Probability for Trivalent Conditionals

Paul Égré Lorenzo Rossi Jan Sprenger

September 16, 2024

Abstract

This paper presents a unified theory of the truth conditions and probability of indicative conditionals and their compounds in a trivalent framework. The semantics validates a Reduction Theorem: any compound of conditionals is semantically equivalent to a simple conditional. This allows us to validate Stalnaker's Thesis in full generality and to use Adams's notion of p-validity as a criterion for valid inference. Finally, this gives us an elegant account of Bayesian update with indicative conditionals, establishing that despite differences in meaning, it is tantamount to learning a material conditional.

1 Introduction

Consider a simple conditional sentence such as "if the die landed even, the number is greater than 3". Intuitively, its probability is 2/3 assuming the die is fair. This corresponds to the conditional probability of the consequent, given the antecedent (e.g., Adams 1965):

For conditional-free A and B: $p(A \to B) = p(B|A)$ (Adams's Thesis)

This claim, known as "Adams's Thesis", is widely endorsed, and accepted by theories of conditionals that differ substantively otherwise. It is more controversial whether the above equality should also hold for *arbitrary* sentences A and B (e.g., allowing A and B themselves to contain conditionals, see Stalnaker 1970):

For arbitrary A and B: $p(A \to B) = p(B|A)$ (Stalnaker's Thesis)

Accepting Stalnaker's Thesis in full generality poses notable difficulties: starting with the classical triviality results by David Lewis (1976) and Robert Stalnaker (1976), an entire literature of technical results (e.g., Bradley 2000; Milne 2003; Hájek 2011; Fitelson 2015) shows that Stalnaker's Thesis clashes with plausible principles of conditional logic and probabilistic reasoning. Moreover, recent contributions argue that Stalnaker's Thesis *should* fail for specific types of nested conditionals (Goldstein and Santorio 2021; Khoo 2022).

Can an attractive analysis of conditionals validate Stalnaker's Thesis in full generality? Some semantics give up on truth conditions for conditionals and *define* the probability of a simple conditional as the corresponding conditional probability (e.g., Adams 1965, 1975; Edgington 1995). Such accounts accurately predict probabilistic reasoning with simple conditionals, but their scope is too limited: they fail to account for the semantics and probability of nested and compound conditionals (see McGee 1989, p. 485). Consequently, they do not yield Stalnaker's Thesis.

By contrast, truth-conditional accounts of conditionals that save Stalnaker's Thesis either (i) restrict attention to a fragment of the language, excluding left-nested conditionals (e.g., McGee 1989), or (ii) they let the proposition expressed by a conditional depend on the agent's epistemic state (van Fraassen 1976; Kaufmann 2009; Bacon 2015). While the first research program saves only a restricted version of Stalnaker's Thesis, the second buys into a strong form of contextualism about conditionals.

A third option consists in adopting trivalent truth conditions for conditionals (McDermott 1996; Cantwell 2006; Rothschild 2014; Lassiter 2020; Égré, Rossi, and Sprenger 2021a, forthcoming). On this account, pioneered by the work of de Finetti (1936), the truth value of "if the die landed even, the number is greater than three" is *true* if the die landed 4 or 6 (=even and greater than three), *false* if the die landed 2 (=even, but not greater than three), and *void* or nonassertive if the die landed odd. The probability of any conditional sentence can be defined in analogy with classical bivalent probability, by restricting attention to those worlds where the sentence takes a classical truth value.

In this paper, we develop a general theory of compound conditionals and their probability within the trivalent framework. It differs from extant accounts in three ways: First, while trivalent accounts typically rely on de Finetti's truth conditions, our account uses Cooper's truth conditions for the conditional (Cooper 1968), making distinct and arguably better predictions in several places. Second, it is not limited to special classes of conditionals, but handles arbitrary combinations of conditionals and other sentential connectives. This is because our semantics validates a Reduction Theorem: *any* nested conditional or compound of conditionals is semantically equivalent to a simple conditional. Third, we define trivalent *conditional* probability by the equation $p_A(B) := p(A \to B)$ (for arbitrary A and B) and show that this expression behaves in almost all respects like a classical, bivalent conditional probability.

These results have far-reaching consequences. Most importantly, we obtain Stalnaker's Thesis in full generality, and we extend Bayesian conditionalization to the trivalent setting and to a procedure for rational update on conditionals. Moreover, we can rephrase inferences with conditionals in a language containing at most simple conditionals, and use Adams's notion of p-validity as a criterion for valid inference. This allows us to deal with compound conditionals and their probability in a way not explored by Adams, or more recent accounts. All this goes significantly beyond past explorations of the logic and probability of trivalent conditionals.[1]

Summing up, we develop a fully compositional, truth-functional semantics of conditionals that respects Stalnaker's Thesis and unifies the semantic layer with probabilistic reasoning and Bayesian learning. The price to pay for such a unified theory is the non-classical behavior of conjunction (and disjunction) and the violation of the product rule in probabilistic reasoning, but these features can be motivated independently, and classical laws in inferences with non-conditional sentences are preserved.

The paper proceeds as follows: Section 2 explains our basic idea for trivalent truth conditions and the probability of indicative conditionals. Section 3 presents the Reduction Theorem, namely how nested conditionals and compound conditionals are, on this semantics, equivalent to simple conditionals. Section 4 defines logical consequence relations for probabilistic inference with trivalent conditionals along the lines of Égré, Rossi, and Sprenger (forthcoming) and shows that Adams's p-validity criterion is sufficient for deciding which inferences are valid, and which ones aren't. The section concludes with the discussion of an impossibility result es-

[1] In particular, the trivalent account by Lassiter (2020) only validates specific cases of Stalnaker's Thesis and it is based on different definitions of truth-functional connectives and conditional probability. And while Égré, Rossi, and Sprenger (forthcoming) lay the semantic foundations of the present account, they mainly explore valid and invalid inferences of the resulting logics. They do *not* address the validity of Stalnaker's Thesis or Bayesian learning for trivalent propositions.

tablished by McGee (1981) and Schulz (2009) for trivalent representations of probabilistic inference with conditionals. Section 5 introduces a suitable notion of conditional probability for complex sentences that avoids Lewis-style triviality results and yields Stalnaker's Thesis in full generality. Here and in the next section, we also respond to recent counterexamples against the unrestricted form of Stalnaker's Thesis. Section 6 generalizes Bayesian conditionalization to the trivalent setting. In particular, we obtain that updating on a simple indicative conditional amounts to updating on the corresponding material conditional. Section 7 concludes.

2 Outline of Trivalent Semantics

It is controversial whether indicative conditionals have factual truth conditions and whether they express propositions just like conditional-free and modal-free sentences do (e.g., see the dialogue in Edgington 1991; Jeffrey 1991). However, even a defender of a non-truth-conditional view such as Adams (1965, p. 187) admits that we feel compelled to say that a conditional "if A, then B" has been *verified* if we observe both A and B, and *falsified* if we observe A and $\neg B$". For example, take the sentence "if it rains, the match will be cancelled"; it seems to be true if it rains and the match is in fact cancelled, and false if the match takes place in spite of rain. No similarly strong intuitions apply to the case where the antecedent is false (i.e., in case it does not rain).

This observation motivates the treatment of the indicative conditional "if A, then B" as a *conditional assertion*—i.e., as an assertion about B upon the supposition that A is true. On this account, when the antecedent is false, the speaker is committed to neither truth nor falsity of the consequent (e.g., de Finetti 1936; Quine 1950; Belnap 1973). There is simply no factual basis for evaluating the assertion. Therefore the conditional assertion is classified as neither true nor false. See Table 1.

Truth value of $A \to B$	B true	B false
A true	true	false
A false	neither	neither

Table 1: Partial truth table for a conditional $A \to B$ analyzed as a conditional assertion.

The question is whether "neither" should be understood as a truth-value gap, i.e., whether the valuation function for $A \to B$ should be a *partial* *function*, or whether we should treat "neither" as a third semantic value ("void" or "nonassertive") that can freely interact with the two standard semantic values. In the remainder, we will pursue the second option and develop it systematically.

Suppose that we have a set of possible worlds W, where each possible world corresponds to a complete trivalent valuation function over all sentences of our language $\mathcal{L}^{\rightarrow}$. (Of course, we have not yet specified *which* valuation functions we select for $\mathcal{L}^{\rightarrow}$; we will do this shortly, but the general shape of our approach can be presented without committing ourselves to specific valuation functions). Sentences of $\mathcal{L}^{\rightarrow}$ are interpreted as propositions, i.e., as functions from W to a set of truth values $\{1, 0, 1/2\}$ (=true, false, or neither). In addition, we assume that atomic sentences, and in general all conditional-free sentences, only take classical truth values (in Humberstone (2011)'s terminology, that valuations are atom-classical). The guiding idea, going back to Cooper (1968), is that $\mathcal{L}^{\rightarrow}$ is an *extension* of the Boolean propositional language $\mathcal{L}$ to a language with a conditional connective, and so $\mathcal{L}$-sentences should receive classical truth values.

Suppose further that we have a credence function $c : \mathcal{A} \mapsto [0, 1]$ on the measurable space of possible worlds $(W, \mathcal{A})$, where $\mathcal{A}$ is an algebra defined on subsets of W, representing the subjective plausibility of a particular set of possible worlds. Moreover, we assume that any algebra $\mathcal{A}$ includes the singletons of worlds, i.e., for every $w \in W$, $\{w\} \in \mathcal{A}$. We assume that the credence function c is a finitely additive probability function, i.e., for all singleton worlds $\{w\} \in \mathcal{A}$, we have $c(\{w\}) \geq 0$, whereas $c(\emptyset) = 0$, $c(W) = 1$, and $c(X \cup Y) = c(X) + c(Y)$ whenever $X \cap Y = \emptyset$.

We can then define a (non-classical) probability function $p : \mathcal{L}^{\rightarrow} \mapsto [0, 1]$ on the language $\mathcal{L}^{\rightarrow}$, taking into account that sentences of $\mathcal{L}^{\rightarrow}$ can receive three values: true, false, or neither ("void").[2] For convenience, define

$$A_T = \{w \in W \mid v_w(A) = 1\} \qquad\qquad A_V = W \backslash (A_F \cup A_T)$$
$$A_F = \{w \in W \mid v_w(A) = 0\}$$

where v_w is the valuation function associated with the world w (recall, worlds essentially are valuation functions). In other words, A_T, A_F, and A_V are the sets of possible worlds where A is true, false and neither true nor false, respectively. What should then be the probability of A? Traditionally, the probability of a sentence A is simply the credence assigned to

[2]For a survey of non-classical probability functions, see Williams (2016).

all possible worlds where A is true: $p(A) = c(A_T)$. But when A involves a conditional, it is possible that $c(A_F \cup A_T) < 1$, so using that approach we may end up assigning a small probability to A, even though A is much more likely to be true than false.

