Three-Valued Plurivaluationism of Vague Predicates

Wen-fang Wang, wfwang@ym.edu.tw
Institute of Philosophy of Mind and Cognition
National Yang Ming University, Taipei, Taiwan

Abstract. Disagreeing with Williamson on vagueness, the author proposes a solution that he calls "three-valued plurivaluationism" to the age-old sorites paradox. In essence, it is a three-valued semantics for a first-order language with identity with the additional suggestion that a vague language has more than one correct interpretation. Unlike the traditional three-valued approach to a vague language, the so-called three-valued plurivaluationism, so the author argues, can accommodate the phenomenon of higher-order vagueness. And, unlike the tr`aditional three-valued approach to a vague language, the so-called three-valued purivaluationism, so the author argues, can also accommodate the phenomenon of penumbral connection when equipped with "suitable conditionals". The author also shows that this three-valued purivaluationism is a natural consequence of a restricted form of Tolerance Principle (T_R) and a few related ideas, and argues that (T_R) is well-motivated by considerations of how we learn, teach, and use vague predicates.

Keywords: vagueness, sorites paradox, tolerance principle, three-valued semantics, plurivationism, conditionals.

1. Vague Predicates and the Sorites Paradox

A vague predicate is a predicate that has *possible* borderline cases, i.e., possible cases such that it is semantically indeterminate whether or not the predicate applies. Examples of vague predicates abound in natural languages. Here is just a short list of examples in English: "bald", "heap", "tall", "red", "table", "a small portion of C" (where C is a class of, say, 50 students), "similar to", "identical with", and so on.

A problem about vague predicates is that they give rise to the sorites paradox. Let "a_i" be a name of someone with i hairs. Then from the apparently plausible premises "a_o is bald" and "if a_n is bald, so is a_{n+1}, for whatever number n", one can infer the absurd conclusion that "$a_{100,000}$ is bald". Or, to take another example from [18], let b_0 be me, and suppose that there are n molecules in my body. Let b_0, b_1, ..., b_n be a sequence of objects each of which is obtained from its predecessor by replacing one molecule of me with a molecule of scrambled egg, so that b_n is all scrambled egg. Let β_i be the statement that "$b_{i-1} = b_i$". Then each β_i, where $1 \leq i \leq n$, seems to be true. Yet, by n applications of the rule of transitivity of identity, we reach the absurd conclusion that I am scrambled egg. From the fact that one can "prove" almost everything s/he wants to prove by a sorties argument, of course we should conclude that some of these sorites arguments must be unsound. However, it has been proved very difficult both to pinpoint the problem(s) of these arguments and to give a plausible explanation of why we are taken in.

In the past 40 years, philosophers have witnessed a bunch of theories aiming at solving the sorites paradox.[1] A benefit/cost analysis of even a small portion of these theories will be an impossible task for a short paper like the present one. This paper suggests that we start from scratch to re-think about how we learn, teach, and use vague predicates and hopes that we will gain some insights from such an inspection.

2. Start from Scratch

Before we start, however, let me give a few preliminary comments. In the very beginning of this paper, I defined "a vague predicate" to be a predicate "having *possible* borderline cases", but why possible cases? Why not define a vague predicate in terms of its actual borderline cases? Here is the reason. If we call a predicate "vague" only when it *actually* has some borderline cases, then some predicates that are intuitively vague will not be "vague" in the defined sense, and this seems undesirable. For example, if we define an F-snail to be a snail that walks much faster than most slow turtles, then intuitively this notion of *F-snails* is a vague one so long as notions of *much faster than*, *most*, and *slow* are. Yet, surely nothing in the world is an F-snail (or, if this is not the case, replace the word "turtle" by "panther"), so there is no actual borderline case for this intuitively vague predicate. As a result, the notion of *F-snails* turns out not to be "vague" in the new, defined sense, and this seems undesirable. Here is another example. If we define a baldsome male to be a male who is both very bald and very handsome, then, again, intuitively this notion of *baldsome males* is a vague one so long as both the notion of *very bald* and that of *very handsome* are. However, it may happen that there are a few men that are clearly both very bald and clearly very handsome while all others are either clearly not-very-bald (though some of them may be vaguely very handsome) or clearly not-very-handsome (though some of them may be vaguely very bald), so there will be, in this case, no actual borderline case for this intuitively vague predicate.[2] As a result, the notion of *baldsome males* turns out not to be "vague" in the new, defined sense either, and this seems equally undesirable. So, if we want to characterize vague predicates as predicates having borderline cases, it seems better that we take into account all possible cases as well as all actual ones of these predicates. Or, I may even put things in this way: what I will call "the extension" of a vague predicate in this paper may be called by other philosophers its "intension", but I don't think that there will be anything important that hinges on this difference.

There are different kinds of vague predicates; especially, there are primitive ones as well as defined ones, and, among each category, there are perceptual ones as well as non-perceptual ones. So which ones will I be talking about in this paper? I intend the semantics proposed in sections 3 and 4 to be applicable to vague predicates in general, but I will restrict my discussion in this section to primitive ones to make my exposition simpler. As examples of primitive vague predicates, I take "red", "bald", "soft" (these are

---

1 To name just a few: epistemicism proposed by [3], [2], [28], and [23], gap theories proposed by [8] and [13], glut theories proposed by [18] and [10], supervaluationism proposed by [6], [14], [11], [4], [19], [1], and [12], fuzzy theories proposed by [17] and [22], plurivaluationism proposed by [27] and [16], and contextualist theories proposed by [25], [20], [5], and [21].

2 On the other hand, it is not clear that predicates like "bald but not self-identical", "bald or self-identical", and "tall or greater-than-or-equal-to exactly four feet in height" are vague predicates, for it is impossible for these predicates to have borderline cases.

vague perceptual predicates), "a small portion of C" (where C is a class of 50 students), and "identity" (these are non-perceptual ones). Examples of non-primitive vague predicates include, on the other hand, "4-tall" (exactly 4 feet in height or is tall), "baldsome" (very bald but very handsome), and "F-snails" (snails that walks much faster than most slow turtles). If you think that some of these examples are wrongly classified, be my guest to adjust the classification by yourself. Again, I do not think that there will be anything important that hinges on the choice.

The notion of *F-relevant respects* of a vague predicate F will play an important role in what follows, so we'd better get a good grip of it now. Let me begin with the notion of *determination*. A set of respects (properties or relations) R_1, ..., R_n determines a certain respect R (property or relation) iff, for all possible objects (or sequences of possible objects) α and β, it is necessarily the case that if α and β are exactly the same with respect to R_1, ..., R_n then they are exactly the same with respect to R (and so the two sentences "*Ra*" and "*Rb*" will have the same truth value (either both are true, or both are false, or both are neither), where "*a*" and "*b*" are names of α and β (or are sequences of names of objects in α and β)). It can easily be proved that if a set S of respects determines a respect R, so does any superset S' of S, so the notion of determination is not a very useful one. For the purpose of defining "F-relevant respects" of a vague predicate F, we need a tighter notion than that of *determination*: *m-determination* (short for "minimal determination"). A set S of respects (properties or relations) R_1, ..., R_n m-determines a certain respect R (property or relation) iff (a) S determines R, and (b) for any set $T = \{R'_1, ..., R'_m\}$ that also determines R, the set T "entails" the set S in the sense that, for any possible object (or sequence of possible objects) α, it is necessarily true that if α has every respect in T then it also has every respect in S. An F-relevant respect of a vague predicate F is then a respect in any set S that m-determines whether something is F. Given this definition of a relevant respect of a vague predicate, it is easy to see that even a primitive vague predicate, such as "red", may have multiple relevant respects, such as hue, value, and chroma. A semantically primitive vague predicate may therefore stand for an ontologically complex property, i.e., a property whose existence ontologically depends on the existence of several simpler properties.

