

Plan for today:

1. Comparatives.
2. Epistemic contradictions.
3. Truth and content.

## 1 Unifying probability talk with a general account of comparatives

### 1.1 A really obvious point

First appearances suggest that **probability operators** (*is as likely as*, *probably*, *is more likely than*, etc) should be treated together with what we might call **pure modal** talk (*might*, *may*, *ought*, *must*, *possibly*, *necessarily*, etc) because:

- They are non-truth-functional operators on **propositions**.
- They interact logically with epistemic modals:

R3. *Must to probably*.

$$\frac{\text{It must be that } \phi}{\text{Probably } \phi}$$

R4. *Probably to might*.

$$\frac{\text{Probably } \phi}{\text{It might be that/It is possible that } \phi}$$

You may, however, have noticed that:

- Comparative forms are not as easy/natural with straight modal talk:  
*? is as possible as*, *? is as necessary as*. Hopeless with modal auxiliaries.
- The various comparative probability operators (*is as likely as*, *is (much) more/less likely than*, etc) obviously have a bunch of a further decomposable structure. The range of morphology involved in this further structure occurs in comparative discourse generally (this is what Kennedy [2006] calls **degree morphology**):

**comparative morphemes:** more, less, as

**intensifiers:** very, quite, rather

**sufficiency morphemes:** too, enough, so

and obviously, all this morphology is playing the same semantic role.

That last observation tells us very clearly that **our semantics for probability operators is really going to have to play nice with the true semantics for degree morphology**.

And you might think *that* recommends the following methodological disposition for doing semantics for probability operators:

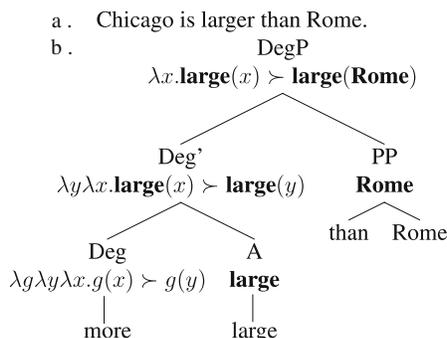
Seek the gradable adjectives literature for help in doing the semantics for probability operators.

To some extent this is Portner [2009]’s attitude. (Also my attitude at SuB in 2006.)

## 1.2 Kennedy 2007

Portner [2009] on probability operators is basically applied Kennedy [2006]. Some of the basic features of Kennedy’s picture:

1. We assume things called **dimensions**. *Expensive* and *cheap* both involve the dimension **cost**; *tall* and *short* both involve the dimension **height**, etc. Dimensions come in (totally orderable) magnitudes measurable by **degrees**. A dimension and its corresponding total order of degrees is a **scale**.
2. Sentences with gradable adjectives (*expensive*, *tall*... most other adjectives!) **express relations between degrees**. Rough pattern of analysis:  
‘The coffee in Rome is expensive’ = 1 iff the degree of cost possessed by coffee in Rome is greater than a contextually determined standard of cost.
3. From a conceptual point of view it is the comparative idea (*is more expensive than*) that is fundamental. Basically all constructions with gradable adjectives express a comparative meaning, if not overtly. The big challenge in the literature is to say exactly how the covert comparatives—the superficially simplest constructions like ‘John is tall’—work in detail.
4. Kennedy’s proposal about this—one among many out there!—is that gradable adjectives express functions from objects to degrees on the appropriate scale (objects of type  $\langle e, d \rangle$ ): they map objects to the degree to which they have the property in question. Easier to see this motivation for this approach in the context of an analysis of a full comparative construction:



Kennedy’s basic project in the paper is to say what ‘Chicago is large’ means (the **unmarked positive case**), given this semantic analysis of the comparative.

5. We want ‘is large’ in ‘Chicago is large’ to give us something of type  $\langle e, t \rangle$ , but the semantics of ‘large’ gives us something of type  $\langle e, d \rangle$ , so we need something that will operate on the latter to give us the former. One idea is a general type-shifting rule. Kennedy prefers hypothesizing a null morpheme in the LF, what he calls **pos**. Here’s his first pass semantics:

$$\llbracket \text{pos} \rrbracket^c = \lambda g. \lambda x. g(x) > d_c$$

where  $d_c$  is the appropriate degree of comparison at  $c$ . Ultimately he prefers this:

$$\llbracket \text{pos} \rrbracket^c = \lambda g. \lambda x. g(x) \succeq s_c(g)$$

with  $s$  “a context-sensitive function that chooses a standard of comparison in such a way as to ensure that the objects that the positive form is true of ‘stand out’ in the context of utterance, relative to the kind of measurement that the adjective encodes” (17).