Therefore, for trivalent valuations, the relative weight of truth and falsity is a better indicator of the probability of a sentence than the credence that it is true (the same point is made by Bradley 2002, p. 363). Apart from its intuitive appeal, this idea is supported by Bayesian accounts that explicate probability by means of fair *betting odds* (de Finetti 1974; Ramsey 1929/1990). In particular, the subjective probability of A is the inverse of the decimal betting odds on A: $p(A) = 1/O(A)$, e.g., a probability of 1/3 corresponds to 3:1 betting odds. These odds specify the factor by which the bettor's stake is multiplied in case A occurs and she wins the bet. When A is a conditional assertion, bets are naturally generalized as follows: they are settled in the ordinary way if A takes a classical truth value, and they are *called off* otherwise (i.e., the original stakes are returned). Indeed, it is hard to imagine how we should declare a bet on a conditional assertion like "if it rains on Saturday, the match will be cancelled" as won or lost unless it actually rains on Saturday.

The relation between probabilities and fair betting odds helps us to show why only the relative weight of $c(A_T)$ and $c(A_F)$ should affect the probability of A. Suppose that the bettor is betting on A at stake $S > 0$ and odds $O(A) > 0$ and that we are sampling possible worlds at random, according to a credence function c. Since the bet on A will be called off in case A does not take a classical truth value, with the stake returned to the bettor, her long-run net gain will in the limit approach the following quantity:

$$\begin{aligned} G &= -S + c(A_T) \times S \times O(A) + c(A_F) \times 0 + c(A_V) \times S \\ &= S \times (-1 + c(A_T) \times O(A) + c(A_V)) \end{aligned}$$

If bettor and bookie agree on credence function c, the bet is fair if and only if $G = 0$, i.e., neither side is supposed to have a long-run advantage. This can be shown to be equal to

$$c(A_T) \times O(A) = 1 - c(A_V)$$

or, equivalently,

$$O(A) = \frac{c(A_T) + c(A_F)}{c(A_T)},$$

which agrees with the definition of decimal betting odds in the standard bivalent case. Moving back from conditional bets to probabilities for sentences with trivalent valuations, and defining probability as the inverse of fair betting odds (as in the bivalent case), we obtain:

$$p(A) := \frac{c(A_T)}{c(A_T) + c(A_F)} \qquad \text{if } c(A_T) + c(A_F) > 0 \qquad \text{(Probability)}$$

(If $c(A_T) + c(A_F) = 0$, then we set $p(A) = 1$ by convention.) For a Boolean, conditional-free sentence A, this will collapse to the standard definition of probability as the credence in the set of possible worlds where A is true (since $A_T \cup A_F = W$ and hence $c(A_T) + c(A_F) = 1$). Whereas for a simple conditional "if A, then B", with conditional-free sentences A and B, we obtain, given Table 1,

$$p(A \to B) = \frac{c(A_T \cap B_T)}{c(A_T)} = \frac{p(A \wedge B)}{p(A)} = p(B|A) \qquad \text{(Adams's Thesis)}$$

where the probability function in the last two expressions refers to standard bivalent probability on $\mathcal{L}$. In other words, Adams's Thesis—the equality of probability of conditionals and conditional probability for simple conditionals—follows as a *corollary* of the trivalent semantics and need not be stipulated as a definition, as in Adams's own account.

These considerations do not yet show how we can assign a probability to Boolean compounds of conditionals, nested conditionals, and other complex sentences of $\mathcal{L}^{\to}$. We need to extend Table 1 and the truth tables for the regular Boolean connectives to cases where both the antecedent and the consequent can take the third semantic value, which we call "void". For the conditional, there are two main options: the de Finetti conditional $\to_{DF}$ where void antecedents are grouped with false ones (e.g., de Finetti 1936; Belnap 1970; Dubois and Prade 1994; McDermott 1996; Rothschild 2014) and the Cooper conditional $\to_C$ where they are grouped with true ones (e.g., Cooper 1968; Belnap 1973; Cantwell 2008; Olkhovikov 2002/2016).

See Table 2.[3] Semantic values are represented by numbers, with 1, 1/2 and 0 standing for true, void and false.

$\to_C$	1	1/2	0
1	1	1/2	0
1/2	1	1/2	0
0	1/2	1/2	1/2

$\to_{DF}$	1	1/2	0
1	1	1/2	0
1/2	1/2	1/2	1/2
0	1/2	1/2	1/2

	$\neg$
1	0
1/2	1/2
0	1

$\wedge$	1	1/2	0
1	1	1	0
1/2	1	1/2	0
0	0	0	0

$\vee$	1	1/2	0
1	1	1	1
1/2	1	1/2	0
0	1	0	0

Table 2: Truth tables for the trivalent conditional according to Cooper ($\to_C$) and to de Finetti ($\to_{DF}$), and the Boolean connectives in our trivalent logic.

Reasons for preferring the Cooper conditional to the de Finetti conditional, i.e., treating a void antecedent like a true one, are discussed in detail by Égré, Rossi, and Sprenger (2021a, forthcoming), as are those for deviating from the Strong Kleene connectives for conjunction and disjunction. We just mention one important motivation: as argued by Bradley (2002), we want to be able to consider "partitioning" sentences like

(1) If the sun shines tomorrow, John goes to the beach; and if it doesn't, he goes to the museum.

as true when one of two conditional predictions is verified (e.g., the sun shines and John goes to the beach). But if we used Strong Kleene connectives for conjunction, (1) would always be void or false. This means that as long as John's actions are uncertain, (1) will have probability zero. This would be an unwelcome prediction since (1) is certainly more assertable than "If it rains tomorrow, John goes to the beach; and if not, he goes to the museum.". Therefore, we consider the conjunction of a true and a void sentence to be true. In other words, the third truth value acts as a chameleon: the conjunct or disjunct with classical truth value determines the truth value of the compound.[4] In spite of their non-standard nature,

[3] Negation is Strong Kleene negation and conjunction and disjunction happen to coincide with those of Sobociński (1952). The combination of these Boolean connectives and the connective $\to_C$ was first advocated by Cooper (1968).

[4] Another unwelcome consequence of the Strong Kleene tables is that regardless of whether we follow Cooper or de Finetti in the interpretation of the conditional, $v((A \to B) \vee (B \to A)) \geq 1/2$ for all valuations v. In other words, such sentences—encoding the so-called *linearity* principle—cannot be false.

the truth tables for conjunction and disjunction interact with each other, and with Strong Kleene negation, in the usual way, respecting the de Morgan rules and the distributivity laws. We use them to define the family of valuation functions we will employ:

Definition 1 (Cooper valuations). *A valuation function $v : \mathcal{L}^{\rightarrow} \mapsto \{1, 1/2, 0\}$ that assigns semantic values to all formula of $\mathcal{L}^{\rightarrow}$ is called a* Cooper valuation *if and only if (i) it assigns classical values to all atomic formulas and (ii) respects all constraints in the truth tables of Table 2 with the $\rightarrow_C$ interpretation of the conditional connective.*

In other words, Cooper valuations combine the "ideal" configuration of truth tables with the aforementioned idea that the conditional connective extends a language with classical semantic values. Specifically, assuming Cooper valuations, our trivalent probability function obtains more structure, and satisfies the following three principles:

(1) $p(\top) = 1$ and $p(\bot) = 0$.

(2) $p(A) = 1 - p(\neg A)$ (unless $c(A_F \cup A_T) = 0$, in which case $p(A) = p(\neg A) = 1$).

(3) If $A_T \cap B_T = \emptyset$ (i.e., A and B cannot both be true) and $A_T \cup A_F = B_T \cup B_F$ (i.e., they take classical truth values in the same conditions), then $p(A \vee B) = p(A) + p(B)$.

These axioms are similar to the standard axioms of probability, where the additional requirement $A_T \cup A_F = B_T \cup B_F$ for the Additivity axiom reflects that the probability of a disjunction reduces to the probability of its disjunctions only if the disjunctions take classical truth values in the same worlds.

Notably, our conjunction is non-classical and allows for violations of the standard probability law $p(X \wedge Y) \leq p(X)$. This behavior can be justified as follows: suppose that A, B and C are conditional-free sentences, with A and B false and C true. Then the bet on $(A \rightarrow B) \wedge C$ yields a positive return (because the expressed proposition is true for Cooper valuations), while the bet on $A \rightarrow B$ is called off. Betting on X is therefore not always safer than betting on X and Y. Since these properties of betting odds transfer to probabilities, some probability functions will have the feature $p((A \rightarrow B) \wedge C) > p(A \rightarrow B)$, in notable difference to standard bivalent probability. Of course, as long as X and Y are conditional-free, the standard law $p(X \wedge$

$Y) \leq p(X)$ will continue to hold. Section 4 will get back to this point in more detail.

3 Simplifying Complex Conditionals

This trivalent account deals with nested and compound conditionals in a particularly pleasant way. It is a well-known drawback of Adams's account that complex conditionals (and compounds of conditionals) such as $A \rightarrow (B \rightarrow C)$ or $(A \rightarrow B) \wedge (C \rightarrow D)$ with $A, B, C, D \in \mathcal{L}$, cannot be directly analyzed; instead they have to be *rephrased* by means of a formula in $\mathcal{L}_1^{\rightarrow}$ that approximates their meaning. For instance, letting $A \supset B$ stand for $\neg A \vee B$, Adams *defines* the meaning of a conjunction of conditionals $(A \rightarrow B) \wedge (C \rightarrow D)$ as $(A \vee C) \rightarrow ((A \supset B) \wedge (C \supset D))$ (Adams 1975, 1998, pp. 164-165). This move, however, is not independently justified.[5]

The advantage of a fully compositional trivalent semantics is that Adams's "translations" from $\mathcal{L}^{\rightarrow}$ to its flat fragment $\mathcal{L}_1^{\rightarrow}$ cease to be *definitions*. Instead, they are *statements* whose truth or falsity can be decided in our semantics, in line with the following definition.

Definition 2. *For two formulas $\varphi, \psi \in \mathcal{L}^{\rightarrow}$, we write $\varphi \approx \psi$ if and only if for all Cooper valuations $v : \mathcal{L}^{\rightarrow} \mapsto \{1, 1/2, 0\}$, $v(\phi) = v(\psi)$.*

We can now state some equivalence results for nested conditionals and Boolean compounds of conditionals.

