Church's Set Theory with a Universal Set

Thomas Forster

Contents

Abstract

A detailed and fairly elementary introduction is given to the techniques used by Church to prove the consistency of his set theory with a universal set by constructing models of it from models of *ZF*. The construction is explained and some general facts about it proved.

Introduction

In 1974 Alonzo Church and Urs Oswald simultaneously and independently lit upon a refinement of Rieger-Bernays permutation models which enabled them to give elementary proofs of consistency for some set theories with a universal set. My own interest began much later, when I made the aquaintance of Flash Sheridan, a former student of Church's who was then writing an Oxford D.Phil.

thesis on Church's work in this area. I am greatly endebted to Flash for kindling my interest in this engaging bywater, and for showing me some relevant and (even now) unpublished material. This essay is substantially the same as the version that has appeared in the Church festschrift: I have corrected a few errors, reversed the American spellings introduced by the editors of that volume, and updated the notation and the bibliography.

Church was not generally known for having an interest in set theory, and one might wonder what moved him to write this late and isolated piece. An interest in set theory with a universal set was even less mainstream twenty years ago than it is now, but the consistency question for such theories is something that those whose interest in set theory was purely philosophical might well have mused about at any stage in its history, and I suspect that Church's motive was to make a small polemical point to the effect that there are consistent set theories with a universal set. It is a point worth making to students even now. Against this view is that fact that there were already two papers making the same point, and we know that one of them (Jensen [1969]) was known to Church. (The manuscript of Ward Henson's review of Jensen [1969] in *The Journal of Symbolic Logic* archives has comments on it in Church's hand.) However it is possible that Church felt that this paper did not make the point convincingly, since the system whose consistency is proved therein is not a pure set theory, but admits *urelemente*. The other, Grishin [1969], (which proves the consistency of NF_3, a pure set theory with a universal set and no *urelemente*) was apparently not known to Church. I am endebted to Herb Enderton for historical details about *The Journal of Symbolic Logic*.

1 Rieger-Bernays permutation models

We will start by swiftly recapitulating and updating a treatment of Rieger-Bernays models from Forster [1995]. If $\langle V, R\rangle$ is a structure for the language of set theory, and π is any (possibly external) permutation of V, then we say $x\ R_\pi\ y$ iff $x\ R\ \pi(y)$. $\langle V, R_\pi\rangle$ is a *permutation model* of $\langle V, R\rangle$. We call it V^π. Alternatively we could define Φ^π as the result of replacing every atomic wff $x \in y$ in Φ by $x \in \pi(y)$. We do not rewrite equations in this operation: $=$ is a logical constant not a predicate letter. The result of our definitions is that $\langle V, R\rangle \models \Phi^\pi$ iff $\langle V, R_\pi\rangle \models \Phi$.

A wff ϕ is stratified iff we can find a *stratification* for it, namely a map f from its variables (after relettering where appropriate) to $\mathbb{N}$ such that if the atomic wff '$x = y$' occurs in ϕ then $f($'x'$) = f($'y'$)$, and if '$x \in y$' occurs in ϕ then $f($'y'$) = f($'x'$) + 1$. Variables receiving the same integer in a stratification are said to be of the same *type*. Values of f are *types*.

To discuss these topics properly we will also need $j =_{\mathrm{df}} \lambda f \lambda x.(f``x)$. The map j is a group homomorphism: $j(\pi\sigma) = (j(\pi))(j(\sigma))$. A (possibly external) permutation σ of a set X is *setlike* if, for all n, $j^n(\sigma)$ is defined and is a permutation of $\mathcal{P}^n(X)$ (if this last thing exists). *ZF* proves that every definable permutation is setlike. This is actually a formulation of the axiom scheme of replacement! This is related to—but to be distinguished from—the set theorists' concept of an *amenable* class: one all of whose initial segments is a set. We shall start with a lemma and a definition, both due to Henson [1973]. The definition arises from the need to tidy up Φ^τ. A given occurrence of a variable 'x' which occurs in 'Φ^τ' may be prefixed by 'τ' or not, depending on whether or not that particular occurrence of 'x' is after an '$\in$'. This is messy. If there were a family of rewriting rules around that we could use to replace $x \in \tau(y)$ by $\sigma(x) \in \gamma(y)$ for various other σ and γ then we might be able to rewrite our atomic subformulæ to such an extent that for each variable, all its occurences have the same prefix.

Why bother? Because once a formula has been coerced into this form, every time we find a quantifier Qy in it, we know that all occurrences of y within its scope have the same prefix. As long as that prefix denotes a permutation then we can simply remove the prefixes! This is because $(Qx)(\ldots\sigma(x)\ldots)$ is the same as $(Qx)(\ldots x\ldots)$. If we can do this for all variables, then τ has disappeared completely from our calculations and we have an invariance result. When can we do this?

Henson's insight was as follows. Suppose we have a stratification for Φ and permutations τ_n (for all n used in the stratification) related somehow to τ, so that, for each n,

$$x \in \tau(y) \text{ iff } \tau_n(x) \in \tau_{n+1}(y).$$

then by replacing ' $x \in \tau(y)$ ' by ' $\tau_n(x) \in \tau_{n+1}(y)$ ' whenever 'x' has been assigned the subscript n, every occurrence of 'x' in ' Φ^τ ' will have the same prefix. Next we will want to know that τ_n is a permutation, so that in any wff in which 'x' occurs bound—$(\forall x)(\ldots\tau_n(x)\ldots)$—it can be relettered $(\forall x)(\ldots x\ldots)$ so that 'τ' has been eliminated from the bound variables. It is not hard to check that the definition we need to make this work is as follows

DEFINITION **1** $\tau_0 =$ identity, $\tau_{n+1} = (j^n(\tau)) \circ \tau_n$.

This definition is satisfactory as long as $j^n(\tau)$ is always a permutation of V whenever τ is, for each n. For this we need τ to be setlike. The trick of relettering variables that this facilitates is of crucial importance and will be used often.

This gives us immediately a proof of the following result.

LEMMA **2** ***Henson*** **[1973]**. *Let Φ be stratified with free variables 'x_1', ..., 'x_n', where 'x_i' has been assigned an integer k_i in some stratification. Let τ be a setlike permutation and V any model of NF. Then*

$$(\forall \vec{x})V \models (\Phi(\vec{x})^\tau \longleftrightarrow \Phi(\tau_{k_1}(x_1)\ldots\tau_{k_n}(x_n))).$$

In the case where Φ is closed and stratified, we infer that if τ is a setlike permutation, then

$$V \models \Phi \longleftrightarrow \Phi^\tau.$$

This was proved by Scott in [1962]. We can actually prove something with the flavour of a completeness theorem about these things: a formula is equivalent to a stratified formula iff the class of its models is closed under permutation models using setlike permutations. The proof of this is too long to be done here, but see Forster [1995], where it is theorem 3.0.4.

Remark 3 *If $\langle V, \in\rangle \models ZF$ and τ^{-1} is a setlike permutation of V then $\langle V, \in_\tau\rangle \models ZF$.*

Proof: The stratified axioms are no problem. The only unstratified axiom scheme is replacement. It is easy enough to check for any ϕ that if $\forall x \exists ! y \phi$ then $\forall x \exists ! y \phi^\tau$, so that for any set X the image of X in ϕ^τ is also a set. Call it Y. But then $\tau^{-1}(Y)$ is the image-of-X-under-ϕ (in the sense of V^τ). ■

We will also need the following observation, due to Coret.

Lemma 4 *If f is a stratification of Φ thought of as the partial function $\mathbb{N} \to \mathbb{N}$ that sends the variable's* subscript *(rather than the variable itself) to the type then*

$$\Phi(x_1, \ldots, x_k) \longleftrightarrow \Phi((j^{f(1)}(\sigma))(x_1), \ldots, (j^{f(k)}(\sigma))(x_k)).$$

Proof: It might be an idea to have an illustration before a full proof. For example, the lemma tells us that

$$x \in y \longleftrightarrow \sigma(x) \in \sigma``y.$$

The theorem is simply a more general assertion true for the same reasons.

Now for a full proof. By definition of j we have $x \in y$ iff $\tau(x) \in (j(\tau))(y)$ for any permutation τ. In particular if 'x' has been assigned type n and 'y' the type $n+1$, we invoke the case where τ is $j^n(\sigma)$ to get $x \in y \longleftrightarrow (j^n(\sigma))(x) \in (j^{n+1}(\sigma))(y)$. By substitutivity of the biconditional we can do this simultaneously for all atomic subformulae in $\Phi(x_1, \ldots, x_k)$. Variables 'y' that were bound in '$\Phi(x_1 \ldots x_k)$' now have prefixes like '$j^n(\sigma)$' in front of them but, since '$\Phi(x_1, \ldots, x_k)$' was stratified, they will be constant for each such variable 'y'. We then use the fact that $j^n(\sigma)$ is a permutation of V so that any formula $(Qy)(\ldots(j^n(\sigma))(y)\ldots)$ (Q a quantifier) is equivalent to $(Qy)(\ldots y \ldots)$. ■

Definition 5 A $\mathcal{P}$-embedding from $\mathcal{A}$ into $\mathcal{B}$ is an injection $i : \mathcal{A} \to \mathcal{B}$ for which the power set operation is absolute. "No new members or subsets of old sets." If i is the identity we say $\mathcal{B}$ is a **$\mathcal{P}$-extension** of $\mathcal{A}$.

THEOREM **6** *Let $\mathcal{M} = \langle M, \in \rangle$ be a wellfounded model of ZF, and let σ be a setlike permutation of M. Let $i : \mathcal{M} \hookrightarrow \mathcal{M}^\sigma$ be recursively defined by $i(x) =: \sigma^{-1}(i``x)$. Then (i) i is a $\mathcal{P}$-embedding and (ii) i is elementary for stratified formulæ.*

Proof:

(i) If $x \in_\sigma i(y)$ then x is a value of i so i is an end-extension. Suppose $(x \subseteq i(y))^\sigma$: we want x to be a value of i. $(x \subseteq i(y))^\sigma$ is just $\sigma(x) \subseteq \sigma(i(y)) = i``y$ so $\sigma(x)$ is a set of values of i so x is a value of i.

(ii) Let $\phi(\vec{x})$ be a stratified formula whose free variables are precisely the $\vec{x}$, a tuple of length k. Assume

$$\mathcal{M}^\sigma \models \phi(i(x_1), i(x_2), \ldots i(x_k)).$$

We now need the following fact, due to Coret.

Let $M \prec_{\mathcal{P}} K$ be structures for set theory and suppose that for all $x \in K$ there is $y \in M$ such that there is a setlike permutation π of K with $\pi``y = x$. Then the inclusion embedding is elementary for stratified formulæ.

