Intensionality, Invariance, and Univalence

Steve Awodey
Carnegie Mellon University

Midwest PhilMath Workshop
18 November 2020

Dedicated to Mic Detlefsen

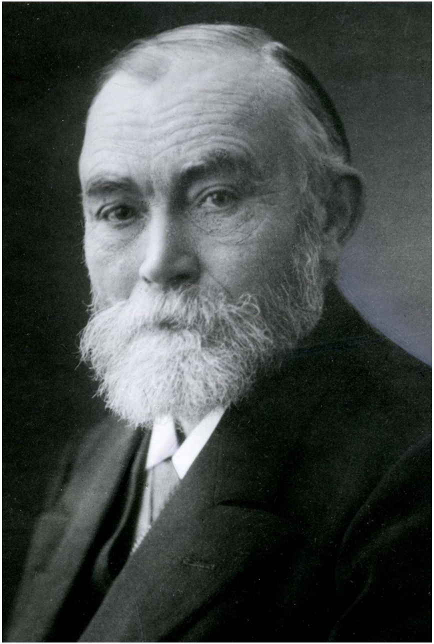
1. Frege's puzzle about equality

What is the meaning of a word or statement?

Frege begins by considering equality:

> *Equality gives rise to challenging questions which are not altogether easy to answer. Is it a relation? A relation between objects? Or between names or signs of objects? If we were to regard equality as a relation between that which the names 'a' and 'b' designate, it would seem that $a = b$ could not differ from $a = a$ (provided $a = b$ is true). A relation would thereby be expressed of a thing to itself, and indeed one in which each thing stands to itself but to no other thing. What is intended to be said by $a = b$ seems to be that the signs or names 'a' and 'b' designate the same thing, so that those signs themselves would be under discussion; a relation between them would be asserted.*

Frege, Über Sinn und Bedeuting, 1892



1. Frege's notions of meaning and sense

- He finally decides that it must be a relation between things, but that every expression (name, predicate, sentence) must have both a meaning (*Bedeutung*) and a sense (*Sinn*).
- The meaning is the thing denoted (*das Bezeichnete*).
- The sense is how the meaning is presented (*Art des Gegebenseins*).
- The meaning of a sentence is called its "truth-value" (*Wahrheitswert*).

  If our conjecture that the meaning of a sentence is its truth-value is correct, the latter must remain unchanged when part of the sentence is replaced by an expression with the same meaning but a different sense. ... What else but the truth-value could be found, that belongs quite generally to every sentence ... and remains unchanged by substitutions of the kind in question?

1. Frege's notions of meaning and sense

> *If now the truth value of a sentence is its meaning, then on the one hand all true sentences have the same meaning and so, on the other hand, do all false sentences.*

This is the conclusion that I would like to avoid: that all true statements mean the same thing, namely True – especially for statements of logic and mathematics.

- For mathematical objects, the role of sense can be played by a *presentation* of the object.
- The meaning of an object or theorem should then not depend on a choice of a presentation.
- But e.g. different theorems may still mean something very different, even though both are true. An arithmetical equality like $5 + 7 = 12$ means something different than a quantified statement like the Commutative Law.

2. Martin-Löf's intensional type theory

One approach that has this character is the intensional type theory of Per Martin-Löf, which uses two different kinds of equality.

- $a \equiv b$ *judgmental* equality,
- $a = b$ *propositional* equality $\big($also written $\mathrm{Id}(a, b)\big)$.

In this system, $(a \equiv b)$ implies $(a = b)$, but not conversely.
So these relations can be used to model Frege's distinction:

- $a \equiv b$ says a and b have the same **sense**:
  roughly "same syntactic presentation in the system".
- $a = b$ says a and b have the same **meaning**:
  roughly "same things being reasoned about by the system".

