The Bases-Theory of Conditionals

UWE MEIXNER

1.

The theory of conditionals I propose, summarily stated, is the following. The meaning of every uttered conditional comprises a propositional basis which grounds the conditioning-relation expressed by the conditional in the utterance. If the conditional is uttered in different contexts, the propositional basis for a syntactically identical conditional may be different, and hence also the (propositional) meaning of the uttered conditional (and consequently its truth-value). For conditionals that differ syntactically, the propositional basis may even be different in the same context of utterance (even if the utterance is by the same speaker). Given the propositional basis for a conditional uttered in a certain context, its meaning (as uttered) can be completely analyzed by using, besides its antecedent and its succedent, merely the operator of analytical necessity (which is the object-language representation of its concept of analytical truth), material implication and a sentence expressing the propositional basis concerned. This analysis can be stated in the modal object-language itself, without any need of going beyond it (to possible worlds and a relation of similarity between them—entities which are invoked in the standard semantics of Stalnaker and Lewis). For this reason the proposed theory of conditionals is ontologically minimal. (Note that in it there is no object-language quantification over propositions either.) And it is very simple: according to it, the logic of conditionals simply is a definitional extension of modal S5-logic. This result may seem to be highly implausible in view of the various peculiarities that distinguish conditionals both from strict and material implications. But the peculiarities can be completely accounted for by shifts in the propositional bases of utterances of conditionals. The demonstration of this will fill by far the largest part of the paper; naturally, the explanation of the logically erratic behavior of conditionals makes or breaks any theory of conditionals.

The proposed theory, call it "bases-theory", has a certain similarity to the so-called metalinguistic theory of conditionals which originally is due to Nelson Goodman; Frank Jackson presents the central tenet of that theory in his introduction to [1991] on p. 5:

> "if A, then B" is true iff *there is a statement S, meeting condition β, such that* "A&S" *logically entails* B.

But the dissimilarities of bases-theory and the metalinguistic theory of conditionals should already be clear, too. First of all, bases-theory is not metalinguistic in a non-trivial sense (it can of course be brought into a metalinguistic form; see below): in it, for stating the truth-conditions of conditionals, there is no need to quantify over object-language statements. Secondly, the metalinguistic theory of conditionals does not adequately represent the context-dependence of conditionals; for bases-theory the very same conditional (syntactically speaking) can be true on one occasion of utterance, and false on another, while the truth-values of A and B, and even the propositions expressed by them, stay the same. (Nor is this context-dependence adequately represented by the object-language counterpart of the metalinguistic theory: with quantifiers binding propositional variables; the meaning that bases-theory ascribes to conditionals cannot be adequately analyzed in this way.)

Bas van Fraassen says in [1985] on p. 29: "a counterfactual can be true, but only because some factual statement is true. Thus, consider 'If I were to open my drawer, I should see a bottle of ink.' This is true because there *is* a bottle of ink in the drawer (and because I have adequate eyesight, and so on). It would be quite difficult to give a general account of the factual conditions that make counterfactuals true or false." Indeed, it would be quite difficult, because these factual conditions are very heterogeneous. Fortunately, there is no need, in presenting a theory of conditionals, to give a general account of them; we can simply take them as being expressed by sentences out of a certain class of statements, say, S1, S2, S3, ...; all these statements express conditions that could be presupposed in utterances of conditionals. Then the adequate metalinguistic statement of the basis-theory of conditionals (*without* quantification over object-language statements) is:

> "if A, n-then B" is true iff Sn *is true and* "A&Sn" *logically* [or better: *analytically*] *entails* B, *or* A *alone logically entails* B.

This is to be understood as giving the truth-condition of an utterance of "if A, then B" on an occasion in which Sn is taken to express the presupposed condition. ("logically entails" is to be broadly construed: in the sense of "analytically entails." The tag "or A alone logically entails B" needs to be added; for sometimes, *pace* van Fraassen, the presupposed condition is of no importance for the truth of the uttered conditional, even if it is a counterfactual one.)

2.

The formal apparatus necessary for a precise formulation of the proposed theory of conditionals is quickly stated. Let L be a language of propositional logic (p, q, r, p', q', r', ... are its propositional variables) whose basic operators are negation ("~") and material implication ("→"). L is being enriched by the operator of analytical necessity ("N") and by infinitely many *basis-variables* (**b1**, **b2**, **b3**, ...) which function syntactically just like propositional variables. Basis-variables are to be thought of as expressing possible bases for uttered conditionals.

The logic of L is stated by the following axiom- and rule-schemata.

A1 $A \rightarrow (B \rightarrow A)$.
A2 $(A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$.
A3 $({\sim}A \rightarrow {\sim}B) \rightarrow (B \rightarrow A)$.
A4 $NA \rightarrow A$.
A5 $N(A \rightarrow B) \rightarrow (NA \rightarrow NB)$.
A6 ${\sim}NA \rightarrow N{\sim}NA$.
R1 $A, A \rightarrow B \vdash B$.
R2 $A \vdash NA$.

"&" ["and"], "v" ["or"], "P" ["possibly"] are defined in the usual way by "~", "→" and "N." Binding strength decreases from left to right in the sequence: ~, &, v, →.

3.

To this familiar logical system three definition-schemata are added.

D1 (if A, **n**-then B) := $\mathbf{bn} \& N(\mathbf{bn} \& A \rightarrow B) \vee N(A \rightarrow B)$.
D2 (if A had been, **n**-then B) := ~A&(if A, **n**-then B).
D3 (if A had been, **n**-then B would have been) := ~B&(if A, **n**-then B).

