A Stochastic Graphs Semantics for Conditionals

Krzysztof Wójtowicz · Anna Wójtowicz

Received: 18 August 2018 / Accepted: 28 June 2019 / Published online: 22 November 2019
© The Author(s) 2019

Abstract
We define a semantics for conditionals in terms of stochastic graphs which gives a straightforward and simple method of evaluating the probabilities of conditionals. It seems to be a good and useful method in the cases already discussed in the literature, and it can easily be extended to cover more complex situations. In particular, it allows us to describe several possible interpretations of the conditional (the global and the local interpretation, and generalizations of them) and to formalize some intuitively valid but formally incorrect considerations concerning the probabilities of conditionals under these two interpretations. It also yields a powerful method of handling more complex issues (such as nested conditionals). The stochastic graph semantics provides a satisfactory answer to Lewis’s arguments against the PC = CP principle, and defends important intuitions which connect the notion of probability of a conditional with the (standard) notion of conditional probability. It also illustrates the general problem of finding formal explications of philosophically important notions and applying mathematical methods in analyzing philosophical issues.

The problem of estimating the probabilities of conditional sentences is interesting for at least two reasons:

First, it seems that in many cases it is not clear which estimation is “The True Estimation” (as different language users give different values).

Second, providing a coherent and robust method of evaluating these probabilities would give us a better understanding of many other problems connected with conditionals—e.g. the problem of counterfactuals.
In this article, we define a semantics for conditionals in terms of stochastic graphs, which gives a straightforward and simple method of evaluating the probabilities of conditionals.\(^1\) It seems to be a good and useful method in the cases already discussed in the literature (cf. e.g. Kaufmann (2004)), and it can easily be extended to cover more complex situations. In particular, it allows us to describe several possible interpretations of the conditional (the global and the local interpretation, and generalizations of them), and yields a powerful method of handling more complex issues (such as nested conditionals\(^2\)). It provides a satisfactory answer to Lewis’s arguments against the PC = CP principle, and defends important intuitions which connect the notion of probability of a conditional with the (standard) notion of conditional probability.

The article has the following structure:

In Part 1 we analyze a simple example, which shows how to define the stochastic graph semantics for conditionals.

In Part 2 we demonstrate how the notions introduced in Part 1 can be used to analyze more complex problems (we examine the case of two urns, which makes the situation non-trivial).

Parts 3 and 4 are devoted to the global and local interpretations of conditionals. They provide a formal framework suitable for their analysis (and we formalize the intuitive computations given by Kaufmann for the global and local interpretations).

In Part 5 we discuss Lewis’s arguments (concerning the PC = CP principle) and show, how to resolve within our framework the problems indicated by Lewis.

A short summary (part 6) follows, in which we also discuss possible generalizations and indicate some results that we have obtained already (but which lie outside the scope of this article).

The Appendix presents some technical tools which are needed in the text.

A “byproduct” of our results is a contribution to the vivid discussion concerning the explanatory (explicatory) role of mathematics. Notions from probability theory (and the theory of Markov processes) have been used in order to give a precise explication of some philosophically important notions. So apart from their interest to the problem of conditionals, our results exemplify the general strategy of finding formal counterparts (explications) of philosophical notions. The notion of probability is a classic example of an explication (the \textit{locus classicus} is Carnap (1950), cf. e.g. Brun (2016) for a contemporary discussion). In a sense, we widen the scope of an explication and show how to give formal, precise explications of the probability of conditionals, taking different interpretations and variants into account.

---

\(^1\) We define the appropriate probabilities, but the model also uses the notion of truth-conditions (expressed in terms of the rules of a game). However, we do not discuss the topic here, as the aim of the article is to present the formal model and to discuss the interpretation of the conditional (and its impact on the discussion about PC = CP).

\(^2\) The problem of nested conditionals is discussed in Wójtowicz and Wójtowicz (2019a).
1 A Simple One-Urn Example

Very often in discussions concerning conditionals (and in particular, when their probabilities are under discussion), problems from natural language are analyzed. For example, suppose we try to estimate the probability that If Reagan worked for the KGB, I’ll never find out (Lewis 1986, 155). It is typically the case that our intuitions drive the analysis, but it is difficult (indeed, impossible) to give a concrete value of this probability and it is often even impossible to determine whether different language users apply the same methods of estimating it. For these reasons, in this article (like in many other works on conditionals, e.g. Kaufmann (2004, 2005, 2009, 2015), Khoo (2016), van Fraassen (1976)) we will investigate examples where the probabilities of conditional sentences in question can be given concrete values. This precise semantics can later be applied to everyday-life examples.

Let us start with a simple one-urn example. A ball is drawn from the urn containing 10 White, 8 Green and 2 Red balls. What probability is a rational subject going to ascribe to the sentence A Green ball has been drawn? It is of course 8/20, which is exactly the probability computed within the classical probability model. It is also equal to the expected value in the game, where the reward is 1 for winning the game (i.e. drawing the Green ball), and 0 for losing it. This is in accord with the intuition that the subjective probability of a sentence is the odds at which you would be willing to place a bet on the outcome or the price you are willing to pay for the game. So, for the same reasons, the sentence A non-White and Green ball has been drawn has probability 8/20; the sentence A White ball has been drawn has probability 10/20 etc. These are the prices to pay for the respective (fair) games.

A rational subject estimates the probabilities (under the threat of a Dutch Book) so as to conform to the Kolmogorov probability axioms. In other words, the subject works within the probability space $$\Xi = (\Omega, M, P)$$, where $$\Omega = \{W, G, R\}$$ are elementary events (i.e. drawing a ball of one of the colors), with probabilities respectively: $$P(W) = r = 10/20$$; $$P(G) = p = 8/20$$; $$P(R) = q = 2/20$$. M is the $$\sigma$$-field of all events, and P is the probability function defined on M.

Let us estimate the probability of the sentence If we draw a non-White ball, we draw a Green ball (or, simpler: If the ball is not White, it is Green). We symbolize it as $$\neg W \rightarrow G$$.

As we have observed before, estimating this probability amounts to estimating the price for playing a certain (fair) game, which we might call “The Conditional Game”. The problem is therefore: how much are we willing to bet on the truth of the

---

3 In order to maintain consistency with the upcoming formalisation, we shall capitalise the three colour adjectives henceforth.
4 This is a much-discussed issue, but we will not go into details here.
5 M is formally very simple: $$M = 2^\Omega$$ (i.e. it is the power set of $$\Omega$$). Of course, we could also use as the model a (more complicated) 20-element space (with 10 White, 8 Green, 2 Red balls, all equally probable), but we prefer the simpler—and more general—interpretation.
conditional \( \neg W \rightarrow G \)? Before we bet, we have to define the game conditions in a precise manner, i.e. we have to stipulate in a clear way how the game should be played and in what circumstances we will consider it as settled. In particular, we have to determine the result for all possible draws (i.e. White, Green, Red—W, G, R).

There is no doubt that we win if we draw a Green ball, and we lose if we draw a Red ball. But the problem arises with what should we do when we draw a White ball.

There are a priori four possibilities. After drawing a White ball:

1. We win the game.
2. We lose the game.
3. The game is undecided and stops.
4. The game is undecided, we put the ball into the urn and draw the ball again.

Choosing option (1) is equivalent to the claim, that we do not really play The Conditional Game \( \neg W \rightarrow G \), but rather The Material Implication Game: we win always when the ball is White or Green.\(^6\) The probability of this event is clearly 18/20. If we chose (2), it means, that we lose after drawing a White or a Red ball and the probability of this event is 12/20. But it is widely agreed that this is not the appropriate way of interpreting conditionals (plenty of discussion can be found in the literature, so there is no need to consider the matter further here).

Choosing (3) (the game stops as undecided) has a major drawback: it “annihilates“ the problem of counterfactuals. So this is not an attractive solution.\(^7\) The conclusion therefore is that the only reasonable definition of The Conditional Game \( \neg W \rightarrow G \) is (4): drawing a White ball does not end the game but extends it: we put the ball back into the urn, and draw again. This choice is a natural one: our bet concerns the case when the ball is not White, and in order to resolve the situation, we have to check what happens if the ball is non-White! If we stipulate the rules of our game in this way, then the outcomes are not single draws, but rather sequences of draws, consisting of a (possibly empty) sequence of White balls (which simply have the effect that the game restarts), followed by a Green or Red ball (either of which settles the game).\(^8\)

---

\(^6\) This reflects the simple tautology of the propositional calculus: \((\neg p \rightarrow q) \leftrightarrow (p \lor q)\). But it is widely agreed that the interpretation of a conditional as the material implication is not appropriate.

\(^7\) We admit, that this choice might be natural if we do not admit „restarting actuality“, and in particular if we do not want to think in terms of possible worlds. If we think of evaluating the probability or truth value of the conditional in terms of the actual world only, without allowing possible worlds (or alternative scenarios) etc. our approach can be considered problematic. However, it is quite standard for other approaches, e.g. of van Fraassen: Imagine the possible worlds are balls in an urn, and someone selects a ball, replaces it, selects again and so forth. We say he is carrying out a series of Bernoulli trials (van Fraassen 1976, 279).

\(^8\) Examples of sequences which settle the game are: WWWG (we win with the fourth move); R (we lose with the first move); WWWWWW (we lose with the sixth move) etc.
This approach to the problem of computing the probability of a conditional means, that we have to define formally a new probability space (different from the simple initial space \( \Xi = (\Omega, M, P) \), but of course intimately connected with it), where computing the probability of the outcome of the game will be possible. Observe, that in our initial space \( \Xi = (\Omega, M, P) \), there is no event corresponding to the conditional \( \neg W \rightarrow G \).\(^9\)

1.1 A Stochastic Graph Representation

We can represent the flow of the game in a very natural and convenient way by means of a stochastic (Markov) graph. We will present the general idea of a Markov graph using a simple example (for the reader’s convenience, technical details and references can be found in the Appendix).

Gambler’s ruin. Consider a coin-flipping game. At the beginning of the game, two players each have a number of pennies, say \( n_1 \) and \( n_2 \). After each flip of the coin, the loser transfers one penny to the winner and the game restarts on the same terms. The game ends when the first player (“The Gambler”) has either lost all his pennies or has collected all the pennies. Assume, that the probability of tossing heads is \( p \), and of tossing tails is \( q \) (\( p + q = 1 \)). The Gambler gets one penny if heads show up.

The simplest non-trivial example is when \( n_1 = 1 \) and \( n_2 = 2 \) (when \( n_1 = n_2 = 1 \), the game stops after the first flip). At each moment, the state of the game is the number of The Gambler’s pennies, so the possible states are 0, 1, 2, 3. The game starts in state 1 and the transitions between states occur as a consequence of one of two actions: heads or tails (H,T). The game finishes when the gambler enters either state 0 or 3—both 0 (loss) and 3 (victory) are absorbing states.

The dynamics of the game is represented in Graph 1 (Fig. 1, next page).

We can naturally track the possible scenarios of the game (i.e. ending either with victory or with loss) as possible paths in the graph. For instance, the path HH (tossing heads twice) leads to victory and the corresponding transitions between the states are 1 \( \rightarrow \) 2 \( \rightarrow \) 3.

There are infinitely many paths (scenarios) that lead to victory: HTHH (1 \( \rightarrow \) 2 \( \rightarrow \) 1 \( \rightarrow \) 2 \( \rightarrow \) 3); HHTHHH (1 \( \rightarrow \) 2 \( \rightarrow \) 1 \( \rightarrow \) 2 \( \rightarrow \) 1 \( \rightarrow \) 2 \( \rightarrow \) 3); HT…HTHH (1 \( \rightarrow \) 2 \( \rightarrow \) … \( \rightarrow \) 2 \( \rightarrow \) 1 \( \rightarrow \) 2 \( \rightarrow \) 3), etc.

The possible paths for losing the game are: T (1 \( \rightarrow \) 0); HTT (1 \( \rightarrow \) 2 \( \rightarrow \) 1 \( \rightarrow \) 0); HT…HTT (1 \( \rightarrow \) 2 \( \rightarrow \) … \( \rightarrow \) 2 \( \rightarrow \) 1 \( \rightarrow \) 0), etc.

