Gradability without Degrees

MSc Thesis (*Afstudeerscriptie*)

written by

Silvia Gaio
(born November 3rd, 1981 in Montebelluna (TV), Italy)

under the supervision of **Prof.dr Frank Veltman** and **Dr Robert van Rooij**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:
June 24th, 2008

Members of the Thesis Committee:
Prof.dr Frank Veltman
Dr Robert van Rooij
Prof.dr Peter van Emde Boas
Prof.dr Jeroen Groenendijk

Institute for Logic, Language and Computation

Contents

Acknowledgements

I am grateful to my supervisors, Frank Veltman and Robert van Rooij, especially for their patience during our long discussions. I started my research having some *vague* ideas about vagueness. They helped me to make them more and more precise, and that was not an easy job.
I warmly thank the people who helped me in the proofreading stage: Tom, Lena, Massimiliano.
I would like to thank the Evert Willem Beth Foundation for having granted me a scholarship for my first year of Master's studies.
From my discussions with many of my Master classmates I have learnt what it means to develop an interdisciplinary research in Logic. Moreover, they supported and encouraged me in several ways during my studies at the ILLC. Thank you so much for that!

Chapter 1

Sketching the Problem of Vagueness

Vagueness is a phenomenon that arises in ordinary language and involves several lexical categories[1]. It is not easy to give a precise definition of what vagueness is. What I will then present here is what vagueness phenomena look like.
Consider a British fellow named John who is 173 cm tall. Suppose an agent, referring to him, says:

(1) "John is tall".

[1] Russell in [36] even argues for the thesis that all the words in a ordinary language are vague, even the words of pure logic such as logical connectives. 'Or' and 'and' seem at first glance to have a precise meaning. That a sentence containing 'or' or 'and' is true or false depends on the truth value of the terms that are connected by the logical connectives. Nevertheless, if such truth-bearers are not precise, then the truth-value of sentences that connect them will not be precise either. Russell's argument is the following:

P1 Non-logical words are vague.

P2 Truth and falsehood, as concepts applied to propositions containing non-logical words, are vague too.

P3 Propositions containing non-logical words are the substructure on which logical propositions (containing logical connectives) are built.

C Logical propositions become vague through the vagueness of truth and falsehood of the propositions containing non-logical words.

There are of course some objections that can be raised. One is the following: vague non-logical words, when connected one to each other by logical connectives, give rise to vague propositions, but logical connectives are by themselves not vague. The source of vagueness is still natural language, and not logic itself, so it is wrong to say that logical connectives *are* vague.

How can we evaluate (1)? Is it true or not? Our indecision about the truth value of (1) might be considered some sort of ignorance. We may say, for example, that we do not know something concerning the discourse itself. That means, we do not know if our statements involving adjectives like 'tall' are true or false. That happens because we do not know if 'tall' can be applied to some cases (the so-called *borderline* cases, as we will see later on). Suppose we are acquainted with the exact measure of John's height and compare it with statistics for the group of people he belongs to, for example to British fellows. That is, we know the facts of the matter, but we might still not be able to say if John is tall or not. So, does our ignorance lie on the fact that we do not know if the adjective "tall" is used appropriately in (1)? If so, it seems that the extension of predicate "tall" is not fixed in a clear way because we are not sure about what the predicate applies to. It seems that we can say then that John is not clearly tall, nor clearly not tall. When we utter a sentence like (1), we use a predicate whose meaning is not well defined. In such a case, the view on vagueness endorsed is close to Bosch's in Bosch [9], according to which vagueness is a case of incomplete definition.
Even if we have precise pieces of information about what the world looks like, we are still ignorant of something, namely, of the truth value of (1). Since, according to Tarski's schema, "John is tall " is true if and only if John is tall, and we do not know if "John is tall " is true, it follows that we do not know if John is tall. Formally speaking:

$True\ulcorner\phi\urcorner \leftrightarrow \phi$

$\neg K(True\ulcorner\phi\urcorner)$

Then: $\neg K(\phi)$.

So, our ignorance does not just depend on a linguistics imprecision concerning the use of a term. The problem is epistemic and concerns the fact that we do not know how many centimetres are enough to say that an individual is tall. We are not able to say where the cut-off point between tall and short individuals lies, if such exists.

However, it seems that we cannot avoid using vague predicates, we communicate through and by them. But what do we communicate?
Barker [2] sketches an interesting theory about what we communicate while

uttering a vague expression. If someone asserts (1), the audience can implicate (as a sort of Gricean implicature) that the standard of tallness for her speaking community is not greater than the degree of John's tallness. Basically, what the speaker does by uttering (1) is update the communicative context. Barker assigns to (1) a metalinguistic use (under some circumstances): the speaker asserting (1) communicates how she uses the word 'tall'appropriately. Barker assigns to the speaker herself the intention to use the vague expression to convey such a metalinguistic information. Nevertheless, it seems to me that even if the speaker does not have such a conscious intention, the result of his utterance of (1) for the hearer is that whatever the context is, the hearer will update her information states: she gets the information that for the speaker the standard of tallness is at least as high as John's tallness.

1.1 Matter of Perception or Matter of Fact

Consider now the adjective 'red'. Imagine a wall that is red on the left and progressively becomes orange on the right. So, we have a progressive series of colour chips. Each chip differs from the next one by an imperceptible change in hue, that is, there is a difference, but we as human beings, with our perceptive apparatus, are not able to see it. So, if we take the first chip from the left side and then the second one, any human being would not notice any difference, so anyone would infer that, if the first chip is red, the second is red as well. And so on, taking into consideration only two chips one next to the other, applying this reasoning we would come up saying that also the last chip, on the right side, is red. But we perceive it as clearly orange. This is a form of the Sorites paradox (see next section) that involves a series of questions on perception. We have to consider "the difference between being red (a question of fact) and seeming red (a question of perception)"[2]. There are precise measurement devices that can determine whether an object is red, for any object you pick. Conversely, our perceptual ability is not great enough to determine with precision whether an object is red.
So, when we have cases where our perceptions are not precise enough to determine the application of a colour predicate, we could stipulate that an expert could decide which chips are red, and which ones are orange. That means, we might use an instrument to get the exact measure of the hue of some colour. The colour expert can use instruments to measure the

[2]Barker [3], p. 3 (online version). See also Raffman [35].

precise amount of hue that there is in a part of our wall. In such a case, it seems that the problems of the speakers concerning the applicability of the colour-predicates can be solved thanks to an appeal to an expert's knowledge. I will call this objection to the standard way of thinking of vagueness (as a problem to be solved exclusively in some theory on natural language) the 'expert objection'. A thesis that a supporter of the expert objection may take on his own is that colour predicates like 'red' are not observational.

There are some possible replies to the expert objection. I rehearse those of R. Parikh[3]. First of all, if we have an instrument that can discriminate differences greater than a certain x, then we can still produce a paradox for differences between two objects smaller than x. Secondly, saying that a colour predicate is not observational seems to go against our intuitions, since we decide whether an object has a colour by looking at it. So, even though we can develop a non-observational theory about colour-predicates, we nonetheless need a theory about how we *use* such predicates. Even if there is a community of experts that are able to determine the exact applicability of some predicate, in our normal use of language we cannot always ask an expert to judge whether a wall is red. Speakers use 'red' according to their colour perception, and not to any precise scale of hues or similar devices. Moreover, what Parikh stresses is that we learn the use of this kind of predicate by ostension, since it is not possible to give them definitions as we do for some other adjectives, for instance 'bachelor'[4]. Imagine a mother showing her child red objects, in order to make it learn what 'red'means. She cannot, though, discuss with it the colours of the objects it will see during his life. So, it might be that the child fixes an extension of 'red'that has much in common with its mothers words extension, but is not exactly the same. Let $X = \{o_1, ..., o_k\}$ be the set of objects described by the mother as red and $Y = \{o'_1, ..., o'_k\}$ the set of objects described as not red. The child learns about R, the set of red objects, the following:

$$X \subseteq R$$

$$R \cap Y = \emptyset.$$

[3] See Parikh [31], pp. 522-523.

[4] We can learn what 'bachelor' means only by looking at a definition: 'bachelor = not married man'. As soon as we learn that, there are not any doubts about when it applies or not. But we cannot do the same for 'red'. We need to *see* some red things and someone telling us that those object are red.

Those two conditions might be satisfied by more than one set R. X might be the smallest R, while Y^- the largest. Now, the child probably will not fix X nor Y^- as the right set R, since the child will recognise other objects than those collected by X that are red, and in the objects contained in Y^- it can find some examples of non-red objects. The child will choose some set similar to the one chosen by its mother as the extension of ‘red’, but nothing guarantees that they are the same. This theoretical argument seems to be confirmed by an experiment based on the Munsell colour chart. Its technical details can be seen in Parikh [31], p. 524. I report only the main result: we have disagreement about category boundaries between colours. Even if we arrive at a definition of ‘red’, it does not mean that we already have such a definition and that it governs our use of the word. As Parikh states, “if we were to *now* define the color red as light of wavelength 6,000 angstrom units, that would not explain how Shakespeare was using the word and why we are able to understand him”[5].

1.2 Characteristics of Vague Expressions

Up to now I mentioned only vague predicates. They are probably the most widely discussed example of vague expressions. I concentrated on them in my research, but in these introductory pages I want to mention five other interesting kinds of linguistic expressions that are considered vague as well:

1. adverbs (‘quickly’, ...)
2. quantifiers (‘almost’, ‘most’, ‘many’,...)
3. modifiers (‘very’, ‘clearly’, ...)
4. relations (‘be friend of...’, ...)
5. proper names of entities (‘Mont Blanc’, ‘Sahara’, ...) or names that refer to phasal sortals (‘embryo’, ‘teenager’, ...) and to specific objects (‘Teseo’s ship’, ...). With this kind of terms we have the so-called problem of indeterminacy of identity.

After having seen possible examples of vague expressions, we try now to characterise vagueness. We need to have a good understand of the concept

[5] Parikh [31], p. 525.

of vagueness that is used in the philosophical and linguistic literature.
Three features are attributed to vague expressions (I follow Keefe-Smith [16] and Keefe [15]):

1. Borderline cases: these are cases where it is not clear whether the expression applies. Some people are borderline tall, as well as a person is a borderline case of child when it is not clear if he is still a child or a teenager, and so on. The indeterminacy concerning the applicability of vague expressions to some cases may allow us to think that the respective sentences are neither true nor false. For instance, if John is a borderline case of tall man, sentence (1) turns out to be neither true nor false. This contrasts with the principle of bivalence. As we will see, some theories of vagueness vow to do away with such a principle.

2. No sharp boundaries: on a scale of heights, what is the exact measure that makes a person tall? What is the amount of time that makes a child become a teenager? With respect to predicates, having no sharp boundaries means lacking definite extensions. This is closely related to the presence of borderline cases. The indecisiveness about borderline cases makes it difficult to draw a sharp line between the extension and the anti-extension of vague predicates. But for there to be no sharp boundary between the elements that have the property P and the elements that do not have P does not coincide with there being a region of borderline cases. If such a region were sharply bounded, then P would have a sharp boundary with its borderline cases. However, fuzzy boundaries affect borderline regions too, so the existence of fuzzy boundaries for P is not exactly the same as the existence of borderline cases of the applicability of P[6].

3. Sorites paradox: vague expressions are susceptible to Sorites paradoxes, known also as 'little-by-little arguments'. Soritical arguments will be considered in detail in the following section.

Only if an expression shows all these three features we can call it vague. Expressions that have only one or two of these features are not properly considered vague. In fact, taking those features into account, we can now consider what vagueness has to be distinguished from:

[6]The fuzziness involving the boundaries of the region of borderline cases is closely related to the problem of higher-order vagueness. For my present purposes, though, I will not deeply analyse this problem.

- Ambiguity: terms like 'bank' have two different main senses. Both senses, moreover, can be considered vague, but vagueness does not concern the indecision over fixing one of those senses as the intended one. That is just ambiguity. Moreover, you can disambiguate ambiguous words by fixing a context and looking at their referents (the objects they refer to). There are no ambiguous objects, and for each context of use we can detect the right referent for the ambiguous word. This is not the case of vagueness: even fixing a context and looking at the objects does not help in giving (or finding) sharp boundaries to vague expressions[7].

- Underspecificity: underspecified terms are not adequately informative for the purposes of the discourse. Consider some expressions as examples of underspecificity:
  "Someone said something": we have no borderline people, nor borderline cases of saying something.
  "x is a natural number greater than 40": the predicate 'being a natural number greater than 40' has no fuzzy boundaries, no borderline cases, and does not give rise to a Sorites paradox.

- Mere context-dependence (usually regarding predicates corresponding to adjectives): that a predicate has different extensions in different contexts is not enough to say that such a predicate is vague. Take a vague predicate and fix a context: the extension of the predicate in that context might be still vague. So, in order to characterise vagueness, it does not suffice to say that a predicate is vague *because* its extension varies from context to context.

1.3 Sorites Paradox

We have seen some conceptual difficulties in considering a series of objects (in our examples, parts of a wall) that differ one from the other only by a small degree in some measurement. The crucial point is that small changes do not seem to bring any big consequence, but we might have a very big difference between from a_1 to a_{100}, even though between any a_n and a_{n+1} there is no big difference. From such a tension between small changes and big consequences the so called Sorites Paradox arises. The paradox is also called 'the paradox of the heap', but it is usually referred to by the name 'Sorites', that comes from the Greek name 'soros' that means, precisely,

[7]See also Prinz [34], paragraph 5.

'heap'. The first formulation of the paradox is attributed to Eubulides of Miletus (belonging to the Megarian logic school, IV. BC), to whom also the first version of the Liar paradox is attributed.
The usual formulation of the argument consists of two premises and a conclusion:

- 1 grain of sand does not make a heap
- adding a single grain of sand to a collection of grains of sand that does not make a heap, does not turn that collection to a heap

If we assume both the premises, then no matter how many grains we have, we never get to have a heap.

The paradox works also in the other way around. Consider the following premises:

1. 10000 grains of sand make a heap
2. taking a single grain of sand off from a heap, we still have a heap

So, if we start with a 10000-grains heap and take one by one 9999 grains off, we will get a heap made by only one grain; but that is against our intuitions: a single grain does not make a heap.
There are at least three ways to formalise the Sorites paradox[8]:

- Consider a series of objects $o_0, ...o_n$ and a predicate F applying to o_0, but not to o_n. We have:

  Premise 1 $F(x_1)$

  Premise 2 $\forall i : F(x_i) \rightarrow F(x_{i+1})$

  Conclusion $F(x_n)$

  If we interpret F as the predicate *short*, and $x_1, ...x_n$ as a series of persons that differ one from the next one for one millimetre with respect to height, the conclusion is false. It must be mentioned also that the second premise is the so-called *inductive step* and here it is formulated by a universal quantification.

[8] For the considerations about the Sorites paradox I refer to the first chapter of Keefe [15].

- Replace the quantified inductive premise with a sequence of conditional premises of the type $F(x_i) \to F(x_{i+1})$, that is:
  $F(x_1) \to F(x_2)$
  $F(x_2) \to F(x_3)$
  $F(x_3) \to F(x_4)$
  ...

- It is not necessary to express the inductive premise with a conditional. Moreover, if it were not interpreted as a material conditional, the paradox would not arise. We can substitute the sequence of conditional premises as follows:
  $\neg(F(x_1) \wedge \neg F(x_2))$
  $\neg(F(x_1) \wedge \neg F(x_3))$
  $\neg(F(x_3) \wedge \neg F(x_4))$
  ...

  Or, in alternative to the sequence, a unified quantified premise:
  $\forall i, \neg(F(x_i) \wedge \neg F(x_i))$

Why do we say that the Sorites argument (in one of its forms) is a paradox? The argument has apparently true premises, the inference rules seem to be well applied, but the conclusion we get is regarded as false (or, at least, counter-intuitive). Since we are not up to accept the conclusion, it seems that there is something wrong, i.e. some mistake occurs either in the premises or in the reasoning. To find such a mistake is doing what is called the *diagnosis* of the paradox[9]. Paradoxes are considered as a kind of pathology that appears in a language (in the case of the Sorites paradox, the language is any natural one). A diagnosis for a paradox has to show what makes the argument at issue paradoxical. Possible kinds of diagnoses are the following[10]:

1. There is a mistake in the reasoning, that is, in the use of inference rules;

2. One of the premises is not true;

[9] For the terminology used in the philosophical discussion on paradoxes I refer to Bradley Armour-Garb [1].

