Tjalling C. Koopmans Research Institute

Discussion Paper Series nr: 09-31

Strategic vagueness and appropriate contexts

Kris De Jaegher
Robert van Rooij

Universiteit Utrecht

Utrecht School of Economics

Tjalling C. Koopmans Research Institute
Utrecht School of Economics
Utrecht University

Janskerkhof 12
3512 BL Utrecht
The Netherlands
telephone +31 30 253 9800
fax +31 30 253 7373
website www.koopmansinstitute.uu.nl

The Tjalling C. Koopmans Institute is the research institute and research school of Utrecht School of Economics.
It was founded in 2003, and named after Professor Tjalling C. Koopmans, Dutch-born Nobel Prize laureate in economics of 1975.

In the discussion papers series the Koopmans Institute publishes results of ongoing research for early dissemination of research results, and to enhance discussion with colleagues.

Please send any comments and suggestions on the Koopmans institute, or this series to J.M.vanDort@uu.nl

ontwerp voorblad: WRIK Utrecht

How to reach the authors

Please direct all correspondence to the first author.

Kris De Jaegher
Utrecht University
Utrecht School of Economics
Janskerkhof 12
3512 BL Utrecht
The Netherlands.
E-mail: k.dejaegher@uu.nl

This paper can be downloaded at: http://
www.uu.nl/rebo/economie/discussionpapers

Utrecht School of Economics
Tjalling C. Koopmans Research Institute
Discussion Paper Series 09-31

Strategic Vagueness and Appropriate Contexts*

Kris De Jaegher[a]
Robert van Rooij[b]

[a]Utrecht School of Economics
Utrecht University

[b]Institute of Logic, Language and Computation
University of Amsterdam

November 2009

Abstract
This paper brings together several approaches to vagueness, and ends by suggesting a new approach. The common thread in these approaches is the crucial role played by context. Using a single example where there is a conflict of interest between speaker and listener, we start by treating game-theoretic rationales for vagueness, and for the related concepts of generality and ambiguity. We argue that the most plausible application of these models to vagueness in natural language is one where the listener only imperfectly observes the context in which the speaker makes her utterances. We next look at a rationale for vagueness when there is no conflict between speaker and listener, and which is an application of Horn's rule. Further, we tackle the Sorites paradox. This paradox apparently violates standard axioms of rational behaviour. Yet, once it is taken into account that vague language is used in an appropriate context, these axioms are no longer violated. We end with a behavioural approach to vagueness, where context directly enters agents' preferences. In an application of prospect theory, agents think in terms of gains and losses with respect to a reference point. Vague predicates now allow agents to express their subjective valuations, without necessarily specifying the context.

Keywords: Vagueness, signalling games, decision theory, prospect theory.

JEL classification: D82, D83.

Acknowledgements
We wish to thank participants of the workshop Vagueness and Decision Theory, 29th May 2009, Ecole Normale Supérieure, Paris and participants of the conference Explaining Vagueness, Games and Decision in Pragmatics III, Centre for General Linguistics Typology and Universals Research (ZAS), Berlin, 23-24 November 2007.

Strategic Vagueness, and appropriate contexts

Kris De Jaegher and Robert van Rooij

1 Introduction

This paper brings together several approaches to vagueness, and ends by suggesting a new approach. The common thread in these approaches is the crucial role played by context. In Section 2, we treat game-theoretic rationales for vagueness, and for the related concepts of ambiguity and generality. Common about these rationales is that they are based on the assumption of a conflict of interest between speaker and listener. We review this literature using a single example. We argue that the most plausible application to vagueness in natural language of these models is one where the listener only imperfectly observes the context in which the speaker makes her utterances. Yet, it is clear that not all vagueness can be accounted for by conflicts of interest. This is why the rest of the paper looks at the case of common interest. Section 3 argues that being vague by saying that someone is bald makes sense in a context where precision is of less importance; in a context where precision is of more importance, one can then refer to someone as completely bald. This make sense because the longer and therefore more costly to utter expression 'completely bald' is then used less often. Vagueness is thus seen as an application of Horn's pragmatic rule that (un)marked states get an (un)marked expression. Section 4 tackles the Sorites paradox, which apparently leads to the violation of standard axioms of rational behaviour, and shows that this paradox arises from the use of vague predicates in an inappropriate context. If, as suggested by the Sorites paradox, fine-grainedness is important, then a vague language should not be used. Once vague language is used in an appropriate context, standard axioms of rational behaviour are no longer violated. Section 5 finally takes a different approach from the previous sections, and following prospect theory assumes that context directly enter agents' utility functions in the form of reference points, with respect to which agents think in gains and losses. The rationale for vagueness here is that vague predicates allow players to express their valuations, without necessarily uttering the context, so that the advantage of vague predicates is that they can be expressed across contexts.

2 Vagueness and games of conflict

Game theorists have tried to justify generality, vagueness and ambiguity in language as ways to solve conflicts of interest between a sender and a receiver. We start by treating such a sender-receiver game with a conflict of interest (a variant of a game treated by Farrell, 1993), and its equilibria in the absence of what is interpreted as generality, vagueness and ambiguity.

Table 1:

$U(t_i, a_j)$	a_1	a_2	a_3
t_1	3,3	1,0	0,2
t_2	1,0	0,3	-1,2

Consider the signalling game with the payoff structure in Table 1. The sender (she) observes whether the state of nature is t_1 or t_2. Each of these states occur with probability $\frac{1}{2}$. The receiver (he) can choose among action a_1, a_2 or a_3. After having observed the state of nature, and prior to the receiver taking one of the actions, the sender can send a signal 's_1', a signal 's_2' to the receiver, or no signal at all. Sending a signal comes at no cost whatsoever to the sender. The meaning of the signals is completely conventional, but we focus on separating equilibria where each signal is sent more often in one particular state, thus justifying the labels of the signals. All aspects of the game are common knowledge among the sender and the receiver. Using a story adapted from Blume, Board and Kawamura (2007), we may consider the sender as Juliet, who may be of two types, namely one who loves Romeo (the receiver), and one who is merely fond of him as a friend. a_1 means that Romeo acts as if Juliet is in love with him, a_2 means that Romeo acts as if Juliet is a friend, and a_3 means that Romeo acts as if Juliet could be either a friend or in love with him. Without any knowledge, Romeo prefers the latter action, but this is Juliet's least preferred outcome. Both types of Juliets prefer a_1 to a_2.

We start by looking at the standard case where any randomisation of the sender in her strategies is unrelated to the randomisation of the receiver (excluding correlated equilibria see below), and where a signal sent by the sender is always received and never misinterpreted by the receiver (excluding noisy equilibria — see below). It is easy to check that the signalling equilibria has two Nash equilibria. In a first equilibrium, the *pooling equilibrium*, Juliet does not send any signal, and Romeo does a_3. This Nash equilibrium is also a perfect Bayesian equilibrium (cf. Fudenberg & Tirole, 1991) because we can find beliefs for each player that underpin this equilibrium. Romeo believes that anything Juliet says does not contain any credible information. His beliefs are not disconfirmed as he never actually observes Juliet talking, and he might as well keep his beliefs. Juliet from her side correctly believes that anything that she might say will be met with action a_3. More correctly, there is in fact a range of pooling equilibria, where given that signals are costless to send, Juliet may talk in any uninformative way, for instance by always saying the same, or mixing between signals in a manner that has little correlation with her own state.

