The Ackermann Arithmetic of a Set Theory (with special attention to stratified set theories)

Thomas Forster

January 19, 2025

Contents

1 Justifying stratified $\in$-induction over V_ω **2**

2 Which Arithmétic Expressions are stratified? **3**

3 A Remark about IO **6**

4 Envoi **7**

5 Work still to be done **7**

ABSTRACT

Ackermann's celebrated encoding of hereditarily finite sets by natural numbers gives a way of extracting an arithmetic from any set theory that proves the existence of all hereditarily finite sets. This derivation is more sensitive to stratification constraints than is (e.g.) the Frege/Russell method of equipollence classes; here we examine how some of those constraints play out.

There are at least four distinct ways in which one can extract an arithmetic from a set theory. The first is the usual, obvious, way: look at the finite cardinals of the theory, what it proves about them (if you are working with a stratified set theory – and such theories will be our main preoccupation here) possibly restrict attention to strongly cantorian finite cardinals in order to be more confident that there is no last natural). There is also Church numerals.

> The device of von Neumann naturals (finite von Neumann ordinals) is also available but this is perhaps better thought of as a way of interpreting arithmetic into a set theory rather than a way of ascertaining what the theory itself believes about natural numbers. We will not have much to say about the von Neumann interpretation: for one thing, it is generally well-understood; for another it is an unnatural device to use with stratified set theories which – as *per* the title – will be our princple focus.

A fourth way – which i should've tho'rt of *ages* ago – is *via* the Ackermann interpretation. Let T be a set theory, and think about T's theory of hereditarily finite sets, and what theory of natural numbers one obtains thence by means of the Ackermann interpretation; call this the *Ackermann Arithmetic* of T. If

T is a set theory of a convenional sort, with full separation/aussonderung, then its Ackermann arithmetic is presumably the same as the arithmetics obtained by the first two methods. However, if T is a stratified set theory such as NF, NFU, KF etc (with only stratified separation) it's presumably going to be a much weaker system than one gets from the first two methods, but it's at any rate something different. An investigation could be enlightening; it could tell us something about the arithmetical meaning of unstratfied assertions in set theory à la [1].

This fourth way is slightly trickier to deal with than the first two (equipollence classes and Church numerals) in that it doesn't appear to be invariant under Rieger-Bernays permutations. It might be, but we haven't got anything that looks remotely like a proof. Let us not allow ourselves to be put off by this discouraging start!

HF is defined as the least fixed point for $x \mapsto \mathcal{P}_{\aleph_0}$, the set of finite subsets of x, so it supports $\in$-induction for whatever properties the enveloping set theory has separation for. We will write 'HF' or 'V_ω' indiscriminately. 'HF' could mean the greatest fixed point, but it is only the least fixed point we are interested in so there is no danger of confusion in writing 'HF'. 'V_ω' is more usual, admittedly.

Let us recover a direct (non-recursive definition) of HF.

$x \in HF \longleftrightarrow (\forall y)(\mathcal{P}_{\aleph_0}(y) \subseteq y \to x \in y)$ Contrapose to get

$(\forall y)(x \notin y \to \mathcal{P}_{\aleph_0}(y) \not\subseteq y)$

$(\forall y)(x \notin y \to (\exists z)(z \in \mathcal{P}_{\aleph_0}(y) \wedge z \notin y))$ Complementation gives[1]

$(\forall y)(x \in y \to (\exists z)(z \in \mathcal{P}_{\aleph_0}((V \setminus y) \wedge z \in y)))$

$(\forall y)(x \in y \to (\exists z \in y)(|z| \in \mathbb{N} \wedge z \cap y = \emptyset))$

1 Justifying stratified $\in$-induction over V_ω

It's routine to check that HF defined in this way supports $\in$-induction, which is to say – since we are working in a theory with only *stratified* separation – that it supports *stratified* $\in$-induction. Routine it may be, but it's probably worth writing out for the satisfaction of the suspicious.

Suppose $\phi()$ is weakly stratified with one free variable, and that

$$(\forall x)(|x| \in \mathbb{N} \wedge (\forall y \in x)(\phi(y)) \to \phi(x))$$

We want $(\forall x)(x \in HF \to \phi(x))$

Suppose not. Then $A = \{x \in HF : \neg\phi(x)\}$ is nonempty; let a be a member of A. This is where we need a nod in the direction of the restrictions on comprehension. Our comprehension is for stratified expressions only, and this is where we need ϕ to be stratified.

