to have the important property of decidability: every term can be reduced to a unique normal form. This crucial property makes the intensional system superior for the purpose of formal verification systems. It is lost when the stronger rules for equality are added to form the extensional system.

We have now considered the origin of the homotopy interpretation of type theory, and the position of the univalence axiom within it, from a mathematical point of view. These origins were mainly practical, having to do with allowing more efficient computer formalizations, which then gave rise to some unexpected mathematical connections. But what about the philosophical justification of this new interpretation? Is there any prinicipled reason why a system of foundations of mathematics should involve the notion of homotopy? Even more fundamentally, why should a system of logic have anything to do with ∞-groupoids?

It turns out that there is indeed a more basic reason for these connections, which is related to some of the oldest and most difficult problems in the philosophy of logic and mathematics — but in order to explain this, we first have to review a bit of the early development of modern logic.

3 Logical Foundations Revisited

The goal of Frege-Russell Logicism, recall, was to reduce mathematics to a basic system of formal logic with no primitive notions apart from relations and some basic logical operations like negation, implication, and quantification. The reduction was lengthy and laborious, with hundreds of pages of formal deduction required just to show that $1 + 1 = 2$.

Moreover, there was the philosophical question of what made the chosen starting point "logical" in the first place. Why not simply include, say, the natural numbers as "logic", as was later proposed by Wittgenstein and Carnap as a solution to certain technical difficulties in the proposed reduction.

An alternative, less laborious, approach was represented by Hilbert's Formalism — aka the axiomatic method — which separated out the "logical" part of the sytem as what we now call FOL, and then used it to axiomatize various branches of mathematics individually, e.g. the natural or real numbers, or Euclidean geometry. This, Russell quipped, had all the advantages of theft over honest toil. But in fairness, without a principled justification for all that toil, it was hard to convince anyone that it was really worth the extra effort.

In the extreme, Zermelo's axiomatic set theory could be used as a framework in which to formalize other mathematical theories, in a sort of "formalist reduction" of mathematics to axiomatic set theory. One problem with this approach is that it leaves the issue of the relation between logic and mathematics unresolved, since set theory itself is still being understood axiomatically. Some philosophers, notably Gödel, have tried to fill this gap by giving conceptual justifications for the axioms of ZFC. The logical character of axiomatic FOL, for its part, is now taken more or less for granted — but of course, one can still reasonably ask what makes, say, negation or quantification "logic", but not the Peano Postulates or, for that matter, the Law of Gravitation.

In the last 100 years, type theory has made some signficant advances, and modern systems allow for a much smoother development than did the old systems of Frege and Russell. As already mentioned, many of the technical improvements have been spurred by the use of modern computers to implement these systems. In these computer implementations, the logical formalization of mathematics is finally becoming not just theoretically possible, but actually feasible for everyday mathematical use. What was once too lengthy or tedious to be done by hand can now be taken over by a computer, and the user can then focus on the higher-level organization of the proof — just as in informal mathematics. The result is then fully formalized mathematics in a logical system that runs like a computer program, and that can be mechanically checked for correctness, then stored alongside previously checked results in an electronic archive for further use.

Eventually, this new kind of computer-assisted formalization could become a practical tool for the working mathematician — just as originally envisaged by Frege, who compared the invention of his Begriffsschrift with that of the microscope. Indeed, the original idea of logical foundations of mathematics fits so well with this use of computer proof systems that future historians of mathematics may well wonder how Frege and Russell could ever have come up with the idea of logical foundations before there were any computers to run them on!

Unlike the purely axiomatic approaches, the modern system of type theory can still claim to be a system of "pure logic" — at least as plausibly as Frege's system could. Indeed, several leading type theorists have developed sophisticated justifications of their systems, mainly on the basis of so-called "meaning explanations" of the type-theoretic constructions, inspired by a combination of the Brower-Heyting-Kolmogorov functional interpretation of constructive logic and Dummet's verificationist reading of intuitionist logic. But before turning to the issue of what makes type theory "logical", I need one more fact about the homotopical character of the particular system of MLTT.

