Conditionals as representative inferences

van Rooij, R.; Schulz, K.

Link to publication

ORIGINAL PAPER

Conditionals As Representative Inferences

Robert van Rooij[1] · **Katrin Schulz**[1]

Abstract

According to Adams (Inquiry 8:166–197, 1965), the acceptability of an indicative conditional goes with the conditional probability of the consequent given the antecedent. However, some conditionals seem to be inappropriate, although their corresponding conditional probability is high. These are cases with a missing link between antecedent and consequent. Other conditionals are appropriate even though the conditional probability is low. Finally, we have the so-called biscuit conditionals. In this paper we will generalize analyses of Douven (Synthese 164:19–44, 2008) and others to account for the appropriateness of conditionals in terms of evidential support. Our generalization involves making use of *Value*, or intensity. We will show how this generalization helps to account for biscuit conditionals and conditional threats and promises. Finally, a link is established between this analysis of conditionals and an analysis of generic sentences.

Keywords Conditionals · Relevance · Biscuit conditionals · Threats

1 Introduction

The (standard and strict) material accounts of indicative conditionals have a well-known problem: the truth (or known truth, for the strict material account) of the consequent is sufficient to warrant the truth/acceptability of the conditional. As such, these theories have a hard time explaining what is wrong with a conditional like

✉ Robert van Rooij
 R.A.M.vanRooij@uva.nl

 Katrin Schulz
 K.Schulz@uva.nl

1   Institute for Logic, Language and Computation, University of Amsterdam, Amsterdam,
 Netherlands

(1)   If it is sunny today, Ajax won the Champions League in 1995.

especially given that Ajax won the Champions League in 1995. More modern theories of conditionals such as the similarity-based account (e.g., Stalnaker 1968), the information-based account (e.g., Veltman 1985), or the probabilistic account (Adams 1965; Stalnaker 1970) still have a similar problem: if the antecedent and consequent are (known to be) true, the conditional is predicted to be true, or acceptable, as well. None of these analyses account for the intuition many people have that the truth or acceptability of a conditional relies on a *dependence* of the consequent on the antecedent. Other theories do demand a link between antecedent and consequent. Relevance theorists (Anderson and Belnap 1962; Urquhart 1972; Restall 1996) claim that the link should be either one of overlapping *aboutness*, or of the use of (part of) the antecedent for the proof of the consequent, while others (Krzyżanowska et al. 2013; Douven et al. 2018) claim that for acceptability of an indicative conditional we have to be able to *infer* the consequent from the antecedent.

In this paper we argue that the problem of the missing link (cf. Douven 2008) is not restricted to 'standard' conditionals like (1). Also for biscuit conditionals and conditional threats and promises, for instance, there should be a link between antecedent and consequent for the conditional to be appropriate. This already suggests that if we want a more uniform analysis of (indicative) conditionals, the above theories that account for a link are not general enough. We propose such a more uniform analysis of (indicative) conditionals that accounts not just for 'standard' (indicative) conditionals, but for other types of indicative conditionals as well. The analysis builds on the notion of 'contingency', or 'relevance', that was already used by Douven (2008) to explain what is wrong with 'missing link' conditionals. The notion of contingency was originally introduced in learning psychology to measure the learnability of a dependence between the two features. We will propose that for the appropriateness of a conditional, the conditional probability of the consequent on the antecedent has to be weighted by their contingency (**basic proposal**). This will be shown to solve the 'missing link' problem discussed above (**application 1**). Drawn on experimental results on learning, we will motivate an extension of contingency to what we will call 'representativenes' of C for A. This extension will allow us to account for biscuit conditionals (**application 2**), conditional threats and promises (**application 3**), and perhaps even anankastic- and 'even if'-conditionals (**applications 4 and 5**). We will consider in how far this proposal can serve as a general analysis of the meaning of conditional sentences. Finally, we will point out the close link the present analysis draws between conditionals and generic sentences (**application 6**). In the final main section, we provide more detail of how our notion of representativeness is related with learning, and explain why representativeness is often confused with probability.

2 From Learning to Representativeness

Within classical learning-by-conditioning psychology, learning a dependency between two events C and A is measured in terms of the <u>contingency</u> ΔP_A^C of one event on the other (cf. Rescorla 1968):[1]

$$\Delta P_A^C = P(C|A) - P(C|\neg A), \quad \text{where } P \text{ measures frequencies.}$$

Contingency not simply measures whether the probability of C given A is high, but whether it is high compared to the probability of C given all other (contextually relevant) cases than A ($\neg A$ abbreviates $\bigcup Alt(A)$). Thus, it is measured how representative or typical C is for A. Rescorla (1968) showed that rats learn a $\langle tone, shock \rangle$ association if the frequency of shocks immediately after the tone is higher than the frequency of shocks undergone otherwise, even if shocks occur only in, say, 12% of the trials in which a tone is present.[2] Gluck and Bower (1988) show that contingency is crucial for human associative learning as well.

Experiments in the aversive (i.e., fear) conditioning paradigms (e.g., Annau and Kamin 1961; Forsyth and Eifert 1998) show that the speed of acquisition and the strength of the association in rats increases with the *intensity* of the shock. Slovic et al. (2004) show, similarly, that people build stronger associations related to events with high emotional impact. To capture this we introduce a new measure, the **representativeness** ∇P_A^C, defined as below, where $V(C|A)$ measures the absolute value (or intensity) of C given A.

$$(RES) \quad \nabla P_A^C = P(C|A) \times V(C|A) - P(C|\neg A) \times V(C|\neg A).$$

The value of C given A measures something like (the absolute value of) a conditional utility, or conditional preference. Although in some applications we assume that $V(C|A) = V(C)$, in other applications the conditionality of the utility is important.[3] Although somewhat unusual, conditional utilities have been used before, e.g., by Armendt (1988). For Armendt, $V(C|A)$ measures the present utility for C under the hypothesis A, which need not be identical to the utility the agent *would have* if A were true, or if (s)he came to believe A. We (mostly) think of (conditional) utilities as *experienced utilities* as originally thought of by Bentham (1824/1987). Although for a long time such experienced utilities were thought of as unmeasurable and thus unscientific, this opinion changed significantly more recently: several measures (some involving dopamine) are used nowadays to measure experienced joy and (experienced) fear within conditioning psychology, while due to the work of Kahneman and his collaborators (e.g., Kahneman et al.

