Modeling vagueness by nonstandardness

Athanassios Tzouvaras

Department of Mathematics
Aristotle University of Thessaloniki
541 24 Thessaloniki, Greece
e-mail: tzouvara@math.auth.gr

Abstract

We are concerned with a mathematical modeling of *vague predicates* and *vague partitions* (or *equivalences* or *similarities*) by the help of nonstandard sets of integers. Nevertheless, it is not applicable to concrete phenomena, since nonstandard numbers are too big to be used in actual counting. The properties of a vague measurable similarity $\sim$ which we consider as most important and which are captured in the present modeling are: To $\sim$ there corresponds an assignment $\mu(x)$ of integral values to the objects of the domain of discourse, and a corresponding distance $d(x, y) = |\mu(x) - \mu(y)|$, such that:

(i) If $x \sim y$ and $d(x, z) \leq d(x, y)$, then $x \sim z$.

(ii) For any two $\sim$-similar objects x, y and for any two $\sim$-non-similar objects x', y', $d(x, y) < d(x', y')$.

(iii) For any x, y such that $x \sim y$, we can find a z in the same class, i.e., $x \sim y \sim z$, such that $d(x, y) < d(x, z)$.

(iv) For any $\sim$-non-similar objects x, y, such that $\mu(x) < \mu(y)$, there is a third object z, non-similar to both x and y such that $\mu(x) < \mu(z) < \mu(y)$.

Further we discuss the issue of what a "right" nonstandard model of numbers would be and some related questions.

Keywords: Vague predicate, nonstandard model of arithmetic, cut, saturation.

1 The aim of this approach and comparison with other ones

In this paper we are concerned with a mathematical representation-interpretation of vagueness. The emphasis is on its theoretical aspects rather than its practical ones. It means that the representation succeeds in modeling the most essential characteristics of the notion, but it fails to be applicable to concrete vague situations. This is because the representation makes use of "generalized" integers (called *nonstandard*) which, though have all the familiar properties with respect to addition, multiplication, etc., in fact *exceed* all standard ones, and hence are absolutely infeasible objects.

Apart from applications however, one might be interested in the *structure* of vagueness, and this is what we are primarily trying to grasp in this paper through mathematical modeling. The main features of this structure are the following:

(i) There are *no jumps* when passing from a vague predicate P to its negation $\neg P$, or from one class of a vague partition $\sim$ to another. The transition is smooth and continuous.

(ii) Despite (i), one always believes that any two objects x, y are *objectively* similar or dissimilar with respect to the $(P, \neg P)$-partition or to $\sim$, as if there were a strict distance relation $d(x, y)$ among them, such that $x \sim y$ if and only if $d(x, y) <$ *something* (this *something*, clearly, cannot be an ordinary number if we want $\sim$ to be transitive).

(iii) In the case of a partition $\sim$, given any two classes of it (colors, species, etc.), we always consider it possible to find an intermediate class for them, lying between, as if the totality of the $\sim$-classes were actually dense with respect to the distance d above, and eventually infinite. (For example, between the zebra and the horse, there had always been room for an intermediate species. And indeed the missing hybrid, called "zebroid" was recently obtained.)

We distinguish two basic manifestations of vagueness:

(a) *Vague predicates* or, equivalently, vague two-class partitions of the universe into objects satisfying P and those satisfying $\neg P$ (which also generalize to n-class partitions for finite n).

(b) *Vague partitions* of the universe (or some totality of objects) with infinitely many classes, which correspond to various taxonomies of objects into "kinds".

Of course, apart from predicates and partitions, there are also *vague objects*. A vague object is one with uncertain boundaries like Atlantic Ocean, or Mount Everest or "that tree". But, clearly, any such object can be identified with the *set* of elementary units of matters (e.g. molecules) comprised within the boundaries, or the *property* $P(x)$: "to be a molecule of". Thus objects exhibit essentially the same type of vagueness as that of (a). So we may confine ourselves to vague predicates and vague partitions.

Of the preceding two manifestations of vagueness, we believe that the second one is by far more genuine and interesting. Vagueness of predicates, on the contrary (hence also vagueness of objects), admits *precisifications.* The ability to make a predicate P more and more precise, by sharpening more and more the boundaries separating P from non-P, in fact invalidates vagueness itself as a notion of theoretical significance, and the issue enters the scope of the usual approximation techniques. This is particularly clear in the case of (physical) objects (seas, mountains, etc.) where the boundaries are certain *lines* drawn by mere convention. Whether a particular molecule of water is or is not part of Atlantic Ocean is what we call in this paper *conventional vagueness.* It has to do mainly with the problem of measurement of continuous magnitudes using reals. Yet, this is the type of vagueness that most authors (e.g. [1], [6]) are primarily engaged in.

According to Kit Fine, vagueness is "*deficiency of meaning. As such, it is to be distinguished from generality, undecidability and ambiguity*" [1, p. 265]. An immediate outcome of this deficiency is of course the existence of truth-value gaps. Vague predicates like "small", "green", "clever", etc., cannot be said true or false for certain borderline arguments. So the approaches taken by many authors aim at remedying this situation by raising the gaps. This "logicistic" approach has two subdirections: (i) the so called "truth-value" approach (Fine's terminology) is attempting to raise the gaps by modifying appropriately the underlying logic (typical work of this kind is [7], using Kleene's three-valued logic to this effect), and (ii) the "super-truth-value" approach, mainly taken by Fine [1].

The present paper does not share any of the aims of logicistic approach. Quite modestly it seeks to describe a kind of "structure" for vagueness, by giving *quantitative* expression to the characteristic features of the notion. If we should give a name to such an approach we might call it "representational".