As we have seen, under the homotopical interpretation an equality is understood as a path $p : a = b$. Although the system is intensional in the sense that $p : a = b$ does not warrant substitutions of b for a everywhere, it is still the case that every construction *respects* equality in the following weakener sense: given any path $p : a = b$ and any "proof" $t : P(a)$ of some property of a, although we do not immediately have $t : P(b)$, we instead have a systematic way of "transporting" t along the path p, which we may write p_*t, and for that object we then have $p_*t : P(b)$.

$$\frac{p : a = b \qquad t : P(a)}{p_*t : P(b)}$$

So given a proof of $a = b$ and one of $P(a)$, we can combine them to get a proof of $P(b)$.

Now, in just the same sense as this, every construction in the system can be seen to respect the relation of *equivalence* of types. In fact, we have the following:

Principle of Invariance Every property or construction in HoTT is "homotopy invariant", in the sense that it respects equivalence of types:

$$\frac{e : A \simeq B \qquad t : P(A)}{e^*t : P(B)}$$

Thus given any equivalence of types $e : A \simeq B$ and a term $t : P(A)$ of any type $P(A)$ constructed from A, there is an associated term $e^*t : P(B)$ of the corresponding construction on B.

A practical consequence of this invariance is that the univalence axiom is consistent! Indeed, if any predicate P of types were able to distinguish between equivalent types $A \simeq B$, then we could not have univalence, since by univalence equivalent types can always be identified by a path $p : A = B$, and so any $t : P(A)$ can be transported along p to give $p_*t : P(B)$.

Indeed, Univalence itself can be seen to be logically equivalent to the Invariance Principle together with a certain "computation rule". For, assuming Invariance, we can transport the property "$A = (-)$" along any equivalence $e : A \simeq B$ in order to infer $e^*t : A = B$ from $t : A = A$.[1]

[1] The Invariance Principle is thus logically equivalent to what may be called *weak univalence*:

$$A \simeq B \to A = B\,.$$

Taking this as an axiom would mean assuming a term $p : A \simeq B \to A = B$ witnessing the inference; so if $e : A \simeq B$ is an equivalence, then $p(e) : A = B$ is a path in U. Since we can transport along this path in any family of types over U, and transport is always an equivalence, there is in particular the transport $p(e)_* : A \simeq B$ in the identity family. The required "computation rule" then states that $p(e)_* = e$.

The Univalence Axiom and the associated invariance of the system of homotopy type theory provide a foundation of mathematics that works more smoothly than conventional foundations, because it allows for a more "structural" style of reasoning that is closer to modern mathematical practice. This permits explicit formalizations of high-level mathematics that would hardly be possible in conventional systems, where everything has to be analysed all the way down to specific, basic elements of set theoretic structures.[2] In this way, the system retains some of the convenience of the axiomatic approach, since objects are entirely determined by their mapping properties. (Modern type-theoretic intro and elim rules for type constructors are essentially just universal mapping properties.)

As useful as invariance may be, however, this is still just a matter of practical mathematical convenience. What I am really aiming at is a more important conceptual aspect of invariance, to which we can now finally turn: it agrees with, and indeed strengthens, a proposed characterization of what it means to be "logical". This I think provides a conceptual justification for the system of homotopy type theory that is more satisfactory than the meaning explanations previously given for MLTT. It also constitutes a solution to a long-standing problem in the philosophy of logic, as I will now try to explain.

4 The Tarski-Grothendieck Thesis

I will call the following principle the *Tarski-Grothendieck Thesis*:

> *If a statement, concept, or construction is purely logical, then it should be invariant under all equivalences of the structures involved. A statement that is not invariant must involve some non-logical specifics, pertaining not to general logical form but to some particular aspects of the objects bearing the structure. If it is the hallmark of a logical concept that it should pertain only to general, formal structure and not to any specific features of the objects bearing that structure, then this formal character may be witnessed by the fact that the concept is invariant under all equivalence transformations.*

This is really a fusion of two different threads: one of them was originally proposed by Frege, and later refined by Russell, Wittgenstein, and Carnap; the other starts with a somewhat later proposal by Tarski, modified in light of the more recent developments in mathematics that we have already mentioned, due mainly to Grothendieck and Voevodsky.

