For citations go to the published version:
2011: *Vagueness and Language Use*
Paul Egré and Nathan Klinedinst (eds.) Palgrave Macmillan

Supervaluationism and Fara's argument concerning higher-order vagueness

Pablo Cobreros
pcobreros@unav.es

January 26, 2011

There is an intuitive appeal to truth-value gaps in the case of vagueness. The idea is that the facts that determine the meaning of vague expressions (facts about use, most likely) left unsettled for a range of cases whether the expression applies. This sort of *semantic unsettledness* has been taken for a long time as a proper motivation for truth-value gaps; since it is unsettled whether the expression applies the correspoding sentence is neither true nor false.

The issue might be formulated in terms of a sort of symmetry in our dispositions to assent to a borderline sentence or to assent to its negation. Suppose Tim is a borderline case of the predicate 'thin'. The semantic unsettledness is supposed to lead to a situation in which competent speakers refuse to classify Tim as thin and refuse to classify Tim as not thin. It looks that any consistent truth-value assignment would arbitrarily break the symmetry of our dispositions, and one way to respect the symmetry is by saying that neither sentence is true.[1]

Another motivation for truth-value gaps is that many have found incredible the idea that we are ignorant in borderline cases. If there is a unique correct bivalent valuation of the language, then there must be a cut off in a sorites series for, say, the predicate 'tall'. That is, there must be an item in the series which is the last tall member of the series, immediately followed by a non-tall member. Since it is pretty clear that we do not know where this cut off lies, then we are ignorant about where exactly it does.[2] Truth-value gap theories claim that we do not have any ignorance of this sort for the simple reason that there is no such cut off point.

[1]Some discussion on this might be found in (Williamson, 1994, page) followed by Burgess (2001) and Weatherson (2003).

[2]A general point might be made using Bivalence, Unknowability (If it is borderline whether φ then we do not know whether φ) and a definition of 'ignorance'.

For citations go to the published version:
2011: *Vagueness and Language Use*
Paul Egré and Nathan Klinedinst (eds.) Palgrave Macmillan

Though the truth-gap view on vagueness has been popular for a time, a number of problems have made many philosophers depart from this position. I take it that Fara's argument concerning higher-order vagueness is one of them. Fara argues that, in order to explain the seeming absence of sharp transitions in sorites series, the truth-gap theorist is committed to an endless hierarchy of *gap-principles*. But then she shows, that given her/his particular commitments on logical consequence, the truth-gap theorist cannot consistently endorse the truth of all these gap-principles for any (finite) sorites series.

The aim of this paper is to provide a way out of this problem for the defender of truth-value gaps (in particular, the supervaluationist). The basic idea is taking a third way between two unsatisfactory options: global and local validity. Local validity is not a satisfactory option for a defender of truth-value gaps because the notion of truth preserved by it does not allow for failures of bivalence. In this respect, global validity looks like a natural option. But the notion of truth preserved by global validity adopts an *external* perspective that makes it impossible the accommodation of higher-order vagueness. The proposal considers these two problems and provides a notion of consequence in which the notion of truth necessarily preserved allows for truth-value gaps, but is formulated from an *internal* perspective. The relevant notion of consequence, which I call *regional validity*, lies strictly between global and local validity.

The paper is divided into three sections. The first section briefly presents the supervaluationist theory and Fara's argument concerning higher-order vagueness. The second section presents an argument for the supervaluationist commitment to regional validity and considers briefly the logic provided by this notion of consequence. The third section shows how we might endorse all the gap-principles appealed to in Fara's argument given the regional notion of consequence.

1 Supervaluationism and Fara's argument

1.1 Supervaluationism

The thought underlying truth-value gap theories of vagueness is that vagueness is a matter of underdetermination of meaning. The facts that determine the meaning of a vague predicate do not determine, for a range of cases, whether the predicate applies. The supervaluationist theory provides a picture for understanding vagueness as a form of underdetermination of meaning. The idea is that a vague predicate such as 'tall' can be made precise in several ways consistent with the use we make of it. If Peter is a borderline case of tallness, the sentence 'Peter is tall' will be true in some ways of making 'tall' precise and false in some others. Since each way of making it precise is consistent with the use we make of the expression, our use does not *decide* between these various ways and so the truth value of 'Peter is bald' is left unsettled.

