Mathematics in Three Types,

or, doing without ordered pairs

M. Randall Holmes

September 12, 2012

Aims of this talk

We plan to discuss mathematical constructions, notably the definition of relations, functions, and cardinals, in the simple theory of types using only three types. This is somewhat tricky because the usual definitions of the ordered pair require at least 3 types to define a pair, and so 4 types before any implementation of a relation or a function as a set of ordered pairs is feasible. Nonetheless, quite a lot can be done.

I have given this talk before (or something like it) at a conference at Boise State several years ago. At that time I was not aware of Allen Hazen's paper "Relations in Monadic Third-Order Logic", which says some closely related things though not always in the same way. Everything here is independent of Hazen though he can certainly claim prority for some of the results.

A few words about "New Foundations": the simple theory of types

This is not a talk about New Foundations, the infamous set theory proposed by W. v.O. Quine in 1937, but I believe I have been sold to this audience as an expert on this topic, and there is some incidental relation to the motivation of what I am actually talking about.

The simple theory of types (TST) is a simple typed theory of sets. It is arguable that it is sketched by Russell in Principles of Mathematics. It is certainly *not* the type theory of Principia Mathematica. It was first formally presented (usually with some extraneous features) by various logicians around 1930.

simple type theory continued...

The objects of TST live in sorts indexed by the natural numbers. TST is a first order theory with equality and membership as primitive relations. The sorts enter in because the sorting of an atomic formula must satisfy one of the templates $x^n = y^n$; $x^n \in y^{n+1}$.

The axioms of TST are extensionality (sets of any positive type with the same elements are the same) and comprehension (for any formula ϕ in the language of the theory, $\{x^n \mid \phi\}^{n+1}$ exists).

simple type theory continued...

Axioms of Infinity and Choice are usually added, but we do not regard these as part of the base theory TST. In the first presentations of this theory, it was usual to add the Peano axioms on type 0 objects, producing the more specific theory "arithmetic of order ω". I have yet to see a presentation of pure TST older than the one in Quine's New Foundations paper of 1937, though I am assured that it was described.

TST is a beautifully simple foundational system. The consistency strength of TST + Infinity + Choice is exactly the same as that of Mac Lane set theory (Zermelo set theory with Δ_0 separation).

For any formula ϕ in the language of TST, let ϕ^+ be the formula obtained by raising the type of every variable appearing in ϕ by one. Clearly this will remain well-formed. Further, it is clear that if ϕ is an axiom, so is ϕ^+, and if ϕ entails ψ by a logical rule, then ϕ^+ will entail ψ^+. From this it follows that if a sentence ϕ is a theorem, ϕ^+ is also a theorem.

It is also the case that for any mathematicial object defined by a set abstract $\{x \mid \phi\}$, there is an exactly analogous object $\{x \mid \phi^+\}$ in the next higher type.

For these reasons it is common on occasions when this theory is actually used to suppress the type indices, because anything proved about one type can be proved about any higher type, and the relative types of variables in a formula can usually be deduced from context. This is related to the "systematic ambiguity" practiced by Russell and Whitehead in Principia, but in TST it is especially simple.

New Foundations

Observing this, Quine proposed the theory New Foundations, which is an unsorted first-order theory with equality and membership as primitives, whose axioms are precisely the formulas which can be obtained by dropping all distinctions of type between variables in an extensionality or comprehension axiom of TST (without introducing identifications between variables of different types). Note that "$\{x \mid x \notin x\}$ exists" (the instance of naive comprehension yielding the Russell paradox) is not an axiom of New Foundations because $x \notin x$ cannot be obtained from a formula of the language of TST by dropping type distinctions. Formulas which can be so obtained are called "stratified" formulas, and it is traditional to give a definition of stratified formula for NF which does not depend on the language of a different theory TST, but we will not do that here.

Results of Specker

Specker showed that the consistency strength of NF is exactly that of the version of TST in which we add the axiom scheme of ambiguity $\phi \leftrightarrow \phi^+$ for each sentence ϕ. In TST we know that ϕ^+ is *provable* if ϕ is provable, but we do not know that ϕ^+ is true if ϕ is true (and under such reasonable further assumptions as the axiom of choice, we can show that this is not always true).

