STUDIA SEMIOTYCZNE (SEMIOTIC STUDIES), 37(1), 45–74
ISSN 0137-6608, e-ISSN 2544-073X
DOI: 10.26333/sts.xxxvii1.04

Article

ANNA WÓJTOWICZ [*], KRZYSZTOF WÓJTOWICZ [**]

THE SUBJECTIVE PROBABILITY OF CONDITIONALS AND ITS FORMALIZATIONS[1, 2]

SUMMARY: The aim of the paper is to discuss the problem of the assignment of the degree of belief to conditional sentences by a rational agent. After presenting the general methodological framework, we analyze two possible approaches. In both cases we assume that we have a non-conditional probabilistic system of beliefs expressed in a language L_0, and modeled in some initial probability space S, which allows us to assign probabilities to sentences from L_0. Our aim is to extend this system of beliefs to a given class of sentences Φ containing conditionals.

The first approach is what we call the "credence-like" approach: for a given class Φ, we define credence as a function defined directly on linguistic objects. The second approach consists in assuming the existence of a standard probability space, in which the sentences from the set Φ are interpreted as events. In this case, the degree of belief of α is defined as the probability of the corresponding event in the probability space S_Φ. We present both of these approaches, indicating what their advantages and disadvantages are. The thesis of the article is that the probabilistic approach to the analysis of degrees of belief of sentences containing the conditional connective is by far the more universal one and preferable method, particularly in the context of more complex conditionals.

KEYWORDS: conditionals, probability of conditionals, credence of conditionals, Markov graphs, PCCP.

[*] University of Warsaw, Faculty of Philosophy. E-mail: amwojtow@uw.edu.pl. ORCID: 0000-0002-6177-964X.

[**] University of Warsaw, Faculty of Philosophy. E-mail: wojtow@uw.edu.pl. ORCID: 0000-0002-1187-8762.

[1] We express our gratitude to the anonymous referee for providing valuable insights on matters discussed in the paper and a very careful reading of the manuscript.

[2] The preparation of this paper was supported by the National Science Centre [pol. *Narodowe Centrum Nauki*] grant 2020/39/B/HS1/01866.

1. Introduction

We are confronted with conditionals virtually everywhere: in everyday situations, in political discussions, in scientific discourse—and of course in philosophical analysis. When discussing technical problems, we make statements like "If the temperature was higher, this piece of metal would melt"; in the context of medicine, we make claims concerning the likely outcomes of a treatment that was not undertaken. When discussing political fiction, we might like to consider "If Reagan worked for the KGB, I will never find out" (Lewis, 1986, p. 155) or the famous Oswald-Kennedy examples: "If Oswald did not kill Kennedy, someone else" did and "If Oswald had not killed Kennedy, someone else would have" given in Adams' (1970). And when thinking about our fate as philosophers, we might analyze statements like "If I became a football player, I would be happy".

A lively and rigorous discussion is currently in progress concerning the appropriate logical description of conditionals and its scope of applicability. To be specific, an important question is whether it is confined solely to conditionals in the indicative mood or encompasses subjunctive or counterfactual instances. In this paper, we will confine ourselves to conditionals in the indicative mood, in accordance with the approach taken by McGee (1985) and Kaufmann (2004). We do not make strong assumptions except for asserting that in specific situations, sentences containing conditionals are true, while in others, they are false, and that the conditional connective $\rightarrow$ is not reducible to material implication. This renders the problem of the proper logic of conditionals pressing: what are the appropriate axioms and the rules of inference? Is some version of the Law of Excluded Middle true? Is the Import-Export Principle concerning nested conditionals true and valid?[3] Presenting an adequate semantics poses difficulties—it is far from clear what the appropriate structures are and how the problem of the truth values of conditionals should be handled.[4]

The assignment of degrees of belief to conditionals is another intricate problem, and in this paper we will focus on this issue. Sometimes we are only able to give very rough, qualitative estimates (for instance, low versus high) or might only have intuitions concerning their relative likelihood. Indeed, most philosophers would consider the conditional claim "If I had become a lawyer, I would be rich" as more likely than "If I had become a poet, I would be rich"—even if it is quite problematic to assign them precise numerical values.

In many cases, we are also able to give quantitative estimates. Take a fair die as a toy example: we fully agree on the probability of non-conditional propositions like "It is an even number" or "It is less than 4", etc. But we also have

[3] According to the logical version of the Import-Export Principle, the right-nested conditional $A \rightarrow (B \rightarrow C)$ is equivalent to $(A \wedge B) \rightarrow C$.

[4] For instance, Edgington (1995) and Gibbard (1981) claim that conditionals are not factual sentences.

intuitions concerning conditional beliefs: for instance, the chance of "If it was even, it would be a six" is 1/3—not 0.15 or 0.99.

Our degree of belief concerning conditionals depends on our base knowledge, i.e., on the initial, non-conditional probability distribution. Indeed, the likelihood of "If it was even, it would be a six" depends on our knowledge of the probability distribution on the die. For a fair die, it is 1/3, but for a biased die the estimate would be quite different.

Terminological clarification is necessary, as the term "probability" is used in the literature in various senses. Specifically, it is employed in an intuitive sense, referring to our subjective judgments—and in this context, the term "degree of belief" is more appropriate. In general, it might denote any assignment of a numerical parameter to propositions, intended to model the epistemic attitude of the agent in some way (terms such as "likelihood", "credibility", "assertability", or "subjective probability" are used here). In this paper, we use the term "probability" in the orthodox, mathematical textbook sense, i.e., when referring to a probability distribution P defined in a probability space $S = (\Omega, \Sigma, P)$. In other cases, when considering only a numerical assignment intended to express or measure the agents' epistemic attitude, we use the term "credence". The term "the degree of belief" is used to emphasize that this assignment has a pretheoretic, intuitive character. Constructing a formal model that is consistent with these concepts is always challenging because our intuitions are often clear only for simple cases (such as simple conditionals). The problem arises of how—in a consistent manner—to extend them to more complex sentences of a given language.

The general problem can be formulated as follows: let L_0 be the base language concerning non-conditional beliefs, modeled in a probability space $S = (\Omega, \Sigma, P)$. Consider its expansion Φ which contains conditionals (we will present more precise definitions later). Our task is to extend our probabilistic beliefs concerning L_0 in such a way that it also accounts for the new sentences from Φ. Obviously, this is not a purely technical problem. We aim to define this extension in a formally correct fashion, which also takes into account our intuitions concerning the degree of belief in conditionals—and fulfills reasonable methodological criteria.[5]

In the paper we will discuss two possible approaches:

[5] Our intuitions might be vague and imprecise and even misleading. But we definitely need to take them into account—either by incorporating them into the model or by explaining why we should consider them misleading. Explaining sources of misunderstanding is an important point in clarifying notions.

1) The first consists in defining the notion of the degree of belief in a direct way, i.e., we formalize it as the credence function Cr: $\Phi \rightarrow [0,1]$. In this case, no probability space is constructed.
2) The second consists in constructing a probability space S_Φ in which sentences from the set Φ are given interpretation as events. The degrees of belief are identified with the probability function P_Φ in S_Φ.[6]

Both these approaches differ in their presuppositions, and of course the technical "implementation" of our intuitive beliefs also differ. We will consider these two cases separately, indicating what their advantages and disadvantages are.

In our opinion, the probabilistic approach has a fundamental advantage over an approach based on the notion of credence.

The structure of the paper is as follows:

In Section 2, *Meaning Postulates Concerning* →, we address the problem of meaning postulates characterizing the conditional connective—and the problem of incorporating them into the formal model. We also give a formal definition of a hierarchy of languages $L_0 \subseteq L_1 \subseteq L_2 \subseteq L_3 \subseteq \ldots$, containing conditionals. L_0 is the base language, in which non-conditionals claims are expressed.

In Section 3, *The Credence-Like Approach*, we discuss the idea of formalizing the notion of degree of belief as a function defined on the class Φ, without assuming the existence of a probability space.

In Section 4, *PCCP: How Should Credences Be Evaluated?*, we briefly present the claim that the probability of conditionals is conditional probability (in the initial probability space $S = (\Omega, \Sigma, P)$). PCCP is a much-debated topic and is also an important example in our presentation.

In Section 5, *The Case of Complex Conditionals*, we indicate that, for instance, Adams' solution to the problem of the credence of conditionals has a very limited range, and we need to address also more complex conditionals, which are themselves also natural.

In Section 6, *The Probability-Space Approach*, we discuss the alternative approach in which a probability space is constructed, where there is a standard probability space in which sentences from Φ are interpreted as events.

In Section 7, *A Brief Comparison and Conclusion*, we compare these two approaches and argue that, for methodological reasons, the probability-space approach is much better.

The *Appendix* contains a brief overview of the general idea of the Markov graph model.

[6] To be precise, the degree of belief of α is formalized as the probability P_Φ of the event $[\alpha]$ in S_Φ, which is the semantic counterpart of α.

2. Meaning Postulates Concerning →

Regardless of which solution (i.e., the "credence-like" versus the probability space approach) we are going to choose to model degrees of belief, it is natural to think of some universal constraints which any such function should satisfy. An obvious requirement is that the credence or probability assignment on conditionals is an extension of the base system of beliefs. We will also impose some restrictions on the interpretation of the connective →. This is natural: we intend to model conditionals, not just some arbitrary two-argument connective.

Consider the following meaning postulates for conditionals:

(1) The truth of $A \wedge B$ guarantees the truth of the conditional $A \rightarrow B$.[7]

(2) The truth of $A \wedge \neg B$ implies the falsity of the conditional $A \rightarrow B$.

