Counterepistemic indicative conditionals and probability

J. R. G. Williams
Very drafty: some references missing

(November 19, 2007)

1 Introduction

Two major themes in the literature on indicative conditionals are (1) that the content of indicative conditionals typically depends on what is known;[1] (2) that conditionals are intimately related to conditional probabilities.[2]

In possible world semantics for counterfactual conditionals, a standard assumption is that conditionals whose antecedents are *metaphysically impossible* are vacuously true.[3] This aspect has recently been brought to the fore, and defended by Tim Williamson, who uses it in to characterize alethic necessity by exploiting such equivalences as:

$$\Box A \Leftrightarrow \neg A \mathrel{\Box\!\!\rightarrow} A.$$

One might wish to postulate an analogous connection for indicative conditionals, with indicatives whose antecedents are (in some relevant sense) *epistemically impossible* being vacuously true: and indeed, the modal account of indicative conditionals of Brian Weatherson has exactly this feature.[4] This allows one to characterize an epistemic modal $\boxdot$ by the equivalence

$$\boxdot A \Leftrightarrow \neg A \rightarrow A.$$

For simplicity, in what follows we write $\boxdot A$ as KA and think of it as expressing that subject S knows that A.[5]

The connection to probability has received much attention. Stalnaker (1970) suggested, as a way of articulating the 'Ramsey Test', the following very general schema for indicative conditionals relative to some probability function P:

$$P(A \rightarrow B) = P(B|A)$$

[1]For example, Nolan (2003); Weatherson (2001); Gillies (2007).

[2]For example Stalnaker (1970); McGee (1989); Adams (1975).

[3]Lewis (1973). See Nolan (1997) for criticism.

[4]'epistemically possible' here means incompatible with what is known (where 'what is known' is to be cashed out in some relevant sense).

[5]This idea was suggested to me in conversation by John Hawthorne. I do not know of it being explored in print. The plausibility of this characterization will depend on the exact sense of 'epistemically possible' in play—if it is compatibility with what a single subject knows, then $\boxdot$ can be read 'the relevant subject knows that p'. If it is more delicately formulated, we might be able to read $\boxdot$ as the epistemic modal 'must'. See Gillies and von Fintel (2007) for general discussion of epistemic modals.

Slogan: the probability of a conditional is the conditional probability.[6] The principle—or more carefully, *simple* cases of it—are strongly backed by intuition and by empirical evidence on folk judgements of the probability of conditionals.[7] As importantly, it arguably articulates a central theoretical role for indicative conditionals: to render propositional our dispositions to change our factual beliefs on receipt of new factual information.[8]

The general criterion has become notorious, due the impossibility results formulated by Lewis (1976) and developed in various ways by many others.[9] It turns out (a) that if the conditional is reasonably well-behaved, the above characterization is simply not satisfiable.[10] And interpreting probabilities as credence or evidential probability, even single instances of it run into trouble when we consider how the probability evolves under the impact of new information.[11]

But two things should be borne in mind. First, there are no impossibility results that show that we can't have all *simple* instances of the equation above satisfied in a single probability function—and it is these that intuition and empirical evidence for the equation supports, and which are primarily emphasized by the theoretical role for conditionals as capturing dispositions to change factual beliefs. Indeed, van Fraassen (1976) and McGee (1989) have shown that no such impossibility results will be forthcoming.[12] In an important sense, many instances of the equation are safe, at least 'statically'. Second, it is itself philosophically controversial whether and how belief-update (or the evidential analogue) should be represented probabilistically. The most impressive 'dynamic' cases against instances of the equation above rely on Bayesian assumptions (and assumptions about the nature of the information we are conditionalizing on) which someone sympathetic to instances of the equation should try to resist.[13]

Intuition, empirical evidence and theoretical rationale support *simple instances* of the equation despite the extant impossibility results. What we shall see here—without any assumptions about how beliefs evolve over time—is that single instances of the equation have dramatic results, when combined with a conditional which makes counterepistemic conditionals vacuously true.

---

[6] For related discussion, see Adams (1975).

[7] See Evans and Over (2004, ch. X). Evidence for instances of the equation for non-simple conditionals—conditionals with conditionals in their antecedents or consequents—are surprisingly little discussed. And in fact, the most tractable such utterances (so called 'right-nested' conditionals of the form $A \to (B \to C)$) are such that folk judgements depart markedly from the predictions of the equation. This is one moral to take from McGee's putative 'counterexamples to modus ponens' (McGee, 1985); and is useful to bear in mind in what follows.

