Handout (Week 3)

Truth and Validity

Isaac Wilhelm

Goals of this handout: give an account of truth in L, and use that account to define validity for arguments in L.

Reading guide for completing the homework

- To solve homework problems 1-5, you will probably have to read all of sections 1.1, 1.2, and 1.3.
- To solve homework problems 6-10, you will probably have to read all of sections 2.1 and 2.2.

Summary of handout

- Section 1.1: I present and explain the truth tables for the five connectives.
- Section 1.2: I present a summary of those truth tables, so that you can locate them quickly when doing the homework problems.
- Section 1.3: I work through—in full detail—an example of the sort of problem in the first half of the homework.
- Section 1.4: I discuss a confusion that students sometimes have, when they first learn about truth tables.
- Section 1.5: I present more examples of the sorts of problems in the first half of the homework.
- Section 2.1: I present five key definitions, including the definition of valid arguments in L.

- Section 2.2: I work through—in detail—some examples of the sorts of problems in the second half of the homework.
- Section 2.3: I explain how the definition of valid arguments in L is similar to the definition—presented in the week 1 handout—of valid arguments in English.
- Section 2.4: I explain the difference between validity and truth.

1 Truth

In this section, I give a complete account of truth in L. That is, I say exactly what it takes for any given sentence in L to be true. The basic idea is as follows. Sentence letters in L do not have fixed truth values: they can be true, or false, depending on what they represent. But once the truth values of the sentence letters are fixed, the truth values of *all other* sentences get fixed as well. In other words, the truth values of more complicated sentences—sentences, that is, which have been built out of sentence letters and logical connectives—are completely determined by the truth values of the simplest sentences: namely, sentence letters.

For convenience, we will use what are called 'truth tables' to study truth in L. For any given truth table, and for each way of assigning truth values to certain sentences, the table specifies what the truth value of a certain complicated sentence must be. So truth tables provide a clear, concise specification of *all* the ways in which the truth values of more complicated sentences depend on the truth values of simpler sentences.

1.1 Truth for the Five Connectives

In this section, I give an account of truth for sentences built using various symbols in L.

To start, consider sentence letters. Each sentence letter can be either true or false. The truth value of a sentence letter is not really determined by anything. In particular, the truth value of a sentence letter is not determined by the truth values of any simpler sentences which that sentence letter contains. The reason is simple: sentence letters do not contain simpler sentences at all. They are the simplest sentences of L. Because of this, there are no truth tables for sentence letters.

For example, take the sentence letter 'p'. This sentence letter can be either true or false, depending on what we take it to represent. For instance, if we use 'p' to represent the English sentence "It is the 21st century", then 'p' is true: for it is, indeed, the 21st century. If we use 'p' to represent the English sentence "It is the 20th century", then 'p' is false: for it is not the 20th century.

Now for negation. For any sentence 'ϕ', '$\neg\phi$' is true if and only if 'ϕ' is false, and '$\neg\phi$' is false if and only if 'ϕ' is true. Here is the truth table.

ϕ	$\neg\phi$
T	F
F	T

Read this truth table as follows. The left-most column lists all possible truth values of 'ϕ': it can be true or false. 'T' represents 'true', of course, and 'F' represents 'false'; 'T' and 'F' are called 'truth values'. For each truth value of 'ϕ', the right-most column says what the truth value of '$\neg\phi$' is. For example, the first row of truth values—that is, the second row in the truth table—says the following: if 'ϕ' is true, then '$\neg\phi$' is false. And the second row of truth values—that is, the third row in the truth table—says the following: if 'ϕ' is false, then '$\neg\phi$' is true.

So for example, suppose 'ϕ' is the sentence 'p'. And suppose that 'p' is true. Then by the second row in the truth table, the sentence '$\neg p$' is false. Or suppose 'ϕ'

is the sentence '$p \vee q$'. And suppose that this sentence is false; I will explain how disjunctions can be false below. Then by the third row in the truth table, '$\neg(p \vee q)$' is true.

The truth table for '$\neg$' is pretty intuitive. Roughly, it says that a sentence of the form "It is not the case that X" is true just in case 'X' is false. And that seems pretty plausible. For example, consider the English sentence "It is not the case that violets are neon green". That sentence seems true. Why? Because "Violets are neon green" is false. The above truth table basically says the same thing, just for sentence of L rather than sentences of English.

Now for conjunction. For any sentences 'ϕ' and 'ψ', '$\phi \wedge \psi$' is true if and only if both 'ϕ' is true and 'ψ' is true. Otherwise, '$\phi \wedge \psi$' is false. Here is the truth table.

ϕ	ψ	$\phi \wedge \psi$
T	T	T
T	F	F
F	T	F
F	F	F

Read this truth table as follows. Taken together, the two columns on the left list all possible *combinations* of truth value assignments to 'ϕ' and 'ψ'. There are four:

1. 'ϕ' could be true and 'ψ' could be true,
2. 'ϕ' could be true and 'ψ' could be false,
3. 'ϕ' could be false and 'ψ' could be true, and
4. 'ϕ' could be false and 'ψ' could be false.

