Strict finitism, feasibility, and the sorites

Walter Dean
Department of Philosophy
University of Warwick
e-mail: W.H.Dean@warwick.ac.uk

This paper bears on three different topics: observational predicates and phenomenal properties; vagueness; and strict finitism as a philosophy of mathematics. Of these three, only the last requires any preliminary comment. Dummett (1975), p. 302

Why are mathematicians so convinced that exponentiation is total? Because they believe in the existence of abstract objects called numbers. What is a number? Originally, sequences of tally marks were used to count things. Then positional notation – the most powerful achievement of mathematics – was invented. Nelson (1986), p. 173

Abstract. This paper bears on four topics: observational predicates and phenomenal properties, vagueness, strict finitism as a philosophy of mathematics, and the analysis of feasible computability. It is argued that reactions to strict finitism point towards a semantics for vague predicates in the form of nonstandard models of weak arithmetical theories of the sort originally introduced to characterize the notion of feasibility as understood in computational complexity theory. The approach described eschews the use of non-classical logic and related devices like degrees of truth or supervaluation. Like epistemic approaches to vagueness, it may thus be smoothly integrated with the use of classical model theory as widely employed in natural language semantics. But unlike epistemicism, the described approach fails to imply either the existence of sharp boundaries or the failure of tolerance for soritical predicates. Applications of *measurement theory* (in the sense of Krantz et al. 1971) to vagueness in the nonstandard setting are also explored.

Dummett's (1975) "Wang's Paradox" is commonly cited as the origin of contemporary philosophical interest in vagueness. It is thus notable that in this paper Dummett's point of departure was not a proposal in philosophy of language, but rather one about the foundations of mathematics – i.e. Yessenin-Volpin's (1961; 1970) exposition of the view Dummett calls *strict finitism*. Proponents of this standpoint seek to ground mathematics in operations which are performable "in practice" as opposed to merely "in principle". The formulation of strict finitism thus requires that we take seriously notions such as *feasibly constructible number* or *surveyable proof* which Dummett famously argued are susceptible to versions of the sorites paradox.

The aim of this article will not be to rehabilitate strict finitism itself. However I will argue below that Dummett's repudiation of such a view is based on assumptions which Yessenin-Volpin and other strict finitists would almost certainly have rejected. Among these are several suppositions about how we use number systems to count or measure magnitudes and how we use numerals to denote numbers – e.g. that mathematical induction holds for arbitrary predicates of natural numbers, that for every decimal numeral there is a co-denoting unary numeral, or that the fields of numbers which may legitimately be used to measure empirical or psychological magnitudes must always be Archimedean.

One might, of course, be reluctant to abandon such assumptions as part of one's chosen foundation for mathematics or theory of scientific measurement. But the examples which have traditionally been used to motivate strict finitism have typically concerned predicates $F(x)$ like *feasibly constructible number* (or *small number*) whose mathematical or scientific status may initially be unclear. Such predicates are typically assumed to hold of 0 – i.e. $F(0)$ – to be closed under successor – i.e. $\forall x(F(x) \rightarrow F(x+1))$ – but fail to hold of certain *infeasible* (or *large*) numbers – e.g. $\neg F(1000000)$ or $\neg F(2^{50})$. Part of the proposal I will develop below involves the fact that it is possible to provide a consistent interpretation of these premises by treating the syntactic expressions which are typically employed to denote such values as referring to "infinite integers" in nonstandard models of arithmetical theories which are so weak that they fail to prove the totality of functions (such as exponentiation) which are required to prove the existence of co-denoting unary numerals.

The origin of such theories within mathematical logic can be traced to attempts to come to terms with strict finitism as a foundational standpoint. But I will also suggest below that they are applicable as part of a theory of vague predicates within natural language. In this setting, the use of formal arithmetic and nonstandard methods more generally is almost completely foreign.[1] I thus face the uphill battle of convincing readers whose primary interests lie in philosophy of language or linguistics that there is something to be gained from reconsidering the sorites in light of such technicalities.[2]

The exposition which follows is intended to speak directly to this concern. In §1, I will begin by isolating a class of sorites arguments which are formulated in terms of what I refer to as *ordinal predicates*. Such predicates are distinguished by the fact that they are associated with a *sortal unit* – e.g. a *grain* of wheat, sand, etc. in the case of *heap* or a single *hair* in the case of *bald*. As such units can be counted in the same manner as natural numbers, I will argue that the formulation of the corresponding soritical arguments relies on mathematical premises related to how we are able to refer to the members of the corresponding sorites sequences.

Spelling out these details will provide a useful context in which to consider the historical connections between the sorites and strict finitism in §2. In §3 I will consider a related proposal of Parikh (1971) for responding to Dummett's critique of strict finitism using so-called *almost consistent theories*. While inconsistent in the traditional sense, an appropriately formulated almost consistent theory has the property that any proof of a contradiction from its axioms must be "infeasibly long". This observation serves as the basis for what I will refer to as the *feasibilist theory of vagueness* – i.e. the view that

[1] The only exception to this of which I am aware is work on vagueness in the tradition of Vopěnka's (1979) Alternative Set Theory [AST]– e.g. (Hájek, 1973), (Novák, 1992), (Tzouvaras, 1998). Many of the observations framed below could also be formulated in this setting. I have opted for a presentation based on classical first-order arithmetic and analysis both in virtue of their greater familiarity and also because the current proposal does not depend on the critique of Cantorian set theory which Vopěnka used to motivate AST.

[2] Of paramount concern to such readers is likely to be the question of whether nonstandard interpretations can plausibly be regarded as providing faithful models of how speakers understand the meanings of vague predicates in everyday use. Although this issue will often be in the background below, I have postponed direct consideration until §7. For as I will suggest there, the question of whether a nonstandard interpretation may plausibly be regarded as "psychologically real" is intertwined with a number methodological issues in semantic theory which are often overlooked. Among these are the status of the idealization which comes along with the use of formal languages in model theoretic semantics (as understood within linguistics) as well as the interaction of this subject with results in model theory (as understood within logic).

although the premises of the sorites are inconsistent in principle, this does not pose a threat to the consistency of everyday reasoning as we are unable to derive a contradiction from its salient instances in practice.[3]

Although such a view is in keeping with the general spirit of my proposal, I will argue that its plausibility is substantially diminished by results which demonstrate that certain forms of soritical reasoning may be "sped up" using a proof-theoretic technique first described by Solovay (1976) and later rediscovered by Boolos (1991). I will also suggest that feasibilism runs afoul of the goal of providing a semantics for vague predicates compatible with classical model theory. Such observations motivate the adoption of what I will refer to as the *neo-feasibilist theory of vagueness*. The application of this view to instances of the sorites based on ordinal predicates will be developed in §4 using nonstandard models of arithmetic of the sort described above. In §5 I will also propose that nonstandard models of analysis can be employed to provide a related interpretation of cases of the so-called *phenomenal sorites* which was originally described by Dummett (1975) using the predicate *looks red*.

As instances of the phenomenal sorites often do not appear to depend on the use of terms denoting large numbers, the relevant application of nonstandard techniques does not in this case draw on the same metaphors involving the notion of feasibility just described. But it remains to be seen what can be said for or against neo-feasibilism relative to the desiderata which are used to evaluate mainstream approaches to vagueness. I will address these issues in §6. Therein I will consider how neo-feasibilism relates to supervaluationism and epistemicism and also describe how it interacts with phenomena such as borderline cases and higher-order vagueness. Finally in §7, I will discuss some remaining methodological issues about the role of mathematical representation and the notion of *finitude* in the semantics of vague predicates.

§1. On the logical formulation of the sorties Before it is acknowledged that the sorites is indeed a *paradox* it is reasonable to demand that it be formulated in the manner exemplified by the semantic or set theoretic paradoxes – i.e. as a fixed formal derivation leading from apparently true premises to a contradiction. It is thus notable that the sorites is often treated as possessing several distinct forms. A useful guide is provided by Hyde (2011) who distinguishes between the *conditional*, *inductive*, *line drawing*, and *phenomenal* variants of the sorites argument. I will focus on the first two forms in §1 through §4 before returning to the latter, respectively in §5 and §6.1.

Suppose that $P(x)$ is a soritical predicate and that $\mathsf{a}_0, \mathsf{a}_1, \ldots$ are constant symbols denoting the items purported to comprise a sorites sequence. Hyde's formulations of the conditional and inductive forms of the sorites are as follows:

1.1 The conditional sorites

$$\begin{array}{l} P(\mathsf{a}_0) \\ P(\mathsf{a}_0) \rightarrow P(\mathsf{a}_1) \\ P(\mathsf{a}_1) \rightarrow P(\mathsf{a}_2) \\ \quad\vdots \\ P(\mathsf{a}_{n-1}) \rightarrow P(\mathsf{a}_n) \\ \hline \therefore\ P(\mathsf{a}_n) \text{ (where } n \text{ can be arbitrary large)} \end{array}$$

[3] Such a view was originally put forth by Parikh (1983) and has been developed more recently by Sazonov (1995), Carbone (1996), and Magidor (2011). As I will explore further in notes 25 and 37, feasibilism also has affinities with more traditional contextualist approaches to vagueness such as those of Kamp (1981) and Gaifman (2010).

$P(x)$	type	unit	n_P	reference
walking distance	ratio	foot	6000	Gaifman (2010)
non-heap	ordinal	grain	10000	Barnes (1982), Sainsbury (1995)
few	ordinal	grain	10000	Barnes (1982), Williamson (1994)
noonish	ratio	second	10000	Sorensen (2001)
height of a tall person	ratio	.01 inch	91200	Graff (2001)
small town	ordinal	inhabitant	50000	Gómez-Torrente (2010)
bald	ordinal	hair	1000000	Fine (1975), Tye (1994)
age of a child	ratio	second	10^9	Shapiro (2011)
heartbeat in childhood	ordinal	heartbeat	2.5×10^9	Dummett (1975)
poor	ratio	€.01	10^{10}	Kölbel (2010)
age of an old person	ratio	nanosecond	3×10^{18}	Field (2008)

Table 1. *Examples of numbers used as extremal values in the formulation of the sorites.*

1.2 The inductive sorites

$$\begin{array}{l} P(\mathsf{a}_0) \\ \forall i(P(\mathsf{a}_i) \rightarrow P(\mathsf{a}_{i+1})) \\ \hline \therefore\ \forall i P(\mathsf{a}_i) \end{array}$$

My goal in this section will be to argue that our willingness to regard instances of these argument schema as paradoxical is mediated by mathematical assumptions which are not typically stated among their premises. This is immediate in the case of the inductive sorites, as the conclusion is understood to be obtained by applying *modus ponens* to the displayed premises together with an unstated instance of the mathematical induction scheme for the predicate $P(x)$ – a point to which I will return below. But there are several other ways in which derivations 1.1 and 1.2 fall short of comprising formally correct derivations of a contradiction from the displayed premises. Note, for instance, that bringing the conclusions $P(\mathsf{a}_n)$ or $\forall i P(\mathsf{a}_i)$ into logical conflict with the premises of 1.1 and 1.2 requires that we also include a premise of the form $\neg P(\mathsf{a}_n)$.[4] This in turn requires a paradox monger to explicitly designate a term of the form a_{n_P} for which the predicate $P(x)$ does not hold.

Two observations come into view at this point: 1) the presentations just supplied are typical not only in their use of constant symbols to denote the elements of a sorties series for $P(x)$ but in their use of *natural numbers* to subscript them; 2) the value which is supplied for n_P – which I will refer to as the *extremal value* for $P(x)$ – will typically be a *large* number. Some published examples are provided in Table 1.

In order to distinguish the cases to which the feasibilist and neo-feasibilist theories to vagueness are most directly applicable, it will be useful to begin by introducing a distinction between what I will refer to as *ordinal* and *ratio* predicates. An ordinal predicate $P(x)$ is characterized by the existence of an associated sortal predicate $S_P(x)$ such that the determination of whether a given object or aggregate o falls under $P(x)$ depends on the number of items satisfying $S_P(x)$ which comprise o. For instance, of the predicates listed in Table 1, *non-heap*, *bald*, and *small town* are ordinal predicates and have the associated sortal units *grain* (of sand, wheat, etc.), *hair*, and *inhabitant*. There will typically be a close semantic relationship between $P(x)$ and $S_P(x)$ – e.g. while the predicate *bald* may not have a sharp boundary, it is still plausible to assume that whether a man is bald is

[4] Expressions like $\forall i P(\mathsf{a}_i)$ and $\forall i(P(\mathsf{a}_i) \rightarrow P(\mathsf{a}_{i+1}))$ (i.e. the so-called *tolerance principle* for $P(x)$) are often used to formulate the inductive sorites. Strictly speaking, however, such formulas are disallowed by the syntax of first-order logic as quantifiers cannot bind the indices of constants as they are not variables in the object language. This provides further motivation for either adopting an arithmetical language to formalize the sorites or to treat tolerance as a primitive relation as discussed in §5.

determined by the number of hairs on his head. Thus although an extremal value for a given ordinal predicate will typically be determined empirically, such a value does not depend on a further choice of unit in terms of which the items falling under $P(x)$ must be counted or measured.[5]

The situation is more complex in the case of ratio predicates. This class is exemplified by predicates like *walking distance*, *noonish*, or *lightweight* which apply to empirical magnitudes such as lengths, times, or masses. To linguistically formulate a sorites argument for such predicates, a unit of measure – e.g. feet, second, or pounds – is commonly chosen in order to systematically assign denotations to the constants a_i appearing in derivation 1.1. But since the choice of such a unit is generally arbitrary, the relationship between ratio predicates and plausible choices of extremal values will typically not be as direct as in the case of ordinal predicates.[6] Similar remarks apply to phenomenal predicates such as *looks red*, *tastes sour*, and *sounds louder* whose degrees we also often measure indirectly using empirical units of measure (e.g. nanometers, pH units, or decibels).

As I will discuss further in §5, the terms "ordinal" and "ratio" are chosen to reflect the sorts of numerical scales which would be most naturally employed to represent the extensions of the predicates in question relative to the conventions of *measurement theory* in the tradition of (Krantz et al., 1971). The potential applicability of measurement theory to vagueness has been widely explored in linguistics in the study of gradable adjectives – e.g. (Sassoon, 2010), (van Rooij, 2011), (Lassiter, 2011). I will return to discuss such predicates in §5 and §7. But as their treatment introduces a number of complications into the logical formulation of the sorites, I will concentrate initially on simpler cases involving ordinal predicates which may be formulated without explicitly introducing measurement-theoretic apparatus.

Letting $P(x)$ be an ordinal predicate, next note that the expressions $\mathsf{a}_0, \ldots, \mathsf{a}_n$ appearing in derivation 1.1 are traditionally understood as constant symbols denoting objects which are in its field of application. Let us call the set of these objects $O_P = \{o_0, \ldots, o_{n_P}\}$. Traditional presentations of the sorites also assume that such a set comes equipped with a discrete linear order $\prec_P$ which relates the o_i such that the constant symbol a_i denotes the $i+1$st object with respect to the order $\prec_P$. It is then conventional to classify a predicate $P(x)$ as *soritical* just in case it satisfies the following properties: a) it appears that $P(x)$ is true of the object o_0 denoted by a_0; b) it appears that $P(x)$ is false of the last object o_{n_P} denoted by a_{n_P}; and c) each pair of adjacent objects o_i and o_{i+1} $(i < n_P)$ appear sufficiently similar so as to be indiscriminable with respect to $P(x)$.

Note, however, that at least in the case of ordinal predicates such as *bald* or *non-heap*, the world rarely presents us directly with a structure $\mathcal{P} = \langle O_P, \prec_P \rangle$ satisfying the properties just described. In order to describe an instance of the conditional sorites we must thus construct (either in reality or via a thought experiment) a sequence of collections of grains, men, etc. such that the first member is comprised of a single item falling under

[5] Although I will adopt the former claim as a simplifying assumption below, it is clear that the latter remains true even if we acknowledge that factors other than the number of items following under $S_P(x)$ determine whether a given object satisfies $P(x)$ – e.g. the arrangement or density of hairs and the size of a given man's scalp in the case of *bald*. More generally, it should be clear how the neo-feasibilist proposal can be integrated with various contextual and multi-dimensional views of vagueness by adding appropriate parameters to the sorts of interpretations described below. According to the neo-feasibilist, however, such additions are not required for the resolutions of the relevant instances of the sorites themselves.

[6] For instance, although Gaifman (2010) uses the number 6000 to identify a *non-walking distance* measured in feet, he might also have used the numbers 1828800, 182880, 1828.8, or 1.8288 had he elected to describe the same magnitude in millimeters, centimeters, meters, or kilometers.

$S_P(x)$, the second member is comprised of two such items, etc. Thus in order to formulate a soritical argument of the form 1.1. or 1.2, we must additionally be in possession of terms which allow us to refer to the members of O_P in the relevant order.

It is at this point where the significance of the magnitudes of the values of n_P reported in Table 1 begins to come into sharper focus. For note that at least in the case of instances of the inductive or conditional sorites involving ordinal predicates, most published examples involve cases where n_P is taken to be 10000 or greater.[7] On the other hand, natural languages do not typically contain sufficiently many *syntactically primitive* expressions by which we might systematically make reference to the items in a sorites sequence whose length is measured by the sorts of values in Table 1.[8] This is not, of course, to deny

[7] Once such a claim has been made using a specific value, one is, of course, drawn to search for compelling counterexamples. One source of examples which do not appear to involve large numbers is the phenomenal sorites which will be discussed in §5. But another likely objection is that it is possible to construct instances of the conditional sorites involving what might be termed "medium numbers" which are "significantly smaller" than those depicted in Table 1. Suppose, for instance, that a teacher states that a "fairly large" number of students were absent from an exam. If the class consisted of 25 students, then although a collection of 0 or 1 or 2 ... students presumably do not fall under the intended sense of "fairly large", a collection of 15 presumably does. Of course simply pointing to such cases does not itself suggest a principled means by which they can be distinguished from those involving the numbers given in Table 1. Nonetheless, four additional points may be noted in this regard. First, it seems that the sort of cases in question are indeed most naturally formulated using modifiers such as "a fairly large number of" or "many" to obtain what is purported to be a soritical predicate of the form $Q(x) =$ *a fairly large number of/many* $P(x)$. But in such case it may be possible to provide an independent contextual account of how such modifiers contribute to the meaning of $Q(x)$ which allows for another means of resolving the "short" sorites arguments in question (see, e.g., Fernando and Kamp, 1996). Second, even if this step is not taken, nothing about the view developed in this paper imposes a lower bound on the extremal value which a paradox monger might attempt to use to formulate an instance of the inductive or conditional sorites (inclusive of $n_P = 15$). Third, as we will see in §3, it is possible to derive a contradiction from the premises of the conditional sorites with $n_P = 10000$ in a "medium" number of steps (≤ 90) which is not only "significantly smaller" than 10000 but is also in the range which a human subject could reliably carry out. (Although this fact was observed by Boolos (1991), it appears to have had little subsequent influence on the literature on vagueness in philosophy or linguistics. Nonetheless, I will suggest below that the phenomenon in question provides a fundamental motivation for moving from a feasibilist to a neo-feasibilist view of vagueness.) Fourth, observe that in the case described at the beginning of this note we are asked to divide a sharply defined set X (the students in the class) into a subclass Y with a vague boundary (a "fairly large" subset of absent students) – i.e. a *proper semiset* in the terminology of Vopěnka (1979). In such cases the division arises in virtue of the *relative cardinality* of X to Y. This suggests that the possibility of making real (or at least rational) valued *measurements* enters into how we conceive of such cases – a point which becomes more apparent when we consider similar instances involving "fairly tall" members of a team or "fairly warm" days in a month. This in turn suggests that aspects of the measurement theoretic approach to the sorites described in §5 are often applicable to cases involving "medium" numbers as well. For more on the specific numbers involved, see notes 31 and 53 below.

[8] This point can be made more vivid by considering what would happen if we attempted to draft ordinary given names to denote the members of O_P. We might, of course, elect to employ the lexicographic features of such expressions to keep track of the order $\prec_P$ – e.g. by taking `Adam` to denote o_0, `Benjamin` to denote o_1, `Charles` to denote o_2, ... But if we are unable to rely on a formally defined ordering of strings in this manner, it seems likely that we would quickly lose track of the order in which they are intended

that we possess *syntactically complex* terms which allow us to refer to the items in a sorites sequence $\langle O_P, \prec_P \rangle$. A strategy which is closer to how we appear to understand the reasoning of derivation 1.1 in practice takes inspiration from the fact that the use of numerals as the subscripts to the constants a_i serves both to distinguish these expressions orthographically and also as an external notational device which allows us to keep track of the order in which we wish to denote the elements of O_P.[9] One way of exploiting this observation is to define an operation which for each ordinal predicate $P(x)$ returns a predicate $P^*(x)$ of natural numbers defined by

$$P^*(n) \text{ iff a collection of } n \text{ items of type } S_P(x) \text{ comprises an object satisfying } P(x)$$

I will refer to $P^*(x)$ as the *arithmetization* of the predicate $P(x)$.

If we are to reason deductively about such predicates, it would seem that we must also acknowledge that the language in which appropriately corrected versions of derivations 1.1 and 1.2 are formulated overlaps at least partially with the language $\mathcal{L}_a = \{\mathsf{0}, ', \mathsf{+}, \times, \mathsf{<}\}$ of first-order arithmetic over which formal theories such as Robinson or Peano arithmetic (i.e. Q and PA) are formulated. For if we wish to reformulate these derivations using $P^*(x)$ instead of $P(x)$ we will need to employ singular terms which allow us refer to an initial segment of the natural numbers and also to speak of the successor of an arbitrary number. A familiar way of doing this is to employ the constant $\mathsf{0}$ to denote 0 and the symbol $'$ to denote the successor function. This allows us to formulate terms $\mathsf{0}', \mathsf{0}'', \ldots$ denoting the numbers $1, 2, \ldots$ I will refer to such expressions as *unary numerals* and employ the standard abbreviation $\overline{n} = \mathsf{0}'^{\cdots\prime}$ (n times) to denote the corresponding terms of $\mathcal{L}_a$.

At this point one might begin to worry that there is nothing about our "naive" understanding of the soritical premises which uniquely determines which form of arithmetical notation must be used to formulate them linguistically. Note, however, that we have yet to confront the fact that derivation 1.1 makes use of an ellipsis to abbreviate the steps intervening between the conditionals $P(\mathsf{a}_0) \to P(\mathsf{a}_1)$ and $P(\mathsf{a}_{n_P-1}) \to P(\mathsf{a}_{n_P})$. The use of such a device in the formulation of the conditional sorites is, of course, eliminable in favor of an explicit listing of the conditionals $P(\mathsf{a}_i) \to P(\mathsf{a}_{i+1})$ for $0 < i < n_P$. But at least for the sorts of values considered in Table 1, it will generally be beyond our practical abilities to either write down or survey the resulting derivation in its entirety.

This familiar observation also does not depend on whether we use numerals or some other sorts of terms in order to denote the members of a sorites sequence. But it does highlight another respect in which our apprehension of the conditional sorites appears to implicitly depend on arithmetical assumptions. We have just observed that an explicit formulation of such a derivation may be exceedingly long in its "vertical" dimension – i.e. as measured in the number of *steps* it contains. But we also expect that we should be able to survey not only expressions like a_{10000} which appear in the final line of such a derivation, but also that we could construct any of the elided steps – e.g. $P(\mathsf{a}_{4913}) \to P(\mathsf{a}_{4914})$ – if called upon to do so. This feature of the derivation appears to require that these steps are formulated so that they are also reasonably short in their "horizontal" dimension – e.g. as measured in terms of the number of *symbols* they contain.

to denote the o_i. And in this case it seems that we would accordingly lose faith in the acceptability of statements which would replace the premises formulated above using subscripted constants – e.g. $P(\mathtt{Matthew}) \to P(\mathtt{Mark})$, $P(\mathtt{Luke})$, or $\neg P(\mathtt{John})$.

[9] Another option would be to make use of the fact that since we have assumed that $\prec_P$ is a discrete linear order with a left endpoint, we may refer to *the first element of* $\prec_P$, *the second element of* $\prec_P$, ... But although we could then make use of a Russellian ι-operator (or a similar device) to formalize the reasoning in question, it is evident that in order to do so we would need to make use of an indexed family of variables $x_0, x_1, x_2, \ldots$ similar to the indexed constants $\mathsf{a}_0, \mathsf{a}_1, \mathsf{a}_2, \ldots$

These observations suggests that whatever system of notation we adopt for formalizing the conditional sorites must satisfy two criteria. First, it must employ some form of compositional device which allows us to inductively construct a notation denoting the object o_{i+1} in the sorites sequence given a notation denoting the object o_i. And second, it must also provide sufficiently compact notations to allow us to write down terms of surveyable length which denote the sorts of numbers exemplified by those in Table 1.