The previous picture assumes that the (intuitive or philosophical) notion of truth allows for failures of bivalence. This is often conveyed by the supervaluationist slogan that 'truth is supertruth': a sentence is true just in case it is true in every way of making precise the vague expressions contained in it. We might extend our language with a *definitely* operator ('$\mathcal{D}$' henceforth) which mirrors the notion of supertruth in the object language. The extended language is amenable to a semantics analogous to the possible-worlds semantics for a simple modal language. An interpretation for a language with '$\mathcal{D}$' is an ordered triple $\langle W, R, \nu \rangle$ where W is a non-empty set of *admissible precisifications* (intuitively, admissible ways of making precise all the expressions of the language), R is an admissibility relation between precisifications (wRw' will be read as w' is admitted [deemed admissible] by w) and ν a function assigning truth-values to sentences at precisifications.[3] Classical operators have their usual meaning (relative to precisifications); the definition of $\mathcal{D}$ is analogous to the definition of the modal operator for necessity:

$\varphi \rightarrow \psi$ takes value 1 at w just in case at w: either φ takes value 0 or ψ takes value 1.

$\neg\varphi$ takes value 1 at w just in case φ takes value 0 at w.

$\mathcal{D}\varphi$ takes value 1 at w just in case φ takes value 1 at every w-admitted precisification.

Constraints on R will depend on the informal reading of the semantics and on questions concerning higher-order vagueness, but it is almost universally admitted in the literature that R should at least be reflexive.

Given this sort of semantics the next question concerns how to define logical consequence. Since logical consequence is a matter of necessary preservation of truth, the commitment to a particular notion of consequence will follow from the commitment to a particular notion of truth. We consider in the first place *local validity* which is the standard way to define logical consequence in modal semantics:[4]

Definition 1. (Local validity) A sentence φ is a local consequence of a

[3] I shall shall often talk about *points* instead of *precisifications* to remain neutral about the informal reading of the semantics.

[4] See, for example, (Blackburn et al., 2001, 31)

set of sentences Γ, written $\Gamma \vDash_l \varphi$, iff for every interpretation and any point w in that interpretation: if all the $\gamma \in \Gamma$ take value 1 in w then φ takes value 1 in w.

Local validity preserves *local truth*, where a sentence φ is locally true at w just in case φ takes value 1 at that point. Local validity is a natural way to interpret logical consequence in modal semantics since under this reading of the semantics, being true (in the intuitive or philosophical sense) is being locally true. But local truth does not allow for failures of bivalence in the sense that given an interpretation and a point w, any sentence of the language will be either locally true or locally false at w (and so, there are no interpretations with points at which some sentence is neither locally true nor locally false).[5] For this reason, given that validity is a matter of necessary preservation of truth, local validity cannot be adequate for a defender of truth value gaps. It is usually assumed in the literature that the supervaluationist is committed to something like *global validity*:[6]

Definition 2. (Global validity) A sentence φ is a global consequence of a set of sentences Γ, written $\Gamma \vDash_g \varphi$, iff for every interpretation: if all the $\gamma \in \Gamma$ take value 1 in every point then φ takes value 1 in every point.

Global validity preserves *global truth*, where a sentence φ is globally true at w just in case it takes value 1 at every point. Global truth allows for failures of bivalence since if a sentence φ takes value 1 at a point and value 0 at some other point in the same interpretation, φ is neither globally true nor globally false in any point in that interpretation. In this sense, global truth looks like the natural option for the supervaluationist. As we shall see later, however, global truth is not fully adequate in a different sense.

A particular distinctive feature of global validity is the validity of the following rule:

$\mathcal{D}$-introduction $\Gamma \vdash \varphi \Longrightarrow \Gamma \vdash \mathcal{D}\varphi$

which has as a particular case the inference from φ to $\mathcal{D}\varphi$. While the inference is globally valid (if φ takes value 1 at every w so does $\mathcal{D}\varphi$), it is not locally valid (since φ might take value 1 at w whereas $\mathcal{D}\varphi$ does not). The inference plays a key role in Fara's argument below.

[5] I equate *falsity* with *truth of the negation*.

[6] Fine implicitly assumes the global notion of consequence in his classical paper (Fine, 1975, 290). Keefe assumes also the global notion in (Keefe, 2000b, 176) and Keefe (2000a).