For example, for the type $\mathbb{N}$ of natural numbers we have

$$5 + 7 \equiv 12\,,$$

but only

$$\forall n\,(m + n = n + m)\,.$$

2. Martin-Löf's intensional type theory

In addition to equality, the basic operations of type theory are:

$$0,\ 1,\ A+B,\ A\times B,\ A\to B,\ \Sigma_{x:A}B(x),\ \Pi_{x:A}B(x)\,.$$

These correspond to the logical operations:

$$\bot,\ \top,\ p\vee q,\ p\wedge q,\ p\Rightarrow q,\ \exists x\, p(x),\ \forall x\, p(x)\,.$$

- But unlike in predicate logic where one is only concerned with entailment $p \vdash q$, in type theory one also has terms $x : A \vdash t : B$ which can be regarded as proofs, computations, witnesses, grounds of truth, etc.
- Mere provability $\vdash p$ is replaced by having a term $\vdash t : A$.
- Thus the *meaning* of a type A is not just true or false, but the collection of all its "proofs" $\vdash t : A$.

3. Propositions as types

This is also called the Curry-Howard correspondence.

Proofs	**Constructions**
proof : Proposition	term : Type
assumption of A	$x : A$
$\wedge$-intro	$\langle a, b \rangle : A \times B$
$\Rightarrow$-intro	$\lambda x.t(x) : A \to B$
...	...

3. Propositions as types

There are, at first blush, two kinds of construction involved: constructions of proofs of some proposition and constructions of objects of some type. But I will argue that, from the point of view of foundations of mathematics, there is no difference between the two notions. A proposition may be regarded as a type of object, namely, the type of its proofs. Conversely, a type A may be regarded as a proposition, namely, the proposition whose proofs are the objects of type A. So a proposition A is true just in case there is an object of type A.

W.W. Tait
The law of excluded middle
and the axiom of choice (1994)

4. Equality types

- Under propositions as types, the meaning of a proposition is its not just its truth-value, but the collection of its proofs.
- This is already a better notion of meaning than just the truth-values derived from logical equivalence of propositions.
- But there is an even richer notion of meaning in mathematics, related to isomorphism of algebraic structures.
- This can also be captured in type theory, using the already mentioned **equality type** $(a = b)$.

4. Equality types

- For any type A and terms $a, b : A$, there is a type $(a = b)$.
- For any $a : A$, there is a term $r(a) : (a = a)$.
- Two terms $a \equiv b$ are always interchangeable, so if $a \equiv b$ holds, then there is a term $t : (a = b)$.
- Thus $a \equiv b$ implies $a = b$.
- But the rules for $a = b$ **do not** imply that $a \equiv b$.

Thus *different* "presentations" a, b may *mean* the same thing.

4. Equality types

- Moreover, every property $x : A \vdash P(x)$ of A-objects respects this notion of meaning in the following sense: given a proof $p : (a = b)$, and one $t : P(a)$, there is an associated one $p_* t : P(b)$.
- This is Frege's observation that the truth-value of $P(a)$ does not change when one substitutes a component by another one with the same meaning.
- But here we see that such a substitution also acts on the PAT meaning, i.e. the set of all proofs of the proposition.
- Indeed, a proof $p : (a = b)$ induces a function p_* from terms $t : P(a)$ to terms $p_* t : P(b)$.

5. The homotopy interpretation

But more is actually true: equality endows each type with a structure that is respected by all constructions.

Suppose we have terms of ascending equality types:

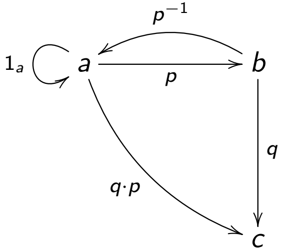
$$
\begin{aligned}
a,\ b &: A \\
p,\ q &: (a = b) \\
\alpha,\ \beta &: (p =_{(a=b)} q) \\
\ldots &: (\alpha =_{(p =_{(a=b)} q)} \beta)
\end{aligned}
$$

Consider the interpretation:

$$
\begin{array}{rcl}
\text{Types} & \rightsquigarrow & \text{Spaces} \\
\text{Terms} & \rightsquigarrow & \text{Maps} \\
a : A & \rightsquigarrow & \text{Points } a : 1 \to A \\
p : (a = b) & \rightsquigarrow & \text{Paths } p : a \sim b \\
\alpha : (p =_{(a=b)} q) & \rightsquigarrow & \text{Homotopies } \alpha : p \approx q \\
& \vdots &
\end{array}
$$

6. The fundamental groupoid of a type

In topology, the points and paths in any space bear the structure of a **groupoid**: a category in which every arrow has an inverse.

