

DALE JACQUETTE

CONUNDRUMS OF CONDITIONALS IN CONTRAPOSITION

A previously unnoticed metalogical paradox about contraposition is formulated in the informal metalanguage of propositional logic, where it exploits a reflexive self-non-application of the truth table definition of the material conditional to achieve semantic diagonalization. Three versions of the paradox are considered. The main modal formulation takes as its assumption a conditional that articulates the truth table conditions of conditional propositions in stating that if the antecedent of a true conditional is false, then it is possible for its consequent to be true. When this true conditional is contraposed in the conclusion of the inference, it produces the false conditional conclusion that if it is not the case that the consequent of a true conditional can be true, then it is not the case that the antecedent of the conditional is false.

1. THE LOGIC OF CONDITIONALS

A conditional sentence is the *literal contrapositive* of another conditional if and only if the antecedent of one is the negation of the consequent of the other. The sentence $\neg q \supset \neg p$ is thus ordinarily understood as the literal contrapositive of $p \supset q$. But the requirement presupposes that the unnegated antecedents of the conditionals are identical in meaning to the unnegated consequents of their contrapositives. The univocity of 'p' and 'q' in $p \supset q$ and $\neg q \supset \neg p$ can usually be taken for granted within a single context of application in symbolic logic, but in ordinary language the situation is more complicated. To appreciate the difficulties, consider the use of potentially equivocal terms in the following conditionals whose reference is specified in particular speech act contexts:

(1.1) If the money is in the bank, then the money is safe.

(1.2) If it is not the case that the money is safe, then it is not the case that the money is in the bank.

Here if 'bank' in (1.1) refers to a financial institution, and in (1.2) to the edge of a body of water, then (1.1) and (1.2) will not be literal contrapositives of one another. Similar problems can occur if indexical terms are introduced.

(2.1) If I [Jones] leave today, then I [Jones] will return tomorrow.

(2.2) If it is not the case that I [Smith] will return tomorrow, then it is not the case that I [Smith] leave today.

2. INDEXICAL SELF-REFERENCE IN CONTRAPOSITION

The conditionals (2.1) and (2.2) for obvious and philosophically uninteresting reasons cannot be considered literal contrapositives. This is such a familiar phenomenon of colloquial language that the application in these illustrations is hardly worth mentioning. The requirement gains importance only when the indexicals in question self-referentially designate the conditionals in which they occur.

(3.1) If this conditional is true, then this conditional is true.

(3.2) If it is not the case that this conditional is true, then it is not the case that this conditional is true.

The indexical 'this' in (3.1) refers to a different conditional than in (3.2). By the above definition of a literal contrapositive, it follows that (3.1) and (3.2) are not literal contrapositives, even though (3.1) and (3.2), like (1.1) and (1.2) and (2.1) and (2.2), are uniform syntactical substitution instances in ordinary language of the paradigm symbolic contrapositive pair, $p \supset q$ and $\neg q \supset \neg p$. The difference is that in (3.1) and (3.2), the conditionals wear the disabling equivocation on their sleeves by speaking explicitly self-referentially of what are manifestly different conditional types and tokens involving syntactically different antecedents and syntactically different consequents. Of course, (3.1) and (3.2) are truth functionally equivalent despite not being literal contrapositives. Yet the failure of the conditional in (3.1) to imply a literal contrapositive in (3.2) suggests a limitation in the application of contraposition to ordinary language expressions involving indexicals.

Here is a dilemma about the logic of everyday linguistic usage. There are only two possibilities of resolving the impasse. We can either decide that: (i) nonliteral contraposition is sometimes though not typically deductively valid in the logic of ordinary language; or (ii) nonliteral contraposition is never deductively valid. If nonliteral contrapo-

sition is sometimes even if not typically deductively valid in ordinary language, then, contrary to the assumption of modern symbolic logic, contraposition, insofar as it is supposed to apply to everyday discourse, cannot be correctly defined as a purely syntactical transformation. If, on the other hand, nonliteral contraposition is never deductively valid, then the counterexample proves that ordinary language tolerates deductively invalid contrapositions. In that case, the principles of deductively valid inference elaborated by modern symbolic logic once again fail to apply universally to the semantics of deductively valid inference structures in ordinary language. It follows in either case, for reasons involving specific inadequacies of applied formal symbolic logic in ordinary language, that the logic of everyday thought and language is not the logic of formal symbolic logic.

