A critique of The Vastness of Natural Languages by Langendoen and Postal

Hart, K.P.

Publication date
2021

Document Version
Final published version

Citation (APA)
Hart, K. P. (2021). *A critique of The Vastness of Natural Languages by Langendoen and Postal*.
https://ling.auf.net/lingbuzz/006052

A CRITIQUE OF 'ON THE VASTNESS OF NATURAL LANGUAGES' BY LANGENDOEN AND POSTAL

KLAAS PIETER HART

To the memory of Ken Kunen, who taught me Set Theory

Abstract. This paper looks at an argument in *On the Vastness of Natural Languages* by D. T. Langendoen and P. M. Postal.

The conclusion is that it does not pass mathematical muster.

Introduction

In the autumn of 2019 I got into a discussion, on Twitter, on the nature of books. In [8] Marc van Oostendorp wrote about [9] by Paul M. Postal wherein books are likened to numbers because they are 'merely' sequences of symbols. Given that we generally make no qualms about writing things like "let L be the set of sequences of symbols from the alphabet A (plus spaces, interpunction, etc) of length n" we may have inadvertently created Borges' *Library of Babel* and hence all books — past, present, and future — of 410 pages. The question then becomes whether a book-as-a-sequence is invented by the author or merely discovered among all those sequences. That is what the discussion on Twitter was about. But it is not what this paper is about.

In Postal's article I found a sentence that piqued my mathematical interest:

> Then, appealing to the reasoning of Langendoen and Postal [6], one can show further that the universe of books is truly vast, amounting to what is called a *proper class* in some varieties of set theory.

This I needed to know more about and I tried to get hold of the book, *The Vastness of Natural Languages*, cited in this quote. This proved harder than I expected but some searching led me to many reviews of it and one of these reviews pointed out that the paper [7] (*Sets and Sentences*) by the same authors contains the main mathematical arguments of the book. It is freely available from Langendoen's website and a footnote on its first page says it is adapted from the book, so I decided to make do with the paper and see what was going on.

What I found surprised me and I decided to write a note about my experiences as a cautionary tale, because it shows what can happen when non-mathematicians try to apply some non-trivial mathematics without some of the rigor that we are used to.

I wish I could have talked to Ken about this; I'd like to think he would have enjoyed it and would have offered some insightful comments that would have improved this note considerably.

Date: Friday 25-06-2021 at 15:05:46 (cest).

2020 *Mathematics Subject Classification.* Primary 03E75; Secondary 03E10 03E20.

Key words and phrases. natural language, set theory.

Overview. While I was working on this note I discovered a digital copy of *The Vastness of Natural Languages* on `archive.org`; it turned out that Section 2 of [7], which is the mathematical core of the paper covers roughly Chapter 4 of the book. The third section of the paper corresponds to Chapter 5 and uses the mathematical results to argue that many theories about natural languages are invalid.

We shall begin in section 1 by looking at the mathematical goings on in Section 2 of the paper, indicating the minor differences with what is in the book. In section 2 we look at some other parts of the book and how Set Theory is treated there.

Finally, section 3 contains a summary and some set-theoretical remarks that shed further light on the mathematical deficiencies of the book and the paper.

1. The analogy with Cantor's results

Section 2 of [7], from which we took the title for the present section, starts with the announcement that we shall see a 'strict parallelism' between the sentences of a natural language and the collection of all sets.

The results of Cantor alluded to in the title stem from the diagonal argument in [2]: the set of all sets, the set of all cardinal numbers, etc., do not exist. The paper aims to show that the collection of sentences of a natural language is a similar 'megacollection' (the authors' term for what in Set Theory would be called a proper class).

1.1. **Sets.** The section begins with a (parenthetical) definition of a set: "a collection with fixed magnitude, finite or transfinite".

