Anti Counterfactuals

Eric C.R. Hehner

1985 May 13

An implication is an expression of the form

antecedent ⇒ *consequent*

A counterfactual is an implication whose antecedent is false. Matt Ginsburg of Stanford University began an interesting seminar last week by saying that counterfactuals are difficult to understand. He said that people are uneasy with the idea that false implies anything, and so it must be unnatural. He then introduced a new kind of implication: counterfactual implication. On the left below, using $\top$ for true and $\bot$ for false, we have the old, ordinary (material) implication; on the right, Ginsburg's new counterfactual implication.

ordinary implication

a	b	$a \Rightarrow b$
$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\top$
$\bot$	$\bot$	$\top$

counterfactual implication

a	b	$a \Rightarrow b$
$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$
$\bot$	$\top$	?
$\bot$	$\bot$	?

The value ? stands for “unknown”, a third logical value; Ginsburg says this is more intuitive than either $\top$ or $\bot$ where it appears. According to Ginsburg, when someone makes a statement of the form “if a then b”, and a is false, the statement should (often? always?) be formalized as counterfactual implication. I believe that Ginsburg has misunderstood the purpose and use of logic. When someone makes a statement, they are saying that the statement is true. (Whether it really is true or not is not at issue.)

Suppose that a speaker states an implication “if a then b”, and a is true. We can then infer that b is true by looking at the truth table for ordinary implication: find all rows in which a is $\top$ and $a \Rightarrow b$ is $\top$; there is just one such row, and in it b is $\top$.

Suppose that someone states an implication “if a then b”, and a is false. Look again at the truth table for ordinary implication. Look at all rows in which a is $\bot$ and $a \Rightarrow b$ is $\top$ (because the speaker is telling us that the implication is true). There are two such rows, and in one b is $\top$ and in the other b is $\bot$. Hence we do not know whether b is true or false. It is the consequent that is unknown, not the implication. Counterfactual implication does not represent what was being said.

Here is what is strange about an implication whose antecedent is known to be false: the speaker seems to be trying to tell us something, but is not telling us anything new. The implication was already deducible from its antecedent. We ask ourselves: what is the speaker really trying to say? Let me illustrate using Ginsburg's own example. Suppose we know

(a) Dinner is late.
(b) The electricity is off.

and the speaker says

(c) If the electricity were not off, dinner would not be late.

With the obvious translation to formal logic, the antecedent of (c) is false, so we can deduce that the implication (c) is true, so the speaker needn't tell us that (c) is true. So perhaps we have not fully understood what the speaker meant. Perhaps the formalization should be

(a') $late\ (today)$
(b') $off\ (today)$
(c') $\forall date \cdot\ \neg off\ (date) \Rightarrow \neg late\ (date)$

Here, (c') is indeed additional information, not implicit in (a') and (b'). It says that whenever the electricity is not off, dinner is not late. Or, equivalent to (c'), we could write

(c'') $\forall date \cdot\ late\ (date) \Rightarrow off\ (date)$

which says that whenever dinner is late, (it is because) the electricity is off. Version (c'') is clearly not counterfactual. One might be reluctant to admit that (c') and (c'') are equivalent because in English, "if a then b" sometimes means "a causes b". In logic, implication does not mean cause. If the speaker meant that the electricity being off is the cause of dinner being late, we must invent a *cause* relation, and formalize (c) as

(c''') $cause\ (off, late)$

Whether (c) is true is not our concern; we are being told that it is true, so we take it as being true. We formalize our understanding as (c') or (c''') using ordinary logic; counterfactual implication cannot express it.

Ginsburg has done some very nice work, but it has nothing to do with counterfactual implication. In its general setting, it is the following. Suppose we have a knowledge base of consistent axioms (in ordinary logic). And then we are presented with one more axiom that is not consistent with the knowledge base (it need not be an implication). Suppose further that we trust the new information more than we trust the knowledge base. What is the minimal change we must make to the knowledge base to allow us to add the new axiom without inconsistency? Ginsburg has defined "minimal change" in a very nice way, and has a procedure for finding a minimal change. It is unfortunate that it is not unique, and that finding the "right" one may involve a dialogue, but that's life. I think this is an important piece of work. It should be presented without the confusing and wrongheaded counterfactual implication.