Let's concentrate on the hard case of the existential quantifier. We want to show that if $(\exists y)\phi(\vec{x}, y)$ where ϕ is stratified and the $\vec{x}$ are in M, then there is $y \in M$ witnessing the quantifier. We will use the lemma which in Forster [1992] I called "Boffa's lemma on n-formulae" but which I suspect is really due to Coret. To keep things readable, let us suppose there are only two x variables, that y is of type 5, and that x_1 is of type 2 and x_2 of type 4. To invoke Boffa's lemma we must find a permutation π such that $(j(\pi))(x_1) = x_1$, $(j^3(\pi))(x_2) = x_2$ and $(j^4(\pi))(y)$ is something wellfounded. We must think of the action of π on the things in $\bigcup^2 x_1$ and $\bigcup^4 x_2$ and $\bigcup^5 y$. π must fix everything in $\bigcup^2 x_1$ and $\bigcup^4 x_2$ and must send everything in $\bigcup^5 y$ to something wellfounded. It will be sufficient for $\bigcup^5 y$ to be the same size as something in M.

Actually this is easy, and appeals to a sort of inside-out *pigeonhole principle*, which says something like: when there is enough room, you can do whatever you like. We have to be confident that once we have specified π at least to the extent of saying it must fix everthing in $\bigcup^2 x_1$ and $\bigcup^4 x_2$ there are nevertheless still enough M-sets that can be moved for us to be able to send all the non-M members of $\bigcup^5 y$ to them. (Notice that in ZF no illfounded set can be symmetrical, since (Boffa again) if x is symmetrical and illfounded, then $\bigcup^n x = V$ for some n, and of course this cannot happen in ZF.)

The instance of Coret's lemma that is of interest to us is the case where K is V^σ and M is V. Notice that in V^σ every set is (externally) the same size as a wellfounded set, since the members (in V^σ) of x are the members (in V) of $\sigma^{-1}(x)$, and $\sigma^{-1}(x)$ is certainly in V. But in *ZF* any bijection between sets can be extended to a setlike permutation of the universe. (This relies on replacement and is not true of weaker theories!)[1]

[1]The proof of this result in the version of this paper appearing in the Church festschrift is flawed: thanks to Randall Holmes for spotting the error.

COROLLARY **7** *The axiom of foundation is independent of ZF.*

Proof: Use the transposition $(\emptyset, \{\emptyset\})$. In the new model the old empty set has become a **Quine atom**: an object equal to its own singleton. ■

The completeness theorem for setlike permutations and stratified formulæ is a powerful and satisfying piece of 1950's-style model theory, not unlike Birkhoff's theorem in flavour, but it is actually a nuisance. It is all very well if one has a model of a stratified theory, and want to generate further models of it, but it does mean that if one starts with a model of a theory that does not have a universal set, then the result will not have a universal set either. (After all, the assertion that there is no universal set is stratified, and if it starts off false will remain false). It gives us the independence of the axiom of foundation, but not the relative consistency of a universal set. If we want a universal set we have to enrich the construction slightly.

[*HOLE Think in terms of di Giorgi models. It's obvious that if there is a bijection between X and some ideal in $\mathcal{P}(X)$ then there will be a bijections between X and the union of that ideal and its dual filter—at least as long as $|X| = 2 \cdot |X|$. Less than blindingly obvious that this bijection can be coded in the di Giorgi model arising from the bijection with the ideal. However, we can make this explicit.*]

2 Church-Oswald models

Let $\langle V, \in\rangle$ be a model of *ZF* and let k (for '*k*ode') be a bijection between V and $V \times \{0,1\}$. Next we define

$$x \in_{\mathbf{co}} y \ \text{ iff } \left\{ \begin{array}{l} 1.\ k(y) = \langle y', 0\rangle \text{ and } x \in y', \textit{ or} \\ 2.\ k(y) = \langle y', 1\rangle \text{ and } x \notin y'. \end{array} \right.$$

This is the simplest version of the construction, and is the one in Oswald [1976]—though Oswald considered specifically the case where V was V_ω and presented it very differently. (Oswald's model is presented more fully in section 2.1.1.)

The first thing to notice is that every set has a complement in the sense of $\in_{\mathbf{co}}$. (Perhaps the first thing of all to notice is that $\in_{\mathbf{co}}$ really is extensional!). In fact the resulting model is a model of a theory known as NF_2, which Oswald was studying at the time. The axioms of NF_2 are (i) Extensionality (ii) complementation ($-x$ is a set, always) (iii) $x \cap y$ and (iv) existence of $\{x\}$. This axiomatisation is finite but it is perhaps more convenient to replace (iv) by the scheme giving the existence, for each $n \in \mathbb{N}$ and for all n-tuples $\vec{x}$, of the set $\{x_1, x_2, \ldots x_n\}$.

We took k to be a bijection between V and $V \times \{0,1\}$, but it could have been $V\times$ anything K we like ('K' for '*K*ode'), as long as $0 \in K$ and we have suitable clauses for $\in_{\mathbf{co}}$ x where the second component of $k(x)$ is not 0. The idea is that in general we define $x \in_{\mathbf{co}} y$ if either (i) $\mathtt{snd}(k(y)) = 0$ and $x \in \mathtt{fst}(y)$ or[2] (ii) various other clauses concerning (new) membership in sets y such that $\mathtt{snd}(k(y)) \neq 0$. The idea is that other entries in K will correspond to other operations on sets (in the case we have just seen 1 corresponds to complementation). In general there is a problem of extensionality. There is of course no difficulty in showing that x and x' satisfy extensionality as long as $\mathtt{snd}(k(x)) = \mathtt{snd}(k(x')) = 0$, but there are other cases to consider. The task of verifying extensionality in the new structure will be much easier if all x and y such that $\mathtt{snd}(k(x)) \neq \mathtt{snd}(k(y))$ are so different that the possibility of them having the same members in the new sense simply never arises. For example, in the case we have just seen, if $\mathtt{snd}(k(x))$ is 0 then x has only a set (in the old sense) of members-in-the-new-sense, whereas if $\mathtt{snd}(k(x))$ is 1 it has a proper class (in the old sense) of members-in-the-new-sense. This tells us that if the theory for which we are trying to obtain a model by this construction is T, then the extensionality problem for the model is deeply related to the word problem for T. In particular if we have a good notion of normal forms for T words over a set of generators then we will be able to take K to be (roughly) a set of such normal words. *What this reveals is that this technique is not a great deal of use for constructing models of a theory T unless T has an easy word problem.* Set theories with an easily solvable word problem are unlikely to be of interest.

For the moment we will—in the name of generality—ignore the question of what the remaining components of $k(x)$ can be. We will assume that K is some arbitrary collection such that $0 \in K$ and there are rules to ensure that $\in_{\mathbf{co}}$ is extensional and that when $\mathtt{snd}(k(y)) = 0$ then $x \in_{\mathbf{co}} y \longleftrightarrow x \in \mathtt{fst}(k(y))$. We will try to prove some general results about this construction. Let us call structures built in this way **CO structures**. ("Church-Oswald").

To summarise, there are three parts to a **CO** construction over a model $\langle V, \in \rangle$. There is a collection K of objects available to be used as second components of ordered pairs; there is a bijection k between V and $V \times K$, and finally there is a family of rules telling us how possible membership (in the new sense) of x in y depends on the second component of $k(y)$. Rieger-Bernays permutation model constructions can be seen as those special cases of **CO**-constructions where K is a singleton. Accordingly the only thing one is free to play with is the bijection k.

Since we are assuming that one of the clauses in the definition of $\in_{\mathbf{co}}$ is always $\mathrm{snd}(k(x)) = 0 \rightarrow (\forall y)(y \in_{\mathbf{co}} x \longleftrightarrow y \in \mathtt{fst}(k(x)))$ we can make the following definition.

Definition 8 A **low set** is a set x such that $\mathtt{snd}(k(x)) = 0$.

This is at variance with other definitions in the literature, but it is the most useful. This is not to be confused with the notion of a **hereditarily low** set,

[2] $\mathtt{fst}(x)$ and $\mathtt{snd}(x)$ are the first and second components of the ordered pair x.

for this will turn out to be important as well.

DEFINITION 9 H_{low} is the greatest fixed point for the function that sends an argument x to the set of low subsets of x. That is to say, H_{low} is the collection of sets x such that everything in the transitive closure of x is low.

(Notice that the least fixed point must consist entirely of wellfounded sets.)

DEFINITION 10 The Axiom scheme of **Low Comprehension** states that, for any formula $\phi(x, \vec{y})$, and for all $\vec{y}$, if the collection of all x such that $\langle V, \in_{\mathbf{co}}\rangle \models \phi(x, \vec{y})$ is a set of the original model, then $\langle V, \in_{\mathbf{co}}\rangle \models$ "$\{x : \phi(x, \vec{y})\}$ is a (low) set".

Thus it is by no means obvious that low comprehension is axiomatisable. For this reason the most profitable approach to this topic is to think of the **CO** constructions as things that give us models, rather than to attempt to be specific in saying what the axioms are of the theory whose consistency we have proved.

The following triviality is central to what is to come.

THEOREM 11 *All* **CO** *structures satisfy low comprehension.*

Proof: Let $\phi(x, \vec{y})$ satisfy the antecedent, and consider the class of all x such that $\langle V, \in_{\mathbf{co}}\rangle \models \phi(x, \vec{y})$ which is a set of the original model, X, say. Then $\{x : \phi(x, \vec{y})\}$ in the sense of the new model is simply $k^{(-1)}(\langle X, 0\rangle)$, and is of course low. ■

In particular we have an axiom of **pairing**.

In a typical **CO** construction there will be plenty of new sets containing all low sets: V for one. However

PROPOSITION 12 *No new set containing all low sets can be low.*

Proof: Suppose there were a low set containing all low sets. Then, by low comprehension, the collection of all low sets is low. That is to say, there is an x such that $(\forall y)(y \in_{\mathbf{co}} x \longleftrightarrow \mathtt{snd}(k(y)) = 0)$. But there is certainly a proper class of y such that $\mathtt{snd}(k(y)) = 0$, so this x has a proper class of members and is therefore not low. ■

Attempts to reconstruct the usual paradoxes in this new context give rise to demonstrations that certain sets are not low. Take Russell's paradox for example. The new model cannot contain the set of all things that are not members of themselves. If the collection of x such that $\neg(x \in_{\mathbf{co}} x)$ were a set of the original model then the Russell class would be a set of the new model. So there is a proper class of x such that $\neg(x \in_{\mathbf{co}} x)$.

COROLLARY 13 *Every surjective image of a low set is low and every subset of a low set is low.*

Proof: The first part follows from replacement in the original model and the second from comprehension. ■

The first follows from the second by AC but a proof without AC is preferable since we will not otherwise be making any use of AC in this development.