It is a common misconception that sets come with an inherent magnitude, or cardinality, or potency, or number of elements, or ... They do not. The misconception probably has its origins in our youth where we learn to count and associate natural numbers with sets of objects. Also, in elementary Set Theory we freely compare cardinalities and write stuff like $|X| \leqslant |Y|$ and $|X| = |Y|$ without reminding the reader that $|X|$ in isolation has no meaning. And of course, the use of the symbols $\aleph_0$ and $\mathfrak{c}$ only reinforces the idea that 'magnitude' or 'cardinality' is a thing.

If only more people would follow Ken Kunen's example, and write $X \preccurlyeq Y$ and $X \approx Y$ rather than $|X| \leqslant |Y|$ and $|X| = |Y|$ to mean "there is an injection from X to Y" and "there is a bijection from X to Y", respectively; *and* only use $|X|$ once it has been properly defined, see [5, page 27], and also [1, 5.43 and 4.2].

1.2. **The workings of languages.** The next two pages describe a particular way in which sentences can be formed out of other pieces of languages that can be sentences themselves.

The basic idea is that a sentence is formed by taking two or more units, join each unit with a connective, and laying out the results of these joinings in some order. See Figure 1 for an example: in this both C_1 and C_2 are called *conjuncts*,

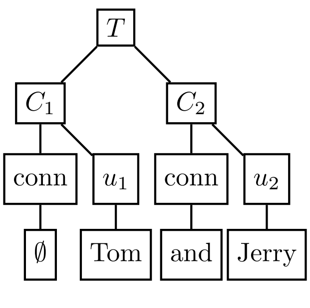

Figure 1. A simple sentence

where C_1 is made up of the empty connective $\emptyset$ and the *subconjunct* 'Tom', and C_2 is made up of the connective 'and' and the subconjuct 'Jerry'. The empty connective is used to make a common construction as in the figure fall into this general scheme. The result, T, is called a *co-ordinate compound constituent*, later shortened to 'co-ordinate compound'.

1.3. **Compounding and projecting.** Next we get a general description of how co-ordinate compounds and sets of constituents interact.

We adopt some category Q of constituents. Assume we have a set U of constituents with at least two elements and a co-ordinate compound T, all of category Q. Then T is a *co-ordinate projection* of U and U is the *projection set* of T if

(1) each conjunct of T has an element of U as a subconjunct,
(2) each element of U is a subconjuct of a conjunct of T
(3) no element of U appears more than once as a subconjunct of any conjunct of T
(4) if two elements of U occur as subconjuncts of conjuncts C_i and C_j of T, then C_i and C_j occur in a fixed order. Where C_i and C_j are of distinct lengths, assume the shorter precedes; where C_i and C_j are the same length, assume some arbitrary order. Really? How?

Thus {Tom, Jerry} is the projection set of T in Figure 1, and T is a co-ordinate projection of that set, but so is "Jerry and Tom".

The authors go on to argue from (4) that we can say that T is *the* co-ordinate projection of U, because: "it insures that different orders of conjuncts are irrelevant".

I would disagree. First: (4) is badly written as it is. Conjuncts should occur in a fixed order and that 'fixed order' amounts to 'some arbitrary order' if they have the same length. That would warrant a "Really? How?" written in red pen in the margin. Either you define that 'fixed order' properly or you leave the set unordered; "just do something" is not an option.

And second: no matter how you look at it {Tom, Jerry} has two co-ordinate projection; that "Tom and Jerry" is ingrained in our minds does not make "Jerry and Tom" any less valid.

To make even more certain that the co-ordinate projections are unique the authors simplify the language to use just one connective (or rather *co-ordinating particle*). Thus we'd get "and Tom and Jerry" and "and Jerry and Tom" as possible projections. This would make the projection unique *given an ordering of the projection set* because the choice of connectives was the only variability left.

1.4. **Existence.** Next comes an interesting bit of mathematical sleight of hand when the authors set themselves the task to show that every set of constituents has a co-ordinate projection (back to the indefinite article?). A task which they deem to be straightforward.

They take a set U of constituents and denote its cardinality by k (finite or infinite). I'll list the steps in the argument and then comment on them.