[10] See Armour-Garb [1], p. 116 and Keefe [15], pp. 19-20.

3. The conclusion is only apparently false;
4. There is no mistake in the reasoning, nor in the premises, and the conclusion is actually false: so, some concept involved in the premises is either incoherent or has limited applicability. For instance, in the case of Sorites paradox, the vague predicate involved might turn out to be incoherent, whenever we accept premises, reasoning and conclusion of the argument.

Mostly, the proposed solutions for the Sorites paradox focused on the second diagnosis. The inductive premise (in one of its possible formulations) is the target usually chosen. For example, the epistemic theories deny the inductive premise and state there is a i such that $F(x_i) \wedge \neg F(x_{i+1})$, even if we do not know which i is.
A rather extreme (usually called *nihilist*) position supported by Peter Unger embraces the fourth kind of diagnoses: there is no mistake in the argument and the problem is the incoherence of some predicates. There is no extension at all for vague predicates, because they are intrinsically incoherent (see Unger [41]).
However, Unger's view is not the unique way to take if you accept the soundness of the Sorites argument. The problem can be seen from a pragmatic point of view. Dummett [11], for instance, argues for the thesis that the *use* of (at least some) observational predicates (like colour predicates) is inconsistent. Usually, we face situations where we are able to use observational predicates without difficulties and there is no inconsistency in those cases. Some problems arise when we have a Sorites series. In such a case, vague predicates are too coarse to be fruitfully used. We need more precise tools to describe the situation and detect the differences between objects. A pragmatic solution on the basis of Dummett's observations was developed by Veltman and Muskens later on[11]. Considerations of pragmatic nature like Dummett's ones are also the theoretical basis of the model proposed in chapter 4.

[11] See Veltman & Muskens [45].

Chapter 2

Approaches to Sorites: an Overview

Theories of vagueness usually aim to provide an account of borderline cases, to explain why vague expressions (apparently) have no sharp boundaries, to supply a logic and semantics to vague languages, and to deal with the Sorites paradox. Some theories deal only with some, but not all, those aspects. In this chapter I focus on the main approaches in the philosophical literature that deal with all of them and that, in particular, offer a solution to the Sorites paradox.
There are four approaches to vagueness, and in each of them different kinds of solutions are developed. The four approaches are: semantic, epistemic, realistic, contextualist. I briefly sketch those accounts in the following lines.

2.1 Semantic approach

Vagueness is due to the way the natural language is related to the world. When we have a sentence involving a vague predicate applied to a borderline case, we are not willing to give a determinate truth-value to that sentence. To provide a reason of that, two theories refuse the principle of bivalence:

Fuzzy Logic : it is a kind of logic that allows for degrees of truth-value. The theories that appeal to Fuzzy Logic are often called *Degree theories*[1]. A predicate is no longer a function from objects to the set

[1]It has to be noted, though, that in linguistic literature degree-based theories developed for vague adjectives do not make use of Fuzzy Logic. While Fuzzy Logic provides a scale of degrees of truth-values ranging from 0 to 1, a linguistic degree-based theory considers

$\{0, 1\}$ (with 1 standing for *true* and 0 for *false*), but a function from objects to the values in the unit interval [0,1]. So, an object might be .2-red, and not either red or not red. Such a continuous series of degrees of truth is theoretically supported by the observation that the Sorites Paradox is developed on a continuous chain of cases.
The degree of truth of a compound formula is given by the degrees of truth of its components. Consider the constraints for the evaluations of negation, disjunction, conjunction and conditional:

$$\mathcal{V}(\neg\phi) = 1 - \mathcal{V}(\phi)$$

$$\mathcal{V}(\phi \vee \psi) = max(\mathcal{V}(\phi), \mathcal{V}(\psi))$$

$$\mathcal{V}(\phi \wedge \psi) = min(\mathcal{V}(\phi), \mathcal{V}(\psi))$$

$$\mathcal{V}(\phi \rightarrow \psi) = \begin{cases} 1 & \text{when } \mathcal{V}(\psi) \geq \mathcal{V}(\phi) \\ 1 - \{\mathcal{V}(\phi) - \mathcal{V}(\psi)\} & \text{when } \mathcal{V}(\phi) > \mathcal{V}(\psi) \end{cases}$$

The constraint for the conditional can be intuitively read in this way: if the antecedent has a higher degree of truth than the consequent, the conditional cannot be entirely true. Hence, given a conditional in a sorites argument:

$$F(x_n) \rightarrow F(x_{n+1}),$$

if x is a borderline case of F, the antecedent has a degree of truth less than 1 and greater than 0, and higher than the consequent. So, the conditional is not entirely true. Call $F(x_0), F(x_1), ...F(x_n)$ a sorites series. Intuitively, the degree of truth decreases as we go on along the sorites series. This seems to be plausible; it also reflects the intuition of the gradual passage from F to $\neg F$ without any sharp boundary.
This theory presents some problems, though. First of all, one of the reasons to adopt Fuzzy Logic is to be closer to speakers' intuitions about the truth-value of sentences containing vague expressions. It

adjectives as functions from objects to degrees on a scale. Such a scale does not consist of truth-values degrees and, moreover, the range is from 0 to infinity rather than from 0 to 1. Degree-based theories for vague adjectives will be treated in chapter 3 and not in the present chapter, since they do not present a solution to the Sorites paradox.

seems that a standard two-valued approach is not good because it cannot provide an agreement about the applicability of a predicate to some objects. But a plural valued approach like Fuzzy Logic does not seem to produce a better agreement. Consider a situation where it is lightly raining. Since it might seem not completely true to assert "It rains", if we were to take such an approach seriously, we could claim "It rains at 0.45 degree". However, such a claim is certainly not intuitive or natural at all. Secondly, it is not clear where these degrees come from and how we can state that an object is exactly .33-red and not .34-red. A third problem and, in my opinion, the most serious, arises from the semantics given for Fuzzy Logic. Recall the constraints for negation, conjunction and disjunction stated above. Consider an object o that has a degree of redness equal to .5. The sentence "o is red" has truth-degree of .5. But then, also its negation, "o is not red", will have degree equal to .5. If we take now the conjunction of the two sentences, we get something we do not want, that is: "o is red and is not red" has degree .5, and not 0 as we would expect, since it is contradictory. That means, if we get rid of bivalence and accept a degree-based theory of truth, then we can get unwanted results, such as a contradictory statement that is not (completely) false.

Supervaluationism : we have precisifications, that is, partial assignments of truth values to statements. Formally, given a language $\mathcal{L}$ of predicate logic, a supermodel $\mathcal{M}$ for $\mathcal{L}$ is an ordered pair $\langle \mathcal{D}, \mathcal{J} \rangle$ where

1. $\mathcal{D}$ is the domain of $\mathcal{M}$, namely a non empty-set;
2. $\mathcal{J}$ is a non empty set of partial interpretations such that
   - each $\mathcal{I} \in \mathcal{J}$ assigns to each constant a an element d of $\mathcal{D}$ such that $\mathcal{I}(a) = d$;
   - each $\mathcal{I} \in \mathcal{J}$ assigns to each n-ary predicate P a partial function $\mathcal{I}(P)$ from $\mathcal{D}^n$ into $0, 1$. This partial function is what is called *interpretation.*

The sentences that are true (or false) in a model have to remain true (or false) when the model is refined.

Consider what a refinement relation between models consists of. Take two models $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$ and $\mathcal{M}' = \langle \mathcal{D}, \mathcal{I}' \rangle$. They both have the same domain, but the interpretation function varies.

Definition 1 *$\mathcal{M}'$ is a refinement of $\mathcal{M}$ iff:*

1. *If P is n-ary and $\mathcal{I}(P)(\langle d_0...d_n\rangle) = 1$, then $\mathcal{I}'(P)(\langle d_0...d_n\rangle) = 1$;*
2. *If P is n-ary and $\mathcal{I}(P)(\langle d_0...d_n\rangle) = 0$, then $\mathcal{I}'(P)(\langle d_0...d_n\rangle) = 0$.*

The definition intuitively tells us that if we take an interpretation function $\mathcal{I}'(P)$ that is a refinement of $\mathcal{I}(P)$, $\mathcal{I}'(P)$ preserves the same truths and the same falsities as $\mathcal{I}(P)$, and maps at least one more element from $\mathcal{D}^n$ into $0, 1$. In the philosophical literature on vagueness a model $\mathcal{M}'$ that is a refinement of $\mathcal{M}$ is called a *sharpening* or *precisification* of $\mathcal{M}$.
Other constraints are the following:

- for all $\mathcal{I}, \mathcal{I}' \in \mathcal{J}$ and each individual constant a: $\mathcal{I}(a) = \mathcal{I}'(a)$
- there is a $\mathcal{I}_0 \in \mathcal{J}$ such that for all $\mathcal{I} \in \mathcal{J}$: $\mathcal{I}$ is a refinement of $\mathcal{I}_0$
- for all $\mathcal{I} \in \mathcal{J}$ there is a $\mathcal{I}' \in \mathcal{J}$ such that $\mathcal{I}'$ is a refinement (precisification) of $\mathcal{I}$

In such a framework, sentences that are true in all the precisifications are said to be *supertrue* (hence, the name *supervaluationism*). Sentences that are false in all the precisifications are said to be *superfalse.* But sentences describing borderline cases for some vague predicate are true with regard to some precisifications, and false regarding to others. They are said to be neither true nor false, that is, to be lacking in truth values. Supervaluationism is then committed to a non-bivalent logic and specifically a logic that admits truth-value gaps. Such gaps are meant to capture the idea of the indeterminate application of vague predicates to borderline cases.
Logical truths turn out to be supertrue, while falsities (contradictions) are superfalse. Here we get then an improvement over Fuzzy Logic: a sentence like "*o* is red and is not red" will be false in all the specifications, and "*o* is red or is not red" will be true in all of them. One thing to notice is that if "*o* is red" is true in some precisifications, false in others, also "*o* is not red" will turn out to be true in some precisifications, false in others. In such a case "*o* is red and is not red" remains false in all the precisifications, while "*o* is red or is not red" remains true in all of them. That means, supervaluationist semantics is not truth-functional, that is, the truth value of a compound sentence is not determined by truth values of its components.
What solution for the Sorites paradox is provided by supervaluationism? The supervaluationist logical machinery makes one of the

conditional premises neither true nor false (consider sofar the second version of the Sorites paradox). Take for instance the conditional $F(x_n) \rightarrow F(x_{n+1})$. There are some precisifications of the interpretation of F that make both antecedent and consequent true, some precisifications that make both of them false, and one precisification that makes the antecedent true, but the consequent false. The Sorites reasoning is then blocked if in the series there is a conditional premise with those semantic features.
The problem that we have with supervaluationist theories is what is called *higher-order vagueness*. There is no distinction between borderline and not borderline cases. Another problematic issue for supervaluationist theories is that they refuse the truth of the induction premise. But if this means that they negate it, then they should accept that there is a i such that $F(x_i)$ and $\neg F(x_i + 1)$, as the epistemicist does, as we will see later. However, they do not accept that F has a clearly determined extension. In other words, the following formula turns out to be true in all precisifications:

(SB) $\exists n(F(x_n) \wedge \neg F(x_{n+1}))$.

(SB) claims is that there is a sharp boundary between F and $\neg F$. Even if (SB) is supertrue, supervaluationism keeps it true that there is no particular number satisfying the existential quantification in (SB). That means, it is supertrue that there is a n $F(x_n)$ and $\neg F(x_{n+1})$, but the formula is not true nor false for any n you consider. That seems a counter-intuitive result[2].

There is another semantic theory, that does not discuss the principle of bivalence, nor refuse the inductive premise. On the contrary, it assumes that the Sorites argument is sound: the premises are true, the reasoning right. The conclusion, then, is also true, but it sounds wrong to us: that happens because the vague expressions involved in the argument are incoherent. So, basically, the whole Sorites argument is sound because, trivially, there is no such a ordinary thing like a heap. This theory is called *nihilist* and is supported by Unger (See Unger [41]).

Some other philosophers carry on a non-classical reasoning. Consider borderline cases of predicate ‘tall’: they are borderline cases both of ‘tall’and

[2]See Olin [30].

'not-tall'. According to Max Black [8], for example, if John is such a borderline case, (1) is both true and false. In other terms, instead of speaking of truth-value gaps, such an approach considers truth-value gluts. Some of its supporters are dialetheists or paraconsistent logicians: Graham Priest, JC Beall and Marc Colyvan (see, for example, Beall&Colyvan [4]).

2.2 Epistemic approach

Roy Sorensen and Timothy Williamson are the most well-known supporters of epistemic theories for vagueness[3]. Briefly sketched, the main thesis is that vague predicates have determined extensions and the sentences containing them are either true or false. The problem is that we do not know where the border of the extension is, nor the right truth value. That is, vagueness is due to our ignorance, not to our use of language.

The diagnosis of the Sorites paradox provided by the epistemic theory detects the problem in the inductive premise (premises). One instance of the universal inductive premise, or equivalently one of the conditional premises, is considered as not true. We do not know which specific instance this is, because our knowledge is limited, but the only way to get out of the paradox is to admit the existence of such a false instance.

In this approach logic is still classical, the principle of bivalence valid. We can consider the burden of Williamson's work to show that classical logic can still be considered as the logic of natural language. Nevertheless at the same time, the question that arises is: why should we keep classical logic for a theory that dogmatically says that there are cut-off points, without us being able to know them?

Moreover, another objection that can be raised against the epistemic view comes directly from its core. The statements about the existence of unknowable and mind-independent truths sound implausible: if there are unknown truths about the use of some words, how can we generally know what we are saying? According to Williamson, the theoretical understanding of words come together with linguistic practice. But I cannot see how this can justify the claim that we know some mind-independent truths, while we do not some others (relative to borderline cases), and moreover, that the latter are not knowable at all.

[3]To give some samples of their work, see their outstanding monographs on vagueness: Williamson [47], Sorensen [40].

2.3 Realist approach

The realist approach supports a rather extreme ontological thesis, since it claims that vagueness is not a problem of our language, nor of our cognitive states, but it is a feature of the world. The problem is brought to ontology. At the present time, T. Parsons and P. Woodruff stand out for their peculiar theory about indeterminacy of identity: even the relation of identity between objects is affected by vagueness[4]. The vagueness of our language reflects the vagueness of states of affairs. Russell's criticism towards a realistic view of vagueness became famous. Russell [36] attributed a fallacy of verbalism to realistic theories of vagueness. According to him, it is mistaken to attribute a property of language (vagueness) to the world. In fact, in his words:

> Vagueness and precision alike are characteristics which can only belong to a representation, of which language is an example.

So, vagueness and precision are properties of words and cannot be considered as properties of things.

A realist approach is usually invoked in the issues concerning the identity relation. The claim is that some objects (like Mt. Everest) are vague because they lack of determinate boundaries. In other terms, the referents of singular terms are considered vague. But there are also predicates that seem to generate the same problem. Take the predicate 'cloud' and one of its instances: how can we determine the exact boundaries of the object we want to consider? It seems that the evanescent nature of clouds prevent us from finding their boundaries. But can we think of ontological vagueness when we consider predicates like 'tall', 'red', ...? What does it mean to say that the vagueness of such predicates has an ontological nature? The idea is not that those predicates correspond to vague properties existing in the world, as though these were objects. The idea is rather the following: a sentence containing a vague predicate applied to some borderline case corresponds to a state of affairs. It is such a state of affairs that happens to be vague, according to the realist approach to vagueness. Put in different terms, such an approach claims that there is no determinate fact of the matter whether a borderline case of tall man has the property of *being tall*, the latter conceived as something you can find in the world. The corresponding proposition is thought of as lacking truth value.

