Linguistics 280: Language and Formal Reasoning

Propositional logic: truth tables vs. inference

Robert Levine

Autumn Quarter, 2010

Truth tables for complex formulæ

In the preceding file, we introduced truth tables as, in effect, definitions of the logical connectives. These connectives are defined so as to model—in simplified, standardized form—elements of natural language vocabulary that are crucially involved in formal reasoning. And formal reasoning, as noted in the Overview file, serves as a model of the computation that ordinary, informal reasoning carries out in going from particular assumptions to the necessary consequences of those assumptions. It follows that truth tables play a central role in establishing the connection between natural language on the one hand and our reasoning ability on the other. But the relationship is not a simple one.

To see what the actual relationship is, it's useful to start with a formula which is just a bit more complex than the ones which we used to define the connectives in the previous file. Take something simple, such as $p \wedge (q \vee \neg p)$. Under what conditions is a sentence of a human language that has this logical form true? The way to see this is to work out a truth table for the formula and see when it shows up with the truth value 1.

It might seem to you that the task here is considerably harder than it was to state the truth tables for $\wedge$, $\vee$ and so on, because we are dealing with many more symbols than appear in the previous truth tables. In particular, truth values are defined in truth tables only for, at most, *pairs* of logical statements. How do we develop truth tables for formulae which appear to have so many more 'moving parts'? But if you look at the statement $p \wedge (q \vee \neg p)$, you'll see that it actually consists of two formulæ linked by a connective; things look more complex only because one of these two formulæ is itself complex:

$$\underbrace{p}_{p} \wedge \underbrace{(q \vee \neg p)}_{q \vee \neg p}$$

So if we have a truth value for p, on the one hand, and for $q \vee \neg p$ on the other, then we can combine these two truth values by consulting the truth table for $\wedge$. This observation gives a key to proceding: set out the truth values for p and q, and for each combination of the two truth values, work out $\neg q$ and combine that result with p to get the truth value of the second conjunct $q \vee \neg p$. The first conjunct is just p, which will have a truth value in each case. So, with values for both p and $q \vee \neg p$, we can refer to the rule for $\wedge$ given in the previous file to determine the truth value of $p \wedge (q \vee \neg p)$.

If we follow this procedure, we'll need a column for $p, q, \neg q, q \vee \neg p$, and, finally, $p \wedge (q \vee \neg p)$, with this last column showing in each case how the particular combination of truth values for p and q determines the truth value of the whole formula.

To see how this works out in practice, we write down the columns just listed for the truth table for $p \wedge (q \vee \neg p)$.

p	q	$\neg p$	$q \vee \neg p$	$p \wedge (q \vee \neg p)$

You should be able to see that the truth value of the formula heading each column can be determined by the truth values for the formulæ to the left of that column, as determined by the particular conjunction involved. To start with, we pick truth values for p and q:

p	q	$\neg p$	$q \vee \neg p$	$p \wedge (q \vee \neg p)$
1	1			

We can supply the value for the $\neg p$ column by consulting the truth table for $\neg$: when the truth value of p is 1, the truth value of $\neg p$ is 0:

p	q	$\neg p$	$q \vee \neg p$	$p \wedge (q \vee \neg p)$
1	1	0		

We now have the information we need to supply a truth value for $q \vee \neg p$: when one of the disjuncts is 1 and the other is 0, the truth table for $\vee$ specifies the value of the disjunction as 1:

p	q	$\neg q$	$q \vee \neg p$	$p \wedge (q \vee \neg p)$
1	1	0	1	

Finally, we can take the value of p and the value of $q \vee \neg p$, consult the truth table for $\wedge$, and specify the value for the conjunction of these two formulæ, i.e., $p \wedge (q \vee \neg p)$:

p	q	$\neg p$	$q \vee \neg p$	$p \wedge (q \vee \neg p)$
1	1	0	1	1

When both p and q have the truth value 1, so does the formula $p \wedge (q \vee \neg p)$.

Carrying out the same set of steps for the other three possible combinations of p and q, we wind up with the following truth table:

p	q	$\neg p$	$q \vee \neg p$	$p \wedge (q \vee \neg p)$
1	1	0	1	1
1	0	0	0	0
0	1	1	1	0
0	0	1	1	0

The upshot is that the statement $p \wedge (q \vee \neg p)$ is true under exactly one condition: when both premises are true. A sentence such as *Robin is a spy and either Leslie is a saboteur or Robin is a spy*, which has the logical form $p \wedge (q \vee \neg p)$, will be true if and only if *Robin is a spy* is true and *Leslie is a saboteur* is true.