The second Nash equilibrium is a *separating equilibrium*, and takes the form of a mixed equilibrium. Juliet always sends signal 't_1' ('I love you') in state t_1; however, in state t_2, she sends 't_2' ('I like you') with probability $\frac{1}{2}$, and 't_1' ('I love you') with probability $\frac{1}{2}$. Put otherwise, Juliet is always honest when she is in love, but is half of the time honest and half of the time dishonest when she is merely fond of Romeo. Romeo always acts as if Juliet is a friend when she says that she likes him; however, when she says 'I love you', with probability $\frac{1}{2}$ he acts as if she loves him, and with probability $\frac{1}{2}$ he acts as if she could either be a friend or someone who loves him. To see that these are mutual best responses, note first that Juliet's expected payoff of telling that she loves Romeo when she is merely fond of him is now $\frac{1}{2} \times 1 + \frac{1}{2} \times (-1) = 0$; 0 is also her expected payoff of sending signal 't_2'. It follows that Juliet is indifferent between saying 'I love you' and 'I like you'. Second, note that when Romeo sees Juliet telling that she loves him by Bayes' rule will (somewhat prosaically) estimate the probability that she indeed does to be $\frac{\frac{1}{2}}{\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}}$, and will estimate that in fact she is merely fond of him to be $\frac{\frac{1}{2}\times\frac{1}{2}}{\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}}$. It follows that Romeo's expected utility of acting as if Juliet loves him when she tells him she does equals $\frac{\frac{1}{2}}{\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}} \times 3 + \frac{\frac{1}{2}\times\frac{1}{2}}{\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}} \times 0 = 2$. This equals Romeo's payoff 2 of acting as if Juliet might either love him or like him. Similarly, his expected payoff of acting as if Juliet is merely a friend equals $\frac{\frac{1}{2}}{\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}} \times 0 + \frac{\frac{1}{2}\times\frac{1}{2}}{\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}} \times 3 = 1$. It follows that Romeo is indifferent between acting as if Juliet loves him, and acting as if she might either love him or like him. In these circumstances, Romeo might as well take each action half of the time.

As argued by Lipman (2006), one could see Juliet as being vague, as 'I love you' does not always mean then that Juliet really does love Romeo. But a particular feature of such a mixed equilibrium is that Juliet is completely indifferent about what to say when she is merely fond of Romeo, and Romeo is indifferent about what to do when Juliet tells him that she loves him. And yet, each player is assumed to mix between his or her strategies in a very particular way, in order to keep the other player indifferent. A justification given for the mixed equilibrium (Harsanyi, 1973) is that we can see this as a simple representation of a more complex model where in fact we have a population of Juliets, whose payoffs differ so that they can be interpreted as varying according to their degree of trustworthiness; and a population of Romeos, whose payoffs differ so that they can be interpreted as varying according to their degree of trustfulness. Each Juliet knows her own trustworthiness, but does not observe the trustfulness of the Romeo facing her; she merely knows how trustfulness is distributed among Romeos. Similarly, Romeo only knows how trustworthiness is distributed among Juliets. So, rather than mixing taking place, the most trustworthy Juliets in the population are honest, and the most trustful Romeos trust the Juliets. Yet, such a justification of the mixed equilibrium adds new aspects to the game. Sticking to the original game between the individual Romeo and Juliet in the game in Table 1, we now investigate several ways in which the players could still communicate, without the unsatisfactory aspects of a mixed equilibrium.

2.1 Strategic generality

Let us say that a sentence is specific when it is true only in a few circumstances. A sentence is general when it is true in many circumstances. It is standard to assume in pragmatics that it is better to be more specific. However, in case of a conflict of interest, it can be beneficial to be general. To see why, take the example in Table 1, and introduce a third state t_3 about which the interests of sender and the receiver fully coincide (namely, a_4 is then the best action), as represented in Table 2. Concretely, in terms of the above story, in state t_3 Juliet neither likes nor loves Romeo, and action a_4 means that Romeo acts as if she neither likes him nor loves him. Then a separating equilibrium exists where in both state t_1 and state t_2, Juliet tells Romeo 'Either I love you or I like you', upon which Romeo acts as if she might either love him or like him; in state t_3 Juliet tells Romeo that she does not like him and she certainly does not love him.

Table 2:

$U(t_i, a_j)$	a_1	a_2	a_3	a_4
t_1	3,3	1,0	0,2	-2,-2
t_2	1,0	0,3	-1,2	-2,-2
t_3	2,2	2,2	2,2	1,1

A more sophisticated version of this argument is found in Crawford and Sobel (1982). In their model, the sender observes one state out of a continuous range of states of nature, and the receiver can pick an action from a continuous range of actions. The discrepancy between the sender's optimal action and the receiver's optimal action, where the sender always prefers a higher action, measures the degree of conflict between the two players. Crawford and Sobel (1982) show that communication can still take place between sender and receiver in spite of the conflict of interest between them if the sender uses a finite number of signals, where each signal is used for a range of the continuum of states of nature, so that signals partition the continuum of states. In the efficient separating equilibrium, this partition is less fine-grained the higher the degree of conflict between sender and receiver. Intuitively, as her signals become less and less finely tuned, eventually the sender prefers to tell the truth.

2.2 Strategic vagueness and strategic ambiguity

Yet, it is sometimes possible to do better still than with strategic generality, and in the game of Table 2 let the sender and receiver still communicate about t_1 and t_2 as well. In particular, this is possible when the sender uses what could be described either as vague or as ambiguous sentences rather than general sentences. To show this, it suffices to show that players can still communicate in the game in Table 1 without playing a mixed equilibrium.

2.2.1 Noisy signalling interpreted as vagueness/ambiguity

We first look at noisy signalling (Farrell, 1993; Myerson, 1991; De Jaegher, 2003a,b; Blume, Board and Kawamura, 2007; Blume and Board, 2009). Juliet sends noisy signals, which may simply remain unheard (*errors of detection*) or may be misinterpreted (*errors of discrimination*). Denote by $\mu(\tilde{i}|\tilde{j})$ the probability that Romeo perceives signal $\tilde{i}$ when the sender sent signal $\tilde{j} = \tilde{t_1}, \tilde{t_2}$, where $\tilde{i} = \tilde{t_1}, \tilde{t_2}$, 0 (where 0 denotes not receiving any signal). Consider a noisy signalling system where $\mu(\tilde{t_1}|\tilde{t_2}) = \mu(0|\tilde{t_2}) = \frac{1}{4}$, $\mu(\tilde{t_2}|\tilde{t_2}) = \frac{1}{2}$, and $\mu(\tilde{t_1}|\tilde{t_1}) = \mu(0|\tilde{t_1}) = \frac{1}{2}$. It is easy to check now that a separating equilibrium exists where Juliet honestly send signal $\tilde{t_1}$ ($\tilde{t_2}$) in state t_1 (t_2). Romeo does a_1 (a_2) when perceiving signal $\tilde{t_1}$ ($\tilde{t_2}$), and does a_3 when not receiving any signal. To see this, note that '$\tilde{t_2}$' is only received in state t_2, so that Romeo indeed does a_2. When perceiving signal '$\tilde{t_1}$', Romeo's expected payoff of doing a_1 equals $\frac{\frac{1}{2}\times\frac{1}{2}}{\frac{1}{2}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{4}}\times 3+\frac{\frac{1}{2}\times\frac{1}{4}}{\frac{1}{2}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{4}}\times 0 = 2$. This equals his fixed payoff 2 of taking action a_3. Romeo's payoff of taking action a_2 equals $\frac{\frac{1}{2}\times\frac{1}{2}}{\frac{1}{2}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{4}}\times 0+\frac{\frac{1}{2}\times\frac{1}{4}}{\frac{1}{2}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{4}}\times 3 = 1$. It follows that Romeo is just still willing to act as if Juliet loves him when receiving noisy signal '$\tilde{t_1}$'. Juliet's expected payoff when sending noisy signal '$\tilde{t_2}$' in state t_2 equals $\frac{1}{4}\times 1+\frac{1}{2}\times 0+\frac{1}{4}\times(-1) = 0$. Juliet's expected payoff when sending noisy signal '$\tilde{t_1}$' in state t_2 equals $\frac{1}{2}\times 1+\frac{1}{2}\times(-1) = 0$. It follows that, when she is merely fond of Romeo, Juliet is just still willing to send the noisy signal '$\tilde{t_2}$'. Juliet's expected payoff when sending noisy signal '$\tilde{t_1}$' in state t_1 equals $\frac{1}{2}\times 3+\frac{1}{2}\times 0 = 1.5$; her expected payoff of sending noisy signal '$\tilde{t_2}$' equals only $\frac{1}{4}\times 3+\frac{1}{2}\times 1+\frac{1}{4}\times 0 = 1.25$. But note now that this noisy signalling equilibrium perfectly replicates the outcome of the mixed equilibrium (with noiseless signalling) described above, and suffers from similar drawbacks: both

Romeo and Juliet are in fact indifferent about what to do.