The first thread ends where the second one begins, namely with the question of which concepts in a formal language are to be regarded as logical, and taking "logical" to mean something like "purely formal" — as opposed to "contentful" or "empirical". This was the main unresolved problem of Carnap's research into the foundations of logical theory in the 1930s and '40s. It was motivated by Fregean and, later, Wittgensteinian concerns about the nature of logical truth, which had been central to the main doctrines of the Vienna Circle in the 1920s. (Recall Russell's quip about theft versus honest toil, and the formalist's retort.)

[2] E.g. Lumsdaine's formalization of the Blakers-Massey theorem, or Brunerie's calculation of $\pi_4 S^3$.

Today, one regards this line of thought to have ended with the famous Quine-Carnap debate over the theoretical possibility of striking a clean line between the analytic and the synthetic. Quine's decisive challenge rested on Carnap's ultimate inability to distinguish precisely the logical from the descriptive terms of a given formal language, in any but a merely ad hoc, unprincipled way. This difficulty was said by Carnap in later years to be "among the most important problems in all of theoretical philosophy". What made it so important was the supposed refutation of the doctrine of synthetic a priori knowledge, which was to follow from the claim that all logical and mathematical truths are analytic. Without a principled characterization of the logical terms, there was apparently no way to make a principled distinction between the analytic and the synthetic. The *philosophical* program of Logicism (as opposed to the associated mathematical program) — the one begun by Frege in the *Grundlagen*, promoted by Russell in the *Principles of Mathematics*, advanced by Wittgenstein in his *Tractatus*, and layed down by the Vienna Circle as the cornerstone of a new kind of scientific philosophy — depended crucially on being able to characterize what it means to be a logical concept.

Tarski's proposal — which he made in a lecture in 1966, too late for Carnap himself to make any use of it — was to take isomorphism invariance as the defining criterion of "logicality" (as it is now unfortunately called). This approach draws on much earlier ideas of Klein, Dedekind, and others, and indeed has an inherent plausibility. The goal of Klein's Erlangen Program had been to characterize different geometries in terms of the kinds of transformations that they admit. Eilenberg and Mac Lane's invention of category theory was explicitly presented as an extension of this idea, and made the notion of a structure-preserving transformation one of the basic, primitive concepts in the definition of a category. Tarski's idea was to make logic something like the "most general geometry", by taking as its characteristic class of transformations the most general one of all, consisting of all bijective transformations. As Tarski knew [**?**], Carnap had already shown much earlier that all logical concepts in the simple theory of types are invariant under isomorphisms, so it was not such a radical step to consider this property as the defining characteristic of logicality [strangely, Carnap himself never seems to have considered this option, despite having been very close to it several different times].

The general, model-theoretic reasoning behind using invariance to characterize logicality can be explained as follows: Given a notion of elementary structure, determined by some first-order operations and relations, and some first-order axioms in these (called a "theory"), one can consider the class of all such structures, and the subclass of all models of the theory. The important concept of logical entailment $\psi \models \varphi$, and the associated notion of logical validity $\models \varphi$, are then determined in terms of these models: for example, a first-order formula φ is defined to be valid (in this logical "theory") just if it is true in all the models. This determines a concept of "logical truth" for first-order formulas in the given language, but it does not tell us why these formulas in particular should be the candidates for expressing the logical truths, nor that there could not be others.

But if we now consider, not just the class of all models, but the *groupoid* consisting of all models and their isomorphisms, then we can also characterize the logical formulas themselves (up to logical equivalence) as those concepts that are invariant under all the isomorphisms of models. This holds for formulas in higher-order logic, and to a lesser extent for first-order formulas. (I can say more about this if you like.) The main point is that, by considering not just the class of models, but also their isomorphisms, we can determine not only which statements out of some predetermined class are the valid ones, but we can also determine the relevant class of statements, at least to some extent.

