Evans Tolerated*

Elia Zardini

LOGOS, Logic, Language and Cognition Research Group

Department of Logic, History and Philosophy of Science

University of Barcelona

Northern Institute of Philosophy

Department of Philosophy

University of Aberdeen

elia.zardini@ub.edu

October 2, 2013

Abstract

It is first argued that there are strong reasons for accepting borderline identity *in re*. Gareth Evans' influential argument to the effect that such identity is impossible is then rehearsed and extensions thereof are presented that would equally well show that borderline cases *in re* in general and indeed borderline cases in general (whether *in re* or not) are impossible. The naive theory of vagueness (holding that there is no sharp boundary between positive and negative cases of application of a vague predicate) and its accompanying tolerant logics (which are not unrestrictedly transitive), developed in earlier work by the author, are then introduced. Two specific tolerant logics, basically differing on their treatment of definiteness, are offered. Generally, in both logics Evans' argument fails because 'It is definite that P' is in some sense

*An earlier version of the material in this paper has been presented in 2013 at the PERSP Metaphysics Seminar (University of Barcelona). I'm grateful to the members of the seminar and to an anonymous referee for very stimulating discussions and comments. I'm also grateful to the editors Ken Akiba and Ali Abasnezhad for inviting me to contribute to this volume and for their support and patience throughout the process. During the writing of the paper, I have benefitted from the FP7 Marie Curie Intra-European Research Fellowship 301493, as well as from partial funds from the project FFI2011-25626 of the Spanish Ministry of Science and Innovation on *Reference, Self-Reference and Empirical Data*, from the project FFI2012-35026 of the Spanish Ministry of Economy and Competition on *The Makings of Truth: Nature, Extent, and Applications of Truthmaking*, from the project CONSOLIDER-INGENIO 2010 CSD2009-00056 of the Spanish Ministry of Science and Innovation on *Philosophy of Perspectival Thoughts and Facts* (PERSP) and from the FP7 Marie Curie Initial Training Network 238128 on *Perspectival Thoughts and Facts* (PETAF).

not weaker than 'P'. More specifically, in the first tolerant logic 'It is definite that x is identical with x' is in the relevant sense not weaker than 'x is identical with x' and Evans' argument fails because it illegitimately suppresses a logical truth from a valid argument; in the second tolerant logic, 'It is definite that x is identical with y' is in the relevant sense not weaker than 'x is identical with y' and Evans' argument fails because it implicitly appeals to principles that are untenable in a naive theory adopting a tolerant logic.

KEYWORDS

definiteness; Evans' argument; identity; naive theory of vagueness; tolerant logics

1 Borderline Identity *in Re*

Consideration of *vagueness* and *soritical series* strongly suggests that there have to be cases in which it is *borderline* whether an object x is *identical* with an object y. For example, it is a very plausible metaphysical assumption that we can go from a human being to a dumpling through a very long but finite stepwise process, where at each step we simply move one single atom in the whole universe merely within a nanometre's distance. Consider then Greg's seamless transformation into a dumpling, dividing it into a very long finite sequence of subsequent times $t_0, t_1, t_2 \ldots, t_k$ and, for every i such that $0 \leq i \leq k$,[1] calling 'Greg$_i$' the Greg-related substance (Greg, the dumpling, sort of Greg and what have you) wholly present at t_i. Given the other very plausible, mildly essentialist metaphysical assumption that nothing that is human at a time can be a dumpling at the same or other time, Greg$_0$ (i.e. Greg) is distinct from Greg$_k$ (i.e. the dumpling). However, since the transformation of the former into the latter is vague, it cannot be the case that, for every i, either it is definite[2] that Greg$_0$ is identical with Greg$_i$ or it is definite that Greg$_0$ is distinct from Greg$_i$, for whence could the vagueness of the transformation then arise? Thus, it has to be the case that, for some i, it is borderline whether Greg$_0$ is identical with Greg$_i$ (let such i be l).[3] And that is really a question about whether Greg$_0$ is identical with the *object itself* that is Greg$_l$, rather than really a question about the *referent* of 'Greg$_l$': for that referent has in effect been introduced as "the Greg-related substance wholly present at t_l", and where are the two or more objects that are equally

[1]Henceforth, this or similar qualifications on the values of 'i' will be implicitly understood.

[2]Throughout, I follow the established practice in the literature of using 'It is definite that P' and its like in such way that 'It is borderline whether P' is *strongly equivalent* with 'It is neither definite that P nor definite that it is not the case that P' (so strongly as to guarantee, for example, that they are *fully intersubstitutable* even in logics—which will be prominent in this paper—in which mere logical equivalence does not guarantee that).

[3]The text has in effect just moved from a claim of the form 'It is not the case that, for every x, P' to a claim of the form 'For some x, it is not the case that P'. That move is notoriously *intuitionistically* invalid and admittedly less than self-evident in contexts in which reference is made to "*unsurveyable*" totalities like soritical series. However, for the purposes of this paper, the move will be taken as granted (see Zardini [2013c] for some discussion).

good candidates for being such a substance?[4] An analogous consideration holds of course for the referent of ‘Greg_0’,[5] with the upshot that the question facing us is really whether an object x (i.e. Greg_0) is identical with an object y (i.e. Greg_l).

Once we grant ourselves a liberal enough ontology of properties (according to which, roughly, for every ordinary predicate ‘F’, there is such an object as the property of being F, and according to which properties are wholly present in each of the objects that exemplify them), cases of borderline identity multiply (as sketched by Priest [1991], p. 294, who attributes the point to Jean Paul van Bendegem). Consider a standard soritical series for baldness, with a seamless transition from bald people to non-bald people, dividing it into a very long finite sequence of people $b_0, b_1, b_2 \ldots, b_k$ (with, for every i, b_i having exactly i hairs on his scalp) and, for every i, calling ‘baldness_i’ the baldness-related property (baldness, non-baldness, sort of baldness and what have you) exemplified by, and wholly present in, b_i. Given that b_0 is bald while b_k is not, baldness_0 (i.e. baldness) is distinct from baldness_k (i.e. non-baldness). However, since the transformation of the former into

[4] As the phrase ‘wholly present’ indicates, I’m assuming a plausible *three-dimensionalist* view of *persistence* of substances across time (according to which, roughly, substances are typically present at different times and may be so present without different parts of them existing at different times). But, contrary to what some authors seem to think (see e.g. Noonan [1982]), borderline identity *in re* (henceforth, simply ‘borderline identity’) cannot be avoided by simply switching to a *four-dimensionalist* view (saying, for example, that, although it is definite that ‘Greg_l’ refers to a substance having as part the temporal slice present at t_l, it is borderline whether it refers to a substance having also as part the temporal slice present at t_0 or to a substance having also as part the temporal slice present at t_k). For an analogous example could be given in *modal* rather than *temporal* terms, and few of us would be tempted by the idea—discussed e.g. by Weatherson [2013]—that a typical substance is not wholly present in all the worlds in which it is present (nor would many of us be tempted by the idea—defended e.g. by Lewis [1986]—that, in a fundamental sense, a typical substance is present in only one world). In fact, given the plausible assumptions that there could be extended simples (as argued e.g. by Markosian [1998]) and that these could have fuzzy boundaries, an analogous example could be given in *spatial* rather than *modal* or *temporal* terms (contrary to the suggestion of Williamson [1994], pp. 255–256 to the effect that fuzzy spatial boundaries—which are certainly one of the prominent ways in which objects themselves can be thought to be vague—do not imply borderline identity). Notice that the thrust of these points consists in foreclosing the existence of (temporal, modal, spatial) *parts* that would give rise to several candidates for being the referent of the relevant expression. However, the foe of borderline identity might try to conjure up such candidates in other ways. For example, in the case of Greg_0, she might claim that there are at least two three-dimensional substances, Greg′ and Greg″, such that, while Greg′ has ceased to exist at t_l, Greg″ still exists at t_l. She might then hold that it is borderline whether Greg is Greg′ or Greg″, so that there is nothing ‘Greg’ definitely refers to. Unfortunately for the foe of borderline identity, the metaphysics required by this non-mereological move—implying that there are indefinitely many human beings, with slightly different persistence conditions, all of which are spatially co-located at t_0 and collectively undergo a dramatic mass death as Greg’s transformation unfolds—is still in these respects wildly implausible (notice that some of the problematic features just alluded to are shared by the mereological move, which can however at least provide a ready explanation of the different persistence conditions and of the spatial co-location at a time). I’m grateful to Aurélien Darbellay, Dan López de Sa and Giovanni Merlo for urging clarifications of an earlier version of this fn.

[5] For which we could also adduce the additional consideration that, extremely plausibly, there is something (i.e. Greg) that is definitely Greg, from which it follows that there is something that ‘Greg’ definitely refers to (since it is definite that being Greg implies being referred to by ‘Greg’), and so that there is something that ‘Greg_0’ definitely refers to (since it is definite that being referred to by ‘Greg’ implies being referred to by ‘Greg_0’).

the latter is vague, it cannot be the case that, for every i, either it is definite that baldness$_0$ is identical with baldness$_i$ or it is definite that baldness$_0$ is distinct from baldness$_i$, for whence could the vagueness of the transformation then arise? Thus, it has to be the case that, for some i, it is borderline whether baldness$_0$ is identical with baldness$_i$ (let such i be l). And that is really a question about whether baldness$_0$ is identical with the *object itself* that is baldness$_l$, rather than really a question about the *referent* of 'baldness$_l$': for that referent has in effect been introduced as "the baldness-related property exemplified by, and wholly present in, b_l", and where are the two or more objects that are equally good candidates for being such a property?[6] An analogous consideration holds of course for the referent of 'baldness$_0$',[7] with the upshot that the question facing us is really whether an object x (i.e. baldness$_0$) is identical with an object y (i.e. baldness$_l$). Thus, once we grant ourselves a liberal enough ontology of properties, every case of a soritical series can be turned into a case of borderline identity. Although *predication* is distinct from *identification*, there is arguably enough of a connection between the two to turn cases of borderline predication into cases of borderline identification.

2 Evans' Argument and Extensions Thereof

The considerations in section 1 notwithstanding, Evans [1978] heroically argued that borderline identity is impossible (and went even so far as to suggest that this would show that vague objects are impossible).[8] Evans' argument in favour of the *definiteness* of identity is really just a rewrite of the argument in favour of the *necessity* of identity popularised by Kripke (see e.g. Kripke [1971], p. 136), and, freely adapted, goes as follows (letting $\mathcal{D}\varphi$ mean 'It is definite that φ'):

[6] It is common to think that such candidates are easily available. For example, in the case of 'baldness$_0$', it is common to think that such candidates could be: the property of having at most l hairs on one's scalp, the property of having at most $l+1$ hairs on one's scalp, the property of having at most $l+2$ hairs on one's scalp etc. Unfortunately, this common thought forgets that the stuff about hairs on one's scalp is just a useful oversimplification of what baldness consists in: it is definite that a man with l hairs on his scalp that are however much thicker than normal and uniformly distributed so as to cover the whole of his scalp with a bushy mane is not bald, and so, after all, the simple property of having at most l hairs on one's scalp is not a good candidate for being the referent of 'baldness$_0$'. And, as in so many other cases of conceptual analysis, the prospects of coming up with a series of more complex precise properties strong enough as to rule out all definite cases of non-baldness and at the same time weak enough as to rule in all definite cases of baldness are bleak. Even setting this aside, points similar to those raised at the end of fn 4 apply here: it is metaphysically wildly implausible to think that there are indefinitely many trichological disorders, with slightly different exemplification conditions, all of which conspire to affect poor b_0.

[7] For which we could also adduce the additional consideration that, extremely plausibly, there is something (i.e. baldness) that is definitely baldness, from which it follows that there is something that 'baldness' definitely refers to (since it is definite that being baldness implies being referred to by 'baldness'), and so that there is something that 'baldness$_0$' definitely refers to (since it is definite that being referred to by 'baldness' implies being referred to by 'baldness$_0$').

