Why those biscuits are relevant and on the sideboard

van Rooij, R.; Schulz, K.

DOI
10.1111/theo.12309

Publication date
2021

Document Version
Final published version

Published in
Theoria

License
CC BY-NC-ND

THEORIA

THEORIA, 2021, **87**, 704–712
doi:10.1111/theo.12309

Why Those Biscuits Are Relevant and on the Sideboard

by

ROBERT VAN ROOIJ

and

KATRIN SCHULZ

Institute for Logic, Language and Computation (ILLC), University of Amsterdam

Abstract: In this paper, we explain why the antecedent of a biscuit conditional is relevant to its consequent by extending Douven's evidential support theory of conditionals making use of utilities. By this extension, we can also explain why a biscuit conditional gives rise to the inference that the consequence is (most likely) true. Finally, we account for the intuition that (indicative) biscuit sentences are false when the antecedent is false and allow for counterfactual biscuits.

Keywords: biscuit conditionals, probability, relevance

1. Introduction

Consider Austin (1961) biscuit conditional:

(1) There are biscuits on the sideboard, if you want some.

Three things are remarkable about this type of conditional: (i) the truth of the consequent is *independent* of the truth of the antecedent; (ii) notwithstanding this independence, the antecedent still seems *relevant* to the consequent; and (iii) we conclude from the appropriate use of (1) that the *consequent is true*, that is, that there are biscuits on the sideboard.

Assumption (i) is uncontroversial. A number of people have tried to account for (ii). Intuition (iii) has been explained either truth-conditionally or pragmatically by making use of intuition (i), typically without taking care of intuition (ii). In this paper, we want to explain intuitions (ii) *and* (iii), making use of intuition (i).

Why do we conclude from (1) that there are biscuits on the sideboard, that is, how can one account for intuition (iii)? This does not follow immediately from any standard semantic analysis of conditionals. According to one type of response (e.g., Predelli (2009), Ebert et al. (2014)), this shows that conditionals like (1) should be treated quite differently from standard conditionals. Others (e.g., Franke (2007)) feel that such a move is ad hoc and seek to account for the inference in pragmatics using a qualitative notion of *independence*. Of course, this works as well for a quantitative notion of independence. Let us assume with Adams (1965) that we accept an indicative conditional of the form "If A, then C" only if $P(C| A)$ is high. We want to explain why what is communicated with a biscuit conditional is that $P(C)$ is high. But that is straightforward: making use of intuition (i) that the truth of the consequent is *independent* of the truth of the antecedent, that is, that $P(C| A) = P(C)$, it follows immediately that if $P(C| A)$ is high, also $P(C)$ is high.

Such an explanation of intuition (iii) is certainly appealing. Such an analysis by itself, however, still leaves open what it is that makes the antecedent *relevant* to the consequent (cf. Lauer, 2015), for the analysis assumes that the antecedent is *irrelevant* to the consequent, at least on the standard notion of relevance according to which A is relevant to C if $P(C| A) > P(C)$. Thus, how to account for intuition (ii)?

Suppose that there are biscuits on the sideboard. What is the point of the antecedent? Iatridou (1991) and Predelli (2009) state the almost obvious: 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, 1936/1995; Belnap, 1970), the conditional "If A, C" states that C is true, if A holds, and does not say anything otherwise. Notice that on this account, also intuition (iii) is immediately explained. 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. But how then should we account for intuition (ii), that the antecedent is *relevant* for the consequent?

In the following, we propose a utility-based analysis of why the antecedent of (1) is relevant to the consequent. Making use of independence, we will then also explain why we infer from (1) that there (probably) are biscuits on the sideboard. The utility-based analysis of biscuit-conditionals will be a generalization of a simpler analysis of (indicative) conditionals that tries to explain why conditionals of the form "If A, then C" are inappropriate if there exists no link between A and C. Thus, we argue that there is not so much special about biscuit conditionals: the antecedent should always be relevant to the consequent, but the appropriate notion of relevance might depend on the circumstances.