Specker also showed that the Axiom of Choice can be refuted in NF, from which it follows that the Axiom of Infinity is a theorem.

One reason to be interested in the mathematics of three types is that one of the fragments of Quine's "New Foundations" which is known to be consistent is NF_3, shown to be consistent by Grishin, in which only those instances of comprehension are used which would make sense in the theory of types with just three types.

For any n (for example, $n = 3$), TST_n is the theory obtained from TST by cutting down the language to those formulas mentioning only the first n types (0 to $n - 1$) and the axioms to the axioms of TST expressible in this language.

TST_3

The theory of types with three types is a 3-sorted theory with sorts called "type 0", "type 1", and "type 2". Where i is 0 or 1, $x^i \in y^{i+1}$ is a well-formed membership sentence. Where i is 0,1,2, $x^i = y^i$ is a well-formed identity statement. We will not as a rule actually put type indices on variables; they will usually be deducible from context.

The axioms are Extensionality:

$$(\forall A.(\forall B.(\forall x.x \in A \leftrightarrow x \in B) \to A = B))$$

where the type of A, B is one higher than the type of x, and Comprehension:

$$(\exists A.(\forall x.x \in A \leftrightarrow \phi)),$$

where ϕ is any formula in which A does not occur free, the type of A being one higher than the type of x. The object A whose existence is asserted (unique by Extensionality) is called $\{x \mid \phi\}$.

This is a first order version of what is commonly called "third-order logic".

NF_n is obtained from TST_n in the same way that NF is obtained from TST: the comprehension axioms of NF_n are those which can be obtained by dropping all type distinctions from a comprehension axiom of TST_n.

It is a result of Grishin that NF_4 is the same theory as NF. The most elegant way to show this (though it is not how Grishin did it) is to present a finite axiomatization of NF comprehension in which no axiom used more than four types. This has been done by Yasuhara, I believe. The consistency problem for NF has not yet been solved, and has acquired a reputation as a fearsomely difficult problem.

By contrast, every externally infinite model of TST_3 satisfies the ambiguity scheme $\phi \leftrightarrow \phi^+$ (where of course ϕ will mention only two types) and can be shown to have the same theory as a model of NF_3. To my mind, this makes NF_3

a quite interesting theory (it is found everywhere!) and an interest in finding out what mathematics can be done in NF_3 is a large part of the motivation for the work I report here, though I actually present it in a typed context.

As noted briefly earlier, we cannot define a relation or function as a set of ordered pairs in TST_3, because the ordered pair $\langle a, b\rangle = \{\{a\}, \{a, b\}\}$ is only defined for a and b of type 0, and since it is a type 2 object it cannot be a member of any set in this theory.

We could attempt to define relations using sets of *unordered* pairs $\{a, b\}$; of course, only symmetric relations could be defined in this way.

If R is a symmetric relation ($x\,R\,y \leftrightarrow y\,R\,x$) then we can define R^* (the set implementing R) as $\{\{a, b\} \mid a\,R\,b\}$. Of course, only relations on type 0 objects can be implemented in this way (and as a rule it is only structures on type 0 objects that we will attempt to implement).

We can implement any partial order $\leq$ as the collection of its (weak) segments: if R is a reflexive, antisymmetric, transitive relation, define R^* as the collection $\{\{y \mid x\,R\,y\} \mid x = x\}$. The weak segment is preferred because it is possible to distinguish between an ordering of a set with one element and an ordering of the empty set.

Quasi-Orders

More generally, any quasi-order (reflexive, transitive relation) can be implemented. If R is a reflexive, transitive relation, define R^* as $\{\{y \mid x \, R \, y\} \mid x = x\}$, as above.

If R is a set, define $x \, R \, y$ as $(\forall A \in R. x \in A \rightarrow y \in A)$.

Quasi-orders include equivalence relations and partial orders, including linear orders and well-orderings.