Yet, other levels of noise can be found such that each player strictly prefers to follow the separating equilibrium. For instance, take the case $\mu(\tilde{t_1}|\tilde{t_2}) = 0.2$, $\mu(0|\tilde{t_2}) = 0.28$, $\mu(\tilde{t_2}|\tilde{t_2}) = 0.52$, and $\mu(0|\tilde{t_1}) = 0.55$, $\mu(\tilde{t_1}|\tilde{t_1}) = 0.45$. Note that when perceiving signal '$\tilde{t_1}$', Romeo's expected payoff of doing a_1 equals $\frac{\frac{1}{2}\times 0.45}{\frac{1}{2}\times 0.45+\frac{1}{2}\times 0.2} \times 3 + \frac{\frac{1}{2}\times 0.2}{\frac{1}{2}\times 0.45+\frac{1}{2}\times 0.2} \times 0 = 2.08$. His payoff when taking action a_2 equals $\frac{\frac{1}{2}\times 0.45}{\frac{1}{2}\times 0.45+\frac{1}{2}\times 0.2} \times 0 + \frac{\frac{1}{2}\times 0.2}{\frac{1}{2}\times 0.45+\frac{1}{2}\times 0.2} \times 3 = 0.92$, his payoff when taking action a_3 is fixed at 2. It follows that Romeo strictly prefers to do action a_1. When not receiving any signal, Romeo's expected payoff of doing a_1 equals $\frac{\frac{1}{2}\times 0.55}{\frac{1}{2}\times 0.55+\frac{1}{2}\times 0.28} \times 3 + \frac{\frac{1}{2}\times 0.28}{\frac{1}{2}\times 0.55+\frac{1}{2}\times 0.28} \times 0 = 1.99$. His expected payoff of doing a_2 equals $\frac{\frac{1}{2}\times 0.55}{\frac{1}{2}\times 0.55+\frac{1}{2}\times 0.28} \times 0 + \frac{\frac{1}{2}\times 0.28}{\frac{1}{2}\times 0.55+\frac{1}{2}\times 0.28} \times 3 = 1.01$. This is both smaller than the fixed payoff 2 of doing a_3.

Juliet's expected payoff when sending noisy signal '$\tilde{t_2}$' in state t_2 equals $0.2 \times 1 + 0.52 \times 0 + 0.28 \times (-1) = -0.08$. Her expected payoff when sending noisy signal '$\tilde{t_1}$' in state t_2 equals $0.45 \times 1 + 0.55 \times (-1) = -0.1$. It follows that she strictly prefers to send noisy signal '$\tilde{t_2}$'. Juliet's expected payoff when sending noisy signal '$\tilde{t_1}$' in state t_1 equals $0.45 \times 3 + 0.55 \times 0 = 1.35$; her expected payoff from sending noisy signal '$\tilde{t_2}$' in this state equals only $0.2\times 3+0.52\times 1+0.28\times 0 = 1.12$. In general, a range of noisy signalling systems exist that allow for pure-strategy, strict separating equilibria. Note that, as is the case in the treated examples, these noisy signalling systems may leave the players better off than with an equilibrium without noisy signalling as treated in Section 1. Noisy signalling systems are found by plugging general noise levels $\mu(\tilde{i}|\tilde{j})$ into the linear constraints telling that Juliet should be honest, and that Romeo should follow the lead of the perceived signals (or do a_3 when not receiving any signal). Given that $\mu(\tilde{t_1}|t_i)) + \mu(\tilde{t_2}|t_i) + \mu(0|t_i) = 1$, and given that all constraints are linear, the set of noisy signalling systems allowing for a separating equilibrium can be represented as a polyhedron in a four-dimensional space.

A more sophisticated version of this argument is found in Blume, Board and Kawamura, (2007) who extend Crawford and Sobel's (1982) model (see above) to the case of noise. The authors show that noisy communication allows for separating equilibria that are Pareto superior to those presenting strategic generality. In such a model with a continuous range of actions, appropriate noise leads the receiver to revise the action response to any given signal downwards, thus aligning the sender's and the receiver's interest.

This argument, and the example above, merely show that if players happen to use one out of a particular set of noisy signalling systems, they can still effectively communicate. Yet, if multiple noisy signalling systems are available, communication can only be effective if Romeo observes that Juliet is using a particular noisy signalling system. Otherwise, Juliet could pretend to be using such a system, and still always send a clear 'I love you' signal. Put otherwise, the noisy signalling system must ostentatiously be such that the sender cannot control how the receiver will interpret her signals. Blume and Board (2009) assume that the receiver observes the degree of noisiness chosen by the receiver, and that the sender maximises her utility with respect to the level of noise. In terms of the example above, given that the space of noise levels allowing for communication takes the form of a polyhedron, and given the linear structure of Juliet's expected payoffs, she will pick one of the corner points of the polyhedron. In particular, it can be shown that Juliet will pick a noisy signalling system as close as possible to the one replicating the mixed equilibrium.

De Jaegher (2003a) argues that the level of noise in the noisy signalling system can itself be considered as a signal, in constituting a handicap signal. When Romeo receives a noisy 'I love you' signal, Romeo makes the reasoning that only a Juliet who really loves him would be willing to incur the cost of her message sometimes getting lost. For an example, take a case where there are only errors of detection, with $\mu(0|\tilde{t_2}) = \frac{1}{3}$, $\mu(\tilde{t_2}|\tilde{t_2}) = \frac{2}{3}$, and $\mu(0|\tilde{t_1}) = \frac{2}{3}$, $\mu(\tilde{t_1}|\tilde{t_1}) = \frac{1}{3}$. Obviously, Romeo will always take action a_1 (a_2) upon a $\tilde{t_1}$ ($\tilde{t_2}$)) signal. When not receiving any signal, Romeo's expected payoff of doing a_1 equals $\frac{\frac{1}{2}\times\frac{2}{3}}{\frac{1}{2}\times\frac{2}{3}+\frac{1}{2}\times\frac{1}{3}}\times 3+\frac{\frac{1}{2}\times\frac{1}{3}}{\frac{1}{2}\times\frac{2}{3}+\frac{1}{2}\times\frac{1}{3}}\times 0 = 2$. His expected payoff of doing a_2 equals $\frac{\frac{1}{2}\times\frac{2}{3}}{\frac{1}{2}\times\frac{2}{3}+\frac{1}{2}\times\frac{1}{3}}\times 0+\frac{\frac{1}{2}\times\frac{1}{3}}{\frac{1}{2}\times\frac{2}{3}+\frac{1}{2}\times\frac{1}{3}}\times 3 = 1$. The fixed payoff of doing a_3 equals 2. It follows that Romeo is just still willing to do a_3.

Juliet's expected payoff when sending noisy signal '$\tilde{t_2}$' in state t_2 equals $\frac{2}{3}\times 0+\frac{1}{3}\times(-1) = -\frac{1}{3}$. Her expected payoff when sending noisy signal '$\tilde{t_1}$' in state t_2 equals $\frac{1}{3}\times 1+\frac{2}{3}\times(-1) = -\frac{1}{3}$. It follows that she is just still willing to send noisy signal '$\tilde{t_2}$'. Juliet's expected payoff when sending noisy signal '$\tilde{t_1}$' in state t_1 equals $\frac{1}{3}\times 3+\frac{2}{3}\times 0 = 1$; her expected payoff from sending noisy signal '$\tilde{t_2}$' in this state equals only $\frac{2}{3}\times 1+\frac{1}{3}\times 0 = \frac{2}{3}$. Thus, not receiving a sufficient amount of information leads Romeo to taking a costly action. By making her signals more noisy the 'higher' is her type, Juliet makes it more likely that Romeo takes an action that is costly to her the higher is her type. This again aligns the players' interests.

How can such noisy talk be linked to natural language? In the words of Blume et al. (2007): "In a given context, the meaning, and hence the correct interpretation of a vague word may depend on the language habits of the utterer. Here, when Juliet says 'I love you', Romeo understands that she is probably in love with him, but also that these words are occasionally used to express mere fondness." But this rationale very much resembles the one for a mixed equilibrium: Juliet sometimes says 'I love you' when she is merely fond of Romeo. Yet what we are looking for is an utterance '$\tilde{t_1}$' sent only in state t_1, but which Romeo sometimes mistakenly interprets as referring to state t_2. Perhaps ambiguity is a potential example. Suppose that Juliet when she loves Romeo ironically tells him 'Dreadful guy!'. The idea is then that such irony is risky, and that outside of Juliet's control it will sometimes be misunderstood. Unfortunately, this example could be referred to as one of strategic ambiguity rather than strategic vagueness. In general, it is very difficult to make any concrete link between noisy signalling and natural language. After all, in Farrell's original example (1993), one way in which Juliet could still credibly communicate with Romeo is by sending an unreliable carrier pigeon which sometimes fails to arrive. In the next section, we argue that correlated equilibria give a more plausible account of vagueness.

