The Adams Two-Switch Example

Vann McGee, MIT (Draft)

Logic has an anomalous status. It’s an normative science, and we don’t ordinarily look to scientists to tell us what to do. The sciences tell us how things are, not how things ought to be, and yet logic is the science that tells us how we ought to reason. The standard response is to say that logic gives us conditional norms, telling us how we ought to reason if we want to reach conclusions that are true, without telling us whether or why we should seek the truth. Purpose-relative norms are within the jurisdiction of science. We can tell a similar story about other normative sciences, like decision theory, which is reluctant to tell us what our desires should be but eager to give advice on how to achieve them.

Ernest Adams disagrees with the standard response. Based on an exacting study of the ways indicative-mood conditionals are used in English, he concludes that good reasoning with conditionals cannot be understood as the quest for true conclusions. Indeed, he says, conditionals aren’t either true or false, and yet there are norms that govern arguments with conditional conclusions.

Adams began with a withering attack on the then-prevalent truth-functional theory, which regarded a conditional as true just in case either its consequent is true or its antecedent false. He gave examples of inferences that the classical theory regards as valid but that are utterly repugnant to English speakers. Here is one of them:

> The engine will start is Switch A and Switch B are both in the ON position.
>
> Therefore, either the Switch will start if Switch A is in the ON position or the engine will start if Switch B is in the ON position.

Here is another:

> It is not the case that, if John passes history, he will graduate.

> Therefore, John will pass history.

A more memorable argument with the same formal structure bewildered Grice, a man not bewildered easily:

> It is not the case that, if God exists, we are free to do whatever we like.
>
> Therefore, God exists.

Now that the discussion has turned theological, we should mention Dorothy Edgington's example:

> If God does not exist, then it's not the case that, if I pray, my prayers will be answered.
>
> I do not pray.
>
> Therefore, God exists.

We expected a disparity between the ways people reason and they ways they ought to reason. If speakers always spontaneously made good inferences and refused bad ones, we would have no need for a normative science of good inference. But for conditionals, the discrepancy between the inferences people accept and the inferences the theory tells us they ought to accept is so wide that one hesitates to blame it entirely on speaker stupidity. Repeat the switches inference slowly and loudly, carefully enunciating every syllable, and speakers still refuse either to repudiate the premise or to embrace the conclusion. An explanation we surely want to avoid is that English speakers don't understand plain English. Adams proposed instead that we look for an alternative to the classical theory.

Adams' search begin with a complaint Ramsey makes against traditional logical theory. Traditional logic instructs us how we can draw conclusions of which we can be fully certain from premises of which we are fully certain. But we mortals almost never find ourselves in the happy position of reasoning from premises of which we are entirely certain. Here on earth, we have to cope with uncertainty, and we'd like a logic that tells us how to cope well.

The search for a logic of high probability, according to which the near-certainty of the premises ensures the near-certainty of the conclusions, runs into an immediate obstacle. If we understand "near certainty" to mean probability above a certain threshold, so that propositions that are above the threshold are sure enough that we are willing to accept them, we find that such a basic rule as conjunction introduction isn't near-certainty-preserving. If we set the threshold at 0.99, we find that the near-certainty of p and of q only guarantees that the probability of $(p \wedge q)$ is above 0.98. We'll have the same problem with any inference with more than one premise, as long as the inference really requires all its premises.

Adams found an elegant way out of this problem by proposing a continuity condition: You can force the conclusion to be an highly probable as you like by ensuring that the premises are sufficiently probable. An argument is probabilistically valid iff, for any positive ε, there is a positive δ such that any probability assignment that assigns a value greater than $1 - \delta$ to each of the premises assigns a value greater than $1 - \varepsilon$ to the conclusion. For arguments just involving the truth-functional connectives, "$\vee$," "$\wedge$," "$\sim$," "$\supset$," and "$\equiv$," an argument is probabilistically valid iff it's classically valid. Things get more interesting when we introduce the connective "$\Rightarrow$," which is intended to be the formal-language counterpart to the ordinary language conditional.