So, the space of all possible paths that settle the game consists of sequences \( \{ (HT)^nHH, (HT)^nT: n \in \mathbb{N} \} \) \((HT)^n\) means the sequence HT repeated \( n \) times, i.e. HT…HT). They will form the set of elementary events \( \Omega^* \) in a new probability space that describes all the scenarios. It is straightforward to compute their probabilities. If \( P(H) = p \); \( P(T) = q \), then

---

\(^9\) In the space \( \Omega = \{ W, R, G \} \) there are 8 events (as \( M = 2^\Omega \)), which correspond to Boolean combinations of the elementary events. But none of these events corresponds to the conditional \( \neg W \rightarrow G \).
The probability of winning the game is
\[ \sum_{n=0}^{\infty} (pq)^n p^2 = \frac{p^2}{1 - pq} \]

For \( p = 1/2 \) (a fair coin) it is 1/3.

The probability can be computed in a more direct way by examining the graph and writing down a **simple system of equations**. Let \( P(n) \) denote the probability of winning the game, which starts in state \( n \). Obviously, \( P(0) = 0 \) (if we are in state 0, we have already lost), and \( P(3) = 1 \) (if we are in state 3, we have already won). What are the probabilities \( P(1), P(2) \)? We reason in an intuitive way: being in state \( n \), we can either:

- toss heads (with probability \( p \)) — then we are transferred to the state \( n + 1 \), and our chance of winning is \( P(n + 1) \);
- toss tails (with probability \( q \)) — then we are transferred to the state \( n − 1 \), and our chance of winning is \( P(n − 1) \).

So, \( P(n) = pP(n + 1) + qP(n − 1) \). In our example this means that

\[
\begin{align*}
P(0) &= 0 \\
P(1) &= pP(2) + qP(0) \\
P(2) &= pP(3) + qP(1) \\
P(3) &= 1
\end{align*}
\]

The solution is \( P(1) = \frac{p^2}{1 - pq} \), which is not a surprise. As our example is very simple, the difference in the complexity of computations (i.e. within the probability

---

Fig. 1 (Graph 1): Gambler’s ruin
A Stochastic Graphs Semantics for Conditionals

A crucial feature of the process is its memorylessness (Markov property): the coin does not remember the history of tosses. Moreover, the probabilities of actions (H, T) are fixed throughout the whole history (homogeneity). Generally speaking, the future depends only on the present state of the process, not on the past. This applies to our ball-drawing game: anytime we restart the game, the probabilities remain the same as history does not matter. If our Conditional Game concerns the conditional $\alpha$ (e.g. $\neg W \rightarrow G$), then $G_{\alpha}$ denotes the corresponding graph. We might say that the appropriate class of graphs defines the semantics for the conditionals. In general, there are many possibilities of representing a certain stochastic process by Markov graphs; usually we are looking for the simplest representation.

The dynamics of the game is represented in Graph 2 (Fig. 2, next page).

It consists of three states: START, WIN, LOSE. According to our definition, the game proceeds as follows:

We start (obviously) in START. Three moves are possible:

- If we draw a White ball, we get back to START (and the game restarts);
- If we draw a Green ball, we go to WIN;
- If we draw a Red ball, we go to LOSE.

WIN and LOSE are absorbing states: if we get into them, we remain there and the game stops.

The transition probabilities of getting from one state to another are given by (we use the standard notation here)\(^{11}\):

\[
\begin{align*}
P_{\text{START, START}} &= P(W) = r = 10/20 \\
P_{\text{START, WIN}} &= P(G) = p = 8/20 \\
P_{\text{START, LOSS}} &= P(R) = q = 2/20 \\
P_{\text{LOSS, LOSS}} &= 1 \\
P_{\text{WIN, WIN}} &= 1.
\end{align*}
\]

---

\(^{10}\) In the general case, the gambler’s ruin process is defined for arbitrary $N$, which makes the process much more complex (think of describing the possible paths, for instance when $N = 10$ and $n_1 = 1$!). The gambler’s ruin process is a particular case of a more general random walk process; for a survey of applications see for instance chapter 4.5.1. of Ross (2007).

\(^{11}\) In the general case, when there are $n$ possible states $s_i$, for $i = 1, \ldots n$;

- $p_{ij}$ denotes the transition probability from the state $s_i$ to the state $s_j$;
- $P(i)$ denotes the probability of finally (i.e. possibly after a sequence of steps) reaching the winning state starting from the state $s_i$.

The states in this article have their names, so we use symbols like $p_{\text{START,WIN}}$ or $P(\text{START})$. 
Fig. 2 (Graph 2): the dynamics in the one urn case

(All the other transition probabilities are equal to 0).\(^\text{12}\)

This is the simplest (and standard) representation of this random process as a Markov graph. The labels for the states indicate the current state of the game (i.e. whether we have won, lost, or perhaps the game (re)starts), and the edges are labeled by the drawn balls (they correspond to the actions: a White/Green/Red ball was drawn from the urn).

A big advantage of such a representation is that it shows the possible scenarios in the game in a straightforward way: we start in START, and we simply travel through the graph until we get into one of the absorbing (deciding) states WIN, LOSS. We can also compute the probability of winning (losing) the game in a very simple way. The graph representation allows us to understand the dynamics of the game—and this will be particularly important in more complex situations, when we distinguish between the local and global interpretation of conditionals (cf. Kaufman 2004).

Let \( P(\text{START}) \) denote the (as yet unknown) probability of winning the game (i.e. the probability of reaching WIN starting from START). We write down the respective equation for this simple graph\(^\text{13}\):

\[
P(\text{START}) = p + rP(\text{START})
\]

We solve it, obtaining:

\[
(1 - r)P(\text{START}) = p
\]

\(^{12}\) This is the transition probability matrix for the graph 2 (Fig. 2):

|       | START | WIN | LOSS |
|-------|-------|-----|------|
| START | \( r \) | \( p \) | \( q \) |
| WIN   | 0     | 1   | 0    |
| LOSS  | 0     | 0   | 1    |

\(^{13}\) The general theory is presented in the Appendix, but we can also justify the equation in an intuitive way. Assume that we are in the state START. What are our chances \( P(\text{START}) \) of winning? There are two ways of winning: (1) we can win in just one move, with probability \( p \); (2) we will be in the starting position again, with probability \( r \), but then the game is restarted, so the probability of winning is \( P(\text{START}) \). So, the “contributions” of these variants to \( P(\text{START}) \) are (1) \( p \) and (2) \( rP(\text{START}) \), which gives the equation \( P(\text{START}) = p + rP(\text{START}) \).
So finally:

**Fact 1.1**

\[
P(\text{START}) = \frac{p}{1-r} = \frac{p}{p+q}
\]

So the probability \(P(\neg W \rightarrow G)\) is \(\frac{p}{p+q}\). In our particular case, it is 8/10 (which is also the intuitive answer).

### 1.2 The Probability Space

There is a unique probabilistic space corresponding to the graph \(G_\alpha\) (where \(\alpha = \neg W \rightarrow G\)) that describes the possible “travel histories” (or paths) in this graph.\(^{14}\)

The possible histories are of two kinds: \(W^nG\) (we win the game), and \(W^nR\) (we lose the game), where \(n \in \mathbb{N}\). These paths will serve as elementary events in the constructed space.

The probability space we construct will therefore satisfy the following conditions:

i. It ascribes a certain probability to every possible course of The Conditional Game \(\neg W \rightarrow G\) which settles the game. In particular, it will be possible to compute the probability of our victory (and loss).

ii. It takes into account the initial probabilities of drawing a White, Green or Red ball from the urn, i.e. its definition reflects the probabilities from the space \(\Xi = (\Omega, \mathcal{M}, P)\).

iii. As the aim is to compute the probability of a particular conditional \(\alpha\) (e.g. \(\alpha = \neg W \rightarrow G\)), it is important to make our space as convenient as possible for this particular task (in particular, it is a minimal space).\(^{15}\)

Let \(\Xi = (\Omega, \mathcal{M}, P)\) be a probability space, where \(\Omega\) is the set of elementary events: \(\Omega = \{W, G, R\}\) with probabilities respectively \(P(W) = r, P(G) = p, P(R) = q\); \(\mathcal{M}\) is the \(\sigma\)-field of all events, and \(P\) is the probability function defined on \(\mathcal{M}\).

**Definition 1.2** The probability space \(\Xi_\alpha\) corresponding to the conditional \(\alpha = \neg W \rightarrow G\) is the triple \(\Xi_\alpha = (\Omega_\alpha, \mathcal{M}_\alpha, P_\alpha)\), where:

1. \(\Omega_\alpha\) is the set of elementary events corresponding to sequences of events from the space \(\Xi\), which decide The Conditional Game \(\alpha\);
2. \(\mathcal{M}_\alpha\) is the \(\sigma\)-field of all events;
3. \(P_\alpha\) is the probability function defined on \(\mathcal{M}_\alpha\).

\(^{14}\) We can think of them as paths within the graph which end in one of the states WIN or LOSE. Each of these paths has a definite probability which is computed in a very intuitive way: as the draws are independent of each other, we simply multiply the probabilities; for instance, the probability of \(WWWG\) is \(rrrp\).

\(^{15}\) Observe, that our approach is different in this respect from Kaufmann’s approach, which is based on Stalnaker Bernoulli models. Kaufmann presents the construction of a universal probability space, where all possible conditionals (based on the initial space \(\Omega\)) can be represented. However, formal complications are the price to pay for the generality of this model.
Formally:

\[
\Omega_\alpha = \{W^nG, W^nR: n \in \mathbb{N}\}. \tag{16}
\]

\[
M_\alpha = 2^{\Omega_\alpha},
\]

\[
P_\alpha(W^nG) = r^n p \ (\text{for} \ n \in \mathbb{N});
\]

\[
P_\alpha(W^nR) = r^n q \ (\text{for} \ n \in \mathbb{N}).
\]

Observe that our probability space was defined with respect to a particular conditional \(\alpha: \neg W \rightarrow G\), which is not expressible in the language corresponding to the initial probability space \(\Xi\). The definition is formally correct; in particular, the condition \(P_\alpha(\Omega_\alpha) = 1\) is fulfilled (which can be checked by an easy computation\(^{17}\)). \(\Omega_\alpha\) is a countable set, so (due to \(\sigma\)-additivity), the probability measure \(P_\alpha\) extends to \(M_\alpha\), i.e. \(P_\alpha\) is defined for any set \(A \subseteq \Omega_\alpha\):

\[
P_\alpha(A) = \sum_{\omega \in A} P_\alpha(\omega)
\]

It should be noted that, from a formal point of view, speaking of the probability of the sentence \(\neg W \rightarrow G\) is a kind of abuse of language, because de facto we compute the probability \(P_\alpha\) of the event \([\neg W \rightarrow G]_\alpha\), which is the interpretation of the sentence \(\neg W \rightarrow G\) in the probability space \(\Xi_\alpha\). But we will use this expression freely, as it does not lead to misunderstanding (and in the literature it is standard to speak of the probability of sentences). But we stress here that there is always an interpretation of the sentence defining the game in an appropriate probability space: this interpretation makes clear the sense in which we talk of the probability and how we defined the game (i.e. how we defined the conditions for our victory). Being meticulous in identifying the set of events in the concrete probability space, which is the interpretation of the sentence \(\neg W \rightarrow G\), will be of particular importance when the context (i.e. the rules of the game) changes (we will discuss it later). In such situations, the interpretation of the sentence \(\neg W \rightarrow G\) will differ in different probability spaces, corresponding e.g. to different understandings of the conditional (local and global).

We want to compute the probability of the sentence \(\neg W \rightarrow G\). It has its counterpart \(\text{WIN}_\alpha\) in the probability space \(\Xi_\alpha = (\Omega_\alpha, M_\alpha, P_\alpha)\), consisting of the possible scenarios leading to victory:

\[
\text{WIN}_\alpha = [\neg W \rightarrow G]_\alpha = \{W^nG: n \in \mathbb{N}\},
\]

by \(\text{LOSS}_\alpha\) we denote the event leading to our defeat in the game:

\[
\text{LOSS}_\alpha = [\neg W \rightarrow R]_\alpha = \{W^nR: n \in \mathbb{N}\}.
\]

\(^{16}\) The probability of drawing an infinite sequence of White balls (treated as a sequence of events from the initial probability space \(\Xi\)) is 0 and so we do not need to include it in our definition of \(\Xi_\alpha\).