It will also come as no surprise that while unary numerals possess the first property, they fail to possess the second. For given a unary numeral v with *value* n (notation: $[\![\overline{v}]\!]_u = n$), we can obtain a unary numeral denoting $n+1$ by simply concatenating another token of the symbol $'$. But on the other hand, the *length* n of a unary numeral (notation: $|v| = n$) is directly proportional to its value (i.e. $[\![v]\!]_u = |v| - 1$). A consequence of this is that we will be incapable in practice of either writing down unary numerals denoting numbers on the order of (say) 10000 or of dependably distinguishing a unary numeral of this magnitude from that denoting its successor. But it is presumably just such an ability which stands behind the faith which we express in conditionals such as $P^*(\overline{4913}) \to P^*(\overline{4914})$.[10]

It will also not come as a surprise that the familiar system of *decimal numerals* satisfies both properties simultaneously. For although the length of such a numeral does not directly reflect the position of the number it denotes in the natural number sequence, we are well aware of a procedure for constructing a numeral denoting $i+1$ given a numeral denoting i – i.e. by incrementing the value of its last digit by one and carrying as needed. It is a similarly familiar observation that decimal numerals are a form of *positional notation*. This is to say that such a numeral has the form of a finite sequence $\pi = \mathtt{d}_0, \ldots, \mathtt{d}_{k-1}$ of *digits* drawn from a finite set of symbols $\{s_0, \ldots, s_{b-1}\}$ (where $b \geq 2$ is the *base*) such that $[\![s]\!]_b = i$. Although the length of such an expression is still given by the number of symbols it contains – i.e. $|\pi| = k$ – its value is now given indirectly as $[\![\pi]\!]_b = \sum_{i=0}^{|\pi|-1} [\![\mathtt{d}_i]\!]_b \cdot b^i$.[11] As a consequence, the length of the decimal numeral denoting n will be given by $\lfloor \log_{10}(n) \rfloor + 1$. It is this feature of decimal numerals which allows us to concretely write down expressions which denote not only the sorts of values depicted in Table 1, but also "astronomically" larger ones as well – e.g. a decimal numeral of 20 digits denotes a number greater than the estimated age of the universe in seconds.[12]

Suppose we now assume that $\mathcal{L}_P$ contains a *constant symbol* $\underline{n}$ corresponding to each decimal numeral $n \in \mathbb{N}$.[13] An improved version of the conditional sorites can now be

[10] Both points would, of course, still apply if we continued to use constant symbols $\mathtt{a}_i$ or definite descriptions using subscripted variables x_i in the formulation in the case that the subscripts were themselves unary numerals.

[11] For instance $|\underline{4913}| = 4$ and $[\![\underline{4913}]\!]_b = 4 \cdot 10^3 + 9 \cdot 10^2 + 1 \cdot 10^1 + 3 \cdot 10^0$.

[12] Needless to say, much the same would be true if we employed (e.g.) binary rather than decimal notation. In general, since the length of the base-b numeral representing $x \in \mathbb{N}$ is given by $\lceil \log_b(x) \rceil$ it follows that for arbitrary bases $b_1, b_2 \geq 2$, the length of the representation of x in base b_1 will be bounded by a scalar multiple $\alpha = \log_{b_1}(b_2)$ of the length of its representation in base b_2 (e.g. for $b_1 = 2$ and $b_2 = 10$, $\alpha \approx 3.3219$). As scalar differences of this magnitude will generally not affect the considerations discussed below, the reader is free to think of positional numerals as being given for an arbitrary base $b \geq 2$.

[13] As the syntactic form of various sorts of numerical expressions will often be at issue below, the following notational conventions should be underscored: i) v will be used as a meta-variable ranging over unary numerals which are comprised of the symbol $\mathtt{0}$ followed by a finite number of stroke symbols $'$; ii) the unary numeral for $n \in \mathbb{N}$ is of the form $\mathtt{0}^{'\cdots'}$ (n times) which should be understood as corresponding to a (possibly complex) term of the language $\mathcal{L}_a$ and will be abbreviated by $\overline{n}$; iii) π will be used as a metavariable over positional numerals which – when viewed externally – are finite sequences of symbols over an alphabet containing at least two symbols; iv) given a

formulated in Fitch-style natural deduction as follows:

1.3 The numerical conditional sorites

$$
\begin{array}{l|l}
1 & P^*(\underline{0}) \\
2 & P^*(\underline{0}) \rightarrow P(\underline{1}) \\
\vdots & \vdots \\
n_P & P^*(\underline{n_P - 1}) \rightarrow P^*(\underline{n_P}) \\
n_P + 1 & \neg P^*(\underline{n_P}) \\
\cline{2-2}
n_P + 2 & P^*(\underline{1}) \\
\vdots & \vdots \\
2n_P + 1 & P^*(\underline{n_P}) \\
2n_P + 2 & \bot
\end{array}
$$

Derivation 1.3 differs from 1.2 in that each of its steps will now contain a number of symbols bounded by a scalar multiple of $\log_{10}(n_P)$ rather than by n_P itself. But the question remains as to whether we ought to regard such a presentation as a genuine demonstration of a contradiction or merely as a template for deriving a contradiction which may still depend on some additional premises. For *qua* formal proof, 1.3 is still deficient in the sense of containing an ellipsis – a feature which will not be true of familiar formulations of the set theoretic or semantic paradoxes, even when they are written out in full detail. Of course we understand that this gap may be filled in by using decimal numerals to write out the suppressed premises explicitly and then repeatedly applying *modus ponens* as above. But nothing about 1.3 depicts this explicitly.

Several options seems to be available to address this concern. For instance, we can acknowledge that although 1.3 is a sufficient indication of the *structure* of a derivation which yields a contradiction from the displayed premises, whatever faith we have that the elided steps can be filled in to achieve this result is grounded in our acceptance of an auxiliary theory which both described how we are able to construct the appropriate decimals numerals and perform the requisite propositional reasoning. I will have more to say about the specific form which such a theory might take in §4 and Appendix A. But while the length of the terms appearing in 1.3 are bounded in the manner described above, the fact remains that the number of steps which must be constructed is still proportional to the *value* (and hence the *length*) of the unary numeral $\overline{n}_P$.

This in turn suggests that if our hope is to find an adequate logical formulation of the conditional sorites, then we must take the relevant theory to be sufficiently powerful to formalize and derive the following principle about the relationship between unary and positional numerals:

(D) For every base $b \geq 2$ positional numeral π, there exists a unary numeral υ such that $[\![\pi]\!]_b = [\![\upsilon]\!]_u$.

---

choice of base $b \geq 2$, the positional numeral for $n \in \mathbb{N}$ will be abbreviated by $\underline{n}$. In §1-§6 the latter sort of expression should be understood as a *primitive constant symbol* which is being adjoined to a one-sorted language extending $\mathcal{L}_a$ such that object-language variables $x, y, z, \ldots$ can be substituted with either terms of the form $\overline{n}$ or $\underline{n}$ (see also note 40). In Appendix A, these simplifying assumptions will be altered so that $\underline{n}$ can be viewed as terms over a two-sorted language extending $\mathcal{L}_a$ with variables $X, Y, Z, \ldots$ intended to range over the denotations of positional numerals.

Such a principle might at first seem like a necessary truth of either mathematics or a related theory of formal syntax. Observe, however, that we can readily write down or survey concrete inscriptions of decimal numerals for which we have no hope of concretely constructing or surveying a unary numeral denoting the same value. This in turn highlights a sense in which our apprehension of positional numerals is often not fully *transparent* – i.e. there are many instances in which we can fully survey a positional numeral π *qua* formal term without at the same time being able to grasp its reference under the "canonical" presentation provided by a unary numeral whose length coincides with the value $[\![\pi]\!]$.[14]

Such observations may in turn sow seeds of doubt as to whether we are in fact obliged to accept (D) in virtue of accepting the minimal amount of number theory which I have just argued is required to apprehend the premises of a derivation like 1.1. It will be the burden of §4 to illustrate that it is indeed mathematically coherent to acknowledge these principles while remaining agnostic about or even rejecting (D).

On the other hand, if our underlying aim is to uphold the paradoxicality of familiar instances of the sorites, another option would be to acknowledge – appearances potentially to the contrary – that the inductive form 1.2 is in fact a more basic representation of how we apprehend the argument. Note, however, that mathematical induction is presumably a principle which applies to natural numbers as opposed to men, collections of wheat grains, etc. Such an approach thus already seems to be premised on the assumption that the sorites is most appropriately formulated in an arithmetized language such as $\mathcal{L}_{P^*} = \mathcal{L}_a \cup \{P^*(x)\}$. Given such an understanding, it is then straightforward to present derivation 1.2 as follows:

1.4 The numerical inductive sorites

1	$P^*(\mathtt{0})$
2	$\forall x(P^*(x) \to P^*(x'))$
3	$P^*(\mathtt{0}) \wedge \forall x(P^*(x) \to P^*(x')) \to \forall x P^*(x)$
4	$\neg P^*(\underline{n}_P)$
5	$P^*(\mathtt{0}) \wedge \forall x(P^*(x) \to P^*(x'))$
6	$\forall x P^*(x)$
7	$P^*(\underline{n}_P)$
8	$\bot$

Unlike 1.3, 1.4 is a logically correct and fully explicit derivation of a contradiction from the displayed premises. These include an instance of the induction schema for $\mathcal{L}_{P^*}$ – i.e.

$\mathrm{Ind}(\mathcal{L}_{P^*})$ For all $\mathcal{L}_{P^*}$-formulas $\varphi(x)$, $\varphi(\mathtt{0}) \wedge \forall x(\varphi(x) \to \varphi(x')) \to \forall x\varphi(x)$.

Another means of arguing that the the elided steps in derivation 1.3 can be filled in is to argue for the applicability of mathematical induction to the predicate $P^*(x)$. In

[14] Of course one might also take the view that for many practical purposes, positional numerals provide not only our most practical way of referring to and calculating with natural numbers, but are also the most basic medium by which we are able to entertain what we might otherwise be inclined to describe as *de re* beliefs about individual numbers. (Such a view is described in the unpublished Whitehead Lectures of Saul Kripke wherein decimal numerals are described as referntial "buck stoppers". See (Steiner, 2011) for a partial reconstruction of this proposal.) This point notwithstanding, the conditional sorites still seems like a context in which we are forced to simultaneously consider the mechanisms by which positional and unary numerals refer. I will develop this point further in §4.2.

particular, the instance of induction used in 1.4 can be understood as recording precisely the assumption that we are justified in concluding $P^*(t)$ for an arbitrary $\mathcal{L}_{P^*}$-term t whenever we are disposed to also accept $P^*(\underline{0})$ and that for an arbitrary x we can conclude $P^*(x')$ from the assumption $P^*(x)$. But this is exactly the pattern of reasoning we presumably invoke when we conclude that we could in principle derive a statement such as $P^*(\underline{1000000})$ by carrying out the "forced march" reasoning of the elided steps in 1.3 despite the fact that we have not actually done so in practice.

§2. On strict finitism and feasibility Prior to the upsurge in philosophical interest in vagueness ushered in by Dummett's "Wang's Paradox", the significance of large finite numbers and our various means of denoting them was discussed intermittently by mathematicians and philosophers during the late 19th and early 20th centuries. The observation that such numbers pose a challenge to certain views about the foundations of arithmetic is at least as old as Frege's use of numbers such as 135664 and $1000^{1000^{1000}}$ to critique Kant's view that natural numbers must have intuitable representations.[15] Bernays (1935) made a related point about the unintuitabilty of $67^{257^{729}}$ in the context of arguing that no sharp boundary demarcates the sort of "intuitive evidence" which is admissible in intuitionistic mathematics. Similar points have been voiced by Mannoury (1931), Borel (1952), van Dantzig (1955), and Wang (1958).

Such observations provide the immediate context for the view which Kreisel (1958) and Dummett (1959) labeled *strict finitism*. More specifically, Dummett used this term to describe a heterodox proposal about the foundations of mathematics which was promoted by the Russian mathematician and dissident Alexander Yessenin-Volpin starting in the late 1950s.[16] One tenet of the traditional form of finitism associated with the Hilbert program is that natural numbers are to be regarded as symbol types akin to unary numerals in a manner which does not presuppose that we understand them as comprising a completed infinite totality. Yessenin-Volpin went one step further in questioning our grounds for accepting the existence of natural numbers whose representations as unary numerals are too large to be constructed in practice.

Yessenin-Volpin's own presentation of strict finitism was informed not only by such observations about large numbers but also by various assumptions which he understood to have become unjustifiably entrenched in mathematical logic. For instance, he begins (1961, p. 202) by identifying four problematic aspects of what he dubs "traditional mathematics":

(T1) The categoricity of the natural number series $0', 0', 0'', \ldots$ (i.e. the assertion that this series is defined uniquely up to isomorphism).
(T2) The existence of the values of primitive recursive functions such as $x \cdot y$ or x^y for all arguments.

---

[15] See (Frege, 1884, §5, §89). See also (Frege, 1884, §7) and (Frege, 1903, §123) for Frege's use of large numbers in the formulation of his critiques of empiricist and formalist views of number.

[16] Yessenin-Volpin's (1961) own name for his program was *ultra-intuitionism*. Although Yessenin-Volpin is the rhetorical target of (Dummett, 1975), Dummett suggests that he was originally made aware of the paradox which he takes to refute strict finitism by an unnamed paper of Hao Wang. (See (Wang, 1990, p. xvii.) for one of Wang's own formulations.) On the other hand, most of what Dummett says to motivate strict finitism appears to have been informed by his prior work on Wittgenstein's philosophy of mathematics (cf., e.g., Dummett, 1959, pp. 181-183). Another of Dummett's apparent motives was to counter the impression that the instability he took to plague Yessenin-Volpin's notion of *constructibility in practice* might also infect the notion of *constructibility in principle* which figures in most presentations of intuitionism, inclusive of his own (e.g. Dummett, 1959, pp. 183-185, Dummett, 1975, p. 302).

(T3) The principle of mathematical induction.
(T4) The following form of *modus ponens*: from $\vdash \varphi$ and $\vdash \varphi \to \psi$ derive $\vdash \psi$.

In (Yessenin-Volpin, 1961) he then attempted to make use of the *denial* of these theses in the course of achieving his overall mathematical goal – i.e. that of providing a consistency proof for Zermelo-Fraenkel set theory. The details of his purported proof remain murky.[17] But what matters for present purposes is that Yessenin-Volpin's discussion of the notion of feasibility appears to have been motivated by the hope that such numbers could be used to define novel mathematical structures which he in turn hoped to employ in his consistency proof. He introduces such structures by first characterizing an *infeasible number* as one "up to which it is not possible to count" (1970, p. 5). He then introduces what he calls a *Zenonian situation* as a structure "in which the events of an infinite process are to be identified with the parts of a finite object" (1970, p. 8).

To illustrate how infeasible numbers might be understood to give rise to such structures, consider the predicate $F(x) = x$ *is a feasible number*. Suppose further that we regard natural numbers in the manner of traditional finitists – i.e. as symbol types corresponding to the unary numerals $O_F = \{\mathtt{0}, \mathtt{0}', \mathtt{0}'', \ldots\}$. Such expressions are naturally ordered by the relation $\prec_F$ induced by the process of "counting" in the sense of starting with 0 and iterating the operation $\sigma \mapsto \sigma'$. It then follows immediately that

(F1) $F(\mathtt{0})$

But if we have counted up to n by constructing a unary numeral $\overline{n}$ we could presumably also count up to $n+1$ by simply adjoining another token of $'$. This suggests that $F(x)$ also satisfies

(F2) $\forall x(F(x) \to F(x'))$

O_F is an example of what Dummett (1975) calls a *weakly infinite totality* meaning that the ordering $\prec_F$ on this set does not possess a last member. But it seems reasonable to say that O_F is also *weakly finite* in the sense that relative to this same ordering there exists a natural number which is not in O_F. Yessenin-Volpin proposed 10^{12} as an example of such a number.[18] Assuming that we are working over a system which includes terms for decimal numerals, then we can express the infeasibility of 10^{12} as

(F3) $\neg F(\underline{1000000000000})$

It would then appear to follow that although the sequence $\mathcal{F} = \langle O_F, \prec_F\rangle$ can be embedded into the finite initial segment $\langle\{0, \ldots, 10^{12}\}, <\rangle$ of the natural numbers in virtue of satisfying (F1) and (F2), it does not contain 10^{12} in virtue of (F3).

The crux of Dummett's critique of strict finitism is the claim that the existence of totalities like $\mathcal{F}$ which are simultaneously weakly finite and weakly infinite is logically incoherent. There can be no question whether Yessenin-Volpin embraced the existence of such structures.[19] But Dummett's subsequent argument amounts to little more than the

[17] For reconstructions and critiques see (Kreisel, 1967b), (Kreisel and Ehrenfeucht, 1967), letters 44, 45, 47, and 48 in the Gödel-Bernays correspondence (Feferman et al., 2003), (Geiser, 1974), and (Gandy, 1982). Although several of these sources suggest Yessenin-Volpin's proposal might be developed using the techniques of nonstandard analysis, he appears not to have discussed this in any of his published papers accessible in English.

[18] For as he observed, if we started counting from $\mathtt{0}$ in the manner envisioned above, it would take us over 20000 years to construct a unary numeral denoting this value were we able to adjoin one token of $'$ per second.

[19] E.g. "Let us consider the series F of feasible numbers, i.e. of those up to which it is possible to count. The number 0 is feasible and if n is feasible then $n < 10^{12}$ and then so n' also is feasible. And each feasible number can be obtained from 0 by adding $'$; so F forms a natural number series. But 10^{12} does not belong to F." (Yessenin-Volpin, 1970, p. 5)

observation that since $\mathcal{F}$ forms a sorites sequence, the premises (F1) - (F3) may be shown to be inconsistent in virtue of either the conditional or inductive sorites.

A systematic appraisal of how Yessenin-Volpin might have replied to this challenge is beyond the scope of the current paper. But even on the basis of the foregoing sketch, it should be evident that he recognized the relationship of his view to the sorites.[20] And it is equally clear that he would have been in a position to offer a principled reply to Dummett. He might, for instance, have resisted the reasoning of the inductive sorites in virtue of his rejection of (T3) and of the conditional form in virtue of his rejection of (T4). But he might also have attempted to argue directly for the cogency of Zenonian situations directly on the basis of his rejection of (T1). In particular, such structures might be understood to arise naturally as nonstandard models of theories in which, per Yessenin-Volpin's doubts about (T2), expressions such as x^y abbreviating the conventional definition of exponentiation fail to denote total functions.

In §3 I will further discuss the exigencies of attempting to develop a semantics for vague predicates based on the rejection of (T4) – i.e. the validity of *modus ponens* or related deductive principles. This approach appears most faithful to Yessenin-Volpin's own attempts in (1970; 1981) to develop a form of proof theory which directly takes into account the role of time, modality, and grammatical mood in mathematical practice.[21] As I will suggest, however, attempting to provide a faithful representation of such exigencies at either the level of the object language or the metatheory of a deductive system imposes severe limitations on the use of what we normally understand as formal methods.[22] On the other hand, the proposed semantics for vague predicates developed in §4 - §5 should be understood as inspired by the recognition that even if we are willing to accept the idealization embodied by (T4), it is still possible to account for Yessenin-Volpin's other doubts about (T1), (T2), and (T3) using classical model theory.

§3. The feasibilist theory of vagueness The focus of this section will be a philosophical view about vagueness which I will refer to as *feasibilism*. Such a view is motivated by the observation that the number of steps required to derive a contradiction from the premises of the conditional sorites in the manner of 1.3 is proportional to the magnitude of the number n_P. Feasiblists continue by pointing out that for the sorts of values of n_P cited in Table 1 the task of constructing a derivation like 1.3 will indeed be infeasible for the same reasons which appear to have led Yessenin-Volpin to identify 10^{12} as an infeasible number – i.e. human (or possibly even mechanical) agents cannot carry them out subject to practical limitations on resources such as time, memory, and attention.

The original formulation of such a view was motivated by early developments in computational complexity theory, some of which are linked historically to strict finitism.[23] Complexity theory seeks to provide a characterization not of *feasible numbers* but rather of *feasibly computable functions* – i.e. those whose values can be explicitly computed "in practice" as opposed to only in the "in principle" sense of classical computability theory in the sense of (e.g.) (Rogers, 1987). One motivation for such a study is to compare

[20] For instance in a 1958 letter to Brouwer, Yessenin-Volpin (2011) explicitly acknowledged that unless (T1)-(T4) are modified, (F1)-(F3) will lead to the "paradox of the heap".

[21] See, e.g., (Isles, 1981) for a partial reconstruction.

[22] These include not only the application of deductive rules such as *modus ponens*, but also the treatment of substitution, unification, and deciding the identity of terms which must be carried out algorithmically in a computational implementation of a proof system for first-order logic.

[23] See in particular the sequence (Kreisel, 1958), (Wang, 1958), (Yessenin-Volpin, 1961), (Parikh, 1971), (Cook, 1975), and (Buss, 1986, §6).

the difficulty of computing the values of different effectively computable (i.e. recursive) functions. For instance addition, multiplication, and exponentiation are all equally effective in the sense of computability theory. But although we would face little practical difficulty in constructing the decimal numerals representing the value of the sum $314159265358979+271828182845904$ or of the product $314159265358979 \times 271828182845904$, explicitly constructing a numeral representing the value of the power $314159265358979^{271828182845904}$ is beyond the computing abilities of humans and current (or even *foreseeable*) electronic computers. Complexity theory seeks to provide a mathematical framework for making such contrasts precise.[24]

Feasibilists draw on these observations by suggesting that the task of deriving a contradiction in the manner of derivation 1.3 may be likened to a computation whose intermediate steps correspond to the derivation of the statements $P^*(\underline{i})$ for $1 \leq i \leq n_P$. As the length of such a derivation is proportional to the extremal value n_P, one might reasonably suspect that the difficulty of deriving a contradiction from the soritical premises is proportional to that of constructing a unary numeral of length n_P. On the basis of this analogy, feasibilists go on to propose that we develop a theory of *feasible reasoning* – i.e. one which takes into account resource bounds which are faced by human reasoners the same way in which computational complexity theory takes into account the bounds faced by mechanical computing devices. They then suggest that the sorites does not represent a threat to the consistency of everyday reasoning since everyday reasoners will rarely if ever be capable of deriving a contradiction in practice in cases where the value of n_P is on the order of numbers which philosophers typically use to illustrate the paradox.[25]

An explicitly feasibilist theory of vagueness appears to have been first put forth by Parikh on the basis of the following observation:

> It is rare in ordinary life for people to make arguments which take a thousand or more steps. Perhaps this is the explanation of why we use vague predicates in daily life without any serious problems and still avoid difficulties which a logician might run into . . . ¶ . . . It is clear then that vague predicates will not have a logic which satisfies the following

[24] To briefly clarify the case just described in relation to this development, note that the number of digits in the base-b representation of n^m is $\lfloor m \log_b(n) \rfloor$. From this follows it that it would take more than 12000 years simply to even *write down* the values of the displayed power at 1000 digits per second. On the other hand, the intractability of many problems studied in complexity theory derives not from the infeasibility of simply reading the input or inscribing the output but rather due to the difficulty of the intermediate calculations which must be undertaken to yield a solution. This is evidenced, e.g., by the fact that there are 0-1-valued functions on the natural numbers which cannot be computed by a conventional Turing machine in time less than an exponential function of the size of their input. The existence of such functions can be demonstrated by the so-called *Time Hierarchy Theorem* in conjunction either with diagonalization of with the exponential-time completeness of various combinatorial decision problems – see, e.g., Arora and Barak, 2009 or Dean, 2015.

[25] Although feasibilism is not typically included amongst the major contemporary approaches to vagueness, it is grounded in concerns similar to those which have traditionally informed contextualist proposals. For instance Kamp (1981) proposes that the sentences comprising the intermediate steps in derivations like 1.3 should be evaluated not in isolation but relative to a context which takes into account the statements which have been evaluated earlier in the argument or have otherwise been made salient. A contextual parameter of this sort can be understood as a counter which dynamically moves along the derivation envisioned by 1.3 and whose position can be understood to measure a computational resource similar to that invoked by feasibilists. Such an interpretation is, for instance, made explicit by the use of dynamic semantics by Barker (2002).

> requirement: "If a rule of inference is permitted at all, then an arbitrarily large number of applications of such a rule is also allowed".
> (Parikh, 1983, p. 259)

Theories of vagueness based on similar considerations have been proposed more recently by Sazonov (1995), Gaifman (2010), and Magidor (2011). But Parikh's original formulation is of particular interest because the prior passage was informed by an earlier paper (Parikh, 1971) in which he not only anticipated Dummett's critique of strict finitism, but also defines the notion of an *almost consistent theory* which can be understood to contain the germ of a reply.

