A Puzzle about Knowing Conditionals[i]

Daniel Rothschild
University College London

and

Levi Spectre
The Open University of Israel

Abstract: We present a puzzle about knowledge, probability and conditionals. We show that in certain cases some basic and plausible principles governing our reasoning come into conflict. In particular, we show that there is a simple argument that a person may be in a position to know a conditional the consequent of which has a low probability conditional on its antecedent, contra Adams' thesis. We suggest that the puzzle motivates a very strong restriction on the inference of a conditional from a disjunction.

Keywords: conditionals, knowledge, Adams' thesis, conditional probability

Words: 3130

One thousand fair coins were flipped one by one yesterday. You have no information about how they landed but, in fact, not all the coins landed heads. It is tempting to think:

(**Anti-skepticism)** You know that not all the coins landed heads.

We take the name from a related thesis in Dorr et al. (2014). The name is apt because if we deny it then we would most probably need to discount any of our knowledge that has a probabilistic evidential basis, which results in a wide-ranging skepticism.

Here is another attractive principle:

(Independence) You should treat each of the coin flips as probabilistically independent.

Independence is meant to be a constraint on your probabilistic beliefs about the coins: the probability function representing your credences in the coin flips should make each flip probabilistically independent of each other. This hardly needs motivation: after all, they are causally independent by assumption and you have no special information that would break independence. However, it is worth noting, as Bacon (2014) does, that **Independence** is incompatible with assigning probability 1 to the proposition that one coin will land tails.[ii] So much the worse, we think, for the idea that knowledge requires assigning a proposition probability 1.[iii]

Here are some more general principles:

> **(Restricted Adams' Thesis)** Where A and B are non-conditional statements about coin-flips in the setup, you should assign a conditional statement of the form *If A then B* as its probability the conditional probability of B given A.[iv]

This is just an instance of Adams' Thesis (Adams, 1975), which itself puts no restrictions on A and B. Adams' thesis assumes that conditionals are not material, as the material conditional can often have a different probability from the conditional probability of B given A. However, Adams' thesis is widely assumed to accurately characterize our reasoning and talk with natural language conditionals. For example, saying that *it's likely that if A then B* seems to be just the same as saying that it's likely that A conditional on B. This observation is explained by the **Restricted Adams' Thesis**. The main source of trouble for the unrestricted version of Adams' Thesis stem from Lewis's triviality results (1976) and a certain class of cases where the thesis seems unintuitive (e.g., Kaufmann, 2004). We think these issues are orthogonal to those we are discussing here, and in particular do not apply when A and B are restricted to being statements about coin flips in our setup.[v]

> **(Restricted or-to-if)** If you know a statement of the form *A or B* but you do not know that A is true or false or that B is true, then you are in a position to know that *if not A then B*.

This is a famous and much discussed inference pattern (e.g., Stalnaker, 1975). Note that **Restricted or-to-if** is only a substantive hypothesis if the conditional is not the material conditional (as Adams' Thesis implies), since otherwise the disjunction and conditional are logically equivalent. An unrestricted version of the or-to-if principle is more problematic: Suppose you know that it is raining, then you can (perhaps) infer that either it's raining or there's a Martian invasion. In this case, you can reason from or-to-if to infer that if it's not raining then there's a Martian invasion. The restricted version of the or-to-if principle, however, is extremely attractive. It explains many cases of conditional knowledge from inferences. For example, I know Cathy is either in Hong Kong or Sao Paolo, so I know that if she's not in HK, she's in SP.

> **(Knowledge & Probability)** If you are in a position to know something then you cannot assign it a probability of one-half or less.

This is an uncontroversially weak link between one's probabilities and one's knowledge (much weaker than the doctrine that you can only know things you assign probability 1 to).

These principles are in tension. Here is the argument:[vi] There must be a least number *n* such that you know that the first *n* coins did not all land heads. This follows immediately from the setup and **Anti-Skepticism** (as well as principles of classical logic, which we will consider later). Assuming knowledge to be closed under (known) logical equivalence you know on the current setup the disjunction: *either the first n-1 flips did not all land heads or the nth flip landed tails*. Since you do not know either disjunct in this case (the first you don't know by the choice of *n*, the second by the setup and **Knowledge & Probability**) then by **Restricted or-to-if** you know that *if the first n-1 flips all landed heads then the nth flip landed tails.* However, by **Independence** and **Restricted Adams' Thesis** you assign this conditional probability .5. So by **Knowledge & Probability** you do not know this conditional. Contradiction.

