The Paradoxes of Material Implication, Truth Tables, and the Bare Logical Rules

Kent Slinker

1. The paradoxes

A common list of the paradoxes of material implication include the following argument forms:1

1) p ⊃ (q ⊃ p)
2) p ⊃ (q ⊃ q)
3) (p & ~p) ⊃ q
4) p ⊃ (q v ~q)
5) p ⊃ (~p ⊃ q)

I will first argue that truth table analysis shows that forms three and four are impossible logical sentences. Indeed, I will maintain that truth table analysis clearly shows that no single logical operator may join a variable and it's negation, such as (p&~p), (~p ⊃ p) and even (p v ~p). I will contend that the remaining argument forms can not properly be understood as forms that are capable of expressing full entailment where the truth or falsity of the consequent is determined by logical rules alone. The error which gives rise to the paradoxes lies not in logic but with the logician for trying to insist they must somehow express entailment. To see why this is so, a detailed analysis of the truth table is required.

2. The truth table

The foundation of a truth table rests on two important assumptions. The first is the bivalence of the variables, and the second is the independence of the variables. From this, the truth table is grounded in the listing of all possible combinations of truth values for the variables in the given argument. I will call this exhaustive combination the foundation of the truth table.

Let us examine the foundation of a common truth table, expressed in the familiar way, except where the customary t's and f's are replaced by p, ~p, q, ~q. The reason for this will become apparent in just a moment. A truth table with two variables would look like this:

<table>
<thead>
<tr>
<th>p</th>
<th>q</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>~q</td>
</tr>
<tr>
<td>~p</td>
<td>q</td>
</tr>
<tr>
<td>~p</td>
<td>~q</td>
</tr>
</tbody>
</table>

1 According to Urquhart (1972) page 166, and Mares (2002), page 611.
This matrix shows the assumptions previously listed: variable bivalence and truth functional independence. Each row represents a unique combination of values for the given truth table and taken together they express all possible combinations, hence *it can never be the case that one has the same variable and that variable’s negation in the same row* (e.g. never both p and \(-p\) ). Each column, on the other hand, contains a variable and that variable’s negation in equal numbers (e.g. p or \(-p\)). Hence logical truth \((p \lor \neg p)\) together with logical falsehood \((p \land \neg p)\) are already shown by the very construction and rules which govern the truth table.

Let us look at the simplest table, using a single variable p and \(-p\) and consider how to define the most simple logical operator:

<table>
<thead>
<tr>
<th>Logical Operator to be defined (#)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p ((p # p) = \text{ truth value}_1 )</td>
</tr>
<tr>
<td>(-p ) ((-p # \neg p) = \text{ truth value}_2 )</td>
</tr>
</tbody>
</table>

When the only variable is p, the values passed to any logical operator can not be both p and \(-p\) at the same time! This is true for any number of variables! It does not matter whether the operator is \(\lor\), \(\supset\) or \(\land\), or any other logical operator, all violate the rules which make the truth table meaningful! *This means that \((p \land \neg p)\), \((p \lor \neg p)\) etc. are all instances of impossible logical sentences* (henceforth wffs), since they take on values not given by the foundation of the truth table - which lists all possible values! This should be welcome news for the logician who is worried about \((p \land \neg p)\) entailing anything. It consigns that combination of variables to an impossible status. Contradictions as wffs can still occur, but under this view they will always be the result of *more than one* logical operation or the negation of a tautology, and never the result of trying to pass a set of contradictory values from the truth table for the same variable to a single logical operation.\(^2\)

One may counter at this point that \((p \land \neg p)\) and \((p \lor \neg p)\) are just expressions of the mutual exclusion and logical necessity already contained in the construction of a truth table - therefore there is no harm in allowing such assumptions to reappear as wffs. My response is good enough, just remember this fact when puzzling over the paradoxes that seem to result by allowing such wffs into the system! Moreover, what logical presupposition does \((\neg p \supset p)\) express? I think the clear lesson is that such expressions should be avoided altogether - especially, when at best they just express assumptions already contained in the truth table, and at worst they produce puzzling wffs which distract logicians to the point of

\(^2\) See Endnote A for a fuller treatment.
creating elaborate systems aimed at avoiding the pitfalls of such wffs. There is one more reason for restricting their formation, which I will explore more thoroughly below.