We are not free to legislate indexicals out of colloquial language in the way that we may want to eliminate them from specialized formal symbolic languages like first-order propositional and predicate-quantificational logic. So we cannot hope to avoid the dilemma by analyzing away the indexicals in everyday discourse in an idealized reformulation of its syntax or grammar. We can only conclude either that ordinary language sustains deductively invalid contrapositions, or that, contrary to symbolic logic, contraposition is not a purely syntactical transformation. The same point could be made in many different ways. But the preceding discussion sets the stage for a more compelling argument by softening expectations about the validity of contraposition as a purely syntactical operation, especially in cases where a conditional conditionally describes its own truth conditions.

3. METALOGICAL PARADOX ABOUT CONDITIONALS IN CONTRAPOSITION

We now complicate the issue by considering a metalogical paradox about a special set of conditionals in literal contraposition. The paradox involves no self-referential indexicals, nor arithmetization of logical syntax, naive comprehension or abstraction principles, nor violations of type or object and metalanguage distinctions.

The paradox is formulated in the informal metalanguage of propositional logic. It achieves a metalogical contradiction by applying a metalogical conditional formulation of the definition of the material conditional to a particular object language conditional. I assume only the standard truth table interpretation of the conditional and the transformation rule for contraposition for the deductively valid interderivability of $\varphi \supset \psi$ and $\neg\psi \supset \neg\varphi$. The paradox arises by invoking the definition

of the conditional to three different metalogical conditional expressions of the truth conditions for the conditional. The strategy is to use conditionals metalogically to talk about themselves in what look to be innocuous ways that finally contradict the definition of the conditional when they are supposed to be validly transformed by contraposition.

Consider the following inferences, in which generally throughout (n.2) is the literal contrapositive of (n.1), and in which the truth evaluations of assumptions and conclusions are indicated in brackets below each step of the deduction. They are these:

A. Categorical Formulation of Paradox with Nonpreservation of Modality of Truth Conditions

(4.1) If the antecedent of true conditional $p \supset q$ is false, then the consequent of true conditional $p \supset q$ is true. [possibly true]

(4.2) If it is not the case that the consequent of true conditional $p \supset q$ is true, then it is not the case that the antecedent of true conditional $p \supset q$ is false. [necessarily false; not possibly true]

B. Modal Formulation of Paradox with Nonpreservation of Categorical Truth Conditions

(5.1) If the antecedent of contingently true conditional $p \supset q$ is false, then the consequent of contingently true conditional $p \supset q$ can be true. [true]

(5.2) If it is not the case that the consequent of contingently true conditional $p \supset q$ can be true, then it is not the case that the antecedent of contingently true conditional $p \supset q$ is false. [false]

C. Categorical Formulation of Paradox with Nonpreservation of Categorical Truth Conditions

(6.1) If the antecedent of true conditional $p \supset q$ is false, then the consequent of true conditional $p \supset q$ is true or false indifferently. [true]

(6.2) If it is not the case that the consequent of true conditional $p \supset q$ is true or false indifferently, then it is not the case that the antecedent of true conditional $p \supset q$ is false. [false]

In the first argument, the definition of the material conditional implies that it is not necessarily false that if the antecedent of a true

conditional is false, then the consequent of the true conditional is true. The assumption is not necessarily false, because the material conditional is standardly defined as true just in case either the antecedent is false or the consequent is true.

The assumption makes the truth of the consequent of a true conditional depend on the falsehood of the conditional's antecedent. The proposition is not necessarily true and not necessarily false. By the standard truth table definition of the conditional, the consequent of a contingently true conditional with a false antecedent could be either true or false. But the literal contrapositive of the assumption in the conclusion of the first argument, that if it is not the case that the consequent of the true conditional is true, then it is not the case that the antecedent of the true conditional is false, is necessarily false.

