Introduction to Conditionals, Paradox, and Probability: Themes from the Philosophy of Dorothy Edgington.

Lee Walters, University of Southampton.

Dorothy Edgington's work has been at the centre of a range of ongoing debates in philosophical logic, philosophy of mind and language, metaphysics, and epistemology. This work has focused, although by no means exclusively, on the overlapping areas of conditionals, probability, and paradox. In what follows, I briefly sketch some themes from these three areas relevant to Dorothy's work, highlighting how some of Dorothy's work and some of the contributions of this volume fit in to these debates.

1.1 Conditionals

Often we face deep-rooted uncertainty, and so the best we can do is to estimate the probabilities involved, rather than making outright judgments as to the truth or falsity of a claim. For example, there are ten balls in a bag, five red and five white; Priya picks an unseen ball from the bag at random; has Priya picked a red ball? The prudent answer is not to affirm or deny outright that Priya has picked a red ball, but rather to say that it is 50% likely that she has picked a red ball (and consequently, 50% likely that she has not).

This retreat to probabilistic judgements from outright affirmations and denials is not limited to categorical claims such as Priya picked a red ball. It is also present in our consideration of conditional statements. So, adding to the previous example, let's say that three of the red balls have black spots. What should our attitude be to the claim that if Priya picked a red ball, it had a black spot? Well, as three of the five red balls have black spots, the appropriate answer seems to be that it is 60% likely that if Priya picked a red ball, it had a black spot.

Considerations of this sort make attractive the claim that the probability of a conditional is equal to the conditional probability of its consequent on its antecedent.

The Equation: p(if A, C) = p(C/A),

where p(C/A) = p(A&C)/p(A), when p(A)≠0.

So far, so good. The problem is that David Lewis (1976) proved that there could be no proposition expressed by ‘if A, C’ that satisfied The Equation, and so conditionals were not evaluable as true or false, if The Equation holds. So shocking was this result, given the intuitiveness of The Equation, that Robert Stalnaker described Lewis’s result as a ‘bombshell’ in a letter to Bas van Fraassen. This is because Stalnaker was trying to give an account of the propositions expressed by conditional statements that respected The Equation. In the face of this dilemma, Stalnaker rejected The Equation and maintained that that conditionals express propositions (see, inter alia, his contribution to this volume). Others, like Edgington, held on to The Equation and denied that conditionals expressed propositions.

The Equation is extremely intuitive, so in the face of Lewis’s proof, why not give up the assumption that conditionals express propositions as Edgington does? First, we can ask what are people doing when they accept and put forward conditionals, if not believing and asserting them? The non-propositionalist can respond that we conditionally believe and conditionally assert conditionals, where these notions may not be reducible to further mental states or speech acts. From a propositionalist perspective this response may seem ad hoc, but, as Edgington points out, there are not only what, from her perspective, are conditional assertions, but also conditional commands (if it rains, take in the sun loungers) and conditional questions (if Liverpool score first, will they win?), where these are not obviously equivalent to outright imperatives and questions.

A second objection to non-propositionalism states that whereas one might be able to accept that indicative conditionals (if Priya took a red ball, it has a black spot) do not express propositions, it is too much to accept that counterfactual conditionals (if Priya had taken a red ball, it would have had a black spot) do not express conditionals. But whereas Stalnaker reasons from the claim that counterfactuals express propositions to the claim that indicatives also express propositions, Edgington reasons in the opposite direction providing analogous reasons for endorsing non-propositionalism about counterfactuals as she does for indicatives.

Perhaps the main reason for accepting that conditionals express propositions is that the alternative appears to be subject to a version of the Frege–Geach problem that plagues expressivism[1]. That is, whatever non-propositional meaning we assign to conditionals has to be consistent with the meaning of conditionals when they are embedded in more complex linguistic forms. We can, for example, not only embed categorical statements, statements that straightforwardly express propositions, in the consequents of conditionals, we can also seemingly embed conditionals in the consequents of other conditionals, resulting in structures such as ‘if A, then if B, then C’. But if what is required to be the consequent of a conditional statement is a proposition when the consequent is a categorical claim, presumably a proposition is required when the consequent is itself a conditional. Similar points can be made with other operators such as conjunction, disjunction, negation, and modal operators.

A simple way of accounting for the embedding of conditionals in more complex linguistic structures is in terms of propositional contents. But if we reject that conditionals express propositions, we need to find some other way of accounting for embedding conditionals. As it stands, this is a challenge to those who want to hold on to The Equation, rather than a proof that The Equation has to be rejected. Still, the onus appears to be on those who maintain The

[1] See Williamson's chapter in this volume for discussion.

