Erkenntnis
https://doi.org/10.1007/s10670-023-00731-1

ORIGINAL RESEARCH

A Generic Solution to the Sorites Paradox

Based on an Extension of the Modal Logic S4.1

Susanne Bobzien[1,2]

Received: 29 January 2022 / Accepted: 30 July 2023

Abstract
This paper offers a generic revenge-proof solution to the Sorites paradox that is compatible with several philosophical approaches to vagueness, including epistemicism, supervaluationism, psychological contextualism and intuitionism. The solution is traditional in that it rejects the Sorites conditional and proposes a modally expressed weakened conditional instead. The modalities are defined by the first-order logic **QS4M+FIN**. (This logic is a modal companion to the intermediate logic **QH+KF**, which places the solution between intuitionistic and classical logic.) Borderlineness is introduced modally as usual. The solution is innovative in that its modal system brings out the semi-determinability of vagueness. Whether something is borderline and whether a predicate is vague or precise is only semi-determinable: higher-order vagueness is columnar. Finally, the solution is based entirely on two assumptions. (1) It rejects the Sorites conditional. (2) It maintains that if one specifies borderlineness in terms of the −suitably interpreted− modal logic **QS4M+FIN**, then one can explain why the Sorites appears paradoxical. From (1)+(2) it results that one can tell neither where exactly in a Sorites series the borderline zone starts and ends nor what its extension is. Accordingly, the solution is also called agnostic.

1 Introduction

This paper offers a generic revenge-proof solution to the Sorites paradox that is based on an extension of **S4.1** (=**S4M**), namely the normal modal system **QS4M+FIN** and shows that this solution is consistent with a variety of philosophical approaches to vagueness, including epistemicism, supervaluationism, psychological contextualism and intuitionism.

✉ Susanne Bobzien
susanne.bobzien@philosophy.ox.ac.uk

1 Department of Philosophy, University of Oxford, Oxford, UK

2 Princeton, NJ, USA

Published online: 05 June 2024

The proposed solution is in some respects *traditional*. It rejects the universal conditional premise of the Sorites, or Sorites conditional, and proposes a modally expressed weakened conditional to replace it. The modalities are defined by a first-order extension of a normal modal logic which is complete with regard to a possible world semantics and in which borderlineness is introduced in modal terms in the standard way. The solution is in other respects *innovative*. The modal system proposed (system **QS4M+FIN**, Cresswell, 2001) has as yet not been used to solve the Sorites paradox, or in fact to do anything else. The system is employed to bring out the semi-determinability of vagueness: it is a crux of the proposed solution that whether something is borderline or non-borderline and whether a predicate is vague or precise are matters of semi-determinability: higher-order vagueness is columnar. The use of the logic in the solution is somewhat similar to that made of the Grzegorczyk logic **Grz** in provability logic; semi-decidability being a species of semi-determinability. Like **Grz**, **S4M** is a modal companion of intuitionistic sentential logic. This links the sentential part of the solution to intuitionistic theories of vagueness (Bobzien & Rumfitt, 2020). Finally, the solution displays an *unusual parsimony*, as it is based in its entirety on the following two assumptions: (i) it rejects the Sorites conditional; and (ii) it maintains that if one specifies borderlineness in terms of the factive-cognitively interpreted modal logic **QS4M+FIN**, then one can explain why the Sorites paradox seems paradoxical. From these two assumptions it results that one cannot tell for sure where exactly in a Sorites series the borderline zone starts and ends and what the extension of the borderline zone is. Accordingly, the solution is called an *agnostic solution*.

Unlike other solutions that propose a modally weakened ersatz premise for the Sorites conditional, the present one is immune to higher-order vagueness paradoxes. This notwithstanding, it fully accounts for higher-order vagueness (Bobzien 2015). In addition, the agnostic solution neither encounters the problems typically levelled against proposals that include modal axiom **4** (Zardini, 2006), nor the various obstacles faced by prior solutions based on a normal modal logic (e.g. Williamson 1999).

Before I get going, a remark about what I take to be the aim of a solution to a genuine paradox. Such a solution is not to make the paradox disappear or to dissolve it into trivialities. Rather, its purpose is to show that the incoherence in the paradox is apparent only and that a consistent and coherent representation of the apparent incoherence can be given in conjunction with an explanation why it appears incoherent. This is what the present paper undertakes to do.

The paper is structured as follows. ***Part I*** introduces the building blocks for the solution. These include Sorites series, tolerance, and the paradox itself, as well as the modal system and its factive-cognitive interpretation. ***Part II*** presents the solution. It introduces its two assumptions and develops an account that explains why the Sorites conditional appears true, although it is not. ***Part III*** zooms in on the structural characteristic of the interpreted modal system that sets it apart from other modal solutions: semi-determinability. It explains how, due to semi-determinability, the agnostic solution avoids sharp boundaries between the borderline and non-borderline cases in a Sorites and is immune to higher-order vagueness paradoxes. It also remarks on the connections between the factive-cognitively interpreted **QS4M+FIN** and provability logic. ***Part IV*** shows how the agnostic solution can be supplemented with a variety of semantic theories, thus giving rise to epistemicist agnosticism with non-arbitrary cut-off points, supervaluationist-style agnosticism with random cut-offs, psychological-contextualist-style agnosticism with random semantic value distribution. It briefly touches upon an intuitionist-style agnosticism with bivalence rescinded and on the non-modal intermediate logic to which **QS4M+FIN** is a modal companion, i.e. **QH+KF**.

I conclude this introduction with a note of caution. The general logical apparatus used in this paper is simple and familiar from other theories of vagueness (a first-order extension of a normal modal logic). Yet the use made of it is one only **S4M** and some of its extensions permit. It differs fundamentally from the various ways modal logic has been applied in theories of vagueness to date. Readers should be mindful of the fact that the ways in which modal logic can be used as a structural foundation for a philosophical theory are numerous, and that the only constraints on such use is its explanatory power.

Part I: Building Blocks

2 Sorites Series, Tolerance and Sorites Paradoxes

I start with an account of Sorites series whose elements should find general approval. A *Sorites series* w.r.t. some given predicate F is (i) a finite sequence of objects a_1 to a_n that is ordered with respect to some dimension (e.g. height, numbers of grains), with the ordering being total and strict,[1] for which (ii) the principle Polar and (iii) the principle Monotonicity$_{\Box}$ hold, and which (iv) displays tolerance.

The first principle, **Polar**, concerns the polar (that is, first and last) cases of Sorites series $a_1 \ldots a_n$ with regard to some F. It states that the polar cases are non-borderline or clear cases of F and $\neg F$ respectively. Formally, for a finite sequence of objects a_1 to a_n ordered w.r.t. some arbitrary predicate F:

2.1 $\Box Fa_1 \wedge \Box\neg Fa_n$ Polar

Here the (boldface) '**□**' stands in for 'it is clear/definite/determinate/non-borderline-that' as is familiar from the literature. Note that in this entire section the modal operator □ is not just uninterpreted, but also not fully defined. All that is assumed is that syntactically □ is modelled on the necessity operator and that it is factive (or veridical, if the operator is meta-linguistic). What further axioms may be employed is deliberately left open.

[1] Here a predicate 'F' forms a sentence with a complex singular term 'a_i' ('the ith object in the sequence'); numerals stand for the numbers of the members of the sequence and the quantifiers $\forall i$, $\exists i$ range over the numbers of the members of the sequence. Thus, the notion of a Sorites series involves a successor function. The solution offered can be adapted to continuous and topological Sorites, with Sorites continua and Sorites spaces in place of sequences, but not here.

The second principle, **Monotonicity$_\Box$**, expresses monotonicity for non-borderline cases in a finite ordered sequence: any a_k with a lower index than an a_m that is clearly F is itself clearly F; and any a_k with a higher index than an a_m that is clearly not F is itself clearly not F. Formally, for arbitrary predicates φ, and with a_i being the *i*th member of the series:

2.2 $$\forall_i\, ((\Box Fa_{\mathrm{i}} \rightarrow \Box Fa_{\mathrm{i}-1}) \wedge (\Box\neg Fa_{\mathrm{i}} \rightarrow \Box\neg Fa_{\mathrm{i}+1}))$$ Monotonicity$_\Box$

Call a sequence that satisfies conditions (i) to (iii) a *Polar-Monotonic series*, or short a *PM-series*.

The following *borderline-as-buffer* principle holds for all PM-series. It is relevant to the proposed solution:

2.3 $$\exists i\, (\neg\Box Fa_{\mathrm{i}} \wedge \neg\Box\neg Fa_{\mathrm{i}}) \leftrightarrow \neg\exists i\, (\Box Fa_{\mathrm{i}} \wedge \Box\neg Fa_{\mathrm{i}+1})$$ Borderline-as-buffer

It is common convention to call cases that satisfy $\neg\Box Fx \wedge \neg\Box\neg Fx$ borderline cases of F. This granted, *2.3* says that there is a borderline case of F, if and only if no clear case of F is adjacent to any clear case of $\neg F$. Borderline cases are buffers between the clear cases. One can show that *2.3* holds as follows: By the Monotonicity$_\Box$, of any PM-series with F, any a to the left of any $\Box Fa_{\mathrm{i}}$ is itself clearly F and any a to the right of any $\Box\neg Fa_{\mathrm{j}}$ is itself clearly $\neg F$. By Polar, any PM-series with F starts with $\Box Fa_1$ and ends with $\Box\neg Fa_{\mathrm{n}}$. Thus, any a_{i} which is a borderline case (i) must be to the right of the last case that is clearly F, because of its first conjunct and (ii) must be to the left of the first case that is clearly $\neg F$, because of its second conjunct. Hence in a PM-series with F, to be an a that satisfies $\neg\Box Fx \wedge \neg\Box\neg Fx$ is to be *between* the last case that is clearly F and the first that is clearly $\neg F$, so establishing the biconditional. This argument also works 'backwards'. So, for any PM-series with F, if there is an a that satisfies $\neg\Box Fx \wedge \neg\Box\neg Fx$, then a is located between the cases that satisfy $\Box Fx$ and the cases that satisfy $\Box\neg Fx$, and vice versa, which is another way of stating *2.3*.

Finally requirement (iv), which is **tolerance**. I treat tolerance as a phenomenon, that is, as *how things seem to be*, or as epistemic tolerance, as some call it. I do not discuss what the roots of tolerance are.[2] Moreover, I look at tolerance *only* with respect to PM-series with F. There are PM-series with F in which it seems compelling that of any two adjacent objects a_{i}, $a_{\mathrm{i}+1}$ one can't have one be F without the other being F, since the objects are with regard to F indiscriminable, or insufficiently distinguishable, perceptually or otherwise. Of such sequences I say that they display tolerance. Views differ about the precise logical structure of tolerance (see Sect. 8). Here I offer the formalization of a *weak tolerance principle* for PM-series with F.

[2] Among other things, perception, semantics and assertion have been suggested. See Raffman (1994, 2005) for a perceptual interpretation, Gaifman (2010) for a semantic one. Others express tolerance in terms of judgement, assertion and similar: e.g. if one asserts that Fa_{i} one cannot reasonably deny that $Fa_{\mathrm{i}+1}$. In this paper I stay away from questions of assertibility.

2.4 $\forall_i \neg\Box\neg(Fa_{\mathrm{i}} \leftrightarrow Fa_{\mathrm{i+1}})$ TOLERANCE$_{\neg\Box\neg}$

2.4 says that, for PM-series with F, it is not clear (definite, …) that it is not the case (and so one cannot rule out that it *is* the case) that if F applies to a_{i}, it applies to $a_{\mathrm{i+1}}$, and vice versa. *2.4* should suffice as a condition for a PM-series with F to be a Sorites series. Additionally, and consistent with this formalization, where there is full-blown tolerance, we *cannot* rid ourselves of the view that, of a_{i}, $a_{\mathrm{i+1}}$, one cannot have one be F without the other being F—as we *can* rid ourselves for example of the view that the stick in the glass is bent by an explanation that it is in water: it still looks bent, but we now know that it is straight. Similarly, we may successfully rid ourselves of the view that there are fewer than 600 people in the concert hall, once we have had the opportunity of counting them. In contrast, when a PM-series displays tolerance, there is no way for us to know or tell of any two adjacent objects that one can be F without the other. Tolerance is in this sense a *tenacious* phenomenon.[3] Sequences with tolerance differ from other PM-series in having their objects closer together with regard to F—close enough that the objects are with regard to F indiscriminable, or insufficiently distinguishable, perceptually or otherwise. This is why such sequences generally allow the construction of Sorites paradoxes, or are *Sorites susceptible*, and accordingly are called *Sorites series*.

The requirements (i) to (iv) of this account of Sorites series should find general approval. Minor discrepancies and notational variants such as the use of the successor function '′' instead of a_{i}, $a_{\mathrm{i+1}}$ and alternative formulations of *2.1* and *2.2* should not matter in what follows. Neither should it matter whether one defines Sorites susceptibility via full-blown tolerance or via some functionally comparable relation that expresses the required closeness of the adjacent pairs of the sequence.

Sometimes stronger constraints are put on Sorites series. Some (e.g. Wright, 2019) prefer the monotonicity relation in a Sorites series to be, with Fa_1 and $\neg Fa_{\mathrm{n}}$ assumed,

2.5 $\forall i\ ((Fa_{\mathrm{i}} \rightarrow Fa_{\mathrm{i-1}}) \wedge (\neg Fa_{\mathrm{i}} \rightarrow \neg Fa_{\mathrm{i+1}}))$ MONOTONICITY$_{\mathrm{F}}$

rather than *2.2*. In Sect. 3 I say why this may be too strong. Others (e.g. Gaifman, 2010) believe that tolerance of PM-series is not just epistemic, and instead of *2.4*, offer

2.6 $\forall i\ (Fa_{\mathrm{i}} \leftrightarrow Fa_{\mathrm{i+1}})$

This belief cannot be shown to be correct, and I think it is far too strong. In any event, those who favour *2.5* or *2.6* are usually happy with (i) to (iv), or their analogues, as *necessary* conditions for a Sorites series.

[3] The tenacity of tolerance in vague predicates F *generally* can be expressed thus: For F there is a measure m such that if objects differ by m with regard to F, it seems that of any two adjacent objects a_{i}, a_{i+1} one can't have one be F without the other being F, since a_{i} and a_{i+1} are insufficiently distinguishable w.r.t. F, or since the justice with which F is applied does not change from a_{i} to a_{i+1}.

Any Sorites series regarding some predicate F (abbreviated Ss_F) permits the construction of at least two **Sorites paradoxes** with F. I here express them with a universally quantified conditional, that is, as *mathematical induction Sorites*, which employ the successor function:

$$\frac{\begin{array}{l} Fa_1 \\ \forall i\ (Fa_\mathrm{i} \rightarrow Fa_{\mathrm{i+1}}) \end{array}}{Fa_\mathrm{n}}$$

and its reverse

$$\frac{\begin{array}{l} \neg Fa_\mathrm{n} \\ \forall i\ (\neg Fa_\mathrm{i} \rightarrow \neg Fa_{\mathrm{i\text{-}1}}) \end{array}}{\neg Fa_1}$$

Both these two-premise arguments are *paradoxical*, since their premises appear true and their conclusion appears false, and the reasoning appears impeccable. Since in classical logic structurally either is the reverse of the other, and in this paper I only discuss modal systems that preserve classical logic, I limit myself generally to the first.