Again, the reason goes back to the truth table definition of the conditional. If the consequent of a true conditional is false, then the antecedent must be false, making it necessarily false that it is not the case that the antecedent of the true conditional is false where it is not the case that the consequent of the true conditional is true. Yet assumption (3.1) and conclusion (3.2) are literal contrapositives. Thus, we go via contraposition from a conditional that is possibly true in the assumption to a supposedly logically equivalent conditional in the conclusion that is necessarily false.

Similar reasoning with only minor rewording applies in the case of the remaining two inferences. All three of these arguments are deductively invalid by virtue of violating the requirement that in a valid argument one never goes from true assumptions to a false conclusion, nor by implication from possibly true assumptions to a necessarily false conclusion. But since the inferences in all three cases are just literal applications of contraposition, it appears that contraposition itself, despite its justification by the standard truth table semantics for propositional logic, is deductively invalid.

The first argument represents a categorical formulation of the paradox in which deductive invalidity depends on the modality of the truth conditions of the assumption and conclusion. The content of the assumption and conclusion contains no modal terms. But the modality attaches to the fact that the conditional assumption is possibly true, while the conditional conclusion is necessarily false.

The second argument represents a modal formulation of the paradox. Deductive invalidity is predicated on the categorical truth of the modal-term-laden assumption and the categorical falsehood of the modal-term-laden conclusion. The assumption that if the antecedent of a contingently true conditional is false, then the consequent of the

conclusion can be true is itself simply true. The conclusion that if it is not the case that the consequent of a contingently true conditional can be true, then it is not the case that the antecedent of the true conditional is false, is itself simply false.

The third argument represents a categorical formulation of the paradox. Here deductive invalidity is the product of the categorical truth of the assumption that if the antecedent of a true conditional is false, then the consequent of the conditional is true or false indifferently, and of the categorical falsehood of the conclusion that if it is not the case that the consequent of a true conditional is true or false indifferently, then it is not the case that the antecedent of the conditional is false. To speak of the consequent of a true conditional as being true or false indifferently is tantamount to saying modally that where the antecedent of a true conditional is false, the consequent can or could be or is possibly true or false. But it may be worthwhile to see that the paradox can be formulated without explicitly including modal terms or operators, for the sake of showing that the paradox is not straightforwardly resolvable by appealing to philosophical scruples about the intensionality of modal contexts.

4. FORMALIZATION OF THE PARADOX

It is important to remember that the paradox is supposed to hold specifically in the metalogic of conditionals in contraposition. The problem does not occur in the object language of propositional logic, which for present purposes can be accepted as paradox free.

To understand the paradox more thoroughly, we concentrate on the second, explicitly modal version, which may be thought to be the most interesting. We can reformulate the paradox in a formal metalogical notation. Let us introduce a metalogical truth valuation operator, ‘ \mathbb{V} ’, that takes a proposition into its truth value. The informal modal term ‘can’ in the original argument (5.1) \rightarrow (5.2) is interpreted more precisely as having narrow *de re* scope. The conditional ‘ \supset ’ throughout is the material conditional classically defined by standard truth tables. We begin by explicitly assuming that the conditionals in question are contingently true. Then we make a true conditional metalogical statement about the possible truth value of the consequent of a contingently true conditional, if the antecedent of the contingently true conditional is false. The conclusion of the inference applies literal contraposition to the main assumption, taking the inference from true assumptions to a false conclusion. The general form of the metalogical paradox for conditionals in literal contraposition is formalized in this way:

$$(5.0^*) \quad \mathbb{V}(p \supset q) = T \ \& \ \neg \Box \mathbb{V}(p \supset q) = T \quad [\text{true by stipulation}]$$

$$(5.1^*) \quad \mathbb{V}(p) = F \supset \Diamond \mathbb{V}(q) = T \quad [\text{true}]$$

$$(5.2^*) \quad \neg \Diamond \mathbb{V}(q) = T \supset \neg \mathbb{V}(p) = F \quad [\text{false}]$$

The assumptions in (5.0*) and (5.1*) are true; (5.0*) by stipulation, and (5.1*) by the definition of the truth conditions for the material conditional. (5.2*) is the literal contrapositive of (5.1), but is clearly false. If it is not possible for the consequent of a contingently true conditional to be true, then the consequent is necessarily false. But if so, then the antecedent of the conditional must also be false, which contradicts the consequent of (5.2*).