COROLLARY 14 *Every low set has a power set, which is also low.*

Proof: When x is low, $\mathcal{P}(x$ (in the sense of $\in_{\mathtt{co}}$) must be

$$k^{(-1)}(\langle k^{(-1)}``(\mathcal{P}((\mathtt{fst}(k(x))) \times \{0\}), 0\rangle).$$ ■

COROLLARY 15 *If x is a low set of low sets, then it has a sumset, which is low.*

Proof: The object we need to play the role of $\bigcup x$ is

$$k^{-1}\langle \bigcup (\mathtt{fst} \circ k)``(\mathtt{fst} \circ k)(x), 0\rangle.$$

This depends only on the availability of the axiom of sumset in the model we start with, and doesn't tell us about stronger, less restricted forms of the axiom. ■

This seems to be an argument for setting up this theory with low comprehension in the way I have done it rather than with low replacement in the way Church originally did. Low replacement in the form "The image of a wellfounded set in a function is a set" certainly implies existence of power set for low sets and sumset, but we seem to need AC to deduce sumset for low sets of low sets. This is very messy.

2.1 Applications of the technique

2.1.1 Models of NF_2

Oswald did not set up his first illustration with the apparatus of K and k as here. Instead he defined a binary relation on $\mathbb{N}$ as follows:

$n\ E\ m$ iff either

1. m is even and the nth bit of the binary expansion of $m/2$ is 1;
   or
2. m is odd and the nth bit of the binary expansion of $(m-1)/2$ is 0.

This obviously derives from the old trick (due to Ackermann) of defining $n\ E\ m$ ($n, m \in \mathbb{N}$) iff the nth bit of the binary expansion of m is 1.

It might be an idea to concentrate briefly on the three salient features of this the simplest possible case.

$K = \{0, 1\}$, and the other clause gives us complements;
The $\in$ relation of the original model is wellfounded;
The rank of $\mathtt{fst}(k(x))$ is no greater than the rank of x.

When these conditions are met we can say a lot.

Definition 16 An antimorphism is a permutation π of the universe satisfying $(\forall xy)(x \in y \longleftrightarrow \pi(x) \notin \pi(y))$.

Remark 17 *Under the three assumptions above the new model admits an antimorphism.*

Proof: Declare the following recursive definition

$$\sigma(x) =: k^{-1}(\langle \sigma``(\mathtt{fst}(k(x))), (1 - \mathtt{snd}(k(x)))\rangle).$$

By considering putative counterexamples of minimal rank we can show that this is everywhere defined.

Suppose $\sigma(y) \in_{\mathbf{co}} \sigma(x)$. This is

$$\sigma(y) \in_{\mathbf{co}} k^{-1}(\langle \sigma``(\mathtt{fst}(k(x))), (1 - \mathtt{snd}(k(x)))\rangle).$$

Now either $\mathtt{snd}(k(x)) = 0$ or $\mathtt{snd}(k(x)) = 1$.

0 If $\mathtt{snd}(k(x)) = 0$ the displayed formula becomes

$$\sigma(y) \in_{\mathbf{co}} k^{-1}(\langle \sigma``(\mathtt{fst}(k(x))), 1\rangle),$$

which is

$$\sigma(y) \notin \sigma``(\mathtt{fst}(k(x))),$$

which is $y \notin (\mathtt{fst}(k(x)))$. But if $\mathtt{snd}(k(x)) = 0$ this becomes $y \notin_{\mathbf{co}} x$.

1 On the other hand if $\mathtt{snd}(k(x)) = 1$ then

$$\sigma(y) \in_{\mathbf{co}} k^{-1}(\langle \sigma``(\mathtt{fst}(k(x))), (1 - \mathtt{snd}(k(x)))\rangle)$$

is

$$\sigma(y) \in_{\mathbf{co}} k^{-1}(\langle \sigma``(\mathtt{fst}(k(x))), 0\rangle)$$

This is $\sigma(y) \in \sigma``(\mathtt{fst}(k(x)))$ which of course simplifies to $y \in \mathtt{fst}(k(x))$ and (since $\mathtt{snd}(k(x)) = 1$) this becomes $y \notin_{\mathbf{co}} x$.

So either way we have $\sigma(y) \in_{\mathbf{co}} \sigma(x) \longleftrightarrow y \notin_{\mathbf{co}} x$. ■

There is another nice result we get as a reward for making these assumptions. Consider the game G_x, played as follows, by two players, `I` and `II`. `I` picks $x_1 \in x$, `II` picks $x_2 \in x_1$, `I` picks $x_3 \in x_2$ ..., with the first player unable to move (you can't pick a member of the empty set!) losing. Clearly the axiom of foundation implies that G_x is always determinate in the sense of admitting a winning strategy for one player or the other. This is **$\in$-determinacy**. The axiom of foundation is obviously implicated because if $x = \{x\}$ G_x is clearly not determinate. The converse is not true, however, and we can prove the following.

Remark 18 *Under the three assumptions above, the new model obeys $\in$-determinacy.*

Proof: G_x is a win for player `I` if there is a $y \in x$ s.t. G_y is a win for player `II`. Similarly G_x is a win for player `II` iff for every $y \in x$, G_y is a win for player `I`.

First we prove that there are x of arbitrarily high rank such that x is hereditarily low and G_x is a win for player `II`. (And player `I`, but we don't need that.) This is because under the assumptions given H_{low} is an isomorphic copy of the original model and there are clearly x of arbitrarily high rank with those properties.

Next we prove by induction on rank of x that (in the sense of the new model) G_x is determinate (admits a winning strategy for `I` or `II`). Suppose this is not true, and let x be a counterexample of minimal rank. Then $k(x) = \langle y, 1\rangle$ or $k(x) = \langle y, 0\rangle$, for some y. If $k(x) = \langle y, 0\rangle$ for some y then by induction hypothesis G_z is determinate for every $z \in_{\mathbf{co}} x$ and G_x must be determinate as well. If $k(x) = \langle y, 1\rangle$ then there is some ordinal α (namely rank of $y + 1$) such that, for all z, if z is an element of the original model of rank at least α, then $z \notin y$, so (since $k(x) = \langle y, 1\rangle$) $z \in_{\mathbf{co}} x$. Now we have just proved that at least some of these z give rise to G_z that are wins for player `II`. So there is at least one $z \in_{\mathbf{co}} x$ such that `II` has a winning strategy for G_z. But then `I` has a winning strategy for G_x which he launches by picking z. ■

2.1.2 An elementary example

We are going to consider two models, more complicated than the original Oswald model, in which in addition to complements for all x, we also have $B(x)$ for all x. $B(x)$ is $\{y : x \in y\}$.[3] The theory that has boolean axioms ($\cup$, $\cap$ and $-$), singleton and B is *NFO*. (The 'O' is intended to suggest **O**pen. This theory is more naturally axiomatised as Extensionality plus existence of $\{x : \phi(x, \vec{y})\}$ where ϕ is stratified and quantifier-free. It is a relatively straightforward exercise to check that these axiomatisations are equivalent.) The second model will satisfy all the boolean axioms (and so is a model of *NFO*), but we start with the first, which doesn't.

Consider SS = the set of reduced words in the semigroup-with-unity with two generators `c` and `b`, and the equation $\mathtt{c}^2 = 1_{SS}$. '`c`' is intended to recall "**C**omplement" and '`b`' to recall "$B(x)$". We now let k be a bijection between V and $V \times SS$. We define $\in_{\mathbf{co}}$ by recursion by cases:

Definition 19

1. If $\mathtt{snd}(k(x)) = \mathtt{c}w$ for some word $w \in SS$ then

$$y \in_{\mathbf{co}} x \text{ iff } y \notin_{\mathbf{co}} k^{(-1)}\langle \mathtt{fst}(k(x)), w\rangle.$$

[3] I use this notation because it was Boffa who first impressed on me the importance of this operation. I learned later that the first use of this operation was by Quine, and Whitehead suggested to him that $\{y : x \in y\}$ should be called the *essence* of x.

2. If $\mathtt{snd}(k(x)) = \mathtt{b}w$ for some word $w \in SS$ then

$$y \in_{\mathbf{co}} x \text{ iff } k^{(-1)}(\langle \mathtt{fst}(k(x)), w\rangle) \in_{\mathbf{co}} y.$$

3. If $\mathtt{snd}(k(x))$ is 1_{SS} then $y \in_{\mathbf{co}} x$ iff $y \in \mathtt{fst}(k(x))$.

Proposition 20 *The $\in_{\mathbf{co}}$ of definition 19 is extensional.*

Proof: The proof is an inductive proof (by cases) on the construction of words in SS. The situation we are contemplating is two distinct x_1 and x_2 which have the same $\in_{\mathbf{co}}$-members. Naturally we will be interested in the second components of $k(x_1)$ and $k(x_2)$. There are several cases to consider:

1. The second components are both 1_{SS}, the unit of the semigroup. In this case the first components must be the same, and $x_1 = x_2$ follows.

2. One of the second components is 1_{SS} and the other isn't. Suppose $\mathtt{snd}(k(x_1))$ $= 1_{SS}$ and $\mathtt{snd}(k(x_2))$ is $\mathtt{bw}$ or $\mathtt{cbw}$. (We cannot have two adjacent $\mathtt{c}$'s since the words are reduced and $\mathtt{c}^2 = 1_{SS}$.) Then the collection of y such that $y \in_{\mathbf{co}} x_1$ is $\mathtt{fst}(k(x_1))$, which is a set in the sense of $\langle V, \in\rangle$. In contrast the collection of y such that $y \in_{\mathbf{co}} x_2$ is either (i) the collection of y such that $y \notin_{\mathbf{co}} k^{(-1)}\langle \mathtt{fst}(k(x_2)),\mathtt{b(w)}\rangle$—which is to say, by an application of part 2 of definition 19—the same as the collection of y such that $k^{(-1)}\langle \mathtt{fst}(k(x_2)),\mathtt{w}\rangle \notin_{\mathbf{co}} y$, or (ii) is the collection of y such that $k^{(-1)}(\langle \mathtt{fst}(k(x_2)), w\rangle) \in_{\mathbf{co}} y$. The collection in case (i) cannot be a set because for any object a we can easily find proper-class-many unordered pairs which do not have a as a member. The collection in case (ii) cannot be a set because for any object a we can easily find proper-class-many unordered pairs which *do* have a as a member.