2.2.2 Correlated equilibria and vagueness

Instead of looking at a noisy signalling equilibrium, let us now explore the possibility of correlated equilibria. In a correlated equilibrium, each player observes random events in nature, and lets his or her strategies depend on these events. The random events observed by the players may be correlated, in turn making the players strategies correlated (Aumann, 1974). In the Romeo-Juliet game, consider the following example. Fully independently of whether state t_1 or state t_2 occurs, let Juliet observe with probability $P(t_{I,j})$ an event $t_{I,j}$, and with probability $P(t_{II,j})$ an event $t_{II,j}$, where $P(t_{I,j}) + P(t_{II,j}) = 1$. Before signalling takes place, Romeo and Juliet go on a date, and event $t_{I,j}$ means that Juliet had an excellent time, whereas in event $t_{II,j}$ Juliet merely had a nice time. Denote by $t_{I,r}$ the event that Romeo had an excellent time, and by $t_{II,r}$ the event that Romeo had a nice time. Denote by $\mu(i|j)$ the probability that Romeo perceives event $i = t_{I,r}$, $t_{II,r}$ when Juliet has observed event $j = t_{I,j}, t_{II,j}$. Thus, in this model, it need not always be the case that Romeo had an excellent time when Juliet did.

Consider now an equilibrium where Juliet says 'I love you' in events $(t_1, t_{I,j})$, $(t_1, t_{II,j})$, and $(t_2, t_{I.j})$, and says 'I like you' in event $(t_2, t_{II,j})$. Put otherwise, Juliet always says 'I love you' if she loves Romeo. However, she also says this if she merely likes Romeo but just had an excellent time. Romeo acts as if Juliet loves him (a_1) when observing $(t_1, t_{II,r})$, and acts as if Juliet may either like him or love him (a_3) when observing $(t_1, t_{I,r})$. Also, he acts as if Juliet merely considers him as a friend (a_2) whether observing $(t_2, t_{I.r})$ or $(t_2, t_{II,r})$.

Specifically, consider the case $P(t_{I,j}) = 0.5, P(t_{II,j}) = 0.5, \mu(t_{II,r}|t_{I,j}) = 0.5, \mu(t_{I,r}|t_{II,j}) = 0.5$. When receiving signal 'I love you' and observing $t_{II,r}$, Romeo's expected payoff of doing a_1 equals $\frac{\frac{1}{2}\times(0.5\times0.5+0.5\times0.5)}{\frac{1}{2}\times(0.5\times0.5+0.5\times0.5)+\frac{1}{2}\times(0.5\times0.5)} \times 3 +$

$\frac{\frac{1}{2}\times(0.5\times0.5)}{\frac{1}{2}\times(0.5*0.5+0.5\times0.5)+\frac{1}{2}\times(0.5\times0.5)} \times 0 = 2$. In the same case, his expected payoff of doing a_2 equals $\frac{\frac{1}{2}\times(0.5\times0.5+0.5\times0.5)}{\frac{1}{2}\times(0.5\times0.5+0.5\times0.5)+\frac{1}{2}\times(0.5\times0.5)} \times 0 + \frac{\frac{1}{2}\times(0.5\times0.5)}{\frac{1}{2}\times(0.5\times0.5+0.5\times0.5)+\frac{1}{2}\times0.5\times0.5} \times 3 = 1$. His expected payoff of doing a_3 is fixed at 2. It follows that Romeo weakly prefers to do a_1. When receiving signal 'I love you' and observing $t_{I,r}$, Romeo, similarly weakly prefers to do a_3.

Juliet's expected payoff when sending signal 't_1' in state t_2 upon $t_{II,j}$, equals $0.5\times1+0.5\times(1) = 0$. Her payoff when sending signal 't_1' in state t_2 upon $t_{I,j}$, equals $(0.5)\times1+(0.5)\times(1) = 0$. Juliet's expected payoff when sending signal 't_2' in state t_1 equals 0. It follows that Juliet weakly prefers to send signal 't_1' in state t_2 upon $t_{I,j}$, and signal 't_2' in state t_2 upon $t_{II,j}$. Her payoff of sending signal 't_1' in state t_1 upon cue $t_{I,j}$ or cue $t_{II,j}$ equals $0.5\times3+0.5\times0$. Her payoff of sending 't_2', independent of whether she observes $t_{I,j}$ or $t_{II,j}$, is 1. By checking how often each action gets done in each state, it is easy to see that this correlated equilibrium perfectly replicates the noisy equilibrium with $\mu(t_1|t_2) = 0.25, \mu(0|t_2) = 0.25, \mu(t_2|t_2) = 0.5$, and $\mu(0|t_1) = 0.5, \mu(t_1|t_1) = 0.5$. At the same time, this correlated equilibrium perfectly replicates the mixed equilibrium. In fact, the correlated equilibrium treated here is a limit case where the random events observed by the players are independent. When Romeo has an excellent time, this in fact tells nothing about whether or not Juliet had an excellent time.

It is tempting to infer that there must also be correlated equilibria where Romeo is more likely to have had an excellent time when Juliet had an excellent time. In such a candidate equilibrium, Juliet says that she loves Romeo in state t_1, but she also says this in state t_2 when she had an excellent time ($t_{I,j}$). Because of this, when Romeo receives a signal 't_1' ('I love you') and when he had an excellent time ($t_{I,r}$), Romeo does a_3, as it is then quite possible that Juliet says 'I love you' because she had an excellent time, and not because she loves Romeo. Romeo would only do a_1 when receiving a signal 't_1' ('I love you') if he had merely a nice time ($t_{II,r}$). As it is relatively likely that Juliet also merely had a nice time, the signal 't_1' is now convincing. Juliet must really love Romeo, as it seems she merely had a nice time, and still says 'I love you'. Yet, it can be checked that such an equilibrium cannot exist. For Juliet to tell 'I love you' in state t_2 when having had an excellent time ($t_{I,j}$), it must be that she expects it to be relatively likely that Romeo had a nice time, and will thus do a_1. But if the events observed by Romeo and Juliet are correlated, Romeo will on the contrary be more likely to have had an excellent time as well. For this reason, if each player can observe only two realizations of a random event, then the only correlated equilibrium is an equilibrium that replicates the mixed equilibrium, and is a limit case where the two players random events, and with them their strategies, are fully independent.

At the same time, the analysis of noisy equilibria shows that additional correlated equilibria, where their strategies are truly correlated, must exist. By the so-called revelation principle (Myerson, 1991, Section 6.3), all correlated equilibria of sender-receiver games can be replicated by investigating equilibria where there is a mediator between sender and receiver who possibly garbles the senders messages before sending them on to the receiver.[1] In the presence of such a mediator, an equilibrium is found if the sender honestly reports her state to the mediator, and if the receiver follows the advice of the mediator. The expected payoffs obtained by the players in such a mediated equilibrium can also be obtained in a correlated equilibrium where sender and receiver observe correlated events. The noisy signalling system described in Section 2.1 performs the same role of a mediator. The existence of a strict noisy equilibrium treated in Section 2.1 shows that a strict correlated equilibrium also exists. However, the argument above shows that such an equilibrium takes on quite a complex form, as a random event with two realizations for each player does not suffice.

The interpretation of correlated equilibria in terms of vagueness is the following. Contrary to what is the case in noisy equilibria, the signals used themselves are not noisy. However, the sender uses the signals in a certain context (e.g. Juliet says I love you both if she loves Romeo, and if she does not love him but just had an excellent time), and the receiver only observes an imperfect cue of this context. But it is exactly the asymmetric information about the context in which a signal is used that allow the interests of sender and receiver to be still aligned. Put otherwise, the senders signals are still credible because it is vague in which context they are used. We end this section by noting that it is obvious that not all vagueness can be seen as an attempt to solve conflicts of interest. It can easily be seen that in a sender-receiver game of the type in Table I, but where the interests of sender and receiver fully coincide, the players could not possibly benefit from using a vague rather than a precise language (Lipman, 2006). For this reason, in the rest of this paper we focus on situations where the interests of the sender and the receiver fully coincide.