[1] The legitimacy of this complementation move is discussed below!

We have $a \in A$, whence – since $a \in HF$ –

$$(\exists z \in A)(|z| \in \mathbb{N} \wedge z \cap A = \emptyset)$$

But this is impossible, since any such z would have to satisfy $\neg\phi(z)$ while all its members are ϕ.

The $\in$-induction has to be used with care: even if ϕ is stratified the restriction ϕ^{HF} of ϕ to the hereditarily finite sets might not be. However it will be stratified as long as ϕ is stratified and HF (aka V_ω) is a set and therefore can appear as a parameter. The background assumption when we airily talk here about a set theory T will be that T proves (or is augmented with) the existence of V_ω.

In the corresponding arithmetic we have induction – over the corresponding relation "the nth bit in the binary representation of m is set" – for stratified predicates. But what is the notion of stratified predicates of arithmetic? A: the expressions that are stratified are those that are Ackermann translations of stratified expressions of Set Theory.

2 Which Arithmétic Expressions are stratified?

A key (and *old*) observation is that:

Lemma 1
There is no rigid structure on HF definable by a homogeneous expression.

Proof:
The proof of this fact is lost in the mists of time. André Pétry pointed out to me years ago that there is no total order of V_ω definable by a homogeneous expression. This can be proved as follows. (I am not sure that this was his proof!)

Suppose $\phi(x, y)$ were a homogeneous formula that captured a total order of V_ω. Then, for some concrete n, we would have $\phi(x, y) \longleftrightarrow \phi(j^n\sigma(x), j^n\sigma(y))$, which means that the symmetric group on V acts in an order-preserving way on this total order. But every automorphism of a total order is of infinite order, whereas the symmetric group on V is not torsion-free.

■

Whether Pétry proved this himself or was tipped off by Boffa i cannot say.

The more general result – that there is no rigid structure on V_ω definable by a homogeneous expression – is in some sense even easier, since we don't need the fact that Symm(V) isn't torsion-free.

■

A relation which will be of some importance in what follows is the quasiorder $\leq_\rho$ of relative rank: $x \leq_\rho y$ iff the set theoretic rank of x is no greateer than the set-theoretic rank of y. This quasiorder, seen thru' the Ackermann lens, is the binary relation "the largest power of $2 \leq x$ is leq the largest power of $2 \leq y$". Of course this isn't stratified, since we need the order relation $\leq_{\mathbb{N}}$ and that has no stratified definition. That's not to say that there is no stratified definition of $\leq_\rho$ – there could be one for all we know, since $\leq_\rho$ is not rigid so lemma 1 isn't applicable.

However it turns out that there is in fact no stratified definition of the rank quasiorder $\leq_\rho$. This is because we can define $\leq_{\mathbb{N}}$ in a stratified way in terms of $\leq_\rho$, thus: $n <_{\mathbb{N}} m$ iff $n <_\rho m \vee (n \setminus m) \leq_\rho (m \setminus n)$. We know from lemma 1 that $<_{\mathbb{N}}$ is not definable by a stratified expression since it is rigid, so we know that $\leq_\rho$ cannot be so defined either.

The above doesn't mean that $\leq_{\mathbb{N}}$ isn't definable *at all*; it can be defined by recursion on $\in$, but actually we will define it by first recovering the recursive definition of $\leq_\rho$.

$n \leq_\rho m$ iff

$$(\forall Y)\left\{\begin{array}{l}(\forall k)(\langle 0,k\rangle \in Y)\ \wedge \\ (\forall z_1, z_2)((\forall w \in z_1)(\exists u \in z_2)(\langle w,u\rangle \in Y) \to \langle z_1, z_2\rangle \in Y))\end{array}\right. \to \langle n,m\rangle \in Y$$

This involves a universal quantifier over subsets of V_ω, and we want something first-order. We get the next line - line (**) below – by swapping $\mathcal{P}(V_\omega)\setminus Y$ for Y.