Functions and Cardinality

In one special case the implementation of functions (and of the notion of cardinality) is quite easy. If A and B are *disjoint* sets, we can implement any function $f : A \to B$ using the set $f^* = \{\{x, f(x)\} \mid x \in A\}$. The lack of directionality makes no difference, since we are not in doubt as to what is the domain and what is the range (if we were to interchange domain and range, exactly the same set would serve to represent f^{-1}).

So we can define $f : A \to B$, A and B being disjoint sets, as $(\forall x \in A.(\exists y \in B.\{x, y\} \in f)) \wedge (\forall x \in A.(\forall y, z \in B.\{x, y\} \in f \wedge \{x, z\} \in f \to y = z))$. We can define $A \sim B$ as $(\exists f.f : A \to B \wedge f : B \to A)$.

Pabion used the preceding analysis of cardinality in his proof that NF_3 + Infinity is equiconsistent with second order arithmetic , with the additional observation that for finite sets $A \sim B \leftrightarrow A - B \sim B - A$, so the general case reduces to the disjoint case.

Henrard showed (unpublished work) that cardinality can be defined for all sets, finite and infinite, disjoint and overlapping.

Our idea is to represent a function f using its set of forward orbits $\{\{f^n(x) \mid n \geq 0\} \mid x \in \texttt{dom}(f)\}$. This does not quite work, as we will see, but it does allow for a complete definition of cardinality and an “almost” complete definition of function.

Historical digression about Henrard's approach

Henrard also used a representation of orbits in a bijection: he uses the idea of a "chain", which is the set of unordered pairs $\{x, f(x)\}$ in an orbit in the function. We look at how to express this concept without reference to functions: if A is a set of two-element sets such that no element of $\bigcup A$ belongs to more than two elements of A, then A is a *union of chains*. An element of $\bigcup A$ which belongs to only one element of a union of chains A is called an *endpoint* of A.

Historical digression about Henrard's approach (cont.)

A closed chain is a nonempty union of chains with no endpoints, no proper subset of which is a nonempty union of chains with no endpoints. Any chain has the property that it has no proper subset which has no endpoints. A union of chains which has one endpoint and has no subset which is a closed chain is a chain. A union of chains which has two endpoints and which has no proper subset with no endpoints or one endpoint is a chain. These tools can be used to cover much the same ground as ours, however there is a disadvantage that there is no representation of the distinction between f and f^{-1} (for us, this distinction collapses only in (some) finite cycles in f).

We can then say that $A \sim B$ iff there is a set of chains such that each element of $A \cup B$

participates in exactly one of the chains (and nothing else participates in any of them) and each chain either has one endpoint in $A - B$ and one in $B - A$ or is closed and entirely in $A \cap B$.

The material about Henrard's approach was added here after I had completed the work on the approach I present.

Our approach, continued

If F is any definable function (think of this as implemented by a formula $F(x, y)$ with appropriate properties), define O_x^F as $\{y \mid (\forall A.x \in A \wedge (\forall z.z \in A \wedge z \in \mathtt{dom}(F) \rightarrow F(z) \in A) \rightarrow y \in A)\}$. Define F^* as $\{O_x^F \mid (\exists y.y = F(x)) \vee (\exists y.x = F(y))\}$.

Note that I do need orbits (taken to be singletons) for elements of the range of F which are not in the domain.

Orbits O_x^F are of two kinds. There are finite sets (among which the sets of size 1 and 2 are special) and there are infinite sets. If $\{x\}$ is an orbit, then $F(x) = x$. If $\{x, y\}$ is an orbit and $\{x\}$ is not, then $F(x) = y$. From the other finite orbits, we cannot determine a function value.

If O_x^F is not a finite set, then the distinguishing characteristic of $F(x)$ is that $O_x^F - \{x\} = O_y^F$. To identify O_x^F among the sets in F (many of which contain x) observe that it is the intersection of all elements of F that contain x.

It is useful to pause and observe that the notion of finite set is definable in TST_3: the set `Fin` can be defined as the set of all sets which contain $\emptyset$ and contain all sets $x \cup \{y\}$ whenever they contain x.