[8] A related argument is offered in Salmon [2005], pp. 243–246 (whose first edition is from 1982), which however also requires some basic principles from the theory of ordered pairs. *Mutatis mutandis*, all the discussion to follow applies equally well to Salmon's argument.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\varnothing \vdash x = x}\text{reflexivity}}{\varnothing \vdash \mathcal{D}(x = x)}\text{simple definitisation} \qquad \dfrac{}{x = y, \mathcal{D}(x = x) \vdash \mathcal{D}(x = y)}\text{indiscernibility of identicals}}{x = y \vdash \mathcal{D}(x = y)}\text{transitivity}}{x = y, \neg\mathcal{D}(x = y) \vdash \varnothing}\neg\text{-L}}{\neg\mathcal{D}(x = y) \vdash x \neq y}\neg\text{-R}$$

(let's call this argument 'argument E').[9]

Without further justification, Evans claims that $x \neq y$ is inconsistent with $\mathcal{B}(x = y)$ (letting $\mathcal{B}\varphi$ mean 'It is borderline whether φ'), presumably on the grounds of the general principle of *inconsistency*:

(INC) φ is inconsistent with $\neg\mathcal{D}\varphi$,

and, I suppose, implicitly assumes that, if $x \neq y$ is inconsistent with $\neg\mathcal{D}(x \neq y)$, since argument E shows that $\neg\mathcal{D}(x = y)$ entails $x \neq y$ it follows that $\mathcal{B}(x = y)$ is inconsistent, presumably on the grounds of the general principle of *transitivity of adjunction of inconsistents*:

($\mathrm{TR}^{\mathrm{ADJ}^{\mathrm{INC}}}$) If φ entails ψ, φ entails χ and ψ is inconsistent with χ, then φ is inconsistent.

Even without (INC) in its full generality, if the logic of definiteness is as strong as **KB**, rewriting the standard argument for the necessity of *distinctness* we could indeed show that $x \neq y, \neg\mathcal{D}(x \neq y) \vdash \varnothing$ holds. In any event, even without assuming that the logic of definiteness is as strong as **KB** (see Field [2000] for an objection against the **B**-axiom for $\mathcal{D}$), and, more generally, even without going for a ($\mathrm{TR}^{\mathrm{ADJ}^{\mathrm{INC}}}$)-route, the conclusion of argument E would still seem to suffice for committing one thinking that $\neg\mathcal{D}(x = y)$ holds to think that $x \neq y$ holds, presumably on the grounds of the general principle of *closure of commitment to thinking under logical consequence*:

[9] Contrary to other discussions of the same topic, ours is conducted having in view as paradigmatic candidate cases of borderline identity mainly cases of *cross-temporal* and *cross-modal* identity (see section 1 and in particular fn 4). But, as is well-known, in such cases indiscernibility of identicals has to be handled with some care: for example, we don't want to conclude from the fact that in a world Greg is German and in another world Greg is not German that Greg (in the former world) is distinct from Greg (in the latter world). The properties discernibility across which entails cross-temporal and cross-modal distinctness are naturally thought of as properties exemplified *in an atemporal and amodal way*, like, for example, the property of being human, the property of being concrete and the property of being self-identical. We do want to conclude from the fact that Greg is human and a dumpling is not human that Greg (at least in any world in which he exists) is distinct from the dumpling (at least in any world in which it exists), no matter precisely in which worlds they happen to exemplify these properties (for example, Greg may not be human in worlds in which he does not exist). Since the property of being definitely identical with Greg_0 is arguably one of the properties that are exemplifiable in an atemporal and amodal way, the application of indiscernibility of identicals required by argument E in cases of borderline cross-temporal or cross-modal identity is legitimate. I'm grateful to Aurélien Darbellay for raising this issue.

(CCTLC) If φ entails ψ and one thinks that φ holds, then one is committed to thinking that ψ holds.

And that may seem to be rebarbative, since, if one thinks that $\mathcal{B}(x = y)$ holds, it may seem that one would precisely like to avoid commitment to thinking that $x \neq y$ holds.

Instead of making the simple point made at the end of the previous paragraph, Evans puzzlingly suggests instead that the logic of definiteness may be as strong as **S5**. If that were the case, given the conclusion of argument E, since in **S5** $\neg\mathcal{D}(x = y) \vdash \mathcal{D}\neg\mathcal{D}(x = y)$ holds it would follow, by *single-premise closure of definiteness under logical consequence*:

(M) If $\varphi \vdash \psi$ holds, then $\mathcal{D}\varphi \vdash \mathcal{D}\psi$ holds

and transitivity of logical consequence, that $\neg\mathcal{D}(x = y) \vdash \mathcal{D}(x \neq y)$ holds, and so it would follow, by $\neg$-L, that $\mathcal{B}(x = y)$ is indeed inconsistent. Unfortunately for Evans' suggestion, because of *higher-order vagueness* it is widely assumed that the logic of definiteness cannot be as strong as **S5**. However, given simply (M), from the conclusion of argument E it follows that $\mathcal{D}\neg\mathcal{D}(x = y) \vdash \mathcal{D}(x \neq y)$ holds, from which one would expect it to follow that $\mathcal{D}\mathcal{B}(x = y)$ is indeed inconsistent.

Thus, the conclusion of argument E provides the materials for a variety of concerns about the possibility of borderline identity (let's use 'Evans' argument' to refer to argument E plus whatever is one's favoured way of arguing that the conclusion of argument E creates problems for the possibility of borderline identity). In the remainder of this paper, I'll show that there is at least one independently motivated logic of vagueness and definiteness in which argument E fails, and at least one independently motivated logic of vagueness and definiteness in which the conclusion of argument E is harmless. Either way, Evans' argument will fail.

Before seeing that, however, it'll be helpful to reach a more comprehensive perspective on Evans' argument, in particular on argument E and on the style of reasoning it exemplifies. To understand the real punch of argument E, it is useful to ask first why an argument using an equivalence relation weaker than identity like *material equivalence* would not go through. Letting $\equiv$ and B express material equivalence and baldness respectively, since Bb_0 holds, Bb_l is tantamount to $Bb_0 \equiv Bb_l$; therefore, since $\neg\mathcal{D}Bb_l$ holds, so should $\neg\mathcal{D}(Bb_0 \equiv Bb_l)$. But one could then have thought that one can create problems for borderline cases in general *via* an argument analogous to argument E, where the formulas $x = x$ and $x = y$ are replaced by $Bb_0 \equiv Bb_0$ and $Bb_0 \equiv Bb_l$ respectively and the principles of reflexivity of identity and indiscernibility of identicals are replaced by *reflexivity of material equivalence* and *intersubstitutability of materially equivalents* respectively. However, of course, on minimal assumptions about definiteness the last principle fails in those cases in which the envisaged substitution is within a $\mathcal{D}$-context, at least if the ensuing entailment is then used as an input for $\neg$-L and $\neg$-R (in other words, the last principle fails at least in its contrapositive form of *material non-equivalence of non-intersubstitutables*). The real punch of argument E is then that, contrary to material equivalence, identity does seem to be a strong enough relation as to validate intersubstitutability of its terms also within a $\mathcal{D}$-context, even if the ensuing entailment is then used as an input for ¬-L and ¬-R (in other words, indiscernibility of identicals does not seem to fail even in its contrapositive form of *distinctness of discernibles*, see also fn 24).

But, if that is the real punch of argument E, one should indeed be able to extend it so as to create problems for borderline cases in general, by using not of course material equivalence but a stronger equivalence relation such as *semantic equivalence* (in the sense of identity of semantic value). Letting $\cong$ express semantic equivalence, since Bb_0 holds, Bb_l is tantamount to $Bb_0 \cong Bb_l$ (for it is tantamount to its semantic value being the result of applying the semantic value of 'bald' to a bald man); therefore, since $\neg\mathcal{D}Bb_l$ holds, so should $\neg\mathcal{D}(Bb_0 \cong Bb_l)$. But one can indeed then create problems for borderline cases in general *via* an argument analogous to argument E, where the formulas $x = x$ and $x = y$ are replaced by $Bb_0 \cong Bb_0$ and $Bb_0 \cong Bb_l$ respectively and the principles of reflexivity of identity and indiscernibility of identicals are replaced by *reflexivity of semantic equivalence* and *intersubstitutability of semantically equivalents* respectively. Notice in particular that intersubstitutability of semantically equivalents is, if anything, even more compelling than indiscernibility of identicals, since it is a straightforward consequence of *compositionality of semantic value*; for that reason, semantic equivalence has an even better claim than identity to be a strong enough relation as to validate intersubstitutability of its terms also within a $\mathcal{D}$-context, even if the ensuing entailment is then used as an input for ¬-L and ¬-R (in other words, intersubstitutability of semantically equivalents has an even better claim than identity not to fail even in its contrapositive form of *semantic non-equivalence of non-intersubstitutables*).

Relatedly, even without talk of semantic equivalence, one should be able to extend argument E to cover at least borderline cases *in re* in general, whose existence would seem undeniable on the strength of the considerations developed in section 1 (especially in fns 6 and 7). As a warm-up argument, focus first on borderline *parthood in re* (henceforth, simply 'borderline parthood'). Obviously, having something as a part does not imply having all its properties *of every kind whatsoever*; yet, it does seem to imply having all its properties of *definite positive partial location*:[10] how can a place be a place of definite location for a part without being a place of definite location for every larger whole of which it is a part?[11] We thus seem to have the principle of *monotonicity of definite positive location over parthood*:

(MDPLP) If x is part of y and it is definite that x is at p, then it is definite that y is at p.

[10]Henceforth, 'location' and its relatives will be understood as 'partial location' and its relatives.

[11]To elaborate a bit, if x is part of y, y is identical with $y + x$. But, if it is definite that x is at p, surely it is definite that $y + x$ is at p. That is not only intuitively compelling; if the logic of definiteness is as strong as **K**, it follows from the uncontroversial fact that it is definite that, if x is at p, $y + x$ is at p. An application of indiscernibility of identicals that is as legitimate as the one made in argument E (since it only involves expressions—'y' and '$y + x$'—such that there is something they definitely refer to) yields then the desired conclusion.

But, letting $\trianglelefteq$ and L and express parthood and location respectively, one can indeed then create problems for borderline parthood *via* an argument similar to argument E:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{L(x,p) \;\vdash\; \mathcal{D}L(x,p)}\,\text{fact entails definite fact} \qquad \dfrac{}{x \trianglelefteq y, \mathcal{D}L(x,p) \;\vdash\; \mathcal{D}L(y,p)}\,\text{(MDPLP)}}{x \trianglelefteq y, L(x,p) \;\vdash\; \mathcal{D}L(y,p)}\,\text{transitivity}}{x \trianglelefteq y, L(x,p), \neg\mathcal{D}L(y,p) \;\vdash\; \varnothing}\,\neg\text{-L}}{L(x,p), \neg\mathcal{D}L(y,p) \;\vdash\; x \ntrianglelefteq y}\,\neg\text{-R}$$

(let's call this argument 'argument F'). Now, in the cases of interest, while $L(x,p)$ is uncontroversial $x \trianglelefteq y$ is tantamount to $L(y,p)$, and so $\neg\mathcal{D}(x \trianglelefteq y)$ is tantamount to $\neg\mathcal{D}L(y,p)$. Therefore, the conclusion of argument F provides the materials for a variety of concerns about the possibility of borderline parthood analogous to the materials for a variety of concerns about the possibility of borderline identity provided by argument E.[12]

Given this warm-up argument, it's easy to see how, appealing to the liberal ontology of properties introduced in section 1 and letting E express exemplification, one can create problems for borderline cases *in re* in general *via* an argument analogous to argument F, where the formula $x \trianglelefteq y$ is replaced by $E(x,y)$ and (MDPLP) is replaced by the equally compelling principle of *monotonicity of definite positive location over exemplification*:

(MDPLE) If x exemplifies y and it is definite that x is at p, then it is definite that y is at p.[13,14]

[12] It's worth mentioning that there is a relative of argument F which even more closely resembles argument E and which is thereby more general (in that it does away with talk of location). That relative can be got from argument E by replacing the formulas $x = x$ and $x = y$ with $x \trianglelefteq x$ and $x \trianglelefteq y$ respectively and by replacing the principles of reflexivity of identity and indiscernibility of identicals with *reflexivity of parthood* and *monotonicity of definite parthood over parthood* (if x is part of y and it is definite that z is part of x, then it is definite that z is part of y) respectively. Notice that the latter principle, although it may initially come across as slightly less plausible than (MDPLP), can be supported in a way analogous to how (MDPLP) has been supported in fn 11.