(3) The conditional $A \rightarrow \neg B$ is true precisely when the conditional $A \rightarrow B$ is false.

(4) If $\neg A$ obtains, our degree of belief as to $A \rightarrow B$ does not change.

These postulates guarantee us that the conditional is surely true in some situations and surely false in others (postulates 1 and 2), that the negation of a conditional is another conditional (postulate 3), and also that the conditional connective is neither the conjunction nor the material implication (postulate 4).[8] These

[7] This is a special case of a general rule, which states that if we accept $\alpha \wedge \beta$, then we accept $\alpha \rightarrow \beta$ (for any two sentences α and β—also containing conditionals). The principle mentioned here concerns the simplest case when A and B are sentences from L_0. The general version is discussed, for instance, in Egré and Rott's (2021), Cruz, Over, Oaksford, and Baratgin's (2016), Berto and Özgün's (2021), and is known under the names *And-to-If* (this is the terminology we prefer) or Conjunctive Sufficiency.

Not all authors accept it, even in its simplest version. Their concern is that the information contained in the antecedent and consequent in the conditional should be somehow mutually relevant. It is true that Berlin is the capital of Germany, and it is also true that people are mortal. However, as opponents of *And-to-If* would claim, one is not justified in accepting the conditional "If Berlin is the capital of Germany, then people are mortal": people are mortal not *in virtue* of Berlin being the capital. And a conditional worthy of its name should take conceptual dependencies into account.

We will not discuss this principle here. Our opinion is that it is justified—it also holds in many formal models. For an interesting discussion, see, for instance, Berto and Özgün's (2021). They discuss the logical version, i.e., $A \wedge \beta \vDash A \rightarrow \beta$, involving the notion of logical consequence. They contend that:

> A number of mainstream theories of indicatives validate And-to-If: the material conditional view (Grice, 1989; Jackson, 1987) and the probabilistic-suppositional view (Adams, 1975; Edgington, 1995; Evans, Over, 2004), for instance, have it. (Berto, Özgün, 2021, p. 3701)

[8] If we interpreted → as a conjunction, then the left side of the equation would always be 0, and the right side not always so. If → were the material implication, then the left side would be the degree of belief in $\neg A$, and the right side not necessarily so.

postulates establish the basic relations between the classical Boolean connectives and the conditional connective $\rightarrow$. They can be regarded as certain minimum requirements that should be reflected in the definitions of credence and probability. However, our aim is not to discuss the general problem of logical rules which might be accepted in different logical systems formalizing conditionals (a survey can be found, for instance, in Egré, Rott 2021) but rather to indicate how such constraints are "implemented" in formal models. We have chosen these particular postulates because they are referred to by the authors cited in the present text. Our aim is to discuss two different approaches to assessing the degrees of belief of conditionals, rather than debating a specific set of postulates. So our analysis operates somewhat at a meta-theoretical level in relation to these specific meaning postulates characterizing $\rightarrow$.[9]

In our considerations we start with the base language L_0, which is closed under $\neg$ and $\wedge$ (the other Boolean functors are definable). The underlying propositional logic is classical. The language L_0 is used to express non-conditional propositions.

By induction, we define a hierarchy of languages:

$$L_{2n+1} = L_{2n} \cup \{(\varphi \rightarrow \psi): \varphi, \psi \in L_{2n}\};$$
$$L_{2n+2} = L_{2n+1} \cup \{\neg(\varphi), (\varphi \wedge \psi): \varphi, \psi \in L_{2n+1}\}, \text{ for } n = 0, 1, \ldots$$

[9] An interesting example of a different system of such postulates has been given by van Fraassen (1976). Van Fraassen considers it to be "the minimal logic of conditionals suitable for probabilification" (van Fraassen, 1976, pp. 277–278)—i.e., in any structure where probability is defined, these axioms must hold. Here they are:

(I) $(A \rightarrow C) \wedge (A \rightarrow B) \Leftrightarrow (A \rightarrow (C \wedge B))$;

(II) $((A \rightarrow C) \vee (A \rightarrow B)) \Leftrightarrow (A \rightarrow (C \vee B))$;

(III) $(A \wedge (A \rightarrow B)) \Leftrightarrow (A \wedge B)$;

(IV) $(A \rightarrow A) = K$ (the set of all possible worlds);

Van Fraassen also assumes PCCP (i.e., $P(A \rightarrow B) = P(B|A)$) and he presents a construction in which all these assumptions hold. Here we use the symbol $P(B|A)$ for the conditional probability of B given A, which is formally defined as:

$$P(B|A) = \frac{P(A \cap B)}{P(A)},$$

for $P(A) > 0$. Here A, B are events in a probability space, and $\cap$ is the set-theoretic intersection. However, in the literature, the "linguistic" notation is also used, where P refers directly to sentences, and in this case what is meant is:

$$P(B|A) = \frac{P(A \wedge B)}{P(A)}.$$

In relation to the van Fraassen system, one can also contemplate the question of whether the credence-like or probability-like approach is superior. Van Fraassen's set of assumptions is satisfied in two of the models mentioned in Section 6 (we give a sketchy presentation of one model there).

In this way, we obtain an ascending chain of languages $L_0 \subseteq L_1 \subseteq L_2 \subseteq L_3 \subseteq \ldots$[10] At the odd steps, we add all conditionals of the form $\varphi \rightarrow \psi$. At the even steps, we add their Boolean combinations. For example, L_1 contains every simple conditional (but no other combinations of them). Language L_2 contains also all their Boolean combinations, for instance $(A \rightarrow B) \wedge \neg(C \rightarrow D)$. However, it does not contain nested conditionals like $A \rightarrow (C \rightarrow D)$ or $(A \rightarrow B) \rightarrow (C \rightarrow D)$, which appear in L_3. In L_4 we have $(A \rightarrow (C \rightarrow D)) \wedge \neg(E \rightarrow D)$— and so on.

In this definition, we do not pay attention to the probabilities of sentences in the original space S. So in particular we include conditionals of the form $A \rightarrow B$ also when $P(A) = 0$.[11]

We are now ready to formulate the main problem of the paper. Consider L_0, i.e., a language built only by using the conjunction ($\wedge$) and negation ($\neg$), and a probability space $S = (\Omega, \Sigma, P)$, which allows us to ascribe probability to every sentence $A \in L_0$. We enrich the language by introducing the new connective $\rightarrow$, satisfying postulates (1)–(4). Consider a set of sentences Φ in this richer language (so Φ is located somewhere in the hierarchy $L_0 \subseteq L_1 \subseteq L_2 \subseteq L_3 \subseteq \ldots$) and two possible ways of expanding the function P, so as to ascribe degrees of belief to sentences from Φ:

(1) By a direct definition of a function Cr: $\Phi \rightarrow [0,1]$. In this case no probability space is constructed and the function Cr ascribes degrees of belief (i.e., real numbers from the interval [0,1]) directly to sentences from the set Φ.

(2) By first constructing a new probability space $S_\Phi = (\Omega_\Phi, \Sigma_\Phi, P_\Phi)$ in which sentences from the set Φ are given interpretations as events. The probability of a sentence α is given as the probability P_Φ of the corresponding event.

As both approaches are present in the literature, the natural question arises as to what their advantages and disadvantages are, and which one should be considered better. Obviously, the problem posed in this way requires the formulation of criteria against which both extensions can be assessed. The following are the criteria which, in our opinion, should be used in the discussion of the pros and cons of the solutions:

[10] Such hierarchies might be defined in different ways (for instance, we can perform the relevant closure operations in one step, so that the step $2n + 1$ and $2n + 2$ "merge together"), but the general idea is similar: we arrive at the next level by some closure operations.

[11] If we consider such conditionals, we need to be careful when defining their credences and also the interpretations of such sentences in probability spaces. We also might exclude such conditionals from our considerations. If we decide to do so, we need to make some change in the definition of the hierarchy of the languages.

(a) how a given solution incorporates the semantic postulates imposed on the conditional connective →;
(b) what the ontological commitments involved in adopting a given solution are;
(c) how the solution allows one to deal with more complex cases (in particular with the higher levels of the hierarchy $L_0 \subseteq L_1 \subseteq L_2 \subseteq L_3 \subseteq \ldots$).

3. The Credence-Like Approach

The credence-like approach consists—generally speaking—in defining a credence function Cr: $\Phi \rightarrow [0,1]$, which assigns credence $Cr(\alpha)$ to sentences $\alpha \in \Phi$. The essential feature of this approach is that no probability space is defined. $Cr(\alpha)$ is assigned directly to sentences from the set Φ and it represents the epistemic attitude of the agent towards α. This does not mean that the agent thinks of a "proportion of the number of circumstances which make α true", as this notion does not even come up in this approach. We can bracket the problem of the truth conditions of conditionals; in fact, we can even deny that they have truth conditions altogether.[12]

This approach does not involve any additional ontological commitments, as compared to, for example, the probability-space approach. Here, we do not assume the existence of extra-linguistic objects, such as truthmakers, regardless of their essence. [13]

The rational agent assigns degrees of belief to sentences using the function Cr: $\Phi \rightarrow [0,1]$, and this is done without assuming any extralinguistic formal structures. It is sufficient to impose some formal conditions on the credence function, treating this as its axiomatic definition. It may be likened to the purely syntactic approach to logical investigations.[14] So both assigning degrees of belief

---

[12] The "credence-like" approach seems natural when we make intuitive assessments even without knowing what the underlying mathematical structure is. Indeed, the subjective probability of *rolling a six, and then a tossed coin coming up heads, and then again heads—when rolling a fair die, and subsequently tossing a fair coin twice* equals 1/24. We simply multiply the probabilities and do this without knowing what the proper probabilistic model is (which, in this case is a product space containing 24 equiprobable elementary events, the aforementioned sequence being one of them).