This paper bears a certain analogy with Parikh's [5] in the following sense. Parikh's idea is that the main source of vagueness is measuring by using reals, namely the discrepancy between the "real sizes" of things, expressed by real numbers, and the outcomes of our actual measurements which are always rational numbers. So he proposes a theory of "vague real numbers" expressed as intervals $[p, q]$ of rationals. In accordance with this he devices a specific truth-value assignment for propositions containing claims of that type. The analogy lies in that we also shift from measurement with standard integers to that with nonstandard ones, and further to measurement with "cuts", which may be thought of as "generalized" numbers. But the analogy stops here. As already said above, we do not think that the kind of vagueness manifested in measurement impreciseness is of real interest. The question e.g of whether a certain piece of matter belongs to Mount Everest or not (the example is extensively treated in [7]) is of conventional character. There is no compelling similarity among pieces of matter belonging to Mount Everest, which breaks down when a piece belongs and another does not to the Mount.

In contrast, the question of whether a given citrus is a lemon or not is not conventional; it is a matter of taxonomy obeying objective criteria. In terms of the distance d, the conventionality of the first partition is revealed by the fact that we can find a point x belonging to Everest and a point y not belonging to it, *arbitrarily close* to each other. While the citrus c will be a lemon if and only if, given a lemon x and a non-lemon y, $d(c,x) < d(c,y)$. In other words, a lemon and a non-lemon are always *farther apart* than any two lemons. It is these natural vague predicates and partitions, not produced by conventional divisions of the reality, that we shall be mainly interested in here.

More related to the present work seems to be Novák's work [4], but this is only superficially true. Novák is working in P. Vopeěnka's alternative set theory (AST) (a theory designed to deal with *semisets*) and is concerned with a very special approximation of semisets by certain sets that he calls *fuzzy*. As a consequence, his results are technical contributions to the theory AST rather than contributions to the phenomenon of fuzziness itself. Of course the universe of AST is a *nonstandard* one, in the sense precisely used here, hence suitable for representing vague collections. However, as we shall explain in the last section, "semisets" are not entirely new mathematical objects. They are in fact nonstandard segments of numbers dressed in set-theoretic covers. Therefore, instead of exploring the connection of vagueness with semisets, it seems much more natural to go deeper and explore properly the connection of vagueness with the very source of semisets, namely nonstandardness itself.

2 Vague predicates and their measures

A typical vague predicate (concerning numbers) is $P(n)$: *"n is small"*, or (concerning people) $B(x)$: *"x is bald"*.

In the first example, the meaning of the predicate $P(n)$ is complete (and the truth-value "true") for as long as we keep down to admittedly small numbers. The deficiency of meaning begins and grows larger and larger, as we are going upward to bigger and bigger numbers. As the meaning is gradually obscured, the truth-value gaps emerge. But if we continue to move farther, and when we will have been adequately far away, the meaning will be gradually restored and soon will become clear again (together with the truth-value "false"). If we denote by $P(n)\uparrow$ the fact that "$P(n)$ is neither true nor false", the above situation concerning the predicate P can be illustrated in Fig. 1.

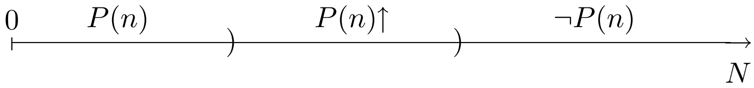

Fig. 1

The sign ")" used to denote the corresponding separation of the set N of integers into intervals, does not indicate a cut-point but just a *gap*, an ordinary order-gap, i.e., a cut of the number line into two segments of which the first one has no last and the second one has no first point. This is because if n is small, then apparently $n+1$ must be small too. The preceding separation consists of three intervals of N but the number of intervals is not a key-factor of the situation. On the contrary, it may prove misleading. The intermediate interval $P(n)\uparrow$ simply represents the agent's period of uncertainty with respect to the supposed objective fact of $P(n)$ or $\neg P(n)$. We similarly might assume another period of uncertainty between $P(n)$ and $P(n)\uparrow$, as well as between $P(n)\uparrow$ and $\neg P(n)$, and so on ad infinitum. Instead, it is much simpler to ignore intervals of subjective uncertainty and admit just one vague transition, namely from $P(n)$ to $\neg P(n)$, the latter supposed to be objective facts (Fig. 2).

Figs. 1 and 2 already depict a rather strange situation, since N appears to contain bounded subsets without greatest element, a fact contradicting the induction principle. (In fact this is the Sorites or Wang's Bald Man Paradox.)

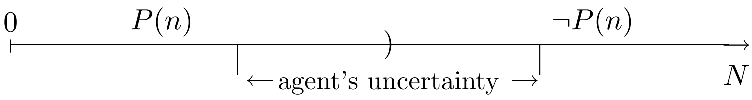

Fig. 2

Concerning now the predicate $B(x)$ of baldness, we have two options: The hairy part of a man's head can be considered to be either a continuum, and thus its *area* should be measured by the help of reals, or a discrete set, in which case we have to *count* the hairs one by one by the help of integers. In both cases we assign to each man x a measure $\mu(x)$, where $\mu(x) \in R$ or $\mu(x) \in N$, which "measures" baldness in the following sense:

(1) If x is bald and $\mu(y) \leq \mu(x)$, then y is bald too.

(2) If x is not bald and $\mu(x) \leq \mu(y)$, then y is not bald too.

As a result of (1) and (2) we have immediately that

(3) For every bald x and for every nonbald y, $\mu(x) < \mu(y)$.

(1)-(3) above hold independently of whether truth-value gaps exist or not. This, in turn, depends on how we measure baldness. If we measure it as an area, i.e., $\mu(x) \in R$, then the only thing we can do is specify some real a, and *define* (by convention) x to be bald if the hairy part of his head is less than or equal to a; else he is nonbald. In such a case, and in view of the agreements (1) and (2), the only uncertainty stems from the process of measuring and concerns just men of hairy area very close to a. We have, in fact, a quite thin gap of meaning, created by the one-point gap of the dense ordering of R as shown in Fig. 3.

