Vagueness, tolerance and non-transitive entailment

Robert van Rooij *

1 Tolerance and vagueness

Vagueness is standardly opposed to *precision*. Just as gradable adjectives like 'tall' and a quantity modifier like 'a lot' are prototypical vague expressions, mathematical adjectives like 'rectangular', and measure phrases like '1.80 meter' are prototypically precise. But what does it mean for these latter expressions to be precise? On first thought it just means that they are precise, because they have an exact mathematical definition. However, if we want to use these terms to talk about observable objects, it is clear that these mathematical definitions would be useless: if they exist at all, we cannot possibly determine what are the rectangular objects in the precise geometrical sense, or objects that are exactly 1.80 meters long. For this reason, one allows for a margin of measurement error, or a threshold, in physics, psychophysics and other sciences. Assuming that the predicates we use are observational predicates gives rise to another consequence as well. If statements like 'the length of stick S is 1.45 meters' come with a large enough margin of error, the circumstances in which this statement can be made appropriately (or truly, if you don't want the notion of truth to be empty) might *overlap* with the circumstances in which the statement 'the

*The main ideas of this paper were first presented in a workshop on vagueness at Pamplona, Spain in June, 2009. Paul Egre acted as a commentator on this paper and soon 'joined' the project. Shortly after, Pablo Cobreros and Dave Ripley joined the project as well, and thanks to them I now have a much better understanding of what I was actually proposing in section 4 of this paper. I thank them for this, but in this paper I tried to stay as close as possible to my original contribution to the Pamplona workshop. Nevertheless, I still got rid of some needless complications, and used already some terminology that is used in our joint work as well. We hope to report soon on the progress (technically and conceptually) made on our joint project in another paper. The original idea of section 4 came up during a talk of Elia Zardini, when I was trying to understand in my own terms what he was proposing.

length of stick S is 1.50 meters'. Thus, although the predicates 'being a stick of 1.45 meters' and 'being a stick of 1.50 meters' are inconsistent in case no errors are made, the predicates might well be applicable to the same object in case a margin of error is taken into account, i.e., in case the predicates are interpreted tolerantly.[1] Thus, although the standard, i.e. precise, semantic meanings of two predicates might be *in*compatible, when one or both of these observational predicates are more realistically be interpreted in a tolerant way, they might well be *com*patible.

A traditional way of thinking about vagueness is in terms of the *existence of borderline cases*. John is a borderline case of a tall man, if the sentence 'John is a tall man' is neither (clearly) true nor (clearly) false. As a result, predicates like 'tall' and 'bald' do not give rise to a two-fold, but rather to a three-fold partition of objects: the positive ones, the negatives ones, and the borderline cases. Authors like Dummett (1975), Wright (1975), Kamp (1981), and others have argued that the existence of borderline cases is inadequate to characterize vagueness. Instead, what we have to realize is that these predicates are *observational predicates* that give rise to *tolerance*: a vague predicate is insensitive to very small changes in the objects to which it can be meaningfully predicated.

If being tolerant to small changes is indeed constitutive to the meaning of vague predicates, it seems that most approaches to vagueness went wrong. Trying to account for the Sorites paradox, these approaches mostly claim that the inductive premise is false. But it is exactly this inductive premise that states that the relevant predicate is tolerant. In this paper I would argue, instead, that the tolerance principle is valid with respect to a natural notion of truth and consequence. What we should give up is the idea that this notion of consequence is transitive. In this paper I will first introduce semi-orders and non-transitive similarity relations. In terms of that, I first discuss traditional approaches to vagueness, before I introduce my own account. Later, I show that my analysis is still closely related to other analyses. In the last section I will relate what I do with more general theories of concept-analysis in cognitive theories of meaning.

[1] Wheeler (2002) rightly argues, in my opinion, that allowing for measurement errors is perhaps the most natural way to motivate paraconsistency in logic.

2 The Sorites and semi-orders

Consider a long series of people ordered in terms of their height. Of each of them you are asked whether they are tall or not. We assume that the variance between two subsequent persons is always *indistinguishable*. Now, if you decide that the first individual presented to you, the tallest, is tall, it seems only reasonable to judge the second individual to be tall as well, since you cannot distinguish their heights. But, then, by the same token, the third person must be tall as well, and so on indefinitely. In particular, this makes also the last person tall, which is a counterintuitive conclusion, given that it is in contradiction with our intuition that this last, and shortest individual, is short, and thus not tall.

This so-called Sorites reasoning is elementary, based only on our intuition that the first individual is tall, the last short, and the following inductive premise, which seems unobjectable:

> (**P**) If you call one individual tall, and this individual is not visibly taller than another individual, you have to call the other one tall too.

Our above Sorites reasoning involved the predicate 'tall', but that was obviously not essential. Take any predicate P that gives rise to a complete ordering 'as P than'. Let us assume that '$\sim_P$' is the indistinguishability, or indifference, relation between individuals with respect to predicate P. Now we can state the inductive premise somewhat more formally as follows:

> (**P**) For any $x, y \in X : (Px \wedge x \sim_P y) \rightarrow Py$.

If we assume that it is possible that $\exists x_1, \ldots, x_n : x_1 \sim_P x_2 \wedge \ldots \wedge x_{n-1} \sim_P x_n$, but $P(x_1)$ and $\neg P(x_n)$, the paradox will arise. It immediately follows that the relation $\sim_P$ cannot be an equivalence relation. It is natural to define the indifference relation $\sim_P$ from an ordering relation 'P-er than', $\succ_P$. For many purposes it is natural to let the relation $\succ_P$ be a strict weak order:

Definition 1 *A (strict) weak order is a structure $\langle D, R\rangle$, with R a binary relation on D that is irreflexive (IR), transitive (TR), and almost connected(AC):*
(IR) $\forall x : \neg R(x, x)$.
(TR) $\forall x, y, z : (R(x, y) \wedge R(y, z)) \rightarrow R(x, z)$.
(AC) $\forall x, y, z : R(x, y) \rightarrow (R(x, z) \vee R(z, y))$.

If we now define the indifference relation, 'I', or in our case '$\sim_P$', as follows: $x \sim_P y$ iff$_{def}$ neither $x \succ_P y$ nor $y \succ_P x$, it is clear that '$\sim_P$' is an equivalence relation. But this means that strict weak orders cannot be used to derive the relevant indifference relation for vagueness.

Fortunately, there is a well-known ordering that does have the desired properties: what Luce (1956) calls a *semi-order*. Semi-orders were introduced by Luce in economics to account for the intuition that the notion of 'indifference' is not transitive:

> A person may be indifferent between 100 and 101 grains of sugar in his coffee, indifferent between 101 and 102, ..., and indifferent between 4999 and 5000. If indifference were transitive he would be indifferent between 100 and 5000 grains, and this is probably false. (Luce, 1956)

Luce's argument fits well with Fechner's (1860) claim, based on psychophysics experiments, that ability to discriminate between stimuli is generally not transitive. Of course, the problem Luce discusses is just a variant of the Sorites paradox. Luce (1956) introduces semi-orders as an order that gives rise to a non-transitive similarity relation. Following Scott & Suppes' (1958) (equivalent, but still) simpler definition, a structure $\langle D, R\rangle$, with R a binary relation on D, is a semi-order just in case R is irreflexive (IR), satisfies the interval-order (IO) condition, and is semi-transitive (STr).[2]

Definition 2 *A <u>semi-order</u> is a structure* $\langle D, R\rangle$*, with* R *a binary relation on* D *that satisfies the following conditions:*
(IR) $\forall x : \neg R(x, x)$.
(IO) $\forall x, y, v, w : (R(x, y) \wedge R(v, w)) \rightarrow (R(x, w) \vee R(v, y))$.
(STr) $\forall x, y, z, v : (R(x, y) \wedge R(y, z)) \rightarrow (R(x, v) \vee R(v, z))$.