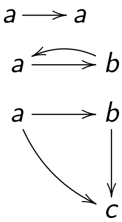


In the same way the **terms** $a, b, c : X$ and **equality terms** $p : (a = b)$ and $q : (b = c)$ of any type X also form a groupoid.

6. The fundamental groupoid of a type

The usual laws of equality provide the **groupoid operations**:

$r : (a = a)$	reflexivity	
$s : (a = b) \to (b = a)$	symmetry	
$t : (a = b) \times (b = c) \to (a = c)$	transitivity	



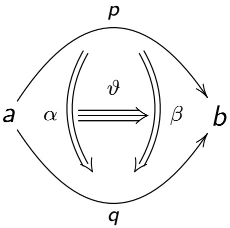
And as in topology, the **groupoid equations**:

$p \cdot (q \cdot r) = (p \cdot q) \cdot r$	associativity
$p^{-1} \cdot p = 1 = p \cdot p^{-1}$	inverse
$1 \cdot p = p = p \cdot 1$	unit

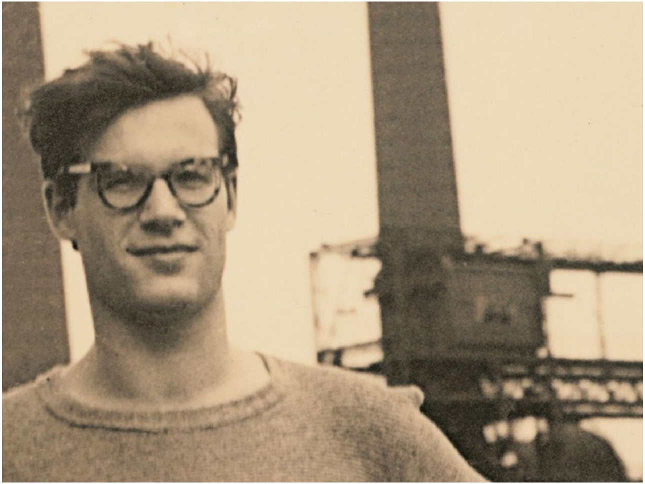
hold only "**up to homotopy**", i.e. up to higher $=$-terms.

6. The fundamental ∞-groupoid of a type

In this way, each type in the system is endowed with the structure of an ∞-**groupoid**, with terms, equalities between terms, equalities between equalities, ...



Such structures already occur elsewhere in Mathematics, e.g. in Grothendieck's famous **homotopy hypothesis**.



7. The hierarchy of *n*-types

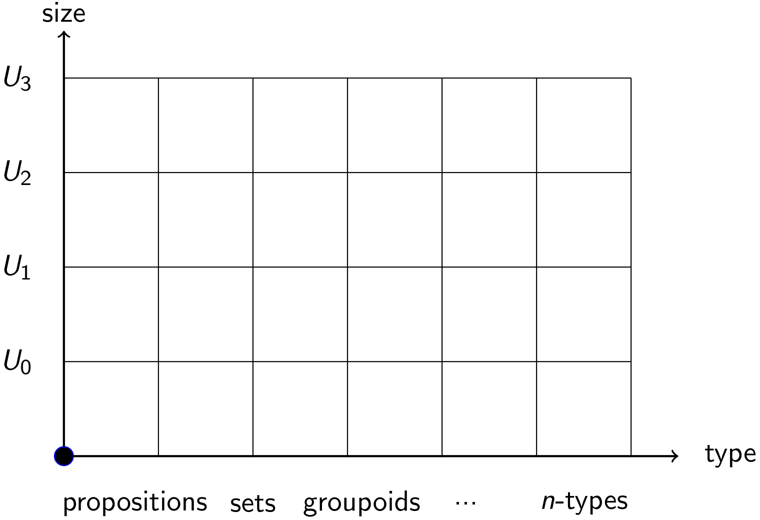
The universe of all types is naturally stratified by the level at which the fundamental ∞-groupoid becomes trivial (if it ever does).