A mathematical induction Sorites can be regarded as an abbreviated form of a *conditional Sorites paradox*. This is a chain argument of the kind Fa_1; $Fa_1 \rightarrow Fa_2$; …; $Fa_{\mathrm{n\text{-}1}} \rightarrow Fa_\mathrm{n} \vdash Fa_\mathrm{n}$. Such an argument can in turn be considered an abbreviation of a *step-by-step Sorites*. These are argument chains that consist of $n-1$ applications of *modus ponens*, where the conclusion of each argument functions as a premise of the next (Fa_1, $Fa_1 \rightarrow Fa_2 \vdash Fa_2$; Fa_2, $Fa_2 \rightarrow Fa_3 \vdash Fa_3$; …; $Fa_{\mathrm{n}-1}$, $Fa_{\mathrm{n}-1} \rightarrow Fa_\mathrm{n} \vdash Fa_\mathrm{n}$).[4]

The universally quantified Sorites conditional and tolerance are logically interrelated. The exact nature of the relation depends on how tolerance is defined (Sect. 8).

3 Whittling Away at the Penumbra and Introducing Assessment Sensitivity

Before I present the normal modal system on which the agnostic solution is based, I draw attention to several facts about Sorites series that pave the way for an understanding of the philosophical elements behind the formal presentation of the solution (Sects. 4–10).

[4] Where a generalization with a universally quantified conditional is not possible, one can use chain-arguments instead. The rest of the proposed solution then needs to be—and can be—adjusted accordingly. The version of the Sorites that leads to the so-called Unpalatable Existential $\exists i\ (Fa_i \wedge \neg Fa_{i+1})$ (e.g. Wright, 2010) can also be given a solution based on the logic and interpretation provided in this paper, with some simple adjustment.

Modal solutions to the Sorites are generally motivated by the fact that Sorites series display a kind of grey area somewhere in the middle where even speakers who are in no way handicapped with respect to assessing whether any of the $a_1 \ldots a_n$ are *F* hesitate, or show some other kind of hedging behaviour, and may disagree with each other or even with their earlier selves, when asked about the objects in this area. Philosophers who propose a modal solution assume that there is something to this grey area (or penumbra) that is pertinent to solving the Sorites. In particular, if there is a change from *a*s that are *F* to *a*s that are not *F*, such a change, it is assumed, would occur in that grey area. This seems to me correct. Philosophers often refer to this grey area as the borderline zone of the Sorites series, and to the objects in that zone as its borderline cases. My preference is to use 'grey area' for the general phenomenon described and 'borderline zone' and 'borderline case' as precise terms, defined differently in different theories, with '*borderline zone*' referring collectively to the borderline cases of a Sorites series. Here it is important to be aware that the introduction of a modal logic is motivated *jointly* by the following: (i) a problem (the paradox) (ii) an associated phenomenon (the grey area) and (iii) the objective to solve the paradox by means of getting a grip on the grey area. The *natural language*[5] expressions 'borderline', 'borderline case', 'clear', 'determinate', etc., and their semantics are not—or at least not directly—germane in this context. If the grey area can be gotten under control modally in such a way that the paradox is solved, we are good.

It will be helpful to note a number of factors that contribute to the size of the grey area, but whose removal leaves the paradox in place. These include the fact that different groups of speakers may use *F* in statistically significantly different ways. In this case we can relativize to those linguistically diverse groups. Also, individuals may have perceptual or linguistic shortcomings, such as colour blindness or a lack of familiarity with *F* ('vermilion', 'gaggle'). We can exclude such individuals from considerations regarding *F*. Moreover, the extension of *F* tends to vary with context of use (or context of evaluation). We can fix the context of use, as far as feasible. There are additional factors, such as the possibility of a gradual shift of meaning of expressions by a change of their use, *et sim*. These can be discounted since they tend to lose significance once a context is fixed.

In this paper, I consider the case in which all such factors that contribute to the extension of the grey area have been eliminated. Instances of the Sorites paradox will still abound, and we can examine what it is that makes these paradoxical unobstructed by the factors mentioned. In this sense alone, I consider an idealized

[5] By 'natural language' I here mean a language that has evolved naturally in humans through use and repetition without conscious planning or premeditation. Scientific and technical expressions that have been introduced by definition for a specific purpose are not natural-language expressions in this sense unless over time they are integrated into natural language by repeated use by speakers who are unaware of or disregard the original scientific definitions.

situation. Such an elimination of contaminating factors is commonplace.[6] Now consider the following fact: with the removal of each of these factors, the grey area shrinks. If one combines them all, it may shrink considerably. This suggests that the grey areas that are relevant to the solution of the uncontaminated Sorites can be quite small and are in any case smaller than is often assumed.

A further factor that may artificially inflate the perceived grey area needs special mention, viz. the fact that people frequently envisage a grey area as between contraries (being blue and being green) rather than contradictories (being blue and not being blue). This encourages thinking of the grey area as a zone in which objects are neither blue nor green but blue-green or teal, say. And depending on the sequence, there can be a rather extended area in which things look neither blue nor green, since they *are* neither, but are blue-green or teal. However, Sorites paradoxes run with contradictories, such as being blue (or a heap) and not being blue (or a heap). And although there is an area with regard to which speakers hedge or voice disagreement, I believe it is mistaken to think that this is because there is a detectable *third kind of colour condition*, between being blue and not being blue, which they may have to detect, or perhaps discover, and label. Compare: ‘let’s call these cases blue–green (or teal)’ with ‘let’s call these cases blue-and-not-blue (or Graham-coloured)’. The first case may involve giving a name to something previously not differentiated but now told apart. Such increased differentiation would indicate an achievement. The second case amounts de facto to relinquishing bivalence and introducing what Crispin Wright has called a *Third Possibility* (Wright, 2003). This, if anything, indicates failure. Despite the temptation of visualizing the grey area in terms of contraries, the grey area that is *relevant to solving the Sorites* and that manifests itself as hedging or lack of agreement results from a sequence that runs from objects that are F to objects of which it is not the case that they are F. The term ‘grey area’ is thus used strictly metaphorically. (For a detailed treatment of the point made in this paragraph see Bobzien 2013) .

So far, this section has provided reasons for thinking that one can chisel away substantial parts of the grey area in a Sorites series without making the paradox vanish. Here is an observation about the remaining part. Although by definition the objects in the series are ordered with regard to F, where there is hedging and lack of agreement, we may be unable to establish *monotonicity regarding F* other than by invoking some semantic relation between vague predicates and degree adjectives (or some kind of meaning postulate that connects ‘blue’ with ‘bluer’, perhaps). In

[6] Where some of these factors are invoked as the sole explanation in a solution to the Sorites, the question remains, what about *those* Sorites paradoxes that remain when these factors are jettisoned? Furthermore, for vague predicates one usually can reintroduce a paradox by further diminishing the distance between the objects.

particular, with regard to the $a_k \ldots a_m$ in the grey area of any Sorites series with F we may not be able to *establish* that, with Fa_1 and $\neg Fa_n$ assumed,

2.5 $$\forall i\ ((Fa_i \rightarrow Fa_{i-1}) \wedge (\neg Fa_i \rightarrow \neg Fa_{i+1}))$$ MONOTONICITY$_F$

There is some empirical evidence that even qualified speakers seem inconstant and capricious in their assessment of such objects with regard to F. They may judge a_{k+n} to be F but $a_{k+(n-1)}$ not to be F, and even may judge the same object first F and shortly after $\neg F$, if they are unaware that it is the same—or even if they *are* cognizant of this fact (e.g. Raffman, 1994). For this reason, I refrain from making MONOTONICITY$_F$ part of the *necessary conditions* for a Sorites series. (The agnostic solution is compatible with *2.5* and its negation.) On the other hand, the weaker

3.1 $$\forall i\ \neg\Box\neg((Fa_i \rightarrow Fa_{i-1}) \wedge (\neg Fa_i \rightarrow \neg Fa_{i+1}))$$ MONOTONICITY$_{\neg\Box\neg}$

is acceptable. (In the modal system **S4**, *3.1* is entailed by *2.2*, that is by MONOTONICITY$_\Box$.)

Overall, the grey area that is relevant to a solution of the Sorites thus emerges as being both smaller and more chaotic in appearance than is often assumed. This fact is taken up and utilized in the agnostic solution. Here is not the place to elaborate on the philosophical theories that may go with it. As an aid to readers, I offer just the very roughest formulation of such a theory and how it relates to the grey area where the borderline cases live. The theory maintains that in certain cases even with the context of use (context of evaluation) fixed, there is still variation in how things appear to us, variation that hinges on the perspective or viewpoint one takes when *assessing a* with regard to F. Thinking of a grey-area a as being F may be one such viewpoint. Focusing in one's mind intensely on something that is purely and distinctly F may be another. Even without details given, it should be plausible that some such viewpoints cannot reasonably be made part of the context of use regarding F, but can still in some cases be decisive as to whether a does or does not appear F.[7] Borderline cases a of F are then cases in which, even with the context fixed, maximally relevantly qualified individuals can take a viewpoint from which a looks just like things that are non-borderline F and another viewpoint from which a looks just like things that are non-borderline $\neg F$. In either case there would appear to be no need to seek out another viewpoint for confirmation. In other words, I suggest that vague predicates display a kind of unsavoury assessment sensitivity in the grey area of Sorites series. (*Unsavoury*, since it seems not to yield sufficient reason for the assertion either of F or of $\neg F$.) Again, the present sketch of the assessment sensitivity is given solely to facilitate understanding of how the semantics of

[7] Recent research by philosophers that work on perception, e.g. Susanna Siegel, provides empirical data about perception that support this suggestion. In the context of vagueness, some of the data used by Diana Raffman in her work can also be seen to back it up. (The two illustrations given in the text would generally not be accepted as part of a *context of use*. Other examples abide.)

QS4M+FIN may relate to a philosophical theory of vagueness—not to persuade the reader that the philosophical theory is true. The latter is an independent undertaking.

4 The Logical Structure of Vagueness: The System FIN

The logic that characterizes the structure of the agnostic solution is a sentential normal modal logic with a first-order extension, set out here as an axiomatic modal system. ('Axiom' and 'theorem' will be used as short for 'axiom schema' and 'theorem schema'.) The sentential portion of the system is **S4M** (also known as **S4.1** or **K1** or **KT4M** or **KT4G$_c$**). Two modal operators are used, one for clarity, the other for borderlineness. Neither is taken as metalinguistic: it would be 'it is borderline whether Sloane is slow' not 'it is borderline whether "Sloane is slow" is true', 'it is borderline true whether "Sloane is slow"' or the like.

The syntax uses p, p_1, p_2 … for atomic sentences; the classical connectives ¬, ∧, ∨, → and ↔; and parentheses ((,)) as punctuation symbols. The clarity modal operator □ ('box') is modelled on the necessity operator (read: 'it is clear that'). Its syntax is the usual one.

A (sentential) normal modal system is defined as a class S of well-formed formulae (*wff*) of a sentential modal logic in which all valid wffs of the classical sentential calculus are axioms (**PC**); the rules of *modus ponens* (**MP**) and of necessitation (**N**) hold; in which

4.1	$\Box(A_1 \rightarrow A_2) \rightarrow (\Box A_1 \rightarrow \Box A_2)$	axiom **K**

is an axiom; and in which the wffs are of a language $\mathcal{L}$ of modal **PC** (e.g. Hughes & Cresswell, 1996, 111). The system that satisfies precisely these conditions is system **K**. System **S4M** extends system **K** by three axioms:

4.2	$\Box \mathrm{A} \rightarrow \mathrm{A}$	axiom **T**
4.3	$\Box \mathrm{A} \rightarrow \Box\Box \mathrm{A}$	axiom **4**
4.4	$\Box\neg\Box\neg \mathrm{A} \rightarrow \neg\Box\neg\Box \mathrm{A}$	axiom **M**

Axiom **T** warrants that non-borderlineness is factive. Axioms **T** and **4** added to system **K** produce system **S4**. In **S4**, $\neg\Box A \rightarrow \Box\neg A$ is not a theorem. This ensures that the existence of borderlineness is not logically precluded. **M** is the McKinsey axiom. Its function is explained below. **S4M** is complete with respect to the class of transitive, reflexive and final Kripke frames. A final frame is one in which every world can access at least one world that cannot access any world besides itself.

A second operator, for borderlineness, ∇ ('nabla'), is modelled on the contingency operator (read: 'it is borderline whether'). It is defined in terms of □ thus: For an arbitrary formula A

4.5 $\nabla A \leftrightarrow \neg\Box A \quad \wedge \neg\Box\neg A$ df ∇[8]

The operators ∇ and $\Box$ obtain their full definitions by the syntax, rules, and axioms of **S4M** as a whole.

Philosophically, the central notions that underlie the sentential logic of vagueness **S4M** are borderlineness and non-borderlineness (as explained in Sect. 3), where non-borderline (or clear) cases are the norm, borderline cases the exception. One way to think about the $\Box$-operator is that it is clear that A precisely if it is both the case that A and not borderline that A.[9] Again, here I have no interest in the semantics of the *natural language* expressions 'it is clear', 'it is definite', etc. (for which see e.g. Barker, 2002, §3, also Raffman, 2014) or 'borderline'. The latter is used in natural language in several incompatible ways (Bobzien 2013, 4-10, 2015, 77-80, also Raffman, 2014, ch.2).

Here are three theorems of **S4M** for the ∇-operator that are germane to the agnostic solution. First the mirror theorem, which states that it is borderline whether A precisely if it is borderline whether $\neg A$:

4.6 $\nabla A \leftrightarrow \nabla\neg A$ MIRROR

The core of MIRROR is captured in the formulation 'borderline *whether*'. The proof is trivial. Second, the equivalence theorem:

4.7 $\Box(A_1 \leftrightarrow A_2) \rightarrow (\nabla A_1 \leftrightarrow \nabla A_2)$ **E**

If it is clear that A_1 if and only if A_2, then A_1 is borderline if and only if A_2 is borderline. Relatedly, borderlineness is closed under logical equivalence, in the sense that if $A_1 \leftrightarrow A_2$ is a theorem of **S4M**, then $\nabla A_1 \leftrightarrow \nabla A_2$ is a theorem of **S4M**. A third theorem, one specific to systems that include axioms **T** and **M**, is

4.8 $\nabla A \rightarrow \nabla\nabla A$ **V**

Theorem V expresses the point that if something is borderline, it is borderline borderline. It is the distinctive theorem of columnar higher-order vagueness (Bobzien, 2015, 65-8, 74-6). In system **T**, and hence in system **S4**, theorem **V** is logically equivalent to axiom **M** (proof in Bobzien, 2015, 84-5).[10] The relevance of **M** to borderlineness may not be immediately evident, while the relevance of **V** can be quite readily gauged. For one thing, **V** squares neatly with how some philosophers have characterized borderline cases as being themselves borderline (e.g.

[8] An alternative axiomatization with ∇ instead of $\Box$ as the primitive operator is possible. $\Box$ is chosen as primitive because it facilitates comparison with other theories of vagueness. Philosophically, borderlineness is the basic notion. For an axiomatization of the contingency logic of **S4** see e.g. Montgomery and Routley (1966), Kuhn (1995).

[9] This reading in terms of ∇ requires axiom **T**, see e.g. Segerberg (1982).

[10] More precisely, in system **T**, and hence in system **S4**, *4.8* (which in **S4M** with the ∇ operator is theorem **V**) is logically equivalent to *4.4* (which in **S4M** is axiom **M**).

