Probabilities of Conditionals: Three Easy Pieces

Bas C. van Fraassen

October 2024

In this paper I will address three topics in the logic of conditionals. The first is the question whether the class of 'reasonable' probability functions must be closed under conditionalization. The second topic is the character of logical consequence when probabilities of conditionals come into play. The third is more specific: I want to present a challenge to the possible worlds approach in formal semantics, in favor of an algebraic approach. For this I will use as a case study Alan Hajek's views on counterfactual conditionals, and its problems with infinity. Included in this will be reasons to expect algebras of propositions to be incomplete algebras.

Throughout I will use as foil what is known variously as *Stalnaker's Thesis*, or the *Conditional Construal of Conditional Probability* (CCCP). That is the thesis that the probability of a conditional A → B is the conditional probability of B given A, when defined. That the CCCP is tenable for a reasonable logic of conditionals I will presuppose in the body of the paper, but I will present its credentials in the Appendix.

The CCCP is to be distinguished from the Extended Stalnaker's Thesis, or Extended CCCP, that the conditional probability of A → B given C equals the conditional probability of B given A and C. That extended thesis has been demolished again and again, and will appear here only in a note, to be dismissed.[1]

1. Probability functions, the CCCP, and Moore's Paradox

Are there statements on which we cannot conditionalize, even though they have positive probability?

This question is akin to the question posed by Moore's Paradox: are there statements that could not be rationally believed, even though they could be true? Yes, like "It is snowing and I do not believe that it is snowing". When we replace the intuitive notion of belief by subjective probability, we find two new forms:

> An **Moore Statement** is one that could be true, but could not be believed
>
> A **Weak Moore Statement** is one that could be true, but could not have probability one. A **Strong Moore Statement** is one that could have positive probability, but could not have probability one.

The CCCP leads to Moore Paradoxes that involve conditionals (first noticed in Hajek (2011), see Appendix 1).

Example 1. Imagine the following situation:
1. The match is not struck
2. The match is wet
3. It is not the case that if the match is struck, it will burn.

That would seem to be a realistic situation, quite easily imagined. Symbolize these sentences in an obvious way, to arrive at this statement:

[~S & ~(S →B)]: *the match is not struck, and it is* <u>*not the case*</u> *that if the match is struck then it will burn.*

For a probability function P that respects the CCCP, this statement cannot have probability 1.

For if P(~S & ~(S →B)) = 1, then P(~S) = 1 and P(~(S →B)) = 1. But it follows that P(S) = 0, so P(B|S) is not defined. Therefore P does not assign 0 to (S→B), hence does not assign 1 to ~(S →B).[2]

Therefore this statement is an Ordinary Moore Statement.

Could P be conditionalized on [~S & ~(S →B)]? On the face of it, certainly, for at first blush we would give that statement a positive probability. But the result of this conditionalization would then be a function that does not respect the CCCP. Therefore, without any special assumptions about the logic of →, we conclude:

Theorem 1. The class of probability functions that respects the CCCP is not closed under conditionalization.

While the argument so far already sufficed, a complete calculation should show a bit more. To display an example of a Strong Moore Statement, we need to show something which can have positive probability. For this we can use a numerical example.

Example 2. Tosses with a fair die.

The basic statements involved are just about the outcome of a toss, and each outcome has probability 1/6. Define:

A = the outcome is either two or six. True in possibilities {2, 6}

~ A = the outcome is neither two nor six = the outcome is either odd or 4. True in possibilities {1, 3, 5, 4}

B = the outcome is six. True in possibilities {6}

Y = (~ A & ~(A → B)): *the outcome is neither two nor six, and it is not the case that(if the outcome is two or six, then it is six).*

The probability of ~A is 4/6.

By the CCCP, the probability of the conditional (if A than B), equals P(B | A) = 1/2 = 3/6.
So the negation of that conditional also has probability three out of six: P(~(if A then B)) = 3/6.

The probability of the disjunction of ~ A and ~(if A then B) is the sum of P(~A) and P(~(A → B) minus P(Y). This disjunction cannot have a probability greater than 1. So

$$0 < (3/6) + (4/6) - P(Y) \leq 1$$

Thus Y has a probability greater than or equal to 1/6. But it cannot have probability 1, by the same argument as in the preceding example. For if P(Y) = 1 then P(~A) = 1, and so P(~(A →B)) is undefined. Therefore Y is a Strong Moore Statement.

Again we note that Y can play a perfectly good role as antecedent in a conditional. But the result of conditionalizing P on Y will be a function that does not respect the CCCP.

2. Conditionalization and the concept of logical consequence

The notion of valid argumentation is surely the most basic of all in logic. Puzzles about conditionals overturned many previous ideas about that very notion of valid argument. If familiarity had not bred obliviousness by now, we would still see the impact as revolutionary.

Counterfactuals

Before the war (that is, WWII) there were two good theories of conditionals, Arend Heyting's intuitionistic implication (1930) and C. I. Lewis' strict implication (1932).[3] Both had the guiding idea that the conditionals' behavior must mimic the principles of valid argumentation. Asserting $A \rightarrow B$ must have something like the force of the assertion that A implies B. The principal principle for the logic of conditionals must therefore be the great law of implication:

X, A entails B if and only if X entails $A \rightarrow B$.

That has several clear consequences for the logic of conditionals, such as *Weakening* ($(X \rightarrow B)$ entails $(X \& A) \rightarrow B$) and *Transitivity* for the arrow.

With Goodman (1947) and Chisholm (1953) it became clear that this is wrong for conditionals in natural language, especially counterfactuals. The paradigm counter-example, that "this match will (would) light if struck" does not imply "this match will (would) light if wet and struck", is by now the stuff of folklore.

Probability's radical impact

When we then introduce probability, there is a further sea change for our concept of valid reasoning.

