

CONSISTENT SET THEORIES WITH UNIVERSAL SET

On occasion of the 1971 Berkeley symposium celebrating Alfred Tarski's achievements in logic and algebra Alonzo Church, who otherwise did not work much in set theory, presented a new system of set theory (Church 1974).

Church saw the two basic assumptions of post-naïve set theories in a restriction of comprehension to a form of separation (as in **ZFC**) and in a limitation of size (as in **NBG**). Similar to other criticism levelled against *Limitation of Size* Church regarded *Limitation of Size* as *ad hoc* (against the antinomies) and 'never well supported' as it proclaims a stopping point of further structures although classes are introduced (in **NBG** and **MK**) as collections, which can be quantified over. Church's set theory – let us call it “**CST**” here – follows **ZFC** in its idea of separation, but allows for collections that are 'large' in a way that even the larger transfinite sets of **ZFC** are not. **CST** does *not* introduce classes, but introduces a distinction *within* the area of sets. It allows even for $U = \{x \mid x=x\}$.

CST distinguishes between 'low sets', which have a 1:1-relation to a well-founded set, 'high sets', which are (absolute) complements of low sets and 'intermediate' sets which are neither. These labels pertain to the cardinality of sets. High sets are in 1:1-correspondence to the universal set U , low sets never. Because of the **CST** version of the *Axiom of Choice* for a set x which is not low *every* ordinal has a 1:1-relation to some subset of x .

So, the universe of **CST** may consist only of sets, but not all are well-founded. Obviously $U \in U$. U is the complement of $\emptyset$, so U is the paradigmatic high set. $\emptyset$ is well-founded, and, of course, $\emptyset \in U$. The constant

predicate “wf()” expresses the property of BEING WELL-FOUNDED, defined in the usual sense (using some order relation “<”):

$$(\forall x)(wf(x) \equiv x = \emptyset \vee$$

$$(\forall y)(y \subseteq x \supset (y \neq \emptyset \supset (\exists z)(z \in y \wedge (\forall w)(w \in y \supset z < w))))).$$

CST can be phrased as a second order system, quantifying over single- or two-argument open formula ϕ . One could understand this second order quantification as using classes, but only given a full-blown second order semantics. One could use schemata instead (i.e. having only free variables for open formula), as in **ZF**. We follow the first option here and make **CST** a second order theory.¹

The **CST** axioms are:

- Extensionality, Pair Set, Sum Set, Infinity as in **Z**
- Choice: $(\forall R)((\forall x)(\exists y)R(x,y) \supset (\exists f)(\forall x)R(x,f(y)))$
- Product Set: $(\forall y,z)(y \in z \supset (\exists u)(\forall x)(x \in u \equiv y \in z \supset x \in y))$
- Separation, Powerset, Replacement *restricted* to a condition “wf(x)“,
e.g. Axiom of Powerset: $(\forall x)(wf(x) \supset (\exists y)(\forall z)(z \in y \equiv z \subseteq x))$,

$$\text{Axiom of Separation: } (\forall x,F)(wf(x) \supset (\exists y)(\forall z)(z \in y \equiv z \in x \wedge F(z)))$$

[where “y” is not free in “F”]

The Axiom of Product Set allows having a substitute for separation in high sets. What is missing is the Axiom of Foundation of **ZF**.

These axioms of **CST** are strong enough to yield **ZF**. Dropping the non well-founded sets one gains a **ZF** universe. The two theories are equi-consistent

¹ Remember that in a full blown second order setting the Axiom of Choice is *not* equivalent to the Well-Order Principle. In fact the Well-Order Principle is *not* a theorem of **ZFC2** (cf. Shapiro 1991, pp.106-108). Neither the presence of the Axiom of Choice nor the presence of Foundation implies that all sets can be well-ordered in **ZFC2**.

(cf. Church 1974, §5). And without violating this relative consistency **CST** can be extended by axioms which go beyond **ZFC**. These are: *Strong Choice* (that U can be well-ordered), *Cardinality Axioms* (that there are cardinal numbers in the sense of Frege and Russell for all well-founded sets) and especially the *Axiom of Complements*:

$$(\forall x)(\exists y)(\forall z)(z \in y \equiv z \notin x).$$

The existence of absolute complements and the existence of U make **CST** a more natural set theory than **ZFC**, one may argue.

The argument in **ZFC** from the *Axiom of Separation* to the non-existence of U and *Cantor's Theorem* pose no problem for U and $\wp(U)$ for the same reason: the *Axiom of Separation* and the *Axiom of Powerset* are restricted to well-founded sets.

The argument to the non-existence of U [cf. Chap. I] now establishes that U *is not a well-founded set*. We knew that before. As U is not well-founded we *do not have* $\wp(U)$ in the first place.