Wright, 2003). Second, **V** goes some way explaining why, independently of whether bivalence holds, we encounter difficulties when trying to pinpoint where the borderline zone of a predicate starts and where it ends (relative to dimension and context). Third, **V** enables us to provide a straightforward semantics for borderlineness that embodies the unsavoury assessment sensitivity of vague expressions described in Sect. 3. With a semantics that takes *viewpoints* as worlds: If it is borderline whether A, then there is not just a *viewpoint* v that can 'see' a *viewpoint* at which A and a *viewpoint* at which $\neg A$, so that borderline A at v. *Viewpoint* v can also 'see' some *viewpoint* that can 'see' only *viewpoints* at which A; or some *viewpoint* that can 'see' only *viewpoints* at which $\neg A$. (These latter can be envisaged as viewpoints where things look no different from non-borderline A or from non-borderline $\neg A$, respectively, so that there appears to be no reason to seek out another viewpoint for confirmation.) Thus, if it is borderline whether A, it is borderline whether it is borderline whether A.[11] Via its logical equivalent **V** (in system **T**), axiom **M** thus offers distinctive benefits for a theory of borderlineness.

Next, the quantified modal system for borderlineness. The agnostic solution offered has its foundation in the triad of transitivity, reflexivity and finality as defined for the possible world semantics of **S4M**. Accordingly, *the relevant first-order extension of normal columnar higher-order vagueness is the extension that corresponds to the possible world semantics with respect to which **S4M** is complete*. This is the quantified modal logic **QS4M+FIN**—a logic introduced in Cresswell (2001) and without any previous philosophical application.

For this first-order system, the syntax is expanded: F, G are used for n-place vague predicates of a natural language with $n \geq 1$; a, a_1, a_2 … for individual constants; x, y, x_1, x_2 … for variables. (p, q, … are redefined as zero-place predicates.) The wffs are now wffs of a language $\mathcal{L}$ of a quantified modal logic. Complementing this syntax, the rules and axioms for **QS4M+FIN** are:

S4M′ If A is a substitution instance of a theorem of **S4M**, then A is an axiom of **QS4M+FIN**.

∀1 If A is any wff and x and y are variables and $A(x/y)$ is A with free y replacing every free x, then $\forall xA \rightarrow A(x/y)$ is an axiom of **QS4M+FIN**.

N If A is a theorem of **QS4M+FIN**, then so is $\Box A$.

[11] **V** (or **M**) does not entail that no differentiations are possible within a borderline zone. We can distinguish three situations: (i) one in which it is clear that one cannot rule out that A; (ii) one in which it is clear that one cannot rule out that $\neg A$; and (iii) a third in which neither is clear. What one cannot have are situations in which both are clear. Which of the cases (i), (ii), (iii) are manifested in a borderline zone and in what order is an empirical question and may differ from case to case. Those who think that anything more precise can be said about borderline cases, are, I believe, confused about something. They may confound orders of borderlineness with degrees of F-ness; or they may confound higher-order vagueness with (what in Bobzien 2013 I have called borderline nestings, that is) the partial precisifications of vague predicates brought about by the introduction of a new predicate (like 'blue-green' or 'teal') that covers the area between those of two contrary predicates like 'blue' and 'green'.

MP If A_1 and $A_1 \rightarrow A_2$ are theorems of **QS4M+FIN**, then so is A_2.

∀2 If $A_1 \rightarrow A_2$ is a theorem of **QS4M+FIN** and x is not free in A_1, then $A_1 \rightarrow \forall x A_2$ is a theorem of **QS4M+FIN**.

FIN $\neg\Box\neg\forall x_1, \ldots, \forall x_n(A \rightarrow \Box A)$ the finality axiom[12]

Hereafter I refer to this first-order modal system as *the finality logic* (***FIN***) or as *(modal) system* ***FIN*****.** In the finality logic, with the given interpretation, ∇Fa is to be read as 'a is (a) borderline F' or 'a is a borderline case of being (an) F'. Higher orders of borderlineness are expressed thus: a is a *first-order* borderline case of F, written $\nabla^1 Fa$, iff ∇Fa. And a is an $(n+1)$*th-order* borderline case (for $n \geq 1$), written $\nabla^{n+1} Fa$, iff $\nabla\nabla^{n} Fa$.

Since system **FIN** is complete with respect to the quantificational possible world semantics with the class of transitive, reflexive, and final frames (Cresswell, 2001, 159–64), the notion of borderlineness as defined by the finality logic is consistent.

For the proposed solution of the Sorites, axiom **FIN** is important, since in **QS4M** it is equivalent to the ∇-operator theorem

4.10 $\exists x_1, \ldots, \exists x_n \nabla A \rightarrow \nabla \exists x_1, \ldots, \exists x_n \nabla A$ **∇∃**

A proof of the equivalence is offered in the "Appendix". Of this general theorem ∇∃, the one-place predicate version is

4.11 $\exists x \nabla Fx \rightarrow \nabla \exists x \nabla Fx$ **∇∃(x)**

The solution relies on **∇∃(*x*)**, with its modal operator '∇' interpreted as 'it is borderline whether'. 4.11 is plausible, since the fact that there are Sorites paradoxes with universal premises suggests that it is not clear whether vague predicates have borderline cases. For additional philosophical justification of *4.11* thus interpreted, and hence of the use of system **FIN** as the logic of vagueness, see the end of Sect. 5.

[12] Cresswell calls the finality axiom **FINAX** or **FIN** (Cresswell, 2001, 160).

BF $\forall x \Box A \rightarrow \Box \forall x A$ the Barcan formula

yields the stronger system **FIN+BF,** which provides a variant of the solution offered in this paper.

5 The Factive-Cognitive Interpretation of the Operators

The agnostic solution is based on system **FIN** in tandem with an interpretation of the operators $\Box$ and ∇ as 'can tell that' and 'cannot tell whether'. This interpretation is factive-cognitive (epistemic in a wider sense),[13] as opposed to semantic or ontic. A factive-cognitive interpretation is apt, since every solution to the Sorites must explain hedging behaviour and potential disagreement in the grey area of a Sorites series, and this requires a cognitive element. It is not thereby precluded that the factive-cognitive interpretation has its ultimate justification in some underlying ontic or semantic theory.

On the factive-cognitive reading, borderlineness of some a regarding some F is cashed out as a kind of cognitive inaccessibility, expressed here as 'one cannot tell whether'. Some a is borderline F precisely when *things are such that* relevantly qualified individuals cannot tell whether Fa. In line with common opinion, the relevantly qualified individuals are those who are in no way handicapped with regard to assessing whether Fa. Hence, when it is borderline whether Fa, the reason does not lie in any shortcomings of the individuals, but in F, a, and possibly in how they relate. The reasons for the cognitive inaccessibility are in this sense *extra-mental*. The phrase 'one cannot tell whether' is used as a natural language *stand-in* for borderlineness as defined modally and interpreted factive-cognitively. I have no more an interest in providing a semantics for the natural language expression 'can tell' than for the natural language expression 'it is clear'. However, I do rely for general understanding on the natural-language connotations of 'can tell' when used to indicate recognition. ('Can you tell whether it's Alex?'—'Is it Alex? Do you recognize her?'). Specifically, the tell-ability at issue is of the kind that, if one can tell, then one can tell for sure; and if one can tell that not, then one can tell for sure that not, that is, then one can rule out for sure. The finality logic supplies the structure for this kind of tell-ability and this kind of ruling out.

The factive-cognitive interpretation reflects the unsavoury assessment sensitivity of borderline cases described in Sect. 3. One cannot tell whether a is F precisely when, even with the context fixed and to maximally relevantly qualified individuals, either there is a viewpoint from which a looks w.r.t. F just like things that are non-borderline F (i.e. that are clearly F), or there is a viewpoint from which a looks w.r.t. F just like things that are non-borderline $\neg F$ (i.e. that are clearly $\neg F$), or both. A lot more needs to be said about the factive-cognitive interpretation, and it will be, albeit elsewhere. The topic of the present paper is—primarily—the structural properties of a solution to the Sorites that is based on a normal modal logic, for which no further interpretational details are required. In line with the tell-ability interpretation of borderlineness, interpreted, the ∇-operator then reads 'one cannot tell whether',

[13] I borrow the term 'factive-cognitive' from linguistics. The interpretation is factive, since $\Box$ is non-metalinguistic and if one can tell that A, then A. The interpretation is factive-cognitive, since the ability to tell is a cognitive ability. Recall that 'epistemic' is not the same as 'epistemicist'. By 'epistemic' I mean 'relating to knowledge', and by 'epistemic in a wider sense' I mean 'relating to cognition'. For epistemicism, see Sect. 15.

the □-operator 'one can tell that'. I abbreviate 'one cannot tell that it is not the case that' (¬□¬) as 'one cannot rule out that'. These two expressions are henceforward used synonymously.

Using the philosophical interpretation sketched in Sect. 3, axiom **FIN** and theorem **V∃(x)** (*4.11*) can now be elucidated further as follows. When there is a borderline case a_i of F, then there is a viewpoint at which a_i is non-borderline F or there is a viewpoint at which a_i is non-borderline $\neg F$, so that in either case there appears to be no need to consult a further viewpoint for confirmation. Since this is so for *every* borderline case of F, one cannot rule out that there are viewpoints at which every a_i of the section of the sequence that is borderline F is *non-borderline* F, so that none of them appears to require consultation of another viewpoint for confirmation. It results that whenever there is a borderline case a_i of F, it is not clear that there is a borderline case of F (i.e. one cannot rule out that there is *no* borderline case of F), since there is a viewpoint from which there is no borderline case of F. This is what the interpreted theorem **V∃(x)** expresses.[14]

The exact relation between the possible world semantics of **FIN** and the unsavoury assessment sensitivity of viewpoints described in Sect. 3 is the topic of a separate paper. Here is only the briefest and somewhat simplified exposition, so that readers can take in the general idea. (i) It is the space of viewpoints that is transitive, reflexive and final. (ii) On this basis, the formulas that are valid in all models over all **S4M** frames (i.e. are *valid tout court*) define the structure of borderlineness. (iii) Every world in a model is a rational *viewpoint of assessment*. Simplified, the *accessibility relation* can be envisaged as a rational way of collecting viewpoints as to whether a is F, where a viewpoint becomes part of one's collected viewpoints if one

[14] To support axiom **FIN**, I here restrict myself to the situation in which **S4M** has been accepted as a plausible candidate for a sentential logic of vagueness. Then, the relevant **FIN**-denying position is the following. In a Sorites series there are no *clear* borderline cases but it is clear that there *are* borderline cases in it. In the absence of **FIN**, this position is consistent. (However, for those who wish to avoid higher-order-vagueness paradoxes, this position is unsuccessful, see e.g. Zardini (2006) for a formal argument that confirms this. Adding **FIN** brings success.) In terms of tell-ability, its proponents may claim that sometimes one can tell that some Sorites series contains borderline cases without being able to tell of any one of the cases in the series that *it* is a borderline case. How can one tell this? I suggest that the rational answer is one cannot. Consider the most common defences of this **FIN**-less position. They all seem to fail. (a) "Some of the cases in the series could (or may) be borderline cases, therefore it is clear that there are borderline cases." Not a valid inference. (b) "There must be borderline cases, since otherwise we do not have a Sorites series." *Petitio principii*. (c) "It is evident, i.e., to some, that there are borderline cases in a Sorites series. Hence one can tell that there are such borderline cases even if one cannot tell which." Subjective opinion in the premise. *To me* this is not evident, but only that we cannot rule out that there are such cases. (d) "When assessing a Sorites series one can always tell that there *are* borderline cases, it is just that when one zooms in on them, they move out of reach." This is perhaps the most interesting case. Still, I reject that in this scenario one can *tell* that there are borderline cases. It is more that one is convinced that one could tell if only one could get sufficiently close. I take this lack of convincing positive reasons for the **FIN**-less position as equally good reasons for accepting **FIN**: If you cannot tell of any one case in a Sorites series that it is borderline, it is plausible that you also cannot tell whether there exist borderline cases in the series. Bear in mind here that the question is not whether there *are* borderline cases in a Sorites series. I believe that there are, just that one cannot tell (or otherwise prove) this. Consequently, I make the reasonable *assumption* that there are borderline cases (Sect. 7). From such an assumption it just does not follow that one can tell that there are borderline cases.

has taken it, i.e. gotten to have it. Then (a) one can always take the viewpoint one has. (b) If, from one's viewpoint, one can take another viewpoint, then one can get to have the viewpoints one can take from that viewpoint. (c) There always is a viewpoint at which a is non-borderline F or at which a is non-borderline $\neg F$ (and so for all objects), and at which in line with this there is no rational motivation to seek out any further viewpoints. Such a viewpoint is a final viewpoint. (iv) The semantic values at viewpoints are truth or falsehood relative-to-a-viewpoint, not truth and falsehood *tout court*. (v) As is generally agreed, truth *tout court* of sentences with vague predicates is relative to a fixed context of evaluation c. The possible world semantics can reflect this if one puts a suitable constraint on the models under consideration, and indexes them to such a context c. Truth *tout court* can then be represented as truth relative to a model of the resulting subclass of models and indexed to a context c. The models do not determine the semantic values *tout court* of borderline sentences. How one is to think about these values, and what constraints one puts on the models, depend on what kind of agnosticism one favours (below Sects. 15–17). (vi) A sentence ∇A is *true* relative to a model indexed to a context, if in the relevant model there is a viewpoint relative-to-which A is true and a viewpoint relative-to-which A is false. (vii) We have no reliable meta-perspective on viewpoints. The reason for (vii) is the following. The closer we get to the borderline zone the smaller a shift in context can make the difference between an a's being borderline or not being borderline; hence, once we are in the grey area, we may be unable to discern whether the difference between two assessments of the same case is due to a shift in context or to a difference in viewpoint (see e.g. Bobzien 2010).[15]

[15] If axiom **5** ($\neg\Box\neg A \rightarrow \Box\neg\Box\neg A$) were added to system **FIN**, the logic would collapse to the trivial modal logic. The proposed interpretation of the modal operator ∇ as cognitively inaccessible cognitive inaccessibility or absolute cognitive inaccessibility helps explain why axiom **5** *should* not hold. In epistemic logic, axiom **5** is sometimes said to express negative introspection (e.g., Bonnay & Égré, 2009): in tell-ability terminology, if one cannot tell that A, one can tell that one cannot tell that A (with substitution of $\neg A$ for A and double negation elimination on **5**). And, applying some trivial logical steps, in terms of ∇, if one cannot tell whether A, one can tell that one cannot tell whether A. For many kinds of sentences negative introspection seems an obvious principle: If one cannot tell which box contains the ball, one generally can tell that one cannot tell this. However, negative introspection does not work reliably for sentences A in the borderline zone, since we do not have full access to whether and how much our viewpoint and our –assumed-to-be-fixed– context may change when we reflect upon A. From our viewpoint we may be unable to tell whether in context c A. By the time we consider whether we can tell that we are unable to tell whether A, we *may* have unintentionally moved from context c to a close context c^*. Such small unintentional changes can be semantically significant in the borderline zone. So, axiom **5** needs to be rejected. Note that system **FIN** does not *preclude* the possibility of negative introspection. Rather for any borderline case, whether we have a case of negative introspection is itself something to which we lack cognitive access. The argument against the addition of axiom **B** would be similar.

