PHIL 50 – INTRODUCTION TO LOGIC

MARCELLO DI BELLO – STANFORD UNIVERSITY

DERIVATIONS IN PROPOSITIONAL LOGIC – WEEK #3

1 LOGICAL CONSEQUENCE V. DERIVATION

So far we have worked with the truth values of formulas, and the reference to truth values has made our treatment semantic. You should be familiar with the semantic notion of **logical consequence**:

$\{\varphi_1, \varphi_2, \ldots \varphi_k\} \models \psi$ **iff** all valuations V's that make true $\varphi_1, \varphi_2, \ldots, \varphi_k$ also make true ψ.

If we put $\{\varphi_1, \varphi_2, \ldots\ \varphi_k\} = \Gamma$, we can define the notion of logical consequence as follows:

$\Gamma \models \psi$ **iff** all valuations V's that make true all formulas in Γ also make true ψ.

I shall now introduce the syntactic notion of **derivation**, as follows:

$\Gamma \vdash \psi$ **iff** there is a derivation with assumptions the formulas in Γ and conclusion ψ.

Make sure you do not confuse the symbol $\models$ with the symbol $\vdash$. The former has to do with *semantic* logical consequences, and the latter has to do with *syntactic* derivations.

We shall see that a derivation always proceeds from a finite number of assumptions, so the formulas in Γ should be finite in number. Whenever a derivation of ψ rests on no assumption whatsoever, we can write $\vdash \psi$, which is the syntactic correlate of $\models \psi$.

But what is a derivation, exactly? In these notes, we shall go through the **derivation rules** that should be followed to construct derivations. So, we might say that a derivation is a "syntactic construction" which begins with a finite set of assumptions $\varphi_1, \varphi_2, \ldots, \varphi_k$ and arrives at a conclusion ψ through the application of some derivation rules. These rules are purely syntactical and based on symbol manipulation. The rules should be applied mechanically so that nothing is left for the imagination, so to speak. And yet, some imagination will be needed to decide which rules to apply and when and how as we are trying to construct a derivation. Although a computer should be able to check whether a derivation is correct or not—i.e. whether it follows the rules or not—some creative intelligence is needed to construct the derivation itself.

Before expounding on the derivation rules, it is important to grasp the distinction between $\models$ and $\vdash$. Later on we shall see that, within the system of propositional logic, $\models$ and

$\vdash$ are equivalent, as follows:

$$\Gamma \models \psi \textbf{ iff } \Gamma \vdash \psi$$

The left-to-right direction—i.e. if $\Gamma \models \psi$ then $\Gamma \vdash \psi$—is called **completeness**. This means that given any semantic relation of logical consequence from Γ to ψ, there is a corresponding derivation whose assumptions are the formulas in Γ and whose conclusion is ψ. Conversely, the right-to-left direction—i.e. if $\Gamma \vdash \psi$ then $\Gamma \models \psi$—is called **soundness**. This means that given any derivation, there is a corresponding relation of logical consequence. Now, if the set Γ is empty, then we can write:

$$\models \psi \textbf{ iff } \vdash \psi$$

The left-to-right direction tells us that given any semantically valid formula ψ—i.e. given any ψ such that $\models \psi$—there is a derivation of ψ from an empty set of assumptions. Conversely, the right-to-left direction tells us that if ψ can be derived from an empty set of assumptions, then ψ is semantically valid.

2 RULES R, $\wedge I$, $\wedge E$, $\rightarrow I$, $\rightarrow E$, $\bot$ AND RAA

The first set of rules we examine deals with reiteration, conjunction, implication, and the connective $\bot$. One feature of the propositional language we are using is that $\neg\varphi$ is simply an abbreviation for $\varphi \rightarrow \bot$. This means that our system we will not have rules for negation, but simply rules for $\bot$; more on this later.