[4] See in particular: Parsons [32] and Parsons-Woodruff [33], [48].

2.4 Contextualist approach

There is another way of approaching vagueness, that is, to consider the context-dependence a central feature of the vagueness phenomena. Far from identifying vagueness with mere context-dependency (since, as we have seen, the two phenomena are not the same), what is underlined is that in any case vague predicates are also context dependent. Contextualism focuses more on predicates than on other linguistic expressions. But then, all the debate concentrates especially on vague predicates. Many theorists from Philosophy and Linguistics support contextualism, following Hans Kamp and Peter Bosch. We cautiously avoided identifying vagueness with context dependence. However, within this research work we are concerned with a specific kind of vague linguistic expression, that is, adjectives. Many of them are context-dependent, that is, their extension varies from context to context. So, *context-dependence* is recognised to be a feature of some predicates. For instance, the predicate 'being tall' applied to the context of human beings has a different extension than in the domain of equatorial trees. Far from saying that vagueness of adjectives like 'tall' is solved by saying that their meaning depends on the context taken into account, a contextualist approach to vague predicates wants to include also a treatment of context-dependence in a theory of vagueness.
According to Delia Graff Fara, many contextualist theorists (from Linguistics and Philosophy) have the goal of providing an answer to the question she calls *Psychological Question*[5]:

> If the universally generalized sorites sentence is not true, why were we so inclined to accept it in the first place? In other words, what is it about vague predicates that makes them seem tolerant, and hence boundaryless to us?

The context-dependence of vague predicates consists of two factors (I refer here to Barker [3]):

- dependence on a contextual standard: Bill is tall if he is tall at least to some degree d, where d is a "threshold for tallness provided by the context"[6]. Such a threshold for 'tall' can be thought of as the measure that is the average of the heights of the individuals in the context. This is not enough though. For example, consider the following example

[5] Fara [12], p. 50.
[6] Barker [3], p. 6 (on-line version)

(given by Fara [12], p. 55-56): suppose we want to state the standard of height for basketball players. Suppose that all the tall basketball players are killed under some tragic circumstances. Now, the average height relevantly decreases. But in such a case, it does not seem to be true to say that the tallest surviving basketball player is tall for a basketball player. So, to make an average is not always intuitively enough to find the threshold. Consider also a second example, given again by Fara. By some bizarre coincidence, all the basketball players happen to be also golfers. Now, it seems rather unfair to say that who is tall for a golfer is tall for a basketball player[7]. We have to consider the *typical* height for things of some specific kind. I doubt there can be exact criterions to do that: we have to trust our intuitions on the specific kind of things we compare. And here we get on to the second feature:

- reference to a comparison class: the threshold is often provided with respect to a set of objects and the predicate at issue is judged with respect to such a set. Such a set is called *comparison class*. The denotation of a predicate is then relativized to a class of objects; to give an example, 'tall' can be relativized to 'basketball players'. In sentences containing vague predicates, the comparison classes can also be specified ("Maria is tall for an Italian woman"). However, 'tall for an Italian woman' is equally as vague as 'tall'. We can also use the vague adjective in an attributive function: "Maria is a tall Italian woman".
  The relativisation of predicates to a comparison class brings us to think that adjectives like 'tall' are not intersective, that is, the extension of 'tall man' is not given by the intersection between the extension of 'tall' and 'man'. Namely, consider the case if John is a tall man and a basketball player. If we intersected the extensions of 'tall', 'man' and 'basketball player', we would get that John is also a tall basketball player, but this is a conclusion we cannot make, if our information is just that John is a man, tall, and plays basketball. Namely, he might well be a tall man, but a short basketball player.
  But what should we consider as comparison classes? For example, we cannot say "Bill is tall for the objects in this room": it is unnatural to say that there is a standard, a typical height among the objects present in a room. That means, we have no notion of what 'being tall' for the things in the room means, nor of what the typical degree of tallness for

[7]Fara [12], p. 56.

such set of objects must be. Fara [12] claims that comparison classes have to be *natural kinds*. Whether one object considered in a certain natural kind has a property depends on the typical standard of that natural kind, and we have some sort of idea of what such a degree is. But, as we have already noticed, the notion of natural kind refers to our intuitions, and is not strongly-characterised. A further objection to the theory of natural kinds can also be the following: consider a room containing only two distinct objects, a chair and a cupboard. It seems plausible to claim something like "the cupboard is tall for the objects in this room": we are comparing the tallness of the chair with the tallness of the cupboard and that seems correct. Even if there are objections to the theory of natural kinds, an advantage of the latter is that you can accurately predict that when someone says "John is tall", she means that John is tall compared with the natural kind of men, for instance. In such a case, it would be unnatural to think that the speaker aimed to compare John with the objects in a room. So, at least for sentences where a comparison class is not explicitly mentioned, it is useful to refer to natural kinds to make the comparison class explicit.

The two factors detected by Barker [3] are not enough to characterise context-dependence. There is a second sense of context-dependence that has to do with the domain of the discourse. Suppose you have a class of eight-years-old children. All of them are part of the natural kind of eight-years-old children, but still you want to compare each one of them with all the others and say who is tall. It might happen that some children that are classified as tall through such a comparison are not tall if considered in the natural kind of eight-years-old children (or some tall eight-years-old children might turn out to be short in the context of the class.).
You can compare objects in a Sorites series taking them in larger or smaller contexts. Dummett made some interesting remarks about observational predicates. If you consider each time only a single pair of objects in a Sorites series generated by an observational predicate (like, for instance, 'red'), one of which is next to the other in the series and they are not observably distinguishable, you accept each sentence of the type $F(x_i) \rightarrow F(x_{i+1})$. But still you do not accept all the sentences taken together, that is, considering all the objects at once. Such an idea to consider in a different way the pairs of objects of a Sorites series and the series itself as a whole has been formalised by Veltman and Muskens[45].
At first glance, this kind of contextual solution to the Sorites paradox seems to be peculiarly faithful with respect to the intuition that in natural

language we give meaning to sentences considering them within a context. It is common practice to consider the meaning of predicates with respect to the individuals in a specific context, that does not always coincide, though, with a comparison class (natural kind).

2.5 Focus on Adjectives

The approaches mentioned in this chapter provide a general treatment for vague expressions. In my research I want to focus only on a specific sort of linguistic expression: adjectives. In order to be able to approach the problem of vagueness of adjectives it seems to be prerequisite work to understand how we use this kind of adjectives and how they can be modelled. First of all, the features of vague adjectives have to be underlined: authors that deal with vagueness problems quite often refer to polarity and gradability as instances of features of vague adjectives. We need to understand whether and how those notions are connected. This is the aim of the present research: to make clear the features of vague adjectives and to provide a model that shows how we actually make use of them in our ordinary speech.

In the next chapter the most widely known linguistic theories for vague adjectives are presented, and the problematic assumptions underlying these theories are analysed. In chapter 4, then, we propose an alternative view on gradable and polar adjectives, that tries to explain how we use these adjectives in cases that are unproblematic, taking into account two other factors related to vagueness: the notion of comparison class and the notion of granularity. Chapter 5 collect some further results and remarks of a theoretical nature.

Chapter 3

Theories for Gradable Adjectives

In this chapter, I present the features of the so-called gradable polar adjectives, and then the most prominent linguistic theories that account for them.

As we have seen, a class of vague predicates is that of predicates generated by adjectives such as *tall, long, expensive, ...*. Those are called *gradable* adjectives because they have the following features:

- they can occur in a predicative position, that is, after verbs such as 'be', 'become', 'seem', ...;
- they can be preceded by degree modifiers such as 'very', 'clearly', ...;
- they can be made into comparatives and superlatives.

Most of them express some properties that can be measured on a scale of size or value.
On the other hand, non-gradable adjectives are usually not vague. Examples of non-gradable adjectives are *married*, *female*, *bachelor*. If they are modified by degree adverbs, the effect you get is just an emphasis (see, for instance: "It is very true"). Someone can argue that non-gradable adjectives do not give rise to vagueness because they have a fixed and defined meaning, that is, they do not vary their meaning along a scale of degrees. But more important than that, non-gradable adjectives cannot be made into comparatives nor superlatives. It does not make any sense, for example, to utter something like:

(a) *"Mary is more female than Lucy"

nor

(b) *"Mary is the most female in the country".

If someone utters (a) or (b), she is using 'female' for some specific pragmatic purpose, that is not the ordinary use of the adjective. As adjectives of this kind do not show borderline cases nor generate a Sorites paradox, they are not vague.
Two theories account for the differences between gradable and non-gradable adjectives[1]:

- The domains of gradable adjectives are ordered according to some property that is usually measurable and allows grading (such properties usually correspond to dimensions). For example, the domain of 'tall' is weakly or partially ordered according to the property of *height*; the domain of 'warm' according to the property of *temperature*.
- According to the theory of adjectives developed by Klein[2], non-gradable adjectives are represented by complete functions from individuals to truth values, while gradable adjectives are represented by partial functions from individuals to truth values. That is, in the latter case, the functions can give value 0, 1 or no value at all for some objects in their domains.

Gradable adjectives, moreover, can be distinguished between relative and absolute[3]:

Absolute Adjectives They have positive forms that relate objects to maximal or minimal degrees, and are not affected by the Sorites paradox, nor do they have borderline cases. They differ, though, from the non-gradable ones, since they demonstrably have all the features of gradable adjectives: they require their arguments to possess some minimal or maximal degree. Moreover, absolute gradable adjectives can acceptably form comparatives and superlatives, can be modified by some degree modifiers, and can occur in a predicative position. Consider some examples. 'Wet' requires its argument to have a *minimal* degree of the

[1] See Kennedy [19].
[2] See, for instance, Klein [23], that recalls the theory by Kamp [14].
[3] See Kennedy [22].

property it describes (adjectives such as 'wet' or 'open' are called *minimum standard* absolute adjectives). The polar counterparts of 'wet' and 'open' are respectively 'dry' and 'closed'. As they require their argument to possess a *maximal* degree of the property in question, they are called *maximum standard* absolute adjectives. Consider the following examples to verify the acceptability of comparatives raised by absolute adjectives, and their sensitivity to degree-modifiers:
"The platinum is less impure than the gold";
"The table is wetter than the floor";
"The door is closed enough to keep out the light".

Relative Adjectives Examples of relative gradable adjectives are: "tall", "big", "expensive". The features of relative gradable adjectives in their positive form are:

- Context-sensitivity: the extension of the predicates generated by relative adjectives changes from context to context. This means also that a sentence containing a relative gradable adjective can get a different truth value depending on the context of utterance. For example, sentence (1)
  (1) John is tall,
  can be true in the comparison class of men, but false in the comparison class of basketball players. Context-sensitivity can be thought of also as the problem of the shifting standards from context to context (see previous chapter).
- Borderline cases: there are cases where it is difficult to determine whether an adjective can be attributed to some object. And moreover, there is no clear sharp boundary between a positive and a negative polar relative adjectives, as we will see below.
- Sorites-sensitivity: every relative gradable adjective can give rise to a Sorites paradox.

It seems clear that relative gradable adjectives turn out to have all the characteristics of vague expressions. Among the whole class of adjectives, then, the ones that generate vague predicates are the relative gradable adjectives. I will thus focus on this class of adjectives.

Sentences containing relative gradable adjectives can get different truth values in different contexts. In order to determine the truth value of "x is ϕ", with ϕ a relative gradable adjective, we have to determine the meaning of x, the features of the utterance context and make a "judgment of whether x counts as ϕ in that context"[4]. A semantic analysis of relative gradable adjectives will then make such a judgment possible giving to the sentence a definite interpretation and at the same time ensuring difference of interpretations across contexts.

Most of (or probably all) relative gradable adjectives have polar counterparts. Adjectives of that kind can be classified as positive or negative. Such a classification is based on some empirical characteristics demonstrated by the adjectives themselves. Measure phrases can be associated with positive adjectives, but not with negative ones (you can say "John is 178 cm tall" but not "John is 178 cm short"). Negative ones allow downward entailments, while positive ones allow upward entailments. Consider, for instance, the pair 'safe'/'dangerous', such that 'safe' is negative, 'dangerous' positive. From "It is dangerous to drive in Paris" you infer "It is dangerous to drive fast in Paris" (and also "It is dangerous to drive slow in Paris") but not the reverse, and from "It is safe to drive fast in Des Moines" (or from "It is safe to drive slowly in Des Moines") you infer "It is safe to drive in Des Moines", but not the reverse. Examples of polar pairs are: 'tall'/'short', 'expensive'/'cheap', 'big'/'small', 'clever'/'stupid'.

Just in order to capture some intuitions about how positive and negative polar adjectives are related, consider what the degree-based theory says (see following section): usually relative gradable adjectives are measurable, therefore they are related to a scale of degrees. Degrees of positive adjectives range from the lower to the upper end of a scale, while the degrees of negative adjectives range from the upper to the lower end of a scale.

3.1 Degrees

One approach to gradable adjectives is to analyse them as relations or functions from objects to degrees on a scale. We have an abstract representation (*scale*) that is a set of elements under a total ordering. Each of those elements is a *degree*. So, when we have a sentence like "x is ϕ", this is true iff the degree to which x is ϕ is at least as great as the degree on the same scale that represents the standard of ϕ-ness.

Comparatives seem to get a simple treatment within a degree-approach. Comparatives define ordering relations between degrees on some scale. A sentence like

[4] Kennedy [21], p. 34.

(2) John is taller than Mary

is analysed as follows:

(3) $\exists d([d > \iota d'.tall(Mary, d')] \wedge [tall(John, d)])$.

Take an antonymous pair of adjectives, like *tall* and *short*: they both define relations (or functions) on the same scale, that is, the scale of height. The ordering relations they give rise to are reversed. Let ϕ_{pos} and ϕ_{neg} be respectively the positive and negative polar adjectives associated with a scale S. Assume ϕ_{pos} denotes a relation between objects and $\langle S, <_\delta \rangle$ and ϕ_{neg} a relation between objects and $\langle S, >_\delta \rangle$. The set of positive degrees and the set of negative degrees on S stand in dual relation and, since there is a bijection between the two sets (namely, the identity function), they are isomorphic (in other terms, positive degrees are the same objects as negative degrees). So, for all degrees $d_1, d_2 \in S$, the following holds:

(4) $d_1 >_{\phi_{pos}} d_2 \Leftrightarrow d_2 >_{\phi_{neg}} d_1$

Therefore, (2) and (5) are equivalent:

(5) Mary is shorter than John.

The first formalisations of such an approach go back to Seuren [37] and Cresswell [10][5]. The intuition behind the degree-based approach is well expressed by Cresswell[6]:

> when we make comparisons we have in mind points on a scale.

A degree-based account is not able, though, to explain the anomaly of the so-called cross-polar phenomenon. One instance of this phenomenon is (6):

(6) * John is taller than Mary is short.

Consider (6). The degrees of tallness are the same objects as the degrees of shortness. (6) is true whenever the degree of John's height exceeds the degree of Mary's height on a scale of height. The undesirable result is that

[5] For a general discussion, see Klein [24].

[6] Cresswell [10], p. 266.

(6) turns out to be interpretable and, even worse, logically equivalent to (2) and (5), which are not anomalous but perfectly acceptable.

3.2 Intervals

Kennedy in [21], [18], [20] criticises the degree based approach because of its incapacity to explain the cross-polar anomaly as in (6). He proposes another approach, based on degrees not taken as points, but as intervals, or extents, on a scale. His theory is based on the works by Seuren [38], von Stechow [46], Löbner [25].
Focusing on the problem of cross-polar anomaly, the initial assumption Kennedy makes is [7]:

> Comparative are semantically well-formed only if they define ordering relations between the same sort of degrees: between positive degrees, between negative degrees, or between degrees that measure divergence from a reference point.