We can conditionalize the restriction to contingently true conditionals by making it the antecedent of a single conditional assumption and conditional conclusion, in which literal contraposition is applied only to the consequent of the assumption to derive the consequent of the conclusion. The counterexample then has this condensed form:

$$(5.1^{**}) \quad (\mathbb{V}(p \supset q) = T \ \& \ \neg \Box \mathbb{V}(p \supset q) = T) \supset$$

$$\quad (\mathbb{V}(p) = F \supset \Diamond \mathbb{V}(q) = T) \quad [\text{true}]$$

$$(5.2^{**}) \quad (\mathbb{V}(p \supset q) = T \ \& \ \neg \Box \mathbb{V}(p \supset q) = T) \supset$$

$$\quad (\neg \Diamond \mathbb{V}(q) = T \supset \neg \mathbb{V}(p) = F) \quad [\text{false}]$$

The assumption in (5.1**) is also true. If the conditional is contingently true, then, if the antecedent is false, it remains possible for the consequent to be true—that is, the consequent could be contingently false. But the conclusion in (5.2**) is false, because, again, where the conditional is contingently true, if the consequent is not possibly true but necessarily false, the antecedent is not, and cannot even contingently be, false. The literal contrapositive of (5.1**) is a true conditional sentence with a false antecedent:

$$(5.3^{**}) \quad \neg(\mathbb{V}(p) = F \supset \Diamond \mathbb{V}(q) = T) \supset$$

$$\quad \neg(\mathbb{V}(p \supset q) = T \ \& \ \neg \Box \mathbb{V}(p \supset q) = T)$$

The paradox is produced only by either separating the requirement that the conditional in question is contingently true as in (5.0*) + (5.1*) → (5.2*), and applying literal contraposition only to (5.1*) to produce (5.2*), or applying literal contraposition only to the consequent of (5.1**) to produce (5.2**).

The argument is not that (5.2), (5.2*) or (5.2**) are universal or logically necessary falsehoods. Nor is it necessary to claim that (5.1), (5.1*) or (5.1**) are universal or logically necessary truths, although

intuitively they appear to be. The paradox minimally requires only that there is at least one, whereas ultimately there are unlimitedly many, instance(s) in which (5.1), (5.1*), or (5.1**) is true and (5.2), (5.2*), or (5.2**) is false. Here is one such application, where the truth values of propositions in the antecedent and consequent of the conditional and its literal contraposition need to be understood hypothetically, in keeping with the metalogical formulation of the paradox. A concrete example may help to clarify the objection. We uniformly instantiate the propositional variables in (5.0*), (5.1*) and (5.2*), and (5.1**) and (5.2**), with sentences of unknown truth value, where we do not know who John and Mary are, nor whether or not they are actually happy. Then we have:

$$\begin{array}{ll}
 (5.0a) & \forall(\textit{John is happy} \supset \textit{Mary is happy}) = T \\
 & \& \neg \Box \forall(\textit{John is happy} \supset \textit{Mary is happy}) = T & [\text{true}] \\
 (5.1a) & \forall(\textit{John is happy}) = F \supset \Diamond \forall(\textit{Mary is happy}) = T & [\text{true}] \\
 \hline
 (5.2a) & \neg \Diamond \forall(\textit{Mary is happy}) = T \supset \neg \forall(\textit{John is happy}) = F & [\text{false}]
 \end{array}$$

The parallel conditional formulation like that for (5.1**) \rightarrow (5.2**) has the following instantiation:

$$\begin{array}{ll}
 (5.1b) & (\forall(\textit{John is happy} \supset \textit{Mary is happy}) = T \\
 & \& \neg \Box \forall(\textit{John is happy} \supset \textit{Mary is happy}) = T) \supset \\
 & (\forall(\textit{John is happy}) = F \supset \Diamond \forall(\textit{Mary is happy}) = T) & [\text{true}] \\
 \hline
 (5.2b) & (\forall(\textit{John is happy} \supset \textit{Mary is happy}) = T \\
 & \& \neg \Box \forall(\textit{John is happy} \supset \textit{Mary is happy}) = T) \supset \\
 & (\neg \Diamond \forall(\textit{Mary is happy}) = T) \supset \neg \forall(\textit{John is happy}) = F) & [\text{false}]
 \end{array}$$