It is natural to think, and indeed provable in a 'classical' context, that the following three entailment relations coincide:

- For all truth-value assignments v, if $v(A) = T$ then $v(B) = 1$
- For all probability functions p, if $p(A) = 1$ then $p(B) = 1$
- For all probability functions p, $p(A) \leq p(B)$

I will call the conviction that these three relations coincide the *classical conception of logical consequence*, or of *valid entailment*. It is appropriate to call this conception "classical" for, as I shall show in a moment, the coincidence of those three relations is provable from very minimal characteristics of classical logic and probability theory.

But if we accept the CCCP this classical conception will not remain tenable.

Probability spaces

Geometry is not, from a modern point of view, the study of what theorems follow from such as Euclid's or Lobachevsky's postulates. It is instead the theory of geometric spaces, that is, of structures in which those postulates hold. Similarly, the theory of probability is not the deductions from the three or four equations that define a probability function.[4] It is the study of probability spaces.

I will assume here that in a semantic analysis of a relevant sort of language the propositions form a Boolean algebra, with or without additional operators.

Definition. A *probability space* PP is a triple <K, F, P> where K is a non-empty set, F is a Borel field of subsets of K, and P is a family of probability measures with domain F. The members of F, the 'measurable sets', we call *propositions.*

When P has just one member I will call the probability space *simple.* Simple spaces may have taken most of our attention, but that is a case of misplaced emphasis. The classical conception of valid entailment requires us to look beyond simple probability spaces. For the principle, for example, that P(A & B) ≤ P(A) does not refer to a single probability assignment – it is an assertion about how things are no matter how probabilities are assigned.

Definition. A probability space PP = <K, F, P> is *closed under conditionalization* iff for all p in P and all elements e of F, p(-|e) is in P if p(e) > 0.

Theorem 2. If PP = <K, F, P> is closed under conditionalization then the following relationships coincide for all elements a, b of F:

1. [**certainty**] for all p in P, if p(a) = 1 then p(b) = 1
2. [**ordering**] for all p in P, p(a) ≤ p(b)

These are two *consequence relations* on the family of propositions. I will relate them to the familiar one of truth-preservation below.

Proof. (a) Suppose that for some p in P, p(a) = 1 but p(b) < 1. Then it is not the case that for all p in P, p(a) ≤ p(b. For the converse suppose that for some p in P, p(e) > p(g). Since F is a Borel, hence Boolean, algebra, p(e) = p(e ∩ g) + p(e – g), and p(e ∩ g) cannot be greater than p(g). So p(e – g) > 0. Define p* to be the conditionalization of p on (e – g). Then p*(e) = 1 and p*(g) = 0.

Definition. A probability space PP = <K, F, P> is *replete* iff for every non-empty element e of F there is a probability measure p in P such that p(e) > 0.

Theorem 3[5]. If PP = <K, F, P> is a replete probability space which is closed under conditionalization then the following three relationships coincide for all elements a, b of F:

1. [**inclusion**] a ⊆ b
2. [**certainty**]for all p in P, if p(a) = 1 then p(b) = 1
3. [**ordering**]for all p in P, p(a) ≤ p(b)

The relationship **inclusion** is the familiar consequence relation among propositions: if the first is true then so is the second, always.

Proof. First, by the preceding theorem, **certainty** and **ordering** coincide.

If a ⊆ b then p(a) ≤ p(b), and hence also if p(a) = 1 then p(b) = 1. So **inclusion** entails both **certainty** and **ordering.**

Second, suppose that for specific propositions a, b, **inclusion** does not hold: it is not the case that a ⊆ b. Then (a ∩ −b) is a non-zero element. By repleteness, there is probability measure q in P such that q(a ∩ −b) > 0. Clearly q(a) > 0 in that case, so the function q*, which is q conditionalized on a, is well defined.

Since PP is closed under conditionalization, q* is in P. Since q*(a) = 1 and q*(b) = 0, it follows that the relations of **certainty** and **ordering** do not hold for a, b.

Failure of the classic conception of entailment

These arguments for the classic conception of logical consequence assumed closure under conditionalization. And we saw in the previous section that the class of probability functions which respect the CCCP is not closed under conditionalization. That raises the suspicion that here the classic conception will fail. And it does.

There has been some discussion of the putative validity of the 'Or to If' inference.[6] There are familiar, plausible sounding examples:

> Peter is either in France or in Italy. Therefore, if he is not in France then he is in Italy.

Such examples seem to make it inescapable that "if … then" is the material conditional.[7] At the same time, there are examples where the inference appears invalid.

I have no wish to either advocate or deny the validity of 'Or to If'. Instead it functions as an effective counterexample to the classical conception of valid entailment, for this inference is valid by [**certainty**] but invalid by [**ordering**].

Let A and B be statements:

a) In all cases, for all probability functions p, if $p(\sim A \vee B) = 1$ then $p(A \rightarrow B) = 1$
b) In many cases, for some probability functions p, $p(\sim A \vee B) > p(A \rightarrow B)$

Both a) and b) are true if we equate $p(A \rightarrow B)$ with $p(B \mid A)$. For the first note that if $p(\sim A \vee B) =1$ then $p(A \& \sim B) = 0$, so $p(A \& B) = p(A)$. Therefore $p(B|A) = p(B \& A)/p(A) = p(A)/p(A) = 1$. For the second, imagine we are going to toss a fair die. The probability that the outcome will be *either odd or six* equals 4/6 = 2/3. But the probability of *six*, given an even (not odd) outcome, equals 1/3.

Theorem 4. If PP = <K, F, P> is a probability space such that F is a Boolean algebra and is closed under an additional binary operator →, and for all elements A, B of F and all probability functions p in P, it is the case that $p(A \rightarrow B) = p(B|A)$ when defined, then PP is not closed under conditionalization.

This follows at once from **Theorem 2** and the finding that the consequence relations **certainty** and **ordering** do not coincide in this sort of probability space.[8] It is an alternative proof of **Theorem 1**, but this proof is more enlightening. For it shows that our most fundamental concept in logic breaks into several parts. It shows that a language with conditionals has, in effect, more than one logic.[9]

Culprit not conditionals but probability of probabilities

David Lewis explicitly took it as an objection to any construal of the conditional that would not include closure under conditionalization. Recently, in correspondence, Alan Hajek suggested that the above results should be taken as a major objection to the CCCP.[10]

But in fact the lack of closure under conditionalization does not come specifically from the construal of the conditional. It is characteristic of discourse in which (some) probabilities are themselves probabilistically assessed. Given the CCCP, conditional statements encode information about probabilities. But the probabilities themselves already involve obstacles to conditionalization, in a very ordinary, every-day context.