[13] (MDPLE) can be supported in a way analogous to how (MDPLP) has been supported in fn 11. Letting $\text{pla}(x)$ be the sum of the places at which x is located (and assuming that there is something 'pla' definitely refers to), if x exemplifies y, $\text{pla}(y)$ is identical with $\text{pla}(y) + \text{pla}(x)$. But, if it is definite that p is part of $\text{pla}(x)$, surely it is definite that p is part of $\text{pla}(y) + \text{pla}(x)$. That is not only intuitively compelling; if the logic of definiteness is as strong as **K**, it follows from the uncontroversial fact that it is definite that, if p is part of $\text{pla}(x)$, p is part of $\text{pla}(y) + \text{pla}(x)$. An application of indiscernibility of identicals that is as legitimate as the one made in argument E (since it only involves expressions—'$\text{pla}(y)$' and '$\text{pla}(y) + \text{pla}(x)$'—such that there is something they definitely refer to) yields then that it is definite that p is part of $\text{pla}(y)$, which is tantamount to its being definite that y is at p.

[14] It's worth mentioning that there is a relative of the argument considered which even more closely resembles argument E and which is thereby more general (in that it does away with talk of location). That relative employs plural talk and assumes that there are some things 'the bald people' definitely refers to. (Notice that this is a fair assumption given that, at this point in the text, we're working under the hypothesis that there are borderline cases *in re*: if there is something 'baldness' definitely refers to, the things exemplifying it surely are the things 'the bald people' definitely refers to. Notice also that, independently of these issues, the arguments to follow can be reworked as arguments targeting

To sum up, Evans' argument can be extended in various ways so as to cover borderline cases *in re* in general and indeed borderline cases in general (whether *in re* or not). Given the extensions to borderline cases *in re* in general, and *contra* the approach exemplified for instance by Salmon [2005], *any theorist admitting such cases must find fault with Evans-style reasoning* (rather than accepting Evans' argument as a sound demonstration that borderline identity, contrary to other kinds of borderline cases *in re*, is impossible). And, given the extension to borderline cases in general (whether *in re* or not), and *contra* the approach exemplified for instance by Lewis [1988], *any theorist whatsoever must find fault with Evans-style reasoning, and must do so at some step other than the one at which equivalent expressions are substituted.*

3 The Naive Theory of Vagueness and Tolerant Logics

In earlier work (Zardini [2006a]; [2006b]; [2008a]; [2008b]; [2009]; [2013a]; [2013b]; [2013c]; [2013d]), I've developed and defended a *naive theory of vagueness*, according to which, very roughly, the vagueness of an expression consists in its *tolerance* (see Wright [1975])—that is, in its *not drawing a sharp boundary between positive and negative cases of application.* Thus, for example, going back to the second soritical series mentioned in section 1 the vagueness of 'bald' (in that situation) consists in the fact that, for every i, if b_i is bald, so is b_{i+1}. The theory applies straightforwardly also to soritical series involving identity. Thus, for example, going back to the first soritical series mentioned in section 1 the vagueness of 'identical with Greg_0' (in that situation) consists in the fact that, for every i, if Greg_i is identical with Greg_0, so is Greg_{i+1}.

The naive theory of vagueness arguably enjoys a number of important advantages over its rivals (see Zardini [2008b], pp. 15–16, 21–71; Sweeney and Zardini [2011]; Zardini [2013d]). It however faces a significant problem that has usually been taken to be fatal: as shown by *standard soritical reasoning*, it is *inconsistent in almost any logic of vagueness.* The naive theory thus requires a novel logic, and I do provide a family of logics—*tolerant logics*—that are hospitable to the theory in Zardini [2008a]; [2008b], pp. 93–173; [2009];

at least borderline being-some-of *in re*.) Letting $\sqsubseteq$ and bb express being-some-of and the bald people respectively, the argument in question can be got from argument E by replacing the formulas $x = x$ and $x = y$ with $xx \sqsubseteq xx$ and $xx \sqsubseteq bb$ respectively and by replacing the principles of reflexivity of identity and indiscernibility of identicals with *reflexivity of being-some-of* and *monotonicity of definite being-some-of over being-some-of* (if the xx are some of the yy and it is definite that the zz are some of the xx, then it is definite that the zz are some of the yy) respectively. Notice that the latter principle, although it may initially come across as slightly less plausible than (MDPLP) (or (MDPLE)), can be supported in a way analogous to how (MDPLP) has been supported in fn 11. Indeed, letting $\doteqdot$ and ll express plural identity and the plurality formed by the bald people together with b_l respectively, taking argument E, replacing the formulas $x \doteqdot x$ and $x \doteqdot y$ with $ll \doteqdot ll$ and $ll \doteqdot bb$ respectively and replacing the principles of reflexivity of identity and indiscernibility of identicals with *reflexivity of plural identity* and *indiscernibility of identical pluralities* respectively, we obtain $\neg\mathcal{D}(ll \doteqdot bb) \vdash \neg(ll \doteqdot bb)$, which is tantamount to $\neg\mathcal{D}Bb_l \vdash \neg Bb_l$ (since $ll \doteqdot bb$ is tantamount to Bb_l).

[2013b]; [2013c]. The main feature of tolerant logics is that they place restrictions on *transitivity* of logical consequence, a restriction which proves crucial in blocking standard soritical reasoning and, more generally, in restoring the consistency of the naive theory.

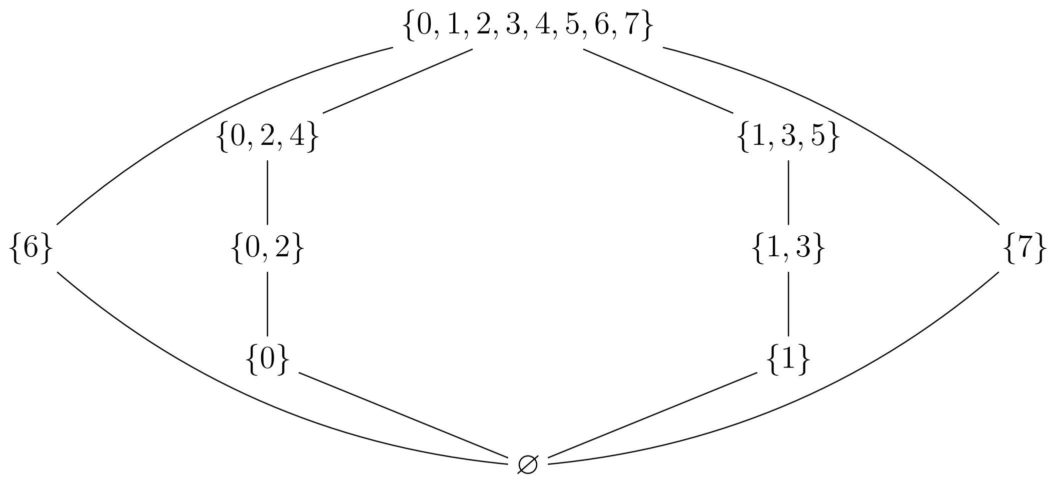
Instead of offering the full framework in which tolerant logics can be developed, in this paper I'll simply introduce my favoured broad kind of tolerant logic, with the main aim of using it to offer an analysis of where and why Evans' argument fails. I'll offer two versions of that kind of logic, basically differing on the treatment of definiteness. In this regard, I should stress that the focus of this paper is on using those two tolerant logics to explore the issues raised by Evans' argument rather than on offering two philosophically and formally complete theories of definiteness. As a consequence, I'll only give as much philosophical and formal detail about those two theories as is necessary for such an exploration, leaving it to another occasion to offer a philosophically and formally adequate presentation and comparison of them (notice that, among other things, I'll mostly ignore the considerable philosophical and formal complications required by higher-order vagueness). Also, although, in both cases, the analysis of where and why Evans' argument fails will straightforwardly cover also the extensions of Evans' argument developed in section 2, I'll save the reader the details of this.

4 The Tolerant Logic $\mathbf{V_0}$

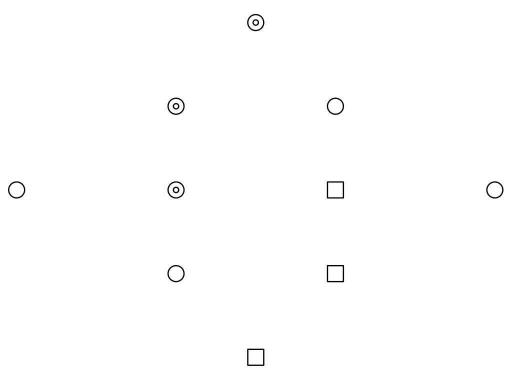
Let's start then with the logic $\mathbf{V_0}$, which treats definiteness as obeying a relatively standard modal logic. For our purposes, $\mathbf{V_0}$ and its like can be introduced *model-theoretically*. Informally, we consider a certain class of *lattice-theoretic* models rich enough as to include operations sufficient for interpreting a standard first-order modal language. Such models include a set of *designated* values and also a set of *tolerated* values which is a *proper superset* of the former set. Rather than defining, as usual, logical consequence in terms of *preservation* in every model of designated (or tolerated) value from the premises to the conclusions, we define it instead as *connection* in every model of designated value in the premises with tolerated value in the conclusions. It is this style of definition of logical consequence which will give rise to failures of transitivity for $\mathbf{V_0}$ and its like.

Definition 1. A $\mathbf{V_0}$-*model* $\mathfrak{M}$ is a 9ple $\langle U_{\mathfrak{M}}, V_{\mathfrak{M}}, \preceq_{\mathfrak{M}}, D_{\mathfrak{M}}, T_{\mathfrak{M}}, \mathrm{neg}_{\mathfrak{M}}, \mathrm{def}_{\mathfrak{M}}, \mathrm{id}_{\mathfrak{M}}, \mathrm{int}_{\mathfrak{M}}\rangle$. $U_{\mathfrak{M}}$ is a *domain* of objects. $V_{\mathfrak{M}}$ is a set of *values* representable as: $\{X : X \in \mathrm{pow}(\{i : 0 \leq i \leq 7\})$ and, if $X \neq \{i : 0 \leq i \leq 7\}$, either, [[for every $i \in X$, i is even][15] and, [for every i and j, if $i \in X$ and ≤ 4 and j is even and $< i$, $j \in X$] and, [for every i and $j \in X$, $|i-j| < 6$]] or, [[for every $i \in X$, i is odd] and, [for every i and j, if $i \in X$ and ≤ 5 and j is odd and $< i$, $j \in X$] and, [for every i and $j \in X$, $|i-j| < 6$]]$\}$. $\preceq_{\mathfrak{M}}$ is a *partial order* on $V_{\mathfrak{M}}$ representable as: $\{\langle X, Y\rangle : X \subseteq Y\}$. Thus, $V_{\mathfrak{M}}$ and $\preceq_{\mathfrak{M}}$ jointly constitute the lattice depicted by the following Hasse diagram:

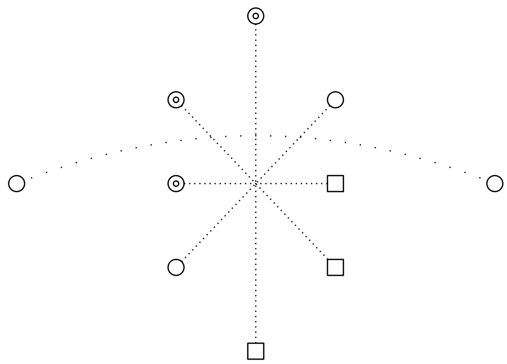
[15]Throughout, I use square brackets to disambiguate constituent structure.