A is a **proposition**:	*A has at most one term.*
A is a **set**:	*identity on A is always a proposition.*
A is a **groupoid**:	*identity on A is always a set.*
A is an $(n+1)$-**type**:	*identity on A is always an n-type.*

$$\vdots$$

7. The hierarchy of n-types

This gives a new view of the mathematical universe in which types also have intrinsic higher-dimensional structure.



8. Equivalence of types

- The idea of n-types refines the *propositions as types* conception: types are now also *higher structures*, rather than mere *propositions* (truth-values) or *sets* of proofs.
- **Logical equivalence** of types ("same truth-values") is defined as usual:

$$A \leftrightarrow B \ =_{df} \ (A \to B) \times (B \to A)$$

  This is fine for propositions, but it is too coarse for sets.
- **Isomorphism** of types ("same set of terms") is also defined as expected:

$$A \cong B \ =_{df} \ \sum_{f:A \to B} \sum_{g:B \to A} (g \circ f = 1_A) \times (f \circ g = 1_B)$$

  This is fine for sets, but it is too coarse for higher types.

8. Equivalence of types

There is a notion of **equivalence of types** $A \simeq B$, that is finer than isomorphism, and takes the higher structure into account.

Equivalence $A \simeq B$ specializes to:

- *logical equivalence* $A \leftrightarrow B$ for propositions,
- *isomorphism* $A \cong B$ for sets,
- *categorical equivalence* $A \simeq B$ for 1-types (groupoids),

  ...
- *homotopy equivalence* $A \simeq B$ for spaces (∞-groupoids).

8. Equivalence of types

- For a family of types $x : A \vdash B(x)$, we saw that a term $p : (a =_A b)$ and one $t : B(a)$ determine a term $p_* t : B(b)$.
- In fact, the map $p_* : B(a) \to B(b)$ is always an equivalence of types $B(a) \simeq B(b)$.
- Thus equivalence $A \simeq B$ provides a finer notion of meaning than the truth-values derived from logical equivalence $A \leftrightarrow B$.
- This is a better answer to Frege's question "what else but the truth-value could be found ...?"

9. A Fregean test

> *The supposition that the truth-value of a sentence is its meaning shall now be put to further test. We have found that the truth-value of a sentence remains unchanged when an expression is replaced by another having the same meaning: but we have not yet considered the case in which the expression to be replaced is itself a sentence. Now if our view is correct, the truth-value of a sentence containing another sentence as a part must remain unchanged when that part is replaced by another sentence having the same truth-value.*
>
> Frege, Über Sinn und Bedeuting

Of course, we must now replace the truth-value in this test by our new proposal for the meaning of a sentence, namely the equivalence class under $A \simeq B$, which we call the **homotopy type**.

9. A Fregean test

Let $\Phi[X]$ be a type expression containing a type variable X, and consider whether the homotopy type of $\Phi[X]$ respects the homotopy type of X in Frege's sense:

$$(A \simeq B) \to \big(\Phi[A] \simeq \Phi[B]\big)$$

This is an **invariance principle** for the type theoretic language with respect to homotopy equivalence.
It can indeed be shown to hold, e.g. by an induction on the construction of $\Phi[X]$.