One might now argue: So, in **CST** as well, there are some collections which are *there* – in this case *inter alia* the collection of all subsets of U , which are obviously existent if U is – but cannot be collected into a set; some collections which should exist, like $\wp(U)$, do not exist, because they are too large, just as in **NBG** set building operations cannot be applied to classes. Like **NBG** and **MK** tell us not much about classes, **CST** uses ‘large’ sets, but large sets cannot do much, since they are not subject to Separation or other set building principles.

This criticism, however, should be kept apart from a similar criticism levelled against **NBG**. In fact, all subsets of U are collected into *a set* – U itself. What cannot be done is separating a set $\wp(U)$ from U . The same applies to the ordinals: U is *a set* in which all ordinals are collected, but – on pains of re-introducing the Burali-Forti antinomy – we cannot separate a set which collects only the ordinals. So, does **CST** make some progress in

comparison to **NBG**? On the one hand **CST** can avoid using the second ontological category of classes, with all its problems. On the other hand **NBG** can collect just the ordinals into *their own* proper collection, albeit a class. **NBG** cannot introduce $\wp(V)$, as **CST** cannot introduce $\wp(U)$.

It is true that although the principal idea behind **CST** was rejecting Limitation of Size, **CST** exhibits some shadow of Limitation of Size: All high sets are by definition of the same size as U , just as in **NBG** all classes are of the same size as the class of sets.

*

The widest known set theory with a universal set U – for which we have $U \in U$ – is Quine's **NF** (from his paper "New Foundations of Mathematics", 1937). **NF** is Quine's set theory with a universal set that is not just virtual is **NFU**.

NF works by *Extensionality* and a *Comprehension Scheme* that is not as restricted as in **ZFC**.

Quine's **NF** Comprehension Scheme uses the idea of *stratified formula* (similar to the simple theory of types):

$$(\exists y)(\forall x)(x \in y \equiv \phi(x))$$

where " y " is not free in ϕ and ϕ is stratified. A formula is stratified if the set on the right-hand side of " $\in$ " is of a higher level than that on the left, and its definition does not include that on the left. A test for stratification consists in level assignments for sets or in trying a translation into the simple theory of types. The language of **NF** itself is not typed, thus avoiding duplication of structurally identical sets at different levels, but **NF** uses the stratification test to avoid the antinomies. **NF** is equiconsistent with the simple theory of types supplemented with the full ambiguity scheme, which asserts that a

formula φ is equivalent to formulas $\varphi^1, \varphi^2 \dots$ structurally similar to φ with type levels uniformly raised by 1, 2 ...

NF has the power to introduce *Pairing*, (Absolute) *Complement*, *Powerset*, and *Union* as instances of *Comprehension*.

NF allows $U = \{x \mid x = x\}$, since “ $x = x$ ” is a stratified formula. $\{x \mid x \in y\}$ can be stratified, defining the so-called ‘essence’ of an object (the collection of all its properties). “ $|x| = |\{\emptyset, \{\emptyset\}\}|$ ” is stratified as well, so natural numbers (in this case: 2) can be understood in Frege’s way. Stratification excludes, however, the usual definition of an infinite set (as in the Axiom of Infinity in **Z**). One meets in proofs and constructions in **ZFC** many unstratified set definitions. **NF** has to forsake these sets or has to introduce workarounds. Cardinal numbers and ordinal numbers come apart (cardinal numbers are not special ordinal numbers as in **ZFC**).

Non-stratified formula can be *used* in **NF** (this is different to Type Theory), but they cannot be used to define sets. Since non-stratified formulas can be used in **NF** one does not need a universal set or an empty set for every level (as in Russell's Type Theory) to have well-formed formula.

The antinomies – especially the (original) Russell Set – are avoided, since the corresponding open formulas in the *Comprehension Scheme* are not stratified. **NF** itself is not known to be consistent. **NF** with the *Axiom of Counting*, which says that a cardinal number is equal in cardinality to its singleton image, can prove the consistency of **Z**. No relative consistency proofs to **ZF** are available. Note that instances of *Replacement* are not stratified. Some subsystems of **NF** have been shown to be consistent (cf. Forster 1992, Holmes 1999).

