Paraconsistent Vagueness: A Positive Argument

June 17, 2010

Abstract

Paraconsistent approaches have received little attention in the literature on vagueness (at least compared to other proposals). The reason seems to be that many philosophers have found the idea that a contradiction might be true (or that a sentence and its negation might both be true) hard to swallow. Even advocates of paraconsistency on vagueness do not look very convinced when they consider this fact; since they seem to have spent more time arguing that paraconsistent theories are at least as good as their *paracomplete* counterparts, than giving positive reasons to believe on a particular paraconsistent proposal. But it sometimes happens that the weakness of a theory turns out to be its mayor ally, and this is what (I claim) happens in a particular paraconsistent proposal known as subvaluationism. In order to make room for truth-value *gluts* subvaluationism needs to endorse a notion of logical consequence that is, in some sense, weaker than standard notions of consequence. But this *weakness* allows the subvaluationist theory to accommodate higher-order vagueness in a way that it is not available to other theories of vagueness (such as, for example, its paracomplete counterpart, *supervaluationism*).

The subvaluationist theory of vagueness is the dual theory of the well-known supervaluationist theory. Where the supervaluationist reads 'truth' as 'supertruth' (truth in every precisification) the subvaluationist reads 'truth' as 'subtruth' (truth in some precisification). This dual reading of the notion of truth leads to a theory of vagueness in which borderline sentences give raise to *gluts* of truth-value (by contrast to supervaluationism in which borderline sentences give raise to *gaps* in truth-value). Subvaluationism is a paraconsistent theory in the sense that a sentence might both be true and false without triviality (that is, the set of sentences $\{p, \neg p\}$ is subvaluationist-satisfiable); it is weakly paraconsistent in the sense that classical contradictions are not subvaluationist-satisfiable (the analogous *dual* remarks apply to supervaluationism).

The subvaluationist theory have been defended by Dominic Hyde (in Hyde (1997) and in the more recent Hyde and Colyvan (2008)) who exploits the duality between subvaluationism and supervaluationism to argue that the first is at least as good as the second and, consequently, the neglect of paraconsistent theories in the literature lacks a justification. Commenting on Hyde's 1997 paper Beall and Colyvan (2001) point out that Hyde could have gone further arguing that truth-value gluts seem to have the upper hand in the case of paradoxes other than the sorites, such as the paradoxes of self-reference in which truth-value gluts, unlike truth-value gaps, do not succumb to strengthened versions.

Even if truth-value glut theories have the upper hand in the case of self-referential paradoxes, this might not constitute enough justification for a glut solution in the case of vagueness. There are other phenomena that suggest a *gappy* treatment and there's no forthcoming argument for the claim that gluts solve things everywhere. It seems to me that paraconsistent proposals on vagueness, subvaluationism in particular, deserve a positive argument, a justification on its own not parasitic on a paracomplete dual. This paper provides such an argument based on Fara's (so-called) *paradox of higher-order vagueness*. Fara (2003) shows that if the supervaluationist is committed to a rule of inference known as $\mathcal{D}$-introduction, then she/he cannot endorse the complete hierarchy of *gap-principles* needed to explain the seeming absence of sharp transitions in sorites series. This paper argues that these *gap-principles* are equally compelling to other theories of vagueness in which the notion of a borderline case plays a key role. Then it shows that these theories, if committed to a notion of logical consequence as strong as local consequence, cannot endorse a *strengthened* version of the paradox. But the subvaluationist can.

The paper is divided into three sections. The first one describes in a general way what it is understood by a *borderline-based* theory of vagueness and proposes a general setting to define a notion of definiteness for this sort of theories. Different informal readings of the general setting will render different notions of truth for each theory and, consequently, different notions of logical consequence. The paper considers three alternatives: supervaluationist, local and subvaluationist consequence. The second section presents Fara's paradox of higher-order vagueness as applied to the supervaluationist theory. The last section considers a strengthened version of Fara's paradox and explains why theories committed to a notion of logical consequence as strong as local consequence cannot handle this version of the paradox while the subvaluationist theory can.

1 A general setting for borderline-based theories

1.1 Borderline-based theories of vagueness

Most theories of vagueness take the notion of *borderline case* as a central one in the explanation of the phenomenon of vagueness.[1] The general idea is that a vague predicate such as 'bald' lead to a situation in which competent speakers refuse to classify certain people as 'bald' and refuse at the same time to classify them as 'not bald'. In this sense, borderline cases of a given vague predicate are objects to which the predicate meaningfully applies but such that competent speakers manifest a kind of symmetry in their dispositions to apply the predicate or its negation. This rough characterization of borderline cases intends to remain neutral among several different interpretations. We might read borderline cases in epistemic terms, in which case the symmetry in our dispositions to assent to the sentence 'Tim is thin' and to assent to the sentence 'Tim is not thin' is a manifestation of a particular sort of ignorance associated to vague expressions. In the truth-value gap reading, the symmetry manifests the fact that each sentence lacks a truth-value and in the subvaluationist the fact that both sentence are equally true (and false).

An adequate theory of vagueness should provide an explanation of the sorites paradox. Consider a long sorites series for the predicate 'tall'. The first member of the series is 2.5 meters tall, the last is 1.5 tall and each member differs from its successor in the series by less than a millimeter. The paradox originates because though the first element is clearly tall and the last is clearly not tall, there seems to be no sharp transition between the elements that are tall to those that are not tall. At this point a borderline-based theory will make use of borderline cases (reading this notion in the particular way of the theory). The supervaluationist will explain that there's actually no such sharp transition, since among the poles of the series there are cases that are neither truly tall nor falsely tall (that is, there are borderline cases, reading this notion in supervaluationist terms). The epistemicist bites the bullet claiming that there's such a sharp transition, but explains that we cannot know where the transition lies since there are cases in the series of which we cannot know whether they are tall (that is, there are borderline cases, reading this notion in epistemic terms).

[1] Remarkable exceptions are Fara (2000) and Zardini (2008).