(1) "Clearly, from the purely formal point of view, there is a co-ordinate compound W belonging to the category Q."
(2) "Since there are no size restrictions on co-ordinate compounds, W can have any number, finite (more than one) or transfinite of immediate constituents."
(3) "W can then, in particular have exactly k such constituents."

(4) "To show that W is a co-ordinate projection of U, it then in effect suffices that there exist a one-to-one mapping from U to the set V of constituents of W.
(5) "But this is trivial, since the two sets have the same number of elements."

Step (1). This is a non-argument; "a purely formal point of view" carries no weight at all.

Granted, we are talking about languages so we should expect there to be some (compound) sentences but not because of some "purely formal point of view". There are generally two options: actually construct an example of a W that belongs to category Q. The second is to explicitly state this as an axiom, much like in Set Theory, where $(\exists x)(x = x)$ is sometimes treated as a logical axiom and sometimes as a proper axiom of ZFC, for emphasis [5, Axiom 0].

Step (2). I should mention that this uses a basic assumption of the paper and also of the book: there is no restriction on the possible size of sentences. We shall come back to this later when we discuss some portions of the book.

Nevertheless this is again a non-argument. The statement is not a simple consequence of the first claim, it goes way beyond simply asserting that there is a co-ordinate compound. It suddenly states that there are very many such things, without any justification. Let me be clear: the fact that one does not impose size restrictions does not entail that suddenly all kind of arbitrarily large examples spring into being.

To illustrate: Axioms 0, 1 and 3 of [5] impose no size restrictions on sets but they hold in the universe with just one member, $\{\emptyset\}$, and hence there is no proof from these axioms that there is a set with at least one element.

Also there is the matter of mathematical style: in the first claim W stands for just one object; in the second claim W seems to stand for many objects simultaneously.

This step and the next also sow the seeds of circular reasoning in the proof of the main result of the paper as we shall see later.

Step (3). With no argument at all the arbitrary W has been transformed into one that suits the argument. The points in Subsection 1.3 then do indeed imply that "The subconjuncts of W form a set V of cardinality exactly k" but a good argument this is not.

Step (4). This is where the argument truly goes off the rails. The co-ordinate compound W started out arbitrary, unrelated to U, then its projection set gained the correct cardinality, that is, that of U, and now has become a co-ordinate projection of U even. However the 'proof' of that last assertion is simply false: a bijection does not make sets equal.

Indeed, this argument establishes, in particular, that "and Tom and Jerry" is a co-ordinate projection of the set {Laurel, Hardy} because there is a bijection between that set and {Tom, Jerry}. This seems, to borrow a word used by the authors a lot, absurd.

Step (5). This is a circular conclusion. Already in Step (3) the set V is asserted to have cardinality k. That is *equivalent* to there being a bijection between U and V. So this step boils down to "there is a bijection because there is a bijection".

1.5. **Closure Principles.** Having established[1] that "each subset U of constituents of category Q has a co-ordinate projection" the authors turn to closure under compounding. After noting that the co-ordinate projection may not necessarily

[1] It should be clear that I do not agree, but let us continue reading anyway.

well-formed in the natural language under consideration they restrict their attention to cases where the projections are well-formed and rightfully use the word 'axiom'.

The Closure Principle for Co-ordinate Compounding. If U is a set of constituents each belonging to the collection S_w of (well-formed) constituents of category Q of any natural language, then S_w contains the co-ordinate projection of U. □

The authors admit that it is not clear to what categories this principle applies but they feel confident in adopting it for the category S of sentences of a natural language:

Closure under Co-ordinate Compounding of Sentences. (a) If U is a set of constituents each belonging to the collection, S_w, of (well-formed) constituents of the category S of any natural language, then the co-ordinate projection of U belongs to S_w.

More precisely, (a) can be stated as in (b):

(b) Let L be the collection of all members of the category S of a natural language and let $\mathrm{CP}(U)$ be the co-ordinate projection of the set of sentences U. Then

$$(\forall U)(U \subset L \longrightarrow \mathrm{CP}(U) \in L)$$ □

This subsection ends with a justification of this principle by means of some examples of (very finite) sets of English sentences and well-formed co-ordinate projections thereof.