Equation, like Edgington, to provide a compositional account of the meanings of conditionals in non-propositional terms that allows them to be embedded.

The above construal of the dialectic rests on the assumption that we can unproblematically embed conditionals in more complex linguistic forms. But this is far from obvious. For example, (i) 'if *A*, then if *B*, *C*' seems equivalent to (ii) 'if (*A*&*B*), *C*'. But if (i) and (ii) are equivalent, then, as Gibbard (1981) has shown, 'if A, C' is equivalent to ~A v C and so the falsity of a conditional's antecedent is sufficient for its truth. But this consequence appears unacceptable. To take one of Edgington's examples, from the fact that The Queen is not at home, it does not follow that if The Queen is at home, she is waiting for me to telephone![2]

Moreover, as McGee (1985) argued, treating embedded conditionals as expressing propositions seems inconsistent with modus ponens. For example,

1. If a Republican wins, then if Reagan does not win, Anderson will.
2. A Republican will win.

Therefore,

3. If Reagan does not win, Anderson will.

(3) seemingly follows from (1) and (2) by modus ponens. But whereas (1) and (2) both seem true, (3) does not (the Democrat Carter was running second in the polls, with Anderson a distant third). McGee takes such cases to show that modus ponens is indeed invalid, but such a conclusion is hard to swallow. The non-propositionalist can note that she does not run into this difficulty, for on her construal, the argument above should be replaced with:

---

[2] Gibbard takes his proof to show that indicative conditionals do not express propositions. In the same paper, he also presents his example of Sly Pete, a so-called 'Gibbard case' to reason to the same conclusion. Edgington (1997) argues, however, that the proper lesson of such cases is that indicative conditionals are somehow epistemic, rather than that they do not express propositions. See also Rothschild (Chapter 3, this volume).

4. If a Republican wins and it is not Reagan, then Anderson will win.

5. A Republican other than Reagan wins.

6. Anderson will win.

And this argument is valid. The propositionalist, then, needs to explain why we are inclined to accept that (i) and (ii) are equivalent when, for him, they are not.

More generally, Edgington has argued that embeddings of conditionals are problematic, a result that is surprising, if they express propositions. Angelika Kratzer (Chapter 4, this volume) furthers the debate on embeddings of conditionals by considering what account we can give of quantified conditionals such as ‘no one will pass, if they goof off’. Kratzer argues that such conditionals do represent a problem for a propositional account of conditionals and that there is no general account of the embeddings of conditionals. This is music to Edgington’s ears, but rather than rejecting propositionalism, Kratzer appeals to pragmatics to address the problem (see also Chapter 3, this volume, where Rothschild argues that by adopting a Kratzer-style treatment of ‘if’ as a restrictor, propositionalists have the resources to respond to a number of arguments for non-propositionalism).

1.2 The Paradox of Vagueness

A second literature to which Edgington has made important contributions and which is the focus of several chapters in this volume, is the paradox of vagueness. Many concepts expressed in natural language appear to be vague, in the sense that they appear to lack precise application conditions and admit of borderline cases, cases where it neither seems right to say the concept applies nor to say that the concept does not apply. For example, consider the concept *bald*. Although there are people who are clearly bald and people who are clearly not bald, there are people who are in-between, neither clearly bald nor clearly not bald. It will not do, apparently, to say that such people are neither bald nor not bald, since, plausibly, this is to contradict oneself. The trouble does not end there, however.

It is characteristic of vague concepts, unlike precise concepts, that they appear to obey a principle of tolerance, the iterated application of which leads to absurdities such as that someone with a million hairs on their head is bald.[3] The paradoxical reasoning only requires two premises and multiple applications of modus ponens:

7. A person with zero hairs on their head is bald.

8. If a person with n-1 hairs on their head is bald, a person with n hairs on their head is bald.

Therefore,

9. A person with one hair on their head is bald.

(7) is an obvious application of the concept *bald*; (8) is justified by the principle of tolerance that appears to be characteristic of vague concepts; and (9) follows from (7) and (8), by modus ponens. By itself, (9) does not represent a problem. But once we have (9), we can combine this with (8) to derive that a person with two hairs on their head is bald, and so, by repeated applications of (8), we can deduce that someone with a million hairs on their head is bald!

How should we respond to such paradoxical reasoning? Edgington (1996) approaches the paradox of vagueness by arguing that there is analogy between it and the preface paradox. In particular, she thinks that we can learn lessons about vagueness by considering the degree of belief response to the preface paradox.

