(CHICAGO 2003: ASL Annual Meeting)

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 $(G, H) \longmapsto G \square H$ as follows.
Underlying set:

$$|G \square H| = \begin{cases} |G| \cap |H| & \text{if this} \neq \emptyset \\ \{1\} & \text{otherwise} \end{cases}$$

The operation on $G \square H$ is defined as …"

BAD definition:

1) elements of G and H are abstract, and (should) have nothing to do with each other;
2) operation $\square$ should be, and it isn't, invariant under isomorphisms.

(Invariance under isomorphism:

$$G \cong G' \ \& \ H \cong H' \Rightarrow G \square H \cong G' \square H')$$

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

<u>AIM</u>: a "comprehensive" set theory, of sets of urelements <u>only</u>, in which it is <u>impossible</u> to define the above operation $(G, H) \longmapsto G \square H$; in fact, <u>all</u> such operations on groups, say, <u>are</u> invariant under isomorphism.

NOTE: ⊙ There <u>will</u> be things other than sets but: <u>all</u> groups will have underlying 'sets': abstract sets.

⊙ Isomorphism: $G \cong G'$

$$\Longleftrightarrow$$

$\exists$:

$$fg = \mathrm{Id}_G \circlearrowleft G \underset{f}{\overset{g}{\rightleftarrows}} G' \circlearrowright \mathrm{Id}_{G'} = gf$$

In the System:

equality is restricted to
one set a time.

A: set $\rightsquigarrow$ $=_A$ equality on A

$a =_A b$ is meaningfull 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

type: SET

$A \in SET$: declaration of variable A as being of type SET

Note: $A \in SET$ is not a proposition; e.g., $\neg(A \in SET)$ cannot be formed.

Under $A \in SET$, 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:

$A \in SET;\ a \in E(A)$

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

Example for a grammatical sentence:

$$\forall A \in SET.\ \forall a \in E(A).\ \forall b \in E(A) \left(a =_A b \Leftrightarrow b =_A a \right)$$

In general: FOLDS =

First Order Logic with Dependent Sorts

(or: 'Types' here)

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

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/makkai/

$\boxed{\text{FOLDS has:}}$

- variable signatures:

  a FOLDS <u>signature</u> L is a special kind of category ⊛

- an L-structure is a functor

$$M : L \longrightarrow \mathcal{S}et$$

- Str_L = the category of L-stuctures

$$= \mathcal{S}et^{L}$$

- concept of identity: $\simeq_L$: L-equivalence

$$M \simeq_L N \underset{\text{def}}{\Longleftrightarrow} \exists : \begin{tikzcd} & P \arrow[dl, "m"'] \arrow[dr, "n"] & \\ M & & N \end{tikzcd} \quad \underset{\sim}{\text{in}} \ \mathrm{Str}_L$$

where m and n are fiberwise surjective (fs)

- $P \xrightarrow{m} M$ is fs if

$$\forall : \begin{tikzcd} Q \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\circ"] & P \arrow[d, "m"] \\ R \arrow[r] & M \end{tikzcd} \quad \exists : \begin{tikzcd} Q \arrow[r] \arrow[d, tail] & P \arrow[d] \\ R \arrow[r] \arrow[ur] & M \end{tikzcd}$$

$\boxed{\text{in words:}}$ m has the Right Lifting Property with respect to 1-1 maps $Q \rightarrowtail R$.

A category L is a FOLDS signature

$\underset{\text{def}}{\Longleftrightarrow}$

1) has finite fan-out:

$$\forall A \in \mathrm{Ob}(L).\ \{f \mid \exists B.\, f : A \to B\} \text{ is finite}$$

2) reverse well-founded:

no infinite ascending sequence

$$A_0 \xrightarrow{f_0} A_1 \to \cdots \to A_n \xrightarrow{f_n} A_{n+1} \to \cdots$$

with all $f_n \neq \mathrm{id}_{A_n}$.

For FOLDS signature L, define $L_n \subseteq \mathrm{Ob}(L)$:

$$A \in L_n \Leftrightarrow A \notin \bigcup_{k<n} L_k$$

$$\& \ \forall f : A \to B,\ B \in \bigcup_{k<n} L_k$$

We have:

$$\mathrm{Ob}(L) = \dot{\bigcup_{n\in\mathbb{N}}} L\ ; \quad \dim A = n \underset{\text{def}}{\Longleftrightarrow} A \in L_n$$

Have:

$$A \xrightarrow{f} B \ \& \ f \neq \mathrm{id} \Rightarrow \dim A > \dim B$$

The syntax and the 'identity-semantics' of FOLDS are tightly linked:

the statements expressible in FOLDS over L are exactly the Σ^1_1-statements over L that are invariant under $\simeq_L$.