There are two things about this justification that are worth remarking. Again there is the switching between indefinite and definite articles: "any set has a well-formed co-ordinate compound" and in the examples we see "the co-ordinate projection of each set". The second is that the examples show two connectives rather than the unique one that was postulated a few pages earlier: there is both the comma and the connective 'and'.

1.6. **A hierarchy.** The next thing in the paper is the construction of a hierarchy of sets of sentences in a natural language that contains a countably infinite set S_0 of non-compound sentences.

For example, as a variation on the authors' theme, we can generate an infinite sequence of English sentences by recursion, as follows:

- s_0: The real line is uncountable
- s_1: I know that the real line is uncountable
- s_2: I know that I know that the real line is uncountable
- $\cdots$
- s_{n+1}: I know that s_n
- $\cdots$

If we were talking about the English language then we could let the set $\{s_n : n \in \omega\}$ be our starting point S_0.

The construction of the hierarchy $\langle S_n : n \in \omega\rangle$ of sets of sentences then proceeds recursively, as follows

$$S_{n+1} = S_n \cup K_n$$

where

$$K_n = \{x : (\exists y)(y \subseteq S_n \land x \text{ is the co-ordinate projection of } y\}$$

In the paper it takes over a page to get to the step from S_0 to S_1.

First we get a description in words, where much is made of the point that co-ordinate projections/compounds are only defined for sets of more than one constituent (L is the natural language under consideration):

> "L also contains a set S_1 made up of all the elements of S_0 together with all and only the co-ordinate projections of every subset of S_0 with at least two elements"

This is illustrated with an S_1 — "S_1 can be taken" — based on our example S_0; the members of S_1 that are exhibited are all finite, though it is suggested that S_1 contains many infinite sequences of members of S_0 as well. The finite compounds seem indeed sorted by length but there is no clear indication of how the infinite compounds are ordered — ellipses can only suggest so much.

Then we get a very verbose determination of the cardinality of S_1, which boils down to saying that $y \mapsto \mathrm{CP}(y)$ is a bijection between $\{y \in \mathcal{P}(S_0) : |y| \geqslant 2\}$ and the then still unnamed set K_0. The result of this argument is surprising: that cardinality "is of the order of the continuum, that is $\aleph_1$, the cardinality of S_1 is $\aleph_1$". One would expect to see $2^{\aleph_0}$ here but the book has a footnote at this point: "Our notation assumes purely for convenience, the 'generalized continuum hypothesis', according to which $2^{\aleph_\alpha} = \aleph_{\alpha+1}$". One could, for notational convenience, also have used $\beth_n$ to denote the cardinality of S_n.

The intermediate conclusion is that the natural language L contains at least as many sentences as there are in S_1, so at least $\aleph_1$.

The step from S_1 to S_2 is spelled out as well, leading to a set of cardinality $\aleph_2$.

At this point the recursion is made explicit, the cardinalities of the sets S_n are not given.

It is not clear whether the authors think that the sequence $\langle S_n : n \in \omega\rangle$ is the end of the story, so to speak. The sentence

> "At no point can a set of sentences be obtained that exhausts a natural language having sentence co-ordination governed by the closure law."

seems to point to this. On the other hand they offer a separate proof that "no set can exhaust a natural language", and this does not use the sequence $\langle S_n : n \in \omega\rangle$.

A comment. Since the whole paper is about languages one would expect to see some rules that govern the co-ordinate compounding. The "assume some arbitrary order" in the definition allows for quite lax rules it seems. If one looks at the hierarchy some problems do crop up.

The set S_0 is countable and naturally ordered by length. The elements of S_1 of infinite length all have the same length and can be ordered using the lexicographic order. However it is consistent to assume that S_2, which is essentially the power set of the real line, has no linear order, see [3, page 142]. Of course this happens in a model where the Axiom of Choice fails but this implies that one needs this very non-constructive principle when making co-ordinate projections. We will have more to say about this in section 3.