3. $\mathtt{snd}(k(x_1)) = \mathtt{c}(\mathtt{w}_1)$ and $\mathtt{snd}(k(x_2)) = \mathtt{c}(\mathtt{w}_2)$ where the first letter of both $\mathtt{w}_1$ and $\mathtt{w}_2$ is $\mathtt{b}$. Assume $(\forall y)(y \in_{\mathbf{co}} x_1 \longleftrightarrow y \in_{\mathbf{co}} x_2)$ with a view to deducing $x_1 = x_2$. We can expand this in accordance with part 2 of definition 19:

$$(\forall y)(y \notin_{\mathbf{co}} k^{(-1)}\langle \mathtt{fst}(k(x_1)),\mathtt{w}_1\rangle \longleftrightarrow y \notin_{\mathbf{co}} k^{(-1)}\langle \mathtt{fst}(k(x_2)),\mathtt{w}_2\rangle),$$

which is to say

$$(\forall y)(y \in_{\mathbf{co}} k^{(-1)}\langle \mathtt{fst}(k(x_1)),\mathtt{w}_1\rangle \longleftrightarrow y \in_{\mathbf{co}} k^{(-1)}\langle \mathtt{fst}(k(x_2)),\mathtt{w}_2\rangle).$$

That is to say, if we have distinct x_1 and x_2 with the same members-in-the-sense-of-$\in_{\mathbf{co}}$, where the first letter of $\mathtt{snd}(k(x_1))$ is the same as the first letter of $\mathtt{snd}(k(x_2))$, namely $\mathtt{c}$, then there are distinct y_1 and y_2 (namely $k^{(-1)}\langle \mathtt{fst}(k(x_1)),\mathtt{w}_1\rangle$ and $k^{(-1)}\langle \mathtt{fst}(k(x_2)),\mathtt{w}_2\rangle$) satisfying $(\forall z)(z \in_{\mathbf{co}} y_1 \longleftrightarrow z \in_{\mathbf{co}} y_2)$ and $\mathtt{snd}(k(y_1))$ and $\mathtt{snd}(k(y_2))$ are shorter than $\mathtt{snd}(k(x_1))$ and $\mathtt{snd}(k(x_2))$—indeed they are terminal segments.

This clearly reduces to the case where the two words in question are $\mathtt{w}_1$ and $\mathtt{w}_2$ both beginning with $\mathtt{b}$, which we now treat.

4. The second components of $k(x_1)$ and $k(x_2)$ are words $\mathtt{b}(\mathtt{w}_1)$ and $\mathtt{b}(\mathtt{w}_2)$. We have $(\forall y)(y \in_{\mathbf{co}} x_1 \longleftrightarrow y \in_{\mathbf{co}} x_2)$. Expanding this by clause (2) in definition 19 we obtain

$$(\forall y)(k^{(-1)}(\langle \mathtt{fst}(k(x_1)),\mathtt{b}(\mathtt{w}_1)\rangle) \in_{\mathbf{co}} y \longleftrightarrow k^{(-1)}(\langle \mathtt{fst}(k(x_2)),\mathtt{b}(\mathtt{w}_2)\rangle) \in_{\mathbf{co}} y).$$

   Now as long as

$$k^{(-1)}(\langle \mathtt{fst}(k(x_1)),\mathtt{b}(\mathtt{w}_1)\rangle) \neq k^{(-1)}(\langle \mathtt{fst}(k(x_2)),\mathtt{b}(\mathtt{w}_2)\rangle)$$

   we can falsify this biconditional by taking y to be the singleton of one of these two. Singletons exist by low comprehension: the singleton of x (in the sense of $\in_{\mathbf{co}}$) is just $k^{(-1)}(\langle \{x\}, 0\rangle)$.

5. $k(x_1)$ has second component $\mathtt{cbw}_1$ and $k(x_2)$ has second component $\mathtt{b}(\mathtt{w}_2)$. This is really the case where one word begins with a $\mathtt{c}$ and the other begins with a $\mathtt{b}$ but since the words are all reduced we also know that the letter following the $\mathtt{c}$ in the word beginning with a $\mathtt{c}$ must be a $\mathtt{b}$. If we have

$$(\forall y)(y \in_{\mathbf{co}} x_1 \longleftrightarrow y \in_{\mathbf{co}} x_2),$$

   this becomes

$$(\forall y)(y \notin_{\mathbf{co}} k^{(-1)}(\langle \mathtt{fst}(k(x_1)),\mathtt{b}(\mathtt{w}_1)\rangle) \longleftrightarrow k^{(-1)}(\langle \mathtt{fst}(k(x_2)),\mathtt{w}_2\rangle) \in_{\mathbf{co}} y).$$

   This is the same as

$$(\forall y)(y \in_{\mathbf{co}} k^{(-1)}\langle \mathtt{fst}(k(x_1)),\mathtt{b}(\mathtt{w}_1)\rangle \longleftrightarrow k^{(-1)}(\langle \mathtt{fst}(k(x_2)),\mathtt{w}_2\rangle) \notin_{\mathbf{co}} y).$$

   Now $y \in_{\mathbf{co}} k^{(-1)}\langle \mathtt{fst}(k(x_1)),\mathtt{b}(\mathtt{w}_1)\rangle$ iff (by clause (2) in definition 19)

$$k^{(-1)}\langle \mathtt{fst}(k(x_1)),\mathtt{w}_1\rangle \in_{\mathbf{co}} y$$

   and substituting this for '$y \in_{\mathbf{co}} k^{(-1)}\langle \mathtt{fst}(k(x_1)),\mathtt{b}(\mathtt{w}_1)\rangle$' in

$$(\forall y)(y \in_{\mathbf{co}} k^{(-1)}\langle \mathtt{fst}(k(x_1)),\mathtt{b}(\mathtt{w}_1)\rangle \longleftrightarrow k^{(-1)}(\langle \mathtt{fst}(k(x_2)),\mathtt{w}_2\rangle) \notin_{\mathbf{co}} y)$$

   we obtain

$$(\forall y)(k^{(-1)}(\langle \mathtt{fst}(k(x_1)),\mathtt{w}_1\rangle) \in_{\mathbf{co}} y \longleftrightarrow k^{(-1)}(\langle \mathtt{fst}(k(x_2)),\mathtt{w}_2\rangle) \notin_{\mathbf{co}} y).$$

   Now anything like $(\forall y)(a \in y \longleftrightarrow b \notin y)$ must always be false because of the existence of the empty set. ∎

The rest is easy:

PROPOSITION 21 *With $\in_{\mathbf{co}}$ as in definition 19, $\langle V, \in_{\mathbf{co}}\rangle \models (\forall x)(-x$ exists) and $(\forall x)(B(x)$ exists).*

Proof: The complement of x will be $k^{(-1)}\langle \mathtt{fst}(k(x)), \mathtt{csnd}(k(x))\rangle$, for, by clause (1) in definition 19, we have

$$\begin{array}{l} y \in_{\mathbf{co}} k^{(-1)}\langle \mathtt{fst}(k(x)), \mathtt{csnd}(k(x))\rangle \text{ iff} \\ y \notin_{\mathbf{co}} k^{(-1)}\langle \mathtt{fst}(k(x)), \mathtt{snd}(k(x))\rangle \text{ iff} \\ y \notin_{\mathbf{co}} x. \end{array}$$

$B(x)$ will be $k^{(-1)}\langle \mathtt{fst}(k(x)), \mathtt{bsnd}(k(x))\rangle$, for, by clause (2) in definition 19,

$$\begin{array}{l} y \in_{\mathbf{co}} k^{(-1)}\langle \mathtt{fst}(k(x)), \mathtt{bsnd}(k(x))\rangle \text{ iff} \\ k^{(-1)}\langle \mathtt{fst}(k(x)), \mathtt{snd}(k(x))\rangle \in_{\mathbf{co}} y \text{ iff} \\ x \in_{\mathbf{co}} y. \end{array}$$

■

2.1.3 $\mathcal{P}$-extending models of Zermelo to models of *NF*O

Before we contemplate the second construction (which gives us a model containing $x \cap y$ and $x \cup y$ for all x and y) we had better ask ourselves why we didn't get it last time. After all, these axioms hold when $K = \{0, 1\}$. The point is that in that case everything is the same size as an old set or the complement of an old set. Also if x and y are both the same size as an old set or the complement of an old set, then so are $x \cap y$ and $x \cup y$, and so $\cap$ and $\cup$ do not construct anything that isn't already there. Once we have $B(x)$ this breaks down, and if we want $x \cup y$ and $x \cap y$ in general we have to construct them specially. This makes the task of constructing models of *NF*O altogether more daunting.

THEOREM 22 *Every model of ZF+ foundation has a $\mathcal{P}$-extension that is a model of NFO + low comprehension.*

Proof: (We certainly need not restrict ourselves to models of *ZF*, for this is certainly true for well-founded models of *Z* and presumably weaker theories as well.)

The idea of the construction was originally that $k(x)$ is to be a pair $\langle y, w\rangle$ where y is a set and w is a reduced word in some algebra with operations that correspond to the operations we want the universe of the new model to be closed under. Although this can be made to work, the approach it gives is very much less smooth than an approach that creates a gigantic free *NF*O model over a proper class of generators where there is one generator for each set of the old model. Unfortunately this second, smoother approach is not really a **CO** construction, and so strictly doesn't belong here as an illustration. The excuse for putting it in here is that the construction is, in spirit, very close to the **CO** constructions that precede it and will succeed it, and its presence here will help.

In the last case the only operations in which we were interested were unary (complementation and B) so the case had a spurious simplicity. Recall that the axioms of *NF*O are (apart from extensionality) existence of $\{x\}$ and closure under B, $\cup$ and $\cap$. If this were a **CO** construction *au pied de la lettre* we would get closure under singleton free because of low comprehension, so we could explicitly forget about it. Here too we can forget about it, and will indeed do so—for the moment. Later we will see why this is all right after all. We should be able to make do with only one of $\cap$ and $\cup$, but conjunctive and disjunctive normal forms for boolean words are so useful that we will retain both.

Our language of terms has a constant term g_x for each old set x. The constants will eventually correspond to low sets of the new model. We also have function letters $\cup$, $\cap$ $-$ and B. We have to do a bit of work to find the correct notion of *reduced word* for this algebra. There is the irritating feature that we do not want to have both $a \cap b$ and $b \cap a$ but we can get round that by wellordering the alphabet and extending the order lexicographically to the words. It will turn out that we will want to augment the language by adding Δ (symmetric difference).

Definition **23** A **restricted word** is either a constant, or is $W \Delta g$ where W is a boolean combination of Bs of restricted words, and g is a constant. If g is the empty set we can drop it and speak of a **pure** restricted word.

Now we have to show that everything that we wish to construct can be denoted by a restricted word.

We can think of a word `w` as a union of intersections, where the things being intersected are constants or Bs and complements of either. Consider an intersection like

$$a \cap b \cap c \cap d \ldots.$$

If even one of these is a constant (i.e., will correspond to an low set) then the whole intersection can be represented by just one (new) constant. Intersections of any number of complements of low sets can be represented as one complement of a low set. Thus the intersections are either constants, or intersections of Bs and $-B$s with the complement of a low set, which is to say, an intersection of Bs and $-B$s minus some low set.