Why is this OK? Any Y that is closed under the operation in the lower clause is going to be infinite; The thought is that, if there is a $Y \subseteq \mathcal{P}(V_\omega)$ that excludes a tuple that we don't want, then there is a cofinite $Y \subseteq \mathcal{P}(V_\omega)$ that excludes that same tuple. We cross our fingers behind our backs and hope that if we restrict the quantifier '$\forall Y$' to '$\forall$ **cofinite** Y' we get the same result. That is to say, we rely on the intersection of all (infinite) $Y \subseteq \mathcal{P}(V_\omega)$ such that

$$(\forall k)(\langle 0,k\rangle \in Y) \wedge (\forall z_1, z_2)((\forall w \in z_1)(\exists u \in z_2)(\langle w,u\rangle \in Y) \to \langle z_1, z_2\rangle \in Y)$$

being the same as the intersection of all (cofinite) $Y \subseteq \mathcal{P}(V_\omega)$ such that

$$(\forall k)(\langle 0,k\rangle \in Y) \wedge (\forall z_1, z_2)((\forall w \in z_1)(\exists u \in z_2)(\langle w,u\rangle \in Y) \to \langle z_1, z_2\rangle \in Y).$$

Now we have no right to quantify over infinite collections, even if we restrict ourselves to cofinite collections. However we definitely *are* allowed to quantify over *finite* collections, so we replace a quantifier over a domain of cofinite sets with a quantifier over the domain of their complements. And hope for the best. We will get away with it as long as the binary relation that we end up defining behaves like the comparative rank quasiorder. Does it support stratified induction??

This rewriting gives us

$$(\forall Y)(\left\{\begin{array}{l}(\forall k \leq n)(\langle 0,k\rangle \notin Y)\ \wedge \\ (\forall z_1, z_2)((\forall w \in z_1)(\exists u \in z_2)(\langle w,u\rangle \notin Y) \to \langle z_1, z_2\rangle \notin Y))\end{array}\right. \to \langle n,m\rangle \notin Y)(**)$$

Contrapose

$$(\forall Y)(\langle n,m\rangle \in Y \to \neg\left\{\begin{array}{l}(\forall k \leq n)(\langle 0,k\rangle \notin Y)\ \wedge \\ (\forall z_1, z_2)((\forall w \in z_1)(\exists u \in z_2)(\langle w,u\rangle \notin Y) \to \langle z_1, z_2\rangle) \notin Y\end{array}\right.)$$

It may be worth noting (tho' we will not develop this point here) that the use of contraposition at this point means that this treatment is not constructively secure.

$$(\forall Y)(\langle n,m\rangle \in Y \to \left\{ \begin{array}{l} \neg(\forall k \le n)(\langle 0,k\rangle \notin Y) \ \vee \\ \neg(\forall z_1,z_2)((\forall w \in z_1)(\exists u \in z_2)(\langle w,u\rangle \notin Y) \to \langle z_1,z_2\rangle \notin Y) \end{array} \right.)$$

$$(\forall Y)(\langle n,m\rangle \in Y \to \left\{ \begin{array}{l} (\exists k \le n)(\langle 0,k\rangle \in Y) \vee \\ (\exists \langle z_1,z_2\rangle \in Y)(\forall w \in z_1)(\exists u \in z_2)(\langle w,u\rangle \notin Y)) \end{array} \right.)$$

Once we have $<_{\mathbb{N}}$ we can define `succ` and then we are away laughing.

Another old NF result – it's in [2] and probably in [1] – is that AxCount$_\le$ is equivalent[2] to the assertion that there is a permutation model in which V_ω and the graph of $\le_\rho \restriction V_\omega$ are sets. This has the rather neat consequence that (roughly)

AxCount$_\le$ implies that the Ackermann Arithmetic interprets Peano Arithmetic.

The story reads as follows. Work in a theory with AxCount$_\le$; move to the permutation model in which V_ω and the graph of $\le_\rho \restriction V_\omega$ are sets. But if $\le_\rho \restriction V_\omega$ is a set then it is allowed to appear as a parameter, and as a result we have stratified definitions of $<_{\mathbb{N}}$ and `succ`, and then we have access to the routine definitions of recursive functions and everything that follows from that.