This idea has generated much recent work in philosophical logic, and it has been shown to be promising, but not entirely adequate as stated. For one thing, it relies on an external notion of model and transformation, usually characterized in terms of models in set theory, and this then involves various set theoretic issues such as the GCH in the characterization of the logical concepts. Another problem is that the class of all set theoretic isomorphisms does not characterize exactly the logically definable concepts, but only those relative to some previously given, isomorphism-invariant, basic constructions, such as products and powerset types — and this already tips the scales in a certain direction, like simple type theory or higher-order logic.

Moreover, there are reasonable concepts that are invariant, but are not given by any logical formula, and we may choose to regard these, too, as "logical" in a generalized sense. For example, there are much more general notions of structure than those initially determined in this special way by first-order operations and relations: one might start with a groupoid of such first-order models, but then add some special conditions on the maps in order to cut down to some structure that is not definable by elementary means (see for example continuous functions). Or one might specify a groupoid directly by some means other than as the models of a single, underlying logical theory (e.g. smooth manifolds). This, too, can be regarded as a system of "structured objects", but now there is no common structure to be described uniformly by logical formulas. However, we can still consider the statements, properties, and constructions that are invariant under the given isomorphisms. These are then the generalized "formulas", or "logical" statements in an extended sense; and the ones that hold universaly are then reasonably regarded as the "valid" ones. In this way, we extend the notion of a "logical concept" from one that can be expressed by a formula in a particular logical language, to one that is invariant with respect to a given notion of isomorphism.

What is needed then for a general characterization of the "logical" concepts — broadly construed — is, first of all, a broader class of maps under which all intuitively logical concepts are invariant. Moreover, since the reliance on set-theoretic models apparently undercuts the methodology, an intrinsic or independent development is also needed, if the characterization is to play the desired role as a general, account of logicality in the broader sense.

Now I hope you can see where this is heading: such a broader notion than isomorphism is exactly what is provided by the notion of equivalence in HoTT. Tarksi's proposal of isomorphism invariance was too weak; but in HoTT we can instead use invariance under *homotopy equivalence* — the relation of "same shape", rather than "same size". This is of course a much wider class of maps — equating e.g a solid ball and a single point, which are certainly not isomorphic. So Tarski was right that invariance may be taken as the hallmark of logicality; he just failed to take the widest possible class of transformations. Note that intensional MLTT *is* invariant under the wider class of (homotopy) equivalences, while HOL is not, owing to the occurrence of powersets in the latter.

It is worth noting at this point that, under Grothendieck's homotopy hypothesis — associating homotopy with ∞-groupoids — the notion of homotopy equivalence is associated to that of ∞-*isomorphism*, which is a natural generalization of isomorphism: roughly speaking, it is arrived at by *defining* an ∞-isomorphism between two objects to consist of maps back and forth, the respective composites of which are ∞-isomorphic to identities:

$$A \cong_\infty B \ =_{df} \ \text{there are } f : A \leftrightarrows B : g \text{ with } g \circ f \cong_\infty 1_A \text{ and } f \circ g \cong_\infty 1_B$$

That's not a mistake, but rather what is known as a "co-inductive" definition — it is indeed "circular", but not viciously so. The formal concept of equivalence in HoTT captures exactly this informal intuition of "isomorphism up to isomorphism up to isomorphism ..." (but in a single, finite expression). And indeed, every concept definable in homotopy type theory is therefore also "∞-isomorphism invariant" — generalizing Carnap's above result about the isomorphism invariance of the concepts in simple type theory.

Now the Univalence Axiom rests squarely on the fact that all concepts are invariant — as we have seen, its formal consistency requires that this be so. But taking UA as a new foundational axiom goes a step further: it in effect *asserts* that all concepts that can appear in the system of foundations are invariant, and therefore broadly "logical" according to the above Tarski-Grothendieck thesis.

Note by the way that this is the same sort of thing that Russell did when he added the axiom of extensionality to the system of PM in the second edition. That axiom identified extensionally equivalent propositional functions in order to define sets, eliminating at least some uses of the notorious Axiom of Reducability.

Russell's justification for adding this axiom was an argument that he attributed to Wittgenstein, who formulated it in typically cryptic fashion as the slogan: "a function can only appear in a proposition through its values". What this means is that every propositional function that one can actually write down in logical symbols is extensional, and so one might as well assume that *all* functions are extensional (according to Wittgenstein's syntactic account of what it means to be logical).