It is important to see that if we interpret the relation '$\succ_P$' as a semi-order, it is irreflexive and transitive, but need not be almost connected. Intuitively, this means that according to this ordering the statement '$x \succ_P y$' means that x is *significantly* or *noticeably* P-er than y. The fact that semi-orders are irreflexive and transitive but not almost connected, is important for us. The reason is that in terms of '$\succ_P$' we can define our desired similarity

[2] Any relation that is irreflexive and satisfies the interval-order condition is called an *interval order*. All interval orders are also transitive, meaning that they are stronger than strict partial orders.

relation '$\sim_P$' as follows: $x \sim_P y$ iff neither $x \succ_P y$ nor $y \succ_P x$. The relation '$\sim_P$' is reflexive and symmetric, but need not be transitive. Thus, '$\sim_P$' does not give rise to an equivalence relation. Intuitively, '$x \sim_P y$' means that there is no significant, or noticiceable, difference between x and y. I believe semi-orders capture most of our intuitions about vagueness.[3] Semi orders can be given a measure-theoretical interpretation in a weak sense. '$x \succ_P y$' is true iff there is a real valued function f_P and some fixed (small) real number ϵ (the *margin of error*) such that $f_P(x) > f_P(y) + \epsilon$ (see Luce, 1956). In the same way '$x \sim_P y$' is true if the difference in P-ness between x and y is less than or equal to ϵ, $|f_P(x) - f_P(y)| \leq \epsilon$. In case $\epsilon = 0$, the semi-order is a weak order.

3 Solving the Sorites by weakening (P)

The standard reaction to the Sorites paradox is to say that the argument is valid, but that the inductive premise (**P**) (or one of its instantiations) is *false*. The question that arises then is why it does at least *seem* to us that the inductive premise is true. It is here that the different proposals differ.

According to supervaluation theory, (**P**) seems true because none of the instantiations of its negation is supertrue. According to proponents of degree theories such as fuzzy logic, the inductive premise, or *principle of tolerance* seems true because it is *almost* true.

Many linguists and philosophers do not like the fuzzy logic approach to vagueness, for one thing because it is not really clear what it means for a sentence to be true to degree $n \in [0, 1]$. For another, the approach seems to over-generate, certainly if one seeks to account for comparative statements in terms of degrees of truth. First, it has been argued that an adjective like 'clever' is multidimensional, and thus that the 'cleverer than'-relation gives rise only to a partial order. But fuzzy logicians have to say it gives rise to a weak, or linear order. Second, if all sentences have a degree of truth, it remains unclear why 'The temperature here is much higher than Paul is tall' is so hard to interpret.[4] The treatment of vagueness and

[3] Kamp (1981), Pinkal (1984), Veltman (1987), van Deemter (1995), and Gaifman (1997) all make implicitly or explicitly use of semi-orders. I argue in van Rooij (2010) that Graff (2000) does the same. Also Williamson's (1994) accessibility relation between worlds, used to represent epistemic indistinguishability, can be defined in terms of a corresponding semi order relation R between these worlds.

[4] Linguists and philosophers have given many other reasons why they don't like a fuzzy logic approach to vagueness. I have to admit that I don't find most of these reasons very

the Sorites paradox in supervaluation theory is not unproblematic either, however. The selling point of supervaluation theory is that it preserves all classical validities. Thus, or so it is claimed, logically speaking there is no difference between classical logic and supervaluation theory. But the non-standard way of accounting for these validities still comes with its *logical prize*. Proponents of supervaluation theory hold that although there is a cutoff-point — i.e. the formula $\exists x, y[Px \wedge x \sim_P y \wedge \neg Py]$ is supertrue —, still, no one of its instantiations itself is supertrue. This is a remarkable logical feature: in classical logic it holds that $A \vee B \models A, B$ (meaning that at least one of A and B must be true in each model that verifies $A \vee B$). In supervaluation theory this doesn't hold anymore; $\exists x Px \not\models_{supv} Px_1, ..., Px_n$. Another problem is of a more conceptual nature. Supervaluation theory makes use of complete refinements, and supervaluation theory assumes that we *can* always make sharp cutoff-points: vagueness exists only because in daily life we are *too lazy* to make them. But this assumption seems to be wrong: vagueness exists, according to Dummett (1975), because we *cannot* make such sharp cutoff-points even if we wanted to.[5]

For a while, the so-called 'contextualist' solution to the Sorites paradox was quite popular (e.g. Kamp, 1981; Pinkal, 1984; Veltman, 1987, Raffman, 1996, van Deemter, 1995, Graf, 2000). Kamp (1981) was the first, and perhaps also the most radical contextualist. He proposed that each instance of the conditional '$(Px \wedge x \sim_P y) \rightarrow Py$' is true, but that one cannot put all these conditionals together into a true universal statement. Most proponents of the contextuallist solution follow Kamp (1981) in trying to preserve (most of) (**P**), and by making use of a mechanism of *context change*.[6] They typically propose to give up some other standard logical assumption. One way of working out the contextual solution assumes that similarity depends on context, and that this context changes in a Sorites sequence. Making the similarity relation context dependent, means to make it a *four*-place relation. One way to do so is to assume that the similarity relation is of the form '$\sim^z_P$', and that $x \sim^z_P y$ is defined to be true iff $x \sim_P z$ and $y \sim_P z$ (and only defined in case either $x \sim_P z$ or $y \sim_P z$). Notice that $x \sim_P y$ iff $x \sim^x_P y$ iff $x \sim^y_P y$, and that the paradox could be derived as usual in case (**P**) would

convincing.

[5]Other problems show up if we want to account for higher-order vagueness in terms of a definiteness operator. See Williamson (1994), and Varzi (2007) for discussion of the problem, and Keefe (2000) and Cobreros (2008) for replies.

[6]For a discussion of this mechanism of context change, see Stanley (2003) and papers that followed that. See also Keefe (2007).

be reformulated as $\forall x, y : (Px \wedge x \sim_P^x y) \rightarrow Py$. Thus, this principle is still considered to be false, though almost all of its instantiations are considered to be true. How, then, is the paradox avoided? Well, observe that $\sim_P^z$ is an equivalence relation, and thus that the relation is transitive after all. Notice that if the contextual tolerance principle ($\mathbf{P}_{c1}$) is formulated in terms of a fixed '$\sim_P^z$' relation,

$$(\mathbf{P}_{c1}) \quad \forall x, y : (Px \wedge x \sim_P^z y) \rightarrow Py,$$

it is unproblematic to take the principle to be valid. As a consequence it has to be assumed, however, that $x \sim_P^z y$ is false for at least one pair $\langle x, y \rangle$ for which $x \sim_P y$ holds: in contrast to '$\sim_P$', '$\sim_P^z$' gives rise to a clear cut-off point. Thus, ($\mathbf{P}_{c1}$) is a weakening of ($\mathbf{P}$). However, the idea of contextualists is that this unnatural fixed cutoff point is avoided, because in the interpretation of a Sorites sequence the relevant individual z that determines the similarity relation changes, and the extension of P with it. Thus, although every context gives rise to a particular cutoff point, context change makes that we are unable to find the cutoff point between P and $\neg P$.