$U \in U$ means that *Cantor's Theorem* does not hold (in general) in **NF**; but the set of unit sets of its elements is smaller in cardinality than U itself! The usual proof of *Cantor's Theorem* defines a set $y \in \wp(x)$ relative to a supposed

bijection f between x and $\wp(x)$ as $y = \{z | z \in x \wedge z \notin f(z)\}$, which is not stratified. In **NF** one can define, however, $y = \{z | z \in x \wedge z \notin f(\{z\})\}$, which is stratified with f being a supposed bijection between the set of singletons of $z \in x$ and $\wp(x)$. By the usual indirect argument one sees that there is no bijection between $\{z | z = \{w\} \wedge w \in x\}$ and $\wp(x)$, which means $|\wp(x)| > |\{z | z = \{w\} \wedge w \in x\}| = |\wp_1(x)|$. As $\wp(U) \subseteq U$ we have $|U| \geq |\wp(U)|$ and thus with the previous inequality we know that the cardinality of the set of singletons of elements of U is *smaller* than $|U|$: $|U| > |\wp_1(x)|$. That could mean that not all elements of U possess a singleton: although we have the set of all singletons in **NF**, as $\{x | (\exists y)(x = \{y\})\}$ is stratified, not all objects seem to have a singleton. This cannot be the solution. *Comprehension* provides a singleton for any object z : “ $x \in y \equiv x = z$ ” can be stratified. Contradiction is avoided finally by the *non-existence of the function* which maps any object to its singleton. This function does not exist *even though* every object has its singleton!

NF has other highly controversial features like the existence of infinite descending chains of cardinals, what conflicts with the *Axiom of Choice*, which thus doesn't hold in **NF**, which again implies – even in the absence of an *Axiom of Infinity* – that the universe of **NF** has to be infinite, since all finite sets can be well-ordered. The universe of **NF**, supposedly U , cannot be well-ordered then. **NF** might be consistent, though, with the claim that all well-founded sets can be well-ordered. Some functions (like the successor function) are *not* part of the universe – so where are they?

One can extend **NF** by introducing classes. In fact, if one *denies* in **NF** Rosser's *Axiom of Counting* [see above] one can prove the existence of non-set collections which are finite! Indeed a ‘strange landscape’ (Forster 1992, pp.29-32).

If one restricts the set building axioms to sets and uses unrestricted class comprehension one arrives at Quine's system **ML** (cf. Quine 1963, §§40-42). The problems the system **NF** has with unstratified induction and its incompatibility with the *Axiom of Choice* are resolved then. There is a class of all sets, $\cup U$, and *Cantor's Theorem* does not apply to it, as it is no set (i.e. has no powerset at all). Apart from resolving these problems classes play no constructive role in **ML**. Obviously we find ourselves in a system very similar to **MK**, and corresponding criticism applies here.

*

A version of **NF** that tries to avoid many of the peculiarities of **NF** is **NFU** (**NF** with urelements), developed by Randall Holmes (2005). **NFU** is built from **NF** by adding urelements and restricting extensionality to non-empty sets, introducing $\emptyset$ by an axiom. One may think of it as a subsystem of **NF** that allows only such models which contain the urelements.

NFU is consistent! It is consistent with the *Axiom of Choice*!

NFU is an extremely strong set theory. It can provide models for **ZFCU** by having *very* large cardinals (strongly inaccessible cardinals).

By working with a longer (*finite*) list of simple axioms:

Extensionality, Complements, Unions, Singletons, Cartesian Products, Converses, Domains, Projections, Singleton Image of a Set, Choice (!), Infinity...

Stratified Comprehension can be proven as a theorem!

The universal set is provided by its own axiom:

(U) $\{x \mid x = x\}$ exists.

U contains all sets as elements. All sets can be well-ordered, which is equivalent to the *Axiom of Choice*. So, U can be well-ordered, in contrast to

U in **NF**. By the existence of U , the *Axiom of Complements* and *Stratified Comprehension* absolute complements exist.

NFU also has its peculiarities: urelements, atoms and ordered pairs (!) are taken *as primitive*. One wonders what is supposed to be ‘in’ the ordered pair $\langle x, y \rangle$ if not x and y . And the **NFU Axiom of Ordered Pairs** contains the standard identity condition on ordered pairs *without* saying that an ordered pair is a set. The *Axiom of Pairing* of **Z** looks more natural.

More problematically: $[\in]$, $\{\langle x, y \rangle \mid x \in y\}$, does *not* exist, the existence of $[\in]$ leading to antinomies.² Whereas we thus have a set U , which may stand in as the extension of the predicate “set”, we have no extension for the membership relation, although $[\in]$ looks as natural as U . Interestingly $[\subseteq] = \{\langle x, y \rangle \mid x \subseteq y\}$ exists. So in many cases “ $\{x\} \subseteq y$ ” may substitute for “ $x \in y$ ”.

The non-existence of $[\in]$, which, of course, entails $[\in] \notin U$, entails that the membership relation is not modelled by U (U has no element corresponding to it), so that U cannot be a model for **NFU** itself. So as with the standard set theories **ZF**, **ZFC**, although working with a universal set, **NFU** has to look *outside of itself* for models. If the universal set is really *universal*, where should that outside be? We seem to be back to larger cardinals or similar collection-like entities or some hierarchy.