---

[1] For a causal derivation of ΔP_A^C see Pearl (2000).

[2] Cheng and Holyoak (1995) point out that it is important that the background conditions should be kept constant when measuring contingency.

[3] Although many times $V(C|A) = V(C)$ or $V(A \wedge C)$, this is clearly not always the case. As a result, $V(C|A)$ can be very different from $V(A|C)$. An example of Armendt (1988) illustrates this: my utility for having medical insurance, under the hypothesis that I am hospitalized, is considerably greater than my utility for being hospitalized, under the hypothesis that I have insurance. We will also assume (with Armendt 1988), that even if $P(C|A) = 1$, it still might be that $V(C|A)$ can be very high.

1997), experienced utility became a respectable notion even in economics. By making use of experienced instead of revealed utilities, we propose to make a link between standard decision theory and the use of intensity in learning-by-conditioning psychology. We will assume that in many circumstances, or that per default, $Value(C|A) = 1 = Value(C|\neg A)$, meaning that under normal circumstances our notion of representativeness reduces to contingency, $\nabla P_A^C = \Delta P_A^C$.

3 Conditionals As Representative Inferences

ΔP_A^C is a measure of the probabilistic dependence between C and A. To overcome the missing link problem of approaches to indicative conditionals of the form $A \Rightarrow C$, one might therefore suggest to use ΔP_A^C or (RES) to check the acceptability of a conditional sentence. Indeed, Douven (2008) uses the measure $P(C|A) - P(C)$ for these purposes, and it is easy to prove that $P(C|A) > P(C)$ iff $\Delta P_A^C > 0$.[4] An advantage for using ΔP_A^C is that this measure has the maximal value, i.e., 1, if and only if $P(C|A) = 1$ and $P(C|\neg A) = 0$. But this holds exactly whenever 'If A, then C' is strengthened to 'A if and only if C', a strengthening often observed for indicative conditionals under the name of 'conditional perfection' (cf. Geis and Zwicky 1971). However, Skovgaard-Olsen et al. (2016) show that although $\Delta P_C^A > 0$ is a necessary condition for acceptability of indicative conditionals, it is not a sufficient one: it is also demanded that $P(C|A)$ is high. To account for that, one can make use of the following condition:[5]

$$(CON') \quad A \Rightarrow C \text{ is acceptable} \quad \text{iff} \quad \frac{\Delta P_A^C}{1 - P(C|\neg A)} \text{ is high.}$$

This latter measure is known in the literature as the measure of *relative difference* (Shep 1958). Cheng (1997) uses it to measure causal strength and shows that for this measure, $P(C|A)$ counts for more than $P(C|\neg A)$. This captures part of the intuition

---

[4]

$P(C|A)$	$>$	$P(C)$	iff
$P(C|A)$	$>$	$(P(A) \times P(C|A)) + (P(\neg A) \times P(C|\neg A))$	iff
$(1 - P(A)) \times P(C|A)$	$>$	$(1 - P(A)) \times P(C|\neg A)$	iff
$P(C|A)$	$>$	$P(C|\neg A)$	iff
ΔP_A^C	$>$	$0.$	

Note that $P(C|A) - P(C)$ and ΔP_A^C might still have different numerical values. Skovgaard-Olsen et al. (2017) found experimentally that the measure ΔP_A^C accounts better for the perceived degrees of relevance of the participants than $P(C|A) - P(C)$.

[5] There are two things noteworthy about (CON') that we won't discuss in this paper: (i) this rule talks about acceptability, and not about truth, and (ii) it doesn't say that a conditional is more acceptable with increasing $\frac{\Delta P_A^C}{1 - P(C|\neg A)}$. As for (i), we think that it is natural that our measure only accounts for the acceptability of indicative conditionals, not for their truth values (if they have that). With respect to (ii), based on experimental data of Skovgaard-Olsen (2016) and others, it seems that acceptability is graded. But we won't argue for that here. Our intuition that acceptability is graded influenced us enough, however, not to be more specific on what it means to be 'high'.

that for $A \Rightarrow C$ to be acceptable, it should (normally) be the case that $P(C|A)$ is high. (see Sects. 4 and 5 for more on this).

For the general case, however, we should not look only at informational value: utility, or emotional value, counts as well. Therefore, we propose the following generalization of (CON') as our general condition (with $EV(C|\neg A)$ as an abbreviation for $P(C|\neg A) \times V(C|\neg A)$):

$$(CON) \quad A \Rightarrow C \text{ is acceptable} \qquad \text{iff} \qquad \frac{\nabla P_A^C}{\max\{1, V(C|A)\} - EV(C|\neg A)} \text{ is high.}$$

Notice that if *Value* is irrelevant (meaning that $V(C|A) = V(C|\neg A) = 1$), for acceptability it is a *necessary* condition that $\Delta P_A^C > 0$. Moreover, under these circumstances, (CON) comes down to the simpler condition (CON') above.