REITERATION

(R)

$$\frac{\varphi}{\varphi}\ R$$

RULES FOR $\wedge$

($\wedge I$)

$$\frac{\varphi \quad \psi}{\varphi \wedge \psi}\ \wedge I$$

($\wedge E$)

$$\frac{\varphi \wedge \psi}{\varphi}\ \wedge E$$

$$\frac{\varphi \wedge \psi}{\psi}\ \wedge E$$

RULES FOR $\to$

$(\to I)$

$$\dfrac{\begin{matrix}[\varphi]^i \\ \vdots \\ \psi\end{matrix}}{\varphi \to \psi} \to I^i$$

$(\to E)$

$$\dfrac{\varphi \to \psi \quad \varphi}{\psi} \to E$$

RULES FOR $\bot$

$(\bot)$

$$\dfrac{\bot}{\psi} \bot$$

(RAA)

$$\dfrac{\begin{matrix}[\neg\varphi]^i \\ \vdots \\ \bot\end{matrix}}{\varphi} RAA^i$$

REMARKS AND EXAMPLES

The reiteration rule R should be self-explanatory. If we have a derivation of φ, then φ can be repeated in the next line of the derivation. Pay attention to how the rule is stated. The rule R only allows us to repeat the same formula in the next line of the derivation, not further down in the derivation. So the following is a wrong application of rule R:

WRONG!

$$\dfrac{\begin{matrix}\varphi \\ \vdots\end{matrix}}{\varphi} R$$

The rule for $\wedge I$ and $\wedge E$ should be self-explanatory, as well. The **rule for** $\wedge I$ codifies the idea that if we have a derivation of φ and a derivation of ψ, we also have a derivation of $\varphi \wedge \psi$. The **rule for** $\wedge E$ codifies the idea that if we have a derivation of $\varphi \wedge \psi$, we also have a derivation of either of the conjuncts.

Note that the rules for $\wedge$ are either for the introduction of $\wedge$ or for the elimination of $\wedge$. This is also how the derivation rules will be subdivided relative to the other connectives, such as $\rightarrow$ and $\vee$. The case of $\neg$ is somewhat special as we shall see.

The **rule for** $\rightarrow E$ is *modus ponens*, and it should be obvious. We should guard, however, against possible misuses. For instance, the following application of $\rightarrow E$ is wrong:

WRONG!

$$\dfrac{\begin{matrix}\varphi \rightarrow \psi \\ \vdots \quad\quad \\ \quad\quad \varphi\end{matrix}}{\psi} \rightarrow E$$

Both $\varphi \rightarrow \psi$ and φ should occur on the same line in the derivation and should also occur on the line immediately preceding the formula derived through the application of the rule $\rightarrow E$. More generally, when you construct a derivation, you cannot use formulas which have been derived earlier and which occur somewhere earlier in your derivation. So this is a wrong application of $\wedge I$:

WRONG!

$$\dfrac{\begin{matrix}\varphi \quad\quad \\ \vdots \quad\quad \\ \quad\quad \psi\end{matrix}}{\varphi \wedge \psi} \wedge I$$

The rules that we have seen so far apply only to formulas that are "immediately preceding" as you can see by looking at the formulation of rules R, $\wedge I$, $\wedge E$, and $\rightarrow E$. Not all the rules are like that, however. As we shall see shortly, the rules $\rightarrow I$ and RAA allow you to use formulas that occur earlier in the derivation.

So let's have a look at the **rule for** $\rightarrow I$. The rule says that if we assume φ and manage to derive ψ later on, we can cancel the original assumption φ and derive $\varphi \rightarrow \psi$. Importantly, we can cancel the assumption φ because φ has been "incorporated" as the antecedent of the formula $\varphi \rightarrow \psi$. Whenever we assume a formula we shall write it between square brackets and label it with a number, e.g. $[\varphi]^2$. A new assumption should be assigned a new number.