A somewhat more general way to work out the contextualist idea is to assume that similarity is context dependent, because similarity depends on a contextually given *comparison class* c (cf. Veltman, 1987; van Deemter, 1995). Say that $x \sim_P^c y$ iff $\neg \exists z \in c : x \sim_P z \succ_P y$ or $x \succ_P z \sim_y$. Thus, x and y are similar with respect to comparison class c if x and y are not (even) *indirectly* distinguishable w.r.t. elements of c.[7] The inductive premises are reformulated in terms of the new context dependent similarity relation. The idea is that c contains only the individuals mentioned before, and in the sentence itself. Notice that this is fine if one just looks at specific conditionals of the form '$(P(x, c) \wedge x \sim_P^c y) \rightarrow P(y, c)$': c consists just of $\{x, y\}$. A major problem of this approach, however (and shared by the original contextual solutions of Kamp, 1981, and Pinkal, 1984), shows up when we look at the inductive premise as a quantified formula:

$$(\mathbf{P}_{c2}) \quad \forall x, y[(P(x, c) \wedge x \sim_P^c y) \rightarrow P(y, c)].$$

In this case, c must be the set of all individuals, some of which are considered to have property P and some do not. Notice that the relation

[7] This notion was defined previously by Goodman (1951) and Luce (1956).

$\sim^c_P$ is an equivalence relation. But this means that it thus gives rise to a fixed cutoff-point for what counts as P. Notice that ($\mathbf{P}_{c2}$) is again a weakening of ($\mathbf{P}$). Thus, contextualists succeed in making a weakened version of ($\mathbf{P}$) valid, but do so for a surprising reason: ($\mathbf{P}_{c2}$) is valid because for some x and y for which $x \sim_P y$, it holds that the antecedent of ($\mathbf{P}_{c2}$) is false because $x \not\sim^c_P y$.[8]

How good is contextualist solution to the Sorites? As we saw, it comes with two proposals: (i) the inductive premise of the Sorites seems to be valid, because a close variant of it, i.e., ($\mathbf{P}_{c1}$) or ($\mathbf{P}_{c2}$) is valid, and (ii) context change. Both proposals have been criticized. The first because the 'natural' notion of similarity is replaced by an unnatural notion of indirect distinguishability (see, e.g. Williamson, 1994). The contextualist realizes this unnaturalness, and claims that she can avoid the unnatural consequences of making use of this indirect notion by an appeal to context change. But either context change is pushed up until the last pair in a Sorites sequence, and we have a contradiction after all, or it stops at one point, and we still have an unnatural cutoff point (a cutoff-point between x and y, even though $x \sim_P y$).

A more recent contextual solution to the paradox was proposed by Gaifman (1997/2010) (see also Pagin, 2010, and van Rooij, 2010a,b).[9] The idea is that it only makes sense to use a predicate P in a context – i.e. with respect to a comparison class –, if it helps to clearly demarcate the set of individuals that have property P from those that do not. Thus, c can only be an element of the set of *pragmatically appropriate* comparison classes C_A just in case the gap between the last individual(s) that have property P and the first that do(es) not must be between individuals x and y such that x is clearly, or significantly, P-er than y. This is not the case if the graph of the relation '$\sim_P$' is closed in $c \times c$. Indeed, it is exactly in those cases that the Sorites paradox arises. Notice that also Gaifman's solution comes down to weakening inductive hypothesis ($\mathbf{P}$). This time it is by quantifying only over the *appropriate* comparison classes:

[8]Still, Graff (2000) claims that this is the way it should be. According to her, c would (or could) rather say that c just contains those individuals focussed on. Suppose we have the following ordering: $v \sim_P w \sim_P x \sim_P y$ such that $v \succ_P x$ and $w \succ_P y$. Suppose now that $c = \{v, y\}$. In that case, she would claim, it is natural that the cutoff-point between P and $\neg P$ occurs between w and x.

[9]Though in van Rooij (2010a,b) it is claimed that the solution is actually very much in the spirit of the later philosophy of Wittgenstein.

$$(\mathbf{P}_g) \quad \forall x, y \in D, c \in C_A : (P(x,c) \wedge x \sim_P y) \rightarrow P(y,c)$$

Another solution is closely related with recent work of Raffman (2005) and Shapiro (2006).[10] Shapiro states it in terms of three valued logic, and Raffman in terms of pairs of contrary antonymns. The idea is that predicate P and its antonym $\overline{P}$ do not necessarily partition the set of all objects, and there might be elements that neither (clearly) have property P nor property $\overline{P}$, but are somewhere 'in the middle'. Once one makes such a move it is very natural to assume that the inductive principle (**P**) is not valid, but a weakened version of it, ($\mathbf{P_s}$), is. This weakened principle says that if you call one individual tall, and this individual is not visibly/relevantly taller than another individual, you will/should not call the other one short/not tall.

$$(\mathbf{P_s})\ \forall x, y : (Px \wedge x \sim_P y) \rightarrow \neg\overline{P}y.$$

Of course, principle ($\mathbf{P_s}$) can only be different from the original (**P**) if $\neg\overline{P}y$ does not come down to the same as Py. Thus, a gap between the sets of P- and $\overline{P}$-individuals is required. Notice that the Sorites paradox can now be 'solved' in a familiar way: Px_1 and $\overline{P}x_n$ are true, and modus ponens is valid, but the inductive hypothesis, or (all) its instantiations, are not. However, because we adopt ($\mathbf{P_s}$) as a valid principle of language use, we can explain why inductive hypothesis (**P**) *seems* so natural. To illustrate, if $D = \{x, y, z\}$, it might be that $I(P) = \{x\}, I(\overline{P}) = \{z\}$, and $x \sim_P y \sim_P z$. Notice that such a models satisfies ($\mathbf{P_s}$) but not (**P**).

A final proposal I will discuss here was made by Williamson (1994). It is well-known that according to Williamson's epistemic approach, predicates do have a strict cutoff-point, it is just that we don't know it. As in other approaches, also for Williamson it is clear that adopting (**P**) immediately gives rise to paradox. To explain why we are still tempted to accept it, Williamson (1994) offers the following weakening of (**P**) that doesn't give rise to paradox:

$$(\mathbf{P}_\Box) \quad \forall x, y \in I : (\Box Px \wedge x \sim y) \rightarrow Py$$

[10] Shapiro (2006) argues that his solution is closely related to Waisman's notion of 'Open Texture'. For what it is worth, I believe that Waisman's notion is more related to the previously discussed 'solution' of the Sorites.

Thus, if x is *known* to have property P, and x is similar to y, y will actually have property y. Notice that this is also a weakening of (**P**) because the $\Box Px$ entails Px. Williamson's proposal is closely related to the 'three-valued' one discussed above. Suppose we redefine the objects that do not have property $\overline{P}$ as the objects that might have property P (in model M): $M \models \Diamond Px$ iff$_{df}$ $M \not\models \overline{P}x$. In that case, ($\mathbf{P_s}$) comes down to ($\mathbf{P_{s'}}$) : $\forall x, y \in I : (Px \wedge x \sim y) \rightarrow \Diamond Py$. If '$\Diamond$' is the dual of '$\Box$', this, in turn, comes down to ($\mathbf{P}_\Box$). The notion of duality will play an important role in the following section as well.

4 Tolerance and non-transitive entailment

In the previous section we have seen that it is standard to 'tackle' the Sorites paradox by weakening the inductive premise (**P**) in some way or other. But there exists an interesting other view, which perhaps goes back to Kamp (1981), and has recently been defended by Zardini (2008). According to it, the tolerance principle is true, but the Sorites-reasoning is invalid because the *inference relation* itself is not transitive. Zardini's way of working out this suggestion into a concrete proposal is rather involved. In this section I will work out the same suggestion in an (I think) much simpler and more straightforward way. My aim in this section is to make this rather non-standard approach more plausible on independent grounds. In the next section I will try to seduce proponents of other views, by showing that this solution is actually closely related to (some of) the approaches discussed in the previous section.