For each of those combinations, the right-most column of the truth table says what the truth value of '$\phi \wedge \psi$' is. For example, the first row of truth values says the following: if 'ϕ' is true and 'ψ' is true, then '$\phi \wedge \psi$' is true. The second row of truth values says the following: if 'ϕ' is true and 'ψ' is false, then '$\phi \wedge \psi$' is false. The third row of truth values says the following: if 'ϕ' is false and 'ψ' is true, then '$\phi \wedge \psi$' is false. And the fourth row of truth values says the following: if 'ϕ' is false and 'ψ' is false, then '$\phi \wedge \psi$' is false.

So for example, suppose 'ϕ' is the sentence 'p' and 'ψ' is the sentence 'q'. Furthermore, suppose that 'p' is true and 'q' is false. Then by the second row of truth values, the sentence '$p \wedge q$' is false. Or suppose 'ϕ' is the sentence '$p \vee q$', and 'ψ' is the sentence 'r'. Suppose that '$p \vee q$' and 'r' are both true. Then by the first row of truth values, '$(p \vee q) \wedge r$' is true.

Truth tables are helpful and convenient, because they compress all that information—all the information in the paragraph above—into a very small space. And note what the truth table is doing: it says how the truth value of the more complicated sentence '$\phi \wedge \psi$' is built out of the truth values of (i) the less complicated sentence 'ϕ', and (ii) the less complicated sentence 'ψ'.

The truth table for '$\wedge$' is pretty intuitive. Roughly, it says that a sentence of the form "X and Y" is true just in case 'X' is true and 'Y' is true. If either 'X' or 'Y' is false, then 'X and Y' is false. And that all seems pretty plausible. For example, consider the English sentence "Roses are red and violets are blue". That sentence seems true. Why? Because "Roses are red" is true, and "Violets are blue" is true. As another example, consider the English sentence "Roses are red and violets are neon green". That sentence seems false. Why? Because "Violets are neon green" is false. The above truth table basically says the same thing, just for sentence of L rather than sentences of English.

Now for disjunction. For any sentences 'ϕ' and 'ψ', '$\phi \vee \psi$' is true if and only if either 'ϕ' is true or 'ψ' is true. Otherwise, '$\phi \vee \psi$' is false. Here is the truth table.

ϕ	ψ	$\phi \vee \psi$
T	T	T
T	F	T
F	T	T
F	F	F

As before, read this truth table like so: the two columns on the left list all possible combinations of truth value assignments to 'ϕ' and 'ψ', while the column on the right lists the corresponding truth values of the more complicated sentence '$\phi \vee \psi$'.

So for example, suppose 'ϕ' is the sentence 'p' and 'ψ' is the sentence 'q'. Furthermore, suppose that 'p' is false and 'q' is true. Then by the third row of truth values, the sentence '$p \vee q$' is true. Or suppose 'ϕ' is the sentence '$p \to q$', and 'ψ' is the sentence 'r'. Suppose that '$p \to q$' and 'r' are both false. Then by the fourth row of truth values, '$(p \to q) \vee r$' is false.

The truth table for '$\vee$' is pretty intuitive. Roughly, it says that sentences of the form "X or Y" are true just in case either 'X' is true or 'Y' is true. If both 'X' and 'Y' are false, then 'X and Y' is false. And that all seems pretty plausible. For example, consider the English sentence "Roses are red or the moon is made of cheese". That sentence seems true. Why? Because "Roses are red" is true. Of course, "The moon is made of cheese" is false. But the sentence "Roses are red or the moon is made of cheese" is still true, because one of its disjuncts—namely, the disjunct "Roses are red"—is true.

Now for the conditional. For any sentences 'ϕ' and 'ψ', '$\phi \to \psi$' is true if and only if either 'ϕ' is false or 'ψ' is true. Otherwise, '$\phi \to \psi$' is false. Here is the truth table.

ϕ	ψ	$\phi \to \psi$
T	T	T
T	F	F
F	T	T
F	F	T

As before, read this truth table like so: the two columns on the left list all possible combinations of truth value assignments to 'ϕ' and 'ψ', while the column on the right lists the corresponding truth values of the more complicated sentence '$\phi \to \psi$'.

So for example, suppose 'ϕ' is the sentence 'p' and 'ψ' is the sentence 'q'. Furthermore, suppose that 'p' is true and 'q' is true. Then by the first row of truth values, the sentence '$p \to q$' is true. Or suppose 'ϕ' is the sentence '$\neg q$', and 'ψ' is the sentence '$r \wedge p$'. Suppose that '$\neg q$' is true while '$r \wedge p$' is false. Then by the second row of truth values, '$\neg q \to (r \wedge p)$' is false.