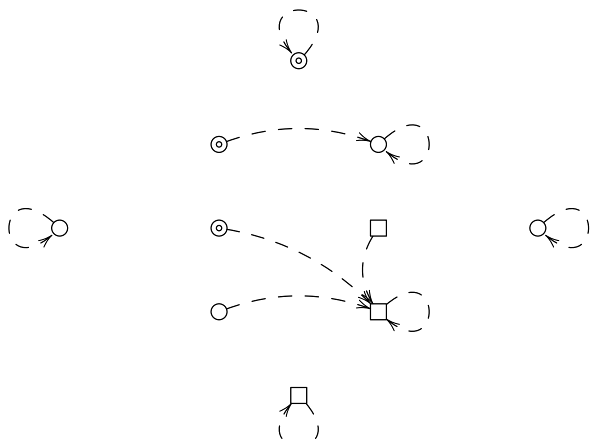
$D_{\mathfrak{M}}$ is a set of *designated* values while $T_{\mathfrak{M}}$ is a set of *tolerated* values, with $D_{\mathfrak{M}} \subset T_{\mathfrak{M}}$. Indicating designated values with doubly circular nodes, tolerated but not designated values with simply circular nodes and neither designated nor tolerated values with square nodes, they can be depicted as:



$\mathrm{neg}_{\mathfrak{M}}$ is a *negation* operation on $V_{\mathfrak{M}}$. Indicating it with pointed edges, it can be depicted as:

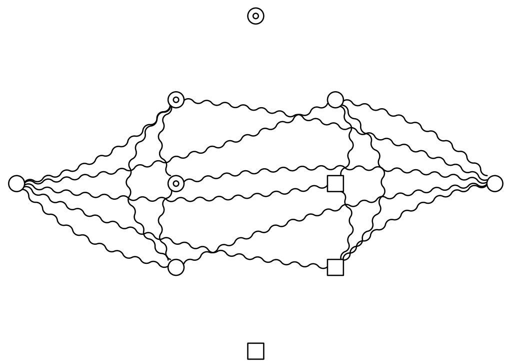


$\mathrm{def}_{\mathfrak{M}}$ is a *definitisation* operation on $V_{\mathfrak{M}}$. Indicating it with dashed arrows, it can be depicted as:



$\mathrm{id}_{\mathfrak{M}}$ is an *identification* function from the Cartesian product of $U_{\mathfrak{M}}$ with itself to $V_{\mathfrak{M}}$. In addition to being commutative in its arguments, it can be represented as being such that, for every $u \in U_{\mathfrak{M}}$, $\{0, 2, 4\} \subseteq \mathrm{id}_{\mathfrak{M}}(u, u)$ and as being such that, if $\mathrm{id}_{\mathfrak{M}}(u, v) \in D_{\mathfrak{M}}$, for every atomic $i$ary predicate Φ^i (including $=$) $\mathrm{val}_{\mathfrak{M},\mathrm{ass}}(\Phi^i)(u_0, u_1, u_2 \ldots, v \ldots, u_{i-1})$[16] either is identical with $\mathrm{val}_{\mathfrak{M},\mathrm{ass}}(\Phi^i)(u_0, u_1, u_2 \ldots, u \ldots, u_{i-1})$ or at most deviates from it in the way depicted by the squiggle dashes:

[16] $\mathrm{val}_{\mathfrak{M},\mathrm{ass}}$ is the model- and assignment-relative valuation function that will be defined below in terms of $\mathfrak{M}$ (and in particular in terms of $\mathrm{id}_{\mathfrak{M}}$). The circularity is not vicious as we're here only constraining rather than defining $\mathrm{id}_{\mathfrak{M}}$. It is of course not guaranteed that such constraint is satisfiable. Models in which it is satisfied will be sketched in theorems 5 and 6.



$\mathrm{int}_{\mathfrak{M}}$ is an *interpretation* function from the union of the Cartesian products of the sets of $i$ary atomic predicates (minus the identity predicate) with the sets of $i$tuples of members of $U_{\mathfrak{M}}$ to $V_{\mathfrak{M}}$.

$\mathrm{int}_{\mathfrak{M}}$ can be extended to a full *valuation* function $\mathrm{val}_{\mathfrak{M}}$ (relative to assignments) in the usual way (using glb in $\preceq_{\mathfrak{M}}$ for interpreting conjunction and universal quantification, $\mathrm{neg}_{\mathfrak{M}}$ for interpreting negation, $\mathrm{def}_{\mathfrak{M}}$ for interpreting definiteness and $\mathrm{id}_{\mathfrak{M}}$ for interpreting identity).[17] With such function in place, the advertised style of definition of logical consequence becomes available:

Definition 2. $\Gamma \vdash_{\mathbf{V_0}} \Delta$ holds iff, for every $\mathbf{V_0}$-model $\mathfrak{M}$ and assignment ass, if, for every $\varphi \in \Gamma$, $\mathrm{val}_{\mathfrak{M},\mathrm{ass}}(\varphi) \in D_{\mathfrak{M}}$, then, for some $\psi \in \Delta$, $\mathrm{val}_{\mathfrak{M},\mathrm{ass}}(\psi) \in T_{\mathfrak{M}}$.

In Zardini [2008a]; [2008b], pp. 93–173; [2009]; [2013c], I explore in great detail the properties of what is essentially the $\mathcal{D}$-free fragment of $\mathbf{V_0}$. Here, I only pause to record the main properties of $\mathcal{D}$ and the other properties that are relevant for our purposes:

Theorem 1. *In* $\mathbf{V_0}$*, definiteness obeys a modal logic* at least as strong as $\mathbf{CN^{\neg,\mathcal{D}}DT^{\subset}B45}$*, in the sense that, in* $\mathbf{V_0}$*, the principles:*

(C) $\mathcal{D}\varphi, \mathcal{D}\psi \vdash \mathcal{D}(\varphi \wedge \psi)$ *holds;*

($\mathrm{N}^{\neg,\mathcal{D}}$) *If* φ *does not contain either* $\wedge$ *or* $\forall$ *and* $\varnothing \vdash \varphi$ *holds, then* $\varnothing \vdash \mathcal{D}\varphi$ *holds;*[18]

(D) *If* $\varphi \vdash \varnothing$ *holds, then* $\mathcal{D}\varphi \vdash \varnothing$ *holds;*

($\mathrm{T}^{\subset}$) $\varnothing \vdash \mathcal{D}\varphi \supset \varphi$ *holds;*

(B) $\varnothing \vdash \varphi \supset \mathcal{D}\neg\mathcal{D}\neg\varphi$ *holds;*

(4) $\varnothing \vdash \mathcal{D}\varphi \supset \mathcal{D}\mathcal{D}\varphi$ *holds;*

[17] In turn, we can assume that disjunction, implication and particular quantification are defined in the usual way.

[18] ($\mathrm{N}^{\neg,\mathcal{D}}$) was more informally labelled 'simple definitisation' in section 2.

(5) $\varnothing \vdash \neg\mathcal{D}\varphi \supset \mathcal{D}\neg\mathcal{D}\varphi$ *holds*

hold.[19]

Theorem 2. *In* $\mathbf{V_0}$*, negation is* exclusive *and* exhaustive*, in the sense that, in* $\mathbf{V_0}$*, the principles:*

(EXC) *If* $\Gamma \vdash \Delta, \varphi$ *holds, then* $\Gamma, \neg\varphi \vdash \Delta$ *holds;*

(EXH) *If* $\Gamma, \varphi \vdash \Delta$ *holds, then* $\Gamma \vdash \Delta, \neg\varphi$ *holds*[20]

hold.

Theorem 3. *In* $\mathbf{V_0}$*, identity is* reflexive *and entails* indiscernibility across non-composite properties*,*[21] *in the sense that, in* $\mathbf{V_0}$*, the principles:*

(R) $\varnothing \vdash \tau = \tau$ *holds;*

(II) *If* φ *does not contain either* $\wedge$ *or* $\forall$*, then* $\tau = \sigma, \varphi \vdash \varphi[\tau/\sigma]$[22] *holds*[23]

hold.

[19]Notice that, with *classical* logic as *background* logic for the *modal* logic, the list of modal principles given in the text is multiply redundant. But this need no longer be so if a *non-classical* logic is used as background logic for the modal logic, in particular if the non-classical logic in question is in the relevant respects so weak as $\mathbf{V_0}$ is. Notice also that, with the exception of (N$^{\neg,\mathcal{D}}$), the listed principles will not be directly concerned by our discussion. However, the validity of principles (C) and (N$^{\neg,\mathcal{D}}$) serves to mark the extent to which definiteness is still *closed under logical consequence* in $\mathbf{V_0}$ and its like (see the discussion in section 6), while the validity of principles (D)–(5) serves to show that Evans' argument fails in $\mathbf{V_0}$ and its like even under *very strong assumptions* about definiteness (setting aside that, as I've noted in section 2, at least (5) is widely rejected because of higher-order vagueness).

[20](EXC) and (EXH) were more neutrally labelled '¬-L' and '¬-R' respectively in section 2.

[21]A property is *non-composite* iff it is neither conjunctive nor disjunctive. Arguably, a composite property is exemplified in virtue of a non-composite property being exemplified and of logical facts about conjunction and disjunction. That in turn plausibly suggests that, if identicals are indiscernible across composite properties, they are such in virtue of their being indiscernible across non-composite properties and of logical facts about conjunction and disjunction. But, for reasons I'll partially adumbrate in section 6, the required logical facts about conjunction and disjunction do not obtain in $\mathbf{V_0}$ and its like, and in fact cannot possibly obtain in any naive theory of vagueness adopting as tolerant logic $\mathbf{V_0}$ or one of its like. Therefore, (II) is arguably the proper formulation of the principle of indiscernibility of identicals in $\mathbf{V_0}$ and its like. Even more strongly, for reasons I cannot go into in this paper, identicals cannot possibly be indiscernible across composite properties in any naive theory adopting as tolerant logic $\mathbf{V_0}$ or one of its like, and, on a natural way of making metaphysical sense of any such theory, this is so because indiscernibility across composite properties may concern the same object in different circumstances (and it is uncontroversial that the same object may exemplify different properties in different circumstances, see fn 9).

[22]With the usual proviso that τ be free for σ in φ.

[23](R) and (II) were more informally labelled 'reflexivity' and 'indiscernibility of identicals' respectively in section 2.

Corollary 1. *In* $\mathbf{V_0}$*, identity is* transitive *and distinctness is entailed by* discernibility across non-composite properties*, in the sense that, in* $\mathbf{V_0}$*, the principles:*

(TR$^=$) $\tau = \sigma, \sigma = \upsilon \vdash \tau = \upsilon$ *holds;*

(DD) *If* φ *does not contain either* $\wedge$ *or* $\forall$*, then* $\varphi, \neg\varphi[\tau/\sigma] \vdash \tau \neq \sigma$ *holds*[24]

hold.

Theorem 4. $\mathbf{V_0}$ *is* non-transitive*, in the sense that, in* $\mathbf{V_0}$*, the principle:*

(TR$^\vdash$) *If, [for every* $\varphi \in \Theta$*,* $\Gamma \vdash \Delta, \varphi$ *holds] and* $\Lambda, \Theta \vdash \Xi$ *holds, then* $\Lambda, \Gamma \vdash \Delta, \Xi$ *holds*[25]

does not hold.[26]

5 The Failure of Argument E in $\mathbf{V_0}$

Argument E employs (TR$^\vdash$), and so it fails in $\mathbf{V_0}$ (indeed, since, in $\mathbf{V_0}$, all the other principles employed by argument E hold, in a good sense in $\mathbf{V_0}$ argument E fails *exactly* at (TR$^\vdash$)). However, the fact that argument E in its specificity fails in $\mathbf{V_0}$ is no great reassurance that its conclusion does not hold in $\mathbf{V_0}$, since, for all that has been shown, there could be more ingenious arguments that establish that conclusion and that are valid in $\mathbf{V_0}$. What needs to be shown is that the conclusion does not hold in $\mathbf{V_0}$, which is to say that there is a $\mathbf{V_0}$-model $\mathfrak{M}$ such that $\mathrm{val}_{\mathfrak{M},\mathrm{ass}}(\neg\mathcal{D}(x = y)) \in D_{\mathfrak{M}}$ and $\mathrm{val}_{\mathfrak{M},\mathrm{ass}}(x \neq y) \notin T_{\mathfrak{M}}$. Indeed, to make sure that the borderline identity between x and y does not have problematic consequences for the naive theory of vagueness, and that, even together with it, it remains silent on whether x is identical with y, we should show that, in $\mathbf{V_0}$, $\mathcal{B}(x = y)$ together with the relevant fragment of the naive theory entails neither $x \neq y$ nor

[24] Many theories of borderline identity block Evans' argument by rejecting (DD) while preserving (II), claiming that (DD) is only a degenerated principle resulting from the compelling (II) plus questionable assumptions about negation like (EXC) and (EXH) (see e.g. Parsons [1987], pp. 9–11). However, contrary to such assessment, (DD) strikes me as equally compelling as (II): if x and y are discernible with respect to a property, how could they fail to be distinct? Moreover, the move in question does not apply at least to the semantic-equivalence extension of Evans' argument developed in section 2.

[25] (TR$^\vdash$) was more informally labelled 'transitivity' in section 2.