\(^{17}\) \(P_\alpha(\Omega_\alpha) = P_\alpha(\{W^nG, W^nR: n \in \mathbb{N}\}) = P_\alpha(\{W^nG: n \in \mathbb{N}\}) + P_\alpha(\{W^nR: n \in \mathbb{N}\}) = \sum_{n=0}^{\infty} pr^n + \sum_{n=0}^{\infty} qr^n = \frac{p}{1-r} + \frac{q}{1-r} = 1.\)
Fact 1.3
\( P_\alpha(\text{WIN}_\alpha) = \sum_{n=0}^{\infty} pr^n = \frac{p}{1-r} = \frac{p}{p+q} \)

Fact 1.4
\( P_\alpha(\text{LOSS}_\alpha) = \sum_{n=0}^{\infty} qr^n = \frac{q}{1-r} = \frac{q}{p+q} \)

Fact 1.5
\( P_\alpha(\text{WIN}_\alpha) + P_\alpha(\text{LOSS}_\alpha) = \frac{p}{p+q} + \frac{q}{p+q} = 1 \)

Obviously, the following two facts are true:

Fact 1.6
\( P_\alpha(\text{WIN}_\alpha) = P(\text{G|\neg W}) \)

Fact 1.7
\( P_\alpha(\text{LOSS}_\alpha) = P(\text{R|\neg W}) \)

where \( P \) is the probability measure in the initial probability space \( \Xi = (\Omega, \mathcal{M}, P) \), which was the basis for our construction of the space (of possible game scenarios) \( \Xi_\alpha = (\Omega_\alpha, \mathcal{M}_\alpha, P_\alpha) \).

To sum up: we have given a semantics for conditional sentences \( \alpha \) (consisting of atomic sentences in the initial probability space), which allows us to compute their probabilities in a straightforward way. This semantics is given by an appropriate graph \( G_\alpha \) and the corresponding (unique) probability space \( \Xi_\alpha \). The correlation between the values of the probability \( P_\alpha \) of the conditional and the conditional probability \( P \) in the space \( \Xi \) is given by the Facts 1.6 and 1.7.\textsuperscript{18}

1.3 A Comparison with Kaufmann’s (Stalnaker Bernoulli) Space

A similar approach to the problem of computing the probability of conditionals is presented in Kaufmann (2004, 2005, 2009, 2015). It is based on the Stalnaker Bernoulli model (cf. also van Fraassen (1976); this terminology is used there and in Kaufmann’s papers). Here we indicate the most important similarities and differences between this model and our stochastic graphs semantics.

For the sake of this comparison, consider Kaufmann’s construction of the probability space for the conditional \( \neg W \rightarrow G \). Using a notation similar to ours, the appropriate space would have the form:

\( \Xi^* = (\Omega^*, \mathcal{M}^*, P^*) \), where
\( \Omega^* \) is the set of infinite sequences consisting of the events W, G, R (Kaufmann speaks of worlds being elements of these sequences);
\( \mathcal{M}^* \) is the \( \sigma \)-field defined over \( \Omega^* \);
\( P^* \) is a probability measure defined on \( \mathcal{M}^* \) in an appropriate way.

\textsuperscript{18} By the way, observe, that this semantics allows us also to compute the probabilities of sentences like I will win The Conditional Game within an even number of moves, which is a nice feature of the model.
An important feature of the definition of $P^*$ is the fact, that any particular (infinite) sequence of balls has null probability (we consider it to be a sequence of events coming from the space $\Xi$), but the appropriate “bunches” of such sequences have non-zero probabilities. Considering the sentence $\neg W \rightarrow G$, such bunches are constructed from sequences starting with an initial finite (perhaps empty) sequence of White balls, followed by a Green or Red ball, after which an arbitrary infinite sequence of balls follows.

In this way, Kaufmann is able to give a method of computing the probability of victory in The Conditional Game, i.e. $P^*(\neg W \rightarrow G)^*$. It corresponds exactly to our intuitions: this computation amounts to estimating the fraction of sequences starting with $W^nG$ (i.e. the winning sequences) within all the sequences settling the game (i.e. starting with $W^nG$ or $W^nR$). And of course, this value is exactly the same as the value computed here.

Specific features of Kaufmann’s construction become even more clearly visible when we try to give it a graph-semantical presentation. First, Kaufmann’s method leads to all possible graphs at once: every such graph corresponds to a possible conditional in the language in question. In our graph (and the corresponding probability space) we can only play the game $\neg W \rightarrow G$ (and—after an obvious relabeling WIN/LOSS—also the “dual game” $\neg W \rightarrow R$). Secondly, in our approach the game terminates after entering WIN or LOSS (so the sequences, being elementary events in the probabilistic space $\Xi_\alpha$ are finite); but under Kaufmann’s approach, the game (formally speaking) lasts infinitely long. Both these approaches lead to the same results. The main difference lies in a certain redundancy appearing in Kaufmann’s model, which might be important when more complicated cases are discussed.

To sum up, the advantage of stochastic-graph semantics is its intuitiveness (it is connected with the particular conditional in a natural way) and it gives us an easy method of computing the probability of the conditional. It is formally correct (there is a corresponding formally defined probability space), but this semantics reveals its full power in more complicated cases in which the use of Kaufmann’s semantics might face technical complications of a combinatorial character (the probability space might turn out to be very intricate). We suppose that in some cases these technical complications might even make the treatment impossible.

19 Usually, when we speak of fractions, we have finite sets in mind—but the sets here are infinite (and even uncountable). So perhaps the word “proportion” would be more appropriate.

20 Kaufmann’s model using Stalnaker Bernoulli spaces covers right-nested conditionals $B \rightarrow (C \rightarrow D)$; left-nested conditionals $(B \rightarrow C) \rightarrow D$; conjoined conditionals $(A \rightarrow B) \land (C \rightarrow D)$ and conditional conditionals $(A \rightarrow B) \rightarrow (C \rightarrow D)$. The model makes it possible to give explicit formulas for these probabilities: \[ P^*(B \rightarrow (C \rightarrow D)) = P(CD | B) + P(C^c | B)P(D | C), \] where $C^c$ is the complement of $C$, and $CD$ is the conjunction (intersection) of events $C$ and $D$. $P^*$ is the probability in the Bernoulli Stalnaker space. The formula for the conjoined conditional (Kaufmann 2009, 12, see also McGee 1989, 500) is:

\[
P^*((A \rightarrow B) \land (C \rightarrow D)) = \frac{P(ABCD) + P(\neg ACD)P(A \rightarrow B) + P(AB\neg C)P(C \rightarrow D)}{P(A \lor C)}
\]

These are very important achievements and a perfect motivation and justification for introducing and investigating the Bernoulli Stalnaker model. The graph model covers all these cases as well but simplifies matters and allows for generalizations (see also footnote 37).
2 Two-Urn Examples: Global and Local Conditional Probabilities

Consider the following example (which is similar to the example analyzed in Kaufman (2004) and Khoo (2016)):

We have two urns I, II. In urn I there are 2 White, 9 Green and 1 Red ball. In the urn II there are 50 White, 1 Green and 9 Red balls. Assume that urn I can be chosen with probability 1/4 and urn II with probability 3/4.21

We have chosen these particular values to simplify the comparison of our results with the results given in Kaufman (2004). But as our considerations have a general character, we use the following symbols: $\lambda_1, \lambda_2$ are the probabilities of choosing urn I and II; $r_1, p_1, q_1$ and $r_2, p_2, q_2$ are the probabilities of a White/Green/Red ball within the respective urns I and II.

Again, we want to compute the probability of the sentence If not-White, then Green (i.e. $\neg W \rightarrow G$). To do this, we consider The Conditional Game and try to estimate the chance of winning it. This requires a precise description of the rules of the game. An urn must first be chosen and then a ball from this urn has to be drawn. If we draw a Green ball (from any of the urns I, II), we win; if we draw a Red ball (for any urn), we lose. But what happens when we draw a White ball?

A priori, the following decisions are possible:

1. We win the game.
2. We lose the game.
3. The game is undecided and stops.
4. The game is undecided. We put the ball back (to the urn it was drawn from), and restart the game: we draw the urn and the ball from it.
5. The game is undecided. We put the ball back (to the urn it was drawn from), but remain within the chosen urn (i.e. we keep drawing the ball from the same urn until the game is decided).

Just like we did in the one-urn model, we can agree that the options (1), (2) and (3) are not attractive. This means that we have (4) and (5) at our disposal: after drawing a White ball we can restart the game from the beginning or remain within the drawn urn and continue the game within it. A priori none of these variants is superior and so we have to make a decision. This means that we have two possible variants of The Conditional Game—and in particular two possible interpretations of the sentence in question. Depending on the chosen variant, we have to adjust the formal constructions of it. Both these variants of the game correspond to the global and local interpretations of the conditional discussed in the literature (Kaufmann 2004; Khoo 2016). We will use this terminology, in particular denoting global as the

---

21 In Kaufmann’s model we have a red ball with a spot instead of green and a red ball without a spot for red. Choosing a red ball (with or without a spot) corresponds to choosing a non-White ball. Kaufmann considers the sentence: If I pick a red ball, it will have a black spot, which corresponds exactly to ours If the ball is non-White, then it is Green. As the examples are entirely isomorphic, it is easy to compare the results.
conditional defined (interpreted) by (4), and local as the conditional defined (interpreted) by (5). We adopt the terminology from Kaufmann (2004, 2005, 2009, 2015), leaving aside the discussion concerning the detailed relationships between this classification, and other possible ones, like the standard subjunctive/indicative classification of conditionals (as exemplified by the famous Oswald–Kennedy sentences (Adams 1970)) or epistemic/metaphysical distinction from Khoo (2015).

When discussing these two possible variants we will see the advantage of the graph semantics—it shows clearly the differences and allows us to compute the corresponding probabilities in a simple (and mathematically correct) way.

## 3 The Global Interpretation

Consider the game with a global interpretation of the conditional (i.e. given by the condition (4)). In this case, we can give several graph representations of the game where the one we give is the simplest one.

### 3.1 A Stochastic Graph Representation

Graph 3 (Fig. 3, next page) illustrates the mechanism in the Global Conditional Game.

The transition probabilities in this graph are given in the following way:

\[
\begin{align*}
\text{P}_{\text{START, I}} &= \lambda_1 = 3/4 \\
\text{P}_{\text{START, II}} &= \lambda_2 = 1/4 \\
\text{P}_{\text{I, START}} &= r_1 = 2/12 \\
\text{P}_{\text{I, WIN}} &= p_1 = 9/12 \\
\text{P}_{\text{II, LOSS}} &= q_1 = 1/12 \\
\text{P}_{\text{II, START}} &= r_2 = 50/60 \\
\text{P}_{\text{II, WIN}} &= p_2 = 1/60 \\
\text{P}_{\text{II, LOSS}} &= q_2 = 9/60 \\
\text{P}_{\text{WIN, WIN}} &= 1 \\
\text{P}_{\text{LOSS, LOSS}} &= 1.
\end{align*}
\]

Let us (as before) denote by \(P(\text{START})\) the probability of getting (finally) to the state WIN from the state START. In addition, let \(P(\text{I}), P(\text{II})\) denote the probabilities

\[\begin{array}{cccccc}
\text{GLOBAL} & \text{START} & \text{I} & \text{II} & \text{WIN} & \text{LOSS} \\
\hline
\text{START} & 0 & \lambda_1 & \lambda_2 & 0 & 0 \\
\text{I} & r_1 & 0 & 0 & p_1 & q_1 \\
\text{II} & r_2 & 0 & 0 & p_2 & q_2 \\
\text{WIN} & 0 & 0 & 0 & 1 & 0 \\
\text{LOSS} & 0 & 0 & 0 & 0 & 1
\end{array}\]
of getting (finally) to the state WIN from the states I, II respectively (perhaps after several moves). Just as in the case of the one‑urn model, we compute P(START) using the appropriate system of equations for the Markov graph:

\[
P(\text{START}) = \lambda_1 P(I) + \lambda_2 P(\text{II})
\]

\[
P(I) = p_1 + r_1 P(\text{START})
\]

\[
P(\text{II}) = p_2 + r_2 P(\text{START})
\]

After solving this system, we obtain the formula:

**Fact 3.1**

\[
P(\text{START}) = \frac{\lambda_1 p_1 + \lambda_2 p_2}{\lambda_1 p_1 + \lambda_2 p_2 + \lambda_1 q_1 + \lambda_2 q_2} = 0.6
\]

Here we have made use of the fact that \(1 - r_1 = (p_1 + q_1)\); \(1 - r_2 = (p_2 + q_2)\).