Of all the truth tables for all the logical connectives of L, this one is probably the weirdest. There are several reasons why: here, I will discuss only one. The third row of truth values says that in order for a conditional to be true, it suffices for the antecedent of that conditional to be false. But that is not really how the English construction corresponding to the symbol '$\to$'—the construction "If ..., then ..."—seems to work. To see why, consider the sentence "If the moon is made of cheese, then roses are red". It is not clear whether that sentence is true or false: some people think it is intuitively false, some think it is intuitively true, and some think that intuitively, it does not have a truth value. But according to the truth table above, the corresponding sentence of L—the sentence '$p \to q$', where 'p' represents "The moon is made of cheese" and 'q' represents "Roses are red"—is true, since 'p' is false and 'q' is true. So the symbol '$\to$', in L, is quite different from the corresponding English expression "If ..., then ...".

Many philosophers, linguists, and psychologists study this. In particular, many philosophers, linguistics, and psychologists explore different ways of setting up formal languages, so that (i) the resulting formal language has lots of the nice properties that L has, but (ii) the resulting formal language has a symbol which does a better job than '$\rightarrow$' of representing the English construction "If ..., then ...". For lack of space, I will not investigate that here. But it gets really, really interesting.

Now for the biconditional. For any sentences 'ϕ' and 'ψ', '$\phi \leftrightarrow \psi$' is true if and only if 'ϕ' and 'ψ' have the same truth value. In particular, here is the truth table for '$\leftrightarrow$'.

ϕ	ψ	$\phi \leftrightarrow \psi$
T	T	T
T	F	F
F	T	F
F	F	T

As before, read this truth table like so: the two columns on the left list all possible combinations of truth value assignments to 'ϕ' and 'ψ', while the column on the right lists the corresponding truth values of the more complicated sentence '$\phi \leftrightarrow \psi$'.

So for example, suppose 'ϕ' is the sentence 'p' and 'ψ' is the sentence 'q'. Furthermore, suppose that 'p' is false and 'q' is false. Then by the fourth row of truth values, the sentence '$p \leftrightarrow q$' is true. Or suppose 'ϕ' is the sentence '$\neg q$', and 'ψ' is the sentence '$r \wedge p$'. Suppose that '$\neg q$' is true while '$r \wedge p$' is false. Then by the second row of truth values, '$\neg q \leftrightarrow (r \wedge p)$' is false.

The truth table for '$\leftrightarrow$' is perhaps more intuitive than the truth table for '$\rightarrow$', but it is still less intuitive than the other truth tables. Roughly, it says that a sentence of the form “X if and only if Y” is true just in case either (i) ‘X’ is true and ‘Y’ is true, or (ii) ‘X’ is false and ‘Y’ is false. So “X if and only if Y” is false whenever the truth values of ‘X and Y’ differ. And that all seems at least somewhat plausible. For example, consider the English sentence “Roses are red if and only if the moon is made of cheese”. That sentence seems false. Why? Because “Roses are red” is true, but “The moon is made of cheese” is false. Or consider the sentence “Roses are red if and only if violets are blue”. Since both sides of the ‘. . . if and only if . . . ’ construction are true, that sentence seems—at least to some extent—true as well.

1.2 Summary: Truth Tables for Five Connectives

Here is a succinct summary of the truth tables for all five connectives. In what follows, ‘ϕ’ and ‘ψ’ are sentences.

ϕ	$\neg\phi$
T	F
F	T

ϕ	ψ	$\phi \wedge \psi$
T	T	T
T	F	F
F	T	F
F	F	F

ϕ	ψ	$\phi \vee \psi$
T	T	T
T	F	T
F	T	T
F	F	F

ϕ	ψ	$\phi \to \psi$
T	T	T
T	F	F
F	T	T
F	F	T

ϕ	ψ	$\phi \leftrightarrow \psi$
T	T	T
T	F	F
F	T	F
F	F	T

1.3 Truth for Sentences of L: A Detailed Example

Now for the first big upshot of the above theory of truth: using the truth tables for the five connectives, one can construct truth tables for *any sentence of L whatsoever*. In this subsection, I explain how. I focus on a single example: '$p \vee \neg q$'. In later subsections, I work through a few more examples.

Consider the sentence '$p \vee \neg q$'. When is it true, and when is it false? Or a little more precisely: under what assignments of truth values to the sentence letters in '$p \vee \neg q$' does that sentence come out true, and under what assignments of truth values to the sentence letters in '$p \vee \neg q$' does that sentence come out false? The truth table for '$p \vee \neg q$' provides a complete answer. Let us see how to construct it.

The construction proceeds in roughly two parts. The first part is pretty simple. The second is quite complex.

The first part is as follows. To start, determine which sentence letters are in '$p \vee \neg q$'. As you can see, there are two: the sentence letter 'p', and the sentence letter 'q'.