We have already noted that the notion of equinumerousness of finite sets is definable, so the cardinal of each type 1 finite set is already definable as a type 2 set.

Strictly speaking, one does not need to allude to the notion of finite set in the definitions which follow.

We define a first approximation to function application. Where F is a set and x is an element of $\bigcup F$, we define $F[x]$ as

x, in case $\{x\} \in F$

y, in case $\{x, y\} \in F$ and $\{x\} \not\in F$

For the next case, we need to define O_x^F as $\bigcap\{A \in F \mid x \in A\}$:

the unique y such that $O_y^F = O_x^F - \{x\}$, if this exists

else x, when none of the special conditions above hold.

If F is a definable function and F^* is defined as above, $F^*[x] = F(x)$ is true except in two special cases:

If x is in the range of F but not in the domain of F then $F^*[x] = x$ will hold: knowledge of the intended domain and range of F makes this harmless.

More annoyingly, if x is in a finite orbit in F with three or more elements, $F^*[x] = x$ rather than $F(x)$. This is an essential obstruction to defining functions in three types which we cannot entirely overcome.

What is an obstruction to defining functions in general is not an obstruction to defining cardinality. If F is a definable bijection from A to B, then F^* with application defined as above will also be a bijection from A to B. The fact that $F^*[x]$ is defined as x on $B-A$ is harmless. Less obviously, the fact that $F^*[x]$ is defined as x on finite orbits in F is also harmless: the reason that this is not a problem is that a finite orbit in F clearly must lie in $A \cap B$, and changing this to the identity, while it does change what bijection from A to B we consider, does not change the fact that it is a bijection from A to B.

So we can define $A \sim B$ in a quite standard way: $(\exists F \mid (\forall x \in A.x \in \bigcup F \wedge F[x] \in B \wedge (\forall y \in A.F[x] = F[y] \rightarrow x = y) \wedge (\forall x \in B.(\exists! y \in A.F[y] = x))))$.

It is important to consider whether we have the theory of cardinality that we expect. Is $\sim$, thus defined, an equivalence relation? Can we prove the Schröder-Bernstein theorem? The answer to both of these questions is yes, though the proofs are slightly different from the usual ones.

Equinumerousness is an Equivalence Relation

We prove that $\sim$ is an equivalence relation in more usual contexts by observing that the identity function on A is a bijection witnessing $A \sim A$ (this works here), the fact that the inverse of a bijection from A to B is a bijection from B to A (this works here: if F is a (set) bijection from A to B, the relation $F[y] = x \wedge y \in A$ is bijective and (because F is coded by a set) has no finite cycles of length greater than 2, so it is represented by a set).

The proof of transitivity uses the fact that the composition of a bijection from B to C with a bijection from A to B is a bijection from A to C. This works, but not quite painlessly. Let F be a bijection from A to B and G be a bijection from B to C, both coded by sets. Let $\mathcal{H}(x, y)$ be defined as $y = G[F[x]] \wedge x \in A$. This relation is bijective, and so $\mathcal{H}^*$ is a bijection witnessing $A \sim C$; but it is not necessarily the composition of G and F (it may be corrected to eliminate finite cycles).

We can define $|A| \leq |B|$ as "there is a subset C of B such that $A \sim C$". An important result in the usual theory of cardinals is that $|A| \leq |B| \wedge |B| \leq |A| \rightarrow |A| = |B|$ (where $|A| = |B|$ is of course synonymous with $A \sim B$). This is the Schröder-Bernstein theorem.

The proof has the same flavor as the last clause of the previous proof: a bijective relation is defined in the manner of the usual proof, but the function obtained in the end is not necessarily the expected function.

Suppose that $f : A \to B' \subset B$ and $g : B \to A' \subseteq A$ are sets coding bijections. For any set A and function f whose domain includes A, define $f``A$ as $\{f[x] \mid x \in A\}$. Define $P$ as the intersection of all sets which contain every element of $A - g``B$ and contain $g[f[z]]$ whenever they contain z. Define $\mathcal{H}(x, y)$ as $(x \in P \wedge f[x] = y) \vee (x \in A - P \wedge g[y] = x)$. This is a bijective relation from A to B, and $\mathcal{H}^*$ will code a bijection from A to B (but not necessarily the expected one).