Another comment. As noted above the authors do not quite make clear whether the sequence $\langle S_n : n \in \omega\rangle$ builds up (or exhausts) the whole language. Readers of Ken's book will know that assuming that the sequence does indeed do so is no big restriction. The mathematical community is still busy working Exercise IV.9 in [5]: verify that in ZC (and hence in $V_{\omega+\omega}$) one can develop 99% of modern mathematics.

1.7. **The main result.** Without further ado we get the main theorem.

The Natural Language Vastness Theorem. Natural languages are not sets (are megacollections).

Proof. Let L be the collection of all sentences of some natural language and assume L is a set.

Then L has a fixed cardinality, $\#L$.

By the closure assumption L contains the set Z of all co-ordinate projections of all subsets of L that have at least two elements.

The set Z is a proper subset of L as there are (many?) sentences that are not co-ordinate compounds. (This remark is superfluous. It is not needed to justify the coming 'Hence'.)

Hence $\#Z \leqslant \#L$. (This would have been true also if $Z = L$.)

But $\#Z$ is of the order of the power set of L. (In short $\#Z = 2^{\#L}$.)

Hence, by Cantor's Theorem, $\#Z > \#L$.

Contradiction, so the assumption that L is a set is untenable. □

1.8. **The argument does not constitute a proof.** Let us look back at the argument for the main theorem; its main steps are

(1) Every set has a co-ordinate projection
(2) The assumption/axiom that language is closed under co-ordinate compounding *and* a set leads to a contradiction with Cantor's theorem that $X \prec \mathcal{P}(X)$ for all sets X.

The second step is not problematic; it is indeed parallel to the argument that yields a contradiction when one applies Cantor's theorem to a purported 'set of all sets'.

It is the first step that is truly problematic. Its proof is full of holes.

The biggest hole can be found in steps (2) and (3) in Subsection 1.4. From the absence of assumptions that limit the cardinality the authors deduce that there are indeed no limits: there are compounds of arbitrarily large cardinality. However that deduction seems to be based on nothing other than the apparent idea that dropping an assumption means adopting its negation. That is not how mathematics works; dropping an assumption is just that: you don't let yourself apply it anymore.

In fact the hole is so big that it invalidates the proof. To repeat, step (2) says "Since there are no size restrictions on co-ordinate compounds, W can have any number, finite (more than one) or transfinite of immediate constituents." This means in effect that there are co-ordinate compounds of every possible cardinality. But this already means that the co-ordinate compounds do not constitute a set. If it were a set then an application of the Axiom Scheme of Replacement, [5, Axiom 6], shows that the cardinalities form a set, which they do not by Cantor's arguments.

Whether one uses Replacement or follows the authors' longer route the result is not a valid proof of the main result. This is because what Step (2) says, is in effect

> The collection of co-ordinate compounds is at least as large as the collection of cardinalities.

This already asserts that the co-ordinate compounds form a 'megacollection'.

Thus the second section of the paper can then be summarized as establishing the following tautology:

> If we assume that a natural language forms a megacollection then we can prove that it forms a megacollection.

2. About the book

The previous section is based largely on the reading of the article *Sets and Sentences* [7], which, as noted before, corresponds roughly to parts of Chapters 4 and 5 of [6]. There are a few things to be said about the rest of the book and that will occupy us in this section.

2.1. **Set-Theoretical background.** The first chapter intends to give the necessary set-theoretical background. The approach is largely the 'naïve' one we see in first-year courses on mathematical methods; definitions are by synonym, for example the very first sentence opens: "Collections are ensembles . . . ". Other words in use are 'class', 'aggregate', and of course 'set'. The authors opt to use 'collection' throughout, because in some older writings 'class' and 'set' mean the same thing.

Rather than present the tools of the trade the authors concentrate on the problems that arose when Cantor's theorem, that $X \prec \mathcal{P}(X)$ for all sets, was applied to the 'set of all sets' and its proof was used by Russell in [10] to formulate his paradox.