Proposition 1. *For all sentences $A, B, C, D \in \mathcal{L}$:*

$$\neg(A \rightarrow B) \approx A \rightarrow \neg B \qquad \text{(Negation Commutation)}$$
$$(A \rightarrow B) \wedge (C \rightarrow D) \approx (A \vee C) \rightarrow ((A \supset B) \wedge (C \supset D)) \qquad \text{(Conjunction)}$$
$$(A \rightarrow B) \vee (C \rightarrow D) \approx (A \vee C) \rightarrow ((A \wedge B) \vee (C \wedge D)) \qquad \text{(Disjunction)}$$
$$A \rightarrow (B \rightarrow C) \approx (A \wedge B) \rightarrow C \qquad \text{(Import-Export)}$$
$$(A \rightarrow B) \rightarrow C \approx (A \supset B) \rightarrow C \qquad \text{(Left-Nesting)}$$

For Negation Commutation, and Import-Export, the identities hold even for $A, B, C \in \mathcal{L}^{\rightarrow}$.

[5]Actually, Adams assigns "ersatz truth values" to the conjunction of conditionals $P_1 \wedge \ldots \wedge P_n$, classifying it as false if any of the P_i is false, as true if none of the P_i is false and at least one of them is true, and as being neither true nor false otherwise. This description agrees exactly with Cooper's conjunction as described in Table 2.

Proof: By inspection of the (trivalent) truth tables. □

The equivalences expressed by Conjunction and Disjunction of conditionals are intuitive and correspond to Adams's translation procedures (e.g., his "quasi-conjunction" of conditionals). Negation Commutation seems intuitive in natural language: when we negate sentences such as

(2) If Oswald didn't kill Kennedy, Jack Ruby did.

we seem to assert that under the supposition that Oswald didn't kill Kennedy, someone other than Jack Ruby did (Cantwell 2008, p. 246). Santorio (2022, pp. 60-61) notes that this phenomenon persists under the scope of epistemic operators such as "doubt" and "believe". Also from a logical point of view, the principle has been frequently defended (Cooper 1968; Adams 1975; Ramsey 1929/1990).

Of course, Negation Commutation in trivalent semantics implies that some sentences of the form $A \wedge \neg A$ cannot be false. (Take $A := (\bot \to B)$, for any B, then both A and $\neg A$ take semantic value 1/2.) However, this would only be a problem if logics for trivalent conditionals satisfied the principle of explosion, according to which any B follows from a sentence of the form $A \wedge \neg A$. As we will see in the next section, they don't.[6]

The Import-Export principle means that right-nested conditionals can be simplified: the right-nested conditional $A \to (B \to C)$ expresses the simple conditional $(A \wedge B) \to C$. Import-Export enjoys strong theoretical support (McGee 1985, 1989; Ciardelli 2020; Ciardelli and Ommundsen 2024) and it is also supported by empirical work in cognitive psychology (van Wijnbergen-Huitink, Elqayam, and Over 2015). Even accounts that reject Import-Export, like Khoo and Mandelkern (2019) and Mandelkern (2020), concede that these two expressions have the same meaning in many contexts, and specifically when A, B and C themselves do not contain conditionals.

Notably, on our account Left-Nesting yields a distinct simplification rule, for it means that left-nested conditionals like $(A \to B) \to C$ express an assertion of the consequent C conditional on the *material* conditional $(A \supset B)$. While the indicative conditional $A \to B$ and the material conditional $A \supset B$ do not have the same *meaning*, in the antecedent of a conditional they express the same *supposition*—either because they have the same value,

[6]More details on negation commutation, possible trade-offs and the relationship to Conditional Excluded Middle can be found in Égré, Rossi, and Sprenger (2021a,b).

or because $A \supset B$ gets the value 1 when $A \to B$ gets the value 1/2. This observation will be important in Section 6, where we develop a Bayesian theory of learning conditionals.[7]

Finally, note the important restriction to Boolean sentences for Left-Nesting, Conjunction, and Disjunction: these equivalences do not hold for arbitrary sentences, unlike Negation Commutation and Import-Export. For example, for Left-Nesting, counterexamples to the scheme $(F \to G) \to H \approx (F \supset G) \to H$ require $v(G) = 1/2$ (and $v(F) = 1$ and $v(H) \neq 1/2$), i.e., G is a conditional with false antecedent. Still, the equivalences are practically useful since natural language conditionals will rarely be more complex than the ones displayed in Proposition 1.

The above observations allow us to derive a general result:

Theorem 1 (Reduction Theorem)**.** *For every $X \in \mathcal{L}^{\to}$ there is an $X_1 \in \mathcal{L}_1^{\to}$, i.e., the fragment of $\mathcal{L}^{\to}$ containing at most simple conditionals, such that $X \approx X_1$.*

While the formal proof by induction is in the appendix, we can summarize the argument informally: Suppose $A \to B$ and $C \to D$ are simple, non-nested conditionals, i.e., $A, B, C, D \in \mathcal{L}$. Then each compound of these sentences is semantically equivalent to another simple conditional. For conjunction, disjunction, and negation, this is immediate from Proposition 1, while the $\mathcal{L}_1^{\to}$-sentence corresponding to $(A \to B) \to (C \to D)$ is $((A \supset B) \wedge C) \to D$. Hence, compounds of $\mathcal{L}_1^{\to}$-sentences are equivalent to $\mathcal{L}_1^{\to}$-sentences, regardless of the chosen connective. Now we gradually simplify any complex $\mathcal{L}^{\to}$-sentence X, starting with the innermost elements, until X can be rewritten as containing at most a simple conditional.

This means that Adams was right on an important point: restricting the language of interest and the logic of conditionals to $\mathcal{L}_1^{\to}$, i.e., the flat fragment of $\mathcal{L}^{\to}$, does not lose anything logically nor semantically essential because complex conditionals are *equivalent* to simple conditionals. We can consider them with Adams, if we want, as convenient linguistic abbreviations. The next section explores the consequences of the Reduction Theorem for a theory of valid inference with conditionals.

[7]Actually, several combinations of truth tables—de Finetti or Cooper conditional, Strong Kleene connectives or quasi-connectives for conjunction and disjunction—validate Import-Export and Negation Commutation (see Égré, Rossi, and Sprenger 2021a). However, only Cooper valuations validate the entire set of equivalences. Specifically, Strong Kleene connectives fail Conjunction and Disjunction of conditionals (e.g., if $v(A) = 0$, $v(C) = c(D) = 1$), while de Finetti's conditional fails Left-Nesting: it satisfies $(A \to B) \to C \approx (A \wedge B) \to C \approx A \to (B \to C)$, i.e., it does not distinguish between left-nested and right-nested conditionals. This difference is relevant for the discussion of Bayesian learning in Section 6.

4 Inferences with Conditionals

So far, we have dealt with the semantics and probability of trivalent conditionals; we have not said much about inference, except indirectly for the equivalences stated in Proposition 1. Following Adams, we consider the consequence relation $\models_{\mathsf{U}}$ introduced by Égré, Rossi, and Sprenger (forthcoming) for valid inference in $\mathcal{L}^{\to}$ in terms of non-increasing uncertainty: the conclusion must not be less probable than the conjunction of the premises.

Definition 3 (Valid Inference in U)**.** *For a set Γ of formulas of $\mathcal{L}^{\to}$ and a formula $B \in \mathcal{L}^{\to}$: $\Gamma \models_{\mathsf{U}} B$ if and only if there is a finite subset of the premises $\Delta \subseteq \Gamma$ such that for all probability functions $p : \mathcal{L}^{\to} \mapsto [0,1]$:*

$$p\Big(\bigwedge_{A \in \Delta} A\Big) \leq p(B)$$

As a limiting case of U, there is the logic C of *reasoning with certain premises* where probability 1 is preserved:

Definition 4 (Valid Inference in C)**.** *For a set Γ of formulas of $\mathcal{L}^{\to}$ and a formula $B \in \mathcal{L}^{\to}$: $\Gamma \models_{\mathsf{C}} B$ if and only if for all probability functions $p : \mathcal{L}^{\to} \mapsto [0,1]$:*

if for all $A \in \Gamma, p(A) = 1,$ *then* $p(B) = 1.$

Both notions of probabilistic validity have been characterized semantically in a trivalent setting (see Égré, Rossi, and Sprenger forthcoming for proofs and further analysis of the properties of C and U):

Proposition 2 (Trivalent Characterization of C)**.** *For a set Γ of formulas of $\mathcal{L}^{\to}$ and a formula $B \in \mathcal{L}^{\to}$: $\Gamma \models_{\mathsf{C}} B$ if and only if for all Cooper valuations $v : \mathcal{L}^{\to} \mapsto \{0, 1/2, 1\}$:*

if for all $A \in \Gamma, v(A) \geq 1/2,$ *then* $v(B) \geq 1/2.$

In other words, preservation of probability 1 is equivalent to preservation of designated values $D = \{1, 1/2\}$ in passing from the premises to the conclusion. For uncertain reasoning in U, we can derive a similar result:

Definition 5 (Consistent and Inconsistent Sets)**.** *A set of formulas Γ is consistent if it has no subset $\Delta \subseteq \Gamma$ such that $p\left(\bigwedge_{A \in \Delta} A\right) = 0$ in all probability functions $p : \mathcal{L}^{\to} \mapsto [0,1]$. It is inconsistent if such a subset exists.*

Proposition 3 (Trivalent Characterization of U)**.** *For a consistent set of formulas Γ of $\mathcal{L}^{\to}$ and a formula $B \in \mathcal{L}^{\to}$ with $\not\models_{\mathsf{C}} B$, $\Gamma \models_{\mathsf{U}} B$ if and only if there is a finite subset $\Delta \subseteq \Gamma$ such that for all Cooper valuations $v : \mathcal{L}^{\to} \mapsto \{0, 1/2, 1\}$:*

$$v(\bigwedge_{A_i \in \Delta} A_i) \leq v(B).$$

This means that an inference in U is valid if and only if both sets of designated values $D = \{1, 1/2\}$ and $D' = \{1\}$ are preserved in passing from the premise from the conclusion. Equivalently, the semantic value of the conclusion must never drop below the semantic value of the conjunction of (a subset of) the premises.[8] These results connect the (trivalent) truth conditions of conditionals to probabilistic reasoning and answer one of the questions raised above at the opening of this paper: valid *uncertain* reasoning preserves both truth and non-falsity, and valid *certain* reasoning preserves non-falsity.

