A new set theory

M. Makkai (McGill University)

1. Abstract sets

Naive set theory: a mix of ideas:

1) the cumulative hierarchy
   (Cantorian universe)
   formalized by ZFC (+ ...)

versus

2) sets with urelements.

"Let: G and H be arbitrary groups. Define \((G, H) \mapsto G \circ H\) as follows.

Underlying set:

\[ |G \circ H| = \begin{cases} |G| \cap |H| & \text{if this } \neq \emptyset \\ \{1\} & \text{otherwise} \end{cases} \]

The operation on \(G \circ H\) is defined as ...
"

BAD definition:

1) elements of \(G\) and \(H\) are abstract, and (should) have nothing to do with each other;
2) operation \(\circ\) should be, and isn't, invariant under isomorphisms.
**Invariance under isomorphism:**

\[ G \cong G' \quad \& \quad H \cong H' \quad \Rightarrow \quad G \circ H \cong G' \circ H' \]

Statement 1) takes 'set' to mean 'abstract set', or, 'set of urelements' plus: it takes the sets \(|G|\) and \(|H|\) to be "independent".

**Aim:** a "comprehensive" set theory, of sets of urelements only, in which it is impossible to define the above operation \((G, H) \mapsto G \circ H\); in fact, all such operations on groups, say, are invariant under isomorphism.

**Note:** ☋ There will be things other than sets but: all groups will have underlying 'sets': abstract sets.

**Isomorphism:** \[ G \cong G' \]

\[ \iff \]

\[ E: \quad \begin{array}{ccc}
G \circ \text{Id}_G & = & G' \circ \text{Id}_{G'} \\
\circ & = & \circ \\
\text{Id}_G & \text{Id}_{G'} & \\
\end{array} \]
In the System:

equality is restricted to one set at a time.

\( A : \text{set} \mapsto =_A \text{equality on } A \)

\( a =_A b \) is meaningful only if we already know that \( a \in A, b \in A \).

Untyped \( = \) is not used.

Grammar: a theory of types, in fact dependent types

In one version: we have

\[
\begin{aligned}
type : \quad & \text{SET} \\
A \in \text{SET} : & \quad \text{declaration of variable } A \text{ as being of type SET} \\
\text{Note: } A \in \text{SET} & \text{ is not a proposition; e.g., } \neg (A \in \text{SET}) \text{ cannot be formed.}
\end{aligned}
\]

Under \( A \in \text{SET} \), we can form the dependent type \( E(A) \), and then declare the variable \( a \) as being of type \( E(A) \):
\[ a \in E(A) \]

Full variable declaration:

Example:

\[ A \in \text{SET}; a \in E(A) \]

(\( a \in E(A) \) can be abbreviated to \( a \in A \))

Example for a grammatical sentence:

\[ \forall A \in \text{SET}. \forall a \in E(A), \forall b \in E(A) (a =_A b \iff b =_A a) \]

In general:

\[ \text{FOLDS} = \]

First Order Logic with Dependent Sorts

(or "Types here")

I came out with this at the 1995 Haifa Logic Colloquium; papers:

Towards a Categorical Foundation of Mathematics

in: Logic Colloquium '95, Lecture Notes in Logic 11
1998; pp. 153-190

Monograph: First Order Logic with Dependent Sorts

Unpublished; available at: www.math.mcgill.ca/makai/
FOLDS has:

- variable signatures:
  a FOLDS signature $L$ is a special kind of category

- an $L$-structure is a functor $M : L \to \text{Set}$

- $\text{Str}_L = \text{the category of } L$-structures
  $= \text{Set}^L$

- concept of identity: $(\cong_L : L$-equivalence $M \cong_L N \iff \exists \theta : M \to N$ in $\text{Str}_L$)
  where $m$ and $n$ are fiberwise surjective (fs)

