A challenge for the type system

Wiles' proof of Fermat's Last Theorem uses many sources, some of which rely on advanced set theory. Most theorem provers use a type system to prohibit undefined expressions. If we want to formalize Wiles' proof without modification, we therefore need a theorem prover with a powerful type system.

The following is a short introduction to category theory followed by an excerpt from chapter 0 of [2] (EGA I for the insiders).

Naive category theory

Recall that a *category* is a class of objects together with a prescription that assigns to every pair of objects of the class, say X and Y, a set of morphisms from X to Y, and a composition operator $\circ$ between morphisms.

If C is a category, we write $Ob(C)$ to denote the class of objects of C. We write $C(X,Y)$ to denote the set of morphisms from X to Y. For every triple of objects X, Y, Z, the composition gives a function $C(Y,Z) \times C(X,Y) \to C(X,Z)$: if $f \in C(X,Y)$ and $g \in C(Y,Z)$ then $g \circ f \in C(X,Z)$.

This composition is required to be associative: if $f \in C(X,Y)$ and $g \in C(Y,Z)$ and $h \in C(Z,W)$, then $(h \circ g) \circ f = h \circ (g \circ f)$. It is also required that every object X has an identity morphism $1_X \in C(X,X)$ such that $g \circ 1_X = g$ for every $g \in C(X,Y)$ and $1_X \circ h = h$ for every $h \in C(Y,X)$.

Let **Set** be the category with all sets as objects and with functions as morphisms. A category C is called *small* if $Ob(C)$ is a set (not only a class). The category **Set** is not small.

If $(P, \leq)$ is a partially ordered set, we can make it into a small category in the following way. Put $Ob(P) = P$. If $x \leq y$, let $P(x,y)$ be the singleton set that only contains the pair $\langle x, y \rangle$, otherwise $P(x,y)$ is empty.

If G is a monoid (or a group), we can make it into a small category, by giving it a single object, say $*$, and defining $G(*,*) = G$ with the composition equal to the monoid operation.

Top is the category of the topological spaces as objects and the continuous functions between them as morphisms.

If C is a category, we define the *dual* or *opposite* category C^o as follows. $Ob(C^o) = Ob(C)$. For every pair of objects X and Y of C, we take $C^o(X,Y) = C(Y,X)$. The composition of C^o is the transpose of the composition of C.

Functors and natural transformations

If C and D are categories, a *functor* $F : C \to D$ is a prescription that assigns to every object X of C an object $F(X)$ of D, and to every morphism $f \in C(X,Y)$ a morphism $F(f) \in D(F(X), F(Y))$, in such a way that identity morphisms are mapped to identity morphisms and F distributes over composition of morphisms.

If $F, G : C \to D$ are two functors from C to D, a *natural transformation* from F to G is a prescription, say t, that assigns to every object X of C a morphism $t_X \in D(F(X), G(X))$ such that, for every morphism f of category C, say $f \in C(X,Y)$, we have $t_Y \circ F(f) = G(f) \circ t_X$ in $D(F(X), G(Y))$.

Now *Funct*(C,D) is the category with functors from C to D as objects and natural transformations from F to G as morphisms.

Presheaves on a topological space

Let X be a topological space. We write $|X|$ to denote the set of open subsets of X, partially ordered by inclusion. Let $|X|^o$ be the opposite category, so that open subsets $U \subseteq V \subseteq X$ induce a morphism $\langle U, V \rangle$ in $|X|^o(V, U)$. A presheaf on X with values in a category C is a functor from $|X|^o$ to C. If P is such a presheaf, the morphism $P(\langle V, U \rangle) \in C(P(V), P(U))$ is called the restriction from V to U.

If C is the category of groups, rings, etc., then presheaves with values in C are called presheaves of groups, or rings, etc.

Sheaves are presheaves that satisfy certain additional conditions.

A *ringed space* is defined as a topological space together with a sheaf of rings. There is a natural way to form a category **RiSp** of ringed spaces, with a natural "forgetful" functor to **Top**.

Yoneda's Lemma

For any category C, we define $C^@ = \mathit{Funct}(C^o, \mathbf{Set})$. We define a functor $h : C \to C^@$ in the following way. For every object X of C, we should get a functor $h(X) : C^o \to \mathbf{Set}$; for every object Y of C, we define the set $h(X)(Y) = C(Y, X)$. If $f \in C(Y, Z)$, then $f \in C^o(Z, Y)$ and we define $h(X)(f) \in \mathbf{Set}(h(X)(Z), h(X)(Y)) = \mathbf{Set}(C(Z, X), C(Y, X))$ by $h(X)(f) = (\lambda g \cdot g \circ f)$.

It is a standard verification to see that this makes $h(X)$ to an object of $C^@$. There is only one natural way to define h on morphisms of C, namely with $h(f)_Z = (\lambda u \cdot f \circ u)$. The verification that this makes h into a functor from C to $C^@$ is standard.

Lemma 1 *For every object X of C and every object F of $C^@$, there is a natural bijection $F(X) \to C^@(h(X), F)$. If $F = h(Y)$, this bijection equals the corresponding branch of the functor h.*

Proof. To find a bijection $t : F(X) \to C^@(h(X), F)$, one proceeds as follows. Let $f \in F(X)$. We have to construct some $t(f) \in C^@(h(X), F)$. Let Y be an object of C. We have to construct some function $t(f)_Y : h(X)(Y) \to F(Y)$. Let $g \in h(X)(Y) = C(Y, X)$. We have to construct $t(f)_Y(g) \in F(Y)$. Since F is a contravariant functor from C to **Set**, we have a function $F(g) : F(X) \to F(Y)$. This implies that we can define $t(f)_Y(g) = F(g)(f)$. Abstraction gives $t(f)_Y = (\lambda g \cdot F(g)(f))$. We can then verify that the family of functions $(t(f)_Y)_Y$ is a natural transformation $h(X) \to F$. This gives us a function $t : F(X) \to C^@(h(X), F)$.

Conversely, $s = (\lambda u \cdot u_X(1_X))$ is a function $C^@(h(X), F) \to F(X)$ and a straightforward computation gives that $s \circ t$ and $t \circ s$ are both identity functions. Therefore t is a bijection. The special case of $F = h(Y)$ may be left to the reader. $\square$

This result implies that the category C can be identified in a natural way as a nice subcategory of $C^@$. An functor in (or object of) $C^@$ is called *representable* iff it is isomorphic in $C^@$ to an object of the form $h(X)$.

This Lemma is applied in various forms in the sources used in Wiles' proof, usually, with category C not small. In other words, we need this result in a form where C can be instantiated with an arbitrary (not necessarily small) category.

Strictly speaking, all this requires a hierarchy of universes, and the theory may become parametrized by a universe, see [1] p.1–8.

References

[1] M. Artin, A. Grothendieck, and J.L. Verdier. *Théorie des topos et cohomologie étale des schémas (SGA4)*. Springer V., Berlin, etc., 1972. LNM 269.

[2] A. Grothendieck and J.A. Dieudonné. *Eléments de Géométrie Algebrique I.* Springer V., Berlin, etc., 1971.