D1 is appropriate for *indicative conditionals*. Every conditional that is neutral as to the truth-value of both its antecedent and its succedent is here considered to be an indicative conditional. (Sometimes indicative conditionals in this sense are also formulated in the mode "if A were, B would be".) D2 is appropriate for *semi-counterfactual conditionals*: the definiens of D2 logically implies only the falsity of the antecedent, not also the falsity ~~of the succedent.~~

D3 is fitting for *fully counterfactual conditionals*: the definiens of D3 logically implies both the falsity of the antecedent and the falsity of the succedent. (Sometimes fully counterfactual conditionals are also formulated in the mode "if A were, B would be".)

As to the justification of D1: given an utterance of "if A, then B", the conditional uttered can always be disambiguated according to the basis used for the utterance. If, for example, this basis is the proposition expressed by "**b1**", the disambiguation of "if A, then B" is "if A, **1**-then B." But the latter, I claim, precisely amounts to "**b1**&N(**b1**&A→B)∨N (A→B)"; this is what is meant, when it is said that the conditioning relation between the proposition expressed by A and the proposition expressed by B holds *in virtue of* the proposition expressed by "**b1**" (if there is any need at all of an external support for that relation; the eventuality that there is no need is taken care of by the clause "∨N(A→B)").

4.

The following two important theorem-schemata are easily proved:

T1 (If A, **n**-then B)→(A→B).
T2 N(A→B)→(if A, **n**-then B).

It is quite obvious that neither the converse of T1 nor the converse of T2 can be proved; in logical strength, "if,**n**-then" is *between* strict and material implication (for all **n**).

If we add to the axiom-schemata above "**bn**→N**bn**", then the converse of T2 becomes provable; and if the converse of T2 is assumed as an axiom-schema, we can prove "**bn**→N**bn**": an instance of the converse of T2 is "(if ~**bn**, **n**-then **bn**)→N(~**bn**→**bn**)",and this, given D1, amounts to "**bn**→N**bn**".

If, however, we add "**bn**&(N(**bn**→C)∨N(**bn**→~C))∨NC∨N~C" to the axiom-schemata above, then the converse of T1 becomes provable: assume A→B; we have (1) **bn**&N(**bn**→B) ∨ **bn**&N(**bn**→~B) ∨ NB ∨ N~B, and (2) **bn**&N(**bn**→A) ∨ **bn**&N(**bn**→~A) ∨ NA ∨ N~A; let the four disjuncts of (1) be designated by (a), (b),(c), (d), those of (2) by (a′), (b′), (c′), (d′); given (a) or (c): if A, **n**-then B (by D1); given (b′) or (d′): if A, **n**-then B (by D1); given (d) and (c′): this contradicts the assumption; given (d) and (a′): this contradicts the assumption; given (b) and (c′): this contradicts the assumption; given (b) and (a′): this contradicts the assumption.

Conversely, if the converse of T1 is assumed as an axiom-schema, we can prove "**bn**&(N(**bn**→C)∨N (**bn**→~C)) ∨ NC ∨ N~C":
"(~C→C) → **bn**&N(**bn**&~C→C)∨N (~C→C) is an instance of that converse (expanded according to D1), and so is
"(C→~C)→**bn**&N(**bn**&C→~C)∨N (C→~C); the rest is obvious. The principle just deduced from the converse of T1 is more concisely formulated as "N**n**C∨N**n**~C"; see D4 in section 5 below. Of course, neither "**bn**→N**bn**" nor "N**n**C∨N**n**~C" is at all plausible as a general logical principle for all **n**; this is the measure of the implausibility that "if,then" coincides logically with material or strict implication.

We can, however, also prove the theorem schemata

T3(a) (If A, **n**-then B)&(if B, **n**-then C)→(if A, **n**-then C),
T3(b) (If A had been, **n**-then B would have been)&(if B had been, **n**-then C would have been)→(if A had been, **n**-then C would have been),

and their corollaries (since "If A&D, **n**-then A" is trivially provable)

T4(a) (If A, **n**-then B)→(if A&D, **n**-then B),
T4(b) (If A had been, **n**-then B would have been)→(if A&D had been, **n**-then B would have been).

And these results seem to demonstrate the inadequacy of the proposed theory of conditionals. Apparently, there are obvious counter-examples to T3 and T4 in ordinary language:

(1) "If the match is struck, it lights" is true, but "if the match is struck and wet, it lights" is not true.

I answer: The two sentences must be analyzed with reference to a context of utterance in which the first is true and the second false—a context concerning normal matches. But then the propositional basis for the first conditional must differ from the propositional basis for the second conditional. Let the propositional basis (in the context of utterance) for the first conditional be expressed by "**b1**". Clearly, since "if the match is struck, it lights" is true, "N(**b1**→ the match is not wet)" needs to be true also, or in other words, the propositional basis for the first conditional needs to comprise the proposition *that the match is not wet*; otherwise it could not support the truth of the conditional. Since "N(**b1**→ the match is not wet)" is true, "if the match is struck and wet, the match lights" would have to be trivially true, if "**b1**" *also* expressed the propositional basis for the latter conditional; for then this conditional (as meant in the context of utterance) would have to be disambiguated by "if the match is struck and wet, *1-then* the match lights", and according to D1 this is trivially true, if "N(**b1**→ the match is not wet)" is true (given the truth of "**b1**", whose truth is implied, according to D1, in the assumption that the first conditional is true; there is no basis-free analytic connection between its antecedent and succedent). But on the contrary, "if the match is struck and wet, then it lights" is not true. Therefore, "**b1**" does not express the basis for that conditional, and that basis is different from the basis for "if the match is struck, it lights." Let the basis for "if the match is struck and wet, it lights" be expressed by "**b2**." (Most likely that basis is the proposition expressed by "**b1**" *minus* the proposition that the match is not wet.) Then the two conditionals with the propositional meaning they have in the supposed context of utterance can be formalized like this: "if p, **1**-then q", "if p&r, **2**-then q"; but "(if p, **1**-then q)&not(if p&r, **2**-then q)" is not a counter-example to T4(a).

(2) "If the match had been struck, it would have lighted" is true, but "if the match had been struck and wet, it would have lighted" is not true.