Thus `w` can be rewritten in the form

$$[(w_1 \cap \overline{g_1}) \cup (w_2 \cap \overline{g_2}) \cup (w_3 \cap \overline{g_3})] \cup g_{n+1} \quad (1)$$

where the w_i are intersections of values of B or complements of values of B, and the g_i are low sets. We will work through this in the case where $n = 3$, so that the reader than see how to do the general case. (This is probably more helpful than a rigorous proof of the general case would be!) For the moment we are interested only in the stuff inside the square bracket of formula 1:

$$(w_1 \cap \overline{g_1}) \cup (w_2 \cap \overline{g_2}) \cup (w_3 \cap \overline{g_3}).$$

This expands to an intersection of $2^3 = 8$ subformulæ as follows:

$$(w_1 \cup w_2 \cup w_3) \cap \text{ seven other terms.}$$

A typical example of these other terms is $w_1 \cup w_2 \cup \overline{g_3}$. It is typical in that it contains at least one entry of the kind $\overline{g_i}$. Now $w_1 \cup w_2 \cup \overline{g_3}$ is the same as

$$-(\overline{w_1} \cap \overline{w_2} \cap g_3).$$

This is the complement of a low set so we can think of this as $\overline{g_{\text{novel}_1}}$ for some novel low set g_{novel_1}. This can be done to all the six remaining terms so the stuff-inside-the-square-bracket in formula 1 now looks like

$$(w_1 \cup w_2 \cup w_3) \cap \overline{g_{\text{novel}_1}} \cap \overline{g_{\text{novel}_2}} \cap \overline{g_{\text{novel}_3}} \cap \ldots \overline{g_{\text{novel}_7}}$$

subtracting finitely many low sets is the same as subtracting one, so the stuff-inside-the-square-bracket in formula 1 is now

$$(w_1 \cup w_2 \cup w_3) \cap \overline{g_{\text{novel}}}\ .$$

So, in the general case, we have reduced `w` to

$$(\bigcup_{i \leq n} w_i \setminus g_{\text{novel}}) \cup g_{i+1}.$$

This is actually

$$(\bigcup_{i \leq n} w_i) \Delta G,$$

where G is

$$((\bigcup_{i \leq n} w_n) \cap g_{\text{novel}}) \cup (g_{n+1} \cap -(\bigcup_{i \leq n} w_i)),$$

which is a low set.

Notice that we started with `w` as a boolean combination of B's and constants and have ended up with something rather simpler: $\bigcup_{i \leq n} w_i$ is a boolean combination of B's and G can be taken to be a constant.

This tidying-up process has not turned `w` into a restricted word, but if we invoke a notion of rank (where the rank of a word is simply the maximal depth of nesting of Bs in it) then we can see that none of the manipulations involved in the foregoing increases the rank of `w`, so that is we perform these manipulations successively on words of increasing rank, every word will eventually be manipulated into a restricted word.

Every word is equivalent to a restricted word. Verifying extensionality in this model will depend on our being able to show that it is equivalent to a *unique* restricted word. (note that it has not yet been made clear quite *how* it will depend on this uniqueness!) To do this it will be sufficient to show that if $\mathtt{w}_1$ and $\mathtt{w}_2$ are two pure restricted words then $\mathtt{w}_1 \ \Delta \ \mathtt{w}_2$ is not low. Suppose this were not so. Then there is some word W which evaluates to a low set. Without loss of generality we can think of W as a union of lots of terms each of which of course must themselves evaluate to something low and each such term is of the form $B(w_1) \cap B(w_2) \cap \ldots B(w_n) \cap - B(u_1) \cap - B(u_2) \cap \ldots - B(u_k)$. (where the $\vec{u}$ and the $\vec{w}$ are all distinct) which is supposed to be a low set. (Here "$B(x)$" means the set (in the sense of $\langle V, \in \rangle$) of those y such that $x \in_{\mathbf{co}} y$). Now $\{w_1, \ldots w_n\} \cup \{x\}$ belongs to this collection as long as x is not a u or a w, and there is certainly going to be a proper class of such x so it is clear that this collection must be a proper class (in the sense of $\langle V, \in \rangle$).

Now we have to define a new membership relation $\in_{\mathbf{co}}$. There is a bijection between the words of this new algebra and the sets of our old model. We may as well call this k as before. The elements of the model will be the words of the algebra. We will have to define $\in_{\mathbf{co}}$ by recursion on the structure of the terms.

If w is a molecular term of rank n then, in full generality, it is $w'\Delta g$ for some constant g and some boolean combination w' of $B(z_i)$ for various i. We then say $x \in_{\mathbf{co}} w$ iff $(x \not\in_{\mathbf{co}} g) \longleftrightarrow$(the obvious boolean combination of things like $z_i \in_{\mathbf{co}} x$)[4] This is a recursive definition that appeals to a rank function on *NF*O terms, where rank is the depth of nesting of Bs. What are we to make of '$x \in_{\mathbf{co}} g_y$', where g_y is a constant? This is defined to be $k^{-1}(x) \in y$. This ensures that every constant corresponds to a low set and *vice versa*. As before we can define an injection i from the old model into the new, by recursion. This time the recursion is $i(x) = g_{(k^{(-1)} \circ i)``x}$.

We had better verify that this does actually define a $\mathcal{P}$-extension. Suppose $y \in_{\mathbf{co}} i(x)$. Then $y \in_{\mathbf{co}} g_{(k^{(-1)} \circ i)``x}$ which is to say $k^{-1}(y) \in (k^{(-1)} \circ i)``x$ whence $y \in i``x$ and $y$ is a value of $i$. So the range of $i$ is transitive. Next suppose $i(y) \in_{\mathbf{co}} i(x)$. This is $i(y) \in_{\mathbf{co}} g_{(k^{(-1)} \circ i)``x}$ which in turn is $k^{-1}(i(y)) \in (k^{(-1)} \circ i)``x$ which is just $y \in x$. So $i$ is an isomorphism. Finally let $y$ be a low set of values of $i$. Then $i(y)$ will have to be $g_{(k^{(-1)} \circ i)``y}$ if this last thing is defined. But $(k^{(-1)} \circ i)``y$ is certainly a set so $g_{(k^{(-1)} \circ i)``y}$ is defined and is available to be $i(y)$.

■

One corollary of this is that any fragment of Z strong enough to execute this construction proves every $\Pi^{\mathcal{P}}_1$ theorem of *NF*O. See Forster-Kaye [1991] for an explanation of $\Pi^{\mathcal{P}}_1$ expressions.

[4] where the $\vec{z}$ are of course the free variables occurring in words of rank $n-1$ in w.

2.1.4 Church's Model

Church's models are all rough **CO** constructions in which every hereditarily low set has an n-cardinal. What is an n-cardinal? The 1-cardinal of x is the cardinal of x in the usual sense. And for larger n? Actually it doesn't matter a great deal what happens for larger n, since whatever we decide to mean by it it can be made to work. Sheridan is developing a construction that accomodates a version different from the one here (with the effect that the singleton function, considered as a set of Wiener-Kuratowski ordered pairs, is a union of finitely many of these cardinals). One natural version of the idea of n-cardinal is quite well-developed in *NF* studies. Recall that j is the operation on permutations defined by $(j(\pi))(x) = \pi``x$. Then we say x is n-equivalent to y if there is some permutation π of V so that $(j^n(\pi))(x) = y$. See Forster [1995] for more on this sort of n-cardinal. These ideas are anticipated in the eminently readable Tarski [1986].

In the version used here, two sets have the same $n+1$-cardinal iff there is a bijection between them such that the elements paired by the bijection have the same n-cardinal. We start off with the case $n = 1$ for simplicity's sake.

As with the example of the last section, this is not a strict **CO** construction. We need a bijection (k, as ever) between V and a set of codes for objects. We will use the notation '$|x| = |y|$' to mean that x and y are the same size. We will see very soon how these fake terms ('$|x|$' etc.) can be treated as genuine denoting terms.

Definition **24** The things that are are values of k are either:

1. ordered pairs $\langle x, i\rangle$ where x is an arbitrary set and i is 0 or 1 (this will provide low sets and complements of low sets as usual), or
2. ordered pairs $\langle i, \kappa\rangle$ where κ is a cardinal (other than 0) and i is either I or II. I and II are two unspecified distinct objects that weren't in the original ground model. The idea is that these objects are to be cardinals (and complements of cardinals) in the new model.

Now we say $y \in_{\mathbf{co}} x$ iff

1. $\texttt{snd}(k(x)) = 0$ and $y \in \texttt{fst}(k(x))$, *or*
2. $\texttt{snd}(k(x)) = 1$ and $y \notin \texttt{fst}(k(x))$, *or*
3. $\texttt{fst}(k(x)) = \mathrm{I}$ and $\texttt{snd}(k(y)) = 0$ (so y is low) and $|\texttt{fst}(k(y))| = \texttt{snd}(k(x))$, *or*
4. $\texttt{fst}(k(x)) = \mathrm{II}$ and $\texttt{snd}(k(y)) \neq 0$ (so y is low) or $|\texttt{fst}(k(y))| \neq \texttt{snd}(k(x))$.

Clauses 1 and 2 make sure that every low set has a complement. Notice that nothing has been said about what cardinal numbers *are*. Notice also that this does not matter! All we need is that there should be a definable class $\mathcal{C}$ and a definable relation `belongs-to` between sets and members of $\mathcal{C}$ satisfying

$$(\forall x)(\exists! y \in \mathcal{C})(x \texttt{ belongs-to } y),$$

$$(\forall x \forall y)(|x| = |y| \longleftrightarrow (\forall z \in \mathcal{C})(x \texttt{ belongs-to } z \longleftrightarrow y \texttt{ belongs-to } z)).$$

The term $|x|$ can then be taken to denote the appropriate member of $\mathcal{C}$. We do not need the axiom of choice to define cardinal numbers since as long as we have foundation (which we are assuming here) we can use Scott cardinals. The Scott cardinal of x is the set of all things the same size as x that are of minimal rank with this property.

Clause 3 will ensure that every low set has a cardinal in the new model (in the strong sense that for every low set x, the collection of all sets that have the same cardinal as x is a set of the new model). We stipulate that cardinals used do not include 0. We do this for two reasons: (i) to keep 0 free to signal low sets as usual, also (ii) because the extension of the cardinal number 0 (the set of all empty sets) is a set by low comprehension anyway, and we do not wish it to make difficulties for ourselves with extensionality by manufacturing it twice. Clause 4 ensures that the complement of every such cardinal is a set.