The very idea that degrees of certainty can be measured numerically isn't something we should take for granted. That our beliefs come in degrees, so that we regard some hypotheses as more likely that others, is patently obvious, but it is by no means evident that the strengths of our beliefs can be measured by numbers. We are constantly called upon to compare the likelihoods of rival hypotheses, as when we come to a fork in the road and ask which path is more likely to take us home, so you'd expect us to pretty good at making such comparisons. Numerical measurement requires us to go beyond this to compare the likelihoods of compatible beliefs, which we might have thought incomparable. Which is more likely, that the path to the left will take us home or that the cafeteria will serve spaghetti on Thursday? The way to provoke unnatural comparisons is to force unnatural choices, and a way to get people to make unnatural choices is to appeal to their gambling instincts. Which would I prefer, a bet that I win if the left fork takes me home, or a bet with the same stakes at the same odds that's won just in case they serve spaghetti in the cafeteria?

To get from pairwise comparisons to numerical evaluations is easy, if we take advantage of the tools we can find in a casino. In the casino, we find mechanical devices – roulette wheels and dice, for example – specially designed so that their operation produces a range of possible outcomes that reasonable people regard as equiprobable. If r is a number between 0 and 1 with a finite base-six expansion, we can contrive a dice game with a probability r of being won, and we can assess an agent's subjective probability that the cafeteria will serve spaghetti on Thursday by asking whether she'd prefer the bet on the spaghetti or the bet on the dice. We can also inquire about the odds at which the agent is willing to bet that the cafeteria will serve spaghetti, but that method will only work after we have established the agent's utility scale, which we can do by looking at her preferences among gambling-hall bets.

In order to sensibly bet that p, the agent has to be suppose that it is possible to settle the bet, which means that it's possible, and not prohibitively expensive, to determine whether p is true. Once we allow for the possibility of bets that can't be settled, we get distortions. You should always bet against an open-ended universal generalization, even if you think it's likely, since if an unexpected counterexample comes along, you'll collect, whereas if the generalization is true, you can stave off payment forever by saying, "The counterexample hasn't arrived yet. Maybe tomorrow.” (Dummett’s example, “A city will never be built here.”) Even for simpler bets, if you stubbornly refuse to pay off until it’s completely certain that you’ve lost your bet, you’ll never have to pay, since we can’t exclude the possibility that an evil demon has made the spots on the die look different from what they are. We can partially silence this sort of noise if we are willing to engage in a little pretense, asking the agent hypothetically whether she would be willing to bet that *p*, on the assumption that it will be costlessly and certainly possible to determine the outcome of the bet. I think it’s not unreasonable to begin an inquiry into probabilistic logic with the assumption that we have a store of what Adams calls “factual” questions, whose answers can be found out unproblematically, although it would be clearly unreasonable to suppose that this assumption is anything more than a convenient fiction. The factual sentences are assumed to be closed under the Boolean connectives.

“The degree of a belief,” Ramsey says, “is a causal property of it, which we can express vaguely as the extent to which we are prepared to act on it.” We focus our attention on betting behavior because it’s easily observed, but our main aim in doing so it to understand how beliefs guide behavior outside the gambling hall. As Ramsey observes, when we go to the railway station, we are, in effect, betting that the trains are going to run.

There is, as far as I am aware, no reason at all to suppose that, once they’ve left the casino, the choices made by flesh-and-blood human beings are accurately described by the probabilistic theory of expected-utility maximization. What the theory can tell us is how ideally rational agents would cope with uncertainty in making decisions. For the purposes of logic, this idealization is actually welcome. The ways we ought to reason are the ways ideally rational agents reason. The normativity here is instrumental. The agent has desires, described by her utility function. Her beliefs guide her to act in such a way as to optimize her prospects of fulfilling her desires. It is traditional to tie the story together with the Dutch book argument, hinted at by Ramsey and developed in detail by de Finetti. Adams himself has significant misgivings about the Dutch book, but they needn't detain us here.