1.2 Fara's argument concerning higher-order vagueness

Take a long series of men. The first man in the series is 2.5 meters tall and the last is 1.5. Each man in the series differs from his successor by less than a millimetre. It seems that there is no sharp transition from the members of the series that are tall to those that are not tall. Truth-value gap theories explain this fact by appealing to the presence of borderline cases in between. There is actually no sharp transition from the members of the series of which it is true that they are tall to those of which it is true that they are not tall (= falsely tall), because there are some members in between that are neither truly tall nor truly not tall (falsely tall). Thus, what explains the seeming absence of a sharp transition is the fact that there is no x in the series such that x is truly tall but its successor in the series is truly not tall (falsely tall). Or equivalently, for any x in the series, if x is truly tall, it is not the case that its successor in the series is truly not tall (falsely tall). This *gap principle* for the predicate 'tall' is what explains, according to the truth-value gap theorist, the seeming absence of sharp transitions from the members of the series that are tall to those that are not. Given that, for the truth-value gap theorist, '$\mathcal{D}$' is an object language expression of the theory's notion of truth, we might express this gap principle in the object language:

$$(\textsf{GP for '}T\textsf{'})\ \mathcal{D}T(x) \rightarrow \neg\mathcal{D}\neg T(x'),$$

where 'T' stands for 'tall' and x' is the successor of x in the series.

However, the seeming absence of a sharp transition in the series cuts deeper than that. We are just as unable to find a sharp transition from the definite cases of 'tall' to its non-definite cases, as we were with respect to the simple positive and negative cases of the predicate. Since the phenomenon seems to be no different in kind, the truth-value gap theorist is compelled to explain this second seeming absence of a sharp transition in analogous terms. Thus, what explains the seeming absence of a sharp transition from the members of the series that are definitely tall to those that are not definitely tall is the truth of a gap principle, this time for 'definitely tall':

$$(\textsf{GP for '}\mathcal{D}T\textsf{'})\ \mathcal{D}\mathcal{D}T(x) \rightarrow \neg\mathcal{D}\neg\mathcal{D}T(x')$$

There seems to be no reason to claim that higher-order vagueness stops at some finite level. Once one accepts an account of vagueness based on borderline cases, one should be prepared to endorse the hierarchy of borderline cases all the way up. If the theory cannot treat non-terminating higher-order vagueness at least as a logical possibility, one might well start doubting the whole picture. If the truth-value gap theorist endorses the last claim then there might be no sharp transitions in suitably long sorites series, and explaining this fact requires endorsing the truth of all the gap principles of this form:

(GP for '$\mathcal{D}^n T$') $\mathcal{D}\mathcal{D}^n T(x) \rightarrow \neg\mathcal{D}\neg\mathcal{D}^n T(x')$

In her 2003 paper, Delia Fara argues that truth-value gap theorists cannot appeal to the truth of all these gap-principles to explain the seeming absence of sharp transitions in sorites series. Fara points out that when we read $\mathcal{D}$ as a truth-operator, then one seems to be committed to the rule of $\mathcal{D}$-introduction (and in particular to the inference from φ to $\mathcal{D}\varphi$)(Fara, 2003, 199-200). But Fara shows that, given $\mathcal{D}$-intro, the truth of all the gap principles is inconsistent for finite sorites series. The proof can be represented with the next diagram:[7]