The preface paradox is as follows. A careful author believes each of the claims she makes in her book, but, acknowledging her fallibility, she states in the preface that some claims she

[3] In our toy example, I'm ignoring the fact that whether someone is bald depends not only upon how many hairs they have on their head, but also upon the distribution of those hairs.

makes in the book are bound to be false. The author appears to be rational, but we can reason from a number of claims that she takes to be true (each claim she makes in the book) to a claim that she rejects, namely that all the claims in the book are true. What to do? The degree-theoretic response is to say that the author believes each individual claim she makes in the book to some high degree less than certainty, but that when she considers the book as a whole her individual doubts add up, so that she believes the entire book only with a low degree of certainty. This approach can be formalized by modelling degrees of belief probabilistically to show that the probability of the conclusion of a valid argument can be lower than the probability of any of the premises. Certainty, unlike truth, is not preserved by valid reasoning.

Edgington's account of vagueness, like the above approach to the preface paradox, employs a probabilistic degree-theoretic structure that she calls 'verities' or 'degrees of closeness to clear truth'. How does this help with the paradox of vagueness? In the paradoxical reasoning above, we move from a clear case of baldness, a statement with verity 1, to a clear case of non-baldness, a statement with verity 0. The argument is valid, and the principle of tolerance looks in good shape. Edgington's idea is that as we move along the sequence of persons each with a single more hair on their head than the previous one, we gradually move away from people who are clearly bald (verity 1), through the people where it is completely unclear whether they are bald (verity 0.5), to the people who are clearly not bald (verity 0). But at every point in the sequence, the relevant conditional that is an instance of the principle of tolerance is extremely plausible, because it has a verity just short of 1. For Edgington, the verity of a conditional just is the conditional verity of the consequent given the antecedent. And the conditional verity of x is bald, given that y is bald, is the value to be assigned to x is bald on the hypothetical decision to count y as bald, a decision that is not clearly wrong, given that y is a borderline case of baldness. By employing conditional verities, then, Edgington hopes to explain the plausibility of the instance of the principle of tolerance used in the paradoxical reasoning.

Edgington's approach to the paradox of vagueness is intriguing, but it raises many questions (as do all approaches to vagueness). One set of questions concerns the nature of verities themselves. Edgington is clear that verities are not degrees of truth and so are not intended to replace the classical bivalent truth-values. In Chapter 15, Nicholas Jones argues that Edgington is mistaken to think that verities and the classical truth-values are not in competition because classical semantics is incompatible with plausible principles concerning the relationship between the two approaches. Jones also casts doubt on Edgington's claim that verities are not in fact truth-values.

However we ultimately understand verities, Edgington's approach is motivated by what she takes to be analogies between the paradox of vagueness and the preface paradox. In Chapter 14 Alan Hájek argues that Edgington was correct to draw parallels between reasoning with uncertainty and reasoning with vague concepts. Hájek points to experiments in which subjects are taken along a series of coloured patches, where such subjects display so-called *reverse hysteresis* in their responses. In the experiment Hájek discusses, subjects are presented with a series of colour patches ranging from clearly blue, through bluey-green patches to those that are clearly green. What happens is that there are patches that subjects label as green when approaching from the blue-end of the range that they label as blue when approaching from the green-end of the range. Hájek takes such judgments to be rational and argues that the best explanation of this is that this is a version of the Preface Paradox, the Progressive Preface Paradox.

Rosanna Keefe, whilst not wanting to deny some analogies between the preface paradox and the paradox of vagueness, argues in Chapter 13 that there are important disanalogies between reasoning with vague concepts and the preface paradox and that this constitutes a case against Edgington's treatment of vagueness. In particular, Keefe argues that whereas in the preface paradox we believe all of the premises individually, but not their conjunction or universally quantified form, in the case of the paradox of vagueness we believe both the individual premises and their conjunction and universally quantified form. Keefe, instead, argues that a supervaluationist treatment of vagueness is better equipped to take account of these facts.

1.3 The Paradox of Knowability

Edgington's final contribution to be discussed here, is her novel take on the so-called paradox of knowability. It is now widely known that a weak form of verificationism *classically* entails an absurdly strong form of verificationism, given certain seemingly minimal assumptions. In particularly, reading 'Kp' as p is known by someone at some time or other, the weak form of verifictionism:

Knowability $\forall p(p \rightarrow \Diamond Kp)$

entails the implausibly strong version of verificationism:

Known$\forall p(p \rightarrow Kp)$.