With these preliminaries in mind, we are now in a position to investigate how we learn, teach, and use vague predicates. In general, I believe that the following story is a rough but faithful picture about how we learn and teach vague predicates. When learning or teaching how to use a primitive vague predicate F, we do so by means of ostension (what else can we do?), i.e., by giving or by being given examples or "paradigms", both positive and negative ones, of F. Moreover, some paradigms are introduced explicitly by pointing to them or by showing them, while others are introduced implicitly by hints or by implicatures. For examples, we may point to a few heads and call them "bald", at the same time implicitly implying or implicating that heads with fewer hairs or heads whose numbers of hairs are between those of the paradigms are also paradigms of bald heads. Another example: we may illustrate the use of "a small portion of C", where C is a class of 50 students, by saying loud that "a subset of C with 5 or less members is a small portion of C", at the same time implicitly implying or implicating that a set of C with 45 or more members are negative paradigms of the predicate. Because we all teach and learn a primitive vague predicate F in this standard ostensive way, each competent speaker of F, i.e., one who understands how to use F correctly, will have both some positive paradigms and some negative paradigms of F in his or her mind. Moreover, it seems that nothing in the process of teaching and learning F can be both a positive and a negative paradigm of F on pain of confusion.

However, in order for the teaching and learning process of a vague predicate F to be successful, the difference in F-relevant respects between any positive paradigm and any negative paradigm of F must be "salient" to the learner. Otherwise, it is hard to imagine how the learner can even re-identify a positive (or negative) paradigm of F as a positive (or negative) paradigm of F again, let alone has an idea about how to make further applications of the predicate F. We say that two paradigms of F differ saliently in F-relevant respects to a subject S in an occasion O iff the overall dissimilarity between them in F-relevant respects is easily observable for S in O or is intellectually significant for S in O. The requirement that it is easily observable for S in O is tailored especially for vague perceptual predicates, such as "red", so that, according to this requirement, the overall difference in red-relevant respects between a paradigm red patch and a paradigm not-red one must be easily observable to the learner when the predicate is learned. The requirement that it is intellectually significant for S in O, on the other hand, is tailored especially for non-perceptual vague predicates, such as "is a small portion of C", so that, according to this requirement, the overall difference in a-small-portion-of-C-relevant respects between, say, a 5-membered subset of C and a 45-membered subset of C must be intellectually significant to the learner, and presumably the intellectual significance in this case may simply consist in the fact that the difference between the ratios of the two subsets to C is close to 1 or at least much greater than a half.

So far, there is no guarantee that two competent speakers of a vague predicate F will have any common positive paradigm or any common negative paradigm in their minds, and this seems to make the publicity of a vague language problematic. Fortunately, because people have roughly, though not exactly, the same perceptual and intellectual capacities, and also because many positive and negative paradigms of a vague predicate F are implicitly introduced when teaching or learning F, all competent speakers of F ultimately share at least some common paradigms, both positive and negative ones, of F in their minds. This is not to deny that the perceptual and intellectual capacities that one has differ from person to person and from occasion to occasion. But this fact should not lead us to overlook the equally important fact that our perceptual and intellectual capacities are very similar after all. (Another important fact about our perceptual and intellectual capacities is this: we are all limited creatures; our abilities of discernment and our intellectual swiftness and astuteness are all very limited, so that, for example, no one can really discriminate a large number of "border-line shades of colors" between a positive paradigm and a negative paradigm of redness. I will not emphasize this important fact here, but I will come back to it when I consider the problem of "gradual transition" in next section.) Thus, if it makes sense at all to assign an extension F^+ and an anti-extension F^- to a vague predicate F, at least these common positive paradigms of F should be included in the extension F^+, and at least these common negative paradigms of F should be included in the anti-extension of F^-. And, from what we have said two paragraphs ago, it is also reasonable to assume that these two extensions of a vague predicate are mutually exclusive.

As I see it, the most distinguished feature of any vague predicate, in contrast with a precise predicate, is the existence of a "sorites sequence" for the predicate: for any occasion O and any two paradigms a_1 and a_n of a vague predicate F, and for any competent speaker S of F, there always is a sequence of possible cases $<a_1, ..., a_n>$ between a_1 and a_n such that any two adjacent cases in the sequence are "very similar" to S in F-relevant respects in O in the sense that the overall dissimilarity in F-relevant respects between them is not observable or is intellectually insignificant for S in O.[3] Now, a vague predicate F must allow its competent users to be able to apply and re-apply it, not only to those positive and negative paradigms that are introduced in the learning process, but also to possible cases beyond these paradigms (this is also true of most precise predicates), otherwise, it will not be a vague predicate at all but belongs to a very special kind of precise predicates. (Consider Fine's example: a number is an F if it is smaller than or equal to 13 and is not an F if it is greater than or equal to 15. Defined in this way, this predicate F will have no further possible application beyond those paradigms that are introduced in this definition, but it will not be regarded as a vague predicate by most philosophers either.) For precise predicates, the possibility of further applications is given by their definitions. But this is obviously not the case for primitive vague predicates. So, by what rule (or rules) does a competent speaker of a vague predicate F extend its use to cases other than those introduced in the learning process? To this question, I suggest[4] the following answer: every competent speaker S of a vague predicate F tacitly accepts the following "restricted tolerance principle" (T_R):

(T_R): If it is correct for a subject S to classify x as a member of F^+ (or F^-) in an occasion O and y and x are "very similar" for S in O, then it is also correct for S to classify y as a member of F^+ (or F^-) in O, so long as, after so classified, the difference in F-relevant respects between any member of F^+ and any member of F^- remains observationally or intellectually salient for S in O.

In short, I believe that the following statements 1-7 jointly constitute a roughly true story about how we learn, teach and use primitive vague predicates:

---

3 This feature does not seem to me to be owned by any precise predicate, perhaps because the F-relevant respects of a precise predicate F are just those respects specified in the definition of F, so everything falling within F differs saliently, observationally or intellectually, in F-relevant respects from everything falling out of F. (Consider the case of the precise predicate "is an even number".) As a result, even if one can find a sequence of possible cases $<a_1, ..., a_n>$ between a positive paradigm a_1 and a negative one a_n such that any two adjacent cases in it are very similar in some respects, there still will be two adjacent cases in the sequence that differ saliently in F-relevant respects.

4 I do not just suggest (T_R), but also think that it is supported by at least three arguments. First, not only is (T_R) true of vague predicates, it is also true of precise predicates if we interpret "very similar" in it as "having or lacking the same defining properties". So (T_R) seems to be a principle for predicates in general. Second, (T_R) is a logically weaker principle than Wright's tolerance principle (T): If it is correct for S to classify x as a member of F^+ (or F^-) in O and y and x are "verysimilar" in F-relevant respects, then it is also correct to classify y as a member of F^+ (or F^-) in O. So evidences for (T) are automatically evidences for (T_R), and [30] did provide a few good evidences for (T). Finally, I believe that (T_R) can better explain, while (T) cannot, the phenomenon that is found in the "forced march sorites paradox" in [9], but I will leave the justification of this explanatory power of (T_R) to another paper due to its complicated nature.