1.2 A general setting

Since the notion of a borderline case plays a key role in these theories, it is natural for them to consider a notion of *definiteness* to speak about borderline cases. Several of these proposals share the common framework of a possible-worlds semantics in which we can define the notion of definiteness.[2] An interpretation for a propositional language with an operator for definiteness ('$\mathcal{D}$' henceforth) is a triple $\langle W, R, \nu \rangle$ where, W is a non-empty set of *worlds*, R is a relation in W and ν is a truth-value assignment to sentences at worlds such that classical operators are defined classically (though things are relative to worlds) and the $\mathcal{D}$-operator is defined as the modal operator for necessity, that is,

$$\nu_w(\varphi \to \beta) = 1 \text{ iff either } \nu_w(\varphi) = 0 \text{ or } \nu_w(\beta) = 1$$

$$\nu_w(\neg\varphi) = 1 \text{ iff } \nu_w(\varphi) = 0$$

$$\nu_w(\mathcal{D}\varphi) = 1 \text{ iff } \forall w' \text{ such that } wRw'\ \nu_{w'}(\varphi) = 1^3$$

At this stage the difference between theories concerns the different informal readings of the semantics. Epistemicism will read worlds as some sort of *epistemic possibilities*, contextualism as contexts of utterance and supervaluationism and subvaluationism as precisifications. Each informal reading of the semantics motivates, however, different readings of the relevant notion of truth for the theory. Supervaluationism is associated to *supertruth* where φ is supertrue in an interpretation at a world w just in case it takes value 1 at every world. Epistemicism and contextualism are associated to *local truth* where φ is locally true in an interpretation at a world w just in case it takes value 1 at w. Finally, subvaluationism is committed to *subtruth* where φ is subtrue in an interpretation at a world w just in case it takes value 1 at some w. Since logical consequence is a matter of necessary preservation of truth, different commitments on the notion of truth render different commitments on logical consequence.

Definition 1 (Local consequence)**.** A sentence φ is a local consequence of Γ, written $\Gamma \vDash_l \varphi$, just in case for every interpretation and world w in the interpretation: if every member of Γ is true at w then φ is true at w.

Definition 2 (Supervaluationist consequence)**.** A sentence φ is a supervaluationist consequence of Γ, written $\Gamma \vDash_{SpV} \varphi$, just in case for every interpretation: if every member of Γ is true at every world then φ is true at every world.

[2]The use of possible world semantics for different theories of vagueness is now completely standard. See for example Williamson (1994) and **?**

[3]For readability I will often write 'φ is true at w' or 'φ holds at w' instead of '$\nu_w(\varphi) = 1$'.

Definition 3 (Subvaluationist consequence)**.** A sentence φ is a subvaluationist consequence of Γ, written $\Gamma \vDash_{SbV} \varphi$, just in case for every interpretation: if every member of Γ is true at some world then φ is true at some world.

Note that the standards for satisfaction set by each notion of truth differ in strength. It is *harder* to supervaluationist-satisfy a set of sentences Γ than to locally satisfy it. Since if Γ is supervaluationist-satisfied in an interpretation, then it certainly is locally satisfied in that interpretation (if every $\gamma \in \Gamma$ hold everywhere in the interpretation then, for any w in that interpretation Γ holds locally). In turn, it is *harder* to locally satisfy a set of sentences Γ than to subvaluationist-satisfy it. Since if Γ is locally satisfied in an interpretation, it certainly is subvaluationist-satisfied in that interpretation (if Γ holds at a particular world w in an interpretation, then each $\gamma \in \Gamma$ hold somewhere in that interpretation). So subtruth is the weakest standard for satisfaction. This fact is crucial for our discussion on the satisfiability of gap-principles below.

Local consequence preserves local truth in every interpretation; it is, therefore, a well-defined notion of consequence for theories committed to local truth. In a similar manner, supervaluationist consequence is well defined for supervaluationism if this theory is committed to supertruth[4] and subvaluationist consequence for the subaluationist theory.

Local consequence is the standard definition of logical consequence for the simple modal language.[5] Every locally valid argument is supervaluationist-valid but not the other way (in particular, $\varphi \vDash_{SpV} \mathcal{D}\varphi$ but $\varphi \nvDash_{l} \mathcal{D}\varphi$). On the other hand not every locally valid argument is subvaluationist-valid (since $\{\varphi, \neg\varphi\} \vDash_{l} \bot$ but $\{\varphi, \neg\varphi\} \nvDash_{SbV} \bot$) nor every subvaluationist-valid argument

[4]The use of the expression 'supervaluationist consequence' might look tendentious since not all authors agree on whether supervaluationism is committed to this form of consequence. *Canonical* supervaluationism identifies truth with supertruth and is (naturally) committed to this notion of consequence (for example (Fine, 1975, 290) and (Keefe, 2000, 176)). Some authors, however, hold a supervaluationist-like treatment of vagueness but endorse local consequence instead (for example Varzi (2007) and Asher et al. (2009)). The author of this paper holds that supervaluationism is committed to a third alternative between local and global consequence (in XXX and XXX). In order to concentrate on the arguments in this paper I leave the discussion of this issue for a better occasion, but just two quick remarks on the strategy of linking supervaluationism to local consequence. First, supertruth allows for truth-value gaps while local truth does not (in each world in every interpretation a sentence is either locally true or locally false); thus, linking supervaluationism to local consequence jeopardizes either the truth-value gap explanation of vagueness or the idea that logical consequence is a matter of necessary preservation of truth. Second, as we shall see, local consequence does not escape from the *strengthened version* of Fara's paradox, so even if supervaluationism is committed to local consequence, it is still in trouble to handle the issue of higher-order vagueness.