4 Applications

4.1 Application 1: The Missing Link Problem

Already contingency accounts for conditionals like (1) (cf. Douven 2008; Skovgaard-Olsen 2016; Skovgaard-Olsen et al. 2016). If antecedent and consequent are probabilistically independent, we get $\Delta P_C^A = P(C|A) - P(C|\neg A) = 0$. If *Value* doesn't count, it follows from independence that $\Delta P_A^C = \nabla P_A^C = 0$. Hence, we predict that even in case $P(C) = 1$ (and perhaps $P(A) = 1$) and, therefore, $P(C|A) = 1$, the conditional (1) is not appropriately acceptable. As noted above, we believe that contingency is not the appropriate measure to account for indicative conditionals: $P(C|A)$ should count for more than $P(C|\neg A)$. For this reason, (CON') seems to be preferred to contingency. But there is more that speaks in favor of (CON'): As shown by Cheng (1997) and Pearl (2000), $\frac{\Delta P_A^C}{1 - P(C|\neg A)}$ follows from a causal analysis (under some natural assumptions). Cheng calls the measure 'causal strength', while Pearl (2000) refers to the measure as the 'probability of causal sufficiency'. By thinking of things in this way, what is missing in missing link conditionals, is a causal connection between antecedent and consequent, or so van Rooij and Schulz (2019a) argue. van Rooij and Schulz (2019a) use this causal view behind the measure $\frac{\Delta P_A^C}{1 - P(C|\neg A)}$ also to show that under various natural circumstances (e.g., if A is (thought to be) the only cause of C, or if the potential causes of C are mutually inconsistent), acceptability of conditionals can be measured by conditional probability, suggesting that the original proposals of Adams (1965) and Stalnaker (1970) were not far off.[6]

---

[6] In fact, also $\frac{\Delta P_A^C}{1 - P(C|\neg A)}$ comes down to conditional probability if $Alt(A) = \emptyset$, at least if one stipulates that $P(C|\bigcup Alt(A)) = 0$, if A has no alternative.

4.2 Application 2: Biscuit-Conditionals

To account for missing-link conditionals we argued that the value of $P(C|A)$ should be higher than that of $P(C|\neg A)$ (or of $P(C)$). But there are obvious exceptions to this. Most prominently: Austin's (1961) *biscuit conditionals*:[7]

(2) a. There are biscuits on the sideboard, if you want some.
     b. If you are interested, there's a good documentary on BBC tonight.
     c. If you need help, my name is Sue.

Iatridou (1991) and others claim that in a biscuit conditional, the if-clause specifies the circumstances in which the consequent is relevant. DeRose and Grandy (1999) seek to account for this by proposing a conditional assertion analysis of biscuit conditionals. According to such an analysis (cf. de Finetti de Finetti 1936/1995; Belnap 1970), the conditional 'If A, C' states that C is true, if A holds, and doesn't say anything otherwise. Belnap (1970) himself, however, already argued against such an analysis for biscuit conditionals:

> But I do know that "There are biscuits on the sideboard if you want some" is *not* generally used as a conditional assertion; for if there are no biscuits, even if you don't want any, it is plain false, not nonassertive. (Belnap 1970, p. 11).

We agree with Belnap's intuition. Franke (2007) argues that semantically speaking, biscuit conditionals could just be analyzed as material or strict implications. He proposes to use pragmatics (using a qualitative or quantitative notion of independence), instead, to explain why (2-a), for instance, entails that there are biscuits on the sideboard. This proposal is certainly appealing, but as noted by Lauer (2015), this analysis by itself still leaves open what it is that makes the antecedent relevant to the consequent. Indeed, what we need is both (i) *epistemic* independence (e.g., $P(C|A) = P(C)$ and thus $\Delta P_C^A = P(C|A) - P(C|\neg A) = 0$), without giving up that (ii) the antecedent is still of value to the consequent. Our analysis (CON) captures this.

To see this, notice that in the relevant situation the biscuits are on the sideboard, independently of whether you want some or not. Thus $\Delta P_A^C = P(C|A) - P(C|\neg A) = 0$. What makes the antecedent still of value for the consequent? Right, high $V(C|A)$! If you want biscuits, it is important to know that the biscuits are easy to take: they are just there on the sideboard. Similarly for (2-b)–(2-c). Thus, for biscuit conditionals the *Value* in the definition of representativeness ∇P_A^C matters. In (2-a)–(2-c), learning the truth of the consequent is of little or no value if the antecedent is false, but this value is high if the antecedent is true. Hence, $V(C|A) >> V(C|\neg A) \approx 0$. As a result, $\nabla P_A^C = P(C|A) \times V(C|A) - P(C|\neg A) \times V(C|\neg A)$ will be high, and this explains the appropriateness of the conditional.

---

[7] A number of authors (e.g., Geis and Lycan 1993) have argued that biscuit conditionals are not 'real' conditionals, but only share the surface structure of conditionals. We are not convinced by these arguments and take it that a uniform analysis is preferred, if possible. Our goal is to show that such a (more) uniform analysis is, indeed, possible.

Notice that in (CON) we used $\max\{1, V(C|A)\} - EI(C|\neg A)$ in the denominator, and not simply $1 - P(C|\neg A)$. Although the former comes down to the latter in natural circumstances—i.e., when $V(C|A) = V(C|\neg A) = 1$—, it is crucial for biscuit conditionals that we used the more general formula. The reason is that for biscuit conditionals $P(C|\neg A) = 1$, meaning that $1 - P(C|\neg A) = 0$ and thus that the fraction would not be defined if we used $1 - P(C|\neg A)$ as denominator. As we noticed above, for biscuit conditionals it might be that $V(C|\neg A) = 0$ and $V(C|A) = 1$, meaning that $\frac{\nabla P_A^C}{\max\{1, V(C|A)\} - EI(C|\neg A)}$ reduces in those cases to $\frac{P(C|A) \times V(C|A)}{V(C|A)} = P(C|A)$, which, in turn, typically will have value 1 for a good biscuit conditional. We have seen already that if $V(C|A) = V(C|\neg A)$, it will be the case that $\nabla P_A^C = 0$, because for biscuit conditionals $P(C|A) = P(C|\neg A)$, and thus that the conditional is unacceptable.