Let us start with a (set of) semi-order(s) $\langle X, \succ_P \rangle$, one for each vague predicate P (that holds in all models), that thus gives rise to a similarity relation '$\sim_P$' that is reflexive, symmetric, but not transitive. I would like to propose now that given this similarity relation, we can interpret sentences in at least two different ways: in terms of '$\models$' as we normally do, but also in a *tolerant* way in terms of '$\models^{t'}$'. Take a standard first order model $M = \langle X, I \rangle$ and a fixed similarity relation $\sim_P$ (for each P), and define '$\models^{t'}$' (and '$\models$') recursively as follows (for simplicity I use the substitutional analysis of quantification and assume that every individual has a unique name):

$M \models \phi$ defined in the usual way.
$M \models^{t'} P(\underline{a})$ iff $\exists d \sim_P a : M \models P(\underline{d})$, with $\underline{d}$ as name for d
$M \models^{t'} \neg\phi$ iff $M \not\models^{t'} \phi$

$M \models^{t'} \phi \wedge \psi$ iff $M \models^{t'} \phi$ and $M \models^{t'} \psi$
$M \models^{t'} \forall x \phi$ iff for all $d \in I_M : M \models^{t'} \phi[^x/_{\underline{d}}]$.

Now we can define two *tolerant* entailment relations, '$\models^{tt}$' and $\models^{ct}$, as follows: $\phi \models^{tt} \psi$ iff $[\![\phi]\!]^{t'} \subseteq [\![\psi]\!]^{t'}$, and $\phi \models^{ct} \psi$ iff $[\![\phi]\!] \subseteq [\![\psi]\!]^{t'}$, where $[\![\phi]\!]^{(t')} = \{M : M \models^{(t')} \phi\}$. We will say that ϕ is tolerance-valid iff $M \models^t \phi$ in all models M with an indistinguishability relation. Although the first tolerant entailment relation is defined rather classical, I will be mostly interested in the second entailment relation. This second entailment relation is *not transitive*: from $\phi \models^{ct} \psi$ and $\psi \models^{ct} \chi$ it doesn't follow that $\phi \models^{ct} \chi$. Assume, for instance, that for all models $a \sim_P b \sim_P c$, but that $a \succ_P c$. Now $P(\underline{a}) \models^{ct} P(\underline{b})$ and $P(\underline{b}) \models^{ct} P(\underline{c})$, but not $P(\underline{a}) \models^{ct} P(\underline{c})$: there might be a model M such that $I_M(P) = \{a\}$.

Material implication doesn't mirror '$\models^{ct}$', but we can define a new conditional connective, i.e. '$\rightarrow^{ct}$', that does. Say that $M \models \phi \rightarrow^{ct} \psi$ iff$_{def}$ if $M \models \phi$, then $M \models^t \psi$. Notice that $(\mathbf{P_t})$ $\forall x, y : (P(x) \wedge x \sim_P y) \rightarrow^{ct} P(y)$ is classically valid. This is not problematic to account for the Sorites, because the hypothetical syllogism is not valid when formulated in terms of '$\rightarrow^{ct}$.[11] Nice as it may seem, this is just a roundabout way of doing things. Instead of reinterpreting implication, it is much more natural to interpret negation differently. I will show that with negation defined in this way, $(\mathbf{P})$ itself is tolerance-valid.

In the following I will define the meaning of negation used to interpret tolerance-truth in terms of a new notion of *strict* truth: '$\models^s$'. In fact, we have to define '$\models^t$' and '$\models^s$' simultaneously.

$M \models^t P(\underline{a})$ iff $\exists d \sim_P a : M \models P(\underline{d})$
$M \models^t \neg\phi$ iff $M \not\models^s \phi$
$M \models^t \phi \wedge \psi$ iff $M \models^t \phi$ and $M \models^t \psi$
$M \models^t \forall x \phi$ iff $\forall d \in I_M, M \models^t \phi[^x/_{\underline{d}}]$

$M \models^s P(\underline{a})$ iff $\forall d \sim_P a, M \models P(\underline{d})$
$M \models^s \neg\phi$ iff $M \not\models^t \phi$
$M \models^s \phi \wedge \psi$ iff $M \models^s \phi$ and $M \models^s \psi$
$M \models^s \forall x \phi$ iff $\forall d \in I_M, M \models^s \phi[^x/_{\underline{d}}]$

The connectives '$\vee$' and '$\rightarrow$' are defined in terms of '$\neg$' and '$\wedge$' as usual.

[11] A similar story would hold for conditionals like a$\rightarrow^{sc}$' and '$\rightarrow^{st}$'.

Notice that $P(\underline{a}) \vee \neg P(\underline{a})$ is tolerant-valid, can be strictly-true, but is not a strict-tautology. $P(\underline{a}) \wedge \neg P(\underline{a})$, on the other hand, cannot be strictly true, but can be tolerantly true.[12] We will assume that the similarity relation should be interpreted in a fixed way, and cannot have a separate strict or tolerant reading: $M \models a \sim_P b$ iff $M \models^s a \sim_P b$ iff $M \models^t a \sim_P b$, and $M \models a \not\sim_P b$ iff $M \models^s a \not\sim_P b$ iff $M \models^t a \not\sim_P b$. The most appealing fact about this system is that the original tolerance-principle, (**P**) $\forall x, y[(Px \wedge x \sim_P y) \rightarrow Py]$ is tolerance-valid! This is easy to see because for this sentence to be tolerance-true in M it has to be the case for any a and b such that $a \sim_P b$ that either $\exists d \sim_P a : M \not\models P(\underline{d})$ or $\exists d' \sim_P b : M \models P(\underline{d'})$. But this is always the case. Thus, on the present analysis we can say that the original (**P**) is, though not classically valid, still tolerantly valid. Notice that (**P**) is neither classically, nor strictly valid.[13]

Thus, what we have now, finally, is a notion of validity according to which the original (**P**) is valid. This does not give rise to the prediction that all objects have property P in case the entailment relation is $\models^{tt}$. For that relation, modus ponens is not valid. However, we have opted for entailment relation $\models^{ct}$ (with t' replaced by t everywhere) according to which modus ponens *is* valid. Still, no reason to worry, because this relation is non-transitive. We can conclude that it does not follow that all objects have property P.[14]

.

5 Comparison with other approaches

Although our approach seems rather non-standard, it is closely related with other approaches. Consider, for instance super- and subvaluationalism.[15]

[12]Dave Ripley (p.c.) pointed out that my notions of tolerant and strict truth in fact correspond with the notions of truth in Priest's (1979) logic of paradox (LP) and Kleene's system K3, respectively. These connections will be proved and worked out further in Cobreros et al (manuscript).

[13]It was also not valid in our earlier fomulation of tolerant truth.

[14]Although it is widely acknowledged that one can 'solve' the Sorites by assuming that the entailment relation is non-transitive, it is hardly ever seriously defended (if at all). The reason for this, it seems, is Dummett's (1975) claim that one cannot seriously deny the ability to chain inferences, because this principle is taken to be essential to the very enterprise of proof. To counter this objection, in .... et al. (manuscript) we provide a proof theory that corresponds to $\models^{ct}$ (and, in fact, many other non-classical inference relations).

[15]This connection was pointed out by Paul Egre in his comment on my Pamplona paper.