Proof:

10. $q \wedge \neg Kq$	Assume there is an unknown truth for reductio;
11. $(q \wedge \neg Kq) \rightarrow \Diamond K(\mathrm{q} \wedge \neg Kq)$	Instance of Knowability;
12. $\Diamond K(q \wedge \neg Kq)$	10, 11, modus ponens;
13. $\Diamond(Kq \wedge \mathrm{K}\neg Kq)$	12, Knowledge distributes over conjunctions;
14. $\Diamond(Kq \wedge \neg Kq)$	13, The Factivity of knowledge.

But no contradiction is possible, contra (14), so our original assumption, (10), is false and there are no unknown truths. So, Knowability entails Known—that all truths are known. But since the latter is unacceptable—no one knows how many hairs were on my head twenty years ago—so is Knowability. Knowability, then, has to go.

As presented here, the paradox of Knowability (or Fitch's Paradox, or the Church-Fitch Paradox) is first taken as a proof that Knowability entails Known, and second, given that Known is false, that Knowability is also false. But where is the paradox in that? Rather, it seems that we should take the proof not as a *paradox* but rather as a *result* showing that there are certain structural limitations on knowledge (cf. Williamson, 2000: 271). In particular, the proof shows that where q is an unknown truth, the fact that it is an unknown truth cannot be known. Stated this way, the falsity of Knowability seems obvious.

Nevertheless, Edgington claims that '[t]hat truths which are [in principle] unknowable … should abound in the form of the most ubiquitous and mundane facts, such that no one noticed a fly on the ceiling, or when this leaf fell from this tree, strikes many as paradoxical' (2010: 42). Paradoxical or not, Edgington argues 'that there is a sense in which one can know that, as things actually are, p and it is not known that p, but from a counterfactual perspective—as it were, from a modal distance' (2010: 42). Edgington's thought is that the possible situation of the knower need not be identical to the possible situation of the unknown truth. Rather than vindicating Knowability, what Edgington argues for is the following (where s and s^* range over possible situations):

E-Knowability $\forall p\ \forall s\ ((\text{in } s{:}\ p) \rightarrow \exists s^*\ (\text{in } s^*{:}\ K(\text{in } s{:}\ p)))$.

Williamson (2000, and Chapter 12, this volume) raises a challenge for Edgington's defence of E-Knowability, namely, how is the situation of the truth, s, specified? It cannot be specified by a subject of some other situation, s*, by use of the phrase 'the actual situation', since this phrase will pick out s* not s. More generally, there seems to be no way that the potential knower can pick out s via some causal referential chain. A subject in s* can pick out s by description, though: s is the situation in which p, q, r, … But as Williamson notes, this renders E-Knowability trivial, which was not Edgington's intention, since the consequent of E-Knowability is true simply of someone knowing that in the situation in which p, q, r, … obtain, p is true. How, then, to allow that a subject in s* can pick out s, without allowing that this is trivial? Edgington's approach is to invoke counterfactual conditionals, but Williamson argues at length that this approach does not work, pressing, amongst other objections, the Frege-Geach worry for Edgington's view that conditionals are not truth-apt.

***

Dorothy Edgington's work far outstrips the three topics discussed above. And her contributions to the three debates mentioned are more numerous and offer more insight than can be discussed here. Nevertheless, from this brief overview, and from the chapters contained in the rest of the volume, we can see the ingenuity and the breadth of Edgington's work, the difficulty of the problems that she has focused on, and how she has advanced our understanding of them.

References

Edgington, D. (1995). 'On Conditionals', *Mind*, 104: 235–329.

Edgington, D. (1996). 'Vagueness by Degrees' in Rosanna Keefe and Peter Smith (eds), *Vagueness: A Reader*. Cambridge MA: MIT Press, pp. 294–316.

Edgington, D. (1997). 'Truth, Objectivity, Counterfactuals and Gibbard', *Mind*, 106: 107–16.

Edgington, D. (2010). 'Possible Knowledge of Unknown Truth', *Synthese*, 173: 41–52.

Gibbard, A. (1981). 'Two Recent Theories of Conditionals' in W. L. Harper, R., Stalnaker, and C. T. Pearce (eds), *Ifs*, Dordrecht: Reidel.

Lewis, D. (1976). 'Probabilities of Conditionals and Conditional Probabilities', *Philosophical Review*, 85: 297–315.

McGee, Vann (1985). ‘A Counterexample to Modus Ponens’, *Journal of Philosophy*, 82: 462–71.

Williamson, T. (2000). *Knowledge and Its Limits*. Oxford: Oxford University Press.