Probability theory tells how to apportion our beliefs among the factual propositions coherently. It also tells us how to update our beliefs when we acquire new information. If a rational agent starts with probability function *Pr* and she learns the factual proposition *p* with certainty, and she learns nothing else that's relevant, her new probability function, Pr_p, will be given by $Pr_p(q) = Pr(q \mid p) =_{\text{Def}} {}^{Pr(p \wedge q)}/_{Pr(p)}$, for *q* factual, assuming $Pr(p) \neq 0$. This conditional probability $Pr(q \mid p)$ is equal to the fair odds on a conditional bet that is won if $(p \wedge q)$, lost if $(p \wedge \sim q)$, and called off if $\sim p$.

Probabilistic logic doesn't tell you what factual sentences you should regard as likely, any more than traditional logic tells you what sentences you should regard as true. But it puts coherence conditions on what patterns of partial beliefs it is reasonable to hold, much as traditional logic puts consistency constraints on what sets of sentences might be true.

We would like to extend the account to encompass simple conditionals, that is, conditionals with factual components. The key to doing so lies, once again, in a remark of Ramsey, who tells us that, "if two people are arguing 'If *p* will *q*,' and both are in doubt as to *p*, they are adding *p* hypothetically to their stock of knowledge and arguing on that basis about *q*." The likelihood of the conditional $(p \Rightarrow q)$ is the likelihood *q* acquires if we add *p* to the stock of things we know and adjust the rest of our beliefs to accommodate the addition. This revised probability of *q* is designated "$Pr_p(q)$," so we have:

Ramsey Test. $Pr(p \Rightarrow q) = Pr_p(q)$, for *p* and *q* factual.

Combining the Ramsey Test with the Updating Rule we get:

Adams' Thesis. If p and q are factual, with $Pr(p) \neq 0$, $Pr(p \Rightarrow q) = Pr(q \mid p)$.

There was something a bit dodgy in the transition we just made. Ramsey talked about adding p hypothetically to your stock of knowledge, and Dorothy Edgington has warned against taking it for granted that the result of supposing p as an hypothesis is the same as what you'd wind up with if you actually learned that p. Adams' Thesis requires that a conditional be independent of its antecedent, inasmuch as learning the antecedent will have no effect on your credence in the conditional. $Pr_p(p \Rightarrow q) = Pr(p \Rightarrow q)$. But could it not happen that learning the antecedent would either undermine or bolster your reason for accepting the conditional? Here's an example. Xochitl presents herself at the clinic with a sore throat. By listening to her story and looking at her throat, we are almost sure she has a viral infection, but we culture her for strep, just to be safe. If the culture comes back positive, we'll start her on a course of antibiotics, which will surely cure the infection, but if, as expected, her infection is viral, antibiotics wouldn't do her any good. Why don't we start her on the antibiotic right away? Because we are convinced she has a virus, and so we think it very unlikely that, if she takes the antibiotic, it will cure her disease. However, if, a couple of days later, we learn that she is taking the antibiotic, perhaps by seeing a pharmacy receipt, we'll think it likely that the antibiotic will cure her, since she'd only be taking the antibiotic if the culture came back positive. Under these circumstances, since the antibiotic is very effective against bacterial infections, it is substantially more likely that she takes the antibiotic and is cured than that she takes the antibiotic without being cured, so the conditional probability that she is cured given that she takes the antibiotic is high, even though the conditional is unlikely. The conditional is unlikely, but the probability the consequent would acquire if you learned the antecedent isn't close to one. The reason for the discrepancy, I think (I get this idea from Edgington), is that supposing hypothetically that Xochitl takes the antibiotic doesn't require you to suppose that the throat culture came back positive, whereas when you actually discover that she takes the antibiotic, you have strong evidence of a positive test result. So in this case, Ramsey's proposal disagrees with Adams', and the Ramsey gets the right result.

The real cause for dissatisfaction with Adams' proposal isn't the occasional stray counterexample. It's that it offers poor prospects for further advancement. We have a marvelously elegant and nearly flawlessly accurate account of simple conditionals, but it has nothing to say about conditionals whose components aren't factual or about compound sentences with conditional components. This is particularly frustrating because of our starting place. The discussion began with very striking counterexamples to the traditional truth-functional account, the two-switch example and Grice's and Edgington's proofs of the existence of God. These examples impelled us to develop a new and better theory of conditionals. So now we go back and ask how the new theory handles the examples, and we find it has nothing to say. The examples involve Boolean combinations of conditionals, and the theory doesn't assign either probability values or truth values to Boolean compounds of conditionals.