There exists a close relation between our notions of strict and tolerant truth with the notions of truth in supervaluationalism, and subvaluationalism (Hyde, 1997), respectively. Notice in particular that subvaluationalism is paraconsistent, just like our notion of tolerant truth: $P\underline{a}$ can be both true *and* false, without giving rise to catastrophic consequences. Indeed, these theories, just like supervaluationalism and our notion of strict truth, are very similar when we only consider atomic statements: in both cases we define truth in terms of existential and universal quantification, respectively. Moreover, in both cases the two notions are each others duals. But the analogy disappears when we consider more complex statements. The reason is that we make use of this quantificational interpretation at the *local level*, while they only do so only at the *global* level. Although looking at the global level means to give up on the idea that interpretation goes compositional, interpreting globally instead of locally still seems to be advantageous. This is so, because as a result, both $P(\underline{a}) \vee \neg P(\underline{a})$ and $P(\underline{a}) \wedge \neg P(\underline{a})$ are validities, while for us the former can be strictly false, and the latter can be tolerantly true. Moreover, the idea of interpreting globally is crucial for Fine's (1975) analysis of penumbral connections. We have seen already that supervaluationalism is not so classical after all, once one does not limit oneself to single-conclusion arguments. Something similar holds for subvaluationism: as already observed in the original article, and stressed by Keefe (2000), Hyde (1997) makes non-classical predictions once one does not limit oneself to single-premise arguments: $\phi, \psi \not\models_{subv} \phi \wedge \psi$. In .... et al (manuscript) we will show that there is much to say in favor of our notions of truth and entailment. In particular, $\phi, \psi \models^{ct} \phi \wedge \psi$ and $\phi \vee \psi \models^{ct} \phi, \psi$. As for penumbral connections, we admit that $\neg P(\underline{a}) \wedge P(\underline{b})$ can be tolerantly true even if $a \succeq_P b$. In .... et al. (manuscript) we will argue that as far as semantics is concerned, this is, in fact, not a problem. What has to be explained, though, is why it is pragmatically *inappropriate* to utter a statement saying '$\neg P(\underline{a}) \wedge P(\underline{b})$'. The explanation will be that without any further information, a hearer of this utterance will conclude from this that $b \succ_P a$, because this is the only way in which the statement can be true if the statement is interpreted in the strongest possible way.[16] If the speaker knows that $a \succeq_P b$ it is thus inappropriate to make such a statement. See Alxatib & Pelletier (manuscript) for a very similar move to solve the very similar problem of

[16] Interpreting sentences that semantically allow for different interpretations in the strongest possible way is quite standard in pragmatics. Of course, this kind of pragmatic interpretation can be overruled (i.e., behaves non-monotonic) by further information, in our case that $a \succeq b$.

why contradictory attributed can sometimes truly be attributed to the same borderline object.

For another comparison, consider Williamson's approach. Recall that he wanted to 'save' the intuition of tolerance by turning ($\mathbf{P}_\Box$) ($\forall x, y$: $(\Box Px \wedge x \sim_P y) \rightarrow Py$) into a validity. Similarly for our reformulation of the tolerance principle of Shapiro: ($\mathbf{P_s}$) ($\forall x, y : (Px \wedge x \sim_P y) \rightarrow \Diamond Py$). I will show now that by re-interpreting '$\Box$' and '$\Diamond$' in terms of our similarity relation, there is an obvious relation these approaches and mine. The redefinition goes as follows:

$$M \models \Box\phi \text{ iff } M \models^s \phi \text{ and } M \models \Diamond\phi \text{ iff } M \models^t \phi$$

Notice that $\Diamond P(\underline{a}) \wedge \Diamond\neg P(\underline{a})$ is possible, but $\Diamond P(\underline{a}) \wedge \neg\Diamond P(\underline{a})$ is impossible; $\Diamond P(\underline{a}) \vee \neg\Diamond P(\underline{a})$ and $\Box P(\underline{a}) \vee \neg\Box P(\underline{a})$ are tautologies; $\Box P(\underline{a}) \wedge \neg\Box P(\underline{a})$ is impossible, just as $\Diamond P(\underline{a}) \wedge \Box\neg P(\underline{a})$. Observe also that it now immediately follows that $\neg\Box\neg\phi \equiv \Diamond\phi$ and $\neg\Diamond\neg\phi \equiv \Box\phi$: '$\Box$' and '$\Diamond$' are duals of each other. Notice that both $\forall x, y : \Box Px \wedge x \sim_P y \rightarrow Py$ and $\forall x, y : Px \wedge x \sim_P y \rightarrow \Diamond Py$ are valid, and are equivalent to each other. The fact that ($\mathbf{P}$) is tolerantly valid is actually *weaker* than either of them: the reformulation of ($\mathbf{P}$) would be $\forall x, y : \Box Px \wedge x \sim_P y \rightarrow \Diamond Py$. For Williamson (1994) it is only natural to assume that if ($\mathbf{P}_w$) holds, agents *know* that it holds. The corresponding strengthening of ($\mathbf{P}$) in our case, however, doesn't seem natural. Indeed, it certainly is not the case that $\Box\forall x, y[(Px \wedge x \sim_P y) \rightarrow \Diamond Py]$ is valid. Before I suggested to account for a notion of vague inference as follows: $\phi \models^{ct} \psi$ iff $[\![\phi]\!]^c \subseteq [\![\psi]\!]^t$. Alternatively, we could do something else, which sounds equally natural: $\phi \models^{sc} \psi$ iff $[\![\psi]\!]^s \subseteq [\![\phi]\!]^c$. In terms of our 'modal' system, these inference relations can be incorporated into the object-language as follows: $\phi \models^{ct} \psi$ iff for all M, $M \models \phi \rightarrow \Diamond\psi$, and $\phi \models^{sc} \psi$ iff for all M, $M \models \Box\phi \rightarrow \psi$. These notions do not exactly coincide.

Consider, finally, the contextualist solution. Recall that according to Kamp's (1981) solution, each instance of the conditional $(Px \wedge x \sim y) \rightarrow Py$ is true, it is just that we cannot put all these conditionals together to turn them into a true universal statement. Our solution is similar, though we don't talk about truth of conditional statements, but of valid inferences: each inference step is (ct) valid, but we cannot chain them together to a (ct) valid inference. As a second connection, observe that our introduced conditional '$\rightarrow^{ct}$' is very similar to the conditional introduced by Kamp (1981). As a last point of contact, consider the notion of meaning change proposed in contextualist' solutions. Contextualist typically say that the meaning of predicate P changes during the interpretation of the Sorites sequence. It is almost immediately obvious in terms of our framework how this meaning change takes place: First, it has to be the case that $M \models P\underline{a}$. At the second step, the meaning of P changes, and we end up with a new model M' such that $M' \models P\underline{b}$ iff $M \models \Diamond P\underline{b}$ (or $M \models^t P\underline{b}$). At the third step, the meaning of P changes again, and we end up with a new model M'' such that $M'' \models P\underline{c}$ iff $M' \models \Diamond P\underline{c}$ (or $M' \models^t P\underline{c}$). And so on, indefinitely. But do we really need to go to new models every time? We need not, if we can iterate modalities.

6 Similarity and borderlines

Traditional approaches of vagueness start with borderlines. To account for higher-order vagueness, one then needs a whole sequence of higher-order borderlines. In this section I suggest two ways to represent higher order borderlines: one in terms of iteration of 'modalities'; another in terms of fine-grainedness.

6.1 Iteration, and higher order vagueness

Let $\mathbf{B}\phi$ be an abbreviation of $\neg\Box\phi \wedge \neg\Box\neg\phi$. Thus, $\mathbf{B}P(\underline{a})$ means that a is a borderline case of P. Our system allows for first-order borderline cases, but it makes it impossible to account for higher-order borderlines, and thus cannot account for higher-order vagueness. But why don't we just say that a is a second-order borderline case of P if $\neg\Box\Box P(\underline{a}) \wedge \neg\Box\Box\neg P(\underline{a})$. This sounds ok, but the problem is that we cannot yet interpret these types of formulas, because we haven't specified yet how to make sense of '$M \models^s \Box\phi$' or '$M \models^t \Box\phi$'. So let us try to do just that. What we need to do is to interpret formulas with respect to a (perhaps empty) *sequence* of s's and t's, like $\langle s, s, t\rangle$ or $\langle t, t\rangle$. We will abbreviate a sequence by 'σ', and if $\sigma = \langle x_1, ..., x_n\rangle$, then '$\sigma t$' will be $\langle x_1, ..., x_n, t\rangle$ and 'σs' will be $\langle x_1, ..., x_n, s\rangle$. σ^* will just be the same as σ except that all t's and s's are substituted for each other. Thus, if $\sigma = \langle s, s, t\rangle$, for instance, then $\sigma^* = \langle t, t, s\rangle$. Furthermore, we are going to say that if σ is the empty sequence, '$\langle\rangle$', $M \models^\sigma \phi$ iff $M \models \phi$.