1. We learn and teach how to use a vague predicate F by ostension. Some paradigms of F are introduced explicitly by pointing to them or by showing them, while others are introduced implicitly by hints or by implicatures.
2. Due to the way we learn and teach a vague predicate F, each competent speaker of F will have in mind some positive paradigms and some negative paradigms of F.
3. For any competent speaker of F, the difference in F-relevant respects between any positive paradigm and any negative paradigm of F must be either perceptually or intellectually salient.
4. We have roughly, though not exactly, the same perceptual and intellectual capacities.
5. Due to facts 1 and 4, all competent speakers of a vague predicate F share at least some common positive paradigms that belong to the extension F^+ of F and some common negative paradigms that belong to the anti-extension F^- of F.
6. For any occasion O and any two paradigms a_1 and a_n of a vague predicate F, and for any competent speaker S of F, there always is a sequence of possible cases $<a_1, ..., a_n>$ between a_1 and a_n such that any two adjacent cases in the sequence are "very similar" to S in F-relevant respects in O in the sense that the overall dissimilarity in F-relevant respects between them is not observable or is intellectually insignificant for S in O.
7. Every competent speaker S of a vague predicate F tacitly accepts the restricted tolerance principle (T_R): If it is correct for a subject S to classify x as a member of F^+ (or F^-) in an occasion O and y and x are "very similar" for S in O, then it is also correct for S to classify y as a member of F^+ (or F^-) in O, so long as, after so classified, the difference in F-relevant respects between any member of F^+ and any member of F^- remains observationally or intellectually salient for S at O.

However, if statements 1-7 are correct, then it follows that:

8. For any vague predicate F, $F^+ \cup F^-$ is not equal to the set of everything that the predicate F can meaningfully apply, for any competent speaker S of F.
9. Due to 4 and 7, the extension F^+ and the anti-extension F^- of F may be different for different competent speakers of F, though there is a common "core" for all competent speakers.
10. Although 9, so long as one's assignment of F^+ and F^- to F obeys (T_R) and some other "natural restrictions", his or her interpretation of F is correct.
11. Due to 8, a correct interpretation of a vague language *L* must be a three-valued interpretation and there seems to be no reason for having more than three values. Due to 10, there can be more than one correct interpretation of a vague language *L*.

3. Let's Get a Bit Formal

Let *L* be a first-order language with identity sign, vague predicates, and connectives "¬", "∧" and "∨". A model $M=<D_M, VI_M, v_M>$ for *L* is a triple that satisfies the following conditions:

1. D_M is a non-empty set.

2. VI_M is a subset of D_M^2, where (i) for any $<\alpha, \beta>$ that belongs to VI_M, α is not the same as β, (ii) if $<\alpha, \beta>$ belongs to VI_M, so does $<\beta, \alpha>$, and (iii) if $<\alpha_1, ..., \alpha_{i-1}, \alpha, \alpha_{i+1}, ..., \alpha_n> \in F_M^+$ while $<\alpha_1, ..., \alpha_{i-1}, \beta, \alpha_{i+1}, ..., \alpha_n> \notin F_M^-$ for some n-place predicate F, $<\alpha, \beta> \notin VI_M$.
3. v_M assigns to each individual constant of L a member of D_M to be its value and assigns to each n-place predicate F a pair of sets $<F_M^+, F_M^->$ of n-tuples of members of D_M such that $F_M^+ \cap F_M^- = \varnothing$.

Intuitively, VI_M specifies a relation of "vague identity" that is both irreflexive and symmetric on the domain D_M and never invalidates Leibiz's Law. Given a model *M*, we define the concept of true-in-*M* ($v_M(A) = 1$ in symbol), that of false-in-*M* ($v_M(A) = 0$ in symbol), and that of neither-true-nor-false-in-*M* ($v_M(A) = n$ in symbol) in the usual way:

1. $v_M(Fc_1 ... c_n) = 1$ if $<v_M(c_1), ..., v_M(c_n)>$ belongs to F_M^+. $v_M(Fc_1 ... c_n) = 0$ if $<v_M(c_1), ..., v_M(c_n)>$ belongs to F_M^-. Otherwise, $v_M(Fc_1 ... c_n) = n$.
2. $v_M(c_1 = c_2) = 1$ if $v_M(c_1) = v_M(c_2)$. $v_M(c_1 = c_2) = n$ if $<v_M(c_1), v_M(c_2)>$ belongs to VI_M. Otherwise, $v_M(c_1 = c_2) = 0$.
3. Truth-values of compound sentences are determined by the usual strong K_3 charts.
4. $v_M(\forall x_i \phi) = 1$ if $v_M(\phi(c_i)) = 1$ for every constant c_i. $v_M(\forall x_i \phi) = 0$ if $v_M(\phi(c_i)) = 0$ for some constant c_i. Otherwise $v_M(\forall x_i \phi) = n$. (For simplicity, we assume that everything in the domain has a name.)

Again, the notion of validity is defined in the usual way: an argument is valid iff it preserves truth-in-*M* for every model *M*.

However, the present approach differs from most semantic theories of vagueness in that it proposes that there is more than one correct (or intended) interpretation of a vague language *L*, all of which differ *only* in how vague predicates are to be interpreted. (Since they differ only in how vague predicates are to be interpreted, I will assume in what follows that all of them share the same domain, assign the same objects to individual constants, and assign the same extensions and anti-extensions to non-vague predicates.) According to what we have said in the previous section, while these intended interpretations may differ in assigning different pairs $<F_M^+, F_M^->$ to a vague predicate F, these different pairs nevertheless share a "common core", i.e., $\cap\{v_M(F_M^+) \mid M \in S\} \neq \varnothing$ and $\cap\{v_M(F_M^-) \mid M \in S\} \neq \varnothing$, where *S* is the set of all correct interpretations of a vague language *L*.

I think that the following intuition about the set of all correct interpretations of a vague language *L* is quite plausible: if it is correct to interpret an atomic sentence p as a borderline sentence and it is also correct to interpret another atomic sentence q as a borderline sentence, then it is, *ceteris paribus*, correct to interpret both p and q together as borderline sentences. The reason why this intuition is plausible is, I think, that the third interpretation mentioned in it is intuitively "weaker" or "vaguer", therefore less likely to make mistake, than first two interpretations. It is, however, desirable to make this intuitive idea more general and more precise. As a first approximation, I think that it is plausible to assume that the set *S* of all correct interpretation of a vague language *L* is closed under the following relation (call this assumption (A_1)):

Assumption (A_1): Let p and q be any *atomic* sentences of L. If $v_M(A) = n$ and $v_{M'}(B) = n$ for some models M, $M' \in S$, then there is a model $M^* \in S$ such that M^* is both "weaker than" M' and M and $v_{M^*}(A) = v_{M^*}(B) = n$.