[5]See (Blackburn et al., 2001, 31)

is locally valid (since $\neg\mathcal{D}\varphi \vDash_{SbV} \neg\varphi$ but $\neg\mathcal{D}\varphi \nvDash_{l} \neg\varphi$).[6]

2 Fara's paradox of higher-order vagueness

In her 2003 paper Delia Graff Fara presents an argument against truth-value gap theories (supervaluationism in particular) concerning higher-order vagueness. Think again in our sorites series for the predicate 'tall'. The first member of the series is clearly tall and the last is clearly not tall, but there seems to be no sharp transition from the members of the series that are tall to those that are not tall. A truth-value gap theory of vagueness (supervaluationism in particular) explains this fact appealing to the presence of a truth-value gap in between; there seems to be no sharp transition from the tall to the not tall for the simple reason that there's no such a sharp transition. Members of the series that are truly tall are not immediately followed by members that are falsely tall; that is, the truly tall and the falsely tall are separated by a truth-value gap, making true the following *gap-principle*:

(GP for 'tall') $\mathcal{D}\text{tall}(x) \to \neg\mathcal{D}\neg\text{tall}(x')$ (where x' is the successor of x)

However, the problem of vagueness does not stop at that point since there seems to be no sharp transition either from the truly tall to the non-truly tall. Given that the phenomenon seems to be of the same kind, the explanation must be carried out in analogous terms, that is, the truly tall and the non-truly tall are separated by a gap making true the following gap-principle:

(GP for '$\mathcal{D}$tall') $\mathcal{D}\mathcal{D}\text{tall}(x) \to \neg\mathcal{D}\neg\mathcal{D}\text{tall}(x')$

Non-terminating higher-order vagueness should be treated, at least, as a logical possibility rendering all the gap-principles of this form,

(GP for '$\mathcal{D}^n$tall') $\mathcal{D}\mathcal{D}^n\text{tall}(x) \to \neg\mathcal{D}\neg\mathcal{D}^n\text{tall}(x')$

In addition to the truth of each gap-principle, Fara claims that a truth-value gap theory (supervaluationism in particular) is committed to the following inference rule:

$\mathcal{D}$-introduction: $\varphi \vdash \mathcal{D}\varphi$[7]

[6] See the Appendix for a more detailed account of the relations between these notions of consequence.

[7] The rule of $\mathcal{D}$-introduction must not be confused with the *Necessitation* rule of normal modal logics. $\mathcal{D}$-introduction can be generally stated as: $\Gamma \vdash \varphi \implies \Gamma \vdash \mathcal{D}\varphi$ while the *Necessitation* rule can be stated this way: $\Gamma \vdash \varphi \implies \mathcal{D}(\Gamma) \vdash \mathcal{D}\varphi$, where $\mathcal{D}(\Gamma)$ is $\{\mathcal{D}\gamma \mid \gamma \in \Gamma\}$.

The reason for this commitment is the fact that for the truth-value gap theorist '$\mathcal{D}$' means something like 'It is true that', in which case it looks impossible for a sentence φ to be true while the sentence saying that φ is true (i. e. '$\mathcal{D}\varphi$') is not (Fara, 2003, 199-200). But Fara provides an argument to show that one cannot consistently hold both the commitment to gap-principles and the commitment to this rule of inference. Her argument is as follows.

Suppose we have a finite sorites series of m elements. The first element is clearly tall, the last is clearly not tall. The difference between each adjacent pair of members of the series is small enough to justify the truth of each instance of any of the previously mentioned gap-principles. Now from the fact that the last element, m, is not tall it follows, by $\mathcal{D}$-introduction, that it is definitely not tall: $\mathcal{D}\neg\text{tall}(m)$. But consider the following instance of the gap-principle for 'tall':

(GP for 'tall') $\mathcal{D}\text{tall}(m-1) \rightarrow \neg\mathcal{D}\neg\text{tall}(m)$

Making use of the contrapositive form and of *Modus ponens*, from the fact that $\mathcal{D}\neg\text{tall}(m)$ it follows that $\neg\mathcal{D}\text{tall}(m-1)$. Making use of $\mathcal{D}$-introduction again we obtain: $\mathcal{D}\neg\mathcal{D}\text{tall}(m-1)$. And again we use an instance of a gap-principle, this time for '$\mathcal{D}$tall':

$\mathcal{D}\mathcal{D}\text{tall}(m-2) \rightarrow \neg\mathcal{D}\neg\mathcal{D}\text{tall}(m-1)$

As before, we obtain $\neg\mathcal{D}\mathcal{D}\text{tall}(m-2)$ using the contrapositive form of the gap-principle and *Modus ponens*. Making use of the $m-1$ relevant instances of the relevant gap-principles we can construct an argument showing that $\neg\mathcal{D}^{m-1}\text{tall}(1)$. But this contradicts our first assumption that the first element of the sorites is tall since, by $m-1$ applications of $\mathcal{D}$-introduction it should be $\mathcal{D}^{m-1}$ times tall. The argument is graphically explained in Figure 1.

As Fara points out her argument is inspired on an argument of Wright's (Wright (1987) and Wright (1992)). Wright's original argument makes use of a *second-order* gap-principle: $\forall x(\mathcal{D}\mathcal{D}T(x) \rightarrow \neg\mathcal{D}\neg\mathcal{D}T(x'))$ from which he derives $\forall x(\mathcal{D}\neg\mathcal{D}T(x') \rightarrow \mathcal{D}\neg\mathcal{D}T(x))$, with the aid of (a weakened form of) $\mathcal{D}$-introduction. The derived sentence is as paradoxical as the original induction premise of the sorites argument. Thus, the argument poses a difficulty to accommodate higher-order vagueness for theories committed to the rule of $\mathcal{D}$-introduction (or to a slightly weakened version). The problem with Wright's argument is that, though the rule of $\mathcal{D}$-introduction is supervaluationist-valid, the argument itself is not supervaluationist-valid. The reason is that in the presence of $\mathcal{D}$-introduction some classically valid rules of inference are not unrestrictedly valid (Williamson, 1994, 151-2). In particular, we cannot discharge premises with rules like conditional proof

$$
\begin{tikzcd}
2.5\text{ m.} & \cdots & \cdots & \cdots & 1.5\text{ m.} \\
\text{tall}(1) \arrow[ddddd] & & & & \neg\text{tall}(m) \arrow[d] \\
 & & & \neg\mathcal{D}T(m-1) \arrow[d] & \mathcal{D}\neg T(m) \arrow[l] \\
 & & \neg\mathcal{D}^2T(m-2) \arrow[d] & \mathcal{D}\neg\mathcal{D}T(m-1) \arrow[l] & \\
 & & \mathcal{D}\neg\mathcal{D}^2T(m-2) & & \\
 & [\ldots] & & & \\
\mathcal{D}^{m-1}T(1) & \neg\mathcal{D}^{m-1}T(1) & & &
\end{tikzcd}
$$