Let us get back for a moment to Adams. His logic of uncertain reasoning with (simple) conditionals (known as system P) is based on the following consequence relation:

Definition 6 (p-valid inference, Adams 1975)**.** *Suppose $\Gamma \subseteq \mathcal{L}_1^{\to}$, $B \in \mathcal{L}_1^{\to}$. Then $\Gamma \models_p B$ if and only if for all probability functions $p : \mathcal{L}_1^{\to} \mapsto [0,1]$, the uncertainty of the conclusion does not exceed the cumulative uncertainty of the premises:*

$$U(B) \leq \sum_{A \in \Gamma} U(A) \qquad (p\text{-valid inference})$$

where $U(X) := 1 - p(X)$ for any $X \in \mathcal{L}_1^{\to}$.

This definition of valid uncertain reasoning has been highly influential among logicians, computer scientists, and psychologists of reasoning (e.g., Pearl 1988; Goodman, Nguyen, and Walker 1991; Evans and Over 2004; Kleiter 2018); in fact, p-validity often serves as a benchmark for other systems of uncertain reasoning. However, as mentioned in the introduction, it is limited to the fragment of $\mathcal{L}^{\to}$ involving at most simple conditionals. It is therefore a significant result that U extends p-valid inference to all sentences of $\mathcal{L}^{\to}$:

[8] The restriction to consistent premise sets is required to deal with some degenerate cases, but imposes no substantial limitation.

Proposition 4 (U is a conservative extension of P)**.** *Suppose* $\Gamma \subseteq \mathcal{L}_1^{\rightarrow}$*,* $B \in \mathcal{L}_1^{\rightarrow}$*. Then* $\Gamma \models_p B$ *if and only if* $\Gamma \models_{\mathsf{U}} B$*.*

Proof. According to Adams (1986, p. 264), the premises $\Gamma = \{A_1, \ldots A_n\}$ *yield* a conclusion B (where $A_1, \ldots A_n$ and B are sentences of $\mathcal{L}_1^{\rightarrow}$) if and only if (1) any atom-classical valuation that falsifies none of the A_i, and verifies at least one of them, verifies B, too; (2) any atom-classical valuation that falsifies B also falsifies at least one of the A_i. He then shows ("Meta-Metatheorem 3") that $\Gamma \models_p B$ if and only if a subset of Γ yields the conclusion B. It is not difficult to see that Adams's conditions on yielding are equivalent to conditions for valid inference in our system U: for a subset $\Delta \subseteq \Gamma$, if $v(\bigwedge_{A \in \Delta} A) = 1$ then $v(B) = 1$, and if $v(B) = 0$, then $v(\bigwedge_{A \in \Delta} A) = 0$, where the logical vocabulary satisfies the Cooper truth tables (Table 2), as the Cooper connectives behave classically on classical inputs. □

The relation between U und p-valid inference is, however, more complex than extension. Taking into account the Reduction Theorem, we can establish an equivalence between the two logics, in the following sense.

Theorem 2 (U is representable in P)**.** *Suppose* $\Gamma \subseteq \mathcal{L}^{\rightarrow}$ *and* $B \in \mathcal{L}^{\rightarrow}$*. Then there is a function* $f : \mathcal{L}^{\rightarrow} \mapsto \mathcal{L}_1^{\rightarrow}$ *such that*

$$\Gamma \models_{\mathsf{U}} B \qquad \textit{if and only if} \qquad f(\Gamma) \models_p f(B).$$

Proof. Immediate from Theorem 1—the function f is the translation of $\mathcal{L}^{\rightarrow}$-formulas into $\mathcal{L}_1^{\rightarrow}$-formulas in the proof of the theorem—and from Adams's yielding criterion cited in the proof of Proposition 4. □

In other words, all valid or invalid conditional inferences can be expressed by means of valid or invalid inferences with *simple* conditionals. The Representation Theorem thus answers McGee's challenge to Adams from the introduction. It shows why Adams's restriction of the logic of probability conditionals to $\mathcal{L}_1^{\rightarrow}$ is not harmful: p-valid inference can be characterized as inference in trivalent semantics (and vice versa), covering also nested conditionals and Boolean compounds of conditionals.

The Representation Theorem raises interesting puzzles and questions. For instance, Modus Ponens is invalid in U, but p-valid. How can this happen if valid inferences in both logics can be mapped to each other? The reason is that the translation procedure from $\mathcal{L}^{\rightarrow}$ to $\mathcal{L}_1^{\rightarrow}$ changes the logical form of the premises. For conditional-free sentences $A, B, C \in \mathcal{L}$, Modus Ponens fails in U for inferences of the form "$A \to (B \to C)$ and A, therefore $B \to C$" (compare Égré, Rossi, and Sprenger forthcoming, Section 8). However, when we project these formulas to their semantic equivalents in $\mathcal{L}_1^{\to}$, we obtain a different inference scheme, namely "$(A \wedge B) \to C$ and A, therefore $B \to C$". This inference is p-invalid, as predicted by Theorem 2. For $\mathcal{L}_1^{\to}$-premises, however, we retain that $A \to B, A \models_p B$, and similarly in U.

Another interesting question prompted by the Representation Theorem concerns impossibility results for representations of p-validity in many-valued logic. McGee (1981) showed that no consequence relation in many-valued logic defined as preservation of a set of designated values agrees with p-validity on Adams's restricted language. Adams (1995) proved that such a characterization is possible if validity is defined more liberally, basically as we established above for U (namely not in terms of the preservation of a fixed set of designated values, but by quantifying over such sets). However, Schulz (2009, Corollary 3.3) shows that no such characterization is possible for a language admitting compounds of conditionals when the conjunction is supposed to be "classical" with respect to the logical consequence relation, i.e., if the following two conditions are satisfied:

(i) If $\Gamma \models \phi$ and $\Gamma \models \psi$, then $\Gamma \models \phi \wedge \psi$.

(ii) If $\Gamma \models \phi \wedge \psi$, then $\Gamma \models \phi$ and $\Gamma \models \psi$.

Schulz concludes:

> This poses a dilemma for proponents of a three-valued account of compounds of conditionals which is supposed to conform to the thesis that conditionals are evaluated by conditional probabilities. Either they will not succeed in defining a classical conjunction, or their conception of validity will disagree with p-validity on the restricted language. (Schulz 2009, p. 516)

For our consequence relation $\models_U$, condition (i) holds, but condition (ii) doesn't.[9] This means that conjunction in U is indeed not classical, and so we choose the first horn of Schulz's dilemma. However, while a non-classical conjunction may look awkward at first, there are good independent reasons to adopt it.

First, as argued in Section 2 (especially p. 8), the chameleon-like behavior of the third truth value (i.e., $1 \wedge 1/2 = 1$, etc.) is essential for showing

[9] Counterexample: $v(\Gamma) = v(\phi) = 1$, but $v(\psi) = 1/2$. In this case $\Gamma \models \phi \wedge \psi$, but $\Gamma \not\models \phi$.

that specific conjunctions of conditionals, which we often employ in natural language, can be true. Second, there is strong evidence that when conditionals are involved, the probability of conjunctions need not be smaller than the probability of a conjunct. Consider the following example from Santorio and Wellwood (2023):[10]

(3) If the outcome of the die roll was even, it was two; and if it was odd, it was one.

While "the die landed 2 if it was even" and "the die landed 1 if it was odd" both have probability 1/3, their conjunction seems to be true exactly when the die landed one or two, and false in all other cases. So (3) should have probability 1/3, too. This prediction is borne out in our framework, but it is at odds with what we would expect from a classical conjunction. Santorio and Wellwood (2023) back up their argument with experimental data, and so it seems that the non-classical behavior of conjunction in uncertain reasoning has both normative and empirical support.[11]

Hence, we propose to interpret Schulz's result as follows: a non-classical conjunction is the only way of (i) providing a many-valued logic that agrees with Adams's widely accepted p-validity in their common domain, while (ii) yielding an adequate account of the truth values of compounds of conditionals. Some properties of classical conjunction simply do not generalize to more complex languages.

5 Conditional Probability and Stalnaker's Thesis

A workable definition of conditional probability is essential for modeling how we learn propositions and reason under uncertainty. Bayesian conditionalization, for example, expresses the idea that the rational degree of belief in a proposition B after learning proposition A is the conditional probability of B, given A. Unfortunately, the standard analysis of condi-

[10] A similar case is made by Ciardelli and Ommundsen (2024).

[11] We view such examples as providing an answer to Schulz's challenge as set in Schulz (2009, p. 513): "If one designs a semantic theory for the unrestricted conditional language, can there be any doubt that the conjunction should obey the standard introduction and elimination rules? The burden of proof would be on the side of those who think that it should not. Counterexamples would have to be produced".

tional probability

$$p(B|A) := \frac{p(A \wedge B)}{p(A)}, \qquad \text{if } p(A) > 0, \qquad \text{(Ratio Analysis)}$$

fails in the trivalent case. Since the non-classical behavior of conjunction allows for cases where $p(A \wedge B) > p(A)$, we could end up with $p(B|A) > 1$, which is unacceptable for any conditional probability function.

In this section, we develop a surrogate notion of conditional probability, which coincides with standard conditional probability for conditional-free sentences, but extends to sentences which can take all three truth values. Our definition is simple: for all $A, B \in \mathcal{L}^{\rightarrow}$, the conditional probability of B given A is the probability of the conditional $A \rightarrow B$.

Definition 7. *For all $A, B \in \mathcal{L}^{\rightarrow}$, the* trivalent conditional probability *of B given A, in symbols $p_A(B)$, is defined as follows:*

$$p_A(B) := p(A \rightarrow B) = \frac{c(A_{TV} \cap B_T)}{c(A_{TV} \cap (B_T \cup B_F))} \qquad \textit{if } c(A_{TV}) > 0$$
(Trivalent Conditional Probability)

where $A_{TV} := A_T \cup A_V$ (i.e., the set of worlds where A is not false).

It is clear by the truth-table for Cooper's conditional that this definition respects the operational definition of probability: it is the ratio of the weight of worlds where the conditional is true to the weight of worlds where it is defined, i.e. it has a classical value. It is also clear that this definition agrees with standard axioms of conditional probability if A and B are conditional-free: $A_{TV} = A_T$ because $A_V = \emptyset$, and the numerator will thus be equal to $p(A \wedge B)$ and the denominator to $p(A)$ (since $B_T \cup B_F = W$). In this special case, $p_A(B)$ also satisfies the standard axioms for conditional probability.[12] We need to check, however, whether the behavior of $p_A(B)$ agrees with

[12]Using $p(A \rightarrow B)$ as a surrogate notion for $p(B|A)$ has been suggested first by McGee (1989) as a means of introducing conditional probability into a language with a conditional. McGee (1989, p. 504) provides an axiomatic characterization of the function $p(A \rightarrow B)$, and of its interaction with the probability of conditional-free sentences. The main pillar in his edifice is the

(Simple) Independence Principle (McGee 1989, p. 499). For conditional-free sentences A, B, $C \in \mathcal{L}$, and assuming that A and C are logically incompatible and $p(A) > 0$, then

$$p(C \wedge (A \rightarrow B)) = p(C) \cdot p(A \rightarrow B). \qquad \text{(C1)}$$

what we expect from a conditional probability function if A and B can contain conditionals. As before, we assume that propositional atoms receive only classical values, and therefore that conditionals are the only source of the third truth value.