[1] In general, however, the set of correlated equilibria is larger than the set of noisy equilibria with one-sided communication. For instance, the electronic mail game (Rubinstein, 1989), treated to illustrate noisy signalling in De Jaegher (2003b) is a sender-receiver game where the sender not only sends a signal, but also takes an action. In this case, the set of correlated equilibria can only be approached by means of noisy signalling if the sender and the receiver engage in a conversation of noisy signals. Intuitively, the players can only correlate their actions if the receiver sends noisy confirmations of the sender's messages. For when such mechanisms replicate the set of correlated equilibria, see Forges (1986).

3 Absolute adjectives and adverbs

In this section, we assume that sender and receiver have common interests, but that due to computational limitations, the sender finds it advantageous to think and talk in coarse-grained categories. Thus, just as in Crawford and Sobel (1982), the sender partitions the continuum of states, but now due to computational limitations rather than to a conflict of interest. Consider adjectives like 'full', 'flat', or 'straight'. Just like the meaning of 'tall', also the meaning of these adjectives is vague. These adjectives are also perfectly acceptable in comparatives: there is nothing wrong with saying that one surface is *flatter* than another, or that one bottle is *fuller* than another. In this respect they differ from adjectives like 'pregnant' and 'even', and are on a par with other gradable adjectives like 'tall'. However, as observed by Unger (1975) and also discussed by Rothstein & Winter (2004) and Kennedy & McNally (2005), while with relative adjectives one can easily say something like 'John is tall, but not the tallest/ but somebody else is taller', this cannot be done (so naturally) with (maximal) absolute adjectives. What this contrast shows is that sentences with absolute adjectives generate entailments that sentences with relative adjectives lack: it is inconsistent to say that something is flatter than something that is flat. Thus, from 'The desk is flatter than the pavement.' we conclude that the pavement is not flat.

It seems natural to propose that semantically speaking, an object is flat if and only if it is a *maximal element* with respect to the 'flatter than'-relation: $x \in Flat_{\mathcal{M}}$ iff$_{def}$ $\forall y : x \geq^{\mathcal{M}}_{Flat} y$, i.e., $x \in max(>^{\mathcal{M}}_{F})$. But this straightforward analysis immediately gives rise to a problem: it falsely predicts that an absolute adjective like 'flat' can hardly ever be used. It gives rise to another problem as well: it cannot explain why absolute adjectives combine well with adverbs like 'absolutely', 'completely', and 'hardly'. A proposal that is both compatible with the natural way to give meaning to absolute adjectives, but that can still account for both of these problems was suggested by Lewis (1979): what is a maximal element with respect to a comparative relation like 'flatter than' depends on the *level of fine-grainedness*. Of course, once we look at things from a courser grain, we loose some information. But this need not always be a bad thing, and can even be beneficial. This is, for instance, the case when we reduce some information (the noise) in an audio signal. Hobbs (1985) suggests that thinking and talking of objects in terms of coarser granularities can be advantageous given our computational limitations, i.e., given that we are only *boundedly rational*:

> Our ability to conceptualize the world at different granularities and to switch among these granularities is fundamental to our intelligence and flexibility. It enables us to map the complexities of

> the world around us into simple theories that are computationally tractable to reason in.

Suppose we have two models, $\mathcal{M}$ and $\mathcal{N}$, that each give rise to a comparative ordering 'flatter than': '$>_F^{\mathcal{M}}$' and $>_F^{\mathcal{N}}$'. Assume that the domains of these models are the same, but that this ordering is more fine-grained in model $\mathcal{M}$ than in model $\mathcal{N}$, i.e. $\forall x, y \in I : x >_F^{\mathcal{M}} y \rightarrow x >_F^{\mathcal{N}} y$. In that case, the set of maximal elements of this ordering in $\mathcal{M}$, i.e., the flat objects in $\mathcal{M}$, is a subset of the set of maximal elements of this ordering in $\mathcal{M}$, i.e., the flat objects in $\mathcal{N}$. Let us assume that in fact $max(>_F^{\mathcal{M}}) \subset max(>_F^{\mathcal{N}})$.

Let us now assume that although the denotation of 'flat' depends on the fine-grainedness of the model, it still has an independent 'meaning': a function from a level of fine-grainedness/model to the maximal elements of the 'flatter than'-relation. If we would limit ourselves to the models $\mathcal{M}$ and $\mathcal{N}$, the meaning of 'flat' would be just $\{max(>_F^{\mathcal{M}}), max(>_F^{\mathcal{N}})\}$.

If 'flat' just denotes in each model the maximal elements of the 'flatter than'-relation, modification by means of an adverb like 'absolutely' or 'completely' does not seem to make much sense: 'flat' and 'completely flat' would have the same denotation in each model, and thus they would even have the same 'meaning'. But why, then, would we ever use the adverb? The solution to this problem, we propose, also explains why bare adjectives like 'flat' have a 'vague' meaning.

The explanation is really just the same as Krifka's (2004) proposal for why round numbers are more vague than others. His explanation crucially makes use of the so-called *Horn's division of pragmatic labor*: (un)marked forms get an (un)marked meaning. This principle has recently been given a game-theoretic explanation: if we assume that φ_c and φ have the same meaning, i.e. $[\![\varphi_c]\!] = [\![\varphi]\!] = \{m, m'\}$, but that (i) the marked form, φ_c, is slightly more costly than an unmarked form φ, and (ii) state m is more probable than m', it can be explained (by Pareto optimality, Parikh; evolution, van Rooij 2004 and De Jaegher 2008; forward induction, van Rooij, 2008)[2] why the mapping that associates φ with m and φ_c with m' is pragmatically the most natural equilibrium.

Let us see how this explanation works for our case. First we say that 'flat', 'f', and 'completely flat', 'cf', have the same meaning: $[\![f]\!] = [\![cf]\!] = \{max(>_F^{\mathcal{M}}), max(>_F^{\mathcal{N}})\}$. It seems natural to assume that 'completely flat' is the marked form, because it is longer than 'flat'.[3] The difference in probability

[2] For a recent laboratorium experiment, see De Jaegher, Rosenkranz and Weitzel (2008).

[3] Of course, for this explanation to go through, we have to assume that 'flat' is not in competition with another costly expression with the same meaning like 'flat, roughly speaking'.

of the two elements in $[\![f]\!]$ naturally follows from the fact that $max(>^{\mathcal{M}}_F) \subset max(>^{\mathcal{N}}_F)$. But that is all we need: the marked expression 'completely flat' gets the more precise but also the more unlikely meaning $max(>^{\mathcal{M}}_F)$, while the unmarked 'flat' is interpreted as $max(>^{\mathcal{N}}_F)$. As an immediate consequence we also see that the bare adjective 'flat' is thus interpreted pragmatically in a rather vague, or coarse-grained, way.

4 Semi-orders and Bounded Rationality

In this section we will show that the reason why a predicate is vague is closely related with the reason why agents are only boundedly rational (in a certain way). We will see that a particular type of bounded rationality can be modeled in terms of so-called 'semi-orders' and that these orders are also appropriate to model vagueness. In this section we axiomitize such orderings and show how this is related with a pragmatic take on the Sorites paradox, related with the pragmatic approach to vagueness as explained in Section 4.

In the economic theories of individual and collective choice, preference relations are crucial. In the theory of 'revealed preference' it is standard to *derive* a preference relation in terms of how a *rational* agent would choose among different sets of options. Let us define with Arrow (1959) a *choice structure* to be a triple $\langle I, O, C\rangle$, where I is a non-empty set of options, the set O consists of all finite subsets of I, and the choice function C assigns to each finite set of options $o \in O$ a subset of o, $C(o)$. Arrow (1959) stated the following principle of choice (C) and constraints (A1) and (A2) to assure that the choice function behaves in a 'consistent' or 'rational' way:[4]

(C) $\forall o \in O : C(o) \neq \varnothing$.
(A1) If $o \subseteq o'$, then $o \cap C(o') \subseteq C(o)$.
(A2) If $o \subseteq o'$ and $o \cap C(o') \neq \varnothing$, then $C(o) \subseteq C(o')$.

If we say that $x > y$, iff$_{def}$ $x \in C(\{x, y\})$ and $y \notin C(\{x, y\})$, one can easily show that the ordering as defined above gives rise to a *strict weak order*. A structure $\langle I, R\rangle$, with R a binary relation on I, is a strict weak order just in case R is irreflexive (IR), transitive, (TR), and negatively transitive (NTR).