This is wrong!

3. Functions and categories

New primitive: function

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

Grammar:

$$A \in SET;\ B \in SET;\ f \in A(A,B)$$

'arrow' for 'function'

Note: 'relation R between A and B'

$$R \subseteq A \times B$$

can be reproduced as a 'span' of functions:

$$\begin{tikzcd} & R \arrow[dl] \arrow[dr] & \\ A & & B \end{tikzcd}$$

(jointly monic 1-1)

$\boxed{\text{One possible}}$ way of proceeding: use

Kind (= type heading) Apply

subject to:

$$A, B \in SET;\ f \in A(A,B);\ a \in E(A);\ b \in E(B);$$
$$\tau \in Apply(f, a, b)$$

read as:

"τ 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 $A \xrightarrow{f} B$ [i.e., $\langle A, B \in SET; f \in A(A,B) \rangle$]

abbreviate:

$$f'a = b \underset{def}{\Longleftrightarrow} \exists \tau \in Apply(f, a, b).\ T$$

'TRUE'

[Note: this example shows how FOLDS handles relations (operations) in the theory, via types.

We can then <u>define</u> composition

of functions:

$$A \xrightarrow{f} B \xrightarrow{g} C \quad \rightsquigarrow \quad A \xrightarrow[g \circ f]{} C$$

and also, equality of parallel functions:

for $A \underset{g}{\overset{f}{\rightrightarrows}} B$:

$$f = g \underset{def}{\iff} \forall a \in A .\ \forall b \in B\ (f' a = b \leftrightarrow g' a = b)$$

$\boxed{\text{Alternatively:}}$ '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

$$1 \xrightarrow{a} A \xrightarrow{f} B \ : \quad f' a \underset{def}{=} f \circ a$$

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 nth 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

<u>Example</u>: 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: $\mathbf{C}\ (= A \times B)$ comes with $\boxed{\text{projections}}$

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

$$\pi_B : C \to B \quad [(a, b) \mapsto b]$$

with "universal property": the span

$$\begin{tikzcd} & C = A\times B \arrow[dl, "\pi_A"'] \arrow[dr, "\pi_B"] & \\ A & & B \end{tikzcd}$$

is terminal among all such spans:

$$\forall \quad \begin{tikzcd} & D \arrow[dl, "p_A"'] \arrow[dr, "p_B"] & \\ A & & B \end{tikzcd}$$

$$\exists! \quad \begin{tikzcd} & C \arrow[dl, "\pi_A"'] \arrow[dr, "\pi_B"] & \\ A \arrow[r, phantom, "\circ"] & {} \arrow[r, phantom, "\circ"] & B \\ & D \arrow[ul, "p_A"] \arrow[ur, "p_B"'] \arrow[uu, "f"] & \end{tikzcd} \qquad (\pi_A \circ f = p_A,\ \pi_B \circ f = p_B)$$

<u>NOTE:</u> $A\times B$ is not (and, by ‘such means’, cannot) be defined $\boxed{\text{uniquely}}$: with

$$\begin{tikzcd} & C \arrow[dl, "\pi_A"'] \arrow[dr, "\pi_B"] & \\ A & & B \end{tikzcd}$$ one product, and

$$D \cong C \ : \ \begin{tikzcd} D \arrow[loop left, "gf = id_D"] \arrow[r, shift left, "f"] & C \arrow[l, shift left, "g"] \arrow[loop right, "id_C = fg"] \end{tikzcd}$$

$$\begin{tikzcd} & D \arrow[dl, "\pi_A \circ f"'] \arrow[dr, "\pi_B \circ g"] & \\ A & & B \end{tikzcd}$$ is $\boxed{\text{also}}$ a product

Briefly: $A \times B$ is defined at most up to isomorphism.

In fact: $A \times B$ is defined – exactly – up to isomorphism.

<u>NOTE</u>: 1) we do not need the 'usual' equality of sets (via extensionality)

2) the operative concept of identity for sets is isomorphism, (also) $\boxed{\text{defined}}$ 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 <u>set</u>; no (useful) <u>equality</u>.