We have three principles, the Update Rule, the Ramsey Test, and Adams' Thesis, any two of which imply the third. As I'll be using the terms, the Update Rule only applies to factual sentences and the Ramsey Test and Adams' Thesis only applies to simple conditionals. We'd like to extend the theory beyond this restricted core, Wanting to generalize an account that applies to factual sentences and simple conditionals, there are three natural ways we might go. We might allow simple conditional, as well as factual, sentences to appear as antecedents; we might allow simple conditional consequents; or we might allow Boolean combinations of factual sentences and simple conditionals. Each of these steps is intended as only a beginning.

If we extend the Ramsey Test to allow conditionals with simple conditional consequents, we get the following (assuming $Pr(p \wedge q) \neq 0$):

$Pr(p \Rightarrow (q \Rightarrow r)) = Pr_p(q \Rightarrow r)$ [by the extended Ramsey test]

$= Pr_p(r \mid q)$ [by Adams' Thesis, applied to Pr_p]

$= Pr_p(q \wedge r) / Pr_p(q)$ [by definition]

$= Pr((q \wedge r) \mid p) / Pr(q \mid p)$ [by the Updating Rule]

$= Pr(r \mid (p \wedge q))$ [by calculation]

$= Pr((p \wedge q) \Rightarrow r)$ [by Adams' Thesis].

This gives us a satisfying treatment of the Edgington example. The probability that, if God does not exist, then it's not the case that, if I pray, my prayers will be answered will be equal to the conditional probability that my prayers go unanswered given that God does not exist but I pray (assuming this is defined), and this can be high even though it's improbable that I pray and improbable that God exists.

The generalized Ramsey Test has the initially surprising consequence that *modus ponens* isn't probabilistically valid. This seems to accord with ordinary usage. Suppose that the count has been murdered and that you strongly suspect the butler. You have no reason at all to suspect the gardener, who is his only other servant. Then you'll believe that, if one of the servants killed the count, then if it wasn't the butler, it was the gardener – there aren't any other servants – and you'll also believe that one of the servants killed the count, but you won't believe that, if the butler didn't kill the count, the gardener did. This example is a variant on Stalnaker's *direct argument* that $(q \supset r)$ implies $(q \Rightarrow r)$. I want to say that the fallacy in the direct argument is *modus ponens*.

We get less promising outcomes when we extend the Ramsey Test to conditionals with conditional antecedents, like we find in Adams' example:

If John won't graduate if he doesn't pass history, he won't graduate.

Therefore, if John passes history, he won't graduate.

Any change in your beliefs that moves $Pr(p \supset q)$ to 1 also changes $Pr(p \Rightarrow q)$ to 1, and *vice versa,* so we have $Pr_{(p \Rightarrow q)} = Pr_{(p \supset q)}$. The extended Ramsey Test gives us $Pr((p \Rightarrow q) \Rightarrow r) = Pr((p \supset q) \Rightarrow r)$. But that's not right Without knowing much of anything about Mimi, we can say that, if she moves to Aruba if she wins the lottery, she really likes the beach. Take this judgment and materialize the conditional in the antecedent, and we get a conditional we can't accept without knowing about Mimi's recreational preferences. The Ramsey Test looks at what happens when we treat the antecedent as certain, and the logic of certainty can't distinguish ordinary conditionals from material conditionals.

Turning now to the problem of extending the theory to include Boolean compounds, we find that any theory that embraces Adams' Thesis and extends the standard laws of probability to Boolean compounds of simple conditionals will give us the desired result that the classically-valid two-switch argument is probabilistically invalid. $((a \wedge b) \Rightarrow e)$ can be highly probable even though $(a \Rightarrow e)$ and $(b \Rightarrow e)$ are highly improbable. But if $(a \Rightarrow e)$ and $(b \Rightarrow e)$ are both highly improbable, their disjunction can't be very probable, since a disjunction will always have probability less than 2ε if both its disjuncts have probability less than ε.