### 3.2 The Probability Space

We will construct a probability space \(\Xi_\alpha^{\text{glob}} = (\Omega_\alpha^{\text{glob}}, M_\alpha^{\text{glob}}, P_\alpha^{\text{glob}})\) for the game defined in this way.

Let \(W_1, W_2, G_1, G_2, R_1, R_2\) denote the drawing of a White, Green or Red ball from urns I or II respectively. In the case of the global interpretation, the paths in the graph show us which sequences of results (draws) are possible. The set of sequences, representing the possible histories of The Global Conditional Game is: \{\(G_1, G_2, R_1, R_2, W_1G_1, W_1G_2, W_2G_1, W_2G_2, W_1R_1, W_1R_2, W_2R_1, W_2R_2, W_1W_1G_1, W_1W_1G_2, W_1W_2G_1, W_1W_2G_2, \ldots\}\). Our victories are represented by sequences ending with \(G_1\) or \(G_2\) (our losses—respectively—\(R_1\) or \(R_2\)).

The important feature of this case is that after drawing (and replacing) the White ball from urn I, we can in the next step draw any ball from any urn—e.g. Green from
urn II. This is because returning the White ball to any urn restarts the whole procedure—again we select an urn and draw a ball from it. This is indicated by the arrows in the graph leading from I and II to START (and again to any urn and any ball). In particular, this means that in the probability space we will construct, there will be “mixed” sequences: first, there are White balls (from any of the urns), then the decisive ball (from any urn).

Consider the probability space $\Xi_{U2} = (\Omega_{U2}, \mathcal{M}_{U2}, P_{U2})$, where $\Omega_{U2}$ is the set of elementary events: $\Omega_{U2} = \{W_1, G_1, R_1, W_2, G_2, R_2\}$, with probabilities $\lambda_1r_1, \lambda_1p_1, \lambda_1q_1, \lambda_2r_2, \lambda_2p_2, \lambda_2q_2$ respectively; $\mathcal{M}_{U2}$ is the $\sigma$-field of all events; $P_{U2}$ is the probability measure on $\mathcal{M}_{U2}$. We assume that $\lambda_1 + \lambda_2 = 1; p_1 + q_1 + r_1 = 1; p_2 + q_2 + r_2 = 1$.24

**Definition 3.2** The space $\Xi_\alpha^{\text{glob}}$ for the globally understood conditional $\alpha = \neg W \rightarrow G$ is the triple $\Xi_\alpha^{\text{glob}} = (\Omega_\alpha^{\text{glob}}, \mathcal{M}_\alpha^{\text{glob}}, P_\alpha^{\text{glob}})$ where:

1. $\Omega_\alpha^{\text{glob}}$ is the set of elementary events—these are the sequences of events from the space $\Xi_{U2}$, which decide The Global Conditional Game $\alpha$;
2. $\mathcal{M}_\alpha^{\text{glob}}$ is the $\sigma$-field of all events;
3. $P_\alpha^{\text{glob}}$ is a probability measure on $\mathcal{M}_\alpha^{\text{glob}}$.

Formally:

$\Omega_\alpha^{\text{glob}} = \{(W_1/W_2)^n(G_1/G_2), (W_1/W_2)^n(R_1/R_2): n \in \mathbb{N}\}$, where by $(W_1/W_2)^n$ we denote any sequence of White balls of length $n$, coming from any urn; similarly $(G_1/G_2), (R_1/R_2)$ denotes any Green (Red) ball.

So these sequences consist of a (possibly empty) series of White balls coming from any urn, followed by a Green or Red ball (from any urn).

$\mathcal{M}_\alpha^{\text{glob}} = \mathcal{P}(\Omega_\alpha^{\text{glob}})$ (here $\mathcal{P}$ denotes the powerset of $\Omega_\alpha^{\text{glob}}$)

The probability $P_\alpha^{\text{glob}}$ is defined on elementary events in the following way:

$P_\alpha^{\text{glob}}((W_1/W_2)^nG_1) = (\lambda_1r_1)^k(\lambda_2p_2)^{n-k}(\lambda_1p_1)$—when $k$ balls in the sequence $(W_1/W_2)^nG_1$ are of the kind $W_1$, and $n-k$ balls are of the kind $W_2$, for $n,k \in \mathbb{N}$; $0 \leq k \leq n$

$P_\alpha^{\text{glob}}((W_1/W_2)^nG_2) = (\lambda_1r_1)^k(\lambda_2q_2)^{n-k}(\lambda_2p_2)$—when $k$ balls in the sequence $(W_1/W_2)^nG_1$ are of the kind $W_1$, and $n-k$ balls are of the kind $W_2$, for $n,k \in \mathbb{N}$; $0 \leq k \leq n$

$P_\alpha^{\text{glob}}((W_1/W_2)^nR_1) = (\lambda_1r_1)^k(\lambda_2q_2)^{n-k}(\lambda_1q_1)$—when $k$ balls in the sequence $(W_1/W_2)^nR_1$ are of the kind $W_1$, and $n-k$ balls are of the kind $W_2$, for $n,k \in \mathbb{N}$; $0 \leq k \leq n$

$P_\alpha^{\text{glob}}((W_1/W_2)^nR_2) = (\lambda_1r_1)^k(\lambda_2r_2)^{n-k}(\lambda_2q_2)$—when $k$ balls in the sequence $(W_1/W_2)^nR_2$ are of the kind $W_1$, and $n-k$ balls are of the kind $W_2$, for $n,k \in \mathbb{N}$; $0 \leq k \leq n$25

The chance that our game ends at the $(n+1)$-st move, and the last ball is $G_1$ is:

$$(\lambda_1r_1 + \lambda_2r_2)^n(\lambda_1p_1) \quad (\text{for } n \in \mathbb{N})$$

---

24 U2 in the index indicates the fact that we have two urns to draw the balls from.

25 E.g. the probability of $W_1W_1W_2W_1W_1G_1$ equals $(\lambda_1r_1)^3(\lambda_2r_2)^2(\lambda_1p_1)$. 
So:
The chance, that the game lasts (n + 1) moves and the last ball is $G_1$ is $(\lambda_1 r_1 + \lambda_2 r_2)^n(\lambda_1 p_1)$.

Similarly:
The chance, that the game lasts (n + 1) moves and the last ball is $G_2$ is $(\lambda_1 r_1 + \lambda_2 r_2)^n(\lambda_2 p_2)$.

So, the chance that the game ends at the (n + 1)-th move and we win, is:

$$(\lambda_1 r_1 + \lambda_2 r_2)^n(\lambda_1 p_1 + \lambda_2 p_2).$$

We want to compute the probability of the conditional $\neg W \rightarrow G$. It has its counterpart $WIN_\alpha^{\text{glob}}$ in the probability space $\Xi_\alpha^{\text{glob}} = (\Omega_\alpha^{\text{glob}}, M_\alpha^{\text{glob}}, P_\alpha^{\text{glob}})$. This set consists of all game scenarios leading to the victory:

$$WIN_\alpha^{\text{glob}} = [\neg W \rightarrow G]_\alpha^{\text{glob}} = \{(W_1/W_2)^n (G_1/G_2) : n \in \mathbb{N}\}$$

The event *We lose in this game* is the set:

$$LOSS_\alpha^{\text{glob}} = [\neg W \rightarrow R]_\alpha^{\text{glob}} = \{(W_1/W_2)^n (R_1/R_2) : n \in \mathbb{N}\}$$

We compute $P_\alpha^{\text{glob}}(WIN_\alpha^{\text{glob}})$ as a sum of geometric series and obtain the (already known) result:

**Fact 3.3**

$$P_\alpha^{\text{glob}}(WIN_\alpha^{\text{glob}}) = \sum_{n=0}^\infty (\lambda_1 r_1 + \lambda_2 r_2)^n(\lambda_1 p_1 + \lambda_2 p_2) = \frac{\lambda_1 p_1 + \lambda_2 p_2}{\lambda_1 (1 - r_1) + \lambda_2 (1 - r_2)}$$

Similarly:

**Fact 3.4**

$$P_\alpha^{\text{glob}}(LOSS_\alpha^{\text{glob}}) = \sum_{n=0}^\infty (\lambda_1 r_1 + \lambda_2 r_2)^n(\lambda_1 q_1 + \lambda_2 q_2) = \frac{\lambda_1 q_1 + \lambda_2 q_2}{\lambda_1 (1 - r_1) + \lambda_2 (1 - r_2)}$$

**Fact 3.5**

$$P_\alpha^{\text{glob}}(WIN_\alpha^{\text{glob}}) + P_\alpha^{\text{glob}}(LOSS_\alpha^{\text{glob}}) = 1$$
3.3 Kaufmann’s Equation (The Global Version)

In our model, some intuitive considerations concerning the chances to win can be formalized. Kaufmann (2004, 586) examines an equation which is intuitively clear but formally not quite precise:

$$P(\neg W \rightarrow G) = P(\neg W \rightarrow G|I)P(I|\neg W) + P(\neg W \rightarrow G|II)P(II|\neg W) = 0.6$$

Imagine someone who has already observed that the drawn ball is not White (so the game is already decided). What is the chance that it is Green? The knowledge, that the ball is not White gives us some indirect information about the urn it (probably) comes from. This information is contained in the expressions $P(I|\neg W)$ and $P(II|\neg W)$. Now, if the ball is from urn I, the chance of the conditional $\neg W \rightarrow G$ is $P(\neg W \rightarrow G|I)$, if it is from urn II, the chance is $P(\neg W \rightarrow G|II)$. This leads to the equation above.

The weakness of this equation consists in the fact that it is not clear where (i.e. in what probability space) the probability function $P$ is defined. In particular, there should be an event corresponding to the conditional $\neg W \rightarrow G$, but Kaufmann does not define an appropriate probability space.\(^{26}\)

We can give a precise counterpart of this equation in the following way (using the observation, that drawing urn I/II can be written as the event $(W_1 \lor G_1 \lor R_1)$, resp.: $(W_2 \lor G_2 \lor R_2)$):

Instead of $P(I|\neg W)$ we take: $P_{U2}(\neg W_1 \land \neg W_2)$.  
Instead of $P(II|\neg W)$ we take: $P_{U2}(\neg W_1 \land \neg W_2)$.

Instead of $P(\neg W \rightarrow G|I)$ we take: $P_{U2}(G_1)(G_1 \lor R_1)$.
Instead of $P(\neg W \rightarrow G|II)$ we take: $P_{U2}(G_2)(G_2 \lor R_2)$.\(^{27}\)

The conditional probabilities in the (simple) space $\Xi_{U2}$ are given by:

$$P_{U2}((W_1 \lor G_1 \lor R_1)(\neg W_1 \land \neg W_2)) = \frac{\lambda_1(p_1 + q_1)}{\lambda_1 p_1 + \lambda_2 p_2 + \lambda_1 q_1 + \lambda_2 q_2} = 5/8$$

$$P_{U2}((W_2 \lor G_2 \lor R_2)(\neg W_1 \land \neg W_2)) = \frac{\lambda_2(p_2 + q_2)}{\lambda_1 p_1 + \lambda_2 p_2 + \lambda_1 q_1 + \lambda_2 q_2} = 3/8$$

$$P_{U2}(G_1)(G_1 \lor R_1) = \frac{p_1}{p_1 + q_1} = 9/10$$

$$P_{U2}(G_2)(G_2 \lor R_2) = \frac{p_2}{p_2 + q_2} = 1/10$$

\(^{26}\) Khoo observes: “Kaufmann’s theory marks a significant advance in our thinking about conditionals. However, I think it faces some challenges. First, it is incomplete: since Kaufmann never specifies a semantics for conditionals, it remains unclear why conditionals should have local and global interpretations – for instance, is it due to ambiguity or perhaps context-dependence?” Khoo (2016, 8). We remove this weakness.