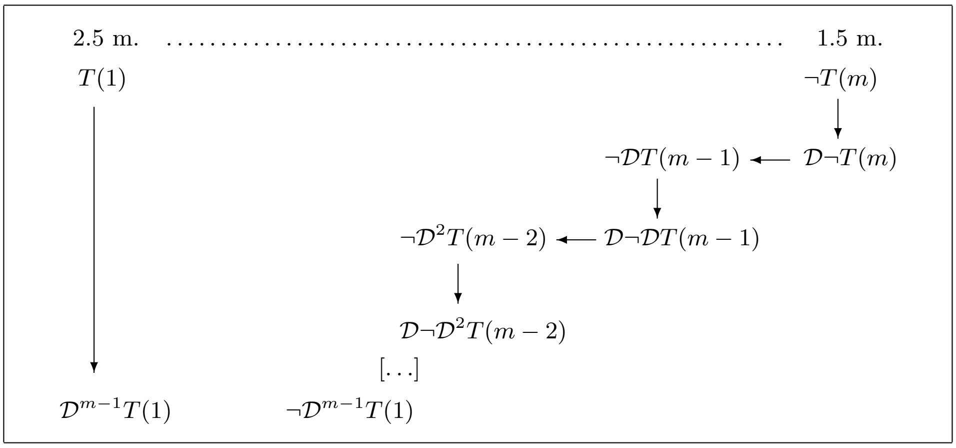

Figure 1: Fara's argument

Each move downwards in the picture represents an application of $\mathcal{D}$-intro, while each move leftwards represents an application of the relevant instance of the relevant gap-principle (in contrapositive form). As the diagram shows, the relevant instance of the relevant gap-principle works in tandem with $\mathcal{D}$-intro, forcing us to move leftwards to pick out a new element in the series. A number $m-1$ of moves leftward suffices to reach the conclusion that $\neg\mathcal{D}^{m-1}T(1)$. But this conclusion contradicts the assumption that the first

[7]A picture of this sort might be found in (Fara, 2003, page). The argument is based on the same idea as an argument of Wright's (Wright (1987) and Wright (1992)). However, the reasoning used in Wright's argument is too strong since it is not globally valid (this observation on Wright's argument is to be found in Heck (1993)). Further discussion of the argument can be found in Edgington (1993) and Sainsbury (1991).

member of the series is tall, since we might infer that $\mathcal{D}^{m-1}T(1)$ from it, by $m-1$ applications of $\mathcal{D}$-intro.

The reasoning employed by Fara in her argument is globally valid. This means that if one is committed to global validity, one might try to disguise the result, but cannot consistently endorse the truth of all the gap-principles. This looks to me as a pretty bad result for global consequence. Still, the supervaluationist might argue that the notion of global validity is not adequate to model the supervaluationist notion of consequence.

2 Regional validity

2.1 From higher-order vagueness to regional validity

As mentioned above, local validity does not constitute a satisfactory notion of consequence for truth-value gap theories in general and for supervaluationism in particular. The reason is that validity is a matter of necessary preservation of truth, but the notion of truth preserved by local validity does not allow for truth-value gaps. In this respect global validity looks like the right choice, since a sentence might be neither globally true nor globally false in a point in an interpretation. Still, global validity is not adequate in a different sense, connected to the problem of higher-order vagueness.

Higher-order vagueness, from a truth-value gap theory's perspective, is linked with vagueness in the notion of truth. Take a vague predicate 'F' and a suitably long sorites series for the predicate. If the truth predicate at work is precise there might be cases in the series of which it is neither true nor false that they are 'F', but there will be a sharp transition from the cases of which the predicate is true to those of which it is not. Likewise, there will be a sharp transition from cases in the series of which the predicate is false to those of which it is not. Now the notion of *global truth* is precise since in any interpretation any sentence will be either globally true in every point or not globally true in every point and so there cannot be borderline cases for 'globally true'.

Following Williamson (1999), the point might be stated in a somewhat more precise way. Second-order vagueness in φ, from a truth-value gap theory's perspective, might be defined as vagueness in a classification according to which either 'It is true that φ' holds or 'It is true that $\neg\varphi$' holds or 'It is neither true nor false that φ' holds. The classification is precise if and only if each member is precise and vague otherwise. A sentence φ is precise in an interpretation just in case it takes value 1 in every point or value 0 in every point in the interpretation (Williamson, 1999, 131). Now a sentence φ is globally true at a point in an interpretation if and only if it is globally true in every point in that interpretation. Likewise φ is globally false if and only if it is globally false in every point in that interpretation and neither globally true nor globally false if and only if neither globally true nor globally false in every point in that interpretation. This means that global truth makes the above classification precise for any interpretation. Thus, global truth does not allow even for second-order vagueness.

The problem with global validity is that the notion of truth preserved by it takes an *external perspective* in the sense that relativity to points plays no role in its definition. Saying that something is globally true at a point w in an interpretation is equivalent to simply saying that it is globally true in that interpretation, since, as pointed out before, φ is globally true at a point if and only if it is globally true at every point. This fact makes it impossible the existence of borderline cases for *global truth*, since this would require something being globally true at a point and not globally true at some other point. The natural way to address this question is to consider a notion of truth defined from an *internal perspective*, that is, a notion of truth in which relativity to points play a substantial role. Say that φ is *regionally true* at w in an interpretation just in case φ takes value 1 at every w-admitted point in that interpretation. The corresponding notion of consequence might be called *regional validity*[8].