Such a theory S_τ is formulated over a language $\mathcal{L}_F = \mathcal{L}_a \cup \{F(x)\}$ extending that of first-order arithmetic with a new primitive predicate $F(x)$. I will assume that S_τ includes an arithmetical base theory Z extending the theory PA^-.[26] Almost consistent theories may also be formulated over stronger arithmetical theories containing induction for a class of formulas Γ – e.g. Parikh originally took Γ to be the class of $\mathcal{L}_a$-formulas (and hence $\mathsf{Z} = \mathsf{PA}$). But we assume that Γ does *not* include any formulas containing the predicate $F(x)$.

S_τ is additionally assumed to contain principles involving $F(x)$ which include at least

(F_0) $F(\mathsf{0})$ $\qquad$ (F_s) $\forall x(F(x) \rightarrow F(x'))$

$(F_<)$ $\forall x \forall y(F(y) \wedge x < y \rightarrow F(x))$ $\qquad$ (F_τ) $\neg F(\tau)$ where τ is a closed term of $\mathcal{L}_\mathsf{Z}$

Parikh's original intention was that $F(x)$ expresses the property of being a feasible number in Yessenin-Volpin's sense. (F_0), (F_s), and $(F_<)$ then respectively express that the class of such numbers includes 0, is closed under successor, and is closed downwards.[27] As $\neg F(\tau)$ expresses that this predicate also fails to hold of the number denoted by τ, it is clear that any theory satisfying the foregoing description will be inconsistent.

Note, however, that Z is formulated so that $F(x)$ cannot occur in its induction schema (presuming that it includes one at all). This can be understood as conforming to the traditional view that mathematical induction applies only to "definite" predicates of natural numbers. The meaning of this (and related) terms have been historically characterized it different ways. But it is presumably common to all parties in the debate that vague predicates – inclusive of both mathematical examples like *feasible number* and non-mathematical ones like *number of hairs on a bald man's head* – lack the relevant property of definiteness. For present purposes, I will take this precedent to serve as an adequate response to the inductive sorites in the form 1.4.[28]

---

[26] In addition to the axioms of Q (Robinson arithmetic), PA^- contains axioms stating that $+$ and $\times$ are associative and commutative, that $+$ distributes over $\times$, and that $<$ is a linear order. These additional assumptions constrain the possible order-types of models in manner which simplifies the formulation of several results below (see Kaye 1991, §2.1).

[27] Parikh (1971) originally also included closure conditions for addition and multiplication. These can be added without affecting any of the results below.

[28] The view that mathematical induction holds only for definite predicates has a long pedigree in mathematical logic which is independent of strict finitism. For instance, Frege (1879, §27) originally proposed that induction only be applied to "determinate" properties precisely to explain why its application to *number of grains in a non-heap* does not lead to a contradiction in the manner of the inductive sorites. (See, e.g., (Kamp, 1981) and (Williamson, 1994) for similar assessments within the philosophical literature on vagueness.) Related arguments to the effect that set-theoretic comprehension should only be applied to "definite" or "unambiguous" predicates can be traced to the concerns raised by Poincaré (1906), Russell (1908), and Zermelo (1929) about the relationship between impredicative definitions and the semantic paradoxes. See (Mostowski, 1950) for an illustration of how this concern bears

If this limitation is accepted, it might then appear that a "short" derivation of a contradiction in the manner of 1.4 is blocked. The question thus arises as to whether the task of deriving a contradiction must inevitably require a derivation of "infeasible length" similar to that envisioned by 1.3. But before this question can be addressed, we must be more specific about the form of the term τ. In §1 I suggested that if we are to have any hope of apprehending the derivation of the conditional sorites for the sorts of values depicted in Table 1, τ ought to be taken to be a positional numeral. Although I will return to the significance of this requirement in §4, it will be useful to begin by considering the case where τ is a unary numeral $\overline{m}$. If we let $\mathsf{S}_{\overline{m}}$ denote the almost consistent theory with $\tau = \overline{m}$, then one way in which a contradiction may be derived in $\mathsf{S}_{\overline{m}}$ is by successively instantiating (F_s) with the numerals $\overline{0}, \overline{1}, \overline{2}, \ldots$ and then applying *modus ponens* in the manner of derivation 1.3. As he length of such a derivation will still be on the order of m steps, a reasonable supposition would be that this is a necessary property of any derivation of a contradiction in $\mathsf{S}_{\overline{m}}$. It turns out, however, that additional insight into the structure of first-order reasoning allows for the derivation of a contradiction to be significantly "sped up".

The degree of speed-up which can be achieved depends to some extent on the arithmetical theory Z on which $\mathsf{S}_{\overline{m}}$ is based. On one extreme, Boolos (1991) observed that if we consider the theory $(F_0) + (F_s) + (F_<) + \neg F(\overline{m})$ with no additional arithmetical axioms, it is already possible to achieve exponential speed up – i.e. it is possible to derive a contradiction in a number of steps proportional to $\log_2(m)$. For instance in the Fitch system it is possible to obtain a contradiction in 90 steps if $m = 10000$ and in 126 steps if $m = 1000000$.[29] But if we assume that Z contains even a modest amount of arithmetic in the vicinity of PA^-, then it is possible to obtain superexponential speed-up of such a derivation in $\mathsf{S}_{\overline{m}}$ by employing the so-called *cut shortening method* introduced by Solovay (1976).[30]

Recall that the superexponential function is defined by $2^x_0 = x$ and $2^x_{y+1} = 2^{2^x_y}$ and that values of this function for small arguments are already astronomically large – e.g. $2^0_6 = 2^{65536}$. The Solovay method proceeds by defining a sequence of predicates $F_i(x)$ which describe subsets of $F(x)$ closed under functions whose rates of growth approach 2^x_y. Reasoning in Z from these definitions it is possible to derive statements of the form $F(\overline{2^j_k})$ in $\mathsf{S}_{\overline{2^j_k}}$ in a number of steps proportional to $c_0 + c_1 j + c_2 k$ for constants c_0, c_1 and c_2 which depend on the details of the proof system but which are independent of j and k themselves. In the Fitch system, for instance, it is possible to derive $F(\overline{2^j_k})$ – and hence also a contradiction – in under 300 steps where the value of $\overline{2^j_k}$ exceeds all of the values in Table 1.[31]

on the range of predicates which can be substituted into the induction schema in a predicative development of set theory.

29 The key observation is that by reasoning from the premises (F_0) and (F_s) alone we are able to uniformly derive $\forall x(F(x) \to F(x+\overline{2^n}))$ (where the latter term abbreviates $x^{\prime\ldots\prime}$ 2^n-times) in a number of steps which is *linearly* proportional to n. This allows us to derive $F(\overline{0}) \to F(\overline{2^m})$ in $6 \cdot m$ steps. It thus follows that by using $(F_<)$ we can obtain a contradiction in $\mathsf{S}_{\overline{m}}$ in at most $6 \cdot \lceil \log_2(m) \rceil + 6$ steps in the Fitch system.

30 The standard textbook presentation is (Hájek and Pudlák, 1998, III.3, V.5c) while an accessible exposition is also provided by (Sheard, 1998). The derivation described by Boolos (1991, pp. 702-704) makes use of a special case of this technique.

31 Approximately 100 of these (corresponding to the value of c_0) are required to derive lemmas about the relationships between the defined predicates $F_i(x)$ which could then be reused in subsequent proofs. It may additionally be observed that although the speed-up techniques currently under consideration reduces the plausibility of the original feasibilist approach to the sorites, the "sped up" proofs which they produce

Such observations attest to the fact that once the premises of the conditional sorites have been arithmetized, it is nowhere near as arduous to derive a contradiction from them as it might at first appear. Nonetheless, the following result demonstrates that there are limits to how much speed up can be achieved by the Solovay technique:

Theorem 3.1 (Parikh, 1971) *Let $m \in \mathbb{N}$ and $\mathsf{S}_{\overline{m}} = \mathsf{Z} + (F_0) + (F_s) + (F_<) + \neg F(\overline{m})$ where $\mathsf{Z} \supseteq \mathsf{PA}^-$. Then there is a primitive recursive function $f(x)$ defined uniformly on the size $|\mathcal{D}|$ of derivations $\mathcal{D}$ in $\mathsf{S}_{\overline{m}}$ such that for any formula not containing the predicate $F(x)$ and any $\mathsf{S}_{\overline{m}}$-derivation $\mathcal{D}$, if $\mathsf{S}_{\overline{m}} \vdash \varphi$ via $\mathcal{D}$ and $f(|\mathcal{D}|) < m$, then $\mathsf{Z} \vdash \varphi$.*

Dragalin (1985) obtained a version of this result based on the sequent calculus where the value of $f(x)$ depends only on the number of sequents $|\mathcal{D}|$ in $\mathcal{D}$. If we define 4^x_y analogously to 2^x_y – i.e. $4^x_0 = x$ and $4^x_{y+1} = 4^{4^x_y}$ – then Dragalin's bound is of the form $f(|\mathcal{D}|) = 4^{39|\mathcal{D}|}_{39|\mathcal{D}|}$. Since $\mathsf{S}_{\overline{m}}$ inconsistent, $\mathsf{S}_{\overline{m}} \vdash \bot$. But supposing we have taken $\bot$ to abbreviate some standard Z-refutable formula (e.g. $0 = 0'$), it follows that unless Z is itself inconsistent, any proof $\mathcal{D}$ of a contradiction in $\mathsf{S}_{\overline{m}}$ must be sufficiently long so that $f(|\mathcal{D}|) \geq m$.

Theorem 3.1 has thus sometimes been described as a means of demonstrating that there exist almost consistent theories which are "concretely consistent" in the sense that they are conservative over Z for "short" or "feasible" proofs.[32] For whatever number k we take to serve as a lower bound on the size of such a derivation, the result can be invoked to show that there exist *specific* natural numbers m such that the result of adjoining $\neg F(\overline{m})$ to the soritical premises $(F_0), (F_s), (F_<)$ does not lead to a contradiction over Z via a proof of size less than k. Note, however, that in order to ensure that no contradiction can be derived in this way in fewer than 1000 steps (per Parikh's original suggestion) requires that we choose $m \geq 4^{3900}_{3900}$. But not only is this number astronomically larger than those appearing in Table 1, it is also close to an optimal lower bound virtue of Theorem 3.1.

This observation in turn casts doubt on whether the form of feasibilism described at the beginning of this section provides a tenable response to many everyday instances of the conditional sorites. For at least for the sorts of values of n_P depicted in Table 1, the speed up techniques allow us to derive a contradiction in $\mathsf{S}_{\overline{n}_P}$ of length which we presuambly *can* construct in practice – e.g. 90 steps if $n_P = 10000$. The more recent feasibilist proposals of Sazonov (1995) and Magidor (2011) have been formulated to take into account the results just surveyed. Both authors contend that feasibilists ought to react to the speed-up results by rejecting the claim that *modus ponens* – a rule of inference which can be seen to figure centrally in the derivations produced by the Boolos and Solovay techniques – is an allowable rule in the course of soritical reasoning.[33] Such proposals may indeed be useful

for the values of n_P appearing in Table 1 still require a "medium" number of steps in the sense discussed in note 7. Although it may be feasible for a human subject to explicitly construct a proof of (say) 90 steps, we still presumably do not possess sufficiently many *non-mathematical* primitive terms to refer to the items even in one of the "sped up" sorites derivations. Many of the points framed in §1 about the necessity of using positional numerals to formulate the sorites thus still apply. This highlights how the central characteristic of neo-feasibilism position which will be developed below is not its reliance on metaphors about feasibility, but rather the proposal that that we acknowledge the non-transparency of positional numerals in the manner discussed in §1 and elaborated further in §4 and Appendix A.

32 See, e.g., (Gandy, 1982) and (Carbone, 1996).

33 It is easy to see that the exponential and superexponential speed-up techniques both make heavy use of the rules of conditional introduction (i.e. $\rightarrow I$) followed by conditional elimination (i.e. $\rightarrow E$ or *modus ponens*). They are thus highly *non-normal* if viewed as natural deduction derivations or far from being *cut-free* if viewed as sequent derivations. It is a consequence of the Normalization Theorem for natural deduction systems or the Cut Elimination Theorem for sequent systems that such "detours"

suggestions if our goal were to develop strict finitism in accordance with Yessenin-Volpin's proposed denial of thesis (T4). But outside the context of his critique of "traditional mathematics" such proposals appear to lack motivation. For as Dummett (1975) (and many others) have observed, not only is *modus ponens* a mainstay of mathematical and everyday reasoning, the fact that we take it to be a legitimate rule of inference seems central to why we regard the reasoning of the sorites as compelling in the first place.[34]

It is also useful to observe that even if feasibilists opt for an almost consistent theory based on a formulation of first-order logic without *modus ponens* or the cut rule, they must still presumably employ the predicate $G(x) = x$ *numbers the steps in a feasibly constructible derivation* as part of their account of feasible reasoning. But it would appear that $G(x)$ inherits whatever vagueness we take to be inherent in Yessenin-Volpin's notion of a feasible number. It thus follows that to the extent that feasibilism can be understood as providing a resolution to the sorites in the case of predicates such as *bald* or *non-heap*, it does so only at the expense of employing another vague predicate – which is itself apparently susceptible to a soritical argument – in its metalanguage.[35]

This is, of course, a feature common to many traditional theories of vagueness. For instance, supervaluationists speak of "admissible precisifications" of vague predicates without also giving a precise account of what makes a particular extension for a predicate like *bald* admissible, degree-theoretists speak of assigning real numbers as the degrees of truth to statements of the form $P(\mathsf{a}_i)$ without giving a precise account of what value should be assigned to the statement expressing that a particular man is bald, etc.[36] Theorists such as Sainsbury (1990) and Tye (1994) go even further in arguing that to the extent which such approaches rely on classical logic in their metatheory, they either misrepresent our understanding of vague predicates or require us to adopt a vague interpretation of the metatheoretic apparatus itself.

The approach developed in the following sections is explicitly intended to counter such claims. But it is also important to realize that feasibilism is unlike most traditional approaches to vagueness in that it seeks to provide a proof-theoretic rather than a model-theoretic account of the meaning of vague predicates. For since theories like $\mathsf{S}_{\overline{m}}$ are inconsistent in the "in principle" sense of ("classical") proof theory, they also do not possess interpretations in the sense of ("classical") model theory. Another drawback to such a view is thus that it rules out the possibility of providing an account of how predicates

in reasoning can be eliminated. But it also follows from well known results that the resulting "direct" proofs will be of size superexponential in the original derivations. Hence although it will still be possible to derive a contradiction via a normal or cut-free proof in $\mathsf{S}_{\overline{m}}$, it can be shown that such a proof will, like derivation 1.3, have length proportional to m. See (Carbone, 1996) for a detailed analysis of the combinatorial structure of the "sped up" proofs illustrating these points.

[34] Of course this claim is not entirely uncontentious in the philosophical literature on vagueness wherein the adoption of various forms of non-classical logics have been repeatedly suggested as the appropriate response to the conditional sorites (see, e.g., Cobreros et al., 2012 for a recent example). But to reiterate: one of the goals of the neo-feasibilist proposal is to demonstrate that a felicitous model of reasoning about vague predicates can be obtained which does not require undertaking the substantial modifications to the paradigm of model theoretic semantics which would be required to accommodate a non-classical definition of deductive or logical consequence.

[35] Although Dummett (1975) failed to anticipate speed up results of the sort surveyed above, he did suggest that the necessity of choosing a particular number k as an upper bound on the numbers falling under $G(x)$ (which he refers to as *apodictic numbers*) renders feasibilism vulnerable to a revenge-style objection in the form of a sortical argument involving $G(x)$. See (Magidor, 2011) for a reconstruction and critique of Dummett's argument from the standpoint of strict finitism.

[36] See Williamson (1994, §4.12) for an extended critique of such views on this basis.

like *bald* or *non-heap* could be assigned denotations in anything like the conventional sense of model theoretic semantics.[37]

§4. The neo-feasibilist theory of vagueness

4.1. A semantic reinterpretation The goal of this subsection is to lay the formal groundwork for the neo-feasibilist account of ordinal predicates and the conditional sorites which I will present in §4.2. A useful point of departure is the following observation which highlights just how close almost consistent theories come to being genuinely consistent:

Proposition 4.1 *Let* Z *be a consistent* $\mathcal{L}_a$*-theory extending* PA^- *and* c *a new constant symbol. Then the following* $\mathcal{L}_F$*-theories are consistent and conservative over* Z*:*

$$\mathsf{T}_\exists = \mathsf{Z} + (F_0) + (F_s) + (F_<) + \exists x \neg F(x); \quad \mathsf{T}_\mathsf{c} = \mathsf{Z} + (F_0) + (F_s) + (F_<) + \neg F(\mathsf{c})$$

$\mathsf{T}_\exists$ and T_c are similar to an almost consistent theory of the form $\mathsf{S}_{\overline{m}}$. But rather than stating that a *particular* unary numeral fails to fall under $F(x)$, they respectively state that *some* number fails to fall under $F(x)$ and that the denotation of a new constant symbol c fails to fall under $F(x)$. The consistency and conservativity of $\mathsf{T}_\exists$ and T_c follow immediately from the Compactness Theorem of first-order logic which can also be used to construct nonstandard models of arithmetical theories like Z.[38]

Recall that such a model is a structure $\mathcal{M} = \langle M, <^M, s^M, +^M, \times^M, 0\rangle$ which satisfies PA^- but which is not isomorphic to the standard model $\mathcal{N} = \langle \mathbb{N}, <, s, +, \times, 0\rangle$. It will also be useful to briefly recall some properties of such structures:

1) A nonstandard model of $\mathcal{M} \models \mathsf{Z}$ contains so-called *nonstandard integers* – i.e. $a \in M$ such that $\mathcal{M} \models \overline{n} < a$ for all $n \in \mathbb{N}$.

---

[37] Or alternatively, one might take the foregoing observations as a challenge to develop a semantics which is compatible with at least some of the aspects of the feasibilist approach to vagueness. This is the approach adopted by Gaifman who describes feasibilism (or as he labels it, the "brute force" approach to the sorites) as "indicting the right direction ... ¶ ... [but] unsatisfactory since it imposes a restriction on proofs, without making explicit the underlying contextual element of local usage" (Gaifman, 2010, pp. 16-17). He then develops a formal semantics for vague predicates which incorporates a notion he refers to as a *feasible context* – i.e. a classification on a given occasion of a subset of the objects in a potential sorites sequence for $P(x)$ which is sufficiently small (or sparse) such that it does not include a subsequence which spans the boundary between the extension of $P(x)$ and its anti-extension. Presuming that the soritical premises are always evaluated with respect to such contexts (which Gaifman takes to be induced by the use of terms similar to the constants a_i which appear in derivations like 1.1) the tolerance principle is thereby validated with respect to his semantics. Like the proof-theoretic approaches just surveyed, however, the notion of feasibility is itself used to formulate the semantics of the system he proposes, thus leaving the proposal susceptible to the revenge-style objection just discussed. But perhaps more seriously, Gaifman's proposal also fails to take into account that the speed up arguments allow us to derive a contradiction in a theory like $\mathsf{S}_{\overline{m}}$ in a manner which does not require us to evaluate the truth of $P(x)$ for a number primitive terms which might otherwise be judged infeasible relative to the size of m.

[38] The observation that a theory formed by adjoining axioms stating the existence of a nonstandard element to a standard system of arithmetic such as PA (or even True Arithmetic) results in a conservative extension appears to owe to Kreisel (1967a, pp. 166-168). Such results illustrate that the postulation of such elements has no effect on the properties of the ordinary arithmetical operations which can be proven to hold in such a theory. I will not make use of this fact until §7 below.

2) Although all nonstandard $\mathcal{M} \models \mathsf{Z}$ contain an initial segment $\overline{0}^M <^M \overline{1}^M <^M \ldots$ isomorphic to $\langle \mathbb{N}, < \rangle$, $<^M$ itself is not a well-ordering. In particular, below each nonstandard element $a_0 \in M$, there exists an infinite descending chain of nonstandard elements $\ldots <^M a_2 <^M a_1 <^M a_0$.
3) All nonstandard $\mathcal{M} \models \mathsf{Z}$ also contain so-called *proper cuts* – i.e. initial segments $I \subseteq M$ containing 0 (i.e. $0^M \in I$) which are closed under successor (i.e. $\forall x \in M(x \in I \rightarrow x' \in I)$), and closed downward (i.e. $\forall x(x \in I \wedge y <^M x \rightarrow y \in I)$ but for which $I \neq M$. It is easy to see that if I is a proper cut in $\mathcal{M}$ and $a \in M - I$, then a must be nonstandard.
4) Although there are uncountably many pairwise non-isomorphic nonstandard models of PA^- for all cardinalities $\kappa \geq \aleph_0$, the order-type of $<^M$ of such a model will always be of the form $\omega + (-\omega + \omega) \cdot \eta$ where η is the order-type of a dense linear ordering without endpoints. In the case where $\mathcal{M}$ is countable (i.e. $|M| = \aleph_0$), $<^M$ will thus consist of an initial segment isomorphic to $\mathbb{N}$, followed by countably many copies of the integers $\mathbb{Z}$ which are themselves ordered in the manner of the rationals $\mathbb{Q}$.

Let us now see how these properties may be used to prove Proposition 4.1:

Proof. To see that $\mathsf{T}_\exists$ and T_c are consistent, let $\mathcal{M} \models \mathsf{Z}$ be nonstandard and let $I \subsetneq M$ be a proper cut. Now define an expansion $\mathcal{M}^F$ of $\mathcal{M}$ to $\mathcal{L}_F$ by letting $F^M = I$. Note that I contains 0^M, is closed downward, and closed under successor. Thus $\mathcal{M}^F$ satisfies (F_0), (F_s) and $(F_<)$. Also note that since I is proper, there must exist $a \in M - I$. Since $a \notin F^M$, we hence also have that $\mathcal{M}^F \models \exists x \neg F(x)$. And if we additionally interpret $\mathsf{c}^{M_F} = a$, we have $\mathcal{M}^F \models \neg F(\mathsf{c})$ as well.

To see that these theories are also conservative over Z, suppose that φ were an $\mathcal{L}_\mathsf{Z}$-sentence such that $\mathsf{T}_\exists \vdash \varphi$ or $\mathsf{T}_\mathsf{c} \vdash \varphi$ but $\mathsf{Z} \nvdash \varphi$. By the Compactness Theorem there exists a nonstandard model $\mathcal{M} \models \mathsf{Z} + \neg\varphi + \{\overline{n} < \mathsf{c} \mid n \in \mathbb{N}\}$. We can extend $\mathcal{M}$ to a model $\mathcal{M}^F$ of $\mathsf{T}_\exists$ or T_c by again letting F^M be a proper cut I in $\mathcal{M}$ and interpreting c as $\mathsf{c}^{M_F} \in M - I$. As φ is an $\mathcal{L}_\mathsf{Z}$-sentence, we must also have $\mathcal{M}^F \models \neg\varphi$, contradicting the assumption that φ is provable in $\mathsf{T}_\exists$ or T_c. □

4.2. The neo-feasibilist account of the conditional sorites Proposition 4.1 suggests another way in which the linguistic formulations of sorites arguments are sensitive to how we designate an extremal value for an ordinal predicate $P(x)$. For the proof just given shows that inconsistency is not achieved by simply asserting that there exists a value of which $P^*(x)$ fails to hold. Nor is it even achieved by stating that $P^*(x)$ fails to hold of a number denoted by a fixed term c unless $\mathsf{c} = t$ can be proven in our background theory for some $\mathcal{L}_a$-term t.

This again highlights the role which the values of n_P given in Table 1 play in our apprehension of the conditional sorites. Although the magnitude of plausible extremal values may depend on the predicate $P(x)$ under consideration, it is evident that the *specific* numbers which are chosen typically do not represent the outcome of anything like a process of careful deliberation or experimentation. For on the one hand, it is evident that the soritical reasoning does not depend on the number theoretic properties particular to n_P (e.g. whether it is even or odd, prime or composite, etc.). And on the other, such values rarely seem to be chosen with any regard to finding a *least* upper bound on the number satisfying $P^*(x)$ (presuming that one exists).[39] Rather they appear to be selected to serve

[39] Dummett (1975) attempted to find such a value in the course of formulating his original instance of the phenomenal sorites. But as I will argue in §5, the specification of an extremal value in such cases is generally much less straightforward than in the simple cases of the conditional sorites considered in §1.

as upper bounds which are "safe" in the sense that they are unlikley to be contested – a fact which is attested to by the preponderance of "round numbers" in decimal notation in Table 1.