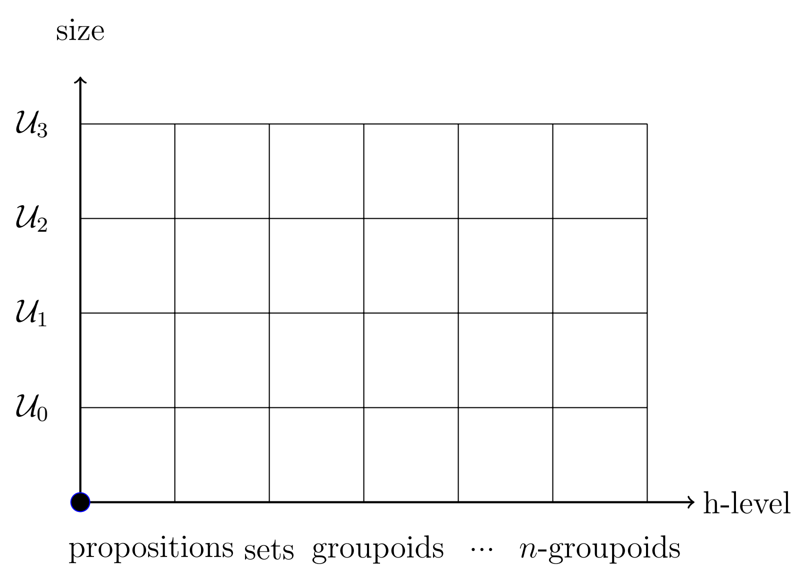
We can justify the UA in much the same way: since every construction that one can actually write down in HoTT respects equivalence, i.e. is invariant, we might as well assume that all constructions are invariant — not because we believe in Wittgenstein's theory about the syntactic nature of logic, though, but because we accept invariance as a reasonable formal model — an explication, if you will — of what it means for a statement, property, or construction to be "logical" in the broad sense (or, if you prefer, "mathematical"). That is to say, we decide to accept the above Tarski-Grothendieck Thesis.

Making this move of adding UA to HoTT, essentially asserting that the objects of the system are something like "pure logical structures", then has many fascinating consequences. It results in a coherent picture of the mathematical universe, but one that is rather different from that of classical set theory:

- Sets, as we have said, can be "identified" if they are isomorphic — which is to say that they are only determined up to isomorphism, so they are more like the cardinal numbers defined by Frege (or the sets as defined by category theorists). This agrees with an older conception of sets as bare "multiplicities" or "Mannigfalltigkeiten". In particular, the elements of a set are themselves unstructured points, or "lauter Einsen", to speak with Cantor, and not themselves sets of sets, etc. (cf. [**?**]).

- Similarly, algebraic structures like groups and rings also appear as abstract objects that are "determined only up to isomorphism", in the sense that any property or structure on one can be transferred to any isomorphic one.
- As already mentioned, the formerly separate logical and mathematical propositions are now also included at the base of the system, as objects of a special kind. For them, the notion of equivalence specializes to conventional logical equivalence.
- The type of all propositions is then a set, but the type of all sets is an object of a "higher" kind — namely a groupoid. The notion of equivalence for such groupoids is just the one familiar from category theory.
- But the mathematical universe does not stop there: there are also higher forms of objects, with more subtle forms of equivalence.

The universe of all types in HoTT with UA contains many different kinds of objects, which are arranged in a newly discovered hierarchy of "homotopical dimension" or "h-level", in addition to the hierarchy of "size" familiar from set theory. The hierarchy of size is driven by concerns of consistency, self-reference, self-membership, etc. But the system of h-levels is determined by the "internal structure" of the objects.



The system of HoTT with the UA presents a new and fascinating picture of the mathematical universe, and it may also turn out to be a useful, practical foundation for everyday mathematics. I hope to have convinced you that it is also philosophically well-motivated, and that it may also be useful for philosophical investigations into the nature of logic and mathematics.