



Dutch Book against Lewis

Anna Wójtowicz¹ · Krzysztof Wójtowicz¹

Received: 29 September 2020 / Accepted: 5 May 2021 / Published online: 20 May 2021
© The Author(s) 2021

Abstract

According to the PCCP thesis, the probability of a conditional $A \rightarrow C$ is the conditional probability $P(C|A)$. This claim is undermined by Lewis' triviality results, which purport to show that apart from trivial cases, PCCP is not true. In the present article we show that the only rational, "Dutch Book-resistant" extension of the agent's beliefs concerning non-conditional sentences A and C to the conditional $A \rightarrow C$ is by assuming that $P(A \rightarrow C) = P(C|A)$ (i.e., in accord with PCCP). In other cases a diachronic Dutch Book against the agent can be constructed. There is a tension between our findings and Lewis' results, which needs to be explained. Therefore, we present a probability space which corresponds in a natural way to the diachronic Dutch Book—and which allows the conditional $A \rightarrow C$ to be interpreted as an event in a mathematically sound way. It also allows to formalize the notion of conditionalizing $A \rightarrow C$ on $\neg C$ which plays a crucial role in Lewis' proof. Our conclusion is that Lewis' proof is circular, so it cannot be considered to be a sound argument against PCCP.

In both scientific discourse and natural language communication, when thinking about history and in everyday situations we are confronted with conditionals such as *If it is even, then it is a six* (when commenting on a die roll) or *If Reagan worked for the KGB, I'll never find out* (Lewis, 1986, p. 155) or *If I had drunk my morning coffee, I would not have headache*. Estimating or computing the probabilities of such claims is a notorious problem. In particular, a much-discussed issue is whether the

✉ Anna Wójtowicz
amwojtow@uw.edu.pl

Krzysztof Wójtowicz
wojtow@uw.edu.pl

¹ Department of Philosophical Logic, Faculty of Philosophy, University of Warsaw, Wydział Filozofii UW, Krakowskie Przedmieście 3, 00-927 Warszawa, Poland

probability of a conditional sentence equals conditional probability (in the appropriate probability space). This is Adams' definition: the probability of a conditional $P(A \rightarrow C)$ is $P(C|A)$ (Adams, 1965, 1970, 1975, 1998).¹ We shall use the standard acronym PCCP for this thesis.²

This solution is simple, attractive and coherent with our intuitions in many cases. The intuitive probability of the conditional *If it is even, then it is a six* is $1/3$ (for a fair die), which is exactly the conditional probability $P(\text{It is a six} | \text{It is even})$.³ However, it is far from obvious that the rule according to which the probability of the conditional is conditional probability should be accepted in full generality, and there is intense discussion going on. Lewis' triviality results (Lewis, 1976) (which purport to show that this equality holds in very special cases only) are very important for this discussion as they indicate that the problem has a fundamental character. In a nutshell, they suggest that it is not possible to give a sound formal argument for PCCP (apart from trivial cases), which—in Lewis' words—leads to absurd results.⁴

In the present article we propose an original solution to this problem. Our starting point is the construction of a diachronic Dutch Book for the conditional $A \rightarrow C$. We analyze the betting behavior of a rational agent whose beliefs are expressed in a propositional language that—apart from the Boolean connectives—also contains conditionals. This agent ascribes credence to their beliefs, and a minimal condition for the agent being rational is the coherence of their beliefs (in particular their credence assignments).⁵ In particular, this means that a Dutch Book **cannot** be constructed against a rational agent. We will use the term “DB-resistant” for such a system of beliefs.

¹ In Adams (1965) the term “assertability” is used. However, we will not discuss the problem of the relationship between assertability, credence, probability etc.

² It is important to stress Stalnaker's contribution (Stalnaker, 1968): indeed, PCCP is also called „Stalnaker's Thesis“. Stalnaker formulated his thesis in terms of conditional degrees of belief, while Adams originally formulated his claims in terms of assertability, and for simple conditionals only, i.e. $A \rightarrow C$, with A, C Boolean sentences (i.e. not containing the conditional operator). Adams even claimed that his approach does not extend to compound conditionals: “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). However, in its full generality, Stalnaker's thesis applies also to compound conditionals. Khoo and Santorio (2018) give a comprehensive discussion, Stalnaker (2009) gives a non-technical presentation in a broad philosophical context.

³ *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).

⁴ An extensive discussion of the PCCP problem (indicating, for instance, its importance for decision theory) can be found in Hajek (2011, 2012). Hajek also presents elegant generalizations of Lewis' results. Hajek and Hall (1994) discuss several versions of PCCP, in particular they discuss Lewis' arguments, present strengthening of triviality results and argue that PCCP cannot be upheld as a general rule. See also Fitelson (2015) for version of a triviality result.

⁵ We shall use the term “credence” to make clear that it need not be a formally defined probability function. The formal counterpart (i.e. a genuine probability function on a space) is defined later.

The role of Dutch Book analysis for the rationality of agents was stressed in Lewis (1999) and we accept his point of view (see also footnote 18).⁶ This is a fairly standard approach when considering beliefs expressed as ordinary Boolean sentences, i.e. **without** the conditional. In this paper we propose an extension of this approach to language that also contains conditionals. This will allow the following questions to be answered:

1. What is the coherent, DB-resistant extension of these beliefs to the conditional sentence $A \rightarrow C$ —assuming that the agent has a coherent, DB-resistant system of beliefs concerning the Boolean part of their language? Namely, what credence should be assigned to $A \rightarrow C$, given $P(A)$, $P(C)$ and $P(A \wedge C)$?
2. What is the appropriate (and, if possible, the simplest) probability space $S^* = (\Omega^*, \Sigma^*, P^*)$ which allows the DB-resistant probability of the conditional $A \rightarrow C$ to be computed in a formally sound way and which provides a mathematical underpinning for our DB results?
3. How should we interpret Lewis' triviality proof in the light of our findings? Is it true that (apart from trivial cases) there is no sound mathematical argument in favor of PCCP?⁷

Providing answers to these questions will give a strong argument in favor of PCCP, in particular by identifying some weaknesses in Lewis' original reasoning (Lewis, 1976).⁸ To make the presentation of our arguments lucid and illustrative, we use a simple urn example (a counterpart of the example from Edgington (1995) or Kaufmann (2004)). The numerical calculations are therefore very simple, but without loss of generality they exhibit the crucial phenomena and illustrate the problematic aspects of the intuitive, informal argumentation. The general formal construction is given in the “Appendix”.

The structure of the paper is as follows:

In Sect. 1, *An introductory example*, we show how to define a diachronic Dutch Book against an agent who has inconsistent beliefs about the credence of the conditional. Accordingly, we show that the DB-resistant credence of the conditional $A \rightarrow C$ is given as conditional probability $P(C|A)$. We show this without (yet) invoking any formally defined probability space.

In Sect. 2, *The integration of partial information*, we examine the paradoxical consequences of a straightforward application of the Law of Total Probability to

⁶ We believe that the incoherence in the agents' views can be identified by Dutch Book arguments, and that this incoherence reveals irrationality. However, there is a lively discussion going on—for example concerning the pragmatic aspects, or the problem whether there is an intimate connection between beliefs and bets. Discussing these general status of Dutch Book arguments exceeds the scope of the present study.

⁷ In short: if it was true that in the light of a Dutch Book argument PCCP must be true—what should we do with Lewis' triviality results?

⁸ The title of the paper deserves a short comment: the Dutch Book argument we present is not directly aimed at any of Lewis' theses, but rather against the rule (DB-LTP-C) used by Lewis in his proof as step (11)—see Sect. 3.

conditionals (a real-life example is also presented). The Law of Total Probability is an elementary result in probability theory, but it is a very subtle question how it works for conditional sentences.

In Sect. 3, *Lewis' reasoning*, we give an analysis of Lewis' triviality proof from Lewis (1976) and identify the problematic assumptions in this proof.

In Sect. 4, *The space S^** , we define in a mathematically sound way the probability space $S^* = (\Omega^*, \Sigma^*, P^*)$, which allows an independent mathematical underpinning to be given for the obtained results. In particular, it allows the probability of the conditional $P^*(A \rightarrow C)$ to be computed in a mathematically sound way—so that it coincides with the DB-resistant value for bets that is obtained with the aid of the analysis of optimal betting behavior. This explains the source of the misunderstandings and identifies the proper version of the Law of Total Probability for conditionals.

In Sect. 5, *Lewis' dilemma in S^** , we analyze two possible interpretations of conditionalizing the conditional $A \rightarrow C$ on B (i.e. of $P(A \rightarrow C|B)$). We show that neither of these interpretations allows two of Lewis' essential assumptions to hold simultaneously in the properly defined probability space. The constructed space $S^* = (\Omega^*, \Sigma^*, P^*)$ allows to explain where Lewis' assumptions are problematic.

We conclude with a short *Summary*.

In the “[Appendix](#)” we present the general construction of a diachronic Dutch Book for the conditional $A \rightarrow C$.

1 An introductory example

Consider an urn containing 100 White balls, 80 Green balls and 20 Red balls, all of which are equally likely to be drawn (in short: 100 W, 80G, 20R).⁹ This is modeled in a natural way by the sample space $S = (\Omega, \Sigma, P)$, in which:

$$\begin{aligned}\Omega &= \{W, R, G\}; \\ P(W) &= 100/200 = 0.5; \\ P(G) &= 80/200 = 0.4; \\ P(R) &= 20/200 = 0.1.\end{aligned}$$

Σ is the σ -field, which consists of all subsets of Ω , i.e. $\Sigma = 2^\Omega$. We will make use of this sample space $S = (\Omega, \Sigma, P)$ throughout this paper.¹⁰ The corresponding language is formed from three atomic sentences: W, G, R . In this case, we restrict ourselves to one particular conditional: $\neg W \rightarrow G$. This means that in our language we have the Boolean combinations of W, G, R and $\neg W \rightarrow G$.¹¹

⁹ In Edgington (1995) and Kaufmann (2004), there are non-Red balls, and Red balls which might or might not have a black spot. We have White and Red balls, and instead of Red spotted balls we have Green balls. We think that this makes the example simpler.

¹⁰ This level of generality is sufficient to exhibit the relevant phenomena. In particular the fact that the set of Green balls is included in the set of non-White balls does not matter.

¹¹ We do not discuss compound conditionals in this paper. In general, sentences from our language are in italics, and other objects (like elementary events, balls etc.) are not. So, W stands for the sentence “*The ball is White*”, and W is the corresponding event in the probability space.

We are interested in the probability of the conditional:

(*) *If the ball is non-White, it is Green.* (i.e. $\neg W \rightarrow G$).

None of the 8 events in $S = (\Omega, \Sigma, P)$ is a counterpart of $\neg W \rightarrow G$.¹² So, if we want to define the probability of $\neg W \rightarrow G$ in the mathematically standard way, we need to construct an appropriate probability space. Before we do this, we will find the DB-resistant credence of $\neg W \rightarrow G$ by means of a Dutch Book analysis.