In order to do so, I will now turn to the definition of all logical operators for a table with a single variable. There are four possibilities, illustrated below:

<table>
<thead>
<tr>
<th>Operator (#1)</th>
<th>(#2)</th>
<th>(#3)</th>
<th>(#4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>p (p) = t</td>
<td>p (p) = f</td>
<td>p (p) = t</td>
</tr>
<tr>
<td>~p</td>
<td>~p (p) = t</td>
<td>~p (p) = t</td>
<td>~p (p) = f</td>
</tr>
</tbody>
</table>

The four logical operations listed are exhaustive and the results follow from the general formula: \(2^n\), which relates the number of variables \((n)\) to the number of distinct logical combinations obtainable from \((n)\) variables. This is the second reason why wffs containing a variable and its negation on either side of a single logical operator should be eliminated, since if they were included the total number of logical operators would not follow the \(2^n\) rule.\(^3\)

Now a brief note on why I have chosen to use \(p, \neg p, q, \neg q\) instead of the customary \(t's\) and \(f's\). This is a reminder that the truth of the variables is determined by a criteria outside the particular truth table itself. That criteria could be empirical, or even logical, either way, it is not part of the role of a truth table to differentiate.\(^4\) This is also a reminder that truth comes in (at least) two different types, one which is purely logical, and is a result of rules which take on two values given by the foundation of the truth table and return a result for each unique combination of values. The failure to make this distinction helps give rise to the paradoxes relating to material implication.

Now, if we want to expand the truth table to include two variables the result will be sixteen unique logical operators. From these sixteen, we can chose which operator to represent implication. However, since the variables are logically independent we must firmly resist the temptation to conclude any one of the sixteen logical operators is an example of implication - since no \(q\) follows from any \(p\), nor does any \(p\) follow from any \(q\)! These sixteen logical operators are the bare logical rules - the rules which take two truth functionally independent

---

\(^3\) See Post (1921)

\(^4\) I will not pause to discuss object language and meta-language here, rather I will use this system to represent the values of variables when I deem it important to clearly distinguish between the truth of the variables and the resulting truth of a logical operator. Sometimes I will use the customary \(t\) and \(f\) in both cases, when that distinction is not necessary to avoid confusion. See Endnote A.
variables and return sixteen unique logical operators. The sixteen different logical possibilities for \( p \) and \( q \) are:

\[
\begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}
\hline
p & q & #16 & #15 & #14 & #13 & #12 & #11 & #10 & #9 & #8 & #7 & #6 & #5 & #4 & #3 & #2 & #1 \\
\hline
p & q & t & f & t & f & t & f & t & f & t & f & t & f & t & f & f \ \\
p & ~q & t & f & t & f & t & f & t & f & t & f & t & f & f & f & f \ \\
~p & q & t & t & t & t & f & f & f & f & t & t & f & f & f & f & f \ \\
~p & ~q & t & t & t & t & t & t & t & t & f & f & f & f & f & f & f \ \\
\hline
\end{array}
\]

The truth table is replete with information, and information is key to logical analysis! We can at this point, ask certain questions. For example, when \( p \) is true (rows one and two), which logical operations are also true? The table shows the answer to be logical operators 16, 12, 8, 4. The table also shows that logical operator 4 is identical in truth value to \( p \) itself.

Which logical operators are always true when \( q \) is true (rows one and three)? Operators 16, 14, 8, and 6. Operator 6 having the same value as \( q \) itself.

These connections are purely rule based. They follow from the requirement to take all possible truth values for two truth functionally independent variables and list all possible logical operations on those two values. Looking at the connection between \( p \) and logical operator 12, one can say that when \( p \) is true, so must be \((p \subset q)\), where \((p \subset q)\) is a logical truth. Logicians are tempted to express this relationship as \( p \supset (q \supset p) \) and subsequently use that expression as a paradigm.

---

5 While it is true that the Sheffer stroke (15) or the Peirce arrow (9) are sufficient to represent all logical operators, they can not do so in one single operation between two variables.
6 Here I include, when possible, the logical symbols given in Church (1956) page 37 or Wittgenstein (1922) sec. 5.101 (The exceptions are #16, #9, #2, and #1)
7 Since \((p \subset q) \equiv (q \supset p)\)
case of a paradox of material implication. However, my analysis shows that this
can not be seen as an instance of true entailment, as the relationships here are
purely logical, due to the reasons stated above. The truth of p itself is determined
by a criteria outside the system, the truth of (q ⊃ p) follows from the logical rules.

The relationship between variable truth and logical operator truth are difficult to
see, hence I suggest the following be adopted as an alternative way to write truth
tables. Let the truth values of the variables be written differently than the truth
values of the resulting logical operation on the truth value of those variables.
While this is not absolutely necessary, and is some cases proves to be
undesirable (e.g. when a single variable also appears as a premise), doing so
helps keep the distinction between the criteria which determine the values found
in each column of the table up front, and helps greatly in avoiding the many
confusions which follow when the same symbol is used to represent both
variable truth and truth which is the result of logical operations. Hence the above
table could be re-written (looking at a few familiar operators):