- $P \xrightarrow{m} M$ is fs if $\exists \theta : Q \to P$ such that $\theta \circ o = m$

  \[ \begin{array}{c}
    Q \\
    \downarrow o \\
    P \\
    \downarrow m \\
    M
  \end{array} \quad \begin{array}{c}
    Q \\
    \downarrow o \\
    P \\
    \downarrow m \\
    M
  \end{array} \]
A category \( L \) is a FOLDS signature \( \iff \)
\[
\text{def}
\]
1) has finite fan-out:
\[
\forall A \in \text{ob}(L). \ \{ f | \exists B. \ f : A \to B \} \text{ is finite}
\]
2) reverse well-founded:
no infinite ascending sequence
\[
A_0 \xrightarrow{f_0} A_1 \to \ldots \to A_n \xrightarrow{f_n} A_{n+1} \to \ldots
\]
with all \( f_n \neq \text{id}_{A_n} \).

For FOLDS signature \( L \), define \( L_n \subseteq \text{ob}(L) \):
\[
A \in L_n \iff A \not\subseteq \bigcup_{k < n} L_k
\]
& \( \forall f : A \to B, \ B \in \bigcup_{k < n} L_k \)

We have:
\[
\text{ob}(L) = \bigcup_{n \in \mathbb{N}} L_n \ j \ \dim A = n \ \iff \ A \in L_n \ \text{def}
\]

Have:
\[
A \xrightarrow{f} B \land f \neq \text{id} \ \Rightarrow \ \dim A > \dim B
\]
The syntax and the 'identity-semantics' of FOLDS are tightly linked:

- The statements expressible in FOLDS over $L$ are exactly the $\Sigma^1_1$-statements over $L$ that are invariant under $\equiv_L$.

3. Functions and Categories

New primitive: function

For sets $A, B$, we have (the concept of) a function $f$ from $A$ to $B$: $f: A \to B$

Grammar:

$$A \in \text{SET}; \quad B \in \text{SET}; \quad f \in A(A, B)$$

↑

'arrow' for 'function'

Note: 'relation $R$ between $A$ and $B$

$R \subseteq A \times B$

can be reproduced as a 'span' of functions:

$$\text{span} \quad \begin{array}{c} R \quad \begin{array}{c} \downarrow \quad \begin{array}{c} \text{jointly mono} \left\{ \begin{array}{c} 1-1 \end{array} \right. \end{array} \end{array} \end{array}$$

A \quad B$$
One possible way of proceeding: use

Kind (= type heading) Apply

Subject to:

\[ A, B \in \text{SET}; f \in A(A, B); a \in E(A); b \in E(B); \]
\[ \exists \in \text{Apply}(f, a, b) \]

Read as:

"\[ \exists \in \] is a witness to the fact that, applying the function \[ f : A \to B \] to the element \[ a \in A \], we obtain \[ b \in B \]."

For \[ A \to B \] [i.e., \( \langle A, B \in \text{SET}; f \in A(A, B) \rangle \)]

Abbreviate:

\[ f \cdot a = b \iff \exists \in \text{Apply}(f, a, b). \]

Note: this example shows how FOLDS handles relations (operations) in the theory, via types.
We can then define composition of functions:

\[ A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{\mu} A \xrightarrow{g \circ f} C \]

and also, equality of parallel functions:

for \( A \xrightarrow{f} B \):

\[ f = g \iff \forall a \in A. \forall b \in B. \left( f(a) = b \iff g(a) = b \right) \]

Alternatively, 'composition' can be adopted as primitive; 'elements' can be recovered as arrows:

\[ a \in A \iff 1 \xrightarrow{a} A \]

[one-element set \{*\}]

'application': for \( 1 \xrightarrow{a} A \xrightarrow{f} B \):

\[ f(a) = \text{def} \quad g \circ a \]
Topos theory, based on the notion of (elementary) topos (F.W. Lawvere and M. Tierney 1970) shows, although not in the above language, that much set theory, including set order arithmetic for all $n \in \mathbb{N}$, can be developed in the above context.