[26] (TR$^\vdash$) is a version of transitivity for logical consequence with a somewhat *intermediate* strength. It is *strong* in that, for example, it allows one to apply transitivity in the presence of *side premises and conclusions* (contrast with the principle saying that, if $\varphi \vdash \psi$ and $\psi \vdash \chi$ hold, then $\varphi \vdash \chi$ holds). It is *weak* in that, for example, it does not allow one to apply transitivity in order to dispense with intermediate premises taken *distributively* rather than *collectively* (contrast with the principle saying that, if $\Xi \vdash \Lambda, \Theta$ holds and, for every $\varphi \in \Theta$, $\Delta, \varphi \vdash \Gamma$ holds, then $\Delta, \Xi \vdash \Lambda, \Gamma$ holds). In fact, the naive theory of vagueness requires failures of transitivity *even in the absence of side premises and conclusions*, and, unsurprisingly, *even that weak version of transitivity fails in* $\mathbf{V_0}$ *and its like* (see Weir [2005] for a different family of non-transitive logics in which such weak version of transitivity is preserved). However, argument E requires transitivity *in the presence of side premises*, and so we focus on (TR$^\vdash$).

$x = y$, which is to say that there is a $\mathbf{V_0}$-model $\mathfrak{M}$ of the relevant fragment of the naive theory such that $\text{val}_{\mathfrak{M},\text{ass}}(\mathcal{B}(x = y)) \in D_{\mathfrak{M}}$ and $\text{val}_{\mathfrak{M},\text{ass}}(x \neq y) \notin T_{\mathfrak{M}}$ and a $\mathbf{V_0}$-model $\mathfrak{M}'$ of the relevant fragment of the naive theory such that $\text{val}_{\mathfrak{M}',\text{ass}}(\mathcal{B}(x = y)) \in D_{\mathfrak{M}'}$ and $\text{val}_{\mathfrak{M}',\text{ass}}(x = y) \notin T_{\mathfrak{M}'}$.

So let's consider the fragment of the naive theory of vagueness concerning Greg's transformation as described in section 1. Let's assume that $H\tau$ translates 'τ is human', g_i translates 'Greg$_i$' and τ' translates 'the substance canonically present at the time that is the successor of the time at which τ is canonically present' (with the understanding that, for every i, Greg$_i$ is the substance canonically present at t_i). Setting, for merely illustrative purposes, $k = 999,999$, the relevant fragment of the naive theory is then:

(H^p) Hg_0;

(H^n) $\neg Hg_{999,999}$;

(H^t) $\neg\exists x(Hx \wedge \neg Hx')$;

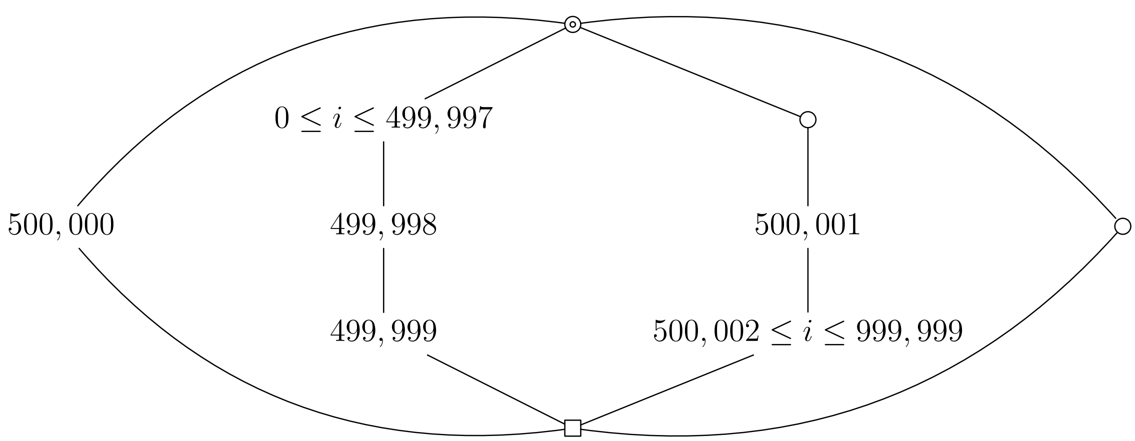
($=^n$) $g_0 \neq g_{999,999}$;

($=^t$) $\forall x(x = x')$,

and we can set the borderline identity to be $\mathcal{B}(g_0 = g_{499,998})$.

Theorem 5. *There is a* $\mathbf{V_0}$*-model* $\mathfrak{C}_0$ *such that* $\text{val}_{\mathfrak{C}_0,\text{ass}}((\text{H}^p))$, $\text{val}_{\mathfrak{C}_0,\text{ass}}((\text{H}^n))$, $\text{val}_{\mathfrak{C}_0,\text{ass}}((\text{H}^t))$, $\text{val}_{\mathfrak{C}_0,\text{ass}}((=^n))$, $\text{val}_{\mathfrak{C}_0,\text{ass}}((=^t))$ *and* $\text{val}_{\mathfrak{C}_0,\text{ass}}(\mathcal{B}(g_0 = g_{499,998})) \in D_{\mathfrak{C}_0}$ *while* $\text{val}_{\mathfrak{M},\text{ass}}(g_0 \neq g_{499,998}) \notin T_{\mathfrak{C}_0}$*, and there is a* $\mathbf{V_0}$*-model* $\mathfrak{C}_0'$ *such that* $\text{val}_{\mathfrak{C}_0',\text{ass}}((\text{H}^p))$, $\text{val}_{\mathfrak{C}_0',\text{ass}}((\text{H}^n))$, $\text{val}_{\mathfrak{C}_0',\text{ass}}((\text{H}^t))$, $\text{val}_{\mathfrak{C}_0',\text{ass}}((=^n))$, $\text{val}_{\mathfrak{C}_0',\text{ass}}((=^t))$ *and* $\text{val}_{\mathfrak{C}_0',\text{ass}}(\mathcal{B}(g_0 = g_{499,998})) \in D_{\mathfrak{C}_0'}$ *while* $\text{val}_{\mathfrak{C}_0',\text{ass}}(g_0 = g_{499,998}) \notin T_{\mathfrak{C}_0'}$.

Proof. Sketch for $\mathfrak{C}_0$. Take $\mathfrak{C}_0$ to be a $\mathbf{V_0}$-model such that, for every i, $\text{val}_{\mathfrak{C}_0,\text{ass}}(g_i = g_{i+1}) = \{0, 2, 4\}$ and, letting a number i occupying the place of a value X mean that $\text{val}_{\mathfrak{C}_0,\text{ass}}(Hg_i) = \text{val}_{\mathfrak{C}_0,\text{ass}}(g_0 = g_i) = X$, such that:



$\square$

Thus, at a *formal* level, one can adopt $\mathbf{V_0}$ and thereby be committed to accepting the claim that it is definite that Greg_0 is identical with Greg_0, accept the (H^p)–$(=^t)$-fragment of the naive theory of vagueness and the claim that it is borderline whether Greg_0 is identical with $\text{Greg}_{499,998}$, accept that that claim together with the previous claim that it is definite that Greg_0 is identical with Greg_0 entails that Greg_0 is distinct from $\text{Greg}_{499,998}$ and yet not be committed to accepting that Greg_0 is distinct from $\text{Greg}_{499,998}$ (or that Greg_0 is identical with $\text{Greg}_{499,998}$).

But how can we make sense of this position at a *philosophical* level? A natural way to do so starts by distinguishing *two grades* in which a sentence ψ can be (strictly or non-strictly) *weaker* than a sentence φ (see Zardini [2013b] for more details). Very roughly, the first grade of being weaker (which I'll denote with single starring) is present as soon as φ *could not hold without* ψ *holding*; the second grade of being weaker (which I'll denote with double starring) is present when ψ *could not be less (epistemically) likely than* φ. *Like* many other logics, tolerant logics are constructed with the aim of forcing the conclusion of a valid argument to be *weaker** than the premise, but, *unlike* many other logics, tolerant logics achieve that without forcing the conclusion to be *weaker*** than the premise—in tolerant logics, φ may entail ψ (and so it may be impossible that φ holds without ψ holding) even if ψ is less likely than φ. For example, in $\mathbf{V_0}$ $(\text{H}^t) \wedge Hg_{499,998} \vdash Hg_{499,999}$ holds (it is impossible that $(\text{H}^t) \wedge Hg_{499,998}$ holds without $Hg_{499,999}$ holding) even if $Hg_{499,999}$ may be less likely than $(\text{H}^t) \wedge Hg_{499,998}$ (for it is natural to think that the former may make a "weightier" claim than the latter, in that, contrary to it, it straightforwardly decides that $\text{Greg}_{499,999}$ is bald). Tolerant logics thus have the best of both the *non-deductive* and the *deductive* world: they allow for valid arguments in which the conclusion goes beyond the premise while ensuring that nevertheless the truth of the premise guarantees the truth of the conclusion. The non-transitivity of tolerant logics can then be seen to arise from this *mismatch* between the relation of being weaker** on the one hand and the relations of being weaker* and of logical consequence on the other hand, for, under these circumstances, chaining together enough valid arguments may eventually lead from a very likely (indeed, certainly true) initial premise to a very unlikely (indeed, certainly false) final conclusion.

Now, logic $\mathbf{V_0}$ encodes the idea that *definitely being is a mode of being that is not weaker than merely being*. For sentences about borderline cases like '$\text{Greg}_{499,998}$ is human', this feature is manifested as a relatively straightforward failure of definitely being to be weaker* than merely being: by the law of excluded middle,[27] either $\text{Greg}_{499,998}$ is human or $\text{Greg}_{499,998}$ is not human, but he is nevertheless neither a definite human nor a definite not human. A borderline case of humanity may be human (non-human), but, if it is human (non-human), it is not so much of a human (non-human) as to be a definite human (non-human). For logical truths like 'Greg_0 is identical with Greg_0', the feature is manifested not as any failure of definitely being to be weaker* than merely being (for every i, it is

[27]Notice that the law of excluded middle does hold in $\mathbf{V_0}$ and its like, being a straightforward consequence of (EXH).

definite^{i}[28] that Greg_0 is identical with Greg_0), but still it is manifested as a failure of definitely being to be weaker** than merely being. Given that it is a logical truth that Greg_0 is identical with Greg_0, by ($\text{N}^{\neg,\mathcal{D}}$) it is definite that Greg_0 is identical with Greg_0, but, although that could not fail to be so, it is nevertheless to some extent less likely than Greg_0's being identical with Greg_0. Similarly, under the assumption that $\text{Greg}_{499,998}$ is identical with Greg_0, by (II) it would follow from its being definite that Greg_0 is identical with Greg_0 that it is definite that $\text{Greg}_{499,998}$ is identical with Greg_0, but, although that could not fail to be so, it would nevertheless be to some extent less likely than its being definite that Greg_0 is identical with Greg_0. And, since it is in fact already less than certain that it is definite that Greg_0 is identical with Greg_0, the further uncertainty generated by (II) determines that, under the assumption that $\text{Greg}_{499,998}$ is identical with Greg_0, it could in fact be so unlikely that it is definite that $\text{Greg}_{499,998}$ is identical with Greg_0 that 'It is definite that $\text{Greg}_{499,998}$ is identical with Greg_0' cannot be taken to follow simply from '$\text{Greg}_{499,998}$ is identical with Greg_0'.

Relatedly, since, on this position, in general φ could hold without $\mathcal{D}\varphi$ holding, the argument from '$\text{Greg}_{499,998}$ is identical with Greg_0' to 'It is definite that $\text{Greg}_{499,998}$ is identical with Greg_0' would indeed be an argument whose conclusion is not only not weaker**, but also not weaker* than its premise, and so an invalid argument—unless it is rescued by a suitable finer-grained analysis of its logical form. What is supposed to come to the rescue is the observation that the premise is an identity, and the assumption that 'It is definite that Greg_0 is identical with Greg_0' as a further premise turning the argument into an instance of (II) can be *suppressed* as it is a logical truth. But, while all logical truths are indeed likely enough as to guarantee that they must hold, as we've seen in the previous paragraph some of them are not so likely that they can be suppressed in a valid argument.[29]

[28]Henceforth, 'definitely^{i}' or its like is the result of concatenating the empty string with i occurrences of 'definitely' or its like.