1.1 Some general assumptions

In order to discuss the rational probability assignment made by the agent, we first formulate some minimal assumptions concerning the agent's decisions, their understanding of the notion of credence, and their interpretation of the conditionals.

1. The bets accepted by the rational agent

The standard way of identifying the subjective probability of a sentence is by analyzing the bets the agent considers to be fair.¹³ We use the symbol P^{DB} for the credence function of the agent. This function leads from the set of sentences in the language of the agent into the interval $[0, 1]$.¹⁴ At this stage we do not yet assume that P^{DB} is a probability function (on some probability space), but we will show later that it does indeed have a formal counterpart.

The betting behavior of the rational agent has the following properties:

1.1. If the agent thinks that $P^{DB}(A) = x$, then they consider both selling and buying bets on A for $\$x$ to be fair:

(a) The Bookmaker sells the bet on A for $\$x$, i.e. they get $\$x$ from the agent. If A happens to be true, the Bookmaker pays $\$1$ to the agent (so in this case the agent's win is $\$(1 - x)$). If A is false, the Bookmaker keeps the $\$x$.

(b) The Bookmaker buys the bet on A from the agent for x , i.e. he pays $\$x$ to the agent. If A happens to be true, the agent pays $\$1$ to the Bookmaker, i.e. the agent's loss is $\$(1 - x)$. If A is false, the agent keeps the $\$x$.

¹² In the probability space $S = (\Omega, \Sigma, P)$ there are three elementary events in Ω . The σ -field Σ is simply the power set of Ω so it contains $2^3 = 8$ events.

¹³ For a presentation of this view see Easwaran (2011a, b).

¹⁴ In the paper we use the notation for probabilities of **sentences**, i.e. we write, for instance, $P(\neg W \wedge G)$ and not $P(W^c \cap G)$, which would be typical of standard textbooks on probability theory (as events are sets). But the standard way of formulating the problem of probabilities of conditionals is in terms of probabilities of **sentences**. Moreover, it is natural to think of the system of beliefs of the agents in terms of sentences and probability/credence assignments made by the agent on these sentences. However, we also give an analysis in terms of standard probability spaces in which the sentences (including the conditionals) have their counterparts as events.

1.2 If the agent considers a bet to be fair, they are willing to repeat it an arbitrary number of times and to make n bets simultaneously (i.e. to buy/sell for $\$nx$ a bet where the win is $\$n$).

2. The credence function P^{DB}

The agent's beliefs concerning atomic sentences are modeled in the sample probability space $S = (\Omega, \Sigma, P)$. We know that if the function P^{DB} violated the rules of probability, it would be possible to construct a Dutch Book against the agent, i.e. a series of bets, each of which the agent considers to be fair but which inevitably lead to the agent's loss.¹⁵

So, for Boolean sentences A and C (i.e. formed from atomic sentences with the use of only \neg, \wedge, \vee but **not** the conditional \rightarrow), the rational agent's credence function P^{DB} is exactly the probability function P , i.e.:

$$2.1 \quad P^{DB}(A) = P(A).$$

Of course, because it is a probability function that is defined on the Boolean part of the language, P^{DB} has the following properties:

$$2.2 \quad P^{DB}(A) = 1 - P^{DB}(\neg A).$$

$$2.3 \quad P^{DB}_A(C) = \frac{P^{DB}(A \wedge C)}{P^{DB}(A)}.$$

Here $P^{DB}_A(\cdot)$ is the credence function which results from a belief revision by conditionalizing on a sentence A (i.e. the agent assumes that A is true and modifies the belief system appropriately).¹⁶

3. Interpretation of the conditional $A \rightarrow C$

The agent has the conditional \rightarrow in their language, so we have to give an outline of how the conditional is interpreted. Of course, $A \rightarrow C$ is **not** identified with the material implication (i.e. with the Boolean sentence $\neg A \vee C$).

3.1 In order to make and settle bets on conditionals, the agent has to assume that there are circumstances in which the conditional is considered to be true, and there are also circumstances in which the conditional is considered to be false. This means that the notion of truth conditions for $A \rightarrow C$ is accepted.

3.2 The agent believes that if both A and C are true, then the conditional $A \rightarrow C$ is true.

3.3 The agent believes that if A is true but C is false, then the conditional $A \rightarrow C$ is false.

¹⁵ See for instance Hajek (2009) or Vineberg (2016) for general presentation.

¹⁶ This is a standard conditionalization rule. For discussion of what happens if the agent violates this rule, see for instance Rescorla (2020) and the extensive bibliography within.

3.4 The agent believes that if A is false, then the conditional $A \rightarrow C$ is neither true nor false.¹⁷

3.5 If $P^{DB}(A \rightarrow C) = x$, then the agent considers the bet on $A \rightarrow C$ for $\$x$ to be fair (cf. 1.1), but assumes that it will be cancelled if A turns out to be false (in this case the $\$x$ is refunded to the buyer of the bet).

These three groups of assumptions each have a different character:

Assumptions (1.1–1.2) describe the decisions that the agent make after ascribing credence to sentences from the language. In particular, there is no risk aversion (and no propensity to risk) and the agent is always ready to engage in actions (bets) which have a non-negative expectancy.

Assumptions (2.1–2.3) have a practical character: the agent knows that if their credence function P^{DB} violates these assumptions, it will be possible to construct a (synchronic or diachronic) Dutch Book against them. In terms of the *practical* character of these assumptions, we mean not the real financial loss¹⁸ but the fact that the agent can identify the desired properties of the function P^{DB} by simply analyzing the safe values of bets which give a coherent, DB-resistant system of beliefs. This argument is independent of the formal construction of any probability space.

Assumptions (3.1–3.4) clarify the understanding of the conditional (in particular, of the truth conditions of the conditionals). They might be viewed as general postulates which are independent of any previous probability considerations.

Assumption 3.5 is the consequence of 3.1–3.4 and the assumptions regarding fair bets.

1.2 A diachronic Dutch Book for the conditional $\neg W \rightarrow G$

What credence should our agent assign to the conditional $\neg W \rightarrow G$? Namely, what is the DB-resistant value of $P^{DB}(\neg W \rightarrow G)$? We strongly believe that the proper answer (for the urn containing 100 W, 80G, 20R) is 0.8, as 80% of the non-White balls are Green. However, the argumentation should be based only on the assumptions 1.1–3.5 and proceeds by defining a Dutch Book against the agent, who accepts a different value.

So, assume that the agent believes the opposite:

¹⁷ There is a tradition dating back at least to De Finetti by which conditionals with a false antecedent are analyzed by assuming an **additional logical value** in the model. [Douven (2016) gives an analysis; for formal results of the trivalent approach to conditionals, see for instance Égre et al. (2019)]. We do not discuss this matter here.

¹⁸ “Note also that the point of any Dutch book argument is not that it would be imprudent to run the risk that some sneaky Dutchman will come and drain your pockets. After all, there aren’t so many sneaky Dutchmen around; and anyway, if ever you see one coming, you can refuse to do business with him. Rather, the point is that if you are vulnerable to a Dutch book, whether synchronic or diachronic, that means that you have two contradictory opinions about the expected value of the very same transaction. To hold contradictory opinions may or may not be risky, but it is in any case irrational”. Lewis (1999, pp. 404–405).

$$P^{DB}(\neg W \rightarrow G) \neq 0.8.$$

We will show that this leads to disaster, i.e. to a diachronic Dutch Book against the agent. In this case, assume that

$$P^{DB}(\neg W \rightarrow G) = 0.875,$$

i.e. that the agent considers \$0.875 to be a fair price for a standard bet on $(\neg W \rightarrow G)$. The number 0.875 is chosen to make the calculations convenient, but the reasoning is similar for any value of $0.8 < p < 1$.¹⁹

We assume that in the course of events, the agent and the bookmaker agree that it is possible to update their knowledge (assumption 2.3) and to place new bets after this update; of course, they agree on the present status of the bet.

We play the role of the Bookmaker; of course, we will only buy/sell bets which are considered fair by the agent. The game consists of drawing a ball from the urn. Before we draw, we make the following bets:

Bet(W). We buy from the agent the bet that the ball is White for \$0.5.²⁰

Bet($\neg W \rightarrow G$). We sell 10 bets on $\neg W \rightarrow G$ for \$0.875 each.²¹

The agent accepts the bets as they consider them to be fair and coherent with their credence assignments (assumptions 1.1 and 1.2).

We start the game, i.e. a ball is drawn. The first step is to identify whether the ball is White or not: this will settle **Bet(W)**. Importantly, at this moment **we only** check whether it is White or not! If it is White, we finish the game and take our money, as we have won **Bet(W)**. But if it is **not** White, we make another bet before we check whether it is Red or Green!²²

So, these are our actions:

If the ball is White:

- The agent loses **Bet(W)** and has to pay \$1 (i.e. the agent's loss is \$0.5);

¹⁹ For $0 < p < 0.8$, the reasoning has a symmetric character: instead of selling bets, we buy bets and vice versa. The construction of the Dutch Book depends on the concrete value of p , but the scheme is the same. In the "Appendix" we present the general construction.

²⁰ We pay \$0.5. If the drawn ball is W, we win, i.e. the agent has to pay 1\$. If the drawn ball is not W, we lose, and the agent keeps the \$0.5.

²¹ We get \$0.875 for each bet on $\neg W \rightarrow G$ from the agent. We keep the money (10x\$0.875) if $\neg W \rightarrow G$ is **not** true, and we pay \$10 if $\neg W \rightarrow G$ is true.

²² Edgington's and Kaufmann's example of non-Red balls and Red balls with or without a black spot (see footnote 9) are perhaps more illustrative in this respect: we already see that the ball is Red, but we still have to check whether it has a black spot or not. Our process of gaining new knowledge extends in time.

- $\text{Bet}(\neg W \rightarrow G)$ is cancelled and \$8.75 (i.e. $10 \times \$0.875$) is given back to the agent (Assumption 3.5).

If the ball is non-White:

- The agent wins $\text{Bet}(W)$ and keeps (i.e. wins) \$0.5.
- We do not yet have knowledge about the outcome of $\text{Bet}(\neg W \rightarrow G)$ as the color of the non-White ball is not yet known.

At exactly this moment, we propose $\text{Bet}(R)$, i.e. the bet that the ball is Red. Both we and the agent have updated our beliefs, conditionalizing on $\neg W$ (according to Assumption 2.3.). So now we use the credence function $P^{\text{DB}}_{\neg W}$. The agent knows that the fair bet value on R (given it is $\neg W$) is 0.2, as.

$$P^{\text{DB}}_{\neg W}(R) = \frac{P^{\text{DB}}(\neg W \wedge R)}{P^{\text{DB}}(\neg W)} = \frac{P(\neg W \wedge R)}{P(\neg W)} = 0.2 \text{ (according to Assumptions 2.1 and 2.3).}$$

A bet is now made:

Bet(R). We sell 10 bets that the ball in question is Red for \$0.2 each.²³

After the bet we check the color of the non-White ball. Of course, being non-White, the ball is either Green or Red.

