If the Moon is Green Cheese, then I'm a Monkey's Uncle

Terry R. McConnell

November 22, 2010

> If that that you swear by is not,
> then you are not foresworn.
>
> Touchstone, *As You Like it*

Introduction

This is an essay about implication, which is the kind of statement we are making when we say "If A, then B." Implication *seems* quite straightforward – the truth of B follows from the truth of A – but enough people are confused by true statements like the one in the title that it seems worth discussing this slippery logical connective in detail.

Some of the confusion arises from misunderstanding of what, exactly, an implication asserts. My elementary logic students invariably smirk when I tell them that "If the moon is green cheese, then I'm a monkey's uncle." It may be that the mere mention of a monkey's uncle prompts them to entertain the possibility that I might be one. More likely, they are confusing the truth of the implication, A implies B, with that of its parts, A and B. For example, the assertion "If it is raining, then it is cloudy," says nothing about whether it is cloudy or raining right now. The statement can be made truthfully on a sunny day, and it is just as true then as on any other day. It says only that a condition of cloudiness *would follow* from a condition of raininess if, indeed, it were raining.

A deeper cause of confusion has to do with the way the truth of the overall implication depends on the truth of its parts. In order to discuss this issue sensibly it is necessary to review some concepts from elementary logic, which we shall do in the following section. After that, we shall discuss the role of implication in mathematics. Finally we will argue that the technical meaning of implication is engineered for mathematical convenience, and that its peculiarities are largely irrelevant in practise.

Sentential Logic

Sentential logic is a stripped-down mathematical model of human thought. It analyzes expressions ("sentences") built using **statement letters** $(A, B, \dots)$ and **connective symbols** ($\vee \ \sim \Rightarrow \ \Longleftrightarrow \ \wedge$.) For example, here is a sentence:

$$(A \wedge \sim B) \Rightarrow C. \tag{1}$$

A serious introduction to logic would address issues of syntax, e.g., how parentheses must be placed in order to resolve ambiguities over which parts of an expression a given connective actually connects. We shall not go into these issues here. Indeed, we assume the reader has at least a passing acquaintance with sentential logic and its connectives already, and we include this brief discussion only to jog the memory.

Sentences like (1) acquire meaning only when the letters are *interpreted*, i.e., when statements about actual physical or mathematical objects are substituted for the letters. For example, A might stand for "I am healthy," B for "I am rich," and C for "I am happy." Then a translation of the entire sentence into English is "If I am healthy and not rich then I am happy." Alternatively, "To be healthy though poor is to be happy nevertheless" expresses the same thing, perhaps more eloquently. English words, especially the tenses and moods of verbs, often have connotations that cannot be rendered in logical notation.

As a result of an interpretation, each statement letter acquires a **truth value,** either T or F. The truth value for an entire sentence can then be deduced using the **truth table** for each connective. The "or" connective, for example, is defined by:

A	B	$A \vee B$
T	T	T
T	F	T
F	T	T
F	F	F

Here are truth tables for some of the other commonly used connectives:

A	B	$A \wedge B$
T	T	T
T	F	F
F	T	F
F	F	F

A	B	$A \Longleftrightarrow B$
T	T	T
T	F	F
F	T	F
F	F	T

A	B	$A \Rightarrow B$
T	T	T
T	F	F
F	T	T
F	F	T

For the most part, these truth tables affirm the everyday interpretations of their respective connectives. For example, the truth table for the "and" connective, $\wedge$, shows that an expression $A \wedge B$ is true when **both** A and B are true, and in no other case. This corresponds closely with the meaning of the word "and" in English.

On the other hand, the truth table for the implication ($\Rightarrow$) connective, once its implications[1] are grasped, seems downright weird. The first row is reasonable enough. When A and B are both true it seems quite natural that $A \Rightarrow B$ should be true. But why should $F \Rightarrow T$ or $F \Rightarrow F$ be true? These lines of the table lead to true statements like the one in the title. They are often summarized as "a false hypothesis implies any conclusion."

People are sufficiently puzzled by this aspect of implication that elementary texts usually try to convince the reader that the truth table for implication is the correct one. Such arguments often take the following form: imagine that your high school guidance counselor has told you that "if you study hard, you will get into college." When you don't study hard, and either do or don't get into college, did the counselor lie to you? Of course not (the texts claim,) so assigning a T as the truth value of the implication in these cases is the correct decision.

We can't be certain exactly what the guidance counselor intended with the above statement, but it is likely that he had in mind something more along the lines of the equivalence ($\iff$) truth table. That is, he probably intended for you to think also that if you don't study hard, then you won't get into college, but that interpretation is not supported by a strict reading of the truth table. "If A then B" in everyday usage often imputes a causal connection between A and B, i.e., B follows from A because A is a cause of B. This interpretation does nothing to explain the troublesome bottom two rows of the implication truth table since a false A cannot be the cause of anything. It seems misguided to justify definitions with everyday statements whose intended meaning does not match the definitions. Rather, we shall argue below that the definition of implication is chosen for mathematical convenience rather than the desire to accurately model its meaning in everyday usage.