Definition 1.
A <u>strict weak order</u> is a structure $\langle I, R\rangle$, with R a binary relation on I that satisfies (IR), (TR), and (NTR):

[4] Arrow (1959) actually stated (C) and (A): If $o \subseteq o'$ and $o \cap C(o') \neq \varnothing$, then $C(o) = o \cap C(o')$. But, obviously, (A) is the combination of (A1) and (A2).

(IR) $\forall x : \neg R(x,x)$.
(TR) $\forall x,y,z : (R(x,y) \wedge R(y,z)) \rightarrow R(x,z)$.
(NTR) $\forall x,y,z : (\neg R(x,y) \wedge \neg R(y,z)) \rightarrow (\neg R(x,z))$.

If we now define the indifference relation, '$\sim$', as follows: $x \sim y$ iff$_{def}$ neither $x > y$ nor $y > x$, it is clear that '$\sim$' is not only reflexive and symmetric, but also transitive and thus an equivalence relation. It is well-known that in case '$>$' gives rise to a (strict) weak order, it can be represented numerically by a real valued utility function u such that for all $x, y \in I$: $x > y$ iff $u(x) > u(y)$, and $x \sim y$ iff $u(x) = u(y)$.

The transitivity of the 'indifference'-relation defined in terms of the preference relation as induced above is sometimes problematic. A famous example to show this is due to Luce (1956):

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

That the observed indifference relation is non-transitive might be interpreted in two different ways: (i) on a *descriptive approach* one can say that the choice behavior simply does not need to satisfy Arrow's constraints, or (ii) alternatively, one takes the axiom to define *rationality*, and an agent who does not seem to obey them can then at best be only *boundedly rational*. Of course, the non-transitivity of the relation '$\sim$' extends beyond the case of preference: sounds a and b can be judged equally loud, just as sounds b and c, but it need not be the case that sounds a and c are judged equally loud. This non-transitivity is a problem, because it immediately leads to the 'Sorites paradox', the famous problem induced by vague expressions.

Consider a long series of cups of coffee ordered in terms of the grains of sugar in it. Of each of them you are asked whether the coffee is sweet or not. We assume that you are always *indifferent* between two subsequent cups. Now, if you decide that the coffee in the first cup presented to you — the one with 100 grains of sugar — does not taste sweet, i.e., taste bitter, it seems only reasonable to judge the second cup of coffee to be bitter as well, since you are indifferent between the two different cups. But, then, by the same token, the third cup should also be bitter, and so on indefinitely. In particular, this makes the last cup of coffee taste bitter, which is in contradiction with our intuition that the coffee in this last cup is sweet.

This so-called Sorites reasoning is elementary, based only on our intuition that the coffee in the first cup is not sweet, the last one is, and the following inductive premise, which seems unobjectable:

> [**P**] If you call one cup of coffee bitter, and this cup is indistinguishable from another cup, you have to call the other cup bitter too. Thus, for any $x, y \in I : (P(x) \wedge x \sim_P y) \rightarrow P(y)$.

If we assume that it is possible that $\exists x_1, \ldots, x_n : x_1 \sim_P x_2 \wedge \cdots \wedge x_{n-1} \sim_P x_n$, but $P(x_1)$ and $\neg P(x_n)$, the paradox will arise. Strict weak orders, however, do not allow for the possibility that $\exists x_1, \ldots, x_n : x_1 \sim_P x_2 \wedge \cdots \wedge x_{n-1} \sim_P x_n$, but $P(x_1)$ and $\neg P(x_n)$. Fortunately, there is a well-known ordering with this property. According to this ordering the statement '$x \succ_P y$' means that x is *significantly* or *noticeably* greater than y, and the relation '$\succ_P$' is irreflexive and transitive, but need not be almost connected. The indistinguishability relation '$\sim_P$' is reflexive and symmetric, but need not be transitive. Thus, '$\sim_P$' does not give rise to an equivalence relation. The structure that results is what Luce (1956) calls a *semi-order*. A structure $\langle I, R \rangle$, with R a binary relation on I, is a semi-order just in case R is irreflexive (IR), satisfies the interval-order (IO) condition, and is semitransitive (STr)

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

As we noted above, semi-orders can be understood as models of bounded rationality or limited perceptive ability. Just as weak orders, also semi-orders can be given a measure theoretical interpretation. If P is the predicate 'bitter', one can think of R as the semi-order relation (significantly) 'more bitter than', and say that '$x \succ_P y$' is true iff the bitterness of x is higher than the bitterness of y plus some fixed (small) real number ϵ. In the same way '$x \sim_P y$' is true if the difference in bitterness between x and y is less than ϵ. In case $\epsilon = 0$, the semi-order is a strict weak order.

Semi-orders are appropriate to *represent* vagueness, but by themselves do not point to a solution to the Sorites paradox at all. The standard reaction to the Sorites paradox taken by proponents of fuzzy logic and/or supervaluation theory is to say that the argument is valid, but that the inductive premise [**P**] (or one of its instantiations) is *false*. But why, then, does it at least *seem* to us that the inductive premise is true? According to the standard accounts of vagueness making use of fuzzy logic and supervaluation theory, this is so because the inductive premise is *almost* true (in fuzzy logic), true in almost all complete valuations (in supervaluation theory), or that almost all its instantiations are true.

Most proponents of the contextuallist solution follow Kamp (1981) in trying to preserve (most of) [**P**] by giving up some standard logical assumptions, and by making use of a mechanism of *context change*. But we do not believe that context change is essential to save natural language from the Sorites paradox. Our preferred 'solution' is radically *pragmatic* in nature and completely in line with Wittgenstein's *Philosofische Untersuchungen*.[5] The first observation is that the meaning of vague adjectives like *small* and *tall* are crucially *context dependent*: for Jumbo to be *a small elephant* means that Jumbo is being small for an elephant, but that does not mean that Jumbo is small. For instance, Jumbo will be much bigger than any object that counts as *a big mouse*. One way to make this explicit is to assume with Klein (1980) that every adjective should be interpreted with respect to a *comparison class*, i.e. a set of individuals. The truth of a sentence like *John is tall* depends on the contextually given comparison class: it is true in context (or comparison class) o iff John is counted as tall in this class. The idea of the radically pragmatic solution to the Sorites paradox is that it only makes sense to use a predicate P in a context – i.e. with respect to a comparison class –, if it helps to clearly demarcate the set of individuals within that comparison class that have property P from those that do not. Following Gaifman (1997), we will implement this idea by assuming that any subset of I can only be an element of the set of *pragmatically appropriate* comparison classes O just in case the gap between the last individual(s) that have property P and the first that do(es) not must be between individuals x and y such that x is clearly P-er than y. This is not the case if the graph of the relation '$\sim_P$' is closed in $o \times o$.[6] Indeed, it is exactly in those cases that the Sorites paradox arises.

How does such a proposal deal with the Sorites paradox? Well, it claims that *in all contexts o in which P can be used appropriately*, [$\mathbf{P}_1$] is true, where [$\mathbf{P}_1$] is $\forall x, y, o : (P(x, o) \wedge x \sim_P y) \rightarrow P(y, o))$. If we assume in addition that the first element x_1 of a Sorites series is the absolute most P-individual, and the last element x_n the absolute least P-individual, it also claims that *in all contexts o in which it is appropriate to use predicate P in combination with* x_1 *and* x_n, '$P(x_1, o)$' is *true* and '$P(x_n, o)$' is *false*.[7] Thus, in all *appropriate* contexts, the premises of the Sorites argument are considered to be true. Still, no contradiction can be derived, because using predicate P when *explicitly*

[5]See in particular sections 85-87.

[6]Notice that also in discrete cases the relation '$\sim_P$' can be closed in $o \times o$. In just depends on how '$\sim_P$' is defined.

[7]But don't we also feel that in case o is an inappropriate context, the first element should still be called a P-individual, and the last one a $\neg P$-individual? According to this account, this intuition is 'accounted' for by noting that this indeed comes out in all appropriate subsets of o.

confronted with a set of objects that form a Sorites series is *inappropriate*. Thus, in contrast to the original contextualist approaches of Kamp (1981), Pinkal (1984), and others, the Sorites paradox is not avoided by assuming that the meaning (or extension) of the predicate changes as the discourse proceeds. Rather, the Sorites paradox is avoided by claiming that the use of predicate P is inappropriate when confronted with a Sorites series of objects.