If the ball is Green:

- The agent wins $\text{Bet}(\neg W \rightarrow G)$ —so wins $10 \times (1 - \$0.875) = \1.25 .
- The agent loses $\text{Bet}(R)$ —so loses $10 \times \$0.2 = \2 .

If the ball is Red:

- The agent loses $\text{Bet}(\neg W \rightarrow G)$ —so loses $10 \times \$0.875 = \8.75 .
- The agent wins $\text{Bet}(R)$ —so wins $10 \times \$0.8 = \8 .

The agent’s wins and losses in these three cases are summarized in the table:

	Bet(W)	Bet($\neg W \rightarrow G$)	Bet(R)	The outcome
W	−\$0.5	0 (the bet was cancelled)	0 (the bet has not been placed)	−\$0.5
G	+\$0.5	+\$1.25	−\$2	−\$0.25
R	+\$0.5	−\$8.75	+\$8	−\$0.25

²³ We get $10 \times \$0.2 = \2 from the agent. We keep the $10 \times \$0.2 = \2 if the ball is Red and have to pay $10 \times \$1$ to the agent if the ball is not Red (the agent wins \$8 in this case).

Regardless of the result (i.e. the color of the drawn ball), the agent loses. This is the consequence of assuming that $P^{DB}(\neg W \rightarrow G) \neq 0.8$, i.e. that $P^{DB}(\neg W \rightarrow G) \neq P(G|\neg W)$. This reasoning shows that it is rational to assume the proper “betting credence” of the conditional $\neg W \rightarrow G$ to be $P^{DB}(\neg W \rightarrow G) = 0.8$. Of course, this is a general rule: it is **not** a numerical artifact related to some special properties of the particular values 0.8, 0.875 etc.

Importantly, we assumed that the agent’s beliefs concerning the Boolean fragment of the language are coherent, which means that no “Boolean Dutch Book” can be constructed. Nevertheless, a diachronic DB can be constructed against a system of beliefs to which the conditional sentence in question is added and in which the agent assumes that $P^{DB}(\neg W \rightarrow G) \neq P(G|\neg W)$. This is the source of the incoherence and the agent’s misfortune.

It is also important to observe that we have constructed the Dutch Book in question based on the internal relationships between the agent’s beliefs: we need not examine whether they are adequate in any sense to the empirical situation. This is a general feature of the Dutch Book construction. Even if the agent believes that regardless of the content of the urn the probabilities are always $P^{DB}(W) = 1/3$; $P^{DB}(R) = 1/3$; $P^{DB}(G) = 1/3$; $P^{DB}_{\neg W}(G) = 1/2$; $P^{DB}_{\neg W}(R) = 1/2$; $P^{DB}(\neg W \rightarrow G) = 1/2$, it will **not** be possible to construct a Dutch Book (even if it is highly probable that the Bookmaker wins the bets in the long run by making use of the law of large numbers). It is the incoherence in the agent’s beliefs (i.e. the fact that the agent believes that $P^{DB}(\neg W \rightarrow G) \neq P(G|\neg W)$) which leads to the diachronic Dutch Book and an inevitable loss.

At this stage we have not yet used any formal model (semantics) for probabilities of conditionals. The Dutch Book reasoning has a very intuitive character: informally speaking, it identifies the rational betting behavior, but does not explain the mathematical (theoretical) reasons. Nevertheless it: (i) gives an elementary argument in favor of PCCP; (ii) helps to identify the problematic places in Lewis’ reasoning; and (iii) provides an important boundary condition: the to-be-constructed probabilistic model must not lead to a Dutch Book disaster!

2 The integration of partial information

The findings from the previous section can be summarized as follows:

the only way to expand the agent’s belief system in a DB-resistant way to include the conditional $\neg W \rightarrow G$ is by assigning:

$$P^{DB}(\neg W \rightarrow G) = P(G|\neg W).$$

This formula can also be written in the form:

$$P^{DB}(\neg W \rightarrow G) = P^{DB}_{\neg W}(G).$$

To justify this claim, we made use only of some non-controversial assumptions concerning the agent’s “betting behavior”. We will use our findings in the discussion concerning Lewis’ triviality results, which—according to Lewis—prove that the PCCP principle, i.e.

$$P(A \rightarrow C) = P_A(C) = P(C|A).$$

can be formally justified only in very special cases, namely in **trivial**²⁴ probability spaces in which at most 4 values of probabilities of events are assumed.²⁵ Our sample space $S = (\Omega, \Sigma, P)$ obviously is not such a special case, but the DB-resistant values of P^{DB} for conditionals obey the PCCP principle. Lewis’ theorem suggests that this result cannot be given a proper mathematical formalization and justification, so a feeling of incoherence arises which needs to be explained.

To clearly exhibit the problematic assumptions in Lewis’ reasoning, we will modify our example. Consider an additional feature of the balls in our sample space (apart from their color): some of them are Heavy (H) and some are Light (L).²⁶ In this case, assume that

- (a) there are 20 White, 60 Green and 20 Red balls in the “Heavy subspace”;
- (b) there are 80 White and 20 Green balls in the “Light subspace”.

Assume that our agent has the following partial information (concerning the Heavy and Light subspaces) at their disposal:

- (i) They know the values $P^{DB}(H)$ and $P^{DB}(L)$, i.e. the chances of drawing a Heavy/Light ball.
- (ii) They know the values $P^{DB}_H(Y)$ and $P^{DB}_L(Y)$ of every **Boolean** sentence Y .

Is this knowledge sufficient to compute $P^{DB}(Y)$? The Law of Total Probability (an elementary theorem in probability theory) states that:

$$(LTP) P(Y) = P(Y|H) P(H) + P(Y|L) P(L) = P_H(Y) P(H) + P_L(Y) P(L),$$

for any event Y and complementary events H, L in the probability space.

²⁴ “If a probability function never assigns positive probability to more than two incompatible alternatives, and hence is at most four-valued, then we may call it a **trivial probability function**” (Lewis 1976, p. 302).

²⁵ Lewis does **not** claim that the genuine probability of, for instance, *If it is Even, it is a Six* (for a fair die), is 1/6. Rather he purports to show that it is highly problematic to justify such probability claims as reasoning that appears to be sound leads to absurd results.

²⁶ This means that in our language there are six sentences that represent all 6 combinations of properties: Heavy White, Heavy Green, Heavy Red, Light White, Light Green, Light Red.

So, the version for P^{DB} is straightforward:

$$(DB-LTP) P^{DB}(Y) = P^{DB}_H(Y) P^{DB}(H) + P^{DB}_L(Y) P^{DB}(L).^{27}$$

The agent can use this formula to integrate their partial knowledge concerning the **proper** values of P^{DB} in the subspaces H and L in order to obtain the **proper** value of P^{DB} in the whole space. Another way of putting it is that someone who is a DB-resistant Y-player in the Heavy and Light subspaces will become a DB-resistant Y-player in the whole urn just by using LTP. Undoubtedly this method works for all **Boolean sentences**. A natural question is whether the same can be done with the conditional $\neg W \rightarrow G$.

Assume now that the agent has the following information:

- (i) they know the proper (DB-resistant) values $P^{DB}(H)$, $P^{DB}(L)$;
- (ii) they know the proper (DB-resistant) values $P^{DB}_H(\neg W \rightarrow G)$ and $P^{DB}_L(\neg W \rightarrow G)$.

The analogue of (DB-LTP) is obtained by substituting $\neg W \rightarrow G$ for Y and has the form:

$$(DB-LTP-C) P^{DB}(\neg W \rightarrow G) = P^{DB}_H(\neg W \rightarrow G) P^{DB}(H) + P^{DB}_L(\neg W \rightarrow G) P^{DB}(L).$$

(The ‘‘C’’ in (DB-LTP-C) is for ‘‘conditional’’). In our example, the proper (i.e. DB-resistant) values in the subspaces are: $P^{DB}_H(\neg W \rightarrow G) = 0.75$ and $P^{DB}_L(\neg W \rightarrow G) = 1$.²⁸ We also have $P^{DB}(H) = P^{DB}(L) = 0.5$. Applying the formula (DB-LTP-C) gives:

$$P^{DB}(\neg W \rightarrow G) = 0.75 \times 0.5 + 1 \times 0.5 = 0.875$$

which is wrong! We have already shown in Sect. 1 that taking this value leads to a Dutch Book against the agent, and that the proper DB-resistant value is $P^{DB}(\neg W \rightarrow G) = 0.8$.

Importantly, we started with the **proper** (DB-resistant) values for $P^{DB}_H(\neg W \rightarrow G)$, $P^{DB}(H)$, $P^{DB}_L(\neg W \rightarrow G)$ and $P^{DB}(L)$. Then we applied (DB-LTP-C) and obtained absurd results! This means that the Law of Total Probability holds for Boolean sentences (which is not surprising, as this is elementary probability theory), but its counterpart (DB-LTP-C) **does not** hold for the conditional $\neg W \rightarrow G$. Using the game metaphor, we can say that LTP is a golden rule for people who know good ‘‘Heavy subspace / Light subspace’’ strategies that can be used to plan the strategy for the

²⁷ To justify it, we only need to use (a) assumption 2.1, i.e. $P^{DB}(Y) = P(Y)$ for any Boolean sentence Y; (b) assumption 2.3, i.e. $P^{DB}_A(B) = \frac{P^{DB}(A \wedge B)}{P^{DB}(A)}$.

²⁸ The reasoning which justifies these claims has exactly the same structure as the reasoning which allows the proper value of $P^{DB}(\neg W \rightarrow G)$ to be identified. This is also intuitive: 75% of the non-White Heavy balls are Green, and 100% of the non-White Light balls are Green.

whole urn; however, this works for **Boolean** sentences only! Importantly, in Sect. 5 we shall see that LTP-C in this misleading form is involved in Lewis’ reasoning.

Remark: We will obtain the proper value of $P^{DB}(\neg W \rightarrow G)$ using a different “information integration formula” for subspaces H and L:

$$(DB-LTP-C^*) P^{DB}(\neg W \rightarrow G) = P^{DB}_H(\neg W \rightarrow G) P^{DB}_H(\neg W) + P^{DB}_L(\neg W \rightarrow G) P^{DB}_L(\neg W).$$

Indeed:

$$P^{DB}_H(\neg W) = P(\neg W|H) = 20/100 = 0.2;$$

$$P^{DB}_L(\neg W) = P(\neg W|L) = 80/100 = 0.8.$$

After substituting these values to the formula (DB-LTP-C*), we obtain:

$$P^{DB}(\neg W \rightarrow G) = 0.2 \times 1 + 0.8 \times 0.75 = 0.8$$

This is not a coincidence but a general phenomenon. We present a proof of (DB-LTP-C*). To begin with, it is a simple exercise to check that:

$$P(G|\neg W) = P_H(G|\neg W) P(H|\neg W) + P_L(G|\neg W) P(L|\neg W)$$

where P is a standard probability function, P(.) is conditional probability, P_H and P_L are conditionalized probability measures.²⁹ We use this formula to justify (DB-LTP-C*).