What Kennedy considers necessary to assume in an approach to polar adjectives is to make a sortal distinction between positive and negative degrees. Put otherwise, the set of positive degrees must be different from the one of negative degrees. But what he wants to assure at the same time is the equivalence between (2) and (5), respecting then one of our strongest intuitions about polar adjectives.
Antonymous pairs of adjectives convey the same kind of information about an object: for instance, *tall* and *short* refer to an object's height. What differs is the perspective from which they consider the projection of any object on some scale. A positive adjective has a 'down-up'perspective towards an object x, negative adjective an 'up-down'perspective towards the same x, so to speak.
Kennedy defines a scale S as a linearly ordered, infinite set of points. Each scale represents some type of measurement indicated: height, length, weight, etc. A degree is not defined as a point on S, but as a convex, nonempty subset of S, in the same way as an interval in a linearly ordered set of points is usually defined, and it is called *extent*[8]:

$$\forall p_1, p_2 \in d \; \forall p_3 \in S \; [p_1 < p_3 < p_2 \rightarrow p_3 \in d].$$

[7] See Kennedy [21], p. 51.
[8] The terminology and the basic idea go back to Seuren [38].

Given a scale S, define the set of positive degrees $POS(S)$ and the set of negative degrees $NEG(S)$ as follows:

$$POS(S) = \{d \subseteq S \mid \exists p_1 \in d\ \forall p_2 \in S\ [p_2 \leq p_1 \rightarrow p_2 \in d]\}$$

$$NEG(S) = \{d \subseteq S \mid \exists p_1 \in d\ \forall p_2 \in S\ [p_1 \leq p_2 \rightarrow p_2 \in d]\}$$

As a consequence of this definition, $POS(S)$ and $NEG(S)$ are disjoint. Let $pos_S(x)$ be the positive projection of an object x on S and $neg_S(x)$ the negative projection of x on S. They are ordered sets. MAX and MIN are functions from ordered sets to their maximal and minimal element, respectively. The relation between $pos_S(x)$ and $neg_S(x)$ is the following:

$$\text{MAX}(pos_S(x)) = \text{MIN}(neg_S(x)).$$

That means, given an antonymous pair, the maximal point of the interval identified by the positive adjective coincides with the minimal point of the interval identified by the negative adjective.
So, the positive and negative projections of x on S are complementary intervals on S.
Gradable adjectives are thought of as functions from objects to intervals. More precisely, positive adjectives denote functions from objects to positive intervals and negative adjectives functions from objects to negative intervals. Two antonymous adjectives, then, have the same domain but different ranges; they map the same objects onto complementary regions of the same scale.
The interval-based theory developed by Kennedy gets the same positive results as the degree-based ones. For instance, (2) and (5) turn out to be logically equivalent. Nevertheless, Kennedy's theory is also able to overcome the difficulties the degree-based approach presents. First of all, cross-polar anomalies are not acceptable: the basic idea is that this happens because *tallness* and *shortness* are not comparable, since they are different sorts of objects. This explanation is accepted in the interval-based approach thanks to the idea that positive and negative extents are disjoint. Informally speaking, (6) is true iff there is an extent that properly includes the extent of Mary's shortness, and John is tall to that extent. Kennedy's account imposes a restriction: the extent argument of a positive adjective must be a positive extent, and positive extents can include only positive extents. Similarly for negative adjectives: their extent arguments must be negative, and negative extents can include only negative extents. So, in

order for (6) to be true, the argument of the positive adjective 'tall' has to be a negative extent. Such a restriction shows that there is a sortal mismatch, and (6) turns out to be anomalous.

There are three main objections that can be raised against the degree-based approach as well as the interval approach: the first concerns the ontological commitments of the approach, the second the multidimensional aspects of some adjectives, the third involves some cognitive aspects.

1. What kind of objects are degrees? Does their use necessarily lead one to some ontological commitments? What is in doubt here is why we should assume a class of abstract objects. If we add a scale of degrees to our ontology, we have to justify it and say why we need to assume the existence of abstract objects. The ontology we have, then, is quite large, especially if we take a dense real-valued scale as the scale of degrees. The assumption of degrees does not simplify the matter: it seems to be a rather nice account for comparatives and vague adjectives, but in fact gives some problems in the ontological assumption. The question is: is it necessary to have such a large ontology and admit infinite abstract object to account for vague adjectives?
   The same question can be raised also for Kennedy's theory: what kind of objects are intervals? Kennedy is aware of this ontological question and claims to address a similar question in the paper published in 2001 (Kennedy [21]). But in that paper what Kennedy does is describe the intervals (or *extents*, that he also calls *degrees*). His argumentation seems to go in this direction: we need to replace degrees with this kind of object if we want to give the right interpretation to gradable adjectives. But what he does not explain is why we need a class of abstract objects in order to account for vague adjectives. He seems to recognise, though, this problem and in footnote 3 he refers to another article by himself (Kennedy [19]), where he tries to show why approaches that do not make use of measure theories fail. In Kennedy [19] the analyses of gradable adjectives in terms of partial functions are said to do a good job in explaining most of the semantic properties of gradable adjectives, but are not able to explain the behaviour of antonymous adjectives in comparatives (neither the anomalies, nor the normal uses). For this reason, that is, showing that the alternative theories fail to grasp some phenomena, he argues for the necessity of an approach modelled on his proposal. However, this kind of argument is not enough. He does not address the ontological problem directly. To

the question why we need abstract objects he replies saying: we need them because all the other alternatives given by now fail to grasp some phenomena that an interval-based theory does. But this is clearly an *ad hoc* argument. Again, what he needs to do is to directly justify his ontological assumptions. That intervals work well is not a *sufficient* reason to make people believe in abstract things.

2. A problem arises also in the treatment of comparatives: adjectives like *white, beautiful, wise, happy*, are different from *tall, long, expensive*. The second group of adjectives might be characterised as one-dimensional: the extension of each of these adjectives is determined by only one measurable aspect, and there is only one contrary adjective for each of them (antonym). The comparative that is made from a one-dimensional adjective is not underspecified. As far as multi-dimensional adjectives are concerned, that is, the adjectives of the first group, their extension depends upon more than one dimension and therefore might have more than one contrary adjective. Moreover, the comparatives they built seem to suffer of underspecification. Take for instance 'clever': x might be clever with respect to some ability, and we cannot usually numerically measure the amount of intelligence a person has, taking into account all the dimensions. Consider the comparative:

   "John is cleverer than Mary".

   With respect to which ability is John cleverer than Mary? We might think of an ordering for each property that determines someone's cleverness. Klein suggests that whoever utters a sentence like "John is cleverer than Mary" has already fixed a dimension, without mentioning it. But it is not clear in sentences like that which property is involved.

3. When children learn to use relative gradable adjectives, they are taught that an individual is tall and another is short while comparing those individuals between them or within a class of individuals that differ from them. A child does not measure the difference between the individuals she sees, nor has she any clue about what a centimetre is, but despite all this she can learn and properly use 'tall' and 'short'. So, it seems that we are able to use relative gradable adjectives without the notion of measurement. On such a notion the degree-based theory

is based. But if we do not need to refer to measures to use vague adjectives, why should we use measures to model our use? Can we do the same without measures and degrees?

3.3 Tropes

An alternative approach to the degree- and interval-based approach for comparatives is presented by Moltmann[9]. The central notion of her proposal is what philosophers have ended up to call *tropes*. Tropes are thought as concrete objects which adjective nominalisations refer to, and which are actually compared. For instance, the nominalisation of 'tall', i.e. 'tallness', is a trope. Hence, (1) is understood as follows:

(7) John's tallness exceeds Mary's tallness.

Let f be a function that map a property and an object to the trope corresponding to that property, relative to a world w and a time i. (1) will be analysed as follows:

(8) $f(John, [tall], w, i) > f(Mary, [tall], w, i)$.

According to Moltmann, the advantages of a trope-based account are mainly the following:

- No abstract and hardly characterisable entities are invoked. Tropes are referents of adjective nominalisations and the speakers themselves refer to them.
- Two tropes of the same type can be ordered (John's tallnes > Mary's tallness, for instance) by an exceed-relation. The ordering depends on the entities involved themselves.
- Incommensurability is naturally accounted. For example, that height cannot be compared with happiness follows directly. Nevertheless, the theory allows comparison between height and width, because they can be viewed as tropes of the same sort.

A problem for this account is the treatment of polar adjectives. If (2) and (5) are equivalent, then John's tallness and John's shortness has the

[9] See Moltmann [29] and especially [28].

same trope as referent. But then we have the following wrong inference:
(A1) John's tallness exceeds Mary's tallness
(A2) Mary's shortness exceeds John's shortness
(A3) Mary's shortness = Mary's tallness
(A4) John's shortness = John's tallness
⇒ (C1) John's shortness exceeds Mary's tallness
⇒ (C2) John's shortness exceeds Mary's shortness.
And also:
⇒ (C3) John's shortness exceeds John's tallness.
Conclusion (C2) contradicts assumption (A2), while (C3) contradicts (A4).

In order to solve this problem Moltmann proposes to consider the ordering among tropes as imposed by the concepts of the adjectives in question. She supports a trope-based account also to answer the question on which way the property expressed by the adjective imposes the ordering. Properties are construed in such a way that they resemble tropes[10],

> more precisely in terms of functions mapping indices to sets of possible tropes that resemble each other. A gradable property, moreover, will be construed as a function mapping indices to ordered sets of tropes [...] If we call tropes as conceived on the standard view *standard tropes*, then the referents of adjective nominalizations should be entities that are standard tropes 'insofar as'they instantiate that property, and this means tropes whose properties should all be based on them playing the role of instances of the property in question. Thus, besides standard another kind of trope is needed for the semantics of nominalizations. The latter should fulfil two conditions that distinguish them from standard tropes: First, only one exceed-relation should be applicable to them. Second, Johns weakness should be distinct from the entity that is Johns strength.

This approach does not seem very convincing either, especially for the not elegant (and unclear) way to accomodate polarity of gradable adjectives in the theory. Moreover, even if Moltmann claims that there is no ontological assumption of hardly characterisable entities, she refers to properties as entities. Again, we have to assume something else than just adjectives. Even if it might be simpler to understand what tropes are, why again should we assume more entities than what we want to explain? That means, we have

[10] Moltmann [28], pp. 19-20 (online version).

objects that have properties, but why should we consider these properties as objects too?

3.4 Any Alternative?

In the degree and interval approaches we considered here all the properties are assumed to be measurable. However, this is maybe an assumption that is not necessary to take. Moreover, consider the behaviour of positive and comparative form of adjectives according to the degree and interval approaches: take, for instance, the predicate 'tall'. First, the meaning of the expression 'tall to degree d' is determined. Then, the comparative 'taller than' is defined over 'tall to degree d'. Finally, the meaning of the positive form 'tall' is defined over the meaning of the comparative. Some objections against this view can be raised, as we have seen.
The trope theory does not solve the problems of abstractions that it wants to address and its assumptions are highly objectionable, as I have briefly sketched.
Now, is it possible to have an alternative theory that explains how we use relative gradable adjectives without assuming degrees, nor intervals, nor tropes? In the following chapter we try to develop an alternative theory that accounts for relative gradable adjectives, both in their positive and comparative forms. The goal is to define the meaning of the comparative over the meaning of the positive form of the adjective. That means, first we determine the meaning of 'tall', then we define the meaning of 'taller than' by using comparison classes and constraints on the behaviour of the adjective functions in comparison classes. The meaning of 'tall to degree d', useful for measure phrases, is then defined on the meaning of the comparative form. To do that, measurement theory can be used. But I do not go into that problem. What I am mostly concerned with, in this research, is how to get the meaning of the comparative form from the positive form of adjectives. Such an attempt goes back to Kamp and Klein, and takes van Rooij's suggestions as reference point[11].

[11]See van Rooij [44], [43].

Chapter 4

A Model for Polar Adjectives

4.1 Aim

The aim of this research is to account for polar relative gradable adjectives such as *tall/short*, *big/small*, and so on.
Each gradable adjective comes with a polar counterpart. Some adjectives can form a polar pair with more than one adjective. For example, *short* can form a pair with *tall* and another with *long*. In the formalisation given below, *short* in the pair *short/tall* will be considered different than *short* in *short/long*. Moreover, there are one- and multi-dimensional adjectives. While *tall* is uniquely used to refer to the distance from the top to the bottom of an object, *clever* can be used to refer to some feature of cleverness that an individual endorses. Here a model for one-dimensional adjectives, and not for multi-dimensional ones, is proposed.
My aim is to give a model to account for our use of vague adjectives. When English native speakers have to judge on non-borderline cases, the use of adjectives such as *tall/short* is not problematic. For example, consider the set of *men* and a subset of it containing three individuals, John, Bill and Marc. John is 190 cm tall, Bill is 188 cm, Marc 165 cm. Speakers that see the three men do not know their precise height, but can observe that John and Bill do not relevantly differ in height, as well as they can detect a big difference between John-Bill and Marc. So, if the agents have to describe the height of John, Bill and Marc, they will naturally say that John and Bill are tall, Marc is short. The natural intuition seems to be this: when there is not a big difference between two objects with respect to some property represented by an adjective, we can appropriately attribute the same adjective to both the objects, but if there is a relevant difference

between them, then we describe them using a pair of polar adjectives. In the first case if we have a Soritical series we can get into trouble, while in the second case we do not have difficulty to use gradable adjectives. The intuition is that observations made for the unproblematic case cast light on the problematic ones. This chapter presents a model that describes the computational operations underlying speakers' decision about applying gradable adjectives to both problematic and unproblematic cases.

In this chapter, after some initial theoretical observations, I define a language $\mathcal{L}$; then a model is built up and applied to relative gradable adjectives. Adjectives are taken as primitive choice functions and the comparative relation is defined over the positive form of adjectives. Two ways of ordering elements in the domain are considered: weak and semi-orders.

4.2 Theoretical Background

In our account, vagueness is considered to be a feature of some expressions of natural language and is related to context-dependency and granularity.

As we have seen, *context-dependency* is a feature of relative gradable adjectives. For instance, the predicate *tall* applied to the context of human beings has a different extension than the same predicate applied to the domain of equatorial trees. However, also within the domain of human beings there might be sub-contexts that influence the interpretation of the predicate. For example, in the context where we consider only Dutch women the predicate has a different extension than in the context where we consider Japanese women. Moreover, suppose an individual named Bill is four years old and 130 cm tall. We can say that he is tall as a child, but short as a human being. In fact, if we compare Bill with the individuals of the set of all human beings, his height turns out to be below the height-average. However, considered as a child, he is quite tall. So, the extensions of vague predicates depend on the valuations that are made each time in a specific comparison class called (or natural kind, following Fara [12]). But within a natural kind, like the class of *children*, we can be interested in a more restricted context. For example, the set of children of the first grade in some primary school in Amsterdam.

Nevertheless, context-dependency is not sufficient for explaining the vagueness of predicates. Given a context, that is, a set of individuals, and

a vague predicate P, there are several ways to consider the differences between the individuals in the context. We can look at them with different standards of precision. In fact, the grain size varies from context to context and the grain size we choose often depends on our interest or our actual purpose[1]. As the grain size changes, we may cover different things under the same label or split meanings in a more refined way. This phenomenon is called *granularity*. Different levels of granularity can be thought of as different standards of precision. Let us try to understand what that means. Take the pair of polar adjectives *tall* and *short* and the example about John, Bill and Marc stated in the previous section. Formally, we have a set $o = \{j, b, m\}$, with j, b, m standing for John, Bill, Marc, respectively. If we look at o from a very coarse point of view, no difference among its elements is detected: Bill, John and Marc are equally tall. A coarse point of view can be given, for example, by a distant point of observation or by some specific purpose (for example, if we have to enlist men shorter than 160 cm, we do not discriminate between the men we have to rule out: they are all equally tall as they equally exceed the cut-off point). From a less coarse point of view, that is, using a finer grain size to discriminate differences, we might say that John and Bill are equally tall, while Marc is short. We can then establish a comparative relation: John (as well as Bill) is taller than Marc. However, with an even finer grain size we can perfectly distinguish the height of all the three men and say that John is taller than Bill and Marc, and Bill taller than Marc. Now, the same ordering between the elements is provided by two models that differ in the extension of *tall* and *short*. According to model 1, both John and Bill are tall and Marc short, according to model 2, only John is tall and both Bill and Marc short. But our intuition is that only model 1 is correct. We want then to find some constraints to rule out models like 2 that do not respect our intuitions.
The phenomenon of vagueness seems then to be captured (also) by the idea of granularity: some words are vague because the degree of specification of their meaning varies.