Our pragmatic solution assumes that the appropriate use of a predicate P must clearly *partition* the relevant comparison class. Moreover, we want 'P-er than' to generate a semi-order rather than a strict weak order. In the following we will suggest how to generate semi-orders in terms of the way predicates partition *appropriate* comparison classes.

Let us start just like Arrow (1959) with a *choice structure*: a triple $\langle I, O, P\rangle$, where I is a non-empty set of options, the set O of comparison classes, and for each predicate a choice function P assigns to each $o \in O$ a subset of o, $P(o)$. We will adopt Arrow's condition (C) and his constraints (A1) and (A2). If we say with Klein (1980) that that $x >_P y$, iff there is an $o \in O$ such that $_{def}$ $x \in P(o)$ and $y \in 0 - P(o)$, it is easy to see that the relation will be a strict weak order *if O consists of all subsets of I* (of cardinality 2 and 3). The idea how to generate a semi-order '$\succ_P$' is to assume that O does *not* consists of *all* these subsets. Instead, we just start with all subsets of I that consist of two elements that are not indifferent and close this set of subsets of O under the following closure conditions:

(P1) $\forall o \in O : \forall x \in \bigcup O : o \cup \{x\} \in GAP \rightarrow o \cup \{x\} \in O$.
with $o^n \in GAP \quad \text{iff}_{def} \quad \exists_{n-1} o' \subset o : card(o') = n - 1 \wedge o' \in O$.
(OR) $\forall o \in O, \{x, y\} \in O : o \cup \{x\} \in O$ or $o \cup \{y\} \in O$.
(P2) $\forall o \in O, x \in I : o \in GAP_2 \rightarrow o \cup \{x\} \in O$.
with $o^n \in GAP_2 \quad \text{iff}_{def} \quad \exists_n o' \subset o : card(o') = n - 1 \wedge o' \in O$.

Constraint (P1) says that to any element o of O one can add any element $x \in I$ to it that is in an ordering relation with respect to at least one other element, on the condition that $o \cup \{x\}$ satisfies the GAP-condition. The intuition behind this condition is that only those subsets of I satisfy GAP if there is at least one gap in this subset w.r.t. the relevant property. It is easy to show that (P1) guarantees that the resulting ordering relation will satisfy transitivity and will thus be a strict partial order.[8] If constraint (OR) is added, the resulting ordering relation also satisfies the interval ordering condition. Constraint (P2), finally, guarantees that the ordering relation also satisfies semi-transitivity, and thus is a semi-order. As far as we can see, the closure conditions stated above

[8]The proof of this result, and the others mentioned below, are all closely related with similar proofs for slightly different choice structured discussed in Van Rooij (2009).

are such that they generate *all and only all appropriate contexts*, i.e., just those subsets of I for which there is such a sufficiently large gap such that also vague predicates can clearly partition the context without giving rise to the Sorites paradox. In particular, O does not necessarily contain all subsets of O. This is essential, because otherwise the resulting ordering relation would also satisfy (NTR) which is equivalent to $\forall x, y, z : x \succ_P y \rightarrow (x \succ_P z \vee z \succ_P y)$. If that were the case, the resulting ordering relation would not only be a semi-order, but also a weak order, which is what we do not want. It suffices to observe that no constraint formulated above forces us to assume that $\{x, y, z\} \in O$ if $x \succ_P y$, which is all that we need.

5 A behavioural approach to vagueness: prospect theory

In the previous sections, context was each time used to account for vagueness. Yet, context did not directly enter the agent's valuations. We now treat a model where context as seen by the agent, or the agent's reference point, directly enters his or her valuation function. Let x be the temperature of a room in degrees Celsius. An agent has a valuation function $v(x)$, with $v'(x) > 0$. If $v \geq V$, the agent considers the room as warm. If $v < V$, she considers it as cold. The problem with such a simple model is that it is highly implausible that there exists such a context-independent cut-off point. This would for instance mean that our agent considers any room of 18^o as cold, whether she has just walked into this room from freezing temperatures outside, or from another room heated at 21^o.

Let us try to extend our model such that the meaning of 'warm' and 'cold' becomes context-dependent. Consider the set of all possible alternatives (= in alternative room temperatures) open to agent, denoted as I, where x denotes a typical alternative in I. Let an agent currently compare the alternatives in a set o, with $o \subseteq I$. Let $|o|$ be the number of elements in o. Define the average value of the alternatives in set o as $\frac{1}{|o|}\sum_{x\in o} v(x)$. Define now as the agent's psychic valuation the value:

$$u(x_i, o) \quad =_{def} \quad v(x_i) + \alpha[v(x_i) - \frac{1}{|o|}\sum_{x\in o} v(x)]$$

The coefficient $\alpha \geq 0$ measures the extent to which the agent is subject to the context in his valuation of x_i. For $\alpha = 0$, we have a standard utility function, leading to a context-independent cut-off point. For α approaching infininity, we have an agent who strictly thinks in gains and losses with respect to a value reference point, which in this case takes on the value $\frac{1}{|o|}\sum_{x\in o} v(x)$.

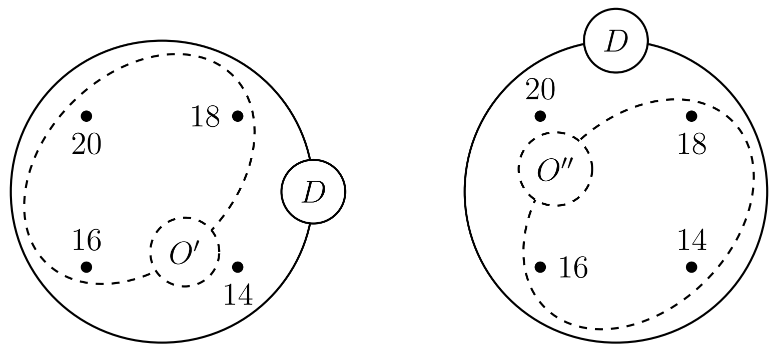

Figure 1: Temperatures and reference points.

The latter is a simplified version of the psychic valuation described by Kahneman and Tversky (1979) in their prospect theory. Assume now that if $u \geq U$, the room is considered as warm. If $u < U$, it is considered cold. Any α with $0 < \alpha < 1$ now suffices to account for a concept 'warm' that has no clear context independent boundaries.

To see why, consider the example in Figure 1. The choice sets under consideration of the agent are first the subset o', and second the subset o'' of I. The numbers are real temperatures measured in Celsius. Let $v(x) = x$, thus $v(18^o) = 18^o$. Consider u as how warm or cold the agent subjectively feels the real temperature to be. Let $u \geq 19^o$ be warm, and let $u < 19^o$ be cold. Let $\alpha = 0.5$. Consider first $u(18^o, o')$. As the average real temperature in this set of alternatives is 18^o, $u(18^o, o') = 18$, and the agent considers 18^o to be cold. The agent may be seen as recently having been in rooms heated at 16^o, 18^o and 20^o. A room heated at 18^o is then felt to be a cold room. Next, consider $u(18^0, o'')$. Here, the average actual temperature in u'' is 16^o. As this average temperature is rather low, $u(18^o, o'') = 19^o$. Thus, when the alternatives under consideration are the ones in o'', 18^o feels like 19^o, and is considered as warm. The agent now may be seen as having recently walked between outside temperatures of 14^o and 16^o, and a room heated at 18^o. The room heated at 18^o now feels warm. These subjective feelings may also be seen as a form of satisficing, where satisficing behaviour depends on one's reference level. When the choice is between 14^o, 16^o and 18^o, 18^o will do. When the choice is between $16^o, 18^o$ and $20^o, 18^o$ is considered as unsatisfactory.