### 1.3 Flatfooted implementation of Kennedy for probability operators

Probability operators are equipped with a scale  $\langle [0,1], \geq, \text{probability} \rangle$ . Give the same semantics as Kennedy on adjectives, but switch out individuals for propositions:

$$\begin{aligned} \llbracket \text{probable} \rrbracket &= \lambda p. Pr(p) \\ \llbracket \text{pos} \rrbracket^c &= \lambda g. \lambda p. g(p) \succeq s_c(g) \end{aligned}$$

Basically, *probable* literally denotes a probability function mapping propositions into their probabilities in the closed interval  $[0,1]$ . We might conjecture the morpheme *-ly* express *pos* above. If so we'd have:

$$\llbracket \text{probably } \phi \rrbracket^c = 1 \text{ iff } Pr(\phi) \geq s_c(Pr)$$

So this would be a probabilistic alternative to the picture of probability operators in Kratzer [1991], which does not use quantitative probabilities but a mere partial order.

Now we might even try to drag the pure epistemic modal down this road. Kennedy tries to build an empirical case for the idea that some gradable adjectives are **absolute**, serving to characterize the degree to which an object has a property by reference to the upper and lower bounds of the relevant scale.

Two varieties:

**1. Minimum Standard:** adjectives which compare the degree of  $x$  to the lower bound of the scale.

The door is open.  
The rod is bent.

Only non-zero door aperture/degree of bend required for truth.

$$\llbracket \text{bent} \rrbracket = \lambda x. \text{bend}(x) > \min(\text{bend})$$

**2. Maximum Standard:** adjectives which compare the degree of  $x$  to the upper bound of the scale.

The door is closed.  
The rod is straight.

Requires completely closed door/completely straight rod.

$$\llbracket \text{straight} \rrbracket = \lambda x. \text{straight}(x) = \max(\text{straight})$$

Now you might go crazy and propose this:

$$\begin{aligned} \llbracket \text{possibly} \rrbracket &= \lambda p. Pr(p) > \min(Pr) \\ \llbracket \text{must} \rrbracket &= \lambda p. Pr(p) = \max(Pr) \end{aligned}$$

## 1.4 Portner

Portner thinks a scale is necessary for *probable* (and *possible*), but he does not favor an analysis of *probable* (or pure epistemic modals) based on numerical probabilities. He makes a few points:

1. *Probable* and *possible* take different modifiers. He thinks this predicts that they need different scales.

Comment: Perhaps, but that doesn't mean probability is not implicated in the scales appropriate to both operators.

2. The interval on our scale is closed, "so we expect a degree specification like *completely* to be acceptable". But:

# It's completely probable that it's raining.

In a footnote (76, fn.15) Portner considers the reply that the scale associated with *probable* uses just the degrees in  $(0,1)$ . He says that this would incorrectly saddle *Probably*  $\phi$  with the presupposition *It is not certain that*  $\phi$ , and cites data that is supposed to support this point. The relation of the data he cites to his point is somewhat mysterious; the data he seems to want is:

A: Probably it's raining.

B: Yes. In fact it's certain that it's raining.

Comments:

Kennedy: "not all modifiers co-occur with all adjectives for apparently idiosyncratic reasons" (34). E.g., the dimension associated with cost has a minimal element (\$0); but it's bad to say, of a free t-shirt, that it is completely inexpensive.

Second, it would be technically trivial to extend the range of a probability measure on the closed interval  $[0,1]$  to one on  $[0,1+r)$  for  $r$  some positive real. The extra values above 1 just go unused. (Not an ideal solution, however.)

3. While *possibly* can be modified by *completely*, it would not get the right meaning; with a  $[0,1]$  scale *completely possible* should mean probability 1.

Comment: This is right: the maximal degree of possibility is not probability 1. Actually it seems like the 'completely' in 'completely possible' is redundant:

? It's possible that Bob is in his office, but not completely possible.

In this respect it is more like 'straight' than 'bent'. But unlike 'straight', it is unclear what the comparative form ('is more possible than') really means. This is not obviously a gradable notion.

But even if *possibly* is gradable and needs a scale, the scale might be derived from a probability measure. (For instance, it might be characterizing the domain of the probability measure.)

4. We shouldn't analyze pure epistemic modals in terms of something numerical, because it is dubious that a numerical reading is implicated in the other readings (deontic, circumstantial, etc.) of this modals, and we want a unified analysis.

Agreed. (Note Yalcin [2007] doesn't do that.) But we very well might need probabilities for probability operators, and in that case the epistemic modals should probably quantify over the domains of the relevant probabilities spaces (since we need to preserve the inference patterns (R3) and (R4) above).

## 2 Moore's paradox and epistemic contradictions

(This puzzle about to be sketched arises for both epistemic modals and probability operators. I want to focus mainly on this puzzle as it arises for probability operators, but things get complicated very fast in the probabilistic case, so let's ease into it with the epistemic modal version first.)