I answer: Changing the supposed counter-example from indicative to fully counterfactual conditionals does not improve it. Simply add to the above analysis for (1) the piece of information that "the match has not lighted" is true, and adjust what has to be adjusted; then the analysis also applies to (2), and (2) is seen not to be a counterexample to T4(b).

(3) "If J. Edgar Hoover were today a communist, then he would be a traitor" is true, and "if J. Edgar Hoover had been born a Russian, then he would today be a communist" is also true; but "if J. Edgar Hoover had been born a Russian, then he would be a traitor" is not true.

I answer: This example, which is due to Robert Stalnaker (in [1968]; see [1991], p. 38), is somewhat outdated; but presumably there once was a context of utterance which yielded precisely the offered distribution of truth-values on the three conditionals. Now, clearly, in that context of utterance the three conditionals did not all have the same basis. Consider the bases used in it. The basis for the first conditional obviously contains the proposition that J. Edgar Hoover is (at the time concerned) an American citizen; the basis for the second conditional obviously contains the proposition that everybody who is born a Russian cannot escape communist indoctrination which will make him a communist for life; the basis for the third conditional, however, does not contain the first proposition, or not the second (most likely not the first). Since the third basis differs from the first and second, (3) presents no counter-example to T3(b).

Another theorem-schema that is easily proved is

T5 (If A, **n**-then ~B)→(if B, **n**-then ~A).

There seem to be obvious counter-examples to T5 (representing all variants of contraposition), too:

(4) "If he has made a mistake, then it is not a big mistake" is true, but "if he has made a big mistake, then it is not a mistake" is not true. (This one is due to Frank Jackson; see [1991], p. 3.)

I answer: Again there is no counter-example. But first of all, the two sentences, if they are to constitute a counter-example to T5 at all, need to be rephrased: "if he has made a mistake, then he has not made a big mistake", "if he has made a big mistake, then he has not made a mistake." Now, the basis for the first conditional in the supposed context of utterance is different from the basis for the second (hence the correct formalization of (4) can only be, say, "(if p, **3**-then ~q)&~(if q, **4**-then ~p)", which obviously does not contradict T5). The obtaining basis for the first conditional contains the proposition that any mistake he has made (at the occasion concerned) is not a big mistake, or in other words, that he has not made a big mistake (at the occasion concerned). The basis for the second conditional, however, does not obtain, or does not contain that proposition. If it did contain it and obtained, the second conditional would turn out to be trivially true, contradicting what is assumed in (4): "if q, **4**-then ~p" is according to D1 short for "**b4**&N(**b4**&q→~p)∨N (q→~p)" ("q" represents "he has made a big mistake", "p" represents "he has made a mistake"); but this, since "**b4**" is supposed to be true, is trivially true, if "N(**b4**→~q)" is true, what is precisely the case, if the basis for the second conditional contains the proposition that he has not made a big mistake.

5.

Whatever the basis for the second conditional is (we have assumed that it is expressed by "**b4**"), the conditional "if he has made a big mistake, **n**-then he has made a mistake" -formalized: "if q, **n**-then p"—will turn out to be true not only for **4**, but for any **n**; this is so simply in virtue of the truth of "N(q→p)."

Suppose “if q, **m**-then ~p”—in contrast to “if q, **4**-then ~p”—is indeed true (in the same context of utterance) *besides* “if q, **m**-then p.” Then, what kind of basis has to be expressed by “**bm**”? In general, we obtain from “if A, **n**-then B” and “if A, **n**-then ~B”: “**bn**&N(**bn**→~A)∨N ~A”. Hence, since “~N~q” is certainly true, we obtain the result that “**bm**&N(**bm**→~q)” needs to be true (given the above supposition); that is, “**bm**” must express an obtaining basis that is incompatible with the proposition that he (the person concerned) has made a big mistake. One such basis is the basis expressed by “**b3**”.

I add the following definitions:

D4 N**n**A := **bn**&N(**bn**→A)∨N A.
D5 P**n**A := ~N**n**~A.

From the preceding paragraph we can then gather that the following theorem is provable:

T6 P**n**A → ((if A, **n**-then B) → ~(if A, **n**-then ~B)).

This corresponds to Stalnaker's axiom (a4) (in [1968]). It is moreover obvious from D1 and D4 that we can also prove

T7 (if A, **n**-then B)↔N**n**(A→B).

This, from the right to the left, corresponds to Stalnaker's axiom (a3). From the left to the right, it is not obtainable in Stalnaker's system. (According to Stalnaker [see [1991], p. 37f.], the conditional connective cannot be analyzed as a modal [necessitation] operation performed on a material conditional, the reasons being the apparent counter-examples (1) - (4); one more problem case is discussed below in section 10.) But the principle corresponding to the weaker theorem “N**n**(if A, **n**-thenB) → N**n**(A→B)” is easily proven in Stalnaker's system from (a6), which corresponds to T1 above, and (a2), which corresponds to the also provable

T8 N**n**(A→B)→(N**n**A→N**n**B).

In the proof of the principle corresponding to “N**n**(if A, **n**-then B)→N**n**(A→B)”, besides *modus ponens*, Stalnaker's *rule of necessitation* is used, which is here represented by the provable rule-schema

T9 A ⊢ N**n**A.

We can also prove the theorems which correspond to Stalnaker's definitions of necessity and possibility:

T10(a) N**n**A ↔ (if ~A, **n**-then A),
T10(b) P**n**A ↔ ~(if A, **n**-then ~A).

6.

