

## The Soundness Theorem for Sentential Logic

These notes contain a proof of the Soundness Theorem for Sentential Logic. The Soundness Theorem says that our natural deduction proofs represent a sound (or correct) system of reasoning. Formally, the Soundness Theorem states that:

If  $\Gamma \vdash \phi$  then  $\Gamma \vDash \phi$ . This says that for any set of premises, if you can prove  $\phi$  from that set, then  $\phi$  is truth-functionally entailed by that set. In other words, if our rules let you prove that something follows from some premises, then it ‘really does’ follow from those premises.

By definition  $\Gamma \vDash \phi$  means that there is no TVA that makes all of the members of  $\Gamma$  true and also makes  $\phi$  false. Note that an equivalent definition would be “ $\Gamma \vDash \phi$  if any TVA that makes all the members of  $\Gamma$  true also makes  $\phi$  true”. By contraposition, the soundness theorem is equivalent to “If it is not true that  $\Gamma \vDash \phi$  then it is not true that  $\Gamma \vdash \phi$ .” Unpacking the definitions, this means “If there is a TVA that makes all the members of  $\Gamma$  true but makes  $\phi$  false, then there is no way to prove  $\phi$  from  $\Gamma$ .” We might call this “The Negative Criterion.” This is important since the standard way of showing that something is not provable is to give an invalidating TVA. This only works assuming that the soundness theorem is true.

Let’s get started on the proof. We are trying to prove a conditional claim, so we assume its antecedent and try to prove its consequent.

**Assumption 1:**  $\Gamma \vdash \phi$  – In other words, there is a proof of  $\phi$  from  $\Gamma$ .

**GOAL:**  $\Gamma \vDash \phi$  – in other words, there is no TVA that makes all of the members of  $\Gamma$  true and makes  $\phi$  false.

Keeping our goal in mind, let’s examine the last line of this proof. We know that the sentence  $\phi$  is on that line and it depends on assumptions only found in  $\Gamma$ . So I will prove of this line, that there is no TVA that makes all of the assumptions written to the left of this line true and makes our conclusion  $\phi$  false. If I do this, I can easily show that my goal is true. Since all of the assumptions written to the left of the last line are contained in the set  $\Gamma$ , if I can’t make all of them true and  $\phi$  false, I can’t make all of  $\Gamma$  true and  $\phi$  false.

Let’s call the assumptions that our last line depends on  $P_1, P_2, \dots, P_n$ . I will call our conclusion  $C$ . (This is sentence  $\phi$ )

**New Goal:** There is no TVA that makes all of  $P_1, P_2, \dots, P_n$  true and which makes  $C$  false.

In order to prove this, I will prove something stronger. I will prove that this connection between assumptions and what is on the line holds of every line of every proof. This would suffice for showing that it holds of the last line.

**New Goal:** For every line of every proof, there is no TVA that makes all of the assumptions written to the left of the line true and which makes the sentence written on the line false.

\*\* - NOTE. I showed in class how to turn a valid sequent into a theorem. One such way is to form the conditional whose antecedent is the conjunction of all the premises and whose consequent is the conclusion. This same reasoning can be used to turn a single line of a proof into a theorem. For example, take the last line of the proof. If you can prove  $C$  from  $P_1, P_2, \dots, P_n$  then  $(P_1 \& P_2 \& \dots \& P_n) \rightarrow C$  is a theorem. If there is no TVA that makes  $P_1, P_2, \dots, P_n$  true and which makes  $C$  false then the sentence  $(P_1 \& P_2 \& \dots \& P_n) \rightarrow C$  is a tautology. The Soundness Theorem is thus equivalent to proving that every theorem is a tautology.

Back to the proof...

In order to prove our new goal, I will use reductio reasoning and assume that it is false. Then I will derive a contradiction from this assumption. This would prove that that our new goal really is true. In other words, I am assuming:

**Assumption 2:** There is at least one line (maybe many) such that there is a TVA that makes all of the assumptions written to the left of that line true and which makes what is written on the line false.