The observation that we tend to invest little in the specific extremal values which we use to formulate the sorites may appear innocuous on its own. But once it is acknowledged, Proposition 4.1 may then be viewed as exhibiting a sense in which the premises of the conditional sorites describe not a contradictory situation but rather a consistent one in which the term denoting the extremal value n_P is assigned an infinite denotation in a nonstandard model and the arithmetization of the soritical predicate is itself interpreted as a proper cut. My aim for the rest of this paper will be to argue that the structural properties of models of this sort provide non-trivial insights into many issues which are often discussed in regard to natural language vagueness.

I argued in §1 that our ability to derive a contradiction from the premises of the conditional sorites in the case of extremal values most often employed to illustrate the paradox depends on certain aspects of how we understand and reason about numbers whose magnitudes are exemplified by those in Table 1. In such cases, the juxtaposition of unary and positional notations seems to be an ineliminable aspect of the formulation of the premises. For note that if we wish to eliminate the ellipsis from the formulation of derivation of 1.3, we must employ a universally quantified statement of the form (F_s) which employs a symbol for the successor function (i.e. $'$). This in turn allows us to generate the unary numerals $\mathtt{0}, \mathtt{0}', \mathtt{0}'', \ldots$ by which we naturally envision counting up to the unary numeral $\overline{n}_P$. But on the other hand, in order to describe the relevant values of n_P linguistically, we must often employ a positional numeral of the form $\underline{n}_P$.

Such a juxtaposition might reasonably lead one to suspect that in order to derive a contradiction from the soritcal presmises we will need to construct a unary numeral $\overline{m}$ such that we can prove $[\![\underline{m}]\!]_b = [\![\overline{m}]\!]_u$.[40] But recall that the value of a positional numeral of the form $\mathtt{d}_0 \ldots \mathtt{d}_{k-1}$ is given by $\sum_{i=0}^{k-1} [\![\mathtt{d}_i]\!]_b \cdot b^i$. Thus if $\underline{m}$ is (e.g.) a decimal numeral of k digits, a co-denoting unary numeral of the form $\mathtt{0}'^{\cdots\prime}$ will contain 10^{k-1} symbols. I labeled the supposition that there will always exist such a sequence as principle (D) in §1.

In order to comprehend the formula giving the value of a positional numeral we must presumably understand the basic properties of addition, multiplication, and exponentiation, as well as the recursion required to handle the summation in the definition just

[40] If we were able to derive a statement of the form $\underline{m} = \overline{m}$ in our background theory Z, then having also arrived at the conclusion $F(\overline{m})$ (say in the manner of derivation 1.3), we would then be able to conclude $F(\underline{m})$ and hence also a contradiction. This possibility is blocked in the current setting by treating $\underline{m}$ as a constant symbol about which no additional mathematical principles are assumed – i.e. as analogous to the symbol $\mathtt{c}$ in Proposition 4.1. (See Nelson, 1986, §32 for a related proposal.) A more realistic formalization is described in Appendix A wherein $\underline{m}$ is taken to denote a member of the domain on binary numerals which provides a better characterization of how we reason conjointly with unary and positional notations. A different approach is adopted by Boolos (1991) who presents a theory in which both binary and unary numerals are treated as terms of the same type. In this theory it is possible to restore a contradiction either by directly proving $\underline{m} = \overline{m}$ or indirectly by developing a formulation of the sorites for binary numerals using a version of the Solovay technique. It should be kept in mind, however, that Boolos's theory is grounded in the assumption that the appropriate means of expressing that a positional numeral π and unary numeral υ denote the same number is via a conventional identity statement of the form $\pi = \upsilon$. Relative to the approach developed in Appendix A, such an equality should instead be expressed by a statement of the form $[\![\pi]\!]_b = [\![\upsilon]\!]_u$ where $[\![\cdot]\!]_b$ and $[\![\cdot]\!]_u$ are the denotation functions respectively for binary and unary numerals of the sort discussed in §1. I will argue below that the other mathematical principles required for formulating the sorites do not commit us to the fact $\forall\pi\exists\upsilon([\![\pi]\!]_b = [\![\upsilon]\!]_u)$.

given. This suggests that (D) already expresses a more sophisticated proposition than it might at first appear. But even if we understand the relevant concepts, it seems like an extra premise is still required to ensure that there will always exist an expression of the form $0^{\prime\cdots\prime}$ whose length corresponds to the value of $\underline{m}$. For as I observed in §1, it is easy to explicitly construct positional numerals for which we have no hope of constructing a co-denoting unary numeral in practice. This in turn might lead one to suspect that the existence of such an expression is not guaranteed by the basic arithmetical principles which I have argued are implicit in our understanding of the other soritical premises. And if it turns out that we are not obliged to accept (D), then one might additionally imagine that a consistent interpretation of the soritical premises can be obtained by taking the constant c in the formulation of $\mathsf{T_c}$ to correspond to a positional numeral $\underline{m}$ for which we are willing to allow that there might not exist a co-denoting unary numeral.[41]

Whether it is in fact coherent to imagine such a situation depends not only on which principles we take to govern our reasoning about positional numerals but also on how we elect to formalize them. For before we can even formulate (D) as a mathematical proposition, we must find some means of formally expressing statements about numerals and their denotations. Unsurprisingly, the relevant notions can be formalized in relatively strong theories like PA wherein we can interpret binary numerals as finite 0-1 sequences via a standard coding function which maps such sequences into single natural numbers. We can then prove that the formula for $[\![\cdot]\!]_b$ given above defines a total function such that $[\![\underline{m}]\!]_b$ gives the length (and hence also the value) of a co-denoting unary numeral $\overline{m}$ which itself can be coded in a similar manner as a finite sequence of 1s.[42]

It is evident, however, that such a treatment collapses the distinction between how we operate with unary and positional numerals in practice which I argued in §1 is crucial to our apprehension of many instances of the sorites. Additionally, nothing which has been said thus far suggests that either our understanding of the soritical premises or our everyday reasoning about numbers is grounded in a theory which is as strong as PA.

At present there is little philosophical consensus as to what kind of theory best reflects the practices of counting, calculation, and measurement which underlie our everyday application of number theory, inclusive of the sort of situations envisioned by instances of the sorites. Similarly, logicians and computer scientists have proposed several sorts of theories for formalizing conjoint reasoning about unary and positional numerals of the sort which commonly arises in our application of numerical algorithms.[43] The task of selecting an arithmetical theory Z on which to base a theory for formalizing the conditional sorites which is felicitous to our everyday reasoning is thus not only delicate but almost certainly underdetermined. It may thus come as a surprise that natural formalizations of (D) are formally independent of a wide range of plausible candidates.

One such theory is the fragment of PA known as $\mathrm{I}\Delta_0$ consisting of PA^- together with the induction scheme for *bounded formulas* – i.e. those containing only bounded quantifiers of the form $\exists x < t\varphi(x)$ and $\forall x < t\varphi(x)$ where t is an $\mathcal{L}_a$-term not containing x. Parikh (1971) originally introduced $\mathrm{I}\Delta_0$ as a candidate for formalizing the "anthropomorphic" standpoint

---

41 In fact such a possibility has repeatedly been contemplated by strict finitists – e.g. (Yessenin-Volpin, 1961, p. 204), (Parikh, 1971, p. 494), and (Nelson, 1986, p. 173).

42 For instance, suppose $(\cdot)$ denotes a function coding finite $\mathtt{0}$-$\mathtt{1}$ sequences $\mathtt{b}_1 \ldots \mathtt{b}_n$ as natural numbers – e.g. $(\mathtt{b}_1 \ldots \mathtt{b}_n) = 2^{b_1}3^{b_2} \ldots p_n^{b_n}$ where p_n is the nth prime. Then it is easy to see that the function defined by $f((\mathtt{0})) = 0$, $f((\mathtt{1})) = 1$, $f((\pi \cdot \mathtt{0})) = 2 \times f((\pi))$, and $f((\pi \cdot \mathtt{1})) = 2 \times (\pi) + 1$ which gives the value of a coded binary numeral (and, e.g., outputs 0 if its input is not the code of a binary sequence) is primitive recursive and hence provably total in PA.

43 Systems which have been considered in these regards range from primitive recursive arithmetic (e.g. Tait 1981), to predicative fragments of Frege arithmetic (e.g. Heck 2014), to systems similar to $\mathrm{I}\Delta_0$ and V^1 discussed in Appendix A (e.g. Ganea 2010).

about the foundations of mathematics described by Wang (1958).[44] A fundamental fact about this theory is the following:

Theorem 4.2 (Parikh, 1971) *Suppose that* $I\Delta_0 \vdash \forall x \exists y \varphi(x,y)$ *where* $\varphi(x,y)$ *is a bounded formula. Then there exists a* $\mathcal{L}_a$*-term* $t(x)$ *such that* $I\Delta_0 \vdash \forall x \exists y < t(x)\varphi(x,y)$.

$\mathrm{I}\Delta_0$ is sufficiently strong to prove simple facts about the properties of positional notation such as the fact that the value of the binary numeral $\mathtt{10\ \ldots 0}$ (n $\mathtt{0}$s) is 2^n. It may also be shown that the graph of the exponentiation function is definable by a bounded formula.[45] A consequence of Theorem 4.2 is thus that $\mathrm{I}\Delta_0$ does not prove the totality of functions such as $\mathsf{2}^x$ whose order of growth is not bounded by a polynomial. It thus also follows that no reasonable formalization of (D) is provable in $\mathrm{I}\Delta_0$.

Similar results hold for a wide range of theories extending PA^- with bounded axioms. In Appendix A I will describe a theory V^1 wherein unary and positional numerals are treated as distinct logical sorts. This theory provides a natural medium for formalizing our practices for computing with such expressions conjointly.[46] But although (D) may be expressed naturally in the language of V^1, it is underivable from its axioms. Nonetheless, the general neo-feasibilist approach to vagueness does not depend on the details of this theory itself. I will thus continue here under the assumption that the arithmetical theory Z relative to which the soritical premises are formulated is a bounded theory to which an appropriate analog of Theorem 4.2 applies.

Suppose we now let δ be the formalization of (D) in the language of Z (an explicit example is again provided in Appendix A). The foregoing observations entail that it is consistent to simultaneously accept Z while remaining agnostic about or even rejecting δ. Such a denial is equivalent to asserting the existence of positional numerals for which there exists no co-denoting unary numeral. In the case of ordinal predicates such as *bald* or *non-heap*, the neo-feasibilist theory of vagueness thus holds that the conditional sorites may be resolved by regarding the terms employed to denote extremal values for the relevant predicates as nonstandard integers in the sense illustrated by Proposition 4.1.

Such a view provides a means of treating our willingness to endorse the soritical premises as an indication that we apprehend them in a manner which does not commit us to a formal contradiction. Neo-feasibilists go on to propose that such an apprehension is grounded in the recognition that no inconsistency would arise in our everyday reasoning were we to treat the terms employed to denote extremal values for many ordinal predicates as if they denoted infinitely large values. Proposals in this vicinity have again been repeatedly explored by strict finitists and proponents of related views.[47] But neo-feasibilists take an additional step in suggesting that we interpret vague predicates in natural language relative to nonstandard models of appropriately weak systems of formal arithmetic.

---

[44] Wang used this term to refer to a foundational standpoint similar to that described by Yessenin-Volpin wherein feasibility is treated as a fundamental notion. He suggested that such a view is motivated by the observations that the reduction of "decimal notation to stroke notation entails . . . a considerable decrease of the range of numbers which we can actually handle" as well as the role of mathematical induction in proving "that there exists a unique decimal or stroke notation for each positive integer" (Wang, 1958, p. 474).

[45] See (Cook and Nguyen, 2010, III.3.3.).

[46] V^1 is a natural theory in which to formalize such reasoning as its provably total functions correspond to those which can be computed in polynomial time. Since polynomial time computability is often taken to coincide with the pre-theoretical notion of feasible computability (see, e.g., Dean, 2015), this further suggests that V^1 is also a reasonable candidate for characterizing Wang's "anthropomorphic standpoint".

[47] See, e.g., (van Dantzig, 1955, p. 275), (Robinson, 1966, §2.10-11, §10.1-5), and (Nelson, 1977).

There are, of course, a number of reasons why a proposal of this sort may still continue to appear ill-motivated. Perhaps most seriously, the approach just described still treats as paradigmatic instances of the conditional sorites with extremal values which are "large" in the sense discussed in §1–§3. Although such values are typical of those employed to motivate strict finitism as a philosophy of mathematics, much of the contemporary philosophical and linguistic literature on vagueness has focused on a form of the sorites which does not appear to involve large numbers at all (at least *prima facie*). Before considering several further consequences and applications of the neo-feasibilist approach to vagueness in §6, it will thus be useful to first consider some of the cases in question in the form of the phenomenal sorites.

§5. The phenomenal sorites and measurement theory An important form of the sorites which we have not yet considered is what Hyde (2011) labels the *phenomenal* variant of the argument. The origins of this form can be traced to Dummett's original discussion of *phenomenal* (or *observational*) *predicates* such as *looks red*, *tastes sour*, or *sounds loud*. Such predicates $P(x)$ are assumed to be associated with phenomenal properties which vary along a continuum $\mathcal{C}_P$ – e.g. of hue, sourness, or volume – which give rise to so-called *P-indiscriminable* elements – i.e. pairs of distinct objects whose gradation with respect to $\mathcal{C}_P$ is so small that they cannot be distinguished perceptually. Such pairs are thus often characterized by the fact that they either both satisfy $P(x)$ or both fail to satisfy $P(x)$. If we let $\sim_P$ denote the relation of *P-indiscriminability*, then the relevant property can be expressed by

$$(\mathrm{Tol}_P)\quad \forall x \forall y[(P(x) \wedge x \sim_P y) \rightarrow P(y)]$$

This is the so-called *tolerance principle* for $P(x)$ and is often presented as being partially constitutive of the meaning of observational predicates – cf., e.g., Dummett (1975), (Graff, 2001).

In order to formulate an instance of the phenomenal sorites we must identify a set of objects $O_P = \{o_0, \ldots, o_{n_P}\}$ with the following properties: i) O_P is linearly ordered by the relation $\prec_P$ corresponding to the position of the objects in the continuum $\mathcal{C}_P$ (e.g. of *looking more red*, *tasting more sour*, or *sounding louder*); ii) o_0 clearly falls under $P(x)$; iii) o_{n_P} clearly fails to fall under $P(x)$; and iv) for all $0 \leq i \leq n_P$, objects o_i and o_{i+1} are indiscriminable with respect to $P(x)$ i.e. – $o_i \sim_P o_{i+1}$. The underlying idea is, of course, that we can identify such a sequence of objects by subdividing $\mathcal{C}_P$ into a sequence of P-indiscriminable points or regions ordered by $\prec_P$ so that the point or region on one end of the ordering definitely satisfies $P(x)$ and the point or region on the other end definitely fails to satisfy $P(x)$.

One way of describing such a situation linguistically is to again introduce a sequence of subscripted constant symbols $\mathsf{a}_0, \ldots, \mathsf{a}_{n_P}$ to denote the objects o_i and also treat $\sim_P$ as a symbol in the object language. Upon taking these steps, Hyde's (2011) formulation of the the phenomenal sorites takes the following form:

5.1 The phenomenal sorites

$$\begin{array}{l} P(\mathsf{a}_0) \\ \mathsf{a}_0 \sim_P \mathsf{a}_1 \\ \mathsf{a}_1 \sim_P \mathsf{a}_2 \\ \vdots \\ \mathsf{a}_{n_P-1} \sim_P \mathsf{a}_{n_P} \\ \hline \therefore \quad P(\mathsf{a}_{n_P}) \end{array}$$

Although this argument schema is again not itself a formally correct derivation of a contradiction, it may be transformed into one by adjoining the premise (Tol_P) and the supposition $\neg P(\mathtt{a}_{n_P})$ stating that the $\prec_P$-greatest member of the sequence O_P does not fall under $P(x)$. As in the case of the conditional sorites, the formulation of such a derivation again requires the introduction of constant symbols $\mathtt{a}_i$ denoting the objects $\mathtt{o}_i$ in the order $\prec_P$. And again, the length of such a derivation will again be proportional to the value n_P. One might thus think that an adequate neo-feasibilist response to the phenomenal sorites could be developed in parallel to the strategy described in §4 – i.e. by maintaining that we can consistently interpret $\mathtt{a}_{n_P}$ (or a corresponding numeral) as a nonstandard element of the sequence $o_0, o_1, \ldots$ which cannot be reached as the result of a feasible number of indiscriminable transitions from o_i to o_{i+i}.

Such an approach might be effective if the sorts of values of n_P for which the phenomenal sorites was typically formulated were similar to those used for the ordinal predicates given in Table 1. But when philosophers describe instances of this argument, they often employ much smaller numbers for the extremal value n_P. For instance Graff (2001), Raffman (1994), and Keefe (2000) respectively use 30, 50, and 100 for *looks red*, and in his original example Dummett (1975) used the value 4 (see note 49 below). Such numbers clearly encroach into the class of those up to which we can explicitly construct and survey unary numerals, count, or dependably perform the relevant number of applications of *modus ponens*. Likening such numerals to terms denoting nonstandard integers would thus strain the analogies on which the original feasibilist proposal seeks to build past the point of collapse.

Such observations notwithstanding, it is evident that mathematical principles still play a substantial role in how we apprehend common instances of the phenomenal sorites. In order to appreciate why this is so, first note that predicates like *looks red* are unlike ordinal predicates such as *bald* in that they lack sortal units which may be used to identify a sorites sequence by a simple process of counting.[48] But in order to formulate the above argument, we must still identify a sufficiently long sequence of regions or objects which differ imperceptibly in hue in order to describe a sorites sequence for this predicate. By assumption, this is not a task which can be performed by unaided visual inspection alone. For if we cannot uniformly distinguish the objects O_P will have no means of dependably ordering them in accordance with $\prec_P$. Nonetheless, we are all familiar with the sorts of auxiliary processes which might be employed to construct such a sequence – e.g. we can superimpose a grid of uniformly sized squares onto a color spectrum or use a spectrophotometer to sort paint chips by hue.

Not only do these particular processes involve empirical measurement, but this seems to be a central feature of the examples of the phenomenal sorites which are most often discussed. For instance, a ruler is required to construct the sort of grid just described or even to uniformly point to equidistant (but otherwise phenomenally indistinguishable) patches along a projected spectrum in the appropriate order. And similar points would seem to apply (*mutatis muntandis*) to other cases of gustatory, olfactory, and auditory perception – e.g. in order to prepare an appropriate sequence of sample items for *tastes sour* we need to measure their pH, for *sounds loud* we need to measure the amplitude of an appropriate class of sound waves, etc.[49]

---

[48] On this point see, e.g., Frege (1884, §54).

[49] An additional case in point is provided by the example Dummett (1975) originally used to illustrate the phenomenal sorites. In this instance, $P(x)$ is intended to hold of time intervals (measured in seconds) such that the minute hand of a clock appears at an angle perceptually indistinguishable from vertical. Dummett imagines that the movement of the minute hand is not constant but rather that it suddenly "jumps" $0.25°$ at the end of every second within an interval of 10^{-5} seconds. As he also supposes that the value of the just noticeable difference for angular position is $1°$, he thereby describes a version of the phenomenal sorites with $n_P = 4$. (Dummett offers no empirical justification for his choice of these values. But the set up of the phenomenal sorites allows even for cases

As noted in §1, the importance of measurement to the semantics of vague predicates is now widely recognized in the linguistic literature on vagueness wherein the techniques of *measurement theory* (in the sense of Krantz et al., 1971) are standardly employed. This framework provides an account of the circumstances under which we are justified in using a given mathematical structure $\mathcal{A}^* = \langle A^*, R_1^*, \ldots, R_n^* \rangle$ (where R_i is an r_i-ary relation on A) to represent an empirical structure $\mathcal{A} = \langle A, R_1, \ldots, R_n \rangle$ in terms of the properties of mappings $f : A \to A^*$ which preserve the structure of the empirical relations $R_i(\vec{x})$. Adapting standard terminology slightly, I will say that such a mapping is a *measurement function* for $\mathcal{A}$ with respect to the *scale* $\mathcal{A}^*$ just in case for all $a_1, \ldots, a_n \in A$, if $R_i(a_1, \ldots, a_{r_i})$ then $R_i^*(f(a_1), \ldots, f(a_{r_i}))$.

Consider, for instance, the familiar case where we start out by taking the empirical domain A to be an appropriately prepared set of paint chips varying in hue between orange and red. In this case, we can measure the chips in terms of the *wavelength* of the light which they reflect measured by a fixed unit such as *nanometers* (where 1nm $= 10^{-9}$m). An instrument such as a spectrophotometer for measuring reflected wavelength may thus be mathematically represented as a measurement function $f : A \to \mathbb{R}^+$ which for each chip returns a positive real number which can be interpreted as the wavelength in nanometers of the light which it reflects.[50]

A straightforward means of constructing a sorites sequence for the predicate *looks orange* would thus be to use the instrument to select a sequence of chips $o_1, \ldots, o_{n_P}$ from A such that $f(o_0) < f(o_1) < \ldots < f(o_{n_P-1}) < f(o_{n_P})$, $f(o_0) = 600$nm (for orange), $f(o_{n_P}) = 680$nm (for red), and for all $0 \leq i < n_P$ $|f(o_i) - f(o_{i+1})| < \alpha_P$ where α_P is the so-called *just noticeable difference* for color perception – i.e. the smallest difference in wavelength which is visually discriminable. Widely cited empirical estimates of α_P are in the vicinity of 5nm for the red-orange portion of the spectrum (e.g. Wright and Pitt,

with $n_P = 3$ – i.e. the smallest number of objects required to demonstrate a failure of transitivity for $\prec_P$.) Note, however, that in order to use such a case to formulate an instance of derivation 5.1 still requires that we are able to linguistically describe the magnitudes which are in the field of $P(x)$. But it seems that we would be at a loss as to how to do this in the case Dummett describes were we not able to draw on our prior practices of chronographic and angular measurement (possibly in conjunction with a physical description of the operation of the clock in question) to describe magnitudes which are smaller than our perceptual tolerances allow us to discriminate.

[50] It would appear that the empirical magnitudes like wavelength associated with phenomenal predicates typically (and perhaps always) lead to what are known as *ratio* or *interval* scales. In the former case we require that the empirical domain contain objects $e, u \in A$ respectively corresponding to a zero magnitude for which $f(e) = 0$ and a choice of unit for which $f(u) = 1$ and that it also is possible to define a concatenation-like binary operation $\circ$ definable on $\mathcal{A}$ which (together with $\prec$) satisfies the operations of an ordered abelian group which is preserved under $f(x)$ in the sense that $f(x \circ y) = f(x) + f(y)$. This is known as a case of *extensive measurement*. It is also characteristic of ratio scales that if $f(x)$ is an admissible scale for a structure $\mathcal{A}$, then so is $\beta \cdot f(x)$ for any $\beta \in \mathbb{R}^+$. From this it follows that both differences $|f(x) - f(y)|$ and ratios $f(x)/f(y)$ between the measurements of objects $x, y \in A$ are empirically meaningful – e.g. it is meaningful to say both that the difference between a wavelength of 200nm and 100nm is greater than the difference between wavelengths of 50nm and 10nm and also that the ratio between the latter pair is larger than that between the former. Only the latter is true of interval scales such as temperature which lack an empirically determined zero point. But in both cases the numerical value $f(x)$ itself is determined only up to the choice of unit in the sense illustrated in note 6 above. For complete definitions of these notions and further discussion see (Krantz et al., 1971).

1934). This means that it is indeed possible to prepare a sorites sequence for *looks orange* with $n_P = 20$ by selecting the objects in O_P such that $|f(o_i) - f(o_{i+1})| = 4\text{nm}$.[51]

The precise values of such psychologically determined constants are often not mentioned alongside presentations of the phenomenal sorites. Nonetheless, it seems that anyone familiar with the basic setup of the argument will appreciate that a measurement process similar to the one just described must be undertaken in order to select or prepare the objects denoted by the constants $\mathsf{a}_1, \ldots, \mathsf{a}_n$ and to ensure that they are ordered correctly with respect to $\prec_P$. But once such details are made explicit, it is also reasonable to re-examine some of the presuppositions of the argument, inclusive of the fact that P-indiscriminability should be understood as a primitive rather than a defined relation.

It may be noted in this regard that the predicates which are employed to linguistically formulate the phenomenal sorites are prototypically *gradable adjectives* – i.e. terms such as *hot* or *expensive* which can be modified by adverbs such as *fairly*, *very*, or *extremely*. In the case of phenomenal predicates like *tastes sour* or *sounds loud*, our use of such modifiers presumably tracks how we locate objects relative to perceptual phenomenal $\mathcal{C}_P$ using the relation $\prec_P$ – e.g. if we regard x, y, and z respectively as *fairly*, *very*, and *extremely sour*, then we will presumably order them as $x \prec_P y \prec_P z$. Moreover the gradability of such terms also seems to underlie our willingness to speak of phenomenal predicates as "admitting to degrees" – e.g. of one object looking red or tasting sour to a *greater* or *lesser degree* than another.