We have already justified the claim, that the proper, DB-resistant value $P^{DB}(\neg W \rightarrow G)$ is $P(G|\neg W)$. The DB-argument is general, and it applies to any probability measure, in particular to $P^{DB}_H(\neg W \rightarrow G)$ and $P^{DB}_L(\neg W \rightarrow G)$. We also know, that for Boolean sentences X, $P^{DB}(X) = P(X)$ (this applies to P_H and P_L as well). So, if we substitute:

- $P^{DB}(\neg W \rightarrow G)$ for $P(G|\neg W)$;
- $P^{DB}_H(\neg W \rightarrow G)$ for $P_H(G|\neg W)$;
- $P^{DB}_H(\neg W)$ for $P(H|\neg W)$;
- $P^{DB}_L(\neg W \rightarrow G)$ for $P_L(G|\neg W)$;
- $P^{DB}_L(\neg W)$ for $P(L|\neg W)$;

²⁹ Indeed, this is a special case of a general formula: take a partition $\{C_1, \dots, C_n\}$ and let P_i be the respective conditionalized probability measures, i.e. $P_i(A) = P(A|C_i)$, for $i = 1, \dots, n$. Then $P(G|\neg W) = P_1(G|\neg W) P_1(C_1|\neg W) + \dots + P_n(G|\neg W) P_n(C_n|\neg W)$.

in the formula above, we obtain the required:

$$(DB-LTP-C^*) \quad P^{DB}(\neg W \rightarrow G) = P^{DB}_H(\neg W \rightarrow G)P^{DB}_H(\neg W) + P^{DB}_L(\neg W \rightarrow G)P^{DB}_L(\neg W).^{30}$$

Is vaccination effective? To illustrate the misleading intuitions which suggest the wrong formula (DB-LTP-C), consider a population in which people are:

- (i) type A or B (this might, for instance, be some genetic feature);
- (ii) either Sick or not ($S / \neg S$);
- (iii) either vaccinated or not ($V / \neg V$).

We are interested in the effectiveness of vaccination, i.e. in the probability of a vaccinated person not getting sick. This means that we are interested in the probability of the conditional *If someone has been vaccinated, they won't get sick*:

$$(VAC) \quad V \rightarrow \neg S.$$

Our task is to evaluate $P(V \rightarrow \neg S)$ on the basis of empirical data collected within two **subpopulations** A and B. Experts working in subpopulations A and B know that DB is a good way of producing coherent systems of beliefs, so they use it to estimate $P_A(V \rightarrow \neg S)$ and $P_B(V \rightarrow \neg S)$. Their findings are as follows:

In subpopulation A: $P_A(V \rightarrow \neg S) = 1$ (as 100% of the vaccinated A-people did not get sick).

In subpopulation B: $P_B(V \rightarrow \neg S) = 0.5$ (as 50% of the vaccinated B-people did not get sick).

In this case, assume that the number of A-people and B-people is equal, i.e. $P(A) = P(B) = 0.5$. If we use the (DB-LTP-C) formula to estimate the probability of $V \rightarrow \neg S$, i.e. $P(V \rightarrow \neg S)$ in the whole population, the result is:

$$P(V \rightarrow \neg S) = P_A(V \rightarrow \neg S) \times P(A) + P_B(V \rightarrow \neg S) \times P(B) = 1 \times 0.5 + 0.5 \times 0.5 = 0.75.$$

This is not a reasonable result as it gives the (usually false!) value 0.75 in radically different situations. For instance:

³⁰ It might happen that (DB-LTP-C) and (DB-LTP-C*) coincide: for instance in the trivial case, when $P^{DB}(H) = P^{DB}(\neg W|H)$ and $P^{DB}(L) = P^{DB}(\neg W|L)$. In such cases, the formula (DB-LTP-C) gives good results. However, it cannot be safely used as a general rule. We owe this comment to the anonymous referee, and we also thank for pressing us to give a clear mathematical justification of the formula (DB-LTP-C*).

Situation 1: There are 100 A-people and 100 B-people, but there are only 2 vaccinated A-people, and **neither of them** are sick (so indeed $P_A(V \rightarrow \neg S) = 1$); there are 98 vaccinated B-people, and 49 of them are not sick (so indeed $P_B(V \rightarrow \neg S) = 0.5$). In this case there are 100 vaccinated people, and 51 of them are not sick. So, $P(V \rightarrow \neg S)$ does **not** equal 0.75 but is much lower ($51/100 = 0.51$).

Situation 2: There are 98 vaccinated A-people (none of them are sick, so indeed $P_A(V \rightarrow \neg S) = 1$) and 2 vaccinated B-people (1 of them is not sick, so indeed $P_B(V \rightarrow \neg S) = 0.5$). In this case $P(V \rightarrow \neg S)$ does **not** equal 0.75 but is much higher ($99/100 = 0.99$).

$P(V \rightarrow \neg S)$ is computed by means of the Law of Total Probability as a weighted sum. But (DB-LTP-C) uses the wrong weights, i.e. the relative size of the samples.³¹ The proper weights should be **the relative frequency of the vaccinations** in the sub-populations—exactly as in the **true** formula (DB-LTP-C*).

3 Lewis’ reasoning

Let us recall an important fragment of Lewis’ argumentation from Lewis (1976):

“(7) $P(A \rightarrow C|B) = P(C|AB)$, if $P(AB)$ is positive.

[...]

we have:

(8) $P(A \rightarrow C) = P(C|A)$

By (7), taking B as C or as \bar{C} and simplifying the right-hand side, we have:

(9) $P(A \rightarrow C|C) = P(C|A C) = 1$

(10) $P(A \rightarrow C|\bar{C}) = P(C|A \bar{C}) = 0$

For any sentence D, we have the familiar expansion by case:

(11) $P(D) = P(D|C) \times P(C) + P(D|\bar{C}) \times P(\bar{C})$.

In particular take D as $(A \rightarrow C)$. Then we may substitute (8), (9), and (10) into (11) to obtain:

(12) $P(C|A) = 1 \times P(C) + 0 \times P(\bar{C}) = P(C)$ ” (Lewis, 1976, p. 300)

Lewis describes this result as “absurdity, but not quite a contradiction” (Lewis, 1976, p. 300).

In our terminology, Lewis divides the set of possible events into two (virtual) subspaces C and $\neg C$. Then, in steps (9) and (10) of his reasoning, he claims that:

³¹ This reminds of to Simpson’s paradox in statistics, where seemingly compelling reasoning concerning the presence of a trend in different groups leads to false results when these groups are combined (“amalgamated”).

- (i) in subspace C: $P_C(A \rightarrow C) = 1$;
- (ii) in subspace $\neg C$: $P_{\neg C}(A \rightarrow C) = 0$.

In step (12), in order to compute the total probability of the conditional $A \rightarrow C$, Lewis uses formula (11), which is an exact counterpart of the formula (DB-LTP-C). But we have already seen that this formula is wrong: depending on the division of the sample space into subspaces, it leads to inconsistent results.

To give an illustration of the weak point in Lewis’ reasoning, we show its exact counterparts in our vaccination and urn examples. The first seems to be more intuitive, while the second allows to perform exact numerical calculations.

The vaccination example. Let us divide the population in the vaccination example into the subpopulations of healthy people (i.e. $\neg S$) and sick people (i.e. S). Undoubtedly

$$P_{\neg S}(V \rightarrow \neg S) = 1;$$

$$P_S(V \rightarrow \neg S) = 0.$$

Integrating these results with the use of Lewis’ formula (i.e. DB-LTP-C) can give wrong results (see Sect. 2).

The urn example. We divide the urn into subspaces of Green and non-Green balls:

Subspace G: 80G;
 Subspace $\neg G$: 100 W, 20R.

The probability of choosing subspace G is $P^{DB}(G) = 80/200 = 0.4$; for subspace $\neg G$ this is $P^{DB}(\neg G) = 120/200 = 0.6$.

The diachronic Dutch Book reasoning within the subspaces G and $\neg G$ shows that:

$$P^{DB}_G(\neg W \rightarrow G) = 1;$$

$$P^{DB}_{\neg G}(\neg W \rightarrow G) = 0.^{32}$$

After applying the formula (DB-LTP-C):

$$P^{DB}(\neg W \rightarrow G) = P^{DB}_G(\neg W \rightarrow G) P^{DB}(G) + P^{DB}_{\neg G}(\neg W \rightarrow G) P^{DB}(\neg G)$$

we obtain the wrong value:

$$P^{DB}(\neg W \rightarrow G) = P^{DB}(G) = 0.4 \text{ (instead of } 0.8\text{).}$$

³² This is absolutely obvious even without any Dutch Book analysis: if there are **only** Green balls, you can bet all your money on $\neg W \rightarrow G$. Similarly, if there are **no** Green balls, you should not bet anything on $\neg W \rightarrow G$!

$P^{DB}(\neg W \rightarrow G)$ depends therefore on the division into subspaces: (i) for our H/L division from Sect. 2, we obtain 0.875; (ii) for Lewis’ division $G/\neg G$, we obtain 0.4; (iii) for the division $R/\neg R$, we obtain 0.9.

Observe also that if we apply the formula (DB-LTP-C*) we get:

$$P^{DB}(\neg W \rightarrow G) = P^{DB}_G(\neg W \rightarrow G) P_{\neg W}(G) + P^{DB}_{\neg G}(\neg W \rightarrow G) P_{\neg W}(\neg G) = 1 \times 0.8 + 0 \times 0.2 = 0.8,$$

i.e. the proper result.

So far, we have arrived at the following conclusions.

Observation 1. The possibility of producing a Dutch Book against the agent arises when their systems of beliefs are incoherent. If the agent’s credence function P^{DB} for the Boolean part of the language is modeled within the probability space $S = (\Omega, \Sigma, P)$, the **only coherent** (i.e. DB-resistant) way of expanding it to the conditional $A \rightarrow C$ is by setting $P^{DB}(A \rightarrow C) = P^{DB}_A(C) = P(C|A)$.

Observation 2. The integration of the partial information concerning the **proper** (sic!), i.e. DB-resistant values, of $P^{DB}_H(\neg W \rightarrow G)$ and $P^{DB}_L(\neg W \rightarrow G)$, by means of the formula (DB-LTP-C) gives the **wrong** value of $P^{DB}(\neg W \rightarrow G)$. We have shown that in the general case:

$$P^{DB}(\neg W \rightarrow G) \neq P^{DB}_H(\neg W \rightarrow G) P^{DB}(H) + P^{DB}_L(\neg W \rightarrow G) P^{DB}(L).$$

As a consequence, this opens the way to constructing a diachronic DB against the agent. This means that the straightforward, “mechanical” generalization of the Law of Total Probability does not work for conditionals.

Observation 3. The reasoning we analyzed is based on the misleading formula (DB-LTP-C) and leads to a contradiction: we obtain different values for $P^{DB}(\neg W \rightarrow G)$ depending on the division of the sample space. But this reasoning is similar to Lewis’ argumentation. So, this suggests that it is based on implausible assumptions. We think that the source of the problems are as follows:

1. Lewis did not specify any probability space on which the conditional $A \rightarrow C$ is interpreted as an event.
2. Lewis does not explain what the term $P(A \rightarrow C|B)$ really means, but the assumption that $P(A \rightarrow C|\neg C) = 0$ is crucial in his reasoning.