To derive a contradiction, I will begin by applying the Least Number Principle (every non-empty set of positive integers has a least element). Let us call a line where there is a TVA that makes all the assumptions written to the left of the line true and the sentence written on the line false a BAD line. Our assumption (2) says that there exists at least one BAD line in some proof. Hence, the set of BAD lines in such a proof is non-empty. So, by the Least Number Principle:

**FACT 3:** There is a lowest numbered BAD line. (I call this a fact rather than an assumption because it comes from using assumption two but is not itself a new assumption.)

I will now show that (3) reduces to absurdity. This lowest numbered BAD line must have a line number. Let's call it line  $k$ . It must look like this:

$a_1, a_2, \dots, a_n$      $k$ .     $X$     some rule

Since this is a line in a proof (done correctly we are assuming) it is justified by one of our rules. We can assume that the proof is done using only the primitive rules ( $A, \rightarrow E, \rightarrow I, \& E, \& I, \vee E, \vee I, \leftrightarrow E, \leftrightarrow I, RAA$ ). If we want to extend the Soundness Theorem to derived rules it can easily be done later.

Line  $k$  must be justified using one of these ten rules. I will now complete the proof by examining each of these ten rules and arguing that a contradiction will follow no matter

which rule justifies line k. In effect, I will be arguing that none of our rules can introduce the first bad line in a proof.

**CASE 1.** Line k is justified by the rule of Assumptions. Then line k looks as follows:

k      k. X    A

But then line k cannot be a BAD line; for the definition of a BAD line is that there is a TVA that makes all of the assumptions written to the left of the line true and which makes the sentence written on the line false. But in this case, the assumption written to the left of the line is the very same sentence as the sentence written on the line. So, they cannot be assigned different truth-values by any TVA.

**CASE 2:** Line k is justified by the &E rule. Then there must be a line j earlier in the proof where the sentence occurring on line j is a sentence whose main connective is &, and where the sentence on line k is one of the conjuncts of j. The same assumptions are written to the left of lines j and k. The overall picture is thus:

$a_1, a_2, \dots a_n$	j.	X&Y	some rule
$a_1, a_2, \dots a_n$	k.	X	j &E

Officially, line j could also be of the form Y&X, but the same argument will apply in that case. Now, since we are assuming that line k is the first BAD line in the proof, this means that line j is not a BAD line so there is no TVA that makes  $a_1, a_2, \dots a_n$  all true and makes X&Y false. Because line k is BAD, there is some TVA that makes all of  $a_1, a_2, \dots a_n$  true and makes X false. Now using the truth-table for &, we can see that since X is false on this TVA, X&Y will be false also. But then this TVA makes all of all of  $a_1, a_2, \dots a_n$  true and makes X&Y false contradicting the fact that j is not a BAD line. In other words, if the use of &E produces a BAD line, then the line from which it is justified must also be a bad line. So the &E rule can't introduce the first BAD line into a proof.

**CASE 3:** Line k is justified by the &I rule. Then there are two earlier lines i and j where line i is the sentence X, line j is the sentence Y, and line k is the sentence X&Y. Since lines i and j are earlier than line k, neither of them is a BAD line. The assumptions written to the left of line k are the union of the assumptions written to the left of lines i and j. The overall picture is thus:

$a_1, a_2, \dots a_n$	i. X	some rule
$b_1, b_2, \dots b_m$	j. Y	some rule
$a_1, a_2, \dots a_n, b_1, b_2, \dots b_m$	k. X&Y	i,j &I

Since line k is BAD, there is a TVA that makes all of the assumptions written to the left of line k -  $a_1, a_2, \dots a_n, b_1, b_2, \dots b_m$  - true and makes X&Y false. By the truth-table for &, since this TVA makes X&Y false, the TVA must make X false or make Y false. This gives rise to two cases:

Case A. The TVA makes  $X$  false. Then the TVA makes all of  $a_1, a_2, \dots, a_n$  true but makes  $X$  false. This means that line  $i$  would be BAD, contradiction our assumption that  $k$  is the first BAD line in the proof.