However this result may not be as contentful as its arresting sound promises. The point is that the permutation which gives a model of "NF + V_ω and $\le_\rho \restriction V_\omega$ exist" from a model of NF + AxCount$_\le$ uses the Ackermann interpretation, so it's hardly surprising that if we code up the naturals as sets and then decode them we get to where we started – with the arithmetic of NF + AxCount$_\le$.

But some other bits of arithmetic are stratified in a more direct way. $n = 2^m$ is definable in a stratified way: "m is the unique member of n". For a while i tho'rt that "n is of the form $2^m - 1$", too, is definable in a stratified way beco's it's "n contains all subsets of any of its members". But that's not quite right. However, even if we have a stratified way of saying that a set is a power of two, or is an initial segment of $\mathbb{N}$, there is no stratified way of connecting the two. Suppose i could recover $2^n - 1$ from n. Then, from n and m i would be able to recover $2^m - 1$ and $2^n - 1$ and i can certainly capture $\subseteq$ in a stratified way, so we would be able to define $\le_{\mathbb{N}}$ – and we know we can't.

The multiplicative structure of $\mathbb{N}$ is not rigid, so the "no rigid structures" constraint of lemma 1 doesn't exclude the possibility that "x divides y" might be captured by a stratified expression.

There are miscellaneous arithmetical factoids that might come in useful here. For example "$\binom{n}{k}$ *is odd iff every bit set in k is set in n*." (Think of the 7th – or the 15th – row of Pascal's triangle, for example, to see this) It means that "$\binom{n}{k}$ is odd" is stratified – indeed homogeneous. (There is also this amazing

[2] AxCount$_\le$ is the assertion that $n \le Tn$ for all natural numbers n.

fact that if p^c is the largest power of p that divides $\binom{a+b}{a}$ then c is the number of carries performed in the addition of a and b to base p. Is that any use?) All the fun with `nim` addition is stratified, because `nim` addition is `XOR`.

Not sure about `nim` multiplication. Depends on whether we can define multiplication-by-2 (which of course is `rotate-left`) in a stratified way ... and that doesn't look terribly likely. However there may yet be a way in. The manual says that if $x < y = 2^{2^n}$ then x `nim-mult` y is plain old $x \cdot y$. Now altho' we don't seem to be able to give a cute definition of $x \cdot y$ in general, we can if y is a power of 2 much bigger than x. Think of x as $\sum_i 2^{x_i}$; if y is a power of 2 much bigger than x then y is 2^{y_0} with $y_0 > x_i$ for all i. Then $y \cdot x$ is $\sum_i 2^{y_0+x_i}$ which corresponds to $\{x' \cup \{y_0\} : x' \in x\}$ and that is a set by stratified comprehension.

However y `nim-mult` y (for y of the form 2^{2^n}) might be an obstacle, since it is $3y/2$, and we don't know how to do multiplication. Now $3y/2$ is $y + y/2$, otherwise known as $2^{2^n} + 2^{2^n-1}$. So this is a pair, one member of which is $\bigcup y$ and the other is ... somehow to be obtained from y. But we know that we cannot recover $2^n - 1$ from 2^n in any stratified way. Bummer.

Does this arithmetic enable us to define primes in a stratified way? Does it perhaps prove Unique Factorisation...?? It certainly seems to obey *exp* beco's HF is closed under singletons.

3 A Remark about IO

There was a time when i felt that IO ought to mean something in the Ackermann Arithmetic. It looked as if it should mean something like *exp* but we will show now that you get it free so it doesn't do anything special.

I am now reasonably confident of the following:

THEOREM 1 *KF proves that if V_ω is a set then $V_\omega \models IO$.*

Proof:

We use $\in$-induction. Suppose x is a finite subset of V_ω each of whose members is the same size as a set of singletons in V_ω. We will show that x, too, is the same size as a set of singletons in HF.

Wellorder x somehow and perform the following recursion. Associate to each $y \in x$ a set of singletons in V_ω the same size as y *and have not been used for any earlier* $y' \in x$. Pick one and call it s_y. (This doesn't need AC beco's we (*)
are making only finitely many choices[3].) Send $y \mapsto \{\bigcup s_y\}$. This function is injective (and type-level so) its graph is a set. It makes x the same size as a set of singletons $\subseteq V_\omega$.

This establishes the induction step. "x is the same size as a set of singletons in V_ω" is stratified (since V_ω is a parameter) so we infer the conclusion by stratified $\in$-induction.