Part II: The Solution to the Sorites

6 The Two Basic Assumptions of the Solution, and the First Step of the Solution

The agnostic solution is based on two assumptions. The purpose of the first assumption is to remove the apparent contradiction. The purpose of the second assumption is to introduce an explanation why there seems to be a contradiction, and to do so without creating any new paradox.

A couple of remarks on notation: hereafter, unless noted otherwise, all formulae, proofs and arguments are relative to arbitrary Sorites series with *F*, with 'Sorites series' understood as introduced in Sect. 2. ('Sorites series with *F*' is short for 'Sorites series with regard to being (an) *F*'.) In the formulation of principles this is indicated by the phrase 'for arbitrary Ss_F'. The *name* of a principle is given in underlined capitals (XY). '*The XY-formula*' refers to the part in XY after 'For arbitrary Ss_F'. For example, the mathematical-induction Sorites conditional will be expressed as

6.1 For arbitrary Ss_F $\forall i\ (Fa_i \rightarrow Fa_{i+1})$ SC

and $\forall i\ (Fa_i \rightarrow Fa_{i+1})$ will be called the SC-formula. A name '*XY', where '*' is a logical operator or a combination of logical operators, names a principle obtained by prefixing '*' to the XY-formula.

The first basic assumption is common. It determines what kind of solution to the Sorites the theory suggests:

Assumption 1 If a PM-series (above, Sect. 2) is a Sorites series with *F*, then it is not the case that for all adjacent pairs of objects it holds that, if the first is *F*, so is the second.

Formally, Assumption 1 negates the Sorites conditional; i.e. negating the SC-formula for Sorites sequences yields:

6.2 For arbitrary Ss_F $\neg\forall i\ (Fa_i \rightarrow Fa_{i+1})$ ¬SC

The reasons for rejecting SC are familiar. In brief, used for PM-series, the combination of bivalence, classical first-order logic and the Sorites conditional leads to paradox. The dispensability of these three is taken to be in reverse order: SC can be eliminated without major disruption to everyday and scientific reasoning. Additional independent reasons for the expendability of SC were given in Sect. 3. This suggests that—within the bounds of classical logic—discarding SC is the right move.[16] Assumption 1 is thus the first step of the proposed solution. It repudiates the Sorites conditional SC and thereby removes the apparent inconsistency from the Sorites. (The agnostic solution

[16] The bounds are in fact wider than classical logic. As the No-Sharp-Boundaries Paradox shows, discarding the universal premise may be just as reasonable for super-intuitionistic logics with double negation elimination (e.g. Bobzien & Rumfitt, 2020, §1).

does not include a proof of ¬<u>SC</u>. Rather, the reasons for the rejection of <u>SC</u> make up part of the reasons for the acceptance of the agnostic solution.)

Neither the Sorites conditional <u>SC</u> nor Assumption 1 (or ¬<u>SC</u>) is modalized. The second basic assumption relates the interpreted modal system **FIN** to Sorites series, to tolerance, and to the Sorites paradox. More specifically, it directly connects the logic to the generally agreed statement that Sorites series have borderline cases, that is, to

6.3 For arbitrary Ss_F $\exists i\ \nabla Fa_i$

where the (boldface) ∇-operator is taken to be defined in terms of the unspecified (boldface) □-operator from above (Sect. 2) in the usual way ($\nabla A =_{df} \neg\Box A \wedge \neg\Box\neg A$). Here is the second assumption:

Assumption 2 If one specifies borderlineness in the statement that there are borderline cases in Sorites series (i.e. in *6.3*) by means of the factive-cognitively interpreted modal system **FIN**, then one can describe the grey areas in Sorites series in a way that explains why the Sorites paradox seems paradoxical.

The antecedent of Assumption 2, i.e. the specification of *6.3* by means of the interpreted **FIN**, will be expressed (with normal font ∇) as

6.4 For arbitrary Ss_F $\exists i\ \nabla Fa_i$ <u>∃∇</u>

In a nutshell the solution is this: If one assumes that the Sorites conditional is false and that borderline cases of Sorites series are defined —in the usual way— by the factive-cognitively interpreted modal logic **FIN**, then the paradox can be resolved. The following sections illustrate the link Assumption 2 provides between system **FIN**, the Sorites and tolerance.

7 The Existence of Borderline Cases as a Buffer between the Non-borderline Cases and the Relation between Borderline Cases and the Grey Zone

A solution to the Sorites that defines borderlineness in terms of modalities—among other things—for the purpose of explaining the grey area and hedging behaviour, must have room for the kind of borderline cases it defines. In specifying *6.3* as *6.4* in the antecedent of Assumption 2, the agnostic solution preserves (as part of the assumption) the existence of borderline cases in Sorites series: it is part of the solution that in every Sorites series there exists at least one borderline case of the kind defined by the cognitive-factively interpreted finality logic, that is, *a borderline case of which one cannot tell that it is borderline*.

Generally, philosophers consider the purpose of the introduction of borderline cases in modal terms to be this: that it warrants that there is no sharp boundary between the non-borderline F and the non-borderline $\neg F$ cases in a Sorites series, i.e.

7.1 For arbitrary Ss_F $\neg\exists i\,(\Box Fa_i \wedge \Box\neg Fa_{i+1})$

The specification of *6.3* by *6.4* provides this warrant:

For arbitrary Ss_F

(1)	$\exists i\ \nabla Fa_i$	assumption
(2)	$\exists i\,(\neg\Box Fa_i \wedge \neg\Box\neg Fa_i)$	(1) df ∇
(3)	$\exists i\,(\neg\Box Fa_i \wedge \neg\Box\neg Fa_i) \leftrightarrow \neg\exists i\,(\Box Fa_i \wedge \Box\neg Fa_{i+1})$	Borderline-as-buffer (*2.3*) for **FIN**[17]
(4)	$\exists i\,(\neg\Box Fa_i \wedge \neg\Box\neg Fa_i) \rightarrow \neg\exists i\,(\Box Fa_i \wedge \Box\neg Fa_{i+1})$	(3) left-to-right, *2.3*
(5)	$\neg\exists i\,(\Box Fa_i \wedge \Box\neg Fa_{i+1})$	(2), (4) **MP**

Hence $\exists\nabla$ produces a buffer between the non-borderline cases:

7.2 For arbitrary Ss_F $\neg\exists i\,(\Box Fa_i \wedge \Box\neg Fa_{i+1})$

The fact that modally expressed borderline cases provide this buffer between the non-borderline cases in a Sorites series furnishes a theory of vagueness that has the means for a suitable description of its grey area. The specification of *6.3* by *6.4* aids in the description of the grey area as follows. The non-borderline cases on either side of the borderline cases are such that one can tell (for sure) that they are non-borderline. By the definition of tell-ability, they provide areas at either end of a Sorites series in such a way that the grey area plausibly falls somewhere in between. The remaining middle section is such that the grey area can be assumed to be somewhere there. And given the empirical, and necessarily in part arbitrary, demarcation of the grey area, this is all anyone can hope for. (In Sect. 14, the question of the grey area is revisited from a different angle.) With these two points settled, I continue with the details of the solution.

8 Why the Sorites Conditional Appears True although it is Not: The Weakened Conditional

As the first step of the solution, Assumption 1 repudiates the Sorites conditional <u>SC</u>, and thus removes the element of paradox from the Sorites. An explanation is still needed why it nonetheless appears to be the case that <u>SC</u>. (It is on this point that modal theories are most at variance.) The following sections provide such an explanation. Since the Sorites conditional may appear true to different people for different reasons—or even to the same person for multiple reasons, I offer several distinct

[17] The argument for *2.3* (Borderline-as-buffer) from Sect. 2 transfers directly to the cognitive-factively interpreted borderline cases of system **FIN**.

plausible explanations of why people should think that it is true. Each explanation tallies with Sorites agnosticism.

The first explanation why SC appears true is traditional in kind. The interpreted system **FIN** serves to replace the SC formula in the false SC by a *weakened conditional* (the WC formula), resulting in the true

8.1 For arbitrary Ss_F $\forall i\ (\Box Fa_i \rightarrow \neg\Box\neg Fa_{i+1})$ WC

WC has these three features: (i) it is sufficiently similar to SC that it is plausible that people confuse it with SC, (ii) unlike SC, it is true and (iii) unlike SC, it does not lead to paradox. In natural language WC might be written

$\underline{\mathrm{WC}}_{\mathrm{NL1}}$ For any pair of adjacent objects in a Sorites series with *F*, if one can tell (for sure) that the first is *F*, then *one can't rule it out (for sure)* that the second is also *F*.

This formulation is somewhat akin to Williamson's 'If we know that the first is *F*, then the second is *F*' ($KFa_n \rightarrow Fa_{n+1}$, for example, with heaps, in Williamson, 1994, p. 232). Both formulations use an impersonal factive-cognitive modal expression and move from a modally stronger to a modally weaker formula. Alternatively, WC could be written

$\underline{\mathrm{WC}}_{\mathrm{NL2}}$ For any pair of adjacent objects in a Sorites series with *F*, if the first is *clearly F*, then *it is not clear* that the second is *not F*.

Questions of meta-language versus object-language operators aside, $\underline{\mathrm{WC}}_{\mathrm{NL2}}$ is akin to formulations like 'If the first is definitely true, then it is not definite that the second is not definitely true' (which corresponds to Wright, 1992, p. 130, (iii)) or generally to $D^m Fa_i \rightarrow \neg D\neg D^{m-1} Fa_{i+1}$, with $m \geq 1$. With the modal logic **FIN** one obtains an analogous generalization by allowing for modalized predicates $\Box^n F$ with $n \geq 0$, in lieu of *F*. Thus WC holds up fine compared with ersatz conditionals suggested by other well-known theories.

As regards the truth of WC, it can readily be shown that, given the antecedent of Assumption 2, it is true and not trivially so. The proof in first-order logic is:
For arbitrary Ss_F

(1)	$\neg\exists i\ (\Box Fa_i \wedge \Box\neg Fa_{i+1})$	(*7.2*)
(2)	$\forall i\ \neg(\Box Fa_i \wedge \Box\neg Fa_{i+1})$	(1) double negation introduction, duality of $\exists,\forall$
(3)	$\forall i\ (\Box Fa_i \rightarrow \neg\Box\neg Fa_{i+1})$	(2) df$\rightarrow$

(3) is WC. So given the antecedent of Assumption 2, the weakened conditional WC is true. The non-Soritic PM-sequences show that the WC formula is not trivially true. A non-Soritic sequence with *F* is a PM-series for which it holds that

8.2 $$\exists i\ (\Box Fa_{i} \wedge \Box \neg Fa_{i+1})$$

8.2 is incompatible with the WC formula $\forall i\ (\Box Fa_{i} \rightarrow \neg \Box \neg Fa_{i+1})$.[18] So the WC formula does not hold trivially for all PM-series but specifically for Soritic ones.

The considerations about the borderline areas of vague predicates from Sect. 3 adduced empirical reasons why it may be acceptable to replace SC by WC: the hedging behaviour of even fully competent speakers may reflect the fact that it is not clear that there are sufficient reasons for asserting instances of SC for borderline cases. Further, in order for WC to be an acceptable stand-in for arguments that use SC, it also must not itself lead to Sorites-like paradoxes or be otherwise inconsistent. Indeed, WC does not permit the construction of a chain of arguments that corresponds to the chain of arguments in a step-by-step Sorites. In a step-by-step Sorites, *modus ponens* arguments with particular Sorites conditionals chain together, with the conclusion of the first argument providing the atomic premise for the next, etc., until the paradoxical conclusion is reached. One cannot construct a similar chain of arguments with the WC formula. This is easy to see. The particular conditionals would all have the form $\Box\varphi(a_{i}) \rightarrow \neg\Box\neg\varphi(a_{i+1})$. Using these as conditional premises in *modus ponens* arguments, the conclusion of the first *modus ponens* argument, for example, would be $\neg\Box\neg Fa_{2}$. However, the required premise for the construction of the next *modus ponens* argument would be $\Box Fa_{2}$. So the conclusion of one argument cannot serve as a premise for the next and no paradox ensues with WC. Nor, as far as I am aware, is there any other analogous kind of argument chain.

9 Why the Sorites Conditional Appears True although It is Not: Tolerance Tangle

A second reason why it is believed that SC is true has to do with tolerance. Since SC cannot be proved directly, people frequently try to justify it by supporting it with a principle meant to capture tolerance. They may say things like (*a*) "no-one can reasonably deny (or say that it's not the case) that in Sorites series for adjacent pairs, if one is *F*, so is the other", or (*b*) "for adjacent pairs in a Sorites series, if one is *F*, the other must be, too", or (*c*) "any reasonable person must admit that in Sorites series, of adjacent pairs, either both are *F* or neither is". It is not at all obvious what the logical structure of such sentences is assumed to be; and it may well be unclear what

[18] This becomes obvious once one takes into account that the steps from (1) to (2) and from (2) to (3) are first-order equivalence transformations and that (1) is the negation of *8.2*. Proof:

For arbitrary $Ss_{F.}$

(1)	$\forall i\ (\Box Fa_{i} \rightarrow \neg\Box\neg Fa_{i+1})$	WC (*8.1*)
(2)	$\forall i\ \neg(\Box Fa_{i} \wedge \Box\neg Fa_{i+1})$	(1) df$\rightarrow$
(3)	$\neg\exists i\ (\Box Fa_{i} \wedge \Box\neg Fa_{i+1})$	(2) duality of $\exists,\forall$ ((3) = *7.2*)
(4)	$\exists i\ (\Box Fa_{i} \wedge \Box\neg Fa_{i+1})$	*8.2*
(5)	$\neg\exists i\ (\Box Fa_{i} \wedge \Box\neg Fa_{i+1}) \wedge \exists i\ (\Box Fa_{i} \wedge \Box\neg Fa_{i+1})$	(3), (4) $\wedge$-introduction

someone who comes out with (*a*), (*b*) or (*c*) intends its precise logical structure to be. Any of the following formalizations might be intended. Here, as above in Sect. 2, the (boldface)$\Box$ is not afforded any specific clarity/determinacy/etc. interpretation.

9.1	For arbitrary Ss_F	$\forall i\ (Fa_i \leftrightarrow Fa_{i+1})$
9.2	For arbitrary Ss_F	$\Box\forall i\ (Fa_i \leftrightarrow Fa_{i+1})$
9.3	For arbitrary Ss_F	$\neg\Box\neg\forall i\ (Fa_i \leftrightarrow Fa_{i+1})$

Principles *9.1*, *9.2* and *9.3* each lead to Sorites-like paradoxes, and I believe them to be false. SC would follow directly from *9.1* and from *9.2*. For *9.3* one would appeal to MONOTONICITY$_F$, POLAR, and the impossibility of disproving SC.[19] Those who express tolerance principles in natural language easily confuse *9.1*, *9.2*, and *9.3*, and may leave it underdetermined which of them is intended. Moreover, *9.3* is easily mistaken for the very similar

2.4	For arbitrary Ss_F	$\forall i\ \neg\Box\neg(Fa_i \leftrightarrow Fa_{i+1})$	TOLERANCE$_{\neg\Box\neg}$

English sentences like (a) above might be supposed to express *9.3* or *2.4*, and confusions of $\forall x\Diamond A(x)$ and $\Diamond\forall xA(x)$ are common even among the brainier of humans. But whereas *9.3* leads to paradox, *2.4* does not. In fact, when $\Box$ is defined by system **FIN** and is given the tell-ability interpretation, *2.4* becomes

9.4	For arbitrary Ss_F	$\forall i\ \neg\Box\neg(Fa_i \leftrightarrow Fa_{i+1})$	TOL

and TOL entails the weak conditional WC (*8.1*).[20]

This then is the second explanation: when trying to justify SC by principles of tolerance, people confuse versions of *9.1* or *9.2* with *9.3* and *9.3* (or *9.1* or *9.2* directly) with *2.4*. They ride on the modal ambiguities in these sentences, moving between readings that justify SC but are themselves paradoxical and readings that do not justify SC and are not paradoxical—or in any case not with the modal system **FIN**. The complexity of the semantics of epistemic modals is very likely to add to the muddle. To give an example, the 'must' in (b) may be construed as weaker than *9.2* and even weaker than the naked universal quantifier in *9.1*.