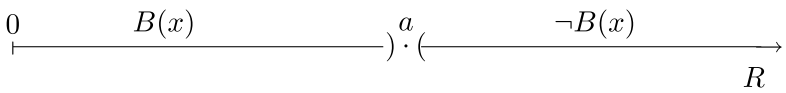

Fig. 3

The most striking defect of measuring baldness by reals is that we lose the sense of "internal affinity" between members belonging to the same side of the boundary a. Specifically, one expects any two bald men or any two nonbald men to be closer to each other (w.r.t. this characteristic of course), than a bald and a nonbald man. But once nearness is measured by the distance $|\mu(x) - \mu(y)|$, any notion of affinity collapses, since the distance between objects from both sides of a can be arbitrarily small, hence smaller than the distance between objects on one side. We see this in Fig. 4 below, where we find the bald man y to be "farther" from the bald man x than from the nonbald z. Fig. 4 exemplifies what we previously called "conventional vagueness".

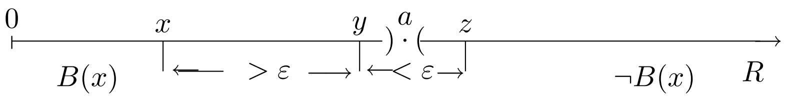

Fig. 4

If, on the other hand, $\mu(x)$ counts the number of hairs of a man's head as discrete items, the situation is similar to that of the predicate $P(n)$. There are objectively bald man (for $\mu(x)$ small) and nonbald men (for $\mu(x)$ big) only, though the observer may have difficulty to decide (for $\mu(z)$ such that $P(\mu(z)) \uparrow$) (see Fig. 5).

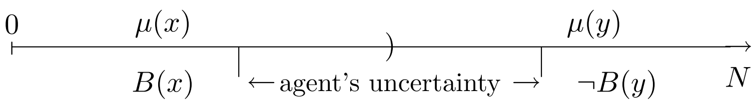

Fig. 5

The situation of having bounded subsets of N with no last elements appears again (whence the name "Bald Man Paradox").

The preceding examples motivate the following definition.

Definition 2.1 A predicate $P(x)$ on a set X is said to be *measurable* if there is a function $\mu : X \to N$ such that either for all $x, y \in X$,

(i) $\mu(x) \leq \mu(y) \ \&P(y) \ \Rightarrow P(x)$,

or for all $x, y \in X$,

(ii) $\mu(x) \leq \mu(y) \ \&P(x) \ \Rightarrow P(y)$.

(Since, clearly, P satisfies (i) if and only if $\neg P$ satisfies (ii), P is measurable if and only if $\neg P$ is measurable.) A measurable predicate is said to be *vague* if either $\{\mu(x) : P(x)\}$ or $\{\mu(x) : \neg P(x)\}$ is bounded and contains no greatest element.

A great deal of commonly used vague predicates, like "tall", "far", "heavy", "big", "strong", etc., are obviously measurable. But even "red", "dark", "soft", "smooth", etc., may become measurable by choosing more refined techniques of measurement. Others, like "nice", "ugly", "happy", "fair", "moral" and so on, may remain for ever essentially nonmeasurable, but to the extend they do so, we are allowed to consider them *ambiguous* rather, than simply vague, i.e., as exhibiting a high degree of personal or social taste and habit.

3 Vague partitions and their measures

The case of partitions differs from that of predicates in the number of equivalence classes produced. Instead of having just two ($\{x : P(x)\}$ and $\{x : \neg P(x)\}$) as in the case of predicates, we now have, in general, infinitely many classes $[x] = \{y : y \sim x\}$, for x belonging to the ground set. For most partitions $\sim$, we can again attribute values $\mu(x)$'s to the values x's (preferably integers), and the distance $d(x, y) = |\mu(x) - \mu(y)|$ represents the nearness relation corresponding to $\sim$. For instance, if $\sim$ is the sameness of color, $\mu(x)$ could be the wavelength measured in an appropriate scale. Then the desired properties that d should satisfy in order for $\sim$ to be a *vague* partition are the following:

(i) If $x \sim y$ and $d(x, z) \leq d(x, y)$, the $x \sim z$ (simply by the fact that d "measures" $\sim$).

(ii) For any two $\sim$-similar objects x, y, and for any two $\sim$-nonsimilar objects x', y', $d(x, y) < d(x', y')$. This captures the intuitive idea that any two objects within the same color shade are "closer" to each other than any two of distinct shades.

(iii) For any x, y such that $x \sim y$, we can find a z in the same class, i.e., $x \sim y \sim z$, such that $d(x, y) < d(x, z)$. We can always move farther and farther in the same class.

(iv) For any $\sim$-nonsimilar objects x, y, there is always a third one z non-similar to both x and y but "lying between" them, e.g. $d(x, z) = d(y, z) = d(x, y)/2$. This represents the fact that the classes of $\sim$ are densely spread, or that no class has an immediate neighbor. This is clearly the case of colors, species, etc.

Typical of this kind are the partitions (equivalences) $x \sim y$: "*x, y are of the same color*", or "*x, y are of the same species*". The symbol $\sim$ will be used for any such partition. Another name for this relation is "similarity", in which case the equivalence class is called "kind". The notion of similarity lies at the very basis of our knowledge about the world, since every cognitive activity comprises classification of things into kinds. As Quine puts it in [6]:

> "The notion of a kind and the notion of similarity or resemblance seem to be variants or adaptations of a single notion. Similarity is immediately definable in terms of kind; for, things are similar when they are of a kind. (...) We cannot imagine a more familiar or fundamental notion than this, or a notion more ubiquitous in its applications. On this score it is likely the notions of logic: like identity, negation, alternation and the rest. And yet, strangely, there is something logically repugnant about it. One's first hasty suggestion might be to say that things are similar when they have all or most or many properties in common. But any such course only reduces our problem to the unpromising task of settling what to count as a property [6, p. 117].