4.3 Application 3: Conditional Threats and Promises

Our analysis works, or so we think, also for conditional threats, promises and warnings:

(3) a. If you won't give me your wallet, I will kill you.
  b. If you give me 10.000 euros, I will destroy the (for you hazardous) tapes.
  c. If you go to New York, watch out for the taxi drivers.

We take it (following Schelling, 1960 and many others) that conditional threats and promises are used strategically in order to influence the hearer's behaviour: the speaker wants the addressee to give him (or her) the wallet or the 10.000 euros, and the threat and promise states what the speaker will 'offer' in return. What needs to be explained for such conditionals is that addressees many times 'accept' them, although these threats and promises are not very credible (cf. Schelling, 1960; Hirschleifer 1991). Would it really be rational for the threatener to kill the addressee if the latter doesn't give the former his or her wallet? And once (s)he has the 10.000 euros in his pocket, why would the promiser still destroy these valuable tapes? Thus, although the speaker of (3-a) and (3-b) seems to commit him or herself to a particular action conditional on the antecedent, why should (s)he stick to his or her commitment?

Indeed, for the addressee $P(C|A)$ is typically not very high.[8] However, for both (3-a) and (3-b), the probability of the consequent given $\neg A$, $P(C|\neg A)$ will certainly not be higher than given A (certainly if the speaker is, or pretends to be, desperate or irrational enough). As a result, $P(C|A) - P(C|\neg A) > 0$. On our analysis this is not enough for the conditionals to be acceptable. What we need for that is that the value

---

[8] One reviewer claimed that promises and threats are not truth apt, and wonders how, then, $P(C|A)$ should be determined. But we don't see why, e.g., 'I will kill you' cannot have a truth-value in the circumstances where you do, or do not, give me your wallet.

of C (given A), $V(C|A)$, is high.[9] It is natural to assume that in these conditionals, the emotional impact of the consequent is independent of the antecedent. Thus, representativeness reduces to $\Delta P_A^C \times V(C)$. Given that in these cases $V(C)$ is extremely high for the addressee, it follows that ∇P_A^C will be high, even if $P(C|A) - P(C|\neg A)$ is low. Thus, these conditional threats/promises are accepted, as long as the stakes communicated in the consequent are high enough.[10]

The reader must have noticed that for our analysis of conditional threats and promises, it is the *addressee's* probabilities and utilities that count, not those of the speaker, as is normally assumed for analyses of indicative conditionals. Indeed, we think that in contrast to standard (indicative) conditionals, the addressee's attitudes are crucial to account for the acceptability of conditional threats and promises. One might wonder[11] to what extent one can then still speak of a (more) 'uniform' analysis? We think that our analysis of conditional threats and promises is still part of a uniform analysis, if we take seriously the use of 'you' in the antecedent of the conditionals. What this indicates, or so we would like to propose, is that the perspective is shifted from the speaker to the addressee. We don't have a worked-out theory of when and how such a shift of perspective will take place, but it seems natural to us that such a shift is needed to account for conditional speech acts like (3-a) and (3-b).

What about conditional warnings? For these, it seems it is the difference between $V(C|A)$ and $V(C|\neg A)$ that counts. The speaker of (3-c) seems to intend to communicate that it is useful for the addressee to know that taxi drivers are more dangerous in New York city that in the addressee's hometown.

4.4 Applications 4 and 5: Anankastic and Even-if Conditionals

According to Kratzer (1991) (following Lewis 1975), conditional sentences of the form 'If A, then C' should be represented logically by 'Quantifier + if A, C'. A logical form like 'Most + If A, then C' and 'Must + If A, then C' are then interpreted roughly as follows: 'for most of the (selected) worlds in which A is true, C is true as well', and 'in all (selected) worlds in which A is true, C is true as well', respectively. One serious challenge for this analysis are so-called 'Anankastic' conditionals like the following:

---

[9] For conditional promises, like 'If you go out with me, I will buy you a drink' (provided by a reviewer), the use of *Value*$(C|A)$ is much less important. Notice that in these examples $P(C|A) - P(C|\neg A) \approx 1$. Thus, such a conditional promise can be acceptable even if the addressee is not interested in the speaker buying him or her a drink, meaning that $V(C|A) = 1 = V(C)$. According to one reviewer, (3-a) and (3-b) can be accounted for as well without making use of *Value*. The reason is that according to this reviewer once a speaker is credible, after the use of these conditionals the difference between $P(C|A)$ and $P(C|\neg A)$ will be close to 1. We don't agree: we would give our wallet even if the the probability of C would be increased in any significant way far below 1.

[10] We believe that these type of conditionals are actually more basic than the standard 'uninterested' ones. Notice that psychological research involving the Wason selection task clearly shows that participants perform much better for such 'real life' conditionals than for standard ones (cf. Johnson-Laird et al. 1972).

[11] As did some reviewers.

(4) a. If you want to go to Harlem, take the A-train.
     b. If you want sugar in your soup, ask the waiter.

Intuitively, (4-a) is true, or appropriate, just in case taking the A-train is the best, or most useful, way to go to Harlem. This intuition is captured to a large extent by saying that in all (selected) worlds in which you *go* to Harlem, you take the A-train. Unfortunately, this is not what Kratzer's analysis predicts. Her theory predicts that (4-a) is true just in case in all (selected) worlds in which you *want to go* to Harlem, you take the A train. Thus, for a Kratzer-like analysis of conditionals, the problem is one of compositionality: how to 'get rid' of the 'want' in the antecedent of the conditional (cf. Saebo 2001)? There is no shortage of proposals of how this should be done, but only seldomly, if ever, the similarity is observed, or made use of, between anankastic conditionals, on the one hand, and biscuit conditionals like (2-a), on the other.

