

## Vagueness, Interpretation, and Probabilistic Belief

Daniel Lassiter (NYU Linguistics)

In many areas of natural language semantics it is useful to idealize away from the complexities of interpretation in context. In investigating quantification, for example, we can learn a lot about the meaning of the phrase *most dogs have tails* without resolving the precise denotation of *dog* or *tail*: for these purposes, it is usually enough to say that *dog* denotes the characteristic function of the set of dogs, etc. As Lewis (1969, 1970) claimed, vague terms are one place where this idealization breaks down: we cannot understand the phenomenon of vagueness without paying close attention to the process by which speakers interpret each others' utterances in concrete contexts. Interpreters, being epistemically limited agents, can never know precisely what their interlocutors' utterances are intended to communicate; as a result, even if they have access to languages as precise as our logical languages, it is not possible for an interpreter to settle on a single such language as the current language of conversation.

Lewis (1970) suggests that we can reconcile the precision of formal syntax and semantics with the uncertainty of communication in context if we allow that "languages themselves are free of vagueness but ... the linguistic conventions of a population, or the linguistic habits of a person, select not a point but a fuzzy region in the space of precise languages". The approach to vagueness presented here is essentially a Bayesian implementation of Lewis' idea. Bayesianism has been employed with great success in cognitive science, and recently probabilistic theories of language acquisition, representation, and processing have also been explored (see in particular Yang 2002 and the papers in Bod et al. 2003). I argue here that a Stalnakerian picture of linguistic knowledge together with probabilistic modal logic yields a novel account of vagueness.

Stalnaker (1978) gives a theory of assertion in which the role of an assertion is to eliminate certain possibilities from the common ground, construed as a set of worlds considered by the conversational participants as live possibilities for how the actual world might be. This theory is not restricted to update of non-linguistic beliefs: among the assumptions that speakers bring to a conversation are beliefs about the common ground, which can also change in the course of a conversation. In Stalnaker's example, the assertion *An optometrist is an eye doctor* has the effect of eliminating worlds from the common ground in which the current language of conversation does not map *optometrist* and *eye doctor* to the same extension.

Probabilistic variants of the possible-worlds model of belief are motivated by the desire to account for the graded nature of belief; Fagin and Halpern (1994) and Halpern (1997) show that these are logically well-behaved. In this model a speaker's belief-set is a set of pairs of a world and a real number. For example, an individual A's belief-set might look like this:

$$B_A = \{ \langle w_1, .1 \rangle, \langle w_2, .4 \rangle, \langle w_3, .3 \rangle, \langle w_4, .2 \rangle \}$$

This model becomes relevant to vagueness when we realize that, if we follow Stalnaker in supposing that linguistic and non-linguistic beliefs are represented in similar fashion, we can model uncertain *linguistic* belief using the same technology. Just as a possible world is supposed to be a perfectly precise, fully specified state of affairs, in each possible world there is a unique, perfectly precise language that is being spoken. (This latter assumption is not actually necessary, but it simplifies the model and will be adopted here for this reason.) Just as two worlds might be identical in every detail except for the value of a single sentence, so two languages can be identical in every detail except for the value of a single term: for instance, whether an object *x* counts as being in the extension of a particular term of the language.

This approach allows us to treat vagueness as linguistic uncertainty, as Lewis suggested – though not as a 'fuzzy region', but as a *probability distribution* over the space of precise languages. If a speaker A utters *U*, for all that an interpreter B can know, *U* could in principle designate an unlimited number of distinct properties  $P_1 \dots P_n$  each of which divides the continuum in a different way. Each of these properties can be the designatum of *U* in a range of

distinct possible languages, which can be as similar as you like except for the value of  $U$ . So there are languages in which the dividing line between ‘tall’ and ‘not tall’ is ‘6’0”, ‘6’1”, ‘6’1.1”, etc.

The defining characteristic of vague terms on this approach is that when a term is vague there is no point at which a possible language  $L_1$  yields a plausible interpretation of “tall”, and a neighboring possible language  $L_2$  which resolves “tall” in a similar fashion yields an implausible interpretation. Rather, the plausibility of neighboring interpretive theories declines gradually. In terms of a probabilistic model, vague terms like “tall” are characterized by the fact that the probability distribution over languages that resolve them in similar ways is a continuous function.