2. Why the Antecedent Is Relevant

2.1 Truth-related relevance

What is communicated by an (indicative) conditional of the form "If A, then C"? Intuitively, it communicates that there exists a connection, link, or relevance relation between the antecedent A and the consequence C. Unfortunately, most standard analyses of indicative conditionals do not demand such a connection: the material and strict implication accounts, Stalnaker (1968) similarity account, and Adams (1965) probabilistic account all predict that "If A, then C" follows from "A and C," meaning that the truth of "A and C" suffices to make "If A, then C" true as well. Intuitively, however, (2),

(2) If Trump was the president of the United States, Ajax Amsterdam won the Champions League several times,

is just inappropriate, even though Trump was the president of the United States and Ajax Amsterdam won the Champions League several times. The inappropriateness is easily explained: the truth of the antecedent is *irrelevant* to the truth of the consequent, or almost equivalently the truth of the consequent is *independent* of the truth of the antecedent.

This suggests that an indicative conditional is appropriate only if the truth of the consequent is (positively) *dependent* on the truth of the antecedent. Making use of the standard probabilistic notion of (in)dependence, we can demand that "If A, then C" is appropriate only if $P(C|\ A) > P(C)$; that is, when A is taken to be positively relevant to C according to the standard notion of probabilistic relevance. Indeed, Douven (2008) demands this to explain the inappropriateness of (2). On top of that, however, he demands something extra because the above demand is *too weak*:

My colleague Henry's quitting his job is evidence that I shall teach next year's introductory course in social philosophy because conditional on the former the latter is a bit more probable than it is unconditionally. But even the conditional probability is still exceedingly low, given that I simply lack the requisite background for teaching such a course. It is in effect much more likely that Tom, my other colleague, who has a specialization in social philosophy, will teach the course if Henry quits his job. Thus,

(3) If Henry quits his job, I shall teach next year's introductory course in social philosophy.

is not acceptable to me, notwithstanding that its antecedent is evidence for its consequent. (Douven, 2016, pp. 107–108.)

Douven (2008, 2016) demands that for a conditional "If A, then C" to be acceptable, it should not only be the case that $P(C|\ A) > P(C)$ but also that $P(C|\ A)$ is high. Skovgaard-Olsen et al. (2016) gives additional experimental evidence for Douven's two requirements.

It turns out that there is an interesting way to "implement" this. To see this, note first (the well-known fact) that $P(C|\ A) > P(C)$ if $P(C|\ A) > P(C|\ \neg A)$. Now we demand that the conditional is acceptable if $P(C|\ A) - P(C|\ \neg A)$ is close to

$1 - P(C|\neg A)$, meaning that the measure $\frac{P(C|A) - P(C|\neg A)}{1 - P(C|\neg A)}$ must be high. But this can be the case only if both $P(C|A) - P(C|\neg A)$ and $P(C|A)$ are high and thereby deriving Douven's demands. Thus, we come to the following acceptability condition for indicative conditionals denoted by $A \Rightarrow C$ (where we abbreviate $P(C|A) - P(C|\neg A)$ by ΔP_A^C):

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

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 that for $A \Rightarrow C$ to be acceptable, it should be the case that $P(C|A)$ is high. In particular, $\frac{\Delta P_A^C}{1 - P(C|\neg A)}$ has its *maximal value* 1 only if $P(C|A) = 1$.[1] For these and other reasons, van Rooij and Schulz (2019) have argued that (CON') is the appropriate measure to account for the acceptability of many indicative conditionals.

Although (CON') accounts in a natural way for why (2) is inappropriate, it makes completely wrong predictions for our biscuit conditional (1). By using (CON'), we would falsely predict that (1) is *in*appropriate because the truth of the consequent of (1) is *independent* of the truth of the antecedent. But that means that $P(C|A) = P(C) = P(C|\neg A)$, with the result that both ΔP_A^C and $\frac{\Delta P_A^C}{1 - P(C|\neg A)}$ have the value 0.