The principle corresponding to Stalnaker's controversial axiom (a5) is not provable in our system. Nor should it be; there are very clear counter-examples to the *conditional distribution principle* "(If A, **n**-then B∨C) → (if A, **n**-then B)∨(if A, **n**-then C))", which, since it is valid for material implication, is actually one of the less known *paradoxes* of material implication. Consider a container filled with fifty white and fifty black balls (and no other balls) of the very same size and weight which are being constantly mixed. Let "**b5**" express the (hypothetical) fact just described (and some other facts: concerning colour-constancy, and so on). Then the conditional "if George draws a ball from the container, **5**-then he draws a white ball or a black one" is true; but neither the conditional "if George draws a ball from the container, **5**-then he draws a white ball" is true, nor the corresponding conditional ending with "a black ball." The basis for these conditionals, the (hypothetically true) proposition expressed by "**b5**", is only sufficient for the truth of the first conditional, but not for the truth of the second or the third. This situation does not change at all, if we consider counterfactual conditionals instead of indicative conditionals. Suppose George, in fact, never draws a ball from the container. "If George had drawn a ball from the container, **5**-then he would have drawn a white ball or a black one" is true (according to the definitions given); but neither "if George had drawn a ball from the container, **5**-then he would have drawn a white ball" is true, nor "if George had drawn a ball from the container, **5**-then he would have drawn a black ball." (This counter-example can easily be adapted to providing also a counter-example to the *principle of conditional excluded middle*, "(If A, **n**-then B)∨(if A, **n**-then ~B)", which is an obvious logical consequence of the conditional distribution principle.)

7.

Stalnaker's last axiom (a7) corresponds to

T11 (If A, **n**-then B)&(if B, **n**-then A) → ((if A, **n**-then C)→(if B, **n**-then C)).

The proof of this, given T3(a), is immediate (Since (a1) in Stalnaker's system merely captures the tautologies of propositional logic, (a5) is the only Stalnaker-principle not acceptable for the bases-theory of conditionals.) Stalnaker's (a7) is obviously intended to be the (weaker) substitute for the *transitivity principle* (here represented by T3). But if the apparent counter-examples to the transitivity principle are considered to be convincing (they are not), then it is not very difficult to produce counter-examples of the same ilk to (a7) as well:

(5) "If Oswald didn't shoot Kennedy, somebody else shot Kennedy" is true; "if somebody else shot Kennedy, Oswald didn't shoot Kennedy" is also true, and so is "if Oswald didn't shoot Kennedy, he was not [at the time of the shooting] in the place from where Kennedy was shot"; but "if somebody else shot Kennedy, Oswald was not in the place [at the time of the shooting] from where Kennedy was shot" is not true.

In fact, (5) does not contradict T11; for the four conditionals do not all have the same basis (in the supposed context of utterance). The first three conditionals can be assigned a common basis, expressed by "**b6**"; this obtaining basis contains (analytically implies) the propositions that precisely one person shot Kennedy, and that everybody in the place [at the time of the shooting] from where Kennedy was shot shot Kennedy, hence that there was at most one person in the place [at the time of the shooting] from where Kennedy was shot. (It can easily be checked that the first three conditionals turn out to be true according to D1, if "**b6**" is true). But if the fourth conditional is to be *not true*, the basis expressed by "**b6**" cannot be the basis for it; for if it were, the fourth conditional would be true, too. (Suppose somebody other than Oswald shot Kennedy; hence Oswald didn't shoot Kennedy [because precisely one person shot Kennedy]; hence Oswald [at the time of the shooting] was not in the place from where Kennedy was shot [because everybody in that place at the time of the shooting shot Kennedy].) On the contrary, the basis for the fourth conditional obviously does not preclude the possibility that there was more than one person in the place from where Kennedy was shot [at the time of the shooting].

8.

How can the proposed theory of conditionals account for the fact (or the intuition of the majority of people) that (a) "if Oswald didn't shoot Kennedy [in Dallas, on that day in 1963], somebody else did" is true, while (b) "if Oswald had not shot Kennedy [in Dallas, on that day in 1963], somebody else would have" is false? The reason for this is not that fully counterfactual conditionals are essentially different from indicative conditionals, as so many suppose. The two sentences—in the given context of utterance in which the one is true, the other false—simply have different bases. The basis for (a) may be taken to be expressed by "**b6**", although a much weaker basis would suffice for the truth of (a): a basis that consists merely in the proposition that somebody shot Kennedy. But the basis for (b) cannot also be expressed by "**b6**"; else (b) would turn out to be true (according to D1 and D3, assuming that in fact nobody other than Oswald shot Kennedy). Rather, the basis for (b) (in the given context of utterance) quite obviously comprises the proposition that for some reason Kennedy *was bound* to be shot by somebody (as a matter of personal doom, historical necessity, or whatever); since most of us do not believe this (while believing that somebody shot Kennedy!), and since also, as every reasonable person knows, Kennedy's being shot by somebody else is not analytically implied by his not being shot by Oswald, most of us judge (b) to be false (in accordance with D1 and D3).

The relationship between indicative conditionals and fully counterfactual conditionals is the one codified by D3: if the succedent of an indicative conditional is false, then the corresponding fully counterfactual conditional has the same truth-value as the indicative conditional, *if there is the same basis for both conditionals*. Let me illustrate this further. Suppose that George is looking for Jim; George wants to know from Jim how old Jim is. While looking for Jim, George learns that everybody in room 101 is older than thirty. On the basis of this fact, he truly asserts "if Jim is in room 101, he is older than thirty", which assertion can be disambiguated by "if Jim is in room 101, 7-then he is older than thirty", "**b7**" expressing the proposition that everybody in room 101 is older than thirty. But a little bit later George discovers that Jim is younger than 25. In an appropriate sense, George *can* now assert truly "if Jim had been in room 101, he would have been older than thirty", namely in the sense of "if Jim had been in room 101, 7-then he would have been older than thirty".(This shows that not only lawlike statements can sustain a counterfactual conditional.) But of course there are other senses of "if Jim had been in room 101, then he would have been older than thirty"—in other words, other possible bases for this sentence in the supposed context of utterance—according to which it is false. Let "**b8**" express the proposition that *necessarily* (as determined by the laws of nature) everybody in room 101 is older than thirty; "**b8**" is (I presume) a false sentence, and this makes "if Jim had been in room 101, 8-then he would have been older than thirty" *false* (according to D1, D3); but it also makes the corresponding indicative conditional "if Jim is in room 101, **8**-then Jim is older than thirty" *false* (according to D1). Or let "**b9**" express the proposition that at least 90% of the persons in room 101 are older than thirty. "**b9**" is a true sentence; nevertheless, "if Jim had been in room 101, **9**-then he would have been older than thirty" is *false*, since it is analytically possible that Jim is in room 101 and at least 90% of the persons in room 101 are older than thirty, but Jim isn't; and for the very same reason the corresponding indicative conditional "if Jim is in room 101, **9**-then he is older than thirty" is also *false*.