If I am asked to bet on the tossing on a die, I may easily say something like this to myself:

> It seems to me only as likely as not, that this die is fair
>
> *equivalently*:
>
> My subjective probability that the objective chance of each outcome is 1/6 equals ½.

This sort of talk which is surely quite ordinary, generates Moore statements, and lack of closure under conditionalization.

Consider first the following (not a Moore statement) said when about to toss a die:

> [1] The number six won't come up, but the chance that six will come up is 1/6.

On this occasion both conjuncts can be true. The die is fair, so the second conjunct is true, and when we have tossed the die we may verify that our prediction (the first conjunct) was true as well.

Moreover, [1] can be believed and equally, it can have a positive subjective probability. For example, if it is known that the die is fair, then the probability that [1] is the case equals 5/6.

In this sort of example we express two sorts of probability, one subjective and one objective. Are there some criteria to be met? Is there to be some harmony between the two?

There are some controversies about how they ought to be related to each other. I propose what I take to be an absolutely minimal constraint:

> **Minimal Harmony.** $P(ch(A) > 0) = 1$ implies $P(A) > 0$ *If I am sure that there is some positive chance that A then it seems to me at least a little likely that A.*

Could someone seriously, and rationally, violate this? What of the gambler who feels lucky, and says "Certainly there is some chance that the six will come up, but I am sure it won't!" Well, good luck!

To construct a Moore Statement we only need to modify [1] a little:

> [2] The number six won't come up, but the chance that six will come up is not zero
>
> ~Six & ~[ch(Six) = 0]

That [2] could be true we can argue just like we did for [1]. But [2] is a Moore Statement for it could not have subjective probability 1, by the following argument.

Assume that P([2]) = 1. Then:

1. $P(\sim Six) = 1$
2. $P(Six) = 0$
3. $P(\sim[ch(Six) = 0]) = 1$
4. $\sim[ch(Six) = 0]$ is equivalent to $[ch(Six) > 0]$
5. $P(ch(Six) > 0) = 1$

Here 2. and 5. together are a violation of Minimal Harmony.

This means also that [2] is a statement on which *you cannot conditionalize* your subjective probability, in the sense that if you do, your posterior opinion will violate Minimal Harmony.

So we have here another case where the space of admissible probability functions is not closed under conditionalization, and it involves no conditionals.

It appears to be the characteristic of probabilistic assessments of probabilities that they are subject to Moore's Paradox, which makes lack of closure under conditionalization ubiquitous.

Presumptions that bedevil us

How does "Or to If" get to tug at our heart strings at all, how does it get to feel plausible?

There may be a a major defect in the use of natural language examples of counterfactual conditionals. To provide an example of a conditional that others will accept as true, we provide reasons that bring it close to certainty. For example, to motivate "If Hoover had been a Communist, he would have been a traitor" we may assert that in those days, the enemy of the USA was, precisely, Communism.

However, the better the reason, the closer the conditional comes to a necessary implication, to C. I. Lewis' curly arrow, 'necessarily (not A or B)'. Is it possible that we will then tend to confuse the natural language arrow with the curly arrow?

If Peter must be in France or Italy then he must be in Italy if he is not in France – sure!

But now, try to do this: suppose that Peter is not in France, *without supposing* that Peter is certainly not in France. Can you do it? The instruction sounds rather like "Suppose that (Peter is in France and he might not be in France)". What do you do? Some views now current about epistemic modals would see this as the request to suppose the impossible.

There is a more immediate moral for issues in the logic of conditionals. When David Lewis addressed Stalnaker's Thesis and proved his celebrated triviality results, he began its motivation with

> presumably our indicative conditional has a fixed interpretation, the same for speakers with different beliefs, and for one speaker before and after a change in his beliefs. Else how are disagreements about a conditional possible, or changes of mind? (Lewis 1976: 301).

And he spelled out how this brings him to closure under conditionalization:

> Our question, therefore, is whether the indicative conditional might have one fixed interpretation that makes it a probability conditional for the entire class of all those probability functions that represent possible systems of beliefs.
>
> This class, we may reasonably assume, is closed under conditionalizing. (ibid. 302)

But as we have just seen, the relevant probability space cannot be closed under conditionalization. If it were, the classical conception of valid entailment would not be violated.

This result calls Lewis' philosophical presumptions, and not just the significance of his triviality result, into question.[11] It is far from evident that there will be a single algebra of propositions, expressed by the sentences in a common language, with these sentences having the same content for speakers whose subjective probabilities are different. Indeed, it is plausibly a function of conditionals to convey aspects of a speaker's doxastic attitudes.[12]

What about conditionals' context-dependence?[13] I submit that when the relevant probability is subjective, to model a person's opinion, the conditional statement and the conditional probability are context-dependent in the same way. Witness the parallel examples:

(A) Peter: If you jumped from here you would die.
Paul: No, if I jumped from here I would not die. For I would not jump without a safety net.

(B) Peter: You would most probably die if you jumped from here.
Paul: No, I would most probably not die if I jumped from here. For I would not jump without a safety net.

Here we can see in which sense disagreement is possible, the sense in which two speakers express disagreement by one saying "I think so!" and the other "I don't think so!". The context-dependent conditionals express the context-dependent subjective probabilities.

4. On Hajek's 'Chancy' Theory of Conditionals

When probability is introduced into the theory of conditionals, there are many places where problems about infinity can enter. My main purpose here is to argue that they can practically force us to leave possible world semantics for a more algebraic approach. The point is general, but I will focus on those problems as they appear for Alan Hajek's provocative views on counterfactuals.