Now put those sentence letters into a truth table, like so.

p	q	

Then list out all the combinations of truth value assignments to those sentence letters. Note that for each sentence letter, there are two possible assignments of truth values to that sentence letter: it could be assigned 'T', and it could be assigned 'F'. So for two sentence letters, there are four different truth value combinations: they could both be true, the first could be true while the second is false, the second could be false while the first is true, and they could both be false. So there are four combinations of assignments—or for short, there are four 'assignments'—of truth values to 'p' and 'q' taken together. List them in the truth table, like so.

p	q	
T	T	
T	F	
F	T	
F	F	

Now take the sentence in question—that is, '$p \vee \neg q$'—and put it in the upper right corner, like so.

p	q	p	$\vee$	$\neg$	q
T	T				
T	F				
F	T				
F	F				

This completes the first part of the construction. It will soon become clear why I created so much space in the right-most column.

The second part is as follows. Look at *each* assignment of truth values to sentence letters. That is, look at *each* row in the above truth table. For each such assignment, use it to determine the truth value of the sentence on the right.

Let us begin with the first assignment: this is the second row in the above table. According to that assignment, 'p' is true and 'q' is true. What is the truth value of '$p \vee \neg q$', given that assignment? To answer this question, let us break it up into pieces.

For starters, what is the truth value of '$\neg q$', given that assignment? The truth table for negation provides the answer: since 'q' is true on this assignment, the second row of the truth table for negation implies that '$\neg q$' is false. And so what is the truth value of '$p \vee \neg q$'? The third row of the truth table for disjunction provides the answer: since 'p' is true on this assignment, and '$\neg q$' is false on this assignment, the third row of the truth table for disjunction implies that '$p \vee \neg q$' is true.

All that reasoning, in the paragraph above, can be carried out using the truth table. Let us see how. And remember that right now, we are only looking at the first assignment of truth values; that is, we are only looking at the second row in the truth table.

To start, under each sentence letter in the right-most sentence, list that sentence letter's truth value *for that first assignment.*

p	q	p	$\vee$	$\neg$	q
T	T	T			T
T	F				
F	T				
F	F				

This amounts, basically, to just taking the truth values of the sentence letters on the left—for the first assignment—and copying them under the sentence letters on the right.

Now ask a somewhat vague but very important question: given that 'p' is true and 'q' is true—on this first assignment—what truth values of what sentences can be immediately determined? Note that to determine the truth value of '$p \vee \neg q$', we need to know the truth value of 'p' and *also* the truth value of '$\neg q$'. On this first assignment, the truth value of 'p' is clear: it is 'T'; I just typed that in the right-most column of the table above. We do not yet know the truth value of '$\neg q$', however. So before determining the truth value of '$p \vee \neg q$', we first need to determine the truth value of '$\neg q$'.

And what is that truth value? Recall that on this first assignment, 'p' is true and 'q' is true. We can use this, it turns out, to determine the truth value of '$\neg q$'. For since 'q' is true, the second row of the truth table for negation implies that '$\neg q$' is false. So put an 'F' under the '$\neg$' in the second row of the above truth table, like so.

p	q	p	$\vee$	$\neg$	q
T	T	T		F	T
T	F				
F	T				
F	F				

The 'F' under the '$\neg$' represents the fact that on the first assignment, '$\neg q$' comes out false.

Now ask: given that 'p' is true and '$\neg q$' is false—on this first assignment—what is the truth value of '$p \vee \neg q$'? By the third row of the truth table for disjunction, '$p \vee \neg q$' true. So put a 'T' under the '$\vee$' in the second row of the above truth table, like so.

p	q	p	$\vee$	$\neg$	q
T	T	T	**T**	F	T
T	F				
F	T				
F	F				

Note that the 'T' is bolded. This makes clear to everyone involved—namely, you and I—exactly which truth value in the right-most column is **the** truth value for the sentence at the top of that column.

We are still not done with the second part of the construction. But before continuing, here is a summary of what we have accomplished. We asked the following question: what is the truth value of '$p \vee \neg q$' on the first assignment of truth values to the sentence letters in that sentence? We answered this question in step-by-step fashion. We used that first assignment to determine the truth value of '$\neg q$' (for that assignment), and then we used *that* to determine the truth value of '$p \vee \neg q$' (for that assignment).

To complete the truth table—and thus, the second part of the construction—we repeat the above reasoning for *all* the truth value assignments. So we do it for the second assignment – which takes 'p' to be true and 'q' to be false – the third assignment – which takes 'p' to be false and 'q' to be true – and the fourth assignment – which takes 'p' to be false and 'q' to be false. For each assignment, we first determine the truth value of '$\neg q$' (for that assignment), and then we use that to determine the truth value of '$p \vee \neg q$' (for that assignment). As a result, we get the following truth table.

p	q	p	$\vee$	$\neg$	q
T	T	T	**T**	F	T
T	F	T	**T**	T	F
F	T	F	**F**	F	T
F	F	F	**T**	T	F

This truth table provides a *complete* answer to the question with which we began. Recall the question: when is '$p \vee \neg q$' true, and when is '$p \vee \neg q$' false? Or a little more precisely: under what assignments of truth values to the sentence letters in '$p \vee \neg q$' does that sentence come out true, and under what assignments of truth values to the sentence letters in '$p \vee \neg q$' does that sentence come out false? The above truth table provides the answer. The sentence '$p \vee \neg q$' is true on the first, second, and fourth assignments. On all other assignments—namely, the third one—'$p \vee \neg q$' is false.