If we start out from this perspective, then it is also natural to treat our acknowledgement of the existence of pairs x, y of P-indiscriminable objects not as a brute fact but rather one which arises because we intuitively understand that the "distance" between x and y relative to $\mathcal{C}_P$ may be "smaller than" an appropriate perceptual threshold u. Our intuitive understanding of $\mathcal{C}_P$ *as a continuum* – i.e. a structure which shares at least some of the order or topological properties of the real numbers – presumably also disposes us to acknowledge the possibility of regarding u as a point or interval within $\mathcal{C}_P$ itself. For although we may have no means of ostending such an item, its existence is presumably entailed by (and perhaps even constitutive of) our recognition of the finite sensitivity of our perceptual faculties relative to $\mathcal{C}_P$.

It is, of course, precisely intuitions of this sort which the application of measurement theory to the semantics of phenomenal predicates is intended to capture. Suppose, for instance, that we start out by assuming that there exists a measurement function $f_P : A \to \mathbb{R}^+$ which maps the field A of $P(x)$ into the positive real numbers in a manner which is compatible not only with our use of the associated relation $\prec_P$ but also with respect to the rest of our other informal talk of "degrees of P-ness". In this case, we may posit the existence of a just noticeable difference for the $P(x)$ as a value $\alpha_P \in \mathbb{R}^+$.

Upon so doing, it then becomes possible to regard the phenomena of indiscriminability as arising in the manner described above – i.e. in virtue of the existence of pairs of distinct

[51] When described in such simplistic terms, it may appear that it is being claimed here that our ability to set up instances of the relevant argument scheme requires that we identify phenomenal color properties with reflectance properties of physical objects (and similarly for other phenomenal properties). Indeed, it seems as though talk of "imperceptible differences" on which the phenomenal sorites trades presuppose a broadly non-dispositional view of phenomenal properties. But further reflection on the cases in question suggests that all that is required is that we are able to isolate some continuous magnitude m such that reported phenomenal changes in $P(x)$ are correlated with changes in m under appropriate "standard conditions". In particular, preparation of the envisioned scenarios does not seem to require that we subscribe to a view on which it is possible to "reduce" color properties to physical ones.

objects x, y such that the absolute difference of $f(x)$ and $f(y)$ is less than α_P. But in this case, we may then define the relation of P-indiscriminability as follows:[52]

(ID) $x \sim^*_P y$ if and only if $|f_P(x) - f_P(y)| < \alpha_P$

In addition to allowing for the formalization of degree-theoretic talk about phenomenal predicates, another benefit of introducing terms denoting measurement functions into the language in which we formalize the phenomenal sorites is that it allows us to avoid the need to introduce constant symbols to refer to the members of a potential sorites sequence.[53] For suppose we take $\mathtt{a}$ to refer to an object which clearly falls under $P(x)$ (e.g. an object which definitely looks orange), $\mathtt{b}$ to refer to an object which clearly does not fall under $P(x)$ (e.g. an object which definitely looks red) and α_P to refer to a real number which serves as a just noticeable difference for $P(x)$ with respect to the measurement function $f(x)$. Then if $\mathcal{A}^* = \langle A^*, \prec^*_P \rangle$ is a scale for $\langle A, \prec_P \rangle$ with respect to $f(x)$, we may refer to the members of a sorites sequence for the predicate $P^*(x)$ defined by $\forall x(P(x) \leftrightarrow P^*(f(x)))$ by choosing a real number $0 < \epsilon < \alpha$ such that the terms $f(\mathtt{a}), f(\mathtt{a}) + \epsilon, f(\mathtt{a}) + \overline{2}\epsilon, \ldots$ (where $\overline{n}\epsilon$ abbreviates $\epsilon + \ldots + \epsilon$ n-times) denote an increasing sequence of real numbers representing the degrees of a sequence of objects which are linearly ordered by $\prec_P$ and such that adjacent pairs are P-indiscriminable.

If we additionally assume that our object language contains a predicate $A(x)$ to denote the domain of $\mathcal{A}$ and a function symbol $f(x)$ to denote the measurement function itself, we can now also reformulate (Tol_P) in terms of our definition of $\sim^*_P$ via (ID) as follows:

($\mathrm{Tol}^{\alpha_P}_P$) $\forall x \in A \forall y \in A[(P^*(f(x)) \wedge x \sim^*_P y) \rightarrow P^*(f(y))]$

We can now also formalize the assumptions enumerated above as $P^*(f(\mathtt{a})), \neg P^*(f(\mathtt{b})), f(\mathtt{a}) < f(\mathtt{b}), \mathtt{0} < \epsilon < \alpha$. In parallel to principle ($F_<$) for ordinal predicates, it is also reasonable to assume that $P^*(x)$ is closed downwards – i.e.

($P^*_<$) $\forall x \in A \forall y \in A[(f(x) < f(y) \wedge P^*(f(y))) \rightarrow P^*(f(x))]$

At this point, one might think that we have assumed enough to mimic the reasoning of derivation 5.1 in the mathematical domain. For by successively instantiating (Tol^{α}_P) with adjacent pairs of values in the sequence $f(\mathtt{a}), f(\mathtt{a}) + \epsilon, \mathtt{a}, f(\mathtt{a}) + \overline{2}\epsilon, \ldots$ we may conclude $P^*(f(\mathtt{a})), P^*(f(\mathtt{a}) + \epsilon), P^*(f(\mathtt{a}) + \overline{2}\epsilon), \ldots, P^*(f(\mathtt{a}) + \overline{n}\epsilon)$ for any $n \in \mathbb{N}$. It might then be thought that in order to derive a contradiction, it is sufficient to choose n sufficiently large

---

[52] Independent motivation for adopting such an account of indiscriminability derives from the fact that the relations defined in this manner correspond to those generated by defining $x \sim_R y =_{\mathrm{df}} \neg(R(x, y) \vee R(y, x))$ where $R(x, y)$ satisfies the axioms of a finite *semi-order* (cf. Scott and Suppes 1958). Such indiscriminability relations – which are reflexive and symmetric but not transitive – were originally introduced by Luce (1956) in order to study failures of preference transitivity generated by cases similar to the phenomenal sorites. The applications of semi-orders to the linguistic study of vagueness has recently been pursued by van Rooij (2011).

[53] For note that even though the length of sorites sequences for a phenomenal predicate like *looks red* need not be of "infeasible length", we still presumably do not possess (say) 20 proper names for shades of red or orange which allow us to uniformly and inter-subjectively refer to the members of the sort of "medium length" (cf. note 7) sequence described above in the appropriate order. Thus if we were not able to employ a measurement function to refer to the members of the sequence in the manner described above, we would be forced to result to referring to them by descriptions such as "the ith chip from the left" or similar demonstrative expressions (whose applications would presumably need to be subscripted). So even in this case it would seem that at least some of the mathematical principles which I described in §1 must still be counted as presuppositions of the phenomenal sorites.

such that $f(\mathtt{b}) < f(\mathtt{a}) + \overline{n}\epsilon$. For we could then conclude $P^*(f(\mathtt{b}))$ by $(P^*_<)$ and hence also a contradiction relative to our premise $\neg P^*(f(\mathtt{b}))$.

But not only does the reasoning just described contain an important lacuna, it is also possible to consistently adjoin the axioms just described to the theory of first-order analysis A consisting of the axioms of an ordered field, together with the schema of Dedekind Completeness for $\mathcal{L}_A = \{\mathtt{0}, \mathtt{1}, +, \times, <\}$ formulas.[54] Before presenting a formal statement and proof of this fact it will be useful to observe where the gap in the foregoing argument occurs. Note in particular that in order to conclude that there is a natural number n such that the value denoted by $f(\mathtt{b}) < f(\mathtt{a}) + \overline{n}\epsilon$ exceeds that denoted by $f(\mathtt{b})$ tacitly assumes that the field of numbers $\mathcal{A}^*$ into which the measurement function $f(x)$ maps the constants $\mathtt{a}$ and $\mathtt{b}$ is *Archimedean* – i.e. that for every $x, y \in A^*$, such that $x <^* y$, there exists an $n \in \mathbb{N}$ such that $y <^* x +^* \ldots +^* x$ (n-times).

It is well known that A possesses nonstandard models in which this property fails. This suggests that it should be possible to find a consistent interpretation of the premises just described by interpreting the just noticeable difference α_P as a number α^*_P which is *infinitesimally small* with respect to the distance between $f(\mathtt{a})$ and $f(\mathtt{b})$.

Proposition 5.1 *Let $\mathcal{L} = \mathcal{L}_A \cup \{A(x), P(x), P^*(x), \prec_P, f(x), \mathtt{a}, \mathtt{b}, \alpha\}$ and consider the $\mathcal{L}$-theory $\mathsf{U}^\alpha_\mathtt{b}$, consisting of the axioms of A, axioms asserting that $\prec_P$ is a linear order, $(\mathrm{Tol}^{\alpha_P}_P), (P^*_<)$, and the following additional principles involving the interpretation of the measurement function $f(x)$: $\forall x \in A \forall y \in A(x \prec_P y \leftrightarrow f(x) < f(y)), \forall x \in A(P(x) \leftrightarrow P^*(f(x))$, $P^*(f(\mathtt{a}))$, $\neg P^*(f(\mathtt{b}))$, $f(\mathtt{a}) < f(\mathtt{b})$, $\mathtt{0} < \alpha$. Then U_α is consistent and conservative over A.*

Proof. Let $\mathcal{R}^* = \langle \mathbb{R}^*, \mathtt{0}^*, \mathtt{1}^*, +^* \times^*, <^* \rangle$ be a nonstandard model of A.[55] In order to show that $\mathsf{U}^\alpha_\mathtt{b}$ is consistent, it suffices to show how we can extend this model to a structure $\mathcal{R}^\dagger$ in which the symbols $P(x)$, $P^*(x)$, $A(x)$, $f(x)$, $\mathtt{a}, \mathtt{b}$ and α_P are interpreted to satisfy its non-mathematical axioms and in which the mathematical symbols of $\mathcal{L}$ retain their interpretation from $\mathcal{R}^*$. Let $A^\dagger$ be the non-negative part of $\mathbb{R}^*$, $f^\dagger(x)$ be the identity function on $A^\dagger$, and $\prec^\dagger_P$ be the restriction of $<^*$ to $A^\dagger$. Next let $\mathtt{a}^\dagger$ and $\mathtt{b}^\dagger$ be standard real numbers in $\mathbb{R}^+$ – i.e. $\mathtt{a}^\dagger, \mathtt{b}^\dagger <^* \mathtt{n}^*$ for some $n \in \mathbb{N}$ – for which we also assume that $\mathtt{a}^\dagger <^* \mathtt{b}^\dagger$ and $|\mathtt{a}^\dagger -^* \mathtt{b}^\dagger| >^* \mathtt{1}^*$. Let $P^\dagger$ and $P^{*\dagger}$ both be the interval $\{x \in \mathbb{R}^\dagger : \mathtt{a}^\dagger \leq x <^*$

[54] The latter states that each first-order formula $\varphi(x)$ (with parameters) which defines a bounded set possesses a least-upper bound. Note that in the theory U_α considered below, the completeness schema is assume to *not* extended to formulas including the predicate $P^*(x)$. This conforms with the policy of excluding vague terms from mathematical schema whose history and motivation were discussed in note 28.

[55] Such a model can be obtained by either applying the Upwards Löwenheim-Skolem theorem to the standard model $\mathcal{R}$ of A with domain $\mathbb{R}$ or a familiar ultraproduct construction (see, e.g., Kanovei and Reeken, 2004). Note that any such model contains both nonstandard integers $a >^* \overline{n}$ for all $n \in \mathbb{N}$ and also so-called *infinitesimals* $b <^* 1/\overline{n}$ for all $n \in \mathbb{N}$ – e.g. the latter can be regarded as the reciprocals of the former. Another feature of these construction of $\mathcal{R}^*$ is that they validate the so-called *Transfer Principle* – i.e. every first-order statement about real numbers which can be formulated in $\mathcal{L}_\mathsf{A}$ which holds in the standard model $\mathcal{R}$ also holds in $\mathcal{R}^*$. This includes all instances of Dedekind Completeness where the set in question is defined by a formula of $\mathcal{L}_\mathsf{A}$. But this property does *not* hold in $\mathcal{R}^*$ for so-called "external sets" of the sort exemplified by the interpretation of the predicate $P^*(x)$ below which is bounded but does not possess a least upper bound in $\mathcal{R}^*$. The possibility of using non-Dedekindian extensions of the reals to interpret indiscriminable differences in this manner suggests that there is also a straightforward extension of the approach described in this section to treat cases of what Weber and Colyvan (2010) refer to as the *continuous* and the *topological sorites* (both of which depend on the existence of appropriate least upper bounds).

$(\mathtt{b}^\dagger -^* 1/\overline{n}^*)$ for some $n \in \mathbb{N}\}$. Finally, let $\alpha_P^\dagger >^* \mathtt{0}^*$ be an *infinitesimal* – i.e. $\alpha_P^\dagger <^* 1/\overline{n}^*$ for all $n \in \mathbb{N}$. It is evident that the model $\mathcal{R}^\dagger$ defined in this manner satisfies all of the additional axioms of $\mathsf{U}_{\mathtt{b}}^\alpha$ with the potential exception of $(\mathrm{Tol}_P^{\alpha_P})$. To see that this principle is also satisfied, suppose that $u, v \in \mathbb{R}^*$ are such that $\mathcal{R}^\dagger \models P^*(f(u)) \wedge u \sim_P^* v$. The truth of the second conjunct implies that $|u - v|^* <^* \alpha_P^\dagger$ and since $\sim^*$ is symmetric we may assume without loss of generality that $u <^* v$. In this case, the truth of the first conjunct implies that there is an $n \in \mathbb{N}$ such that $u <^* (\mathtt{b}^\dagger -^* 1/\overline{n}^*)$. But since $v - u <^* \alpha_P^\dagger$, it hence follows that $v <^* u + \alpha_P^\dagger <^* (b - 1/\overline{n}^*) + \alpha_P^\dagger$. But then $v <^* b - 1/\overline{2n}^*$ since $\alpha_P <^* 1/\overline{n}$. But then $\mathcal{R}^\dagger \models P^*(f(v))$ as desired. The proof of the conservativity of $\mathsf{U}_{\mathtt{b}}^\alpha$ over A proceeds similarly to Proposition 4.1. □

Much like Proposition 4.1, Proposition 5.1 attests to the fact that the premises of the phenomenal sorites form a consistent set when they are formulated in a manner which does not insist that the constant symbol α_P which we have introduced to denote a just noticeable difference for $P(x)$ is interpreted as a *specific* real number. For as the preceding proof attests, even if we insist that the images of the objects denoted by $\mathtt{a}$ and $\mathtt{b}$ themselves denote fixed real numbers, we may still interpret α_P as a number which is infinitesimally small with respect to the difference between these values.

In the case of the conditional sorties, I defended the use of a similar kind of nonstandard interpretation for the extremal value $\underline{n}_P$ in virtue of the fact that paradox mongers typically attach little significance to their precise values. On the other hand, the constant α_P has been introduced without any prior discussion as to how we might go about identifying and denoting a real number corresponding to a just noticeable difference for a given phenomenal predicate.

But this assumption seems appropriate in the case of phenomenal predicates. On the one hand, our willingness to speak of *degrees* of orangeness, sourness, loudness, etc. is at least suggestive that some aspects of the measurement framework are implicit in our everyday reasoning about phenomenal predicates. But on the other hand, our use of such terminology does not on its own equip us with knowledge of the *specific* numerical values which can measure the left and right endpoints of a potential sorites sequence for such a predicate or that of the relevant just noticeable difference – e.g. of the values 600, 680, and 5 in the case of *looks orange*. For not only will these values depend on an arbitrary choice of unit – e.g. nanometers, pH units, or decibels – they will also depend on empirical and psychophysical facts about our perceptual faculties of which we will typically be unaware.

The proof of Proposition 5.1 makes clear that the mere supposition that there is a means of interpreting our comparative and degree theoretic discourse about such predicates in this manner does not commit us to the fact that the standard real numbers are the only mathematical structure which can serve this purpose. In particular, the structure $\mathcal{R}^*$ employed in the foregoing proof is a *non-Archimedean field* – i.e. it is not the case for all $x, y \in \mathbb{R}^*$, such that $x <^* y$, there exists an $n \in \mathbb{N}$ such that $y <^* x +^* \ldots +^* x$ (n-times). It is thus not isomorphic to the standard model of analysis $\mathcal{R} = \langle \mathbb{R}, <, +, \times, 0, 1 \rangle$.

As noted above, this observation explains the gap in the failed derivation of a contradiction in the theory $\mathsf{U}_\alpha^{\mathtt{b}}$. But Proposition 5.1 also illustrates how the factors which initially seem to motivate us to accept the tolerance principle (Tol_P) do not require us to accept that there is any *particular* real number which measures a just noticeable difference for $P(x)$. For in parallel to the formulation of the theory T_{E} in Proposition 4.1, it is also evident that the proof of Proposition 5.1 demonstrates that the variant of (Tol_P) which merely asserts the existence of such a value – i.e.

$$(\mathrm{Tol}_P^\exists)\ \exists z > 0 \forall x \in A \forall y \in A[(P^*(f(x)) \wedge |f(x) - f(y)| < z) \rightarrow P^*(f(y))]$$

– may be consistently adjoined to the other axioms of $\mathsf{U}_\alpha^{\mathtt{b}}$ even in the case where we have assumed that the values $f(\mathtt{a})$ and $f(\mathtt{b})$ are fixed.

This observation does not contradict the fact that it is possible to describe concrete cases of the perceptual phenomenon underlying the phenomenal sorites using a "small" value for n_P – e.g. a sequence of 20 paint chips as described above. But it does suggest another means of replying to the paradox monger's insistence that such situations give rise to linguistically mediated contradictions which our acceptance of the soritical premies obliges us to confront. For as we have seen, the adoption of the measurement theoretic understanding of phenomenal predicates already provides a principled reason to reject the use of a primitive indiscriminability relation to describe our reasoning about such terms. But this framework also enables us to acknowledge that we understand such predicates to be tolerant in the sense of ($\mathrm{Tol}_P^{\exists}$) while also failing to affirm ($\mathrm{Tol}_P^{\alpha_P}$) for any specific value of α_P. This remains so even if a non-infinitesimal value for α_P has been determined by a prior psychophysical experiment. For if we are unaware of this value ourselves, we will have no reason to accept ($\mathrm{Tol}_P^{\alpha_P}$) and thus be under no compulsion to endorse the reasoning standing behind the corresponding version of derivation 5.1 in which $\sim_P$ is understood as defined by (ID).[56]

Viewed in this light, the premises of the phenomenal sorites may again be understood as an implicit description of a nonstandard model in which we understand just noticeable differences to correspond to infinitesimal magnitudes. The availability of such an interpretation also demonstrates that it is consistent with the other premises of the phenomenal sorites to assume that the indiscriminability relation $\sim_P^*$ defined by (ID) may be a transitive relation when α_P is infinitesimal.[57] As a consequence, the model described may fail to validate another principle which one might argue is implicit in our apprehension of phenomenal predicates – i.e. that the "distance" between any pair of distinct items along a given phenomenal continuum may always be bridged by finitely many transitions between indiscriminable items of the appropriate sort.

Since that the sort of interpretation under consideration fails to allow a finite number of indiscriminable changes in hue, sourness, loudness, etc. to "add up" to a discriminable change, it might still be claimed the extension of the neo-feasibilist approach to the phenomenal sorites fails to provide a felicitous account of our apprehension of the situations in question. Note, however, that the formulation of this complaint itself requires the use of the notion *finite*. But as finitude is presumably a *mathematical* notion, this in turn suggests that the assumptions underlying the potential objection may not be of the same "naive" character as the other soritical premises. After first surveying several additional features and applications of neo-feasibilism in §6, I will return to address this issue in §7.

§6. Applications and comparisons

6.1. The line drawing sorites The *line drawing sorites* is another argument scheme which is often presented as an alternative form of the conditional sorites. Hyde's (2011) presentation is as follows:

5.1 The line drawing sorites

$P(\mathsf{a}_0)$

[56] The question remains what human subjects would do if they were confronted with the fact that a particular non-infinitesimal value of α_P describes their own perceptual tolerances. But although this is clearly an empirical matter, there seems to be no abiding reason to assume that they would continue to accept ($\mathrm{Tol}_P^{\alpha_P}$) rather than using the reasoning of derivation 5.1 to conclude via *modus tollens* that *this* principle is indeed *not* constitutive of how they understand $P(x)$.

[57] This can be seen, for instance, by observing that if $\sim_P^*$ is defined by (ID), then the interpretation of this symbol in the model constructed in the proof of Proposition 5.1 is indeed transitive.

$$\frac{\neg\forall i P(\mathsf{a}_i)}{\therefore \exists i(P(\mathsf{a}_i) \wedge \neg P(\mathsf{a}_{i+1}))}$$

Rather than constituting a contradiction, the line drawing sorites is supposed to eventuate in the counterintuitive consequence that there is a sharp boundary between the extension of a soritical predicate $P(x)$ and its anti-extension. Such a boundary corresponds to the existence of an object o_i in the envisioned sorites sequence $o_o, \ldots, o_{n_P}$ such that $P(x)$ holds of o_i and $\neg P(x)$ holds of o_{i+1} for some $0 \leq i \leq n_P$. But presented in the form just given the argument again relies on a suppressed mathematical premise – i.e. the *least number principle* for $\mathcal{L}_{P*}$

LNP($\mathcal{L}_{P^*}$) For all $\mathcal{L}_{P^*}$-formulas $\varphi(x)$, $\exists x\varphi(x) \rightarrow \exists y(\varphi(y) \wedge \forall z < y\neg\varphi(z))$

If this principle is adopted, a replacement for 5.1 can now be constructed which us to shows that $\exists x(P^*(x) \wedge \neg P^*(x'))$ is derivable in the theory $\mathsf{PA}^- + \mathrm{LNP}(\mathcal{L}_{P^*}) + P^*(0) + \neg\forall x P^*(x)$. But unlike the repair of derivation 1.2 using the induction principle $\mathrm{Ind}(\mathcal{L}_P)$, the envisioned derivation relies on several other arithmetical axioms of PA^- – e.g. the fact that 0 doesn't have a predecessor. It is also easy to see that over such a base theory the least number principle for $\mathcal{L}_{P*}$ is *equivalent* to $\mathrm{Ind}(\mathcal{L}_P)$ – i.e. the theories $\mathsf{PA}^- + \mathrm{LNP}(\mathcal{L}_{P^*})$ and $\mathsf{PA}^- + \mathrm{Ind}(\mathcal{L}_{P^*})$ have the same theorems. Thus if we continue to hold that mathematical induction may not be applied to vague predicates, then we should also presumably hold that the least number principle is also not applicable to them as well. This blocks the derivation of $\exists x(P^*(x) \wedge \neg P^*(x'))$ from $P^*(0) + \neg\forall x P^*(x)$ over PA^- in much the same way that rejection of $\mathrm{Ind}(\mathcal{L}_{P^*})$ blocks derivation 1.4.

The question thus arises whether the theory $\mathsf{U}_\exists = \mathsf{Z} + P^*(0) + \neg\forall x P^*(x) + \neg\exists x(P^*(x) \wedge \neg P^*(x'))$ is consistent. One might conjecture that since $P^*(x)$ is stipulated to hold of 0 but to fail to hold of some other number, then it must be possible to find the point of transition from the extension to the anti-extension of $P^*(x)$ by searching through the numbers $0, 1, 2, \ldots$ so as to find a witness demonstrating $\exists x(P^*(x) \wedge \neg P^*(x'))$. But it is easy to see that this idea cannot be turned into a formal proof. For observe that it is possible to construct a model of $\mathsf{U}_\exists$ in the manner of Proposition 4.1 – i.e. by letting $\mathcal{M} \models \mathsf{Z}$ be nonstandard, and then interpreting $P(x)$ in $\mathcal{M}$ as a proper cut $I \subsetneq M$.[58]

Such a model satisfies $P^*(0)$ because all cuts contain 0 and it satisfies $\neg\forall x P^*(x)$ because $M - I$ is non-empty. What is more interesting is why $\mathcal{M}$ fails to satisfy $\exists x(P^*(x) \wedge \neg P^*(x'))$. Suppose we define the sets $P^+ =_{\mathrm{df}} \{a \in M \ : \mathcal{M} \models P^*(a)\} = I$ – i.e. the extension of $P(x)$ in $\mathcal{M}$ and $P^- =_{\mathrm{df}} \{a \in M \ : \mathcal{M} \models \neg P^*(a)\} = M - I$ – i.e. the anti-extension of $P(x)$ in $\mathcal{M}$. Since I is closed under successor, it follows that P^+ contains no $<^M$-greatest element. Similarly, P^- will be closed under predecessor and thus contain no $<^M$-least element. And from this it follows that there does not exist an element $a \in M$ such that $a \in P^+$ and $a +^{\mathcal{M}} 1 \in P^-$, or equivalently $\mathcal{M} \not\models \exists x(P^*(x) \wedge \neg P^*(x'))$.