9.

Consider what Stalnaker calls the "direct argument" (in [1975]; see[1991], p. 136):

(6) Either the butler or the gardener did it. Therefore, if the butler didn't do it, the gardener did.

The obvious validity of this inference is often taken to be very good evidence (especially in view of T1) for the hypothesis that the truth-conditions for indicative conditionals are those for material implications. I will argue that this is not so.

For understanding an uttered conditional to the extent of being able to assign a truth-value to it, it is all important to determine *the basis for it* in the given context of utterance. In practice, we are very good at accurately infering that basis without one additional word spoken from the evidence presented by the context of utterance. But if we have difficulties in determining it, we can always ask, and sometimes, if we care enough, we actually do: "What are your reasons for this conditional judgment?" (Certainly, the person asserting the conditional, while being quite positive in its assertion, will not always be able to give a sufficiently precise answer to this question. This phenomenon is no point against the bases-theory of conditionals, but simply an example of semantic vagueness, which in other instances, too, is paired with positive assertion.) Indeed, sometimes, the speaker is even so kind as to state his or her reasons for asserting a conditional *before* even asserting the conditional itself. This, typically, is done, if it is feared that the truth of the conditional will be doubtful for the hearer; and, typically, it is done by stating an inference of the conditional (as conclusion) from a sentence expressing the intended basis for it (as premise). (6) is a fitting example. Let "**b10**" express the proposition that either the butler or the gardener did it. It is incomplete, and hence misleading, to formalize (6) as "$p \vee q$ [$\sim p \rightarrow q$]; therefore: if $\sim p$, then q"; its correct formalization is "p∨q; therefore: if ~p, **10**-then q." Since "**b10**" is analytically equivalent to "either the butler or the gardener did it", (6) is immediately seen to be a valid inference on the basis of D1 (its premise cannot possibly be true without its conclusion being also true). But it is not valid in virtue of being an instance of a general rule of inference "A∨B [~A→B]; therefore: if ~A, [n-]then B" which would allow to deduce indicative conditionals from material implications. The proposed inference-schema is not correct; this can even be seen, if we stick to "p" ("the butler did it") and "q" ("the gardener did it"); for we can find a possible context of utterance in which the inference "p∨p; therefore: if ~p, then q" is *not* valid. In that context of utterance simply a stronger basis for "if ~p, then q" is used than is expressed by "p∨q" (or by "**b10**")—a basis that does not obtain in that context of utterance, thereby falsifying "if ~p, then q", although "p∨q" is true in it. Suppose the speaker in asserting "if the butler didn't do it, the gardener did" grounds this—seriously believing in a literary stereotype—on the basis that the murderer is *always* (in murders involving as suspects a gardener and a butler) either the butler (concerned) or the gardener (concerned), and not merely on the fact that in the particular murder at hand the murderer is either the butler or the gardener.

10.

Material implications are sometimes used as bases for conditionals; this accounts for the illusion that indicative conditionals have the same truth conditions as material implications. Moreover, sometimes even the succedent of a conditional, or rather the proposition expressed by its succedent, is used as a basis for it (or a basis is used which analytically implies that proposition). This, by generating an apparent lack of logical connection between antecedent and succedent, creates the illusion that T7, *read from the left to the right*, is not a correct principle for conditionals. Consider:

(7) "If the earth is a cube, the moon is not a cube" is true; but "it is not impossible that both the earth and the moon are cubes" is true, too.

I answer: Of course it is not *analytically* impossible that both the moon and the earth are cubes; but this does not give us a counter-example to T7 (from left to right). The basis for the conditional in the supposed context of utterance is quite obviously the proposition that the moon is not a cube. (Imagine the context of utterance: Hobbes obstinately maintains that both the moon and the earth are cubes; Calvin, however, while being uncertain about the correct shape of the earth, is absolutely sure that the moon is not a cube. In exasperation he exclaims "If the earth is a cube, the moon is *not* a cube!") Let this proposition be expressed by "**b11**"; then we have: "if the earth is a cube, **11**-then the moon is not a cube" is true; but so is "N**11**(the earth is a cube → the moon is not a cube)", and hence also "~P**11**(the earth is a cube & the moon is a cube)." Therefore, (7) is either irrelevant for T7, or untenable.

11.

What is deviant about the conditional "if Jones lives in London, he lives in Scotland"? Nothing whatever, except that it would be rather difficult for us to say something true, if we uttered it. This tells against the truth-conditions of indicative conditionals being the truth-conditions of material implications; for if the former were the latter, our example-sentence (which is due to Frank Jackson; see [1991], p. 2) would, on the contrary, have a very good chance of turning out true on various occasions of utterance. *Material implicationists* reply that what is deviant about the sentence is not its unlikeliness of being true when uttered, but rather its very low *assertability* which, they say, is only to some degree determined by its chances of being true.