Yet a paradox is lurking here. The term "similarity" is commonly attributed to relations weaker than equivalences, namely those which are reflexive and symmetric but miss transitivity. If $\mu(x)$'s are as before measures assigned to the objects x's of our domain, then typical similarity relations between them are the relations $\sim_n$, for $n \in N$, where

$$x \sim_n y \text{ iff } |\mu(x) - \mu(y)| \leq n.$$

The failure of transitivity is thought not only to be in accordance with vagueness, but even to capture it. The classical view holds that transitivity and vagueness are incompatible. If we wish to save transitivity, then we must drop vagueness and vice-versa. The paradox here is that at the same time, when talking about kinds, we think of them as normal non-overlapping classes which partition the domain of discourse, though such classes can be produced by nontransitive relations. When talking about colors, for example, we implicitly assume that colors are distinct non-overlapping and exhaust the range of colorful objects. There is no pink which is, at the same time, red, and no yellow which is also orange. Yet the boundary between yellow and orange remains unclear.

A way out of the paradox is to show that vagueness and transitivity *are* in fact compatible. That the relation $\sim$, though a real equivalence, *can* produce classes with unclear boundaries. Such relations are made possible in the nonstandard approach, if we replace the similarity $\sim_n$ by $\sim_I$:

$$x \sim_I y \text{ iff } |\mu(x) - \mu(y)| \in I,$$

where I is a cut (cf. section 5), and show $\sim_I$ to have the right properties. The properties we want a vague partition to satisfy have already been mentioned in the beginning of this section (while their motivation was also discussed in section 1). Let us restate them here in the form of an official definition.

Definition 3.1 An equivalence $\sim$ on a set X is said to be *measurable,* if there is a function $\mu : X \to N$ such that, if we put $d(x, y) = |\mu(x) - \mu(y)|$, then the distance d has the following properties:

(i) If $x \sim y$ and $d(x, z) \leq d(x, y)$, then $x \sim z$.

(ii) For any x, y such that $x \sim y$ and x', y' such that $x' \not\sim y'$, $d(x, y) < d(x', y')$.

The measurable equivalence $\sim$ is said to be *vague* if, in addition, it holds that:

(iii) For any (or, at least, some) x, the set $\{d(x, y) : y \sim x\}$ has no greatest element, and

(iv) For any x, y such that $x \not\sim y$, and $\mu(x) < \mu(y)$, there is a z such that $x \not\sim z$ and $y \not\sim z$ and $\mu(x) < \mu(z) < \mu(y)$.

The problem both with vague predicates and partitions as defined in Definitions 2.1 and 3.1 is just (!?) their existence. Specifically, it is clear from the examples cited that vague partitions in the sense of Definition 2.1 exist if and only if bounded sets of integers without last element, like those in Figs. 1 and 5, exist. We will also see that if such sets exist, then we can easily find partitions satisfying the properties of Definition 3.1. Thus the crucial factor is proved to be the structure of the system of numbers by which we measure the objects of our domain of discourse. It is therefore necessary to give some background information on such systems without heavy technical details. This will be done in the next two sections, where also the existence of relations defined in Definitions 2.1 and 3.1 is shown. The reader interested in more technical details of the facts mentioned may consult, say, [2].

4 Basic facts about models of arithmetic and correlations with vague predicates

The set of integers $\omega = \{0, 1, 2, \ldots\}$, with its operations $+$, $\cdot$ and its ordering $<$, is the least controversial object in the entire mathematics. The word "arithmetic", however, or "number theory", is more ambiguous. The vast majority of people mean by that the "true facts of ω", i.e., the totality of propositions of the language of numbers that hold true in the structure ω. This (semantic) interpretation of the term is also called “complete arithmetic” and is denoted $Th(\omega)$, “the theory of ω”. In logical notation, $Th(\omega) = \{\phi \in L : \omega \models \phi\}$, where $L = \{0, ', +, \cdot\}$ is the “language of arithmetic”, and $\models$ means satisfaction.

Early in this century, on the other hand, Peano isolated a few basic truths about numbers and proposed them as axioms of the system known today as “Peano Arithmetic”, or PA. PA is also called “formal arithmetic” and contains precisely the propositions of the language of numbers that can be *proved* from the axioms of PA by purely syntactic means. Since ω is the archetype structure for numbers, the axioms of PA must of course hold in ω, hence PA $\subseteq Th(\omega)$. But Gödel in 1931 showed that PA is a theory strictly weaker than $Th(\omega)$, and Skolem in 1934 found that there is an immense variety of systems $M, N, \ldots$ which “look like” ω, but are not isomorphic to that or to each other, and which satisfy the axioms of PA. If $M \models$ PA is any such model, it is an easy exercise to show that M contains (an isomorphic copy of) ω as an *initial segment*, i.e., $\omega \subseteq M$, $+$, $\cdot$ and $<$ of M and ω agree and all elements of $M - \omega$ lie *after* all elements of ω, i.e.,

$$(\forall m \in \omega)(\forall n \in M - \omega)(m < n).$$

The elements of ω are called *standard*, those of $M - \omega$ *nonstandard* and the structure $M \models$ PA is said to be a *nonstandard model* of (formal) arithmetic, the standard model being, of course, ω. Therefore the picture of a nonstandard model is as in Fig. 6. (For a brief account of the construction of a concrete nonstandard model see Appendix A).