\(^{27}\) $P(\neg W \rightarrow G|I)$ is replaced by $P_{U2}(G_1)(G_1 \lor R_1)$. $P(\neg W \rightarrow G|II)$ expresses our intuitive degree of belief that we will win the Conditional Game $\neg W \rightarrow G$, under the assumption, that we are in urn I. But to win the game $\neg W \rightarrow G$ amounts to drawing a Green ball, provided it is Green or Red, i.e. not White (and everything is going on within urn I). And this is exactly $P_{U2}(G_1)(G_1 \lor R_1)$.
So, finally:

\[ P(\neg W \rightarrow G) = \frac{\lambda_1 q_1 + \lambda_2 q_2}{(\lambda_1 p_1 + \lambda_2 p_2 + \lambda_1 q_1 + \lambda_2 q_2)} = 0.6 \]

which is formally correct, identical with the results known from the literature (Kaufmann 2004; Khoo 2016) and already established by Fact 3.3. We have here just made the intuitive considerations mathematically precise.

Two facts are of particular importance for the ongoing analyses:

**Fact 3.6**

\[
P_{\alpha}^{\text{glob}}(\neg W \rightarrow G) = P_{\Pi}(G_1 | (G_1 \lor R_1))P_{\Pi}((W_1 \lor G_1 \lor R_1)(\neg W_1 \land \neg W_2)) \\
+ P_{\Pi}(G_2 | (G_2 \lor R_2))P_{\Pi}((W_2 \lor G_2 \lor R_2)(\neg W_1 \land \neg W_2)) \\
= \frac{p_1}{p_1 + q_1} \frac{\lambda_1 (p_1 + q_1)}{(\lambda_1 p_1 + \lambda_2 p_2 + \lambda_1 q_1 + \lambda_2 q_2)} \\
+ \frac{p_2}{p_2 + q_2} \frac{\lambda_2 (p_2 + q_2)}{(\lambda_1 p_1 + \lambda_2 p_2 + \lambda_1 q_1 + \lambda_2 q_2)} \\
= \frac{\lambda_1 q_1 + \lambda_2 q_2}{(\lambda_1 p_1 + \lambda_2 p_2 + \lambda_1 q_1 + \lambda_2 q_2)}
\]

**Fact 3.7**

\[
P_{\alpha}^{\text{glob}}(\neg W \rightarrow R) = \frac{q_1}{p_1 + q_1} \frac{\lambda_1 (p_1 + q_1)}{(\lambda_1 p_1 + \lambda_2 p_2 + \lambda_1 q_1 + \lambda_2 q_2)} \\
+ \frac{q_2}{p_2 + q_2} \frac{\lambda_2 (p_2 + q_2)}{(\lambda_1 p_1 + \lambda_2 p_2 + \lambda_1 q_1 + \lambda_2 q_2)} \\
= \frac{\lambda_1 q_1 + \lambda_2 q_2}{(\lambda_1 p_1 + \lambda_2 p_2 + \lambda_1 q_1 + \lambda_2 q_2)}
\]

The structure of these formulas is very similar to the structure of the formulas for the probabilities of \textsc{WIN}_\alpha and \textsc{LOSS}_\alpha in The (One-Urn) Conditional Game \neg W \rightarrow G (cf. Facts 1.3 and 1.4). They correspond precisely to the situation where we have just one urn (call it \textit{U}_{\text{new}}), where the probabilities of drawing a White, Green, or Red ball are respectively:

\[
P_{\text{new}}(W) = \lambda_1 r_1 + \lambda_2 r_2 \\
P_{\text{new}}(G) = \lambda_1 p_1 + \lambda_2 p_2 \\
P_{\text{new}}(R) = \lambda_1 q_1 + \lambda_2 q_2.
\]

In this urn:

\[
P_{\text{new}}(G | \neg W) = \frac{\lambda_1 p_1 + \lambda_2 p_2}{(\lambda_1 p_1 + \lambda_2 p_2 + \lambda_1 q_1 + \lambda_2 q_2)}
\]

Consider the probability space \( \Xi_{\text{new}} = (\Omega_{\text{new}}, M_{\text{new}}, P_{\text{new}}) \), where \( \Omega_{\text{new}} \) is the set of elementary events: \( W, G, R \) with probabilities respectively \( \lambda_1 r_1 + \lambda_2 r_2, \lambda_1 p_1 + \lambda_2 p_2, \lambda_1 q_1 + \lambda_2 q_2; M_{\text{new}} \) is the \( \sigma \)-field of all events and \( P_{\text{new}} \) is the probability measure defined on \( M_{\text{new}} \).
The following fact takes place:

**Fact 3.8**

\[ p_{\text{glob}}(\neg W \rightarrow G) = p_{\text{new}}(G|\neg W) \]

It is important to observe that the probability space \( \Xi_{\text{new}} \) has a universal character which means that it allows us to compute the probabilities of other conditionals as conditional probabilities. For example, the probability of \( \beta = \neg G \rightarrow W \) can be computed within \( \Xi_{\text{new}} \) simply as \( P_{\text{new}}(W|\neg G) \). Of course, the appropriate probability space \( \Xi_{\beta}^{\text{glob}} = (\Omega_{\beta}^{\text{glob}}, \Sigma_{\beta}^{\text{glob}}, P_{\beta}^{\text{glob}}) \) will be different than \( \Xi_{\alpha}^{\text{glob}} = (\Omega_{\alpha}^{\text{glob}}, \Sigma_{\alpha}^{\text{glob}}, P_{\alpha}^{\text{glob}}) \), as it will consist of sequences of the form \( G^n W, G^n R \). So \( P_{\beta}^{\text{glob}}(\neg G \rightarrow W) = P_{\text{new}}(W|\neg G) \).

We summarize these observations as:

**Fact 3.9**

For any probability space \( \Xi = (\Omega, \Sigma, P) \), where \( \Omega = \{ W, G, R \} \), there is another probability space \( \Xi_{\text{new}} = (\Omega_{\text{new}}, \Sigma_{\text{new}}, P_{\text{new}}) \), where \( \Omega_{\text{new}} = \{ W, G, R \} \) is the set of elementary events with probabilities \( \lambda_1 r_1 + \lambda_2 r_2, \lambda_1 p_1 + \lambda_2 p_2, \lambda_1 q_1 + \lambda_2 q_2 \) respectively; \( \Sigma_{\text{new}} \) is the \( \sigma \)-field of all events, and \( P_{\text{new}} \) is the probability measure defined on \( \Sigma_{\text{new}} \), such that:

For any globally interpreted conditional \( \neg X \rightarrow Y \), where \( X, Y \in \{ W, G, R \} \) and \( X \neq Y \), and the space \( \Xi_{\neg X \rightarrow Y}^{\text{glob}} = (\Omega_{\neg X \rightarrow Y}^{\text{glob}}, \Sigma_{\neg X \rightarrow Y}^{\text{glob}}, P_{\neg X \rightarrow Y}^{\text{glob}}) \):

\[ P_{\neg X \rightarrow Y}^{\text{glob}}(\neg X \rightarrow Y) = P_{\text{new}}(Y|\neg X) \]

Let us once again take a look at the formula in Fact 3.6. The terms \( \frac{p_1}{p_1 + q_1} \) and \( \frac{p_2}{p_2 + q_2} \) are just the probabilities of the conditional \( \neg W \rightarrow G \) computed separately in the two urns I and II (or: the probabilities of winning The Conditional Game within these single urns). The more complicated terms \( \frac{\lambda_1 (p_1 + q_1)}{\lambda_1 (p_1 + q_1) + \lambda_2 (p_2 + q_2)} \) and \( \frac{\lambda_2 (p_2 + q_2)}{\lambda_1 (p_1 + q_1) + \lambda_2 (p_2 + q_2)} \) are conditional probabilities of being in a certain urn, provided we know that a non-White ball has been drawn. So we have a mathematically precise expression which corresponds exactly to Kaufmann’s intuitive formulas. Remember that The Global Conditional Game in two urns can be simulated with one urn, where the probabilities of White, Green and Red balls have been fixed in an appropriate way. This urn is universal (or: stable) in the sense that it can also be used for computing the probabilities of all globally interpreted conditionals formulated in the language with atomic sentences \( W, G, R \).\(^{28}\)

### 4 Local Interpretation

#### 4.1 The Stochastic Graph Representation

If we chose the local interpretation, i.e. option (5), we remain in the urn drawn in the first move until the game is decided. This means that we have different scenarios

---

\(^{28}\) It does not matter that we have three atomic sentences – the generalization to \( n \) atomic sentences \( X_1, \ldots, X_n \) is straightforward. It will also work for a countably infinite language (set of elementary events).
leading to the decision compared to the global case. We can give the simplest graph representation of the game (just like in the global case).

Graph 4 (Fig. 4, next page) corresponds to graph 3. It has the same states (START, I, II, WIN, LOSS), but its structure (i.e. edges) is different.

The difference with respect to graph 3 (Fig. 3—illustrating the global case) is that it is no longer possible to get back to START from the states I, II. This is because in the local interpretation, we remain within an urn after we get into it. In particular, if we are already in one of the urns i (i = I, II), we have 3 possible actions (Fig. 4):

- After drawing a White ball we are again in the same state i (with probability \( r_i \))
- After drawing a Green ball we go to WIN (with probability \( p_i \))
- After drawing a Red ball we go to LOSS (with probability \( q_i \))

WIN and LOSE are absorbing states: if we get in any of them, we remain there and the game stops.

The transition probabilities in the graph are given in the following way:

\[
\begin{align*}
P_{\text{START}, I} &= \lambda_1 = 1/4 \\
P_{\text{START}, II} &= \lambda_2 = 3/4 \\
P_{I, I} &= r_1 = 2/12 \\
P_{I, \text{WIN}} &= p_1 = 9/12 \\
P_{I, \text{LOSS}} &= q_1 = 1/12 \\
P_{\text{II}, II} &= r_2 = 50/60 \\
P_{\text{II}, \text{WIN}} &= p_2 = 1/60 \\
P_{\text{II}, \text{LOSS}} &= q_2 = 9/60 \\
P_{\text{WIN}, \text{WIN}} &= 1 \\
P_{\text{LOSS}, \text{LOSS}} &= 1
\end{align*}
\]

This Markov graph exhibits the dynamics of the game and allows for a very simple computation of the probability of victory. The respective system of equations is:

\[
\begin{align*}
P_{\text{START}} &= \lambda_1 P(I) + \lambda_2 P(II) \\
P(I) &= r_1 P(I) + p_1 \\
P(II) &= r_2 P(II) + p_2
\end{align*}
\]

After solving it, we obtain:

\[29\] The transition probability matrix for this graph (Fig. 4) in the general case is:

| LOCAL    | START | I   | II  | WIN | LOSS |
|----------|-------|-----|-----|-----|------|
| START    | 0     | \( \lambda_1 \) | \( \lambda_2 \) | 0   | 0    |
| I        | 0     | \( r_1 \) | 0   | \( p_1 \) | \( q_1 \) |
| II       | 0     | \( r_2 \) | 0   | \( p_2 \) | \( q_2 \) |
| WIN      | 0     | 0   | 0   | 1   | 0    |
| LOSS     | 0     | 0   | 0   | 0   | 1    |
Fact 4.1

\[
P(\text{START}) = \lambda_1 \frac{p_1}{(p_1 + q_1)} + \lambda_2 \frac{p_2}{(p_2 + q_2)}
\]

The paths in the graph show us, which sequences of draws are possible. Of course, in the local situation, the sequences are homogeneous: only one index appears in them as a green/red ball from urn I (II) can be preceded only by white balls from the same urn I (II). So the sequences form the following set:

\[
\{G_1, G_2, R_1, R_2, W_1G_1, W_1R_1, W_2G_2, W_2R_2, W_1W_1G_1, W_1W_1R_1, W_2W_2G_2, \ldots\}
\]

4.2 The Probability Space

We construct the probability space \(\Xi_\alpha^{\text{loc}} = (\Omega_\alpha^{\text{loc}}, M_\alpha^{\text{loc}}, P_\alpha^{\text{loc}})\) for the Local Conditional Game.