Figure 1: Fara's argument

and *reductio* when we have used $\mathcal{D}$-introduction in the sub-proofs (that is, when we have applied $\mathcal{D}$-introduction to the premises or to anything deduced from them), as Wright does.[8]

What makes Fara's argument particularly interesting is that the rules used in her argument are always supervaluationist-valid. In particular, Fara makes use of the contrapositive forms of the relevant instances of the relevant gap-principles but by supervaluationist standards, a conditional and its contrapositive form are logically equivalent. Then Fara combines $\mathcal{D}$-introduction with *Modus ponens* but (unlike conditional proof or *reductio*) this rule remains supervaluationist-valid in the presence of $\mathcal{D}$-introduction. In short, what makes Fara's argument interesting is that it is actually a *a proof* showing that gap-principles are supervaluationist-unsatisfiable for finite sorites series.[9] Thus, Fara's argument shows that the supervaluationist has problems with his own explanation of the seeming absence of sharp transitions in sorites series. If the supervaluationist is really committed to what we characterized as supervaluationist consequence above, then she/he is in a rather desperate position to respond to Fara's challenge.

[8]This observation concerning Wright's argument is to be found in Heck (1993). Further discussion on the argument in Edgington (1993) and Sainsbury (1991)

[9]The Appendix outlines a tableaux-based proof system for the notions of logical consequence discussed in this paper and shows, making use of this brute-force procedure, that gap-principles are supervaluationist-unsatisfiable.

3 Strengthening the paradox

3.1 Local consequence and gap-principles

Though Fara's objection is intended against the supervaluationist theory, the principles appealed to in the paradox are compelling to other borderline-based theories of vagueness (reading '$\mathcal{D}$' in the particular way of the theory). The general idea is that for a borderline-based theory of vagueness the presence of borderline cases among, say, the clearly tall and the clearly not tall is a constitutive part of the phenomenon of vagueness (reading the notion of *borderlineness* in the particular way of the theory). But when one considers the vagueness of '$\mathcal{D}$-tall' one is compelled to a borderline-based explanation claiming that there are borderline cases among the clearly $\mathcal{D}$-tall and the clearly not $\mathcal{D}$-tall (reading the notion of *borderlineness* in the particular way of the theory). And so the story goes, justifying the truth of each gap-principle, reading the notion of definiteness involved in those principles in each theory's preferred way. For example, for the epistemicist there is a sharp transition from tall members to non-tall members of the series. However, we cannot know where this transition lies because there are members of the series of which we cannot know whether they are tall (that is, there are borderline cases of 'tall', reading the notion of *borderlineness* in epistemic terms). In a similar way, we cannot know where the transition from the definitely tall to the non-definitely tall lies, since there are borderline cases, this time for 'definitely tall'...

The situation for theories committed to local consequence is not, perhaps, as desperate as the situation for theories committed to supervaluationist consequence since we might consistently handle gap-principles given local consequence. Fara's argument works, as pointed out in section 2, based on the relevant instances of the relevant gap-principles. To take a pocketsize example, imagine a sorites series of 3 elements: $\langle 1, 2, 3\rangle$. The first element is clearly tall, the last is clearly not tall. The relevant instances of the relevant gap-principles,

(GP 1*) $\mathcal{D}\text{tall}(2) \to \neg\mathcal{D}\neg\text{tall}(3)$

(GP 2*) $\mathcal{D}\mathcal{D}\text{tall}(1) \to \neg\mathcal{D}\neg\mathcal{D}\text{tall}(2)$

should be true. Diagram in Figure 2 depicts an interpretation in which all this holds.

The first element of our "sorites" series is clearly tall and the last is clearly not tall. (GP 1*) takes value 1 at w_0 since the antecedent takes value 0 at w_0. (GP 2*) takes value 1 at w_0 since the consequent takes value 1 at w_0. This example shows that the relevant instances of relevant gap-

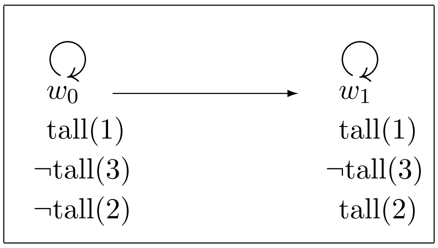

Figure 2: Local satisfaction of gap-principles

principles are locally satisfiable for any given finite sorites series (of at least three elements). Note, however, that (GP 1*) takes value 0 in w_1. This is not a particular feature of this model; rather, Fara's argument shows that gap-principles are not supervaluationist-satisfiable and so any interpretation showing the local satisfiability of gap-principles will contain some world at which some gap-principle is false. We exploit this fact in the following section to formulate a strengthened version of the paradox.

3.2 The strengthened paradox

Definition 4 (Absolute definiteness)**.** The absolute definitization of a sentence φ is the set $\{\mathcal{D}^n\varphi \mid n \in \omega\}$.