$M \models^\sigma \Box\phi$ iff $M \models^{\sigma s} \phi$ and $M \models^\sigma \Diamond\phi$ iff $M \models^{\sigma t} \phi$

$$
\begin{array}{lll}
M \models^{\sigma t} P(\underline{a}) & \text{iff} & \exists d \sim_P a : M \models^{\sigma} P(\underline{d}) \\
M \models^{\sigma t} \neg\phi & \text{iff} & M \not\models^{\sigma^* s} \phi \\
M \models^{\sigma t} \phi \wedge \psi & \text{iff} & M \models^{\sigma t} \phi \text{ and } M \models^{\sigma t} \psi \\
M \models^{\sigma t} \forall x \phi & \text{iff} & \forall d \in I_M, M \models^{\sigma t} \phi[^x/_{\underline{d}}]
\end{array}
$$

$$
\begin{array}{lll}
M \models^{\sigma s} P(\underline{a}) & \text{iff} & \forall d \sim_P a : M \models^{\sigma} P(\underline{d}) \\
M \models^{\sigma s} \neg\phi & \text{iff} & M \not\models^{\sigma^* t} \phi \\
M \models^{\sigma s} \phi \wedge \psi & \text{iff} & M \models^{\sigma s} \phi \text{ and } M \models^{\sigma s} \psi \\
M \models^{\sigma s} \forall x \phi & \text{iff} & \forall d \in I_M, M \models^{\sigma s} \phi[^x/_{\underline{d}}]
\end{array}
$$

To see what is going on, let us assume a domain $\{u, v, w, x, y, z\}$ such that $u \sim_P v \sim_P w \sim_P x \sim_P y \sim_P z$ and $u \succ_P w$, $v \succ_P x$, $w \succ_P y$, and $x \succ_P z$ together with the assumption that '$\succ_P$' is a semi-order. Let us now assume that $I_M(P) = \{u, v, w\}$. If we build the complex predicate '$\Box P$' and say that this holds of a in M iff $M \models \Box P(\underline{a})$, it follows that $I_M(\Box P) = \{u, v\}$, and $I_M(\Box\Box P) = \{u\}$. Similarly, it follows that $I_M(\Box\neg P) = \{y, z\}$, and $I_M(\Box\Box\neg P) = \{z\}$. The first-order borderline cases of P, $\mathbf{B}^1 P$, are those d for which it holds that $\neg\Box^1 P(\underline{d}) \wedge \neg\Box^1\neg P(\underline{d})$. Thus, $I_M(\mathbf{B}^1 P) = \{w, x\}$. Similarly, $I_M(\mathbf{B}^2 P) = \{d \in X : M \models \neg\Box^2 P(\underline{d}) \wedge \neg\Box^2\neg P(\underline{d})\} = \{v, w, x, y\}$ and $I_M(\mathbf{B}^3 P) = \{u, v, w, x, y, z\}$.[17]

Our analysis of higher-order vagueness is similar to Gaifman's (1997/2010) treatment. Both start with a standard two-valued logic and build higher-order vagueness in terms of it. What would happen if our basic logic was not two-valued, but three-valued instead? Very little, except that n-order borderlines are now defined 'one step behind'. Suppose we take the same domain as above, giving rise to the same semi-order, but assume that $I_M(P) = \{u, v\}$ and $I_M(\overline{P}) = \{y, z\}$. One proposal would be to say that $I_M(\mathbf{B}^n P) = \{d \in X : M \models \neg\Box^{n-1} P(\underline{d}) \wedge \neg\Box^{n-1}\overline{P}(\underline{d})\}$. Thus $I_M(\mathbf{B}^1 P) = X - (I_M(P) \cup I_M(\overline{P})) = \{w, x\}$, while $I_M(\mathbf{B}^2 P) = \{v, w, x, y\}$ and $I_M(\mathbf{B}^3 P) = \{u, v, w, x, y, z\}$. Perhaps more in accordance with tradition would be to define $I_M(\mathbf{B}^n P)$ as follows: $I_M(\mathbf{B}^n P) = \{d \in X : M \models \neg\Box^{n-1} P(\underline{d}) \wedge \neg\Box^{n-1}\overline{P}(\underline{d}) \wedge \neg\mathbf{B}^{n-1} P(\underline{d})\}$. But to make sense of this, we have to know what things like $M \models^t \mathbf{B}P(\underline{d})$ mean. A natural definition goes as follows:

[17] Alternatively, we might define the nth order borderline cases of P as those d for which it holds that $\neg\Box^n P(\underline{d}) \wedge P(\underline{d}) \wedge \Diamond^n P(\underline{d})$. In that case, $I_M(\mathbf{B}^1 P) = \{w\}$, $I_M(\mathbf{B}^2 P) = \{v, w\}$ and $I_M(\mathbf{B}^3 P) = \{u, v, w\}$.

$$M \models \mathbf{B}P(\underline{d}) \quad \text{iff} \quad d \notin I_M(P) \cup I_M(\overline{P})$$
$$M \models^{\sigma s} \mathbf{B}P(\underline{d}) \quad \text{iff} \quad \forall d' \sim_P d : M \models^{\sigma} \mathbf{B}P(\underline{d'})$$
$$M \models^{\sigma t} \mathbf{B}P(\underline{d}) \quad \text{iff} \quad \exists d' \sim_P d : M \models^{\sigma} \mathbf{B}P(\underline{d'})$$

Shapiro's (2006) weakened version of (**P**), i.e. ($\mathbf{P}_s$), could now perhaps best be stated as follows: $\forall x, y : (\Box^n Px \wedge x \sim_P y) \rightarrow (\Box^n P(y) \vee \mathbf{B}^{n+1}P(y))$.

6.2 Borderlines and fine-grainedness

In natural language we conceptualize and describe the world at different levels of granularity. A road, for instance, can be viewed as a line, a surface, or a volume. The level of granularity that we make use of depends on what is relevant (cf. Hobbs, 1985). When we are planning a trip, we view the road as a line. When we are driving on it, we view it as a surface, and when we hit a pothole, it becomes a volume to us. In our use of natural language we even employ this fact by being able to describe the same phenomenon at different levels of granularity within the same discourse. Thus, we sometimes explicitly shift perspective, i.e., shift the level of granularity to describe the same situation. This is perhaps most obviously the case when we talk about time and space: "It is two o'clock. In fact, it is two minutes after two." In this sentence we shift to describing a time-point in a more specific way. Suppose that we consider two models, M and M' that are exactly alike, accept that they differ on the interpretation of a specific ordering relation, such as 'earlier than', or 'taller than'. When can we think of the one model as being finer-gained than the other? The only reasonable proposal seems to be to say that M' is a refinement of M with respect to some ordering $\geq$, $M \sqsubseteq M'$, only if $\forall x, y, z \in I$: if $M' \models x \geq y \wedge y \geq z$ and $M \models x \sim z$ (with $x \sim y$ iff $x \not> y$ and $y \not> x$), then $M \models x \sim y \wedge y \sim z$. This follows if we define refinements as follows: M' is a refinement of M with respect to $\geq$ iff $V_M(>) \subseteq V_{M'}(>)$.