Our (rather provisional, to be honest) analysis is different from Kratzer's. We would analyse anankastic conditionals similarly as we treated biscuit conditionals: the consequent is relevant for the hearer only in case the antecedent holds: if you want to go to Harlem, or sugar in your soup. Thus $V(C|A)$ should be high, or at least much higher than $V(C|\neg A)$. Of course, on our analysis we should take $P(C|A)$ and $P(C|\neg A)$ into account as well. But then, anankastic condtitionals are typically, if not always, used to give an *advice*. Typically, (4-a) is given as answer to a question like 'Which train should I take if I want to go to Harlem?' A questioner like that has little or no idea what is the best train to take, so the difference between $P(C|A)$ and $P(C|\bigcup Alt(A))$ is rather small, where $Alt(A)$ are the alternative destinies.[12] Thus, $P(C|A) - P(C|\neg A)$ is small, that is, not high enough for making the conditional acceptable. What makes the conditional acceptable is the difference between $V(C|A)$ and $V(C|\neg A)$, just like in the case of biscuit conditionals.

Skovgaard-Olsen et al. (2016)'s experiments suggest that relevance, or positive ΔP_A^C, is necessary for 'ordinary' indicative conditionals, but not for so-called 'even if'-conditionals like

(5) Mary comes, even if John comes.

According to them, the acceptability of 'even if' conditionals 'goes with' the corresponding conditional probability. We have argued above that under specific conditions our general measure $\frac{\nabla P_A^C}{\max\{1, V(C|A)\} - EI(C|\neg A)}$ comes down to the conditional probability $P(C|A)$. The most relevant case for our purposes seems the case where $V(C|\neg A) = 0$ (and $V(C|A) = 1$). Perhaps this is what is going on in 'even if'-conditionals like (5): we don't care whether Mary comes if John doesn't come, presumably because we know already that she would come in that case anyway.[13] The only interesting case is the one where John comes. Thus, under this proposal,

---

[12] Recall that we used $\neg A$ as an abbreviation for $\bigcup Alt(A)$.

[13] To be sure, we don't want to be committed to this suggestion, but we think it is an interesting idea to explore.

'even if'-conditionals have a lot in common with biscuit conditionals, although it doesn't have to be the case that $P(C|A) = 1$.

4.5 Application 6: Generics

Generics and conditionals are much alike. They both have at least the following purposes: (i) to state (inductive) *generalizations* ('Tigers are striped', 'If you push this button, the lamp will light'); (ii) to express (perhaps desired) *norms* ('Boys don't cry', 'If you see a general, you salute him'), and (iii) to express threatening cases like (iii) 'Pit bulls are dangerous dogs' and 'If you don't give me your wallet, I will kill you'. This suggests that they should be given very similar analyses. Indeed, just like there exists the missing-link problem for conditionals, generics of the form 'As are C' also seem to be acceptable (under normal conditions) only if being an A is *relevant* for having feature C. To show this, the following generic is generally taken to be inappropriate, because Germans are not special in terms of right-handedness:

(6)   ?Germans are right-handed.

As it turns out, in van Rooij (2017) and van Rooij and Schulz (2019b) (building on Cohen (1999) and Leslie (2008)) an analysis of generic sentences in terms of representativeness was indeed proposed: A generic sentence of the form 'As are C' was proposed to be true, or acceptable, iff C is a representative feature of As. It is shown that in terms of this analysis quite a number of examples can be accounted for that are problematic for more standard semantic analyses of generic sentences making use of *conditional probability* or *normality*. For instance, this analysis immediately accounts for generics like 'Ticks carry the Lyme disease' or 'Sharks attack swimmers' that are problematic for default-based approaches (e.g., Asher and Morreau 1995) and called '*striking generics*' by Leslie (2008), who notes that 'striking' often means 'horrific or appalling'. Observe that in case all features are equally important, it is predicted that a generic of the form 'As are C' is true iff ΔP_A^C is high, from which it follows that $\Delta P_A^C > 0$, which is exactly what Cohen (1999) demands for so-called 'relative generics' (e.g., 'Dutchmen are good sailors') to be true. Making use of ΔP_A^C one can explain, for instance, why the generic 'Ducks lay eggs' is predicted to be ok, although the majority of ducks don't lay eggs, and why (6) is a questionable generic, although most germans are right handed.

However, this analysis accounts as well for the intuition that standard generics like 'Birds fly' and 'Birds lay eggs' are acceptable and true (because 'flying' and 'laying eggs' are among the most distinguishable features for birds). Our weak analysis of generics also explains examples paradoxical for many other theories: First, although only (adult) male lions have manes, 'Lions have manes' is an accepted generic, but 'Lions are male' is not.[14] Our analysis thus correctly predicts that 'As are C' can be true and 'As are D' false, although $P(D|A) > P(C|A) < \frac{1}{2}$. Second, it explains why 'Peacocks lay eggs' and 'Peacocks have beautiful feathers'

---

[14] This example is strikingly similar to Kahneman and Tversky's (1972) famous *conjunction fallacy*. We will show that our analysis accounts for the two 'paradoxes' in the very same way, in terms of our implementation of Tversky and Kahneman's *representativeness*-heuristics making use of relevance.

are both considered true, although no peacock lays eggs (female) *and* has beautiful feathers (male). Both generics are predicted to be true simply because *relative to other animals (in general)*, many peacocks have the relevant features.

Thus, our proposal provides a *uniform analysis* of all types of examples discussed in this paper, including various types of indicative conditionals and generics. What this analysis of generics does not yet explain is why people typically *interpret* generics of the form '*As are C*' as saying that (almost) all *As* are *C*. In van Rooij and Schulz ([2019b](#)) it was argued that this was due to the fact that people confuse representativeness for conditional probability, and accounted for this making use of Tversky and Kahneman's ([1974](#)) 'heuristics and biases'-program. At this point, however, we think that the strong interpretation of generics can better be explained in terms of how we learn generalizations.[15]

5 Representativeness As Expectation

Even if hearers *accept* conditionals of the form 'If *A*, then *C*' due to our proposed weak acceptance rules, hearers still *interpret* conditionals typically in a much stronger way: the likelihood of *C* given *A* is high (Adams [1965](#)). Why? We think it has something to do with how we *learn* generalizations.