The idea is that the absolute definitization of φ is the set containing φ and $\mathcal{D}\varphi$, and $\mathcal{D}\mathcal{D}\varphi$ etc. In the present setting, for example, the absolute definitization of a classically valid sentence is valid in either reading of our three notions of consequence. Making use of this notion of absolute definiteness, we can establish the following connection between supervaluationist and local consequence:

Claim 1. If $\Gamma \vDash_{SpV} \varphi$ then $\{\mathcal{D}^n\gamma \mid \gamma \in \Gamma, n \in \omega\} \vDash_l \varphi$

That is, φ is a supervaluationist consequence of Γ just in case φ is a local consequence of the absolute definitization of the elements in Γ.[10]

For any finite sorites series of m elements, let Λ be the set of premises of Fara's argument for that series. That is, for any such series, Λ contains 'tall(1)', '¬tall(m)' plus all the instances of the relevant gap-principles used in the argument. We will call to the absolute definitization of the elements of Λ, that is $\{\mathcal{D}^n\lambda \mid \lambda \in \Lambda, n \in \omega\}$, the *strengthened paradox of higher-order vagueness*. Since for a finite sorites series of m elements $\Lambda \vDash_{SpV} \bot$, then, by Claim 1 above, the absolute definitization of the elements in Λ *locally* entail a contradiction. That is, theories committed to a notion of logical

[10]A sketch of the proof of Claim 1 might be found in the Appendix.

consequence as strong as local consequence succumb to the strengthened version of Fara's paradox of higher-order vagueness.[11]

How bad is this last result? Surely, it is not as bad as lacking even the possibility of accepting gap principles plainly (as it happens in the case of supervaluationist consequence). But it seems to me that the result is bad enough. As pointed out before, gap principles are compelling for each borderline-based theory of vagueness (reading '$\mathcal{D}$' in the preferred way of the theory). Now since for each theory the justification of the truth of gap-principles follows from considerations of the theory itself, claiming that gap-principles are not absolutely definite amounts to claiming that the theory itself is not absolutely definite (reading 'definite' in each theory's preferred way).[12] For example, for the epistemicist, it wouldn't be absolutely knowable whether the theory is right (and the theory going wrong should be always an epistemic possibility). For the contextualist reading, more dramatically I think, there must be contexts where some gap-principles are false, and thus, there are contexts where the theory itself goes wrong.

At this point, the weakness of subvaluationist's logic turns out to be a great advantage. In order the make room for truth-value *gluts*, logical consequence in the subvaluationist theory must be weaker than either local or supervaluationist consequence (since $\{p, \neg p\} \vDash_l \bot$ but $\{p, \neg p\} \nvDash_{SbV} \bot$). The set of sentences $\{p, \neg p\}$ is $\vDash_{SbV}$-satisfiable for, in a given interpretation p might hold in w and $\neg p$ might hold in a different w' (and, of course, this is what $\vDash_{SbV}$-satisfiability requires). We already know that the set of premises in Fara's argument supervaluationist-entail a contradiction; by the duality of $\vDash_{SpV}$ and $\vDash_{SbV}$ (see the Appendix) a valid sentence subvaliationist-entails the negation of some of the premises. This means that for any interpretation, some gap-principle will be *subfalse*. Still, since for the supervaluationist 'being true' means 'being true somewhere' this does not show that gap-

[11] Wright's argument in (Wright, 2010, 537-9) is based on similar grounds. In their recent paper Asher et al. point out to the fact that we can construct an argument not appealing to $\mathcal{D}$-introduction but to the definiteness of gap-principles instead (Asher et al., 2009, 915). However, this conflicts with their observation according to which 'gap-principles can be made determinately-to-the-n true for any n' (Asher et al., 2009, 924) (in fact, the second diagram in p. 924 fails to show that the definitization of the first gap-principle (their **G1**) holds). The tableaux-based proof of the supervaluationist-unsatisfiability of gap-principles in the Appendix shows that there is a systematic connection among *how definite* we need gap-principles to be to run the argument and the length of the sorites series (for a sorites series of m elements we need just a limited number of iterations of $\mathcal{D}$ attached to each gap-principle for the argument to work).

[12] It is commonly accepted that definiteness is closed under logical consequence in the sense that if $\Gamma \vDash \varphi$ then $\{\mathcal{D}\gamma \mid \gamma \in \Gamma\} \vDash \mathcal{D}\varphi$. Thus, contrapositively, if φ is not definite, then some of the γ's is not definite. The justification of the truth of gap-principles is not, perhaps, a strict logical consequence from the theory itself; but these theories aim to provide a logic of definiteness for our everyday notion of consequence and in this sense the claim in the text is well motivated.

principles are not $\vDash_{SbV}$-satisfiable (a principle might be subfalse at a world but subtrue at some different world). In fact the absolute definitization of the premises in Fara's argument is $\vDash_{SbV}$-satisfiable. Consider again our pocketsize example with a sorites series of three elements $\langle 1, 2, 3 \rangle$:

(GP 1*) $\mathcal{D}$tall(2) $\rightarrow \neg\mathcal{D}\neg$tall(3)

(GP 2*) $\mathcal{D}\mathcal{D}$tall(1) $\rightarrow \neg\mathcal{D}\neg\mathcal{D}$tall(2)

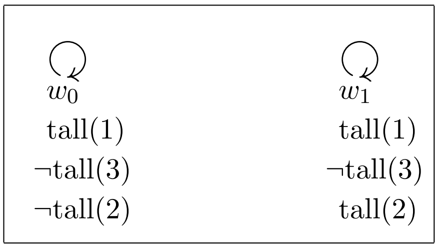

Figure 3: Sub-satisfaction of absolutely definite gap-principles

Figure 3 shows that (GP 1*) holds in w_0 since, as before, the antecedent is false at w_0. Now this time, the absolute definitization of (GP1*) holds in w_0 since it holds in every w accessible from w_0 (namely, w_0 itself). Similarly, the absolute definitization of (GP2*) holds in w_1 since it holds in every w_1-accessible (namely, w_1 itself). Thus, the absolute definitization of the relevant instances of relevant gap-principles for a sorites series of three elements plus the absolute definitization of the assumption that the first element is tall and that the last is not is subvaluationist-satisfiable. More generally, we might show that the absolute definitization of the relevant instances of relevant gap-principles for a finite sorites series is subvaluationist-satisfiable constructing a model in which there is a world for each gap-principle and where these worlds relate only to themselves. At each of these worlds, of course, the absolute definitization of some other gap-principle will take value 0 but, again, showing that the absolute definitization of the premises is $\vDash_{SbV}$-satisfiable requires just showing that, for each φ in that set, there is at least a w at which φ takes value 1.