Definition 3. (Regional validity) A sentence φ is a regional consequence of a set of sentences Γ, written $\Gamma \vDash_r \varphi$, iff for every interpretation and any point w in that interpretation: if all the $\gamma \in \Gamma$ take value 1 in every w-admitted point then φ takes value 1 in every w-admitted point.

A sentence φ might take value 1 at a w-admitted point and value 0 at a different w-admitted point, in which case φ will be neither regionally true nor regionally false at w. Unlike local truth, regional truth allows for failures of bivalence and in this sense regional validity is acceptable for the supervaluationist. But unlike global truth, regional truth does not rule out the possibility of second or higher-order vagueness, since a sentence might be regionally true at a point and not regionally true at a different point in the same interpretation.

In a sense, the notion of regional truth provides a natural strategy for the supervaluationist to accommodate higher-order vagueness. We said before that for the supervaluationist higher-order vagueness is interpreted as vagueness in the notion of truth. Now the idea behind regional truth is that this notion might vary from one precisification to another and, thus, that

[8]As far as I know the first author to consider this sort of approach is Williamson in (Williamson, 1994, p.159, footnote 32). The term *regional validity* is introduced in Cobreros (2008)

the relevant notion of truth for the theory can be made precise in several ways (which is what it is characteristic of vague predicates according to supervaluationism).

The supervaluationist might look at the situation as follows. If supervaluationism is really committed to global validity, then the discussion concerning higher-order vagueness stops here: the theory cannot accommodate even second-order vagueness. If we give her/him the chance to endorse regional validity instead, then the discussion goes on and we'll see whether she/he can address Fara's argument.

2.2 Logic with regional validity

The question about which logic flows from the regional notion of consequence depends on restrictions on the admissibility relation. It is usually assumed that factivity ($\mathcal{D}\varphi \rightarrow \varphi$) is a minimal requirement[9] and where admissibility is reflexive regional consequence is stronger than local consequence. For assume that $\Gamma \nvDash_r \varphi$; in that case there is an interpretation and a precisification w such that every member of Γ takes value 1 at every w-admitted precisification and φ value 0 at some. The precisification at which φ takes value 0 shows that $\Gamma \nvDash_l \varphi$. On the other hand, where admissibility is reflexive, not every regionally valid argument is locally valid. In particular $\{\varphi, \neg\mathcal{D}\varphi\} \vDash_r \bot$ (given reflexivity, these two sentences cannot be both regionally true at a precisification w since this would require φ taking value 0 at some w-accessible and φ value 1 at every w-accessible) but $\{\varphi, \neg\mathcal{D}\varphi\} \nvDash_l \bot$.

Without any other constraint on admissibility, regional validity is weaker than global validity. In particular, the inference from φ to $\mathcal{D}\varphi$ no longer holds, since a sentence φ might be regionally true at w_0 while $\mathcal{D}\varphi$ is not. Graphically explained:

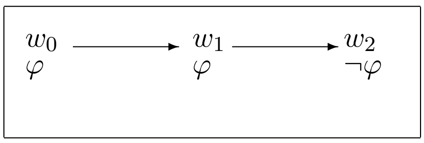

Figure 2: Failure of $\mathcal{D}$-intro

The diagram shows that things change if admissibility is required to be transitive; the inference from φ to $\mathcal{D}\varphi$ is regionally valid for transitive interpretations.[10] But there is a reason to reject that admissibility is transitive

[9] See for example (Williamson, 1999, 130) or the discussion between (McGee and McLaughlin, 1998, 227) and (Williamson, 2004, 119).

[10] The two notions actually collapse for reflexive and transitive interpretations (Cobreros, 2008, 305).

based on the vagueness of *supertruth* that is quite natural for the supervaluationist. The idea is that in the same way in which the supervaluationist rejects the validity of the schema '$Fa \to \mathcal{D}Fa$' based on the vagueness of 'F', the supervaluationist rejects the validity of the schema '$\mathcal{D}Fa \to \mathcal{D}\mathcal{D}Fa$' based on the vagueness of 'supertruth' (recall that $\mathcal{D}$ is an object language expression of *supertruth*).[11] Since the transitivity of admissibility guarantees the validity of the schema, the rejection of the schema entails the rejection of the transitivity of admissibility.

