Probabilities of counterfactuals are counterfactual probabilities[1]

Abstract. Suppose that, yesterday at noon, Maria considered flipping a fair coin, but didn't. What probability do you assign to "If Maria had flipped the coin, the coin would have landed heads"? Now suppose that, contrary to fact, Maria did indeed flip the coin. In that counterfactual scenario, what is the probability of "The coin will land tails"? The two questions sound strikingly similar. I argue that they sound similar because they are equivalent. The chance of a counterfactual "If A, would C" equals the chance of C, in the counterfactual scenario that A (and a similar principle holds for credence). This principle does better than similar principles that have been defended (like Skyrms' Thesis), avoids triviality, and gives us important clues for a semantics for counterfactuals.

1 Introduction

Suppose that, yesterday at 1pm, Maria considered flipping a fair coin. Eventually, she didn't. Consider the following counterfactual:

(1) If Maria had flipped the coin, it would have landed heads.

Here is a question: what is the probability of (1)? (For the moment, it won't matter whether we interpret probability as credence or chance.) Now suppose that, contrary to fact, Maria did indeed flip the coin. Consider the following claim, referring to this counterfactual scenario:

(2) The coin would have landed heads.

Here is a second question: in the counterfactual scenario where Maria does flip the coin at 1pm, what is the probability of (2)?

For now, I'm not interested in the specific answers to these questions. But I want to notice that they sound strikingly similar. In fact, they seem to be intuitively equivalent. If an interlocutor asked you the two in succession, you would feel as if they are asking you the same thing twice.

This paper investigates the relation between counterfactuals like (1) and probability. A number of theorists, starting with Ernest Adams (1976) and passing through

[1] *Acknowledgments suppressed for anonymous review.*

Brian Skyrms (1980b), have proposed bridge principles connecting the two. Here I defend a new bridge principle, based directly to the intuition elicited above. In short, this principle is: the probability of a counterfactual A $\Box\!\!\rightarrow$ C equals the probability of C, under the counterfactual supposition that A. In a simple slogan: probabilities of counterfactuals are counterfactual probabilities. The notion of probability in play can be understood both as chance and as rational credence, *modulo* appropriate qualifications. I show that this principle has several advantages over competing accounts: (i) it handles without effort the context dependence of counterfactuals, including so-called backtracking readings; (ii) it accommodates counterfactuals of any complexity; (iii) perhaps most importantly, it avoids a vexing triviality result for counterfactuals.

Let me also clarify what this paper does *not* do. It does not develop a new semantics for counterfactuals that accommodates the new bridge principles. As a result, I will not offer a tenability result, i.e. I do not show that the new bridge principle is vindicated by a semantics. But establishing the correct bridge principle, and showing that this principle avoids existing triviality results, is an important step in this direction.

I also have two additional goals.

First, I clarify the general status of of probability-counterfactuals bridge principles. These principles are generally formulated as concerning rational credences in counterfactuals (and, in particular, subjects' expectations of objective chances). I show that, once we distill their specific contribution, these principles turn out to concern chances of counterfactuals. For example, Skyrms' Thesis, which asserts that one's rational credence in a counterfactual A $\Box\!\!\rightarrow$ C should equal one's expectation of the conditional chance of C, given A, is equivalent (*modulo* the Principal Principle) to the claim that chances of counterfactuals equal the conditional chance of the consequent, given the antecedent. Similarly for other principles of this sort. So the debate about bridge principles between probability and counterfactuals concerns, at its core, constraints on chances.

Second, I link the debate on probabilities of counterfactuals to the recent debate on probabilities of indicatives. A tradition of work dating back to Ramsey (1926) links the rational credence in an indicatives conditionals to the conditional credence in the consequent, given the antecedent. But this intuitive view appears to be untenable, given so-called triviality results.[2]

One line of response to triviality for indicatives shows that we can preserve Stalnaker's Thesis in full generality if we replace the standard Bayesian operation for

[2]For references on Stalnaker's Thesis and the probabilities of indicative conditionals, see (among many) Lewis 1976, Van Fraassen 1976, Hájek & Hall 1994, Bradley 2012.

updating credences, i.e. conditionalization, with a new update operation (*[reference omitted]*). The new update operation agrees with conditionalization for all non-conditional propositions, but produces different results for conditionals and modal claims in general. This paper pursues a similar idea for chances. When we update a chance function with a counterfactual hypothesis, conditionalization is not, in general, the right procedure. The correct procedure involves rather shifting the so-called grounding argument of the chance function. The two procedures almost always agree, but end up producing different results in some key cases, which are the ones that give rise to triviality.

I proceed as follows. In §2, I provide some background about chance, the Principal Principle, and Skyrms' Thesis. In §3, I point out that Skyrms' Thesis is, at its core, a claim about chance. In §4, I give an informal overview of the ideas behind the project. In §5, I show that Skyrms' Thesis falls victim to a number of problems. Variants of Skyrms' Thesis that have been put forward in the literature (see Schulz 2017, Schultheis 2022, Khoo 2020) each avoid some of these problems; but none of them avoids all problems. In §6, I introduce my alternative principle, the Counterfactual Chance Thesis. In §7, I show how that the CCT avoids the problems that affect Skyrms' Thesis.

A point about notation: I use sans-serif uppercase letters ('A') as metavariables over both sentences and propositions. Context will disambiguate.

2 Background: chance and Skyrms' Thesis

This section introduces some basic ideas and principles. All my assumptions are grounded in classical work on chance (in particular, Lewis 1980). But I will make some departures in notation that are useful for my purposes.

2.1 Chance

I take chance to be a kind of objective probability, which is independent of particular subjects' epistemic states. I will be neutral on the question whether chances require indeterminism, though at times I might slip into indeterministic talk for simplicity.

I represent chances formally via chance functions. Following Meacham 2010, I take a chance function to have two arguments. The **object argument**, A, is simply the proposition that is assigned a chance. The **grounding argument**, G, is a proposition that specifies the information on the basis of which a probability value is assigned to the object argument. Following Lewis, I take the grounding argument to consist of a conjunction of two propositions.

(i) A complete history of the world up to a time t, understood as a proposition specifying all matters of particular fact in w up to t, $H_{w,t}$.

(ii) A complete theory of chance T_w. This is a theory that assigns a chance to every proposition at every nomologically possible history.[3] (Technically, a theory of chance is determined by a function T from worlds to theories of chance.)

Hence, the general form of a chance claim is, schematically:

$$ch_{T_w H_{w,t}}(\mathsf{A}) = x$$

Here is a simple example. Suppose that Maria will flip a fair coin today at 1pm. We have:

$$ch_{T_@ H_{@,now}} \text{ (Heads)} = 1/2$$

where: (i) $H_{@,now}$ is a proposition summing up actual history, up to now; (ii) $T_@$ is a theory that chance that entails that, at $H_{@,now}$, the chance of the coin landing heads is 1/2.

Later on in the paper, I will adopt a more liberal take on the history element of the grounding argument. In particular, I will allow that incomplete histories (what I call 'historical propositions') may enter the grounding argument. I will say more in due course.

On a more traditional construal, chance functions are one-place functions, indexed to a time. This construal is equivalent to the present one (provided that the role of time-indexing is understood correctly). But making the grounding argument explicit will be particularly helpful in stating my view.

Let me also highlight a connection between conditional chances and the grounding argument. Suppose that A is compatible with a history proposition $H_{w,t}$. On a natural view, the chance of C relative to the grounding proposition $T_w H_{w,t}$, conditional on A, equals the unconditional chance of C relative to the conjunction of A and $T_w H_{w,t}$ (which I will denote as '$T_w H_{w,t} + \mathsf{A}$'). It is helpful to make this principle explicit:

Limited Equivalence

$ch_{T_w H_{w,t}}(\mathsf{C} \mid \mathsf{A}) = ch_{T_w H_{w,t}+\mathsf{A}}(\mathsf{C})$, for all A compatible with $T_w H_{w,t}$

[3] Lewis (1980) characterizes a complete theory of chance T as a series of history to chance conditionals, i.e. conditionals of the form $\ulcorner$ If H_i, then the chance of A would be $x\urcorner$.