■

[3]This is crucial: if we needed AC we'd be buggered, beco's by lemma 1 there is no total order – let alone a wellorder – of HF definable by a stratified expression.

How do we know that there will always be a set of singletons the same size as y satisfying the italicised condition at (*)? Answer: because if there is even one set of singletons of members of HF the same size as y (which is what the induction hypothesis gives us) then there will be infinitely many (I'm hoping this is too obvious to need a proof) and only finitely many of them can have been used up by earlier members of x.

Where have we used the assumption that HF is a set? We don't need it to prove the induction step; we need it when it comes to ensuring that the property we are proving by induction (namely "x is the same size as a set of singletons in HF") is stratified, and we need that to be stratified since we have $\in$-induction over HF for (weakly) stratified expressions only; it would not be weakly stratified if we had to write out "$x \in HF$" in primitive notation. Since HF exists it is available as a parameter.

Observe that the ruse we used for avoiding AC in proving that $HF \models IO$ relies on everything in HF being finite. If we try to construct an analogous proof that the world of wellfounded sets satisfies IO we will need some extra assumptions.

There was a time when i felt that IO ought to mean something in the Ackermann Arithmetic. It looked as if it should mean something like *exp* but the above result shows that it doesn't ... since IO holds anyway, and we get *exp* in the Ackermann Arithmetic even if we don't start with IO.

4 Envoi

The question remains: what is the Ackermann arithmetic of any of the systems NF, NFU, KF plus "HF exists"? Central to answering the question is the endeavour to ascertain which arithmetical properties can be simply defined in terms of "the nth bit of m is set".

My guess would be that the theories NF + "HF exists", NFU + "HF exists" and KF + "HF exists" all have the same Ackermann Arithmetic; my hope would be that this theory should turn out to be some natural fragment of PA long known to the *cognoscenti*.

5 Work still to be done

I think this may be the time and the place to sort out quite what one means by saying that "Arithmetic is stratified". It might help with this task to keep in mind Paris-Harrington and Finite Ramsey, and find a language in which we can express the two of them and then explain both how Paris-Harrington is *prima facie* not stratified and also how it nevertheless turns out to be stratified after all.

Sex up $\mathcal{L}(PA)$ with a new unary function symbol 'T' for an automorphism. Let's call this new language $\mathcal{L}(PA)^T$. Naturally we add axioms to say that T is an automorphism (tho' it is presumably suff to say merely that it's an automorphism of $<_{\mathbb{N}}$). We can then further enhance the axioms in two ways:

(i) Add all instances of induction for the new language; or

(ii) Add all stratified instances of induction for the new language.

Presumably (ii) is a conservative extension of PA. In fact, presumably (i), too, is a conservative extension of PA, since T could be the identity. NF interprets (ii); NFC (= NF + AxCount) interprets (i). Are either of these theories interpretable in PA?

We need to think about interpretations of $\mathcal{L}(PA)^T$ into $\mathcal{L}(PA)$. There is a fairly easy interpretation of the stratified fragment of $\mathcal{L}(PA)^T$ into $\mathcal{L}(PA)$.

But actually perhaps we want Weak-second-order Arithmetic. We can code finite sets of naturals using Ackermann so we can get a weak higher-order arithmetic, the theory of sets hereditarily finite over $\mathbb{N}$. Its language is $(PA) \cup \{\in\}$. It's one-sorted but it admits a stratification algorithm. I'm not entirely sure what set existence scheme it hath ... presumably the Kaye-Wong theory ... ?[4] Clearly we can interpret it − not just the stratified part − back into PA. We now add a function symbol: '$| - |$', for cardinal-of. We might want to add it as a primitive and adopt axioms for it. However we can also give a recursive definition of it. And i think this is the point at which to add T - beco's once we've got $\in$ and $| - |$ we can define T!! T of course raises by one type. It doesn't seem to matter by how many types '$| - |$' raises as long as it's always the same. I have the feeling that we might want to tighten up the stratification algorithm so that we accept a formula as stratified only if it can be stratified *whatever* type-difference is imposed by '$| - |$' ... but that may be too strict, beco's i think we want to be able to make "n is relatively large" stratified by putting 'T's in the correct places.