In behavioral psychology, the learning of generalizations, or expectations, was studied in classical conditioning (or Pavlovian conditioning). What is the expectation that the $n + 1$th cue a will be accompanied with consequence c?[16] The perhaps most natural idea would be that it is just the times that cue a was accompanied with consequence c divided by the times that cue a was given at all. If we say that $O_i(c|a) = 1$ if at the ith exposure cue a is accompanied with consequence c, and that $O_i(c|a) = 0$ if at the ith exposure cue a is not accompanied with consequence c, the expectation that the $n + 1$th cue a will be accompanied with consequence o, i.e., $P_{n+1}^*(c|a)$, can be stated as follows:

$$(RF) \quad P_{n+1}^*(c|a) = \frac{O_1(c|a) + \cdots + O_n(c|a)}{n}$$
$$= \frac{1}{n} \sum_{i=1}^{n} O_i(c|a)$$

It can be shown, however, that for the calculation of $P_{n+1}^*(c|a)$ it is not needed to maintain a record of all cases where cue a was accompanied with consequence c. One can calculate $P_{n+1}^*(c|a)$ incrementally as well, by constantly changing the expectations. This can be shown as follows (adapted from a very similar proof by Sutton and Barto ([2016](#))):

---

[15] We think that the notion of 'representativeness' that plays such an important role in Tversky and Kahneman ([1974](#)) comes about, for a large part, through learning, so we feel that the proposals of van Rooij and Schulz ([2019b](#)) and the current one are not incompatible at all.

[16] To be sure, 'consequence' should here not (by necessity) be given a causal interpretation.

$$P^*_{n+1}(c|a) = \frac{1}{n}\sum_{i=1}^{n} O_i(c|a)$$

$$= \frac{1}{n}\left(O_n(c|a) + \sum_{i=1}^{n-1} O_i(c|a)\right)$$

$$= \frac{1}{n}\left(O_n(c|a) + (n-1)\frac{1}{n-1}\sum_{i=1}^{n-1} O_i(c|a)\right)$$

$$= \frac{1}{n}\left(O_n(c|a) + (n-1)P^*_n(c|a)\right)$$

$$= \frac{1}{n}\left(O_n(c|a) + nP^*_n(c|a) - P^*_n(c|a)\right)$$

$$= P^*_n(c|a) + \frac{1}{n}\left(O_n(c|a) - P^*_n(c|a)\right)$$

Notice that the last incremental learning rule always gives rise to the relative frequency observed, with small demands on memory and computation power. It turns out that the form of this incremental learning rule is very common. It is known as learning by *expected error minimization* and is used in almost all modern methods of machine learning. The general form of such rules is as follows:

$$NewExpectation \quad \longleftarrow \quad OldExpectation + Stepsize[Target - OldExpectation]$$

The *Stepsize* is also know as the learning rate. In the case above this was $\frac{1}{n}$, but many times this is taken to be a small constant. The *Target* is the value of the new observation, $O_i(c|a)$. Above, the target was 1 or 0, but this could in general be anything you want. In particular, it could depend on the intensity of the consequence. Indeed, because $P^*_{n+1}(c|a) = \frac{1}{n}\sum_{i=1}^{n} O_i(c|a)$, if $O_i(c|a)$ is high for each $i \leq n$ where a is accompanied with c, $P^*_{n+1}(c|a)$ will clearly be high as well, and much higher than the conditional frequency, in particular.

As we saw in Sect. 2, Rescorla (1968) observed that rats learn a tone (cue/cause)-shock (outcome/consequence) association if the frequency of shocks immediately after the tone is higher than the frequency of shocks undergone otherwise. This holds, even if in the minority of cases a shock actually follows the tone. Gluck and Bower (1988) and others show that humans learn associations between the representations of certain cues (properties or features) and consequence (typically another property or a category prediction) in a very similar way. Thus, we associate consequence c with cue a, not so much if $P(c|a)$ is high, but rather if $\Delta P^c_a = P(c|a) - P(c|\neg a)$ is high.[17] How can this be explained? Rescorla and Wagner (1972) show that this can be explained by an error–based learning rule very similar to the one above. The only thing that really changes is that this time the learning rule is also *competition-based*. The idea is that a cue can also be taken as a *combination* of separate cues: if a_1 and a_2 are cues, $a_1 a_2$ is taken to be a cue as well, and they all could be accompanied with the same outcomes. According to Rescorla and Wagner

---

[17] A reviewer of this paper noted that this connection between relevance-based accounts of conditionals and Rescorla and Wagner's (1972) was briefly noted by Skovgaard-Olsen, (2015) as well.

([1972](#)), we should keep track of expectations, or associations, for cue-action pairs for all primitive cues, i.e., a_1 and a_2. For the calculation of this expectation $E_{n+1}^*(c|a_1)$ after the nth trial, however, we should also look at $E_{n+1}^*(c|a_2)$ in case the actual cue at the nth trial is the combined cue $a_1 a_2$. The famous Rescorla–Wagner learning rule (RW) for each primitive cue a_i is stated as follows, if at the nth exposure (perhaps complex) cue a^* is given of which a_i is 'part' (where $j \preceq a^*$ holds if a_j is part of the (perhaps) complex cue a^*):

$$(RW) \quad E_{n+1}^*(c|a_i) = E_n^*(c|a_i) + \lambda \left(O_n(c|a^*) - \sum_{j \preceq a^*} E_n^*(c|a_j) \right)$$