We also get a satisfying response to the Grice argument. Assuming the probability that God exists is nonzero, the probability that it's not the case that, if God exists, we are free to do whatever we like will be equal to the probability that, if God exists, we aren't free to do whatever we like. $Pr(\sim (g \Rightarrow f)) = 1 - Pr(g \Rightarrow f) = 1 - Pr(f \mid g) = Pr(\sim f \mid g) = Pr(g \Rightarrow \sim f)$ (assuming $Pr(g) \neq$

0), which can be high even if $Pr(g)$ is low. As Ramsey notes, "in a sense, 'If *p, q*' and 'If *p*, ~*q*' are contradictories."

We get a reasonable response to the examples from any account that assigns values to Boolean combinations of factual sentences and simple conditionals that upholds Adams' Thesis and the standard laws of probability. That is, it assigns values in the interval [0,1] in such a way that truth-functional tautologies get value one and that the value assigned to a disjunction with truth-functionally incompatible disjuncts is equal to the sum of the values given to the disjuncts. The trouble is that we don't have any credible, principled way to choose such an assignment.

In extending the theory to encompass Boolean compounds, the simplest case to look at will be sentences of the form $((p \Rightarrow q) \wedge r)$, with *p, q,* and *r* factual. Extending the Updating Rule to allow the sentence being updated to be a simple conditional gives us an answer:

$$Pr((p \Rightarrow q) \wedge r) = Pr((p \Rightarrow q) \mid r) \cdot Pr(r) \text{ [by definition]}$$

$$= Pr_r(p \Rightarrow q) \cdot Pr(r) \text{ [by the extended Updating Rule]}$$

$$= Pr_r(q \mid p) \cdot Pr(r) \text{ [by Adams' Thesis, applied to } Pr_r\text{]}$$

$$= {}^{Pr_r(q \wedge p)}/_{Pr_r(p)} \cdot Pr(r) \text{ [by definition]}$$

$$= {}^{Pr((q \wedge p \mid r)}/_{Pr(p \mid r)} \cdot Pr(r) \text{ [by the Ramsey Test]}$$

$$= Pr(q \mid (p \wedge r)) \cdot Pr(r) \text{ [by calculation],}$$

(assuming $Pr(p \wedge r) \neq 0$).

Unfortunately, it's the wrong answer. Setting *r* equal to ~ *q*, we get $Pr((p \Rightarrow q) \wedge \sim q) = 0$. However, "I won't buy a car next year" and "If I win the lottery, I will buy a car next year" can both be highly probable, and so, by the rule that a conjunction has probability greater than 1 - 2ε if it conjuncts both have probability greater than 1 - ε, we know that the conjunction needn't be highly improbable. The rule that tells us to update by conditionalizing gives us incorrect results when we are updating our beliefs in conditionals. When we learn $\sim s$ with certainty, our new probability of $(r \Rightarrow s)$, $Pr_{\sim s}(r \Rightarrow s)$, should be equal to zero, as Adams' Thesis (applied to $Pr_{\sim s}$) tells us, but the extended Updating Rule tells us that $Pr_{\sim s}(r \Rightarrow s)$ should be equal to $Pr((r \Rightarrow s) \mid \sim s) = Pr(\sim s \wedge (r \Rightarrow s)) / Pr(\sim s)$, which in nonzero, because its numerator is nonzero.

This disturbing calculation is the crux of the Lewis Triviality Theorem, which shows that the combination of the extended update rule with the original Adams' Thesis, applied to both *Pr* and the Pr_ps, can't be satisfied by any probability measure on the Boolean compounds of factual sentences and simple conditionals, except for degenerate measures that take on at most four values.