There may be some connection here with Pabion's work.

Paris-Harrington is straightforwardly expressible in the language of the theory of sets hereditarily finite over $\mathbb{N}$, since it does not have quantifiers over arbitrary finite sets. Finite Ramsey does have such a quantifier, but it can be eliminated: we simply instantiate the variable 'x' in "(For all sets x of size k) $\phi(x)$" to '$[1, k]$'. And the language of sets-hereditarily-finite-over-$\mathbb{N}$ really is the right language to use, not second-order arithmetic. All our sets are finite, but they can be sets of sets of sets (a colouring is a set of k-tuples) and Ƅ"$[X]_n$ − which we need in the statement of Paris-Harrington − is further up still.

Euler's totient function is inhomogeneous but it is primitive recursive:

The relation "n divides m" is primitive recursive, since it is $(\exists k < n)(k \cdot m = n)$. So "$n$ is prime" is a primitive recursive predicate. "m and n are coprime" is a primitive recursive predicate, being the negation of the primitive recursive predicate "$(\exists k < m)(k|m \wedge k|n \wedge S(0) < k)$". We can now obtain $\phi(n)$ by bounded summation thus

$$\sum_{m<n} \texttt{if } m \text{ and } n \text{ are coprime } \texttt{then } 1 \texttt{ else } 0.$$

This is a homogeneous definition. How do we prove that it obeys the inhomogeneous equation? By an (unstratified?) induction. . . ? I think this reduces to

[4]It's easy to verify TCo co's $2^n - 1$ is always transitive!

the problem of showing, for any (primitive recursive, to keep it simple) predicate ϕ that

$$\sum_{m<n} (\mathtt{if}\ \phi(m)\ \mathtt{then}\ 1\ \mathtt{else}\ 0) = |\{m < n : \phi(m)\}|$$

Or – simpler still –

$$\sum_{m<n} (1) = n$$

But it's obvious that the sum of all the naturals in a multiset is the same type as the naturals in the multiset (+ is homogeneous) and so of lower type than the cardinality of the multiset.

Hmmm ... worth noting that when $f : V \to \mathbb{N}$,

$$\sum_{x \in X} (f(x)) \neq \sum f``X$$

unless $f \restriction X$ is injective, but i think that is a sideshow. In contrast $\bigcup_{i \in I} X_i$ is the same as $\bigcup\{X_i : i \in I\}$. This may be no more than the fact that $\cup$ is idempotent but + is not.

One can also consider the Ackermann Arithmetic of more mainstream (unstratified) set theories. The Ackermann bijection directs our attention to what one might call *Weak Higher-Order* Arithmetic, which is the study of the sets hereditarily-finite-over-$\mathbb{N}$. This theory is the natural context for Paris-Harrington and Finite Ramsey (and the Euler totient function, and other functions that give the number of wombats of size n, e.g sequence A000041 in the OEIS, that sort of thing). Since it is a set theory one can easily deploy the idea of stratification in it. That would provide the other half of a paper of which the foregoing is the first half.

Well actually it's not that simple. PH and the totient function both make use of the function `cardinal-of` which (strictly) does not belong to weak higher-order arithmetic. However we can define cardinal-of in a homogeneous way by

$|n| = \sum_{m<n}(\mathtt{if}\ m\,E\,n\ \mathtt{then}\ 1\ \mathtt{else}\ 0)$

Using n as the bound for the summation in this way works beco's m E n $\to$ $m < n$. And '$m\ E\ n$' is primrec beco's it is '$\binom{n}{2^m}$ is odd'.

$\binom{|n|=\sum_{m<n}(n}{2^m}\ \mathtt{mod}\ 2)$

All primitive recursive functions are homogeneous. It's easy: we prove it by induction on the recursive datatype of primitive recursive functions. So – in particular – $n \mapsto 2^n$ is homogeneous!!

This might be worth dredging up for counting.tex. A type-level cardinality function!

So i think this is what is going on. Consider the weak higher-order theory of arithmetic, the theory of the sets hereditarily finite over $\mathbb{N}$. The Ackermann bijection gives us an interpretation of this theory into PA. (They're presumably synonymous). Key banality: every primitive recursive function is homogeneous.