Something has to give. The only plausible candidates to us seem to be: **Anti-skepticism**, **Restricted Adams' Thesis*,* Restricted or-to-if** and perhaps the background classical logic that we used to derive the contradiction. A few thoughts on these: **Restricted Adams' thesis** might seem the softest target as the unrestricted thesis is independently problematic and known to have apparent counterexamples. Nonetheless the restricted version of Adam's thesis does not obviously on its own lead to any paradoxical results and we could further restrict it to just the one instance used in the previous paragraph. This use of Adam's thesis does not look anything like the standard apparent counterexamples. Indeed, it seems intuitive to us that the probability of the conditional *if the first n-1 flips did all land heads then the nth flip land landed tails* is just .5 as Adams thesis states.

A natural reaction to **Anti-skepticism** is to think that the coin proposition, *the first 1000 flips did not all land heads*, looks like a lottery proposition, e.g., *this ticket will lose the NY State lottery.* Many epistemologists think lottery propositions are not knowable (in absence of direct evidence) so we might think that the coin propositions are also not knowable and reject **Anti-skepticism**.[vii] However, it would be a mistake to think that theoretical consistency requires us to take the same attitude toward coin propositions as to lottery propositions. There are many non-lottery propositions that we think have a small probability of being false that we nonetheless want to say we know. For example, reading in the local newspaper that you lost the local lottery with 1/1000 chances of winning, would seem to give you knowledge that you lost, even if the probability that the paper made a printing error and that you have actually won equals the probability of winning the New York State lottery. More tendentiously, you might think that you know that you will not win the next 100 local 1/1000 lotteries even though the probability of this combination of events can be higher than that of winning an exceptionally large lottery. The present coin case seems much more like the former than the latter.[viii] Similar things can be said about an example proposed by Vogel (1990). It seems we know that not all 50 beginner golfers will get a hole-in-one on the Heartbreaker, even if the chance of such an event isn't 0. If every golfer's probability of getting a hole-in-one is stipulated to be independent (perhaps they play on different days and have no knowledge of the others' success, for instance) knowledge does not seem to disappear. In fact, the independence assumption only seems to make us more confident that we know.

Thinking lottery propositions are unknowable, then, doesn't force you to reject **Anti-skepticism**. There are positive reasons to accept **Anti-skepticism** as well. As Dorr et al. (2014) show, it is easy to transform skepticism about coin toss cases into skepticism about everyday propositions about the future. Suppose that in each one-hour period in autumn there is an independent chance of 1/2 that a leaf will fall off the tree. If you know the leaf will fall off the tree by the end of Autumn, you would seem to need to accept **Anti-skepticism.** Lottery propositions, as single events, do not have an analogous probabilistic structure. So it seems that we can't untangle the rejection of **Anti-skepticism** from skepticism about the future.[ix]

The argument for the inconsistency of the premises depends—as many arguments do—on assumptions in classical logic. Most obviously, the law of excluded middle (**LEM**) is necessary to establish the claim that figured in the argument for inconsistency above that there is a least number *n* such that you know that exactly *n* coins won't all land heads.[x] Many think that the **LEM** should not be accepted for vague statements, and the relevant knowledge ascriptions do seem vague. We can however, give, another, slightly more cumbersome version of the argument that doesn't rely on the **LEM**.

Given **Knowledge & Probability** and the fact that you know that all the coins are fair using modus tolens we can infer that you don't know that any of the coins will land heads (or tails). We can also derive these 1000 conditionals from the **Restricted or-to-if** principle for each *n* between 1 and 1000.