[58] A related structural feature of such interpretations may also be taken to bear on the view about vagueness traditionally referred to as *nihilism*. Nihilists such as Unger (1979) argue that since we are often inclined to accept the soritical premises $P^*(0)$ and $\forall x(P^*(x) \rightarrow P^*(x'))$, we are hence obliged to accept the conclusion that $P^*(\overline{n})$ for arbitrary $n \in \mathbb{N}$ as well. On this basis they argue that since all objects in a potential sorties sequence for $P(x)$ are such that the corresponding number falls under $P^*(x)$, no objects can fall under $\neg P(x)$. Nihilists thus treat the paradox as a *reductio* of the premise that (e.g.) there are heaps or non-bald men. Note, however, in the model $\mathcal{M}$ of $\mathsf{U}_\exists$ just described the extension of $\neg P^*(x)$ must be non-empty, from which it follows that $\mathcal{M} \models \exists x\neg P^*(x)$. For this reason neo-feasibilists should not be characterized as "embracing" the sorites along with nihilists.

Such a model thus provides a consistent interpretation of the premises of derivation 5.1 in which there is not a sharp boundary between between P^+ and P^- in the sense sense defined above. On the other hand, P^+ and P^- are clearly defined so that they partition the domain of $\mathcal{M}$ disjointly – i.e. $P^+ \cup P^- = M$ and $P^+ \cap P^- = \emptyset$. It hence follows that $\mathcal{M}$ satisfies $\forall x(P(x) \vee \neg P(x))$ in accordance with classical logic while also satisfying the tolerance principle $\forall x(P(x) \rightarrow P(x'))$ and thus $\neg\exists x(P(x) \wedge \neg P(x))$. On the other hand, any proper initial segment of the standard model $\mathcal{N}$ must have the form $\{0, 1, 2, \ldots, k\}$ for some $k \in \mathbb{N}$. It thus follows that if we used such a set to interpret $P(x)$, $k+1$ must fall in the corresponding anti-extension for $P(x)$ from which it follows that $\mathcal{N} \models \exists x(P(x) \wedge \neg P(x'))$.

It may also be noted that in $\mathcal{M}$ the transition from P^+ to P^- will occur after infinitely many discrete steps along the ordering $<^M$. This is possible because the order-type of a nonstandard $\mathcal{M} \models \mathsf{PA}^-$ is $\omega + (-\omega + \omega) \cdot \eta$. If P^+ is interpreted as a proper cut, then the structure $\langle P^+, <^M \rangle$ will thus be isomorphic to either ω or to $\omega + (-\omega + \omega) \cdot \zeta$ where ζ is the order-type of a proper initial segment of a dense linear order without endpoints. But in neither case will there be a $<^M$-greatest element of P^+ or an $<^M$-least upper bound of P^+ in M. This illustrates in a more precise way how the order-structure of a nonstandard model can serve to explicate the sense in which it is commonly maintained that vague predicates lack sharp boundaries.[59]

6.2. Borderline cases Another phenomena which is widely discussed in relation to vagueness is that of *borderline cases* – i.e. objects to which a given predicate $P(x)$ neither definitely applies nor definitely fails to apply. Theorists such as Fine (1975) take the existence of such objects to be criterial of what it means for $P(x)$ to be vague. But even for ordinal predicates of the sort we have been considering, it is difficult to find published values analogous to those given in Table 1 which are taken to be characteristic examples – e.g. of the number of hairs on the head of a borderline bald man or the number of grains which comprise a borderline heap.

The paucity of numerical data notwithstanding, it still appears that arithmetical principles play an important role in our understanding of borderline cases. Suppose, for instance, that m serves as an accepted extremal value for $P(x)$ – e.g. $n_P = 10000$ in the case of *non-heap*. Since in such a case we will presumably accept $P^*(\mathsf{0})$ and reject $P^*(\underline{n}_P)$, one obvious place to look for borderline cases is the range of numbers which are simultaneously as far from 0 and n_P as possible – i.e. in *the middle* of the sequence $0, \ldots, n_P$. One way in which we might refer to such values is via the use of fractions such as $n_P/2$ or $n_P/2 \pm j$ which denote the values in the vicinity of half of n_P. For note that it would seem that there is an intrinsic instability in using values of the forms k or $n_P - k$ when k is "small" in comparison to n_P – say $k = 50$ in the case $m = 10000$.

In order to see why this is so, suppose that we adopt the standard assumption that we can express what it means for a number to be a borderline case of $P^*(x)$ by using a

[59] The sense in which P^+ and P^- fail to have a sharp boundary in $\mathcal{M}$ may be partially likened to the reason why the intersection of the sets $A_0 = \{x \in \mathbb{Q} : x \leq \sqrt{2}\}$ and $A_1 = \{x \in \mathbb{Q} : \sqrt{2} \leq x\}$ is empty – i.e. in neither case is there an object in the domain of the structures which falls on the border. In the case of A_0 and A_1, however, this gap is naturally "filled in" by moving to the Dedekind completion of $\mathbb{Q}$ – i.e. $\mathbb{R}$. In the nonstandard case, the analogous step would be to attempt to regard the cut P^+ itself as a new *object* $[P^+]$ corresponding to the least upper bound of its members in an expanded model. But in this case, we cannot have $[P^+] \in P^+$ as otherwise $[P^+] + 1 \in P^+$ in contradiction to $\forall x(x \in P^+ \rightarrow x \leq [P^+])$. This suggests that the use of nonstandard interpretation to model the non-existence of sharp boundaries should be regarded as a *sui generis* proposal which is not readily accounted for using structures with standard order-types as analogies.

propositional operator $\mathbb{D}\varphi$ with the intended interpretation φ *is definitely true*. We may now introduce another operator $\mathbb{B}P^*(x) =_{df} \neg\mathbb{D}P^*(x) \wedge \neg\mathbb{D}\neg P^*(x)$ which expresses that for x to be a borderline case of $P^*(x)$ is for neither $P^*(x)$ nor $\neg P^*(x)$ to be definitely true. It is traditionally taken to be counterintuitive that the transition from being a definite non-heap to being a borderline non-heap could occur with addition of a single grain. Thus when $P(x)$ is vague, it is also conventional to assume that the property expressed by $\mathbb{D}P^*(x)$ is itself tolerant – i.e. $\forall x(\mathbb{D}P^*(x) \rightarrow \mathbb{D}P^*(x'))$. But now suppose that we are willing to assert that 50 grains form a borderline heap – i.e. $\mathbb{B}P^*(\underline{50})$. From this it follows that $\neg\mathbb{D}P^*(\underline{50})$. It will then follow that we can formulate another instance of the conditional sorites for $\mathbb{D}P^*(x)$ based on the premises $\mathbb{D}P^*(\mathsf{0})$ and $\forall x(\mathbb{D}P^*(x) \rightarrow \mathbb{D}P^*(x'))$ with $n_{\mathbb{D}P} = 50$.

But note that it is also often assumed in such cases that $\mathbb{D}\neg P^*(x)$ will itself be tolerant. Parallel observations thus apply to values expressed in the form $n_p - k$ in virtue of our ability to formulate a "short" sorites sequence for this predicate downwards from n_P. Other factors equal, it would thus appear that we reach the conclusion that it will be most stable to locate borderline cases for $P(x)$ at a point which is simultaneously as far away from 0 and n_P as possible – i.e. approximately halfway between 0 and n_P.[60]

These observations again do not entail that there is a unique form of numerical expression which we must use for designating borderline cases. But what they do suggest is that once we have fixed an extremal value n_P for a particular soritical predicate $P(x)$, it is natural to locate such cases by speaking of various *fractional* values of the number n_P. The specific language $\mathcal{L}_a$ over which we have been working does not contain a functional symbol for division which allow us to talk about such values directly. However, for each $k > 0$ we can readily express in $\mathcal{L}_P$ that $F(x)$ holds of the result of rounding x/k down to the next smallest integer by a formula of the form $F(\lfloor x/k \rfloor) =_{\mathrm{df}} \exists q < x \exists r < \overline{k}(\overline{k} \times q + r = x \wedge F(\overline{k} \times q))$.

We can now record the following.

PROPOSITION 6.1 *Let $\mathsf{T}_{\underline{m}} = \mathsf{Z} + (F_0) + (F_s) + (F_<) + \neg F(\underline{m})$ where $\mathsf{Z} \supseteq \mathrm{I}\Delta_0$, $\underline{m}$ is treated as constant symbol, and $k > 1$. Then presuming Z is consistent, $\mathsf{T}_{\underline{m}} \nvdash F(\lfloor \underline{m}/k \rfloor)$ and $\mathsf{T}_{\underline{m}} \nvdash \neg F(\lfloor \underline{m}/k \rfloor)$ – i.e. the statement $F(\lfloor \underline{m}/k \rfloor)$ is* formally independent *of $\mathsf{T}_{\underline{m}}$.*[61]

Proof. It suffices to construct models $\mathcal{M}_0 \models \mathsf{T}_{\underline{m}} + \neg F(\lfloor m/k \rfloor)$ and $\mathcal{M}_1 \models \mathsf{T}_{\underline{m}} + F(\lfloor m/k \rfloor)$. To see that this is possible, observe that $\mathsf{T}_{\underline{m}} + \neg F(\lfloor \underline{m}/k \rfloor)$ is satisfied in any model $\mathcal{M}_0$ in which $F^{\mathcal{M}_0} = \mathbb{N}$ – i.e. the *standard cut* – and $a = \underline{m}^{\mathcal{M}_0} \in M_0 - \mathbb{N}$ – i.e. any nonstandard integer. Since in this case $\lfloor a/k \rfloor^{\mathcal{M}_0}$ is nonstandard as well, $\mathcal{M}_0 \models \neg F(\lfloor m/k \rfloor)$.[62] To construct $\mathcal{M}_1 \models \mathsf{T}_{\underline{m}} + F(\lfloor m/k \rfloor)$ take any nonstandard model of $\mathcal{M} \models \mathsf{Z}$ and extend it to models of $\mathsf{T}_{\underline{m}}$ in which F is interpreted as a proper cut $I \neq \mathbb{N}$. Now consider an element

[60] Given the tendency of paradox mongers to choose "safe" extremal values (such as $n_P = 1000000$ for bald), one might be dubious as to whether genuine borderline cases only start to arise only in the vicinity of $n_P/2$. As we will soon see, however, what is crucial to the following account is not that such cases occur approximately *halfway* along sorites sequences, but rather that once a realistic value of n_P has been fixed, we will tend to locate its borderline cases by dividing the range $0, \ldots, n_P$ into two or more parts rather than by counting upwards or downwards from one of its endpoints.

[61] Using either variants of the following construction or the related technique of realizing types, this result can be generalized to show that it is possible to construct models of $\mathsf{T}_{\underline{m}}$ in which the denotation of $\underline{m}$ in $\mathcal{M}$ is (e.g.) even or odd, prime or composite, etc. This further illustrates the extent to which neo-feasibilism is compatible with various other constraints which might be imposed on the interpretation of vague predicates or extremal values provided they are consistent with the axioms of $\mathsf{T}_{\underline{m}}$.

[62] Note that taking $F^{\mathcal{M}_0} = \mathbb{N}$ suffices because $\mathcal{M}_0$ only needs to satisfy Overspill for $\mathcal{L}_a$-formulas.

$b \in I - \mathbb{N}$. Since b is non-standard, $c = \overline{k}^{\mathcal{M}} \times^{\mathcal{M}} b$ is as well and also $b +^{\mathcal{M}} i \neq c$ for any $i \in \mathbb{N}$. It thus suffices to let $F^{\mathcal{M}_1}$ be the cut $\{d \in M \ : \ d < b +^{\mathcal{M}} i$ for some $i \in \mathbb{N}\}$. □

As we have seen, the standard means of explaining what it means for t to denote a borderline case of $P(x)$ is that this predicate neither determinately holds nor determinately fails to hold of t. But one way of understanding such indeterminacy is to maintain that in accepting the sortical premises embodied by $\mathsf{T}_{\underline{m}}$ we do not thereby commit ourselves to whether objects which we take to correspond to borderline cases fall under $P(x)$. But if we let $F(x)$ be the arithmetization of $P(x)$, then this is exactly what Proposition 6.1 appears to demonstrate subject to the proviso that in practice we will tend to employ expressions such as $\lfloor n_P/k \rfloor$ denoting fractional values of n_P to describe borderline cases of $P^*(x)$.[63]

6.3. Supervaluationism and epistemicism The foregoing observations may initially appear to suggest that neo-feasibilism is most naturally viewed as compatible with the view that vagueness should be regarded as a form of semantic underdetermination. For it seems plausible to suppose that there will be situations in which the axioms of a theory like $\mathsf{T}_{\underline{m}}$ will constitute the only principles we will be willing to assert about a given soritical predicate $P(x)$. But although Proposition 6.1 demonstrates that $\mathsf{T}_{\underline{m}}$ does not decide the truth values of borderline cases denoted by expressions of the form $\lfloor m/k \rfloor$, this theory still possesses classical models in which $P^*(x)$ receives a definite extension in virtue of the consistency of $\mathsf{T}_{\underline{m}}$.

We have seen that in such a model the interpretation assigned to $P^*(x)$ partitions the domain sharply into an extension and an anti-extension. According to the view just described, however, vague predicates are neither true nor false of borderline cases. Supervaluationists such as Fine (1975) suggest that such gaps reveal a semantic deficiency in the meaning of vague predicates which may be removed by "precisifying" their interpretations. But since any such precisification would be arbitrary in how it handles borderline cases, such theorists accordingly hold that a semantics for vague predicates should take *all* admisible

63 It is also possible to build on this account to show how a nonstandard model $\mathcal{M} \models \mathsf{T}_{\underline{m}}$ can be used to provide an interpretation of the definiteness operator $\mathbb{D}$ which allows for arbitrary degrees of higher-order vagueness. Suppose for instance that we begin by identifying $\lfloor n_P/2 \rfloor$ as a borderline case of $P^*(x)$. At the first stage we can begin by interpreting $\mathbb{D}P^*(x)$ in $\mathcal{M}$ as a proper cut $I_1 \neq \mathbb{N}$ not containing $\lfloor n_P/2 \rfloor^{\mathcal{M}}$ and $\mathbb{D}\neg P^*(x)$ as $\{\underline{n}_P^{\mathcal{M}} - i \ : \ i \in \mathbb{N}\} \cup \{a \in M : \ n_p^{\mathcal{M}} \leq^{\mathcal{M}} a\}$. It is generally acknowledged that the property of being a borderline case is itself tolerant in both the positive and negative directions – i.e. if x is a borderline case, then so are $x+1$ and $x-1$, etc. This suggests taking the extension of $\mathbb{B}P^*(x)$ to be the set $\{\lfloor n_P/2 \rfloor^{\mathcal{M}} - i \ : \ i \in \mathbb{N}\} \cup \{\lfloor n_P/2 \rfloor^{\mathcal{M}} + i \ : \ i \in \mathbb{N}\}$. If we continue to follow the classical description of higher-order vagueness provided by Sainsbury (1991), we should then also acknowledge that there exist borderline cases of $\mathbb{B}P^*(x)$. This in turn suggest that we would naturally tend to locate such objects in regions surrounding the denotation of the terms $\lfloor n_P/4 \rfloor$ and $\lfloor 3n_P/4 \rfloor$. And this suggests taking the extension of $\mathbb{B}\mathbb{B}P^*(x)$ to be $\{\lfloor n_P/4 \rfloor^{\mathcal{M}} - i \ : \ i \in \mathbb{N}\} \cup \{\lfloor n_P/4 \rfloor^{\mathcal{M}} + i \ : \ i \in \mathbb{N}\} \cup \{\lfloor 3n_P/4 \rfloor^{\mathcal{M}} - i \ : \ i \in \mathbb{N}\} \cup \{\lfloor 3n_P/4 \rfloor^{\mathcal{M}} + i \ : \ i \in \mathbb{N}\}$. Since the order-type of $\mathcal{M}$ is $\omega + (-\omega + \omega) \cdot \eta$ where η is a dense linear-ordering without endpoints, it follows that this process can be continued indefinitely to assign non-overlapping extensions to the iterates $\mathbb{B}\mathbb{B}P^*(x), \mathbb{B}\mathbb{B}\mathbb{B}P^*(x) \ldots$ of the borderline operator. Similarly, since we have interpreted $\mathbb{D}P^*(x)$ as a proper cut $I \supsetneq \mathbb{N}$, we can additionally interpret iterations of this operator $\mathbb{D}\mathbb{D}P^*(x), \mathbb{D}\mathbb{D}\mathbb{D}P^*(x), \ldots$ as a non-well-founded sequence of proper cuts $I_1 \supsetneq I_2 \supsetneq I_3 \subseteq \ldots \subsetneq \mathbb{N}$. Detailed consideration of whether this construction provides a felicitous model of reclassifications which may be induced by the iteration of the borderline and definiteness operators or whether it avoids the other difficulties customarily associated with higher-order vagueness (cf., e.g., Wright, 2010) will have to await another occasion.

precisifications of $P(x)$ into account. In particular, they hold that a sentence should be regarded as true just in case it is true on all admissible precisifications (i.e. *super-truth*), false just in case false on all admissible precisifications (i.e. *super-falsity*), and neither true nor false otherwise.

Some familiar consequences of replacing truth simplicter with super-truth in a semantics for vague predicates are as follows: i) all classically valid schema are still validated; ii) the super-truth of $P^*(t) \vee \neg P^*(t)$ notwithstanding, neither $P^*(t)$ nor $P^*(t)$ will itself be super-true when t denotes a borderline case; iii) $\exists x(P^*(x) \wedge \neg P^*(x'))$ is super-true since $P^*(x)$ will possess a sharp boundary for each precisification. Since supervaluationists will consequently count $\forall x(P^*(x) \rightarrow P^*(x'))$ as super-false, such theorists are hence committed to regarding both the conditional and inductive sorites as unsound. Such a view thus preserves classical logic at the schematic level and provides a rationale for rejecting the reasoning of derivations like 1.3 and 1.4. But it does so at the expense of abandoning classical logic at the substitutional level[64] and also appearing to assert the existence of sharp boundaries for soritical predicates (from which the failure of tolerance also follows).

A counterpoint to supervaluationism is provided by the epistemic views of vagueness associated with Williamson (1994) and Sorensen (1988). Rather than regarding vagueness as a form of semantic underdetermination, epistemicists take it to be a form of ignorance. Such theorists assume all predicates – inclusive of those traditionally regarded as vague – have sharp boundaries which are determined by a combination of the world and our linguistic practices, possibly in a complex manner.[65] As such they are able to preserve classical logic in the sense of both i) and ii). But while epistemistists agree with supervaluationists about the existence of sharp boundaries in the sense of iii), they offer a different account of their status. For on the one hand, supervaluationists explain the sense in which they take $\exists x(P^*(x) \wedge \neg P^*(x'))$ to be true by pointing out that it may be made true by different witnesses in different precisifications. On the other hand, epistemicists matter-of-factly accept the existence of an n such that (e.g.) a man with n hairs is bald and a man with $n+1$ hairs is not, while offering auxiliary accounts of why we cannot come to know the value of n.[66]

Neo-feasibilists are in a position to chart a different route through these concerns which may plausibly be taken to combine several of the insights traditionally associated with supervaluationism and epistemicism while also avoiding some of their evident weaknesses. As we have seen, for instance, theories like $\mathsf{T}_{\underline{m}}$ possess classical models in which the predicate $P^*(x)$ receives a definite interpretation. But since such models satisfy $\forall x(P^*(x) \rightarrow P^*(x'))$ – and thus fail to satisfy $\exists x(P^*(x) \wedge \neg P^*(x'))$ – neo-feasibilists are not committed to the failure of tolerance or the existence of sharp boundaries for soritical predicates. Such theorists also appear to be better positioned than epistemicists in accounting for how vague predicates can possess such definite extensions. For whereas the sorts of auxiliary explanations just alluded to are often regarded with suspicion, the existence of models $\mathcal{M} \models \mathsf{T}_{\underline{m}}$ in which $P^*(x)$ is assigned an extension is a simple consequence of the Compactness Theorem for first-order logic given the consistency of $\mathsf{T}_{\underline{m}}$.

---

64 E.g. in the sense that $\exists x\varphi(x)$ may be super-true without it being the case that $\varphi(x)$ is "supersatisfied" by the same object in all admissible precisifications.

65 See, e.g., (1994, §7.5) for Williamson's defense of the view that epistemicism is compatible with the view that meaning supervenes on language use, albeit in a manner which may be "unsurveyably chaotic".

66 For instance, Williamson (1994, ch. 8) proposes that knowledge of statements of the form $P(t)$ must be the product of belief-forming mechanisms which allow for margins of error which would not be satisfied in the case where t denoted a borderline case of $P(x)$. On the other hand, Sorensen (1988) offers a complex set of analogies between our apparent inability to classify borderline cases and other epistemic "blindspots".

On the other hand, it is a consequence of Proposition 6.1 that the Compactness Theorem does not itself tell us whether terms of the form $\lfloor n_P/k \rfloor$ by which we might plausibly denote various borderline cases fall within the extension of $P^*(x)$. More generally, the following results suggest that although we are able to prove that models of $\mathsf{T}_{\underline{m}}$ exist, the only characteristics we can know about the extensions which they assign to $P^*(x)$ are essentially *negative*:

Proposition 6.2 *Let $\mathsf{T}_{\underline{m}}$ be as defined above where $\mathsf{Z} \supseteq I\Delta_0$ and $\mathcal{M} \models \mathsf{T}_{\underline{m}}$.*

i) $\mathcal{M}$ is not a recursive model – i.e. were $\mathcal{M}$ to be mapped isomorphically on to a model $\mathcal{N}$ with domain $\mathbb{N}$ (which is always possible if $\mathcal{M}$ is countable), then the induced extensions of the functional symbols $+$ and $\times$ on $\mathcal{N}$ would not be recursive.

ii) If $\mathcal{M} \models \mathrm{Ind}(\Sigma_n)$ (i.e. induction for Σ_n-formulas), then the extension of $P^*(x)$ in $\mathcal{M}$ is not definable by a Σ_n-formula of $\mathcal{L}_a$ in $\mathcal{M}$.

Since any model $\mathcal{M}$ of $\mathsf{T}_{\underline{m}}$ must be nonstandard, the first of these observations is simply a recapitulation of Tennenbaum's Theorem.[67] This result can be taken to suggest that metatheoretic results such as the Compactness Theorem should not be viewed as mechanisms for explicitly constructing nonstandard models $\mathcal{M} \models \mathsf{T}_{\underline{m}}$, but merely as a means of inferring their existence on the basis of other mathematical or set theoretic assumptions.[68] In virtue of this, little can be said about the the extension of $P^*(x)$ absent an explicit specification of $\mathcal{M}$ which, per i), cannot be given constructively. For per ii), such an extension cannot coincide with that of any $\mathcal{L}_a$-formula for which $\mathcal{M}$ satisfies induction.[69]

We have already seen that one advantage neo-feasibilism has over both supervaluationalism and epistemicism is that it provides a means of resolving the conditional and inductive sorites in a manner which allows us to retain the intuition that vague predicates do not possess sharp boundaries. With respect to preserving classical logic, the neo-feasibilist has two options. On the one hand, he can adopt a supervalutational view of truth according to which an $\mathcal{L}_{P^*}$-sentence is true just in case it is true in all models of $\mathsf{T}_{\underline{m}}$, false if it is false in all such models, and neither true nor false otherwise. On the other hand, he can hold that the extension of $P^*(x)$ is fixed with respect to some particular model $\mathcal{M} \models \mathsf{T}_{\underline{m}}$ obtained in the manner described above. But in so doing, he can also appeal to Proposition 6.2 to explain why we are not thereby provided with an explicit characterization of this extension.

I will adopt the latter strategy here not only in virtue of its simplicity, but also because it allows for the formulation of several potentially salutary refinements to epistemicism. Note first if we assume that the extension of $P^*(x)$ is fixed relative to a particular model $\mathcal{M} \models \mathsf{T}_{\underline{m}}$, classical logic is preserved in the sense of both i) and ii) above. On the other hand,

---

[67] See (Kaye, 1991, §11.3).

[68] It is now standard to account for the sort of non-constructiveness at issue using the framework of Reverse Mathematics – see, e.g., (Simpson, 2009, §IV.3, §VIII.2) or (Dean and Walsh, 2017, §4).