This connection captured by Limited Equivalence is highly intuitive, but it cannot be vindicated if we assume that the grounding argument has to include a full history, since in general the conjunction $H_{w,t} \wedge \mathsf{A}$ does not specify a full history. As I mentioned, in what follows I will relax exactly this requirement. As a result, I will endorse Limited Equivalence.[4]

I also assume the Principal Principle (PP). The PP imposes a constraint on the priors of a rational subject: roughly, it says that the priors of the subject, conditional on the information about the chance of a proposition A (and non-offending evidence) should assign A a credence equal to the chance. More formally: let *up* be a credence function that captures the ur-priors of a rational subject, i.e. their priors before they receive any information. Then the Principal Principle says:

Principal Principle. $up(\mathsf{A} \mid ch(\mathsf{A}) = x \wedge E) = x$, with E admissible

In words, this says that a rational subject's ur-prior in A, conditional on the information that the chance of A is x and some admissible evidence E, should also be x. The notion of admissible evidence has generated substantial discussion. For current purposes, suffice it to say that inadmissible evidence is evidence that swamps information about the chances.[5]

2.2 Counterfactuals and Skyrms' Thesis

The term 'counterfactual' is notoriously problematic. Here I use it to pick out all *would*-conditionals, including *would*-conditionals in which the reference time of the antecedent and the consequent are in the future (so-called 'future-less-vivid' conditionals).[6]

Following standard accounts (Stalnaker 1968, Lewis 1973), I assume that counterfactuals have a semantics based on **selection functions**. Formally, selection functions work slightly differently on Stalnaker-style and Lewis-style semantics. For Stalnaker, a selection function is a function $s : W \times \mathcal{P}(W) \mapsto W$ from a pair of a world and a proposition to a world. For Lewis[7], a selection function is a function $\mathcal{W} \times \mathcal{P}(\mathcal{W}) \rightarrow \mathcal{P}(\mathcal{W})$ that maps a world and a proposition to a proposition. In both cases, the intuition behind the semantics is that a counterfactual $\mathsf{A} \mathrel{\Box\!\!\rightarrow} \mathsf{C}$

[4]Notice that the requirement that the proposition A is compatible with $H_{w,t}$ is crucial. Things will work very differently in other cases. See §6 for discussion.

[5]See Meacham 2010 for an excellent analysis of admissibility and a discussion about how to reformulate the Principal Principle without mentioning admissibility.

[6]So-called future-less-vivid conditionals have been introduced as a semantic category by Iatridou (2000). See, among others, Arregui 2007 for extended discussion of conditionals of this kind.

[7]Or at least, for the version of Lewis that accepts the limit assumption.

selects a set of (most similar) worlds that verify the antecedent, and evaluates that consequent at that set.

I will come back to the precise definition of selection functions in §6. For the moment, I don't need to choose between the two formulations, and can rely on an informal statements of the truth conditions of counterfactuals. I represent counterfactuals schematically using the traditional boxarrow '$\Box\!\!\rightarrow$' symbol. I also add an 's' subscript to mark the fact that the truth conditions of counterfactuals depend on a choice of selection function. (The subscript will be suppressed where it won't matter.)

> **Selection semantics for counterfactuals (informal)**
> $\mathrm{A} \Box\!\!\rightarrow_s \mathrm{C}$ is true, relative to w and s, iff the worlds in $s(w, \mathrm{A})$ verify C

Giving a precise semantics for counterfactuals goes beyond the purposes of this paper. So I will stick with these informal truth conditions throughout.

A question that has gained traction in recent literature is how counterfactuals relate to probability (see e.g. Moss 2013, Schulz 2017, Khoo 2020, Schultheis 2022). All contemporary views start from a classical proposal, originally put forward by Bryan Skyrms (1980b). Informally, Skyrms' Thesis says: your credence in $\mathrm{A} \Box\!\!\rightarrow \mathrm{C}$ should equal your expectation of the past chance of C, given A. Formally:

> **Skyrms' Thesis.**
> Let $ch_{T_w H_{w,t}}$ be the chance function at w and t. For all A, C, and for all rational credence functions cr_{E_t} such that E is the subject's total evidence at t:

$$cr_{E_t}(\mathrm{A} \Box\!\!\rightarrow \mathrm{C}) = \sum_{w \in W} cr_E(w) \times ch_{T_w H_{w,t^-}}(\mathrm{C} \mid \mathrm{A})$$

Notice that Skyrms' Thesis involves a shifted time-index on the chance function: we should consider not the current chances, but rather the chances that obtained at some point in the past. In particular, we should consider the chances that obtained 'just before' the truth status of the counterfactual antecedent was settled. This introduces an element of vagueness in Skyrms' Thesis—exactly what time should we pick out?—but this vagueness is generally taken to be tolerable.

Let me emphasize the driving idea behind Skyrms' Thesis. To evaluate a counterfactual, we 'rewind' the actual course of history, up to a point at which the counterfactual antecedent has probability greater than zero. When we reach that point, we conditionalize the chance function at that time on the counterfactual antecedent, and assess the chances of the consequent. Crucially, the evaluation of counterfactuals goes by chances that obtain *on the actual timeline*. This idea will be the source of some of the problems with the Thesis.

For illustration, consider again (1), again uttered after Maria failed to flip a coin yesterday at 1pm:

(1) If Maria had flipped the coin, it would have landed heads.

We are certain (suppose) that the coin is fair. So we are certain that, at the relevant time (i.e. a time 'just before' Maria decided not to flip) the chance of heads conditional on flipping was 1/2. As a result, in this case Skyrms' Thesis gives us:

$$cr_{E_{today}}(\text{Flip} \mathrel{\Box\!\!\rightarrow} \text{Heads}) = ch_{T_@ H_{@,<1pm}}(\text{Heads} \mid \text{Flip}) = 1/2$$

I.e., our credence in (1) should equal the chance of heads, conditional on the coin being flipped, that obtained yesterday just before 1pm. Hence we get the intuitive verdict that our rational credence in (1) should be 1/2.

3 A the core of Skyrms' Thesis: the Chancy Equation

Before proceeding with my criticism of Skyrms' Thesis, it is useful to notice the relationship between this Thesis and a principle about chances.

Fact 1. Given the Principal Principle, Skyrms' Thesis is equivalent to:[8]

Chancy Equation (CE). $ch_{T_w H_{w,t}}(\text{A} \mathrel{\Box\!\!\rightarrow} \text{C}) = ch_{T_w H_{w,t^-}}(\text{C} \mid \text{A})$

(The proof of **Fact 1** is in the appendix.)

Informally, the Chancy Equation says that the chance of a counterfactual A $\Box\!\!\rightarrow$ C (at w, t) equals the conditional chance of C given A (at w, t^-). Given that the Principal Principle is widely accepted on independent grounds[9], we can think of the Chancy Equation as capturing the distinctive content of Skyrms' Thesis. So, even though Skyrms' Thesis bridges credence and chance, at its heart there is a principle entirely about chance.

[8] The label 'Chancy Equation' comes from Williams 2012. Williams also points out one direction of the equivalence: given the Principal Principle, we can get to the Chancy Equation.

[9] Several theorists hold that the Principal Principle should be replaced by a different principle, usually called the 'New Principle' (Hall 1994, Lewis 1994). Adopting the New Principle here would not make a difference.

4 The project: representing chances under a hypothesis

Before delving into the technical details, it is useful to state, in nontechnical term, the general idea of this paper.

A lot of our reasoning with chances involves a notion of chance under a hypothesis, or **hypothetical chance**. To illustrate what I have in mind, consider the following case:

> **Three coins.** Tomorrow, Sylvia will toss one of the three coins from her pocket, chosen at random via a chancy process. Coin 1 is fair; coin 2 has a 75-25 bias towards heads; coin 3 has a 75-25 bias towards tails.