10. Tarski's invariance proposal

Now suppose we ... consider still wider classes of transformations. In the extreme case, we would consider the class of all one-one transformations of the space, or universe of discourse, or 'world', onto itself. What will be the science which deals with the notions invariant under this widest class of transformations? Here we will have very few notions, all of a very general character. I suggest that they are the logical notions, that we call a notion 'logical' if it is invariant under all possible one-one transformations of the world onto itself.

Tarski, What are logical notions? (1966)

10. Tarski's invariance proposal

Tarski observes that all concepts Φ that are definable in Russell's theory of types (higher-order logic) are invariant under *isomorphism*, and proposes this as an explication of the notion of a "logical concept".

We have just seen that all concepts in Martin-Löf type theory are invariant under an even wider class than isomorphism, namely *homotopy equivalence*.

(Cardinality is an example of a concept definable in HOL that is not invariant under homotopy equivalence.)

11. Internalizing invariance

We can state this invariance principle *in type theory* by adding a universe of types U, so that we have type variables $X : U$.

We then replace the schematic type $\Phi[X]$ by a family of types

$$X : U \vdash P(X)\,.$$

Finally, we consider the type

$$(A \simeq B) \to \big(P(A) \simeq P(B)\big)$$

which formulates the invariance principle internally.

11. Internalizing invariance

Given the invariance principle

$$(A \simeq B) \to \big(P(A) \simeq P(B)\big)\,,$$

take $P(X)$ to be the equality type on the universe $(A = X)$ to get

$$(A \simeq B) \to \big((A = A) \simeq (A = B)\big).$$

Since $(A = A)$ we then get

$$(A \simeq B) \to (A = B).$$

Thus *equivalent types are equal*.

The celebrated **Univalence Axiom** of Voevodsky says something even stronger, namely that equivalence is *equivalent* to equality:

$$(A \simeq B) \simeq (A = B)\,.$$

12. Invariance and Univalence

Univalence in the form

$$(A \simeq B) \simeq (A = B)$$

is thus an *internalization* of the principle of invariance: it says that all concepts respect equivalence.

Applied to *itself*, univalence becomes

$$(A \simeq B) = (A = B)\,.$$

Now let us reconsider our original proposal regarding the meaning of a mathematical statement.

13. Univalence and Intensionality

$$(A \simeq B) = (A = B)$$

Equality, recall, is *sameness of meaning* for type expressions A, B.

Such expressions $A, B, \ldots$ are *presentations* of mathematical structures and propositions, and they present *the same* mathematical object if $A = B$.

Thus univalence says that two type expressions A, B present the same homotopy type just in case they mean the same thing – indeed it says something stronger: to say that A, B present the same homotopy type *means the same thing* as to say that they mean the same thing.

The meaning of a mathematical statement is its homotopy type.

References

Awodey, S., A proposition is the (homotopy) type of its proofs. In: *Logic, Philosophy of Mathematics, and their History: Essays in Honor of W.W. Tait*. E. Reck (ed.), College Publications, 2018. arXiv:1701.02024

Shulman, M., Homotopy type theory: The logic of space. In: *New Spaces in Mathematics and Physics*, M. Anel and G. Catren (ed.s), Cambridge University Press, 2020. arXiv:1703.03007

The Univalent Foundations Program, *Homotopy Type Theory: Univalent Foundations of Mathematics*, Institute for Advanced Study, 2013. `www.homotopytypetheory.org/book`

PS. Univalence implies Strong Invariance

Univalence in the form

$$(A \simeq B) \simeq (A = B)$$

implies invariance.

- Given an equivalence $e : A \simeq B$, by univalence we get an equality $\overline{e} : (A = B)$.
- Then the equality $\overline{e}$ acts on any type family $X : U \vdash P(X)$ to give an equivalence $\overline{e}_* : P(A) \simeq P(B)$.
- So univalence implies the invariance principle

$$(A \simeq B) \to \big(P(A) \simeq P(B)\big) .$$

- The homotopy invariance of type theory (with universes!) thus follows from the consistency of univalence.