This is why our aim in the next section is to construct a probability space $S^* = (\Omega^*, \Sigma^*, P^*)$ in which the conditional $\neg W \rightarrow G$ is interpreted as an event, so that $P^*(\neg W \rightarrow G)$ is a mathematically well-defined term. This will allow us to understand (i) why the Law of Total Probability in the form (DB-LTP-C) fails and how it should be applied properly; (ii) whether it is justified to assume that $P^*(\neg W \rightarrow G|\neg G) = 0$.

4 The space S^*

The Dutch Book-resistant value for $P^{DB}(A \rightarrow C)$ is $P(C|A)$. This formula works, but why? Can we consider $P^{DB}(A \rightarrow C)$ to be a genuine probability, or is it just some real number that only accidentally coincides with $P(C|A)$ and has nothing to do with any genuine probabilistic structure? What are the reasons for the failure of the generalization of LTP for P^{DB} ?

We want to understand the logical structure of the problem, i.e. we want to describe the situation in a mathematically sound model. The model should give a mathematical underpinning for the formula $P^{DB}(A \rightarrow C) = P(C|A)$ so that we can apply standard mathematical theorems (in particular LTP) to P^{DB} . This means that we need to construct a probability space $S^* = (\Omega^*, \Sigma^*, P^*)$ in which the conditional $A \rightarrow C$ will have an interpretation as an event so that $P^*(A \rightarrow C)$ is a well-behaved probability function. In the following presentation, we keep matters as simple as possible.

Undoubtedly, the to-be-constructed probability space $S^* = (\Omega^*, \Sigma^*, P^*)$ has to fulfill some requirements:

- $S^* = (\Omega^*, \Sigma^*, P^*)$ is associated with the sample space $S = (\Omega, \Sigma, P)$, which models the Boolean fragment of the agent's beliefs. $S^* = (\Omega^*, \Sigma^*, P^*)$ should allow the Boolean sentences to be interpreted, and the probabilities from the sample space $S = (\Omega, \Sigma, P)$ should be preserved in $S^* = (\Omega^*, \Sigma^*, P^*)$.
- The conditional $\neg W \rightarrow G$ should have a representation as an event, i.e. a subset $[\neg W \rightarrow G] \subseteq \Omega^*$. To avoid confusion, we will use square brackets to distinguish sentences from their counterparts (events) in the probability space S^* . But this means that any elementary event $\omega^* \in \Omega^*$ is either the element of $[\neg W \rightarrow G]$ or its complement, i.e. it confirms or disconfirms the conditional $\neg W \rightarrow G$. The construction is therefore based on the assumption that there are circumstances in which $\neg W \rightarrow G$ is true (false), i.e. we accept the notion of truth conditions for conditionals (assumptions 3.1–3.4).
- As it is convenient to think of probabilities of conditionals in terms of bets, it is natural to describe their truth conditions in terms of game scenarios (which lead to either a win or a loss). The games have a dynamic character, i.e. there is always some unfolding scenario, and bets are placed at appropriate moments of the game. S^* should therefore model this fact and naturally correspond to the gambling situation.

Imagine that we start the “conditional game” for $\neg W \rightarrow G$: the rules are fixed (and conform to assumptions 3.2–3.5), the appropriate bets are made, and we draw the first ball. Obviously, if the first ball is Green or Red, the conditional $\neg W \rightarrow G$ is decided. If the first ball is White, $\neg W \rightarrow G$ is undecided, and we have to draw a ball again. This process is repeated until either R or G appears, at which point the conditional is shown to be true or false.

So, the scenarios confirming the conditional $\neg W \rightarrow G$ are:

- G, WG, WWG, WWWG, ...

And the scenarios disconfirming the conditional $\neg W \rightarrow G$ are:

- R, WR, WWR, WWWR, ...

We will use W^nG (*resp.* W^nR) to denote the sequence consisting of n White balls followed by a Green ball (*resp.* a Red ball). These scenarios are natural candidates for elementary events in S^* .

4.1 Formal definition of $S^* = (\Omega^*, \Sigma^*, P^*)$ ³³

We start with our toy-example sample space $S = (\Omega, \Sigma, P)$, in which $\Omega = \{W, G, R\}$ with $P(G) = p$, $P(R) = q$, $P(W) = r$. The conditional in question is $\neg W \rightarrow G$. The associated probability space $S^* = (\Omega^*, \Sigma^*, P^*)$ is defined in the following way:

$$\begin{aligned} \Omega^* &= \{W^nG, W^nR : n \in \mathbb{N}\}; \\ \Sigma^* &= 2^{\Omega^*} \text{ (i.e. the power set of } \Omega^*); \\ P^*(W^nG) &= r^n p \text{ (for } n \in \mathbb{N}); \\ P^*(W^nR) &= r^n q \text{ (for } n \in \mathbb{N}). \end{aligned}$$

In $S^* = (\Omega^*, \Sigma^*, P^*)$, elementary events are not balls but sequences of balls (intuitively speaking, to “draw” an event from Ω^* is to draw a “game scenario”). The stipulations $P^*(W^nG) = r^n p$ and $P^*(W^nR) = r^n q$ are straightforward.

Ω^* is a countable set, so for any set $A \subseteq \Omega^*$ we set:

$$P^*(A) = \sum_{\omega \in A} P^*(\omega).$$

So, P^* is indeed a probability function properly defined on the σ -field Σ^* .

We can interpret the sentences from our language (i.e. the Boolean sentences **and** the conditional $\neg W \rightarrow G$) in the probability space $S^* = (\Omega^*, \Sigma^*, P^*)$ as events, i.e. subsets of Ω^* . In particular, $\neg W \rightarrow G$ has its counterpart $[\neg W \rightarrow G] \subseteq \Omega^*$, which consists of scenarios confirming the conditional:

$$[\neg W \rightarrow G] = \{W^nG : n \in \mathbb{N}\}.$$

Of course

$$P^*([\neg W \rightarrow G]) = \sum_{n=0}^{\infty} P r^n = \frac{p}{1-r} = \frac{p}{p+q},$$

i.e.

$$P^*([\neg W \rightarrow G]) = P(G|\neg W) = P^{DB}(\neg W \rightarrow G)$$

³³ The construction of this space is based on the construction given in Wójtcowicz and Wójtcowicz (2019, 2021). However, we give independent motivations and do not use the Markov chains formalism.

where P is the probability from the sample space $S = (\Omega, \Sigma, P)$.³⁴ This means that the probability P^* models the DB-resistant value P^{DB} .³⁵

4.2 The probability of Boolean sentences in S^*

The space $S^* = (\Omega^*, \Sigma^*, P^*)$ is designed to offer interpretations for Boolean sentences and the conditional $\neg W \rightarrow G$ from our language.³⁶ To give an interpretation of a given sentence α is to define the set $[\alpha] \subseteq \Omega^*$. This means that for any event $\omega^* \in \Omega^*$ we have to decide whether $\omega^* \in [\alpha]$, i.e. whether it makes the given sentence α true or false.

Consider $\alpha = \textit{The ball is Green}$. Which of the elementary events $\omega^* \in \Omega^*$, i.e. sequences W^nG and W^nR , make α true? We start the game, and obviously we claim that α is true (“We win the game”) when we see a Green ball as the first ball. We lose when we see a Red or White ball as the first ball. And the fact that later—perhaps after drawing 50 White balls—a Green ball might appear does not matter: we are **not** taking about the conditional *If it is non-White, it is Green* but the simple sentence *The ball is Green*. So:

$$[\textit{The ball is Green}] = \{G\};$$

$$\text{and } P^*([G]) = r^0p = p.$$

So, the event $[\textit{The ball is Green}]$ contains only **one sequence** of length 1 consisting of one Green ball (as elementary events in Ω^* are sequences)! Similarly:

$$[\textit{The ball is Red}] = \{R\}$$

$$\text{and } P^*([R]) = r^0q = q.$$

Now consider *The ball is White*. Obviously, we win when we see a White ball and lose when we see a Green or Red ball. So, the truthmakers for *The ball is White* are sequences beginning with W:

$$[\textit{The ball is White}] = \{W^nG, W^nR: n > 0\}$$

³⁴ Similarly: $[\neg W \rightarrow R] = \{W^nR: n \in \mathbb{N}\}$; $P^*([\neg W \rightarrow R]) = \sum_{n=0}^{\infty} qr^n = \frac{q}{1-r} = \frac{q}{p+q}$ i.e. $P^*([\neg W \rightarrow R]) = P^{DB}(\neg W \rightarrow R) = P(R|\neg W)$. So, $P(R|\neg W) = 1 - P(G|\neg W)$.

³⁵ The argumentation might also be based on the Bernoulli-Stalnaker space (Kaufmann, 2004, 2005, 2009, 2015; van Fraassen 1976), which is slightly more complex, but leads to similar formal results. A feature of such models is that they are construed, generally speaking, in terms of certain relationships between possible worlds – and an important inspiration is Stalnaker’s idea of closest or most similar world (Stalnaker, 1968). Also McGee’s model from (1985) is based directly on Stalnaker’s idea of selection function (on possible worlds); in McGee’s model the probability of the conditional in question is computed by using probability distribution on such selection functions. Khoo and Santorio (2018) give a comprehensive discussion of Stalnaker’s thesis and triviality results. (We thank the anonymous referee for the suggestions).

³⁶ It is associated with the particular conditional $\neg W \rightarrow G$, but the construction for other conditionals (for instance $\neg G \rightarrow R$) is similar.

and $P^*([W]) = r$.

This means that

$$P^*([G]) = p = P(G);$$

$$P^*([R]) = q = P(R);$$

$$P^*([W]) = r = P(W).$$

The probabilities of the simple sentences W, G, R are preserved in $S^* = (\Omega^*, \Sigma^*, P^*)$, as required.

4.3 Heavy and light balls

In order to give an analysis of Lewis’ reasoning, we divided the urn in our example into Heavy and Light balls. So, assume the division of the space $S = (\Omega, \Sigma, P)$ into subsets H and L (which, in the general case, we assume to be independent of the division W, G, R). It is convenient to think of our sample space as consisting of six events, HG, HR, HW, LG, LR, LW, corresponding to six types of balls: Heavy Green, Heavy Red etc. The sample space $S_{HL} = (\Omega_{HL}, \Sigma_{HL}, P_{HL})$ is specified as follows:

$$\Omega_{HL} = \{HG, HR, HW, LG, LR, LW\} \text{ (i.e. there are 6 elementary events);}$$

$$\Sigma_{HL} = 2^{\Omega_{HL}}.$$

For notational convenience, we shall use the following symbols:

$$p_H = P_{HL}(HG); \quad q_H = P_{HL}(HR); \quad r_H = P_{HL}(HW); \quad p_L = P_{HL}(LG); \quad q_L = P_{HL}(LR);$$

$$r_L = P_{HL}(LW).^{37}$$

We shall also use the obvious symbols W, G, R, H, L for White, Green, and Red (regardless of weight) and Heavy and Light (regardless of color), and we set $p = p_H + p_L; q = q_H + q_L; r = r_H + r_L$.³⁸

With this sample space $S_{HL} = (\Omega_{HL}, \Sigma_{HL}, P_{HL})$ we shall associate a probability space $S^* = (\Omega^*, \Sigma^*, P^*)$ ³⁹ which should also allow the situation to be modeled when the additional property (the weight of the balls) is taken into account.