According to the theory of conditionals here proposed, a conditional is assertable *in a context of utterance*, only if the speaker believes it to be true (only if the subjective probability of the conditional is 1 or almost 1); but there may be other requirements it must meet, too (see below in section 12). Being assertable in a context of utterance has to be distinguished from assertability *simpliciter*; only the latter can be *high* or *low*. Believing a conditional to be true, usually (but not always) involves believing that the basis for it in the given context of utterance is an *obtaining* proposition, something which is the case. Thus, the [absolute] *assertability* of a conditional is measured by the number of believable bases for it that make the conditional true, if they obtain. Every possible basis for the utterance of a conditional (whether it obtains or not) will "make" a conditional true whose antecedent analytically implies its succedent; hence such conditionals are assertable in the highest degree. But the assertability of "if Jones lives in London, he lives in Scotland" is indeed low, because there are not so many believable bases for it that make the conditional true, if they obtain. In fact, given a certain construal of "lives in", they must be found among the obtaining bases that are analytically incompatible with the proposition expressed by its antecedent. Given the mentioned construal of "lives in", "if Jones lives in London, he does not live in Scotland" is assertable in the highest degree, because its antecedent analytically implies its succedent, or in other words, because "N(Jones lives in London → Jones does not live in

Scotland)" is true on that construal. Hence, according to D1, "if Jones lives in London, **n**-then he does not live in Scotland" is true for every **n**. Hence, if the conditional "if Jones lives in London, **m**-then he lives in Scotland" is to be made true by a believable basis **m** if it obtains, we must conclude, because of T6, that "~P**m**Jones lives in London" is true for it; hence, by D4 and since "~NJones lives in London" is true: **m** is an obtaining basis that is analytically incompatible with Jones' living in London.

More interesting believable bases that make the conditional true, if they obtain, can only be had, if "lives in" is construed differently, more liberally: in such a manner that it is analytically possible for one and the same person to live in different places (as everybody construes it, if somebody says "I live in Sweden and in New Zealand"). But if "lives in" is construed in that manner, then but a little strain to the imagination will produce a believable basis that makes the conditional true, if it obtains, *and* that is not incompatible with the proposition expressed by its antecedent. Consider the proposition that Jones, a bigamist, is living precisely where his two wives live, that each wife is living at one place in Britain, that none of the wives knows that he has another wife, that Jones wants to keep it that way, that, therefore, he has put a distance between his two wives that is as long as it can be, given that both wives are living on the mainland of Britain and given that the woman he married first lives in the south of England. A private eye, determined to find out where Jones lives, might well come to believe this in the course of his investigations, and on the basis of this and some well known geographical facts, he might well say to his secretary "If Jones lives in London, then he lives in Scotland"—and this is true, if the basis is in fact the case. (The initial astonishment of the secretary at this enigmatic utterance will quickly disappear, if the detective cares to explain.)

Suppose now this situation is real, and the detective finds out that Jones is *not* living in Scotland. Then, holding on to the basis for his utterance of the indicative conditional, he can assert the fully counterfactual conditional "If Jones lived [had lived] in London, then he would live [would have lived] in Scotland", and this also is true, if the basis is in fact the case. Let's assume that the basis in fact obtains. According to the Stalnaker-semantics for counterfactual conditionals, the truth of "if Jones lived in London, then he would live in Scotland" implies that the world most similar to the actual world which is a world in which Jones is living in London is a world in which he is also living in Scotland. According to the Lewis-semantics for counterfactual conditionals (restated in [1979]; see [1991], p. 56), its truth implies that some world in which Jones is living in London and in Scotland is more similar to the actual world than any world in which Jones is living in London and not in Scotland. Both similarity statements are *prima facie* highly implausible; they need explanation, and the explanation consists precisely in recounting the piece of information about the actual world that constitutes the basis for the detective's utterance. Therefore, why not leave merely picturesque—and sometimes positively misleading—talk about possible worlds and their countless similarities, why not straightforwardly consider the propositional bases for the utterances of conditionals?

The same point can be made with reference to Kit Fine's example "if Nixon had pressed the button, there would have been a nuclear holocaust." According to Lewis, the truth of this—it is agreed on all hands that the conditional is true—implies that some world in which Nixon presses the button and there is [ensues] a nuclear holocaust is more similar to the actual world than any world in which Nixon presses the button and there is no nuclear holocaust. Again, this is *prima facie* highly implausible, and it can be made plausible only, if similarity is spelled out in terms of a *propositional basis* for the utterance of the conditional that consists mainly of laws of nature and certain contingent matters of fact, but *not* of propositions about the *overall phenomenal* appearance of the earth (compare [1991], p. 58ff., especially p. 64). *That* propositional basis is all that needs to be referred to in evaluating the conditional, and, of course, it need not be interpreted as being a set of possible worlds.

But to suit everyone's taste and for reasons of better comparability with the rivalling approaches, here is the orthodox semantic statement of the truth-condition for conditionals according to bases-theory:

V_i(if A, **n**-then B)=1 iff: $i \in V(n,A,B)$ and $V(n,A,B) \cap [A] \subseteq [B]$, or $[A] \subseteq [B]$; where i is a possible world out of a non-empty set I of possible worlds, n is the index out of a non-empty set of indices C that corresponds to **n**, V(n,A,B) is a subset of I: the proposition which is the conditioning basis relative to A and B in context n, [A] (respectively [B]) is a subset of I: the proposition expressed by A [respectively B], the set of possible worlds j such $V_j(A)=1$. (If one doesn't want context-indices attached to "if, then" in the object-language, then "if,then" has to be treated like an indexical expression: the indices in C serving as semantic parameters in addition to the elements of I.)

Before moving on to the subject of probability and conditionals, consider how easy it is, in the bases-theory of conditionals, to solve the conundrum that "if Verdi and Bizet were countrymen, then Bizet would have been Italian" and "if Verdi and Bizet were countrymen, then Verdi would have been French" each appears to be true separately, although they cannot very well be true together. The first counterfactual is true on the basis of "Verdi is Italian"; the second counterfactual is true on the basis of "Bizet is French"; both of them are indeed true on the basis of "Verdi is Italian, and Bizet is French" (and any basis that is stronger); but this is a basis that *contradicts* the antecedent of the counterfactual (*assuming* that being Italian implies not being the countryman of any French). While bases that are inconsistent with the antecedent of a conditional are not forbidden by the definitions, they are certainly not normal. (One might consider adding to the first disjunct of the definiens of D1 "&P(**bn**&A)"; then there is no basis on which both counterfactuals are true.)