In the special case that the ordering relation is a weak order, this way to relate different models in terms of a coarsening relation made use of a standard technique. Recall, first, that the relation $\sim$ is in that case an equivalence relation. In a coarser-grained model M we associate each equivalence class in the finer-grained model M' via an homomorphic function f with an equivalence class of the coarse grained model M, and say that $M \models x > y$ iff $\forall x' \in f^{-1}(x), y \in f^{-1}(y) : M' \models x' > y'$. But observe that only a slight extension of the method can be used for other orders as well, in particular for semi-orders (recall that a weak order is a special kind of semi-order). Thus we say that M' is a refinement of M with respect to $\succ$ iff $V_M(\succ) \subseteq V_{M'}(\succ)$. Notice that if $V_M(\succ) \subset V_{M'}(\succ)$, it means that in M more individuals are $\sim$-related than in M'. In measure theoretic terms, it means that the margin of error ϵ, is larger in M than it is in M', which is typically the case if in M' more is at stake.[18] Similarly, we say that $M \models x \succ y$ iff $\forall x' \in R(x), \forall y' \in R(y) : M' \models x' \succ y'$, where R is a relation between elements of M and M' that preserves $\succ$.[19] Suppose that the ordering is the ordering '(observably) P-er than'. Notice that at M it only makes sense to say that $Px \wedge \neg Py$ in case $M \models x \succ_P y$. Suppose that in M the last individual in the extension of P is x, while y is the first individual in its anti-extension. Does that mean that we have a clear cutoff-point for the extension of P? It does not, if we are allowed to look at finer-grained models, where the domain of such a finer-grained model might be bigger than the domain of M.

One can image a whole sequence of refinements of a model M_0: $M_0 \sqsubset M_1 \sqsubset ... \sqsubset M_n$....[20] It terms of it, we might define a *definiteness*-operator to account for higher-order vagueness.[21] Say that $M_i \models \mathbf{D}Px$ iff $\forall x' \in \{y \in D_{M_j} : xR_{ij}y\} : M_j \models Px'$ (where M_j is the immediate refinement of M_i, and R_{ij} is a relation with domain M_i and range M_j respecting the ordering relations $\succ$ in their respected models). Similarly, we might define a to be a borderline-case of P in M_i, $M_i \models \mathbf{B}P\underline{a}$, if it holds that $M_i \models \neg\mathbf{D}P\underline{a} \wedge \neg\mathbf{D}\neg P\underline{a}$. Similarly for higher-order borderline cases.

Recall that $M_i \models^s P\underline{a}$ iff $\forall d \sim_P a : M_i \models P\underline{d}$. Observe that there exists a relation between $P\underline{a}$ being *stricty* true in M_i, and $P\underline{a}$ being *definitely* true in in M_i: $M_i \models^s P\underline{a}$ iff $M_i \models \mathbf{D}P\underline{a}$ iff $\forall d \in \{x \in D_{M_i} : aR_{ij}x\} : M_j \models P\underline{d}$. Similarly, $M_i \models^t P\underline{a}$ iff $M_i \models \mathbf{D}P\underline{a} \vee \mathbf{B}P\underline{a}$, i.e., if $\exists d \in \{x' \in D_{M_j} : aR_{ij}x'\} : M_j \models P\underline{d}$. Notice that $M_i \not\models^s P\underline{a}$ does not correspond with $M_i \models \neg\mathbf{D}P\underline{a}$, but rather with $M_i \models \mathbf{D}\neg P\underline{a}$.

[18] I believe that much of what Graff (2000) discusses as 'interest relative' can be captured in this way.

[19] Meaning that if $M \models x \succ y$, then $\forall x' \in \{z \in D_{M'} : xRz\}, \forall y' \in \{z \in D_{M'} : yRz\} : M' \models x' \succ y'$.

[20] Perhaps there is no most fine-grained model.

[21] For what it is worth, I feel that this is at least in the spirit of what is proposed by Fine (1975) and Keefe (2000).

7 Clusters, prototypes, and defining similarity

In section 2 we started with an ordering relation and defined a similarity relation in terms of it. But this is obviously not crucial for thinking about similarity, or resemblance. Suppose we start out with a primitive similarity relation, $\sim$, that is reflexive and symmetric, but not necessarily transitive. We can now think of a *similarity class* as a class of objects S such that $\forall x, y \in S : x \sim y$. A maximal such similarity class might be called a *cluster*. Clusters hardly play a role in categorization when starting out with one-dimensional ordering relations like 'taller than', or 'earlier than'. But they play a crucial role in categorization when more dimensions are at stake. Clusters can be tolerant, or strict. Let us say that a cluster is tolerant in case $C^t =_{df} \{x \in D | \exists y \in C : x \sim y\} \neq C$, and strict otherwise. If predicate P is interpreted by cluster C_P, it holds that $M \models^t P\underline{a}$ iff $a \in C^t_P$. We can also define the strict version of a cluster, $C^s_P =_{df} \{x \in C_P | \forall y \in D : y \sim x \rightarrow y \in C_P\}$. It follows immediately that $M \models^s P\underline{a}$ iff $a \in C^s_P$. Notice that if C_P is strict, $C_P = C^s_P$. In general, the classical interpretation of P, $[\![P]\!] = C_P$, should be such that $C^s_P \subseteq [\![P]\!] \subseteq C^t_P$. Notice that by definition it holds that $C^t_P \subseteq C_Q$ iff $C_P \subseteq C^s_Q$, which is again an interesting relation between our dual concepts.

In terms of a similarity relation and a cluster, we can define a notion of a prototype. First, define '$\preceq$' as follows: $x \preceq y$ iff$_{df}$ $\{z \in D : z \sim x\} \subseteq \{z \in D : z \sim y\}$.[22] Now suppose that for a cluster C, there exists an element $x \in C$ such that $\forall y \in C : x \preceq y$. In such a case it makes sense to call this element a *prototype* of C. Notice that it is well possible that a cluster C has more than one prototype. It is useful to have such prototypes, because it is taken to be much more expensive to represent meanings in terms of their extensions than in terms of their prototypes. There need not be any loss involved: in Gärdenfors' (2000) geometrical approach to meaning, for instance, the extension of a (set of) term(s) can be derived from the prototype(s). But if a cluster has a prototype, something similar (actually, something stronger) can be done here as well: the cluster C associated with prototype x_C is determined as *the* unique cluster such that x is an element of it. Notice that being a prototype is something special, because there might well be two clusters, C_1 and C_2, such that $\exists x \in C_1, \exists y \in C_2 : x \sim y$, i.e. $C^t_1 \cap C_2 \neq \emptyset$ (or equivalently, $C_1 \cap C^t_2 \neq \emptyset$).

[22] Notice that if we define $x \approx y$ as true iff $x \preceq y$ and $y \preceq x$, it immediately follows that '$\approx$' is an equivalence relation, and, in fact, the *indirect* indistinguishability relation as defined by Goodman (1951) and Luce (1956).

'Similarity' is not an absolute notion: one pair of objects can be more similar to each other than another pair. In geometrical models of meaning, similarity is measured by the inverse of a distance measure d between two objects. In Tversky's (1977) contrast model, the similarity of two objects is determined by the primitive features they share, and the features they differ on. Say that object x and y come with sets of primitive features X and Y. If we only consider the features they share, the similarity of x and y can be measured in terms of $X \cap Y$: x is more similar to y than v is to w iff $f(X \cap Y) > f(V \cap W)$, with f some real valued function monotone on '$\supseteq$'.[23] Clusters as determined above now depend on when we take two objects similar enough to be called 'similar'. If we fix this, we can determine what a cluster is, and what a tolerant cluster is. If C is a cluster, there still might be some elements in C that are more similar to all other elements of C than just 'similar'. Following Tversky, we can measure the prototypicality of each $x \in C$ as follows: $p(x, C) = \sum_{y \in C} f(X \cap Y)$. A prototype of C is then simply an element of C with the highest p-value.