[19] E.g. 'one cannot rule out that $(Fa_1 \leftrightarrow Fa_2)\wedge(Fa_2 \leftrightarrow Fa_3)\wedge\ldots\wedge(Fa_{n-1} \leftrightarrow Fa_n)$, but by POLAR one can rule out that $Fa_1 \leftrightarrow Fa_n$, with $\neg\Box\neg$ for 'one cannot rule out that'.

[20] For arbitrary Ss_F.

(1)	$\forall i\ \neg\Box\neg(Fa_i \leftrightarrow Fa_{i+1})$	TOL
(2)	$\forall i\ \neg\Box\neg(Fa_i \rightarrow Fa_{i+1})$	(1) df $\leftrightarrow$, $\wedge$-elimination, substitution into modal context
(3)	$\forall i\ (\Box Fa_i \rightarrow \neg\Box\neg Fa_{i+1})$	(2) **K7**, i.e. $\neg\Box\neg(A \rightarrow B) \leftrightarrow (\Box A \rightarrow \neg\Box\neg B)$, **∀2**

Formula (3) is the WC-formula (= *8.1*).

The third explanation I offer also concerns the use of tolerance in purporting to justify SC. Some may erroneously overgeneralize the phenomenon of tolerance. When they consider tolerance in its *9.1*, *9.2* or *9.3* versions, it appears true to them (in each version), since they de facto imagine only non-borderline cases. That is, the objects they think of, and think of as F (or as $\neg F$), are cases that, first, are themselves non-borderline (either non-borderline F or non-borderline $\neg F$) and that, second, are in the Sorites series sufficiently close to a polar end that the adjacent cases towards the middle are also still non-borderline. In contrast, we can assume that, when people are invited to start with an object of which they are unable to say whether it is F, they display greater hesitation in saying of an object that appears just like it (but is known to differ slightly) that it must be F-wise whenever the other is F-wise, despite the fact that it is unclear to them what it is (whether it is F or not, or something else still).[21] So here the error is overgeneralization. This is an unsurprising blunder, since—as I emphasized earlier—non-borderline cases are the norm, borderline cases the exception, and it seems unachievable for people to agree upon actual examples of borderline cases of the kind relevant to a Sorites.

In sum, the three—non-exclusive—explanations why SC appears true even though it is not are (i) that the Sorites conditional SC is taken for the weakened conditional WC, (ii) that a false stronger tolerance principle is taken for a true weaker one (e.g. a version of *9.1* or *9.2* or *9.3* for a version of *2.4*), (iii) that, by overgeneralization, the false tolerance principles *9.1*, *9.2* and *9.3* are taken for similar true biconditionals for non-borderline cases. In each instance, the misstep is to take a false and paradox-inducing principle for a very similar true and non-paradox-inducing one.

10 Summary of the Solution

This completes the agnostic solution to the Sorites in its elementary form. To sum up: the agnostic solution makes two assumptions. Assumption 1 negates the Sorites conditional and removes the seeming contradiction in the paradox. Assumption 2 states how one obtains an explanation of why there seems to be a contradiction, namely on the basis of $\underline{\exists\nabla}$ (= *6.4*). If it is agreed that there are borderline cases in Sorites series and these borderline cases are defined by the cognitive-factively interpreted finality logic, then one can explain why the Sorites appears paradoxical. In more detail: $\underline{\exists\nabla}$ entails the weakened conditional WC; and for PM-series the weak tolerance TOL entails WC. WC and TOL each can explain why the Sorites conditional SC appears true even though it is not. This vindicates the choice of Assumption 2.

The relations between $\underline{\exists\nabla}$, WC and TOL are in fact even closer. Over PM-series, the formulae of $\underline{\exists\nabla}$ and WC are materially equivalent. Combined, proofs in Sects. 7

[21] This is implied for instance in Raffman (1994).

In system FIN, for arbitrary Ss_F

$\underline{\mathrm{TOL}_1}$ $\rightarrow$	$\underline{\mathrm{TOL}_2}$ $\rightarrow$	$\underline{\mathrm{TOL}_3}$ $\rightarrow$	$\underline{\mathrm{TOL}}$	(if true, apparent truth of Sorites can be explained)		
	$\downarrow$		$\downarrow$			
	$\underline{\mathbf{SC}}$ (assumed not true) $\rightarrow$		$\underline{\mathrm{WC}}$	$\leftrightarrow$	$\underline{\boldsymbol{\exists\nabla}}$	(if true, apparent truth of Sorites can be explained)

Fig. 1 Diagrammatic representation of the chief logical relations between principles relevant to the proposed solution

and 8 establish, via *7.1*, that $\underline{\exists\nabla}$ entails $\underline{\mathrm{WC}}$. The converse is also readily shown.[22] So we have

10.1 $\exists i\ \nabla Fa_i \leftrightarrow \forall i\ (\Box Fa_i \rightarrow \neg\Box\neg Fa_{i+1})$ $\underline{\exists\nabla} \leftrightarrow \underline{\mathrm{WC}}$

Since, over PM-series, $\underline{\mathrm{TOL}}$ entails $\underline{\mathrm{WC}}$,[23] $\underline{\mathrm{TOL}}$ also entails $\underline{\exists\nabla}$.

For the readers' benefit **Fig. 1** (next page) provides a 'flow diagram' that represents the chief logical relations between the principal sentences, *even if one disregards their assumed truth-values,* i.e. in particular the falsehood of $\underline{\mathrm{SC}}$, $\underline{\mathrm{TOL}}_1$, $\underline{\mathrm{TOL}}_2$ and $\underline{\mathrm{TOL}}_3$. Here '$\rightarrow$' is used for **FIN**-implication over Sorites series and '$\leftrightarrow$' for **FIN**-equivalence over Sorites series. The two phrases in brackets after the two names in bold indicate how these two formulae relate to the two basic assumptions of the solution. (Note that in **Fig. 1** the names of the principles stand in for their formulas. This is *solely* for convenience, to allow one to take in the relations at one glance. TOL_1, TOL_2 and TOL_3 correspond to *9.2*, *9.1* and *9.3*, except that here they are logically fully specified by **FIN**.)

In the agnostic solution offered, $\underline{\mathrm{SC}}$, $\underline{\mathrm{TOL}}_1$, $\underline{\mathrm{TOL}}_2$ and $\underline{\mathrm{TOL}}_3$ are not true, and consequently the Sorites is not paradoxical, and $\underline{\mathrm{TOL}}$ and/or $\underline{\mathrm{WC}}$ and $\underline{\exists\nabla}$, if true, provide the backbone of an explanation why the Sorites *appears* paradoxical.

[22] For arbitrary Ss_F.

(1)	$\forall i\ (\Box Fa_i \rightarrow \neg\Box\neg Fa_{i+1})$	$\underline{\mathrm{WC}}$ (= *8.1*)
(2)	$\forall i\ \neg(\Box Fa_i \wedge \Box\neg Fa_{i+1})$	(1) df $\rightarrow$
(3)	$\neg\exists i\ (\Box Fa_i \wedge \Box\neg Fa_{i+1})$	(2) df $\exists$, double negation elimination
(4)	$\exists i\ (\neg\Box Fa_i \wedge \neg\Box\neg Fa_i) \leftrightarrow \neg\exists i\ (\Box Fa_i \wedge \Box\neg Fa_{i+1})$	Borderline-as-buffer (= *2.3*)
(5)	$\neg\exists i\ (\Box Fa_i \wedge \Box\neg Fa_{i+1}) \rightarrow \exists i\ (\neg\Box Fa_i \wedge \neg\Box\neg Fa_i)$	(4) df $\leftrightarrow$, $\wedge$-elimination
(6)	$\exists i\ (\neg\Box Fa_i \wedge \neg\Box\neg Fa_i)$	(3), (5) **MP**
(7)	$\exists i\ \nabla Fa_i$	(6) df ∇

[23] See above, Sect. 9.

Part III: The Sorites and the Semi-determinability of Vagueness

This part gets to the heart of the solution. It assumes familiarity with the notion of semi-decidability.

11 Semi-Decidability and the Semi-determinability of Vague Expressions

Sections 6–10 set out the basic structure of the agnostic solution. They did not broach the questions of (i) how the solution avoids detectable sharp boundaries between the borderline and the non-borderline cases and (ii) how the solution stays clear of higher-order-vagueness paradoxes. To see how the finality logic provides the logical foundation for an answer to these two questions, one needs to reflect on the structural property of vague predicates that it brings out. This is the property—I argue—that makes vague predicates susceptible to paradox and as such helps explain the apparent inconsistency in the Sorites. It can be best explained by using the standard informal definition of semi-decidability, or more precisely of *proper* semi-decidability (by which I mean semi-decidability *without decidability*):

11.1 A class of questions is *properly semi-decidable* if and only if there is a procedure that comes to a halt and says 'yes' iff the answer is positive, but there is no procedure that comes to a halt and says 'no' iff the answer is negative.

Here I am solely interested in the *generic property* defined by *11.1*. In particular, there is no assumption that the accept-reject method of *11.1* is restricted to mathematical procedures that make the definition one of recursive enumerability. For now, any method counts that involves a class of questions 'A?' over the expressions of a language $\mathcal{L}$ such that the method outputs 'yes' for cases with a positive answer, but does not output 'no' for (all) cases with a negative answer, and that for systematic reasons leaves it open for some A whether the answer is positive or negative however often the question is asked. To pre-empt misunderstandings, I use the term semi-*determinability* (with the 'proper' being understood) from here on.[24]

How is semi-determinability exemplified in the case of vague predicates? What is the class of questions? The semi-determinable class of questions would be of the form 'A (for sure)?' for each sentence A of a vague language $\mathcal{L}$ that includes the modal operators $\Box$ and ∇. The bracketed expression '(for sure)' indicates that the responses invited are reasoned and without doubt. (The questions can be imagined as similar to questions asked under oath.) The method which manifests this semi-determinability would run over all sentences A of $\mathcal{L}$, with the same context fixed for all expressions. Individuals fully competent with regard to A (rather than a Turing computer) would function as assessors of A whose responses 'yes' or 'no' bring the

[24] So, (proper) semi-decidability entails (proper) semi-determinability, but not vice versa.

procedure to a halt, which means that the response is recorded and the question is not going to be asked again.

The epistemically (in the wider sense) interpreted modal logic **FIN** can be utilized to bring out this semi-determinability, to which the following two meta-principles of **FIN** correspond: for any n, $\Box A \leftrightarrow \Box^n A$ and for any n, $\nabla A \leftrightarrow \nabla^n A$. In the interpreted logic, for any n and any sentence A, we have either $\Box^n A$ or $\Box^n \neg A$ or $\nabla^n A$. Then (i), halting corresponds to the cognitively accessible cognitive accessibility (*absolute* cognitive accessibility) which yields $\Box^n$. This yields **Rule 1**: "When the procedure halts for A, the cognitive accessibility that makes it halt, being absolute, will make $\Box A$ halt, too."[25] So, in the first case above ($\Box^n$), the answer to the question 'A (for sure)?' is positive; in the second ($\Box^n\neg$), the answer to '$\neg A$ (for sure)?' is positive; and so subsequently for $\Box A$, $\Box\Box A$, etc. (ii) The indefinite looping that prevents halting and characterizes semi-decidability in Turing procedures has its semi-determinability analogue in the cognitively inaccessible cognitive inaccessibility (*absolute* cognitive inaccessibility) of the third case (∇^n). This yields **Rule 2**: "Since cognitive inaccessibility is *absolute*, when the procedure halts for any modalization of A with $\Box$ or ∇, it will thereafter also halt for A." Consequently, if A is cognitively inaccessible, so are any modalizations of A with $\Box$ or ∇. So, in the case of cognitive inaccessibility, no part of the procedure ever yields information that leads to a halt for either A or any of its modalizations, however often any of these recur as a question. This results in never-ending loops of coming up for assessment, no assessment, coming up for (re)assessment, no assessment …, for all relevant sentences. (iii) Finally **Rule 3**: "A response 'no' can be given to questions 'A (for sure)?' with negative answers if a response 'yes' has been given to '$\neg A$ (for sure)?'."

The following analogy with provability logic may be helpful. The assessments of sentences by qualified individuals can be compared to the proofs of sentences in provability logic (e.g. Boolos, 1993). The epistemically interpreted modal systems **S4M** and **FIN** attest the semi-determinability of the class of the first-order fragment with $\Box$-operator of the sentences of a vague language $\mathcal{L}$ with regard to cognitive accessibility in a way similar to how the interpreted Grzegorczyk modal system **Grz** attests the semi-decidability of the class of sentences of Peano Arithmetic (*PA*) with regard to truth-cum-provability. (For the latter, see again Boolos, 1993.) So, instead of truth-cum-provability for sentences of *PA*, for the sentences of a vague language we have truth-cum-non-borderlineness (or absolute cognitive accessibility). *Formally*, the assessments by competent individuals are of utter simplicity, compared with the complexity of proofs. (In substance, of course, the competence of the individuals is highly complex and sensitive to detail.) But this is not the point of comparison here.[26]

[25] See Haldèn (1963) and Burgess (1999) for related suggestions for axiom **4**.

[26] Systems **Grz** (Grzegorczyk, 1967) and **S4M** are very close. As shown by Esakia (1976), both are modal companions of the intuitionistic sentential calculus, with **Grz** being the largest such system and an extension of **S4M.** Both are reflexive, transitive, and final. **Grz** exceeds **S4M** by one axiom, as yet not discovered, which makes **Grz** Noetherian, i.e., makes it contain neither non-trivial cycles nor infinite ascending chains. Cf. Wolter and Zakharyaschev (2014) for some details. **S4M** has no non-trivial cycles but allows for infinite ascending chains.

The significance of the semi-determinability of the class of sentences of a vague first-order modal language $\mathcal{L}$ can be discerned at different levels. Sects. 12 and 13 present the two that are directly relevant to the Sorites.

12 Semi-determinability and the Borders of the Borderline Zone

Usually, in modal theories of vagueness, the interpreted $\Box$-operator (or analogue) demarcates the extensions of the clear (definite, determinate …) cases of some predicate F. As a result, with the context fixed and all, in a Sorites series with F there are so-called *sharp boundaries* between the non-borderline F and the borderline F on either side of the borderline zone, delimiting three *extensions*. There are a last non-borderline-F and a first borderline-F as well as a last borderline-F and a first non-borderline-not-F. Such a set-up leads to several problems. The first is that it introduces two sharp boundaries that appear to have no counterpart in the phenomenon the theories seek to capture: the grey zone, the hedging behaviour, *et sim*.