[13] Undoubtedly, there are mathematical notions involved, prompting an intriguing inquiry into whether the use of mathematical tools in the analysis of philosophical issues entails ontological commitments, in the manner of the famous Quine-Putnam indispensability argument. The philosophy of mathematics is currently witnessing an intense ongoing discussion, and providing even a brief overview is challenging. Here, the focus is not on these "theoretical environment" commitments but rather on the "object-level commitments": if we directly refer to truth-conditions conceived as objects, we need to acknowledge their existence.

[14] The situation is analogous to the characterization of logical systems either in the syntactic mode (by giving axioms and rules of inference) or by presenting the semantics, i.e., the set of structures in which the language is interpreted. In the case of classical logic,

to conditionals and even some calculations are made without referring to the probability space in question: they have an intuitive character and rely on assumptions the credence function should satisfy.[15] On this approach, it is natural to formulate two types of conditions the function Cr should satisfy: both general (concerning any credence function) and related to the specific features of the conditional connective $\rightarrow$.

The general conditions guarantee that Cr is an extension of the initial probability function P and that it satisfies a version of Kolmogorov's axioms (so that the agent using it will not be exposed to a Dutch Book-type argument).[16]

(i) $Cr(\neg\alpha) = 1 - Cr(\alpha)$.

(ii) $Cr(T) = 1$, if T is a tautology.

(iii) $Cr(\alpha \vee \beta) = Cr(\alpha) + Cr(\beta)$ if $Cr(\alpha \wedge \beta) = 0$.[17]

(iv) $Cr(A) = P(A)$ for every $A \in L_0$.[18]

The specific conditions refer to how Cr should take into account the interpretation of the conditional connective $\rightarrow$. Here we use the example of postulates (1)–(4), which imply the following conditions concerning credence:

(Cr-1) $Cr(A \wedge B) \leq Cr(A \rightarrow B)$.

(Cr-2) $Cr((A \wedge \neg B) \wedge (A \rightarrow B)) = 0$.

(Cr-3) $Cr(A \rightarrow \neg B) = 1 - Cr(A \rightarrow B)$.

(Cr-4) $Cr(\neg A \wedge (A \rightarrow B)) = Cr(\neg A) \cdot Cr(A \rightarrow B)$.

---

they coincide, but in general this is an interesting problem of the relationship between these approaches.

[15] For instance, the arguments of Edgington, Lance, McDermott, Cantwell, et. al. presented in Section 4 and 5 do not refer to any probability space. However, they do discuss the probability of complex conditionals.

[16] In general, if the system of beliefs of an agent violated the rules of probability, it would be possible to construct a Dutch Book against it. A Dutch Book is—generally speaking—a system of bets with the property that each single bet is considered by the agent to be fair. However, accepting the whole system leads inevitably to the agent's loss—and in this way reveals the incoherence of the agents' views. There is an interesting discussion concerning this type of argumentation going on, but it exceeds the scope of the present study to present it. See, for instance, Hajek (2009) or Vineberg (2016) for general presentation.

[17] We might also formulate it as a stronger claim:

$$Cr(\alpha \vee \beta) = Cr(\alpha) + Cr(\beta) - Cr(\alpha \wedge \beta)$$

for all α, β.

[18] We do not consider infinite conjunctions or disjunctions. This means that finite additivity is sufficient for our purposes, and we do not need σ-additivity when defining the credence directly on the language.

It is worth noting that the very formulation of postulates (Cr-2) and (Cr-4) requires that Cr is defined on at least some formulas from the language L_2.

4. PCCP: How Should Probabilities Be Evaluated?

The simplest case is when we want to add just one single conditional $A \rightarrow B$ to our system of beliefs, so that $\Phi = L_0 \cup \{A \rightarrow B\}$. In this case, $\Phi \subseteq L_1$. We want to extend our probabilistic beliefs from L_0 to Φ, i.e., to assign credence or probability to the conditional $A \rightarrow B$. This assignment should take into account our base knowledge expressed in L_0 and modeled in the probability space S and also the accepted meaning postulates concerning conditionals.

Of course, the postulates (Cr-1)–(Cr-4) themselves here do not give us any concrete numerical values that should be assigned to the degree of belief as to conditionals of the form $A \rightarrow B$. We have, by assumption, the numerical values of the function Cr given only for non-conditionals sentences from L_0: these are the values of the function P from the space S. The question is how to use this knowledge in order to identify the value $Cr(A \rightarrow B)$.

A simple solution has been proposed by Adams (1965; 1970; 1975; 1998), who defined the notion of the probability of $A \rightarrow B$ by setting $P(A \rightarrow B) \coloneqq P(B|A)$, i.e., the standard conditional probability from the probability space S, i.e., $P(B|A) = \frac{P(A \cap B)}{P(A)}$, for $P(A) > 0$.[19] Adams' thesis is often referred to as PCCP (which stands for "Probability of Conditionals is Conditional Probability"). Since Adams does not define the probability space in which $A \rightarrow B$ is interpreted (i.e., to which the hypothetical event corresponding to the sentence $A \rightarrow B$ belongs), in our terminology Adams' definition should be formulated rather as $Cr(A \rightarrow B) = P(B|A)$.[20] Many examples confirm that this is very natural. Indeed, consider our toy example: *If it is Even, then it is a Six*. Intuitively, its probability is 1/3, which is the conditional probability of rolling a 6, given that an even number was rolled, which we symbolize as $P(\textit{It is a Six} \mid \textit{It is Even})$.[21]

Because Adams defined the numerical value of $Cr(A \rightarrow B)$ as $P(B|A)$, PCCP might be considered to be an analytic claim, true by definition. But obviously, it

---

[19] Intuitively, this is the proportion of B-objects/events within the class of A-objects/events.

[20] In the discussion, the term "probability" is used very often, regardless of the formal details. Adams' original formulation is in terms of assertability. We might also consider terms like "reliability", "credibility", or "acceptability" (and others). However, we will not discuss this issue and generally use the term "degree of belief" when informal judgments are in question, and "credence" and "probability" when thinking about the formal models.

[21] "What is the probability that I throw a six if I throw an even number, if not the probability that if I throw an even number, it will be a six?" (van Fraassen, 1976, p. 273).

is also a substantial claim concerning our pretheoretic assignment of degrees of belief to conditional sentences.[22]

An important feature of Adams' approach is that his definition works for simple conditionals only, i.e., for $A \rightarrow B$, when A, B are sentences from the base language, not containing the conditional connective $\rightarrow$. In doing this, we are not confronted with the conceptual and technical problem of constructing an appropriate probability space. However, there is also a price to pay. Regardless of whether we accept Adams' proposal as the appropriate solution of the problem of simple conditionals, it is clear that it cannot be applied to more complex propositions. Adams even declared that "we should regard the inapplicability of probability to compounds of conditionals as a fundamental limitation of probability, on a par with the inapplicability of truth to simple conditionals" (Adams, 1975, p. 35). Indeed, even if we agree that the probability of *If it is Even, then it is a Six* is 1/3, and similarly *If it is a Prime, it is a Three* is 1/3 (prime numbers are 2, 3, 5), it is not intuitively obvious what the probability of *If it is Even, then it is a Six and if it is a Prime, it is a Three* should be. Is it 1/9? Or 1/3? Or perhaps 0, as a six and a three cannot both occur? And in more complex cases (for instance, when we have nested conditionals), the situation is even more problematic. Even if we could make intuitive judgments in simple cases, this would rather be an ad hoc procedure. It is risky to use the notion of credence in a purely intuitive fashion, without having any idea of how the appropriate mathematical model looks. Making intuitive judgments is often fraught with various conceptual traps, difficulties, and even paradoxes: we are not good intuitive statisticians, as many empirical results show.

Adams' solution is therefore of very limited use—it is only effective if the set Φ contains only simple conditionals of the form $A \rightarrow B$. In other words, the Cr function is defined only on formulas from language L_1. In particular, this means that with Adams' approach the conditions (Cr-2) and (Cr-4) are not even expressible. This can hardly be considered a satisfactory solution.

---

[22] When discussing PCCP, it is impossible not to mention Stalnaker's contribution (e.g., Stalnaker, 1968). Indeed, PCCP is also known as "Stalnaker's Thesis". There are differences between their original versions, for instance Stalnaker speaks of conditional degrees of belief, while Adams originally formulated his claims in terms of assertability. However, our aim is to discuss the approaches to formalization, and not the details of the formulations.

There is an intense debate on the plausibility of PCCP and its diverse variants and modifications (for instance, Bennett, 2003; Edgington, 1995; 2020; Hájek, 2011; 2012; Khoo, 2016; Khoo, Santorio, 2018; Rehder, 1982; Stalnaker, 2009; 2019; van Fraassen, 1976—to name just a few). PCCP is valid in McGee's model (1989), in Bernoulli-Stalnaker spaces (Kaufmann, 2004; 2005; 2009; 2015; 2022; van Fraassen, 1976), in the Markov graph model (Wójtowicz, Wójtowicz, 2021; 2022), in the model of Węgrecki and Wroński (2023), and in the minimal model (Wójtowicz, Wójtowicz, 2023).

Triviality results purport to show that PCCP is not reliable as a general rule, the seminal paper being Lewis (1976; 1986). See, for instance, Hajek (2011) for discussion and generalization, and Khoo and Mandelkern (2019) for discussion concerning linguistic practice.