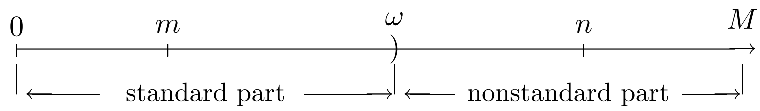

Fig. 6

The multiplicity of models of PA is due of course to the specific axioms of the latter, which fail to characterize uniquely the worlds they describe. PA consists of a finite part PA$^-$ containing trivial truths, like “0 is the first number”, “$x + 0 = x$”, etc., and the Induction Scheme, denoted Ind, consisting of all formulas of the (syntactic) form

$$(I_\phi) \qquad [\phi(0) \;\&\; (\forall x)(\phi(x) \Rightarrow \phi(x+1))] \Rightarrow (\forall x)\phi(x),$$

where $\phi(x)$ ranges over the formulas of the first-order language L in one free variable x. ("L is first-order" means that each sentence of L talks only about numbers, neither about *sets* of numbers, nor sets of sets of numbers, etc.) The scheme Ind $= \{Ind_\phi : \phi \in L\}$ is the heart of the system PA. It tells us how a certain collection of "good" sets behave in a structure.

Given a model $M \models$ PA (either standard or nonstandard), each formula $\phi(x)$ of L produces a set $X \subseteq M$, which is *defined* by ϕ in M, namely the set of elements of M which satisfy $\phi(x)$, i.e., $X = \{m \in M : M \models \phi(m)\}$. Each such set is called *definable* and let $Def(M)$ be the collection of all definable subsets of M. In contrast to the arbitrary $X \subseteq M$, each element of $Def(M)$ has a "description" in L, hence $Def(M)$ is only *countable*, since the descriptions $\phi(x)$ available in the language are only countably many. Now it is a simple exercise to show that the induction scheme Ind can be informally and equivalently rephrased as follows:

(Ind) *Every definable set of numbers has a least element.*

Complete arithmetic differs from the formal one essentially in the fact that it replaces Ind above by the much stronger, obviously *second-order*, principle:

(Ind*) *Every set of numbers has a least element.*

Ind* says that the set of integers is *well-ordered.* And, indeed, ω is such by its very construction, as the least set which contains an initial object 0 and whenever contains x it also contains $x \cup \{x\}$. Conversely, if a number structure M satisfies Ind*, the $M \cong \omega$. Consequently, Ind* characterizes completely the world of numbers.

Thus, to choose between formal and complete arithmetic is to choose between Ind and Ind*. From a mathematical point of view PA$^-$ + Ind* does not constitute a theory, since the latter always means an axiomatized set of claims, and Ind* is, clearly, not (first-order) axiomatizable.

A strong evidence against Ind* is the existence of various *sorites* occurring very frequently in all contexts of real life. A heap of sand considered as a set of grains, the set of hairs of a nonbald man, the set of units of a big number, the set of moments of one's childhood, etc., are all sorites. The stepwise passing from the property to be a sorite to its negation remains mysterious (this is the Sorite Paradox), in so far as we measure such collections with standard numbers. In standard mathematical formalizations such collections simply do not exist. Ind* does not permit their existence.

If we consciously reject Ind* in favor of Ind, then it would be more consistent to choose to live in a nonstandard world than in ω, since the latter still satisfies Ind*. (For the problem of *how* to pick such a world out of a multitude of candidates, see section 6.) A nonstandard model M always leaves room for "intractable", "uncontrolled", vague sets of numbers. For instance, within such an M, the set $\omega \subseteq M$ of standard elements is non-definable (otherwise the set $M - \omega$ would have a least element, hence ω would have a greatest one). Consequently, ω is one of those "vague" subsets, apparently the simplest one.

To return to the situation of the examples of Figs. 2 and 5 and to the Definition 2.1 of vague predicate, a glance comparing Figs. 2 and 5 with 6 suffices to convince us that vague phenomena of this type occur if and only if the structure of numbers used to measure them contain initial segments with no greatest element. Such segments have traditionally been called *cuts*. ω is one such but in fact there exist many many more. So, as a conclusion of what has been said so far, we can state the following.

Proposition 4.1 *The following statements are equivalent:*
(i) There are measurable vague predicates.
(ii) There are proper cuts of integers.
(iii) Axiom Ind is false.*

Proof. (i)$\Rightarrow$(ii): Let $P(x)$ be measurable and vague and let $\phi(x)$ be a measure such that $\mu(x) \leq \mu(y)$ and $P(y)$ imply $P(x)$. Then $\{\mu(x) : P(x)\}$ obviously forms a proper cut of N.

(ii)$\Rightarrow$(iii): Let $I \subset N$ be a proper cut. Then the set $N - I$, clearly does not have a least element. (If it did have a least element n, then $n - 1$ would be the greatest element of I, a contradiction.) Thus Ind* fails.

(iii)$\Rightarrow$(i): Let Ind* fail in N. It follows there is a set $X \subseteq N$ with no least element. Putting $Y = \{y \in N : (\forall x \in X)(y < x)\}$, clearly Y is a proper cut of N. If we define $P(x) := x \in Y$, then $P(x)$ is measurable with measure $\mu(x) = x$, and vague, so we are done. $\dashv$

5 Some facts about cuts of numbers and existence of vague partitions

In the preceding section we showed that existence of vague predicates (in the sense of Definition 2.1) is equivalent to simple existence of cuts. For the existence of vague partitions as defined in Definition 3.1, a little closer examination of cuts is needed.

For the rest of the paper we fix some nonstandard model N of PA. We do not assume any special properties of N for the time being. (But for the extra nice properties that N might have, see the last section.) Let the letters $I, J, \ldots$ range over cuts. Always we refer to *proper* cuts, i.e., I such that $I \neq N$. ω is the least cut of N. Once a cut exists, numerous other appear too. There are cuts in N arbitrarily long. For instance if $k \in N$, then the set $\omega \cdot k = \{x : x \leq n \cdot k, \textit{ for some } n \in \omega\}$ is a cut produced by the sequence $k, 2k, 3k, \ldots$. Cuts are linearly ordered in the obvious way ($I < J$ if $I \subset J$), as well as w.r.t. the numbers ($k < I$ if $k \in I$, and $I < k$ if $n < k$ for every $n \in I$). Thus, cuts behave in a sense like generalized numbers. Addition, multiplication etc. can also be defined between cuts, as well as between cuts and numbers, though not in a unique way. For example, there are two versions of $I + J$: the "lower" one

$$\{x : (\exists m \in I)(\exists n \in J)(x \leq m + n)\},$$

and the "upper" one

$$\{x : (\forall m \notin I)(\forall n \notin J)(x < m + n)\}.$$

Similarly for the other operations.