How are we to cash out the idea of "weaker than"? Define a relation ≤ between models as such: $M_1 \leq M_2$ iff $F_{M1}^+ \subseteq F_{M2}^+$ and $F_{M1^-} \subseteq F_{M2^-}$ for every predicate F^n of L. Then the sense of "weaker than" that I have in mind in Assumption (A_1) is simply the relation ≤. With this identification, Assumption (A_1) now becomes:

Assumption (A): Let p and q be any *atomic* sentences of L. If $v_M(p) = n$ and $v_{M'}(q) = n$ for some models M, $M' \in S$, then there is a model $M^* \in S$ such that $M^* \leq M$, $M^* \leq M'$, and $v_{M^*}(p) = v_{M^*}(q) = n$.

In words: if it is correct to classify an atomic sentence p as a borderline sentence and it is also correct to classify an atomic sentence q as a borderline sentence, then it is correct to classify both atomic sentences p and q as borderline sentences in an interpretation that does not assign classical truth-values, i.e., truth and falsity, to more sentences than the previous two do. I think that Assumption (A) is intuitively plausible and will assume it in what follows. If we further assume that every object in the common domain of S has a name in the language L, Assumption (A) can also be put in a slightly different but equivalent way as (A*):[5]

Assumption (A*): Let F^n and G^m be any n-place and m-place predicates of L. If there are models M, $M' \in S$ and sequences of objects $<a_1, \ldots, a_n>$ and $<b_1, \ldots, b_m>$ such that $<a_1, \ldots, a_n> \notin F^n{}_M{}^+ \cup F^n{}_{M^-}$ and $<b_1, \ldots, b_m> \notin G^m{}_{M'}{}^+ \cup G^m{}_{M'^-}$, then there is a model $M^* \in S$ such that $M^* \leq M$, $M^* \leq M'$, $<a_1, \ldots, a_n> \notin F^n{}_{M^*}{}^+ \cup F^n{}_{M^{*-}}$, and $<b_1, \ldots, b_m> \notin G^m{}_{M^*}{}^+ \cup G^m{}_{M^{*-}}$.

For the record, I also state below one more assumption that I will make about the set S of all correct interpretations of a vague language L:

Assumption (B): For any *atomic* sentence p, if there is a model $M \in S$ such that $v_M(p) \neq 1$ and there is a model $M' \in S$ such that $v_{M'}(p) \neq 0$, then there is a model $M^* \in S$ such that $v_{M^*}(p) = n$.

Assumption (B) says that, in terms of the terminologies that I am about to introduce, if an atomic sentence is neither true *simpliciter* nor false *simpliciter*, then it is correct to interpret it as having no truth value. With the assumption that every object in the common domain of S has a name in the language L, Assumption (B) can also be put in a slightly different but equivalent way as (B*):[6]

Assumption (B*): For any n-place predicate F, if some sequence of objects $<a_1, \ldots, a_n>$ is such that there is a model $M \in S$ such that $<a_1, \ldots, a_n> \notin F_M{}^+$ and there is also a model $M' \in S$ such that $<a_1, \ldots, a_n> \notin F_{M'^-}$, then there is a model $M^* \in S$ such that $<a_1, \ldots, a_n> \notin F^n{}_{M^*}{}^+ \cup F^n{}_{M^{*-}}$.

5 That Assumption (A) and Assumption (A*) are equivalent, giving that every object in the common domain has a name, is not difficult to prove and is left as an exercise for readers.

6 Again, the equivalence between Assumption (B) and Assumption (B*), giving that every object in the common domain has a name, is not difficult to prove and is left as an exercise for readers.

I think that Assumption (B) (or (B*)) is self-evident if we take the predicate F in it to be a vague predicate, but I also think that it is justifiable by what we have said in the previous section. In Appendix, I will appeal to both Assumption (A) and Assumption (B) to prove that the current approach preserves a very important advantage of the traditional three-valued approach to vague predicates.

Given a vague language *L* and the set *S* of all its correct interpretations, we can now define the important notions of "true *simpliciter*" and "false *simpliciter*" as follows:

A sentence is true *simpliciter* iff it is true-in-*M* for every *M* in *S*.
A sentence is false *simpliciter* iff it is false-in-*M* for every *M* in *S*.

Borderline sentences are then sentences that are neither true *simpliciter* nor false *simpliciter*. (For a further classification of borderline sentences, see below.)

Notice that, even though the definition of the notion of truth *simpliciter* (or falsity *simpliciter*) is superficially similar to that of the notion of supertruth (or superfalsity) of supervaluationism, these two notions differ significantly in at least two respects. First, the former does not, while the latter does, appeal to the notion of a classical precisification of a three-valued model for its definition. Second and more importantly, with assumptions (A) and (B), we can prove that all operators that we have met so far are truth-functional in the sense that, e.g., a disjunction is true *simpliciter* iff one of its disjunct is true *simpliciter*. (The proof of this claim is given in the appendix.) As a further result, the definition of truth (or falsity) *simpliciter* given here does not, while the notion of supertruth (and superfalsity) does, suffer from the problem of missing witness. For example, with assumptions (A) and (B) in hands, we can show that an existential statement is true *simpliciter* iff one of its instance is true *simpliciter* and that a conjunction is false *simpliciter* iff one of its conjunct is false *simpliciter*. In short, the current approach enjoys a very important advantage of the traditional three-valued approach to vague predicates, i.e., truth-*simpliciter*-functionality.

We can, if we want to, make a further distinction among borderline sentences. It may or may not happen that a borderline sentence is neither-true-in-*M*-nor-false-in-*M* for every *M* of *S*. When it happens in this way, we call such a sentence "a pure borderline sentence" and the object it mentions "a pure borderline case" of the vague predicate. We say that the kind of vagueness that these sentences and cases have is first-order. However, it may also happen that a borderline sentence is true-in-*M* for some but not all *M* of *S*, or false-in-*M* for some but not all *M* of *S*, or both. When a sentence is true-in-*M* for some but not all *M* of *S*, or false-in-*M* for some but not all *M* of *S*, or both, we call such a sentence "an impure borderline sentence" and the object it mentions "an impure borderline case". We also say that the kind of vagueness that these sentences and cases have is higher-order. Of course, we can make a further distinction among sentences of higher-order vagueness according to their fate in the set *S*, but there is no need to pursue this line of thought here.

The following, then, is my formal "solution" to the sorites paradox, and I suggest the name "three-valued plurivaluationism" for it. In short, three-valued plurivaluationism asserts that a vague language *L* has more than one correct three-valued interpretation, and it diagnoses the fallacy of a paradoxical sorites argument as follows: in each correct interpretation *M* of *L*, there is a premise in the sorites argument that is neither-true-nor-false-in-M; so one of the premises of the sorites argument is not true *simpliciter*. The argument is still valid, as one can easily verify, but it is unsound. Why are we taken in by a paradoxical sorites argument? Traditionally, the reply to this question from a three-valued theorist is mainly this: even though one of its premises is not true *simpliciter*, none of its premises are false *simpliciter* either. Because none of the premises of a paradoxical sorites argument is false *simpliciter*, we are thereby led to think that all of them are true *simpliciter*, and this is how we are taken in. A three-valued plurivaluationist would agree with this reply, but s/he would also add to it: we are led to take the premises of a paradoxical sorites argument to be true *simpliciter*, not only because none of them is false *simpliciter*, but also because they are often true in some, perhaps even in many or most though not in all, correct interpretations of the language L.