With the class of vague expressions understood as semi-determinable and borderlineness defined by the cognitive-factively interpreted finality logic the situation is rather different. For a Sorites series with F, the logic offers the following picture. From left to right, for any n, a sequence of $\Box^n F$ cases is followed by a sequence of $\nabla^n F$ cases, followed by a sequence of $\Box^n \neg F$ cases. However, the *point* of the logic is not to delimit extensions that correspond to these three sequences. The point of the logic is to explain why there are no (or no accessible) fully determined extensions.

From Monotonicity$_{\Box}$ and Polar one can derive that for any Sorites series with F it is clear that there is a last clear case. But with the $\Box$-operator defined by the factive-cognitively interpreted **FIN**, it is not clear which case this is: let b be the last $\Box^n F$ case of the series for some n; then it is not clear that b is the last $\Box^n F$ case. We cannot tell that b is the last clear case. Since any borderline case a is $\nabla^{\mathrm{n}} F$ for any n, one can never rule out that for a there is an $n+1$ such that a is $\nabla^{\mathrm{n+1}} F$. So, one cannot rule out that a is the last clear case of F, nor that a is the last clear case of $\Box^{\mathrm{n}} F$ for any n.

In terms of the unsavoury assessment sensitivity, the reason for this is that for any borderline case a there is an accessible viewpoint ('world') v at which a is a non-borderline case, i.e. a case that is $\Box^{\mathrm{n}} Fa$ for any n or a case that is $\Box^{\mathrm{n}} \neg Fa$ for any n. We cannot tell that a is not $\Box F$, since there always is a *final viewpoint* v at which a is $\Box F$ and we cannot tell that a is not $\Box^2 F$, since at this *viewpoint* a is also $\Box^2 F$, and so on, and the same for $\Box \neg F$.

In terms of semi-determinability, in a Sorites series up to the last case of which we can tell that it is clear, the response to 'Fa for sure?' is 'yes'. For the sequence of borderline cases, there is no response to 'Fa for sure?', '$\Box Fa$ for sure?', $\Box \neg Fa$ for sure?', '$\Box^2 Fa$ for sure?', '$\Box^2 \neg Fa$ for sure?', and so on. At no level n is there a response to whether $\Box^n Fa$, and the possibility is always open that the response to '$\Box^{n+1} Fa$ for sure?' will be 'yes'. The method does not lead to a halt for any of these sentences. Each of them will come up for assessment again and again in accordance with some suitable algorithm. The method establishes tell-ability, and there is no other way to establish whether *A* than that relevantly competent individuals can tell that *A*.[27]

The fact that the same cases come up repeatedly, and that each time the competent speakers cannot respond with either 'yes' or 'no', does not enable the speakers to tell that such cases are borderline cases. Recall that the speakers are taken not to have a reliable meta-perspective and that this assumption reflects the fact that once we are in the grey area, we may always be unable to discern whether the difference between two assessments of the same case is due to a shift in context or to a difference in viewpoint (above Sect. 5), and hence that prior assessments of the same questions are not reliable.

So, although there is a last clear case of *F*, the finality logic brings out how it cannot be established which case in a Sorites series this last case is. (And this is the point of the logic at the level of sentences in a Sorites.) *The reason why* this cannot be established, that is, whether the reason is epistemic, semantic or ontic in nature, or any combination of these, is side-stepped here and picked up in *Part IV*. The solution offered separates these two questions, (i) the question regarding the logical structure and the cognitive element of the solution and (ii) the metaphysical question about *the reason why* this is the structure of vague expressions.

The finality logic accurately represents the absence of sharp boundaries between the non-borderline cases and the borderline cases that has its origin in the semi-determinability of vague predicates and that matches the empirical datum that it seems impossible to establish where exactly the hedging behaviour starts and where it ends. (The two meta-principles of **FIN** that, for any n, $\Box A \leftrightarrow \Box^n A$ and $\nabla A \leftrightarrow \nabla^n A$, indicate this, too.) The existence of last clear cases is thus not the same as the existence of sharp boundaries between the borderline and non-borderline cases.

At this level, it also becomes clear how system **FIN** makes it possible to stake out the grey area in a Sorites series. The affirmative responses to literals *Fa*, ¬*Fa* on either side of the borderline zone will correspond roughly to the sections that enclose the grey area. (A grey area can never be exactly delineated because it is established empirically.) Whenever a fully informed individual can respond to '*A* (for sure)?' with 'yes', other individuals' hedging behaviour can likely be put down to lack of qualification or some other factor not essential to the Sorites (context not sufficiently specified, etc.). Thus the grey area does not correspond to a *third kind* of cases. It corresponds to those cases that evade assessment, but for which one cannot rule out that they are of one or the other kind.

The last two paragraphs indicate a similarity between the proposed agnostic solution and Crispin Wright's claim that vagueness leaves us in a quandary (Wright,

[27] Competent individuals need not be silent on every question regarding borderline cases. For some borderline cases, they may be able to respond 'yes' to 'is it the case that you can't rule out that *Fa*?'; for others they may be able to respond 'yes' to 'is it the case that you cannot tell that *Fa*?'.

2001, 2003). Where Wright seems to hold that, for borderline *a* with regard to *F* it is not knowable whether it is knowable whether *Fa* (Wright, 2001) or that, if it is indeterminate whether *Fa*, then it is indeterminate whether it is indeterminate whether *Fa* (Wright, 2003), the agnostic solution offered here maintains that if it is cognitive-factively inaccessible whether *Fa*, then it is cognitive-factively inaccessible whether it is cognitive-factively inaccessible whether *Fa*, or that, if *a* is borderline *F*, it is borderline borderline *F*. Unlike Wright's theory, this solution includes a coherent normal modal system that can be taken to express this fact. (For the logical relation between **FIN** and intuitionistic theories of vagueness see Sect. 17).

13 Semi-determinability and the Borderline Existence of Borderline Cases

The second level at which semi-determinability is relevant to the Sorites is that of quantified modalized literals. The relevant class of questions here is whether (for sure) some predicate *F* has no borderline cases. The relevant sentences *A* are of the form $\neg\exists x\nabla(\varphi)x$. Where at the first level the semi-determinability class of questions is whether some objects are *F* (for sure), at this second level the relevant class of questions is whether a predicate has no borderline cases (for sure)—or whether a predicate is *precise*, for theories that take the vagueness of an expression to be tantamount with it having borderline cases. Now distinguish the following two cases.

First assume that there *are* borderline cases of some *F*, i.e. $\exists x\nabla Fx$. Then *F* is viewpoint-sensitive and one cannot rule out that there is a borderline case. One also cannot rule out that there is no borderline case, and that is, one also cannot rule out that every object is either $\Box^n F$ or $\Box^n\neg F$ for any *n*. So, on the assumption that $\exists x\nabla Fx$, there is no answer to the question whether there exist borderline cases of *F*. In the interpreted logic **FIN**, whenever there exists a borderline case of *F*, one cannot tell (for sure) whether there exists a borderline case of *F*. Formally, this can be expressed, via theorems V and V∃, with the meta-principle

13.1 for any *n* $\exists x\nabla Fx \rightarrow \nabla^n\exists x\nabla Fx$

Second assume that in finite PM series (see sect. 2) *F* has no borderline cases, i.e. $\neg\exists x\nabla Fx$. Then the predicate *F* is viewpoint-insensitive. In this case one can tell that there is no borderline case of *F*. Of every object up for assessment one can tell that it is either clearly *F* or clearly not *F*. (Consider 'is a prime number under 1000'.) Thus, on the assumption that in a PM series *F* has no borderline cases, or $\neg\exists x\nabla Fx$, the response to the question if (for sure) there exist no borderline cases of *F* is always 'yes'. In the interpreted logic **FIN**, whenever in a PM series *F* has no borderline cases, one can tell (for sure) that *F* has no borderline cases. If one adds the Barcan formula to **FIN** (Cresswell 2001), this generalizes to the meta-principle

13.2 for any n $\neg\exists x\nabla Fx \rightarrow \Box^{n}\neg\exists x\nabla Fx$[28]

This *lack of determinability of the existence of borderline cases* has significance for the solution of the Sorites. Since an explanation of the paradox requires that there are borderline cases in a Sorites series, in system **FIN** it also requires that it is borderline whether there are borderline cases—another consequence of the definition of borderlineness by the finality logic. Semantically, the borderlineness of whether there are borderline cases results directly from the assessment sensitivity of vague predicates. As said above, the underlying philosophical theory, in crudest outline, is that for every borderline case in a Sorites series with F, even with the context fixed, there is a viewpoint from which it looks non-borderline F or a viewpoint from which it looks non-borderline $\neg F$. Semantically, this introduces the general possibility that for any sub-sequence of borderline cases $a_m, \ldots, a_{m+n}$, of a Sorites series, there is a viewpoint regarding $Fa_m, \ldots, Fa_{m+n}$ at which there are no borderline cases.[29] Given that, for any arbitrary Sorites series with F, there is at least one such viewpoint, *one cannot rule out* that there are no borderline cases in the Sorites series—however unlikely this may be. Nor, of course, can one rule out that there are borderline cases (Sects. 3 and 7). It is in keeping with this theory, that the agnostic solution employs the principle that it is borderline whether there are borderline cases.

Modally, the borderlineness of the existence of borderline cases in a Sorites series can be expressed with the ∇ operator as

13.3 For arbitrary Ss_F $\nabla\exists i\nabla Fa_i$ $\nabla\underline{\exists}\nabla$

[28] Proof sketch:

(1)	$\neg\exists x\nabla Fx$	assumption
(2)	$\forall x\neg\nabla Fx$	(1) df $\exists$
(3)	$\forall x\neg(\neg\Box Fx\wedge\neg\Box\neg Fx)$	(2) df ∇
(4)	$\forall x(\Box Fx\vee\Box\neg Fx)$	(3) DeMorgan
(5)	$\forall x(\Box\Box Fx\vee\Box\Box\neg Fx)$	(4) axiom **4**, constructive dilemma
(6)	$\forall x\Box(\Box Fx\vee\Box\neg Fx)$	(5) $\vee$-agglomeration for the $\Box$-operator
(7)	$\Box\forall x(\Box Fx\vee\Box\neg Fx)$	(6) **BF**
(8)	$\Box\neg\exists x\nabla Fx$	(7) DeMorgan, df $\exists$, double negation elimination, df ∇ (substitution into modal context)
(9)	$\neg\exists x\nabla Fx\rightarrow\Box\neg\exists x\nabla Fx$	(1), (8) assumption discharged

One obtains the meta-principle *13.2* by induction via adding n applications of axiom **4** on (7) and modifying (8), (9) accordingly. (Bobzien 2015 considers mainly **FIN** with the Barcan formula added. However, this restriction to single-domain Kripke semantics is not necessary for the Sorites solution.)

[29] More specifically, there are a number of such viewpoints, where the number of viewpoint-conjuncts is that of the borderline cases and the number of viewpoints is restricted by POLAR and MONOTONICITY$_C$ to that of the number of the borderline cases plus one. For example, with three borderline cases a_m, a_{m+1}, a_{m+2}: viewpoint (i) $\Box Fa_m$, $\Box Fa_{m+1}$, $\Box Fa_{m+2}$, viewpoint (ii) $\Box Fa_m$, $\Box Fa_{m+1}$, $\Box\neg Fa_{m+2}$, viewpoint (iii) $\Box Fa_m$, $\Box\neg Fa_{m+1}$, $\Box\neg Fa_{m+2}$, viewpoint (iv) $\Box\neg Fa_m$, $\Box\neg Fa_{m+1}$, $\Box\neg Fa_{m+2}$.

In system **FIN**, $\nabla\underline{\exists\nabla}$ follows from the assumption of $\underline{\exists\nabla}$, by theorems $\underline{\nabla\exists}$ and **MP**. $\nabla\underline{\exists\nabla}$ is part of the full explanation of the paradoxicality of the Sorites (Sect. 14).

14 Lack of determinability, Higher-Order Vagueness Paradoxes and Provability Logic

The manifestation of semi-determinability with regard to the borderline zones of Sorites series served to explain the absence of cognitively accessible sharp boundaries of the grey areas. Of equal significance is the lack of determinability in the sub-class of questions whether $\neg\exists x\nabla FxF$ for arbitrary F: It renders the agnostic solution immune to all known higher-order vagueness paradoxes. This is a major advantage of the theory. The key element of the argument that shows the immunity uses the fact that in system **FIN**, since ∇A entails $\neg\Box A$, we also have entailed by $\underline{\exists\nabla}$

14.1	For arbitrary Ss_F	$\neg\Box\exists i\nabla Fa_i$	$\neg\Box\underline{\exists\nabla}$

(Details are set out in Bobzien 2015, 76-80.) This fact allows us to extend the comparison of the agnostic solution with provability logic. Looking at **Grz** and truth-cum-provability (or alternatively **GL** and provability), there are sentences A in the language of *PA* that are true in the standard model but not provable. These sentences all contain the provability predicate of the logic. Similarly, with the factive-cognitively interpreted **FIN** there are sentences A in the language $\mathcal{L}$ that are true but not cognitively accessible, and they all contain the modal operator ∇ of the logic. (Their most basic form is $\exists x\nabla\varphi(x)$). Just as in the case of provability logic the unprovable true A can be proved in a stronger theory (which will contain unprovable true sentences B, for which there will be an even stronger theory, and so forth), so in the language $\mathcal{L}$ of **FIN** the true but cognitively inaccessible existential modal sentences A would be both true and cognitively accessible in a stronger theory, but that theory would itself contain new existential modal sentences B that are true but cognitively inaccessible, and so forth. (Bobzien 2013, 4-7, 17-30, provides an informal template for constructing such enriched more precise languages.)

We can now also see why the second assumption of the agnostic solution introduced $\underline{\exists\nabla}$ conditionally on explaining the paradoxicality of the Sorites: $\underline{\exists\nabla}$, though assumed to be true, is not (fully) epistemically accessible. (In some theories of knowledge and assertion, the assumptions in a theory must be knowable—since (i) the assumptions of a theory must be assertible and (ii) one must only assert what one knows; I happen to disagree with both (i) and (ii). However, the theory I am proposing—*as* I am proposing it—is independent of my view on these points.)