4.3 Language and Interpretation

To give a model to account for gradable adjectives, let us fix first a language and then an interpretation for it.

[1]See Hobbs [13] and Mani [27].

4.3.1 Language

Let $\mathcal{L}$ be a formal language through which we can represent English expressions. $\mathcal{L}$ consists of:

- individual constant symbols (that represent proper names: *John, Mary, Sue,...*): j, m, s, ...
- individual variable symbols: x, y, z...
- monadic predicates (representing common nouns like *pig*, *man*, *winner*): A, B, C, ...
- functions (representing adjectives): $P, Q, R, ...$, standing for *tall*, *big*, *fat*, ... We will later define the polar counterparts of such functions: $\overline{P}, \overline{Q}, \overline{R}, ...$, standing for *short*, *small*, *thin*, ...
- usual logical connectives with identity, quantifiers.

4.3.2 Interpretation of $\mathcal{L}$

We have a domain $\mathcal{D}$ of objects.
Monadic predicates select some objects of $\mathcal{D}$. Their extensions are called *natural sets*. I am assuming here that it is always possible to give a precise extension for each monadic predicate. For example, I do not consider the problem concerning the extension of *child*, that is known as being a vague predicate too. I ignore this kind of problems now because I am concerned with polar vague adjectives. So, I assume the domains to which such adjectives apply to be precise.
$I(A)$ is a natural set, for A a monadic predicate. Call it s. Let NS be the set of natural sets s.

First, the comparison classes within which the individuals are compared are selected; then, the values of the functions corresponding to adjectives within each comparison class are obtained. So, first we have the comparison classes (natural sets), then we can apply the functions to the domain of each of those comparison classes. In this way, we can predict that an individual x is P in some comparison class, $\overline{P}$ in other comparison classes: for example, when Bill is said to be tall as a child, but short as a man, it means that in the natural set of children, Bill is within the extension of *tall*, while in the set of men, he is not.
But, as we saw, we need more than comparison classes in order to evaluate

the assignment of an adjective to an individual. We need contexts. A context o is a subset of a natural set.

Definition 2 *Let O_s be the set of all contexts in some natural set s: $O_s = \wp(s)$.*

4.4 Context Structures and Weak Orderings

In this section contexts structures are defined and, then, given some cross-contextual constraints, the comparative relation is also defined. The ordering relation between the elements of each natural set obtained will turn out to be a weak ordering. More constraints will be given in order to remove the context structures that do not make a right attribution of gradable adjectives to elements of contexts.

As Luce [26] himself highlights, the non-transitivity of indifference relations reflects human inability to discriminate with precision among things that do not differ much one from the other. Luce's consideration on this point perfectly fits our problem with vague predicates. We cannot make precise distinctions between two objects with respect to some observable property. That is why we get into trouble with Sorites series. Nevertheless, if we have some more precise standard of precision or a better way of measurement we can detect more differences between the elements we consider. That is nothing else than the concept of granularity, as we saw above. According to different standards of precision, we can have different models that give rise to different orderings of the objects of our domain. Let us see how this works in a more detailed way.

Let $M = \langle \mathcal{D}, I_{NS}, P \rangle$ be a fixed model, or context structure. $\mathcal{D}$ is the whole domain, I_{NS} the set of natural sets, P a choice function that maps the individuals of context o to $P(o)$.
Then, we define the polar counterpart $\overline{P}$ of P as a function that applies to the elements in a context to which P does not apply:

Definition 3 $\overline{P}(o) = \{x \in o : x \notin P(o)\}$.

Here we consider P and $\overline{P}$ as contradictories[2].

[2] A further development can consist in improving the model in order to treat P and $\overline{P}$ as

I want to make possible that the meaning of an adjective changes over contexts: some object that has a property in a context might not have that property in an enlarged domain. This leads to the construction of a comparative relation[3].

How to account for this cross-contextual change of meaning?
First of all, assume that P is a choice function that takes elements form each finite set of options o: $P(o)$ is a subset of o, that is, the subset that contains the elements that have the property represented by P. Some constraints can be put on such a function in order to make it behave in a different way in each set and produce an ordering relation.

4.4.1 Van Benthem's Constraints

Consider the cross-contextual constraints that van Benthem introduced in [42]. They are based on the concept of difference pair (DP):

Definition 4 *Two elements x, y form a difference pair in a context o iff x is in the extension of P and y in the extension of $\overline{P}$, that is:*
$\langle x, y\rangle \in DP(o)$ *iff*$_{def}$ $x \in P(o)$ *and* $y \in \overline{P}(o)$.

The first constraint is the so called **Upward Difference** (UD):

(UD) Let $\langle e, e'\rangle$ be a difference pair in a context o. In each context o' containing o, there exist different pairs.

Put otherwise, if in a context o one element is tall, another short, (UD) makes sure that all the supersets of o will contain at least one element that is tall and one that is short. Those elements are not necessarily e, e'. Consider the following example: o contains an individual e that is 200 cm tall and an individual e' that is 180 cm tall, and other individuals whose height is somewhere between 180 and 200 cm. We can state that e, e' form a difference pair in o. But now, take a context o' that have all the same

contraries, and not contradictories, allowing in this way a function that maps individuals also to the set of elements that are neither P nor $\overline{P}$.

[3] I will consider here only direct comparatives, not indirect ones, as in the example: "Compared to Mary, John is tall". Different considerations are needed to account for the semantics of this kind of comparison. To give an example of a treatment of direct vs. indirect comparison, see Kennedy [17].

individuals as o, plus an individual e'' that is 150 cm tall. Now, e, e' will not any longer be a difference pair in o', even though some other difference pairs can be detected (for instance, $\langle e, e'' \rangle \in DP(o')$). Moreover, the converse of (UD) does not hold: a subset of o does not necessary contain a difference pair; for instance, it may contain only elements that are all equally tall.

Van Benthem proposes other two constraints on the behaviour of the elements of a difference pair in o contained in different contexts than o.
The second is **No Reversal**:

(NR) Let $\langle e, e' \rangle$ be a difference pair in a context o. There is no context o' such that $\langle e', e \rangle \in DP(o')$.

If in a context o one element e is tall and another e' short, in any other context o' the reverse cannot be the case. Maybe both e and e' are tall, or short, but it can never be the case that e' is tall and e short.

The third constraint is **Downward Difference**:

(DD) Let $\langle e, e' \rangle$ be a difference pair in a context o. In each context o' contained in o which includes e, e', there exist some difference pairs.

If e is tall and e' short in a large context o, in a smaller context o' containing e, e' there will be difference pairs too.

4.4.2 Comparative Relation

Given the three constraints (NR), (UD), (DD), we can define first the comparative relation $>_P$ (to read: "more P than"):

Definition 5 $x >_P y$ *iff* $x \in P(\{x, y\}) \wedge y \notin P(\{x, y\})$.

The relation $>$ is defined with respect to a predicate P and gives rise to a *weak order*. A weak order is a structure $\langle I, R \rangle$ with R a binary relation on I that is irreflexive, transitive and almost-connected:
(IR) $\forall x : \neg R(x, x)$
(TR) $\forall x, y, z : (R(x, y) \wedge R(y, z)) \rightarrow R(x, z)$
(AC) $\forall x, y, z : R(x, y) \rightarrow (R(x, z) \vee R(z, y))$

Define now the relations 'being as P as' (i.e. the similarity relation $\sim_P$) and 'being at least as P as' ($\geq_P$) as follows, respectively:

Definition 6 $x \sim_P y$ *iff*$_{def}$ *it is not the case that* $x >_P y$ *nor* $y >_P x$.

Definition 7 $x \geq_P y$ *iff*$_{def}$ $x >_P y$ *or* $x \sim_P y$.

4.4.3 On the Set of Context Structures

From what has been discussed in the previous section, we can make another consideration. The conditions for comparatives do not uniquely determine the behaviour of property P across comparative classes. Let M be a context structure of type $\langle \mathcal{D}, I_{NS}, P \rangle$. Different context structures can give rise to different $>_P$ orderings for the same set of contexts. That means that some context structures detect more differences between the elements in the contexts than other context structures; this fact corresponds to the intuition of granularity. Different context structures can have the same level of granularity.

Take a context $o \in O_s$. Since any context structure provides us with an equivalence relation $\sim_P$, equivalence classes partitioning the context are obtained. Equivalence classes are groups of objects, that according to that specific context structure turn out to be indistinguishable:

Definition 8 *Let* $a \in o$. *Define the equivalence class of* e *under* $\sim_P$ *as follows:*
$[e]_{\sim_P} =_{def} \{x \in o : x \sim_P e\}$.

Now we can define also a comparative relation $>^*_P$ between equivalence classes, as follows:

Definition 9 $[x]_{\sim_P} >^*_P [y]_{\sim_P}$ *iff*$_{def}$ $\exists x \in [x]_{\sim_P} \exists y \in [y]_{\sim_P} (x >_P y)$.

Different context structures can give rise to different partitions, and therefore to different ordering between objects. Consider the following example.

Example 1. Let o be a context containing with three elements ($o = \{a, b, c\}$) and the predicate T. Consider the possible orderings we can have on o:

- $a \sim_T b \sim_T c$. The context structure that models such an ordering gives only one partition, $[a]_{\sim_T}$. This ordering is the coarsest. There is only one equivalence class, all the elements are considered equal with respect to P, so we cannot have a distinction between objects that are P and objects that are $\overline{P}$. We do not want this kind of ordering, because it does not help us to solve the soritical problem.

- $a >_T b >_T c$. This is the finest ordering we can have. The context structure that models such an ordering gives three partitions: $[a]_{\sim_T}$, $[b]_{\sim_T}$, $[c]_{\sim_T}$.

- Between the coarsest and the finest orderings, there can be a third ordering. Either $a \sim_T b >_T c$ or $a >_T b \sim_T c$. If a context structure gives rise to the first ordering, it gives two partitions: $[a]_{\sim_T}$ and $[c]_{\sim_T}$. If the second ordering is modelled, then we have again two partitions, but different ones: $[a]_{\sim_T}$ and $[b]_{\sim_T}$.

We can partially order the context structures from the coarsest to the finest with respect to any context $o \in O_s$. First, we need to define the relation $\geq^c$ between cardinality of sets:

Definition 10 *For all the equivalence classes $[x]$ in M and all the equivalence classes $[y]$ in M', $|[y]^{M'}| \geq^c |[x]^M|$ iff the number of $[y]^{M'}$ is greater than or equal to the number of $[x]^M$ in o.*

Now we can define the following relation (*finer than*) between context structures:

Definition 11 *M' is finer than M iff the number of equivalence classes of M' is larger than the number of equivalence classes of M, that is:*
$M' \leq^{**} M$ iff $|[y]^{M'}| \geq^c |[x]^M|$.

However, we are more interested in a refinement relation defined as follows:

Definition 12 *M' is a refinement of M iff each equivalence classes in M' is a (not necessarily proper) subset of an equivalence class in M, that is:*
$M' \sqsubseteq M$, iff $\forall [y]^{M'} \exists [x]^M : [y]^{M'} \subseteq [x]^M$.

From definitions 8 and 9 follows:

$\forall M(M' \sqsubseteq M \rightarrow M' \leq^{**} M)$.

A natural constraint on the refinement ordering between context structures is that the ordering between individuals given in the most fine-grained context structure cannot be reversed in the coarse-grained ones. Formally:

(RfM) $\forall x, y, z \in s$: if $M' \models x \geq_P y \wedge y \geq_P z$ and $M \models x \sim_P z$, then $M \models x \sim_P y \wedge y \sim_P z$.

Consider again Example 1. We can claim that a context structure that models $a \sim_T b >_T c$ and one that models $a >_T b \sim_T c$ are both finer than a context structure that models $a \sim_T b \sim_T c$, but coarser than one that models $a >_T b >_T c$. However, if we want to consider a refinement ordering between all those context structures, we start with the coarsest one, that is, one that models $a \sim_T b \sim_T c$. Then, we take either a context structure that models $a \sim_T b >_T c$ or one that models $a >_T b \sim_T c$ (they are alternative and incompatible; none is a refinement of the other). Finally, the finest context structure for o is one that models $a >_T b >_T c$ and it is the refinement of either the context structure that models $a \sim_T b >_T c$ or the one that models $a >_T b \sim_T c$. Put otherwise, we can also say that once we have a coarse context structure, we can refine it and get several granular levels, that is, levels of refinement.

An important observation has to be made at this point. For contexts with more than two equivalence classes, even if two context structures give rise to the same $>_P$ ordering, they might have different choice functions of type P and $\overline{P}$. In general, for contexts with more than two equivalence classes there are at least two context structures that give rise to the same ordering.

Consider some examples of context structures for the context $o_1 = \{a, b, e, f\}$ that order o_1 as follows: $a >_P b >_P e >_P f$ and that agree on the behaviour of P on the pairs $\{a, b\}$, $\{b, e\}$, $\{e, f\}$. These are:
$M_1 \models P(o_1) = \{a\}, \overline{P}(o_1) = \{b, e, f\}$
$M_2 \models P(o_1) = \{a, b\}, \overline{P}(o_1) = \{e, f\}$
$M_3 \models P(o_1) = \{a, b, e\}, \overline{P}(o_1) = \{f\}$

But we do not want to have all these models. We want to rule out some of them, exactly those that predict wrongly which elements are P and which are $\overline{P}$. To do that, we suggest this: given a context $o \in O_s$, consider also the context o^s, defined as follows:

Definition 13 $o^s =_{def} \{z \in s \mid \exists x, y \in o : x \geq_P z \geq_P y\}$.

That is, any context o contains some elements of s. Given a set of context structures that give rise to the same ordering for all contexts in O_s, when we consider some context $o \in O_s$ we consider also the elements in the domain of s that are 'in-between' the elements of o given the order raised by the set of context structures considered.
Here a problem arises: how can we guarantee that there are enough elements 'in-between'? I try to explain this problem in the following lines. The intuition is the following. When we consider contexts, we look at real objects. Namely, when we use gradable adjectives we want to judge on some situation in the world. However, to correctly use gradable adjectives to describe objects, we can think that we add possible objects that are 'in-between' the real objects according to the comparative relation. That means, if John and Bill are men that relevantly differ along their height, we say that John is tall and Bill short because there can be other men, whose height differs and is less than John's but more than Bill's. This intuition goes together with the fact that all vague relative adjectives suffer for the Sorites Paradox: the crucial point of Sorites paradox is that we have a series of objects, such that there are small differences between every two objects that are contiguous in the series. We want to model vague relative adjectives that give rise to Sorites series: so, we need to assume that we *can* have a domain of individuals that are "equally distributed" with respect to a property, and such that each element of any $s \in NS$ is indistinguishable from at least two other from an observational point of view. Put otherwise, what we want is each set to have possible objects that form a Sorites series, and each real object to correspond to one of the possible objects of the domain. So, we need one restriction concerning the domain of individuals of each natural set.

Here I state the condition that needs to be imposed to each natural set. Take a natural set s: each element of s is observably indistinguishable from at least two other. It might happen only for two elements that each of them is indistinguishable from another element[4]. So, every natural set s comes with an indistinguishable relation $\approx_{s_P}$ with respect to a property P,

[4] These last two elements are predicted to be the minimal and the maximal element of the set of the individuals when ordered.

such that:

(**SC**) $\exists u \exists v(u \neq v \wedge \forall x((x \neq u \wedge x \neq v) \rightarrow \exists y \exists z(x \approx_{s_P} y \wedge x \approx_{s_P} z \wedge \neg(y \approx_{s_P} z))) \wedge \exists z(z \neq u \wedge u \approx_{s_P} z) \wedge \exists z(z \neq v \wedge v \approx_{s_P} z))$.