Consider again the probability space \(\Xi_{\text{U2}} = (\Omega_{\text{U2}}, M_{\text{U2}}, P_{\text{U2}})\), where \(\Omega_{\text{U2}}\) is the set of elementary events \(\Omega_{\text{U2}} = \{W_1, W_2, G_1, G_2, R_1, R_2\}\) with probabilities \(\lambda_1r_1, \lambda_2r_2, \lambda_1p_1, \lambda_2p_2, \lambda_1q_1, \lambda_2q_2\), \(M_{\text{U2}}\) is the \(\sigma\)-field of all events, and \(P_{\text{U2}}\) is the probability measure on \(M_{\text{U2}}\) (cf. Definition 3.2.)

Definition 4.2 The space \(\Xi_\alpha^{\text{loc}}\) for the locally understood conditional \(\alpha = \neg W \rightarrow G\) is the triple \(\Xi_\alpha^{\text{loc}} = (\Omega_\alpha^{\text{loc}}, M_\alpha^{\text{loc}}, P_\alpha^{\text{loc}})\), where:

1. \(\Omega_\alpha^{\text{loc}}\) is the set of elementary events—these are the sequences of events from the space \(\Omega_{\text{U2}}\) which decide The Local Conditional Game \(\alpha\);
2. \(M_\alpha^{\text{loc}}\) is the \(\sigma\)-field of all events;
3. \(P_\alpha^{\text{loc}}\) is a probability measure on \(M_\alpha^{\text{loc}}\).

Formally:

\[
\Omega_\alpha^{\text{loc}} = \{W_1^nG_1, W_1^nR_1, W_2^nG_2, W_2^nR_2 : n \in \mathbb{N}\}
\]
\[ M_{\alpha}^{\text{loc}} = \mathcal{P}(\Omega_{\alpha}^{\text{loc}}) \] (here \( \mathcal{P} \) denotes the powerset of \( \Omega_{\alpha}^{\text{loc}} \))

The probability measure \( P_{\alpha}^{\text{loc}} \) is defined on the elementary events in the following way:

\[ P_{\alpha}^{\text{loc}}(W_1^nG_1) = \lambda_1(r_1)^n p_1, \text{ (for } n \in \mathbb{N}) \];

\[ P_{\alpha}^{\text{loc}}(W_1^nR_1) = \lambda_1(r_1)^n q_1, \text{ (for } n \in \mathbb{N}) \];

\[ P_{\alpha}^{\text{loc}}(W_2^nG_2) = \lambda_2(r_2)^n p_2, \text{ (for } n \in \mathbb{N}) \];

\[ P_{\alpha}^{\text{loc}}(W_2^nR_2) = \lambda_2(r_2)^n q_2, \text{ (for } n \in \mathbb{N}) \].

So there are four types of elementary events in this space, corresponding to possible scenarios of The Local Conditional Game:

- Urn I has been drawn, and then \( n \) balls \( W_1 \), finally \( G_1 \);
- Urn I has been drawn, and then \( n \) balls \( W_1 \), finally \( R_1 \);
- Urn II has been drawn, and then \( n \) balls \( W_2 \), finally \( G_2 \);
- Urn II has been drawn, and then \( n \) balls \( W_2 \), finally \( R_2 \);

The counterpart of \( \neg W \rightarrow G \) in this probability space (i.e. the set of scenarios leading to victory) is:

\[ \text{WIN}_{\alpha}^{\text{loc}} = \{ W \rightarrow G \}_{\alpha}^{\text{loc}} = \{ W_1^nG_1, W_2^nG_2 : n \in \mathbb{N} \} \]

The event we lose is:

\[ \text{LOSS}_{\alpha}^{\text{loc}} = \{ W_1^nR_1, W_2^nR_2 : n \in \mathbb{N} \} \]

The probabilities of victory (and defeat) have been computed in a simple way with the use of the stochastic graph. We can also (as in the case of the global interpretation) compute it in the straightforward way: after summing up appropriate geometric series, we obtain the familiar results:

**Fact 4.3**

\[ P_{\alpha}^{\text{loc}}(\text{WIN}_{\alpha}^{\text{loc}}) = \lambda_1 \frac{p_1}{(p_1 + q_1)} + \lambda_2 \frac{p_2}{(p_2 + q_2)} \]

**Fact 4.4**

\[ P_{\alpha}^{\text{loc}}(\text{LOSS}_{\alpha}^{\text{loc}}) = \lambda_1 \frac{q_1}{(p_1 + q_1)} + \lambda_2 \frac{q_2}{(p_2 + q_2)} \]
4.3 Kaufmann’s Equation (The Local Version)

We can again compare these formalisms and computations with the intuitive (but informal) computations from Kaufman (2004, 586):

\[ P(\neg W \rightarrow G) = P((\neg W \rightarrow G) \mid I)P(I) + P((\neg W \rightarrow G) \mid II)P(II). \]

In the local interpretation, we have \( P(I), P(II) \)—not the conditional probabilities \( P(I \mid \neg W), P(II \mid \neg W) \)! We can formalize Kaufmann’s formulas in the following way:

\[
P(I) = P(U_2(G_1 \lor R_1 \lor W_1)) = \lambda_1 \\
P(II) = P(U_2(G_2 \lor R_2 \lor W_2)) = \lambda_2 \\
P((\neg W \rightarrow G) \mid I) = P(U_2(G_1 \mid G_1 \lor R_1)) = \frac{p_1}{(p_1 + q_1)} \\
P((\neg W \rightarrow G) \mid II) = P(U_2(G_2 \mid G_2 \lor R_2)) = \frac{p_2}{(p_2 + q_2)}
\]

Finally, we have two facts, which are exact counterparts of Kaufmann’s intuitive formulas:

**Fact 4.5**

\[
P_{\alpha}^{loc}(\neg W \rightarrow G) = P(U_2(G_1 \mid (G_1 \lor R_1))) P(U_2(G_1 \lor R_1 \lor W_1) + P(U_2(G_2 \mid (G_2 \lor R_2))) P(U_2(G_2 \lor R_2 \lor W_2)) = \lambda_1 \frac{p_1}{(p_1 + q_1)} + \lambda_2 \frac{p_2}{(p_2 + q_2)}
\]

**Fact 4.6**

\[
P_{\alpha}^{loc}(\neg W \rightarrow R) = P(U_2(R_1 \mid (G_1 \lor R_1))) P(U_2(G_1 \lor R_1 \lor W_1) + P(U_2(R_2 \mid (G_2 \lor R_2))) P(U_2(G_2 \lor R_2 \lor W_2)) = \lambda_1 \frac{q_1}{(p_1 + q_1)} + \lambda_2 \frac{q_2}{(p_2 + q_2)}
\]

To summarize: the computation of the probability of the locally-interpreted conditional is based on the probabilities of the sentence \( \neg W \rightarrow G \) computed separately in urn I and urn II (which are \( \frac{p_1}{(p_1 + q_1)} \) and \( \frac{p_2}{(p_2 + q_2)} \)). They are then summed up with weights equal to the probabilities of drawing the urn I or II (i.e. \( \lambda_1 \) and \( \lambda_2 \) respectively).

4.4 The Global Versus the Local Interpretation

We have shown that the conditional can be formally interpreted in at least two ways. Both these interpretations rely on different assumptions concerning the rules of The Conditional Game (Fig. 5).

To stress the difference, consider a generalization of the two-urn example, namely assume that we have \( n \) different urns. The main question to decide is: what should
we do when the antecedent of the conditional is not fulfilled—i.e. when a White ball has been drawn from urn $i$. According to the rules, we repeat the draw, but we have at our disposal only those urns which are somehow connected with the urn $i$ via a certain relation $R$. If $R$ is the identity relation, it means that every urn $i$ is connected only with itself. So, after drawing White from urn $i$, we put the ball back and repeat the draw within the urn $i$. So in this case we have the local interpretation of the conditional.

If $R$ is the full relation, it means that all urns are connected with each other. Whereas previously (in the local case), after drawing a white we returned the ball to the urn $i$ and restarted the game within that particular urn $i$—here we select an urn again and draw the ball from this urn. In this case, we have the global interpretation of the conditional.

These two approaches are in a sense two extremes (and they are based on the “granularity” of the division of the set of urns), and many intermediate approaches are possible.

What are the important consequences of distinguishing between the local and global interpretation of the conditional? We will indicate only a few of them:

Under the local interpretation, the number of White balls in the urns does not matter. It means that the probability of the conditional is not dependent on the fact of how probable (or improbable) the antecedent is (i.e. what the probability is of

---

30 Observe that in the case, when there are only non-White balls in the urns (i.e. when $r_1 = 0$ and $r_2 = 0$) this distinction would make no sense: the probability of the conditional $\neg W \rightarrow G$ could be computed in one way only and it would be equal to $\lambda_1 p_1 + \lambda_2 p_2$ (the formulas from the facts 3.3, 4.1 both reduce to this simple form when $r_1 = r_2 = 0$). This is obvious, because in this case the game would be settled after drawing the first ball from the chosen urn.

31 In the literature concerning counterfactuals, such a relation $R$ is often interpreted as a relationship occurring between the nearest or neighbouring worlds (this can be traced back at least to Stalnaker (1968)).
drawing a non-White ball). It follows from the fact that the probability of the conditional in any of the urns depends not on the number of White balls but only on the proportion of Green balls to non-White balls (which is expressed by the formula \( \frac{p_i}{p_i + q_i} \)). Imagine that we put \( 10^{10} \) additional White balls into the urn: the probability of the conditional will not change (of course, on average, we will have to wait longer for the result). In other words: the degree of counterfactuality of a sentence has no influence on its logical (mathematical) probability but it influences the magnitude which we might call its “practical probability”. This issue is discussed elsewhere.\(^{32}\)

Under the global interpretation, the probability of the antecedent is important for estimating the probability of the conditional (again, we mean the probability of the event \( \text{non-White was drawn} \) in the initial space \( \Xi_{U2} = (\Omega_{U2}, M_{U2}, P_{U2}) \)). The following example will illustrate this: urn I contains 100 Green balls, 1 Red ball, and \( 10^{10} \) White balls. The probability of drawing this urn is \( 99/100 \). Urn II contains 1 Green ball, 100 Red balls and 1 White ball, and the probability of drawing urn II is \( 1/100 \). The probability of the (global) conditional \( \neg W \rightarrow G \) is approximately \( 1/100 \), because we have virtually no chance to decide the game within urn I. The game will therefore (with probability almost 1) be decided in urn II. This is the consequence of the fact that drawing a non-White ball from urn I is highly improbable.

In particular, it follows from the considerations above that The Global Conditional Game can be reduced to The Conditional Game in one urn \( U_{\text{new}} \). This new urn is universal in the sense characterized by the Fact 3.9. This is not possible for The Local Conditional Game. It follows from the fact that the probability of the locally interpreted conditional is independent from the number of white balls in the urns.

### 5 Lewis’s Triviality Result: PC = CP?

In the discussion concerning the probabilities of conditionals we have to take into account the PC = CP thesis, and the results of Lewis (1976) (and the ensuing discussion).\(^{33}\)

Observe first, that this thesis was considered to be the expression of certain intuitions which are usually presented by the following, oft-quoted claims of Ramsey and van Fraassen:

*If two people are arguing ‘If p will q?’ and both are in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q… We can say that they are fixing their degrees of belief in q given p* (Ramsey 1929, 247).