It looks a bit like cherry-picking; there is not much more to be found than what is used in Chapter 4 and [7]: the distinction between sets and proper classes (sets and megacollections) and how this came about. There is a curious quote from [4], which I could not verify alas

> the totality of all classes is partitioned into two subtypes, sets, and the collections with the size of the collection of all sets ('the universal class')

Since the ordinals do not form a set we have here as a corollary that there is a global well-order of V.

The references in the chapter are all over the place; the authors rely on lots of sources for remarks about the nature of sets, classes, magnitudes cardinality, etc. One is, for example, directed to different books for different axioms, and, the cherry-picking feeling remains: we don't get to see a full set of axioms. The authors state

> We will adopt, although entirely informally, the general viewpoint of the second approach to set theory, one distinguishing class and set.

That second approach is that of Gödel-Bernays-von Neumann set theory.

I would have like to have seen a more concrete foundation. Preferably based on a single source; Ken's book, [5], would have been perfect. It discusses the pitfalls of the overly naïve approach, has all the axioms in one place, and shows in the first chapter how set theory works.

2.2. **Chapters 2 and 3.** I will treat these chapters in one go because their common theme is to keep convincing the reader that bounds on the sizes of natural languages are absurd (a word used quite often) and violate Occam's razor. The latter gets invoked a lot and I do not want to go into a discussion whether the authors apply it correctly (there are many web-pages that will tell you what the 'correct' interpretation is) but often it seems like it is the only argument.

However, as I mentioned before, the authors do not so much argue for "not assuming a size law" but for "assuming the negation of a size law". For example, the rules (if any) of English do not stipulate a maximum finite length of sentences; one can easily break such a stipulation by prefixing a maximum length sentence with "I know that". The rules of English also do not explicitly state that sentences should be finite; one can add "All English sentence should be finite in length" to the rules or not.

The authors argue, quite vociferously at times, against adding that condition mostly on the grounds that it is not a purely linguistic one. However, and this is where I disagree, they then conclude that, somehow, necessarily there should be sentences of infinite length.

Some of the arguments for this are quite advanced: one invokes a compactness argument (not by name but it is there) to go from "arbitrarily long and finite" to "infinitely long". It treats sentences as models of the theory that is a natural language, without actually verifying that this is possible. Another is quite analogous but in different terms: if the rules can distinguish good from bad sentences even though they are impracticably long then they can also distinguish good from bad sentences that are infinitely long; the difference here is that the infinitely long sentences should already exist.

2.3. **Ontological escape hatches.** This is Chapter 6, which is a philosophical extension of what went before and argues, from various points of view, why infinite sequences are real and belong to natural languages.

The conclusion of the chapter is that

> not only is The Natural Language Vastness Theorem a mathematically valid proof, its premises, in particular that natural languages are closed under coordinate compounding, yielding transfinitely long sequences, are all true, and hence its conclusion is a genuine truth about natural languages.

Leaving aside the confusion of theorem and proof, this conclusion does not hold water as we will see when we summarize everything at the end.

The chapter and indeed the rest of the book does use the term 'transfinite sequence' quite freely, and the last chapter has something to say about these.

2.4. **Characterization of transfinite sequences.** This is Chapter 7 (the last one in the book); it deals with transfinite sequences in languages. For the convenience of the reader in recognizable languages such as English or French. The examples give a very limited view of transfinite sequences; the sentences/sequences that we see all look like

> Jack_1 and his father_2 and his father's father_3 and ... and his ... $\text{father}_{\aleph_0}$ are visiting relatives.

The authors consider this a real mathematical object and worthy of inclusion in a natural language but I do not see how this object is defined. What happens during the ellipses is unclear.

In fact, all through the book the authors talk about the length of sequences without ever actually defining what it is. Sequences like the one above are said to have length $\aleph_0$, which is not what a set-theorist would say; we use ordinal numbers, rather than cardinals to measure lengths.