Finally, to conclude Part III, an update to the ‘flow diagram’ from Sect. 10 (next page). Due to the equivalence in **FIN** between $\underline{\exists\nabla}$ and $\underline{\text{WC}}$, on the assumption of $\underline{\exists\nabla}$, we also have

In system FIN, for arbitrary SsF

$$\begin{tikzcd}
\underline{\mathrm{TOL}_1} \arrow[r] & \underline{\mathrm{TOL}_2} \arrow[r] \arrow[d] & \underline{\mathrm{TOL}_3} \arrow[r] & \underline{\mathrm{TOL}} \arrow[d] & \\
& \underline{\mathbf{SC}} \text{ (assumed to be not true)} \arrow[rr] & & \underline{\mathrm{WC}} \arrow[r, leftrightarrow] \arrow[d] & \underline{\exists\nabla} \text{ (if true, the apparent truth of the Sorites can be explained)} \arrow[d] \\
& & & \nabla\underline{\mathrm{WC}} \arrow[r, leftrightarrow] \arrow[d] & \nabla\underline{\exists\nabla} \arrow[d] \\
& & & \neg\Box\underline{\mathrm{WC}} & \neg\Box\underline{\exists\nabla}
\end{tikzcd}$$

Fig. 2 Refined diagrammatic representation of the chief logical relations between principles relevant to the proposed solution

14.2 For arbitrary Ss_F $\nabla\forall i\,(\Box Fa_i \rightarrow \neg\Box\neg Fa_{i+1})$ $\nabla\underline{\mathrm{WC}}$[30]

Philosophically, $\nabla\underline{\mathrm{WC}}$ is explained in much the same way as $\nabla\underline{\exists\nabla}$. Moveover, since by definition of ∇, ∇A entails $\neg\Box A$, from the assumption of $\underline{\exists\nabla}$, in addition to $\neg\Box\underline{\exists\nabla}$ (*14.1*), we also obtain $\neg\Box\underline{\mathrm{WC}}$.[31] Fig. 2 supplements Fig. 1 with the results of Part III. (Once again, the diagram represents the relations that hold *even if the assumed truth-values of the formulae are disregarded*, in particular the falsehood of the formulae $\underline{\mathrm{SC}}$, $\underline{\mathrm{TOL}_1}$, $\underline{\mathrm{TOL}_2}$ and $\underline{\mathrm{TOL}_3}$.)

Part IV: Different kinds of agnostic solutions

This part explains how the agnostic solution can be supplemented with a variety of semantic theories. Depending on whether one preserves any of a non-arbitrary cut-off, $\mathrm{M{\scriptsize ONOTONICITY}}_F$, bivalence, and *tertium non datur*, the agnostic solution gives rise to epistemicist agnosticism, supervaluationist-style agnosticism, chaotic (psychological-contextualist) agnosticism, polar agnosticism or intuitionism-style agnosticism.

[30] Informal proof of *14.2* For arbitrary Ss_F

(1)	$\nabla\exists_i\nabla Fa_i$	assumption $\nabla\underline{\exists\nabla}$
(2)	$\exists i\nabla Fa_i \leftrightarrow \forall i\,(\Box Fa_i \rightarrow \neg\Box\neg Fa_{i+1})$	$\underline{\mathrm{WC}} \leftrightarrow \underline{\exists\nabla}$ (= *10.1*)
(3)	$\nabla\exists_i\nabla Fa_i \leftrightarrow \nabla\forall_i\,(\Box Fa_i \rightarrow \neg\Box\neg Fa_{i+1})$	(2), **E** (= *4.7*)
(4)	$\nabla\forall_i\,(\Box Fa_i \rightarrow \neg\Box\neg Fa_{i+1})$	(1), (3), df$\leftrightarrow$, $\wedge$-elimination, **MP**

[31] Moreover, if one adds the Barcan formula to **FIN**, the converse of $\underline{\nabla\exists}$ also holds (See appendix of Bobzien 2015), and hence so does the converse of $\underline{\mathrm{WC}} \rightarrow \nabla\underline{\mathrm{WC}}$.

15 Epistemicist Agnosticism with a Non-arbitrary Sharp Cut-Off

The bivalence-preserving agnostic solution as presented so far does not entail that there is in a Sorites series with *F* a (and that is precisely one) sharp cut-off point between the true and the false cases of *F* (see above p.25). Hence it does not entail Williamson-style epistemicism (epistemicism$_w$).[32] One obtains such a sharp cut-off, if one adds MONOTONICITY$_F$ (*2.5*) as a third assumption. This results in an epistemicist agnosticism in which the borderline zone *appears* fairly chaotic (Sect. 3) but is in fact governed by a rule that entails a sharp cut-off between the true and the false cases. In this section I compare epistemicist agnosticism with epistemicism$_w$. Suppose for the sake of argument that we adopt as a fourth assumption Williamson's—semantic—explanation for why there is such an epistemically inaccessible non-arbitrary cut-off point (i.e., that the linguistic meaning of vague terms perpetually vacillates). Epistemic agnosticism can then be seen to replace Williamson's combination of his logic of clarity and margin-for-error theory by system **FIN** in its tell-ability interpretation, with its possible-world semantics given substance by the unsavoury assessment sensitivity of vague expressions.[33] Both these theories share—in addition to being classical, bivalent, epistemicist and having a sharp true/false cut-off point—the further elements that their solution introduces an ersatz conditional and that one cannot identify with certainty any specific borderline cases in any Sorites series and that in this sense there are no clear borderline cases.

The differences between epistemicist agnosticism and epistemicism$_w$ include the following. (1) In epistemicist agnosticism, the borderline zone in a Sorites series corresponds to the empirically determinable grey area of hedging behaviour and potential disagreement, with allowances made for factors like small handicaps of individual assessors, variants in linguistic usage, etc. (Sect. 3 above). In epistemicism$_w$, it is not obvious how the theory relates to the *phenomenon* of a grey area. (2) In epistemicist agnosticism, the polar cases are not borderline cases of any order. In epistemicism$_w$ the polar cases are borderline cases of some order. (3) Epistemicism$_w$ requires at least two borderline cases per Sorites series. (This follows from Williamson, 1994, 232.) Epistemicist agnosticism can do with—the possibility of—one. (4) Epistemicist agnosticism does not encounter the conjunction-agglomeration problem for non-vague sentences that epistemicism$_w$ faces (Williamson, 1999, 139–140): in epistemicism$_w$, the conjunction of two clear borderline cases A_1 and A_2 can imply the *n*th-order borderlineness for any *n* of their conjunction $A_1 \wedge A_2$. In system **FIN**, this cannot happen. (5) In the logic of epistemicist agnosticism non-borderlineness (precision) and borderlineness (vagueness) are closed under uniform substitution, whereas in epistemicism$_w$ this is not the case (Williamson, 1999, 132–133). So on balance epistemic agnosticism has the edge over Williamson-style epistemicism.

[32] By Williamson-style epistemicism I mean a theory of vagueness that maintains (i) that both classical logic and bivalence are preserved, (ii) that there is a (and that is precisely one) sharp cut-off point between the true and the false cases in Sorites series, and (iii) that it is epistemically inaccessible where this cut-off is situated.

[33] I do not deny that there is such a thing as a margin for error for knowledge. I suggest that it is not an essential part of a successful solution to the Sorites.

16 Bivalent Agnosticism with Arbitrary Sharp Cut-Offs and without Sharp Cut-Offs

To my mind, the assumption of a *non-arbitrary* cut-off between the true and the false cases in a Sorites series is too counterintuitive for a viable theory of vagueness. The agnostic solution also works with several alternatives. One preserves classical logic, bivalence and MONOTONICITY$_F$ (= *2.5*). One preserves classical logic, bivalence and surrenders MONOTONICITY$_F$. The third preserves MONOTONICITY$_F$ and classical logic and surrenders bivalence, a fourth preserves MONOTONICITY$_F$ and bivalence and surrenders classical logic. In this section, I set out the first two options.

If we remove the assumption of a (= precisely one) *non-arbitrary* cut-off but retain bivalence and MONOTONICITY$_F$, we obtain an agnosticism that preserves the standard penumbral connections. It resembles supervaluationism (e.g. Fine, 1975) in that it has the same set of possible truth-value distributions in the borderline zone. Unlike supervaluationism, it is not just classical but also preserves bivalence. Where supervaluationism has admissible truth-value distributions, non-epistemic bivalent agnosticism has actual truth-value distributions. In this second kind of agnosticism, there is in a Sorites series precisely one transition from one truth-value to the other, and this *semantic switch* occurs at or in the borderline zone. This is all one can say about the truth-value distribution in the borderline zones. Recall the philosophical theory sketched in Sect. 3: for any borderline case of *F* there are viewpoints from which it looks *F* and viewpoints from which it looks ¬*F* and viewpoints from which it looks non-borderline *F* and/or non-borderline ¬*F*, so that it is impossible to say of any borderline case that it is borderline. In this respect there is a homogeneity holding between the borderline cases that provides no reasons for the jump from true to false to occur at one point rather than another. This suggests that a plausible way to think of the actual truth-value distribution in the borderline zone is as being ***random***. The immediate reason why it is unpredictable what truth-value a borderline-case sentence may have in a context *c* is then that there is no recognizable pattern (beyond the stipulated MONOTONICITY$_F$) to the actual truth-values in the borderline zone. If MONOTONICITY$_F$ is considered indispensable, this seems to offer an accurate representation of how things are in the grey areas as far as we can possibly tell. The described randomness of the location of the semantic switch leaves both bivalence and classical logic in place. Since the location of the change from truth to falsehood in a semantic series is random here, the implausibility that some think comes with such borders in epistemicist theories is alleviated.

A more radical non-epistemicist agnosticism is one that rejects MONOTONICITY$_F$ and thereby a mandatory sharp cut-off. In this case, the fact that in the borderline zone speakers do not seem to uphold MONOTONICITY$_F$ (Sect. 3) is taken to echo the actual semantic situation in the borderline zone. Here, it is not the placement of the truth-value switch that is random, but the entire distribution of semantic values: the truth-value distribution in the borderline zone is ***chaotic***. This notwithstanding, both bivalence and classical logic are preserved. This variety of agnosticism, too, matches the philosophical theory sketched in Sect. 3: For any borderline case of *F* there are viewpoints from which it looks *F* and viewpoints from which it looks ¬*F*.

Nothing indicates any direction (from F to $\neg F$ or vice versa) or provides reasons for any particular truth-value distribution in the borderline zone. In this *chaotic non-epistemic agnosticism* without MONOTONICITY$_{\mathrm{F}}$ there is in a Sorites series at least one transition from one truth-value to the other, and there can be at least as many as there are borderline cases in the sequence—for n borderline cases n transitions if n is odd, $n+1$ if n is even. This set-up would be apt for psychological contextualism (e.g. Raffman, 1994). Since there can be a plurality of cases in which a true sentence of the series is followed by a false one, the implausibility that comes with such borders in epistemicist theories is removed.

17 Outlook: Generalizing Revenge-Proof Agnosticism to Non-bivalent and Non-classical Theories of Vagueness

The theories of vagueness considered in the last two sections contain classical normal modal logic and bivalence. However, the agnostic solution offered has more general relevance. What I have in mind is a generalization from classical normal modal logic to classical systems with a three-valued semantics and to systems without *tertium non datur*. In this way the scope of the solution offered is widened to theories that contain two truth-values—e.g. definitely true and definitely false—and allow for a third semantic status that is not itself a truth-value but some kind of indeterminacy with respect to truth-value. There are again several cases.

Use of classical logic without bivalence has been suggested by Ian Rumfitt to formally capture Mark Sainsbury's *polar theory of vagueness* (Rumfitt, 2015; Sainsbury, 1996). In essence, the boundary between, say, blue and not blue colour patches in a Sorites series is blurred, since although 'blue or not blue' is satisfied by all colour patches of the series, the fact that the semantic value of sentences expressing borderline cases is 'Indeterminate' prevents us from inferring where that boundary between the blue and the not blue cases is located. In order to avoid higher-order vagueness paradoxes, this kind of theory requires that, where the semantic value of a sentence is 'Indeterminate', a sentence expressing this has itself the value 'Indeterminate'; and that a sentence that expresses that there exist certain sentences with the value 'Indeterminate' itself has the value 'Indeterminate'. (This is *not* stated in Rumfitt, 2015.) How such a theory relates structurally to the agnostic theory with bivalence intact then becomes apparent. The **FIN** modal triad $\Box F$, ∇F, $\Box\neg F$ corresponds to Rumfitt's vague semantic triad True, Indeterminate and False. To Rumfitt's classical *tertium* correspond the *tertium* plus bivalence. (This preserves the T-schema for vague sentences.) The fourfold distinction of Rumfitt's pairs F & True(F), F & Indeterminate(F), $\neg F$ & Indeterminate($\neg F$) and $\neg F$ & False(F) is matched by $F\wedge\neg\nabla F$, $F\wedge\nabla F$, $\neg F\wedge\nabla F$ and $\neg F\wedge\neg\nabla F\neg F$. (Mormann, 2020 shows that at the sentential level Rumfitt's topological semantics entails higher-order vagueness that is columnar.)

The relation of revenge-free agnosticism to theories of vagueness without the *tertium*, in particular intuitionistic ones, is *prima facie* different than those discussed so far. It is grounded in the fact that **S4M** is a *modal companion* of intuitionistic logic. I briefly mention how some core ideas of intuitionist theory of vagueness are preserved in the generic solution to the Sorites put forward here. Crispin Wright's agnostic intuitionist theory will serve as the paradigm theory (Wright, 2007, 2019). In it, MONOTONICITY$_{\mathrm{F}}$ is retained. Via the McKinsey–Tarski translation **S4M** translates into the non-modal *sentential* logic of intuitionistic theories of vagueness. So, the sentential part of the intuitionistic theory is preserved in the proposed solution; and on account of axiom **M** it is preserved without higher-order vagueness paradoxes. (For details see Bobzien & Rumfitt, 2020, Sects. 1–4.) The first-order modal logic **FIN** however is not a modal companion of intuitionistic logic, and so the agnostic solution as a whole is not intuitionistic. It accompanies a logic stronger than intuitionistic logic, but weaker than classical logic. System **FIN** is a modal companion of the intermediate logic **QH+KF**, where **QH** stands for first-order intuitionistic logic and **KF** for the axiom $\neg\neg\forall x\ (Fx \vee \neg Fx)$. (The logic **QH+KF** is discussed in Gabbay, Skvortsov, and Shehtman (2009) 138, 157-8, 515-6.) Axiom **KF** ensures that the negation of $\forall x\ (Fx \vee \neg Fx)$ is not provable. So **QH+KF** 'hovers' between intuitionistic and classical logic similar to the way in which the interpreted **FIN** 'hovers' between the existence and non-existence of borderline cases. One might say that as **FIN** is the first-order modal logic of vagueness, so **QH+KF** is the first-order *non-modal* logic of vagueness.[34]

In conclusion of *Part IV*, it can be said that the finality logic, that is, the normal modal logic **FIN**, can serve as the basis for a considerable variety of revenge-free agnostic solutions to the Sorites paradox, ranging from epistemicist approaches, via supervaluationist-tinted agnosticism, 'chaotic' or psychological-contextualist approaches and polar theories, to approaches based on core ideas of intuitionistic theories of vagueness. **FIN** seems to be a deserving candidate for the title *first-order modal logic of vagueness*. Moreover, since **FIN** is a modal companion of the intermediate logic **QH+KF**, indirectly it also introduces a non-modal logic of vagueness.

Appendix

(i) Proof sketch that in system QTM, FIN(x) entails $\nabla\exists(x)$

(1)	$\neg\Box\neg\forall x(Fx \rightarrow \Box Fx)$	**FIN(*x*)**
(2)	$\forall x(Fx \rightarrow \Box Fx) \rightarrow (\exists xFx \rightarrow \exists x\Box Fx)$	theorem of **QTM**
(3)	$\neg\Box\neg\forall x(Fx \rightarrow \Box Fx) \rightarrow \neg\Box\neg(\exists xFx \rightarrow \exists x\Box Fx)$	(2) **DR3**, i.e. if $\vdash A \rightarrow B$, $\vdash \neg\Box\neg A \rightarrow \neg\Box\neg B$
(4)	$\neg\Box\neg(\exists xFx \rightarrow \exists x\Box Fx)$	(1), (3) **MP**
(5)	$\Box\exists xFx \rightarrow \neg\Box\neg\exists x\Box Fx$	(4) **K7**, i.e. $\neg\Box\neg(A \rightarrow B) \leftrightarrow (\Box A \rightarrow \neg\Box\neg B)$

[34] Another variant of a revenge-free agnosticism that preserves some elements of intuitionistic theories of vagueness would go deeper and subordinate the modal expressions ∇ and $\Box$ themselves to the **KF**-intermediate logic, something outside the scope of this paper. Here I only note that at the sentential level, the principle that borderline cases are themselves borderline appears to be preserved, because of **S4M**'s role as a modal companion of the intuitionistic sentential calculus. For details of this latter point see Bobzien and Rumfitt (2020, §5). Please note that the present paper was written *before* Bobzien and Rumfitt (2020).