\(^{32}\) Here we observe only that it is easy to construct an example where the probability of the conditional \( \neg W \rightarrow G \) is approximately 1 but at the same time we can claim that the game will last almost forever. It is enough to consider one urn containing 1 Red ball, \( 10^{10} \) Green balls, and \( 10^{1000000} \) White balls. The chance of winning is almost 1, but we will have to wait very long for our victory (on average approx. \( 10^{100000} \) moves). The details are given in Wójtowicz and Wójtowicz (2019b).

\(^{33}\) We do not discuss the historical details and the discussion. The thesis has been also known as “Stalnaker’s Thesis” or “Adams’ Thesis” (Stalnaker 1970; Adams 1965, 1966, 1975). Hajek and Hall (1994)
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, 273).

Generally speaking, they express the view that the probability of the conditional is somehow dependent on the conditional probability. Applying it to the example of drawing balls from one urn (paragraph 1): if we want to know the probability of the conditional If the ball is not White, then it is Green, we have to compute the (usual) conditional probability Green was drawn under the condition non-White was drawn.

According to the widely-accepted interpretation of the PC = CP thesis, this statement should be written as:

\[ P(\neg W \rightarrow G) = P(G|\neg W) \]

And of course, in this form it is immediately subject to Lewis’s criticism: accepting standard assumptions concerning the properties of the probability measure and of the counterfactual leads us to the (obviously) false claim, that:

\[ P(\neg W \rightarrow G) = P(G)^{34} \]

But in the light of the semantics for conditionals proposed here it becomes obvious that even the very formulation of the thesis is not correct. The probability of the conditional \( \neg W \rightarrow G \) can only be defined in the probability space \( \Xi_{\alpha} \), as there is no event, corresponding to this sentence within the initial probability space \( \Xi \). Hayek’s numerical arguments concerning a similar example show it clearly enough (Hajek 2011, 4). So, if we are going to formulate the thesis PC = CP concerning the conditional \( \neg W \rightarrow G \) in a formally proper way, it surely cannot be:

\( (PC = CP) : P(\neg W \rightarrow G) = P(G|\neg W), \) where \( P \) is the probability measure in the initial probability space \( \Xi \).

But notice that in fact, the following equality holds (Fact 1.6):

\( (PC = CP) : P_{\alpha}(\neg W \rightarrow G) = P(G|\neg W), \) where on the left side we have the probability function \( P_{\alpha} \) from the probability space \( \Xi_{\alpha} \) (\( \alpha = \neg W \rightarrow G \)) and on the right side we have the probability function \( P \) defined in the initial probability space \( \Xi \).

Two remarks follow:

1. This formulation corresponds to the intuitions of Ramsey and van Fraassen—it shows how to reduce the difficult problem of estimating the probability of conditionals to the simpler problem of computing the conditional probabilities in the initial probability space \( \Xi \);
2. This thesis can be formally proven—it follows trivially from Fact 1.6.

Footnote 33 (continued)

and Hall (1994) use the term “Conditional Construal of Conditional Probability”; Hajek and Hall (1994) discuss several versions of it. Using their terminology we can say, that what we defend is closest to the universal version and the universal tailoring version Hajek and Hall (1994, 76).

\(^{34}\) Lewis’s result can be formulated as: there is no connective * such that \( P(X^*Y) = P(Y|X) \) (for every \( P \)).
What about the more complicated two-urn example? In this case, the problem of the dependence of the probability of the conditional $\neg W \rightarrow G$, and the conditional probabilities can be given a weaker and a stronger interpretation.

Under the **weaker** interpretation, we expect only that calculating the probability of the counterfactual $\neg W \rightarrow G$ is possible if we know the conditional probabilities $P(G|\neg W)$, calculated separately in urns I and II. This means that we expect that there is a formula giving the probability of the conditional, taking the conditional probabilities as arguments.

Under the **stronger** interpretation, we expect that the probability of the counterfactual $\neg W \rightarrow G$ is equal to a certain conditional probability $P_x(G|\neg W)$ in a suitably-constructed probability space $X$.

Facts 3.6 and 4.3 allow us to formulate the proper versions of the PC = CP principle under the weaker interpretation: on the left side of the equality (equalities) there is the value of the probability of the conditional $\alpha$ in appropriate probability spaces $\Xi^{glob}_\alpha$ and $\Xi^{loc}_\alpha$, and on the right side there is a certain function taking the appropriate conditional probabilities in the space $\Xi_{U2} = (\Omega_{U2}, M_{U2}, P_{U2})$ as arguments. The formulas for the conditional $\neg W \rightarrow G$ will have the following forms:

$$P^{glob}_{\neg W \rightarrow G} C = CP_{II}: \quad P^{glob}_{\neg W \rightarrow G} (\neg W \rightarrow G) =$$
$$P_{U2}(I)(\neg W_1 \land \neg W_2)P_{U2}(G_1|(G_1 \lor R_1)) + P_{U2}(II)(\neg W_1 \land \neg W_2)P_{U2}(G_2|(G_2 \lor R_2))$$

$$P^{loc}_{\neg W \rightarrow G} C = CP_{II}: \quad P^{loc}_{\neg W \rightarrow G} (\neg W \rightarrow G) =$$
$$P_{U2}(I)P_{U2}(G_1|(G_1 \lor R_1)) + P_{U2}(II)P_{U2}(G_2|(G_2 \lor R_2))$$

where I and II are abbreviations for $(W_1 \lor G_1 \lor R_1)$ and $(W_2 \lor G_2 \lor R_2)$ respectively.$^{35}$

Notwithstanding the fact that the right sides of these equalities are quite extended, we must agree that they conform to Ramsey’s intuitions: they reduce the probability of the conditional to the conditional probabilities $P_{U2}(G_1|(G_1 \lor R_1))$ and $P_{U2}(G_2|(G_2 \lor R_2))$ summed up with certain weights (which depend on the interpretation of the conditional as local or global). This is obvious for the local interpretation; for the global interpretation, it is evident from Fact 3.6. And it is clear that the probability of the conditional $\neg W \rightarrow G$ (interpreted globally) is equal to the conditional probability in the new urn $U_{new}$ (Fact 3.8).

For the global interpretation of the conditional, an even stronger version can be formulated in the form: $P^{glob}_{\neg W \rightarrow G} C = CP_{new}$, where $P_{new}$ is the probability in the suitable probability space $\Xi_{new} = (\Omega_{new}, M_{new}', P_{new})$ (cf. Fact 3.8). It should be stressed that this space (according to the remarks at the end of paragraph 3) has a stable

---

$^{35}$ $P_{U2}(G_1|(G_1 \lor R_1))$ and $P_{U2}(G_2|(G_2 \lor R_2))$ are arguments of these functions!
character: it can be used to compute the probabilities of other conditionals built from the atomic sentences $W, G, R$.\(^{36}\)

Observe, that a natural counterpart of the space $\Xi_{\text{new}} = (\Omega_{\text{new}}, M_{\text{new}}, P_{\text{new}})$ for the local interpretation does not exist. Of course, we can trivially define an artificial probability space $\Xi^# = (\Omega^#, M^#, P^#)$ with events dubbed $W, G, R$ in such a way that the probability of the locally interpreted conditional $\neg W \rightarrow G$ equals $P^#(G | \neg W)$. But the stability condition will not be fulfilled: this space will not work for other conditionals, e.g. for $\neg G \rightarrow W$. This is because it is only the globally interpreted conditional which has the Ramseyian character.

To summarize: when analyzing the relationships between the probability of the conditional and conditional probability, we have to be aware of the fact that the notion of probability can have different meanings (in the sense that it can be defined differently in different probability spaces). Lewis’s arguments are directed only against a particular version of the PC = CP thesis, where there is the same probability measure $P$ on both sides of the equality. But it need not have this form; in particular, this form is not enforced by the task it was meant to fulfill, i.e. the task of reducing the probabilities of the conditionals to (functions of) the conditional probabilities. But exactly this task is fulfilled (without any fear of Lewis’s arguments) by the principles we have given here:

- $(P_{\neg W \rightarrow G} C = CP)$: concerning The Conditional Game in one urn, where we have the strong interpretation.
- $(P_{\neg W \rightarrow G} C = CP_{U2})$: concerning The Two Urn Global Conditional Game—where the probability of the conditional $\neg W \rightarrow G$ is computed using conditional probabilities in the space $\Xi_{U2} = (\Omega_{U2}, M_{U2}, P_{U2})$ with appropriate weights—this is the weak interpretation.
- $(P_{\neg W \rightarrow G} C = CP_{\text{new}})$: concerning The Two Urn Global Conditional Game—where the probability of the conditional $\neg W \rightarrow G$ is computed using conditional probabilities in the space $\Xi_{\text{new}} = (\Omega_{\text{new}}, M_{\text{new}}, P_{\text{new}})$—i.e. we have the strong interpretation here.
- $(P_{\neg W \rightarrow G} C = CP_{U2})$: concerning The Two Urn Local Conditional Game—where the probability of the conditional $\neg W \rightarrow G$ is computed using conditional probabilities in the space $\Xi_{U2} = (\Omega_{U2}, M_{U2}, P_{U2})$ with appropriate weights—of course, they are different than in the global case—so this is the weak interpretation.

### 6 Summary

The semantics for conditionals given in terms of stochastic graphs (and the corresponding probability spaces) has several advantages:

- It allows us to exhibit the structure of the problem in a very clear way.
- It reveals the differences between the global and local interpretations of the conditional.

\(^{36}\) It can also be generalized in a straightforward way to the case of finitely (or countably infinitely) many atomic sentences.
• It is mathematically precise and allows for a straightforward computation of the
needed probabilities.

But this model reveals its full power when we turn to more complex issues. In our
opinion, promising areas of research (where we already have obtained some results)
are:

• Generalizing the notions of the local and global interpretation (we consider a more
general case of a certain accessibility relation between the worlds).
• The description of nested conditionals.37
• Making the models more realistic by considering not only mathematical prob‑
ability, but also “practical probability, where the time factor is taken into account
(we will not bet on a game which we will almost surely win—but only after a bil‑
lion years!).

Additionally, there is also a clear metaphilosophical gain: our models show how
certain vaguely understood philosophical notions can be given formal, fully precise
counterparts. They can be called explications (in Carnap’s sense). When turning to
more complex issues, more powerful mathematical methods will be used and they
will provide additional explanatory power for these problems. Our results can be
also viewed as a contribution to the idea of mathematical philosophy (Leitgeb 2013).
So—in a sense—this paper contributes also to the more general discussion concern‑
ing the role of mathematical methods in analyzing philosophical problems.

Acknowledgements The preparation of this paper was supported by National Science Centre (NCN)
grant 2016/21/B/HS1/01955. For helpful discussions and critical comments, we would like to thank the
audience in MCMP in Munich. We also thank the anonymous referee for insightful comments which
allowed to remove some weaknesses of the paper.

37 We can model right‑nested conditionals $B \rightarrow (C \rightarrow D)$ with graphs in a straightforward way. Think of
a toy example of drawing colorful balls which have a heavy or light mass. The conditional If the ball
is Heavy, then if it is not White, it is Green has two interpretations (deep and shallow), and the difference can by represented by the corresponding graphs, which exhibit the dynamics of the game. The Stal‑
naker Bernoulli spaces account imposes the shallow interpretation, which is unnatural in some cases: for example, the model gives very counter‑intuitive results for sentences like If the match is wet, then if you
strike it, it will light. The causal random variable model in Kaufmann (2009) covers the deep interpreta‑
tion (this interpretation agrees with McGee’s account from McGee (1989); without going into details,
the deep interpretation imposes the Import–Export principle). But they are two different models and an
advantage of our model is that it can account in a uniform way for both the deep and shallow interpreta‑
tion of right‑nested conditionals. The method can also be extended to longer right‑nested conditionals
$A_1 \rightarrow (\ldots (A_{n‑1} \rightarrow A_n))$, and also for cases when conditional operators that occur in the long conditional
need different interpretations (i.e. some of them are deep and some are shallow). The rules for construct‑
ing appropriate graphs are straightforward; this is impossible to achieve within the Stalnaker Bernoulli
model (which by definition works only for the shallow interpretation), and we suppose that mixed condi‑
tionals are not easy to cover within the causal model (which is designed to cover the deep interpretation).
It is also possible to explain the interplay between the local/global and deep/shallow interpretations of
the compound conditional. Due to limitations of space it is not possible to give the technical details here;
they can be found in Wójcikowicz and Wójcikowicz (2019a).
Appendix

The Definition of a Markov Chain

Markov chains are random processes with discrete time (formally: sequences of random variables, indexed by natural numbers), which satisfy the property of memorylessness, i.e. the Markov property. It means, that predictions on the future of the process depend only on the present state, not on the full history of the process—i.e. on the particular path, by which the system got to its present state from the initial state.