Conclusion

The capability of endorsing the absolute definitization of gap-principles looks like an appealing feature for any borderline-based theory of vagueness; but this capability is restricted to theories committed to a weak enough notion of logical consequence. If vagueness is to be explained in terms of borderline cases, the foregoing results constitute a good argument in favor of the subvaluationist theory of vagueness. Some philosophers will still find the commitment to parconsistency as something hard to swallow and will probably consider that the result speaks against the whole borderline-based approach to vagueness. These philosophers think that, as Williamson says, 'dialetheism is a fate worse than death' (Williamson, 2006, p. 387)). To these I find appropriate Priest's own response to Williamson: 'I haven't died yet, so I'm not in a position to judge.' (Priest (2007)).

References

Asher, N., Dever, J., and Pappas, C. (2009). Supervaluations debugged. *Mind*, 118(472):901–933.

Beall, J. C. and Colyvan, M. (2001). Heaps of gluts and Hyde-ing the sorites. *Mind*, 110(438):401–408.

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

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

Fara, D. G. (2000). Shifting sands: An interest-relative theory of vagueness. *Philosophical Topics*. Orignally published under the name 'Delia Graff', 28(1):45–81.

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

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

Heck, R. (1993). A note on the logic of (higher-order) vagueness. *Analysis*.

Hyde, D. (1997). From heaps and gaps to heaps of gluts. *Mind*, 106(424):641–660.

Hyde, D. and Colyvan, M. (2008). Paraconsistent vagueness: Why not? *Australasian Journal of Logic*, 6:107–121.

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

Priest, G. (2007). Review of *Absolute Generality*. *Notre Dame Philosophical Reviews*.

Priest, G. (2008). *An Introduction to Non-Classical Logic: From If to Is*. Cambridge University Press.

Rayo, A. and Uzquiano, G. (2006). *Absolute Generality.* Oxford University Press.

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

Varzi, A. (2007). Supervaluationism and its logic. *Mind*, 116(463):633–676.

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

Williamson, T. (2006). Absolute identity and absolute generality. En Rayo y Uzquiano (eds.).

Wright, C. (1987). Further reflections on the sorites paradox. *Philosophical Topics*, 15(11):227–290.

Wright, C. (1992). Is higher-order vagueness coherent? *Analysis*, 52(3):129–139.

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

Zardini, E. (2008). A model of tolerance. *Studia Logica*, 90(3):337–368.

Appendix

Connection between supervaluationist and local consequence

We give now a sketch of the proof for Claim 1:

Proposition 1. If $\Gamma \vDash_{SpV} \varphi$ then $\{\mathcal{D}^n\gamma \mid \gamma \in \Gamma, n \in \omega\} \vDash_l \varphi$

The proof is based on the general fact that for any interpretation showing $\{\mathcal{D}^n\gamma \mid \gamma \in \Gamma, n \in \omega\} \nvDash_l \varphi$ there is a corresponding *generated submodel* showing $\Gamma \nvDash_{SpV} \varphi$ (see (Blackburn et al., 2001, 56) for details on generated submodels).

Proof sketch. Assume that $\{\mathcal{D}^n\gamma \mid \gamma \in \Gamma, n \in \omega\} \nvDash_l \varphi$; then there is a model with a world w_0 such that φ takes value 0 at w_0 and every member of Γ takes value 1 at every w accessible from w_0 *in any number of R-steps* (including w_0 itself). The only way in which this model might fail to show that $\Gamma \nvDash_{SpV} \varphi$ is by containing worlds that are not accessible from w_0 in any number of R-steps at which some of the members of Γ take value 0. However, since these worlds are not accessible from w_0 in any number of R-steps, they are irrelevant for the truth-value of sentences in w_0. So define a model $\langle W', R', \nu' \rangle$ such that W' is $\{w | w_0 R^m w\} \cup \{w_0\}$ and R', ν' the restrictions of R, ν to W'. It can be proved by induction over the set of wff that ν and ν' agree in the truth-values of sentences in w_0. Since $\{\mathcal{D}^n \gamma \mid \gamma \in \Gamma, n \in \omega\}$ holds in w_0 each γ in Γ take value 1 at *any* world in W'. In turn, since φ takes value 0 at w_0, there is at least a world at which it takes value 0. So the new model is an interpretation showing $\Gamma \nvDash_{SpV} \varphi$. $\square$

Some relations between $\vDash_l$, $\vDash_{SpV}$ and $\vDash_{SbV}$

For the classical propositional language (without modal operators) and for single-conclusions $\vDash_l$ coincides with $\vDash_{SpV}$. Since for the classical propositional language the valid arguments of $\vDash_l$ are those of classical logic, this result amounts to the coincidence between supervaluationist and classical logic for a language without $\mathcal{D}$ or *similar operators* (see (Fine, 1975, 283-4) and (Keefe, 2000, 175-6)). The thing changes when we consider multiple conclusions since, classically, the truth of $\{\varphi \vee \psi\}$ guarantees the truth of some in $\{\varphi, \psi\}$ (this rule sometimes named 'subjunction') but we have that $\{\varphi \vee \psi\} \nvDash_{SpV} \{\varphi, \psi\}$ (since the disjunction might be supertrue without either disjunct being supertrue).