12.

On the present account of conditionals, what about the (subjective) probability of a conditional and its relation to the conditional probability of its succedent given its antecedent? These two probabilities, as Lewis has argued (in [1976]), cannot always be guaranteed to be the same by the truth-conditions of conditionals—on pain of triviality. Assuming the axiom-schemata of standard probability-theory having been added to the axiom-schemata in section 2 (L being modified accordingly), I add one more axiom-schema,

A7 $N(A\leftrightarrow B)\rightarrow p(A)=p(B)$.

which connects modality and probability, and which is justifiable by considerations of rationality (remember that "N" is supposed to express analytic necessity). An immediate corollary of A7 is "$NA\rightarrow p(A)=1$", since "$NA\leftrightarrow N(A\leftrightarrow(p\rightarrow p))$" is provable.We then obtain easily ("$NA\rightarrow NNA$" is a theorem-schema of S5)

T12 $p(NA)=1 \vee p(NA)=0$.

Given this, the fundamental theorem about the probability of conditionals is

T13 (if A, **n**-then B)=0 ∨ p(if A, **n**-then B)=1 ∨ p(if A, **n**-then B)=p(**bn**).

Proof: Assume p(if A, **n**-then B) is neither 0 nor 1. If $p(N(A\rightarrow B))$ were 1, p(if A, **n**-then B) would also be 1 (by D1 and probability-theory); hence—given the assumption—$p(N(A\rightarrow B))$ is not 1, hence (by T12) $p(N(A\rightarrow B))$ is 0. Hence by D1 and probability theory

(*) p(if A, **n**-then B)=p(**bn**&N(**bn**&A→B)), hence—given the assumption—p(**bn**&N(**bn**&A→B)) is not 0, hence (by probability-theory) p(N(**bn**&A→B)) is not 0, hence (by T12) p(N(**bn**&A→B)) is 1, hence (by probability-theory)

(**) p(**bn**&N(**bn**&A→B))=p(**bn**). Consequently by (*) and (**):
p(if A, **n**-then B)=p(**bn**).

"If A, **n**-then B" is an *average conditional* iff "N(A→B)" is not true, but "N(**bn**&A→B)" is. It can easily be proved:

T14 p(if A, **n**-then B)=p(**bn**), provided "if A, **n**-then B" is an average conditional.

T14 implies that an average conditional is believed precisely, if its basis is believed. (This is fitting, since an average conditional is true precisely, if its basis is true.)

Concerning the relationship of the probability of conditionals to conditional probability, we have:

T15 p(if A, **n**-then B)=1&p(A)>0 → p(B/A)=1.

Proof: Assume p(if A, **n**-then B)=1 and p(A)>0. In case N(A→B), p(A→B)=1 (by the corollary of A7); hence (by probability theory) p(B&A)=p((A→B)&A)=p(A); hence p(B\A)=1 (by the standard definition of conditional probability, p(A)>0). In case ~N(A→B), p(N(A→B))=0 (by the corollary of A7); hence p(**bn**&N(**bn**&A→B))=1 (by probability theory, because of the assumption and D1); hence p(**bn**)=1 and p(N(**bn**&A→B))=1, hence N(**bn**&A→B) [else we would have p(N(**bn**&A→B))=0], hence (by the corollary of A7) p(**bn**&A→B)=1 and p(**bn**)=1; hence
p(B&**bn**&A)=p((**bn**&A→B)&**bn**&A)=p(**bn**&A),p(B&**bn**&A)=p(B&A), p(**bn**&A)=p(A); consequently, p(B&A)=p(A), hence p(B/A)=1.

T16 p(**bn**)=1 → p(B\A)=1, provided p(A)>0 and "if A, **n**-then B" is an average conditional.

Given T14, T16 is an immediate consequence of T15.

T15 does not distinguish conditionals from material implications, because for these we can of course prove the analogous theorem
"p(A→B)=1&p(A)>0 → p(B/A)=1." But it is well known that p(A→B) may be *close* to 1, while p(B\A)=0: let p(~A) be close to 1, but not equal to 1, while p(B)=0. In this respect, however, conditionals act differently from material implications (in *normal* situations of utterance where the speaker *does not* firmly believe that the proposition expressed by the antecedent and the basis for the utterance of the conditional do not both obtain together):

T17 p(if A, **n**-then B) close to 1, but not equal to 1, & p(**bn**&A)>0 → p(B/A) almost equal to 1.

Proof: Assume that p(if A, **n**-then B) is close to 1, but not equal to 1; hence by T13 p(if A, **n**-then B)=p(**bn**), hence p(**bn**) is close to 1, but not equal to 1. Hence also p(N(A→B))=0 (*otherwise* by T12 p(N(A→B))=1, therefore p(if A, **n**-then B)=1—contradicting the assumption). Hence also p(N(**bn**&A→B))=1 (*otherwise* by T12 p(N(**bn**&A→B))=0, therefore, since p(N(A→B))=0, p(if A, **n**-then B)=0—contradicting the assumption). Since p(N(**bn**&A→B))=1, p(**bn**&A→B)=1 [for suppose p(**bn**&A→B) is not 1, then by the corollary of A7 ~N(**bn**&A→B), hence N~N(**bn**&A→B), hence by that same corollary p(~N(**bn**&A→B))=1, hence p(N(**bn**&A→B))=0—contradicting p(N(**bn**&A→B))=1]; hence (*) p(**bn**&A&B)=p(**bn**&A). p(**bn**&A)=p(A)+p(**bn**)-p(**bn**∨A); therefore, since p(**bn**) is close to 1, but not equal to 1, p(**bn**&A) is almost equal to p(A). p(**bn**&A&B)=p(A&B)+p(**bn**)-p(**bn**∨A&B); therefore, since p(**bn**) is close to 1, while not equal to 1, p(**bn**&A&B) is almost equal to p(A&B). Consequently in virtue of (*): p(A&B) almost equals p(A). Therefore, since p(A)>0 [this is a consequence of the *further* assumption p(**bn**&A)>0] and since p(A&B)>0 [this is a consequence of p(**bn**&A&B)>0, which because of p(**bn**&A&B)=p(**bn**&A) is a consequence of p(**bn**&A)>0], p(B/A) is almost equal to 1.