The final part of the chapter deals with some existence issues of orders on sets of linguistic units. Without giving any examples or justification the authors assert that a well-known property of finite sets — every linear order is a well-order — need not follow from any linguistic law.

This then removes many restrictions on what a linear order on a set of linguistic units forming a sentence of transfinite length may look like. The authors acknowledge problems like what densely ordered sentences may look like, and sentences without beginning or end. There is also the problem of how the sets will be ordered.

The authors offer no real help and state that if there are to be lawful orders then the theory of sentences should provide this. They do consider the possibility of well-orders but hesitate because of the non-constructive nature of the well-ordering theorem. Their solution is interesting: identify the linguistic units with ordinals, then the ordering comes for free.

What the authors do not realize, and one may forgive them for not knowing, is that these solutions are not solutions. This is connected to the remark about the non-orderability of the set S_2 in Section 1. We shall come back to this in the next section.

3. Summary and some Set Theory

It may be clear from the previous pages that I am not convinced by the mathematical aspects of the authors' arguments. I will go through these once more and give some extra reasons why they fail to make their case mathematically. There are also a few points that I did not touch upon earlier.

3.1. **Definitions, or lack thereof.** To begin: there is no recognizable definition of what a natural language is. To a mathematician this seems odd: how can you prove statements about undefined entities, in particular a sweeping statement like the Vastness Theorem?

From the book and the paper one comes to the conclusion that such a language is a collection of sequences of some sort, subject to rules of some sort.

Which brings us to the definition of 'sequence', which is also notably absent. Again, one can guess what the authors mean and what all the examples seem to have in common is a linear order, and Chapter 7 makes this more or less explicit.

Probably "function with a linearly ordered domain" comes, mathematically, closest to what the authors have in mind. The length of a sequence, though undefined, is taken to be a cardinal. One would liked to have seen order types, but we shall use the cardinality of the domain as the definition. Also, some of the examples are ill-defined. The finite examples are clear but the infinite examples, like the one from Chapter 7, see subsection 2.4 above, are quite ambiguous.

3.2. **Proof of the vastness theorem.** We analyzed this proof in Section 1.

The main points are

- The definition of co-ordinate projection leaves something to be desired, namely an explicit description of the ordering.
- The existence proof of co-ordinate projections does not deserve that name. It also assumes the conclusion of the vastness theorem, namely that there are sentences of any possible magnitude.
- This makes the whole proof invalid.

Also the claim that there is a strict parallelism with Cantor's results is not true. The proof of the existence theorem hinges on the statement that, given a set, there is a co-ordinate compound, whose set of constituents is of the same magnitude as the given set, in short: for every set there is (another) set of a specific type and of the same magnitude as the given set. This requires proof.

This does not occur in Set Theory; the analogous statement would be for every set there is a set of the same cardinality. This is easy: the set itself will do.

3.3. **Proving an Axiom is difficult.** The authors spent a large portion of the book convincing the reader that imposing size limits on sentences of a natural language is not a linguistic thing and hence that statements like "All sentences have finite length" should not be part of the assumptions when considering such languages. All this with abundant references to Occam's razor.

However, the authors interpret not imposing upper bounds for sentence sizes as stating that there are sentences of arbitrarily large length. That seems to be in conflict with Occam's razor as it introduces all kinds of complications.

The authors should have read Chapter 1 of [5] to see how Set Theory is developed from its axioms and by applying the normal rules of deduction. They would see that the axioms, which were chosen to reflect mathematical practice as much as possible, do not mention size limits at all. Indeed, the idea of cardinality is introduced only after the tools for dealing with it have been developed, and *cardinal numbers* and a *cardinality function* take still more effort.

Consider the case of ZFC. If one drops the Axiom of Infinity (Inf) then one does not replace it with "all sets are finite", one just does not avail oneself of the set that the axiom provides. One loses the power to prove that infinite sets exist; indeed, V_ω, the collection of hereditarily finite sets is a model of "all sets are finite" thus showing that adding it to $\mathsf{ZFC} + \neg\mathsf{Inf}$ does not lead to contradictions.