A further step was taken by McGee, who formulated several axioms concerning the credence distribution.[23] One of them is the independence principle, which—given mild assumptions—allows one to prove PCCP, i.e.:

$$Cr(A \rightarrow B) = P(B|A).$$[24]

In McGee's (1989) we also find the following formula for conjoined conditionals:

$$Cr((A \rightarrow B) \wedge (C \rightarrow D)) = \frac{[P(ABCD) + P(A^cCD)P(B|A) + P(ABC^c)P(D|C)]}{P(A \vee C)}$$

In order to justify this formula, McGee presents a very interesting argument in terms of fair-bet analysis.[25] PCCP is important for McGee's reasoning, as in the fair-bet analysis it is assumed that the credence of $(A \rightarrow B)$ and $(C \rightarrow D)$ is $P(B|A)$ and $P(D|C)$.

However, McGee's formulas are not universally agreed on. The following example has been given by McDermott: "[i]f it is odd it will be below three, and if it is even it will be above three" (1996, p. 26).

If we formalize it, it has the form:

$$(Odd \rightarrow Below\ Three) \wedge (Even \rightarrow Above\ Three)$$

According to McDermott, its meaning and when it is true is intuitively clear: it is true precisely when we see a 1, 4, or 6. So the probability of this sentence being true is 1/2. But, according to McGee's formula, the probability is 2/9.[26]

---

[23] We use the term "credence" in spite of the fact that McGee consistently uses the term "probability". All his arguments concerning conjoined conditionals are formulated in terms of fair bets, not in terms of events in a probability space. So in our presentation of McGee's views, we take some stylistic license in order to maintain coherence with the terminology adopted in this paper.

[24] The simplest form of the Independence Principle is:

$$Cr(C \wedge (A \rightarrow B)) = P(C) \cdot Cr(A \rightarrow B)$$

for A and C being mutually exclusive. McGee accepts the more general form:

$$Cr(C \wedge (A_1 \rightarrow B_1) \wedge (A_2 \rightarrow B_2) \wedge \ldots \wedge (A_n \rightarrow B_n)) = P(C) \cdot Cr((A_1 \rightarrow B_1) \wedge (A_2 \rightarrow B_2) \wedge \ldots \wedge (A_n \rightarrow B_n)),$$

where A_i, B_i and C are Boolean sentences and C excludes A_i, for $i = 1, \ldots, n$.

[25] McGee's fair-bet argumentation concerning $(A \rightarrow B) \wedge (C \rightarrow D)$ does not directly involve any interpretation of $(A \rightarrow B) \wedge (C \rightarrow D)$ as an event in a probability space. This formula coincides with the results obtained with the aid of Stalnaker-Bernoulli spaces (results of Kaufmann, 2004; 2005; 2009; 2015; 2022; van Fraassen, 1976).

[26] Another example is given by Edgington (1991, p. 202). Consider an ordinary fair coin and a claim of the form: *If it is first tossed at t_0, it will land heads, and if it is first tossed at t_1, it will land heads*, i.e., after formalization—$(T_0 \rightarrow H) \wedge (T_1 \rightarrow H)$. According to McGee's formula, the probability of this sentence is 0.25. But according to Edgington,

Neither of the authors argues directly in terms of a probability space. McGee assigns the value 2/9 using fair bet analysis. McDermott (and not only him, see Footnote 26) is convinced that it is obvious—that this probability is 1/2. So it turns out that relying on intuition alone for complex sentences involving conditionals can yield different results.

5. The Case of Complex Conditionals

An additional problem arises when we try want to extend the *Cr* function to certain complex conditionals, i.e., to right-nested conditionals $A \rightarrow (B \rightarrow C)$ and left-nested conditionals $(A \rightarrow B) \rightarrow C$, which—in our hierarchy—appear at level L_3.

Let us see how McGee deals with them. It does not follow from his assumptions that (some form of) PCCP applies to these types of formulas as well. Its (hypothetical) form would be:

$$Cr(A \rightarrow (B \rightarrow C)) = Cr(A \wedge (B \rightarrow C)) \,/\, Cr(A)$$

McGee adopts a different solution—he assumes that the following equivalence is always true:

(EI) $(A \rightarrow (B \rightarrow C)) \Leftrightarrow ((A \wedge B) \rightarrow C)$,[27]

---

the probability of this sentence is 1/2. The rationale behind this stipulation is that the coin can be tossed for the first time exactly once—and obviously in that case (i.e., when the coin is tossed) the probability of its landing heads is ½ (see also Lance, 1991 for a colorful werewolf example).

[27] It is interesting to observe how this principle is justified by McGee:

> It appears to be a fact of English usage, confirmed by numerous examples, that we assert, deny, or profess ignorance of a compound conditional $B \rightarrow (A \rightarrow \varphi)$ under precisely the circumstances under which we assert, deny, or profess ignorance of $(B \wedge A) \rightarrow \varphi$. The assertability conditions for "If you are asked to submit to the 'voluntary' urine test, then if you refuse, you will be under suspicion" and "If you are asked to submit to the 'voluntary' urine test and you refuse, you will be under suspicion" are the same. The best explanation for the fact that $B \rightarrow (A \rightarrow \varphi)$ and $(B \,\&\, A) \rightarrow \varphi$ are equiassertable is that they are believed to the same degree, which is what the Import-Export Principle asserts. (McGee, 1989, pp. 489–490)

Ciardelli (2020) offers the following argument:

> The argument for this desideratum comes from the observation that (8-a) and (8-b) seem to express the same thing, and that such examples can be multiplied without running into counterexamples […].
>
> (8) a If Bob is in Paris, then if he is staying in a hotel, he is at the Ritz.
>
> b If Bob is in Paris and he is staying in a hotel, he is at the Ritz. (Ciardelli, 2020, p. 515)

from which it follows that:

(Cr-EI) $Cr(A \to (B \to C)) = Cr((A \wedge B) \to C)$.

The formula that is the argument of the Cr function on the right side of the equation belongs to language L_2 and we know how to calculate its value (using PCCP). As a result, we are also able to assign credence to the formula on the left side of the equation.

Another example of the argumentation which de facto relies on (Cr-EI) is given in Cantwell's (2022). He considers the sentence "If it is not diamonds ($\neg D$), then if it is red (R) it is hearts (H)" to be obviously true, i.e., its degree of belief is 1. The sentence is $\neg D \to (R \to H)$ and after applying IE becomes $(\neg D \wedge R) \to H$. But $(\neg D \wedge R)$ is equivalent to H, which gives the expected conclusion.

In McGee's (1985), a similar argument was given which was intended to undermine Modus Ponens (MP) when applied to compound conditionals: "If a Republican wins the election, then if it is not Reagan who wins it will be Anderson". According to McGee, this sentence is true regardless of the situation, which means that its degree of belief is 1.

It is not our aim to take a position on the validity of (Cr-EI). There is a lot of discussion concerning it in the literature.[28] But even if we adopt it, it does not solve other problems. For instance, it is not clear how to handle left-nested conditionals $(A \to B) \to C$. Indeed, they are very difficult to analyze in a purely intuitive fashion and it is far from clear what a general formula for the credence of $(A \to B) \to C$ would look like.

McGee deals with this problem decisively. He simply assumes that formulas of this type do not occur in the language. Consequently, instead of the hierarchy of languages $L_0 \subseteq L_1 \subseteq L_2 \subseteq L_3\ldots$ we have been considering, McGee limits himself only to a fragment of this hierarchy, allowing only right-nestings of the form $(A \to \alpha)$.[29]

So Ciardelli's justification has an inductive character. This limits its power and makes is vulnerable to potential counterarguments (by counterexamples).

[28] Take the wet match example (Lewis, 1973; Stalnaker, 1968). We are willing to consider the claim (a) "If the match is wet, then it will light if you strike it" to be equivalent with (b) "If the match is wet and you strike it, then it will light".

An interesting analysis has been given in probabilistic terms (Kaufmann, 2005, p. 206): (a) of the match getting wet = 0.1; (b) that you strike it = 0.5; (c) that it lights given that you strike it and it is dry = 0.9; (d) that it lights given that you strike it and it is wet = 0.1. Moreover, striking the match is independent of its wetness. We expect the probability of the conditional "If the match is wet, then it will light if you strike it" to be 0.1—i.e., to obey the Import-Export principle in this case.

[29] McGee puts it this way:

> Let me adopt a restriction at the outset. I only want to look at conditionals whose antecedents do not contain conditionals. Conditionals with conditional antecedents are used and understood by English speakers, but they occur sufficiently rarely

6. The Probability-Space Approach

This approach has a quite different character. We construct a probability space $S_\Phi = (\Omega_\Phi, \Sigma_\Phi, P_\Phi)$ in which sentences from the set Φ are interpreted as events.[30] This means that every sentence $\alpha \in \Phi$ has an interpretation $[\alpha]_\Phi \subseteq \Omega_\Phi$ as an event in the space $S_\Phi = (\Omega_\Phi, \Sigma_\Phi, P_\Phi)$ and has a well-defined probability $P_\Phi([\alpha]_\Phi)$.[31] The degree of belief of the sentence α is therefore defined as the probability of its counterpart $[\alpha]_\Phi$ in the space S_Φ.

What does space S_Φ look like? In general, the sample space $S = (\Omega, \Sigma, P)$ is designed to model only non-conditional claims—and is not suited to give interpretation to conditionals. In the case of the die, the obvious probability space $S = (\Omega, \Sigma, P)$ has six elementary events, i.e., $\Omega = \{1, 2, 3, 4, 5, 6\}$. But this means that the conditional $Even \rightarrow Six$ has no interpretation as an event within S.[32] We need something else.