The inference certainly looks paradoxical, with the blame falling on metalogical contraposition. The paradox cannot be overturned by saying that if the conditional that *John is happy* \supset *Mary is happy* is contingently true, then the antecedent of the conditional of the main consequent, $\neg \Diamond \forall(\textit{Mary is happy}) = T$, must be false, rendering the entire embedded conditional trivially true. The reason is that for all that (5.1a) or (5.1b) have to say, Mary may be such that necessarily she is not happy. Although (5.1a) or the consequent of the conditional embedded in the main consequent of (5.1b) states that it is possible that Mary is happy, the conditional as a whole can be true if even it is not possible that Mary is happy, provided that the antecedent of the embedded conditional is also false, where John is happy, or where it is not the case that John is not happy.

It might be wondered whether there is something objectionable about formulating the paradox in terms of sentences containing true and false predications, possibly in noncompliance with Tarski's restrictions against the sentences in any language being able to express their own truth value. However, the paradox in all three formulations does not involve the kind of self-denial (or self-affirmation) of truth which Tarski was concerned to prohibit in his solution to the liar paradox in formal languages. Rather, the assumption and conclusion in each version are already metalogical expressions in an informal metalanguage that refer to the truth and falsehood of the antecedent and consequent of a certain true conditional expressed in the object language of standard propositional logic.

Nor need the modality in the first and second versions of the paradox be disquieting. The modality of a proposition's being such that it can be either true or false, whether or not we can reduce it away, or drive it underground by making reference instead to a proposition's being true or false indifferently, is ineliminably part of the metalogic of the definition of the material conditional, just as it is ineliminably part of the higher-order metalogical semantic characterization of the concept of deductive validity. We must be able metalogically to say truly of a true conditional that if its antecedent is true then its consequent must be true or necessarily is true, and that if its antecedent is false then its consequent can be or is possibly either true or false, and of a false conditional that it must be or is necessarily such that its antecedent is true and its consequent is false. For it is by means of these modalities of truth conditions in the metalogic of standard propositional logic that we define the conditional. It is in terms of the truth table requirements that the metalogical paradox about conditional contraposition is proposed, to which the deductively valid transformation rule of conditional contraposition is then applied with metalogical paradoxical implications.

The paradox concerns only the material conditional expression of the truth conditions specifically of the material conditional. The paradox in any of its formulations does not obtain if the material conditional $p \supset q$ is replaced by any of its truth functional equivalents, such as $\neg p \vee q$ or $\neg(p \ \& \ \neg q)$, to which literal contraposition does not apply. But since these equivalences are determined by the truth table definition of the conditional, the fact that the paradox obtains only for the material conditional reinforces the suggestion that there is something logically strange about the truth conditions of conditionals in contraposition.

5. UNCONCLUDING METALOGICAL POSTSCRIPT

In this essay, I have considered a metalogical paradox about literal contraposition. My purpose has been to call attention to a previously unnoticed inconsistency in the metalogic of conditionals in literal contraposition, and not to recommend any solutions or even to propose a philosophical diagnosis.

My instincts are that nothing can really be wrong with contraposition. Its truth table justification is readily surveyable as entailing no parallel antinomy in the object theory of propositional logic. The problem is tantalizing precisely because, despite these considerations, it is hard to discover any mistaken step in the reasoning leading to the paradox. I nevertheless imagine that there might be something interestingly wrong with the argument. But what?

The paradox, if it stands, threatens the semantic integrity of the material conditional, inference by contraposition, and deductive validity. If there is no satisfactory resolution of the paradox, then the foundations of propositional logic in truth table definitions of the conditional and justification of contraposition are seriously compromised in what has otherwise been assumed to constitute one of the least controversial mainstays of contemporary mathematical logic.*

DEPARTMENT OF PHILOSOPHY
THE PENNSYLVANIA STATE UNIVERSITY, USA

*A version of this essay was presented at the American Philosophical Association, Eastern Division, Washington, DC, 27–30 December 1998. I am grateful to The Institute for the Arts and Humanistic Studies, The Pennsylvania State University, for a Term Fellowship during 1997–1999, which made possible completion of this and related research projects.