(6)	$\Box\exists x\nabla Fx\rightarrow\neg\Box\neg\exists x\Box\nabla Fx$	(5) $\nabla F/F$ (substitution into modal context)
(7)	$\neg\neg\Box\neg\exists x\Box\nabla Fx\rightarrow\neg\Box\exists x\nabla Fx$	(6) **PC** (contraposition)
(8)	$\Box\neg\exists x\Box\nabla Fx\rightarrow\neg\Box\exists x\nabla Fx$	(7) **PC** (double negation introduction)
(9)	$\forall x((\Box Fx\vee\Box\neg Fx)\vee\neg(\Box Fx\vee\Box\neg Fx))$	tertium, **∀2** (theorem of **QTM**)
(10)	$\forall x((\Box Fx\vee\Box\neg Fx)\vee(\neg\Box Fx\wedge\neg\Box\neg Fx))$	(9) DeMorgan (theorem of **QTM**)
(11)	$\forall x((\Box Fx\vee\Box\neg Fx)\vee\nabla Fx)$	(10) df ∇
(12)	$\nabla^2 A\leftrightarrow(\neg\Box\nabla A\wedge\neg\Box\neg\nabla A)$	df ∇ (on ∇A)
(13)	$\nabla^2 A\rightarrow(\neg\Box\nabla A\wedge\neg\Box\neg\nabla A)$	(12) df $\leftrightarrow$, $\wedge$-elimination
(14)	$\nabla^2 A\rightarrow\neg\Box\nabla A$	(13) **PC** (conditional $\wedge$-elimination)
(15)	$\forall x(\nabla^2 Fx\rightarrow\neg\Box\nabla Fx)$	(14) **∀2** (theorem of **QTM** with ∇ operator)
(16)	$\forall x(\nabla Fx\rightarrow\nabla^2 Fx)$	theorem V, **∀2** (theorem of **QTM**)
(17)	$\forall x(\nabla Fx\rightarrow\neg\Box\nabla Fx)$	(16), (15) (transitivity of $\rightarrow$)
(18)	$\neg\exists x\neg(\nabla Fx\rightarrow\neg\Box\nabla Fx)$	(17) df $\exists$
(19)	$\neg\exists x(\nabla Fx\wedge\Box\nabla Fx)$	(18) (df $\rightarrow$, double negation elimination)
(20)	$\nabla A\rightarrow(\neg\Box A\wedge\neg\Box\neg A)$	df ∇, left-to-right (**PC**)
(21)	$\nabla A\rightarrow\neg\Box A$	(20) **PC** (conditional $\wedge$-elimination)
(22)	$\Box\nabla A\rightarrow\nabla A$	axiom **T**
(23)	$\Box\nabla A\rightarrow\neg\Box A$	(21), (22) **PC** (transitivity of $\rightarrow$)
(24)	$\neg(\Box A\wedge\Box\nabla A)$	(23) df $\rightarrow$, commutativity of $\wedge$
(25)	$\forall x\neg(\Box Fx\wedge\Box\nabla Fx)$	(24) **∀2** (theorem of **QTM**)
(26)	$\neg\exists x\neg\neg(\Box Fx\wedge\Box\nabla Fx)$	(25) df $\exists$
(27)	$\neg\exists x(\Box Fx\wedge\Box\nabla Fx)$	(26) double negation introduction (**QC**)
(28)	$\nabla A\rightarrow\neg\Box\neg A$	(20) **PC** (conditional $\wedge$-elimination)
(29)	$\Box\nabla A\rightarrow\neg\Box\neg A$	(22), (28) **PC** (transitivity of $\rightarrow$)
(30)	$\neg(\Box\neg A\wedge\Box\nabla A)$	(29) df $\rightarrow$, commutativity of $\wedge$
(31)	$\forall x\neg(\Box\neg Fx\wedge\Box\nabla Fx)$	(30) **∀2** (theorem of **QTM**)
(32)	$\neg\exists x\neg\neg(\Box\neg Fx\wedge\Box\nabla Fx)$	(31) df $\exists$
(33)	$\neg\exists x(\Box\neg Fx\wedge\Box\nabla Fx)$	(32) double negation introduction (**QC**)
(34)	$\neg\exists x\Box\nabla Fx$	(11), (19), (27), (33) **QC**
(35)	$\Box\neg\exists x\Box\nabla Fx$	(34) rule **N** (necessitation)
(36)	$\neg\Box\exists x\nabla Fx$	(8), (35) **MP**
(37)	$\Box\neg\exists x\nabla Fx\rightarrow\neg\exists x\nabla Fx$	axiom **T, ∀2** (theorem of **QTM**)
(38)	$\exists x\nabla Fx\rightarrow\neg\Box\neg\exists x\nabla Fx$	(37) **PC** (contraposition) **∀2**
(39)	$\exists x\nabla Fx\rightarrow\neg\Box\exists x\nabla Fx$	(36) **PC**
(40)	$\exists x\nabla Fx\rightarrow(\neg\Box\neg\exists x\nabla Fx\wedge\neg\Box\exists x\nabla Fx)$	(38), (39) **PC** (conditional $\wedge$-introduction)
(41)	$\exists x\nabla Fx\rightarrow\nabla\exists x\nabla Fx$	(40) **PC**, df ∇

The formula in line (41) is **V∃(*x*)**.

(ii) Proof sketch that in QT V∃(*x*) entails FIN(*x*)

(1)	$\exists x\nabla Fx\rightarrow\nabla\exists x\nabla Fx$	theorem V**∃**(***x***)
(2)	$\exists x\nabla Fx\rightarrow\neg\Box\neg\nabla\exists x\nabla Fx\wedge\neg\Box\exists x\nabla Fx$	(1) df ∇
(3)	$\exists x\nabla Fx\rightarrow\neg\Box\exists x\nabla Fx$	(2) conditional $\wedge$-elimination (**PC**)
(4)	$\Box\exists x\nabla Fx\rightarrow\exists x\nabla Fx$	axiom **T ∀2**

(5)	$\neg\exists x\nabla Fx \to \neg\Box\exists x\nabla Fx$	(4) contraposition (**PC**) **∀2**
(6)	$\exists x\nabla Fx \lor \neg\exists x\nabla Fx$	tertium (**PC**) **∀2**
(7)	$\neg\Box\exists x\nabla Fx \lor \neg\Box\exists x\nabla Fx$	(3), (5), (6) constructive dilemma (**PC**) **∀2**
(8)	$\neg\Box\exists x\nabla Fx$	(7) **PC ∀2**
(9)	$\neg\Box\exists x(\neg\Box\neg Fx \land \neg\Box Fx)$	(8) df ∇
(10)	$\neg\Box\exists x\neg(\neg\Box\neg Fx \to \Box Fx)$	(9) df $\to$
(11)	$\neg\Box\neg\forall x\neg\neg(\neg\Box\neg Fx \to \Box Fx)$	(10) df $\exists$
(12)	$\neg\Box\neg\forall x(\neg\Box\neg Fx \to \Box Fx)$	(11) double negation elimination
(13)	$\Box A \to A$	axiom **T**
(14)	$\Box\neg A \to \neg A$	(13) $\neg A/A$
(15)	$\neg\Box\neg A \to A$	(14) **PC** (contraposition)
(16)	$(\neg\Box\neg A \to B) \to (A \to B))$	(15) **PC**
(17)	$(\neg\Box\neg A \to \Box A) \to (A \to \Box A))$	(16) $\Box A/B$
(18)	$\forall x((\neg\Box\neg Fx \to \Box Fx) \to (Fx \to \Box Fx))$	(17) **∀2**
(19)	$\forall x(\neg\Box\neg Fx \to \Box Fx) \to \forall x(Fx \to \Box Fx)$	(18) **QC** (∀-distribution)
(20)	$\neg\Box\neg\forall x(\neg\Box\neg Fx \to \Box Fx) \to \neg\Box\neg\forall x(Fx \to \Box Fx)$	(19) **DR3**, i.e. if $\vdash A \to B$, $\vdash \neg\Box\neg A \to \neg\Box\neg B$
(21)	$\neg\Box\neg\forall x(Fx \to \Box Fx)$	(12), (20) **MP**

The formula in line (21) is **FIN(*x*)**. Since nothing hinges on the fact that the proof sketches (i) and (ii) use a one-place predicate, we can generalize, that in **QTM**, and hence in **QS4M**, $\exists x_1, \ldots, \exists x_n\nabla A \to \nabla\exists x_1, \ldots, \exists x_n\nabla A$ is logically equivalent to $\neg\Box\neg\forall x_1, \ldots, \forall x_n(A \to \Box A)$.

Acknowledgements Acknowledgements: I thank Øystein Linnebo, Elia Zardini, Ian Rumfitt, Crispin Wright, Timothy Williamson, Michael Beaney, Jennifer Hornsby, Peter Fritz, Guillaume Massas, Gianluca Griletti, Wim Vanrie, as well as the audiences at my talks on how to solve the Sorites with normal modal logic in Lisbon, Oxford and York and at my 2021 Townsend Lectures on vagueness at Berkeley. I am particularly grateful to John MacFarlane and Wesley Holliday for teaching this paper at the preparatory graduate seminar for the Townsend Lectures and sharing most helpful comments with me that arose from that class.

References

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

Bobzien, S. (2010). Higher-order vagueness, radical unclarity and absolute agnosticism. *Philosophers' Imprint*, *10*(10), 1–30.

Bobzien, S. (2012). If it's Clear, then it's Clear that it's Clear, or is it? – Higher-order Vagueness and the S4 Axiom. In: K. Ierodiakonou & B. Morison (Eds.), *Episteme, etc.* (pp. 189-212). Oxford: Oxford University Press.

Bobzien, S. (2013). Higher-Order Vagueness and Borderline Nestings: A Persistent Confusion. *Analytic Philosophy*, *54*(1), 1–43.

Bobzien, S. (2015). Columnar higher-order vagueness or Vagueness is higher-order vagueness. *Proceedings of the Aristotelian Society, Supplementary Volume*, *89*, 61-87.

Bobzien, S., & Rumfitt, I. (2020). Intuitionism and the modal logic of vagueness. *Journal of Philosophical Logic, 49*(2), 221–248.

Bonnay, D., & Égré, P. (2009). Inexact knowledge with introspection. *Journal of Philosophical Logic, 38*(2), 179–227.

Boolos, G. (1993). *The logic of provability*. Cambridge University Press.

Burgess, J. P. (1999). Which modal logic is the right one? *Notre Dame Journal of Formal Logic, 40*(1), 81–93.

Cresswell, M. (2001). How to complete some modal predicate logics. In M. Zakharyaschev, K. Segerberg, M. de Rijke, & H. Wansing (Eds.), *Advances in modal logic* (Vol. 2, pp. 155–178). CSLI Publications.

Esakia, L. L. (1976). On modal companions of superintuitionistic logics. In *VII Soviet symposium on logic*. Kiev **(in Russian)**.

Fine, K. (1975). Vagueness, truth and logic. *Synthese, 30*, 265–300. Reprinted with corrections in R. Keefe & P. Smith (Eds.) (1996). *Vagueness: A reader* (pp. 119–50). Cambridge: MIT Press.

Gaifman, H. (2010). Vagueness, tolerance and contextual logic. *Synthese, 174*(1), 5–46.

Grzegorczyk, A. (1967). Some relational systems and the associated topological spaces. *Fundamenta Mathematicae, 60*, 223–231.

Halldèn, S. (1963). A pragmatic approach to modal theory. *Acta Philosophica Fennica, 16*, 53–64.

Hughes, G. E., & Cresswell, M. J. (1996). *A new introduction to modal logic*. Routledge.

Kuhn, S. T. (1995). Minimal non-contingency logic. *Notre Dame Journal of Formal Logic, 36*, 230–234.

Montgomery, H., & Routley, R. (1966). Contingency and non-contingency bases for normal modal logics. *Logique Et Analyse, 9*, 318–328.

Mormann, T. (2020). Topological models of columnar vagueness. *Erkenntnis*, *86*, 693–716. https://doi.org/10.1007/s10670-019-00214-2

Raffman, D. (1994). Vagueness without paradox. *Philosophical Review, 103*, 41–74.

Raffman, D. (2005). Borderline cases and bivalence. *Philosophical Review, 114*(1), 1–31.

Raffman, D. (2010). Demoting higher-order vagueness. In R. Dietz & S. Moruzzi (Eds.), *Cuts and clouds: Vagueness, its nature, and its logic* (pp. 509–522). Oxford: Oxford University Press.

Raffman, D. (2014). *Unruly words: A study of vague language*. Oxford University Press.

Rumfitt, I. (2015). *The boundary stones of logic*. Oxford University Press.

Sainsbury, R. M. (1991). Is there higher-order vagueness? *The Philosophical Quarterly, 41*, 167–182.

Sainsbury, R. M. (1996). Concepts without boundaries. In R. Keefe & P. Smith (Eds.), *Vagueness: A reader* (pp. 186–205). Harvard University Press.

Segerberg, K. (1982). *Classical propositional operators*. Clarendon Press.

Shapiro, S. (2005). Context, conversation, and so-called "higher-order" vagueness. *Proceedings of the Aristotelian Society, Supplementary, 105*, 147–165.

Williamson, T. (1994). *Vagueness*. Routledge.

Williamson, T. (1999). On the structure of higher-order vagueness. *Mind, 108*, 127–142.

Wolter, F., & Zakharyaschev, M. (2014). On the Blok-Esakia theorem. In G. Bezhanishvili (Ed.), *Leo Esakia on duality in modal and intuitionist logics, outstanding contributions to logic* (Vol. 4, pp. 99–118). Springer.

Wright, C. (1992). Is higher-order vagueness coherent? *Analysis, 52*, 129–139.

Wright, C. (2001). On being in a quandary: Relativism, vagueness, logical revisionism. *Mind, 110*, 45–98.

Wright, C. (2003). Vagueness: A fifth column approach. In J. C. Beall (Ed.), *Liars and heaps: New essays on paradox* (pp. 84–105). Oxford University Press.

Wright, C. (2007). Wang's paradox. In R. E. Auxier & L. E. Hahn (Eds.), *The library of living philosophers: Michael Dummett* (pp. 415–444). Open Court.

Wright, C. (2010). The illusion of higher-order vagueness. In R. Dietz & S. Moruzzi (Eds.), *Cuts and clouds: Vagueness, its nature, and its logic* (pp. 523–549). Oxford: Oxford University Press.

Wright, C. (2019). Intuitionism and the Sorites paradox. In S. Oms & E. Zardini (Eds.), *The Sorites paradox.* Cambridge University Press.

Zardini, E. (2006). Higher-order vagueness and paradox: The glory and misery of S4 definiteness. *The Baltic International Yearbook of Cognition, Logic and Communication, 1*, 203–220.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.