Bad:

1. # Its raining and it might not be raining
2. # Its raining and possibly its not raining
3. # Its not raining and it might be raining
4. # Its not raining and possibly its raining

Call things of this form:

$$(\phi \wedge \Diamond\neg\phi), (\neg\phi \wedge \Diamond\phi)$$

**epistemic contradictions.** Why are epistemic contradictions bad? We could look for parallels. Moore-paradoxical sentences comes to mind:

5. ? Its raining and I don't know that its raining
6. ? Its not raining and for all I know, its raining

Usual explanation of Moore's paradox: pragmatically defective. Why not say the same for epistemic contradictions? Answer: Because epistemic contradictions **behave differently in embedded contexts**. Consider **suppositional** and **antecedent-of-indicative** contexts:

### Moore paradoxical sentences

14. Suppose it is raining and I/you do not know that it is raining
15. Suppose it is not raining and for all I/you know, it is raining
16. If its raining and I do not know it, then there is something I do not know
17. If its not raining but for all I know, it is, then there is something I do not know

### Epistemic contradictions

7. # Suppose its raining and it might not be raining
8. # Suppose its not raining and it might be raining
11. # If it is raining and it might not be raining, then...
12. # If it is not raining and it might be raining, then...

Looks like we have a dilemma.

- $\neg\phi$  and  $\Diamond\phi$  should be modelled as having **incompatible truth-conditions**, in order to explain why it is not coherent to entertain or embed their conjunction; but
- $\neg\phi$  and  $\Diamond\phi$  should be modelled as having compatible truth-conditions, in order to block the entailment from  $\Diamond\phi$  to  $\phi$ .

## 2.1 Semantics for epistemic possibility

### 2.1.1 Traditional relational story

Usual intensional semantics for the epistemic possibility modal:

$$\llbracket \diamond \phi \rrbracket^{c,w} = 1 \text{ iff } \exists w' (wRw' \wedge \llbracket \phi \rrbracket^{c,w'} = 1)$$

$R$  is provided “by context”. (God knows how.) The accessibility relation  $R$  associated with an epistemic modal clause is one which relates the world  $w$  at which the clause is evaluated to a set of worlds not excluded by some body of knowledge or evidence in  $w$ . So we assume something like:

$$wRw' \text{ iff } w' \text{ is compatible with evidential state } S \text{ in } w$$

Problem: this won’t cover the data above. According the basic structure of the account,

#### 3. # Its not raining and it might be raining

has non-empty truth-conditions. It is just the conjunction of a meteorological claim with (roughly) a claim about a contextually determined agent or group’s ignorance of this meteorological claim. More precisely, the sentence in context is true at a world  $w$  just in case, first, it is not raining at  $w$ , and second, there is some world  $w'$  compatible with what some specific contextually determined agent or group in  $w$  knows (has evidence for, etc.) in  $w$  such that it is raining in  $w'$ . Who exactly the agent or group is, and what exactly their epistemic or evidential relation is to the body of information said to be compatible with rain is, we assume, settled in some more detailed way by  $R$ . The point is just that however these details are cashed out, we will have a totally clear, entertainable possibility in (3). That’s the wrong result.

### 2.1.2 New story

Fix: add a shiftable index parameter, call it  $s$ , explicitly representing the domain over which the epistemic modals quantify.

$$\llbracket \diamond \phi \rrbracket^{c,s,w} = 1 \text{ iff } \exists w' \in s : \llbracket \phi \rrbracket^{c,s,w'} = 1$$

Give semantics for ‘suppose’ and the indicative conditional which shift  $s$  in the right way.

$$\mathbf{Suppose.} \llbracket S_x \phi \rrbracket^{c,s,w} = 1 \text{ iff } \forall w' \in S_x^w : \llbracket \phi \rrbracket^{c,S_x^w,w'} = 1$$

$$\mathbf{Conditional.} \llbracket \alpha \rightarrow \psi \rrbracket^{c,s,w} = 1 \text{ iff } \forall w' \in s_\alpha : \llbracket \psi \rrbracket^{c,s_\alpha,w'} = 1$$

I defined  $s_\alpha$  as a function of  $s$  and the semantic value of  $\alpha$  in the paper, but the definition is sort of complicated. Now I would rather assume a similarity order over the space of information states, and just say that  $s_\alpha$  is the **nearest information state** to  $s$  which accepts  $\alpha$ .

Easy to see how this would generate the right results for embedded epistemic contradictions. But need to say more about how to understand consequence on this picture. Basically we should want to say that **what consequence preserves is acceptance**:

$$\phi_c \text{ is } \mathbf{accepted} \text{ in information state } s \text{ iff for all worlds } w \text{ in } s, \llbracket \phi \rrbracket^{c,s,w} = 1$$

That suggests we should say this about the pragmatics of unembedded epistemic possibility sentences:

**Informational view.** To say  $\diamond \phi$  in a context  $c$  is to propose to make  $\diamond \phi$  accepted with respect to the context set of  $c$ .

Challenge now: extend all this to the probabilistic case, in particular the semantics for ‘suppose’ and the conditional. Obvious thing to do: make  $s$  a probability space.

## References

- C. Kennedy. Vagueness and grammar: The semantics of relative and absolute gradable adjectives. forthcoming in *Linguistics and Philosophy*, 2006.
- Angelika Kratzer. Modality. In A. von Stechow and D. Wunderlich, editors, *Semantics: An International Handbook of Contemporary Research*, pages 639–50. de Gruyter, 1991.
- P. Portner. *Modality*. Oxford University Press, 2009.
- Seth Yalcin. Epistemic modals. *Mind*, 116(464):983–1026, Nov. 2007.