In order to construct the probability space S_Φ, we obviously have to determine the set of elementary events Ω_Φ. Every elementary event $\omega \in \Omega_\Phi$ has to decide—for any sentence $\alpha \in \Phi$—whether it supports α or not. In other words, when specifying the suitable Ω_Φ, we have to define the relation "$\omega \vDash \alpha$". (Later we give explicit examples of such constructions, in particular concrete definitions of the relation "$\omega \vDash \alpha$"). Only then will we be able to identify the semantic correlate of α as the set $[\alpha]_\Phi = \{\omega \in \Omega_\Phi : \omega \vDash \alpha\}$. This means that if we want to

---

that it is hard to gather enough examples to get a fix on what is going on with them. (McGee, 1989, p. 486)

[30] In general, there are many spaces of this kind, so we do not really think about "the S_Φ-space", but rather about "a S_Φ-space", so—in a sense—S_Φ is a metatheoretic variable, referring to mathematical objects of a certain type.

[31] Formally, this means that the set $[\alpha]_\Phi$ is measurable, i.e., it is an element of the σ-field Σ_Φ.

[32] Indeed, none of the 64 subsets of $\{1, 2, 3, 4, 5, 6\}$ seem appropriate. Should 1 belong to the interpretation of $Even \rightarrow Six$—i.e., does 1 make this conditional true? That does not seem reasonable, but the claim that 1 makes $Even \rightarrow Six$ false is not reasonable either. The results of a single die roll do not decide the conditional, and we need a different structure. Similarly, the conditional "If it is not One, it is Two" has—intuitively—a probability of 1/5. But there is no event with a probability of 1/5 in the sample space. The following example is even more striking: consider a simple space consisting of three elementary events, i.e., $\Omega = \{X, Y, Z\}$, and take $P(X) = P(Y) = P(Z) = 1/3$. We might think of three balls in an urn, numbered 1, 2, 3. Intuitively, the probability of "If it is odd, it is three" is 1/2. However, the probability space $S = (\Omega, \Sigma, P)$ contains only events with the probabilities 0, 1/3, 2/3, and 1, so it is not suited to give an interpretation for "If it is odd, it is three".

This is not a coincidence: Hajek (1989) shows that *any* non-trivial finite-ranged probability function has more distinct conditional probability values than distinct unconditional probability values. But if PCCP holds, the space needs to give interpretations to all the events with the conditional probability values. So the original probability space is not the right one: there are not enough events to model all the conditionals.

model the conditional α as an event in a probability space, we need to accept some notion of the circumstances in which α is true and in which it is false. Given this, we can say that the probability P_Φ of a conditional is the probability of its truth in the space S_Φ. This is a profound difference when compared with the credence approach.

There are therefore some minimal assumptions that S_Φ must satisfy:

(I) For every $\alpha \in \Phi$, there is an event $[\alpha]_\Phi = \{\omega \in \Omega_\Phi\colon \omega \vDash \alpha\}$.

(II) There is a homomorphic imbedding of S into S_Φ.

The first condition ensures that S_Φ indeed models Φ, i.e., every sentence $\alpha \in \Phi$ has a semantic counterpart within S_Φ. The second condition ensures that it is an appropriate extension of S. Intuitively, an imbedding might be imagined as presenting a copy of the space $S = (\Omega, \Sigma, P)$ within $S_\Phi = (\Omega_\Phi, \Sigma_\Phi, P_\Phi)$, which preserves the essential features of S.

Formally, a homomorphic imbedding of $S = (\Omega, \Sigma, P)$ into $S_\Phi = (\Omega_\Phi, \Sigma_\Phi, P_\Phi)$ is a function $\imath\colon \Omega \to \Omega_\Phi$ satisfying the following conditions (by abuse of language we denote the image of the set $X \subseteq \Omega$ by $\imath(X)$):

a. $\imath(\Omega) = \Omega_\Phi$;

b. $\imath(X \cap Y) = \imath(X) \cap \imath(Y)$, for $X, Y \subseteq \Omega$;

c. $\imath(X^c) = (\imath(X))^c$, for $X \subseteq \Omega$;

d. $P_\Phi(\imath(X)) = P(X)$, for $X \subseteq \Omega$;

e. $[A]_\Phi = \imath([A])$ (we can say that the imbedding $\imath$ is faithful to the interpretations of the factual sentences).

Of course, sentences from L_0 are also interpreted in S_Φ, as $L_0 \subseteq \Phi$. Condition (e) ensures that the direct interpretation of A within S_Φ is compatible with the two-step procedure: (i) interpreting A in S, and then (ii) imbedding S in S_Φ via $\imath$. From the definition it follows that if A, B are factual sentences then $[A \wedge B]_\Phi = [A]_\Phi \cap [B]_\Phi$ and $[\neg A]_\Phi = \Omega \setminus [A]_\Phi$. So $S_\Phi = (\Omega_\Phi, \Sigma_\Phi, P_\Phi)$ preserves the structure of interpretation of L_0 within S. Importantly, the initial probabilities of factual sentences $A \in L_0$ are preserved in S_Φ, i.e., $P_\Phi([A]) = P([A])$. If the space S_Φ satisfies (I) and (II), this means—by definition—that:

(i) $P_\Phi(\neg\alpha) = 1 - P_\Phi(\alpha)$;
(ii) $P_\Phi(T) = 1$, if T is a tautology;
(iii) $P_\Phi(\alpha \vee \beta) = P_\Phi(\alpha) + P_\Phi(\beta)$ if $[\alpha]_\Phi \cap [\beta]_\Phi = \varnothing$;[33]
(iv) $P_\Phi([A]_\Phi) = P(A)$ for every $A \in L_0$.

So we know that P_Φ is suitable for describing the degrees of belief of a rational (Dutch-Book resistant) agent whose initial degrees of belief are described by the P-function.

The probabilistic counterparts of the aforementioned principles (1)–(4) are mostly just simple corollaries of the claims concerning events, like: if $[\alpha]_\Phi \subseteq [\beta]_\Phi$, then $P_\Phi([\alpha]_\Phi) \leq P_\Phi([\beta]_\Phi)$:

(P_Φ-1) $Cr(A \rightarrow (B \rightarrow C)) = Cr((A \wedge B) \rightarrow C)$;
(P_Φ-2) $P_\Phi([(A \wedge \neg B) \wedge (A \rightarrow B)]_\Phi) = 0$;
(P_Φ-3) $P_\Phi([A \rightarrow \neg B]_\Phi) = 1 - P_\Phi([A \rightarrow B]_\Phi)$;
(P_Φ-4) $P_\Phi([A \rightarrow B]_\Phi | [\neg A]_\Phi) = P_\Phi([A \rightarrow B]_\Phi)$.

(P_Φ-4) can also be written as:

$$P_\Phi([\neg A]_\Phi \cap [A \rightarrow B]_\Phi) = P_\Phi([\neg A]_\Phi) \cdot P_\Phi([A \rightarrow B]_\Phi,$$

or

$$P_\Phi([\neg A \wedge (A \rightarrow B)]_\Phi) = P_\Phi([\neg A]_\Phi) \cdot P_\Phi([A \rightarrow B]_\Phi.$$

The important question is whether it is possible to define a mathematical structure satisfying the requirements imposed on $S_\Phi = (\Omega_\Phi, \Sigma_\Phi, P_\Phi)$ given above. The answer is positive, as documented by many constructions present in the literature. The most classic is Stalnaker Bernoulli spaces (Kaufmann, 2004; 2005; 2009; 2015; van Fraassen, 1976). Wójtowicz and Wójtowicz (2021; 2022) present a construction based on the theory of Markov chains: for a given conditional α, a Markov chain (graph) $G(\alpha)$ is defined which gives rise to a probability space $S(\alpha)$. This space gives a natural interpretation for α as an event $[\alpha]$, and its probability is computed very simply.[34] Constructions of probability spaces are also given in Węgrecki and Wroński's (2023) and Wójtowicz and Wójtowicz's (2023).

The mentioned spaces differ in terms of complexity. For instance, elementary events in the Stalnaker Bernoulli space are infinite sequences of possible worlds.

---

[33] We do not need σ-additivity when defining the credence directly on a language, as we do not consider infinite conjunctions or disjunctions, so finite additivity is sufficient for our purposes.

[34] Computations are much simpler than in the Stalnaker-Bernoulli space, as they consist in solving simple systems of linear equations. However, the construction of $S(\alpha)$ is limited to the particular conditional α in question.

This space has the cardinality of the continuum, so it is rather big. The probability measure is defined on a kind of cylindric subset of the set of all sequences.[35] The modification in Bacon (2015) is even more complex: we have transfinite sequences of possible worlds (of length ω_1, i.e., the first uncountable ordinal number) and the probability space is defined in a very interesting, but also mathematically quite complex way.[36] The spaces $S(\alpha)$ generated by the Markov graphs (Wójtowicz, Wójtowicz, 2021; 2022) are countably infinite, with a very simple structure. The probability space in Węgrecki and Wroński (2023) is finite. The permutation model in Wójtowicz and Wójtowicz's (2023) is proven to have a minimal size (in a certain, natural class of models). The Węgrecki-Wroński model employs a form of formal expressions—broadly speaking—over possible worlds. Both models satisfy van Fraassen's conditions—so both provide a solution to the problem of giving partial, "small" models of conditionals.

We will illustrate the general idea of defining the semantic relation "$\omega \vDash \alpha$" in the constructed space $S_\Phi = (\Omega_\Phi, \Sigma_\Phi, P_\Phi)$—and here we use the permutation model from Wójtowicz and Wójtowicz's (2023) as an example.[37] As usual, we start with a probability space S in which the factual language is interpreted. In the simplest nontrivial case the initial space $S = (\Omega, \Sigma, P)$ contains three events, so that $\Omega = \{1, 2, 3\}$. All these events have the same probability, i.e., $P(1) = P(2) = P(3) = 1/3$. It is convenient to identify the subset of Ω with meaningful propositions in the language. This means that we consider propositions like "It is three", "It is a one", "It is either a one or a two", and so on.