Cuts can also be seen as models of (weak) fragments of (formal) arithmetic. For example, every cut is closed under succession $'$. A bit stronger are those closed under $+$. For instance the cuts of the form $\omega \cdot k$ mentioned earlier are easily seen to be such. The cut $k^\omega = \{x : x \leq k^n, \text{ for some } n \in \omega\}$ is closed under $\cdot$, and so on. The strongest closure condition for a cut I is to be a model of PA, like ω. Classification of cuts from this point of view, that was initiated in [3], is an interesting topic that raises deep problems about logical structure of formal arithmetic.

Given a cut I, let us try to simulate the sought vagues equivalence $\sim$ by the relation $\sim_I$ on N defined by

$$x \sim_I \text{ iff } |x - y| \in I.$$

$\sim_I$ is obviously reflexive and symmetric. If, moreover, I is closed under addition, transitivity is guaranteed.

Proposition 5.1 *If the cut I is closed under addition, then $\sim_I$ is an equivalence relation.*

Proof. Let $x \sim_I y$ and $y \sim_I z$. Then $|x - y| \in I$ and $|y - z| \in I$. By closure under addition we have $|x-z| \leq |x-y|+|y-z| \in I$, hence $|x-z| \in I$. Therefore I is transitive. $\dashv$

Closure under addition yields also the following pretty nice phenomenon, which is essential in modeling vagueness: Two numbers within the cut are always less far apart than any two one of which is within and the other is without the cut.

Proposition 5.2 *Let I be closed under $+$, $x, y \in I$ and $z \notin I$. Then $|x-y| < I < |x-z|$.*

Proof. Clearly $|x-y| < x < I$. On the other hand, by assumption $x < I < z$. Now if $z-x \in I$, then by the closure condition, $(z-x)+x = z \in I$, contrary to the assumption. Thus $z - x \notin I$, or $I < |x - z|$. $\dashv$

Fig. 7 pictures the situation described in Proposition 5.2.

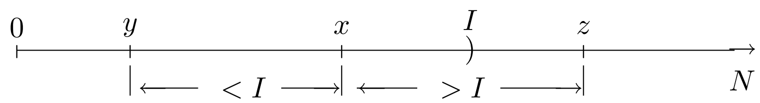

Fig. 7

Notice that the picture above is illusory in the following sense: Although the points x, z *seem* to be able to approach each other as much as we wish, in fact they stand always too far apart, namely $> I$ far. The illusion shows that representing discrete number series by continuous lines my be proved highly misleading. The bizarre properties of cuts cannot be pictured by ordinary geometric figures.

Let x_I be the equivalence class of x w.r.t. $\sim_I$. Then the sets x_I are *intervals* without end-points (i.e., if $y, z \in x_I$ and $y < w < z$ then $w \in x_I$) except from the first interval, of course that has 0 as least element. The classes x_I's are linearly ordered in the obvious way, i.e., $x_I < y_I$ if $x < y$ and $y - x > I$. Recall that a linearly ordered set is dense if between any two of its points there always exist a third one. Hence every dense set is infinite. Density of classes is also a property necessary for simulating vague partitions. The sufficient condition on I in order for $\sim_I$ to be dense is again closure under addition.

Proposition 5.3 *Let I be a cut closed under $+$. Then the set of classes of $\sim_I$ ordered in the natural way is dense.*

Proof. Let $x_I < y_I$. Then $y - x > I$. Let $z = \frac{1}{2}(y+x)$, or $z = \frac{1}{2}(y+x+1)$ (according to whether $y - x$ is even or odd, respectively). Suppose, for simplicity, that $y - x$ is even. Then $|x - z| = |y - z| = \frac{1}{2}(y - x)$. Because of the closure under $+$, $z = \frac{1}{2}(y + x) > I$, since $y - x > I$. Therefore $x_I < z_I < y_I$. This shows that the ordered set of classes is dense (see Fig. 8). $\dashv$

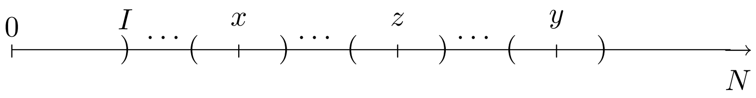

Fig. 8

Using facts 5.1- 5.3 above, we immediately get the following.

Proposition 5.4 *Let I be a cut closed under addition. Then the relation $\sim_I$ is a vague equivalence in the sense of definition 3.1.*

Proof. Recall that the properties required in order for $\sim_I$ to be vague are:

(i) If $x \sim y$ and $d(x, z) \leq d(x, y)$, then $x \sim z$.

(ii) For any x, y such that $x \sim y$ and x', y' such that $x' \not\sim y'$, $d(x, y) < d(x', y')$.

(iii) For any x, the set $\{d(x, y) : y \sim x\}$ has no greatest element, and

(iv) For any x, y such that $x \not\sim y$, and $\mu(x) < \mu(y)$, there is a z such that $x \not\sim z$ and $y \not\sim z$ and $\mu(x) < \mu(z) < \mu(y)$.

In our case $d(x, y) = |\mu(x) - \mu(y)| = |x - y|$, so (i) and (iii) are due to the fact that I is a cut. (ii) and (iv) from propositions 5.2 and 5.3, respectively. $\dashv$

The relations $\sim_I$ may be called *interval partitions* (because of the form of their classes). The simplest of them is of course $\sim_\omega$. Since ω represents the "small" numbers of N and $N - \omega$ the "big" ones, $x \sim_\omega y$ means that $|x - y|$ is a small "feasible" number.