$S^* = (\Omega^*, \Sigma^*, P^*)$ is specified in the following way:

³⁷ In our toy example from Sect. 1, the concrete numerical values are $p_H = 0.3; q_H = 0.1; r_H = 0.1; p_L = 0.1; q_L = 0; r_L = 0.4$.

³⁸ Of course, (1) $P(G) = P(HG) + P(LG) = p_H + p_L = p$; (2) $P(R) = P(HR) + P(LR) = q_H + q_L = q$; (3) $P(W) = P(HW) + P(LW) = r_H + r_L = r$; (4) $P(H) = P(HG) + P(HR) + P(HW) = p_H + q_H + r_H$; (5) $P(L) = P(LG) + P(LR) + P(LW) = p_L + q_L + r_L$.

³⁹ In the remaining part of the chapter, S^* stands for the space in which the balls also have weight, not only color.

$$\Omega^* = \{W^n(HG), W^n(LG), W^n(HR), W^n(LR): n \in \mathbb{N}\};$$

$$\Sigma^* = 2^{\Omega^*}.$$

$W^n(HG)$ is the sequence consisting of n White balls (regardless of weight) and one Heavy Green ball, similarly for $W^n(HR)$ etc. The probabilities of the elementary events are given as follows:

$$P^*(W^n(HG)) = r^n p_H \text{ (for } n \in \mathbb{N}\text{);}$$

$$P^*(W^n(LG)) = r^n p_L \text{ (for } n \in \mathbb{N}\text{);}$$

$$P^*(W^n(HR)) = r^n q_H \text{ (for } n \in \mathbb{N}\text{);}$$

$$P^*(W^n(LR)) = r^n q_L \text{ (for } n \in \mathbb{N}\text{).}$$

Obviously, $[\neg W \rightarrow G] = \{W^n G: n \in \mathbb{N}\}$, where the White balls and the Green ball might be Heavy or Light, and,

$$P^*([\neg W \rightarrow G]) = \sum_{n=0}^{\infty} p r^n = \frac{p}{1-r} = \frac{p}{p+q}.$$

As before, we can interpret the sentences W, R and G in S^* and ascribe probabilities to them. The probabilities of the corresponding events are:

$$P^*([W]) = P_{HL}(W) = r;$$

$$P^*([G]) = P_{HL}(G) = p;$$

$$P^*([R]) = P_{HL}(R) = q.$$

Similarly, we win the bet on *The ball is Heavy* if we draw a Heavy ball (regardless of what happens later). So, *The ball is Heavy* is interpreted as the set of sequences starting with a heavy ball (of any color):

$$[H] = [The\ ball\ is\ Heavy] = \{HR, HG, (HW)W^n(HG), (HW)W^n(LG), (HW)W^n(HR), (HW)W^n(LR): n \in \mathbb{N}\}.$$
⁴⁰

After a short computation we have

$$P^*([H]) = p_H + q_H + r_H = P_{HL}(H).$$

The same applies to $[L] = [The\ ball\ is\ Light]$; its probability is $P^*([L]) = p_L + q_L + r_L = P_{HL}(L)$.

⁴⁰ $(HW)W^n(HG)$ is the sequence consisting of a Heavy White ball followed by n White balls and one Heavy Green ball, similarly for $(HW)W^n(LG)$ etc.

This means that the probabilities of the Boolean part of the language have been preserved in the new space S^* (associated with $S_{H,L}$) as required.⁴¹

4.4 The law of total probability in S^*

According to LTP, for any event Y :

$$P(Y) = P_H(Y) P(H) + P_L(Y) P(L)$$

where $P_H(\cdot)$ and $P_L(\cdot)$ are the probability measures obtained by standard conditionalization on H and L , i.e. $P(\cdot|H)$ and $P(\cdot|L)$.

In Sect. 2 we showed that **in the general case it is not true** that.

$$(DB-LTP-C) P^{DB}(\neg W \rightarrow G) = P^{DB}_H(\neg W \rightarrow G) P^{DB}(H) + P^{DB}_L(\neg W \rightarrow G) P^{DB}(L)$$

even if the inputs are DB-resistant probabilities for $\neg W \rightarrow G$ in the subspaces H and L . However, it is **always true** that.

$$(LTP)^* P^*(\neg W \rightarrow G) = P^*_{[H]}(\neg W \rightarrow G) P^*([H]) + P^*_{[L]}(\neg W \rightarrow G) P^*([L])$$

as it is a simple application of a mathematical theorem (LTP) in the space S^* .

As S^* models the values for P^{DB} in the proper way, observe that the following equalities hold:

$$\begin{aligned} P^{DB}(\neg W \rightarrow G) &= P^*(\neg W \rightarrow G); \\ P^{DB}(H) &= P^*([H]); \\ P^{DB}(L) &= P^*([L]). \end{aligned}$$

So, it cannot be the case that (in addition to these three equalities) the following two equalities also hold in the general case:

$$\begin{aligned} P^{DB}_H(\neg W \rightarrow G) &= P^*_{[H]}(\neg W \rightarrow G); \\ P^{DB}_L(\neg W \rightarrow G) &= P^*_{[L]}(\neg W \rightarrow G). \end{aligned}$$

⁴¹ Observe that the space S can be “immersed” into S^* by treating elementary events as sequences of length 1. It is important to note that one of essential differences between S and S^* consists in the fact, that in the in space S , the two events:

- The **first** element of the sequence is H ;
- **All** elements of the sequence are H ;

coincide – but this phenomenon does not, in general, take place in S^* , as there might be sequences **beginning** with H , but not consisting entirely of H balls.

It will be helpful to remember this fact when analyzing the argumentation in Sect. 5.

Indeed, if they both were true, then the formula (DB-LTP-C) would also hold, which is not the general case. In the next section we will see how this affects Lewis' reasoning.

5 Lewis' dilemma in S^*

Let us examine once again the crucial steps in Lewis' reasoning:

- (6) $P(A \rightarrow C) = P(C|A)$ ⁴²
- (7) $P(A \rightarrow C|B) = P(C|AB)$
- (9) $P(A \rightarrow C|C) = 1$
- (10) $P(A \rightarrow C|\bar{C}) = 0$
- (11) $P(D) = P(D|C) \times P(C) + P(D|\bar{C}) \times P(\bar{C})$.

Under these assumptions, Lewis' corollary is that $P(C|A) = P(C)$, which he considers to be absurd.

A straightforward application of (7) to the conditional $A \rightarrow C$ and the condition $\neg C$ gives $P_{\neg C}(A \rightarrow C) = P_{\neg C}(C|A)$, which is obviously 0. If we apply it to our conditional $\neg W \rightarrow G$, we obtain the result $P_{\neg G}(\neg W \rightarrow G) = 0$. But we have shown in Sect. 2 that if we assume.

$$\begin{aligned} P_{\neg G}^{DB}(\neg W \rightarrow G) &= 0; \\ P_G^{DB}(\neg W \rightarrow G) &= 1 \end{aligned}$$

then "integrating" this information in the form suggested by formula (11) leads to wrong results: in our example we get $P^{DB}(\neg W \rightarrow G) = 0.4$, which is not DB-resistant (and which shows that our beliefs are inconsistent).

There seems to be tension between our findings and the assumptions on which Lewis' proof is based. After having introduced the space S^* , we have tools to offer a better description of the situation.

We think that the key to understanding the problematic steps of Lewis' proof is identifying what $P(A \rightarrow C|B)$ really means: we do not know in which probability space the function P is defined and which events correspond to sentences $A \rightarrow C$ and B . In our example, we are interested in the probability of the conditional $\neg W \rightarrow G$ under the condition H (and its special case when we substitute $\neg G$ for H). The probability space S^* is suited to formalizing the notion of probability of $\neg W \rightarrow G$ in a coherent (DB-resistant) way. So, we will analyze the possible interpretations of conditionalizing the conditional $\neg W \rightarrow G$ on H within S^* . Two of them are most natural:

⁴² To be exact, Lewis assumes that $P_X(A \rightarrow C) = P_X(C | A)$ for any event X , on which P is conditionalized. This is of course a much stronger assumption than just assuming (6) with respect to the given probability measure P !

First interpretation. Conditionalizing $\neg W \rightarrow G$ on H within S^* means considering only games which are played entirely within the Heavy subspace, i.e., formally, sequences consisting **entirely of Heavy balls**. This means that we think only of sequences of the form $(HW)^n(HG)$ and $(HW)^n(HR)$. Similarly, conditionalizing on L means considering only sequences consisting **entirely of Light balls**.

However, in $S^* = (\Omega^*, \Sigma^*, P^*)$ there are not only “pure Heavy” and “pure Light” sequences but also mixed sequences (i.e. in which both Heavy and Light balls occur). Consequently, the universe Ω^* is in the general case divided into **three** subsets, which we shall denote by [Heavy], [Light], [Mixed].⁴³ In this case, the Law of Total Probability in S^* has the form:

$$P^*(\neg W \rightarrow G) = P^*_{[Heavy]}(\neg W \rightarrow G) P^*([Heavy]) + P^*_{[Light]}(\neg W \rightarrow G) P^*([Light]) + P^*_{[Mixed]}(\neg W \rightarrow G) P^*([Mixed]).$$

To make our reasoning analogous to Lewis’ reasoning, we substitute $\neg G$ for H (i.e. [Heavy] = [non-Green]; [Light] = [Green]). In this case, when “translated” to our example, steps (10) and (11) have the following form:

$$(10-1) P^*(\neg W \rightarrow G | [non-Green]) = 0;$$

$$(11-1) P^*(\neg W \rightarrow G) = P^*_{[non-Green]}(\neg W \rightarrow G) P^*([non-Green]) + P^*_{[Green]}(\neg W \rightarrow G) P^*([Green]).$$

(10-1) is true. Indeed, [non-Green] is the set of sequences consisting of non-Green balls only, i.e. $[non-Green] = \{W^nR : n \in \mathbb{N}\}$. Obviously, within this subspace the probability of the conditional $\neg W \rightarrow G$ is 0.

But Ω^* is **not** the union of the two subsets [Green] and [non-Green]. Indeed,

$$[Green] = \{G\};$$

$$[non-Green] = \{W^nR : n \in \mathbb{N}\};$$

which means that we leave out all sequences of the form:

$$\{W^nG : n > 0\}.$$

For this reason, (11-1) is not a proper application of the Law of Total Probability. The term $P^*_{[Mixed]}(\neg W \rightarrow G) P^*([Mixed])$ is missing!