Consider now the proposition expressed by (3), as uttered before it is settled which coin is tossed:

(3) The coin that is tossed will land heads.

When reasoning about this case, we can consider the current chance of the proposition expressed by ((3)). But we can also consider its chance under the hypothesis that coin 1, coin 2, or coin 3 is tossed. This notion of "chance under a hypothesis" is what I call 'hypothetical chance'. To have some precise notation, let us represent the hypothetical chance of B, updated with the information that A, as $ch^{\mathsf{A}}_{t_w H_{w,t}}(\mathsf{B})$.

Standard accounts of chance make no mention of a notion of 'hypothetical chance'. This is because hypothetical chances are simply identified with conditional chances. In compact notation:

$$ch^{\mathsf{A}}_{T_w H_{w,t}}(\mathsf{B}) = ch_{T_w H_{w,t}}(\mathsf{B} \mid \mathsf{A})$$

For example, the chance of (3), under the hypothesis that coin 1 is the coin tossed, is simply identified with the conditional chance of (3), conditional on the information that coin 1 is tossed.

This way of treating hypothetical chances dovetails with a claim that is often made explicitly about chance. This is the claim that the chance function of the world 'evolves by conditionalization' (see e.g. Lewis 1980). This claim concerns not chances under a hypothesis, but rather the connection between the values of the chance function representing the chances of a world at different times. It says that the chance function of a world at a time t_1 equals the chance function at a previous time t_0, conditionalized on the events that have taken place between t_0 and t_1.

How does this connect to conditionals?[10] As we saw, once we distill the distinctive contribution of Skyrms' Thesis, it amounts to the Chancy Equation, i.e. the claim that chances of counterfactuals are conditional chances:

[10]This discussion is indebted to *[reference omitted for blind review]*.

$$ch_{T_wH_{w,t}}(\mathsf{A} \mathrel{\Box\!\!\rightarrow} \mathsf{B}) = ch_{T_wH_{w,t^-}}(\mathsf{B} \mid \mathsf{A})$$

Together with the identification of hypothetical chances and conditional chances, this gives us that the three notions—chance under a hypothesis, conditional chance, and chance of a counterfactual—are equivalent:

$$ch_{T_wH_{w,t^-}}(\mathsf{B} \mid \mathsf{A}) \;=\; ch^{\mathsf{A}}_{T_wH_{w,t^-}}(\mathsf{B}) \;=\; ch_{T_wH_{w,t}}(\mathsf{A} \mathrel{\Box\!\!\rightarrow} \mathsf{B})$$

The main theoretical claim of this paper is that this identification is incorrect. This claim by itself is not too surprising, in the light of the problems that plague Skyrms' Thesis. But, differently from standard critics of Skyrms' Thesis, I take the problem to lie not with the linking of conditionals and hypothetical reasoning, but rather with the identification of hypothetical chances and conditional chances. I deny that updated chances are, in general, accurately captured by conditional chances. So we have:

For some A, B, w, and t:

$$ch_{T_wH_{w,t}}(\mathsf{B} \mid \mathsf{A}) \;\neq\; ch^{\mathsf{A}}_{T_wH_{w,t}}(\mathsf{B})$$

Hence: while useful, conditional chances are not, in general, a good tool to capture chances under a hypothesis. A different notion is needed.

Let me add two qualifications.

First, in *most* cases, the value of updated chances and conditional chances coincides. But there are going to be some cases where the two diverge. Just cases of this sort give rise to triviality results for Skyrms' Thesis.

Second, all that I say is compatible with Lewis's claim that the actual chance function evolves by conditionalization—in fact, I myself endorse this claim. For clarity: when chancy events are actually settled, the way that the chance function updates is always appropriately captured by conditionalization. The divergences between updated chances and conditional chances emerge only when we consider chances under a hypothesis. The cases of divergence are nevertheless significant, since they are the ones giving rise to triviality results and other problems.

In the remainder of the paper, I first rehearse some arguments that show that Skyrms' Thesis is problematic, and then develop my positive theory. After this, I will show how, on the new picture, updated chances and conditional chances can diverge in a restricted range of cases.

5 Three problems with Skyrms' Thesis

In this section, I present three problems for Skyrms' Thesis. Some of these problems are known, and more recent variants of Skryms' Thesis manage to avoid some of them. At the same time, no variant of Skyrms' Thesis manages to address all of them. So it's useful to use Skyrms' Thesis as my foil.

5.1 Problem #1: Morgenbesser cases

The first problem is that (as both Schulz 2017 and Khoo 2020 point out) Skyrms' Thesis appears to be falsified by so-called Morgenbesser scenarios.[11] Here is an example:

> **Coin**. Alice is about to flip a fair coin, and offers Bob a bet on heads. Bob refuses. Alice flips the coin, which lands heads.

In this scenario, (4) is judged to be true (and hence should presumably be assigned credence 1):

(4) If Bob had bet, he would have won.

But Skyrms's Thesis predicts that one's rational credence in (4) should be 1/2. The reason is that, at times before Bob decided whether to make the bet, the chance of Bob winning, conditional on his betting, was 1/2 (and this is known).

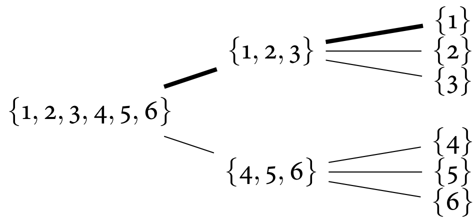
The point generalizes. Morgenbesser scenarios are cases where the antecedent of a counterfactual has a reference time that precedes an indeterministic event that affects the truth value of the consequent. In all these cases, the verdict of Skyrms' Thesis can come apart from intuitive judgments.

5.2 Problem #2: complex counterfactuals

The second problem is, to my knowledge, novel, and involves complex counterfactuals. Consider the following scenario.

> **Game show**. A game show involves randomly selecting a number between 1 and 6. To make things more dramatic, the selection works as follows. First, it is determined (via a chancy process) whether the number will be selected among 1, 2, or 3, or among 4, 5, and 6. Then, the final number is selected (again via a chancy process). Suppose that, in actuality, the number 1 is selected (as indicated by the thick line).

[11] Morgenbesser scenarios are so-called because their discovery is attributed to Sydney Morgenbesser. For discussion, see Slote 1978, Lewis 1979, Hiddleston 2005.



Now, consider the following:

(5) If 5 had been selected, then, if an even number had been selected, it would have been 4.

What credence should we assign to (5)? At least on one salient reading, the answer is: 1/2. (Reason: if 5 had been selected, then on the first round the set including 4, 5, and 6 would have been selected; and, on the assumption that an even number is selected from that, there is a 1/2 chance that 4 is selected.)

Skyrms' Thesis cannot yield this verdict. Since the chances are known in this case, from Skyrms' Thesis we get:

(6) $cr_{E_t}(\text{Five} \mathrel{\square\!\!\rightarrow} (\text{Even} \mathrel{\square\!\!\rightarrow} \text{Four})) = ch_{T_@ H_{@,t^-}}(\text{Even} \mathrel{\square\!\!\rightarrow} \text{Four} \mid \text{Five})$

The key issue is how to assign a conditional chance to the term on the right-hand side. So far as I can see, how this will work formally is not fully settled, given what Skyrms says. In particular, to get a precise verdict we would need to make assumptions about the interaction of the grounding argument, the proposition being conditionalized on, and the time index.[12]

However, even without fixing on a precise verdict, we have enough to show that Skyrms' Thesis won't produce the right outcome. Skyrms' Thesis requires us to consider conditional chances that obtain at some point in the past on the actual timeline. But notice that there is no point in time at which the chance of 4 being picked, conditional on any combination of the two antecedents in (5), is 1/2. To get the right verdict, we have to consider not the chances that obtain on the actual timeline, but the chances that obtain on the counterfactual timeline that leads to the picking of 5.[13]

[12]Since the proposition Five is compatible with the grounding argument $T_@ H_{@,t^-}$, via Limited Equivalence we have that the right-hand side of ((6)) equals $ch_{T_@ H_{@,t^-}+\text{Five}}(\text{Even} \mathrel{\square\!\!\rightarrow} \text{Four})$. From here, applying Skyrms' Thesis again, we obtain $ch_{T_@(H_{@,t^-}+\text{Five})^-}(\text{Four} \mid \text{Even})$, where $(H_{@,t^-}+\text{Five})^-$ is the historical proposition $H_{@,t^-}$ 'rewound' to a slightly earlier time. So far as I can see, nothing that Skyrms says and nothing in Skyrms' Thesis determines what this proposition is. Hence I think Skyrms' Thesis doesn't yield a clear verdict here.