Prospect theory can thus help us to account for the fuzzy boundaries between concepts such as warm and cold. The reason why these boundaries are fuzzy may be that the ad hoc boundary depends on the decision maker's reference point. Consider then a sender and receiver with common interests, in that the receiver would genuinely like to know the extent to which a sender feels something to be warm. Then temperatures will not do, because depend-

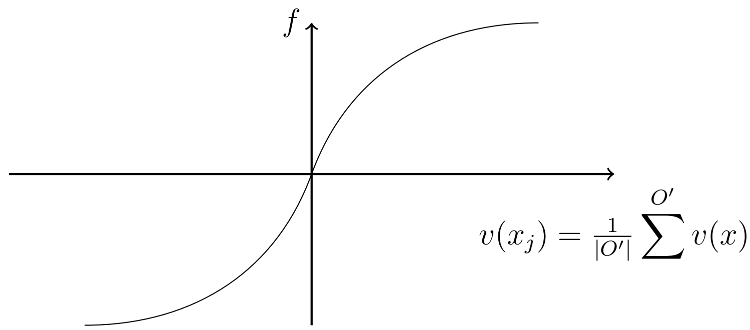

Figure 2: Prospect theory psychic valuation function.

ing on the sender's reference point, he may feel 18^o to be warm or cold. Fuzzy language instead is able to convey the extent to which the sender feels warm.

There is a further aspect of vagueness reflected in the Sorites paradox, namely that small differences (e.g. in temperature) do not matter. This has been theoretically argued by Luce (1956), and has also been observed in psychological experiments. What could prospect theory say about small differences? It seems intuitive that the extent to which small differences are perceived again depends on the agents' reference point. The difference between 18^o and 19^o may seem whimsical to an agent who has just experienced 0^o. But this difference may not seem whimsical to an agent who has only experienced $18^o, 19^o$ and 20^o. Unfortunately, the reference-dependent utility function above cannot account for such an effect. To see this, consider two alternatives x_i and x_j, where first these alternatives are considered within a set o', and next they are considered within a set o''. It is clear that $u(x_i, o') - u(x_j, o') = u(x_i, o'') - u(x_j, o'')$. Because of the linear structure of the reference-dependent utility function, the reference, measured by the average in the set under consideration, cancels out.

Let us instead assume that the consumer does not directly value the difference between v and the reference point (the average of the v's under consideration), but instead values a function f of this difference:

$$u(x_i, o) \quad =_{def} \quad v(x_i) + \alpha f[v(x_i) - \frac{1}{|o|} \sum_{x \in o} v(x)]$$

and let $u(x_i) \approx u(x_j)$ if $|u(x_i)u(x_j)| < \epsilon$. In line with prospect theory, where agents have smaller marginal valuations the further away from the reference point, f is assumed to have an s-shaped form, as represented in Figure 2. The origin in Figure 2 is determined by the reference point. Gains with respect to the reference point are measured above the horizontal axis, losses are measured below the horizontal axis.

As an example, let $\alpha = 0.5$, and let $f(\cdot)$ be equal to $+\surd$ if $v(x)$ is larger than the reference, and equal to $-\surd$ if $v(x)$ is lower. Assume again that $v(x) = x$. Let $\epsilon = 1.2$. Consider now first $u(17^o, o') - u(16^o, o')$ when $o' = \{0^o, 16^o, 17^o\}$. Then $u(17^o, o') - u(16, o') = 1 + \frac{1}{2} \times (6^{0.5} - 5^{0.5}) = 1.11 < 1.2$. When compared to 0^o, the difference between 16^o and 17^o is considered whimsical by the agent. Next, consider $u(17^o, o'') - u(16, o'')$ when $o'' = \{16, 17, 18\}$. Then $u(17^o, o'') - u(16^o, o'') = 1 + \frac{1}{2} \times (1^{0.5}) = 1.5 > 1.2$. When compared to 18^o, the difference between 16^o and 17^o is considered as significant.

The relevance to the Sorites paradox of the point that whether or not small differences matter depends on one's reference point, is the following. When asked to consider a series of Sorites arguments, saying that 0.1^o does not matter, we are contemplating the full range of cups of coffees, from very cold ones to very hot ones. Given this reference point, 0.1^o of difference indeed does not matter. However, when we are making an actual decision on whether a cup of coffee is warm of cold, our reference point will not be the whole range of coffees, but will probably be a small range of coffee temperatures, such as they are produced by our coffee machine. Given such a reference point, small differences may still matter. Thus, the Sorites paradox may arise because the paradox induces us to take e.g. a clearly cold cup of coffee as a reference point, and then to think about the difference between hot and slightly less hot cups of coffee. However, when facing actual cups of coffee, our reference point will never differ that widely from our actual experience, so that small differences may still matter.

6 Conclusion

In Section 2, we have started this paper by reviewing game-theoretic rationales for vagueness (and the related concepts of ambiguity and generality), which see vagueness as solving conflicts of interest between speaker and listener. We conclude from this review that it is difficult to link these game-theoretic rationales for vagueness to concrete instances of natural language. Also, it does not seem plausible that all instances of vagueness would be linked to conflicts of interest. Given these conclusions, two areas for further research arise, and have been treated in this paper. In one approach (Sections 3 and 4) one continues to assume, just as in game theory, that agents are rational, even though they find it difficult to distinguish between different states of the nature. Pragmatically, in contexts where they find this most difficult, it makes most sense to make use of a vague language. Section 5 treats a behavioral approach where contexts, in the form of reference points, directly enter the agents' preferences. Laboratory experiments have shown that agents form such reference points. Analyzing language use under such reference-dependence is an open field for future research.

References

[30] Arrow, K. (1959), 'Rational choice functions and orderings', *Economica*, **26**: 121-127.

[30] Aumann, R. (1974), 'Subjectivity and correlation in randomized strategies'. *Journal of Mathematical Economics*, **1**: 67-96.

[30] Benthem, J. van (1982), 'Later than late: on the logical origin of the temporal order', *Pacific Philosophical Quarterly* **63**: 193-203.

[30] Blume, A., Board, O.J. and Kawamura, K. (2007), 'Noisy talk'. *Theoretical Economics*, **2**: 395-440.

[30] Blume, A. and Board, O. (2008), 'Intentional vagueness', working paper, University of Pittsburgh.

[30] Crawford, V.P. and Sobel, J. (1982), 'Strategic information transmission', *Econometrica* **50**: 1431-1451.

[30] De Jaegher, K. (2003a), 'Error-proneness as a handicap signal', *Journal of Theoretical Biology*, **224**: 139-152.

[30] De Jaegher, K. (2003b), 'A game-theoretic rationale for vagueness', *Linguistics and Philosophy*, **26**: 637-659.

[30] De Jaegher, K. (2008), 'The evolution of Horns rule', *Journal of Economic Methodology*, **15**: 275-284.

[30] De Jaegher, K., Rosenkranz, S. and Weitzel, U. (2008), 'Economic laboratory experiment on Horns rule', working paper, Utrecht University.

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

[30] Farrell, J. (1993), 'Meaning and credibility in cheap-talk games', *Games and Economic Behavior*, **5**: 514-531.

[30] Forges, F.M. (1986), 'An approach to communication equilibria', *Econometrica*, **54**, 1375-1385.

[30] Fudenberg, D. and Tirole, J. (1991), *Game Theory*, MIT Press, Cambridge Ms.

[30] Gaifman, H. (1997), ‘Vagueness, tolerance and contextual logic’, manuscript, Columbia University.

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

[30] Harsanyi, J.C. (1973), ‘Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points’, *International Journal of Game Theory*, **2**: 1-23.

[30] Kahneman, D. and Tversky, A. (1979), ‘Prospect theory: an analysis of decision under risk’, *Econometrica*, **47**: 263-291.

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

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

[30] Keefe, R. (2007), ‘Vagueness without context change’, *Mind*, **116**: 462.

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

[30] Krantz, D., R. Duncan Luce, P. Suppes, and T. Tversky (1971), *Foundations of Measurement, Vol. 1*, Academic Press, New York and London.

[30] Lipman, B.L. (2006), ‘Why is language vague?’, working paper, Boston University.

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

[30] Myerson, R.B. (1991), *Game Theory – Analysis of Conflict*, Harvard University Press, Cambridge, Ms.

[30] Pinkal, M. (1984), ‘Consistency and context change: The Sorites paradox’, in F. Landman & F. Veltman (eds.), *Varieties of Formal Semantics, Proceedings of the fourth Amsterdam Colloquium*, Groningen-Amsterdam Studies in Semantics, Foris Publications, Dordrecht.

[30] Raffman, D. (1994), ‘Vagueness without paradox’, *The Philosophical Review*, **103**, 41-74.

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

[30] Rooij, R. van (2009), ‘Revealed preference and satisficing behavior’, in G. Bonano et. al. (eds.), Proceedings of LOFT 2008, Amsterdam University Press.