The problem to solve is to construct a space $S_\Phi = (\Omega_\Phi, \Sigma_\Phi, P_\Phi)$ in which all simple conditionals and their Boolean combinations have interpretations, so that the semantic relation "$\omega \vDash \alpha$" can be formally defined. This means that we need to interpret sentences like:

$$Odd \to One \text{ (i.e., if it is odd, it is a one)}$$
$$\neg Two \to Three$$
$$(One \vee Three) \to One$$

The corresponding probability space $S_\Phi = (\Omega_\Phi, \Sigma_\Phi, P_\Phi)$ is defined as follows:

(i) Ω_Φ consists of all permutations of set 1,2,3, i.e., $\Omega_\Phi = \{123, 132, 213, 231, 312, 321\}$;

(ii) The σ-field Σ_Φ consists of all subsets of Ω_Φ;

(iii) The probability P_Φ of each of these permutations is set to be 1/6.

[35] Hall (1994) shows that spaces which allow for modeling all conditionals (and satisfy some minimal assumptions, in particular PCCP) must contain events of any probability $p \in [0,1]$.

[36] In fact, transfinite induction up to ω_1 is needed.

[37] We would like to thank the anonymous referee for emphasizing the necessity of providing a concrete example of such constructions here.

In order to define the semantic relation "$\omega \vDash \alpha$" in the space of permutations, a very simple rule is used (X, Y, Z are factual propositions):

$$\omega \vDash X \to Y \text{ iff the first element in the permutation } \omega \text{ which is an } X \text{ is also an } Y.$$
$$\omega \vDash Z \text{ iff the permutation } \omega \text{ begins with } Z.$$

For instance:

$123 \vDash One$ (indeed, it begins with 1),
$123 \vDash Odd \to One$ (the first odd number appearing in 123 is 1),
$213 \vDash Odd \to One$ (the first odd number appearing in 213 is 1),
$231 \vDash \neg Two \to Three$ (the first number appearing in 231 which is not 2 is 3),
$231 \vDash Two \wedge (\neg Two \to Three)$.

If the initial probability space contains n events, we also have the space of all permutations (it has $n!$ elements) with the same heuristic rule.

This demonstrates that this space satisfies van Fraassen's conditions mentioned earlier. The construction can be iterated, so that conditionals of every level of complexity can be interpreted there. Of course, the spaces become more complex. For instance, if we take our "toy permutation space" consisting of the six permutations 123, 132, 213, 231, 312, 321, and perform the next step (in order to be able to interpret nested conditionals, like $Two \to (Odd \to One)$), the "next level space" consists of all the permutations of permutations—and there are $6! = 720$ such objects. For instance, the permutation of the permutations (123)(321)(312)(213)(231)(132) is one of the next-level objects.[38]

How do these spaces deal with PCCP and the interpretation of complex conditionals?[39] Stalnaker-Bernoulli space allows one to give an interpretation of all conditionals (regardless of their complexity) and enables one to extended the degrees of belief to any set Φ. For any sentence α, P_Φ can be calculated in a mathematically sound and unambiguous way. PCCP is satisfied, but (EI) is not. However, it is possible to preserve intuitions which in some cases recommend accepting the equivalence of $A \to (B \to C)$ and $(A \wedge B) \to C$. It is suffi-

[38] If we define the permutation structures in this way, certain "built-in" principles emerge: for instance, the Import-Export principle for conditionals is not universally valid. These spaces might be modified in different ways, but a comprehensive discussion of the technical details exceeds the scope of this study.

[39] PCCP can be formally formulated as $P_\Phi([A \to B]_\Phi) = P_\Phi([B]_\Phi | [A]_\Phi)$—here we have probabilities of events in space S_Φ. (For simple conditionals, as the original space S is imbedded in S_Φ and probabilities are preserved, it is equivalent to: $P_\Phi([A \to B]_\Phi) = P([B] | [A])$). For complex conditionals of the form $\alpha \to \beta$ which have interpretations in S_Φ, the scheme is similar: $P_\Phi([\alpha \to \beta]_\Phi) = P_\Phi([\beta]_\Phi | [\alpha]_\Phi)$. Whether PCCP holds or not in S_Φ is a mathematical problem, not a question of intuitive judgment.

cient to distinguish between those situations in which $A \rightarrow (B \rightarrow C)$ is reducible to $(A \wedge B) \rightarrow C$ and those in which is not, and then to make an internal translation of the language: substituting $(A \wedge B) \rightarrow C)$ for $A \rightarrow (B \rightarrow C)$ where we consider it reasonable.

In probabilistic spaces generated by Markov graphs and permutation spaces, we obtain similar results,[40] but due to the inductive nature of the construction, we can match the size of the model to what sentences appear in the set Φ. The space S_{Φ} in which we calculate all the probabilities which we are interested in depends on the set Φ and is minimal. It can be said that here we achieve an optimal proportion between ontological commitments and model efficiency. In the Appendix we present another probabilistic model designed to manage complex conditionals in a relatively straightforward manner.

7. A Brief Comparison and Conclusion

In the paper we have presented and compared two approaches: one "credence-like" and the other "via probability space". There are some common methodological prerequisites. In both cases, we assume that we have an extension of the non-conditional system of beliefs. This means that the probability assignments from the base space S are preserved. In both cases, we need to incorporate some general postulates concerning conditionals, which put some restrictions on the structures. Of course, the "implementation" of these postulates is different in both cases. However, these approaches differ in character and rest on quite different philosophical assumptions.

The first is attractive to those suspicious of the notion of a conditional being true under some circumstances. The *Cr* function operates directly on sentences, without referring to the existence of any extralinguistic objects. The undoubted disadvantage of this approach is that it is very sensitive to the complexity of the language in which the set Φ is contained. Adams's approach might be considered extreme: the scope is limited to formulas from L_1, as degrees of belief can only be attributed to simple conditionals. The more complex approaches (e.g., Cantwell, 2022; Edgington, 1991; McDermott, 1996; Mc Gee, 1989) extend the function's operation also to formulas from L_2 or even L_3.[41] However, there is no general agreement on what the value of this function actually is, as is documented in the discussions in the literature.

In the case of the probability-space approach, we need to accept the claim that there are entities considered to be semantic correlates of conditionals. Con-

---

[40] PCCP is satisfied and (EI) can be modeled to include only reasonable cases.

[41] In the case of McGee's system (1989), the situation is more complex, as he allows also more complex right-nestings. The details are not relevant here.

structing such spaces involves in particular defining the set of elementary events Ω_Φ and specifying the relation $\omega \vDash \alpha$, for $\omega \in \Omega_\Phi$ and $\alpha \in \Phi$.[42]

Introducing Ω_Φ comes at a price: some ontological commitments must be accepted. Of course—wherever possible—we want to keep them to a minimal level But once we accept this cost, the probability-space approach raises interesting ontological questions concerning the nature and ontological status of the introduced entities. That being said, great methodological advantages. It offers not only a method of formalizing the notion of degree of belief, but also offers a semantic framework. Last—but not least—once we define the probability space $S_\Phi = (\Omega_\Phi, \Sigma_\Phi, P_\Phi)$ in a mathematically rigorous fashion, we have all the methods and theorems of probability theory at our disposal. We can use them to formally analyze conditionals, as a source of inspiration for constructing formal models, and as a methodological framework in which our considerations can be conducted rigorously.[43]

Finally, consider the three criteria formulated in Section 3.

(a) How a given solution incorporates the semantic postulates imposed on the conditional connective $\rightarrow$.

This is possible in both cases—see (Cr-1)–(Cr-4) and (P_Φ-1)–(P_Φ-4).

(b) What the ontological commitments involved in adopting a given solution are.

Obviously, the probability space approach brings in certain ontological commitments: it is necessary to postulate the existence of a structure, where conditionals have semantic correlates. And these commitments might vary as to their complexity. Of course, nothing like this happens in the case of *Cr* functions: no ontological assumptions are necessary, as no extralinguistic objects are postulated.

(c) How the solution allows one to deal with more complex cases (in particular with the higher levels of the hierarchy $L_0 \subseteq L_1 \subseteq L_2 \subseteq L_3 \subseteq \ldots$).

The probability space approach gives many more possibilities here. For instance, it is possible to construct a space in which the principle (EI) is valid only in reasonable cases. In the case of the credence approach, the acceptance of (Cr-EI) is a way of reducing the problem of assignment of *Cr* values to right-nested conditionals. Also, in the probability-space approach, PCCP can be shown to be true not by stipulation, but as a natural result of the way the space is constructed.

---

[42] The core idea is to treat elementary events in the probability space analogously to how we handle possible worlds in truth-conditional semantics. We thank the anonymous referee for helpful comments on these matters.

[43] In the Appendix, we offer a succinct illustration of a model that facilitates the computation of probabilities for more complex cases.

And what seems most important—with the probability space-approach we have no problem computing the probability of an arbitrarily complex conditional. We do not need to impose any rules restricting the class of conditionals (for instance by excluding left-nested conditionals). This is a task that the credence-like approach cannot easily deal with.

All these arguments lead us to consider the probability space approach more fruitful.

APPENDIX

This is a brief overview designed to offer readers a general understanding of the Markov graph model. We focus on conveying the basic idea without delving into technical details and into discussing the philosophical issues involved (for both, we refer readers to the details in Wójtowicz, Wójtowicz, 2021; 2022). It is convenient to view the presented graphs as computational devices, which enable one to compute the probabilities of conditionals in a straightforward manner.