If we have the Axiom of Infinity (which we can express in various forms: $\texttt{Fin} \neq V$ works), we can show that the cardinals of finite sets satisfy the Peano axioms, and define addition and multiplication in sensible ways. We can show that for any finite sets A and B, there is a finite set B' disjoint from A and equinumerous with B, and $|A| + |B| = |A \cup B'|$ (this can be a definition or a theorem if addition is defined in the usual inductive fashion). There is a more complicated way to characterize $|A||B|$, supposing wlog that A and B are disjoint. A set C disjoint from A and B will have this cardinality if there is a set M each element of which is a triple consisting of one element of A, one element of B, and one element of C, such that any two element set with one element of A and one element of B is a subset of exactly one element of M and any element of C belongs to exactly one element of M.

Moreover, although natural numbers are type 2 objects we can nonetheless code any definable class of natural numbers as a set by considering the type 2 set of all type 1 sets which belong to some element of the class of natural numbers. This representation works because the natural numbers are disjoint from one another. $TST_3 +$ Infinity interprets second order Peano arithmetic; in fact it is equiconsistent with second-order arithmetic (and so is $NF_3 +$ Infinity, but this is beyond the scope of this talk).

Having completed the theory of cardinality, we ask whether the theory of functions, which is slightly defective, can be repaired? The answer is that it can be partially but not completely repaired without additional information on the structure of the universe.

Can other applications of the theory of functions be carried out? We will find that we can develop the complete theory of similarity of well-orderings (order types) (and more generally the theory of isomorphism types of linear orders).

Refining the Definition of Function

We show how to refine the definition of function so that it works essentially as often as possible. The difficulty is with finite cycles of length > 2. Suppose that $F(x, y)$ is a functional formula (so $F(x)$ is first-order definable) and $G(x)$ is a formula which is true of exactly one member of each finite cycle of length > 2 in F (and only of elements of such finite cycles). A new class function $F'(x, y)$ is definable as $(\neg G(x) \wedge F(x, y)) \vee (G(x) \wedge x = y)$. The function F' contains no finite cycles, and we define F'^{*} exactly as above. We redefine F^{*} as $F'^{*} \cup \{\{F(x), F(F(x))\} \mid G(x)\}$.

The new elements of F^* tell us where the function "reenters" a finite loop which has been cut in the transition from F to F'. The new two element sets are identifiable, because they are the only two element sets A in F^* which are subsets of an element B of F such that B contains as a subset a singleton element of F disjoint from A. Thus F' can be recovered from F. We define $F(x)$ as $F'[x]$ except when $\{x\} \in F$ and there is a unique y such that $\{y, F'[y]\} \in A$, $\{y, F'[y]\}$ is disjoint from $\{x\}$, and there are elements of A which contain all of $x, y, F'[y]$. In this case $F(x)$ is defined as y.

The difficulty with this definition is that in the absence of a certain amount of Choice, there might be functional relations which did not admit a definable selection of one element from each of their finite cycles. In such a case, this definition might not work.

If one has Choice for collections of disjoint finite sets, this definition will always work. Note that this implies that we can always code functions if Infinity does not hold.

If one has a linear order on the universe (a condition which we can express with our ability to code partial orders as sets) then one will always be able to define functions and moreover one has a much simpler way to do it: let A and B be two disjoint five element sets, let $\leq$ be the linear order, and define $\langle a, b \rangle$ as $\{a, b\} \Delta A$ if $a \leq b$ and $\{a, b\} \Delta B$ otherwise. It will then be possible to define functions as sets of ordered pairs in the usual way.

Functions can always be defined on restricted domains which happen to support linear orders. For example, the statement that two well-orderings (or indeed any two linear orderings) are isomorphic can be stated in the usual way, becausethe linear order on the domain of the first well-ordering can be used to select one element from each cycle in a bijective relation witnessing the isomorphism. That isomorphism is an equivalence relation is provable in exactly the usual way, because compositions of functions with linearly ordered domains can be defined without the finite cycle corrections.