A Markov chain is specified by identifying:

1. A set of states $S = \{s_1, \ldots, s_N\}$;
2. A set of transition probabilities $p_{ij} \geq 0$, for $i, j = 1, \ldots, N$—i.e. the probabilities of changing from state $s_i$ to state $s_j$. The probabilities obey the equations: $\sum_{k=1}^{N} p_{ik} = 1$, for $i = 1, \ldots, N$ (which means, that the probabilities of going from state $s_i$ to any of the possible states sum up to one).

Any Markov chain can be described by a N-by-N matrix, with the probabilities $p_{ij}$ being the respective entries. States with $p_{ii} = 1$ (which means, that the system will never get out of the state $s_i$) are called absorbing.

A very convenient feature of Markov chains is the possibility to give an intuitive picture of the process in form of a transition graph—which exhibits the behavior of the system. This is a simple illustration: imagine a fly is changing its position according to some transition probabilities. There are four states: 0, 1, 2, 3; a spider is lurking at the state 0, while something nice (for the fly) is waiting at state 3. So both 0 and 3 are absorbing states: when the fly gets there, the process terminates. The following matrix specifies the process:

| STATE | 0 | 1 | 2 | 3 |
|-------|---|---|---|---|
| 0     | 1.0 | 0 | 0 | 0 |
| 1     | 0.3 | 0.4 | 0.3 | 0 |
| 2     | 0 | 0.3 | 0.4 | 0.3 |
| 3     | 0 | 0 | 0 | 1.0 |

38 The simplest presentation of the needed notions from Markov chains theory (we know of) is to be found in chapter 7 of Bertsekas and Tsitsiklis (2008). Our spider-fly example is a slight modification of the original one.

39 Formally, a Markov chain is a sequence of random variables $X_0, \ldots, X_n, \ldots$, with values in the set of states $S$, which have the Markov property, i.e.

$P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_n = i) = p_{ij}$

for all times $n$, all states $i, j \in S$, and all possible sequences $i_0, i_1, \ldots, i_{n-1}$ of earlier states.
And the respective graph is Graph 6 (we use the more general version, with \(p, q, r\) instead of 0.4, 0.3, 0.3) (Fig. 6, next page).

In general, it is convenient to think of travelling within the graph, from one vertex to another, until the process stops. The vertices represent the states of the system, and the edges represent transitions: we can think of them as representing actions (like the gambler tossing a coin, the fly choosing what to do next, or—as in this paper—drawing a ball from the urn).

In the fly example, at each state, the fly can either: L—go left; W—wait; R—go right. When starting in 1, the sequence RWLWRR takes the fly to the state 3 (and the game terminates). Observe, that the “go left”, “wait”, “go right” actions for state 1 and 2 have the same probability (this assumption suits the purpose of the paper, as we assume that the respective actions within the games have constant probabilities). For notational convenience, take \(P(R) = \frac{p}{2}\); \(P(L) = \frac{q}{2}\); \(P(W) = \frac{r}{2}\).

The Probability Space \(\Xi^* = (\Omega^*, \mathcal{M}^*, P^*)\) Associated with the Markov Graph

We assume here, that there are two absorbing states in the graph (WIN, LOSS), and the process begins in the initial state START. So, in the spider-fly example: 
\(1 = \text{START}; 0 = \text{LOSS}; 3 = \text{WIN}\).

The \(\Omega^*\) of the constructed probability space consist of all the paths, which can be travelled within the graph, starting from START and ending in an absorbing state. The probability of a particular path is computed by multiplying the respective transition probabilities. For instance, if RWLWRR is one of the paths for the spider-fly example, its probability is given by \(P^*(\text{RWLWRR}) = prqrpp\).

Formally, the set of paths generated by the graph can be defined as the language accepted by the graph treated as a deterministic finite state automaton. The edges of the automaton are labeled by the actions, which lead from one vertex to another. In the gambler’s ruin example these actions are H and T; in the fly example they are R, W, L—and throughout the paper the actions are drawing the balls: \(W(\text{hit})\),

---

\[40\] Indeed, in our case it is natural to think of some very concrete actions leading from state to state. But in general, it is enough only to say, that there occurs a transition with a certain probability.
A Stochastic Graphs Semantics for Conditionals

R(ed), G(reen). As we only want to identify the paths generated by the graph, we can ignore the transition probabilities at this stage.\footnote{So we identify the set of paths generated by the graph by—in a sense—reversing the problem: consider any string of actions, and ask, whether it leads from START to an accepting state. Assume the fly starts in the state 1, and consider four sequences of actions: (1) RWR; (2) RLRLRLWWL; (3) WWWR; (4) RRRRR. We see, that (1) RWR leads to the accepting state 3, so this path is in $\Omega^*$; (2) RLRLRLWWL leads to the accepting state 0, so this path is in $\Omega^*$ as well; (3) WWWR leads to the state 2 (i.e. the game has not finished yet), so this path is not in $\Omega^*$; (4) is not in $\Omega^*$: we land in the accepting state 3 already after two R’s— but there are still two R’s left, so the string is too long. The paths generated by the graph are exactly these sequences of actions, which lead to an absorbing state in their final step. For formal definitions of a deterministic automaton and the language accepted by the automaton see for instance (Anderson 2006) or the classic (Hopcroft et al. 2001).}

The set of elementary events $\Omega^*$ in the constructed space is defined as the set of generated paths—i.e. the language accepted by the automaton. Computing their probabilities is straightforward: for any sequence of actions/labels $A_{i_1j_1} \ldots A_{i_nj_n}$ (which form the accepted word) we multiply the corresponding transition probabilities. If the action $A_{ij}$ leads from the state $s_i$ to the state $s_j$, then the transition probability is $p_{ij}$. So, the probability $P^*$ of the sequence $A_{i_1j_1} \ldots A_{i_nj_n}$ (being an elementary event in the new space) is:

$$P^*(A_{i_1j_1} \ldots A_{i_nj_n}) = p_{i_1j_1} \ldots p_{i_nj_n}$$

We have the $\Omega^*$ and the probabilities of elementary events. As $\Omega^*$ is countable (or finite in trivial cases), we take the respective $\sigma$-field $M^*$ to be the power set, i.e. $M^* = 2^{\Omega^*}$. $P^*$ is defined for elementary events, and it extends in a unique way to all sets $A \subseteq \Omega^*$ by the standard formula $P^*(A) = \sum_{\omega \in A} P^*(\omega)$. Finally, $\Xi^* = (\Omega^*, M^*, P^*)$ is the probability space corresponding to the graph.

Observe, that the probability space $\Xi^* = (\Omega^*, M^*, P^*)$ was constructed at the basis of a simple, initial probability space $\Xi = (\Omega, M, P)$, including the “atomic actions”. The space for the spider-fly example consisted of three elementary events $R, L, W$, with probabilities $P(R) = p$; $P(L) = q$; $P(W) = r$.

Absorbtion Probabilities of the Markov Chain

Consider the general case with $N$ states $s_1, \ldots, s_N$. We assume that there are two absorbing states, say $s_1$ and $s_N$. For the purpose of the presentation, it is convenient to identify $s_N$ with victory, and $s_1$ with loss. The process is described by the transition probabilities $p_{ij}$, for $i,j = 1,2,\ldots,N$. As $s_1$ and $s_N$ are absorbing states, it means that:

$$p_{11} = 1; \quad p_{1,i} = 0 \text{ for } i \neq 1.
\quad P_{NN} = 1; \quad p_{N,i} = 0 \text{ for } i \neq N.$$  

Let $P(i)$ be the probability of eventually reaching state $s_N$, starting from state $s_i$. By definition, $P(1) = 0$ and $P(N) = 1$ (the game is already over).

**Fact** The probabilities $P(i)$ are the unique solution of the system of equations:
\[ P(1) = 0 \]
\[ P(N) = 1 \]
\[ P(i) = \sum_{k=1}^{N} p_{ik} P(k), \text{ for } i = 2, \ldots, N-1 \]

So, we have in general a linear system of \( N \) equations with \( N \) variables, and the solution is straightforward. It is also a basic result, that in such a graph, the probability, that the process will be eventually absorbed is 1. This is enough for our purposes.\(^{42}\)

We end with the observation, that for absorbing Markov chains (like the ones considered here), the probability of absorption is 1. So if only it is possible for the process to be absorbed, it eventually will be absorbed. This means, that we do not have to worry about infinite paths—their probabilities are 0, and they need not be included in our \( \Omega^* \).

### References

Adams, E. W. (1965). On the logic of conditionals. *Inquiry*, 8, 166–197.

Adams, E. W. (1966). Probability and the logic of conditionals. In J. Hintikka & P. Suppes (Eds.), *Aspects of inductive logic* (pp. 265–316). North Holland: Amsterdam.

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.

Anderson, J. A. (2006). *Automata theory with modern applications*. Cambridge: Cambridge University Press.

Bertsekas, D. P., & Tsitsiklis, J. N. (2008). *Introduction to probability* (2nd ed.). Nashua: Athena Scientific.

Brun, G. (2016). Explication as a method of conceptual re-engineering. *Erkenntnis*, 81(6), 1211–1241.

Carnap, R. (1950). *Logical foundations of probability* (2nd ed.). Chicago/London: University of Chicago Press/Routledge and Kegan Paul.

Feller, W. (1968). *An introduction to probability theory and its applications* (3rd ed., Vol. 1). New York: Wiley.

Grinstead, C. M., & Snell, J. L. (1997). *Introduction to probability*. Providence, Rhode Island: AMS.

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

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: Cambridge University Press.

Hall, N. (1994). Back in the CCCP. In E. Eells & B. Skyrms (Eds.), *Probabilities and conditionals: Belief revision and rational decision* (pp. 141–160). Cambridge: Cambridge University Press.

Hopcroft, J. E., Motwani, R., & Ullman, J. D. (2001). *Introduction to automata theory, languages and computation* (2nd ed.). Reading: Addison-Wesley.

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.

---

\(^{42}\) More general results stated in the language of matrix equations can be found for instance in chapter 11 of Grinstead and Snell (1997). The classic reference, with an extensive discussion of the motivations is Feller (1968).
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). New York: Springer.

Khoo, J. (2015). On indicative and subjunctive conditionals. *Philosopher’s Imprint, 15*(32), 1–40.

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

Leitgeb, H. (2013). Scientific philosophy, mathematical philosophy, and all that. *Metaphilosophy, 44*(3), 267–275.

Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. *Philosophical Review, 85*, 297–315.

Lewis, D. (1986). *Philosophical papers* (Vol. 2). Oxford: Oxford University Press.

McGee, V. (1989). Conditional probabilities and compounds of conditionals. *Philosophical Review, XCVIII*, 4, 485–541.

Ramsey, F. P. (1929). General propositions and causality. In Mellor, D. H. (Ed.), (Reprinted in: *Philosophical papers*, pp. 145–163, Cambridge: Cambridge University Press 1990).

Ross, S. M. (2007). *Introduction to probability models* (9th ed.). New York: Academic Press.

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

Stalnaker, R. (1970). Probability and conditionals. *Philosophy of Science, 37*(1), 64–80.

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* (Vol. 1, pp. 261–308)., The University of Western Ontario Series in Philosophy of Science Dordrecht: D. Reidel.

Wójtowicz, A., & Wójtowicz, K. (2019a). A graph model for probabilities of nested conditionals (submitted).

Wójtowicz, A., & Wójtowicz, K. (2019b). Mathematical versus practical probabilities of conditionals: A stochastic graph approach (in preparation).

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