And we can give a primitive recursive formulation of "m is the cardinal of n" where n is thought of an a finite set (so E is the relation between naturals corresponding to $\in$ in the ASckermann interpretation) by

$|n| = \sum_{m<n}(\texttt{if } m\,E\,n \texttt{ then } 1 \texttt{ else } 0)$

Using n as the bound for the summation in this way works beco's m E n $\rightarrow$ $m < n$. And 'm E n' is primrec beco's it is '$\binom{n}{2^m}$ is odd'.

$\binom{|n|=\sum_{m<n}(n}{2^m}\ \texttt{mod}\ 2)$

So here's what i think is going on. $w\mathrm{PA}^\infty$ is Weak Higher-Order arithmetic, the theory of sets hereditarily finite over $\mathbb{N}$. It's a natural environment for much of number theory and combinatorics (We can define Euler's totient function and express P-H once we have enhanced it witha cardinal-of function). Much of what is usually considered Number Theory goes on in this environment. It is a language with a natural stratification discipline, since $\in$ is inhomogeneous. Now PA is in a language whose primitive relations are all homogeneous. Every primrec fn is homogeneous. Now the Ackermann interpretation interprets $w\mathrm{PA}^\infty$ into PA. $\in$ and "cardinal of" both get interpreted by primitive recursive relations/functions in PA.

This (i think!) is the sense in which Arithmetic Is Stratified....

NFistes always say that (the) Arithmetic (of $\mathbb{N}$) is stratified. What exactly might one mean by this, and is it true?? On the face of it the answer is an easy 'yes' since the langauge of arithmetic does not contain any lexical item such as 'cardinal of' or '$\in$' that could give rise to type differences. However this answer is less than satisfactory, since a lot of fairly mainstream Number Theory (one thinks of Euler's totient function, the theorems about the number of ways of partitioning a natural number...) do involve such expressions. To accommodate such material one needs a somewhat wider theory; one that will do is the weak higher-order theory of $\mathbb{N}$, the theory of the sets hereditarily finite over $\mathbb{N}$, at least if we enhance it with a `cardinal-of` function. This language has the inhomogeneous relations $\in$ and `cardinal-of` and so definitely contains unstratified expressions – Paris-Harrington is one example. So the question is rather "Is all the number-theoretic stuff that one wants to use this language to express going to be stratified?"

This is where the Ackermann interpretation comes in. It enables us to interpret the weak higher-order theory of $\mathbb{N}$ into Peano Arithmetic. This interpretation sends the inhomogeneous relation $x \in y$ to the homogeneous relation "the xth bit of y is 1". We can also send "y is the cardinal number of the set x" (where the type difference between the two variables is going to depend on how we implement 'cardinal-of', and cannot be relied upon to be 0) to

$$y = \sum_{m<x}(\texttt{if } m\,E\,x \texttt{ then } 1 \texttt{ else } 0)$$

Using x as the bound for the summation in this way works beco's $m\,E\,x$ implies $m < x$, so all the m we are trying to catch can be found below x. And 'm E n' is primrec beco's it is '$\binom{n}{2^m}$ is odd'. So '$y = |x|$' becomes

$$\binom{y = \sum_{m<x}(x}{2^m\ \texttt{mod}\ 2)}$$

which is homogeneous (the two free variables are of the same type).

Thus in the Ackermann interpretation all the potentially troubling relations that might give rise to type differences get sent to homogeneous expressions. So every formula in weak higher-order arithmetic, stratified or otherwise, gets interpreted by a stratified formula of PA.

I suppose one question is this. Work in primitive recursive arithmetic (something that everyone understands!) If n is a natural number thought of as a set à la Ackermann, then the power set of n is the number

$\sum_{m<n}$(if $\binom{n}{m}$ is odd then 2^m else 0)

and the challenge is to show that the size (defined earlier) of this number is in fact $2^{|n|}$

I am guessing that we can prove this. Does it matter if it isn't? What's bothering me is that this seems to ride roughshod over stratification considerations.

References

[1] T. E. Forster "Permutations and Wellfoundedness: the True Meaning of the Bizarre Arithmetic of Quine's NF". *Journal of Symbolic Logic* **71** (2006) pp. 227–240.

[2] T. E. Forster "Set Theory with a Universal Set, exploring an untyped Universe" Oxford Logic Guides **20** OUP 1992