We got sets coding cardinals of type 0 sets, because their elements are type 1, but we do not get objects coding ordinal numbers or linear order types, because the orderings themselves (always on type 0) are represented by type 2 sets and so cannot belong to further sets.

There can be no reliable notion of function in TST_3

It takes some care to articulate the negative result here.

Suppose that we add a primitive relation $\mathbf{f}(x, y)$ to the language of TST_3, x and y being type 0, about which we assume only that $\mathbf{f}$ is a functional relation (with a full ability to participate in comprehension axioms).

The existence of a definition of function amounts to existence of a formula $\mathbf{f}^*(z)$ (involving $\mathbf{f}$) and a formula F not involving $\mathbf{f}$ such that $(\exists z.\mathbf{f}^*(z)) \wedge (\forall xyz.\mathbf{f}^*(z) \wedge F(x, y, z) \leftrightarrow \mathbf{f}(x, y))$ is a theorem. The ability to prove this theorem for the completely anonymous function $\mathbf{f}$ will give us the ability to specify representatives for any definable functional relation in a uniform way.

We will now exhibit a model of TST_3 in which this situation is impossible.

Consider a model of TST_4 (just add one more type) in which there is a function f (represented by a set of Kuratowski pairs as usual) whose domain is an infinite union of cycles of length 3 and covers type 0.

Let G be the class of permutations of type 0 which are finite products of cyclic permutations of the form $(a, f(a), f^2(a))$ (f restricted to a cycle). We extend a permutation in G to types 1 and 2 by the rule $\pi(A) = \pi``A$. We refer to a set A of higher type as nice iff there is a finite set S of type 0 objects such that any permutation in G which fixes all elements of S fixes A. By standard considerations (the Frankel-Mostowski construction) the hereditarily nice sets make up a model of TST_4.

We will show using the model of TST_3 made up of types 0-2 in the permutation model of TST_4 that there cannot be a definition of function in the sense articulated in the previous slide.

For each type 1 set A with support S, for each a whose orbit does not meet S, the orbit of a is either contained in A or disjoint from A, because A must be fixed by $(a, f(a), f^2(a))$. Thus each set A in the permutation model has finite symmetric difference from a union of orbits. Moreover, any set in the original model of TST_4 which has finite symmetric difference from a union of orbits is also a set of the permutation model.

For each type 2 set A in the permutation model with support S and for each a whose orbit is disjoint from S, and for each Y in the permutation model disjoint from the orbit of a, either 0 or 3 of $\{a, f(a)\} \cup Y$, $\{f(a), f^2(a)\} \cup Y$, $\{f^2(a), a\} \cup Y$ belongs to A and either 0 or 3 of $\{a\} \cup Y$, $\{f(a)\} \cup Y$, $\{f^2(a)\} \cup Y$ belongs to A (again, by considering the action of $(a, f(a), f^2(a))$ on A).

Further, observe from the above that in fact for each a not in the support S of A (of type 1 or 2), A is fixed by $(a, f(a))$.

It then follows that for any first order formula $F(a,b,c)$ with just the given parameters, $F(a,f(a),z)$ will have the same truth value as $F(f(a),a,z)$ for all but finitely many a. [this follows because $\pi(x) = \pi(y) \leftrightarrow x = y$ and $\pi(x) \in \pi(y) \leftrightarrow x \in y$, whence uniform application of permutations preserves any first-order formula]. But this means that if we interpret $\mathbf{f}(x,y)$ as $y = f(x)$, there cannot be any $\mathbf{f}^*$ and F as above such that $(\exists z.\mathbf{f}^*(z)) \wedge (\forall xyz.\mathbf{f}^*(z) \wedge F(x,y,z) \leftrightarrow \mathbf{f}(x,y))$ is a theorem. The supposed z such that $\mathbf{f}^*(z)$ would have the property that $F(a,f(a),z)$ would be true and $F(f(a),a,z)$ would be false for all but finitely many a, and this is impossible.