Conversely, one can understand topos theory as doing the theory of abstract sets and functions.

Example: Cartesian product $A \times B$ of sets $A, B$ cannot be defined as the set of ordered pairs $(a, b)$ for $a \in A$, $b \in B$ because, whatever they are, ordered pairs are not urelements. Rather: $C = A \times B$

comes with projections

$\pi_A : C \rightarrow A \quad [(a, b) \mapsto a]$  

$\pi_B : C \rightarrow B \quad [(a, b) \mapsto b]$
with "universal property": the span

\[ C = A \times B \]

\[ \pi_A \quad \pi_B \]

A \quad B

is terminal among all such spans:

\[ \forall \quad A \quad D \quad B \]

A \quad D \quad B

\[ \pi_A \quad \pi_B \]

C

\[ \pi_A \quad \pi_B \]

A \quad B

\[ \pi_A \quad \pi_B \]

D

(\(\pi_A \circ f = \pi_A\), \(\pi_B \circ g = \pi_B\))

\[ \exists! \quad A \circ f \circ B \]

A \quad B

\[ \pi_A \quad \pi_B \]

C

\[ \pi_A \quad \pi_B \]

D

\[ \pi_A \quad \pi_B \]

A \quad B

\[ \pi_A \quad \pi_B \]

D

\[ \pi_A \quad \pi_B \]

A \quad B

NOTE: \(A \times B\) is not (and, by such means, cannot) be defined uniquely, with

\[ \begin{array}{c}
 C \\
 D \cong C
\end{array} \]

one product, and

\[ \begin{array}{c}
 g f \\
 C 2 \text{id}_C = f g
\end{array} \]

is also a product
Briefly: \(A \times B\) is defined at most up to isomorphism.

In fact: \(A \times B\) is defined exactly up to isomorphism.

**Note:** 1) we do not need the 'usual' equality of sets (via extensionality)

2) the operative concept of identity for sets is isomorphism, (also) defined in terms of (other) primitives.

3) 'extensional' equality cannot be defined in the System.

---

What kind of totality is formed by (all) sets?

Answer: 1) it cannot be a set; no equality.

2) it must be a category!

partly, by the above way of talking about sets (and functions)

& mainly: "category is good."
4. "Reforming" category theory

4.1 Reformulate concept of "category":

Category I

\[
\begin{array}{ccc}
E & \xrightarrow{t_0} & T \\
\downarrow e_0 & & \downarrow t_1 \\
\downarrow e_1 & & \downarrow t_2 \\
A & \xrightarrow{d} & O \\
\downarrow d & & \downarrow c \\
\downarrow c & & \downarrow c \\
I & \xrightarrow{i} & A
\end{array}
\]

Subject to:
\[
\begin{align*}
dt_0 &= dt_2, & ct_0 &= dt_1, & ct_1 &= ct_2 \\
di &= ci \\
d_{e_0} &= de_1, & ce_0 &= ce_1
\end{align*}
\]

Illustration:
\[
M : \mathbb{L} \rightarrow \text{Set}
\]

\[
\begin{array}{ccc}
A & \xrightarrow{f} & B \\
\uparrow \gamma & & \uparrow g \\
C & \xrightarrow{h} & A \\
\end{array}
\]

\[
A, B, C \in M(0), \text{ written: } \mathbb{0} \\
f, g, h \in A \\
\gamma \in T \\
f = t_0 \gamma, \; g = t_1 \gamma, \; h = t_2 \gamma \\
(M_{t_0})(\gamma)
\]
\[ A = df = dh \]
\[ \Rightarrow \quad \Rightarrow \]
\[ dt_0 \alpha = dt_2 \alpha \]

because
\[ dt_0 = dt_2 \]

We write:
\[ \tau \in T(A, B, C; f, g, h) \]
well-defined only if
\[ A, B, C; f, g, h \] fit as in the figure

\[ (A = df = dh, \text{ etc.}) \]

For more, see: [Haifa paper] (above)

Also: For \( X : L \rightarrow \text{Set} \) a category,
\[ X, Y \in X(0), \]

\[ X \cong Y \Leftrightarrow \exists: \]

\[ \begin{array}{c}
\text{ordinary isomorphism in the cat } X \end{array} \]
4.2 Reform 'functor' to 'anafunctor'


[Page 9]

See also:
[Haife paper]

5. The universe

Just as: sets do not form a set, they form a category (they are objects of a category).