> Antecedent: you know (the first *n*-1 coins won't land all heads or the *n*th coin will land tails) *and* you don't know (the first *n*-1 coins won't all land heads)[xi]
> Consequent: you are in a position to know (if the first *n*-1 coins all land heads then the *n*th coin will land tails)

By **Independence** and **Adam's Thesis** the consequents in the 1000 conditionals each have a probability .5. Given **Knowledge & Probability** we can use modus tollens to conclude that you are not in a position to know any of the conditionals in the consequent, so we can derive the negation of each of the consequents. Using modus tollens on the 1000 conditionals we can infer the negation of each of the antecedents. Consider the 1000^{th} antecedent: *You know the first 999 coins won't all land heads or the nth coin will land tails and you don't know the first n-1 coins won't all land heads*. Given **Anti-skepticism** the first conjunct is true, so using the inference rule $\sim(A\&B), A \vDash \sim B$, we can derive that the second conjunct is false. We can now infer, by double negation elimination, that you know that the first 999 coins won't all land heads. By repeating this reasoning we can eventually conclude that you know that the first coin will land tails. This gives us an inconsistency. The proof only relies on Modus Tolens, Double Negation Elimination, and $\sim(A\&B), A \vDash \sim B$, inference rules which a logician who rejects the LEM can still accept.[xii] Of course, a non-classical logician may still find ways to get out of this puzzle, but we have shown that merely eliminating the law of excluded middle is not enough.[xiii]

The **Restricted or-to-if** might seem, then, the better target. However, or-to-if reasoning is a critical way of gaining knowledge of conditionals, so without a better candidate restriction it's unattractive to discard it. One modification that might do the work is to further restrict it to cases where you know you know the disjunction.

> **(Further Restricted or-to-if)** If you know you know a statement of the form *A or B* but you do not know that A is true or that B is true, then you are in a position to know that *if not A then B*.

You might think that considerations along the lines of Williamson's (2000) safety principle (or his margin for error principles) precludes you from knowing that you know that the first *n* coins didn't land heads (where *n* is, again, the least number such that you know that the first *n* coins didn't land head).[xiv] If you don't know you know it, this further restricted or-to-if principle won't apply. Of course there might still be a lowest *m* such that you know that you know the first *m* coins didn't all land heads. By the **Further Restricted or-to-if** principle you are in a position to know that *if the first n-1 coins landed heads, then one of the next n to m coins landed tails* (assuming $m<2n-1$).[xv] The conditional probability of *one of the next n to m coins landed tails* given that *not all of the first n-1 coins landed heads* is just $1-1/2^{m-(n-1)}$. If $n = m-1$, then the conditional probability is .75. In this case knowing the conditional is compatible with **Restricted Adams' Thesis** and **Knowledge & Probability**. However, you might think it plausible that **Knowledge & Probability** is weaker than necessary, and your credence in something should be significantly higher than .75 in order to be in a position to know it. All this shows, though, is that *m* cannot equal *n*+1. In particular the gap between cases in which you know and cases in which you know you know needs to be sufficiently large to satisfy a strengthening of **Knowledge & Probability**. As there is no reason to think such gaps should be small in these cases, this is not a problem with this solution.

References:

Adams, Ernest W. (1975). *The Logic of Conditionals: An application of probability to deductive logic*. Dordrecht.

Bacon, Andrew (2014). Giving your knowledge half a chance. *Philosophical Studies* (2):1-25.

Dorr, Cian; Goodman, Jeremy and Hawthorne, John (2014). Knowing against the odds. *Philosophical Studies* 170 (2):277-287.

Field, Hartry (2008). *Saving Truth From Paradox*. Oxford: Oxford University Press.

Hawthorne, John (2004). *Knowledge and Lotteries*. Oxford University Press.

Hawthorne, Jonathan & Lasonen-Aarnio, Maria (2009). Knowledge and objective chance. Pritchard, Duncan & Greenough, Patrick (eds.)*. Williamson on Knowledge.* Oxford: Oxford University Press. 92-108.

Kaufmann, Stefan (2004). Conditioning against the grain: Abduction and indicative conditionals. *Journal of Philosophical Logic*, 33:583–606

Lewis, David (1976). Probabilities of conditionals and conditional probabilities. *Philosophical Review* 85 (3):297-315.

Sharon, Assaf & Spectre, Levi (2013). Epistemic closure under deductive inference: what is it and can we afford it? *Synthese* 190 (14):2731-2748.

Smith, Martin (2010). What Else Justification Could Be. *Noûs* 44 (1):10-31.