[13]To be sure, there might be ways to save the letter of Skyrms' Thesis by supplementing it with spe-

Of course, one route for the proponent of the driving intuition is simply to reject that Skyrms' Thesis, or any similar claim, should apply to complex counterfactuals. But, as I will show, a comprehensive account that subsumes the results of Skyrms' Thesis and also handles cases like (5) is available.

5.3 Problem #3: triviality

The third problem is that, as Williams 2012 has shown, given minimal assumptions Skyrms' Thesis leads to triviality.

To prove Williams' triviality result, we start by assuming a principle of closure of chances:

> **Closure.** If $ch_{T_wH_{w,t}}(\bullet)$ is a chance function, then, for any A that is compatible with $T_wH_{w,t}$, $ch_{T_wH_{w,t}}(\bullet \mid A)$ is a (conditional) chance function.

Closure says that, by conditionalizing a chance function on a proposition that is compatible with the grounding argument, we still get a chance function. (As Williams points out, the chance functions don't have to be true of the same world. They merely have to be chance functions for *some* world.)

As we pointed out above (and as Williams shows), from the Principal Principle and Skyrms' Thesis, we can derive the following:

> **Chancy Equation (CE).** $ch_{T_wH_{w,t}}(A \mathrel{\Box\!\!\rightarrow} C) = ch_{T_wH_{w,t^-}}(C \mid A)$

As Williams points out, there is a close analogy between CE and Stalnaker's Thesis, i.e. the thesis that credences in indicative conditionals should equal conditional credences. As a result, we can run a proof analogous to Lewis's (1976) original triviality proof to trivialize the Chance Thesis.

Let w and t be an arbitrary world and time, let A and B be a y two propositions, and let ch be any chance function. We have:

i. $ch_{T_wH_{w,t}}(A \mathrel{\Box\!\!\rightarrow} C)$
ii. $ch_{T_wH_{w,t}}(A \mathrel{\Box\!\!\rightarrow} C \mid C) \times ch_{T_wH_{w,t}}(C) + ch_{T_wH_{w,t}}(A \mathrel{\Box\!\!\rightarrow} C \mid \neg C) \times ch_{T_wH_{w,t}}(\neg C) =$

cific assumptions about the grounding argument, the conditionalized proposition, and the time index. In short, the assumptions are: (i) a conditional chance $ch_{T_wH_{w,t^-}}(C \mid A)$ is equivalent to a nonconditional chance with a strengthened grounding argument, where the history proposition is conjoined with the proposition conditionalized on $ch_{T_wH_{w,t^-} \wedge A}(C)$ (see discussion of this point in §2.1); (ii) when applying Skyrms' Thesis to a term whose grounding argument has been shifted in this way, we start shifting the time indexing back from the time A. This will rescue Skyrms' Thesis, but it does so by crucially dropping the idea that we need to consider the past chances on the actual timeline. So far as I can see, this is merely a less formalized variant of my account.

iii. $ch_{T_wH_{w,t}+\mathrm{C}}(\mathrm{A} \mathrel{\Box\!\!\rightarrow} \mathrm{C}) \times ch_{T_wH_{w,t}}(\mathrm{C}) + ch_{T_wH_{w,t}+\bar{\mathrm{C}}}(\mathrm{A} \mathrel{\Box\!\!\rightarrow} \mathrm{C}) \times ch_{T_wH_{w,t}}(\neg\mathrm{C}) =$

iv. $ch_{T_wH_{w,t^-}+\mathrm{C}}(\mathrm{C} \mid \mathrm{A}) \times ch_{T_wH_{w,t}}(\mathrm{C}) + ch_{T_wH_{w,t^-}+\bar{\mathrm{C}}}(\mathrm{C} \mid \mathrm{A}) \times ch_{T_wH_{w,t}}(\neg\mathrm{C}) =$

v. $1 \times ch_{T_wH_{w,t}}(\mathrm{C}) + 0 \times ch_{T_wH_{w,t}}(\neg\mathrm{C}) =$

vi. $ch_{T_wH_{w,t}}(\mathrm{C})$

Skyrms's Thesis is used in the passage from line (iii) to line (iv). After that, we exploit a basic fact about chance functions: if a proposition A is entailed by the grounding argument G of ch_G, then, for any B for which conditional chances are defined, $ch_G(\mathrm{A} \mid \mathrm{B}) = 1$. Assuming the interaction between grounding arguments and chance functions that we have postulated in §2, this fact follows simply by the mathematics of probability.

Williams' triviality argument appears to establish that, for any w, t, and ch, $ch_{T_wH_{w,t}}(\mathrm{A} \mathrel{\Box\!\!\rightarrow} \mathrm{C}) = ch_{T_wH_{w,t}}(\mathrm{C})$: i.e. that the chance of any counterfactual at t and w is equal to the chance of its consequent at t and w. This is absurd. For a simple counterexample: the chance of *The coin would land heads* is not the same as the chance of *If heavily I rigged this coin towards tails, the coin would land heads*.[14]

What has gone wrong? Let me start taking some steps towards a diagnosis. The passage that appears problematic is just the one from line (iii) to line (iv). Consider an example:

> **Two coins.** Sarah is going to flip one of two coins A and B, picked at random. Coin A is fair. Coin B has a 100% bias towards heads.

Let 'A-Flip' stand for the proposition that coin A is flipped, and 'Heads' be the proposition that whichever coin is flipped lands heads. Now, consider the chance of the counterfactual A-Flip $\mathrel{\Box\!\!\rightarrow}$ Heads, under the hypothesis that the flipped coin lands heads. Since the proposition that the coin lands heads is compatible with current history, the relevant chance can be expressed as:

(i) $ch_{T_wH_{w,t}+\text{Heads}}(\text{A-Flip} \mathrel{\Box\!\!\rightarrow} \text{Heads})$

What is this chance? While this judgment might not be easy, I put it to you that one intuition is clear enough: the relevant conditional chance is less than 1. Suppose that the coin will land heads. Even on this supposition, the chance that if A was flipped, it would land heads should be less than 1—since it might be that B is the coin that ends up being flipped in actuality, and that tails would have come up in case of an A-flip.

[14] In addition to being subject to obvious counterexamples, the conclusion of the triviality argument would severely limit the range of the chance function, as pointed out by Lewis in his original triviality proof for credence (1976).

If this is right, then we have an interesting divergence from the predictions of the Chancy Equation. The conditional chance in (i) is intuitively less than 1. But, via the Chancy Equation, (i) equals

(ii) $ch_{T_w H_{w,t^-}+\text{Heads}}(\text{Heads} \mid \text{A-Flip})$

As we saw, (ii) equals 1 by math. Hence chances of counterfactuals and conditional chances don't always line up, contrary to Skyrms' Thesis.

6 The Counterfactual Chance Thesis

6.1 The Counterfactual Chance Thesis, take 1

Recall that, given the Principal Principle, Skyrms' Thesis is equivalent to CE.