The Lewis theorem tells us at least this much: If the probability calculus is extended to take account of Boolean combinations of simple conditionals, the theory will have to treat conditionals differently from the way it treats factual sentences. Adams wants to draw a further conclusion, namely, that we can't identify the probability of a conditional as the probability of its truth. More explicitly, we can't impose truth-in-a-world conditions on conditionals and a finitely additive measure on the worlds in such a way that the probability of a conditional is the measure of the set of worlds in which it is true. The Boolean connectives have clearly established truth-in-a-world conditions, which ensure that the Boolean operations preserve the property that the probability of a sentence is the measure of the set of worlds in which it is true. So if the probability of a conditional were the probability of its truth, the same would hold for Boolean combinations of conditionals, which can't happen, by the Triviality Theorem.

This argument-sketch is not yet a good argument, because it doesn't make all its premises explicit. Lewis's theorem didn't say that we get that we get triviality when we add Adams' Thesis to the laws of probability, with the standard update rule for factual sentences. It says that we get triviality when we combine Adams' Thesis with the laws of probability, the standard update rule, and an extended update rule that sets $Pr_p(s)$ equal to $Pr(s \mid p)$ when s is a simple conditional. We can avoid the Triviality Theorem by refusing to extend the update rule. This move is made palatable by the fact that we already have an alternative extension to the update rule, waiting to take command. Namely, by extending the Ramsey Rule $Pr_p(s) = Pr(p \Rightarrow s)$ by allowing s to be a simple conditional, we get the alternative update rule $Pr_p(q \Rightarrow r) = Pr_{p \wedge q}(r)$, for p, q, and r factual, with $Pr(q \wedge r) \neq 0$.

There are a number of other versions of the Triviality Theorem, by Lewis and others, but all of them require some, usually very natural but not obligatory, generalization beyond the bare minimum requirement that Adams' Thesis and the laws of probability are upheld for Pr and the Pr_ps. That this bare minimum requirement can be met, no matter what values Pr assigns to the factual sentences, is an old result of Bas van Fraassen. Moreover, taking a possible world to be a maximal truth-functionally consistent set of sentences, we can ordain a finitely-additive measure on the worlds so that the probability of a sentence will be the measure of the set of worlds in which it is true.

Our problem isn't that there isn't a way to impose truth-in-a-world conditions in such a way that the probability of a sentence is the probability of its truth. The problem is that there are too many ways to do so. It's the problem we started with: We have no sensible way to assign a numerical probability to the conclusion of the switching example. The laws of probability require that the disjunction of two simple conditionals have probability at least as great as the maximum of the probabilities of the disjuncts, and no greater than the minimum of 1 and the sum of the probabilities of the disjuncts. But I don't know of any other constraints.

I've been tacitly taking it for granted that either conditionals always have truth values or they never do, and I shouldn't be so hasty. The bet settlement conditions for conditional bet suggest truth-value gaps, so that a conditional is true if its antecedent and consequent are both true, false if its antecedent is true and its consequent false, and unsettled otherwise. This proposal is, as far as I can see, harmless but it doesn't let us avoid the unwelcome conclusion that a conditional can be highly improbable even though it's highly improbable that it's true.

We can save the doctrine that only truths are assertable by denying that conditional assertion is a species of assertion. When I make the conditional assertion that, if I win the lottery, I'll buy a new car, I am not attempting to tell you what the world is like, although I am offering you a warrant that will, under suitable circumstances, enable you to draw conclusions about world affairs. A difficulty with this doctrine is that it leaves no room for whatever it is we're doing when we affirm a Boolean combination of conditionals.

"Assertion" is, to a large extent, a term of art of speech act theory. I would have naively supposed that, when I put forward a conditional, I am asserting it, but if our best theory of speech acts says otherwise, I am willing to defer to the experts. I'm not as accommodating about beliefs. I am stubbornly unwilling to deny that I believe that, if I win the lottery, I'll buy an new car.

Adams' Thesis is a statement about probability, and subjective probability is a measure of degrees of belief. There is a connection between high probability and assertability, but no direct correlation. Before I read the lottery outcome in the paper, I can't assert "My lottery ticket didn't win," even though its chance of winning is minuscule. (Williamson makes a big deal of this.) I think the reason is that it's common knowledge we all share that one's odds in the lottery are pitifully slim. If I tell you, "My ticket didn't win," I'm suggesting that I have some reason for thinking my ticket is a loser that goes beyond a general shared understanding of the way lotteries are set up. I am, notice, able to assert that I won't be able to afford a new car next year, in spite of the fact that if my lottery ticket is the winner I will be able to afford one.

