Univalent Foundations of Mathematics

Talk by Vladimir Voevodsky

September 5 , 2011

”Classic” foundations of mathematics

The only formal foundations we had until now are based on the Zermelo-Fraenkel theory with the axiom of choice (ZFC).

It was created in the early 20th century before the results of Goedel at the time when it was hoped that a formal proof of its consistency will be found.

Thus the main goal of the creators of ZFC was to have a theory which is general enough and at the same time as simple as possible to analyze for the purposes of finding a consistency proof. It was not designed with the purpose of making formalization of mathematics convenient.

Problem of equivalence

The key practical problem of ZFC-based formalization of mathematics can be called "*the problem of equivalence*" .

Set-level mathematics is only interested in properties of sets with structures which are invariant under isomorphisms.

Higher level mathematics dealing with categories and their higher analogs is only interested in properties invariant under appropriatley defined higher equivalences.

ZFC does not provide any natural approach to distinguish between constructions and properties which respect equivalences and the ones which do not.

The problems of equivalence and of consistency in the univalent foundations

1. Univalent foundations solve the "problem of equivalence" both for structures on sets and for structures on objects of higher levels by making it syntactically impossible to formulate properties or describe constructions which do not respect equivalences.

2. Formal consistency of the univalent foundations is essentially equivalent to formal consistency of ZFC. However, since the univalent foundations are constructive a possible discovery of formal inconsistency would have a somewhat different effect on the univalent foundations than on the ZFC-based ones.

What is different and what is the same in ZFC-based and univalent foundations

1. As ZFC-based foundations the univalent foundations are "universal" i.e. can be used for systematization and formalization of all areas of mathematics.
2. Unlike ZFC-based foundations which are formulated in the language of predicate calculus the univalent foundations are formulated in languages of a completely different class called Martin-Lof type theories.
3. Unlike ZFC-based foundations the univalent foundations are intrinsically constructive.
4. Univalent foundations are based on axiomatization of the "world" of homotopy types instead of the "world" of sets.

Two main discoveries which made univalent foundations possible are:

1. Grothendieck correspondence between infinity groupoids and homotopy types. It was suggested as an informal conjecture by Alexander Grothendieck in the 80-ies . The first proof was given in early 90-ies by M. Kapranov and myself .

2. Univalent model of Martin-Lof type theory which interprets types not as sets but as homotopy types. I constructed it in 2009 using ideas of Steve Awodey and his students and my earlier concept of the univalent fibrations.

A bit of history

1. The most famous attempt to systematize all of mathematical knowledge was made by Bourbaki group. Their key idea was to define mathematics as the study of various structures on sets. Bourbaki project was highly successful but encountered two major issues:

   (a) Since they used paper publications it was impossible to back-track and make changes in earlier exposition which were called for by new developments in mathematics.

   (b) The appearance of the concepts of categories and functors undermined their main idea of mathematics as the study of structures on sets.

A bit of history (cont.)

2. Later several attempts have been made to use categories as a basis of new foundations of mathematics. These attempts led to some important and interesting discoveries but never succeeded in their original purpose of creating new universal foundations.

   One of the reasons was the introduction of 2-categories and later higher categories which posed the same problems for category-based foundations as categories did for set-based ones .

   Another reason was that it was not recognized that categories are not the "next level" analogs of sets but rather the "next level" analogs of partially ordered sets.

∞-groupoids and homotopy types

As soon as categories are recognized as the next level analogs of partially ordered sets it becomes apparent that the proper next level analogs of sets themselves are objects which are classically know as groupoids - categories where all morphisms are isomorphisms. As one goes to even higher levels one has to consider 2-groupoids, 3-groupoids etc.

Higher groupoids are much more tractable than higher categories due to Grothedieck's insight that the "world" of ∞-groupoids is the same as the world of homotopy types with n-groupoids corresponding to so called n-types.

Summary : homotopy types and foundations of mathematics

1. Proper foundations for contemporary mathematics must provide a way to directly work not only with sets but also with higher analogs of sets.
2. The objects of set-theoretic mathematics which are most likely to correspond to "higher sets" are higher groupoids. By Grothendieck's insight groupoids of all levels may be considered as homotopy types.
3. We conclude that in order to build proper foundations for contemporary mathematics we need to have a formal deduction system which can be used to describe constructions on homotopy types and to prove their properties.

A bit of history (cont.)

In 2005-2006 I started to work on developing such a deduction system using as the basis standard constructions of dependent polymorphic type theory. I understood that universes in of such type theories are to be interpreted as bases of special fibrations which I called univalent fibrations and proceeded to building my own version of constructive type theory which I called "homotopy λ-calculus".