The usual apparatus of low comprehension can now be taken for granted. It should by now be clear that this model is a model of complementation. It is the existence of cardinals that we had better spend a bit of time verifying.

Proposition 25 *The clauses of definition 24 give a model in which every low set x has a cardinal: $\{y : |y| = |x|\}$.*

Proof: Notice that the cardinals that we have created by this means are demonstrably neither low nor are the complements of low sets, which makes life much easier. Let x be any low set. Consider the ordered pair $\langle \mathrm{I}, |\texttt{fst}(k(x))|\rangle$. We will check that $k^{(-1)}(\langle \mathrm{I}, |\texttt{fst}(k(x))|\rangle)$ is the cardinal of x (in the sense that it is the set of all things the same size as x) in the new model. By clause 3 we have $y \in_{\mathbf{co}} k^{(-1)}(\langle \mathrm{I}, |\texttt{fst}(k(x))|\rangle)$ iff y is low and $|\texttt{fst}(k(y))| = |\texttt{fst}(k(x))|$. Since x and y are both low this is the same as saying that the set (in the old sense) of things $\in_{\mathbf{co}} y$ is the same size (in the old sense) as the set of things $\in_{\mathbf{co}} x$, so there is a bijection between these two (old) sets. This bijection is an (old) set of (old) ordered pairs. By low comprehension (theorem 11) the corresponding (new) set of (new) ordered pairs is also a set, so x and y are of the same size in the new sense as well. The other direction is easy. Therefore $k^{(-1)}(\langle \mathrm{I}, |\texttt{fst}(k(x))|\rangle)$ is indeed the cardinal of x in the new model. Correspondingly $k^{(-1)}(\langle \mathrm{II}, |\texttt{fst}(k(x))|\rangle)$ is the complement of that cardinal, which we have to have if complementation is to be true. ■

We have to do a littl bit of work to see how to generalise this correctly to the case $n = 2$, the model where every low set of low sets has a 2-cardinal. What Church actually claims is that for each n his construction gives us a model where every well-founded set has an n-cardinal. I prefer the statement in terms of low sets, low sets of $\ldots^n$ low sets.

For the case $n = 2$ we have to add two more clauses 5 and 6 to definition 24 in the same style. We will need two more novel constants in the style of I and II, which we may as well write 'III' and 'IV'. Objects x s.t. $\texttt{fst}(k(x)) =$ III will be 2-cardinals and objects x s.t. $\texttt{fst}(k(x)) =$ IV will be complements of 2-cardinals. We will need the notation '$\texttt{2-card}`x$' for the 2-cardinal of x, and we will use lower case greek letters to range over 2-cardinals as over cardinals. Then there are to be two further kinds of ordered pairs in the range of k: pairs whose first components are III and pairs whose first components are IV. In both cases the second components are 2-cardinals. We will need the two following new clauses in the definition of $y \in_{\mathbf{co}} x$.

Definition 26

5 $\texttt{fst}(k(x)) = \mathrm{III}$ and y is a low set of low sets and

$$\texttt{2-card}(\{\texttt{fst}(k(z)) : z \in \texttt{fst}(k(y))\}) = \texttt{snd}(k(x)).$$

6 $\texttt{fst}(k(x)) = \mathrm{IV}$ and (y is not a low set of low sets or

$$\texttt{2-card}(\{\texttt{fst}(k(z)) : z \in \texttt{fst}(k(y))\}) \neq \texttt{snd}(k(x))).$$

Clause 6 of course ensures that 2-cardinals, too, have complements. The details will be omitted.

Proposition 27 *The membership relation of definition 26 gives a model in which each low set of low sets has a 2-cardinal.*

Proof: Let x be a low set of low sets. Then the 2-cardinal (in the sense of $\in_{\mathbf{co}}$) of x will be $k^{(-1)}(\langle \mathrm{III}, \texttt{2-card}`\{\texttt{fst}(k(z)) : z \in \texttt{fst}(k(x))\}\rangle)$. We had better check this. Suppose y is a low set of low sets. Then

$$y \in_{\mathbf{co}} k^{(-1)}(\langle \mathrm{III}, \texttt{2-card}`\{\texttt{fst}(k(z)) : z \in \texttt{fst}(k(x))\}\rangle)$$

iff

$$\texttt{2-card}(\{\texttt{fst}(k(z)) : z \in \texttt{fst}(k(x))\}) = \texttt{2-card}(\{\texttt{fst}(k(z)) : z \in \texttt{fst}(k(y))\}).$$

What we actually want is for x and y to have the same 2-cardinal in the new sense. As before, if there is an (old) bijection between $\texttt{fst}(k(x))$ and $\texttt{fst}(k(y))$ there will be a new bijection between x and y by low comprehension. And the same goes, not only for x and y, but for each $x' \in_{\mathbf{co}} x$ and $y' \in_{\mathbf{co}} y$ that are paired by the bijection: if there is an (old) bijection between $\texttt{fst}(k(x'))$ and $\texttt{fst}(k(y'))$ there will be a new bijection between x' and y' by low comprehension as desired. ■

It should now be clear how to tinker with this construction to add simultaneously for all $n \in \mathbb{N}$, the assertion that every (low set of)^n low sets has an n-cardinal. It is perhaps worth noting that it doesn't seem to be necessary to argue that, for each $n \in \mathbb{N}$, we can do this for all $m < n$ and then use compactness.

2.2 Wellfounded sets in CO-structures

The roots that this technique has in Rieger-Bernays permutation models still have fruit to bear, as witness the following two theorems.

THEOREM **28** *For a given choice of K and V and rules, all* **CO** *structures are permutation models of each other.*

Proof: Suppose we have two **CO** structures over a model $\langle V, \in \rangle$, with the same K but different coding functions k and k' respectively. We wish to find $\sigma \in Symm(V)$ such that the first model thinks that $x \in y$ iff the second thinks $x \in \sigma(y)$. σ must be $(k')^{-1} \circ k$. ■

We also have the following:

PROPOSITION **29** *Every permutation model $\langle V, \in_{\mathbf{co}} \rangle^{\sigma}$ of $\langle V, \in_{\mathbf{co}} \rangle$ is obtained from it by replacing k by some k', with a corresponding new membership relation $\in_{\mathbf{CO}'}$. If the permutation is σ, then the new k' is $\sigma^{-1}k$.*

Proof:

$$\langle V, \in_{\mathbf{co}} \rangle^{\sigma} \models x \in y \text{ iff } \begin{cases} \langle V, \in_{\mathbf{co}} \rangle \models x \in \sigma(y), \\ \langle V, \in \rangle \models x \in_{\mathbf{co}} \sigma(y), \\ \langle V, \in \rangle \models x \in_{\mathbf{CO}'} y, \\ \langle V, \in_{\mathbf{CO}'} \rangle \models x \in y. \end{cases}$$

■

THEOREM **30**

1. *H_{low} is always isomorphic to a permutation model of the original universe.*
2. *Whatever K we started with, for any permutation σ of the old universe we can find a coding function k so that $\langle H_{\mathrm{low}}, \in_{\mathbf{co}} \rangle \simeq \langle V, \in_{\sigma} \rangle$.*

Proof:

(1) There will be a bijection $\pi : V \longleftrightarrow H_{\mathrm{low}}$. We seek a σ so that

$$(\forall xy)(x \in \sigma(y). \longleftrightarrow .\pi(x) \in_{\mathbf{co}} \pi(y)).$$

What is $\sigma(y)$? Clearly it has to be $\{x : \pi(x) \in_{\mathbf{co}} \pi(y)\}$. This is a set, since $\pi(y)$ is low. Must check that this definition gives us a σ that is 1-1 and onto. It is certainly 1-1 by extensionality of $\in_{\mathbf{co}}$. Is it onto? Given z we must find a y so that $z = \{x : \pi(x) \in_{\mathbf{co}} \pi(y)\}$. This y must be $\pi^{-1}(k^{-1}(\langle \pi``z, 0 \rangle))$.

(2) We know—however we choose k—that H_{low} is a proper class whose complement is a proper class, so let π be a bijection between V and such a class and let us fasten on that class to be H_{low} and resolve to cook up k so that it actually *is* the H_{low} of the new model. Dugald Macpherson has used the word "moiety" for things that are both infinite and coinfinite: we will borrow it here

to describe proper classes whose complements are proper classes. Let σ be a permutation of V. We want to cook up k so that $\langle H_{\text{low}}, \in_{\mathbf{co}}\rangle \simeq \langle V, \in_\sigma\rangle$. We want

$$(\forall xy)(x \in \sigma(y). \longleftrightarrow .\pi(x) \in_{\mathbf{co}} \pi(y)).$$

The right hand side is $\pi(x) \in \mathtt{fst}(k(\pi(y)))$ (and $\mathtt{snd}(k(\pi)(y)) = 0$ since $\pi(y)$ is a low set). Now $\mathtt{fst}(k(\pi)(y)) = \pi``\sigma(y)$ so we want $k(\pi)(y) = \langle \pi``\sigma(y), 0\rangle$. It is true that this only tells us what k should do to values of π but since the range of π is a moiety and the range of $k \circ \pi$ is also a moiety there will be no problem extending this to a bijection between V and all the ordered pairs we need. ∎

Analogously with the embedding $i : V \hookrightarrow V^\sigma$ defined in theorem 6 we can define a canonical injection from the original model into the new model $\langle V, \in_{\mathbf{co}}\rangle$ defined by recursion on $\in$:

Definition 31 $i(x) =: k^{(-1)}(\langle i``x, 0\rangle)$.

Like the i of theorem 6 this embedding is a $\mathcal{P}$-embedding.

Theorem 32 *If $\langle V, \in\rangle$ is well-founded then i is defined and is a $\mathcal{P}$-embedding from $\langle V, \in\rangle$ into $\langle V, \in_{\mathbf{co}}\rangle$.*

Proof: First we prove by $\in$-induction that i is defined on all sets.

We must next check that i is an isomorphism so we want $i(x) \in_{\mathbf{co}} i(y) \longleftrightarrow x \in y$. Since the second component of $k(i(y))$ is 0, $i(x) \in_{\mathbf{co}} i(y)$ iff $i(x) \in \mathtt{fst}(k(i(y))) = i``y$ iff $x \in y$.

Next we show that the range of i is transitive$^{(\in_{\mathbf{co}})}$. Suppose y is in the range of i, and $y = i(z)$. So $x \in_{\mathbf{co}} i(z) = k^{(-1)}(\langle i``z, 0\rangle)$ iff $x \in i``z$ so x would also be in the range of i.