Chancy Equation (CE). $ch_{T_w H_{w,t}}(\mathrm{A} \mathrel{\Box\!\!\rightarrow} \mathrm{C}) = ch_{T_w H_{w,t^-}}(\mathrm{C} \mid \mathrm{A})$

Given that Skyrms' Thesis is problematic, we should look into replacing CE. After that, we will be in a position to formulate a new principle linking credence and chance.

Consider again the intuition that I elicited in the introduction. Asking for the probability of (1), or asking for the probability of (2) under the supposition that the coin had been flipped, appear to be two ways of asking the same thing.

(1) If Maria had flipped the coin, it would have landed heads.

(2) The coin would have landed heads.

My proposal builds on exactly this idea. I suggest that the chance of a counterfactual A $\Box\!\!\rightarrow$ C equals the chance of its consequent, C, on the counterfactual supposition that A obtains. I am going to call this claim 'Counterfactual Chance Thesis', or CCT.

To implement the idea formally, I have to make some changes in the apparatus developed in §2. I introduce two of these changes in this section, and I formulate a first version of my suggested principle, the Counterfactual Chance Thesis. The third change may or may not be required, depending on one's views about closeness and the metaphysics of chance. I discuss it in §6.3, and I suggest a generalized version of the CCT there.

The first change is that we need a more liberal conception of the grounding argument of chance functions.[15] In §2, I said that the grounding argument involves a

[15] Thanks to *[name omitted for anonymous review]* for discussion here.

history proposition $H_{w,t}$, which I construed as a proposition that specifies all matters of particular fact at world w, up to time t. Here I need a more permissive view. In addition to complete histories, the grounding argument may also include what I call 'historical propositions', i.e. propositions about matters of particular fact at a world and up to a time, which may or may not be complete. This is a small change with respect to the traditional conception (and one that comports with a general way of thinking of chance functions; see Meacham 2010 for discussion). I don't see any drawbacks to it, so I embrace it without qualms.

Second, we need a change in the definition of selection functions. On standard accounts, selection functions map a world (the actual world) and a proposition (the counterfactual antecedent) to a world or a set of worlds (the counterfactual scenario). I require that selection functions take as input not a world, but a historical proposition. So, on the new account, a selection functions is functions $s : \mathcal{P}(W) \times \mathcal{P}(W) \mapsto \mathcal{P}(W)$ from a pair of propositions to a proposition. Also in this case, the change appears harmless, and indeed is a natural move once we are in a framework that allows that it might be indeterminate which world is actual, and hence that there might not be a fully settled actual history. For current purposes, I leave open whether selection functions of the new kind can be defined from old selection functions.[16]

Now I can state a first version of the Counterfactual Chance Thesis, or CCT:

Counterfactual Chance Thesis (first take).
$ch_{T_w H_{w,t}}(\mathsf{A} \mathrel{\Box\!\!\rightarrow}_s \mathsf{C}) = ch_{T_w s(\mathsf{A}, H_{w,t})}(\mathsf{C})$

In words, CCT says that the chance of $\mathsf{A} \mathrel{\Box\!\!\rightarrow}_s \mathsf{C}$ relative to the grounding argument $T_w H_w$ equals the chance of C, relative to a 'shifted' grounding argument $T_w s(\mathsf{A}, H_{w,t})$. (Notice that only the history component of the grounding argument is shifted. The point of contention connected to the third change is exactly whether we need to shift also the theory of chance. See §6.3 for discussion.)

For illustration, consider once more (1), as uttered about the scenario in which Maria refrained from flipping a coin at 1pm yesterday:

(1) If Maria had flipped the coin, it would have landed heads.

According to CCT, the chance of (1) equals $ch_{T_@ s(\mathrm{Flip}, H_{@,\mathrm{today}})}(\mathrm{Heads})$, i.e. the chance

[16] Let me just notice that there is an easy way to do so. Quite simply, we can 'lift' old school selection functions into new ones by requiring that, for any proposition A, $s(\mathsf{A}, H_{w,t})$ returns the union of the values of $s(w, H_{w,t})$, for every w in A.

$$s(\mathsf{A}, H_{w,t}) = \bigcup_{w \in \mathsf{A}} s(w, H_{w,t})$$

of the coin landing heads in the counterfactual scenario that we reach via the selection function, starting from the actual history up to today, and making the counterfactual supposition that Maria had flipped the coin. Skipping over some details (which I discuss below), given that the coin is fair, in the scenario selected the chance of Heads will be 1/2. Hence the CCT's verdict for (1) is 1/2, as is plausible.

6.2 Exploring the Counterfactual Chance Thesis

Let me discuss some noteworthy features of CCT.

Shifting the grounding argument. CCT works by shifting the grounding argument of the chance function. The chance of a counterfactual relative to grounding argument G equals the chance of its consequent, relative to a shifted grounding argument. In a sense, the Chancy Equation also worked by shifting a parameter, since it required to shift backwards the time at which chances are evaluated. But there are two important differences. First, CCT shifts a modal rather than a temporal parameter. Second, as we saw, the temporal shift involved in the CE is not precisely specified. The only constraint is that it has to be a time 'just before' the time of the antecedent. Conversely, CCT provides a precise procedure for selecting a new historical proposition—or at least, a procedure that is no less precise than the selection procedure we use in the semantics of counterfactuals.

Parametrization. CCT is parametrized to a choice of selection function. In particular, we use the same selection function on the two sides of the equal sign.

One desirable side-effect of this is that the context dependence and the vagueness of counterfactuals are not a problem for CCT. Since Lewis 1973, it is acknowledged that many ordinary counterfactuals are context-dependent, or vague, or both. Let's focus on context dependence for the purposes of illustration. One classical example of context dependence is (7):[17]

(7) If Caesar had been in command in Korea, he would have used catapults.

(7) can have a true or a false reading, depending on whether we allow that in the closest worlds where he is in command, Caesar has access to ancient or modern weapons. This kind of context dependence is usually explained via positing that context contributes to fixing the selection function. This context dependence is a hurdle for CE. Theorists often postulate that there is a 'canonical' reading of counter-

[17] The example is discussed by Lewis (1973, p. 66) and attributed to Quine.

factuals, which is more tightly connected to the chances, and claim that CE applies to that reading.

Conversely, context dependence is unproblematic for CCT. Since the selection function appears on both sides of CCT, we are guaranteed that context dependence is resolved in the same way on both sides. Hence CCT applies across the board to all counterfactuals, on all readings. This includes so-called backtracking readings (see Lewis 1979), which we don't need to single out for special treatment.

Parallels and divergences from CE/Skyrms' Thesis. We have seen that, for our toy example (1), CE and CCE give analogous verdicts. The same applies to a wide variety of counterfactuals. Indeed, given a modest assumption about the selection function, we can show that the two Theses are extensionally equivalent for all *would*-conditionals whose antecedents are about the future, like (8).

(8) If Maria flipped a coin tomorrow, it would land heads.

The assumption that we need is:[18]

> **Centering of selection**
> For any selection function s and any propositions A and $H_{w,t}$:
> if A is compatible with $H_{w,t}$, then $s(\mathsf{A}, H_{w,t}) = H_{w,t} + \mathsf{A}$.

Informally, when the antecedent of a counterfactual is compatible with the history proposition, the selected proposition is simply the conjunction of the two. This is in keeping with standard criteria for determining closeness—and in particular, with Lewis's (1979) idea that closest worlds preserve as much overlap with actual history as follows.

Given this assumption, and given the Limited Equivalence from §2.1, for any counterfactual where the antecedent is about the future, we get:

$$ch_{T_w s(\mathsf{A}, H_{w,t})}(\mathsf{C}) = ch_{T_w H_{w,t}+\mathsf{A}}(\mathsf{C}) = ch_{T_w H_{w,t}}(\mathsf{C} \mid \mathsf{A})$$

Hence, when antecedents are about the future, CE and CCT are in full agreement. Conversely, they may diverge in other cases. In particular, they diverge in all the cases that I discussed in §5.