[8] Stalnaker (1984, ch.6?). Note that it is compatible with conditionalization not being a *general* story about belief updating, that it is how we revise our factual beliefs on receipt of factual information. This is important to bear in mind when evaluating the impact of the 'dynamic impossibility results', below.

[9] See in particular Hájek and Hall (2004) for extensive discussion of the results and the literature.

[10] Reference to Stalnaker.

[11] See in particular Hájek (2004).

[12] See also Stalnaker and Jeffrey (1994).

[13] In particular, it can, I think, be shown that the models for simple cases of the equation that van Fraassen describes will still satisfy the equation when updated on a certain probability 0 proposition A^* canonically constructible from any given factual proposition A, and such that conditionalizing by A or by A^* has the same effect on the probabilities of factual propositions. ('Factual'='non-conditional'). Moreover, the friend of the equation might simply reject wholesale the attempt to capture belief evolution by Bayesian means, without abandoning the probabilistic representation of degrees of belief at a time.

McGee (1989) offers a revised sense in which belief-revision might go by 'conditionalization': on receipt of information A, we might set the new $P(B)$ to be the prior probability $P(A \to B)$. This will coincide with the usual understanding of 'conditionalization' when $P(A \to B) = P(B|A)$, which (for McGee) it will for all factual A, B.

2 The result

In this section, we show how to derive the result that an arbitrary proposition A has probability 1. The setting will be a conditional with the Stalnaker logic (Stalnaker, 1968) and with vacuous counterepistemics. Aside from this, the assumptions are simply that $P(KA) \neq 0$ and a single instance of the equation.

We assume first that indicative conditionals with antecedents known to be false are vacuously true.

(VA) $KA \Rightarrow \neg A \rightarrow B$

We shall assume that the logic of conditionals is as Stalnaker (1968) describes it: in particular we have conditional excluded middle (CEM).

(CEM) $(A \rightarrow B) \vee (A \rightarrow \neg B)$

From (VA) and this logic we can derive:[14]

(VC) $KA \Rightarrow B \rightarrow A$

Moving on to probability theory, we shall in the argument that follows assume one instance of the equation mentioned earlier (which one is to be disclosed later):

(PC) $P(X \rightarrow Y) = P(Y|X)$

The following follows from CEM and modus ponens for $\rightarrow$:[15]

(+) $P(A \rightarrow B|A) = P(B|A)$

For the X,Y for which we have (PC), (+) entails:[16]

(∗) $P(X \rightarrow Y|\neg X) = P(X \rightarrow Y)$

If we assume this, and have that $P(KA) > 0$ (so that we may without worry conditionalize upon it), then we can derive $P(A) = 1$ as follows:[17]

---

[14]Proof:

1.	KA	premise	
2.	$\neg(B \rightarrow A)$	supposition for reductio	
3.	$B \rightarrow \neg A$	(2) CEM	
4.	$\neg A \rightarrow B$	VA (1)	On weakened transitivity, see Lewis (1973, pp.33-35)
5.	$\neg A \rightarrow A$	VA (1)	
6.	$B \rightarrow A$	weakened transitivity (3,4,5)	
7.	$B \rightarrow A$	reductio (2-6), DNE	

[15]All $A\bar{B}$ worlds are $\neg(A \rightarrow B)$, else we get a violation of modus ponens. Suppose $(A \rightarrow \neg B)$ at an AB world. Then we'd get a violation of modus ponens. So by reductio, $\neg(A \rightarrow \neg B)$. By CEM, we have $A \rightarrow B$. So the proportion of A worlds where the conditional is true is exactly the proportion of A worlds which are B.

[16]$P(A \rightarrow B) = P(A \rightarrow B|A)P(A) + P(A \rightarrow B|\neg A)P(\neg A) = P(A \rightarrow B)P(A) + P(A \rightarrow B|\neg A)P(\neg A)$ (the last identity from (+). So we get: $P(A \rightarrow B)(1 - P(A)) = P(A \rightarrow B|\neg A)P(\neg A)$. Cancelling $P(\neg A) = 1 - P(A)$, we get our result.