The general idea of the Markov graph model is to identify the conditional with a game—and the probability of the conditional with the probability of winning this game. Consider the illustrative example *Even* → *Six* and envision rolling a fair die. When betting on *Even* → *Six*, it is not controversial to agree that:

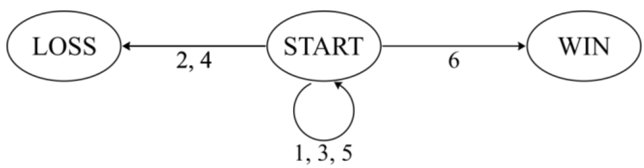
- If a 6 comes up— we win;
- If a 2 or 4 comes up—we lose;
- If a 1, 3, or 5 comes up—the game is undecided.

In the last case, we continue the game, meaning we continue to roll the die until an even number is obtained to determine the outcome. It is conceivable that we might need to roll the die 100 times before an even number appears. (Obviously, in the case of a fair die this is highly unlikely to happen).

The dynamics of the game can be represented by a simple graph:

Figure 1

The Markov graph for Even → Six

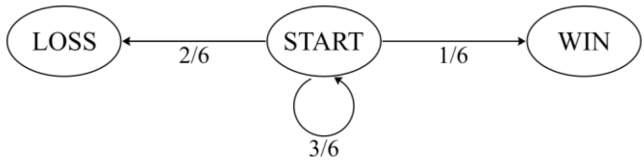


An odd number is "neutral", so the game remains at START. A 6 transfers the game from START to WIN. A 2 or 4 transfers the game from START to LOSS.

We win the game if the process arrives at the state WIN. The following graph depicts the probabilities of the actions in the case of a fair die:

Figure 2

The Markov graph for Even → Six with transition probabilities[44]



The probability of winning the game is the probability that ultimately the process will be absorbed by WIN. Let:

- P_{START} denote the probability of winning the game once you are in the state START;
- P_{WIN} is the probability of winning the game once you are in the state WIN (obviously, it is 1, as you have won already);
- P_{LOSS} is the probability of winning the game once you are in the state LOSS (obviously, it is 0, as you have won already).

The equation that identifies the probability of reaching the state WIN is:

$$P_{\text{START}} = \frac{1}{2} P_{\text{START}} + \frac{1}{6} P_{\text{WIN}} + \frac{2}{6} P_{\text{LOSS}} = \frac{1}{2} P_{\text{START}} + \frac{1}{6}$$

This equation follows from general mathematical theory, but also has an intuitive justification.[45] The result is $P_{\text{START}} = \frac{1}{3}$, as expected.

Intuitively, the graph generates various game scenarios, for instance:

[44] We have slightly simplified the graphs: formally, there should be loops (with a probability of 1) at WIN and LOSS.

[45] You begin the game at START with an initially unknown probability *p* of winning. In this situation, you have the options to (i) loop, i.e., stay at START (with a probability of 1/2) and your chance of winning is still *p*, (ii) advance to WIN (this has a probability of 1/6), which results in a guaranteed chance of winning, i.e., 1, (iii) go to LOSE (this has a probability of 2/6) and then your chance of winning is 0. Combining these cases yields the corresponding equation. A similar intuitive justification can also be given for more complex cases.

- 2 (we lose in the first move);
- 6 (we win in the first move);
- 1356 (we win in the fourth move);
- 55336662 (we lose in the in the eighth move);
  and so on.

The probabilities of these scenarios are obvious, intuitively:

$$P(2) = \frac{1}{6}$$
$$P(1356) = (\frac{1}{2})^3 \cdot \frac{1}{6}$$
$$P(55336662) = (\frac{1}{2})^7 \cdot \frac{1}{6}$$

The general theory of Markov chains assures us that the graph generates a probability space. The elementary events in the space are sequences which start with a series of odd numbers (possibly empty) followed by an even number. It is natural to say that some of these events make the conditional *Even* → *Six* true, while others make it false. This means, that we can straightforwardly define the semantic relation $\omega \vDash Even \rightarrow Six$:

- $\omega \vDash Even \rightarrow Six$ iff the sequence ends with a 6;
- $\omega \vDash \neg(Even \rightarrow Six)$ iff the sequence ends with a 2 or 4.

For factual sentences, like "It is a Five", "It is Even", or "It is not a Prime number", we have the natural stipulation:

- $\omega \vDash X$ iff the sequence starts with an X.

For instance:

- 55336662 ⊨ *It is a Five*;
- 1356 ⊨ *It is not Even*.

It is interesting to consider more complex conditionals; for instance, right-nested conditionals of the form $A \rightarrow (B \rightarrow C)$. Here we briefly present the idea of the construction for $A \rightarrow (B \rightarrow C)$ (the details can be found in Wójtowicz, Wójtowicz, 2022). An important feature of the construction is that it does not satisfy the Import-Export Principle, i.e., the meaning of $A \to (B \to C)$ is—in general—different than the meaning of $(A \wedge B) \to C$.[46]

Consider a colored fair die. Numbers 1, 2, 3 are Red, numbers 4, 5, 6 are Green. Consider the conditional *Green* → (*Even* → *Six*).

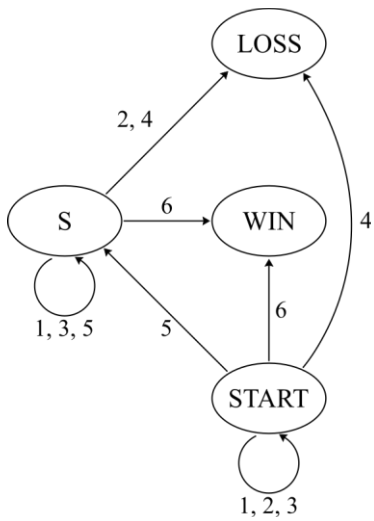
We roll the die and a 2 comes up. It is red, i.e., non-green. This means that the *Green* sentence has not been satisfied, i.e., the antecedent of *Green* → (*Even* → *Six*) is false. This is analogous to the simple conditional (*Even* → *Six*) when an odd number appears: what happens with the successor does not matter. We might think of *Green* as of an activating event, which "opens" the (*Even* → *Six*)-game.

So, after a 2 shows up, we do not lose, and the game is restarted. Assume we see a 6. It is both green and even—we win. If we see a four—we lose. Indeed, it is both green and not a six. If we see a 5, the game is not decided, but the status of the game changes, as the activating event (Green) has already taken place. From now on, we stop paying attention to the colors, and only continue with the (*Even* → *Six*)-game.

This is the graph for the game:

Figure 3

The Markov graph for Green → (*Even* → *Six*)



The state S is the intermediate state: the "Green condition" has already been satisfied, so we continue the game.

The absorption probability is computed by solving a system of equations:

[46] So it will not be accepted by someone who accepts Import-Export as a general principle. However, we can use the model by performing appropriate translations, i.e., $A \to (B \to C)$ is translated to $(A \wedge B) \to C$.

$$P_{\mathrm{START}} = \frac{1}{2}P_{\mathrm{START}} + \frac{1}{6}P_{\mathrm{WIN}} + \frac{1}{6}P_{\mathrm{LOSS}} + \frac{1}{6}P_{\mathrm{S}} = \frac{1}{2}P_{\mathrm{START}} + \frac{1}{6} + \frac{1}{6}P_{\mathrm{S}}$$
$$P_{\mathrm{S}} = \frac{1}{2}P_{\mathrm{S}} + \frac{1}{6}P_{\mathrm{WIN}} = \frac{1}{2}P_{\mathrm{S}} + \frac{1}{6}$$

After solving it, we obtain:

$$P_{\mathrm{START}} = \frac{4}{9}$$

As before, we can define the semantic relation $\omega \vDash Green \rightarrow (Even \rightarrow Six)$. Intuitively, the event ω makes the sentence true iff ω is a scenario that leads from START to WIN. For instance:

- $6 \vDash Green \rightarrow (Even \rightarrow Six)$;
- $121256 \vDash Green \rightarrow (Even \rightarrow Six)$;
- $52 \vDash \neg[Green \rightarrow (Even \rightarrow Six)]$;
- $334 \vDash \neg[Green \rightarrow (Even \rightarrow Six)]$.[47]

In a similar way, we can similarly analyze more complex conditionals, like $A \rightarrow [B \rightarrow (C \rightarrow D)]$. Consider an urn in which we have balls that have three properties:

(i) Color: they are either White, Green, or Red (W, G, R).
(ii) Mass: they are Light or Heavy (L, H).
(iii) Size: they are Big or Small (B, S).

So there are 12 kinds of balls in the urn, for instance: BHG (Big, Heavy, and Green), SHR (Small, Heavy, and Red), and so on.

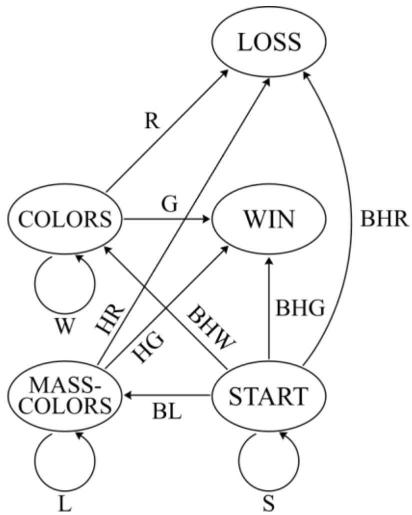
Consider the conditional $\mathrm{B} \rightarrow [\mathrm{H} \rightarrow (\neg\mathrm{W} \rightarrow \mathrm{G})]$. As before, we will consider the events *Big* and *Heavy* as "activating events". For instance:

(i) Drawing a BHG leads to WIN in one move.
(ii) Drawing a BHR leads to LOSS in one move.
(iii) Drawing a BL redirects us to the $\mathrm{H} \rightarrow (\neg\mathrm{W} \rightarrow \mathrm{G})$ game.
(iv) Drawing a BH redirects us directly to the $\neg\mathrm{W} \rightarrow \mathrm{G}$ game.