Before moving on, let me notice that, just thanks to Centering, the CCT vindicates Lewis's principle about the evolution of the world's chance function: chance evolves via conditionalization. At any given time t, how future chancy events are

[18]The name of the principle is due to the analogy between this principle and the centering principles in standard conditional logics.

settled is not entailed by any historical proposition $H_{w,t}$ that is true at t. So propositions that are entirely about future chancy events are always compatible with $H_{w,t}$. Hence, when the relevant event takes place, the new historical proposition of the world is simply the conjunction of the old historical proposition and the proposition describing the new events. The latter is just equivalent to the old chance function, conditional on the new events.

6.3 The Counterfactual Chance Thesis, take 2: expected chances

As we saw, the selection function shifts the value of the historical proposition argument. But it does not shift the value of the theory of chance argument. I.e., on the version of the CCT described in §6.1, the historical proposition relevant for evaluating a counterfactual $\mathsf{A} \mathrel{\Box\!\!\to} \mathsf{C}$ is $s(\mathsf{A}, H_{w,t})$; but the theory of chance is still T_w. As a result, one might worry that the current version of the CCT doesn't quite capture the slogan that "chances of counterfactuals are counterfactual chances".[19] More seriously, one might also worry that this yields some wrong results.

Whether the version of the CCT discussed in §6.1 gets the wrong results or not depends on one's views about closeness and about the metaphysics of chance.

First, on a family of views about the nature of closeness, so-called 'no-miracles' views (see among others Nute 1980, Bennett 1984, Loewer 2007, and Dorr 2016), the worlds that count as closest for the purposes of evaluating counterfactuals are worlds that share the same laws as the actual world. For every world, the theory of chance for that world is entailed by the laws of that world. So counterfactuals invariable take us to worlds with the same theory of chance. Hence, on the 'no miracles' view of closeness, the version of the CCT in §6.1 is perfectly adequate.

Things look different if we allow, with Lewis 1979 and others (see among others Jackson 1977, Lange 2000, Kment 2006, Khoo 2020), that counterfactuals may select worlds that contain so-called small miracles, i.e. small violations of actual laws. These accounts offer no guarantee that the selected worlds will have the same theory of chance as the actual world. To be sure, adopting a miracles account does not dictate that the worlds selected by a counterfactual will have a different theory of chance.[20] It merely allows for it.

[19]Thanks to *[names omitted for anonymous review]* for discussion on this point.

[20]In fact, on a Humean account of laws like Lewis's (1983, 1994), it is plausible that in a lot of cases the presence of small miracles in worlds accessed by counterfactuals will make no difference to the chances. On Humean accounts, laws can be thought of as compact summaries of matters of particular fact; chances are a special case of laws, and are affected (though not determined) by frequencies of events of the relevant kind. In a lot of cases, a small miracle will involve just a small alteration of the relevant frequencies, and will not be sufficient to affect chances.

It is at this point that assumptions about the metaphysics of chance and laws become relevant. In particular, some accounts of laws entail that there are worlds with nonzero chance where the overall theory of chance is different from the actual one. These worlds are known in the literature as 'undermining worlds'.[21] (Roughly, the idea is that the chance function 'undermines itself', by assigning positive chance to a world where it is not the right chance function.) Here is a classical example of undermining, due to Lewis (1994, p. 482):

> [T]here is some minute present chance that far more tritium atoms will exist in the future than have existed hitherto, and each one of them will decay in only a few minutes. If this unlikely future came to pass, presumably it would complete a chancemaking pattern on which the half-life of tritium would be very much less than the actual 12.26 years. . . . Could it come to pass, given the present chances? Well, yes and no. It could, in the sense that there's non-zero present chance of it. It couldn't, in the sense that its coming to pass contradicts the truth about present chances. If it came to pass, the truth about present chances would be different. Although there is a certain chance that this future will come about, there is no chance that it will come about while still having the same present chance it actually has.

This is not the place to establish whether there are undermining worlds. The important point for current purposes is that, if there are undermining worlds, the CCT won't work as it has been formulated in §6.1. The reason is that (at least, plausibly) some counterfactuals will select worlds where the theory of chance is different. So an amendment is in order.

In principle, the amendment is simple. Rather than using the actual theory of chance $T_{@}$, we can use a 'shifted' theory of chance T_{w_i}. But this alone won't work. The reason is that our selection functions return, in general, not a single world, but rather a set of worlds. Hence $T_{s(\mathrm{A},w)}$ is not, in general, well-defined.

The fix for this problem is conceptually simple, though it requires some technical footwork. We need to generalize the function T from worlds to theories of chance, and define a more general function from sets of worlds to a theory of chance. There is a natural way to do this: we define an 'aggregate' theory of chance $\mathcal{T}_S$, where S is a set of worlds, as determine by the weighted average of the theories of chance of the worlds in S—as weighted by the actual chance function.

A theory of chance is not a formal theory. So we cannot apply probabilistic weights directly to theories of chance. But we can define $\mathcal{T}_S$ rigorously by specifying a constraint on the chances we get by using $\mathcal{T}_S$ in the grounding argument of the chance function. This constraint is:

[21]For literature on undermining, see Hall 1994, Lewis 1994, Thau 1994; for more discussion, see Briggs 2009.

$$\text{For any A, } S, w\text{:} \quad ch_{\mathcal{T}_S H_{wt}}(\mathrm{A}) = \sum_{w_i \in S} ch_{T_{w_i} H_{w,t}}(\mathrm{A}) \times ch_{T_@ H_{w,t}}(w_i)$$

Again, the basic idea is that the theory of chance of a set of worlds S will assign to any proposition A a probability that is the weighted average of the chances of A in each of the worlds in S, weighted by the actual chance function.

At this point, we can formulate the new version of the CCT:

Counterfactual Chance Thesis (second take).
$ch_{T_w H_{w,t}}(\mathrm{A} \mathrel{\Box\!\!\rightarrow}_s \mathrm{C}) = ch_{\mathcal{T}_{s(\mathrm{A},T_w H_{w,t})} s(\mathrm{A},T_w H_{w,t})}(\mathrm{C})$

(Notice: since we are now assuming that T_w can in principle shift, I am including it in the information used by the selection function.)

The second version of the CCT is strictly more general, and allows us to accommodate all views on the nature of closeness and on the metaphysics of chance. In particular, if it turns out that there are no undermining worlds, the second version of the CCT simply reduces to the first. At the same time, the second version is relatively unwieldy, so I stick to using the first version in the rest of the paper. Nothing that I discuss there bears on the difference between the two; it's an easy exercise for the reader to swap out the second version for the first.

6.4 From chance to credences

I pointed out in §3 that, given the Principal Principle, Skyrms' Thesis is equivalent to the Chancy Equation. We can establish a similar equivalence between the CCT and a principle about credence.

Fact 2. Given the Principal Principle, CCT is equivalent to:

Credences in Counterfactual Chances Thesis (CCCT)
Let $ch_{T_w H_{w,t}}$ be the chance function at w and t. For all A, B:

$$cr_{E_t}(\mathrm{A} \mathrel{\Box\!\!\rightarrow} \mathrm{C}) = \sum_{w_i \in W} cr(w) \times ch_{T_w s(\mathrm{A},H_{w,t})}(\mathrm{C})$$

In words, CCCT says: S's rational credence in A $\Box\!\!\rightarrow$ C should equal S's expectation of the chance of C, in the counterfactual circumstance that A. The CCCT is the counterpart of Skyrms' Thesis in the current framework. The proof of the equivalence stated in **Fact 2** is in the appendix.

7 Addressing the challenges

I pointed out that CCT and CE are extensionally equivalent in a lot of cases. But, crucially, CCT avoids the problems that are raised by CE and Skyrms' Thesis. In this section, I explain how.

7.1 Morgenbesser cases

Recall the Morgenbesser coin case:

> **Coin**. Alice is about to flip a fair coin, and offers Bob a bet on heads. Bob refuses. Alice flips the coin, which lands heads.