The failure of the inference from φ to $\mathcal{D}\varphi$ might look surprising given the supervaluationist reading of '$\mathcal{D}$'. As Fara points out, for a truth-value gap theorist, '$\mathcal{D}$' means something like 'it is true that', and in that case 'it seems impossible for a sentence S to be true while another sentence, 'it is true that 'S'', that says (in effect) that it's true is not true' (Fara, 2003, 199-200). But the supervaluationist might defend the reasonableness of the failure of that inference in the presence of higher-order vaugeness. On the one hand, the two facts that lead to that failure (the relativity to worlds involved in the notion of regional truth and the failure of transitivity) are well motivated by the problem of higher-order vagueness. It actually seems to me that the source of the intuition appealed to by Fara is the same as the one stating the validity of the schema '$Fa \to \mathcal{D}Fa$', that the supervaluationist rejects. On the other hand, the failure of the inference should not be confused with the idea that the sentences 'φ' and '$\neg\mathcal{D}\varphi$' (the sentence stating φ that is not true) might both be true; the last is not accepted, as it is shown by the fact that $\{\varphi, \neg\mathcal{D}\varphi\} \vDash_r \bot$.

A further motivation for the failure of the inference from φ to $\mathcal{D}\varphi$ might be found in the way the supervaluationist addresses Fara's argument concerning higher-order vagueness, as it will be discussed below.

3 Gap-principles and regional validity

In figure 1, each move leftward represents the use of an instance of a gap principle. The diagram shows that to reach the contradiction for a sorites series of m elements, we need to move $m-1$ times leftwards and, thus, there are $m-1$ relevant instances of a gap-principle:

(GP. $m-1$) $\mathcal{D}\neg T(m) \to \neg\mathcal{D}T(m-1)$
(GP. $m-2$) $\mathcal{D}\neg\mathcal{D}T(m-1) \to \neg\mathcal{D}^2T(m-2)$
⋮
(GP. 3) $\mathcal{D}\neg\mathcal{D}^{m-4}T(4) \to \neg\mathcal{D}^{m-3}T(3)$

[11] An argument of this sort is used in (McGee and McLaughlin, 1998, 224).

For citations go to the published version:
2011: *Vagueness and Language Use*
Paul Egré and Nathan Klinedinst (eds.) Palgrave Macmillan

(GP. 2) $\mathcal{D}\neg\mathcal{D}^{m-3}T(3) \rightarrow \neg\mathcal{D}^{m-2}T(2)$
(GP. 1) $\mathcal{D}\neg\mathcal{D}^{m-2}T(2) \rightarrow \neg\mathcal{D}^{m-1}T(1)$

In order to show that gap principles are consistent under the regional notion of consequence we have to show that there is an interpretation such that the assumptions of the sorites argument (the first element is tall, the last is not tall) and the relevant instances of the relevant gap principles are regionally true. In other words, we have to show that there is an interpretation with a precisification w in it such that the assumptions of the sorites argument plus all the relevant instances of the relevant gap principles take value 1 in every w-admitted precisification. We will give such an interpretation, but before that I would like to consider a remark concerning the existence of absolutely definite cases.

3.1 Absolute definiteness

Suppose that Peter has no hair on his head. Then Peter is bald and, in fact, definitely bald. But is he definitely definitely bald?

Definition 4. (Absolute definiteness) A sentence φ is *absolutely definite* in a precisification w in an interpretation, just in case each member of $\{\mathcal{D}^n\varphi \mid n \in \omega\}$ holds in w. An object a is an absolutely definite positive case of 'F' just in case 'a is F' is absolutely definite; similarly, a is an absolutely definite negative case of 'F' just in case 'a is not F' is absolutely definite.

It seems that, at least for some vague predicates, there are absolutely definite positive and/or negative cases of the predicate. This claim, of course, will depend to some extent on the informal reading of 'definite'. But the claim seems to be fully justified under the supervaluationist reading. In the case of 'bald', for example, it seems that the use we make of the expression forbids us to count as not bald any person without hair on his head. Thus, according to the supervaluationist theory, there is no precisification in which someone without hair on his head is counted as not bald. This means that if Peter has no hair on his head, then the sentence 'Peter is bald' is absolutely definite: you will never reach a precisification through any number of admissibility-steps where this sentence takes value 0. Other examples of predicates that seem to have absolutely definite (positive and/or negative) cases seem to be 'flat', 'empty' or 'open'.