For the purpose of illustration, let's construct a derivation of $(\varphi \wedge \psi) \rightarrow (\psi \wedge \varphi)$ by using $\wedge I$, $\wedge E$, and $\rightarrow I$. In order to prove the implication $(\varphi \wedge \psi) \rightarrow (\psi \wedge \varphi)$, we proceed by assuming the antecedent $\varphi \wedge \psi$ and then derive the consequent $\psi \wedge \varphi$. Once we have done that, the rule $\to I$ allows us to cancel our assumption, and derive $(\varphi \wedge \psi) \to (\psi \wedge \varphi)$. Here is how the derivation will look like:

$$\dfrac{\dfrac{\dfrac{[\varphi \wedge \psi]^1}{\psi}\wedge E \quad \dfrac{[\varphi \wedge \psi]^1}{\varphi}\wedge E}{\psi \wedge \varphi}\wedge I}{(\varphi \wedge \psi) \to (\psi \wedge \varphi)}\to I^1$$

Note that we assumed two instances of $\varphi \wedge \psi$ and later derived $\psi \wedge \varphi$. When we applied rule $\to I$ to derive $(\varphi \wedge \psi) \to (\psi \wedge \varphi)$, we cancelled both instances of $\varphi \wedge \psi$. Is this allowed? Yes, because the derivation leading to $\psi \wedge \varphi$ comprised—or depended upon—both instances of $\varphi \wedge \psi$.

When you assume a formula, say φ, and later on manage to derive another formula, say ψ, there is no requirement that your derivation of ψ should use the assumption φ in any explicit way. You can assume anything and you can assume formulas in any way you want provided you respect the rules. It is perfectly fine to assume ψ and then later on assume φ. By applying $\to I$, we are then able to derive $\psi \to \varphi$. However, at this point our derivation will still hinge on the assumption φ, while the assumption ψ can be cancelled. We can also cancel the assumption φ provided we apply $\to I$ again and derive $\varphi \to (\psi \to \varphi)$. Here is what the derivation would look like:

$$\dfrac{\dfrac{\dfrac{[\psi]^1}{[\varphi]^2}}{\psi \to \varphi}\to I^1}{\varphi \to (\psi \to \varphi)}\to I^2$$

Also, we can easily derive $\varphi \to \varphi$ by assuming φ and then canceling the assumption φ by applying R and $\to I$, as follows:

$$\dfrac{\dfrac{[\varphi]^1}{\varphi}R}{\varphi \to \varphi}\to I^1$$

I shall finally comment on the rules $\bot$ and RAA. The **rule** $\bot$ codifies the idea that from the contradiction anything follows. So, if we manage to derive the contradiction, then any formula can be derived. Rule $\bot$ can also be called an "explosion rule" in the sense that it says that once you have derived the contradiction, then you can derive any formula whatsoever; this is like an explosion. The rule $\bot$ gives us a rationale for why contradictions are so bad and why we should guard against them. Of course, one might say, but what if we just eliminate the rule $\bot$ from our logical system? This is certainly an option. And some logical systems can limit the explosive power of the contradiction precisely by not adopting the rule $\bot$. These logics are called paraconsistent logics. The logical system we are studying, however, does have the rule $\bot$.

Now, deriving the contradiction means deriving a formula, say φ, and its negation, say $\neg\varphi$. At this point, we should note a peculiarity of the propositional language we are using. The formula $\neg\varphi$ is simply an abbreviation for $\varphi \to \bot$. (One can convince oneself that the abbreviation makes sense at the semantic level by constructing a truth table for $\neg\varphi$ and a truth table for $\varphi \to \bot$, where $\bot$ always gets assigned the value 0. As it turns out, the truth values for $\neg\varphi$ and $\varphi \to \bot$ are always the same.) Now, given that $\neg\varphi$ is the abbreviation of $\varphi \to \bot$, the following is a perfectly acceptable application of $\to E$:

$$\dfrac{\varphi \quad \neg\varphi}{\bot} \to E$$

And, again because $\neg\varphi$ is an abbreviation for $\varphi \to \bot$, the following is a perfectly acceptable application of $\to I$:

$$\dfrac{\begin{matrix}[\varphi]^1 \\ \vdots \\ \bot\end{matrix}}{\neg\varphi} \to I^1$$