(4) If Bob had bet, he would have won.

CCT yields the following equation:

$$ch_{T_w H_w}(\text{Bet} \mathbin{\Box\!\!\rightarrow}_s \text{Win}) = ch_{T_w s(\text{Bet}, H_{w,t})}(\text{Win})$$

I.e., the chance of (4) equals the chance of you winning in the counterfactual scenario where you take the bet. Crucially, one of the propositions that we hold fixed in selecting this counterfactual scenario is the proposition that the coin landed heads. The reason is that we use the same selection function as the one we use for the counterfactual on the left-hand side, and we know that, to evaluate that counterfactual we hold fixed that proposition. So the verdict of the CCT for this case matches intuition.

7.2 Complex counterfactuals

Consider again the complex counterfactual in the game show scenario:

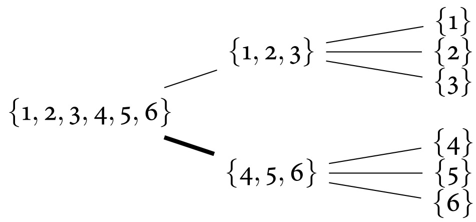
(5) If 5 had been selected, then, if an even number had been selected, it would have been 4.

In this case, the CCT directs us to shift twice the grounding argument of the chance function on the right-hand side.

$$ch_{T_w H_w}(\text{Five} \mathbin{\Box\!\!\rightarrow}_s (\text{Even} \mathbin{\Box\!\!\rightarrow}_s \text{Four})) = ch_{T_w s(\text{Even}, s(\text{Five}, H_{w,t}))}(\text{Four})$$

This says: the chance of $\ulcorner \text{Five} \mathbin{\Box\!\!\rightarrow}_s (\text{Even} \mathbin{\Box\!\!\rightarrow}_s \text{Four}) \urcorner$ is the chance that 4 is selected, in the counterfactual scenario that we reach by supposing that an even number was selected, starting from the counterfactual scenario that we reach by supposing that 5 was selected.

On standard criteria for closeness (for example, following Lewis 1979, maximizing overlap with matters of historical fact in the starting scenario of evaluation), this lands us in a scenario where the set $\{4, 5, 6\}$ has been selected in the first round, and an even number is selected in the second round. This is exactly right, and gets us the right verdict about chance.



7.3 Triviality

Giving a semantics for counterfactuals goes beyond the purposes of this paper. So I cannot give a tenability result for CCT—i.e., I cannot show that CCT is vindicated by a particular theory of counterfactuals. However, I can show that the switch from CE to CCT blocks the key step in Williams' proof.

Recall that, in Williams' proof, the triviality-generating step is the equality in (9):

(9) $$ch_{T_w H_{w,t}+\mathrm{C}}(\mathrm{A} \,\Box\!\!\rightarrow_s \mathrm{C}) = ch_{T_w H_{w,t}+\mathrm{C}}(\mathrm{C} \mid \mathrm{A})$$

As I pointed out, the right-hand side of (9) goes to 1 for mathematical reasons. By replacing CE with CCT, (9) is replaced by:[22]

(10) $$ch_{T_w H_{w,t}+\mathrm{C}}(\mathrm{A} \,\Box\!\!\rightarrow_s \mathrm{C}) = ch_{T_w s(\mathrm{A}, H_{w,t}+\mathrm{C})}(\mathrm{C})$$

Crucially, the right-hand side of (10) need *not* be identical to 1. The reason is that, even if we start from a history that vindicates C, the selection function can take us to a history where C is not true/not settled.

For illustration, consider again the two coins scenario from §5.

[22]This change will be accompanied by a change in the Closure principle in the background of the proof. Williams assumes that chances are closed under conditionalization of propositions with nonzero chance. I assume that chances are closed under update of the grounding argument, as follows:

> **Closure under Update.** If $ch_{T_w H_{w,t}}(\bullet)$ is a chance function, then, for any A that is compatible with $T_w H_{w,t}$, $ch_{T_w H_{w,t}+\mathrm{A}}(\bullet)$ is a chance function.

Without a closure assumption, the triviality proof won't even get started, so here I am granting the best case scenario to the proponent of triviality.

Two coins. Sarah is going to flip one of two coins A and B, picked at random. Coin A is fair. Coin B has a 100% bias towards heads.

The key intuition, recall, concerns the chance of the counterfactual A-Flip $\Box\!\!\rightarrow$ Heads, under the hypothesis that the flipped coin lands heads. Since the proposition that the coin lands heads is compatible with current history, the relevant chance can be expressed as:

(i) $ch_{T_w H_{w,t} + \text{Heads}}(\text{A-Flip} \Box\!\!\rightarrow \text{Heads})$

Given the CCT, (i) equals:

(ii) $ch_{T_w s(\text{A-Flip}, H_{w,t} + \text{Heads})}(\text{Heads})$

Now, crucially, $s(\text{A-Flip}, H_{w,t} + \text{Heads})$ need not be a scenario that makes heads true. In fact, on a plausible choice of selection function, some closest worlds where A is flipped will lead to a tails outcome. Hence the chance of heads will be less than 1, hence the chance of A-Flip $\Box\!\!\rightarrow$ Heads, under the hypothesis that the flipped coin lands heads, is less than 1.[23]

8 Conclusion

I have argued that probabilities of counterfactuals are counterfactual probabilities. This view allows us to avoid a number of counterintuitive consequences of the alternative, and sidestep Williams' triviality results. In addition, as I pointed out in the opening of the paper, it is strikingly intuitive. Sometimes, even in philosophy, the most intuitive view is the one that holds up to scrutiny.

[23]Notice: here it matters that the coinflip temporally precedes the coin landing heads or tails. Otherwise, via the reasoning described in §6.2, the verdict of the CCT would be analogous to the verdict of CE.

Appendix: proofs

Fact 1. Given the Principal Principle, Skyrms' Thesis is equivalent to the Chancy Equation.

Proof. We first derive the Chancy Equation from Skyrms' Thesis. (This part of the derivation builds on Williams 2012.) We start from Skyrms' Thesis:

(i) $$cr_{E_t}(\mathsf{A} \boxright \mathsf{C}) = \sum_{w \in W} cr_E(w) \times ch_{T_w H_{w,t^-}}(\mathsf{C} \mid \mathsf{A})$$

Then we consider a subject who is omniscient about the chances. (For the purposes of the proof, this subject needs to be only logically possible, not actual.) For this subject, (i) simplifies to:

(ii) $$cr_{E_t}(\mathsf{A} \boxright \mathsf{C}) = ch_{T_w H_{w,t^-}}(\mathsf{C} \mid \mathsf{A})$$

Then we notice that, via the Principal Principle, assuming that the subject has no inadmissible evidence, we get:

(iii) $$cr_{E_t}(\mathsf{A} \boxright \mathsf{C}) = ch_{T_w H_{w,t}}(\mathsf{A} \boxright \mathsf{C})$$

From (i)–(iii), via the transitivity of identity, we get the Chancy Equation:

(iv) $$ch_{T_w H_{w,t}}(\mathsf{A} \boxright \mathsf{C}) = ch_{T_w H_{w,t^-}}(\mathsf{C} \mid \mathsf{A})$$

Then, we derive Skyrms' Thesis from the Chancy Equation. Following Lewis 1980, we first notice that the PP entails the following principle (I use 'X_x' as shorthand for the proposition that $ch_{T_w H_w}(\mathsf{A}) = x$, and '$\mathbf{X}$' as the set that contains all such propositions):

(WPP) $$up(\mathsf{A} \mid E) = \sum_{X_x \in \mathbf{X}} up(ch_{T_w H_w}(\mathsf{A}) = x \mid E) \times x$$

Informally, WPP (for 'Weighted Principal Principle') says that one's ur-prior in a proposition A, conditional on E and the information that the chance of A is x, should equal the weighted average of the possible values of the chances of A (weighted by one's ur-prior that that each of those values is actually the value of the chance of A, given E). For clarity, notice also that $ch_{T_w H_w}(\mathsf{A}) = x$ is bound by the subscript of the summation symbol, and that hence WPP could be rewritten as:

(WPP') $$up(\mathsf{A} \mid E) = \sum_{X_x \in \mathbf{X}} up(X_x \mid E) \times x$$

In what follows, I will stick to the original formulation for clarity.