[17]It may not be immediately obvious that (1) is motivated by considerations that lead us to think of counterepistemic conditionals as vacuous. Granted that any conditional whose antecedent is known to be false is itself true. Why should we believe that at some *non-actual* world the conditional is vacuously true *there* just because *there* its antecedent is known false—if in the actual world we have no idea of this. It might be suggested that (VC) be weakened to reflect such concerns: in which case (1) may not follow.

1.	$P(X \to A \mid K(A)) = 1$	By VC
2.	$P(\neg KA \to A \mid KA) = 1$	instance of (1), provided $P(KA) > 0$
3.	$P(\neg KA \to A) = 1$	From (2), $(*)$ at $\neg KA, A$.
4.	$P(\neg KA \wedge \neg A) = 0$	From (3), $\to$ satisfying modus ponens
5.	$P(\neg(KA \vee A)) = 0$	From (4), logical equivalents
6.	$P(KA \vee A) = 1$	From (5), probability theory
7.	$P(A) = 1$	From (6), factivity of K

3 Evaluation

The meaning of this result depends on what interpretation of probability it can be read as invoking (which in turn depends on what notions of probability the equation is plausible for).

Suppose first we interpret probability as credence. Suppose we're not perfectly sure KA fails. Then by the above we must be perfectly sure that A. Contrapositively: if one has any doubt whatsoever concerning A, one should be certain that $\neg KA$. Suppose on the other hand that we interpret probability as evidential probability, and identify our evidence with what we know.[18] Then the result is that unless our evidence supports A conclusively, our evidence rules out our knowing A. Or: if our evidence to any extent supports our knowing A, it conclusively supports A.

Both these results seem very strong. But we can perhaps stabilize them if we think that failures of knowledge are transparent to us: what we fail to know, we know we fail to know. Given this, the constraint on evidential probability will not be a surprise. For if we don't know A, we know we don't know it, and so the evidential probability of $\neg KA$ should be 1, and so the evidential probability of KA should be 0. Contrapositively, to suppose that $P(KA) \neq 0$ is just to suppose that KA. But in such circumstances, A gets evidential probability 1. So the transparency of failures to know would explain why we get the argument.

Can we argue from our result to the transparency of knowledge failures? First make the following assumption:

$$P(X) = 1 \Rightarrow KX \qquad (\dagger)$$

Suppose $\neg KA$. Then it must be that $P(A) \neq 1$, for otherwise (†) gives KA for a contradiction. But by the contrapositive of our argument, this means $P(KA) = 0$, so $P(\neg KA) = 1$, so (by †) $K\neg KA$.

The principle that evidential probability 1 is knowledge is somewhat tempting, but problematic: consider an infinite series of fair coin flips. The evidential probability that it will be heads every time is arguably 0, but intuitively one doesn't know that 'all-heads' won't be the result (and indeed, if it *does* come up heads every time, factivity will ensure we don't know it, despite its evidential probability 1). Despite these worries in the infinitary case, I take the

---

If one has any sympathy for the idea that '*PA*' reports the probability of the *primary intension* of A (at least where the probability involved is epistemic or doxastic) (Chalmers cite) then the issue just raised is not one we need discuss. For the minimal counterepistemic principle just mooted still gives us that a conditional is vacuously true at a world *considered as actual* if the antecedent is known false at that world. And so VC and premise (1) both go through, construed two-dimensionally.

There is a strong way of pursuing the idea that 'only actual world knowledge matters' which I think Weatherson (2007). This says that the semantics the ordering of worlds around *any* world w is determined by what is actually known (compare contextualism in epistemology). On this reading (and setting aside 2D-ist moves), we must add the premise that KA to get our argument to run. But we can in other respects weaken the premises of the argument.

[18] Williamson (2000)...

observations just given to strongly support an S5 logic for K *modulo* our assumptions, and in particular, (PC).[19]

What do these considerations look like in the credence case? Well, the same picture can be propounded. Perhaps, first, knowledge entails subjective certainty; and, second, what we fail to know we know we fail to know. As above, this would explain our result. On the first principle, if even non-zero credence in KA is enough to entail certainty in A, then the principle that KA entails A is not a big step. There's less prospect of an *argument* that the second principle is the only way to go than there is for the evidential setting, if only because (†), read as a principle that subjective certainty is factive, is far less plausible than on the epistemic probability interpretation of P. Nevertheless, in the absence of rival explanations, the S5 model is the obvious way to make sense of our results.