The method of truth table construction for any formula procedes in exactly the same way, no matter how complex the formula is; the only difference is the number of steps involved. To see how this construction works for something a bit more complex, we'll evaluate the logical statement $\neg(p \vee \neg q) \supset (\neg p \vee q)$. In setting up a truth table for a formula, you need to start by identifying the suparts of that formula. The logical statement we're taking on here is fairly complicated-looking—but again, it turns out to be nothing more than several pairs of statements linked by a connective. We begin by identifying the 'top-level' connective. There will always be such a connective, and an easy way to find it is looking for the logical symbol which has expressions of the form $(\ldots)$ or $\neg(\ldots)$ on both sides of it. (You have to be careful here; if negation is the topmost logical operator, then what the formula will look like is $\neg(\ldots)$, so there won't be anything to the left of ¬xs. If you see that, you know negation is 'highest' connective.) Break down each of the formulæ operated on by the top-level connective by, in effect, pretending that each of them is the top-level connective, so that you're again looking for a connective with $(\ldots)$ on both sides. Eventually, you'l get to the point where the only thing on either either side of the connective is a single symbol (or ¬ followed by a single symbol), and at this point you know what columns you need to set up for the truth table.

If we carry out this procedure, this is what we'll wind up with:

$$\underbrace{\neg(p \vee \neg q)} \supset \underbrace{(\neg p \vee q)}$$

The topmost connective must be $\supset$. We can then break down each of the formula on either side of $\supset$ as follows:

$$\neg \underbrace{(p \vee \neg q)}, \underbrace{\neg p} \vee \underbrace{q}$$

On the next cycle, the procedure yields

$$\underbrace{p}, \underbrace{\neg q}$$

and

$$\neg \underbrace{p}$$

This procedure tells us that we will need columns for the atomic symbols p, q; the negations of these symbols $\neg p$, $\neg q$; the complex formulæ $p \vee \neg q, \neg(p \vee \neg q)$ and $\neg p \vee q$, and the combination of these formulæ into the still more complex $\neg(p \vee \neg q) \supset (\neg p \vee q)$. Each of these formulæ will get its own column. To find the truth value of the formula at the top of any given column, we identify the truth values of the two formulæ which combine to form it, and then use the truth table for the connective which joins them to determine what the result of the combination will be.

The following solution follows this procedure exactly. The result is that, to find the truth values of $\neg(p \vee \neg q) \supset \neg p \vee q$ as p and q change their truth values, we need to fill in a truth table whose columns look like the following:

p	q	$\neg p$	$\neg q$	$p \vee \neg q$	$\neg(p \vee \neg q)$	$\neg p \vee q$	$\neg(p \vee \neg q) \supset (\neg p \vee q)$

The next step, as in the earlier example, is to go systematically from column to column, building up the truth values of each column based on the ones to this left, using the truth tables for the connectives as a guide. In the present case, as in the truth table we constructed earlier, we have four possibilities, created by the fact that p and q can each take on the truth values 0 and 1. For any given truth values of p and q, the truth values of $\neg p, \neg q$ follow immediately from the truth table for ¬:

p	q	$\neg p$	$\neg q$	$p \vee \neg q$	$\neg(p \vee \neg q)$	$\neg p \vee q$	$\neg(p \vee \neg q) \supset \neg p \vee q$
1	1	0	0				

Given the value of p and $\neg q$ we're considering, the truth value of $p \vee \neg q$ is immediately given by the truth table for $\vee$:

p	q	$\neg p$	$\neg q$	$p \vee \neg q$	$\neg(p \vee \neg q)$	$\neg p \vee q$	$\neg(p \vee \neg q) \supset \neg p \vee q$
1	1	0	0	1			

and that for $\neg(p \vee \neg q)$ is again given by the truth table for $\neg$:

p	q	$\neg p$	$\neg q$	$p \vee \neg q$	$\neg(p \vee \neg q)$	$\neg p \vee q$	$\neg(p \vee \neg q) \supset \neg p \vee q$
1	1	0	0	1	0		

Similarly, given the values determined for $\neg p$ and q (as identified under their respective columns), we can fill in the value of $\neg p \vee q$ as

p	q	$\neg p$	$\neg q$	$p \vee \neg q$	$\neg(p \vee \neg q)$	$\neg p \vee q$	$\neg(p \vee \neg q) \supset \neg p \vee q$
1	1	0	0	1	0	1	

and, on the basis of the sixth and seventh columns and the truth table for $\supset$, the last column must have the value given:

p	q	$\neg p$	$\neg q$	$p \vee \neg q$	$\neg(p \vee \neg q)$	$\neg p \vee q$	$\neg(p \vee \neg q) \supset \neg p \vee q$
1	1	0	0	1	0	1	1