Start by instantiating 'A $\Box\!\!\rightarrow$ C' for A in WPP. We get:

(i) $up(\text{A} \Box\!\!\rightarrow \text{C} \mid E_t) = \sum_{X_x \in \mathbf{X}} up(ch_{T_w H_{w,t}}(\text{A} \Box\!\!\rightarrow \text{C}) = x \mid E_t) \times x$

To reduce clutter, we start first by rewriting the conditional ur-prior function as a credence function, indexed to evidence E_t:

(ii) $cr_{E_t}(\text{A} \Box\!\!\rightarrow \text{C}) = \sum_{X_x \in \mathbf{X}} cr_{E_t}(ch_{T_w H_{w,t}}(\text{A} \Box\!\!\rightarrow \text{C}) = x) \times x$

At this point, we introduce a second summation, summing credences over worlds in each of the X_x propositions. We get:

(iii) $cr_{E_t}(\text{A} \Box\!\!\rightarrow \text{C}) = \sum_{X_x \in \mathbf{X}} \sum_{w \in X_x} cr_{E_t}(w) \times x$

At this point, we replace the right-hand term of the product on the right-hand side, x, with a function that takes each world in every X_x to the value of the chance function for A $\Box\!\!\rightarrow$ C in that world. That function is, quite simply, $ch_{T_{w_x} H_{w_x,t}}(\text{A} \Box\!\!\rightarrow \text{C})$.

(iv) $cr_{E_t}(\text{A} \Box\!\!\rightarrow \text{C}) = \sum_{X_x \in \mathbf{X}} \sum_{w \in X_x} cr_{E_t}(w) \times ch_{T_{w_x} H_{w_x,t}}(\text{A} \Box\!\!\rightarrow \text{C})$

At this point, we apply the Chancy Equation to get:

(v) $cr_{E_t}(\text{A} \Box\!\!\rightarrow \text{C}) = \sum_{X_x \in \mathbf{X}} \sum_{w \in X_x} cr_{E_t}(w) \times ch_{T_{w_x} H_{w_x,t^-}}(\text{C} \mid \text{A})$

We only need one last step. Notice first that $\bigcup \mathbf{X}$ is just the set of all worlds—the hypotheses about the chance of A $\Box\!\!\rightarrow$ C form a partition over logical space. Notice, moreover, that the elements summed over in the innermost summation form a sub-partition of the partition in $\mathbf{X}$. So we can simplify (v) by removing the two summations in favor of a unique summation over worlds. We get:

(vi) $cr_{E_t}(\text{A} \Box\!\!\rightarrow \text{C}) = \sum_{w_x \in W} cr_{E_t}(w) \times ch_{T_w H_{w_x,t^-}}(\text{C} \mid \text{A})$

Which is Skyrms' Thesis.

Fact 2. Given the Principal Principle, the CCT is equivalent to the CCCT.

Proof. Analogous to the proof of Fact 1, aside from the application of the CCT in place of the Chancy Equation.

References

Adams, Ernest W. (1976). *Prior Probabilities and Counterfactual Conditionals*, Dordrecht: Springer Netherlands, pp. 1–21.

Arregui, Ana (2007). "When Aspect Matters: The Case of Would-Conditionals." *Natural Language Semantics*, 15(3): pp. 221–264.

Bennett, Jonathan (1984). "Counterfactuals and Temporal Direction." *Philosophical Review*, 93(1): pp. 57–91.

Bradley, R. (2012). "Multidimensional Possible-World Semantics for Conditionals." *Philosophical Review*, 121(4): pp. 539–571.

Briggs, Rachael (2009). "The Anatomy of the Big Bad Bug." *Noûs*, 43(3): pp. 428–449.

Dorr, Cian (2016). "Against Counterfactual Miracles." *Philosophical Review*, 125(2): pp. 241–286.

Hájek, Alan, and N. Hall (1994). "The Hypothesis of the Conditional Construal of Conditional Probability." In Ellery Eells, Brian Skyrms, and Ernest W. Adams (eds.) *Probability and Conditionals: Belief Revision and Rational Decision*, Cambridge University Press, p. 75.

Hall, Ned (1994). "Correcting the Guide to Objective Chance." *Mind*, 103(412): pp. 505–518.

Hiddleston, Eric (2005). "A Causal Theory of Counterfactuals." *Noûs*, 39(4): pp. 632–657.

Iatridou, Sabine (2000). "The Grammatical Ingredients of Counterfactuality." *Linguistic Inquiry*, 31(2): pp. 231–270.

Jackson, Frank (1977). "A causal theory of counterfactuals." *Australasian Journal of Philosophy*, 55(1): pp. 3–21.

Khoo, Justin (2020). "The Meaning of *If*." Book draft, MIT.

Kment, Boris (2006). "Counterfactuals and Explanation." *Mind*, 115(458): pp. 261–310.

Lange, Marc (2000). *Natural Laws in Scientific Practice*. Oxford University Press.

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

Lewis, David (1980). “A subjectivist’s guide to objective chance.” In *Ifs*, Springer, pp. 267–297.

Lewis, David (1983). “New Work For a Theory of Universals.” In D. H. Mellor, and A. Oliver (eds.) *Properties*, Oxford University Press.

Lewis, David (1994). “Humean Supervenience Debugged.” *Mind*, 103(412): pp. 473–490.

Lewis, David K. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.

Lewis, David K. (1979). “Counterfactual dependence and time’s arrow.” *Noûs*, 13(4): pp. 455–476.

Loewer, Barry (2007). “Counterfactuals and the Second Law.” In H. Price, and R. Corry (eds.) *Causation, Physics, and the Constitution of Reality: Russell’s Republic Revisited*, Oxford University Press.

Meacham, Christopher J. G. (2010). “Two Mistakes Regarding the Principal Principle.” *British Journal for the Philosophy of Science*, 61(2): pp. 407–431.

Moss, Sarah (2013). “Subjunctive Credences and Semantic Humility.” *Philosophy and Phenomenological Research*, 87(2): pp. 251–278.

Nute, Donald (1980). “Conversational scorekeeping and conditionals.” *Journal of Philosophical Logic*, 9(2): pp. 153–166.

Ramsey, Frank (1926). “Truth and Probability.” In A. Eagle (ed.) *Philosophy of Probability: Contemporary Readings*, Routledge, pp. 52–94.

Schultheis, Ginger (2022). “Counterfactual Probability.” Forthcoming in *Journal of Philosophy*.

Schulz, Moritz (2017). *Counterfactuals and Probability*. Oxford: Oxford University Press.

Skyrms, Brian (1980b). “The prior propensity account of subjunctive conditionals.” In W. Harper, R. C. Stalnaker, and G. Pearce (eds.) *Ifs*, Reidel.

Slote, Michael A. (1978). “Time in Counterfactuals.” *Philosophical Review*, 87(1): pp. 3–27.

Stalnaker, Robert (1968). “A Theory of Conditionals.” In N. Recher (ed.) *Studies in Logical Theory*, Oxford.

Thau, Michael (1994). “Undermining and Admissibility.” *Mind*, 103(412): pp. 491–504.

Van Fraassen, Bas C. (1976). “Probabilities of conditionals.” In *Foundations of probability theory, statistical inference, and statistical theories of science*, Springer, pp. 261–308.

Williams, J. Robert G. (2012). “Counterfactual Triviality.” *Philosophy and Phenomenological Research*, 85(3): pp. 648–670.