2.2 Utility-related relevance

Rather than concluding that the proposed appropriateness conditions of Douven and Van Rooij and Schulz are all wrong, we would like to claim that their analyses are correct for most cases but not general enough to also account for biscuit conditionals. For the general case, we should not look only at informational value: utility counts as well. Here we will make use of Jeffrey (1965) decision theory, where utility functions, V, are functions from worlds to real numbers. In this framework, we can determine the expected utility of any proposition C as below.

$$EU(C) = \sum_{w \in C} P(w) \times V(w).$$

1 More specifically, $\frac{\Delta P_A^C}{1 - P(C|\neg A)} = 1$ if $P(C|A) = 1$ and $P(C|\neg A) \neq 1$.

If we assume that all worlds have the same utility, $EU(C)$ is obviously proportional to $P(C) = \sum_{w \in C} P(w)$. If we assume, moreover, that for each world w, $V(w) = 1$, then $EU(C) = P(C)$. We will mostly be interested in the conditional expected utility of C given A. This is defined as follows:

$$EU(C|A) \;=\; \sum_{w \in C} P(w|A) \times V(w).$$

Just like $EU(C) = P(C)$, if for each world w, $V(w) = 1$, now it holds that under this condition $EU(C|A) = P(C|A)$. Thus, looking at expected utilities is a generalization of "just" looking at probabilities.

For our purpose, however, we are after a generalization of two other notions: ΔP_A^C and $\frac{\Delta P_A^C}{1-P(C|\neg A)}$. A utility-based generalization of ΔP_A^C, denoted by ∇P_A^C, is straightforward:

$$\begin{aligned} \nabla P_A^C &= \sum_{w \in C} P(w|A) \times V(w) - \sum_{w \in C} P(w|\neg A) \times V(w) \\ &= EU(C|A) - EU(C|\neg A) \end{aligned}$$

For a utility-based generalization of $\frac{\Delta P_A^C}{1-P(C|\neg A)}$, we have to decide what stands in the place of "1" in the denominator. For the probabilistic case, "1" stood for the *maximal* probability value C could take. In our case, we have to do something similar: we should now look at the *maximal expected utility value* C could take. But this, obviously, comes down to the maximal utility value a C-world could get. Therefore, we propose the following generalization of (CON') as our general condition:

$$(CON) \;\; A \Rightarrow C \text{ is acceptable} \quad \text{if} \quad \frac{\nabla P_A^C}{max_{w \in C} V(w) - EV(C|\neg A)} \text{ is high.}$$

Notice that if *Value* is irrelevant (meaning that $\forall w : V(w) = 1$), 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.

We wanted to explain why (1) is relevant only if you want some biscuits. But this is now straightforward: you value those worlds where there are biscuits on the sideboard in which you are hungry higher than those in which there are biscuits but in which you are not hungry. If for the latter type of worlds w, $V(w) = 0$, it obviously follows immediately that ∇P_A^C is high and thus that the conditional is acceptable.

If, however, you are known to have no special desire for biscuits, $\nabla P^C_A = 0$, because the consequent C is true, and the conditional becomes non-assertable.[2]

3. Why the Consequent Is True

We have seen in section 1 that on the assumptions that (i) indicative conditionals "go with" conditional probabilities and that (ii) such a conditional is assertable/acceptable with a high such probability, we immediately explain by probabilistic independence why hearers conclude that the consequent is highly probable. Our analysis of indicative conditionals, however, is not based on the assumption that acceptability/assertability of "If A, then C" goes by $P(C|\ A)$. Instead, we say that for biscuit conditionals the assertability/acceptability goes by $\frac{\nabla P^C_A}{max_{w \in C} V(w) - EV(C|\neg A)}$. Can we still explain why a biscuit conditional gives rise to the inference that the consequent is (most likely) true?