[29]It is not unusual to admit that the analogue of this can happen in a *non-deductive* case. Consider for example a non-deductive consequence relation $\Vdash\!\!\!\vdash$ relativised to a biographer's evidence (where $\Gamma \Vdash\!\!\!\vdash \varphi$ holds iff, given the biographer's evidence, the state of information represented by Γ lends credibility to φ), with the biographer's evidence being good enough as to lend credibility both to 'On 25/12/1915, Kafka had ham for dinner' and to 'On 26/12/1915, Kafka had sausages for dinner'. Then, while we have that both $\varnothing \Vdash\!\!\!\vdash$ 'On 25/12/1915, Kafka had ham for dinner' and $\varnothing \Vdash\!\!\!\vdash$ 'On 26/12/1915, Kafka had sausages for dinner' hold, and presumably also that 'On 25/12/1915, Kafka had ham for dinner', 'On 26/12/1915, Kafka had sausages for dinner' $\Vdash\!\!\!\vdash$ 'On 25/12/1915, Kafka had ham for dinner and, on 26/12/1915, Kafka had sausages for dinner' holds, it is a familiar point that 'On 25/12/1915, Kafka had ham for dinner' and 'On 26/12/1915, Kafka had sausages for dinner' cannot be suppressed in the last argument on pain of giving rise to the *preface paradox* (see Makinson [1965]). What is distinctive of the position under discussion is to maintain that the analogue of this can happen in the *deductive* case (I explore further the illuminating similarities between tolerant logics and certain non-deductive consequence relations in Zardini [2013b]).

6 The Conclusion of Argument E and the Properties of Definiteness

As I've already implied, a crucial feature of the construction in section 5 is that a relative of the converse of ($\mathrm{T}^{\subset}$):

($\mathrm{T}^{\vdash}$) $\varphi \vdash \mathcal{D}\varphi$ holds[30]

does not hold.

($\mathrm{T}^{\vdash}$), as well as its strengthening:

($\mathrm{T}^{\vdash^{\omega}}$) For every i, $\varphi \vdash \mathcal{D}^{i}\varphi$ holds,

is arguably necessary for an appealing conception of definiteness according to which *everything is definitely what it is* (although the argument lies beyond the scope of this paper), and, together with (EXC), it does imply (INC). Yet, it *has* to fail in the construction of section 5: for, by theorem 2, (EXC) and (EXH) are both crucial properties of negation in $\mathbf{V_0}$ and its like,[31] and they jointly imply that, if $\varphi \vdash \mathcal{D}\varphi$ holds, so does $\neg\mathcal{D}\varphi \vdash \neg\varphi$, so that ($\mathrm{T}^{\vdash}$) would lead straight away to the conclusion of argument E. But the construction of section 5 is predicated precisely on the assumption that argument E should fail, and, more generally, that its conclusion should not hold.

Recall the reasons in section 2 for that assumption. One reason (the one Evans himself seems to have had in mind) appealed to ($\mathrm{TR}^{\mathrm{ADJ}^{\mathrm{INC}}}$) (and (INC), which does hold if ($\mathrm{T}^{\vdash}$) and (EXC) do). But, just as with ($\mathrm{TR}^{\vdash}$), ($\mathrm{TR}^{\mathrm{ADJ}^{\mathrm{INC}}}$) does not hold in $\mathbf{V_0}$ and its like, and in fact *cannot possibly hold in any naive theory of vagueness adopting as tolerant logic* $\mathbf{V_0}$ *or one of its like*. In its essence, the reason is this (Zardini [2013c] gives more details). Given that, in $\mathbf{V_0}$ and its like, (H^{p}) $\wedge$ (H^{n}) $\wedge$ (H^{t}) $\vdash Hg_1$ and

[30]($\mathrm{T}^{\vdash}$) was more informally labelled 'fact entails definite fact' in section 2. The proper converse of ($\mathrm{T}^{\subset}$) is of course:

($\mathrm{T}^{\supset}$) $\varnothing \vdash \varphi \supset \mathcal{D}\varphi$ holds,

which is typically rejected even by logics of definiteness that accept ($\mathrm{T}^{\vdash}$) (as, by contraposition on material implication and a version of *modus ponens*, it yields the contrapositive of ($\mathrm{T}^{\vdash}$)). Given the usual definition of implication, in $\mathbf{V_0}$ and its like ($\mathrm{T}^{\supset}$) follows from ($\mathrm{T}^{\vdash}$) by (EXH) and the properties of disjunction in $\mathbf{V_0}$ and its like. I think that this is as it should it be, as ($\mathrm{T}^{\supset}$) no less than ($\mathrm{T}^{\vdash}$) is arguably integral to the appealing conception of definiteness that I'm about to introduce in the text. As for the alleged bad consequence of ($\mathrm{T}^{\supset}$) ($\neg\mathcal{D}\varphi \vdash \neg\varphi$), I'm going to argue in this section that, in a naive theory of vagueness adopting a tolerant logic, it is not such.

[31]I'm well aware that (EXH) is rejected by many non-classical logics of vagueness (and that even (EXC) is rejected by some of them). But, setting aside the question of the meaning and logic of the negative constructions ordinarily used in natural language, I think that we clearly do have a notion of a sentence φ *failing to hold* that, on the face of it, is exclusive and exhaustive with respect to φ. $\mathbf{V_0}$ is a theory, among other things, of that arguably theoretically fundamental notion and of the borderline cases that arise with respect to it.

$(\mathrm{H}^p) \wedge (\mathrm{H}^n) \wedge (\mathrm{H}^t) \wedge \neg Hg_2 \vdash \neg Hg_1$ hold (and that, by (EXC), $Hg_1, \neg Hg_1 \vdash \varnothing$ holds), by the properties of conjunction in $\mathbf{V_0}$ and its like and ($\mathrm{TR}^{\mathrm{ADJ}^{\mathrm{INC}}}$) it would follow that, in $\mathbf{V_0}$ and its like, $(\mathrm{H}^p) \wedge (\mathrm{H}^n) \wedge (\mathrm{H}^t) \wedge \neg Hg_2 \vdash \varnothing$ holds, and so, by the properties of conjunction and negation in $\mathbf{V_0}$ and its like, it would follow that, in $\mathbf{V_0}$ and its like, $(\mathrm{H}^p) \wedge (\mathrm{H}^n) \wedge (\mathrm{H}^t) \vdash Hg_2$ holds. Applying analogous arguments another 999,997 times, we would reach the tragic conclusion that, in $\mathbf{V_0}$ and its like, $(\mathrm{H}^p) \wedge (\mathrm{H}^n) \wedge (\mathrm{H}^t) \vdash Hg_{999{,}999}$ holds (and so, since, in $\mathbf{V_0}$ and its like, $(\mathrm{H}^p) \wedge (\mathrm{H}^n) \wedge (\mathrm{H}^t) \vdash \neg Hg_{999{,}999}$ holds, by ($\mathrm{TR}^{\mathrm{ADJ}^{\mathrm{INC}}}$) we would reach the even more tragic conclusion that, in $\mathbf{V_0}$ and its like, $(\mathrm{H}^p) \wedge (\mathrm{H}^n) \wedge (\mathrm{H}^t) \vdash \varnothing$ holds).

Another reason for the assumption that the conclusion of argument E should not hold appealed to (CCTLC). But, just as with ($\mathrm{TR}^{\vdash}$), (CCTLC) cannot possibly hold in any naive theory of vagueness adopting a tolerant logic. For, clearly, it is crucial for such a theory sometimes *to accept* φ*, accept that* φ *entails* ψ *but not to accept* ψ (otherwise, since the naive theory accepts $(\mathrm{H}^p) \wedge (\mathrm{H}^t)$, as well as the relevant logical principles like *modus ponens*, it will be committed to accepting $\neg(\mathrm{H}^n)$). Generally, a naive theory of vagueness adopting a tolerant logic will want to draw a distinction between, very roughly, acceptance for *non-inferential* reasons and acceptance for *inferential* reasons, and will want to hold that, if φ entails ψ, if one accepts φ one is committed [to accepting ψ] only if one accepts φ for non-inferential reasons, while that may not be so if one merely accepts φ for inferential reasons. Similar points apply for attitudes other than acceptance, like conditional acceptance, supposition, desire etc. (see fn 32; Zardini [2013b] offers an extended discussion of the normative import of non-transitive logical consequence).

So, in a naive theory of vagueness adopting a tolerant logic, the two main reasons against the conclusion of argument E no longer have force, this being so for the second reason at least if one does not have non-inferential reasons for accepting $\mathcal{B}(x = y)$. And it does seem to be the case that, *quite generally, one does not have non-inferential reasons for accepting* $\mathcal{B}\varphi$. For such reasons are typically supposed to consist in the fact that, roughly, there is something to be said in favour of φ and also something to be said against φ; but, at least if ($\mathrm{T}^{\vdash}$) holds, if there is something to be said in favour of φ (against φ), there is something to be said in favour of $\mathcal{D}\varphi$ ($\mathcal{D}\neg\varphi$), and so something to be said against $\mathcal{B}\varphi$, which plausibly defeats the putative reason in favour of $\mathcal{B}\varphi$. This claim about borderline cases also dovetails nicely with the way in which the notion has been introduced in section 1. For, on the approach of section 1, our immediate reason in favour of borderline cases is a *general* one concerning overall features of soritical series: on the way from Greg_0 to Greg_k, it cannot be the case that, for no i, it is borderline whether Greg_0 is identical with Greg_i (for whence could the vagueness of the transformation then arise?), and so it has to be the case that, *for some* i, it is borderline whether Greg_0 is identical with Greg_i. Thus, on that approach, our immediate reason in favour of borderline cases is not a *particular* one concerning a specific i to the effect that it is borderline whether Greg_0 is identical with Greg_i, and so, since any non-inferential reason would count as an immediate reason, it follows that, on that approach, for no i does one have non-inferential reasons for accepting that it is borderline whether whether Greg_0 is identical with Greg_i.[32]

[32]I've briefly argued in the text that, for no i, one has non-inferential reasons for accepting $\mathcal{B}(g_0 = g_i)$. If one does not have inferential reasons either for accepting it, that is obviously sufficient for undercutting

It still remains the case that, as I've noted in section 2, the conclusion of argument E would seem to imply that *nothing is a definite borderline case of identity*. However, as I've also noted in section 2, that implication only holds if (M) holds. But, quite obviously, on a view accepting the contrapositive of ($\mathrm{T}^{\vdash}$) just as there are reasons *independent of argument* E for accepting its conclusion—reasons that rely rather on ($\mathrm{T}^{\vdash}$), (EXC) and (EXH)—so there may be reasons *independent of argument* E for rejecting (M) in its full generality—reasons that would basically rely rather on ($\mathrm{T}^{\vdash}$), (EXC), (EXH) and the very plausible claim that there can be definite borderline cases of some kind or other. For, by ($\mathrm{T}^{\vdash}$), (EXC) and (EXH), $\neg\mathcal{D}\varphi \vdash \neg\varphi$ holds, and so, by (M), $\mathcal{D}\neg\mathcal{D}\varphi \vdash \mathcal{D}\neg\varphi$ holds, from which one would expect it to follow, even in a tolerant logic, that $\mathcal{D}\neg\mathcal{D}\varphi \vdash \neg\mathcal{D}\mathcal{B}\varphi$ holds. But one would also expect, even in a tolerant logic, that $\mathcal{D}\mathcal{B}\varphi \vdash \mathcal{D}\neg\mathcal{D}\varphi$ holds, and that it does so in such a way as to combine with the previous valid argument to yield $\mathcal{D}\mathcal{B}\varphi \vdash \neg\mathcal{D}\mathcal{B}\varphi$. Therefore, by (EXC) (and intersubstitutability of a sentence with its double negation), $\mathcal{D}\mathcal{B}\varphi \vdash \varnothing$ holds, and so, by (EXH), $\varnothing \vdash \neg\mathcal{D}\mathcal{B}\varphi$ holds. Thus, on the view under consideration, for reasons independent of argument E (M) implies that *nothing is a definite borderline case of any kind whatsoever*. On the view under consideration then, if there are reasons for rejecting the claim that nothing is a definite borderline case of identity on the grounds that, in general, there can be definite borderline cases, these should be taken to be reasons for rejecting (M) rather than reasons for rejecting the conclusion of argument E.