[69] This follows because if $\varphi(x)$ is such a formula and $\mathcal{M} \models \forall x(P^*(x) \leftrightarrow \varphi(x))$ and also $\mathcal{M} \models \varphi(0) \wedge \forall x(\varphi(x) \rightarrow \varphi(x')) \rightarrow \forall x\varphi(x)$, then $\mathcal{M} \models \forall x P^*(x)$ in contradiction to $\mathcal{M} \models \neg P^*(\underline{m})$. If we assume per §4 that $\mathsf{T}_{\underline{m}}$ extends $\mathrm{I}\Delta_0$, this means that the extension of $P^*(x)$ cannot coincide with that of a Δ_0-formula in $\mathcal{M}$. A sharper computational characterization can be given if we take $\mathsf{T}_{\underline{m}}$ to extend the theory V^1 described in Appendix A – i.e. in this case the extension of $P^*(x)$ cannot coincide with a set decidable in polynomial time with computations relativized to $\mathcal{M}$. Thus not only does $\mathsf{T}_{\underline{m}}$ fail to decide potential borderline cases of $P^*(x)$, such cases would still be hard to decide computationally even if we could perform arithmetical calculations within one of its models.

by taking this step we incur an obligation similar to that of the traditional epistemicist to explain how it is that vague predicates can have fixed extensions which is compatible with otherwise plausible assumptions such as the fact that our usage of them is insufficient to determine their precise extensions.

One possible means of responding to this challenge is to attempt to draw an analogy between the formal undecidability of statements of the form $P^*(\lfloor m/k \rfloor)$ and other mathematical independence results. Note, however, that it is possible to distinguish between several different types of undecidable mathematical statements. For instance there are statements such as the Continuum Hypothesis (CH) which are independent of strong mathematical theories like ZFC and which are occasionally suggested to be "absolutely undecidable". Another relevant class of principles is typified by the statement $\forall x \forall y(x \cdot y = y \cdot x)$ which expresses in the language of group theory that the operation denoted by $\cdot$ is commutative. Although this statement is independent of the group theoretic axioms G, it is a commonplace that neither our understanding of what it is to be a group nor any other mathematical facts make this statement determinately true or false.

Epistemists have occasionally sought to explain the nature of our ignorance about how to classify borderline cases in terms of an analogy with undecidable mathematical statements of the first kind.[70] However a common view (e.g. Koellner, 2009) is that it is difficult to rule out now that future work will cause us to refine our concepts in such a way that statements like CH will come to be regarded as decided on mathematical grounds alone. On the other hand, it seems difficult accept that either further reflection on or application of concepts like *bald*, *heap*, or *looks orange* will leads us to refine our understanding in a manner which would non-stipulatively decide the relevant borderline cases.

On the other hand, it seems that analogy with the second form on incompleteness is more promising. For note that in citing the Completeness Theorem (which is a simple corollary of Compactness) to infer the existence of a model $\mathcal{M}$ of the consistent theory $\mathsf{T}_{\underline{m}}$ we fix the extension of $P^*(x)$ in much the same manner that we would fix the extension of $\cdot$ if we were to employ Completeness to infer the existence of a group G from the proof theoretic consistency of the axioms G. But in this case, the independence of $\forall x \forall y(x \cdot y = y \cdot x)$ illustrates how we can come to know that such a structure exists without thereby coming to know whether its operation is commutative.

When we assert that a theory like $\mathsf{T}_{\underline{m}}$ formalizes the principles which we accept for a vague predicate $P(x)$, we presumably adopt a similar attitude about the factors which can legitimately contribute to its semantic interpretation.[71] Neo-feasibilists are well positioned to explain this. For even if we assume that $P^*(x)$ receives a fixed interpretation in some model $\mathcal{M} \models \mathsf{T}_{\underline{m}}$, Propositions 6.1 and 6.2 intervene to show how little we come to know about its precise extension. Neo-feasibilists can thus point to the non-constructiveness of the metatheoretic principles such as the Completeness and Compactness Theorems underlying these results as providing an epistemological account of the sense in which extensions of vague predicates are underdetermined – i.e. one which allows us to consistently assume that such predicates have fixed extensions while simultaneously accounting for our ignorance of their boundaries.

---

[70] See, e.g., (Williamson, 1994, pp. 203-204).

[71] This is not, of course, to rule out the possibility that such a theory will include additional principles involving $P(x)$ beyond the soritical premises alone – e.g. in the form of what Fine (1975) calls "penumbral connections" to other predicates. For the adoption of such principles will presumably not result in a theory to which Propositions 6.1 or 6.2 cease to apply – e.g. by deciding borderline cases or admitting only standard models.

§7. Vagueness, representation, and finitude Neo-feasibilism differs from most mainstream theories of vagueness in that it seeks to provide an interpretation of vague predicates relative to which the premises of the sorites are classically consistent. I have argued that such an approach derives independent motivation from the strict finitist critique of classical mathematics in §2 - §4 and from the use of measurement theory to represent perceptual continua and indiscriminable differences in §5. In §6 I have also attempted to illustrate how this view provides insight into a variety of phenomena and debates which are often discussed by philosophers in relation to vagueness.

Nonetheless, the suggested interpretations still take the form of nonstandard models of formal theories of arithmetic and analysis originating in mathematical logic. The suggestion that such structures have some bearing on the semantics of natural language may continue to strike some readers as incongruous simply in virtue of their remoteness from our everyday experiences. Perhaps most worryingly, neo-feasibilists seek to block the reasoning of the conditional sorties by interpreting the sorts of expressions we might use to refer to the number of hairs on a bald man's head as denoting numbers with *infinitely* many predecessors and to block the reasoning of the phenomenal sorites by interpreting the phenomenal color spectrum as containing hues which which are *infinitesimally* close to one another. As such representations may seem out of keeping with a everyday understanding of finitude, it will be useful to examine such apparent infelicities in greater detail.

The nature of the alleged misrepresentation may be even more vividly illustrated by the way in which neo-feasibilists will be inclined to interpret several of the predicates included in Table 1 – e.g. *walking distance*, *tall*, or *noonish* – we have yet to discuss. Like the phenomenal predicates discussed in §5, these are all examples of ratio predicates in the sense that their applicability to specific items relies on a conventional but arbitrary choice of empirical unit. On the other hand these predicates apply not to points or regions of perceptual continua, but rather to empirical magnitudes such as *lengths* or *durations* of which we typically have a more robust understanding. For instance, we are well acquainted with the fact that once we have chosen a unit u for length measurement – e.g. a standard foot or meter bar – we may then make systematic use of the real numbers to measure and compare lengths. Moreover, the process of physically concatenating multiple instances of u gives rise to an additive structure on the domain of lengths which we conventionally assume to be isomorphic to that of the non-negative real numbers.

Suppose we now follow Gaifman (2010) in focusing on the predicate $W(x) = x$ *is walking distance measured in feet*. We can then think of the domain of lengths which forms the field of this predicate as comprising a structure $\mathcal{A} = \langle A, \prec, \circ, e, u \rangle$ where A is some sufficiently rich collection of magnitudes which we assume to include the null magnitude e, the unit magnitude u (i.e. a standard foot), $\prec$ is the empirical *longer than* relation, and $\circ$ is the operation of concrete length concatenation. $\mathcal{A}$ is representable by a structure of the form $\mathcal{A}^* = \langle A^*, +, <, +, 0, 1 \rangle$ via a measurement function $f : A \to A^*$ which (for simplicity) I will assume to be injective and to satisfy $f(e) = 0$, $f(u) = 1$, for all $a, b \in A$, $a \prec b$ iff $f(a) < f(b)$, and $f(a \circ b) = f(a) + f(b)$.

In most familiar cases, we would also take $A^* = \mathbb{R}^+$. But once we have fixed such a language for describing our normal practices of length measurement, we can also provide a mathematical interpretation of $W(x)$ by defining $W^*(x)$ iff $f^{-1}(x)$ *is a walking distance*. We may then formulate the relevant sortical premises for $W(x)$ as $W^*(\mathsf{0}), \forall x(W^*(x) \to W^*(x'))$, and – employing Gaifman's choice of extremal value for $W(x)$ – $\neg W^*(\underline{6000})$. If we treat $\underline{6000}$ as a constant symbol, it should now be evident that the theory $\mathsf{U}_{\underline{6000}} = \mathsf{A} + W^*(\mathsf{0}) + \forall x(W^*(x) \to W^*(x')) + \neg W^*(\underline{6000})$ can be shown to be consistent in the manner similar to Propositions 5.1 – e.g. we can interpret $\underline{6000}$ as denoting an infinitely large integer in a non-Archimedean model $\mathcal{R}^\dagger$ of A. Such a model will, of course, felicitously represent the fact that 6000 feet does not comprise a walking distance in the sense that $\mathcal{R}^\dagger \models \neg W^*(\underline{6000})$. But on the other hand if we were to attempt to transfer this fact back into the empirical domain $\mathcal{A}$, we would reach the conclusion that $f^{-1}(\underline{6000})^{\mathcal{R}^\dagger}$ must itself be an *infinite magnitude* – i.e. one which cannot be surpassed by any finite concatenation $e \circ \ldots \circ e$ of empirical foot-long units.[72]

A likely reaction is that a such model *misrepresents* certain of our commonly held beliefs about walking distances and related notions. For although we may accept the soritical premises for $W^*(x)$, we certainly do not think that 6000 feet (≈ 1.136 miles) is an infinite distance. And thus despite the fact that $\mathcal{R}^\dagger$ will satisfy some portion of our practices of length measurement, one might reasonably object that it also distorts other aspects of how we understand the situation envisioned by the sortical premises.[73]

There are several avenues by which a neo-feasibilist might elect to respond to this objection. For instance, he might dwell further on how little stock paradox mongers tend to place in their choice of extremal values in formulating the sorites[74] – e.g. by offering a contextual account of how different circumstances allow us to treat different numbers as "infinite for all intents and purposes". Although such a proposal has an evident affinity with the original feasibilist proposal considered in §3, I will now briefly describe two other possible responses to the envisioned objection. These options – which I will refer to as *radical* and *moderate measurement theoretic anti-realism* – parallel the traditional debate as to whether vagueness resides in the world or in our linguistic and cognitive representations.

According to the radical anti-realist, the envisioned critic does not possess justification for the incredulity with which he regards the neo-feasibilist's claim that $\mathcal{R}^\dagger$ provides a felicitous representation of how we reason about the predicate *walking distance*. Such a view takes inspiration from a tradition within measurement theory which questions whether it is always appropriate to assume that the mathematical structures employed as scales for length measurement must be Archimedean.[75] For note that if A contains infinitely many magnitudes, then the fact that $\mathcal{A}$ satisfies the Archimedean axiom – i.e. that for all $a, b \in A$, if $a \prec b$, then there exists $n \in \mathbb{N}$ such that $b \prec a \circ \ldots \circ a$ (n times) – will generally not be confirmable empirically.

If we lack justification for the assumption that physical space is Archimedean, then in order to rule out the possibility that (e.g.) 6000 feet corresponds to a length that is infinite

---

72 It is also evident that the consistency of $\mathsf{U}_{\underline{6000}}$ can be demonstrated by interpreting $f(u)$ (i.e. the term denoting the image of our standard foot under the measurement function $f(x)$) as an infinitesimal and then interpreting $\underline{6000}$ as denoting the standard integer 6000 in $\mathcal{R}^\dagger$. But since such a model is related to the one just described by an affine transformation, this can be viewed as another description of the same model up to a "re-scaling of units".

73 It should also be evident that similar points apply (*mutatis muntandis*) to our ability to read back the properties attributed to various numbers in models described above for ordinal predicates like *bald* and *non-heap*. For instance, although we are likely to assent to the fact that a man with 1000000 hairs on his head is not bald, we are presumably unwilling to assent to the fact that such an individual has *infinitely* many hairs. The analogous observation with respect to the phenomenal predicate is the fact that the model $\mathcal{R}^\dagger$ constructed in Proposition 5.3 will treat a just noticeable difference for the predicate $P(x)$ as infinitesimally small, from which it follows that a finite number of such differences can no longer be chained together to yield a perceptible one in the manner noted at the end of §5.

74 As exemplified by Gaifman's (2004) own choice on another occassion of $n_W = 50000$ or Dietz's (2011) choice of $n_W = 5280000$.

75 Such doubts can already be found in Hilbert (Hallett and Majer, 2004, p. 171). Similar concerns about the empirical status of the Archimedean axiom have repeatedly surfaced in the subsequent literature on measurement theory – e.g., (Krantz et al., 1971, §1.5), (Skala, 1975), (Narens, 1985), and (Luce et al., 1990, §21.7).

with respect to our chosen unit u, we would need to conduct an experiment which shows that the result of concatenating sufficiently many copies of u yields a magnitude of greater than 6000 feet. But it is unlikely that we will be able to practically carry out either this experiment or a similar one which would be called for had the paradox monger selected a larger value of n_W initially. On this basis the radical anti-realist concludes that the challenge of misrepresentation is grounded in the unjustified assumption that $\mathcal{R}^\dagger$ distorts any essential aspect of our practices of length measurement.

The response offered by the moderate anti-realist is more nuanced. Such a theorist suggests that while structures like $\mathcal{R}^\dagger$ may not accurately represent physical space, it does not follow from this that the most felicitous model of our reasoning about *walking distances* needs to assume that the domain of items to which we apply this notion is comprised of *bona fide* empirical magnitudes. We might, for instance, adopt the approach suggested in §4 and §5 and instead attempt to develop a system of representation which is compatible with both general principles we accept about length measurement and also our judgements about whether specific magnitudes constitute walking distances. As we have seen, this leads naturally to a non-Archimedean scale such as $\mathcal{R}^\dagger$. The moderate anti-realist encourages us to understand the domain of this structure as being comprised of mathematical proxies representing not empirical lengths but rather a generalized notion of spatial magnitude which we have introduced to represent what we colloquially refer to as "walking distances".[76]

It is still a consequence of this strategy that if we were to reason about such magnitude-like items directly using the model $\mathcal{R}^\dagger$ and then attempt to reapply our conclusions to the empirical world, we would be led to the apparently false conclusion that 6000 feet is an infinite distance. This might in turn prompt someone who is using $\mathcal{R}^\dagger$ as a representation to facilitate reasoning about walking distances to acknowledge that such a structure is indeed not an accurate representation of physical space. It would seem, however, that we rarely engage in the pattern of inference just described – i.e. one in which we categorically state a conclusion in non-vague language on the basis of reasoning we have knowingly conducted using vague premises.

Perhaps one reason for this is because we implicitly realize that the empirical world must ultimately be described in a non-vague manner – e.g. in terms of the fundamental non-vague physical magnitudes *length*, *mass*, and *duration*. But in addition to this, the conservativity results recorded in Propositions 4.1 and 5.1 show that whenever we are able to derive a statement formulated in the non-vague portion of the language of a theory such as $\mathsf{T}_{\underline{m}}$ or $\mathsf{U}^{\alpha}_{\mathtt{b}}$, then we will also be able to derive it in the corresponding non-vague mathematical theory – i.e. in Z or A. It thus follows that were it to be possible to derive a proposition expressing that 6000 denotes an infinite distance – say formalized as a statement of the form $\neg\mathtt{Fin}(\underline{6000})$ – in $\mathsf{T}_{\underline{m}}$ or $\mathsf{U}^{\alpha}_{\mathtt{b}}$, then it would already be possible to derive it in Z or A.[77] But this is impossible since if $\neg\mathtt{Fin}(\underline{6000})$ were derivable in one of these theories, it would also be true in their standard models based on the structures $\mathcal{N}$ and $\mathcal{R}$.

This in turn exposes another weakness in the argument of the imagined critic of neo-feasibilism. For the argument just rehearsed shows that in pressing the point that $\underline{6000}$ fails to denote a finite magnitude in $\mathcal{R}^\dagger$, such a theorist must be relying on an auxiliary premise beyond mathematical theories such as Z or A. To the extent that such a premise is available to the critic at all, he must thus have obtained it by some other means – e.g. by "inspecting"

[76] A related approach is often adopted in the construction of appropriate measurement spaces for perceptual magnitudes – e.g. in the use of multidimensional scales for color and pitch measurement (e.g. Suppes et al., 1989, §15).

[77] This supposes that since finitude is a mathematical notion, an adequate definition of $\mathtt{Fin}(x)$ would have to be given in a mathematical language such as $\mathcal{L}_{\mathsf{Z}}$ or $\mathcal{L}_{\mathsf{A}}$.

the nonstandard model $\mathcal{R}^\dagger$ directly. But absent an account of what might be involved in such a process, it would seem that the critic's use of this premise remains unjustified. This in turn illustrates another respect in which debates about vagueness encroach on an issue at the boundary of philosophy of language and philosophy of mathematics – i.e. in what semantic and epistemic relationship do we stand to the mathematical models we introduce in the course of our everyday and scientific reasoning about the empirical world?[78]

A On formulating the sorites in two-sorted bounded arithmetic The goals of this appendix are fourfold: i) to describe a theory of bounded arithmetic V^1 of the sort mentioned in §4 in which it is possible to naturally formalize basic conjoint reasoning about unary and positional numerals;[79] ii) to describe how it is possible to formalize principle (D) in the language of such a theory; iii) to sketch the proof that the resulting sentence is formally independent of V^1; iv) to describe how the premises of the conditional sorites may be interpreted using a nonstandard model of this theory.

Recall that in the relevant case (D) states the following: for every base $b \geq 2$ positional numeral π, there exists a unary numeral υ such that $[\![\pi]\!]_b = [\![\upsilon]\!]_u$. The fact that this statement contains separate quantifiers over unary and positional numerals provides some initial justification for attempting to formulate a two-sorted system with distinct types of variables intended to range over the denotations of such expressions.[80] The languages $\mathcal{L}^2_a$ of the theories V^i which I will now describe thus has *first-sort variables* $x, y, z, \ldots$ and *second-sort variables* $X, Y, Z, \ldots$ First-sort variables are intended to range over $\mathbb{N}$, while second-sort variables are intended to range over the class $2^{<\mathbb{N}}$ of finite subsets of $\mathbb{N}$. Note that a set $X \in 2^{<\mathbb{N}}$ may simultaneously be understood as a finite binary sequence $\mathtt{b}_0, \ldots, \mathtt{b}_{k-1}$ defined by $b_i = 1$ if $i \in X$ and $b_i = 0$ otherwise and also as the natural number that this sequence denotes when viewed as a binary numeral.[81] Via such a correspondence, second-sort variables can also be understood as ranging over the sort of object for which we have previously used υ as a metavariable.

The non-logical symbols of the language $\mathcal{L}^2_a$ include the first-sort constant $\mathtt{0}$, as well as the first sort functions $', +$ and $\times$. Additionally, we include the less than and equality

78 One obvious point of contact is the debate surrounding the so-called *model theoretic argument* of Putnam (1980). Note, however, that nothing which has been said above need be interpreted as suggesting that the mathematical concept of finitude is genuinely indeterminate in order to allow that nonstandard methods may provide felicitous representations of certain aspects of our reasoning about vague predicates. On the other hand, it at least seems reasonable to challenge the imagined critic of neo-feasibilism about the legitimacy of using "more theory" to rule out such interpretations given that the conservativity results show that the formulation of such machinery will have to rely on concepts outside the cluster of everyday notions needed to formulate the sorites itself.

79 V^1 is one of a family of systems of two-sorted bounded arithmetic originally introduced by Zambella (1996) which have recently been presented in slightly different form by Cook and Nguyen (2010). The terminology and notation of this later presentation are adopted in this section.

80 Additional motivation originating in computer science stems from the desire to formalize how we use numerical algorithms such as carry addition and multiplication to compute sum and products using positional notation. For note that although the inputs and outputs of such procedures are positional numerals, their specification also makes reference to other "loop" or "counter" variables whose values are bounded by the length of their inputs and which are only operated upon in the manner of unary numerals by incrementing or decrementing their values by 1.

81 For instance finite the set $X = \{1, 3, 5\}$ corresponds to the binary numeral $\mathtt{101010}$ which in turn denotes the number $42 = 2^5 + 2^3 + 2^1$.

relation symbols $<$ and $=$ for numbers, and the symbol $|\cdot|$ intended to denote a function from second-sort to first-sort objects which returns the successor of the largest element of the set X to which it is applied, or 0 if X is empty. Note that $|X|$ thus corresponds to the *length* of the binary numeral represented by X. Similarly, the predication relation $X(i)$ can be understood as expressing that the ith digit (or *bit*) in X is 1. A model for $\mathcal{L}_a^2$ is thus of the form $\mathfrak{M} = \langle U_1, U_2, <, ', +, \times, |\cdot|, 0\rangle$ where U_1 and U_2 are first- and second-sort domains, $<, ', +$ and $\times$ are interpreted on U_1 as in traditional one-sorted $\mathcal{L}_a$-theories and $|\cdot|^{\mathcal{M}} : U_2 \to U_1$. The standard $\mathcal{L}_a^2$-model $\mathfrak{N}$ is such that $U_1 = \mathbb{N}$, $U_2 = 2^{<\mathbb{N}}$, $\leq, ', +$ and $\times$ are given their standard arithmetical interpretations, and $|X|^{\mathfrak{N}} =$ *the largest element of* $X + 1$.

In order to state the axioms of the theories V^i we must define the classes of second-sort bounded formulas Σ_1^B and Π_1^B. As a first step, we define the *first-sort terms* of $\mathcal{L}_a^2$ in the standard manner to include the constants $\mathsf{0}$, variables $x, y, z, \ldots$, and functional expressions formed from the terms $|X|$ using $+$, $'$, and $\times$. The only second-sort terms of $\mathcal{L}_a^2$ are the variables $X, Y, Z, \ldots$ The atomic formulas of $\mathcal{L}_a^2$ are thus $t = u, X = Y, t \leq u, X(t)$. Formulas are constructed from first- and second-sort quantifiers $\exists x, \exists X, \forall x, \forall X$ and a first-sort quantifier is called *bounded* if it is of the form $\forall x < t\varphi(x) =_{\mathrm{df}} \forall x(x < t \to \varphi(x))$ or $\exists x < t\varphi(x) =_{\mathrm{df}} \exists x(x < t \wedge \varphi(x))$ where x does not occur free in t. We also introduce the abbreviations $\forall X \leq t\varphi(X)$ for $\forall X(|X| \leq t \to \varphi(X))$ and $\exists X \leq t\varphi(X)$ for $\exists X(|X| \leq t \wedge \varphi(X))$ (where in both cases X does not occur in the term t).

We now define the class $\Sigma_0^B = \Pi_0^B$ to be the set of formulas over $\mathcal{L}_a^2$ all of whose first-sort quantifiers are bounded and which contain no second-sort quantifiers (although they may contain free second-sort variables). For $i > 0$, Σ_i^B is defined recursively to be the set of all formulas beginning with a block of zero or more bounded existential second-sort quantifiers followed by a Π_{i-1}^B formula. Σ_1^B is thus the set of formulas that begin with zero or more existential second-sort quantifiers (bounded or unbounded), followed by a Σ_0^B-formula.

All of the theories $\mathsf{V}^0 \subseteq \mathsf{V}^1 \subseteq \ldots$ extend Q and contain the following second-sort axioms:

(i) $\forall y \forall X(X(y) \to y < |X|)$
(ii) $\forall y \forall X(y' = |X| \to X(y))$
(iii) $\forall X \forall Y(X = Y \leftrightarrow \forall x < |X|(X(x) \leftrightarrow Y(x))$

i) and ii) formalize the interpretation of $|X|$ indicated above while iii) expresses that second-sort objects are individuated extensionally. Finally, the theory V^i (for $i \geq 0$) contains the following comprehension scheme for Σ_i^B formulas:

(Σ_i^B-CA) $\exists X \leq y \forall z \leq y(X(z) \leftrightarrow \varphi(z))$ where $\varphi(x) \in \Sigma_i^B$ does not contain X free.

Although none of the theories V^i contain induction axioms or schema, it is not hard to see that all instances of $\mathrm{Ind}(\Sigma_0^B)$ are derivable in V^0. In fact, a simple model theoretic argument shows that V^0 is a conservative extension of $\mathrm{I}\Delta_0$ (Cook and Nguyen, 2010, p. 98-99).