In expressing Adams' Thesis as a judgment about probability, Adams is choosing his words carefully. The conditional probability is a measure of the agent's belief in the conditional, a belief that can be manifested in action. He sometimes spoke of conditional probability as a measure of assertability, but it was a loose way of talking that he came to regret. "Worst mistake I ever made" is what he told me.

Adams thinks that the probability of a conditional is given by the conditional probability, and he thinks that the Triviality Theorem shows that the probability of a conditional isn't the probability of its truth. How we are to understand this contention depends on what understanding of truth we have in the background. Quine advocates a "disquotational" conception, according to which the effect of adding "is true" to the quotation name of a sentence is just to undo the effect of the quotation marks, so that to say that "Snow is white" is true and to say that snow is white are two ways of saying the same thing. On a disquotational conception, $(p \Rightarrow q)$ and $Tr(\ulcorner(p \Rightarrow q)\urcorner)$ mean exactly the same thing, so of course they're equiprobable. Indeed, because of its analyticity, the Tarski (T)-sentence:

$$(Tr(\ulcorner(p \Rightarrow q)\urcorner) \equiv (p \Rightarrow q))$$

will have probability one. On the other hand, the Lewis Triviality Theorem informs us that, provided the factual sentences $(p \wedge q)$, $(p \wedge \sim q)$, and $\sim p$ all have nonzero probability, there can't be any factual sentence r such that the material biconditional

$$(r \equiv (p \Rightarrow q))$$

has probability one. Consequently, the truth attribution $Tr(\ulcorner(p \Rightarrow q)\urcorner)$ can't be a factual sentence.

Adams tell us very little about which sentences are to count as factual, following a sensible methodology of beginning the investigation with a crude notion, waiting to shape and polish the notion until we have a clearer understanding of the theoretical role it will be required to play. One very relaxed understanding of factuality would attempt to isolate the special features of conditionals by stipulating that any sentence counts as factual that doesn't contain any conditional parts. This conception is too liberal. $Tr(\ulcorner(p \Rightarrow q)\urcorner)$ doesn't have any conditional parts, although it contains a name of a conditional as a part, and it isn't factual.

A disquotational conception is committed to both the (T)-sentences:

$$Tr(\ulcorner(p \Rightarrow q)\urcorner) \equiv (p \Rightarrow q)$$

$$Tr(\ulcorner{\sim}(p \Rightarrow q)\urcorner) \equiv {\sim}\,(p \Rightarrow q)$$

By Boolean truth-functional logic, these entail

$$Tr(\ulcorner(p \Rightarrow q)\urcorner) \vee Tr(\ulcorner{\sim}(p \Rightarrow q)\urcorner).$$

Identifying falsity with the truth of the negation, we conclude that the conditional is either true or false. The disquotational conception of truth leaves no room for failures of bivalence.

To allow for failures of bivalence, we need a more robust conception of truth. Non-disquotational theories of truth are often referred to as "correspondence" theories, but the name is misleading. The theory needs to postulate, for any true sentence, a mechanism by which speaker usage establishes truth conditions for the sentence, and features of whatever it is the sentence is about make sure that the conditions are met. But such a mechanism need not involve an ontology of facts or a semantics of correspondence.

With the robust conception of truth, in order to have bivalence, there have to be features of usage that determine a partition of the non-p worlds into $(p \Rightarrow q)$-worlds and non-$(p \Rightarrow q)$-

worlds. We have already seen, using the Triviality Theorem, that it isn't possible to introduce any factual sentence that demarcates the partition. We can say more. There isn't any factual statement we can make about the non-*p* worlds that is even relevant to determining whether a given non-*p* world is a ($p \Rightarrow q$)-world. If *r* is incompatible with *p*, we can't answer the question whether learning *r* would make ($p \Rightarrow q$) more likely, because learning *r* with certainty would render the probability of ($p \Rightarrow q$) undefined. So we have to be more devious. We know that if one proposition is positively relevant to a given proposition, then its negation is negatively relevant, and *vice versa*. We also know that, since ~*r* implies *p*, ~*r* is neither positively nor negatively relevant to ($p \Rightarrow q$). It follows that *r* is neither negatively nor positively relevant to ($p \Rightarrow q$). It isn't credible that our usage somehow introduces an exact partition on the basis of a criterion that is so subtle that it's factually ineffable. The credible hypothesis is that simple indicative conditionals with false antecedents are neither true nor false.