Here below we list Popper's axioms for a conditional probability functions $p(\cdot|\cdot)$ for classical propositional logic (taken from Hawthorne 2016):

1. $0 \leq p(B|A) \leq 1$.
2. If $\models_{\mathsf{CL}} \neg B$ and $\models_{\mathsf{CL}} A$, then $p(B|A) = 0$.
3. If $A \models_{\mathsf{CL}} B$, then $p(B|A) = 1$.
4. Left Logical Equivalence: If $A \models_{\mathsf{CL}} B$ and $B \models_{\mathsf{CL}} A$, then $p(C|A) = p(C|B)$.
5. Additivity: If $C \models_{\mathsf{CL}} \neg(A \wedge B)$, then either $p(A \vee B|C) = p(A|C) + p(B|C)$ or $p(D|C) = 1$ for any D.
6. The Product Rule: $p(A \wedge B|C) = p(A|B \wedge C) \times p(B|C)$.

Before evaluating these axioms with respect to $p_A(B)$, we need to reformulate them, replacing $p(B|A)$ with $p_A(B)$ and classical entailment with its generalization to the language $\mathcal{L}^{\rightarrow}$, i.e., entailment in C (=preservation of certainty).[13] This means that the first four axioms read:

1'. $0 \leq p_A(B) \leq 1$.

2'. If $\models_{\mathsf{C}} \neg B$ and $\models_{\mathsf{C}} A$, then $p_A(B) = 0$.

3'. If $A \models_{\mathsf{C}} B$, then $p_A(B) = 1$.

On our semantics, (C1) is *invalid*: it is immediate from $C \models_{\mathsf{CL}} \neg A$ that the truth value of $C \wedge (A \rightarrow B)$ is identical to the truth value of C. (If C is true, then $A \rightarrow B$ is void and hence the conjunction is true.) Hence $p(C \wedge (A \rightarrow B)) = p(C)$, which is almost always larger than $p(C) \cdot p(A \rightarrow B)$. However, our trivalent probability function satisfies the adequacy conditions C2–C8 that McGee imposes, together with the Independence Principle, as necessary and sufficient conditions for a (conditional) probability distribution on a language with a conditional. While our account is, on the probabilistic level, quite close to McGee's, there are notable differences on the semantic level: McGee uses a Stalnaker-type, bivalent semantics based on possible worlds and selection functions, and he does not consider left-nested conditionals.

[13]Note that C generalizes certain reasoning from a classical Boolean setting to a language with a conditional connective. Therefore C has the same role for trivalent probability as classical logic has for classical, bivalent probability.

4′. Left Logical Equivalence: If $A \models_C B$ and $B \models_C A$, then $p_A(C) = p_B(C)$.

The reader is invited to verify that trivalent conditional probability satisfies (1′)–(4′)—the proofs are simple.[14]

The Additivity axiom (5) has to be modified more substantially, similar to the case of unconditional probability:

5′. Trivalent Additivity: Suppose (i) $c(C_{TV}) > 0$ (ii) $C \models_C \neg(A \wedge B)$ and (iii) if $v(C) \geq 1/2$, then $|v(A) - v(B)| \neq 1/2$. Then:

$$p_C(A \vee B) = p_C(A) + p_C(B).$$

To see that Trivalent Additivity is satisfied by our definition of conditional probability, consider the division of the credence attached to sets of possible worlds in C_{TV} according to the following table:

$W \cap C_{TV}$	B_T	B_V	B_F
A_T	✗	✗	α
A_V	✗	δ	✗
A_F	β	✗	γ

Condition 5′.(ii) ensures that the upper left corner is empty and condition 5′.(iii) ensures that all direct neighbors to the central square are empty. Hence $\alpha + \beta + \gamma + \delta = 1$. We can now calculate the conditional probabilities of A, B and $A \vee B$ according to Definition 7 and obtain

$$p_C(A) = \frac{\alpha}{\alpha + \beta + \gamma} \quad p_C(B) = \frac{\beta}{\alpha + \beta + \gamma} \quad p_C(A \vee B) = \frac{\alpha + \beta}{\alpha + \beta + \gamma}$$

showing that Trivalent Additivity holds.

Notably, trivalent conditional probability does *not* satisfy Popper's axiom 6, the Product Rule. The reason for the failure of the Product Rule is inherited from unconditional probability and the non-classical behavior of conjunction in particular: the term $p(A \wedge B)/p(B)$ can be greater than 1, and so we cannot define conditional probability via the familiar Ratio Analysis (i.e., $p(B|A) = p(A \wedge B)/p(B)$). This feature is inherited from

[14] 1′ is immediate. As to 2′, notice that $B_T = \emptyset$. As to 3′, note that $A_{TV} \cap B_F = \emptyset$, and so numerator and denominator of $p_A(B)$ in (Trivalent Conditional Probability) are equal. As to 4′, notice that given $A \models_C B$ and $B \models_C A$, the conditionals $A \to C$ and $B \to C$ take the same truth values in all Cooper valuations.

the non-classical behavior of conjunction in our semantics, but it does not undermine the status of $p_A(B)$ as the appropriate conditional probability function. Of course, the Product Rule is satisfied if we restrict ourselves to conditional-free sentences.

Let us now look at a famous bone of contention for theories of conditionals: Stalnaker's Thesis $p(B \to C) = p(C|B)$ and its generalization $p(B \to C|A) = p(C|A \wedge B)$. Lewis (1976) has shown that any bivalent semantics of conditionals where Stalnaker's Thesis holds will trivialize the probability function, as long as we assume it to be closed under conditionalization. In the wake of this and successor results (such as Bradley 2000 and Fitelson 2015), many theorists opted either for (i) saving Stalnaker's Thesis at the price of abandoning (full) truth conditions for conditionals (e.g., Adams 1975; McGee 1989; Edgington 1995; Ciardelli and Ommundsen 2024), or (ii) declaring Stalnaker's Thesis to be false (e.g., Fitelson 2015; Khoo and Mandelkern 2019; Goldstein and Santorio 2021). As shown by Lassiter (2020), the triviality proofs rely on *classical* features of probability functions, such as the Product Rule and the Law of Total Probability, which do not generally apply to probability functions over trivalent valuations. Indeed, our trivalent definition of conditional probability yields Stalnaker's Thesis as a mathematical fact:

Theorem 3 (Stalnaker's Thesis, general form)**.** *For any $A, B, C \in \mathcal{L}^{\to}$ with $c(A_{TV}) > 0$ and $c((A \wedge B)_{TV}) > 0$, and any trivalent probability function $p : \mathcal{L}^{\to} \mapsto [0, 1]$:*

$$p_A(B \to C) = p_{A \wedge B}(C) \qquad \text{(Stalnaker's Thesis)}$$

The special case $A = \top$ yields the familiar-looking $p(B \to C) = p_B(C)$.

Proof. Follows immediately from Import-Export and Definition 7:

$$p_A(B \to C) = p(A \to (B \to C)) = p((A \wedge B) \to C) = p_{A \wedge B}(C).$$

□

Recently, Fitelson (2022) has shown that we can obtain Stalnaker's Thesis only at the price of giving up the Product Rule. Whenever conjunction is supposed to behave classically, the failure of the Product Rule looks like an unacceptable price for obtaining Stalnaker's Thesis. But in our framework, we have independent reasons for rejecting classical conjunction, and therefore also for giving up the Product Rule (when conditional sentences are involved). Since our account of conditional probability satisfies axioms 1′-5′ and violates only the Product Rule, it is therefore as similar to standard conditional probability as one can hope for if one wants to obtain Stalnaker's Thesis. The triviality results simply show that no probability of conditionals can be a *fully classical* conditional probability function.[15] But we can reject this requirement on independent grounds. Note that triviality can be avoided *without* giving up plausible logical principles that appear to be in tension with Stalnaker's Thesis, such as $A \to B, B \to A, A \to C \models B \to C$, which is valid in C (compare Khoo 2022, p. 149).

Several recent accounts of conditionals do not validate Stalnaker's Thesis without restrictions. For example, Goldstein and Santorio (2021) deny it for right-nested conditionals. (Objections to Stalnaker's Thesis involving left-nested conditionals are discussed in the next section, since left-nesting is closely related to learning conditional information.) It is instructive to study their counterexample. They consider a fair die and the sentences

A: If the die landed even, then if it didn't land on two or four, it landed on six.

B: If the die did not land on two or four, it landed on six.

C: The die landed even.

We agree with Goldstein and Santorio's premise that an adequate theory of the probability of conditionals should assign the values $p(A) = 1$, $p(B) = 1/4$ and $p(C) = 1/2$. Then they note that C ("the die landed even") is just another way of expressing the *material* conditional corresponding to B ("either the die landed 2 or 4, or it landed on 6"). Since indicative conditionals are, on their account, logically stronger than the corresponding material conditionals, B must entail C. On the other hand, $B \wedge C$ entails each of its conjuncts, and so $B \dashv\models B \wedge C$ and therefore also $p(B \wedge C) = p(B)$. Moreover, by Stalnaker's Thesis and the identity $A = C \to B$, we infer that

[15] For example, both classical triviality results by Lewis (1976) and Stalnaker (1976) rely on the Law of Total Probability $p(A \to B) = p((A \to B) \wedge B) + p((A \to B) \wedge \neg B)$, which is invalid for trivalent conditional probability. Without the Law of Total Probability, Stalnaker cannot derive the essential lemma $p(X \to Y|\neg X) = p(X \to Y)$. Lewis also needs an application of the Product Rule in order to derive that $p(A \to C) = p(C)$. Lassiter (2020) is an excellent survey of how Lewis-style triviality results fail to apply to trivalent semantics for this reason.