Applying exactly this same procedure again to the three remaining combinations of truth values for p and q, we obtain the following complete truth table for $\neg(p \vee \neg q) \supset \neg p \vee q$:

p	q	$\neg p$	$\neg q$	$p \vee \neg q$	$\neg(p \vee \neg q)$	$\neg p \vee q$	$\neg(p \vee \neg q) \supset \neg p \vee q$
1	1	0	0	1	0	1	1
1	0	0	1	1	0	0	1
0	1	1	0	0	1	1	1
0	0	1	1	1	0	1	1

The foregoing illustrates not just the steps involved in filling out the truth table for a complex formula, but a particularly interesting pattern of truth values: the formula $\neg(p \vee \neg q) \supset \neg p \vee q$ is *always true*, no matter what the value of p and q are. This is, on the face of it, quite an unexpected outcome—apparently, there are *no* circumstances under which the formula is ever false. (Such logical statements are called *tautologies*); the opposite sort of statement, which is always false regardless of its atomic formulæ' truth values, is called a *contradiction*). Why should we find things be like that, in the case of this particular formula? Is it sheer 'accident'? The problem with this question is that it assumes some notion of 'logical accident', because of which the formula $\neg(p \vee \neg q) \supset \neg p \vee q$ is *necessarily* true—true, that is, no matter what the truth of its component 'atomic' premises are. But truth values, as we've seen, are determined in a completely systematic fashion, so it's hard to see what the notion 'logical accident' could mean. We have much more reason to suspect that there is some deeper connection between the antecedent and the consequent in this formula that guarantees a value of 1 for the latter, no matter what the values of p and q are. We'll return to this question from a somewhat different angle in the following section.

Exercises:

For each of the following formulæ,

i. provide a truth table reflecting all possible values of its atomic formulæ, using the truth table constructed above for $\neg(p \vee \neg q) \supset \neg p \vee q$ as a guide, and
ii. identify whether or not the formula in question is a tautology.

1. $p \vee \neg p$
2. $p \supset q \supset (q \supset p)$
3. $q \supset (p \supset q)$
4. $p \vee q \supset p \wedge q$
5. $(p \vee \neg q) \wedge (q \vee \neg p)$
6. $p \vee (q \wedge r) \supset (p \vee q) \wedge (p \vee r)$
7. $(p \supset (q \supset r)) \supset ((p \wedge q) \supset r)$
8. $(p \supset q) \supset ((p \vee r) \supset q)$

The limitations of truth tables

In a sense, truth tables would seem to make logic, as a theory of reasoning, unnecessary. After all, if the point of logic is to enable us to identify what validly follows from a true statement, all we need to do is inspect the various truth tables to see what statements are necessarily also true when any given statement is taken to be true. For example, let's consider the two statements $p \supset q$ and $\neg q \supset \neg p$. Does the truth of the latter follow from the truth of the former? If $\neg q \supset \neg p$ is always true just in case $p \supset q$ is true, then the answer is yes. What do the truth tables tell us?

p	q	$\neg q$	$\neg p$	$p \supset q$	$\neg q \supset \neg p$
0	0	1	1	1	1
0	1	1	1	1	1
1	1	0	0	1	1
1	0	1	0	0	0

Sure enough, wherever $p \supset q$ is true, $\neg q \supset \neg p$ is true. The truth of the first apparently guarantees the truth of the second, in the sense that we know that whenever the first is true, we also know that the second is ture, with no exceptions. So all we need are the truth tables. We may as well stop here and all go home, eh?

Or is there something not quite comfortable about this conclusion?

Let's pursue the matter a bit further. You tell me that if Dana turns out to be a spy, we'll lose our jobs at the screening agency where we both work, and where we have recently cleared Dana for access to maximum security level information. We don't lose our jobs at the screening agency—from which, you will reason, we can tell that Dana hasn't been identified as a spy. My reaction is, well, how can you say that? Where did you get *that* idea from? You reply that it follows from the very meaning of the original warning you gave me. If it's true, as you told me, that Dana being a spy means we'll lose our jobs, that just *means* (as you say in some exasperation) that losing our jobs will happen automatically once Dana's activities as a spy are discovered. So keeping our jobs is not consistent with Dana being a spy—the two situations are contradictory—and the way the universe works, that means that you can't have it both ways. Consequently, as long as we're not fired, it can only mean that Dana has failed to be identified as a spy. See?