Here is an example of a derivation using the rule $\bot$ together with other rules, and using the fact that $\neg\varphi$ is an abbreviation of $\varphi \to \bot$:

$$\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^1 \quad [\neg\varphi]^2}{\bot} \to E}{\psi} \bot}{\neg\varphi \to \psi} \to I^2}{\varphi \to (\neg\varphi \to \psi)} \to I^1$$

And here is another example of how the abbreviation under consideration can be used to derive the Principle of Non-Contradiction:

$$\dfrac{\dfrac{\dfrac{[\varphi \wedge \neg\varphi]^1}{\varphi} \wedge E \quad \dfrac{[\varphi \wedge \neg\varphi]^1}{\neg\varphi} \wedge E}{\bot} \to E}{\neg(\varphi \wedge \neg\varphi)} \to I^1$$

Finally, a few words on the **rule** RAA are in order. The rule says that if we assume $\neg\varphi$ and derive a contradiction $\bot$, then we can cancel our assumption $\neg\varphi$ and derive φ. The rule RAA codifies the so-called proof by contradiction. A proof by contradiction is one in which a positive claim, say φ, is established by assuming the negation of φ and by deriving a contradiction from the negation of φ. Importantly, the reasoning encoded by rule RAA is accepted by many mathematicians and logicians, although it is not accepted by the so-called intuitionistic mathematicians and logicians. The rule RAA, in other words, is a rule that belongs to classical logic, and not to intuitionistic logic. To illustrate the difference, let us consider the formula $\varphi \to \neg\neg\varphi$ and the formula $\neg\neg\varphi \to \varphi$. There is an easy derivation of the former formula that does not use RAA, as follows:

$$\dfrac{\dfrac{\dfrac{[\varphi]^1 \quad [\neg\varphi]^2}{\bot} \to E}{\neg\neg\varphi} \to I^2}{\varphi \to \neg\neg\varphi} \to I^1$$

So, the principle that $\varphi \to \neg\neg\varphi$ is accepted by the intuitionistic logician as well. On the other hand, the principle $\neg\neg\varphi \to \varphi$—which basically says that "two negations make an affirmation"—can be derived only by means of RAA (or by some analogous rule). Here is a derivation of $\neg\neg\varphi \to \varphi$:

$$\dfrac{\dfrac{\dfrac{[\neg\varphi]^1 \quad [\neg\neg\varphi]^2}{\bot} \to E}{\varphi} RAA^1}{\neg\neg\varphi \to \varphi} \to I^2$$

In order appreciate the power of RAA, it is important to keep in mind the difference between the following:

$$\begin{array}{c}[\neg\varphi]^i \\ \vdots \\ \dfrac{\bot}{\varphi}\ RAA^i\end{array}$$

and

$$\begin{array}{c}[\neg\varphi]^i \\ \vdots \\ \dfrac{\bot}{\neg\neg\varphi} \to I^i\end{array}$$

Without RAA, if we derive $\bot$ from $\neg\varphi$, we can only derive the double negative claim $\neg\neg\varphi$. With RAA, instead, we are allowed to derive the positive claim φ. Relatedely, one should not confuse the following:

$$\begin{array}{c}[\neg\varphi]^i \\ \vdots \\ \dfrac{\bot}{\varphi}\ RAA^i\end{array}$$

and

$$\begin{array}{c}[\varphi]^i \\ \vdots \\ \dfrac{\bot}{\neg\varphi} \to I^i\end{array}$$

3 RULES FOR DISJUNCTION

We have not yet introduced rules for the connective $\vee$. We shall do so now:

$(\vee I)$

$$\frac{\varphi}{\varphi \vee \psi}\ \vee I$$

$$\frac{\psi}{\varphi \vee \psi}\ \vee I$$

$(\vee E)$

$$\frac{\varphi \vee \psi \qquad \begin{array}{c}[\varphi]^i \\ \vdots \\ \sigma\end{array} \qquad \begin{array}{c}[\psi]^i \\ \vdots \\ \sigma\end{array}}{\sigma}\ \vee E^i$$

Let us begin with the **rule for** $\vee I$. This should be self-explanatory. The rule codifies the idea that if you have a derivation of φ, then you also have a derivation of $\varphi \vee \psi$, where ψ is any formula whatsoever.