For the classical propositional language (without modal operators) and for single-conclusions $\vDash_{SbV}$ is strictly weaker than $\vDash_l$. On the one hand, not every locally valid argument is subvaluationist-valid since $\{\varphi, \neg\varphi\} \vDash_l \bot$ but $\{\varphi, \neg\varphi\} \nvDash_{SbV} \bot$ (since φ and $\neg\varphi$ might both being sub-true in the same interpretation). On the other hand, every subvaluationist-valid argument is locally valid since if $\Gamma \nvDash_l \varphi$ then there is an interpretation with a world w at which every member of Γ takes value 1 and φ value 0; since there are no modal operators, the interpretation consisting of w as the single world in W is an interpretation showing $\Gamma \nvDash_{SbV} \varphi$ (this reason is actually the same invoked to show that supervaluationist and local consequence coincide for the classical vocabulary and single conclusions). Looking at the multiple-conclusions case we have that subvaluationist consequence coincides with classical consequence for single-premise arguments (in the same way in which supervaluationist consequence coincides with classical logic for single-conclusion arguments) but not for multiple-premises arguments; for example $\{\varphi, \psi\}$ classically entails $\varphi \wedge \psi$ (this rule sometimes named 'adjunction') but we have that $\{\varphi, \psi\} \nvDash_{SbV} \{\varphi \wedge \psi\}$ (since each of φ and ψ might be sub-true even in cases in which the conjunction is not). These facts are linked to the *duality* of '$\vDash_{SpV}$' and '$\vDash_{SbV}$'.

For the simple modal language (i. e., for a language with '$\mathcal{D}$') supervaluationist consequence is strictly stronger than local consequence. Every locally valid argument is supervaluationist-valid, but not the other way. In particular, as we already know, the inference from φ to $\mathcal{D}\varphi$ is supervaluationist-valid. Thus, for this language, not every supervaluationist-valid argument is locally valid. In a similar way, for this language, not every subvaluationist-valid argument is locally valid since $\neg\mathcal{D}\varphi \vDash_{SbV} \neg\varphi$ but this is not the case for local consequence.

The duality of '$\vDash_{SpV}$' and '$\vDash_{SbV}$' can be fully expressed extending the definitions of logical consequence to the multiple-conclusions case:

Definition 5 (Supervaluationist consequence: multiple conclusions)**.** A set of sentences Δ is a supervaluationist consequence of a set of sentences Γ, written $\Gamma \vDash_{SpV} \Delta$, just in case for every interpretation: if every member of Γ takes value 1 at every world then some member of Δ takes value 1 at every world.

Definition 6 (Subvaluationist consequence: multiple conclusions)**.** A set of sentences Δ is a subvaluationist consequence of a set of sentences Γ, written $\Gamma \vDash_{SbV} \Delta$, just in case for every interpretation: if every member of Γ takes value 1 at some world, then some member of Δ takes value 1 at some world.

Now for a given set Γ let $\neg(\Gamma)$ be $\{\neg\gamma \mid \gamma \in \Gamma\}$, i. e., the result of attach '$\neg$' to each γ in Γ. Then,

Proposition 2. $\Gamma \vDash_{SpV} \Delta$ *iff* $\neg(\Delta) \vDash_{SbV} \neg(\Gamma)$

Proof. Assume $\Gamma \vDash_{SpV} \Delta$. Then, for any interpretation if all the γ's are true everywhere, then some of the δ's are true everywhere. Contrapositively, if all the δ's are false somewhere, some of the γ's are false somewhere, that is, $\neg(\Delta) \vDash_{SbV} \neg(\Gamma)$ $\square$

Tableaux

Finally, we outline a procedure to extend standard modal tableaux for $\vDash_{SpV}$ and $\vDash_{SbV}$ and show that gap-principles are supervaluationist-unsatisfiable.

Definition 7 (Global modalities)**.** *For any interpretation* $\langle W, R, \nu\rangle$ *and any* $w \in W$, $\nu_w(\Box_g\varphi) = 1$ *iff* $\forall w \in W$ $\nu_w(\varphi) = 1$. *For any interpretation* $\langle W, R, \nu\rangle$ *and any* $w \in W$, $\nu_w(\Diamond_g\varphi) = \nu_w(\neg\Box_g\neg\varphi)$.

Remark: $\Box_g$ and $\Diamond_g$ are *global modalities* in the sense that their truth-conditions depend on what is going on in *every* world (whether accessible or not). Thus, global modalities cancel, so to speak, the local perspective characteristic of modal semantics in the sense that for any sentence φ, interpretation $\langle W, R, \nu\rangle$ and world $w \in W$, $\nu_w(\Box_g\varphi) = 1$ iff for every $w \in W$,

$\nu_w(\Box_g\varphi) = 1$ (the same remark applies '$\Diamond_g$'). Thus, φ holds at every w in an interpretation iff for any w in that interpretation $\nu_w(\Box_g\varphi) = 1$ (this remark is used in the lemma below). We write $\Box_g(\Gamma)$ for $\{\Box_g\gamma|\gamma \in \Gamma\}$ and $\Diamond_g(\Gamma)$ for $\{\Diamond_g\gamma|\gamma \in \Gamma\}$.

Lemma 1. *For any set of sentences* Γ *and* Δ, $\Gamma \vDash_{SpV} \Delta$ *iff* $\Box_g(\Gamma) \vDash_l \Box_g(\Delta)$ *and* $\Gamma \vDash_{SbV} \Delta$ *iff* $\Diamond_g(\Gamma) \vDash_l \Diamond_g(\Delta)$.

Proof. $\Gamma \vDash_{SpV} \Delta$ iff for every interpretation $\langle W, R, \nu\rangle$: if for all $\gamma \in \Gamma$ and for all $w \in W$, $\nu_w(\gamma) = 1$ then there is some $\delta \in \Delta$ such that for all $w \in W$ $\nu_w(\delta) = 1$. Taking into account our previous remark on global modalities, the foregoing hold iff for every interpretation $\langle W, R, \nu\rangle$ and world $w \in W$: if for all $\gamma \in \Gamma$, $\nu_w(\Box_g\gamma) = 1$ then for some $\delta \in \Delta$, $\nu_w(\Box_g\delta) = 1$, that is, iff $\Box_g(\Gamma) \vDash_l \Box_g(\Delta)$. An analogous reasoning shows that $\Gamma \vDash_{SbV} \Delta$ iff $\Diamond_g(\Gamma) \vDash_l \Diamond_g(\Delta)$. $\Box$