There has been a lot of discussion of the semantics of truthvalueless statements, mainly because there has been so much recent interest in vagueness. Quite a few philosophers are cheerfully quick to agree that ordinary speakers don't know how to use vague terms correctly, but those who are not so radically inclined have sought a treatment of vagueness that respects the principle that, as a general rule, ordinary speakers know how to speak ordinary language. In particular, they've looked for a semantic theory according to which the things that careful, thoughtful, sober speakers who aren't factually misinformed say are, by and large, true. If Clare is a border case of "poor," "Clare is poor" will be neither true nor false, and that fits with usage, since well-informed speakers aren't willing either to affirm or deny that Clare is poor. On the other hand, if Clare's sister is noticeably worse off than she, we want to deny that Clare is poor but her sister not, so we'd like that sentence to come out false. The standard way to try to accomplish this is to say that usage doesn't pick our a single interpretation of the language. It picks out a family of acceptable interpretations, and a sentence counts as true if it's true under all of them.

The approach of trying to match truth conditions to usage doesn't work for conditionals. When I tell you, "I'll buy a new car if I win the lottery," I've said something that's perfectly appropriate according to the way ordinary speakers employ conditionals, yet I have said something that isn't true. Adams' rather drastic response to this situation was to replace the traditional norm, which only permits one to assert things that are true, with a norm that confines one's assertions to things that are highly probable. The new norm isn't big enough to fill the place occupied by the old one, for it can't explain why responsible speakers make an effort to make sure that the beliefs they express are appropriately supported by evidence.

The reason Adams' response seems so drastic to me is, while it provides us with a norm of assertion, it doesn't give an account of what makes good reasoning good. Pragmatists have tried to convince us that the virtue of good beliefs is that they lead to successful action. This makes sense for factual beliefs; a paradigm case of acting of the belief is betting on it. Likewise for simple conditional beliefs and conditional bets. But there doesn't appear to be any prudential explanation why we reason as we do when we use compound sentences with conditional components.

We're still stuck at the same place. The probabilistic account give a splendid account of the problem inferences with simple conditionals, and an initially promising account of the inferences with compounds of conditionals, but we aren't able to fill in the details. One, rather defeatist response would be to give up on the attempt to find the "right" way to assign probability values to compounds of conditionals, adopting a probabilistic analogue to supervaluation theory.

Starting with a probability assignment, or a family of acceptable probability assignments, for the factual sentences, we'll count a Boolean combination of factual sentences and simple conditionals as highly probable if it's assigned a value close to one by every probability assignments that extends one of the acceptable factual assignments and obeys Adams' Thesis. We can extend an old result of Adams to show that, once we've made allowance for antecedents with probability zero, the probabilistically valid inferences are precisely the ones endorsed by Stalnaker's logic. This probabilistic analogue to supervaluationism fits with the way that we initially diagnosed what goes wrong in the two-switches example. We don't know how to assign numerical probability value to the disjunction of two conditionals, but however we do it, as long as we respect the probabilistic laws that govern the Boolean connectives, it will be possible for the premises to be highly probable and the conclusion highly improbable.

Defeat needn't be total. It may well happen that there are constraints that rule out some of the probability assignments as unacceptable, analogous to the "penumbral constraints" that forbid an acceptable model from putting Clare in the extension of "poor" and leaving her clearly poorer system outside.

One cannot help being appalled by the inelegance of the probabilistic supervaluational hybrid. It requisitions so much heavy machinery and makes so little use of it. Still, if you aren't too put off by its ungainliness, it seems to fit the data pretty well.