<table>
<thead>
<tr>
<th>p</th>
<th>q</th>
<th>⊃</th>
<th>v</th>
<th>&amp;</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>q</td>
<td>p ⊃ q = t</td>
<td>p v q = f</td>
<td>p &amp; q = t</td>
</tr>
<tr>
<td>p</td>
<td>~q</td>
<td>p ⊃ ~q = t</td>
<td>p v ~q = t</td>
<td>p &amp; ~q = f</td>
</tr>
<tr>
<td>~p</td>
<td>q</td>
<td>~p ⊃ q = t</td>
<td>~p v q = t</td>
<td>~p &amp; q = t</td>
</tr>
<tr>
<td>~p</td>
<td>~q</td>
<td>~p ⊃ ~q = t</td>
<td>~p v ~q = t</td>
<td>~p &amp; ~q = t</td>
</tr>
</tbody>
</table>

3. A model for the conditional as entailment

Now it is desirable to choose one of these sixteen logical operators on which to
base the full conditional that will express entailment. Whichever logical operator
we chose, it should be remembered that it will be a model on which the full
conditional is built - not an expression of that conditional itself (since p and q are
truth functionally independent).

I will take it as self-evident that any logical operation that is to model the
conditional must make (t ⊃ f) be false. There are two other fundamental aspects
of the conditional that help decide the issue, that of tautology (t ⊃ t), and non-
commutation (t ⊃ f) ≠ (f ⊃ t).
From tautology we get:
(t ⊃ t) = t, (f ⊃ f) = t,
This leaves us with:

8 The same reasoning should be used to eliminate “v”, “t”, and “f” (upper or lower case) as letters
for propositional variables!
9 I am using here the symbol “⊃” without prejudice as to which operation should represent the
conditional, as is customary at this point in the argument.
<table>
<thead>
<tr>
<th>p</th>
<th>q</th>
<th>( p \supset q )</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>q</td>
<td>( (t \supset t) = t ) - from tautology</td>
</tr>
<tr>
<td>p</td>
<td>~q</td>
<td>( (t \supset f) = f ) - from the minimum requirements for a conditional</td>
</tr>
<tr>
<td>~p</td>
<td>q</td>
<td>???</td>
</tr>
<tr>
<td>~p</td>
<td>~q</td>
<td>( (f \supset f) = t ) - from tautology</td>
</tr>
</tbody>
</table>

This leaves only the problematic third row. Since conditionals distinguish themselves by having antecedents and consequents (i.e., they do not commute), this means that \( p \supset (\sim q) \neq (\sim q \supset p) \). Since we have already established that \( p \supset (\sim q) \) must be false, then from the logical equivalence of \( (\sim q \supset p) \equiv (\sim p \supset q) \) which can not have the same value as \( p \supset (\sim q) \) [which is false], then \( (\sim p \supset q) \) must be true.10

4. How truth tables resolve the paradoxes of material implication

I have made the claim that truth-table analysis of arguments eliminates many of the problems related to the so called "paradoxes of material implication". To illustrate this lets re-examine the list of paradoxes:

1) \( p \supset (q \supset p) \)
2) \( p \supset (q \supset q) \)
3) \( (p \& \sim p) \supset q \)
4) \( p \supset (q \lor \sim q) \)
5) \( p \supset (\sim p \supset q) \)

I have chosen to use the variables "p" and "q" instead of the capital letters "A", "B", "K" etc. If, as is customary, capital letters stand for logical sentences which can be further analyzed [e.g. \( B = (p \supset (q \supset p)) \)], then truth tables will not represent the information needed to resolve the paradoxes. As a matter of fact, truth tables could not be used at all, since one would not know whether "A" and "B" are logically independent.

First we can start with what Anderson and Belnap, call the "archetype of fallacies of relevance": the logical sentence: \( p \supset (q \supset p) \).
The problem with this expression, according to Anderson and Belnap is that, "It would enable us to infer that Bach wrote the Coffee Cantina from the premise that the Van Allen belt is doughnut shaped - or indeed from any premise you like." 11

This does not follow from a truth table model of entailment.

---

10 Perhaps this is better understood if I revert to the customary way of representing things. Since \( (t \supset f) \) must be false, and \( (t \supset f) \neq (f \supset t) \) [because \( (p \supset q) \neq (q \supset p) \)], then \( (f \supset t) \) must be true.

11 (1975) page 30.
What $p \supset (q \supset p)$ illustrates is the same information already discovered in our examination of the sixteen possible logical operators on two independent variables above. In that case, we can interpret the logical operator #12 as $(q \supset p)$ which is indeed true whenever $p$ is true, but the truth of $p$ is determined by a different criteria than logical operator #12: $(q \supset p)$. At this point, as I have already argued, one can not yet set the stage for full entailment. This is also a case where the distinction between logical truth (the rules that govern the truth output for $(q \supset p)$) and the rules which govern the truth or falsity of $p$ - which lie outside the truth table- helps resolve the issue.