We can. Here is how. Let us assume that for the biscuit conditional "If you are hungry, there are biscuits on the sideboard," you do not care at all whether there are biscuits on the sideboard, if you are not hungry. More generally, for any biscuit conditional of the form "If A, then C," we assume that for all $\neg A$-worlds u, $V(u) = 0$. On this assumption, $EU(C|\ \neg A) = 0$, and thus our measure $\frac{EU(C|A) - EU(C|\neg A)}{max_{w \in C} V(w|A) - EV(C|\neg A)}$ comes down to $\frac{EU(C|A)}{max_{w \in C} V(w)}$. We assume again that the consequent is *independent* of the antecedent, which means that $\sum_{w \in C} P(w|\ A) = P(C|\ A) = P(C) = \sum_{w \in C} P(w)$. Let us now also assume for the topic under discussion the utility function is such that $V(w) = V(v) = \kappa$ for all $A \wedge C$-worlds. By these two assumptions, the measure $\frac{\sum_{w \in C} P(w|A) \times V(w)}{max_{w \in C} V(w)}$ reduces to $\frac{\sum_{w \in C} P(w) \times \kappa}{\kappa}$. This measure has the maximal value 1 just in case $\sum_{w \in C} P(w) = 1$, that is, just in case $P(C) = 1$. This explains why we conclude from a biscuit conditional to the truth of the consequent.

4. Counterfactual Biscuits

What if you do not care about biscuits? We have seen above that Belnap (1970) conditional assertion approach predicts that in that case (1) is nonassertable. But that seems wrong: in case there are no biscuits even if you do not want any,

2 One reviewer wondered whether our analysis is general enough, that is, whether it could also be used to account for the appropriateness of conditionals like "If I may be blunt, what you said was deeply hurtful." We think it can, because it is more valuable to hear that what you said was deeply hurtful, it you do not mind the speaker to be blunt, than if you do mind, and prefer the speaker to be not so blunt. We are not sure how to account for examples like "If you remember, I explained this earlier," also given by this reviewer. For many examples of biscuit-like conditionals, see the appendix of Lycan (2001).

(1) seems false. Notice that our analysis would predict that in case you do not want biscuits, $P(w|A)$ is undefined, and thus the measure $\frac{\nabla P_A^C}{max_{w\in C} V(w) - EV(C|\neg A)}$ as well. This suggests that we should generalise our analysis such that our used measure is also defined in case the antecedent is, in fact, false. The standard way to do that is to make use of a more general counterfactual analysis: do not look at standard conditional probabilities like $P(C|A)$ but rather at the probability that C would have if A would become, or made, true. Let us denote this probability by $P_A(C)$. There are various ways to define $P_A(C)$: one could define it making use of Lewis (1976) *imaging*, which relies on a similarity relation between worlds; or one assumes that $P_A(C)$ is accounted for by means of Pearl (2000) *intervention* and assumes that $P_A(C) = P(C|do(A))$. It does not really matter much how to account for $P_A(C)$ as long as it holds that for propositions C that are not (causally) influenced by A, $P_A(C) = P(C)$, which immediately follows from Pearl's analysis.[3] Notice that if we assume that A does not (causally) influence C, the analysis of the previous section goes through as before: if we use $EU_A(C)$ for the expected utility of C making use of $P_A(C)$ instead of $P(C|A)$, the measure $\frac{EU_A(C) - EU_{\neg A}(C)}{max_{w\in C} V(w) - EV_{\neg A}(C)}$ has the maximal value just in case $P(C) = 1$.

Observe that our new analysis predicts that there are acceptable *counterfactual* biscuit conditionals. And that is at it should be (cf. Swanson, 2013). There is nothing wrong with (4).

(4) There were biscuits on the sideboard, if you had wanted some.

And also from this conditional, you conclude that the consequent is true: there were some biscuits on the sideboard.