Here, $E_{n+1}^*(c|a_i)$ is the agent's expectation after n observations that the $n+1$th primitive cue a_i has outcome c, where λ is a learning rate (typically very small) and where $O_n(c|a^*)$ measures the magnitude of the reinforcement at the nth trial where cue a_i was involved.[18] Notice that the cue at the nth trial could be just a primitive cue, but it could be a combined cue as well. If the nth cue is a combined cue like $a_1 a_2$, $\sum_j E_n^*(c|a_j) = E_n^*(c|a_1) + E_n^*(c|a_2)$, will obviously be larger than $E_n^*(c|a_i)$, and this has interesting consequences. For instance, if our learner is conditioned with the cue-outcome/consequence pairs $a_1 a_2 \to c$ and $a_2 \to \neg c$ that alternate each other, in the long run it will be that $E^*(c|a_1) = 1$ and $E^*(c|a_2) = 0$. Thus, a_1 is associated with consequence c, and cue a_2 is not associated with this consequence at all, although in half of the cases that cue a_2 was involved, consequence c appeared. The opposite is predicted if the learner is conditioned with the cue-consequence pairs $a_1 a_2 \to c$ and $a_2 \to c$ that alternate each other. In that case it will be that in the long run $E^*(c|a_1) = 0$ and $E^*(c|a_2) = 1$. Notice that these predictions are in accordance with what is predicted by the contingency rule, insofar as that in the first case $\Delta P_{a_1}^c = 1$, while in the second case $\Delta P_{a_1}^c = 0$.

More in general, Cheng ([1997](#)) shows that if the alternative cues for c are incompatible with a, $E_{n+1}^*(c|a)$ converges to the actual conditional probability (or relative frequency). If alternative cues are compatible with a, however, $E_{n+1}^*(c|a)$ yields, instead, $\Delta P_a^c = P(c|a) - P(c|\neg a)$ in the long run (see also Danks [2003](#)). If the value $O(c|a)$ is higher than 1 (in terms of the previous sections, this means that $V(c|a) > 1$), $E_{n+1}^*(c|a)$ converges to the actual average conditional impact, $EV(c|a) = P(c|a) \times O(c|a)$, if cues are mutually incompatible, and to something closer to $EO(c|a) - EO(c|\neg a)$ otherwise. Thus, in many cases expectations, or associations, as generated by rule (RW) do not really measure probabilities; they measure something quite different. Still, it is only natural that people take this 'something quite different', i.e., the associations, to be the conditional likelihood. In fact, according to, e.g., Newel et al. ([2007](#)), we can explain many of the problematic probability judgements as found in, e.g., Tversky and Kahneman ([1974](#)) by the assumption that people confuse probabilities with associations as established via associative learning mechanisms like (RW).

---

[18] Take a_i to be a_1. Then it could be that the actual cue was $a_1 a_2$ and that $O_n(c|a_1 a_2) = 1$, although $O_n(c|a_1)$ would be 0.

Rule (RW) is only the simplest associative learning rule, and many variants have been proposed over the years (for instance with time- and cue-dependent learning rates, or where uncertainty of cues is taken into account), variants that give rise to (sometimes slightly) different convergence results. Yuille (2006), for instance, shows that there is a learning rule closely related to (RW) that converges to $\frac{\Delta P_A^C}{1-P(C|\neg A)}$, i.e., the measure we used in acceptability rule (CON'). Most of these alternative learning rules have in common, however, that although they measure expectation, or association, in the long run they don't end up with the relative frequency $P(C|A)$ if there is competition between cues for the same outcome. In this way we explain why hearers *accept* conditionals of the form 'If A, then C' on relatively weak conditions, but that hearers still *interpret* conditionals typically in a much stronger way: the expectation of C given A is high.

6 Conclusion

In this paper we have proposed a uniform analysis of conditionals making use of a notion of 'representativeness'. We have suggested that the proposed analysis can account for many examples standard analyses have problems with. The proposed analysis gives rise to rather weak acceptability conditions. We have suggested that the feeling that conditionals are typically interpreted in a stronger way is due to the way expectations are formed, via something like the Rescorla and Wagner (1972) competition-based learning rule. In this way, the intuition that the acceptance of the conditional 'goes by' conditional expectation can be explained as well.

References

Adams EW (1965) A logic of conditionals. Inquiry 8:166–197
Annau Z, Kamin L (1961) The conditioned emotional response as a function of intensity of the US. J Comp Physiol Psychol 54:428–432
Anderson AR, Belnap N (1962) Tautological entailment. Philos Stud 13:9–24

Armendt B (1988) Conditional preference and causal expected utility. In: Harper W, Skyrms B (eds) Causation in decision, belief change and statistics. Kluwer, Dordrecht, pp 3–25

Asher N, Morreau M (1995) What some generic sentences mean. In: The generic book. University of Chicago Press, pp 300–339

Austin JL (1961) 'Ifs and cans'. In: Philosophical papers. Oxford University Press, Oxford, pp 153–180

Belnap N (1970) Conditional assertion and restricted quantification. Noûs 1:1–12

Bentham J (1824/1987) An introduction to the principles of morals and legislation. In: Mill JS, Bentham J (eds) Utilitarianism and other essays. Harmandsworth, Penguin

Cheng P (1997) From covariation to causation: a causal power theory. Psychol Rev 104:367–405

Cheng P, Holyoak K (1995) Adaptive systems as intuitive statisticians: causality, contingency, and prediction. In: Meyer J, Roitblat H (eds) Comparative approaches to cognition. MIT Press, Cambridage

Cohen A (1999) Think generic! The meaning and use of generic sentences. CSLI Publications, Stanford

Danks D (2003) Equilibria of the Rescorla–Wagner model. J Math Psychol 47:109–121

DeRose K, Grandy RE (1999) Conditional assertions and "Biscuit" conditionals. Noûs 33:405–420