The result of applying a carefully selected (simple) instance of (PC) within a well-motivated general framework for indicatives is dramatic. If PC holds for credences, we have an argument that knowledge entails subjective certainty. More than this, it is strongly suggested that only an S5 logic for our epistemic modality K can work. If K does *not* work this way, we have an overall reductio of the package: holding fixed the logic of the conditional, a new static argument against single, simple instances of the equation of conditional probabilities and the probabilities of conditionals.[20]

References

Adams, E. W. (1975). *The Logic of Conditionals*. Reidel, Dordrecht.

Evans, J. and Over, D. (2004). *If*. Oxford University Press, Oxford.

Gillies, T. (ms. 2007). 'Iffiness: very drafty draft'.

Gillies, T. and von Fintel, K. (MS June 2007). 'Might made right'.

Hájek, A. (2004). 'Triviality on the cheap'. In E. Eels and B. Skyrms, editors, *Probability and Conditionals*.

Hájek, A. and Hall, N. (2004). 'The hypothesis of the conditional construel of conditional probability'. In E. Eels and B. Skyrms, editors, *Probability and Conditionals*.

Lewis, D. K. (1973). *Counterfactuals*. Blackwell, Oxford.

[19] Our transparency principle is simply the modal axiom E for the system governing operator K. But if K can be treated as a factive normal modal operator, this suffices to give S5.

[20] Williamson (2000).... It is important to remember that the result here really concerns the epistemic modal $\boxdot$—which we are only for the sake of argument taking to be K. So we need to be a little careful in reporting consequences as carrying implications for knowledge. For all we have said, $\boxdot$ may be characterized in such a way as to make the S5 principles seem plausible, even if knowledge itself does not satisfy them.

Calling the instance of PC 'simple' may be tendentious. For it might be maintained that K itself is conditional in nature (imagining offering $\neg A \rightarrow A$ as an *analysis* of KA).

Here is one more speculative moral (warning: hunches about to be aired). The formal constructions of van Fraassen (1976) and McGee (1989) give rise to hope that we will be able to find a story for the natural language conditional that vindicates at least paradigmatic instances of PC. A conditional that makes contraepistemic conditionals vacuous can be thought of as living within a space containing only epistemically possible worlds (and given an evidential gloss on probability). So a natural hope is that constructions that may originally have been presented as constructed on the whole of modal space, can now be rerun on epistemic space. But the worry that we've been seeing here is that certain instances of PC force the structure of the underlying modal space to be too neat to plausibly be read epistemically.

Lewis, D. K. (1976). ‘Probabilities of conditionals and conditional probabilities’. *Philosophical Review*, **85**, 297–315. Reprinted with postscripts in Lewis, *Philosophical Papers vol. II* (Oxford University Press, 1986) 133-52. Also reprinted in Jackson (ed) *Conditionals* (Oxford University Press, 1991) 76-101.

McGee, V. (1985). ‘A counterexample to Modus Ponens?’. *The Journal of Philosophy*, **LXXXII**, 462–71.

McGee, V. (1989). ‘Conditional probabilities and compounds of conditionals’. *Philosophical Review*, **98**, 485–541.

Nolan, D. (1997). ‘Impossible worlds: a modest approach’. *Notre Dame Journal of Formal Logic*, **38**, 535–572.

Nolan, D. (2003). ‘Defending a possible-worlds account of indicative conditionals.’. *Philosophical Studies*, **116**, 215–269.

Stalnaker, R. (1968). ‘A theory of conditionals’. In *Studies in Logical Theory*, pages 98–112. Blackwell, Oxford.

Stalnaker, R. (1970). ‘Probability and conditionals’. *Philosophy of Science*, **37**, 64–80.

Stalnaker, R. (1984). *Inquiry*. MIT Press, Cambridge, MA.

Stalnaker, R. and Jeffrey, R. (1994). Conditionals as random variables. In E. Eels and B. Skyrms, editors, *Probability and conditionals*, pages 31–46. Cambridge University Press, New York, NY, USA.

van Fraassen, B. (1976). ‘Probabilities of conditionals’. In W. Harper, R. Stalnaker, and G. Pearce, editors, *Foundations of Probability Theory, Statistical Inference, and Statistical theories of science*, volume 1, pages 17261–308. D. Reidel.

Weatherson, B. (2001). ‘Indicative and subjunctive conditionals’. *Philosophical Studies*, **51**, 200.

Weatherson, B. (MS 2007). ‘Conditionals and indexical relativism’.

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