Finally we must check that any subset of something in the range of i is likewise in the range of i. Suppose x is in the range of i, so that $k(x) = \langle i``z, 0\rangle$ for some $z$. Suppose also that $(\forall w)(w \in_{\mathbf{co}} y \rightarrow w \in_{\mathbf{co}} x)$. Consider the set (in the sense of the original model) of those things that are $\in_{\mathbf{co}} y$. This is indeed a set of the original model, since it is a subset of $i``z$. If it is $i``u$, then $k(y)$ must be $\langle i``u, 0\rangle$ so y is in the range of i. ∎

(Essentially this proof is in Church [1974] though he prefers to say that the well-founded sets form a model of ZF, and does not have the concept of a $\mathcal{P}$-embedding.)

The difference here from theorem 6 is that this time we cannot expect i to be elementary for stratified formulae. (It is a good question which formulæ it is elementary for!) However we can at least derive the following

Corollary 33 *The $\mathcal{P}$-embedding $i : \langle V, \in\rangle \rightarrow \langle H_{\text{low}}, \in_{\mathbf{co}}\rangle$ is elementary for stratified formulæ.*

Proof: Since H_{low} is isomorphic to a permutation model of V (by theorem 30) we can use theorem 6. ■

Does this mean that we can take the well-founded sets of the new model to be the range of i? No, because we can arrange for the range of i to be a set of the new model, and it will be in some sense well-founded.

Proposition 34 *The range of the canonical embedding can be a set of the new model.*

Proof: We use a slight modification of Oswald's original construction from section $\in H$; if x is a finite subset of H then $\Sigma_{n \in x} 2^{n+1} \in H$. Then say $x \in_{\mathbf{co}} y$ iff *either* y is $2n$ and the xth bit of n is 1; *or* y is $2n+3$ and the xth bit of n is 0; *or* $y = 1$ and $x \in H$.

To complete the proof we think of $\mathbb{N}$ as a copy of V_ω by associating with it the Ackermann relation alluded to above. The canonical embedding i is then defined on everything in V_ω which is a set of the new model. ■

However, although it is possible for the range of i to be a set, it isn't always a set.

Proposition 35 *If the three conditions of section 2.1.1 are met the range of i is not a set.*

Proof: We want to show that if X satisfies $(\forall y \in_{\mathbf{co}} X)$($y$ is in the range of i) then X is in the range of i too. The case $\mathtt{snd}(k(x)) = 0$ we have already considered. There remains the case $\mathtt{snd}(k(X) = 1)$.

We will show that this case cannot occur. If it did, $x \in_{\mathbf{co}} X$ iff $x \notin \mathtt{fst}(k(x))$. So X would be so big that there is only a set of things that aren't members$^{(\in_{\mathbf{co}})}$ of it, and all the things that *are* would be values of i. But the collection of values of i is not so big that its complement is a set, since if $x \in$ range i, $\mathtt{snd}(k(x)) = 0$, and there is a proper class of x such that $\mathtt{snd}(k(x)) = 1$. Therefore $\mathtt{snd}(k(x)) \neq 1$.

■

The possibility of adding new well-founded sets in the new model in this *ad hoc* way restricts the things we can say in general about well-founded sets, at least if we take well-founded sets to be those defined in the natural way by means of the inductive definition or as regular sets (definition 36).

The inductive definition is the obvious one: if we think of well-founded sets as those over which we can do $\in$-induction then we are led to the inductive definition:

$$WF(x) \longleftrightarrow (\forall y)((\forall z)(z \subseteq y \to z \in y) \to x \in y). \tag{2}$$

The trouble with this definition is that if we have very little comprehension there may be so few y such that $(\forall z)(z \subseteq y \to z \in y)$ that lots of sets might turn out to be well-founded that shouldn't. In fact, in ZF there are no such sets at all and this definition is completely useless! We talk of **regular** sets instead.

DEFINITION **36** *x is regular iff* $(\forall y)(x \in y \to (\exists z \in y)(z \cap y = \emptyset))$.

If we have an axiom of complementation we can prove:

PROPOSITION **37** *A set is well-founded in the inductive sense iff it is regular.*

Proof: Suppose $(\forall y)((\forall z)(z \subseteq y \to z \in y) \to x \in y)$. Substitute $-y$ for y getting $(\forall y)((\forall z)(z \cap y = \emptyset \to z \notin y) \to x \notin y)$. Contrapose getting $(\forall y)(x \in y \to \neg(\forall z)(z \cap y = \emptyset \to z \notin y))$. This is $(\forall y)(x \in y \to (\exists z)(z \cap y = \emptyset \wedge z \in y))$ which says that x is regular. ■

The inductive definition enables us to prove that every well-founded set has ϕ as long as the extension of ϕ is a set and any set of things-which-are-ϕ is itself ϕ. This is a lot less applicable than it might seem because of the extreme unlikelihood that the extension of ϕ should be a set. We can prove that if $x \in x_1 \in x_2 \ldots x_n \in x$ then $-\{x, x_1 \ldots x_n\}$ is a set y such that $(\forall z)(z \subseteq y \to z \in y)$ and so no well-founded set is a member^n of itself, but one would expect to be able to prove a lot more.

There is also a principle of induction for regular sets. This is standard in modern treatments of set theory, and we do it as follows. Suppose we know $(\forall x)((\forall y \in x)(\phi(y)) \to \phi(x))$, and suppose there is a regular set z such that $\neg\phi(z)$. Let Y be a transitive set containing z. Let $X = \{y \in Y : \neg\phi(y)\}$. $z \in X$ and z is regular, so X must be disjoint from one of its members. But this contradicts the induction hypothesis, so $\phi(z)$.

How does this fare in the Church-Oswald context? If we know that any set of things-which-are-ϕ is itself ϕ and that for any regular z there is a transitive set Y containing z such that $\{y \in Y : \neg\phi(y)\}$ exists, then we can prove that every regular set is ϕ. Can we do this for all regular z and all ϕ? If we could show that for every regular z there is a low transitive Y with $z \in Y$ then we can do the rest by low comprehension. We could do this if we knew that every well-founded set is in the range of i but proposition 34 will tell us that this might not be so.

H_{low}, the hereditarily low sets, might capture better the idea of sets well-founded in the new model. Here we have to remember that foundation fails and that therefore there are several notions of hereditarily low. In particular there may be hereditarily low sets that are not well-founded, for example Quine atoms, which are sets equal to their own singletons. Easy though this is to arrange, it is also easy to avoid. We can arrange that every set that is hereditarily low (in the sense that everything in its transitive closure is low) is also in the range of i. Suppose $\langle x_n : n \in \mathbb{N} \rangle$ is a descending ω-sequence under $\in_{\mathbf{co}}$ and that x_0 is hereditarily low. Then for all n, $\mathtt{snd}(k(x_n))$ is 0, so $x_{n+1} \in_{\mathbf{co}} x_n$ is just $x_{n+1} \in \mathtt{fst}(k(x_n))$, so x_{n+1} is a set of lower rank than $\mathtt{fst}(k(x_n))$. If we set up k so that for all y the rank of $\texttt{fst}(k(y))$ is no greater than the rank of y then there can be no such infinite descending sequence in the new model and all hereditarily low sets are well-founded. In particular this will happen whenever the three conditions of section 2.1.1 are met.

What this says is that we can arrange that the only sets that are illfounded are big, or at least have something big in their transitive closure.[5] There is a similar problem in *NF*: must self-membered sets be big? When I asked this question in the first edition of Forster [1995], Maurice Boffa came up with a permutation model which partly answered my question. The permutation he used (and which he and Pétry reported in [1993]) uses a device a bit like a rank function and it admits a great deal of refinement. It is now possible using results of Körner [1994] to show that it is consistent relative to *NF* that $\in$ restricted to finite sets is well-founded. This is done by Rieger-Bernays permutation models. It sounds a weaker result than arranging for $\in$ restricted to low sets to be well-founded but bear in mind that in the present case all we have to do is prevent the appearance of new ill-founded small sets in the new model, whereas in the *NF* case we not only have to do that but we have to kill off any small ill-founded sets in the model we started with, and—being a model of *NF*—it may have had lots.

Nice though it is to know that the greatest fixed point for "set of all low subsets of" and the least fixed point can be the same, what really matters is that it is the least fixed point we want for our concept of well-founded in the new model. A treatment now follows.

Definition 38 Let us say x is well-founded* if $(\forall X)((\forall y)((low(y) \wedge y \subseteq X) \to y \in X) \to x \in X)$.

In other words the collection of well-founded* sets is the least fixed point corresponding to the greatest fixed point H_{low}.

We will justify a principle of unrestricted $\in$-induction for well-founded* sets.

First we check (i) that the collection of well-founded* sets is transitive and (ii) that everything in it is low.

(i) If x is well-founded*, and $(\forall y)((low(y) \wedge y \subseteq X) \to y \in X)$ then $x \in X$. It will suffice to show that $x \subseteq X$ as well.

Suppose $x \not\subseteq X$. We will show that $(\forall y)(low(y) \wedge (y \subseteq (X \setminus \{x\})) \to y \in (X \setminus \{x\}))$ whence $x \in (X \setminus \{x\})$ (since x is well-founded*). This is impossible.

Suppose $y \subseteq (X \setminus \{x\})$ and is low. Then $y \subseteq X$ and $y \in X$. To deduce $y \in (X \setminus \{x\})$ it will suffice to show $y \neq x$, which would follow from $x \not\subseteq (X \setminus \{x\})$. But we have assumed that $x \not\subseteq X$ so *a fortiori* $x \not\subseteq (X \setminus \{x\})$.

(ii) Suppose x is well-founded* but not low, and that $(\forall y)((low(y) \wedge y \subseteq X) \to y \in X)$. Let y be a low subset of $(X \setminus \{x\})$. Then y is a low subset of X and is therefore a member of X. We want it to be a member of $X \setminus \{x\}$ so we

[5] **Big** sets (French 'gros') are things like complements of singletons. They are not **large** (French 'grand') in the way that, say, measurable cardinals are large.

want $y \neq x$. But y is low and x isn't. Therefore $(\forall y)((low(y) \wedge y \subseteq X) \to y \in (X \setminus \{x\}))$ and $x \in (X \setminus \{x\})$ since x is well-founded*.

Next we need to know that $TC(x)$ exists if x is well-founded*. We prove by induction on $\mathbb{N}$ that if x is well-founded* then $\bigcup^n x$ exists and is low. Low comprehension tells us that a low set of low sets has a low sumset. The induction is good because $\mathbb{N}$ is a low set and so there is enough comprehension for full induction over it. So the collection $\{\bigcup^n x : n \in \mathbb{N}\}$ is a low set of low sets and its sumset—$TC(x)$—is a low set.

Next by relativising the proof of proposition 37 we show that x is well-founded* iff $(\forall y)(x \in y \to (\exists z \in y)(x \cap y = \emptyset \wedge low(z)))$.