[47] The graphs are designed to compute probabilities of specific conditionals, and therefore, they may not always offer an interpretation for the entire language (this is not needed if we are interested in a specific conditional). The general construction, which gives interpretations for all conditionals, is more intricate, see Wójtowicz and Wójtowicz's (2024).

The graph is more complex:

Figure 4
The graph for B → [H → (¬W → G)]



START is obvious. The state MASS-COLORS is the state where B has already been activated (but not H!) and we play the (H → (¬W → G))-game. The state COLORS is where both B and H have been activated, and we play the (¬W → G)-game. The respective probabilities can be computed by solving the system of linear equations with three variables: P_{START}, $P_{\text{MASS-COLORS}}$, P_{COLORS}.

$$P_{\text{START}} = P(\text{S}) \cdot P_{\text{START}} + P(\text{BL}) \cdot P_{\text{MASS-COLORS}} + P(\text{BHW}) \cdot P_{\text{COLORS}} + P(\text{BHG})$$
$$P_{\text{MASS-COLORS}} = P(\text{L}) \cdot P_{\text{MASS-COLORS}} + P(\text{HW}) \cdot P_{\text{COLORS}} + P(\text{HG})$$
$$P_{\text{COLORS}} = P(\text{W}) \cdot P_{\text{COLORS}} + P(\text{G})$$

In general, the absorption probability in the Markov graph corresponding to the conditional is computed by solving a system of linear equations, which is a simple task.

REFERENCES

Adams, E. W. (1965). On the Logic of Conditionals. *Inquiry*, *8*, 166–197.
Adams, E. W. (1970). Subjunctive and Indicative Conditionals. *Foundations of Language*, *6*, 89–94.
Adams, E. W. (1975). *The Logic of Conditionals*. Dordrecht: D. Reidel.

Adams E. W. (1998). *A Primer of Probability Logic*. Stanford California: CLSI, Stanford University.

Bacon, A. (2015). Stalnaker's Thesis in Context. *The Review of Symbolic Logic*, *8*(1), 131–163. doi:10.1017/s1755020314000318

Berto, R., Özgün, A. (2021). Indicative Conditionals: Probabilities and Relevance. *Philosophical Studies*, *178*, 3697–3730. doi:10.1007/s11098-021-01622-3

Bennett, J. (2003). *A Philosophical Guide to Conditionals*. Oxford: Clarendon Press.

Cantwell, J. (2022). Revisiting McGee's Probabilistic Analysis of Conditionals. *Journal of Philosophical Logic*, *51*, 973–1017. doi:10.1007/s10992-022-09657-5

Ciardelli, I. (2020). Indicative Conditionals and Graded Information. *Journal of Philosophical Logic*, *49*, 509–549.

Cruz, N., Over, D., Oaksford, M., Baratgin, J. (2016). Centering and the Meaning of Conditionals. In A. Papafragou, D. Grodner, D. Mirman, J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 1104–1109). Austin: Cognitive Science Society.

Edgington, D. (1991). The Mystery of the Missing Matter of Fact. *Proceedings of the Aristotelian Society Supplementary Volume*, *65*, 185–209.

Edgington, D. (1995). On Conditionals. *Mind*, *104*, 235–329. doi:10.1093/mind/104.414.235

Edgington, D. (2020). Indicative Conditionals. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved from: https://plato.stanford.edu/archives/fall2020/entries/conditionals/

Egré, P., Rott, H. (2021). The Logic of Conditionals. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved from: https://plato.stanford.edu/archives/win2021/entries/logic-conditionals/

Evans, J., Over, D. E. (2004). *If*. Oxford: Oxford University Press.

Gibbard, A. (1981). Two Recent Theories of Conditionals. In W. Harper, R. C. Stalnaker, G. Pearce (Eds.), *Ifs: Conditionals, Belief, Decision, Chance and Time* (pp. 211–247). Dordrecht: Reidel.

Gillies, A. (2009). On Truth-Conditions for *If* (But Not Quite Only *If*). *Philosophical Review*, *118*(3), 325–349.

Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge: HUP.

Hájek, A. (1989). Probabilities of Conditionals—Revisited. *Journal of Philosophical Logic*, *18*, 423–428.

Hájek, A. (2009). Dutch Book Arguments. In P. Anand, P. Pattanaik, C. Puppe (Eds.), *The Handbook of Rationality and Social Choice* (pp. 173–195). Oxford: Oxford University Press.

Hajek, A. (2011). Triviality Pursuit. *Topoi*, *30*(1), 3–15.

Hajek A. (2012). The Fall of "Adams' Thesis"? *Journal of Logic, Language and Information*, *21*, 145–161.

Jackson, F. (1987). *Conditionals*. New York: Blackwell.

Kaufmann, S. (2004). Conditioning Against the Grain: Abduction and Indicative Conditionals. *Journal of Philosophical Logic*, *33*(6), 583–606. doi:10.1023/B:LOGI.0000046142.51136.bf

Kaufmann, S. (2005). Conditional Predictions: A Probabilistic Account. *Linguistics and Philosophy*, *28*(2), 181–231. doi:10.1007/s10988-005-3731-9

Kaufmann, S. (2009). Conditionals Right and Left: Probabilities for the Whole Family. *Journal of Philosophical Logic*, *38*(1), 1–53. doi:10.1007/s10992-008-9088-0

Kaufmann, S. (2015). Conditionals, Conditional Probability, and Conditionalization. In H. Zeevat, H. C. Schmitz (Eds.), *Bayesian Natural Language Semantics and Pragmatics* (pp. 71–94). New York: Springer.

Kaufmann, S. (2023). Bernoulli Semantics and Ordinal Semantics for Conditionals. *Journal of Philosophical Logic*, *52*(1), 199–220. doi:10.1007/s10992-022-09670-8

Khoo, J. (2016). Probabilities of Conditionals in Context. *Linguistics and Philosophy*, *39*, 1–43.

Khoo, J., Mandelkern, M. (2019). Triviality Results and the Relationship Between Logical and Natural Languages. *Mind*, *128*(510), 485–526.

Khoo, J., Santorio, P. (2018). Lecture Notes: Probability of Conditionals in Modal Semantics. Retrieved from: http://paolosantorio.net/teaching.html

Lance, M. (1991). Probabilistic Dependence Among Conditionals. *Philosophical Review*, *100*, 269–276.

Lewis, D. (1973). *Counterfactuals*. Cambridge: Harvard University Press.

Lewis, D. (1976). Probabilities of Conditionals and Conditional Probabilities. *Philosophical Review*, *85*(3), 297–315. doi:10.2307/2184045

Lewis, D. (1986). Probabilities of Conditionals and Conditional Probabilities II. *Philosophical Review*, *95*(4), 581–589. doi:10.2307/2185051

McDermott, M. (1996). On the Truth Conditions of Certain 'If'-Sentences. *Philosophical Review*, *105*, 1–37.

McGee, V. (1985). A Counterexample to Modus Ponens. *Journal of Philosophy*, *82*(9), 462–471.

McGee, V. (1989). Conditional Probabilities and Compounds of Conditionals. *Philosophical Review*, *98*(4), 485–541. doi:10.2307/2185116

Rehder, W. (1982). Conditions for Probabilities of Conditionals to Be Conditional Probabilities. *Synthese*, *53*, 439–443.

Stalnaker, R. (1968). A Theory of Conditionals. *Studies in Logical Theory, American Philosophical Quarterly*, *2*, 98–112.

Stalnaker, R. (2009). Conditional Propositions and Conditional Assertions. In A. Egan, B. Weatherson (Eds.), *Epistemic Modality* (pp. 227–248). Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199591596.003.0008

Stalnaker, R. (2019). *Knowledge and Conditionals*. Oxford: OUP.

van Fraassen, B. C. (1976). Probabilities of Conditionals. In W. L. Harper, R. Stalnaker, G. Pearce (Eds.), *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science* (pp. 261–308). Dordrecht: Reidel.

Vineberg, S. (2016). Dutch Book Arguments. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved from: https://plato.stanford.edu/archives/spr2016/entries/dutch-book/

Węgrecki, J., Wroński, L. (2023). Revisiting the Conditional Construal of Conditional Probability. *Logic and Logical Philosophy*, *32*, 261–268.

Wójtowicz, A., Wójtowicz, K. (2021). A Stochastic Graphs Semantics for Conditionals. *Erkenntnis*, *86*, 1071–1105. doi:10.1007/s10670-019-00144-z

Wójtowicz, A., Wójtowicz, K. (2022). A Graph Model for Probabilities of Nested Conditionals. *Linguistics and Philosophy*, *45*(3), 1–48. doi:10.1007/s10988-021-09324-z

Wójtowicz, A., Wójtowicz, K. (2023). A Minimal Probability Space for Conditionals. *Journal of Philosophical Logic*, *52*, 1385–1415. doi:10.1007/s10992-023-09710-x

Wójtowicz, A., Wójtowicz, K. (2024). Probabilities of Conditionals *via* Labeled Markov Graphs, *Erkenntnis*, doi:10.1007/s10670-024-00896-3