4. Objections and Replies

There are two main objections to a three-valued solution to the sorites paradox. First, it may be argued that a three-valued solution overlooks what [6] called the phenomenon of "penumbral connection": logical relations exist between borderline sentences, as illustrated by the following "intuitively true sentences": "Every head is either bald or not bald", "No head is both bald and not bald", "Every head is such that if it is bald then it is bald, and if it is bald then it is either bald or shining". But this penumbral connection, says the objector, is missing or cannot be asserted in a standard three-valued semantics. Second, it may be said that a three-valued solution faces what [22] called "the jolt problem" or what [29] called "the notorious problem of higher-order vagueness": vague predicates force a "gradual transition" from truth to falsity or have borderline cases of borderline cases, but such a gradual transition cannot be accommodated in a three-valued semantics and such higher-order borderline cases "have never received an adequate treatment". I'll begin with the second objection first.

It is implausible to say that a three-valued semantics cannot accommodate a gradual transition from truth to falsity. After all, there are different orders, i.e., first-order and higher-order, of vagueness between truth *simpliciter* and falsity *simpliciter* as we have seen, so that one cannot directly jump from truth *simpliciter* to falsity *simpliciter* without passing by all these intermediate borderline sentences. But a three-valued plurivaluationist can actually do better than just having a few intermediaries between truth *simpliciter* and falsity *simpliciter*. Let S be the set of all correct interpretations of a vague language L, and let S^{A+} (S^{A-}) be the subset of S containing all and only those models such that A is true (false) in them. We can then define the degree of closeness to truth (or to falsity) *simpliciter* $c^+(A)$ (or $c^-(A)$) of a sentence A simply as $|S^{A+}|/|S|$(or$|S^{A-}|/|S|$). By these definitions, every sentence A will receive a pair of rational numbers $<c^+(A), c^-(A)>$ between 0 and 1 that measure its degree of closeness to truth *simpliciter* and its degree of closeness to falsity *simpliciter* separately. A first-order vague sentence A will then be one such that $c^+(A) = c^-(A) = 0$, while a higher-order vague sentence may receive any rational number between 0 and 1 as its degree of closeness to truth (or to falsity) *simpliciter*. This already gives us both a gradual transition from truth *simpliciter* to pure borderline cases and a gradual transition from the latter to falsity *simpliciter*. However, if one insists that we should have a unique number for "the degree of truth" of a sentence, we may define the degree of truth* of a sentence A to be

$(1+c^+(A)-c^-(A))/2$. By this last definition (or any other equally plausible definition), three-valued plurivaluationism will then allow sentences to have a gradual transition from truth* of degree 1 (positive cases) to truth* of degree 0, i.e., falisity* (negative cases). Either way, we will have an explanation of why some people, such as N. Smith, think that vague predicates force a gradual transition from truth to falsity.

It is true that the above definitions assumes that the set *S* has only finite members; when $|S|$ is some infinite cardinal number, the above definitions may not give the desired results. But the assumption that the set *S* is finite is actually unimportant for a three-valued plurivaluationist, for the whole point of the above definitions is not to define an exact degree of truth (or truth*) for each sentence of *L*, but to show that there is in some sense a gradual transition from truth *simpliciter* to falsity *simpliciter*. Even if $|S|$ is an infinite cardinal number and the above definition will not work properly, it is still possible to find a sequence of sentences $A_1, \ldots, A_n, \ldots$ such that $|S^{A_1+}| \subseteq \ldots \subseteq |S^{A_n+}| \ldots$ (or that $|S^{A_1-}| \subseteq \ldots \subseteq |S^{A_n-}| \ldots$). The existence of such a sequence (or sequences) is enough to show that there is in some sense a gradual transition from truth *simpliciter* to falsity *simpliciter*.

T. Williamson [29, p. 121] complains that "the problem [of higher-order vagueness] has never received an adequate treatment within the framework of three-valued … logic." He imagines that the three-valued theorist defines an operator "Δ", read as "it is definitely that", in the following way: ΔA is true iff A is true and is false if otherwise. He then shows that, while the three-valued theorist can successfully assert of a borderline sentence P that $\neg\Delta P \wedge \neg\Delta\neg P$ (it is neither definitely so nor definitely not so), the three-valued theorist nevertheless has to agree that "$\Delta\Delta P \vee \Delta\neg\Delta P$" (it is definite whether it is definitely so) is also true. But this last assertion, says Williamson, "does not fit the intended interpretation of Δ", because P may be a higher-order borderline sentence and, if so, whether or not it is definitely so will not be definite. I have to admit that I don't quite understand what Williamson's complaint is here. A three-valued theorist is tasked with providing a semantics for vague *predicates*; he is not tasked with providing a semantics for the *operator* "it is definitely that". So long as a three-valued theorist accomplishes his/her task, his/her job is done. Whether or not s/he can provide a further semantics, perhaps by appealing to more complicated structures, for the operator "it is definitely that" is a separate issue, and the failure of the further task cannot diminish a bit of his previous achievement. At any rate, if a semantics for the operator "it is definitely that' is indeed desired, the three-valued plurivaluationist does not have to define the operator in the truth-functional way as Williamson does. He can "modalize" it by stipulating that a sentence $v_M(\Delta A) = 1$ in a correct interpretation *M* if $v_{M'}(A) = 1$ for every correct interpretation *M*' of *S*, $v_M(\Delta A) = 0$ in a correct interpretation *M* if either $v_{M'}(A) = 0$ for every correct interpretation *M*' of *S* or $v_{M'}(A) = n$ for every correct interpretation *M*' of *S*, and $v_M(\Delta A) = n$ in a correct interpretation *M* if otherwise. Defined in this way, we can show that "$\neg\Delta P \wedge \neg\Delta\neg P$" is still true *simpliciter* if "P" is a first-order, i.e., pure borderline sentence,[7] while "$\Delta\Delta P \vee \Delta\neg\Delta P$" is not true *simpliciter* if "P" is a higher-order, i.e., impure borderline sentence.[8]

---

7 This is a satisfactory result, for it is unreasonable to assert of a higher-order borderline sentence that it is neither definitely so nor definitely not so; after all, its being a borderline sentence is

Turning now to the problem of penumbral connection, the first thing to notice is that what seems to be a datum for penumbral connection to K. Fine may not seem so to other philosophers. As Smith points out in [21] p.86:[9]

> *...Consider 'red'. If one indicates a point on a rainbow midway between clear red and clear orange and asks an ordinary speaker the following questions, then in my experience the responses are along the lines indicated:*
> – *"Is the point red?" Umm, well, sort of.*
> – *"Is the point orange?" Umm, well, sort of.*
> – *"But it's certainly not red and orange, right?" Well, no, it sort of is red and orange.*
> – *"OK, well it's definitely red or orange, right?" No, that's what I've been saying, it's a bit of both, the colours blend into one another.*
>
> *These reactions fit with the recursive assignments of truth values, not the supervaluationist assignments.*

The right thing to conclude from these remarks, I think, is that some of the claimed data for penumbral connection, especially those involving truth-functional connectives "and" and "or", are not genuine data at all. But this is not to deny that some data are still genuine, especially those involving conditionals, such as "Every head is such that if it is bald then it is bald" and "Every head is such that if it is bald then it is either bald or shining". However, the fact that these conditionals are indeed true shows only that the connective "if ... then ..." in them should, as many philosophers think that it should, be construed as a non-truth-functional connective for a theorist who prefers a three-valued treatment of a vague language.