Modal tableaux constitute a systematic procedure to check whether a given set of sentences is locally satisfiable.[13] In order to check whether $\Gamma \vDash_l \Delta$ we construct a tableaux to check whether the set $\Gamma \cup \neg(\Delta)$ is locally satisfiable. Given our previous lemma, to check whether $\Gamma \vDash_{SpV} \Delta$ we should check whether $\Box_g(\Gamma) \cup \neg(\Box_g(\Delta))$ is locally satisfiable. The rules for the global modalities are as follows:

$\Box_g\varphi, n$	$\neg\Box_g\varphi, n$
φ, m (for every m in the tableaux)	$\neg\varphi, m$ (for a new m)

The global character is reflected in the rules by the fact that we do not need neither an accessibility node to trigger the $\Box_g$-rule nor to introduce an accessibility node after triggering the $\neg\Box_g$-rule. The rules for $\Diamond_g$ and $\neg\Diamond_g$ are analogous to these ones. Given our previous lemma, soundness and completeness proofs are analogous to those for the standard modal tableaux. A pair of examples.

Example 1 $p \vDash_{SpV} q \to \mathcal{D}p$

[13]Here we assume familiarity with this procedure, for more details see Priest (2008)

$$\Box_g p, 0$$
$$\neg\Box_g(q \to \mathcal{D}p), 0$$
$$\neg(q \to \mathcal{D}p), 1$$
$$q, 1$$
$$\neg\mathcal{D}p, 1$$
$$1r2$$
$$\neg p, 2$$
$$p, 2$$
$$\otimes$$

Example 2 $\neg(q \to \mathcal{D}p) \vDash_{SbV} \neg p$

$$\Diamond_g \neg(q \to \mathcal{D}p), 0$$
$$\neg\Diamond_g \neg p, 0$$
$$\neg(q \to \mathcal{D}p), 1$$
$$q, 1$$
$$\neg\mathcal{D}p, 1$$
$$1r2$$
$$\neg p, 2$$
$$\neg\neg p, 2$$
$$\otimes$$

Let us consider Fara's argument again. Consider a finite sorites series of m elements. The first element of the series is tall, the last element of the series is not tall. There is a number of $m-1$ relevant instances of relevant gap-principles at work in Fara's argument above (we write φ_n instead of tall(n)):

(GP. 1) $\mathcal{D}\neg\mathcal{D}^{m-2}\varphi_2 \to \neg\mathcal{D}^{m-1}\varphi_1$
(GP. 2) $\mathcal{D}\neg\mathcal{D}^{m-3}\varphi_3 \to \neg\mathcal{D}^{m-2}\varphi_2$
(GP. 3) $\mathcal{D}\neg\mathcal{D}^{m-4}\varphi_4 \to \neg\mathcal{D}^{m-3}\varphi_3$
$\vdots$
(GP. $m-2$) $\mathcal{D}\neg\mathcal{D}\varphi_{m-1} \to \neg\mathcal{D}^2\varphi_{m-2}$
(GP. $m-1$) $\mathcal{D}\neg\varphi_m \to \neg\mathcal{D}\varphi_{m-1}$

Lemma 2. *For any finite sorites series of m elements, $\langle 1, ..., m\rangle$, the relevant instances of relevant gap-principles are globally unsatisfiable.*

Proof.

$$\Box_g \varphi_1,\ 1$$
$$\Box_g \neg\varphi_m, 1$$
$$\Box_g \mathcal{D}\neg\mathcal{D}^{m-2}\varphi_2 \to \neg\mathcal{D}^{m-1}\varphi_1,\ 1$$
$$\Box_g \mathcal{D}\neg\mathcal{D}^{m-3}\varphi_3 \to \neg\mathcal{D}^{m-2}\varphi_2,\ 1$$
$$\vdots$$
$$\Box_g \mathcal{D}\neg\varphi_m \to \neg\mathcal{D}\varphi_{m-1},\ 1$$

$\neg\mathcal{D}\neg\mathcal{D}^{m-2}\varphi_2,\ 1$	$\neg\mathcal{D}^{m-1}\varphi_1,\ 1$
$1r2$	$\otimes$
$\mathcal{D}^{m-2}\varphi_2,\ 2$	

$\neg\mathcal{D}\neg\mathcal{D}^{m-3}\varphi_3,\ 2$	$\neg\mathcal{D}^{m-2}\varphi_2,\ 2$
$2r3$	$\otimes$
$\mathcal{D}^{m-3}\varphi_3,\ 3$	
$\vdots$	
$(m-2)r(m-1)$	
$\mathcal{D}\varphi_{m-1}, m-1$	

$\neg\mathcal{D}\neg\varphi_m,\ m-1$	$\neg\mathcal{D}\varphi_{m-1},\ m-1$
$(m-1)rm$	$\otimes$
φ_m	
$\otimes$	

Sentences prefixed with $\Box_g$ must hold everywhere in the tableaux. We trigger the rule for the first instance of a gap-principle. The right branch closes with the fact that $\Box_g\varphi_1$; the left branch is open and lead us to a new world where $\mathcal{D}^{m-2}\varphi_2$ should hold. We trigger the rule for the second instance of a gap-principle in this new world (we can always do this since it is a $\Box_g$-prefixed sentence). The right branch closes and the left branch lead us to a new world where we trigger the rule for the next instance of a gap-principle. Each time we do this, the right branch closes and the left branch lead us to a new world. When we trigger the rule corresponding to the last instance of a gap-principle, the left branch lead us to a world m where φ_m should hold. The branch closes based on the fact that $\Box_g\neg\varphi_m$. $\Box$

The tree reveals that in fact we did not need as much as $\Box_g$ to run the argument: we need just the premises be *definite enough* to reach the appropriate world in the tree (this fact relates to the connection between global and local validity expressed in Proposition 1). Contra (Asher et al., 2009, 924) gap-principles cannot be made definitely n for any n; rather, for each sorites series of m elements, there is a limited number of iterations of '$\mathcal{D}$' that keep the premises locally consistent.