Case B. The TVA makes  $Y$  false. Then the TVA makes all of  $b_1, b_2, \dots, b_m$  true but makes  $Y$  false. This means that line  $j$  would be BAD, again contradicting our assumption.

Thus  $\rightarrow$ I cannot introduce the first BAD line in a proof.

**CASE 4:** Line  $k$  is justified by the  $\rightarrow$ E rule. Then there are earlier lines  $i$  and  $j$ , with line  $i$  containing a sentence  $X$ , line  $j$  containing a the conditional  $X \rightarrow Y$ , and line  $k$  containing the consequent  $Y$ . Since lines  $i$  and  $j$  are earlier than  $k$ , neither of them is a BAD line. The assumptions written to the left of line  $k$  are the union of the assumptions written to the left of lines  $i$  and  $j$ . The overall picture is thus:

$a_1, a_2, \dots, a_n$	$i. X$	some rule
$b_1, b_2, \dots, b_m$	$j. X \rightarrow Y$	some rule
$a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_m$	$k. Y$	$i, j \rightarrow E$

Since line  $k$  is BAD, there is a TVA that makes all of the assumptions written to the left of line  $k$  -  $a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_m$  true and makes  $Y$  false. Now since this TVA makes  $b_1, \dots, b_m$  all true, and line  $j$  must not be a BAD line (since  $k$  is the first BAD line), this TVA must make  $X$  false. If it made  $X$  true, then  $X \rightarrow Y$  would be false making it a bad line. But since this TVA must make  $X$  false, line  $i$  is now a bad line. This shows that  $\rightarrow$ E can't introduce the first bad line in a proof.

**CASE 5:** Line  $k$  is justified by the  $\rightarrow$ I rule. Then there are earlier lines  $i$  and  $j$ , where line  $i$  is an assumption line containing a sentence  $X$ , line  $j$  contains the sentence  $Y$ , and line  $k$  contains the conditional  $X \rightarrow Y$ . Since lines  $i$  and  $j$  are earlier in the proof than  $k$ , neither of them is a BAD line. The assumptions to the left of line  $k$  are those to the left of line  $j$  with the exception of line  $i$ . The overall picture is thus:

$i$	$i. X$	A
$i, b_1, b_2, \dots, b_m$	$j. Y$	some rule
$b_1, b_2, \dots, b_m$	$k. X \rightarrow Y$	$j \rightarrow I (i)$

Note that it is not required for line  $j$  to depend on line  $i$ , but if it doesn't, the same argument will still apply. Now, since line  $k$  is a BAD line, there is a TVA that makes all of  $b_1, b_2, \dots, b_m$  true but makes  $X \rightarrow Y$  false. By the truth-table for  $\rightarrow$ , if  $X \rightarrow Y$  is false,  $X$  must be true and  $Y$  must be false. But now let's examine line  $j$ . Since line  $k$  is bad, we know that the TVA in question makes all of  $b_1, b_2, \dots, b_m$  true. Now, it doesn't matter if line  $j$  also depends on  $i$ , since line  $i$  is just sentence  $X$  which we know must be true. So

all of the assumptions that line  $j$  depends on are true on this TVA. However,  $Y$  is false and so line  $j$  is a BAD line. Thus  $\rightarrow I$  cannot introduce the first BAD line in a proof.

The proof continues by examining all of the remaining rules.

Once we have proven that none of the primitive rules can introduce the first BAD line in a proof, we have contradicted the fact that line  $k$  was the first BAD line of some proof (since it came from one of the rules.) But since this was just an arbitrary line of an arbitrary proof, we have shown that no line of any proof can be a bad line. Thus we have shown in particular, that the last line of any proof will never be a BAD line. Thus there can never be a TVA that makes all of the premises that the last line depends on true while making the sentence on the line false. Thus there is no TVA that makes all of the members of our set  $\Gamma$  true while making our conclusion ( $\phi$ ) false. Thus  $\Gamma \not\models \phi$ . Since we assumed that  $\Gamma \vdash \phi$ , we have now shown that If  $\Gamma \vdash \phi$  then  $\Gamma \models \phi$  which is the Soundness Theorem.