Yes, I see. And it's hard to argue with that kind of reasoning. But suppose you had come to me and said the same thing about Dana and our jobs, and how the fact that we hadn't been fired meant Dana must not have been a spy—and then, when I asked you how you knew, you had said, well, compare the truth tables for a formula corresponding to *If Dana is revealed to be a spy, we'll lose our jobs*, on the one hand, and the the truth table for the formula corresponding to *If we don't lose our jobs, Dana won't have been revealed as a spy*, and if I do that, I'll see that the truth of the first *If... then...* statement always predicts the truth of the second. So the fact that we didn't lose our jobs means Dana wasn't a spy. Would you expect me to find that argument convincing? 'Because the truth table says so?' Yes, it does say so, and the assignments of truth values to each cell of the table are certainly correct—but do those facts by themselves enable us to see exactly *why* the relationship between the two implications reflected in the truth table holds? Would you call this a satisfactory anwer to my question, 'How do you know?' Isn't this approach, rather, something along the lines of, 'It's that way because that's the way it is'?

The fact is that truth tables do not correspond to the reasoning steps that we typically rely on to take us from a true statement to another statement whose truth is guaranteed by those reasoning steps, in a way that convinces us that there could have been no other outcome, given our initial assumptions. Instead, truth tables present us with a simple fait accompli: these are the connectives, this is the formula, these are how the truth values fall out, virtually as the meaning of the connectives. If, as in the case of $p \supset q$ and $\neg q \supset p$, we have two formulæ, call them 1 and 2, where 2 is true in all cases where 1 is, then that shows that the truth of English sentences whose logical form represented by 1 guarantees the truth of an English sentence whose logical form is represented by 2. There's nothing to argue about; the truth table shows that this is the way things are. But no reasoning is involved here. What we have is just a fact. The scenario given above about Dana's spying and our jobs is an illustration of the problem.

What we really need, in order to be intuitively *convinced* that the truth of a certain statement entails the truth of some other statement, is a method of proceeding from the first statement to another whose truth is so self-evidently true that we cannot conceive any alternative in the world as we know it. Such a method must be independent of our factual knowledge, because it's always possible that we're wrong, or mistaken, about what the facts are. Rather, we need a way to proceed so that the logical *structure* of the original statement—determined, admittedly, by its connectives—can be shown to leave us no alternative to the truth of the statement which we're trying to prove on the basis of the first. Your patient efforts to make me see that, given your original premise, the fact that we hadn't been fired could only mean that Dana hadn't turned out to be a spy is an example of what we automatically do in such situations: we attempt to compute the truth of the conclusion based on a series of reasoning steps, each of them claimed to be necessary consequences of the preceding inference, rooted ultimately in our starting assumption.

But what if our original assumption is *wrong?* Well, that's too bad—but in a certain sense, it's not that important. The goal of correct reasoning, after all, can't really be omnicience. There is no *method* for knowing everything. The most we can expect is a way of proceding without error from one logical statement to another which is *necessarily* true if the the preceding statement is true. If we have that, then our reasoning method has done everything we have a right to ask of it. The burden is then on us, to get our facts straight. This distinction corresponds to the distinction that logicians often make between *validity*, on the one hand, and *soundness* on the other. A valid argument is one in which the rules for correct reasoning have been followed. A sound argument is one in which, in addition to being valid, the argument starts from premises which are factually correct. If I start from the premise that *if* Dana owns exactly one car and Terry owns exactly one car, then they jointly own three cars, and then add the information that Dana actually does own exactly one car and Terry owns exactly one car, and conclude that Dana and Terry jointly own three cars, my reasoning is impeccable, and obviously valid. But clearly it isn't at all sound. Soundness is mostly outside what the logician can be held responsible for—our understanding of the facts is always imperfect. But we can *reason* perfectly, based on what we take to be factual.

The bottom line, then, is that (i) we need a method of reasoning from some assumed statement to a conclusion which must be true if the initial assumption is; (ii) that we can't ask for empirical infallibility, but we do have a right to demand logical infallibility; and, finally, (iii) that the reasoning method match the results of the truth tables. This last requirement is non-negotiable. The truth tables reflect the facts of natural language logical vocabulary, standardized and purged of the kinds of pragmatic background assumptions discussed earlier. Inference methods are not in competition with truth tables; the idea is rather that such methods make clear to us *why* the truth tables for two different expressions which are in agreement are, in fact, in agreement. In the next file, the deductive system we will employ this quarter is introduced and justified—and you will see exactly how it is that we get to infer the truth of $\neg q \supset \neg p$ from the truth of $p \supset q$.