Categories do not form a category, they form a 2-category.

Question: What is a 2-category?

In general: \(n\)-categories form (are the objects of) an \((n+1)\)-category. \((*)\)

Question: What is an \((n+1)\)-category even assuming we know what \(n\)-categories are?
(The only?) answer to the Question
Subject to (*)
is in:

[M.M: The multilokic omega-category
of all multilokic omega-categories.

16.1, 16.2

The subject of Higher Dimensional Categories:
(HDC's)

0 largely independent of the Foundations;

0 several different notions of HDC;

[but: they do not (I contend) attain
the level of completeness of the
concept in [loc.cit.]]

0 interest in: algebraic topology
mathematical physics
quantum groups
algebraic geometry
Based on:


Part 1: 153 (2000), 221-246
Available as preprint: Summer 1997]

Gives theory of multitopic sets.

A multitopic (omega-) category of previous [M.M. 99] is a multitopic set with additional properties (only).

The multitopic theory was inspired by,


See also:

[M.M. G99] constructs a (large) multifopic set whose 0-cells are the (small) multifopic categories. Call this \textsc{muf Cat}.

The main theorem, stated but not proved in [M.M. G99] is

\begin{center}

Theorem. \textsc{muf Cat} is a multifopic category.

\end{center}

The notion of multifopic set is given a concise, conceptual characterization in

[V. Harnik, M. M., M. Zawadowski: Multifopic sets are the same as many-to-one computads.
185 pages]

This "monster" paper is to be revised; we now know how to break it into independent pieces.

(back to 16)
Peter May (University of Chicago), in September, 2001, called for an international coordination of effort in the cause of HDC's, and specifically asked for:

- comparing rigorously the various definitions of HDC;
- defining, in whichever context one cultivates, the (large) \((n+1)\)-category of all (small) \(n\)-categories as a test of "completeness" of the definition.

My "multitopic" answer (which certainly incorporates work and ideas by other people—as I will indicate) to: "what is a HDC?" is deeply dependent on FOLDS: not only it is "given" in the framework of FOLDS, but the postulated structure of a HDC is defined in terms of FOLDS-equivalences (identity concepts).
Doctrine of identity

Requirements:

1) We have a language talking about (many different) interdependent types of objects of the universe;

2) Each type $T$ has its associated concept of identity $\approx_T$, a binary relations on objects of type $T$;

3) The two (language and identity) are related by Leibniz's law of indiscernibility of identicals (invariance under identity):

$$A \approx_T B \land P(A) \Rightarrow P(B)$$

for any predicate $P(\cdot)$ [for which both $P(A)$ and $P(B)$ are meaningful] which is expressible in the language;

4) The language is "maximal" with respect to 3) ("expressive completeness").

At: www.math.mcgill.ca/makkai/]}

is another survey paper (like [Haifa paper]), but more sketchy. It was written in response to P. May's attempt at organizing a HDC group. It contains the description (fairly natural on the basis of the capabilities of FOLDS) how to use FOLDS equivalence to state rigorously that two FOLDS-based definitions of HDC are identical.
6. **Strict \( n \)-categories**

With full use of equality, our question has a simple answer, already known in the 1970's (or earlier?)