The point is: with the primitive relation $\approx_{s_P}$ we only constrain the natural sets in order to contain elements that are indistinguishable from at least two other (with two exceptions). We do not impose any order to the sets of possible objects. The ordering between elements is given by the choice function P in each context structure.
An observation about (SC): it commits to the existence of at least two elements for each s.

At this point, we have to redefine context structures. Let $M = \langle \mathcal{D}, I_{NS}, \approx_{s_P}, P \rangle$ be a fixed model, or context structure. $\mathcal{D}$ is the whole domain, I_{NS} the set of natural sets, $\approx_{s_P}$ the indistinguishability relation with respect to a $s \in NS$ and a property P, P a choice function that maps the individuals of context o into $P(o)$.

There is an important observation to make at this point. It concerns the difference between $\sim_P$ and $\approx_P$. They are not the same relation. While $\approx_P$ is given primitively, to describe what each natural set looks like according to our observation, $\sim_P$ is defined on the behaviour of P and $\overline{P}$. Moreover, $\sim_P$ is an equivalence relation, but $\approx_P$ is not: it is reflexive, symmetric, but not necessarily transitive. While $\approx_P$ is used only to assure that we have enough individuals, that is, they form a soritical series, $\sim_P$ is defined on the choice function P within a specific context structure. Each context structure is a way to look at natural sets and their contexts, and to partition each context. That means, according to each context structure we can detect some differences, and make finer partitions also in soritical contexts, as the the grain size becomes finer.

Now, we consider the contexts of type o^s as defined above as elements of O_s:

(Q) $\forall o : o^s \in O_s$.

It has to be noticed that o^s are Sorites series; namely, they contain a sequence of elements that satisfy (SC). Since we have (Q), any context structure $M = \langle \mathcal{D}, I_{NS}, \approx_{s_P}, P \rangle$ gives an ordering to *all* the contexts

$o \in O_s$. So, each M that has a function P that satisfies van Benthem's constraints, gives rise to a weak ordering also in contexts of type o^s. We will consider later, in section 4.5, what happens if we get rid of (Q).

Since with weak orderings we partition all the contexts into equivalence classes, we will consider the elements of contexts as part of some equivalence classes; we consider also the contexts of type o^s as partitioned into equivalence classes.
In such a way, we get that all natural sets are weakly ordered. In fact, if each context structure orders the elements of the subsets (i.e. the contexts) of a natural set, the whole natural set gets an ordering of the same type (weak) within the same context structure. That means, for each $s \in I_{NS}$, we have structures of type $\langle s, > \rangle$.

Assume now that the extensions of each predicate P and of their complements $\overline{P}$ in o^s, for all $o \in O_s$, have to be as follows:

$$|\{[x]_{\sim_P} \in P(o^s)\}| = |\{[x]_{\sim_P} \in \overline{P}(o^s)\}| \pm 1.$$

Then, given a set of context structures that give rise to the same ordering and taking a context $o \in O_s$, we accept the context structures that make the following formula true:

(R) $\forall x \in o : x \in P(o)$ iff $x \in P(o^s)$.

(R) is the constraint that restricts the set of context structures.
Consider again the example mentioned above. Take a predicate P and the context $o_1 = \{a, b, e, f\}$. Let o_1^s be $\{a, b, c, d, e, f\}$. Since I take into consideration only P in the example, I will not mention P in the index of the symbol '$>$' and I will write $[x]$ for the equivalence class $[x]_{\sim_P}$.
Consider the set of models that give rise to the ordering $a > b > e > f$ for o_1. Assume that the same set gives rise to the ordering $a > b > c > d > e > f$. Then, we have: $P(o_1^s) = \{[a], [b], [c]\}$, $\overline{P}(o_1^s) = \{[d], [e], [f]\}$.
Consider the models of our set:
$M_1 \models P(o_1) = \{[a]\}, \overline{P}(o_1) = \{[b], [e], [f]\}$
$M_2 \models P(o_1) = \{[a], [b]\}, \overline{P}(o_1) = \{[e], [f]\}$
$M_3 \models P(o_1) = \{[a], [b], [e]\}, \overline{P}(o_1) = \{[f]\}$

Applying (R) as a restriction on $M_1, ...M_3$, we get only M_2 as acceptable model.

M_1 is ruled out because the equivalence class $[b]$ is mapped into the set $\overline{P}(o_1)$, while our constraint (R) wants it to be mapped into $P(o_1)$. Informally speaking, the extension of $P(o_1)$ given by the model M_1 contains 'too few' equivalence classes.
M_3 is ruled out for the opposite reason: the extension of $P(o_1)$ contains an equivalence class more than what is accepted. According to (R), $[e]$ has to be in the extension of $\overline{P}(o_1)$.

Consider what happens when we have the case when the number of context structures is even.
Let context o be $\{a, c, d, g\}$ and let o^s be $\{a, b, c, d, e, f, g\}$. Consider the set of context structures that model $a > c \sim d > g$ with respect to P. Then, we have a set of two context structures that model $a > c \sim d > g$ and $a \sim b > c \sim d > e \sim f > g$:
$M_1 : P(o) = \{a\} \Rightarrow P(o) = \{[a]_\sim\}$;
$M_2 : P(o) = \{a, c, d\} \Rightarrow P(o) = \{[a]_\sim, [c]_\sim\}$.

By the restriction on the extension of $P(o^s)$ and $\overline{P}(o^s)$, we have:
$P(o^s) = \{[a]_\sim, [c]_\sim\}$
$\overline{P}(o^s) = \{[d]_\sim, [g]_\sim\}$.

So, only M_2 is accepted. Namely, M_1 does not consider the elements of $[c]_\sim$ as part of the extension of $P(o)$, while they are in the extension of $P(o^s)$.

If the number of equivalence classes in o^s is odd, we have an equivalence class that can be either P or $\overline{P}$. No constraint fixes its meaning. So, both M_1 and M_2 are acceptable. That an agent considers $[c]_\sim$ as P or $\overline{P}$ might depend on her conversational purposes, that is, on some pragmatic, and not semantic, factors.

What the present model is now able to do is to predict a correct assignment of the adjective to those elements in a context that are quite different with respect to some property.

Observation

When we have a context o^s with an even number of equivalence classes, the following holds:

$|\{[x]_{\sim_P} \in P(o^s)\}| = |\{[x]_{\sim_P}\overline{P}(o^s)\}|.$

The addition of ± 1 concerns contexts with an odd number of equivalence classes. If o^s is such a case, we can say there is one equivalence class that can be considered either part of the extension of P or of $\overline{P}$. The elements that are in that equivalence class are called *tolerant* elements. We can perhaps think that the tolerant elements are neither P nor $\overline{P}$.

4.4.4 *Tall* and *Short*

Consider now the application of our model to the pair of vague adjectives *tall* and *short*. Let T stand for *tall*, $\overline{T}$ for *short*.
Consider again the example given in the beginning of the chapter. Take the natural set of *men* and a context with three individuals, John, Bill and Marc. John is 190 cm tall, Bill is 188 cm, Marc 165 cm. Speakers that see the three men, do not know their precise heights. However, if they have to describe the height of John, Bill and Marc, they will naturally say that John and Bill are tall, Marc is short. Our model tries to explain why we have such a natural intuition.

Formally, we have the context $o = \{j, b, m\}$, with j, b, m standing for John, Bill, Marc, respectively. Different context structures give rise to different orderings between the elements in o; for example:
$j \sim_T b \sim_T m$
$j \sim_T b >_T m$
$j >_T b >_T m$.

For each set of context structures, we also consider the ordering that it gives to the context o^s. This is obtained by filling o in with all the other individuals of the natural set *men* that are in-between John, Bill and Marc according to the ordering given in each context structure.

Now, we do not consider the coarsest granular models, i.e. the context structures that give rise to only one equivalence class.

Consider then the ordering $j \sim_T b >_T m$. All the context structures that give this ordering will give the following extensions for T and $\overline{T}$ in o:
$T(o) = \{j, b\}$, $\overline{T}(o) = \{m\}$.
Since they detect a difference between j, b and m in o, that is, m is in a different equivalence class than j and b, in o^s also m will be in a different

equivalence class than j and b.

Consider now $j >_T b >_T m$. We can have some context structures that give the following extensions for T and $\overline{T}$ in o:
$T(o) = \{j, b\}$, $\overline{T}(o) = \{m\}$.
But for the same ordering, we can have also other context structures that give the following extensions:
$T(o) = \{j\}$, $\overline{T}(o) = \{b, m\}$.
The latter kind of context structures will be ruled out by our rule (R). In context o^s, we have many individuals, and also many equivalence classes, between b and m. Considering our constraint for the extension of T and $\overline{T}$ in o^s, we have that $T(o^s) = \{j, b\}$, $\overline{T}(o^s) = \{m\}$. So, Bill has to be considered tall. And that is exactly what we intuitively do.

4.5 Context Structures and Semi-Orders

In this section, a change for the model sketched in the previous section is proposed: instead of using a weak ordering relation, it is suggested to make use of semi-orders.

An objection that might be raised against the proposal in the previous section concerns the assumption that O_s contains all the subsets of s. Under the definition for O_s given above we get, as we have seen, that all the contexts containing a soritical series are also ordered by the choice function P within each context structure. The ordering for a context of type o^s, for example, is different depending on the context structure considered. This result might not be accepted if we want to preserve that in a soritical context no assignment of P or $\overline{P}$ to any element of the context is admissible. According to our intuitions, it is very difficult and might seem unnatural to assign P or $\overline{P}$ to individuals belonging to a soritical series in a definite and clear way, as we would be obliged to admit that two objects that taken by themselves are indistinguishable are actually one P, the other $\overline{P}$. We can now see that if we can get rid of the assumption

$$\forall o \in O_s : o^s \in O_s,$$

we obtain that soritical contexts cannot be ordered by the choice function P of contexts structures. They can still be ordered, but the order we get is a semi-order, not a weak order. Let us look at it in more detail[5].

We take $x >_P y$ as irreflexive and transitive, but not necessary almost connected, and $\sim_P$ as reflexive and symmetric, but not necessarily transitive. A set ordered according to the relation $>_P$ gives rise to a semi-order. A *semi-order* is a structure $\langle I, R\rangle$ that is irreflexive (IR), semitransitive (STr) and satisfies the interval-order condition (IO):
(IR) $\forall x : \neg R(x, x)$
(STr) $\forall x, y, z, v : (R(x, y) \wedge R(y, z)) \rightarrow (R(x, v) \vee R(v, z))$
(IO) $\forall x, y, v, w : (R(x, y) \wedge R(v, w)) \rightarrow (R(x, w) \vee R(v, y))$.

Assume now that not all the subsets of s containing at least two elements are proper contexts. In general, as we have seen before, the contexts in which all the elements are indistinguishable with respect to P are not relevant: since there is no distinction to be detected, we do not get any information about the distinction between elements that are P and elements that are $\overline{P}$. This kind of contexts are not proper ones.

If we assume that not all finite subsets of the domain are proper contexts and the subsets to rule out are the ones containing indistinguishable individuals with respect to some property, almost-connectedness of the comparative relation does not hold anymore. Loosing almost-connectedness, though, we loose also transitivity. We can add some closure conditions to get back transitivity.

The strategy here is to start with proper contexts containing only two elements. Then, put some constraints in order to build larger appropriate contexts.
First, we consider the subsets that contain only two elements for which the following holds:

(U) $\emptyset \neq P(o) \neq \emptyset$.

By (U) we want to capture the idea that a context $\{x, y\} \in O_s$, with O_s the set of proper contexts, iff the elements x, y are not indistinguishable with respect to P, that is, intuitively, if there is a gap between the P- and $\overline{P}$-elements in context $\{x, y\}$.

[5] I will mainly refer to van Rooij [43].

Now, the 'new' set of contexts O_s can be closed under the following two conditions (to get appropriate contexts, with cardinality bigger than 2):
(OR1) $\forall o' \in O_s : card(o') = 2 \rightarrow \exists x \in (o' - o) : o \cup \{x\} \in O_s$
(OR2) $\forall x, y, z, v \in s$: if $\{x, y\} \in O_s, \{y, z\} \in O_s, \{x, z\} \in O_s$, then $\{x, y, z, v\} \in O_s$.

These closure conditions are used to make it sure that (IR), (STr) and (IO) are met and to generate sets without Sorites series.

(OR1) says that you can always add one element x from any set o' to o. So you can have contexts with more than two objects. (OR1) guarantees that the comparative satisfies (IO). (OR2) assures that Semi Transitivity (STr) holds.

We can define the comparative and the similarity relations as we have already done:

Definition 14 $x >_P y$ *iff*$_{def}$ $\{x, y\} \in O_s$ *and* $x \in P(\{x, y\})$ *and* $y \in \overline{P}(\{x, y\})$.

Definition 15 $x \sim y$ *iff*$_{def}$ $x \not>_P y$ *and* $y \not>_P x$.

From those definitions and (U) it follows also:

$x \sim_P y$ iff $\{x, y\} \notin O_s$.

Theorem 1 *Any context structure* $M = \langle \mathcal{D}, I_{NS}, \approx_{s_P}, P \rangle$, *where* P *obeys axioms (NR), (DD), (UD) and (U), and where for all* $s \in NS$ *the set* O_s *is closed under (OR1), (OR2), gives rise to a semi-order* $\langle I_{NS}, > \rangle$, *if we define* $x >_P y$ *as* $x \in P(\{x, y\})$ *and* $y \in \overline{P}(\{x, y\})$.

Proof The proof of theorem 1 consists of two parts[6]:

1. To prove: (OR1) guarantees interval-order condition (IO), that is: we assume $\forall o' \in O_s : card(o') = 2 \rightarrow \exists x \in (o' - o) : o \cup \{x\} \in O_s$ and derive $\forall x, y, v, w : (x > y \wedge v > w) \rightarrow (x > w \vee v > y)$.
   Assume (OR1) and $x > y$, $v > w$. We have two cases to consider by

[6] I am adapting the proof provided by van Rooij in [43], pp. 8-9.

(OR1): $\{x, y, v\} \in O_s$ and $\{x, y, w\} \in O_s$.
If $x > y$ and $\{x, y, v\} \in O_s$, then by (UD) there is a difference pair in $\{x, y, v\}$. By (DD) we can have either $v > y$ or $x > v$. $v > y$ is what we want. If $x > v$ is the case, together with the assumption $v > w$ we get $\{x, v, w\} \in O_s$. By (UD) there is a difference pair in $\{x, v, w\}$. Assumed $x > v$ and $v > w$, we have the following possibilities:

- $P(\{x, v, w\}) = \{x, v\}$, $\overline{P}(\{x, v, w\}) = \{w\}$; so, by (DD) $x > w$;
- $P(\{x, v, w\}) = \{x\}$, $\overline{P}(\{x, v, w\}) = \{v, w\}$; so, by (DD) $x > w$.

In both cases, $x > w$ holds.
Assume now $x > y$ and $\{x, y, w\} \in O_s$. By (UD) there is a difference pair in $\{x, y, w\}$. By (DD) we can have either $x > w$ or $w > y$. $x > w$ is what we want. If $x > w$ is the case, together with the assumption $v > w$ we get $\{v, w, y\} \in O_s$. Assumed $v > w$ and $w > y$, there are the following possibilities:

- $P(\{v, w, y\}) = \{v, w\}$, $\overline{P}(\{v, w, y\}) = \{y\}$; so, by (DD) $v > y$;
- $P(\{v, w, y\}) = \{v\}$, $\overline{P}(\{v, w, y\}) = \{w, y\}$; so, by (DD) $v > y$.

In both cases $v > y$ holds.