1.4 Aside: A Common Confusion

In this subsection, I discuss a common confusion that people sometimes have, when they first learn about truth tables. Note that to determine the truth value of '$p \vee \neg q$', we first had to determine the truth values of the simpler sentences—namely, 'p', 'q', and '$\neg q$'—that it contains. This might have been confusing. People sometimes struggle to figure out which simpler sentences' truth values need to be determined first, before determining the truth value of a given larger sentence X. There is always, in fact, a particular order in which you must proceed: first you must determine the truth values of the simplest sentences, then you must determine the truth values of the 'next simplest' sentences, and so on, until finally you can determine X's truth value. But people sometimes struggle to see what the simplest sentences are, what sentences count as 'next simplest', and so on.

If you ever get confused about that, ask the following question: how would I build up X, using simpler sentences? More precisely, how would I build up X, using the Recursive Definition of a sentence of L from the handout from week 2? The answer to that question tells you more than just how X is built from simpler sentences. The answer *also* tells you the order in which you must proceed, to determine the truth value of X. For the order in which you must proceed, to determine the truth value of a sentence, is *exactly the same* as the order in which you must proceed, to build that sentence using the Recursive Definition.

For example, consider the sentence '$p \vee \neg q$'. Here is how you would build that sentence, using the Recursive Definition from the week 2 handout:

1. 'p' is a sentence, since it is a sentence letter;
2. 'q' is a sentence, since it is a sentence letter;
3. '$\neg q$' is a sentence, since 'q' is a sentence;
4. '$p \vee \neg q$' is a sentence, since both 'p' and '$\neg q$' are sentences.

And here is the order in which you must proceed, to determine the truth value of '$p \vee \neg q$' for the first truth value assignment:

1. 'p' is true, since it is true on that first assignment;
2. 'q' is true, since it is true on that first assignment;
3. '$\neg q$' is false—by the second row of the truth table for negation—since 'q' is true;
4. '$p \vee \neg q$' is true—by the third row of the truth table for disjunction—since 'p' is true and '$\neg q$' is false.

As you can see, the order in which we build the sentence '$p \vee \neg q$', using the Recursive Definition, is exactly the same as the order in which we determine the truth values of the simplest sentences, the 'next simplest' sentences, and so on, which '$p \vee \neg q$' contains. So if you ever get confused about the order in which you must proceed, to determine the truth value of some complicated sentence X, just figure out the order in which you must proceed, to build X from simpler sentences. The latter is a guide to the former.

All this, in fact, is why we went through such pains to rigorously define what the sentences of L are. That rigorous definition acts as a guide for determining the truth value of any sentence of L whatsoever. In other words, the Recursive Definition contains the materials needed to give a complete account of truth in L.[1]

1.5 More Examples of Truth Tables

In this subsection, I present two more examples of truth tables for sentences of L. The first is a truth table for the sentence '$p \rightarrow (p \vee \neg p)$'. The second is a truth table for the sentence '$p \leftrightarrow (r \wedge (\neg p \vee q))$'.

To start, consider '$p \rightarrow (p \vee \neg p)$'. Under what assignments of truth values to its sentence letters does this sentence come out true, and under what assignments of truth values to its sentence letters does this sentence come out false? We will answer this question by constructing a truth table. Note that in what follows, I skip several steps; if I wrote out every step, then this handout would become ridiculously long.

[1] And in this respect, L is a much simpler language than, say, English. No one knows an analogously rigorous definition for what the English sentences are, because compared to L, English is extremely complicated. Consequently, no one has been able to propose a rigorous account of what makes any given English sentence true. That is one of the foundational open projects in linguistics, computer science, psychology, and philosophy.

Recall that the construction proceeds in roughly two parts. In the first, we (i) identify all the sentence letters in '$p \to (p \vee \neg p)$', (ii) list those letters in columns of a table, (iii) list all possible assignments of truth values to those letters in those columns, and then (iv) put the sentence '$p \to (p \vee \neg p)$' in a column of its own. The result is below.

p	p	$\to$	$(p$	$\vee$	$\neg$	$p)$
T						
F						

Note that unlike the example from before, there are just two truth value assignments to the sentence letters in '$p \to (p \vee \neg p)$'. The reason is straightforward: there is only one sentence letter in '$p \to (p \vee \neg p)$'—namely, the sentence letter 'p'—so there are only two ways of assigning truth values to the sentence letters which '$p \to (p \vee \neg p)$' contains.