Yet, as already said in the introduction, the above model is far from being applicable to real phenomena. What it represents is the *asymptotic* behavior of the system, when the ordinary bounds on counting and discerning ability of humans are exceeded. To give an example, suppose x, y range over colors and let $\mu(x) = l(x)$ =the wave length of x. The smaller the length unit used to measure $l(x)$ is chosen, the bigger the length $l(x)$ becomes and the finer the partition of the colorful reality into kinds of colors. By repeatedly shortening the unit, we get closer and closer, asymptotically, to the cut situation. Meanwhile, our visual discerning ability fades out and soon comes to an end. Thus, for a sufficiently small unit ε of wavelength (i.e., beyond our visual discerning ability), we can specify two numbers $l < m$ of N such that for all shades x, y,

(i) if $d(x, y) \geq m \cdot \varepsilon$, then the shades x, y are discernibly distinct and

(ii) if $d(x, y) \leq l \cdot \varepsilon$, then the shades x, y are indiscernible.

The pair (l, m) is a real-life substitute for the ideal of the cut ω. If the latter existed, then it would lie somewhere between l and m, i.e., $l < \omega < m$. Hence ω is the theoretical limit of the increasing sequence of the lower bounds l's and the decreasing sequence of the upper bounds m's in the preceding experiment.

6 Are there "standard" nonstandard models?

In this section I would like to comment on a possible natural question that might be raised by the reader, having to do with the tremendous diversity of nonstandard models. How can we choose to work on any particular model as better qualified for a description of vagueness with respect to all the rest? Would not that choice be highly arbitrary? To put it in other words: Can there be "standard" nonstandard models for the treatment of vagueness?

I think the answer to the last question can, in a reasonable sense, be "yes". And the key property required for a nonstandard model to "behave well" is *saturation*. This is a technical model-theoretic term but, intuitively, a saturated model may be thought as a *rich, full* model, one which satisfies *every* property that is consistent with the set of its truths. In fact there are various degrees of saturation. What is appropriate for our discussion here is the so-called $\aleph_1$-saturation. We shall, instead refer to it simply as *saturation*. For its classical definition, as well as some weakenings of it, see Appendix B. Here I want to draw the reader's attention on a simpler equivalent formulation (due to Pabion) that helps to understand why a saturated model is the best for modeling vagueness as a constituent of our world. In conjunction with the fact that for cardinality $\aleph_1$ the saturated model is *unique*, this will provide a satisfactory answer to the question.

Given a nonstandard M, let the letters $m, n, \ldots$ range exclusively over the standard part $\omega \subset M$ and let lower Greek letters $\alpha, \beta, \ldots$ range over the whole M. An ω-sequence of M is any set $X \subseteq M$ enumerated by means of ω, i.e., written in the form $X = \{x_n : n \in \omega\}$. (Clearly, no ω-sequence can be definable.) Recall that a function f from (a subset of) M to M is *definable* (in M), if for some formula $\phi(x,y)$ in two free variables, for all $\alpha,\beta\in M$,

$$f(\alpha)=\beta \text{ iff } M\models\phi(\alpha,\beta).$$

Saturation, then, can be stated a follows (cf. [3] for proofs).

Definition 6.1 A model $M\models \mathrm{PA}$ is said to be *saturated* if for every ω-sequence $(x_n)_{n\in\omega}$, of elements of M there is a definable function f from M to M such that $\omega\subseteq dom(f)$ and $f(n)=x_n$ for all $n\in\omega$.

In order to "translate" the preceding definition into real-life terms, one should bear in mind the following:

(1) When referring to *definable* objects of a model M, it is helpful to think of them as the objects of M which are exactly "seen", or "perceived", or "comprehensible" by the residents of M, i.e., observers located *inside* M.

(2) The length of the cut ω (and possibly other cuts too) represents reasonably the observer's distance from the *horizon* (of whatever kind the latter might be: visual, perceptual, historical, etc.). When watching e.g. from the top of a hill a flat area spreading below across the horizon, then the "vague distance" of our position from the horizon's line can be put equal to ω. Objects lying before the horizon are in a standard distance $n\in\omega$, while objects beyond the horizon are in a nonstandard distance $\alpha>\omega$ from the observer.

If (1) and (2) are agreed upon, then the saturation condition (Definition 6.1) has the following remarkable consequence:

(*) If a perceived (i.e., definable) state of affairs ϕ holds of every element of a sequence $(x_n)_{n\in\omega}$ progressing towards the horizon, then $(x_n)_{n\in\omega}$ is extendible to a sequence $(x_\beta)_{\beta\le\alpha}$ that crosses the horizon and its members also satisfy ϕ.

(Put a bit more naively: Any perceived totality that stretches till the horizon, goes beyond that. When we see e.g. a row of trees, a river, a road, etc., to extend till where our sight reaches, we are sure that this same object continuous even further, where the eye cannot see.)

(*) is a straightforward consequence of saturation. Let us prove it for the sake of completeness.