If the previous remark is correct we should impose a further constraint for an acceptable supervaluationist solution. An interpretation showing that gap principles for a vague predicate 'F' are regionally consistent cannot presuppose that the first element of the series is not an absolutely definite positive case of 'F' or that the last is not an absolutely definite negative case of 'F' (since for the supervaluationist there are cases of this sort, this *solution* would be looking at the wrong place).

3.2 Regional consistency of gap-principles

Given regional validity, the minimum number of elements required to show the consistency of gap principles with absolute definite positive and negative cases is 4. Consider the following pocket-sized example. Take the assumptions,

(1) $\{\mathcal{D}^n T(1) \mid n \in \omega\}$
(2) $\{\mathcal{D}^n \neg T(4) \mid n \in \omega\}$
(GP. 1) $\mathcal{DDD}T(1) \rightarrow \neg\mathcal{D}\neg\mathcal{DD}T(2)$
(GP. 2) $\mathcal{DD}T(2) \rightarrow \neg\mathcal{D}\neg\mathcal{D}T(3)$
(GP. 3) $\mathcal{D}T(3) \rightarrow \neg\mathcal{D}\neg T(4)$

An interpretation showing the regional consistency of all the assumptions might look like this,

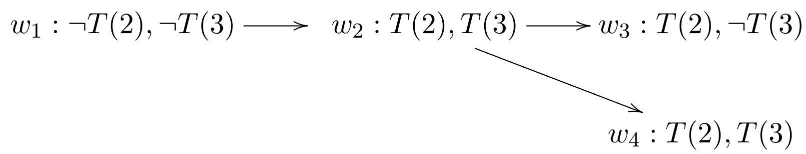

Figure 3: R-consistency of gap-principles.

We assume here (not explicit in the diagram) that $T(1)$ and $\neg T(4)$ hold in every world in the model (this guarantees that (1) and (2) hold in w_1) and also that every world accesses itself (to guarantee the reflexivity of R). (GP. 1) is regionally true at w_1 because the consequent takes value 1 at w_1 and w_2. (GP. 3) is regionally true at w_1 because the antecedent takes value 0 in both w_1 and w_2. Finally (GP. 2) is regionally true at w_1 because the antecedent takes value 0 in w_1 and the consequent value 1 at w_2.

Though the model in figure 3 is just a pocket-sized example, it suffices to show that gap principles can be consistently endorsed under the regional notion of consequence: if we can endorse them for sorites series of just 4 elements, we should be able to do it for larger series.[12]

[12] In fact, a stronger result can be shown. Fara's argument can be run *in the other direction*, this time taking the relevant instances of the schema:

The following question concerns what is the intuition, if any, of why we can accommodate gap-principles given regional validity (and not given global validity). Fara gives the following intuitive explanation of what is going on in her argument (when global validity is at play):

> Grasp the first member of a length-m sorites series for 'tall' in your left hand; grasp the last member in your right hand. To illustrate that there's no 'sharp' boundary between the tall and the not-tall, you want to move your right hand leftward to grasp a different object that is a borderline case of the predicate 'tall' that's true of the object in your left hand but false of the object in your right hand. After one move leftward of your right hand you still have the object in your left hand that is tall, hence definitely tall, and a new object in your right hand that's a borderline case of 'tall', hence not definitely tall. Now to illustrate that there is no sharp boundary between the definitely tall and the not definitely tall, you want to move your right hand leftward again, to grasp an object that's a borderline case of the predicate 'definitely tall' that's true of the object in your left hand but false of the object in your right hand. Each time you do this, you find you have an object in your left hand of which a predicate of the form '$\mathcal{D}^n T(x)$' is true, and an object in your right hand of which that predicate is false. The collection of $m-1$ gap principles appealed to in my argument entail that you can do this at least $m-1$ times. But you cannot do this as many as $m-1$ times; there were only $m-2$ objects between your hands at the start.