Hajek's view of counterfactuals

Alan Hajek has argued that almost all natural language examples of counterfactual conditionals are false. Think of the historically important example of "If J. Edgar Hoover had been Russian, he would have been a Communist". There is much to warrant in this assertion, on the basis of Hoover's known character and ambitions. But is it enough, or even relevant? What are we to imagine about a Hoover born in Russia, growing up in a different social context with different friends, and perhaps even different sexual proclivities? He might have been a faithful Communist, eve a commissar, or he might not.

Hajek, who subscribes to a variant of David Lewis's concept of objective chance, proposes the view that "if A then B" is true (in the case in which A is false) if and only if the objective chance of B, given A, equals 1.

What would the logic be like? Difficulties with infinity

Since objective chance 1 is a form of necessity, we would expect this theory to be akin to C. I. Lewis's theory of strict conditionals, of the form "Necessarily, either not-A or B". But how can we think about nesting of conditionals, when linked to the chance function?

Suppose that the domain of the chance function includes a significant variety of propositions, even if not all.[14] Then such a proposition as that the chance that (if A then B) equals 0.3, may be true. And that would then be understood as the proposition that the chance, that the conditional chance of B on A equals 0.3, equals 1. Does (chance of what the chance is) that make sense?

On the face of it, it does make sense, but perhaps only in unusual conditions. Suppose I have two dice, one loaded and one fair. I have a Geiger counter and some radium, the timing of individual clicks exemplifying an indeterministic process. I decide that if the counter clicks 17 times in the next minute (which is unlikely) then I will toss the fair die, otherwise the loaded one.

In this situation is would seem to be true that there is a low objective chance that the next die toss has objective chance 1/6 of having outcome 6.

It would be interesting at this point to try and devise a semantics for a language with connectors &, ~, →, with models in which those sentences have sets of worlds as semantic values. The admissible valuations (linked to possible worlds) would then be assigning both chances and truth-values to propositions. The constraint to be met on those valuations would then be that, if F is assigned to A then (A → B) receives T if and only if the assigned chances are such that the conditional chance of B given A is defined and equals 1.

I will not speculate further on how this could be done. For as we will see now, possible world models of this sort seem to run into serious difficulties when infinity comes into play.

A chance function is a probability function, and so Hajek's view places conditionals in a dangerous environment. On the face of it, Hajek's theory will have exceptions to Modus Ponens. This was pointed out in discussion at Hajek's 2024 lecture at the University of California, Davis, by Rohan French:

[**One**] If $P(B|A) = 1$ then $P(A \cap -B) = 0$. But probability 0 does not imply falsity. So on this theory both A and (A → B) might be true and B false.

For example, the probability equals 0 that the mass of the moon, in kilograms, is a rational number. But it might still be true.

Perhaps worse still,

[**Two**] Collections of counterfactuals, each of them true on this theory, could be deductively consistent but transfinitely inconsistent.

As an example, let us take the action of placing a Geiger counter near radium sample from time t to t+1. We can refer to this action, for brevity, as the **Action.** The chance that the counter will click during that period, if the **Action** is performed, is according to quantum mechanics, positive but not 1.

Let E be the statement that the counter does not click within [t, t+1], and F(e) the statement that the counter clicks within interval e. In that case

P(If **Action** then F([t, t+1]) or E) = 1

But also, for each number x in [t, t+1]:

P(If **Action** then F([t, t+1] − [x]) or E) = 1

So on Hajek's theory, each of the conditionals

If **Action** then F([t, t+1] − [x]) or E

is true. The intersection of all these true conditionals is:

If **Action** then E

for taking the intersection loses each number in [t, t+1]. But probability of (If **Action** then E) is positive but low**.**

So on Hajek's theory, that conditional is not true. Therefore there is a collection of propositions each of which is true on this theory, but could not possibly be all true.

Try for regularity? At this point it might be tempting to say that the probabilities involved need to be strictly coherent – '*regular*', that is, to assign a positive probability to each non-empty proposition. Models could be modified by reducing the propositions modulo differences of measure 0, so that all propositions with probability 0 would be identified with the empty set. Motivation for this 'solution' could come from philosophy of physics. For example, in their famous quantum logic paper, Birkhoff and von Neumann argued that propositions stating measurement outcomes which differ only by measure 0 should be identified (Birkhoff and von Neumann 1936: 825). Motivation could also come from the practice in mathematics to focus on probability algebras (e. g. Kappos 1969, Birkhoff 1967: 261, example 2).

But alas, for our subject this idea encounters too many, practically insuperable difficulties already noted by Hajek (2012 and ?), and especially violations of symmetry (Parker 2012, 2019).

Proposal for a radical break

On Hajek's view, "the cat is on the mat" and "if the cat purrs then she is on the mat" have very different sorts of truth conditions. That is perhaps one instance of a ubiquitous phenomenon: moral realists, for example, hold that factual and moral judgements both have truth values, but their truth conditions are of a different sort.

To give substance to Hajek's view, or at least its core contention, I propose that we take this difference seriously, and that to do so, we break with possible world semantics.

There is no principled objection to reference to possible worlds in semantics. They acquired their role originally due to Stone's Theorem, that every Boolean algebra is isomorphic to an algebra of sets. The only step needed was to call the elements of those sets "possible worlds". After that, metaphysical intuitions could guide the construction of set theoretic representation of Boolean algebras with operators.

But what about a true break with the way of possible world semantics? *Possible worlds are not even fictions!* If they were fictions, they would have to play the role implied by the fiction, which is, to determine what is true and what is false. No, the only good role for set-theoretic representations will just be this: to show that there exist algebras of propositions of the sort we require for our purpose.

Prolegomenon: CE proposition algebras and the CCCP

We begin with a logic of conditionals that will play a useful role, although it is constructed in a way that does not go with Hajek's view at all.

Definition. A *CE proposition algebra* is a triple <1, F, →>, where F is a Boolean algebra with unit **1**, and → is a binary operator on F such that for all p, q in F:

I. $(p \rightarrow q) \wedge (p \rightarrow r) = (p \rightarrow . \ q \wedge r)$

II. $(p \rightarrow q) \vee (p \rightarrow r) = (p \rightarrow . \ q \vee r)$

III. $p \wedge (p \rightarrow q) = (p \wedge q)$

IV. $(p \rightarrow p) = \mathbf{1}$

I will call I. – IV. the *CE identities*.