Proposition 6.2 *Let M be a saturated model, let $\phi(x)$ be a property in the language of arithmetic, and let $(x_n)_{n\in\omega}$ be an ω-sequence such that $M\models\phi(x_n)$ for all $n\in\omega$. Then there is a nonstandard $\alpha\in M$, (i.e., $\alpha>\omega$) such* *that the sequence extends to a definable one $(x_\beta)_{\beta\leq\alpha}$, such that $M \models \phi(x_\beta)$ for every $\beta \leq \alpha$.*

Proof. Given $(x_n)_{n\in\omega}$, as stated, there is, by saturation, a definable function f such that $f(n) = x_n$ for all $n \in \omega$. Let $\delta > \omega$ be some nonstandard integer. Put

$$Y = \{\gamma \leq \delta : M \models (\forall\beta \leq \gamma)\phi(f(\beta)\}.$$

By hypothesis $M \models \phi(f(n))$ for all $n \in \omega$ and hence $\omega \subseteq Y$. Y is, clearly, a definable and bounded initial segment of M, hence (by induction) it has a greatest element α. Therefore $Y = \{\beta : \beta \leq \alpha\}$, and the required extension of the sequence $(x_n)_{n\in\omega}$ is just the set $X = \{x_\beta : \beta \in Y\}$, where $x_\beta = f(\beta)$. $\dashv$

Condition (*) and its mathematical expression 6.2 has been first explicitly formulated by Vopěnka in his so-called Alternative Set Theory (AST) presented in [10]. This is a theory about finite sets, where, however, "finite" incorporates axiomatically nonstandardness. It is one of the earliest attempts to implement nonstandard into a theory right from the beginning, to the effect that *every* model of the theory contains a nonstandard (and saturated) system of integers. Saturation is introduced there by the so-called "prolongation axiom", which is a disguised form of Definition 6.1. The set-theoretic analog of cuts is called there *semisets*. Semisets and their approximation by certain special sets, called there "fuzzy" (somewhat arbitrarily), is the content of Novák's work [4].

Models satisfying Definition 6.1 are essentially unique (if of minimal uncountable cardinality) and simulate pretty well the way things go in many real situations: The sharp definable properties overspill smoothly and uniformly the horizons of vague partitions. For a concrete application of such models to the study of transformations and identity change of artificial objects the reader may consult [8] and [9].

Appendix A

Concrete nonstandard models can be obtained by the help of "ultraproducts" and "ultrapowers". Let me sketch very briefly for the reader's sake the construction of an ultrapower $M = \omega^\omega/D$. A (nontrivial) *ultrafilter* on ω is a set $D \subseteq P(\omega)$ of subsets of ω such that: (a) If $X, Y \in D$ then $X \cap Y \in D$, (b) If $X \in D$ and $X \subseteq Y$, then $Y \in D$, (c) $\emptyset \notin D$, (d) for every $X \subseteq \omega$, $X \in D$ or $\omega - X \in D$, and (e) $\bigcap D = \emptyset$. The existence of such objects follows easily from the axiom of choice. Let ω^ω be the set of all mappings

$f : \omega \to \omega$. Any ultrafilter D on ω induces an equivalence relation $=_D$ on ω^ω as follows: $f =_D g$ if $\{n \in \omega : f(n) = g(n)\} \in D$. In words, f, g are equal modulo D if they coincide on a set of D. Let $M = \omega^\omega/D$ be the set of equivalence classes with respect to $=_D$. If we identify each $n \in \omega$ with the constant function $n^* \in \omega^\omega$ such that $n^*(m) = n$ for every $m \in \omega$, then $\omega \subseteq M$. Moreover $+$, $\cdot$, $<$, etc., are extended to M in the modulo-D way, i.e., $f + g = h$ iff $\{n : f(n) + g(n) = h(n)\} \in D$, etc. This way M becomes a number structure and it is shown that $M \models \mathrm{PA}$. That M contains nonstandard elements is easily seen from the fact that e.g. $id : \omega \to \omega$ is distinct from all standard n^*. The ultrapower $M = \omega^\omega/D$ is uncountable and, upon assuming the continuum hypothesis, it has cardinality $\aleph_1$.

Appendix B

A model M for the language L is $\aleph_1$-saturated if for every countable set of formulas $\Phi = \{\phi_n(x, \overline{a}) : n \in \omega\}$ in one free variable, containing parameters $\overline{a}$ from a countable set $A \subseteq M$, the following holds:

If every finite subset of Φ is realized in M, i.e., for every $n \in \omega$,

$$M \models (\exists x)(\phi_1(x) \wedge \phi_2(x) \wedge \cdots \wedge \phi_n(x)),$$

then the entire Φ is realized in M, i.e, there is a $c \in M$ such that $M \models \phi_i(c)$ for every $i \in \omega$.

It is easily shown that every $\aleph_1$-saturated model is uncountable. Moreover any two $\aleph_1$-saturated models of cardinality $\aleph_1$ are isomorphic.

If in the above definition we change the set A of parameters from countable to *finite*, we get the notion of $\aleph_0$-saturation. Again all $\aleph_0$-saturated models of PA are uncountable. Strangely, a further weakening of the notion gets us closer to $\aleph_1$-saturation, but now having in addition countable models, namely, in the preceding definition replace the arbitrary set Φ by a *recursive* one and let A be finite again. That way we get the notion of *recursive saturation*, which is a property very close to $\aleph_1$-saturation and, in addition, satisfied by an abundance of countable structures.

References

[1] K. Fine, Vagueness, truth and logic, *Synthese* **30** (1975), 265-300.

[2] R. Kay, *Models of Peano Arithmetic*, Oxford Logic Guides, vol. 15 (Clarendon Press, Oxford, 1991).

[3] M. Kaufman and J. Schmerl, Saturation and simple extensions of Peano Arithmetic, *Ann. Pure and Appl. Logic* **27** (1984), 109-136.

[4] V. Novák, Fuzzy sets-the approximation of semisets, *J. Fuzzy Sets and Systems* **14** (1984), 259-272.

[5] R. Parikh, The problem of vagueness, in: R.S. Cohen and M.W. Wartopsky, Eds., *Language, Logic and Method,* (Reidel, Dordrecht 1983), 241-261.

[6] W. Quine, Natural kinds, in: *Ontological Relativity and Other Essays* (Columbia Univ. Press, New York, 1969), 114-138.

[7] M. Tye, Vague objects, *Mind* **99** (1990), 535-557.

[8] A. Tzouvaras, Significant parts and identity of artifacts, *Notre dame J. Formal Logic* **34** (1993), 445-452.

[9] A. Tzouvaras, Worlds of homoheneous artifacts, *Notre dame J. Formal Logic* **36** (1995), 454-472.

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