Now for the second part: we look at *each* assignment, and determine the truth value of '$p \to (p \vee \neg p)$' under it. As before, we start with the first assignment. And as before, we work in step-by-step fashion. In particular, we (i) write down the truth value of each sentence letter in '$p \to (p \vee \neg p)$' under the instances of that letter in the right-most column, (ii) determine the truth values of the sentences which '$p \to (p \vee \neg p)$' contains—in particular, '$\neg p$', and then '$p \vee \neg p$'—and then (iii) determine the truth value '$p \to (p \vee \neg p)$' itself. The result is below.

p	p	$\to$	$(p$	$\vee$	$\neg$	$p)$
T	T	**T**	T	T	F	T
F						

Here is how I completed the second row in the above truth table. For starters, I determined the truth value of '$\neg p$'. To do so, I observed that since 'p' is true, the second row of the truth table for negation implies that '$\neg p$' is false. So I typed an 'F' under the '$\neg$' in the table above. Then I observed that since 'p' is true and '$\neg p$' is false, the third row of the truth table for disjunction implies that '$p \vee \neg p$' is true. So a typed a 'T' under the '$\vee$' in the table above. Finally, I observed that since 'p' is true and '$(p \vee \neg p)$' is true, the second row of the truth table for the conditional implies that '$p \rightarrow (p \vee \neg p)$' is true. So I typed a bolded 'T' under the '$\rightarrow$' in the table above. And then I was done: for that bolded 'T' represents the fact that on the first assignment of truth values, '$p \rightarrow (p \vee \neg p)$' comes out true.

Finally, we repeat for each other assignment. In this case, there is just one.

p	p	$\rightarrow$	$(p$	$\vee$	$\neg$	$p)$
T	T	**T**	T	T	F	T
F	F	**T**	F	T	T	F

An optional exercise for the reader: check that I correctly filled out the third row of the above truth table.

This truth table provides a complete answer to the question with which we began. Recall the question: under what assignments of truth values to the sentence letters in '$p \rightarrow (p \vee \neg p)$' does that sentence come out true, and under what assignments of truth values to the sentence letters in '$p \rightarrow (p \vee \neg p)$' does that sentence come out false? The above truth table provides the answer. The sentence '$p \rightarrow (p \vee \neg p)$' is true on *each* assignment. It is, in other words, always true.

It is worth doing one more example. This one involves a particularly complicated sentence: '$p \leftrightarrow (r \wedge (\neg p \vee q))$'. Here is the truth table before the truth value of '$p \leftrightarrow (r \wedge (\neg p \vee q))$' has been determined for each assignment.

p	q	r	p	$\leftrightarrow$	$(r$	$\wedge$	$(\neg$	p	$\vee$	$q))$
T	T	T								
T	T	F								
T	F	T								
T	F	F								
F	T	T								
F	T	F								
F	F	T								
F	F	F								

Note that there are eight truth value assignments, because there are eight different combinations of truth value assignments to the three sentence letters ‘p’, ‘q’, and ‘r’ which show up in ‘$p \leftrightarrow (r \wedge (\neg p \vee q))$’.[2]

And here is the completed truth table for ‘$p \leftrightarrow (r \wedge (\neg p \vee q))$’.

p	q	r	p	$\leftrightarrow$	$(r$	$\wedge$	$(\neg$	p	$\vee$	$q))$
T	T	T	T	**T**	T	T	F	T	T	T
T	T	F	T	**F**	F	F	F	T	T	T
T	F	T	T	**F**	T	F	F	T	F	F
T	F	F	T	**F**	F	F	F	T	F	F
F	T	T	F	**F**	T	T	T	F	T	T
F	T	F	F	**T**	F	F	T	F	T	T
F	F	T	F	**F**	T	T	T	F	T	F
F	F	F	F	**T**	F	F	T	F	T	F

So the sentence ‘$p \leftrightarrow (r \wedge (\neg p \vee q))$’ is true on the first, sixth, and eighth assignment; for in the second, seventh, and ninth rows of the above truth table, the

[2] An optional exercise for the reader: prove that if there are n sentence letters in a given sentence, then there are 2^n truth value assignments for that sentence.

bolded letter is 'T'. On all other assignments, '$p \leftrightarrow (r \wedge (\neg p \vee q))$' is false.

Truth tables are among the most important things which you will learn in this course. For they play a crucial role in the definition of validity for arguments in L; I discuss this in the next section. And because they can be used to define validity in L, truth tables also play a crucial role in the definition of validity for arguments in English; I discuss this in the next handout. So make sure that you are very comfortable with truth tables.

2 Validity

The account of truth in L can be used to give an account of valid arguments in L. In this section, I explain how.