Definition. A = <K, F, →, P> is a *CE algebra with probability* iff <K, F, →> is a CE proposition algebra and P is a probability measure whose domain includes F, such that, for all p, q in F:

$$P(p \rightarrow q) = P(q \mid p) \text{ if } P(p) > 0, \text{ and } = 1 \text{ otherwise}$$

Do CE algebras with probability exist? Yes, that there is a large variety of CE algebras with probability was a result proved in possible world semantics, closely related to Stalnaker's semantics for his logic of conditionals(van Fraassen 1976). The corresponding logic CE is a weakening of Stalnaker's, since the semantics lacks the ordering of worlds as a constraint on the selection function which defines the conditionals. For the details see Appendix 2.

Approaching truth algebraically

From now on we will ignore the set-theoretic representations of CE proposition algebras with probability. Their construction had no use except to establish the existence of the relevant algebras for our purpose.

The theory I will present now exemplifies Hajek's view of counterfactual conditionals as true, at least in the main, precisely if the corresponding conditional probability equals 1. But it will not be exactly what I take Hajek to have envisaged, at least in certain particulars. And the intuitions behind it relate to subjective probability rather than objective chance. So I will call this theory the Hajek* Theory.

This theory will pertain to a language with the usual syntax: propositional variables, connectors ~, &, v, →. Its models will be algebraic structures with the admissible interpretations being functions that link the truth of conditionals in the language to probabilities defined on that algebraic structure.

So we begin with a specific, but arbitrary CE algebra with probability **A** = <**1**, F, →, P>, and that syntax. F is a Boolean algebra with unit **1**, operators ∧, ∨ and −, zero element $\mathbf{0} = -1$. An interpretation || ..|| of this syntax in **A** is straightforward:

$$\| \sim A \| = \mathbf{1} - \|A\|$$

$$\|A \ \& \ B\| = \|A\| \wedge \|B\|$$

$$\|A \vee B\| = \|A\| \vee \|B\|$$

$$\|A \rightarrow B\| = \|A\| \rightarrow \|B\|$$

So far, then, interpretation assigns propositions as semantic values to sentences, but without any implication as yet for how truth-values are apportioned.

I will refer to the zero-degree fragment of this syntax as **Lat**, and define **A0** = {[[A]] : A in **Lat**}, which is a Boolean algebra. Intuitively, we can think of **Lat** as the set of empirical statements, like "the cat is on the mat". But as I will take into account later, some may be theory-infected, like "the iron bar is magnetic" which can have some relation to conditionals.

We need recourse to the algebraic notion of a *filter* that corresponds to the logical idea of a consistent theory, and to its dual. The following pertains to algebra **A.**

Definition. A subset X of F is a *proper filter* exactly if, for all elements p, q of F, if p is in X and $p \leq q$ then q is in X, and if p, q are in X then so is $(p \wedge q)$, and **0** is not in X.

Definition. A subset Y of F is a *proper ideal* exactly if, for all elements p, q of F, if p is in Y and $q \leq p$ then q is in Y and if p, q are in Y then so is $(p \vee q)$, and **1** is not in Y.

To complete the interpretation, there must be associated a proper filter – the *truth filter* -- the members of which are designated as the true propositions. This truth filter will have a dual ideal – the *falsity ideal* -- whose elements are designated as the false propositions. And finally, a sentence is true (respectively, false) exactly if its semantic value is a member of the truth filter (respectively, of the falsity ideal).

There is some leeway in what the truth filter may be, it is only strongly constrained but not determined by what we have so far.

Step One. Let **T** = {p in F: P(p) = 1}. It follows that **T** is a proper filter, from the properties of a probability function. **T** has a dual ideal, namely **U** = {**1** − p: p in **T**}, and this is a proper ideal.

Step Two. Let **T0** be the smallest filter which contains a certain subset of **A0**, and be a proper filter, which does not overlap **U**. (The choice of that subset of **A0** is otherwise unconstrained.) **T0** has a dual, the proper ideal **U0** = {**1**− p: p in **T0**}. From the special condition that **T0** does not overlap **U**, it follows that **U0** does not overlap **T**.

Step Three. Let **T*** be the smallest proper filter that contains both **T** and **T0**. Its dual ideal is **U*** = {**1** − p: p in **T***}.

We specify now that *what is true* on this interpretation (consisting of || ..|| and the specified truth filter) are precisely the sentences whose semantic values are in **T***, and *what is false*, precisely the sentences whose semantic values are in **U***. This assignment of truth-values to sentences will in general be a partial function only.

Theorem 6. T* exists and is a proper filter.

Proof. Since **A** is a CE algebra and **A0** is a Boolean subalgebra of **A**, proper filters **T** and **T0** exist.

We may note that filters are closed under finite meets, and can designate the meet of a finite set X as $\bigwedge X$.

Define Y = {p in F: there is a finite subset X of **T** ∪ **T0** such that $\bigwedge X \leq p$}. First, Y is a filter. Secondly, Y does not include **0**. For if it did, there would be elements p of **T** and q of **T0** such that $p \wedge q \leq \mathbf{0}$. Then, since **A** is Boolean, it would follow that $p \leq (\mathbf{1} - q)$ so p is in **U0,** which is ruled out by the non-overlap constraint.

Any filter that contains both **T** and **T0** must include Y. But secondly, Y itself is a proper filter. So Y is the smallest filter that contains **T** and **T0**, which is what **T*** was defined to be.

The Hajek Theory: what it is like*

Theorem 7. Any statement that is true in all interpretations of CE algebras is valid in the Hajek* Theory.

Proof. Since **A** is a CE algebra, the CE identities cannot be violated in **A**. Hence if ||A|| = **1** is a CE identity then ||A|| will belong to **T***.