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 Reform(ulate) concept of "category":

L_{cat}:

$$
\begin{tikzcd}
I \arrow[dr, "i"'] & E \arrow[d, shift right, "e_0"'] \arrow[d, shift left, "e_1"] & T \arrow[dl, shift right=2, "t_0"'] \arrow[dl, "t_1"] \arrow[dl, shift left=2, "t_2"] \\
 & A \arrow[d, shift right, "d"'] \arrow[d, shift left, "c"] & \\
 & O &
\end{tikzcd}
$$

subject to:

$$dt_0 = dt_2 \ , \ ct_0 = dt_1 , \ ct_1 = ct_2$$

$$di = ci$$

$$de_0 = de_1 \ , \ ce_0 = ce_1$$

Illustration: $M : L \longrightarrow Set$

$$
\begin{tikzcd}
 & B \arrow[dr, "g"] \arrow[d, phantom, "\tau"] & \\
A \arrow[ur, "f"] \arrow[rr, "h"'] & {} & C
\end{tikzcd}
$$

$A, B, C \in M(O)$, written: O

$f, g, h \in A$

$\tau \in T$

$f = \underbrace{t_0 \tau}_{(Mt_0)(\tau)} \ , \ g = t_1 \tau , \ h = t_2 \tau$

$$
\begin{array}{ccc}
A = df & \overset{?}{=} & dh \\
\parallel & & \parallel \\
dt_0\tau & \overset{\checkmark}{=} & dt_2\tau
\end{array}
$$

because

$$dt_0 = dt_2 \ \checkmark$$

We write:

$$\tau \in \underbrace{T(A, B, C; f, g, h)}_{\text{well-defined } \underline{\text{only}} \text{ if}}$$

$A, B, C; f, g, h$ fit as in the figure

$(A = df = dh$, e.t.c.)

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

<u>Also</u>: For $\mathbb{X} : L \to Set$ a $\boxed{\text{category}}$

$X, Y \in \mathbb{X}(O)$,

$X \cong Y \iff \exists$:

$$
\left(
\begin{tikzcd}
 & & u \arrow[dll, mapsto] \arrow[drr, mapsto] & & \\
 & & P \arrow[dl, "\mu"'] \arrow[dr, "\nu"] & & \\
X & \mathbb{X} \arrow[r, phantom, "o"] & & \mathbb{X} \arrow[l, phantom, "o"] & Y \\
 & & \mathbb{X} \arrow[ul, "id_{\mathbb{X}}"] \arrow[uu, "i"] \arrow[ur, "id_{\mathbb{X}}"'] & &
\end{tikzcd}
\right)
\quad
\begin{array}{l}
u \in P(O) \\
\mu_O(u) = X \\
\nu_O(u) = Y
\end{array}
$$

ordinary isomorphism in the cat $\mathbb{X}$

4.2 Reform 'functor' to 'anafunctor'

See: [M.M.: "Avoiding the axiom of choice in general category theory". J. Pure and Applied Algebra 108 (1996), 109 – 173]

[p 9]

See also: [Haifa 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 multitopic omega-category
of all multitopic omega-categories.
September, 1999. At: www.math.mcgill.ca/makkai/

! → (*) 16.1, 16.2 —·—

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

⊙ largely independent of the Foundations;

⊙ several different notions of HDC;

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

⊙ interest in: algebraic topology
mathematical physics
quantum groups
algebraic geometry
⋮

(*) Based on:

[C. Hermida, M.M., A.J. Power: On weak higher dimensional categories I. J. Pure and Applied Algebra.

Part 1: **153** (2000), 221–246

Part 2: **157** (2001), 247–277.

Part 3: **166** (2002), 83–104.

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

[J.C. Baez and J. Dolan. Announcement. Nov 30 & Dec 3, 1995; http://math.ucr.edu/home/baez/]

See also:

[J.C. Baez and J. Dolan, Higher-Dimensional Algebra III. n-Categories and the algebra of opetopes. Advances in Math. **135** (1998), 145–206

[M.M. '99] constructs a (large) multitopic set whose 0-cells are the (small) multitopic categories. Call this

mlt Cat

The main theorem, stated but not proved in [M.M. '99] is

$\boxed{\text{Theorem. mlt Cat is a multitopic category}}$

—·—

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

[V. Harnik, M.M., M. Zawadowski: Multitopic sets are the same as many-to-one computads. June 2002. At: www.math.mcgill.ca/makkai/ 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).

$\boxed{\text{Doctrine of identity}}$ 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 $\simeq_T$, 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)

$$A \simeq_T B \;\&\; P(A) \Rightarrow P(B)$$

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").