One virtue of this approach is that it yields a novel account of the sorites paradox. Detailed examination of the paradox would require more space than we have here, but I will attempt to present the basic picture. A probabilistic semantics does not assign truth-values to utterances, but only to utterances *relative to an interpretation*. The standard formulation of the sorites paradox is valid only relative to particular (precise) languages; within these strictures, the inductive premise of the paradox “ $\neg \exists x \exists y [Rxy \ \& \ \neg \text{tall}(x) \ \& \ \text{tall}(y)]$ ” cannot help but be false, as it must within any precise language. Under this reading, the paradox does not arise.

On the other hand, when we refer to utterances which admit of many plausible interpretations, there is no sense in speaking of objects being *tall* (say). Within a probabilistic theory, utterances themselves cannot take objects as arguments. However, we can speak of the probability of an object’s “counting as” *tall*, defined as:

The probability that an object  $o$  counts as an instance of  $U$  is the sum probability of all languages  $L$  which satisfy this condition: The value of  $U$  in  $L$  applied to  $o$  yields 1.

In other words, we take all precise languages that assign to  $U$  a set which contains  $o$  and sum the probability associated with them in the agent’s belief-set. This is the probability that  $o$  counts as  $U$ . Using this definition, the inductive premise can be reinterpreted as follows:

$\neg \exists x \exists y [Rxy \ \& \ p(\text{tall}(x))=0 \ \& \ p(\text{tall}(y))=1]$

This variant of the inductive premise is not paradoxical: it simply says that the probability function does not jump from 0 to 1 at any point along a sorites series, which is presumably a necessary condition of a term’s being vague at all. This version of the inductive premise is compatible with the value of a continuous function increasing gradually from zero to one.

There are many issues left to address, of course; my goal here is simply to make a first case that a probabilistic approach to vagueness is promising, and that it has considerable conceptual and philosophical motivation, as well as close connection to well-explored areas of Bayesian cognitive science and probability logic. A tantalizing area for future research, for example, is the possibility that the probabilistic theory may shed light on the pragmatics of vague terms (Fara 2000, Barker 2002, Kyburg and Morreau 2000). This context-dependence may turn out to be due not to any special semantic mechanisms as these authors claim, but simply run-of-the-mill Bayesian updating. At present, however, the primary attraction of the probabilistic theory is that an apparently well-behaved theory of the semantics of vague terms can be derived entirely from two independently motivated components: Stalnaker’s theory of metalinguistic assertion, and Fagin and Halpern’s logic for uncertain belief and inference.

**References** ♦ **Barker, Chris**. 2002. The dynamics of vagueness. *L&P* 25: 1-36. ♦ **Bod, Rens, et al.** (eds.). 2003. *Probabilistic Linguistics*. MIT Press. ♦ **Fagin, Ronald, and Joseph Y. Halpern**. 1994. Reasoning about knowledge and probability. *Journal of the Association for Computing Machinery* 41(2). ♦ **Fara, Delia**. 2000. Shifting sands: an interest-relative theory of vagueness. *Philosophical Topics* 28(1). ♦ **Halpern, Joseph**. 1997. A logical approach to reasoning about uncertainty. Arrozala et al. eds., *Discourse, Interaction, and Communication*. Kluwer. ♦ **Kyburg, Alice, and Michael Morreau**. 2000. Fitting words: Vague language in context. *L&P* 23(6). ♦ **Lewis, David**. 1969. *Convention*. HUP. ♦ **Lewis, David**. 1972. General semantics. Davidson and Harman eds., *Semantics of Natural Language*. Reidel. ♦ **Lewis, David**. 1973.

*Counterfactuals*. HUP. ♦ **Stalnaker, Robert**. 1978. Assertion. Cole ed., *Syntax and Semantics 9*. Academic Press. ♦ **Yang, Charles**. 2002. *Knowledge and Learning in Natural Language*. OUP.