Theorem 8. The rule of Modus Ponens is valid in the Hajek* Theory.

Proof. Suppose that p and $(p \rightarrow q)$ are both in **T***. Then $p \wedge (p \rightarrow q)$ is in **T***. But $p \wedge (p \rightarrow q) = (p \wedge q)$ is a CE identity. So $(p \wedge q)$,and hence also q, is in **T***.

So the theorems and 'simple' rules of the logic CE are all preserved in the Hajek* Theory. However, since the assignment of truth-values may only be partial, certain natural deduction rules that trade on sub-derivations (e.g. Conditional Proof, Disjunctive Syllogism) may be violated.

Here are some points about how the Hajek* Theory relates to Hajek's views. There certainly are conditionals that are true although the corresponding conditional probability is not 1. To begin, CE has the 'strong centering' principle in common with Stalnaker's and Lewis's logics of conditionals, so that if A and B are both true then $(A \rightarrow B)$ is true. Hajek has objections to strong centering. But those examples are not counterfactuals, so not contrary Hajek's view about counterfactuals and probability. However there can be further exceptions in specific models, due to relations between elements of **A0** and other elements. Intuitive examples would be like: "this iron bar is magnetic", a sentence in **Lat**, which implies "if iron filings are near this iron bar then they move toward it" which, being a conditional, is not in **Lat**. But the implication happens to hold in a specific model where

||this iron bar is magnetic|| ≤ || iron filings are near this iron bar → they move toward it||

so if the former is in the truth filter then so is the latter, regardless of what probabilities are assigned.

Theorem 9. The Hajek* Theory is not subject to transfinite inconsistencies

The proof is simple, but comes from a long story. In model **A** the truth filter is closed under finite meets, but not under arbitrary meets, and does not include **0**. So the sort of example of a transfinite inconsistent family of propositions, such as I gave above, cannot be part of a truth filter in any model.

But there is a more important, more fundamental point to be made. The truth filter, like the entire family of propositions, is part of the domain of the probability function P. When P is a non-trivial probability measure on a domain of the cardinality of the continuum, that domain cannot be a complete algebra: it is not closed under arbitrary meets and joins.

If we accept that propositions must be measurable elements of the algebra of propositions, then we must accept also that the algebra of propositions will in general not be complete. And the reason it generally is not complete is precisely avoidance of transfinite inconsistency.

The history behind this began at the very creation of measure theory. Lebesgue introduce his famous measure on the continuum, as a generalization of length, area, and volume. The question he asked immediately was: *can this measure be extended to all sets of points in the continuum?*

This was answered in the negative, to begin, by Vailati and Hausdorff: if that measure were defined on all subsets of the continuum it would violate geometric invariances.

Did this refute Lebesgue's theory? Not at all! It was accepted as a proof that in general, the domain of the measure cannot be the family of all subsets of a given set. There were more interesting negative results since then, of a more general sort. For example:

Theorem 10. (Birkhoff 1948: 187. Theorem 13)[15] If the Continuum Hypothesis is true, then no non-trivial countably additive measure can be defined for all subsets of the continuum, such that every point has measure 0.

So the domain of a probability measure is in general not a complete Boolean (Borel) algebra. (If the domain is complete then it includes all subsets if and only if it includes all the unit subsets.)

This does not entirely do away with problems about how we are to conceive of truth when infinity is in play. But I would submit that such problems look important only when we focus on set theoretic representations, *aka* possible world semantics.[16]

Consider a model of the Hajek* Theory which is set-theoretic, that is, the unit **1** is a set K (the possible worlds) and F is a special family of subsets of K. Now the truth filter **T*** is part of F, it is closed under finite meets, and does not contain the empty set. But **T*** is after all a family of subsets of K, and the intersection of this entire family may be empty. In that case there is no possible world in which all the true propositions are true!

Well, that is a model, it is a set-theoretic representation of another model of the Hajek* Theory which is just algebraic. Attend to the latter! Don't even think that the set-theoretic models are indispensable, let alone that they are specially important – in fact they are just a *pons asinorem.*

Don't let the familiarity and ease of possible world talk seduce you into allowing it to set the borders of your philosophical thinking, escape its tyranny!

More seriously: the problems that Hajek's view encounters disappear in an algebraic semantics, while they remain (as far as I can see) serious obstacles as long as one attempts to formulate it in possible world semantics.

Appendix 1. Hajek On A Probabilistic Moore's Paradox

In Hajek 2011 he describes his own example as a Moore Paradox in my sense (referring to the CCCP as the PCCP):

> If you are a Bayesian agent who seeks to conform to PCCP at all times, you are apparently unable to revise boldly and moderate your opinions regarding certain propositions. These propositions then have a curious status for you: you give them positive credence, but you can never learn them where learning is modeled by a bold and moderate rule. Borrowing terminology from (van Fraassen, 1984), they are 'Moore propositions' -- propositions that you cannot learn without violating a structural

constraint that is imposed on you (in this case, the upholding of PCCP). (Hájek, 2011, p. 12–13).

However, Hajek argument for this conclusion implicitly assumed that updating must be by conditionalization. Hajek's specific proposal for rational updating is ostensibly weaker. His proposal is that rational updating must be by a rule, and that this must be what he calls a *bold and moderate* rule. This requirement is the following:

Definition. A rule to update prior probability P on evidence E to function P_E is *bold* if and only if, for any P and for any E such that P(E) > 0, $P_E(E) = 1$.

Definition. A rule to update prior probability P on evidence E to function P_E is *moderate* if and only if, for any A that implies E, if P(A) > 0, then $P_E(A) > 0$.

Symmetry arguments establish, however, that if updating P on evidence E is by a rule (that is, depends solely on P and E) and is *bold*, then it is the rule of Bayesian Conditionalization. So the apparently minimal requirement that Hajek imposes is actually closure under conditionalization.

Anna and Krzysztof Wojtowicz (2024) present a critique of Hajek (2011). They note that Hajek's Moore-like counter-example to the CCCP also violates their proposal of a 'minimal meaning postulate':

> (IMP) If A is possible, and B is impossible, then A → B is impossible.