Then in 2009 I discovered that such a deduction system, or rather a class of deduction systems, already exists and moreover that a remarkably sophisticated "proof assistant" based on one of the languages of this class is being taught to undergraduates at Princeton CS department!

Martin-Lof type theories

The deduction systems of the class in question are called Martin-Lof type theories.

The first system of this class was introduced by Per Martin-Lof in the 70-ies as a basis for new foundations of constructive mathematics. Two most important for us features of his theory are:

1. identity types or types of "intensional equality", together with the associated "induction principle" and "computation rules" which are defined for any pair of terms of a given type X,
2. a universe U which is used to quantify over types.

A bit of history (cont.)

It was originally assumed that Martin-Lof theory is something like a constructive set theory. Types were interpreted as sets and constructions on types as corresponding constructions on sets.

It was soon observed however that it is not a very good formalization of the world of sets because many of the natural properties expected from sets were not provable in the Martin-Lof theory. Adding axioms which made the objects of his theory to behave more like sets led to the deterioration of its constructive nature.

Consequently it has not become popular with mathematicians.

A bit of history (cont.)

The ideas of Martin-Lof found their way into theoretical computer science in part through Coquand's Calculus of Constructions and its extension - Calculus of Inductive Constructions .

This later variant of Martin-Lof type theory , more complex and more convenient for practical use because of its sophisticated machinery of inductive definitions became the basis for proof assistant Coq - the proof assiatant which is now taught in the course on programming languages in Princeton University and in many other leading universities.

Martin-Lof theory and homotopy theory

The first hint that Martin-Lof type theory may have something to do with homotopy types appeared in 1996 when Martin Hofmann and Thomas Streicher constructed a new semantics for a version of this theory which interpreted types not as sets but as groupoids.

In 2005 Steve Awodey discovered the connection between the Martin-Lof "induction principle" for the identity types and factorization axioms of the abstract homotopy theory. This led to the interpretation of identity types as path spaces.

At about the same time I understood that the universe U is to be interpreted as the base of a universal univalent fibration.

A bit of history (cont.)

As I mentioned it all came together in the fall of 2009. Combining the ideas of Steve Awodey on the interpretation of the identity types with my ideas on the interpretation of the universes I have constructed the *univalent model* of the calculus of inductive constructions.

This model provides a semantics for the Calculus which allows one to use it to do exactly what was needed from the hypothetical language for new foundations of mathematics which was discussed above.

In February 2010 I started to write a Coq library of formalized mathematics based on the univalent model. See `http://github.com/vladimirias/Foundations/` .

Some of the key univalent concepts

1. There is a filtration on types , or rather on type expressions, by their "h-level".
   (a) Types of h-level 0 are equivalent to the one point type.
   (b) Types of h-level 1 correspond to propositions.
   (c) Types of h-level 2 correspond to sets.
   (d) Types of h-level 3 correspond to groupoids.
   (e) Types of higher levels correspond to higher groupoids or, equivalently, to more general homotopy types.

Some of the key univalent concepts (cont.)

2. Types with decidable equality such as natural numbers, trees etc. have level ≤ 2 e.g. the usual inductive types are sets.
3. Typical examples of types of level > 2 are universes.
4. Constructions translated into CIC using univalent semantics are invariant under *weak equivalences* between types.
5. The univalent model satisfies a new axiom which is called the univalence axiom. It imposes the condition that the identity type between two types is naturally weakly equivalent to the type of weak equivalences between these types.

Some of the key univalent concepts (cont.)

6. The univalence axiom implies the functional extensionality both for "straight" functions and for dependent functions. It also implies that two logically equivalent "propositions" (types of h-level 1) are equal.

7. The univalence axiom implies that the universe of types of h-level n has h-level $n + 1$. In particular, the type of "propositions" is a "set" and the type of "sets" is a "groupoid".

8. The univalence axiom implies similar statements for types with structures e.g. one can prove using the univalence axiom that the identity type between two groups is equivalent to the type of isomorphisms between these groups.

Some of the key univalent concepts (cont.)

9. Unlike many other axioms (e.g. the axiom of excluded middle), the univalence axiom is expected "to have computational content". In other words decidable normalization should be extendable in a certain sense to terms which involve the univalence axiom. For example there is the following precise:

   Conjecture 1. There exists a terminating algorithm which for any term expression t of type [nat] (natural numbers) constructed using the univalence axiom returns a term expression t' of type [nat] which does not use univalence axiom and a term expression of the identity type [Id nat t t'] which may use the univalence axiom.

There will be a full year program on Univalent Foundations topic at the Institute for Advanced Study in 2012-2013 co-organized by Steve Awodey, Thierry Coquand and myself. For information on the program see `http://www.math.ias.edu`.