The only way to save **both** the assumptions (10-1) and (11-1) in Lewis’ proof is to assume that the division of Ω^* into [Green] and [non-Green] is a genuine division. This means that there are no mixed sequences in the universe Ω^* , i.e. no sequences of the form W^nG . But this is only possible when there are no White balls, which makes the conditional $\neg W \rightarrow G$ uninteresting.⁴⁴

⁴³ Importantly, apart from trivial cases, [Heavy] is **not** the set $[H]$ in space S^* specified in Sect. 4.3. Similarly [Light] is **not** $[L]$ as $[H]$ and $[L]$ are the sets of sequences **beginning** with a Heavy/Light ball.

⁴⁴ All balls are non-White. So, $\neg W$ is true of all balls, and the conditional $\neg W \rightarrow G$ really means *The ball is Green*.

Second interpretation. The second possibility is that conditionalizing on H within S^* means considering only sequences **beginning with a Heavy ball**. Similarly, conditionalizing on L within S^* means considering only sequences **beginning with a Light ball**. Indeed, this is the interpretation we used in Sect. 4.3.

In this case, we have a genuine division of Ω^* into two disjointed subsets $[H]$ and $[L]$, and the Law of Total Probability holds:

$$(LTP)^* P^*(\neg W \rightarrow G) = P^*_{[H]}(\neg W \rightarrow G) P^*([H]) + P^*_{[L]}(\neg W \rightarrow G) P^*([L]).$$

Under this interpretation, after substituting $\neg G$ for H in our example, steps (10) and (11) will take the form:

$$(10-2) P^*(\neg W \rightarrow G | \neg G) = 0;$$

$$(11-2) P^*(\neg W \rightarrow G) = P^*_{[\neg G]}(\neg W \rightarrow G) P^*([\neg G]) + P^*_{[G]}(\neg W \rightarrow G) P^*([G]).$$

(11-2) is a proper application of (LTP*) (now the division of Ω^* into $[G]$ and $[\neg G]$ is a genuine division) but in this case (10-2) is **not true**. Indeed, for arbitrary H with positive probability:

$$P^*(\neg W \rightarrow G | [H]) = \frac{P^*(\neg W \rightarrow G) \cap [H]}{P^*([H])} = \frac{p_H}{(p_H + q_H + r_H)} + \frac{r_H}{(p_H + q_H + r_H)} \frac{p}{(p+q)}.$$

As H was arbitrary, we can substitute $H = \neg G$. Obviously, in this case, $p_H = p_{\neg G} = P(G \wedge \neg G) = 0$: the chance of finding a ball that is both Green and non-Green is obviously 0, so the first term $\frac{p_H}{(p_H + q_H + r_H)} = 0$.⁴⁵

However, the number $r_H = r_{\neg G}$ is the probability of finding a ball which is both White and non-Green, but this simply amounts to being a White ball and definitely need not be 0! So, the part $\frac{r_H}{(p_H + q_H + r_H)} \frac{p}{(p+q)}$ need not be 0 at all. This means that:

$$P^*(\neg W \rightarrow G | \neg G) = \frac{r_H}{(p_H + q_H + r_H)} \frac{p}{(p+q)} \neq 0.⁴⁶$$

We can only save the equality (10-2) by assuming that $r_H = 0$, i.e. that there are no White balls, i.e. that $\Omega^* = \{G, R\}$. And, as before, this makes the conditional $\neg W \rightarrow G$ not interesting.

⁴⁵ In our toy example with 100W, 80G, 20R, we have $p_{\neg G} = 0$; $q_{\neg G} = 0.1$; $r_{\neg G} = 0.4$; $p_G = 0.5$; $q_G = 0$; $r_G = 0$, which gives the probability $P^*(\neg W \rightarrow G | \neg G) = 0.64$.

⁴⁶ Apart from the computations in $S^* = (\Omega^*, \Sigma^*, P^*)$, there is a simple intuitive argument that $P^*(\neg W \rightarrow G | \neg G)$ need not be 0: this corresponds to the chance of winning the $\neg W \rightarrow G$ game under the assumption that the **first** ball was not Green. It is obviously **not** zero: the first ball might be White and of course you can still win the game!

We have discussed two possible interpretations of conditionalizing the conditional $\neg W \rightarrow G$ on $\neg G$ within the space S^* .⁴⁷ We have also shown that in each of these interpretations, at least one of Lewis' assumptions is not satisfied: either the formula (11) fails (as the division in question is **not** a genuine division of Ω^*), or assumption (10) fails (as the probability in question is **not** equal 0). The only way to save the proof is by assuming, that there are no White balls in the urn—which makes the conditional uninteresting. Anyway, the triviality proof falls apart.

Finally, observe that in our argumentation there is no need to use nested conditionals, as all expressions of the form $P^*([A \rightarrow B] | [C])$ (for instance $P^*([\neg W \rightarrow G] | [\neg G])$) have a mathematically proper interpretation in the probability space. This means in particular, that conditionalization of the sentence $A \rightarrow B$ by C (with A, B, C Boolean) can be mathematically captured in a direct way, without the need to examine the right-nested conditional $C \rightarrow (A \rightarrow B)$. Therefore, as no nested conditionals are used, the problem of the status of the (probabilistic) Import–Export Principle (IE) does not arise in the direct form at all.

However, it is natural to think of right-nested conditionals in this context, as there are interesting connections between PCCP and IE in the context of triviality results. Fitelson (2015) gives an overview of a general form of the triviality results and points out that triviality results can be justified within a rather uncontroversial (background) theory with a form of Import–Export principle added.

The probabilistic Import–Export Principle in the general case need not hold in the Bernoulli–Stalnaker model. However, it holds in the causal random variable model in Kaufmann (2009), and in McGee's model: indeed, McGee assumes the probabilistic version of the Import–Export Principle as axiom (C7) (McGee, 1989, p. 504).⁴⁸ An important paper discussing the status of IE is (McGee, 1985) (the Reagan–Anderson examples), a vivid discussion followed, see for instance Arlo-Costa (2001) or the recent Mandelkern (2020a) (where the Import–Export

⁴⁷ To achieve this aim, we examined two partitions in S^* :

Partition A: S^* is divided into $[H]$ and $[L]$, i.e.: (1) $[H]$ = sequences **beginning** with H; (2) $[L]$ = sequences **beginning** with L.

Partition B: S^* is divided into $[Heavy]$, $[Light]$, $[Mixed]$, i.e.: (1) $[Heavy]$ = sequences **consisting entirely of Heavy balls**; (2) $[Light]$ = sequences **consisting entirely of Light balls**; (3) $[Mixed]$ = **mixed** sequences.

Depending on whether we choose Partition A or Partition B, The Law of Total Probability will have different forms. After substituting $\neg G$ for H, we could see, that in either of these two forms, at least one of Lewis' assumptions is violated.

⁴⁸ The formula for $P(A \rightarrow (B \rightarrow C))$ in the Bernoulli–Stalnaker model is $P^*(A \rightarrow (B \rightarrow C)) = P(BC|A) + P(\neg B|A)P(CB)$. So it is very different from McGee's formula, i.e. $P_{IE}^*(A \rightarrow (B \rightarrow C)) = P(C|AB)$ (they coincide only in special cases). The Bernoulli–Stalnaker formula gives counterintuitive results in some cases, consider the sentence *If the match is wet, then if you strike it, it will light* (Kaufmann, 2005). For this particular sentence, the IE Principle seems natural. However, there are also many examples in which the Import–Export principle is highly counterintuitive, such examples are presented for instance in Mandelkern (2020a, b). Khoo and Mandelkern (2019) discuss the problem in the context of natural languages and argue that natural language intuitions in favor of IE are really intuitions in favor of a related, but slightly different principle (which they call Sentential Import–Export, and which is expressed in terms of identity of sets of possible worlds modeling propositions). So, there is no reason to accept the IE principle as a general rule. We are grateful to the anonymous referee for the recommendation to discuss the IE principle in this context and also for the bibliographic suggestions.

principle in the context of the interpretation of the conjunction is discussed) or Mandelkern (2020b) (the paper discusses the role of the Import–Export principle in the context of the law of identity, identifying in particular diverging intuitions concerning IE for indicative and subjunctive conditionals). These papers focus on the Import–Export Principle as a logical rule, not on the probabilistic version.

Nevertheless, our reasoning is conducted in a way which **does not** involve nested conditionals, which makes it independent from the problem of the status of IE.

6 Summary

1. The diachronic Dutch Book makes it possible to show that if the agents want to extend their system of beliefs (originally modeled within the sample space $S = (\Omega, \Sigma, P)$) so as to include also the conditional $A \rightarrow C$, then the only DB-resistant extension is by setting $P^{\text{DB}}(A \rightarrow C) = P(C|A)$, i.e. in accordance with PCCP. (At this stage we do not have a model or semantics for the conditional—but have identified the proper value of the confidence level.)

2. We are confronted with Lewis' triviality results, which purport to show that mathematical arguments against PCCP can be formulated. Or, put differently: that formulating a mathematical model for probabilities of conditionals leads to absurd results.

3. The straightforward generalization of the Law of Total Probability (LTP) to conditionals has the form (DB-LTP-C) and is an exact counterpart of Lewis' formula (11) from his proof. But this formula does not work: after we divide the sample space into two disjointed events, H, L, and use the **proper** values of $P^{\text{DB}}_{\text{H}}(A \rightarrow C)$; $P^{\text{DB}}_{\text{L}}(A \rightarrow C)$; $P(\text{H})$ and $P(\text{L})$, we obtain diverse **wrong** results, depending on the division H/L. This is important as Lewis' proof uses conditionalization on C and $\neg C$, which leads to the wrong value of $P^{\text{DB}}(A \rightarrow C)$. The proper formula has the form (DB-LTP-C*). We have justified this conclusion using only a Dutch Book analysis without even mentioning any formally defined probability space.

4. For a given probability sample space $S = (\Omega, \Sigma, P)$ and the given conditional $A \rightarrow C$, we have constructed the probability space $S^* = (\Omega^*, \Sigma^*, P^*)$, which is naturally associated with $S = (\Omega, \Sigma, P)$, such that:

- The conditional $A \rightarrow C$ is interpreted as an event $[A \rightarrow C] \subseteq \Omega^*$;
- Ω^* naturally corresponds to possible scenarios that set the truth value of the conditional $A \rightarrow C$;
- The probabilities of the Boolean sentences from the sample space $S = (\Omega, \Sigma, P)$ are preserved in $S^* = (\Omega^*, \Sigma^*, P^*)$.

We have shown that $P^*([A \rightarrow C]) = P^{DB}(A \rightarrow C)$. This means that we can treat the credence assignments as a mathematically well-defined object. We can in particular analyze the assumptions of Lewis’ Triviality proof within $S^* = (\Omega^*, \Sigma^*, P^*)$.⁴⁹

5. Lewis’ argument against PCCP rests on two crucial assumptions: (i) $P_{\neg C}(A \rightarrow C) = 0$; (ii) a version of the Law of Total Probability. It turns out that—apart from very special situations—at most one of them can be true.

6. We have shown that under one of the possible interpretations of the term $P(A \rightarrow C|B)$, the Law of Total Probability **does not** have Lewis’ form

$$P(A \rightarrow C) = P(A \rightarrow C|C) P(C) + P(A \rightarrow C|\neg C) P(\neg C)$$

so assumption (11) in Lewis’ proof does not hold.