Implication in Mathematics

Mathematical theorems often take the form of an implication: *If* a function is differentiable, *then* it is continuous. *If* x is a natural number, *then* it is a sum of at most 4 squares, etc. To be sure, some theorems assert the equivalence of two statements, but these can be viewed as two implication theorems rolled into one. This is because the sentence $A \iff B$ is identical, in a very strong sense, to the sentence $(A \Rightarrow B) \wedge (B \Rightarrow A)$. Draw the truth tables for the two side by side and you will see that they match up exactly. Such pairs of sentences are said to be **logically equivalent.** For all intents and purposes they can be regarded as identical. Mirroring this, proofs of equivalence theorems are often broken into two "halves," one proving that $A \Rightarrow B$ and the other that $B \Rightarrow A$.

The notion of containment ($\subseteq$) in set theory is a first cousin of the logical concept of implication. Establishing $A \subseteq B$ for sets A and B reduces to establishing the implication $x \in A \;\Rightarrow\; x \in B$. The fact that a false hypothesis implies

[1] Sorry.

any conclusion also has a close relative in set theory: the empty set has essentially any property you might care to name. (Quite a feat, for something that is, effectively, nothing!) For example, the empty set is an arcwise connected, locally arcwise connected Hausdorff topological space. Don't worry about what any of that mumbo-jumbo means, because you can truly replace it with any property of a set that can be expressed in terms of its members. The reason is again the same curious property of implication: since $x \in \emptyset$ is always false, the implication $x \in \emptyset \Rightarrow$ (anything whatsoever) is always true.

These considerations lead naturally to several conventions in mathematics that seem quite puzzling when first encountered. Consider, e.g., the notion of *infimum,* or greatest lower bound. If S is a set of real numbers, its infimum, denoted $\inf(S)$, is that number having the following the two properties: (i) it is less than or equal to every member of S; and, (ii) no larger number has property (i). For example, the closed interval [0,1] and the open interval (0,1) each has infimum equal to zero.

Now we may ask, what should be the infimum of the empty set? We tend to think of an infimum as being something small – it is, after all, a *lower* bound – and the empty set is the smallest possible set. Therefore it may come as a surprise to learn that $\inf(\emptyset) = +\infty$! How can this be so? It is an inevitable consequence of the definition. Note that for any real number y the sentence $x \in \emptyset \Rightarrow y \leq x$ is true. Thus, any real number y qualifies as a *bona fide* lower bound for $\emptyset$. It follows that in order for (ii) to hold for $\inf(\emptyset)$, it must equal $+\infty$.

To illustrate the dangers of failing to appreciate this property of the empty set, consider the old story, perhaps apocryphal, about the graduate student who wrote a dissertation on an arcane class of mathematical objects called *foobers.* It seems that foobers are amazing mathematical objects. The student, in his dissertation defense, went on at length to present some of their wonderful properties, and to give a hint of the clever proofs he had discovered. Late in what was a very long presentation, a voice from the back of the room asked, "Excuse me, but before you go on, could you give us an *example* of a foober?" In the embarassed silence that followed it became clear, first of all, that the student could not *think* of an example; and ultimately, that there *was* no example: the axioms for a foober were self-contradictory! It was especially ironic that the student's proofs were correct in every detail.

The Truth About Implication ?

Most people do not find the equivalence connective $\iff$ troublesome or controversial. Indeed, many seem to want implication to be the same thing as equivalence, or they confuse an implication $A \Rightarrow B$ with its *converse* $B \Rightarrow A$. Consider, e.g., the everyday statement "I will be grounded unless I return by 10:00 p.m." The most straightforward way to render this statement in symbols is $A \Rightarrow B$, where A is "I don't return by 10:00 p.m.", and B is "I am grounded." (The dictionary meaning of "unless" is "if not".) Does it follow then that when I do return by 10:00 p.m. I can expect not to be grounded? No. This would follow from the reverse implication, $B \Rightarrow A$. It may be that my mean parents are resolved to ground me no matter what I do!

It is convenient, indeed indispensible, to be able to express *half* of the $\iff$ relationship. In ignorance, let us name this relationship $\Rightarrow$ and see what comes of it. What should be the truth table of this connective? Certainly, whatever else it may mean, $\Rightarrow$ should not contradict $\iff$: it should be true whenever $\iff$ is. So already, we can fill in the truth table for $\Rightarrow$ this far:

A	B	$A \Rightarrow B$
T	T	T
T	F	?
F	T	?
F	F	T

What should we take for the remaining rows of the table? If we fill in both question marks with a T, we obtain a connective which is always true. Such a connective is clearly not very useful. Both F's gives back $\iff$, so the only real possibility is to make one ? into a T and the other into an F. Which way you do it makes little difference, for one defines $\Rightarrow$ and the other defines $\Leftarrow$, the other "half" of $\iff$.

Perhaps, the entire issue is moot. The last 3 rows of the truth table of implication are largely irrelevant to any constructive purpose in reasoning. A false implication is of no use to anybody, thus eliminating the second line of the table from serious consideration. The only use of the last two lines is in silly examples, like the title of this essay, in which a false assumption is made to truly imply any desired conclusion. So, in the end, the only constructive use of implication is via the first row of its truth table. It is this line which forms the basis of **modus ponens,** the fundamental law of thought which allows us to deduce B when *both* A and $A \Rightarrow B$ are true. The chosen definition of $\Rightarrow$ makes a modus ponens deduction of B from A possible, and it is the least restrictive such definition. In other words, in order to prove the truth of $A \Rightarrow B$ we need to do only a minimum amount of work: we need to verify only that the truth of B follows when A is assumed to be true as a hypothesis. Everything else comes for free.