2. To prove: (OR2) guarantees (STr), that is: we assume (OR2) $\forall x, y, z, v \in s$: if $\{x, y\} \in O_s, \{y, z\} \in O_s, \{x, z\} \in O_s$, then $\{x, y, z, v\} \in O_s$ and derive $\forall x, y, z, v : (x > y \wedge y > z) \rightarrow (x > v \vee v > z)$.
Assume (OR2), $x > y$ and $y > z$. From irreflexivity and interval condition, transitivity follows. So, by transitivity $x > z$. It must then be the case that $\{x, y\} \in O_s$, $\{y, z\} \in O_s$ and $\{x, z\} \in O_s$. So, applying (OR2) we have $\{x, y, z, v\} \in O_s$. Since $x > y$ and $y > z$, we have by (NR) and (UD) the following possibilities:

    - $P(\{x, y, z, w\}) = \{x, y\}$, $\overline{P}(\{x, y, z, w\}) = \{z, v\}$; so, by (DD) $x > v$;
    - $P(\{x, y, z, w\}) = \{x\}$, $\overline{P}(\{x, y, z, w\}) = \{y, z, v\}$; so, by (DD) $x > v$;
    - $P(\{x, y, z, w\}) = \{x, y, v\}$, $\overline{P}(\{x, y, z, w\}) = \{z\}$; so, by (DD) $v > z$;
    - $P(\{x, y, z, w\}) = \{x, v\}$, $\overline{P}(\{x, y, z, w\}) = \{y, z\}$; so, by (DD) $v > z$.

So, in all cases, either $x > v$ or $v > z$ holds. ∎

Each natural set s is then ordered in such a way that it originates a semi-order structure. Any subset of it, such that its elements form a soritical series, is also semi-ordered. We have, then:

$\forall o : o^s \notin O_s$.

Now, we know from Luce [26] that any semi-order can induce a weak order. Let us consider his contribution.

If R is an arbitrary relation on a set s, an indifference relation J can always be defined as follows: for $a, b \in s$: sJb iff neither aRb nor bRa. Then, the relation $(>, \sim)$ induced on s by a given relation R, J on s is defined as follows: $a > b$ if either:
(i) aRb,
(ii) aJb and there exists $c \in S$ such that aJc and cRb, or
(iii) aJb and there exists $d \in S$ such that aRd and dRb.
If neither $a > b$ nor $b > a$, then $a \sim b$.

Theorem 2 *(R,J) is a semi-order if and only if R is transitive and* $(>, \sim)$ *is a weak order.*

For the proof I refer to Luce's paper, Luce [26], pp. 183-185.

Using Luce's theorem, we can claim that whenever we have a semi-order, we can define a relation $>$ (and consequently its symmetric counterpart $\sim$) that generates a weak order. Any semi-order we obtain with the assumption $O_s \neq \wp(s)$ can be mapped then to one induced weak order. In such a way, we can then apply again the machinery we described in the previous section that makes use of equivalence classes.

Chapter 5

Results and Remarks

In this final chapter I show how some interesting results can be proved in the model described in the previous chapter and, then, I state some general observations.

5.1 Further Results

In the model we described by making use of equivalence classes, we are able to prove the following interesting results (for all the relations $>$, $<$, $\leq$ or $\geq$, read them always as $>^*_P$ or $<^*_P$, $\leq^*_P$, $\geq^*_P$, respectively):

(A), (B), (C), (D), (E) hold for each context $o \in O_s$.

(A) $\forall x \in o$, if $|\{[y] \in o^s | [y] > [x]\}| <^c |\{[y] \in o^s | [y] < [x]\}|$, then $x \in P(o)$.

The intuitive reading of (A) is: for all the elements x in some context o such that in context o^s the number of equivalence classes that are 'more P' than the equivalence class $[x]$ is smaller than the number of equivalence classes that are 'less P' than $[x]$, x is in the set given by $P(o)$, that is, x has the property that P represents. Put otherwise, again; for all $x \in o$, if the cardinality of the set of equivalence classes that are 'more P' than $[x]$ is strictly lower than the cardinality of the set of equivalence classes that are 'less P' than $[x]$, then x is in $P(o)$.

(B) $\forall x \in o$, if $|\{[y] \in o^s | [y] < [x]\}| >^c |\{[y] \in o^s | [y] < [x]\}|$, then $x \in \overline{P}(o)$.

The intuitive reading of (B) is: for all the elements $x \in o$ such that in context o^s the number of equivalence classes that are 'more P' than the equivalence class $[x]$ is bigger than the number of equivalence classes that are 'less P' than $[x]$, x is in the set given by $P(o)$, that is, x has the property that P represents. Put otherwise: for all $x \in o$, if the cardinality of the set of equivalence classes that are 'more P' than $[x]$ is strictly greater than the cardinality of the set of equivalence classes that are 'less P' than $[x]$, then x is in $P(o)$.

(C) $\forall x \in o$, if $|\{[y] \in o^s | [y] < [x]\}| = |\{[y] \in o^s | [y] < [x]\}|$, then $x \in P(o)$ or $x \in \overline{P}(o)$.

The intuitive reading of (C) is: for all the elements x in o such that in context o^s the number of equivalence classes that are 'more P' than the equivalence class $[x]$ is equal to the number of equivalence classes that are 'less P' than $[x]$, x is in the set given by $P(o)$ or in the set given by $\overline{P}(o)$, that is, x has the property P or $\overline{P}$. In other terms: for some x and some context o, if the cardinality of the set of equivalence classes that are 'more P' than $[x]$ is equal to the cardinality of the set of equivalence classes that are 'less P' than $[x]$, then x is in $P(o)$ or in $\overline{P}(o)$.
(C) allows for tolerant equivalence classes: whenever the number of equivalence classes in o^s is odd, the equivalence class that is equally distant from the greatest and the lowest equivalence class in the comparative relation can be considered as P or $\overline{P}$ on the basis of speakers' intentions and purposes. Whether to choose P or $\overline{P}$ is not a syntactic nor a semantic matter, but only pragmatic.

(D) $\forall x \in o$, if $x \in P(o)$, then either $|\{[y] \in o^s | [y] < [x]\}| <^c |\{[y] \in o^s | [y] < [x]\}|$ or $|\{[y] \in o^s | [y] < [x]\}| = |\{[y] \in o^s | [y] < [x]\}|$.

The intuitive reading of (D) is: for all the elements $x \in o$, if x is P in context o, then in context o^s the number of equivalence classes that are 'more P' than the equivalence class $[x]$ is either smaller than or equal to the number of equivalence classes that are 'less P' than $[x]$. Put otherwise, for any $x \in o$, if $x \in P(o)$, then the cardinality of the set of equivalence classes that are 'more P' than $[x]$ is lower than or equal to the cardinality of the set of equivalence classes that are 'less P' than $[x]$.

(E) $\forall x \in o$, if $x \in \overline{P}(o)$ then either $|\{[y] \in o^s | [y] > [x]\}| >^c |\{[y] \in$

$o^s|[y] < [x]\}|$ or $|\{[y] \in o^s|[y] > [x]\}| = |\{[y] \in o^s|[y] < [x]\}|$.

The intuitive reading of (E) is: for all the elements $x \in o$, if x is P in context o, then in context o^s the number of equivalence classes that are 'more P' than the equivalence class $[x]$ is either greater than or equal to the number of equivalence classes that are 'less P' than $[x]$. In other terms, for each $x \in o$, if $x \in P(o)$, then the cardinality of the set of equivalence classes that are 'more P' than $[x]$ is greater than or equal to the cardinality of the set of equivalence classes that are 'less P' than $[x]$.

I provide only the proofs of (A) and (D). Since the proofs of (B) and (C) are similar to the one of (A), and the proof of (E) is similar to (D), I shall omit them.

5.1.1 Proof of (A)

(A) $\forall x \in o$, if $|\{[y] \in o^s|[y] > [x]\}| <^c |\{[y] \in o^s|[y] < [x]\}|$, then $x \in P(o)$.

Assume for some $x \in o$: $|\{[y] \in o^s|[y] > [x]\}| <^c |\{[y] \in o^s|[y] < [x]\}|$.

Distinguish two cases:

- $|\{[y] \in o^s\}|$ is even.
  So, $|\{[y] \in P(o^s)\}| = |\{[y] \in \overline{P}(o^s)\}|$.
  Assume: $x \notin P(o)$. So, $x \in \overline{P}(o)$, that means,$[x] \in \overline{P}(o)$. By (R), $[x] \in P(o^s)$, that is, $[x] \in \{[y] \in P(o^s)\}$.
  By the definition of $>$ for equivalence classes, and since $[x] \in \overline{P}(o^s)$, we have the two following facts:

  Fact 1 $\forall [y] \in o^s([y] \in P(o^s) \rightarrow [y] > [x])$

  Fact 2 $\forall [y] \in o^s([y] < [x] \rightarrow [y] \in \overline{P}(o^s))$.

  From Fact 1 we also have: $|\{[y] \in P(o^s)\}| \leq^c |\{[y] \in o^s : [y] > [x]\}|$.
  From Fact 2: $|\{[y] \in o^s : [y] < [x]\}| <^c |\{[y] \in \overline{P}(o^s)\}|$, because $[x] \in \overline{P}(o^s)$, so $|\{[y] \in o^s : [x] < [y]\}| <^c |\{[y] \in \overline{P}(o^s)\}|$ for, at least, one equivalence class, namely $[x]$.
  From these two consequences of fact 1 and 2 and the assumption $|\{[y] \in$

$o^s|[y] > [x]\}| <^c |\{[y] \in o^s|[y] < [x]\}|$, we obtain, by transitivity of $<^c$ and $\leq^c$: $|\{[y] \in P(o^s)\}| <^c |\{[y] \in \overline{P}(o^s)\}|$. We get a contradiction. Then: $x \in P(o)$. $\blacksquare$

- $|\{[y] \in o^s\}|$ is odd.
  Two cases are possible:
  - $|\{[y] \in P(o^s)\}| = |\{[y] \in \overline{P}(o^s)\}| - 1$. So, $|\{[y] \in P(o^s)\}| <^c |\{[y] \in \overline{P}(o^s)\}|$.
    Assume $x \notin P(o)$. So, $x \in \overline{P}(o)$, that means, $[x] \in \overline{P}(o)$. By (R), $[x] \in \overline{P}(o^s)$, that is, $[x] \in \{[y] \in \overline{P}(o^s)\}$.
    Fact 1 and 2, with their consequences, hold here too.
    In this case, if we have an equivalence class $[x]$ such that $\forall[y] \in \overline{P}(o^s) : [y] \leq [x]$, $[x]$ is a tolerant equivalence class, such that $|\{[y] \in o^s|[y] > [x]\}| = |\{[y] \in o^s|[y] < [x]\}|$. We do not consider this case, since it deals with (C).
    Instead, consider the equivalence classes $[x]$ such that: $\exists[y] \in \overline{P}(o^s) : [y] > [x]$.
    We have, then: $|\{[y] \in P(o^s)\}| <^c |\{[y] \in o^s : [y] > [x]\}|$ and $|\{[y] \in o^s : [y] < [x]\}| <^c |\{[y] \in \overline{P}(o^s)\}|$ for more than one equivalence class. Since $|\{[y] \in P(o^s)\}| <^c |\{[y] \in \overline{P}(o^s)\}|$ for only one equivalence class, we have $|\{[y] \in o^s : [y] < [x]\}| <^c |\{[y] \in P(o^s)\}|$. By transitivity, then: $|\{[y] \in o^s : [y] < [x]\}| <^c |\{[y] \in o^s : [y] > [x]\}|$. We get a contradiction. So, $x \in P(o^s)$. $\blacksquare$
  - $|\{[y] \in P(o^s)\}| = |\{[y] \in \overline{P}(o^s)\}| + 1$. So, $|\{[y] \in P(o^s)\}| >^c |\{[y] \in \overline{P}(o^s)\}|$.
    Assume $x \notin P(o)$. So, $x \in \overline{P}(o)$, that means, $[x] \in \overline{P}(o)$. By (R), $[x] \in P(o^s)$, that is, $[x] \in \{[y] \in P(o^s)\}$.
    Fact 1 and 2 hold here too. From $|\{[y] \in o^s : [x] < [y]\}| <^c |\{[y] \in \overline{P}(o^s)\}|$, $|\{[y] \in \overline{P}(o^s)\}| <^c |\{[y] \in P(o^s)\}|$ and $|\{[y] \in P(o^s)\}| \leq^c |\{[y] \in o^s : [y] > [x]\}|$ we get, by transitivity of $<^c$ and $\leq^c$: $|\{[y] \in o^s : [x] < [y]\}| <^c |\{[y] \in o^s : [y] > [x]\}|$, contradiction. Then: $x \in P(o)$. $\blacksquare$

5.1.2 Proof of (D)

(D) $\forall x \in o$, if $x \in P(o)$, then either $|\{[y] \in o^s|[y] < [x]\}| <^c |\{[y] \in o^s|[y] < [x]\}|$ or $|\{[y] \in o^s|[y] < [x]\}| = |\{[y] \in o^s|[y] < [x]\}|$.

Consider an arbitrary $x \in o$: $x \in P(o)$. So, $[x] \in P(o)$. By (R), $[x] \in P(o^s)$, that is, $[x] \in \{[y] \in P(o^s)\}$.
Distinguish two cases:

- $|\{[y] \in o^s\}|$ is even.
  So, we have: $|\{[y] \in P(o^s)\}| = |\{[y] \in \overline{P}(o^s)\}|$.
  By the definition of $>$ as a comparative relation between equivalence classes, and since $[x] \in P(o^s)$ the following facts and their consequences hold:

  Fact 3 $\forall [y] \in o^s([y] \in \overline{P}(o^s) \rightarrow [x] > [y])$

  Fact 4 $\forall [y] \in o^s([y] > [x] \rightarrow [y] \in P(o^s))$.

  Consequence of Fact 3: $|\{[y] \in \overline{P}(o^s)\}| \leq^c |\{[y] \in o^s : [y] < [x]\}|$.
  Consequence of Fact 4: $|\{[y] \in o^s : [y] > [x]\}| <^c |\{[y] \in P(o^s)\}|$, because $[x] \in P(o^s)$, so $|\{[y] \in o^s : [y] > [x]\}| <^c |\{[y] \in P(o^s)\}|$ for at least one equivalence class, $[x]$.