[M.M.: On comparing definitions of "weak n-category. Sept. 2001 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 <u>identical</u>.

6. Strict n-categories

WITH FULL USE OF EQUALITY,

our Question has a simple answer, already known in the 1970's (or earlier?)

Strict $(n+1)$-category $\underset{\text{def}}{=}$

a category enriched in Cat_n

the (ordinary) category of all n-categories

Strict 2-category = a cat enriched in Cat $(= Cat_1)$

Strict 3-category = a cat enriched in Cat_2

(= the category of strict 2-cats)

$\mathbb{X}$ being enriched in $\mathbb{C}$ means:

for objects A and B of $\mathbb{X}$,

$\hom_{\mathbb{X}}(A, B)$ is not just a set,

but an object of $\mathbb{C}$

– with further obvious requirements on the composition structure of $\mathbb{X}$.

Also: Cat_n is enriched over itself

(in fact, Cat_n is Cartesian closed)

$\therefore$ Cat_n is an $(n+1)$-category

— so, what is the problem?:

the <u>over</u>-use of equality

7. $n = 2$

In a strict 2-category $\mathbb{X}$, for objects A and B, arrows $A \to B$ are the <u>objects</u> of the category $\hom_{\mathbb{X}}(A, B)$.

The associative law for composition:

$$\begin{tikzcd} & B \arrow[r, "g"] & C \arrow[dr, "h"] & \\ A \arrow[ur, "f"] \arrow[urr, bend right, "fg"'] \arrow[rrr, bend right=40, "(fg)h"'] & & & D \end{tikzcd} \quad = \quad \begin{tikzcd} & B \arrow[r, "g"] \arrow[drr, bend right, "gh"'] & C \arrow[dr, "h"] & \\ A \arrow[ur, "f"] \arrow[rrr, bend right=40, "f(gh)"'] & & & D \end{tikzcd}$$

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 <u>isomorphism</u>:

$$(fg)h \cong f(gh)$$

of objects of the category $\hom_{\mathcal{K}}(A, D)$.

In fact, we (should) have a <u>specified</u> isomorphism 2-cell

$$(fg)h \xrightarrow[\alpha_{f,g,h}]{\cong} f(gh)$$

and we (should) make the operation

$$(f, g, h) \longmapsto \alpha_{f,g,h}\,.$$

part of the structure of the "2-category". This is how we arrive at the <u>bi</u> 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:

$$
\begin{tikzcd}
 & & (fg)(hi) \arrow[drr, "\alpha_{f,g,hi}"] & & \\
((fg)h)i \arrow[urr, "\alpha_{fg,h,i}"] \arrow[dr, "\alpha_{f,g,h} \circ i"'] & & & & f(g(hi)) \\
 & (f(gh))i \arrow[rr, "\alpha_{f,gh,i}"'] & & f((gh)i) \arrow[ur, "f \circ \alpha_{g,h,i}"'] &
\end{tikzcd}
$$

In a bicategory, this is required to commute, expressing a $\boxed{\text{coherence condition}}$ on the operation $(f, g, h) \longmapsto \alpha_{f,g,h}$. The full definition of "bicategory" involves, in addition, specified isomorphisms λ and ρ, and further coherence conditions involving α, λ, ρ (including 'naturality' rules).

Mac Lane coherence theorem (< 1971):

In a bicategory, 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 further "coherence diagram".

And many more.

The paper

[A. Joyal & R. Street: Braided Tensor Categories. Advances in Math. 102 (1993), 20-78]

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]

(continues slide 24)

The above definition of 'bicategory' conforms to the spirit of contemporary algebra. It's style is a kind of ideal which is followed in the much bigger definition of 'tricategory' in

[R. Gordon, J. Power, R. Street]

But it is reasonable to ask if it is possible to define something essentially equivalent to bicategory in a way that does not explicitly pack in (parts of) coherence, but rather, has coherence as a consequence from something 'conceptual'.

The answer is 'yes'. This is what was suggested and partly done by J. Baez and J. Dolan (see above), and was ~~the suggestions of~~ worked out in my (& my collaborators') work on multitopics.

The key is a return to some key examples, and see why coherence is true in those cases, and make a suitable abstraction of the whole situation of the example. The 'mistake' in the official concept of bicategory is that too much is abstracted away from the examples. o/

It is important to emphasize that the 'right abstraction'

is something that is essentially equivalent

to the classical notion of bicategory.