Next note that if we view a finite set X as a binary sequence with characteristic function $c_X(i)$, then the value of the corresponding binary numeral is given by

$$[\![X]\!]_2 = \Sigma_i^{|X|-1} c_X(i) \cdot 2^i$$

Such an interpretation makes explicit the intention of regarding second-sort objects as the denotations of binary numerals. Note, however, that $\mathcal{L}_a^2$ does not contain a primitive term for the function $[\![\cdot]\!]_2$, nor even for addition or multiplication on second-sort objects. Nonetheless, it is possible to construct formulas $\mathtt{Add}(X, Y, Z)$ and $\mathtt{Mult}(X, Y, Z)$ defining the graphs of second-sort addition and multiplication such that these functions are provably total in V^1.[82]

In order to formalize principle (D) in $\mathcal{L}_a^2$ we must also construct a similar formula $\texttt{Bin}(X, y)$ such that $\texttt{Bin}(X, \overline{n})$ holds in the standard model just in case $[\![X]\!]_2 = n$. This can be accomplished by employing the fact that there is a Δ_0-formula $\varepsilon(x, y)$ defining the graph of the exponential function 2^x in $\mathcal{L}_a$ (Cook and Nguyen, 2010, III.3.3.1). It is then straightforward to formalize the definition of $[\![X]\!]_2$ using facts about the representability of finite sequences similar to those required to define addition and multiplication for second-sort objects in $\mathcal{L}_a^2$. We can then express (D) in $\mathcal{L}_a^2$ via the formula $\forall X \exists y \texttt{Bin}(X, y)$ which states that for every second-sort object X (i.e. object denoted by a binary numeral), there exists a co-denoting first-sort object y (i.e. object denoted by a unary numeral).

But relative to this formalization (D) is underivable in any of the theories V^i – i.e.

PROPOSITION A.1. *For all i, then* $\mathsf{V}^i \not\vdash \forall X \exists y \texttt{Bin}(X, y)$ (presuming, V^i is consistent).

(*Proof sketch*). It is easy to see that that the relation $\texttt{Pow2}(x, Y)$ which holds just in case $[\![Y]\!]_2 = 2^x$ is Σ_0^B definable in a manner which determines a function provably total in V^0.[83] If we let $\texttt{Pow2}(x) : U_1 \to U_2$ abbreviate this function, then it follows that $\mathsf{V}^i \vdash \forall x \forall y[\texttt{Bin}(\texttt{pow2}(x), y) \leftrightarrow \varepsilon(x, y)]$. Now assume for a contradiction that $\mathsf{V}^i \vdash \forall X \exists y \texttt{Bin}(X, y)$ where $\texttt{Bin}(X, y)$ is defined using $\varepsilon(x, y)$ as described above. In this case we would also have $\mathsf{V}^i \vdash \forall x \exists y \varepsilon(x, y)$. Note, however, that since all of the theories V^i are polynomially bounded (i.e. they are axiomatized by a set of formulas all of whose first- and second-sort quantifiers are bounded by $\mathcal{L}_a$ terms), they satisfy the two-sorted analog of Parikh's Theorem (Cook and Nguyen, 2010, V.3.4). In the relevant case, this entails that if $\mathsf{V}^i \vdash \forall x \exists y \varphi(x, y)$ where $\varphi(x, y)$ is a bounded formula, then $\mathsf{V}^i \vdash \forall x \exists y \leq t(x) \varphi(x, y)$ for some $\mathcal{L}_a$ term $t(x)$. But since all such terms are polynomials, it hence cannot be that $\mathsf{V}^i \not\vdash \forall x \exists y \varepsilon(x, y)$. □

The relative directness of the formalizations which stand behind the prior results testifies to the fact that theories in the hierarchy V^i provide a natural framework for representing many aspects of our everyday reasoning involving positional numerals. In fact, V^1 might be taken to be a particularly appropriate choice for such a theory as the class of functions which can be proven to be total relative to Σ_1^1-definitions in V^1 corresponds to those which can be computed in polynomial time. This in turn makes it another plausible candidate for an "anthropomorphic" theory in the sense of Wang (1958), Parikh (1971), and Cook (1975).

Turning now to how Proposition A.1 relates to the formulation of the sorites, first observe that since $\mathsf{V}^1 \not\vdash \forall X \exists y \texttt{Bin}(X, y)$, there exists a model $\mathfrak{M} \models \mathsf{V}^1 + \neg \forall X \exists y \texttt{Bin}(X, y)$. There will hence exist a set C in the second-sort domain U_2 of $\mathfrak{M}$ for which there is no element y in its first-sort domain U_1 such that $[\![C]\!]_2 = y$. Given that $\mathfrak{M}$ satisfies V^1, we will be able to reason with C exactly as if it represented a binary numeral – e.g. by appending or deleting digits, computing its sum or product with other numerals via the carry algorithms, etc. Moreover, since $|\cdot|$ is always interpreted as a total function, the length of C will be measured by a number $|C| = d \in U_1$. But since $[\![Y]\!]_2 < 2^{|Y|}$, it must be the case that U_1 does not contain the number which we would conventionally denote

[82] For instance, the formula for addition must satisfy $\texttt{Add}(X, Y, Z)$ if and only if $[\![X]\!]_2 + [\![Y]\!]_2 = [\![Z]\!]_2$ and be such that $\mathsf{V}^1 \vdash \forall X \forall Y \exists Z \texttt{Add}(X, Y, Z)$. A formula with this property can be constructed by first observing that the formula $\texttt{Carry}(i, X, Y) \leftrightarrow \exists k < i[X(k) \wedge Y(k) \wedge \forall j < i(k \leq j \to (X(j) \vee Y(j)))]$ is true just in case a $\texttt{1}$ would be carried in the ith column were we to perform a carry addition on X and Y. We may now define $\texttt{Add}(X, Y, Z)$ as $(|Z| < |X|+|Y| \wedge \forall i < |X|+|Y|[Z(i) \leftrightarrow (X(i) \oplus Y(i) \oplus \texttt{Carry}(i, X, Y))]$ where $\oplus$ denotes exclusive or.

[83] In fact since the binary representation of 2^x contains a 1 in exactly its ith position, $\texttt{Pow2}(x, Y) \leftrightarrow \forall z < x'(z \in Y \leftrightarrow z = x)$.

by 2^d. And from this it follows that the reduct of $\mathfrak{M}$ to $\mathcal{L}_a$ is a nonstandard model whose domain is not closed under exponentiation.

Generalizing from the construction of §4, it might now appear that a model of $\mathsf{T}_{\underline{m}}$ can also be constructed by letting $C \in U_2$ be the denotation of the constant $\underline{m}$ which lies outside the interpretation of the sortical predicate $F(x)$. Note, however, that since $\mathsf{T}_{\underline{m}}$ is a one-sorted theory, we then must regard $F(x)$ as a predicate of first-sort objects. Taking $\underline{m}^{\mathfrak{M}} = C$ and then asserting $\neg F(\underline{m})$ would thus be tantamount to a *typing error*.[84]

Another approach to applying the two-sort framework to obtain a model of a theory similar to $\mathsf{T}_{\underline{m}}$ is to note that if we adjoin the premises (F_0) and (F_s) to V^1, we may define a second-sort version of the soritical predicate $F(x)$ via the definition

$$F^2(X) = \exists y(\mathtt{Bin}(X, y) \wedge F(y))$$

Note that we may view the objects comprising the second-sort domain of a model V^1 as binary numerals which come equipped with a definable order $<_2$ satisfying $X <_2 Y$ iff $[\![X]\!]_2 < [\![Y]\!]_2$.[85] Relative to this ordering, we can now define a second-sort constant $\mathtt{0}_2$ and a second-sort function symbol $\mathtt{s}_2(X)$ which respectively define the first element and the successor function on this order. It is now possible to use these definitions to show that the second-sort analogues of (F_0) and (F_s)

(F_0^2) $F^2(\mathtt{0}_2)$
(F_s^2) $\forall X(F^2(X) \rightarrow F^2(\mathtt{s}_2(X))$

are provable from the original principles (F_0) and (F_s) over V^1.[86]

It follows from Proposition A.1 that if we replace $\underline{m}$ with a new second-sort constant C, then the theory $\mathsf{R}_{\mathtt{C}} = \mathsf{V}^1 + (F_0) + (F_s) + \neg F(\mathtt{C})$ can be shown to possess a model $\mathfrak{M} = \langle U_1, U_2, \leq, ', +, \times, |\cdot|, 0\rangle$. The reduct $\mathfrak{M}^1$ of $\mathfrak{M}$ to its first-order domain will look like the model of $\mathsf{V}^1 + \neg\forall X \exists y \mathtt{Bin}(X, y)$ constructed above. But we may similarly define an induced $\mathcal{L}_a$-structure $\mathfrak{M}^2 = \langle U_2, <, ', +, \times, 0\rangle$ by taking the interpretations of the formulas defining $\mathtt{s}_2(x), <_2 \mathtt{Add}(x, y, z), \mathtt{Mult}(x, y, z)$ and $\mathtt{0}_2$ on U_2 as the respective interpretations of the symbols of $\mathcal{L}_a$.

The provability of (F_0^2) and (F_s^2) in $\mathsf{R}_{\mathtt{C}}$ will mean that $F^2(X)$ will define a proper cut $I_2 \subsetneq U_2$ in $\mathfrak{M}^2$ containing the binary representations of the numbers which are in the cut defined by $F(x)$ in U_1. Note, however, that $\mathtt{C}$ will now denote a member of $M_2 - I_2$ – i.e. an object which is the denotation of a binary numeral in the sense of $\mathfrak{M}_2$ but which is nonstandard with respect to $<_2$ when viewed externally. It may additionally be shown that a model constructed in this way from a model $\mathfrak{M} \models \mathsf{V}^1$ is also a model of Buss's (1986) first-order theory S_2^1 whose provably recursive functions also correspond to those computable in polynomial time.[87] $\mathfrak{M}^2$ thus has the same structural features as the model obtained in the proof of Proposition 4.1. However, unlike this single-sorted model, $\mathfrak{M}^2$ is

---

84 This observation can also be understood as providing the neo-feabilist with a principled reply to the concern raised in note 40 that in the model $\mathcal{M} \models \mathsf{T}_{\underline{m}}$ constructed in Proposition 4.1 we will have $\mathcal{M} \models \underline{m} \neq \overline{m}$. It is tempting to view such a statement as expressing that the unary and positional numerals denoting m are unequal. But from the perspective of the two-sorted theories now under consideration, such cross-sort identity statements are regarded as reflecting a mistaken understanding of the mechanism by which unary and positional numerals denote numbers.

85 Such an ordering is definable by a Σ_0^B-formula as $X < Y \leftrightarrow (|X| \leq |Y| \vee \exists i[Y(i) \wedge \neg X(i) \wedge \forall j \leq i(X(j) \rightarrow Y(j)]$ formalizing "X is less than Y just in case X is shorter than Y or the first bit i on which X and Y differ is such that $Y(i)$ and $\neg X(i)$".

86 This formulation should be compared to that given by Boolos (1991, p. 702) in his one-sorted system.

87 This is the content of one half of the so-called *RSUV isomorphism*. See (Cook and Nguyen, 2010, VI.2.2 and VIII.8.2).

embedded in $\mathfrak{M}$ in a manner which better accords with the the distinct role played by unary and positional numerals in our reasoning about natural numbers.

Acknowledgements Versions of this paper have been presented in Bucharest, Birmingham, Paris, Helsinki, and Irvine. Thanks are owed to the organizers of these events and to Susanne Bobzien, Michael Detlfesen, Kit Fine, Hidenori Kurokawa, Nick Jones, Richard Kaye, David Miller, Jan van Plato, Penelope Maddy, William Stafford, Kai Wehmeier, and Sean Walsh for comments.

Bibliography

Arora, S. and Barak, B. (2009). *Computational complexity: a modern approach*. Cambridge University Press.

Barker, C. (2002). The dynamics of vagueness. *Linguistics and Philosophy*, 25(1):1–36.

Barnes, J. (1982). Medicine, Experience and Logic. In Barnes, J., Brunschwig, J., Burnyeat, M. F., and Schofield, M., editors, *Science and Speculation*. Cambridge University Press.

Benacerraf, P. and Putnam, H., editors (1983). *Philosophy of Mathematics: Selected readings*. Prentice-Hall, Englewood Cliffs, NJ, second edition.

Bernays, P. (1935). Sur le platonisme dans les mathématiques. *L'enseignement mathematique*, 34:52– 69. Reprinted in Benacerraf and Putnam (1983), pp. 58-71.

Boolos, G. (1991). Zooming down the slippery slope. *Nous*, 25(5):695–706.

Borel, E. (1952). *Les nombres inaccessibles*. Gauthier-Villars.

Buss, S. (1986). *Bounded Arithmetic*. Bibliopolis.

Carbone, A. (1996). Cycling in proofs, feasibility and no speed-up for nonstandard arithmetic. Technical Report IHES/M/96/55, Institut des Hautes Études Scientifiques.

Cobreros, P., Egré, P., Ripley, D., and van Rooij, R. (2012). Tolerant, classical, strict. *Journal of Philosophical Logic*, 41(2):347–385.

Cook, S. (1975). Feasibly constructive proofs and the propositional calculus (preliminary version). In *Proceedings of seventh annual ACM symposium on Theory of computing*, pages 83–97. ACM.

Cook, S. and Nguyen, P. (2010). *Logical foundations of proof complexity*. Cambridge University Press.

Dean, W. (2015). Computational complexity theory. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Fall 2015 edition.

Dean, W. and Walsh, S. (2017). The prehistory of the subsystems of second-order arithmetic. *The Review of Symbolic Logic*, 10(2):357–396.

Dietz, R. (2011). The paradox of vagueness. In *Continuum Companion to Philosophical Logic*. Continuum.

Dragalin, A. G. (1985). Correctness of inconsistent theories with notions of feasibility. In Skowron, A., editor, *Computation Theory: Fifth Symposium, Zaborow, Poland*, volume 208 of *Lecture Notes in Computer Science*, pages 58–79. Springer.

Dummett, M. (1959). Wittgenstein's philosophy of mathematics. *The Philosophical Review*, pages 324–348.

Dummett, M. (1975). Wang's Paradox. *Synthese*, 30(3/4):301–324.

Feferman, S. et al., editors (2003). *Kurt Gödel Collected Works. Vol. IV. Publications Correspondence A-G*. Oxford Univeristy Press.

Fernando, T. and Kamp, H. (1996)). Expecting many. In T., G. and Spence, J., editors, *Proceeding from SALT VI*, pages 53–68. CLC Publications.

Field, H. (2008). *Saving Truth from Paradox*. Oxford University Press, Oxford.

Fine, K. (1975). Vagueness, truth and logic. *Synthese*, 30(3):265–300.

Frege, G. (1879). *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Nebert, Halle.

Frege, G. (1884). *Die Grundlagen der Arithmetik*. Koebner, Breslau.

Frege, G. (1893, 1903). *Grundgesetze der Arithmetik: begriffsschriftlich abgeleitet*. Pohle, Jena.

Gaifman, H. (2004). Non-standard models in a broader perspective. In Enayat, A. and Kossak, R., editors, *Non-standard models of arithmetic and set Ttheory*, pages 1–22. AMS.

Gaifman, H. (2010). Vagueness, tolerance and contextual logic. *Synthese*, 174:1–42.

Gandy, R. (1982). Limitations to mathematical knowledge. In van Dalen, D., Lascar, D., and Smiley, J., editors, *Logic Colloquium 80*, volume 108, pages 129–146. Elsevier.

Ganea, M. (2010). Two (or three) notions of finitism. *The Review of Symbolic Logic*, 3(01):119–144.

Geiser, J. R. (1974). A Formalization of Essenin-Volpin's Proof Theoretical Studies by Means of Nonstandard Analysis. *The Journal of Symbolic Logic*, 39(1):pp. 81–87.

Gómez-Torrente, M. (2010). The sorites, linguistic preconceptions, and the dual picture of vagueness. In Dietz, R. and Moruzzi, S., editors, *Cuts and Clouds: Vaguenesss, Its Nature and Its Logic*. Oxford University Press.

Graff, D. (2001). Phenomenal continua and the sorites. *Mind*, 110(440):905–936.

Hájek, P. (1973). Why semisets? *Commentationes Mathematicae Universitatis Carolinae*, 14(3):397–420.

Hájek, P. and Pudlák, P. (1998). *Metamathematics of First-Order Arithmetic*. Springer.

Hallett, M. and Majer, U., editors (2004). *David Hilbert's Lectures on the Foundations of Geometry, 1891–1902*, volume 1 of *David Hilbert's Foundational Lectures*. Springer.

Heck, R. G. (2014). Predicative Frege Arithmetic and 'Everyday' Mathematics. *Philosophia Mathematica*, 22(3):279–307.

Hyde, D. (2011). The sorites paradox. In *Vagueness: A Guide*, pages 1–17. Springer.

Isles, D. (1981). Remarks on the notion of standard non-isomorphic natural number series. In *Constructive Mathematics*, pages 111–134. Springer.

Kamp, H. (1981). The paradox of the heap. In *Aspects of philosophical logic*, pages 225–277. Springer.

Kanovei, V. and Reeken, M. (2004). *Nonstandard analysis, axiomatically*. Springer.

Kaye, R. (1991). *Models of Peano Arithmetic*, volume 15 of *Oxford Logic Guides*. Oxford University Press.

Keefe, R. (2000). *Theories of vagueness*. Cambridge University Press.

Koellner, P. (2009). Truth in Mathematics: The Question of Pluralism. In Linnebo, O. and Bueno, O., editors, *New Waves in the Philosophy of Mathematics*, pages 80–116. Palmgrave.

Kölbel, M. (2010). Vagueness as semantic. In Dietz, R. and Moruzzi, S., editors, *Cuts and Clouds: Vaguenesss, Its Nature and Its Logic*. Oxford University Press.

Krantz, D., Luce, R., Suppes, P., and Tversky, A. (1971). *Foundations of measurement: Additive and polynomial representations*, volume I. Academic Press.

Kreisel, G. (1958). Wittgenstein's remarks on the foundations of mathematics. *The British Journal for the Philosophy of Science*, 9(34):135–158.

Kreisel, G. (1967a). Informal Rigour and Completeness Proofs. In *Problems in the Philosophy of Mathematics*, pages 138–186. North-Holland, Amsterdam.

Kreisel, G. (1967b). Review of "Le programme ultra-intuitionniste des fondements des mathématiques." (Volpin). *Zentrablatt Math*, 134:9–10.

Kreisel, G. and Ehrenfeucht, A. (1967). Review of "Le programme ultra-intuitionniste des fondements des mathmatiques" (Volpin). *The Journal of Symbolic Logic*, 32(4):p. 517.

Lassiter, D. (2011). Measurement and modality: The scalar basis of modal semantics. *New York: New York University dissertation*.

Luce, D., Krantz, D., Suppes, P., and Tversky, A. (1990). *Foundations of measurement: Representation, axiomatization, and invariance*, volume III. Dover.

Luce, R. D. (1956). Semiorders and a theory of utility discrimination. *Econometrica, Journal of the Econometric Society*, pages 178–191.

Magidor, O. (2011). Strict Finitism and the Happy Sorites. *Journal of Philosophical Logic*, pages 1–21.

Mannoury, G. (1931). *Woord en gedachte: een inleiding tot de signifika, inzonderheid met het oog op het onderwijs in de wiskunde*. Noordhoff.

Mostowski, A. (1950). Some impredicative definitions in the axiomatic set-theory. *Fundamenta mathematicae*, 38:110–124.

Narens, L. (1985). *Abstract measurement theory*. MIT Press.

Nelson, E. (1977). Internal set theory: a new approach to nonstandard analysis. *Bulletin of the American Mathematical Society*, 83(6):1165–1198.

Nelson, E. (1986). *Predicative Arithmetic*, volume 32 of *Mathematical Notes*. Princeton University Press.

Novák, V. (1992). *The alternative mathematical model of linguistic semantics and pragmatics*. Springer.

Parikh, R. (1971). Existence and feasibility in arithmetic. *Journal of Symbolic Logic*, 36(3):494–508.

Parikh, R. (1983). The problem of vague predicates. In *Language, logic and method*, pages 241–261. Springer.

Poincaré, H. (1906). Les mathématiques et la logique. *Revue de métaphysique et de morale*, 14(3):294–317.

Putnam, H. (1980). Models and reality. *The Journal of Symbolic Logic*, 45(3):464–482.

Raffman, D. (1994). Vagueness without paradox. *The Philosophical Review*, pages 41–74.

Robinson, A. (1966). *A Non-Standard Analysis*. North-Holland.

Rogers, H. (1987). *Theory of recursive functions and effective computability*. MIT Press, Cambridge, MA, second edition.

Russell, B. (1908). Mathematical logic as based on the theory of types. *American journal of mathematics*, 30(3):222–262.

Sainsbury, M. (1991). Is there higher-order vagueness? *The Philosophical Quarterly*, pages 167–182.

Sainsbury, R. (1995). *Paradoxes*. Cambridge University Press.

Sainsbury, R. M. (1990). Concepts without boundaries. In Keefe, R. and Smith, P., editors, *Vagueness: A reader*. MIT Press.

Sassoon, G. W. (2010). Measurement theory in linguistics. *Synthese*, 174(1):151–180.

Sazonov, V. (1995). On feasible numbers. In Leivant, D., editor, *Logic and Computational Complexity*, volume 960 of *Springer Lecture Notes in Computer Science*. Springer.

Scott, D. and Suppes, P. (1958). Foundational aspects of theories of measurement. *Journal of Symbolic logic*, pages 113–128.

Shapiro, S. (2011). Vagueness and logic. In *Vagueness: A Guide*, pages 55–81. Springer.

Sheard, M. (1998). Induction the hard way. *The American Mathematical Monthly*, 105(4):pp. 348–353.

Simpson, S. (2009). *Subsystems of second order arithmetic*. Cambridge University Press, Cambridge, second edition.

Skala, H. J. (1975). *Non-Archimedean utility theory*, volume 9. Springer.

Solovay, R. M. (1976). Interpretability in set theories. Letter to Petr Hájek.

Sorensen, R. (2001). *Vagueness and contradiction*. Oxford University Press.

Sorensen, R. A. (1988). *Blindspots*. Clarendon, Oxford.

Steiner, M. (2011). Kripke on logicism, Wittgenstein, and *de re* beliefs about numbers. In Berger, A., editor, *Saul Kripke*. Cambridge University Press.

Suppes, P., Krantz, D., Luce, D., and Tversky, A. (1989). *Foundations of Measurement: Geometrical, threshold, and probabilistic representations*, volume II. Academic Press.

Tait, W. W. (1981). Finitism. *Journal of Philosophy*, 78(9):524–546.

Tye, M. (1994). Sorites paradoxes and the semantics of vagueness. *Philosophical Perspectives*, pages 189–206.

Tzouvaras, A. (1998). Modeling vagueness by nonstandardness. *Fuzzy sets and systems*, 94(3):385–396.

Unger, P. (1979). There are no ordinary things. *Synthese*, 41(2):117–154.

van Dantzig, D. (1955). Is $10^{10^{10}}$ a finite number? *Dialectica*, 9(3-4):273–277.

van Rooij, R. (2011). Vagueness and linguistics. In Ronzitti, G., editor, *Vagueness: A Guide*. Springer.

Vopěnka, P. (1979). *Mathematics in the Alternative Set Theory*. Teubner, Leipzig.

Wang, H. (1958). Eighty years of foundational studies. *Dialectica*, 12(3-4):466–497.

Wang, H. (1990). *Computation, Logic, and Philosophy*. Kluwer, Boston.

Weber, Z. and Colyvan, M. (2010). A topological sorites. *The Journal of Philosophy*, 107(6):311–325.

Williamson, T. (1994). *Vaugeness*. Routledge, London and New York.

Wright, C. (2010). The illusion of higher-order vagueness. In Dietz, R. and Moruzzi, S., editors, *Cuts and Clouds: Vagueness, Its Nature, and Its Logic*, pages 523–549. Oxford Univeristy Press.

Wright, W. and Pitt, F. (1934). Hue-discrimination in normal colour-vision. *Proceedings of the Physical Society*, 46(3):459.

Yessenin-Volpin, A. (1961). Le programme ultra-intuitionniste des fondements des mathématiques. In *Infinitistic Methods, Proceedings of the Symposium on the Foundations of Mathematics*, pages 201–223. PWN.

Yessenin-Volpin, A. (1970). The ultra-intuitionistic criticism and the antitraditional program for the foundations of mathematics. In Kino, A., Myhill, J., and Vesley, R., editors, *Intuitionism and Proof Theory*. North Holland.

Yessenin-Volpin, A. (1981). About infinity, finiteness and finitization (in connection with the foundations of mathematics). In *Constructive mathematics*, pages 274–313. Springer.

Yessenin-Volpin, A. (2011). Letter to Brouwer (dated 23-7-1958). In van Dalen, D., editor, *The selected correspondence of L.E.J. Brouwer*. Springer.

Zambella, D. (1996). Notes on polynomially bounded arithmetic. *The Journal of Symbolic Logic*, 61(3):pp. 942–966.

Zermelo, E. (1929). Über den Begriff der Definitheit in der Axiomatik. *Fundamenta Mathematicae*, 14:339–344.