$p(A) = p(C \to B) = p(B|C)$. But then we obtain a contradiction:

$$1 = p(A) = p(C \to B) = p(B|C) = \frac{p(B \wedge C)}{p(C)} = \frac{p(B)}{p(C)} = 1/2,$$

From our viewpoint, it is tempting to attack the equality $p(B|C) = p(B \wedge C)/p(C)$ since we know that the Product Rule fails in our semantics. But in this concrete case, our account yields $p(B \wedge C) = 1/2$ and the above equality actually holds. Rather, we should reject the equality $p(B \wedge C) = p(B)(= 1/4)$. Due to the non-classical behavior of conjunction, and because B is itself a conditional, $B \wedge C$ can have a greater probability than B. This suffices to block the above argument.[16]

Stalnaker's Thesis is rarely declared invalid because it is believed to be fundamentally mistaken. For example, Goldstein and Santorio agree with the basic intuition about the probability of conditionals as conditional probabilities and they declare Adams's Thesis—the restriction of Stalnaker's Thesis to conditional-free statements—valid. They cannot have Stalnaker's Thesis in full generality because they treat probability as behaving classically with respect to conjunction. Similarly, in the light of the central role of Import-Export both in Gibbard's 1981 collapse result and in the trivialization of Stalnaker's Thesis, Khoo and Mandelkern (2019) give up Import-Export as a logical principle while retaining that two sentences with the logical forms $A \to (B \to C)$ and $(A \wedge B) \to C$ express the same proposition in any context.[17] Such solutions are feasible, but they look to us like workarounds. We find it more attractive and straightforward to preserve Import-Export and an unrestricted form of Stalnaker's Thesis at the price of giving up classical conjunction and the Product Rule, especially since we have independent motivations for such a move.

[16]While the indicative conditional B and its corresponding material conditional C are logically equivalent in the certainty-preserving logic C, the indicative conditional always receives a lower probability. Actually, any simple indicative $X \to Y$ entails $X \supset Y$ in U.

[17]Regarding Import-Export, Ciardelli and Ommundsen (2024) argue, in our opinion convincingly, that the use of Import-Export in the trivialization of Stalnaker's Thesis is innocent from a normative point of view. By contrast, Sanfilippo et al. 2020's treatment of conditionals as conditional random quantities rejects Import-Export on grounds of triviality, despite sharing some features with the trivalent approach. See Castronovo and Sanfilippo 2024 for some comparisons.

6 Bayesian Learning in a Trivalent Setting

We now move on to the *dynamics* of probabilistic reasoning: how should we change our beliefs when we learn conditional propositions such as $A \to B$ (for Boolean A, B), and possibly more complex expressions? The standard theory for inductive learning is Bayesian conditionalization: the rational credence in hypothesis H after learning evidence E should be the rational credence in H conditional on E:

$$p^E(H) := p(H|E) \qquad \text{(Bayesian conditionalization)}$$

This definition does not directly apply to conditionals (e.g., $E = A \to B$) since they are no standard propositions. Extensions of Bayesian conditionalization to $\mathcal{L}^{\to}$ typically rely on a more fine-grained, modal semantics, where worlds are identified with *sequences* of classical valuations (Bacon 2015; Goldstein and Santorio 2021; Khoo 2022). That extension of conditionalization acts on sets of such sequences (e.g., Goldstein and Santorio call it "hyperconditionalization"). At the same time, Bayesian epistemologists have developed various proposals such as updating on the material conditional $A \supset B$ and prior-posterior divergence minimization, without providing a semantic analysis of conditionals (Douven and Romeijn 2011; Eva, Hartmann, and Rafiee Rad 2020).

In this section, we show that Bayesian conditionalization generalizes smoothly from $\mathcal{L}$ to $\mathcal{L}^{\to}$, without having to change the concept of possible world (they remain maximally consistent classical valuations). Specifically, we propose that learning E should change our degree of belief in H from $p(H)$ to $p(E \to H)$, i.e., the trivalent conditional probability of H given E:

Definition 8 (Trivalent Conditionalization)**.** *Suppose we learn $E \in \mathcal{L}^{\to}$ with $c(E_{TV}) > 0$. Then the rational credence in $H \in \mathcal{L}^{\to}$ is the trivalent conditional probability of H, given E:*

$$p^E(H) := p_E(H) = p(E \to H) \qquad \text{(Trivalent Conditionalization)}$$

The first thing to note is that the posterior probability function p^E is indeed a trivalent probability function (proof omitted, but straightforward). Second, sentences that are equivalent in the logic of certain reasoning C produce the same update under trivalent conditionalization. This is not too surprising: learning a proposition E means that E is now a certainty, and if $p(E) = 1$ implies $p(E') = 1$, and vice versa, learning E and learning E' should yield the same updates.

Proposition 5. *For any formulas $E, E', H \in \mathcal{L}^{\rightarrow}$, if $E \Dashv\models_{\mathsf{C}} E'$, then $p^E(H) = p^{E'}(H)$.*

Proof. Follows immediately from the observation that trivalent conditional probability satisfies Left Logical Equivalence with respect to C (Property 4′, p. 19). □

When A and B are conditional-free, we also have $A \supset B \Dashv\models_{\mathsf{C}} A \rightarrow B$, and the above proposition implies the following corollary:

Proposition 6. *For any $H \in \mathcal{L}^{\rightarrow}$, and $A, B \in \mathcal{L}$:*

$$p^{A \rightarrow B}(H) = p^{A \supset B}(H) \qquad \text{(Updating on Simple Conditionals)}$$

In other words, *learning a simple conditional $A \rightarrow B$ is the same as learning the material conditional $A \supset B$.* This prediction is also endorsed by Goldstein and Santorio (2021) and Santorio (2022, Section 7): it is required for explaining why, upon learning a disjunction such as "either the butler or the gardener did it", we also fully accept the sentence "if the butler did not do it, the gardener did it". While Santorio's account does not treat these sentences as logically equivalent, but only as update-equivalent, our account explains the equivalence between learning a simple conditional and learning the corresponding material conditional in terms of their semantic and logical properties in C.

Trivalent conditionalization generalizes Bayesian conditionalization in another important way, too. Conditionalization can be motivated as a special case of a more general updating rule: *minimizing the divergence between prior distribution p and posterior distribution p'.* Indeed, results by Cziszár (1967, 1975) and Diaconis and Zabell (1982) show that for a certain class of divergence functions—the so-called f-divergences—the two following updating policies are equivalent:

(1) Bayesian conditionalization on the event E;

(2) minimizing the f-divergence between the distributions p and p', subject to the constraint that $p'(E) = 1$.

In general, for discrete probability spaces $\Omega = \{\omega_1, \ldots, \omega_n\}$, f-divergences have the form

$$D_f(p, p') = \sum_{i=1}^{N} p(\omega_i) f\left(\frac{p'(\omega_i)}{p(\omega_i)}\right),$$

where $f : \mathbb{R}^{\geq 0} \rightarrow \mathbb{R}$ is a convex and differentiable function satisfying $f(1) = 0$. A well-known f-divergence is the Kullback-Leibler divergence or relative entropy, which is obtained by choosing $f(x) = x \log x$.

The divergence minimization approach not only agrees with Bayesian conditionalization for propositional learning, but it is also independently motivated as a conservative method of belief revision. Incoming evidence should change our beliefs only to the extent that this is strictly required; if possible, we would like to stay close to the original prior distribution. Moreover, in comparison to standard Bayesian conditionalization, it has the advantage of being applicable to a wider variety of constraints on the posterior probability distribution (e.g., constraints which we cannot express in the object language).

We can now show that trivalent conditionalization too minimizes divergence between prior and posterior distribution:

Theorem 4 (Updating Theorem)**.** *The following two procedures for updating credences from a prior probability distribution $p : \mathcal{L}^{\rightarrow} \mapsto [0, 1]$ to a posterior probability distribution $p' : \mathcal{L}^{\rightarrow} \mapsto [0, 1]$ are equivalent:*

(1) trivalent conditionalization on the proposition $E \in \mathcal{L}^{\rightarrow}$;

(2) minimizing the f-divergence between p and p', subject to the constraint that $p'(E) = 1$.

Proof. Suppose we learn $E \in \mathcal{L}^{\rightarrow}$, which, by the Reduction Theorem, can be written as $E \approx A \rightarrow B$, with $A, B \in \mathcal{L}$. Trivalent conditionalization and Proposition 6 yield $p^E = p_{A \rightarrow B} = p_{A \supset B}$, i.e., learning E amounts to conditionalizing on the corresponding material conditional $A \supset B$. Sprenger and Hartmann (2019, Theorem 4.3) show that updating p on the *material* conditional $A \supset B$ is equivalent to minimizing the f-divergence $D_f(p, p')$ subject to the constraint that $p'(B|A) = 1$ (this expression denotes standard conditional probability in $\mathcal{L}$). However, by Adams's Thesis, this constraint on the posterior distribution is equivalent to $p'(E) = p'(A \rightarrow B) = p'(B|A) = 1$. The converse direction makes use of the same identities. □

Since trivalent conditionalization agrees with minimizing f-divergence, it can be defended as a conservative method of belief revision in the Bayesian spirit. And it is much more powerful than standard Bayesian conditionalization: it applies to learning conditionals, and to compounds of conditionals of arbitary complexity. Specifically, the Updating Theorem implies that trivalent conditionalization on $A \to B$, Bayesian conditionalization on $A \supset B$ and minimizing an f-divergence relative to the constraint $p'(B|A) = 1$ are equivalent update procedures.

Some examples in the Bayesian epistemology literature challenge the view that learning a conditional $A \to B$ amounts to learning the material conditional $A \supset B$ (Douven and Dietz 2011; Douven and Romeijn 2011). Moreover, since

$$p((A \to B) \to C) = p((A \supset B) \to C) = p^{A \supset B}(C) = p^{A \to B}(C)$$

such examples immediately affect the tenability of Stalnaker's Thesis for left-nested conditionals. We look at a particularly pressing example (Edgington 1991, p. 202; Khoo 2022, p. 154):

> **Edgington's Coin.** Coin x is either double-headed or double-tailed. Each possibility is equally likely. The coin is flipped with probability 50%, regardless of whether it is double-headed or double-tailed. We learn "if the coin was flipped, it landed heads". What is the now the probability that x is double-headed?

Using the Boolean variables C (coin double-headed), F (flip) and O (outcome heads), we calculate

$$\begin{aligned} p^{F \to O}(C) &= p((F \supset O) \to C) = p(C|F \supset O) = p(C|\neg F \vee O) \\ &= p(C) \times \frac{p(\neg F \vee O|C)}{p(\neg F \vee O)} = 1/2 \times \frac{1}{3/4} \\ &= 2/3 \end{aligned}$$

Edgington and Khoo argue that after learning $F \to O$, we should actually be *certain* that the coin is double-headed. The predicted value of 2/3 is therefore too low.[18] What is driving their intuition? Imagine a reliable source tells us "if the coin was flipped, it landed heads". The most natural

[18]Jeffrey (1991), Kaufmann (2009, 2023) and Khoo (2022) obtain $p(C \to (F \to O)) =$ 3/4. Our calculation differs because trivalent conditional probability does not satisfy the Product Rule and Bayes' Theorem for conditional expressions.

reason for why she can assert $F \to O$ is that she knows that the coin is double-headed. Hence, it is natural to infer that she must know C as well, and we add this proposition to our evidence base. In other words, we learn in this case more than the bare conditional.