There are several well-known theories of conditionals in this direction. For example, [24] and [15] have proposed a very popular way of treating the connective "if ... then ..." as a modal operator. According to this line of treatment, a conditional "if A then B" asserts that, to simplify a bit, every closest A-world is also a B-worlds. Following this line of thought, we can define a model for a vague language L to be a 5-tuple $\langle W_M, D_M, f_M, VI_M, v_M\rangle$, where W_M is a non-empty set of possible worlds and f_M is a selection function from a world and a sentence (or a proposition) to a set of worlds satisfying a few conditions. What kind of logic we will have for conditionals will then depend on the formal properties we impose upon the selection function f_M. In all semantic systems that I have known, sentences of the forms "If A then A" and "If A then A or B" are valid, as desired. However, the three-valued plurivaluationist may offer another simpler suggestion: we may take a conditional to be a claim not about the closest A-worlds but about all correct A-models. According to this suggestion, a claim "if A then B" ("$A \rightarrow B$" in symbol) is true in a correct interpretation M of S if S^{A+} is a subset of S^{B+} and is false in M if S^{A+} is a subset of S^{B-}. Otherwise, "$A \rightarrow B$" is neither true nor false in M. Either way,

---

indeterminate.

[8] "$\Delta\Delta P \vee \Delta\neg\Delta P$" is actually false *simpliciter*. Therefore, its negation, i.e., "$\neg\Delta\Delta P \wedge \neg\Delta\neg\Delta P$" is true *simpliciter*. So, even though we cannot assert of a higher-order borderline sentence that it is neither definitely so nor definitely not so, we can assert of it that it is neither definitely definitely so nor definitely not definitely so.

[9] I also found such a reaction in [26].

sentences of the forms “If A then A” and “If A then A or B” turn out to be true *simpliciter* and we have the desired penumbral connection.

References

1. Bennett, B.: Modal Semantics for Knowledge Bases Dealing with Vague Concepts. In Cohn, A.G., Schubert, L., and Shapiro, S. (eds.) *Principles of Knowledge Representation and Reasoning: Proceedings of the 6th International Conference*, pp. 234-44. San Mateo: Morgan Kaufmann (1998)
2. Campbell, R.: The Sorites Paradox. *Philosophical Studies*. 26, 175-91 (1974)
3. Cargile, J.: The Sorites Paradox. In Keefe, R. and Smith, N.J.J. (eds.) *Vagueness: A Reader*. Cambridge, MIT Press (1997)
4. Dummett, M.A.E.: Wang’s Paradox. *Synthese*. 30, 30124 (1975)
5. Fara, D. G.: Shifting Sands: An Interest-Relative Theory of Vagueness. *Philosophical Topics*. 28, 45-81(2000)
6. Fine, K.: Vagueness, Truth and Logic. *Synthese*. 30, 265300 (1975)
7. Goguen, J.A.: The Logic of Inexact Concepts. *Synthese*. 19, 32573 (1969)
8. Hallden, S.: *The Logic of Nonsense*. Uppsala, Uppsala Universitets Arsskrift (1949)
9. Horgan, T.: Robust Vagueness and the Forced-March Sorites Paradox. *Philosophical Perspective*. 8, 159-88 (1994)
10. Hyde, D.: From Heaps and Gaps to Heaps of Gluts. *Mind*. 108, 641-60 (1997)
11. Kamp, H.: Two Theories about Adjectives. In Keenan, E. L. (ed.) *Formal Semantics of Natural Language*, pp123-55. Cambridge: Cambridge University Press (2000)
12. Keefe, R.: *Theories of Vagueness. Cambridge*. Cambridge University Press (2000)
13. K¨orner, S.: *Conceptual Thinking*. Cambridge, Cambridge University Press (1955)
14. Lewis, D. K.: General semantics. *Synthese*. 22, 1867. Reprinted in his *Philosophical Papers*, vol. 1. Oxford, Oxford University Press (1983)
15. Lewis, D. K.: *Counterfactuals*. Oxford: Basil Blackwell Ltd (1973)
16. McGee, V.: Kilimanjaro. *Canadian Journal of Philosophy*, sup. 23, 141-63 (1997)
17. Priest, G.: Fuzzy Identity and Local Validity. *The Monist*. 81(2), 331-342 (1998)
18. Priest, G.: A Site for Sorites. In Jc Beall (ed.) *Liars and Heaps: New Essays on Paradox*. Oxford: Oxford University Press (2003)
19. Przelecki, M.: Fuzziness as Multiplicity. *Erkenntnis*. 10, 371-80 (1976)
20. Raffman, D.: Vagueness without Paradox. *Philosophical Review*. 103, 41-74 (1994)
21. Shapiro, S.: *Vagueness in Context*. Oxford: Clarendon Press (2006)
22. Smith, N. J. J.: *Vagueness and Degrees of Truth*. Oxford: Oxford University Press (2008)
23. Sorensen, R.: *Vagueness and Contradiction*. Oxford: Clarendon Press (2001)
24. Stalnaker, R.: A Theory of Conditionals. In Rescher N. (ed.) *Studies in Logical Theory*, American Philosophical Quarterly, Monograph: 2, pp. 98-112. Blackwell: Oxford, (1968)
25. Tappenden, J.: The Liar and Sorites Paradoxes: Toward a Unified Treatment. *J. of Philosophy*. 90, 551-77 (1993)
26. van Inwagen, P.: How to Reason about Vague Objects. *Philosophical Topics*. 16, 255-84 (1988)
27. Varzi, A.C. Supervaluationism and Its Logics. *Mind*. 112, 295-9 (2001)
28. Williamson, T.: *Vagueness*. London: Routledge (1994)

29. Williamson, T.: Past the Linguistic Turn? In Brian Leiter (ed.), *The Future for Philosophy*. Oxford: Clarendon Press (2004)
30. Wright, C.: On the Coherence of Vague Predicates. *Synthese*. 30, 325-65 (1973)
31. Wright, C.: Further Reflections on the Sorites Paradox. Philosophical Topics. 15, 227-90 (1987)

Appendix

Our goal in this appendix is to prove Theorem 1, i.e., that a disjunction is true *simpliciter* iff at least one of its disjuncts is true *simpliciter*. I first re-list here two assumptions that I made in section 3:

Assumption (A): $\forall M \in S \forall M' \in S$(if $v_M(\text{p}) = \text{n}$ and $v_{M'}(\text{q}) = \text{n}$, then $\exists M^* \in S(M^* \leq M \wedge M^* \leq M' \wedge v_{M^*}(\text{p}) = v_{M^*}(\text{q}) = \text{n}))$, for any *atomic* sentences p and q.

Assumption (B): If $\exists M \in S(v_M(\text{p}) \neq 1)$ and $\exists M' \in S(v_{M'}(\text{p}) \neq 0)$, then $\exists M^* \in S(v_{M^*}(\text{p}) = \text{n})$, for any *atomic* sentence p.

With these two assumptions, we will prove, as lemmas to Theorem 1, that Assumption (A) and Assumption (B) are not only true of atomic sentences but also true of complex ones. Before giving the proofs, however, we state, without giving the proof, a very famous result about strong K_3 and many other three-valued semantics, i.e., Proposition 1. In Proposition 1, the relation $\leq$ between truth-values is defined as: $\text{n} \leq \text{n}$, $0 \leq 0$, $1 \leq 1$, $\text{n} \leq 0$, and $\text{n} \leq 1$.