  Since $|\{[y] \in P(o^s)\}| = |\{[y] \in \overline{P}(o^s)\}|$, $|\{[y] \in o^s : [x] < [y]\}| <^c |\{[y] \in \overline{P}(o^s)\}|$, and since $|\{[y] \in \overline{P}(o^s)\}| \leq^c |\{[y] \in o^s : [y] < [x]\}|$, we obtain $|\{[y] \in o^s | [y] > [x]\}| <^c |\{[y] \in o^s | [y] < [x]\}|$. ∎

- $|\{[y] \in o^s\}|$ is odd.
  Two cases are possible:

  - $|\{[y] \in P(o^s)\}| = |\{[y] \in \overline{P}(o^s)\}| - 1$.
    So, $|\{[y] \in P(o^s)\}| <^c |\{[y] \in \overline{P}(o^s)\}|$.
    Also in this case, Fact 3 and 4 and their consequences hold.
    By transitivity of $<^c$ and $\leq^c$, from $|\{[y] \in P(o^s)\}| <^c |\{[y] \in \overline{P}(o^s)\}|$ and $|\{[y] \in \overline{P}(o^s)\}| \leq^c |\{[y] \in o^s : [y] < [x]\}|$ we infer $|\{[y] \in o^s : [y] > [x]\}| <^c |\{[y] \in o^s : [y] < [x]\}|$. ∎
  - $|\{[y] \in P(o^s)\}| = |\{[y] \in \overline{P}(o^s)\}| + 1$.
    So, $|\{[y] \in P(o^s)\}| >^c |\{[y] \in \overline{P}(o^s)\}|$.
    Also in this case, Fact 3 and 4 and their consequences hold.
    Consider the equivalence classes $[x]$ such that: $\exists [y] \in P(o^s) : [y] < [x]$.
    Now, we have that $|\{[y] \in \overline{P}(o^s)\}| <^c |\{[y] \in o^s : [y] < [x]\}|$ and

$|\{[y] \in o^s : [y] > [x]\}| <^c |\{[y] \in P(o^s)\}|$ for more than one equivalence class. So, $|\{[y] \in o^s : [y] > [x]\}| <^c |\{[y] \in \overline{P}(o^s)\}|$. By transitivity, then: $|\{[y] \in o^s : [y] > [x]\}| <^c |\{[y] \in o^s : [y] < [x]\}|$.
$\blacksquare$

If we have an equivalence class $[x]$ such that $\forall[y] \in P(o^s) : [y] \geq [x]$, that is, such that $[x]$ is less P than all the P-elements, $[x]$ is a tolerant equivalence class. In such a case, then, $\forall[y] \in o^s([y] < [x] \rightarrow [y] \in \overline{P}(o^s))$ holds. So, we also have: $|\{[y] \in o^s : [y] < [x]\}| = |\{[y] \in \overline{P}(o^s)\}|$. And also: $|\{[y] \in o^s : [y] > [x]\}| = |\{[y] \in P(o^s)\}| - 1$.
Now, $|\{[y] \in P(o^s)\}| - 1 = |\{[y] \in \overline{P}(o^s)\}|$, so: $|\{[y] \in o^s : [y] > [x]\}| = |\{[y] \in \overline{P}(o^s)\}|$. Since $|\{[y] \in o^s : [y] < [x]\}| = |\{[y] \in \overline{P}(o^s)\}|$, we obtain $|\{[y] \in o^s : [y] > [x]\}| = |\{[y] \in o^s : [y] < [x]\}|$.
$\blacksquare$

5.1.3 Theoretical Remark

The results we get with (A), (B), (C), (D), (E) show that, even if we start with a comparative relation as primitive, instead of the choice function P, we can have the same result with respect to the model for polar adjectives described in chapter 4.
There are some (in my opinion, strong) intuitions that might bring us to think that we always assign a polar adjective to an object after some sort of comparison between that object and some comparison class such as a context or a natural kind. Imagine there is only one object $o*$ in the universe: we cannot say if it is big or small, nor if tall or short, etc. We need to have at least another object to be able to properly attribute a property to $o*$. When children learn how to use polar adjectives, they need to see a comparative set among the elements of which some objects are, for example, big, and others small.
It seems that by the means of the primitive choice function P we arbitrarily say (or, we already know) what is big and what is small, and afterwards we build comparatives among the objects characterised as big and as small. But why do we not try to think about the comparative relation as primitive? First, we see how things are in the world and their relations. Then, we can distinguish big from small objects, and so on. It seems to me that an approach of vagueness and of the meaning of adjectives that starts from an analysis of comparatives might be closer to our intuitions. The results I proved show that, if someone wants

to assume such an approach, she can get the same model as I have proposed.

5.2 Problematic Points for Further Development

At this point we are going to underline some problematic points and set some questions.

5.2.1 On Infinity

The model developed and presented here is suitable for domains with a finite number of elements. But what happens to the domains with a countably infinite number of elements? Take the set of natural numbers $\mathbb{N}$ and consider the predicates *small* and *large*. The problem we have is that we cannot draw the line in the context $o{*}^{\mathbb{N}}$ when $|o *^{\mathbb{N}}| = |\mathbb{N}|$. But for all the other contexts $o \in O_{\mathbb{N}}$ we can always fill the correspondent context $o^{\mathbb{N}}$ in and distinguish small from large elements. However, a question might arise: is the inability to treat infinite contexts a real defect of the model? Maybe, such inability is not a mere weakness. To ask whether a certain number is large does not make sense, if the comparison class considered is the whole set $\mathbb{N}$. In normal conversation, we can use the expression “large (small) number” when we have a context with a finite number of elements. The use of gradable adjectives with infinite contexts does not seem to be natural.

5.2.2 On Polarity

Our model gives a plausible explanation why we are able to properly use vague polar adjectives. In our model, though, we do not try to account for a solution to the problem of vagueness, nor do we explain why polar adjectives such as ‘tall’, ‘short’, ‘big’, ‘small’, ... are vague. Instead, what we do is to make some assumptions about soritical series. We use two mathematical structures to deal with soritical domains, that is, to describe them: weak and semi-orders, together with the assumption

$$|\{[y] \in P(o^s)\}| = |\{y \in \overline{P}(o^s)\}| \pm 1$$

that concerns the behaviour of vague predicates on Sorites series. The assumption is that half of the domain has property P, half has not. An intuitive justification for such a strong assumption is based on the interpretation of P and $\overline{P}$ as opposite poles. There is a tension between opposite poles. The objects that are closer to the positive pole P more clearly have property P. As the distance from the positive pole increases, the objects are less clearly P, and the highest grade of tension (and uncertainty) is exactly in the middle of the series. After that point, the objects get closer to the negative pole and so they are more and more clearly $\overline{P}$. The cut-off point between the extension of P and of $\overline{P}$ then lies in the middle of the series. The ideal tension between the two poles as they were magnetic poles is the intuitive justification for our assumption to draw a line in the middle of the series to distinguish P-objects from $\overline{P}$-objects. In fact, the vagueness problem consists exactly in where to draw the cut-off line. With our proposal I do not want to *solve* the problem of vagueness. I refer to the assumption above to explain why we agree on some unproblematic uses of polar adjectives. So, the model does not provide a way to draw the line to discriminate the extensions of P and $\overline{P}$. That is taken as granted and made be the basis for the unproblematic use of polar adjectives. But what is still to be analysed is whether this way of thinking the use of vague adjectives referring to soritical series corresponds to what we also cognitively do: do we refer to a possible soritical series to get the meaning of 'tall' and 'short' when we are using these adjectives in unproblematic contexts?

5.2.3 On Granularity

In the sorites series contiguous elements are indistinguishable with respect to a property P from some point of view (usually, from an observational point of view). Using a more efficient way of measurement, with a standard of precision of a higher level, we can determine a larger number of differences between the objects. The fact that we cannot do it by ourselves is not due to some deficiencies or shortcomings of our cognitive system, but it shows that we have been "attuned to the aspects of our environment that are most likely to be relevant to our interests" (Hobbs [13], p. 433).
This idea might bring about a change in the epistemic conception of vagueness. Under such an idea vagueness is not seen as a mere defect of our epistemic capacities, that is, of our capacity to be acquainted with the world around us. On the contrary, it is positively considered, that is, as the result of human adaptation to the world. If we cannot distinguish some differences is also because, from a pragmatic point of view, we do not usually *need* to do that.

Vagueness and Granularity in Formal Ontology

The idea of connecting granularity to vagueness has been developed in Bittner and Smith's jointed papers, such as [5], [6], [7]. Bittner and Smith worked mainly on the problem of vagueness of proper names in formal ontology. They propose a formal framework connecting the idea of granularity with mereotopology, while what has been proposed in the present thesis is rather a semantic framework connecting granularity with algebra. Sets of fixed models (what I called *context structures*) are taken to correspond to granular levels, and granular partitions are meant to be exactly equivalent classes from an algebraic point of view.
In Smith and Brogaard [39] it is pointed out that the term 'partition' is not used to mean 'equivalence class'. A granular partition is a grid of cells that gives an abstract classification of objects in reality:

> A granular partition is a way of dividing up the world, or some portion of the world, by means of cells.[1]

We can say, then, that while Bittner and Smith's granular partitions are a way to divide things and get several categories of objects related one each other, the granular partitions proposed in the present research are ways to describe objects in the world taken one by one and considered in their similarity relations with the others. In other terms, while granular partitions as systems of cells can be used to give a conceptualisation of the world itself and its constituents from an ontological point of view, granular partitions as equivalence relations can be used to describe single items from a semantic point of view.

5.2.4 On Fine-Grainedness and Precisifications

At first glance, fine-grained levels might closely look like sharpenings in supervaluationist theories. The aim of both the constructions is to obtain more and more precision in the evaluation of vague predicates. But consider the differences between the two views.
In supervaluationism the value of a sentence is given by a quantification over all the precisifications. That means, precisifications are functional to the semantic evaluation of sentences. Instead, there is no quantification over context structures that are of different granular levels. Each of them

[1] Smith and Brogaard [39], p. 6.

represents a way to look at the domain and they are functional to the establishment of the comparative relation. From precisification to precisification, the valuation function of predicates changes. In the model presented here, there is variation of the choice functions representing predicates within *each* granular level. The variation across granular levels concerns primarily the comparative relations, that is, the number of differences detected in the contexts. Precisifications are a semantic device to capture the intuition that sentences concerning borderline cases are problematic to be evaluated. The intuition behind granular levels is different. Those are invoked to reflect the different ways we can look towards a domain with respect to standards of precision. So, the matter here is epistemic.

Bibliography

[1] Armour-Garb, B. Diagnosing Dialetheism. In *The Law of Non-Contradiction: New Philosophical Essays*, G. Priest, J. Beall, and B. Armour-Garb, Eds. Clarendon Press - Oxford, 2004, pp. 113–25.

[2] Barker, C. The Dynamics of Vagueness. *Linguistics and Philosophy 25*, 1 (2002), 1–36.

[3] Barker, C. Vagueness. In *Encyclopedia of language and linguistics.* 2nd ed. Elsevier, Amsterdam, 2006.

[4] Beall, J., and Colyvan, M. Looking for Contradictions. *Australasian Journal of Philosophy 79*, 4 (2001), 564–569.

[5] Bittner, T., and Smith, B. Granular Partitions and Vagueness. In *Proceedings of the international conference on Formal Ontology in Information Systems-Volume 2001*, C. Welty and B. Smith, Eds. ACM Press New York, NY, USA, 2001, pp. 309–320.

[6] Bittner, T., and Smith, B. A Theory of Granular Partitions. In *Foundations of Geographic Information Science*, M. Duckham, M. Goodchild, and M. Worboys, Eds. CRC Press, 2002, pp. 117–151.

[7] Bittner, T., and Smith, B. Vague Reference and Approximating Judgments. *Spatial Cognition and Computation 3* (2003), 137–156.

[8] Black, M. Vagueness. An Exercise in Logical Analysis. *Philosophy of Science 4* (1937), 427–455.

[9] Bosch, P. Vagueness is Context-Dependence. A Solution to the Sorites Paradox. In *Approaching Vagueness*, T. Ballmer and M. Pinkal, Eds. North-Holland, Amsterdam, 1983, pp. 189–210.

[10] Cresswell, M. The Semantics of Degree. In *Montague Grammar*, B. Partee, Ed. Academic Press - New York, 1976, pp. 261–292.

[11] Dummett, M. Wang's Paradox. *Synthese 30*, 3 (1975), 301–324.

[12] Fara, D. Shifting Sands: An Interest-Relative Theory of Vagueness. *Philosophical Topics 28*, 1 (2000), 45–81.

[13] Hobbs, J. Granularity. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, A. Joshi, Ed. Los Angeles, USA: Morgan Kaufmann, 1985, pp. 432–435.

[14] Kamp, H. Two Theories of Adjectives. In *Formal Semantics of Natural Language*, E. Keenan, Ed. Cambridge University Press, Cambridge, 1975, pp. 123–155.

[15] Keefe, R. *Theories of Vagueness*. Cambridge University Press, 2000.

[16] Keefe, R., and Smith, P. Introduction: Theories of Vagueness. In *Vagueness: A Reader*, R. Keefe and P. Smith, Eds. MIT Press, Cambridge (MA), 1997, pp. 1–57.

[17] Kennedy, C. Modes of Comparison. *Proceedings of CLS 43*. forthcoming.

[18] Kennedy, C. Comparison and Polar Opposition. In *Proceedings of SALT*, vol. 7. CLC Publications - Ithaca, 1997, pp. 240–257.

[19] Kennedy, C. Gradable Adjectives Denote Measure Functions, Not Partial Functions. *Studies in the Linguistic Sciences 29*, 1 (1999), 65–80.

[20] Kennedy, C. On the Monotonicity of Polar Adjectives. In *Perspectives on Negation and Polarity Items*, V. S.-V. J. Hoeksema, H. Rullmann and T. van der Wouden, Eds. 2001, pp. 201–221.

[21] Kennedy, C. Polar Opposition and the Ontology of Degrees. *Linguistics and Philosophy 24*, 1 (2001), 33–70.

[22] Kennedy, C. Vagueness and Grammar: the Semantics of Relative and Absolute Gradable Adjectives. *Linguistics and Philosophy 30*, 1 (2007), 1–45.

[23] Klein, E. The Semantics of Positive and Comparative Adjectives. *Linguistics and Philosophy 4* (1980), 1–45.

[24] Klein, E. Comparatives. In *Semantik: Ein internationales Handbuch der zeitgenossischen Forschung*, A. von Stechow and D. Wunderlich, Eds. N. Y. de Gruyter, 1991, pp. 673–691.

[25] Löbner, S. *Wahr neben Falsch: Duale Operatoren als die Quantoren natürlicher Sprache.* Niemeyer, 1990.

[26] Luce, R. Semiorders and a Theory of Utility Discrimination. *Econometrica 24*, 2 (1956), 178–191.

[27] Mani, I. A Theory of Granularity and Its Application to Problems of Polysemy and Underspecification of Meaning. In *Proceedings of the Sixth International Conference on Principles of Knowledge Representation and Reasoning* (1998), L. K. S. A. G. Cohn and S. C. Shapiro, Eds., Morgan Kaufmann, pp. 245–257.

[28] Moltmann, F. Comparatives without Degrees: A Trope-Based Analysis. unpublished.

[29] Moltmann, F. Properties and Kinds of Tropes: New Linguistic Facts and Old Philosophical Insights. *Mind 113*, 449 (2004), 1–43.

[30] Olin, D. *Paradox.* Chesham: Acumen, 2003.

[31] Parikh, R. Vagueness and Utility: The Semantics of Common Nouns. *Linguistics and Philosophy 17*, 6 (1994), 521–535.

[32] Parsons, T. *Indeterminate Identity: Metaphysics and Semantics.* Oxford University Press, 2000.

[33] Parsons, T., and Woodruff, P. Worldly Indeterminacy of Identity. *Proceedings of the Aristotelian Society 95* (1995), 171–191.

[34] Prinz, J. Vagueness, Language, and Ontology. *Electronic Journal of Analytical Philosophy 6* (1998). http://ejap.louisiana.edu/EJAP/1998/prinz98.html.

[35] Raffman, D. Vagueness Without Paradox. *Philosophical Review 103*, 1 (1994), 41–74.

[36] Russell, B. Vagueness. *Australasian Journal of Philosophy 1*, 2 (1923), 84–92.

[37] Seuren, P. The Comparative. In *Generative Grammar in Europe*, F. Kiefer and N. Ruwet, Eds. Dordrecht: Riedel, 1973, pp. 528–564.

[38] SEUREN, P. The Comparative Revisited. *Journal of Semantics 3*, 1-2 (1984), 109–141.

[39] SMITH, B., AND BROGAARD, B. Quantum Mereotopology. *Annals of Mathematics and Artificial Intelligence 36*, 1 (2002), 153–175.

[40] SORENSEN, R. *Vagueness and Contradiction*. Oxford: Oxford University Press, 2001.

[41] UNGER, P. There Are No Ordinary Things. *Synthese 41*, 2 (1979), 117–154.

[42] VAN BENTHEM, J. Later Than Late. On the Logical Origin of the Temporal Order. *Pacific Philosophical Quarterly 63* (1982), 193–203.

[43] VAN ROOIJ, R. Semi-Orders and Satisficing Behaviour. to appear.

[44] VAN ROOIJ, R. Vagueness and Pragmatics. In *The Vagueness Handbook*, G. Ronzitti, Ed. Springer. forthcoming.

[45] VELTMAN, F., AND MUSKENS, R. Lecture Notes on Logical Analyse. unpublished.

[46] VON STECHOW, A. My Reaction to Cresswell's, Hellan's, Hoeksema's, and Seuren's comments. *Journal of Semantics 3* (1984), 183–199.

[47] WILLIAMSON, T. *Vagueness*. Routledge, 1994.

[48] WOODRUFF, P., AND PARSONS, T. Indeterminacy of Identity of Objects and Sets. *Nous (Supplement: Philosophical Perspectives, 11, Mind, Causation and World), 31* (1997), 321–348.