Why does this intuitive explanation no longer work if we assume regional rather than global validity? The reason is (of course) connected to the failure of the inference from φ to $\mathcal{D}\varphi$ in the regional case. In Fara's informal explanation, we grasp in our right hand an object a that is a borderline case of the predicate 'definitely^n tall' to illustrate that there is no sharp boundary between the definitely^n tall and the not definitely^n tall members of the series. Now it does not follow that this object is definitely not definitely^n tall (that is, it does not follow that the predicate 'definitely^n tall' is *false* of this object) and so we are not forced to move leftwards to pick out a new object to illustrate the failure of a sharp transition for the predicate 'definitely^{n+1} tall'. The same object might illustrate both that there is no sharp transition

(GP for '$\mathcal{D}^n \neg T$') $\mathcal{D}\mathcal{D}^n \neg T(x) \rightarrow \neg\mathcal{D}\neg\mathcal{D}^n \neg T(x')$

The previous schema was designed to show that there are no sharp boundaries between the positive cases and the others; this schema is intended to show that there are no sharp boundaries between the negative cases and the others. It can be shown that, given regional validity, the instances of both kind of schemas are consistent for finite sorites series. The minimum number of elements to show this is 5.

for the predicate 'definitely^n tall' and for the predicate 'definitely^{n+1} tall' (this would not have been possible if the object were definitely not definitely^n tall). The general idea is that the borderline cases we take to illustrate the absence of sharp transitions need not be definitely borderline and so the same object might be a borderline case of 'definitely^n tall' and a borderline case of 'definitely^{n+1} tall' (that is, the same object might illustrate both the absence of a sharp transition for 'definitely^n tall' and for 'definitely^{n+1} tall').

The way in which regional consequence allows us to address Fara's argument looks to me intuitively appealing. In order to accommodate higher-order vagueness, the truth-value gap theorist is committed to the extistence of an endless hierarchy of borderline cases arranged in a particular order in suitable long sorites series. Now given regional validity, one might endorse a committment of this sort without endorsing the idea that each borderline case in the series is definite. On this view, not everything that smacks of being a borderline case is treated as a clear borderline case[13] and this looks to me intuitively apealing since the borderline status of an object is, very often, something elusive. When talking about vagueness we say things like 'assume a is a clear borderline case of F-ness'; but the truth is that pinning down a particular case that is a clear borderline case of a predicate is usually trickier than what might seem at first sight. The presence of a borderline case is often better witnessed by our disagreement on whether the predicate applies than by our agreement in its being a borderline case. Regional validity has the virtue of allowing us to endorse that something is a borderline case without the commitment of it being a definite borderline case.

References

Blackburn, P., Rijke, M., and Venema, Y. (2001). *Modal Logic*. Cambridge University Press (Reprinted with corrections: 2004).

Burgess, J. A. (2001). Vagueness, epistemicism and response-dependence. *Australasian Journal of Philosophy*, 79(4):507–524.

Cobreros, P. (2008). Supervaluationism and logical consequence: a third way. *Studia Logica*, 90(3):291–312.

Cobreros, P. (2009). Varzi on supervaluationism and logical consequence. *Mind*, forthcoming.

Edgington, D. (1993). Wright and Sainsbury on higher-order vagueness. *Analysis*, 53(4):193–200.

[13] See (Fine, 1975, 297)

Fara, D. G. (2003). Gap principles, penumbral consequence, and infinitely higher-order vagueness. *Liars and Heaps: New Essays on Paradox*, pages 195–221. Originally published under the name 'Delia Graff'.

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

Heck, R. (1993). A note on the logic of (higher-order) vagueness. *Analysis*, 53(4):201–208.

Keefe, R. (2000a). Supervaluationism and validity. *Philosophical Topics*, 28:93–106.

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

McGee, V. and McLaughlin, B. (1998). Review of *Vagueness*. *Linguistics and Philosophy*, 21:221–235.

Sainsbury, M. (1991). Is there higher-order vagueness? *Philosophical Quarterly*, 41(163):167–182.

Weatherson, B. (2003). Epistemicism, parasites and vague names. *Australasian Journal of Philosophy*, 81(2):276–279.

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

Williamson, T. (1999). On the structure of higher-order vagueness. *Mind*, 108:127–143.

Williamson, T. (2004). Reply to McGee and McLaughlin. *Linguistics and Philosophy*, 27:113–122.

Wright, C. (1987). Further reflexions on the sorites paradox. *Philosophical Topics*, 15:227–290.

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