Until now I started with a specific notion of similarity, perhaps explained in terms of measurement errors, or a primitive idea of what counts as a relevant difference. But Tversky's model suggests that we can explain our similarity relation in terms of shared features. Take any arbitrary n-ary partition Q of the set of all individuals. Which of those partitions naturally classifies those individuals? Take any element q of Q, and determine its *family resemblance* as follows: $FR(q) = \sum_{x,y \in q} f(X \cap Y)$. Categorization Q can now be called 'at least as good' as categorization Q' (another partition of D) just in case $\sum_{q \in Q} FR(q) \geq \sum_{q' \in Q'} FR(q')$. With Rosch (1973) we might now call X a 'basic category' just in case X is an element of the best categorization of D. What is interesting for us is that a best categorization Q can determine a level of similarity to be the 'basic' one, i.e., to be '$\sim$'. But first let us assume that a basic categorization is 'nice' in case $\forall q, q' \in Q : min\{f(X \cap Y) : x, y \in q\} \approx min\{f(X \cap Y) : x, y \in q'\}$.[24] With respect to such a 'nice' categorization Q, we can define the similarity relation as

[23]Tversky's model is much more flexible than this. He even allows for x to be more similar to y than y is to x.

[24]If we assume that $\forall q, q' \in Q : min\{f(X \cap Y) : x, y \in q\} > max\{f(X \cap Z) : x \in q, z \in q'\}$, categorization is clearly analogous to Gaifman's treatment to avoid the Sorites paradox. In fact, this principle is behind most of the hierarchical structuring models: If one starts with a difference measure, one can show that if for all $x, y, z : d(x, y) \leq d(x, z) = d(y, z)$, then the set of objects give rise to a hierarchically ordered tree (cf. .Johnson, 1967).

follows: $x \sim y$ iff$_{df}$ $f(X \cap Y) \geq min\{f(V \cap W)|\ v, w \in q\}$, for any $q \in Q$.

8 Conclusion

In this paper I argued that vagueness is crucially related with tolerant interpretation, and that the latter is only natural for observational predicates. Still, most approaches dealing with the Sorites in the end give up the principle of tolerance. I argued, instead, that once tolerance plays a role, the entailment relation need not be transitive anymore. It was shown how to make sense of this proposal, and how it relates to some of the standard analyses. Finally, I related our analysis to some analyses of concepts in cognitive science.

References

[1] Alxatib, S. and J. Pelletier (2009), 'The psychology of vagueness: Borderline cases and contradictions', manuscript.

[2] Cobreros, P. (2008), 'Supervaluationism and logical consequence: a third way', *Studia Logica*, **90**: 291-312.

[3] Cobreros, P., P. Egre, D. Ripley, and R. van Rooij (manuscript), *Tolerent, classical, strict.*

[4] ... (manuscript), *Tolerent, classical, strict.*

[5] Deemter, K. van (1995) 'The Sorites fallacy and the context-dependence of vague predicates', in M. Kanazawa, et al., *Quantifiers, Deduction, and Context*, CSLI Publications, Stanford, Ca., pp. 59-86.

[6] Dummett, M. (1975), 'Wang's paradox', *Synthese*, **30**: 301-324.

[7] Fechner, G.T. (1860), *Elemente der Psychophysics*, Breitkopf und Härtel, Leipzig.

[8] Fine, K. (1975), 'Vagueness, truth and logic', *Synthese*, **30**: 265-300.

[9] Gaifman, H. (2010), 'Vagueness, tolerance and contextual logic', *Synthese*. (the paper was originally written in 1997, and available ever since)

[10] Gärdenfors, P. (2000), *Conceptual Spaces. The Geometry of Thought*, MIT Press.

[11] Goodman, (1951), *The structure of appearance*, Bobbs-Merill Publ.

[12] Graff, D. (2000), 'Shifting sands: An interest relative theory of vagueness', *Philosophical Topics*, **28**: 45-81.

[13] Hobbs, J. (1985), 'Granularity', *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-85).*

[14] Hyde, D. (1997), 'From heaps and gaps to heaps and gluts', *Mind*, **106**: 641-660.

[15] Johnson, S.C. (1967), 'Hierarchical clustering schemes', *Psychometrika*, **32**, 241-254.

[16] Kamp, H. (1975), 'Two theories of adjectives', in E. Keenan (ed.), *Formal Semantics of Natural Language*, Cambridge University Press, Cambridge, pp. 123-155.

[17] Kamp, H. (1981), 'The paradox of the heap', in U. Mönnich (ed.), *Aspects of Philosophical Logic*, D. Reidel, Dordrecht, pp. 225-277.

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

[19] Keefe, R. (2007), 'Vagueness without context change', *Mind*, **116**: 462.

[20] Klein, E. (1980), 'The semantics of positive and comparative adjectives', *Linguistics and Philosophy*, **4**: 1-45.

[21] Luce, R. D. (1956), 'Semiorders and a theory of utility discrimination', *Econometrica*, **24**: 178-191.

[22] Pagin, P. (2010), 'Vagueness and domain restriction', in N. Clinedinst and P. Egre (eds.), *Vagueness and Language Use*, Palgrave MacMillan.

[23] Priest, G. (1979), 'Logic of paradox', *Journal of Philosophical Logic*, **8**: 219-241.

[24] Raffman, D. (1996), 'Vagueness and Context-Relativity', *Philosophical Studies*, **81**: 175-192.

[25] Raffman, D. (2005), 'Borderline cases and bivalence', *Philosophical Review*, **114**.

[26] Ripley, D. (ms), 'Sorting out the Sorites', Institut Jean-Nicod, ENS, Paris.

[27] Rooij, R. van (2010a), 'Vagueness and linguistics', in Ronzitti (ed.), *Vagueness: A Guide*, Springer.

[28] Rooij, R. van (2010b), 'Implicit versus explicit comparatives', in N. Clinedinst and P. Egre (eds.), *Vagueness and Language Use*, Palgrave MacMillan.

[29] Rosch, E. (1973) 'On the internal structure of perceptual and semantic categories', in T. E. Moore (ed.), *Cognitive development and the acquisition of language*, New York, Academic Press.

[30] Scott, D. and P. Suppes (1958), 'Foundational aspects of theories of measurement', *Journal of Symbolic Logic*, **23**: 113-128.

[31] Shapiro, S. (2006), *Vagueness in Context*, Clarendon Press, Oxford.

[32] Stanley, J. (2003), 'Context, interest relativity, and the Sorites', *Analysis*,

[33] Tversky, A. (1977), 'Features of similarity', *Psychological Review*, **84**: 327-352.

[34] Varzi, A. (2007), 'Supervaluationism and its logic', *Mind*, **116**: 633-676.

[35] Veltman, F. (1987), *Syllabus Logische Analyse 2: Vaagheid*, Universiteit van Amsterdam.

[36] Wheeler, G. (2002), 'Kinds of inconsistency', in W. Carnieli, M. Coniglio and I Loffredo D'Ottaiano (eds.), *Paraconsistency: the logical way to the inconsistent*, LNPAM, New York.

[37] Williamson, T. (1994), *Vagueness*, Rootledge, London.

[38] Wright, C. (1975), 'On the coherence of vague predicates', *Synthese*, **30**: 325-365.

[39] Zardini, E. (2008), 'A model of tolerance', *Studia Logica*, **90**.