Not every supposed set exists, e.g. the set of all Cantorian ordinals. This is not better than in **ZF**. Some collections (like U) which do not exist in **ZFC** can exist in **NFU**, but as **NFU** does not distinguish sets from classes, some collections (like that of Cantorian ordinals) which exist in **NBG** do not exist in **NFU**. **NFU** can be viewed as trying to capture some middle ground between the other systems. This yields its own peculiarities.

² *Proof* (Outline). If $[\in]$ exists, so does its complement $\neg[\in]$ by the Axiom of Complements. $[=] = \{\langle x, y \rangle \mid x = y\}$ exists, since “ $x = y$ ” is stratified. Then the cut of $[=]$ with $\neg[\in]$ exists, and this cut is a cousin of the Russell Set: $\{\langle x, y \rangle \mid x = y \wedge x \notin y\}$.

Especially problematic is that some version of *Cantor's Theorem* is provable. The large sets (like the set of ordinals or U) in **NFU** have the strange property of *not* being equinumerous to their singleton images! How can that be? Although every object *has* a singleton (by the Axiom of Singletons), just as in **NF** the *function* giving the singleton to every object does not exist.³

$|\wp_1(U)| < |U|$ seems to contradict the *Axiom of Singletons*, a *provable* contradiction *only* being avoided by the non-existence of a general singleton function (cf. Quine 1963, p.293). We have – so to say, in analogy to *Separation* restricting *Comprehension* in **Z** – restricted singleton construction!

A corollary is *that the cardinality of the set of singletons of members of U , $|\wp_1(U)|$, is less than the cardinality of U* . This leads in **NFU** to the distinction between ‘Cantorian sets’ x with $|x|=|\wp_1(x)|$ and ‘non-Cantorian sets’, which resembles other limitation of size distinctions.

These cardinality issues lead to *Specker's Theorem* (cf. Holmes 2005, pp.132-34):

$$|\wp(U)| < |U|$$

which is read as proof that there are (many, many) urelements/atoms besides sets in U . Atoms, which have no members, are not elements of $\wp(U)$. If there are no atoms $|\wp(U)|=|U|$. In fact, *most* objects in U have to be atoms or pairs (i.e. ordered pairs not reducible to sets).

Unlike some version of **NF** in (Quine 1963) atoms are *not* identified with their singletons in **NFU**. As mentioned, ordered pairs are also objects in their own right besides sets. Ordered pairs can be taken as atoms in **NFU** as no

³ *Proof* (Outline). If the function $f:U \rightarrow \wp_1(U)$ existed, extensionality of the singleton would yield a function f^{-1} , so $|U| \leq |\wp_1(U)|$, which contradicts the Cantorian argument to $|\wp_1(U)| < |U|$ (given above with respect to **NF**). ■
Holmes (2005, pp. 109-110) provides another proof which relates to stratification.

claim with respect to composition – only with respect to their identity condition – was made. *Specker's Theorem* mirrors this conception.

Again:

- *Specker's Theorem* for **NFU** asserts that many entities in U are not in subsets of U . Moreover, *most* of the universe has to consist of urelements!

So: in all known models of **NFU** $|U| > |\wp(U)|$. All relations are subsets of $U \times U$, and all functions $f: U \rightarrow U$ should, if they are allowed to exist, themselves be sets of ordered pairs, thus be elements of U , thus be available as their own arguments, thus sometimes be forbidden to exist *as sets* at all in a consistent setting (e.g. a function of negative self-application). **NFU** itself can have models only in a realm which possesses properties quite different from what we expect of sets. **NFU**, so, deals not just with sets. In fact, the non-sets *vastly* outnumber the sets in any model of **NFU**. Ideally the non-sets contain just the urelements, but by the argument above concerning the *Axiom of Singletons* we should *expect* there to be a collection containing the ordered pair of any x and its singleton, but this collection cannot be a set in **NFU**. A crucial question is whether **NFU** can at least recapture ordinary sets – **ZFC**-like entities – as a sub-domain. Even though this is possible, however, we regain simply **ZFC** as a sub-universe – and are none the wiser with respect to our universality problem, as U , because of its behaviour in **NFU**, cannot be part of that recaptured realm. U is not ‘Cantorian’. Even the ordinals of **NFU** cannot be well-ordered in a *set model* of **NFU**! Holmes does not introduce proper classes into **NFU**, but admits their existence, supposedly objects of a broader theory (cf. Holmes 2005, p.50). **NFU** is not truly universal, as well.

NFU, thus, seems to be more than lost by the consequences of *Specker's Theorem* and the absence of $[\in]$.