5. Conclusion

In this paper, we have explained why the antecedent of an indicative biscuit conditional is relevant to its consequent by extending Douven's evidential support theory of conditionals making use of utilities. We have shown how we can explain why an indicative biscuit conditional gives rise to the inference that the consequence is (most likely) true. Finally, we have accounted for the intuition that biscuit sentences are not simply undefined in case the antecedent is false and have allowed for counterfactual biscuits.

3 There are many circumstances as well where $P_A(C)$ comes down to $P(C|A)$, if $P(A) > 0$.

Acknowledgements

The idea behind this paper was presented at several workshops and conferences. We thank the audiences for their comments and suggestions. We also would like to thank the reviewers of this paper for their comments and challenging questions. The work for this paper was supported by NWO grant “From Learning to Meaning: A new approach to Generic Sentences and Implicit Biases” (dossiernumber 406.18.TW.007) and by NWO grant “The biased reality of online media - Using stereotypes to make media manipulation visible” (dossiernumber 406.D1.19.059).

References

Adams, E. W. (1965) “A Logic of Conditionals.” *Inquiry* 8: 166–197.

Austin, J. L. (1961) “Ifs and Cans.” In J. O. Urmson and G. J. Warnock (eds), *Philosophical Papers*, pp. 153–180. Oxford: Oxford University Press.

Belnap, N. (1970) “Conditional Assertion and Restricted Quantification.” *Noûs* 1: 1–12.

Cheng, P. (1997) “From Covariation to Causation: A Causal Power Theory.” *Psychological Review* 104: 367–405.

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.

DeRose, K. and Grandy, R. E. (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. (2016) *The Epistemology of Indicative Conditionals*. Cambridge: Cambridge University Press.

Ebert, E., Ebert, C., and Hinterwimmer, S. (2014) “A Unified Analysis of Conditionals as Topics.” *Linguistics and Philosophy* 37: 353–408.

Franke, M. (2007), “The Pragmatics of Biscuit Conditionals,” In M. Aloni, P. Dekker, and F. Roelofsen (eds.), *Proceedings of the 16th Amsterdam Colloquium*, pp. 91–96. ILLC: Universiteit van Amsterdam.

Iatridou, S. (1991), *Topics in Conditionals*, PhD thesis, MIT.

Jeffrey, R. (1965) *The Logic of Decision*. Chicago: University of Chicago Press.

Lauer, S. (2015), “Biscuits and Provisos: Conveying Unconditional Information by Conditional Means.” In E. Csipak and H. Zeijlstra (eds), *Proceedings of Sinn und Bedeutung, 19, Göttingen*.

Lewis, D. (1976) “Probabilities of Conditionals and Conditional Probabilities.” *Philosophical Review* 85: 297–315.

Lycan, W. (2001) *Real Conditionals*. Oxford: Clarendon.

Pearl, J. (2000) *Causality: Models, Reasoning and Inference*. Cambridge: Cambridge University Press.

Predelli, S. (2009) “Towards a Semantics for Biscuit Conditionals.” *Philosophical Topics* 142: 293–305.

Shep, M. C. (1958) “Shall we Count the Living or the Dead?” *New England Journal of Medicine* 259: 1210–1214.

Skovgaard-Olsen, N., Singmann, H., and Klauer, K. C. (2016) "The relevance effect and conditionals." *Cognition* 150: 26–36.

Stalnaker, R. C. (1968) "A Theory of Conditionals." In N. Rescher (ed.), *Studies in Logical Theory, American Philosophical Quarterly Monograph Series, No. 2*. Oxford: Blackwell.

Swanson, E. (2013) "Subjunctive Biscuit and Stand-off Conditionals." *Philosophical Studies* 163: 637–648.

van Rooij, R. and Schulz, K. (2019) "Conditionals, Causality and Conditional Probability." *Journal of Logic, Language and Information* 28: 55–71.