An even more general reason for rejecting (M) that does not rely on ($\mathrm{T}^{\vdash}$) (nor on (EXC) or (EXH)) emerges if we consider the specific situation of a soritical series. Let's focus again on Greg's transformation. Because of higher-order vagueness, it is extremely

---

the *letter* of the second reason against the conclusion of argument E, quite independently of all the previous stuff in the text about (CCTLC). But, even assuming that one does not have inferential reasons either for accepting $\mathcal{B}(g_0 = g_i)$, the richness of articulation afforded by the normative theory which accompanies a naive theory of vagueness adopting a tolerant logic and which leads to the rejection of (CCTLC) is needed in order to undercut subtler renderings of the *spirit* of the second reason against the conclusion of argument E. For example, given that (CCTLC) has deliberately been formulated in terms of the broader notion of *thinking* (which encompasses also attitudes other than acceptance), the foe of the conclusion of argument E could observe that, in $\mathbf{V_0}$ and its like, given the obvious additional constraints on the domain $\exists x\mathcal{B}(g_0 = x) \vdash \mathcal{B}(g_0 = g_0), \mathcal{B}(g_0 = g_1), \mathcal{B}(g_0 = g_2) \ldots, \mathcal{B}(g_0 = g_{999,999})$ holds, and that there should be no objection in a naive theory adopting a tolerant logic to the normative principle concerning multiple-conclusion arguments saying that, if one has non-inferential reasons to accept all the premises of a valid argument, one has inferential reasons for *conditionally accepting each of its conclusions* (that is, *conditionally on the rejection for non-inferential reasons of all the other conclusions*). Letting conditional acceptance be the relevant mode of thinking, the foe of the conclusion of argument E could then use the conclusion of argument E and (CCTLC) to infer that, for every i, one is committed to accepting conditionally $g_0 \neq g_i$ (that is, conditionally, for every other j, on the rejection for non-inferential reasons of $\mathcal{B}(g_0 = g_j)$). These are already in themselves rebarbative commitments. And, if the logic of definiteness also yields $x \neq y, \neg\mathcal{D}(x \neq y) \vdash \varnothing$ (see section 1), the foe of the conclusion of argument E could turn those commitments into downright unacceptable ones by inferring that, for every i, one is committed to accepting conditionally the jointly inconsistent $g_0 \neq g_i$ and $\neg\mathcal{D}(g_0 \neq g_i)$ (again, conditionally, for every other j, on the rejection for non-inferential reasons of $\mathcal{B}(g_0 = g_j)$), from which it follows that, unacceptably, all options open to one are unacceptable options on which one is committed to accepting two jointly inconsistent sentences. This more sophisticated argument is effectively blocked by rejecting (CCTLC) on the grounds adduced in the text.

plausible to think that, for every open formula φ containing only H, $\neg$ and $\mathcal{D}$, φ is vague in that situation. Now, here is, somewhat roughly stated, a plausible requirement of *closure under no definite sharp boundaries* on our theory Γ of Greg's transformation:

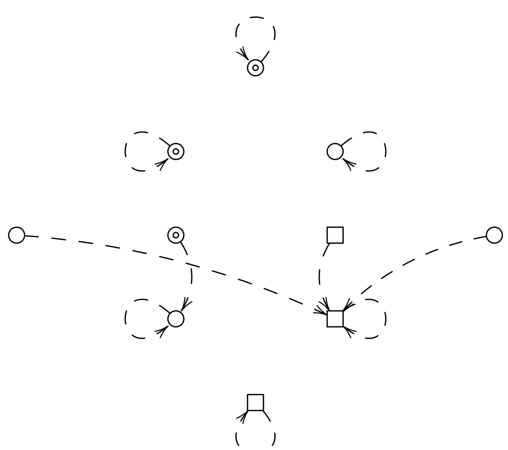
(CNDSB) For every i, if φ is a vague open formula (with respect to ξ) and $\Gamma \vdash \mathcal{D}\varphi[g_i/\xi]$ holds, then $\Gamma \vdash \neg\mathcal{D}\neg\varphi[g_{i+1}/\xi]$ holds.

In other words, for everything that our theory entails to be a candidate for being a point at which the transition from a vague property to its negation is not vague, our theory also entails that it is after all not such a point. However, in virtually every logic of vagueness (including tolerant logics), (CNDSB) is inconsistent with (M), at least given the two extremely plausible assumptions that, for a large "enough" i, $\mathcal{D}^i Hg_0$ as well as $\mathcal{D}^i\neg Hg_{999,999}$ hold and that there is "enough" higher-order vagueness (see in particular Zardini [2013a] for the precise sense of 'enough', for the details of the demonstration and for a discussion of its assumptions).[33]

7 The Tolerant Logic $\mathbf{V_1}$

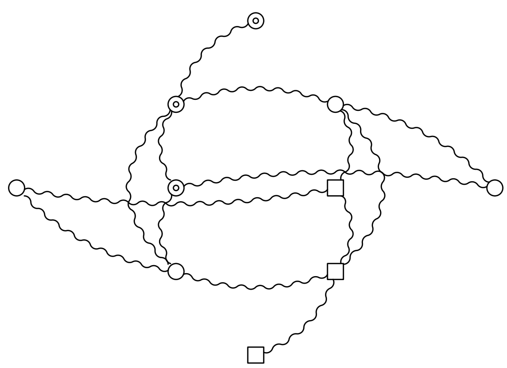
In section 6, I've argued in effect that all of ($\mathrm{T}^{\vdash}$), ($\mathrm{T}^{\vdash^{\omega}}$) and ($\mathrm{T}^{\supset}$) may after all still be acceptable in a naive theory of vagueness adopting a tolerant logic. Let's see then how they can be added to $\mathbf{V_0}$ (thus giving rise to the logic $\mathbf{V_1}$).

Definition 3. A $\mathbf{V_1}$-model $\mathfrak{M}$ is a 9ple $\langle U_{\mathfrak{M}}, V_{\mathfrak{M}}, \preceq_{\mathfrak{M}}, D_{\mathfrak{M}}, T_{\mathfrak{M}}, \mathrm{neg}_{\mathfrak{M}}, \mathrm{def}_{\mathfrak{M}}, \mathrm{id}_{\mathfrak{M}}, \mathrm{int}_{\mathfrak{M}}\rangle$. $U_{\mathfrak{M}}$, $V_{\mathfrak{M}}$, $\preceq_{\mathfrak{M}}$, $D_{\mathfrak{M}}$, $T_{\mathfrak{M}}$, $\mathrm{neg}_{\mathfrak{M}}$ and $\mathrm{int}_{\mathfrak{M}}$ are as *per* definition 1. $\mathrm{def}_{\mathfrak{M}}$ is an operation of the same kind as *per* definition 1 and can be depicted as:



[33](M) may be rescued from the problems discussed in the text if one were to go tolerant *in reasoning about* $\vdash$ *itself*. But, on a natural way of implementing it, this move would equally block as tolerantly invalid the reasoning based on the conclusion of argument E and (M) to the effect that nothing is a definite borderline case of identity.

$\mathrm{id}_{\mathfrak{M}}$ is an operation of the same kind as *per* definition 1, with the modification that, for every $u \in U_{\mathfrak{M}}$, $\{0,1,2,3,4,5,6,7\} = \mathrm{id}_{\mathfrak{M}}(u,u)$ and with the relevant deviations being:



Again, $\mathrm{int}_{\mathfrak{M}}$ can be extended to a full valuation function $\mathrm{val}_{\mathfrak{M}}$ (relative to assignments) in the usual way. With such function in place, the same style of definition of logical consequence yields:

Definition 4. $\Gamma \vdash_{\mathbf{V_1}} \Delta$ holds iff, for every $\mathbf{V_1}$-model $\mathfrak{M}$ and assignment ass, if, for every $\varphi \in \Gamma$, $\mathrm{val}_{\mathfrak{M},\mathrm{ass}}(\varphi) \in D_{\mathfrak{M}}$, then, for some $\psi \in \Delta$, $\mathrm{val}_{\mathfrak{M},\mathrm{ass}}(\psi) \in T_{\mathfrak{M}}$.

$\mathbf{V_1}$ enjoys all the properties recorded for $\mathbf{V_0}$ in section 4 (and we can now explicitly note that (M) fails both in $\mathbf{V_0}$ and in $\mathbf{V_1}$); moreover, contrary to what is the case for $\mathbf{V_0}$, in $\mathbf{V_1}$ ($\mathrm{T}^{\vdash}$), ($\mathrm{T}^{\vdash^{\omega}}$) and ($\mathrm{T}^{\supset}$) hold and the argument $\varnothing \vdash \mathcal{D}(\tau = \tau)$ can no longer give rise to failures of ($\mathrm{TR}^{\vdash}$) (and so $\mathcal{D}(\tau = \tau)$ can be suppressed).[34]

8 The Harmlessness of Argument E in $\mathbf{V_1}$

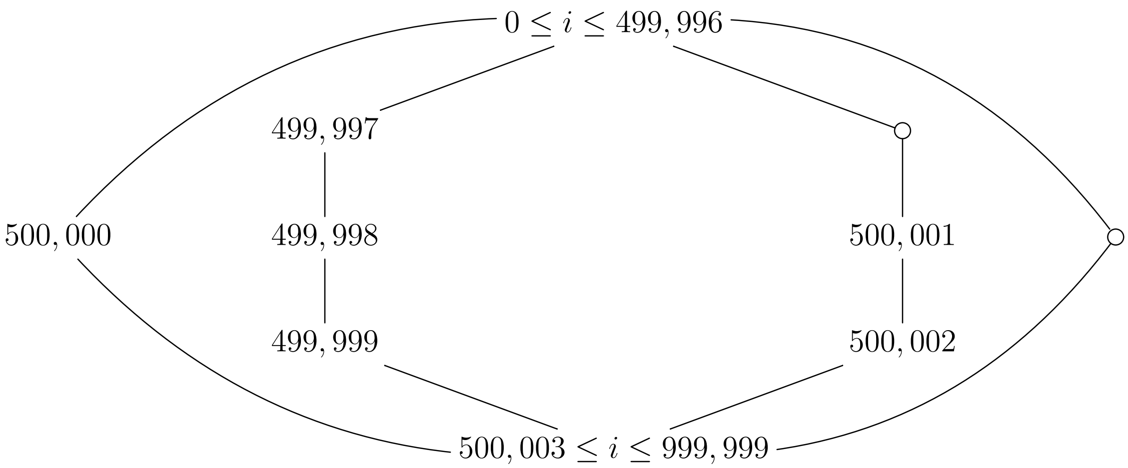
The upshot of section 6 was that, in a naive theory of vagueness adopting a tolerant logic, the two main reasons against the conclusion of argument E no longer have force, and that, very plausibly, in every theory accepting ($\mathrm{T}^{\vdash}$), (EXC) and (EXH) (and, plausibly, in virtually every theory whatsoever), that conclusion does not imply that nothing is a definite borderline case of identity. Focussing now on a naive theory adopting $\mathbf{V_1}$, what needs to be shown is, in line with one of the main contentions of section 6, the consistency in $\mathbf{V_1}$ of the claim that, for some object or other, it is definite that it is borderline whether Greg_0 is identical with it, which is to say that there is a $\mathbf{V_1}$-model $\mathfrak{M}$ such that $\mathrm{val}_{\mathfrak{M},\mathrm{ass}}(\exists x \mathcal{D}\mathcal{B}(g_0 = x)) \in D_{\mathfrak{M}}$. Indeed, and again in line with one of the main contentions of section 6, to make sure that the existence of a definite borderline identity does not have problematic consequences for the naive theory, and that, even together with

[34]All this implies that not only does the conclusion of argument E hold in $\mathbf{V_1}$, but also argument E itself is valid.

it, it remains silent on whether any specific identity is borderline, we should show that, in $\mathbf{V_1}$, $\exists x \mathcal{D}\mathcal{B}(g_0 = x)$ together with the (H^p)–($=^t$)-fragment of the naive theory does not entail $\mathcal{B}(g_0 = g_i)$ for any i, which is to say that, for every i, there is a $\mathbf{V_1}$-model $\mathfrak{M}$ of the (H^p)–($=^t$)-fragment of the naive theory such that $\text{val}_{\mathfrak{M},\text{ass}}(\exists x \mathcal{D}\mathcal{B}(g_0 = x)) \in D_{\mathfrak{M}}$ and $\text{val}_{\mathfrak{M},\text{ass}}(\mathcal{B}(g_0 = g_i)) \notin T_{\mathfrak{M}}$.