\[
\text{Strict } (n+1)-\text{category} \overset{\text{def}}{=} \text{a category enriched in } \text{Cat}_n
\]

the (ordinary) category of all \( n \)-categories

\[
\text{Strict } 2\text{-category} = \text{a cat enriched in } \text{Cat} (= \text{Cat}_1)
\]

\[
\text{Strict } 3\text{-category} = \text{a cat enriched in } \text{Cat}_2
\]

\( = \text{the category of strict } 2\text{-cats} \)

\( X \) being enriched in \( C \) means:

for objects \( A \) and \( B \) of \( X \),

\( \text{hom}_X(A,B) \) is not just a set,

but an object of \( C \)

— with further obvious requirements on the composition structure of \( X \).
Also: $\text{Cat}_n$ is enriched over itself

(in fact, $\text{Cat}_n$ is Cartesian closed)

- So, what is the problem?:
  the over-use of equality

7. $n = 2$

In a strict 2-category $\mathcal{X}$, for objects $A$ and $B$, arrows $A \rightarrow B$ are the objects of the category $\text{Hom}_\mathcal{X}(A, B)$.

The associative law for composition:

$$
(f \cdot g) \cdot h = f \cdot (g \cdot h)
$$

is a BAD THING, from our foundational point of view, — and it just plain DOES NOT HAPPEN in certain important cases (see below)
Rather, we (should) have an isomorphism:

\[(fg)h \cong f(gh)\]

of objects of the category \(\mathcal{F}(A, D)\).

In fact, we (should) have a specified isomorphism 2-cell

\[
\begin{array}{ccc}
(fg)h & \cong & f(gh) \\
\alpha_{f, g, h} & & \\
\end{array}
\]

and we (should) make the operation

\[
(f, g, h) \mapsto \alpha_{f, g, h}
\]

part of the structure of the "2-category". This is how we arrive at the bi-categories of Jean Benabou (1967). Further to the concept of "bicategory": consider now four composable arrows:

\[
A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D \xrightarrow{i} E
\]

They give rise to:
Mac Lane's pentagon:

\[(fg)(hi) \xrightarrow{\alpha_{fg,h,i}} (fg)h \xrightarrow{\alpha_{f,g,h}} f(g(h))i \xrightarrow{\alpha_{f,g,hi}} f((gh)i) \xrightarrow{\alpha_{f,gh,i}} f((gh)i)\]

In a bicategory, this is required to commute, expressing a coherence condition on the operation \((f, g, h) \mapsto \alpha_{f,g,h} \).

The full definition of "bicategory" involves, in addition, specified isomorphisms \(\Delta\) and \(\epsilon\), and further coherence conditions involving \(\Delta, \lambda, \gamma\) (including 'naturality' rules).
Mac Lane coherence theorem (<1971):

In a tricategory, all "coherence diagrams" commute.

Example: Starting with five composable arrows:

\[
A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D \xrightarrow{i} E \xrightarrow{j} F
\]

one can generate a "coherence diagram".

And many more.

The paper


contains a stronger version, and a conceptual proof, of the Mac Lane coherence theorem.

8. \( n = 3 \)

[R. Gordon, A. J. Power and R. Street:

Coherence for Tricategories. Memoirs of the A.M.S. 117 no 558, 1995]
The above definition of (strategy) as a kind of hypocrisy in the spirit of contemporary liberalism is, in itself, a kind of hypocrisy, and it is reasonable to ask if it is possible to define something essentially equivalent to hypocrisy in the context of a coherent conceptual framework. This is what we suggested earlier.

But it does not explicitly state that coherence is not necessary. The concept of strategy is not a merely abstract notion, and the coherence of the whole thing is essential. The concept of strategy is not just a collection of ideas or concepts, but it is a coherent whole.

The key to understanding strategy is the idea of the whole, the whole idea of strategy is not just a series of concepts, but it is a coherent whole, a whole thing. The key is to see why coherence is not just a collection of ideas, but a whole concept that is the whole thing.

P.S. This is my collaboration.
It is important to emphasize that the right abstraction is something that is essentially equivalent to the classical notion of bicategory.