Compare this to a scenario where our source has evidence that the coin landed heads rather than tails if flipped, but admits the possibility that the coin may not have been flipped. In this case, she asserts "if the coin was flipped, it landed heads", but here we should be less than certain that the coin is double-headed: after all, the coin might not have been flipped at all. Therefore we propose to account for the intuition that $p^{F \to O}(C) = 1$ as a pragmatic rather than a semantic phenomenon. Note that this line of response is not *ad hoc*: Eva, Hartmann, and Rafiee Rad (2020) and Sprenger and Hartmann (2019, ch. 4) develop the same strategy and apply it to a variety of puzzling cases of learning conditional information.

A similar attack against Stalnaker's Thesis for left-nested conditionals (and implicitly, against the equivalence between learning a simple conditional and learning the material conditional) has been launched by Justin Khoo (2022, pp. 152-153) on the basis of the following example:

> **Poker Paul.** Paul is playing poker against Nancy. Nancy has a weak hand, but it is still possible Paul's is weaker. It is also possible for Paul to win even if he has the weaker hand, but since Nancy is a good player, that is unlikely. Cheating is a way to increase your success of winning with a weaker hand, and Paul is not opposed to cheating, though it is unlikely he cheated if he had the better hand. But, if Paul won with the weaker hand, it is overwhelmingly likely that he cheated.

The point of contention is the probability of the left-nested conditional

(4) If Paul won if he had the weaker hand, he cheated.

Khoo represents the relevant sentences by H = "Paul has the weaker hand", W = "Paul wins", C = "Paul is cheating". On the basis of his probabilistic model, we assign a rather low probability to (4) ($p((H \to W) \to C) \approx 0.25$)—intuitively way too low according to Khoo. Kaufmann (2023, pp. 216-217) counters that it is overall very *unlikely* that Paul cheated, because probably he has the better hand. Learning that he won if he had the weaker hand raises our confidence that Paul cheated, but not necessarily to a high level. Khoo's intuition is only supported if we interpret (4) along one of the following lines:

(5) If Paul won *with the weaker hand*, he cheated.

(6) If Paul won if he had the weaker hand, he cheated [if he had the weaker hand].

Both (5) and (6) seem intuitively probable, and our semantics confirms this verdict.[19] The reinterpretations may be defended by the observation that if-clauses frequently set topics in a discourse, or restrict the scope of what is said in the main clause (e.g., Lewis 1975; Égré and Cozic 2011; Kratzer 2012; Kaufmann 2023). In this view, "if he had the weaker hand" also restricts the scope of evaluation of the consequent of the outer conditional, and (4) can be rewritten as (6). However, such a reinterpretation would have to be motivated independently, and the burden of showing that (5) or (6) are the right reading of (4) is on their proponents. Our reading, by contrast, follows the original syntactic structure of (4), and does not involve syntactic reanalysis. While these and other data need to be investigated more thoroughly, it seems to us that we can satisfactorily address alleged counterexamples to Proposition 6 and Stalnaker's Thesis for left-nested conditionals.

Summing up, trivalent conditionalization generalizes Bayesian conditionalization from $\mathcal{L}$ to $\mathcal{L}^{\rightarrow}$, without changing its basic mechanism. This result is achieved within the bounds of a purely extensional, truth-functional semantics. Competing accounts which obtain similar results rely on *sequences* of possible worlds as semantic building blocks and probability is defined over such sequences, too. The consequence is a strong form of contextualism: the evidential base and past update operations affect the *meaning* of the conditionals. Accounts along these lines can get extremely complex (van Fraassen 1976; Kaufmann 2009; Bacon 2015), and the less complex and more intuitive ones do not validate Stalnaker's Thesis in general (Goldstein and Santorio 2021; Khoo 2022). Our approach is simpler, conceptually leaner and closer to the standard Bayesian framework.

7 Conclusions

The present paper has proposed a trivalent approach to both the truth conditions and the probability of indicative conditionals, based on the idea of assigning a third truth value when the antecedent is false. We have shown that this semantics, when paired with an appropriate consequence relation

[19]For (5), this is obvious. Moreover, Import-Export implies that (5) and (6) are semantically equivalent: $(H \rightarrow W) \rightarrow (H \rightarrow C) = ((H \rightarrow W) \wedge H) \rightarrow C) = (H \wedge W) \rightarrow C$.

for uncertain reasoning, generalizes Adams's logic of p-valid inference to arbitrary compounds and nestings of conditionals. Moreover, this account (i) shows that all complex conditionals can be rephrased equivalently as simple conditionals, (ii) validates Stalnaker's Thesis for the probability of conditionals in its most general form; (iii) models the learning of (conditional) information by means of generalizing Bayesian conditionalization to updating on conditionals. To our knowledge, no other account offers a unified analysis of the semantics, probability and learning of indicative conditionals based on a simple truth-functional (albeit non-classical) semantics.

In particular, we vindicate Adams's conjecture that compound conditionals represent "linguistic shortcuts". Theorem 1 shows that they are extensionally equivalent to simple conditionals. Therefore we can evaluate any inferences with complex conditionals by translating them to the fragment $\mathcal{L}_1^{\to}$ involving at most simple conditionals, and apply the rules of p-valid inference (Theorem 2). Trivalent semantics provides fully truth-functional, compositional truth conditions *and* an Adams-style theory of probabilistic reasoning. Committing to the view that indicative conditionals do not have truth conditions is unnecessary philosophical baggage for proponents of an Adams-style approach: it bars the road to important insights into how we use complex conditionals, how we learn them, and how valid inference relates to truth preservation.

The second part of the paper amends this semantics with an account of conditional probability and Bayesian learning. Conditional probability $p_A(B)$ is *defined* as the probability of the conditional $A \to B$ (for all $A, B \in \mathcal{L}^{\to}$). We have argued that this is a natural generalization of conditional probability to the trivalent case. Stalnaker's Thesis in its general form, i.e., $p_{A \wedge B}(C) = p_A(B \to C)$, follows, for all sentences of $\mathcal{L}^{\to}$, as an immediate corollary, without falling prey to Lewis-style triviality results or having to restrict the scope of plausible principles such as Import-Export (Theorem 3). Instead, triviality is avoided because trivalent probability does not satisfy the Product Rule. The reason behind this is the (independently motivated) non-classical behavior of conjunction in our semantics. This allows us not only to obtain Stalnaker's Thesis, but also to dodge the impossibility result presented by Schulz (2009) for representing uncertain inference with compounds of conditionals in trivalent logic.

Finally, Theorem 4, the Updating Theorem, establishes that trivalent conditionalization on a proposition E minimizes the f-divergence between prior distribution p and posterior distribution p', subject to the constraint that $p'(E) = 1$. This shows that trivalent conditionalization in a language with conditionals preserves a central epistemological motivation of Bayesian conditionalization. Specifically, updating on a simple, non-nested indicative conditional is equivalent to updating on the corresponding material conditional—a prediction that agrees with the results by Goldstein and Santorio (2021) and Santorio (2022). Despite our agreement with these and other authors on crucial predictions, we consider our account simpler, more unified, and more attractive in its results.

References

Adams, Ernest W. (1965), "The Logic of Conditionals", *Inquiry*, 8, pp. 166-197.

— (1975), *The Logic of Conditionals*, Reidel, Dordrecht.

— (1986), "On the Logic of High Probability", *Journal of Philosophical Logic*, 15, pp. 255-279.

— (1995), "Remarks on a theorem of McGee", *Journal of Philosophical Logic*, 24, pp. 343-348.

— (1998), *A Primer on Probability Logic*, Center for the Study of Language and Information, Stanford/CA.

Bacon, Andrew (2015), "Stalnaker's Thesis in Context", *Review of Symbolic Logic*, 8, pp. 131-163.

Belnap, Nuel D. (1970), "Conditional assertion and restricted quantification", *Noûs*, 4, pp. 1-12.

— (1973), "Restricted quantification and conditional assertion", in *Truth, Syntax and Modality*, ed. by Hugues Leblanc, Elsevier, Amsterdam, vol. 68, pp. 48-75.

Bradley, Richard (2000), "A Preservation Condition for Conditionals", *Analysis*, 60, pp. 219-222.

— (2002), "Indicative Conditionals", *Erkenntnis*, 56, pp. 345-378.

Cantwell, John (2006), "The laws of non-bivalent probability", *Logic and Logical Philosophy*, 15, pp. 163-171.

— (2008), "The logic of conditional negation", *Notre Dame Journal of Formal Logic*, 49, pp. 245-260.

Castronovo, Lydia and Giuseppe Sanfilippo (2024), "A probabilistic analysis of selected notions of iterated conditioning under coherence", *International Journal of Approximate Reasoning*, 165, p. 109088.

Ciardelli, Ivano (2020), "Indicative conditionals and graded information", *Journal of Philosophical Logic*, 49, pp. 509-549.

Ciardelli, Ivano and Adrian Ommundsen (2024), "Probabilities of Conditionals: Updating Adams", *Noûs*, 58 (1 2024), pp. 26-53.

Cooper, William S. (1968), "The propositional logic of ordinary discourse", *Inquiry*, 11, pp. 295-320.

Cziszár, Imre (1967), "Information Type Measures of Difference of Probability Distributions and Indirect Observations", *Studia Scientiarum Mathematicarum Hungarica*, 2, pp. 299-318.

— (1975), "I-Divergence Geometry of Probability Distributions of Minimization Problems", *Annals of Probability*, 3, pp. 146-158.

de Finetti, Bruno (1936), "La logique de la probabilité", in *Actes du congrès international de philosophie scientifique*, Hermann Editeurs, Paris, vol. 4, pp. 1-9.

— (1974), *Theory of Probability*, John Wiley & Sons, New York.

Diaconis, Persi and Sandy L. Zabell (1982), "Updating Subjective Probability", *Journal of the American Statistical Association*, 77, pp. 822-830.

Douven, Igor and Richard Dietz (2011), "A Puzzle about Stalnaker's Hypothesis", *Topoi*, 30, pp. 31-37.

Douven, Igor and Jan-Willem Romeijn (2011), "A New Resolution of the Judy Benjamin Problem", *Mind*, 120, pp. 637-670.

Dubois, Didier and Henri Prade (1994), "Conditional objects as nonmonotonic consequence relationships", *IEEE Transactions on Systems, Man, and Cybernetics*, 24, pp. 1724-1740.