This is then spelled out as:

> It is not rational to hold probabilistic beliefs such that:
>
> (i) P(A) > 0;
>
> (ii) P(B) = 0 and;
>
> (iii) P(A → B) > 0. (Wojtowicz 2024: 8)

Violation of IMP is also a violation of the CCCP, since (i) and (ii) imply that $P(B|A) = 0$.

But IMP is stronger and would need further argument to warrant it. For IMP can be violated if we do not impose the CCCP. For suppose B is not void, just an area with zero probability, and that P(~A) > 1 as well as P(A) > 1. Then it is possible that the selection function s is such that {x: x is in ~A and s(x, A) is in B} has positive probability. It just happens that for a sufficiently large amount of the worlds in ~A, the 'nearest' A world is in B. Then P(~A & (A →B)) > 0, so P(A → B) > 0.

Appendix 2. Existence of CE proposition algebras with probability

The language under consideration, LCE, has as syntax a set of propositional variables and the connectors &, v, ~, →.

Definition. M is a *model* for LCE if M = <K, F, s, P>, where K is a non-empty set, F is a field of subsets of K, s (the *selection function*) is a function of K x F into F such that for all p, q in F

(a) s(x, p) is either a subset of p or Λ
(b) if x is in p then s(x, p) = {x}

and the set $(p \rightarrow q) = \{x \text{ in } K: s(x, p) \subseteq q\}$ is a member of F, and finally, P is a probability function whose domain includes F, and is such that $P(p \rightarrow q) = P(q|p)$ when defined.

This will be recognized as a variant of Stalnaker's semantics for the logic of conditionals, lacking the ordering of worlds that constrains the selection function there. Due to this omission, the triviality results, including Stalnaker's own, do not apply.

If M is a model for LCE then <K, F, →, P> is a CE algebra with probability. More precisely, and more familiarly, when LCE is interpreted in such a model in the usual way, each sentence A receives a proposition ||A|| as its semantic value, and we have:

$\| A\|$ is in F

$\|{\sim} A\| = K - \|A\|$

$\| A \,\&\, B\| = \| A\| \cap \| B\|$

$\| A \vee B\| = \| A\| \cup \| B\|$

$\| A \rightarrow B\| = \| A\| \rightarrow \|B\|$

The range of ||..|| is then included in the domain of probability function P, and is in effect a set-theoretic CE proposition algebra with probability.

Theorem (van Fraassen 1976: 278, 289-291)

If P is a probability measure defined on countable field F of sets on K then there is a model M = <K*, F*, s, P*> for LCE and a one-to-one map f of K into K* which maps F one-to-one into F* such that $P(p) = P^*(f(p))$ for all p in F.

In terms of language: *any coherent probability assignment* to the 0-degree sentences of LCE can be extended to a model of the entire language, while respecting the CCCP. This establishes the existence of a large variety of CE proposition algebras with probability.

REFERENCES

Billingsley, Patrick (1986) *Probability and Measure.* Second edition. New York: John Wiley & Sons.

Birkhoff, Garrett (1948) *Lattice Theory.* Revised edition. New York: American Mathematical Society.

Birkhoff Garrett (1967) *Lattice Theory.* Third edition. New York: American Mathematical Society.

Birkhoff, Garrett and John (1936) "The Logic of Quantum Mechanics". *Annals of Mathematics* Second Series, Vol. 37: 823-843.

Cieśliński, C., Horsten, L., and Leitgeb, H. (2023) „Axioms for typefree subjective probability". *Review for Symbolic Logic*. Online first.

Eells, Ellery and Brian Skyrms (1994) *Probability and Conditionals.* Cambridge: Cambridge University Press.

Fitelson, Branden (2015) “The Strongest Possible Lewisian Triviality Result”. *Thought: A Journal of Philosophy* 4: 69-74.

Goldstein, Simon and Paolo Santorio (2021) “Probabilities for Epistemic Modalities”. *Philosophers’ Imprint* 21 (33): 1-37.

Hajek, Alan (1989) “Probabilities of Conditionals: Revisited”. *Journal of Philosophical Logic* 18:423-428.

Hajek, Alan (2011) “Triviality Pursuit”. *Topoi* 30: 3-15.

Hajek, Alan (2015) “On the Plurality of Lewis’ Triviality Results”. In Loewer and Schaffer (2015: 425 – 445).

Hajek, Alan (2012) “Is Strict Coherence Coherent?” *Dialectica* 66: 411-424.

Hajek, Alan (no date) “Staying Regular?” *ms.*

Hajek, Alan and Ned Hall (1994) “The hypothesis of the conditional construal of conditional probability”. pp. 75-111 in Eels, Ellery and Brian Skyrms (1994) *Probability and Conditionals: Belief Revision and Rational Decision.* Cambridge U Press.

Hall, Ned (1994) “Back to the CCCP”. In Eells and Skyrms (1994: 141- 160).

Halmos, Paul R. (2018) *Lectures on Boolean Algebras.* New York: Dover Publications

Heyting, Arend (1930) “Die formalen Regeln der intuitionistischen Logik I”, *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, *Physikalisch-Mathematische Klasse*, 42–56. English translation “The Formal Rules of Intuitionistic Logic” in Mancosu (1998: 311–327).

Kappos, Demetrios A. (1969) *Probability Algebras and Stochastic Spaces.* New York: Academic Press.

Kaufman, Stefan (2009) “Conditionals, Right and Left: Probabilities for the Whole Family”. *Journal of Philosophical Logic* 38: 1-53.

Kaufman, Stefan (2023) “Bernoulli Semantics and Ordinal Semantics for Conditionals”. *Journal of Philosophical Logic* 52: 199-220.

Khoo, Justin (2022) *The Meaning of If.* Oxford: Oxford University Press.

Khoo, Justin and Matthew Mandelkern (2019) “Triviality Results and the Relation between Logical and Natural Languages”. *Mind* 128: 485-526.