Proposition 1. If $M_1 \leq M_2$, then $v_{M1}(\text{A}) \leq v_{M2}(\text{A})$, for every sentence of L.

We now set out our task. We prove first that Assumption (A) can be generalized to all sentences, i.e., Lemma 2. As a mid-way to Lemma 2, we prove Lemma 1 first.

Lemma 1. Let p be any atomic sentence and B be any sentence. Then $\forall M \in S \forall M' \in S$(if $v_M(\text{p}) = \text{n}$ and $v_{M'}(\text{B}) = \text{n}$, then $\exists M^* \in S(M^* \leq M \wedge M^* \leq M' \wedge v_{M^*}(\text{p}) = v_{M^*}(\text{B}) = \text{n}))$.
Proof. We prove Lemma 1 by induction on the number of connectives in B. For simplicity, we assume that "∧" is defined in terms of "¬" and "∨" and we omit the quantificational case.
Base case. This is automatically true by Assumption (A).
Inductive step. Assume that Lemma 1 is true of sentences C and D whose numbers of connectives are less than that in B. Two cases:

Cases 1: B is "¬C"
Assume that $v_M(\text{p}) = \text{n}$ and $v_{M'}(\text{B}) = v_{M'}(\neg\text{C}) = \text{n}$ for some $M, M' \in S$. Then $v_{M'}(\text{C}) = \text{n}$. By inductive hypothesis, $\exists M^* \in S(M^* \leq M \wedge M^* \leq M' \wedge v_{M^*}(\text{p}) = v_{M^*}(\text{C}) = \text{n})$. But then $\exists M^* \in S(M^* \leq M \wedge M^* \leq M' \wedge v_{M^*}(\text{p}) = v_{M^*}(\neg\text{C}) = v_{M^*}(\text{B}) = \text{n})$.

Case 2: B is "C ∨ D"
Assume that $v_M(\text{p}) = \text{n}$ and $v_{M'}(\text{B}) = v_{M'}(\text{C} \vee \text{D}) = \text{n}$ for some $M, M' \in S$. There are three sub-cases. In each sub-case, we prove that $\exists M^* \in S(M^* \leq M \wedge M^* \leq M' \wedge v_{M^*}(\text{p}) = v_{M^*}(\text{C} \vee \text{D}) = v_{M^*}(\text{B}) = \text{n})$.

Sub-case 2a: $v_{M'}(\text{C}) = \text{n}$ but $v_{M'}(\text{D}) = 0$.
In this case, $v_M(\text{p}) = \text{n}$ and $v_{M'}(\text{C}) = \text{n}$. So, by inductive hypothesis, $\exists M^* \in S(M^* \leq M \wedge M^* \leq M' \wedge v_{M^*}(\text{p}) = v_{M^*}(\text{C}) = \text{n})$. Since $M^* \leq M'$, it follows that $v_{M^*}(\text{D}) \leq v_{M'}(\text{D})$ by Proposition 1. Since $v_{M'}(\text{D}) = 0$ and $v_{M^*}(\text{D}) \leq v_{M'}(\text{D})$, it further follows that $v_{M^*}(\text{D}) = 0$ or $v_{M^*}(\text{D}) = \text{n}$. Either way, $\exists M^* \in S(M^* \leq M \wedge M^* \leq M' \wedge v_{M^*}(\text{p}) = v_{M^*}(\text{C} \vee \text{D}) = v_{M^*}(\text{B}) = \text{n})$.

Sub-case 2b: $v_{M'}(\text{C}) = 0$ but $v_{M'}(\text{D}) = \text{n}$.
The proof of sub-case 2b is similar to that of sub-case 2a.

Sub-case 2c: $v_{M'}(\text{C}) = \text{n}$ but $v_{M'}(\text{D}) = \text{n}$.

The proof of sub-case 2c is similar to that of sub-case 2a. □

Lemma 2. $\forall M \in S \forall M' \in S$(if $v_M(\mathrm{A}) = \mathrm{n}$ and $v_{M'}(\mathrm{B}) = \mathrm{n}$, then $\exists M^* \in S(M^* \leq M \wedge M^* \leq M' \wedge v_{M^*}(\mathrm{p}) = v_{M^*}(\mathrm{B}) = \mathrm{n})$), for any sentences A and B.

Proof. We prove Lemma 2 by induction on the number of connectives in A. Again for simplicity, we assume that "∧" is defined in terms of "¬" and "∨" and we omit the quantificational case.

Base case. This is automatically true by Lemma 1.

Inductive step. Assume that Lemma 2 is true of sentences C and D whose numbers of connectives are less than that in A. Two cases:

Cases 1: A is "¬C"

Assume that $v_M(\mathrm{A}) = v_M(\neg\mathrm{C}) = \mathrm{n}$ and $v_{M'}(\mathrm{B}) = \mathrm{n}$ for some M, $M' \in S$. Then $v_M(\mathrm{C}) = \mathrm{n}$. By inductive hypothesis, $\exists M^* \in S(M^* \leq M \wedge M^* \leq M' \wedge v_{M^*}(\mathrm{C}) = v_{M^*}(\mathrm{B}) = \mathrm{n})$. But then $\exists M^* \in S(M^* \leq M \wedge M^* \leq M' \wedge v_{M^*}(\mathrm{A}) = v_{M^*}(\neg\mathrm{C}) = v_{M^*}(\mathrm{B}) = \mathrm{n})$.

Case 2: A is "C ∨ D"

Assume that $v_M(\mathrm{A}) = v_M(\mathrm{C} \vee \mathrm{D}) = \mathrm{n}$ and $v_{M'}(\mathrm{B}) = \mathrm{n}$ for some M, $M' \in S$. There are three sub-cases. In each sub-case, we prove that $\exists M^* \in S(M^* \leq M \wedge M^* \leq M' \wedge v_{M^*}(\mathrm{A}) = v_{M^*}(\mathrm{C} \vee \mathrm{D}) = v_{M^*}(\mathrm{B}) = \mathrm{n})$.

Sub-case 2a: $v_M(\mathrm{C}) = \mathrm{n}$ but $v_M(\mathrm{D}) = 0$.

In this case, $v_M(\mathrm{C}) = \mathrm{n}$ and $v_{M'}(\mathrm{B}) = \mathrm{n}$. So, by inductive hypothesis, $\exists M^* \in S(M^* \leq M \wedge M^* \leq M \wedge v_{M^*}(\mathrm{C}) = v_{M^*}(\mathrm{B}) = \mathrm{n})$. Since $M^* \leq M$, it follows that $v_{M^*}(\mathrm{D}) \leq v_M(\mathrm{D})$ by Proposition 1. Since $v_M(\mathrm{D}) = 0$ and $v_{M^*}(\mathrm{D}) \leq v_M(\mathrm{D})$, it further follows that $v_{M^*}(\mathrm{D}) = 0$ or $v_{M^*}(\mathrm{D}) = \mathrm{n}$. Either way, $\exists M^* \in S(M^* \leq M \wedge M^* \leq M' \wedge v_{M^*}(\mathrm{A}) = v_{M^*}(\mathrm{C} \vee \mathrm{D}) = v_{M^*}(\mathrm{B}) = \mathrm{n})$.

Sub-case 2b: $v_M(\mathrm{C}) = 0$ but $v_M(\mathrm{D}) = \mathrm{n}$.