2.1 The Key Definitions

In this section, I define five key notions: arguments, conclusions of arguments, premises of arguments, valid arguments, and invalid arguments. All of the definitions concern arguments *in* L. So they are akin to—but ultimately different form—the definitions of analogous notions for arguments in English (see the handout for week 1).

> **Definition: Argument (in L)**
> An 'argument' in L is a sequence of two or more sentences of L.

> **Definition: Conclusion**
> The 'conclusion' of an argument in L is the last sentence in that argument.

Definition: Premise

A 'premise' of an argument in L is any sentence in that argument which is not the conclusion.

Definition: Valid Argument (in L)

A 'valid argument' in L is an argument in L which has the following property: for *each* assignment of truth values to the sentence letters in that argument which makes *all* of the premises true, that assignment *also* makes the conclusion true.

To determine whether or not an argument in L is valid, do the following.

- First, list out all the sentence letters which appear in sentences of that argument.
- Second, list out all combinations of assignments—or for short, list out all 'assignments'—of truth values to those sentence letters.
- Third, for each and every such assignment, check and see if that assignment makes *all* of the premises in the argument true.
- Fourth, collect together all of the assignments which do, indeed, make *all* of the premises in the argument true.
- Fifth, for each of *those* assignments—for each of the assignments collected together, that is, in the fourth step—check and see if that assignment makes the conclusion of the argument true.
- If *every single one* of those assignments makes the conclusion true, then the argument is valid.
- If one of those assignments makes the conclusion *false*, however, then the argument is not valid.

Here is one more key definition.

Definition: Invalid Argument (in L)

An 'invalid argument' in L is an argument in L which is not valid.

In other words, an invalid argument of L is an argument of L with the following property: there is an assignment of truth values to the sentence letters in that argument which (i) makes all of the premises true, but (ii) makes the conclusion false.

To determine whether or not an argument of L is valid, do exactly what you do to determine if the argument is valid. If you find that the argument is not valid, then the argument is invalid.

2.2 Examples

In this section, I give some examples of arguments, non-arguments, valid arguments, and invalid arguments of L.

Here are some examples of arguments and non-arguments.

Argument

1. p
2. $p \to q$
3. q

This is an argument because it is a sequence of three—and so, more than two—sentences of L. Note that the third sentence – namely, 'q' – is the conclusion of the argument, and the first two sentences – namely, 'p' and '$p \to q$' – are the premises of the argument.

Non-Argument

1. p

2. $p \rightarrow q\neg$
3. q

This is not an argument because the second line—namely, '$p \rightarrow q\neg$'—is not a sentence of L.

Now for examples of valid and invalid arguments. Truth tables can be used to check whether a given argument is valid. The procedure, in rough outline, is as follows. First, construct a truth table. Second, use that table to answer the following question: is it the case that for *each* assignment of truth values to sentence letters in the table which makes *all* of the premises in the argument true, that assignment *also* makes the conclusion true? If so, then the argument is valid. If not, then not.

The truth table construction is pretty similar to the constructions discussed in Section 1.3. The main difference is that in truth tables which check for validity, there are *multiple* columns on the right. Each of these columns displays the truth values, under each assignment, of a sentence in the argument.

Let us focus on the following argument as an example.

1. p
2. $p \rightarrow q$
3. q

As before, the construction of a truth table proceeds in roughly two parts. The first part is as follows. To start, identify all the sentence letters which show up in some line or other of this argument. List those letters in columns of a table. Then list all combinations of assignments of truth values to the letters in those columns. Finally, for *each* sentence in the argument, put that sentence in a column of its own. The result is below.

p	q	p	p	$\rightarrow$	q	q
T	T					
T	F					
F	T					
F	F					

Note that on the right side of this table, there is a separate column for each line in the argument: the third column is for the first line in the argument, the fourth column is for the second line in the argument, and the right-most column is for the third line in the argument.

Now for the second part: complete the truth table. To do so, focus on one column at a time. For instance, let us start with the third column. Fill out that column in the way described in Section 1.3. The result is below.

p	q	p	p	$\rightarrow$	q	q
T	T	T				
T	F	T				
F	T	F				
F	F	F				

Now fill out the fourth column, like so.

p	q	p	p	$\rightarrow$	q	q
T	T	T	T	**T**	T	
T	F	T	T	**F**	F	
F	T	F	F	**T**	T	
F	F	F	F	**T**	F	

Finally, fill out the last column, like so.

p	q	p	p	$\to$	q	q
T	T	T	T	**T**	T	T
T	F	T	T	**F**	F	F
F	T	F	F	**T**	T	T
F	F	F	F	**T**	F	F

This is the completed truth table for the argument with which we began.