Stalnaker, Robert (1975). Indicative conditionals. *Philosophia* 5 (3):269-286.

Vogel, Jonathan (1990). Are there Counterexamples to the Closure Principle?. In Michael David Roth & Glenn Ross (eds.), *Doubting: Contemporary Perspectives on Skepticism*. Dordrecht: Kluwer 13-29.

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

Williamson, Timothy (2009). Reply to John Hawthorne and Maria Lasonen-Aarnio. Pritchard, Duncan & Greenough, Patrick (eds.)*. Williamson on Knowledge.* Oxford: Oxford University Press. 313-329.

[i] We greatly benefited from the comments of an anonymous referee which led ~~us~~ to significant alterations. Many thanks also to Harvey Lederman for extensive discussion. Spectre's research was supported by the Israel Science Foundation (grant no. 463/12). Rothschild's work was supported by the UK Arts and Humanities Research Council (grant numbers AH/M009602/1 and AH/ N001877/1).

[ii] To see this consider a two-coin case: if you assign probability 1 to the proposition that both coins won't land heads, then the conditional probability of the second coin landing heads given that the first coin does is 0.

[iii] Or, if there is a notion of probability that does give all knowledge probability 1, such as Williamson's evidential probability (Williamson, 2009), it is not the only relevant notion and not the one we discuss here. For other problems (besides Bacon (2014)) concerning the relation between Williamson's distinction between objective chance and evidential probability, see Hawthorne and Lasonen-Aarnio (2009) and Sharon and Spectre (2013).

[iv] Since the conditional probability is undefined if the probability of A is 0, we take **Restricted Adams' Thesis** not to apply in such cases.

[v] Note also that our claims here are compatible with the idea that conditionals might not express propositions in Lewis's sense and with the view that conditionals might be extremely context-

sensitive, so there are various ways to avoid the problems the triviality results pose for the defender of Adams' Thesis.

[vi] This is inspired by the puzzle presented in Dorr et al. (2014), though only tangentially related to it.

[vii] The knowledge version of the lottery puzzle has two premises, single premise knowledge closure and that at least typically lottery propositions are not known. See Hawthorne's *Knowledge and Lotteries* for the best statement of the lottery puzzle and the attempts to resolve it.

[viii] One recent distinction has been proposed by Martin Smith (2010). His idea is that to know (or to be justified in believing) that a proposition is true, it must be the case that if the proposition were false an explanation would be called for. That one wins the lottery does not raise nearly as much suspicion as having a long sequence of fair coins heads tosses.

[ix] To avoid skepticism by making this distinction between lottery and coin propositions isn't enough of course. One would either need to reject, single-premise closure or make knowledge (ascriptions) somehow sensitive to the situation (of the ascription). Many proposals have been given and we need not rehearse them here since the problem we focus on assumes the weaker principle of closure under equivalences and does not turn on the shifts that would allow for contextualist style resolutions.

[x] The general claim needed was that given a set of statements A_1 to A_n where A_n is true there is a least n such that A_n is true. It is easy to see how to prove this in classical logic with the LEM. If we use a non-classical logic without the LEM there may be no proof. For example, the statement claim could come out neither true nor false in some cases using the Strong Kleene connectives in a three-valued logic.

[xi] We have eliminated the conjunct from the antecedent that you don't know that the *n*th coin will land tails since we have established it is true for all n. This does not materially affect the argument.

[xii] These rules of proof are valid in many non-classical logics which do not validate the LEM such as strong Kleene (if the conditional is taken to be the ``material'' one defined from negation and disjunction), and also the Łukasiewicz logics and the preferred logic of Field (2008) (in these latter cases the conditional may be taken to be the primitive one).

[xiii] We are particularly grateful to an anonymous referee and Harvey Lederman for pressing on this point, which we originally took a different, and incorrect, view on.

[xiv] The safety (or margin for error) principle operates here because though your belief that the first n won't all land heads is safe, your belief that you know they won't all land heads isn't. In an almost identical case the n+1 case is the first sequence of flips you know won't all land heads. So though you know, you *don't* know that you know.

[xv] If there is no such m, then there would be no disjunction that could both be known to be known to be true while its disjuncts are not known.