The proof of sub-case 2b is similar to that of sub-case 2a.

Sub-case 2c: $v_M(\mathrm{C}) = \mathrm{n}$ but $v_M(\mathrm{D}) = \mathrm{n}$.

The proof of sub-case 2c is similar to that of sub-case 2a. □

Lemma 3. If $\exists M \in S(v_M(\mathrm{A}) \neq 1)$ and $\exists M' \in S(v_{M'}(\mathrm{A}) \neq 0)$, then $\exists M^* \in S(v_{M^*}(\mathrm{A}) = \mathrm{n})$, for any sentence A.

Proof. We prove Lemma 3 by induction on the number of connectives in A. Again for simplicity, we assume that "∧" is defined in terms of "¬" and "∨" and we omit the quantificational case.

Base case. This is automatically true by Assumption (B).

Inductive step. Assume that Lemma 3 is true of sentences B and C whose numbers of connectives are less than that in A. Two cases:

Cases 1: A is "¬B"

Assume that $\exists M \in S(v_M(\mathrm{A}) = v_M(\neg\mathrm{B}) \neq 1)$ and $\exists M' \in S(v_{M'}(\mathrm{A}) = v_{M'}(\neg\mathrm{B}) \neq 0)$. Then $\exists M \in S(v_M(\mathrm{B}) \neq 0)$ and $\exists M' \in S(v_{M'}(\mathrm{B}) \neq 1)$. So $\exists M^* \in S(v_{M^*}(\mathrm{B}) = \mathrm{n})$ by inductive hypothesis. So $\exists M^* \in S(v_{M^*}(\mathrm{A}) = v_{M^*}(\neg\mathrm{B}) = \mathrm{n})$.

Case 2: A is "B ∨ C"

Assume that $\exists M \in S(v_M(\mathrm{A}) = v_M(\mathrm{B} \vee \mathrm{C}) \neq 1)$ and $\exists M' \in S(v_{M'}(\mathrm{A}) = v_{M'}(\mathrm{B} \vee \mathrm{C}) \neq 0)$. By the fact that $\exists M \in S(v_M(\mathrm{A}) = v_M(\mathrm{B} \vee \mathrm{C}) \neq 1)$, it follows that $\exists M \in S(v_M(\mathrm{B}) \neq 1)$ and $\exists M \in S(v_M(\mathrm{C}) \neq 1)$. And, by the fact that $\exists M' \in S(v_{M'}(\mathrm{A}) = v_{M'}(\mathrm{B} \vee \mathrm{C}) \neq 0)$, it follows that either $\exists M \in S(v_M(\mathrm{B}) \neq 0)$ or $\exists M \in S(v_M(\mathrm{C}) \neq 0)$. We prove by cases in what follows that either case leads to the conclusion that $\exists M^*\in S(v_{M^*}(A) = v_{M^*}(B \vee C) = n)$.

Case 2a: $\exists M\in S(v_M(B) \neq 0)$

In this case, $\exists M\in S(v_M(B) \neq 1)$ and $\exists M\in S(v_M(B) \neq 0)$. So, by the inductive hypothesis, $\exists M^*\in S(v_{M^*}(B) = n)$. Now, $\exists M\in S(v_M(C) \neq 1)$. So, either $\forall M\in S(v_M(C) = 0)$ or $\exists M\in S(v_M(C) = n)$, otherwise it will contradict with the inductive hypothesis that the lemma holds for C. If $\forall M\in S(v_M(C) = 0)$, then the model $M^*\in S$ s.t. $v_{M^*}(B) = n$ will also be a model in which $v_{M^*}(A) = v_{M^*}(B \vee C) = n$. And if $\exists M\in S(v_M(C) = n)$, then, since $\exists M^*\in S(v_{M^*}(B) = n)$, there will be a model M' s.t. $v_{M'}(B) = v_{M'}(C) = n$ by Lemma 2 and therefore $v_{M'}(A) = v_{M'}(B \vee C) = n$. So, if $\exists M\in S(v_M(B) \neq 0)$, then $\exists M^*\in S(v_{M^*}(A) = v_{M^*}(B \vee C) = n)$.

Case 2b: $\exists M\in S(v_M(C) \neq 0)$

The proof of Case 2b is similar to that of Case 2a. □

Now we prove the main theorem: if a disjunction is true *simpliciter*, then at least one of its disjuncts is true *simpliciter*, i.e.:

Theorem 1. If $\forall M\in S(v_M(A \vee B) = 1)$, then $\forall M\in S(v_M(A) = 1)$ or $\forall M\in S(v_M(B) = 1)$.
Proof: Assuming that $\forall M\in S(v_M(A \vee B) = 1)$, we prove the consequent of the theorem by *reductio*. Suppose that it is neither the case that $\forall M\in S(v_M(A) = 1)$ nor the case that $\forall M\in S(v_M(B) = 1)$. So, $\exists M\in S(v_M(A) \neq 1)$ and $\exists M\in S(v_M(B) \neq 1)$. Four possibilities:

(a) $\forall M\in S(v_M(A) = 0)$ and $\forall M\in S(v_M(B) = 0)$. In this case, $\forall M\in S(v_M(A \vee B) = 0)$, which contradicts with our initial assumption.

(b) $\forall M\in S(v_M(A) = 0)$ but $\neg\forall M\in S(v_M(B) = 0)$. In this case, $\exists M\in S(v_M(B) \neq 0)$. Since $\exists M\in S(v_M(B) \neq 1)$ and $\exists M\in S(v_M(B) \neq 0)$, it follows by Lemma 3 that $\exists M\in S(v_M(B) = n)$. So, $\exists M\in S(v_M(B) = n$ and $v_M(A) =0)$. Therefore, $\exists M\in S(v_M(A \vee B) = n)$, which contradicts with our initial assumption.

(c) $\forall M\in S(v_M(B) = 0)$ but not $\neg\forall M\in S(v_M(A) = 0)$. In this case, $\exists M\in S(v_M(A) \neq 0)$. Since $\exists M\in S(v_M(A) \neq 1)$ and $\exists M\in S(v_M(A) \neq 0)$, it follows by Lemma 3 that $\exists M\in S(v_M(A) = n)$. So, $\exists M\in S(v_M(A) = n$ and $v_M(B) =0)$. Therefore, $\exists M\in S(v_M(A \vee B) = n)$, which contradicts with our initial assumption.

(d) Neither $\forall M\in S(v_M(A) = 0)$ nor $\forall M\in S(v_M(B) = 0)$. So $\exists M\in S(v_M(A) \neq 0)$ and $\exists M\in S(v_M(B) \neq 0)$. It follows from the facts that $\exists M\in S(v_M(A) \neq 1)$ and $\exists M\in S(v_M(A) \neq 0)$, by Lemma 3, that $\exists M\in S(v_M(A) = n)$. Similarly, it follows from the facts that $\exists M\in S(v_M(B) \neq 1)$ and $\exists M\in S(v_M(B) \neq 0)$, again by Lemma 3, that $\exists M\in S(v_M(B) = n)$. So, $\exists M\in S(v_M(A) = n)$ and $\exists M\in S(v_M(B) = n)$. By Lemma2, it follows that $\exists M^*\in S(v_{M^*}(A) = v_{M^*}(B) = n = v_{M^*}(A \vee B))$, which contradicts with our initial assumption. □