Edgington, Dorothy (1991), "Matter-of-Fact Conditionals: Part II", *Proceedings of the Aristotelian Society*, 65, pp. 185-209.

— (1995), "On Conditionals", *Mind*, 104, pp. 235-329.

Égré, Paul and Mikaël Cozic (2011), "If-clauses and probability operators", *Topoi*, 30, pp. 17-29.

Égré, Paul, Lorenzo Rossi, and Jan Sprenger (2021a), "De Finettian Logics of Indicative Conditionals. Part I: Trivalent Semantics and Validity", *Journal of Philosophical Logic*, 50, pp. 187-213.

— (2021b), "De Finettian Logics of Indicative Conditionals. Part II: Proof Theory and Algebraic Semantics", *Journal of Philosophical Logic*, 50, pp. 215-247.

— (forthcoming), "Certain and Uncertain Inference with Trivalent Conditionals", *Australasian Journal of Philosophy*.

Eva, Benjamin, Stephan Hartmann, and Soroush Rafiee Rad (2020), "Learning from Conditionals", *Mind*, 129, pp. 461-508.

Evans, Jonathan and David Over (2004), *If: Supposition, Pragmatics, and Dual Processes*, Oxford University Press.

Fitelson, Branden (2015), "The Strongest Possible Lewisian Triviality Result", *Thought*, 4, pp. 69-74.

— (2022), "Stalnaker's Thesis is Incompatible with the Product Rule", Note dated 12/12/2022, presented at the Berkeley-Paris workshop on Probability and Meaning, UC Berkeley.

Gibbard, Allan (1981), "Two Recent Theories of Conditionals", in *Ifs*, ed. by William Harper, Robert C. Stalnaker, and Glenn Pearce, Reidel, Dordrecht, pp. 211-247.

Goldstein, Simon and Paolo Santorio (2021), "Probability for Epistemic Modalities", *Philosophers' Imprint*, 21.

Goodman, I.R., H.T. Nguyen, and E.A. Walker (1991), *Conditional Inference and Logic for Intelligent Systems*, North-Holland, Amsterdam.

Hájek, Alan (2011), "Triviality Pursuit", *Topoi*, 30, pp. 1-15.

Hawthorne, James (2016), "A Logic of Comparative Support: Qualitative Conditional Probability Relations Representable by Popper Functions", in *The Oxford Handbook of Probability and Philosophy*, ed. by Alan Hájek and Christopher Hitchcock, Oxford University Press, Oxford, pp. 277-295.

Humberstone, Lloyd (2011), *The Connectives*, MIT Press, Cambridge/MA.

Jeffrey, Richard C. (1991), "Matter-of-Fact Conditionals: Part I", *Aristotelian Society Supplementary Volume*, 65, pp. 161-183.

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

— (2023), "Bernoulli Semantics and Ordinal Semantics for Conditionals", *Journal of Philosophical Logic*, 52, pp. 199-220.

Khoo, Justin (2022), *The meaning of "If"*, Oxford University Press, Oxford/new York.

Khoo, Justin and Matthew Mandelkern (2019), "Triviality results and the relationship between logical and natural languages", *Mind*, 128, pp. 485-526.

Kleiter, Gernot (2018), "Adams's *p*-validity in research on human reasoning", *Journal of Applied Logic*, 5, pp. 775-825.

Kratzer, Angelika (2012), *Modals and Conditionals. New and Revised Perspectives*, Oxford University Press.

Lassiter, Daniel (2020), "What We Can Learn From How Trivalent Conditionals Avoid Triviality", *Inquiry: An Interdisciplinary Journal of Philosophy*, 63, pp. 1087-1114.

Lewis, David (1975), "Adverbs of Quantification", in *Formal semantics of natural language: papers from a colloquium sponsored by the King's College Research Centre, Cambridge*, ed. by Edward Louis Keenan, Cambridge University Press, pp. 3-15.

— (1976), "Probabilities of Conditionals and Conditional Probabilities", *Philosophical Review*, 85, pp. 297-315.

Mandelkern, Matthew (2020), "Import-Export and 'And' ", *Philosophy and Phenomenological Research*, 100, pp. 118-135.

McDermott, Michael (1996), "On the truth conditions of certain 'if'-sentences", *The Philosophical Review*, 105, pp. 1-37.

McGee, Vann (1981), “Finite Matrices and the Logic of Conditionals”, *Journal of Philosophical Logic*, 10, pp. 349-351.

— (1985), “A Counterexample to Modus Ponens”, *The Journal of Philosophy*, 82, pp. 462-71.

— (1989), “Conditional probabilities and compounds of conditionals”, *The Philosophical Review*, 98, pp. 485-541.

Milne, Peter (2003), “The Simplest Lewis-Style Triviality Proof Yet?”, *Analysis*, 63, pp. 300-303.

Olkhovikov, Grigory (2002/2016), “On a New Three-Valued Paraconsistent Logic”, *IFCoLog Journal of Logic and Applications*, 3, Originally published in Russian in 2002, in *Logic of Law and Tolerance*, pages 96–113, Yekaterinburg. English translation by T. Ferguson., pp. 317-334.

Pearl, Judea (1988), *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann.

Quine, W. V. O. (1950), *Methods of Logic*, Harvard University Press, Cambridge/MA.

Ramsey, Frank P. (1929/1990), “General Propositions and Causality”, in *Philosophical Papers*, ed. by David H. Mellor, reprint of the original 1929 paper, Cambridge University Press, Cambridge, pp. 237-257.

Rothschild, Daniel (2014), “Capturing the relationship between conditionals and conditional probability with a trivalent semantics”, *Journal of Applied Non-Classical Logics*, 24, pp. 144-152.

Sanfilippo, Giuseppe, Angelo Gilio, David Over, and Niki Pfeifer (2020), “Probabilities of conditionals and previsions of iterated conditionals”, *International Journal of Approximate Reasoning*, 121, pp. 150-173.

Santorio, Paolo (2022), “Path Semantics for Indicative Conditionals”, *Mind*, 131, pp. 59-98.

Santorio, Paolo and Alexis Wellwood (2023), “Nonboolean Conditionals”, *Experiments in Linguistic Meaning*, 2, pp. 252-264.

Schulz, Moritz (2009), “A note on two theorems by Adams and McGee”, *Review of Symbolic Logic*, 2, pp. 509-516.

Sobociński, Bolesław (1952), “Axiomatization of a partial system of three-value calculus of propositions.” *The Journal of Computing Systems*, 1, pp. 23-55.

Sprenger, Jan and Stephan Hartmann (2019), *Bayesian Philosophy of Science*, Oxford University Press, Oxford.

Stalnaker, Robert (1970), “Probability and Conditionals”, *Philosophy of Science*, 37, pp. 64-80.

Stalnaker, Robert (1976), "Letter to Bas van Fraassen", in *Foundations of probability theory, statistical inference, and statistical theories of science*, ed. by W. Harper and C. Hooker, pp. 302-306.

Van Fraassen, Bas (1976), "Probabilities of Conditionals", in *Foundations of probability theory, statistical inference, and statistical theories of science*, ed. by W. Harper and C. Hooker, pp. 261-300.

Van Wijnbergen-Huitink, Janneke, Shira Elqayam, and David Over (2015), "The probability of iterated conditionals", *Cognitive Science*, 39, pp. 788-803.

Williams, Robert G. (2016), "Probability and Nonclassical Logic", in *The Oxford Handbook of Probability and Philosophy*, ed. by Alan Hájek and Christopher Hitchcock, Oxford University Press, pp. 248-276.

A Proof of the Reduction Theorem

We reason by induction on the depth of formulae, assigning formulae of $\mathcal{L}_1^{\rightarrow}$ a depth of 0, and positive depth to compounds of $\mathcal{L}_1^{\rightarrow}$-formulae containing at least one occurrence of $\rightarrow$. This depth $d(X)$ is defined recursively as follows:

$$d(X) = \begin{cases} 0 & \text{if } X \in \mathcal{L}_1^{\rightarrow} \\ d(Y) + 1 & \text{if } X \in \mathcal{L}^{\rightarrow} \setminus \mathcal{L}_1^{\rightarrow} \text{ has the form } \neg Y \\ \max[d(Y), d(Z)] + 1 & \text{if } X \in \mathcal{L}^{\rightarrow} \setminus \mathcal{L}_1^{\rightarrow} \text{ has the form } Y \circ Z \\ & \qquad \text{and } \circ \in \{\wedge, \vee, \rightarrow\} \end{cases}$$

For the base case $d(X) = 0$, i.e., $X \in \mathcal{L}_1^{\rightarrow}$, the theorem holds trivially. We now consider more complex formulae $X \notin \mathcal{L}_1^{\rightarrow}$ case by case, according to the formula's main connective.

Negation Suppose $X = \neg X'$ for some $X' \in \mathcal{L}^{\rightarrow}$. We observe $d(X) = d(X') + 1$, and so, by the induction hypothesis, there are $A', B' \in \mathcal{L}$ such that $X' \approx A' \rightarrow B'$. So $X \approx \neg(A' \rightarrow B')$. The Negation Commutation property stated in Proposition 1 yields $X \approx A' \rightarrow \neg B' \in \mathcal{L}_1^{\rightarrow}$.

Binary Connectives Suppose $X = Y \circ Z$, with $\circ \in \{\wedge, \vee, \rightarrow\}$, and $Y, Z \in \mathcal{L}^{\rightarrow}$. Obviously, $d(Y) < d(X)$ and $d(Z) < d(X)$. Thus we apply the inductive hypothesis to Y and Z and infer $Y \approx A \rightarrow B$ and $Z \approx C \rightarrow D$ with $A, B, C, D \in \mathcal{L}$. We have to consider three cases:

$X = Y \rightarrow Z$

$$\approx (A \to B) \to (C \to D) \qquad \text{(by compositionality)}$$
$$\approx (A \supset B) \to (C \to D) \qquad \text{(by Prop. 1, Left-Nesting)}$$
$$\approx ((A \supset B) \land C) \to D \qquad \text{(by Prop. 1, Import-Export)}$$

$$X = Y \land Z$$
$$\approx (A \to B) \land (C \to D) \qquad \text{(by compositionality)}$$
$$\approx (A \lor C) \to [(A \supset B) \land (C \supset D)] \quad \text{(by Prop. 1, Conjunction)}$$

$$X = Y \lor Z$$
$$\approx (A \to B) \lor (C \to D) \qquad \text{(by compositionality)}$$
$$\approx (A \lor C) \to ((A \land B) \lor (C \land D)) \qquad \text{(by Prop. 1, Disjunction)}$$

In all three cases we have shown that X is semantically equivalent to a $\mathcal{L}_1^{\to}$-formula, completing the proof.