Finally we prove $\in$-induction as follows. Suppose we know $(\forall x)((\forall y \in x)(\phi(y)) \to \phi(x))$, and suppose there is a well-founded* set z such that $\neg\phi(z)$. Let $X = \{y \in TC(\{z\}) : \neg\phi(y)\}$ which exists by low comprehension. $z \in X$ and z is well-founded*, so X must be disjoint from one of its members. But this contradicts the induction hypothesis, so $\phi(z)$. ■

We can prove this unrestricted scheme of $\in$-induction scheme for well-founded* sets despite being unable to prove it for sets well-founded in the inductive sense we began with because there may well be big well-founded sets about which we know nothing.

Armed with this scheme of $\in$-induction for well-founded* sets we can prove that every well-founded* set is in the range of i.

Perhaps the moral of this discussion is that we should restrict ourselves to coding functions k which add no new well-founded sets. In those circumstances "low" is definable in the new language as "same size as a well-founded set" as Church originally intended, and the whole enterprise becomes axiomatisable. It is true that we lose some generality by making this restriction, but we sacrificed generality at the outset of this section by not considering models of *ZF* without foundation. Neither of these seems a harsh sacrifice.

3 Open problems

3.1 The Axioms of Sumset and Power set

At present the only form of the axiom of sumset we can prove for this construction is that a low set of low sets has a sumset (which will be low). Can we tweak this construction to get a less restricted axiom of sumset, dropping the second or perhaps even the first occurence of 'low' in the above? Similarly the only form of the axiom of power set that this construction apparently gives us is power sets of low sets. With neither of these two axioms is there a standard paradox obviously skulking in the wings waiting to cause trouble should we adopt unrestricted forms of them. It is natural to see if we can do better than this. Mitchell's set theory allows power set, but has other disadvantages. For example, $x \cup y$ and $x \cap y$ do not exist in general.

The question is also discussed in Sheridan's thesis.

3.2 Natural strengthenings of theorem 22

This is the most general and the most important of the open problems in this area. Such extensions would be theorems of the form "Every well-founded model of T_1 is the well-founded part of a model of T_2" (where T_1 is a theory consistent with the axiom of foundation and T_2 is a theory with an antifoundation axiom). It is perhaps in this connection that Church's remark in [1974] seems most pertinant: "On source of added axioms to be studied for their consistency with the basic axioms is ... Quine set theory. ... An interesting possibility ... is a synthesis or partial synthesis of *ZF* and Quine set theory." Let *KF* be the subsystem of Zermelo set theory obtained by dropping the axiom of infinity and restricting the assonderung (comprehension) scheme to stratified Δ_0 formulæ. A natural conjecture for *NF*istes to make in this context would be that:

CONJECTURE **39** *If NF is consistent then every* (*well-founded*) *model of KF is the well-founded part of a model of NF.*

Another possibility is:

CONJECTURE **40** *Every model of KF has a $\mathcal{P}$-extension that is a model of NF+ low replacement.*

There is a basic difficulty in the path of anyone trying to prove anything like conjectures 39 and 40. The **CO** construction is a method which—on being presented with (i) a robust method of constructing term models for a trivial subsystem T of *NF*, and (ii) a model $\mathcal{M} \models ZF$—outputs a model of T which has $\mathcal{M}$ as an initial segment of its well-founded part. To make this work, we have to have the method of generating term models for T *in the hand*, as it were. It is not enough to know that T is consistent. (The system *NF*$\forall$ of Forster [1987] has a canonical term model so we might be able to use Church-Oswald constructions to show that every (well-founded) model of *KF* is the well-founded part of a model of *NF*$\forall$.) *A fortiori* we do not have a method of proving conditionals of the kind: if $T \subseteq NF$ is consistent, and $\mathcal{M}$ is a model of ZF then there is $\mathcal{M}' \models T$ of which $\mathcal{M}$ is the well-founded part. There are plenty of interesting assertions of this kind, and it would be very nice to know which of them were true.

3.3 Axiomatisability

We should axiomatise the theories that **CO** constructions give us: low comprehension is obviously important but not obviously axiomatisable. We could do this if we could find in the language of $\in_{\mathbf{co}}$ and $=$ a predicate which meant "low". We can do this if the model we started with satisifes the three condition of section 2.1.1 but it just may be possible to say something even without these extra assumptions.

3.4 Constructive CO constructions

Another question is: is there a sensible constructive treatment of **CO** constructions? The problem is the very classical nature of **CO** constructions: $\texttt{snd}(k(x))$ is always equal to 0, or to something else. This seems to mean that if we want to execute intuitionistic **CO** constructions we will have to assume *tertium non datur* for atomics. It would be quite illuminating to develop **CO** constructions inside intuitionistic *ZF* and see what happens to the transfinite recursive versions of the negative interpretation. It is probably quite hard. Indeed, as unpublished work of Dzierzgowski has shown, it is extraordinarily diffcult to get nontrivial models for intuitionistic versions even of systems as straightforward as NF_2.

3.5 Schröder-Bernstein

The point has been made that in Church's theory and its kin we have no comprehension to speak of for big sets. This means that even apparent banalities can turn out to be hard to prove. For example can we prove Schröder-Bernstein for big sets in Church's set theory? There are various proofs of the Schröder-Bernstein theorem. There is a slick one suitable for use with computer science students and others who have been exposed to the Tarski-Knaster fixpoint theorem for complete lattices. Suppose $f : A \hookrightarrow B$ and $g : B \hookrightarrow A$. Then $\lambda A'.(A \setminus g``(B \setminus f``(A'))$ is a continuous function on the complete lattice $\mathcal{P}(A$ and must have a fixpoint. However this proof depends on the power set axiom, which is not available in Church's theory. There is a much more lo-tech proof *not* using power set for which the outlook is slightly more hopeful. It goes like this.

Consider the two sequences defined by a mutual recursion:

$$b_0 = B \setminus f``A; \qquad b_{n+1} = f``a_n,$$
$$a_0 = A \setminus g``B; \qquad a_{n+1} = g``b_n.$$

Set

$$A' = \bigcup_{n \in \mathbb{N}} a_n.$$

The bijection we want is $f|A' \cup g^{-1}|(A \setminus A')$.

There seem to be two major hurdles to making this work in Church's set theory. (i) Are all the b_n and the a_n sets? If so then $\{a_n : n \in \mathbb{N}\}$ is a set by low comprehension. (ii) If $\{a_n : n \in \mathbb{N}\}$ is a set, is its sumset a set? If A and B are low, then, by low comprehension, all the b_n and the a_n are low sets, so

$\{a_n : n \in \mathbb{N}\}$ is a low set of low sets and its sumset is a low set, and there seems no reason to expect that the bijection $f|A' \cup g^{-1}|(A \setminus A')$ won't also be a low set. However if A and B are merely sets, not low sets, there doesn't seem to be any reason to expect this proof to work. There may be some other way of proving Schröder-Bernstein but there is no reason to expect that either.

One special case can be disposed of easily. In models satisfying the three conditions of section 2.1.1, Schröder-Bernstein will be true, for the following very unsatisfactory reason. In such models every set is either low or co-low. Schröder-Bernstein will certainly hold for the low sets, and no low set will be the same size as a co-low set. If A and B are two co-low sets then the model cannot contain any injections from A into B or B into A since any such injection would be neither low nor co-low (it will be a *moiety*) and must be absent from the model. So Schröder-Bernstein will be vacuously true for non-low sets. (Indeed, because of the absence of moieties, no non-low set can even be the same size as itself!) This argument cannot be used in the case of Church's model, because there are plenty of moieties there.

3.6 Mitchell's set theory

Emerson Mitchell's Ph.D. Thesis [1976] contains a **CO**-like construction. His system has power set for all sets (not just low sets) but does not have closure under (binary) $\cup$ and $\cap$.

3.7 Extensional quotients?

I close by raising an obvious but completely unexplored possibility. The enduring difficulty with these constructions is extensionality. It is easy to show that x and y have the same members (in the sense of $\in_{\mathbf{co}}$) as long as $\texttt{snd}(k(x)) = \texttt{snd}(k(y))$, but if $\texttt{snd}(k(x)) \neq \texttt{snd}(k(y))$ we have our work cut out, and it tends to be feasible only for theories that have an easily solvable word problem. One thing one could naturally do is start by being much more reckless in one's choice of K and membership conditions for non-low sets, and pick up the pieces later by taking an extensional quotient of the result. It is hard to see what might be preserved in a development like this, but that means that there may be many consistency proofs waiting to be revealed by such an extension of the method. My present feeling is that the possibility of developing CO techniques along these lines is their most exciting feature. There is no obvious reason why this should not hold the key that will one day unlock the consistency question of *NF*. There is some work by Jamieson and by Antonelli on extensional quotients which can be found on the NF bibliography.

Bibliography

Antonelli, G. A. Extensional quotients for Type Theory and the consistency problem for NF. JSL 1998 **63** no. 1. pp247-61

Boffa, M. and Pétry, A. [1993] On self-membered sets in Quine's set theory *NF*. *Logique et Analyse* **141-142** pp. 59-60.

Church, A. [1974] Set theory with a universal set. *Proceedings of the Tarski Symposium.* Proceedings of Symposia in Pure Mathematics XXV, ed. L. Henkin, Providence RI pp. 297−308. Also in *International Logic Review* **15** pp. 11−23.

Forster, T.E. [1987] Term models for weak set theories with a universal set. *Journal of Symbolic Logic* **52** pp. 374−87.

Forster, T.E. [1995] Set theory with a universal set. *second edition* Oxford Logic Guides, Oxford University Press.

Forster, T.E. and Kaye, R. [1991] End-extensions preserving power set. *Journal of Symbolic Logic* **56** pp. 323−28.[6]

Grishin, V.N. [1969] Consistency of a fragment of Quine's NF system. *Soviet Mathematics Doklady* **10** pp. 1387−90.

Jamieson, M Set theory with a universal set. Ph.D. thesis, Univ of Florida 1994. 114pp

Jensen, R.B. [1969] On the consistency of a slight(?) modification of Quine's NF. *Synthese* **19** pp. 250−63.

Oswald, U. [1976] Fragmente von "New Foundations" und Typentheorie. Ph.D. thesis, ETH Zürich.

[6]Errata. p 327. Line 11 should read 'and $a \in M$ such that $M \models |\pi`a| = |\mathcal{P}(a|$'. Line 13: the expression following '$M \models$' should be '$|\pi`a| = |\mathcal{P}(a|$'. Line 26: '(not just $\pi`a = \mathcal{P}(a)$' should read '(not just $|\pi`a| = |\mathcal{P}(a|)$'. Line 28: '$\pi`a$' should read '$|\mathcal{P}(\pi`a|$'.