To check for validity, ask the following question: is it the case that for *each* assignment of truth values to sentence letters in the table which makes *all* of the premises in the argument true, that assignment *also* makes the conclusion true? The answer: yes. To see why, note that there is just one assignment of truth values to sentence letters in the table which makes all of the premises in the argument true: the first assignment; that is, the one in the second row. Each other assignment makes at least one premise in the argument false: the second assignment makes the second premise '$p \to q$' false, for instance, and the third and fourth assignments both make the first premise 'p' false. Now ask the question: does that first assignment—the only assignment which makes all of the premises true—also make the conclusion true? The answer: yes. For there is a 'T' in the second row of the right-most column: that represents the fact that on the first assignment, the conclusion of the argument—namely, 'q'—is true. So the argument has the following property: for *each* assignment of truth values to the sentence letters of that argument which makes *all* of the premises true, that assignment *also* makes the conclusion true. And so by the definition of validity, the argument is valid.

Here is an example of an invalid argument.

1. $\neg p$
2. $p \to q$
3. $\neg q$

The truth table for this argument is below.

p	q	$\neg$	p	p	$\rightarrow$	q	$\neg$	q
T	T	**F**	T	T	**T**	T	**F**	T
T	F	**F**	T	T	**F**	F	**T**	F
F	T	**T**	F	F	**T**	T	**F**	T
F	F	**T**	F	F	**T**	F	**T**	F

Once again, to check for validity, ask the question: is it the case that for *each* assignment of truth values to sentence letters in the table which makes *all* of the premises in the argument true, that assignment *also* makes the conclusion true? The answer: no. To see why, note that there are two assignments of truth values to sentence letters in the table which make all of the premises in the argument true: the third assignment – that is, the one in the fourth row – and the fourth assignment – that is, the one in the fifth row. Each other assignment makes at least one premise in the argument false. So consider *only* those assignments which make all of the premises true, and ask the question: do those assignments—that is, both the third assignment and the fourth assignment—*also* make the conclusion true? The answer: no. The third assignment makes the conclusion false, not true: for there is a bolded 'F' in the fourth row of the right-most column. So the argument is invalid.

2.3 Why We Care

Here is the idea behind the formal definition of validity. Recall that according to the intuitive notion of validity discussed in the first handout of this course—the first-pass definition of validity—a valid argument is an argument with the following property: if the premises are all true, then the conclusion must also be true. In other words, according to the intuitive notion of validity, a valid argument is an argument in which the truth of the premises guarantees the truth of the conclusion.

The definition of validity in L does a pretty good job of capturing the intuitive notion of validity. To see why, note that there are many different ways in which all the premises of an argument in L might turn out true. The definition of validity in L provides the resources for figuring out what all those ways are: (i) determine which sentence letters appear in the argument in question, (ii) list out all combinations of assignments of truth values to those sentence letters, and then (iii) for each and every such assignment, check and see whether all of the premises of the argument are true. So the account of truth in L can be used to determine exactly when the premises of an argument are all true, and exactly when some of the premises of an argument are false. And so to determine if an argument of L is valid, just take all those ways for the premises to be true—all those assignments on which the premises are true—and see if for every single one of them, the conclusion is true as well. In other words, to determine if an argument of L is valid, just take all the truth value assignments which make the premises true, and see if they also make the conclusion true. If so, then the argument is valid. If not, then not.

That, according to the definition of validity in L, is what it means to say that the truth of the premises <u>guarantees</u> the truth of the conclusion. The formal theory of truth value assignments, truth tables, and so on, provides L with the resources to make that informal claim—about the premises' truth guaranteeing the conclusion's truth—perfectly precise. In other words, to say

> The truth of the premises guarantees the truth of the conclusion

is, in L, just to say

> Every assignment of truth values to sentence letters which makes all of the premises true *also* makes the conclusion true.

So the definition of validity in L does, indeed, do a pretty good job of capturing the intuitive notion of validity with which this course began. And that is why philosophers care about L in the first place: L can be used to precisify the intuitive, squishy, but extremely important notion of a valid argument, the notion characterized by the first-pass definition of validity on the handout from week 1.

2.4 Distinguishing Validity and Truth

Before closing, it is worth clarifying the distinction between truth and validity. Truth is a property of *sentences*. Validity, in contrast, is a property of *arguments*. So sentences cannot be valid or invalid, and similarly, arguments cannot be true or false.

Of course, sentences in arguments can be true or false. An argument itself is neither true nor false, however; it is either valid or invalid. But the premises of an argument can be true or false, and likewise for the conclusion.

People often confuse the notion of truth with the notion of validity. They say that arguments are 'true', or that sentences are 'valid'. But to be clear: that is not how truth and validity work.

The definition of validity *relies upon* the definition of truth. Recall the definition from Section 2.1: an argument is valid just in case each truth value assignment which makes all of its premises true also makes its conclusion true. In order for this definition to make any sense, we must already have an account of truth. We must already have an account of the conditions under which premises and conclusions—that is, sentences—are true. In a phrase: validity presupposes truth.

So the distinction between truth and validity is important. One notion is used to define the other: in particular, truth is used to define validity. And so these notions are not interchangeable. They are importantly distinct.