7. We have shown that under the second interpretation (which is more natural in the space $S^* = (\Omega^*, \Sigma^*, P^*)$), the Law of Total Probability has the form

$$(LTP)^* P^*([\neg W \rightarrow G]) = P^*_{[G]}([\neg W \rightarrow G]) P^*([G]) + P^*_{[\neg G]}([\neg W \rightarrow G]) P^*([\neg G])$$

which means that Lewis’ assumption (11) holds. However, in this case it is **not true** that $P^*_{[\neg G]}([\neg W \rightarrow G]) = 0$. This means that Lewis’ crucial assumption (10), i.e. $P(A \rightarrow C|\neg C) = 0$, does not hold.

8. Under both discussed interpretations of conditionalizing the conditional $A \rightarrow C$ on $\neg C$, $P(A \rightarrow C|\neg C)$ has both properties which are crucial for Lewis’ proof **only** if we assume that the probability space Ω is trivial, i.e. when $\Omega = \{A, C\}$. This means, that Lewis’ proof does not prove this fact, but rather presupposes it—therefore it is circular.

The presented analysis of Lewis’ Triviality proofs invoking the probability space S^* does not directly use Dutch Book. However, Dutch Book reasoning has two important roles to play in our argumentation: (1) it gives a simple, pedestrian argument for PCCP; (2) it shows that using Lewis’ version of LTP (i.e. (DB-LTP-C), which corresponds directly to assumption (11) in Lewis’ reasoning) will lead to wrong results (and losing against the Bookmaker). This shows that something is wrong with this formula—and the mathematical explanation is provided by the probabilistic model.

⁴⁹ The idea of “lifting” the sample space S to a new “ S^* -like” space is not new: indeed, van Fraassen’s idea of the Bernoulli-Stalnaker space is based on such a sequential construction [see van Fraassen (1976), Kaufmann (2004, 2005, 2009, 2015)]. In Bacon (2015) a mathematically very interesting construction of a similar kind is presented. (However, it uses sequences of worlds of length $\alpha < \omega_1$, i.e. transfinite sequences of any countable length). Wójtowicz and Wójtowicz (2019) present a simpler model based on the theory of Markov chains, in which all the sequences are finite—this is essentially the model presented in this paper. In these models, PCCP (in the suitable formal formulation) holds. In Wójtowicz and Wójtowicz (2021) a model for right-nested conditionals is presented.

In this sense DB shows what boundary conditions have to be met in the model for probabilities of conditionals.

Appendix: Dutch Book for the conditional $A \rightarrow C$

Consider the conditional $A \rightarrow C$. We assume that the Boolean beliefs of the agent are modeled within a sample probability space $S = (\Omega, \Sigma, P)$, and that $P(A) > 0$. This is sufficient to show that the DB-resistant credence of $A \rightarrow C$ is $P(C|A)$. Assume that:

$$\begin{aligned} P(AC) &= p; \\ P(A\neg C) &= q; \end{aligned}$$

(we write $A\neg C$ instead of $A \wedge \neg C$). Obviously:

$$P(A) = P(AC) + P(A\neg C) = p + q.$$

Also, by definition:

$$\begin{aligned} P(C|A) &= P(AC)/P(A) = \frac{p}{p+q} \\ P(\neg C|A) &= P(A\neg C)/P(A) = \frac{q}{p+q} \end{aligned}$$

$$\text{Set } r = 1 - P(A) = 1 - (p + q).$$

Assume that the agent believes that the credence of $A \rightarrow C$ is $p^* > \frac{p}{p+q}$ (i.e. the agent overestimates the credence). The diachronic Dutch Book is defined as follows:

1. The bookmaker buys a $\text{Bet}(\neg A)$ on $\neg A$ for r .
2. The bookmaker sells N bets $\text{Bet}(A \rightarrow C)$ on $A \rightarrow C$ for p^* . (The number N is specified later).
3. The event in question occurs (for instance a ball is drawn).
4. If $\neg A$ is true than the agent loses $\text{Bet}(\neg A)$ and therefore loses $1 - r$. $\text{Bet}(A \rightarrow C)$ has been cancelled. This is the end of the DB-game, and the final outcome is the agent's loss $1 - r$.
5. If A is true then (**before** it is checked whether C is true) the bookmaker sells N bets on $\neg C$ (i.e. $\text{Bet}(\neg C)$) for $\frac{q}{p+q}$ each.
6. Now it is checked whether C or $\neg C$ is true.
7. If C is true, than the agent
 - a. wins $\text{Bet}(A \rightarrow C)$ and therefore wins $N(1 - p^*)$
 - b. loses $\text{Bet}(\neg C)$ and therefore loses $N \frac{p}{p+q}$.
8. If $\neg C$ is true than the agent
 - a. loses $\text{Bet}(A \rightarrow C)$ and therefore loses Np^*
 - b. wins $\text{Bet}(\neg C)$ and therefore wins $N \frac{q}{p+q}$

The agent’s wins and losses in these three cases are summarized in the table (which corresponds to the table for our example in Sect. 1.2):

	Bet($\neg A$)	Bet($A \rightarrow C$)	Bet($\neg C$)	The total outcome
$\neg A$	$-(1-r)$	0 (the bet was cancelled)	0 (the bet has not been proposed at all)	$-(1-r)$
AC	r	$N(1-p^*)$	$-N \frac{q}{p+q}$	$r + N(1-p^* - \frac{q}{p+q})$
$A\neg C$	r	$-Np^*$	$+N \frac{p}{p+q}$	$r + N(\frac{p}{p+q} - p^*)$

Obviously, $-(1-r) < 0$, so if $\neg A$ occurs (and the Bet($A \rightarrow C$) is cancelled) the agent loses.

Of course $1 - \frac{q}{p+q} = \frac{p}{p+q}$. This means, that the outcomes in the cases AC and $A\neg C$ are equal, i.e.:

$$r + N(1 - p^* - \frac{q}{p+q}) = r + N(\frac{p}{p+q} - p^*).$$

N should be a number big enough to make these outcomes negative, i.e. such that $r + N(\frac{p}{p+q} - p^*) < 0$.⁵⁰

This is true, if $r < N(p^* - \frac{p}{p+q})$, i.e. if $N > \frac{r}{p^* - \frac{p}{p+q}}$.

As the agent is willing to buy/sell an arbitrary number N of bets considered to be fair, the bookmaker can construct a Dutch Book.

For the case $p^* < \frac{p}{p+q}$ (i.e. the agent underestimates the credence of the conditional $A \rightarrow C$) the construction is symmetrical.

Acknowledgements The preparation of this paper was supported by the National Science Centre (NCN) grant 2016/21/B/HS1/01955. We would like to thank the anonymous referees whose detailed and insightful comments helped us to improve presentation of the paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

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*. D. Reidel.

⁵⁰ Intuitively, the bookmaker has to make use of the incoherence in the agent’s beliefs and to compensate for the possible loss on Bet(A). If the incoherence (i.e. the agent’s over/underestimation of the credence of $A \rightarrow C$) is small, than the number N needs to be large.

- Adams, E. W. (1998). *A primer of probability logic*. CLSI, Stanford University.
- Arló-Costa, H. (2001). Bayesian epistemology and epistemic conditionals: On the status of the Export-Import laws. *The Journal of Philosophy*, 98(11), 555–593
- Bacon, A. (2015). Stalnaker's thesis in context. *The Review of Symbolic Logic*, 8(1), 131–163
- Douven, I. (2016). On de Finetti on Iterated Conditionals. In C. Beierle, G. Brewka, & M. Thimm (Eds.), *Computational models of rationality*. (pp. 265–279). College Publications.
- Easwaran, K. (2011a). Bayesianism I: Introduction and Arguments in Favor. *Philosophy Compass*, 6, 312–320
- Easwaran, K. (2011b). Bayesianism II: Criticisms and Applications. *Philosophy Compass*, 6, 321–332
- Edgington, D. (1995). On conditionals. *Mind*, 104, 235–329
- Égre P., Rossi L. & Sprenger J. (2019). De Finettian Logics of Indicative Conditionals. <https://arxiv.org/pdf/1901.10266.pdf> (forthcoming in *Journal of Philosophical Logic*)
- Fitelson, B. (2015). The strongest possible Lewisian Triviality Result. *Thought*, 4, 69–74.
- 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 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
- Hájek, A., & Hall, N. (1994). The hypothesis of the conditional construal of conditional probability. In E. Eells & B. Skyrms (Eds.), *Probabilities and conditionals: Belief revision and rational decision*. (pp. 75–110). Cambridge University Press.
- Kaufmann, S. (2004). Conditioning against the grain: abduction and indicative conditionals. *Journal of Philosophical Logic*, 33(6), 583–606
- Kaufmann, S. (2005). Conditional predictions: A probabilistic account. *Linguistics and Philosophy*, 28(2), 181–231
- Kaufmann, S. (2009). Conditionals right and left: probabilities for the whole family. *Journal of Philosophical Logic*, 38, 1–53
- Kaufmann, S. (2015). Conditionals, conditional probability, and conditionalization. In H. Zeevat & H. C. Schmitz (Eds.), *Bayesian natural language semantics and pragmatics*. (pp. 71–94). Berlin: Springer.
- 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. <http://paolo.santorio.net/teaching.html>.
- Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. *Philosophical Review*, 85, 297–315
- Lewis, D. (1986). *Philosophical papers*. (Vol. 2) Oxford University Press.
- Lewis, D. (1999). Why conditionalize? In: *Papers in metaphysics and epistemology*. Cambridge: Cambridge University Press.
- Mandelkern, M. (2020a). Import-export and ‘and.’ *Philosophy and Phenomenological Research*, 100(1), 118–135
- Mandelkern, M. (2020b). A counterexample to modus ponens. *The Journal of Philosophy*, 117(6), 315–331
- 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*, XCVIII, 4, 485–541
- Rescorla, M. (forthcoming). An Improved Dutch Book Theorem for Conditionalization. *Erkenntnis*: 1–29. <https://doi.org/10.1007/s10670-020-00228-1>
- Stalnaker, R. (1968). A theory of conditionals. *Studies in Logical Theory, American Philosophical Quarterly, Monograph*, 2, 98–112
- Stalnaker, R. (2009). Conditional propositions and conditional assertions. In A. Egan & B. Weatherston (Eds.), *Epistemic modality*. (pp. 227–248). Oxford University Press.
- 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. The University of Western Ontario Series in Philosophy of Science*. (Vol. 1, pp. 261–308). D. Reidel.
- Vineberg, S. (2016). "Dutch Book Arguments", *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/spr2016/entries/dutch-book/>.
- Wójtcowicz K., & Wójtcowicz A. (forthcoming). A stochastic graphs semantics for conditionals. *Erkenntnis*, 1–35. <https://doi.org/10.1007/s10670-019-00144-z>

Wójtowicz K., Wójtowicz A. (forthcoming). A graph model for probabilities of nested conditionals. *Linguistics and Philosophy*. <https://doi.org/10.1007/s10988-021-09324-z>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.