The above counter-example which demonstrates that the analogue of T17 for material implications does not hold—p(nonA) close to 1, but not equal to 1, p(B)=0—is incompatible with the antecedent of T17: we can deduce [see above] from that antecedent p(**bn**&A&B)=p(**bn**&A); hence, if p(B)=0, p(**bn**&A)=0—contradicting that antecedent. [Comment: Suppose that "N(A→B↔**bm**)" is true for some **m** (that is, suppose that the material implication "A→B" expresses a basis for conditionals). Then we have as an instance of T17 "p(if A, **m**-then B) is close to 1, but not equal to 1, & p(**bm**&A)>0 → p(B/A) almost equal to 1." But, on the basis of D1, this amounts to "p((A→B)∨N (A→B)) is close to 1, but not equal to 1, & p(A&B)>0 → p(B/A) almost equal to 1." Since we can prove "p(A→B) is close to 1, but not equal to 1 ↔ p((A→B)∨N (A→B)) is close to 1, but not equal to 1", we finally obtain "p(A→B) is close to 1, but not equal to 1, & p(A&B)>0 → p(B/A) almost equal to 1." For material implications expressing bases for conditionals, T17 amounts to a theorem of elementary probability theory.]

T17 and T15 completely fit the fact that the probability of a *fully counterfactual* conditional that is assertable in a certain context of utterance cannot ever be the conditional probability of the succedent given the antecedent. For a counterfactual conditional "if A had been, **n**-then B would have been" can only be [correctly] asserted, if the speaker *firmly believes* that its succedent is false—even if the probability of the counterfactual conditional for him is, albeit not equal to 1, at least close to 1 [that probability being only close to 1 involves, as can be seen from D3 and the proof of T17 above, p(if A, **n**-then B) is almost 1, p(N(A→B))=0, p(N(**bn**&A→B))=1, p(**bn**) is almost 1]. But if p(B)=0, then p(B/A) is also zero, if p(B/A) is defined at all, *while* p(if A had been, **n**-then B would have been) must be 1 or almost 1, given that "if A had been, **n**-then B would have been" is assertable in the context of utterance. However, neither T15 nor T17 are applicable in that case: for we have as theorems: if p(if A had been, **n**-then B would have been)=1, then p(A)=0; if p(if A had been, **n**-then B would have been) is almost 1 and p(B)=0, then p(**bn**&A)=0. (T17 and T15 are also inapplicable in case of assertable semi-counterfactual conditionals; for a semi-counterfactual is only assertable, if the speaker *firmly believes* that its antecedent "A" is false, that is, if p(A)=0.)

13.

Finally, how can causal statements be treated according to the basis-theory of conditionals? Let L be extended to such an extent that it includes singular terms and quantifiers for events, and a monadic predicate "H" such that "H(x)" means "[event] x happens", and a dyadic predicate "<<" such that "(x<<y)" means "[the time of event] x is before [the time of event] y." (For a theory of events and their happening, see my [1994].) Consider then the following definitions:

D6 B, **n**-because A := A&P~B&(if A, **n**-then B).
D7 y **n**-because of x := H(y), **n**-because H(x).

"y **n**-because of x" is not yet "x **n**-causes y." To get there, certain among the bases for the utterance of conditionals have to be distinguished as being *causal bases*. This, I submit, is a very difficult philosophical task (maybe without satisfactory solution); but it is not a task I need to tackle in this paper. I may simply procede on the assumption that we know what makes a basis a causal basis, and that some bases are causal bases. Moreover, x **n**-causes y *does* seem to require that x happens *before* y; however, I do not want to enter into a debate on this. The following definition, then, captures at least a very important conception of causation:

Provided that "**bn**" expresses a causal basis:

D8 x **n**-causes y := (x<<y)&(y **n**-because of x).

Note that "N(x)(y)((x<<y)→P(Hx&~Hy))" seems very much to be true, and so does "(y)P~H(y)." If they are, "x **n**-causes y" amounts to "(x<<y)&H(x)&**bn**&N(**bn**&Hx→H(y))", provided that "**bn**" expresses a causal basis.

References

Jackson, Frank: *Conditionals*, Oxford University Press, Oxford 1991.[1991]

van Fraassen, Bas: *An Introduction to the Philosophy of Time and Space*, Columbia University Press, New York 1985.[1985]

Lewis, David: "Counterfactual Dependence and Time's Arrow", *Nous* XIII (1979), pp. 455-476; reprinted in [1991], pp. 46-75.[1979]

Lewis, David: "Probabilities of Conditionals and Conditional Probabilities", *Philosophical Review* 85 (1976), pp. 297-315; reprinted in [1991], pp. 76-101.[1976]

Meixner, Uwe: "Events and Their Reality", *Logic and Logical Philosophy* 2 (1994), pp. 23-33.[1994]

Stalnaker, Robert: "A Theory of Conditionals", *Studies in Logical Theory, American Philosophical Quarterly*, Monograph, Blackwell, Oxford 1968, pp. 98-112; reprinted in [1991], pp. 28-45.[1968]

Stalnaker, Robert: "Indicative Conditionals", *Philosophia* 5 (1975), pp. 269-286; reprinted in [1991], pp. 136-154.[1975]