Lewis, C.I., and Langford, C.H. (1932) *Symbolic Logic*, New York: Century Company; reprinted, New York: Dover Publications, 2nd edition, 1959, with a new Appendix III (“Final Note on System S2”) by C. I. Lewis.

Lewis, David K. (1976) “Probabilities of Conditionals and Conditional Probabilities”. *The Philosophical Review* 85 (3): 297-315.

Loewer, Barry and Jonathan Schaffer (2015) *A Companion to David Lewis.* New York: Wiley and Sons.

Mancosu, Paolo (ed.) (1998) *From Brouwer to Hilbert. The Debate on the Foundations of Mathematics in the 1920s*, Oxford: Oxford University Press.

Parker, Matthew (2012) "More Trouble for Regular Probabilities". *PhilSci Archive*. http://philsci-archive.pitt.edu/id/eprint/13063.

Parker, Matthew (2019) "Symmetry arguments against regular probability". *European Journal for Philosophy of Science* 9: 8: 1-21.

Santorio, Paolo (2023) Review of Khoo, *The Meaning of If.* Notre Dame Philosophical Reviews 2023.07.2. https://ndpr.nd.edu/reviews/the-meaning-of-if/

Stalnaker, Robert (1975) "Indicative Conditionals". *Philosophia* 5: 269-286.

Van Fraassen, B. C. (1976) "Probabilities of Conditionals", in W. Harper and C.A. Hooker (eds.) *Foundations of Probability and Statistics*, Volume l. Reidel: 261-308.

Wójtowicz, Anna and Krzysztof Wójtowicz (2024) "A Misleading Triviality Argument in the Logic of Conditionals". *Logic and Logical Philosophy.* http://dx.doi.org/10.12775/LLP.2024.015.

NOTES

---

[1] The strongest such triviality proof for the Extended Stalnaker Thesis is Fitelson (2015); see discussion of his results, with reference to many preceding ones, by Khoo and Mandelkern (2019: Appendix and Note 21). .

[2] It is possible to add the convention that P(B|S) = 1 if P(S) = 0. In that case P(~(S →B)) = 1-P(B|S) = 0.

[3] Heyting (1930) presents intuitionistic logic Hilbert-style. Note that his 2.13 (transitivity) and 2.2 together lead to weakening; compare his 2.27 to what I called the great law of implication.

[4] Contrary to the 'axiomatic approach' to probability, including its recent forms (e. g. Cieśliński, Horsten, & Leitgeb 2023).

[5] We can add a fourth condition, equivalent to the other three under these conditions, namely that for all p in P, p(b|a) = 1, if p(a) > 0. This is not needed for my argument below.

[6] Stalnaker (1975) begins with an example of an 'or to if' inference, but argues that it must be dealt with in pragmatics rather than semantics. As Santorio (2023) points out, the 'or to if' inference is closely related to the condition of Weak Sufficiency in Khoo (2022). Stalnaker pointed out that we need to account for the intuitive plausibility one way or another, and I will offer my diagnosis in the sub-section "Presumptions that bedevil us".

[7] The 'Or to If' inference sanctions the inference in general of (A → B) from (~A v B). It is also generally accepted (and I assume it here) that from (A & ~B), equivalent to ~(~A v B) we can infer ~(A → B).

[8] The result is also a corollary to various other results, derived differently. The strongest other such result, which assumes no more than I do here, is Ned Hall's First Result in Hall (1994).

[9] Let me just add a note here about another controversial principle for conditionals, which suggest that **inclusion** does not coincide with the other two consequence relations. I mean the 'And to If' inference: (A & B) implies (A → B). The 'And to If' principle holds for Stalnaker, Lewis, and my logic CE.

Alan Hajek has examples to argue that this inference does not preserve truth, specifically when A and B are not relevant to each other. It may well be true both that I will have a croissant for breakfast tomorrow, and that the sun will rise tomorrow. But the claim that, if so, the sun rises tomorrow if and only if I have a croissant for breakfast, sounds as if I could, like Joshua, get the sun to stay in its path. It sounds like that, but this may or may not be merely an implicature.

However, to support the 'And to If' principle we may note that:

[**certainty**] for all probability functions P, if P(A & B) = 1 then P(B|A) = 1, if defined

[**ordering**] for all probability functions P, P(A & B) ≤ P(B|A), if defined

[10] I imagine that similar sentiments pertain to the Reflection Principle, also much contested, which generates Moore statements similarly.

[11] Lewis' rhetorical question "Else how are disagreements about a conditional possible, or changes of mind?", and its accompanying reasoning, invite our speculations as to what Lewis presupposed. I suggest that first of all he thinks of rational changes of mind as having to be by conditionalization (which is today no longer plausible for reasoning with conditionals). And secondly I suggest that he does not allow for our understanding each other, when we make statements with additional linguistic functions besides fact-stating, such as expressing opinion or making inferential commitments.

[12] As a corollary, the Extended Stalnaker Thesis is in trouble here, for it presupposes closure under conditionalization. But this is just another of the 57 varieties of arguments to that effect.

[13] Alan Hajek (2015: 433) has argued that the CCCP cannot hold because the conditional is context-dependent and probability is not. That may be so for objective probability (chance) but it is in my view not the case for subjective probability.

[14] Hajek tends to limit the domain more than Lewis did, and has argued specifically that propositions that are non-measurable sets of worlds (e.g. in sense of Lebesgue measure) and propositions about free choices must not be assumed to be in the domain.

[15] The Continuum Hypothesis is that the cardinality of the continuum is the second infinite cardinal number. For larger context see the same theorem in Birkhoff (1967), Ch XI, sect. 7 Theorem 13 p, 266).

[16] It might be objected that every Boolean algebra has a(n essentially unique) minimal completion. This follows from Stone's Theorem, see Halmos (2018: 92-97). So we can raise the same problem for that completion of **A**. But, first of all, Stone's Theorem does not apply to Borel algebras (cf. Billingsley 1986: 18-19). Our models, having probability measures involved, must in general be Borel and not just Boolean Secondly, Stone's Theorem, which is also limited in other respects, does not automatically apply to Boolean algebras with additional operators.