Logicians routinely make this distinction in types of truth. For example, if I wished to be a weatherman and always get my forecast right, I could simply say, "if it rains, then it rains", "if it snows, then it snows", "if it breaks 85°, then it breaks 85°". These are examples of logical truths, although they take on as variables empirical truths. Clearly Anderson and Belnap's concern deals with being able to deduce an empirical truth (the Van Allen belt is doughnut shaped) from another (false) empirical truth (Bach wrote the Coffee Cantina).12

Here a very important point needs to be made. One chooses $p$ and $q$ to represent propositions in English only if those propositions are thought be to logically independent. This choice reflects a previous judgment about relevance that occurs outside logic. Logic itself is powerless to dictate this choice, but rather shows the results of that choice.

If one equates the meaning of wffs to the rules that govern their construction and truth evaluation, and their use as knowing when such a wff can stand for a proposition in English, then one can sensibly distinguish between meaning and use. The logical meaning of $(q \supset p)$ is knowing when it is true and when it is false as determined by the rules of logic. Another part of the logical meaning is knowing that one reads, $(q \supset p)$ from left to right, and that $(q \supset p)$ is not the same as $(p \supset q)$. All of this is demonstrated by the truth-table definition of logical operators and the rules which govern proper syntax and the methods used to determine logical equivalence. When $(q \supset p)$ is to be used to substitute for an English conditional is not part of the previously defined logical meaning of $(q \supset p)$!

Classical logic has a determinate meaning in the truth tables, the use of classical logic to represent English sentences that employ phrases that contain, "if . . . then" in a variety of ways is governed by different constraints. Hence, logic will always play a secondary role to the extra-logical requirements in its Philosophical use of argument analysis. The logician can err in insisting that a certain wff must represent an argument in English - however it would be wrong to confuse the error of a logician as a fault or incompleteness in Logic itself. The standard of logical correctness must not be the intuitions of the logician!

---

12 There is not enough space in this paper to explore the consequences for a theory of deduction or a development of relevance and other logical connections based on truth table semantics.
This same reasoning applies to all the other paradoxes of material implication on
the list except number three and four. These forms violate the rule where one
may not form a wff with a variable and that variable’s negation. Hence \((p \& \sim p) \supset q\) and \(p \supset (q \lor \sim q)\) represent an impossible state of affairs where one is just as
justified as saying "nothing follows" as saying that anything follows! \(^{13}\)

Kent Slinker
Pima College
Downtown Campus
RV 150
1255 North Stone Ave.
Tucson, Arizona 85709-3000
kslinker@pima.edu

\(^{13}\) My sincere thanks to Aeyn Edwards Wheat for his careful reading and suggestions.
References


Post, Emil (1921). "Introduction to a General Theory of Elementary Propositions". *American Journal of Mathematics*. Vol. 43, No. 3 (July)


A. Since the restriction of wffs of the form \((p \& \sim p)\) forms a central and novel part of this paper, some further explanation is helpful. First, one might argue that such a restriction would equally forbid going from the valid argument form:

\[
\begin{array}{c}
p \lor q \\
\sim p \\
\therefore q
\end{array}
\]

To the logical sentence:

\[
[(p \lor q) \& \sim p] \supset q
\]

Since the logical sentence has an instance of both \(p\) and \(\sim p\) on the same line. This would indeed be a fatal problem in my analysis. However, as I stated in my explication, this restriction holds only for a single logical operator which takes on values passed to it from the truth table. This is because the truth table represents logical independence of bivalent variables. This independence shows tautology \((p \lor \sim p)\) in the columns and forbids contradiction in the rows \((p \& \sim p)\). The logical sentence:

\[
[(p \lor q) \& \sim p] \supset q
\]

results in a wff whose constituent parts (because of the presence of \((p \lor q)\)) are not logically independent, hence no restriction applies. If we use the symbols "t" and "f" to represent the result of a logical operation on any truth functionally independent variable only, then there is no restriction on a truth value and its negation between single logical operators \([e.g. \ (f \& \sim f)]\). One could add the following rule for truth functional analysis. When a single variable appears as a premise in an argument this occurrence in the truth table results in a logical operation which turns \(p\) to "t" and \(\sim p\) to "f". Hence, disjunctive syllogism would be represented as:

<table>
<thead>
<tr>
<th></th>
<th>((p \lor q))</th>
<th>~p</th>
<th>\therefore q</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p)</td>
<td>q</td>
<td>t</td>
<td>f</td>
</tr>
<tr>
<td>(p)</td>
<td>~q</td>
<td>t</td>
<td>f</td>
</tr>
<tr>
<td>(\sim p)</td>
<td>q</td>
<td>t</td>
<td>t</td>
</tr>
<tr>
<td>(\sim p)</td>
<td>~q</td>
<td>f</td>
<td>t</td>
</tr>
</tbody>
</table>