Theorem 6. *For every i, there is a $\mathbf{V_1}$-model $\mathfrak{C}_1$ such that* $\text{val}_{\mathfrak{C}_1,\text{ass}}((\text{H}^p))$, $\text{val}_{\mathfrak{C}_1,\text{ass}}((\text{H}^n))$, $\text{val}_{\mathfrak{C}_1,\text{ass}}((\text{H}^t))$, $\text{val}_{\mathfrak{C}_1,\text{ass}}((=^n))$, $\text{val}_{\mathfrak{C}_1,\text{ass}}((=^t))$ *and* $\text{val}_{\mathfrak{C}_1,\text{ass}}(\exists x \mathcal{D}\mathcal{B}(g_0 = x)) \in D_{\mathfrak{C}_1}$ *while* $\text{val}_{\mathfrak{C}_1,\text{ass}}(\mathcal{B}(g_0 = g_i)) \notin T_{\mathfrak{C}_1}$.

Proof. Sketch for $i \neq 500,000$. Take $\mathfrak{C}_1$ to be a $\mathbf{V_1}$-model such that, for every j, $\text{val}_{\mathfrak{C}_1,\text{ass}}(g_j = g_{j+1}) = \{0, 1, 2, 3, 4, 5, 6, 7\}$ and, letting a number j occupying the place of a value X mean that $\text{val}_{\mathfrak{C}_1,\text{ass}}(Hg_j) = \text{val}_{\mathfrak{C}_1,\text{ass}}(g_0 = g_j) = X$, such that:



□

Thus, at a *formal* level, one can adopt $\mathbf{V_1}$ and thereby be committed to accepting the claim that it is definite that Greg$_0$ is identical with Greg$_0$, accept the (H^p)–($=^t$)-fragment of the naive theory of vagueness and the claim that, for some i, it is definite that it is borderline whether Greg$_0$ is identical with Greg$_i$, accept that, for every i, the claim that it is borderline whether Greg$_0$ is identical with Greg$_i$ entails that Greg$_0$ is distinct from Greg$_i$ (and entails that Greg$_0$ is identical with Greg$_i$) and yet, for no i, be committed to accepting that Greg$_0$ is distinct from Greg$_i$ (or that Greg$_0$ is identical with Greg$_i$).

But how can we make sense of this position at a *philosophical* level? The position is somehow the opposite of the position discussed in section 5, as it encodes the idea that, by (T$^{\vdash}$) and (T$^{\supset}$), definitely being is a mode of being that is weaker* than merely being (and that, by (T$^{\subset}$), merely being is a mode of being that is weaker* than definitely being). On this position, for example, nothing could be human without being a definite human—there is no limbo of degenerated substances that are human but not so much as to be definite humans. On a natural way to support this intuitively appealing claim, that is so because, to be human, one must be marked as *different enough* from what is not human, and so be a definite human. Every human—no matter how deformed—is a definite human. More generally, and as already suggested in section 6, everything is definitely what it is.

The contrapositive of ($T^{\vdash}$) follows just as naturally from these considerations: for example, if one is not a definite human, one is not marked as different enough from what is not human, and so one is not human after all. And that can in fact most clearly be seen as an application of what it is for 'not human' to be *tolerant*: if x is not human and y is not very different from x, y is also not human. Being thus arguably grounded in tolerance, ($T^{\vdash}$) exhibits the same feature exhibited by other valid arguments that are so grounded: although its conclusion is weaker* than its premise, it is not weaker**, and so ($T^{\vdash}$) can give rise to failures of ($TR^{\vdash}$) (as it should in order for the defence offered in section 6 to be ultimately viable).[35]

All this may sound alright in itself, but can this conception of definiteness still square with the claim with an instance of which this paper began and which passed through the sieve of the criticisms in section 6—that is, with the claim that there are borderline cases? It can. Firstly, the conception of definiteness in question still supports the justification for that claim: if, for every i, either Greg_i is marked as different enough from what is not human or Greg_i is marked as different enough from what is human, whence could the vagueness of human Greg_0's transformation into not human Greg_k arise? Secondly, as was already implicit in our discussion of ($TR^{\text{ADJ}^{\text{INC}}}$) in section 6, the conception of definiteness in question only implies [that, under the assumption that it is borderline whether Greg_i is human, it follows that Greg_i is not human] and [that, under the same assumption, it follows that Greg_i is human]. It does not imply that, under the same assumption, it follows that [Greg_i is not human and Greg_i is human]. Because it does not imply that (nor any other inconsistency), the conception of definiteness in question does not imply that the assumption that it is borderline whether Greg_i is human is inconsistent (which, by (EXH), would in turn imply that it is not borderline whether Greg_i is human). And so it does not imply a sort of "n-inconsistency" in which, although it is asserted that something among a finite number n of things is thus and so, it is also asserted of each of these things that it is not thus and so (a pattern that, under minimal assumptions, would lead to a straightforward inconsistency).[36] Our concept of definiteness pulls in two different directions: on the one hand, everything that is some way is definitely that

[35] Thus, generally, both $\mathbf{V_0}$ and $\mathbf{V_1}$ think that the problem with Evans' argument consists in the fact that $\mathcal{D}\varphi$ is not weaker** than φ. The particular instance on which $\mathbf{V_0}$ focusses is $\mathcal{D}(x = x)$ not being weaker** than $x = x$ while, contrary to what $\mathbf{V_0}$ thinks happens in other instances, being weaker* than it (with the consequence that the former cannot be suppressed in a valid argument although it is a logical truth). On the contrary, $\mathbf{V_1}$ thinks that $\mathcal{D}(x = x)$ is not only weaker*, but also weaker** than $x = x$ (with the consequence that the former can be suppressed in a valid argument), and focusses instead on $\mathcal{D}(x = y)$ not being weaker** than $x = y$ while being weaker* than it. I'm grateful to Krzysztof Posłajko for questions that led to this observation.

[36] Nor does it imply a sort of "supervaluationist thingy" in which, although it is asserted that it is definite that something among a finite number n of things is thus and so, it is also asserted of each of these things that it is not definite that it is thus and so (a pattern that, under minimal assumptions,

way; on the other hand, vague transitions require that there be things that are neither definitely one way nor definitely the other way. And while these two different directions are indeed contradictory in virtually every other logic of vagueness, they are not so in tolerant logics, and can in fact be upheld in a tolerant logic like $\mathbf{V_1}$.

To sum up, on this position 'It is borderline whether Greg_0 is identical with Greg_i' does entail 'Greg_0 is distinct from Greg_i'. Evans was right about that. He was even right for correct reasons: argument E is valid. But, although correct, those reasons are fatally misleading insofar as they suggest that the entailment is due to peculiar features of identity: rather, it is a completely general fact that 'It is borderline whether P' entails 'It is not the case that P' (and that it also entails 'P'). What Evans with his argument was definitely wrong about was to assume, implicitly appealing to principles untenable in a naive theory of vagueness adopting a tolerant logic, that these results show that borderline identity is impossible.

TERMS TO BE INDEXED

acceptance; adjunction; borderline cases; closure of [commitment, definiteness] under logical consequence; closure under no definite sharp boundaries; compositionality of semantic value; conditional acceptance; conjunction; [deductive, non-deductive] consequence relations; contraposition; definiteness; definitisation; [designated, tolerated] values; distinctness; distinctness of discernibles; disjunction; essentialism; [exclusivity, exhaustivity] of negation; Evans' argument; exemplification; extended simples; higher-order vagueness; implication; identification; identity; identity predicate; indiscernibility of [identical pluralities, identicals]; [interpretation, valuation] function; intersubstitutability of [material, semantic] equivalents; KB-logic for definiteness; K-logic for definiteness; lattices; law of excluded middle; likelihood; location; logical truths; [material, semantic] equivalence; modal logic; models; modus ponens; monotonicity of definite being-some-of over being-some-of; monotonicity of definite parthood over parthood; monotonicity of definite positive location over [exemplification, parthood]; naive theory of vagueness; necessity; negation; non-composite properties; [non-inferential, inferential] reasons; [particular, plural, universal] quantification; predication; preface paradox; properties; reference; side premises and conclusions; sorites; soritical series; supervaluationism; tolerant logics; S5-logic for definiteness; substances; reflexivity of [material, semantic] equivalence; reflexivity of [being-some-of, identity, parthood, plural identity]; suppression of logical truths; [temporal, modal, spatial] parts; transformations; transitivity of [identity, logical consequence]; two grades of being weaker; uncertainty; vagueness; vague objects

References

Gareth Evans. Can there be vague objects? *Analysis*, 38:208, 1978.

($\mathrm{T}^{\vdash}$), (EXC) and (EXH), would lead to a straightforward inconsistency). For example, as we've seen, the position under discussion asserts not only $\mathcal{D}\exists x\mathcal{B}Hx$, but also $\mathcal{D}\exists x\mathcal{D}\mathcal{B}Hx$, and, for many is, it does not assert $\neg\mathcal{D}\mathcal{B}Hg_i$ (although for no i does it assert $\mathcal{D}\mathcal{B}Hg_i$).

Hartry Field. Indeterminacy, degree of belief, and excluded middle. *Nous*, 34:1–30, 2000.

Saul Kripke. Identity and necessity. In Milton Munitz, editor, *Identity and Individuation*, pages 135–164. New York University Press, New York, 1971.

David Lewis. *On the Plurality of Worlds*. Blackwell, Oxford, 1986.

David Lewis. Vague identity: Evans misunderstood. *Analysis*, 48:128–130, 1988.

David Makinson. The paradox of the preface. *Analysis*, 25:205–207, 1965.

Ned Markosian. Simples. *Australasian Journal of Philosophy*, 76:213–228, 1998.

Harold Noonan. Vague objects. *Analysis*, 42:3–6, 1982.

Terence Parsons. Entities without identity. *Philosophical Perspectives*, 1:1–19, 1987.

Graham Priest. Sorites and identity. *Logique et Analyse*, 135:293–296, 1991.

Nathan Salmon. *Reference and Essence*. Prometheus Books, Amherst NY, 2nd edition, 2005.

Paula Sweeney and Elia Zardini. Vagueness and practical interest. In Paul Égré and Nathan Klinedinst, editors, *Vagueness and Language Use*, pages 249–282. Palgrave MacMillan, Basingstoke, 2011.

Brian Weatherson. Stages, worms, slices and lumps. ms, 2013.

Alan Weir. Naive truth and sophisticated logic. In Bradley Armour-Garb and JC Beall, editors, *Deflationism and Paradox*, pages 218–249. Oxford University Press, Oxford, 2005.

Timothy Williamson. *Vagueness*. Routledge, London, 1994.

Crispin Wright. On the coherence of vague predicates. *Synthese*, 30:325–365, 1975.

Elia Zardini. Squeezing and stretching: How vagueness can outrun borderlineness. *Proceedings of the Aristotelian Society*, 106:419–426, 2006a.

Elia Zardini. Higher-order vagueness and paradox: The glory and misery of S4 definiteness. In Jurğis Šķilters, Matti Eklund, Ólafur Jónsson, and Olav Wiegand, editors, *Paradox: Logical, Cognitive and Communicative Aspects*, volume I of *The Baltic International Yearbook of Cognition, Logic and Communication*, pages 203–220. University of Latvia Press, Riga, 2006b.

Elia Zardini. A model of tolerance. *Studia Logica*, 90:337–368, 2008a.

Elia Zardini. *Living on the Slippery Slope. The Nature, Sources and Logic of Vagueness*. PhD thesis, Department of Logic and Metaphysics, University of St Andrews, 2008b.

Elia Zardini. Towards first-order tolerant logics. In Oleg Prozorov, editor, *Philosophy, Mathematics, Linguistics: Aspects of Interaction*, pages 35–38. Russian Academy of Sciences Press, St Petersburg, 2009.

Elia Zardini. Higher-order sorites paradox. *Journal of Philosophical Logic*, 42:25–48, 2013a.

Elia Zardini. Breaking the chains. Following-from and transitivity. In Colin Caret and Ole Hjortland, editors, *Foundations of Logical Consequence*. Oxford University Press, Oxford, 2013b. Forthcoming.

Elia Zardini. Getting one for two, or the contractors' bad deal. Towards a unified solution to the semantic paradoxes. In Theodora Achourioti, Kentaro Fujimoto, Henri Galinon, and José Martínez, editors, *Unifying the Philosophy of Truth*. Springer, Berlin, 2013c. Forthcoming.

Elia Zardini. Seconde naïveté. ms, 2013d.