What this all means is that the statement that is used in the proof of the existence theorem, which we could render as

$$(\forall x)(\exists y)\bigl(\Sigma(y) \wedge y \approx x\bigr) \tag{$*$}$$

should be treated as an axiom. Here $\Sigma(y)$ would be a formula that expresses that y is the set of constituents of some co-ordinate compound.

The book can be seen as a long argument that Axiom ($*$) holds in the universe. But there is definitely no *proof* of ($*$) to be found.

3.4. **Orderings.** A final point considers the use of orderings. As mentioned above Chapter 7 addresses some of the problems of that dense orders might pose for sentences and that it is not always clear how to order constituents before forming them into sentences.

The authors offer two ways out: use well-ordered sets of constituents, and leave the orderings to the rules of the grammar. Both lead to problems, not to say impossibilities.

This is best visible in Subsection 1.6 where the hierarchy $\langle S_n : n \in \omega\rangle$ is developed. The initial set S_0 is well-ordered, simply by its listing $\langle s_n : n \in \omega\rangle$.

The set S_1 is linearly ordered by the lexicographic order of infinite sequences, but it cannot be proved to have a definable well-order. This then means that the rules of the natural language cannot be constructive if one wants a well-order for S_1; one will have to invoke an instance of the Axiom of Choice.

The situation is worse for S_2, which we can identify with the power set of $\mathbb{R}$. In [11] Sierpiński proved that *if* that power set has a linear order *then* there is a non-Lebesgue measurable set and mutatis mutandis a set without the Baire property. This means that the rules of the natural language must be quite ugly as they entail the existence of quite ugly sets.

All in all I conclude that the book does not deliver on its promise.

References

[1] Bohuslav Balcar and Petr Štěpánek, *Teorie množin*, Academia [Publishing House of the Czech Academy of Sciences], Prague, 2000 (Czech). second, corrected and extended edition.

[2] Georg Cantor, *Über eine elementare Frage der Mannigfaltigkeitslehre*, Jahresbericht der Deutschen Mathematischen Vereinigung **1** (1890/91), 75–78.

[3] Paul J. Cohen, *Set theory and the continuum hypothesis*, W. A. Benjamin, Inc., New York-Amsterdam, 1966. MR0232676

[4] Murray Eisenberg, *Axiomatic theory of sets and classes*, Holt, Rinehart and Winston, Inc., New York-Montreal, Que.-London, 1971. MR0277373

[5] Kenneth Kunen, *Set theory. An introduction to independence proofs*, Studies in Logic and the Foundations of Mathematics, vol. 102, North-Holland Publishing Co., Amsterdam-New York, 1980. MR597342

[6] D. T. Langendoen and P. M. Postal, *The Vastness of Natural Languages*, Basil Blackwell, Oxford, 1984. Available at `https://archive.org/details/vastnessofnatura00lang`.

[7] ———, *Sets and Sentences*, The Philosophy of Linguistics (Jerrold J. Katz, ed.), Oxford University Press, Oxford, 1985, pp. 225–248.

[8] Marc van Oostendorp, *Boeken zijn als getallen* (15 oktober 2019), `https://www.neerlandistiek.nl/2019/10/boeken-zijn-als-getallen/getallen`.

[9] Paul M. Postal, *Books* (August 2019), `https://ling.auf.net/lingbuzz/004733`.

[10] Bertrand Russell, *On Some Difficulties in the Theory of Transfinite Numbers and Order Types*, Proceedings of the London Mathematical Society **4** (1907), 29–53, DOI 10.1112/plms/s2-4.1.29.

[11] W. Sierpiński, *Sur une propriété de la décomposition de M. Vitali*, Mathematica, Cluj **3** (1930), 30–32 (French).

Faculty EEMCS, TU Delft, Postbus 5031, 2600 GA Delft, the Netherlands
Email address: k.p.hart@tudelft.nl
URL: http://fa.ewi.tudelft.nl/~hart