Douven I (2008) The evidential support theory of conditionals. Synthese 164:19–44

Douven I, Elqayam S, Singmann H, van Wijnbergen-Huytink J (2018) Conditionals and inferential connections: a hypothetical inferential theory. Cogn Psychol 101:50–81

de Finetti B (1936/1995) La logique de la probabilité. Actes du congrès international de philosophie scientifique. Sorbonne, 1935. IV: induction et probabilité, 31–39. Paris: Hermann. English translation (1995): 'The logic of probability'. Philosophical Studies 77:181–190

Forsyth J, Eifert G (1998) Response intensity in content-specific fear conditioning comparing 20% versus 13% CO_2-enriched air as unconditioned stimuli. J Abnorm Psychol 107:291–304

Franke M (2007) The pragmatics of biscuit conditionals. In: Aloni M, Dekker P, Roelofsen F (eds) Proceedings of the 16th Amsterdam colloquium. ILLC, Universiteit van Amsterdam, pp 91–96

Geis M, Zwicky A (1971) On invited inferences. Linguist Inq 2:561–566

Geis M, Lycan W (1993) Nonconditional conditionals. Philos Top 21:35–56

Gluck MA, Bower GH (1988) From conditioning to category learning: an adaptive network model. J Exp Psychol Gen 117:227–247

Hirschleifer J (1991) The paradox of power. Econ Polit 3:177–200

Iatridou S (1991) Topics in conditionals. Ph.D. Thesis, MIT

Johnson-Laird P, Legrenzi P, Legrenzi M (1972) Reasoning and a sense of reality. Br J Psychol 63:395–400

Kahneman D, Tversky A (1972) Subjective probability: a judgment of representativeness. Cognit Psychol 3:430–454

Kahneman D, Wakker P, Sarin R (1997) Back to bentham? explorations of experienced utility. Q J Econ 112:375–405

Kratzer A (1991) Conditionals. In: von Stechow A, Wunderlich D (eds) Semantik: ein internationales Handbuch der zeitgenössischen Forschung/Semantics: an international handbook of contemporary research. De Gruyter, Berlin

Krzyżanowska K, Wenmackers S, Douven I (2013) Inferential conditionals and evidentiality. J Logic Lang Inform 22:315–334

Lauer S (2015) Biscuits and provisos: conveying unconditional information by conditional means. In: Csipak E, Zeijlstra H (eds) Proceedings of Sinn und Bedeutung, vol 19. Göttingen

Leslie SJ (2008) Generics: cognition and acquisition. Philos Rev 117:1–47

Lewis DK (1975) Adverbs of quantification. In: Keenan E (ed) Formal semantics of natural language. Cambridge University Press, Cambridge, pp 3–15

Newel J, Lagnado D, Shanks D (2007) Straight choices: the psychology of decision making. Psychology Press, Hove and New York

Pearl J (2000) Causality: models, reasoning and inference. Cambridge University Press, Cambridge

Rescorla R (1968) Probability of shock in the presence and absence of CS in fear conditioning. J Comp Physiol Psychol 66:15

Rescorla R, Wagner A (1972) A theory of Pavlovian conditioning: the effectiveness of reinforcement and non-reinforcement. In: Black A, Prokasy W (eds) Classical conditioning II: current research and theory. Appleton-Century-Crofts, New York, pp 64–69

Restall G (1996) Information flow and relevant logics. In: Seligman J, Westerstahl D (eds) Logic, language and computation (Volume 1). CSLI Publications, Stanford, pp 463–478

Saebo KJ (2001) Necessary conditions in a natural language. In: Féry C, Sternefeld W (eds) Audiatur vox sapientiae: a festschrift for Arnim von Stechow. Akademie Verlag, Berlin, pp 427–449

Schelling (1960) The strategy of conflict. Harvard University, Cambridge

Shep MC (1958) Shall we count the living or the dead? N Engl J Med 259:1210–1214

Skovgaard-Olsen N (2015) Ranking theory and conditional reasoning. Cogn Sci 40:848–880

Skovgaard-Olsen N, Signmann H, Klauer KC (2016) The relevance effect and conditionals. Cognition 150:26–36

Skovgaard-Olsen N (2016) Motivating the relevance approach to conditionals. Mind Lang 31:555–579. https://doi.org/10.1111/mila.12120

Skovgaard-Olsen N, Singmann H, Klauer K (2017) Relevance and reason relations. Cogn Sci 41:1202–1215

Slovic P, Finucane M, Peters E, MacGregor DG (2004) Risk as analysis and risk as feelings: some thoughts about affect, reason, risk, and rationality. Risk Anal 24:1–12

Stalnaker RC (1968) A theory of conditionals. In: Studies in logical theory, american philosophical quarterly monograph series, No. 2. Blackwell, Oxford

Stalnaker R (1970) Probability and conditionals. Philos Sci 37:64–80

Sutton R, Barto A (2016) Reinforcement learning: an introduction. MIT Press, Cambridge

Tversky A, Kahneman D (1974) Judgment under uncertainty: heuristics and biases. Science 185:1124–1131

Urquhart A (1972) Semantics for relevant logics. J Symb Log 37:159–169

van Rooij R (2017) Generics and typicality. In: Proceedings of Sinn und Bedeutung, vol 22. Berlin

van Rooij R, Schulz K (2019a) Conditionals, causality and conditional probability. J Logic Lang Inform 28:55–71. https://doi.org/10.1007/s10849-018-9275-5

van Rooij R, Schulz K (2019b) Generics and typicality: a bounded rationality approach. Linguist Philos. https://doi.org/10.1007/s10988-019-09265-8

Veltman F (1985) Logics for conditionals. Ph.D. Dissertation, University of Amsterdam

Yuille A (2006) Augmented Rescorla–Wagner and maximum likelihood estimation. In: Advances in neural information processing systems, pp 1561–1568