To illustrate, together with other rules, rule $\vee I$ allows us to derive that $(\varphi \to \psi) \to (\varphi \to (\psi \vee \sigma))$, as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^1 \quad [\varphi \to \psi]^2}{\psi}\ \to E}{\psi \vee \sigma}\ \vee I}{\varphi \to (\psi \vee \sigma)}\ \to I^1}{(\varphi \to \psi) \to (\varphi \to (\psi \vee \sigma))}\ \to I^2$$

We can also derive the Principle of Excluded Middle in a somewhat cumbersome way as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^1}{\varphi \vee \neg\varphi}\ \vee I \quad [\neg(\varphi \vee \neg\varphi)]^2}{\bot}\ \to E}{\neg\varphi}\ \to I^1}{\varphi \vee \neg\varphi}\ \vee I \qquad [\neg(\varphi \vee \neg\varphi)]^2}{\bot}\ \to E}{\varphi \vee \neg\varphi}\ RAA^2$$

Note that the above derivation of $\varphi \vee \neg\varphi$ crucially rests on RAA. Without RAA, we would only be able to derive $\neg\neg(\varphi \vee \neg\varphi)$. The double negative claim $\neg\neg(\varphi \vee \neg\varphi)$ is accepted by the intuitionistic logician, but the positive claim $\varphi \vee \neg\varphi$ is not.

I shall now comment on the **rule for** $\lor E$. The rule says that if you have a derivation of σ from φ and if you have a derivation of σ from ψ, then you have a derivation of σ from $\varphi \lor \psi$. Note that if φ and ψ are assumptions, these assumptions are canceled through the application of $\lor E$. (If φ and ψ are not assumptions, then it is pointless to apply the rule in question. It is pointless, because you would be able to derive σ right away.) We should distinguish two case here.

In one case, $\varphi \lor \psi$ is a new assumption. If so, this new assumption should be assigned a different number from φ and ψ. The following derivation can illustrate the point:

$$\dfrac{\dfrac{\dfrac{[\varphi\lor\sigma]^3 \quad \dfrac{\dfrac{[\varphi]^1 \quad [\psi]^2}{\varphi\land\psi}\land I}{(\varphi\land\psi)\lor\sigma}\lor I \quad \dfrac{[\sigma]^1}{(\varphi\land\psi)\lor\sigma}\lor I}{(\varphi\land\psi)\lor\sigma}\lor E^1}{(\varphi\lor\sigma)\to((\varphi\land\psi)\lor\sigma)}\to I^3}{\psi\to((\varphi\lor\sigma)\to((\varphi\land\psi)\lor\sigma))}\to I^2$$

It can also be the case that there is an independent derivation for $\varphi \lor \psi$ (possibly resting on some other assumption), so in that case $\varphi \lor \psi$ won't be a new assumption. This derivation can illustrate the point:

$$\dfrac{\dfrac{\dfrac{[(\varphi\lor\sigma)\land\psi]^1}{\varphi\lor\sigma}\land E \quad \dfrac{\dfrac{[\varphi]^2 \quad \dfrac{[(\varphi\lor\sigma)\land\psi]^1}{\psi}\land E}{\varphi\land\psi}\land I}{(\varphi\land\psi)\lor\sigma}\lor I \quad \dfrac{[\sigma]^2}{(\varphi\land\psi)\lor\sigma}\lor I}{(\varphi\land\psi)\lor\sigma}\lor E^2}{((\varphi\lor\sigma)\land\psi)\to((\varphi\land\psi)\lor\sigma)}\to I^1$$

The rule $\lor E$ should be intuitive by now, although it is not intuitive why it is called rule for the *elimination* of $\lor$ given that the rule introduces $\lor$ rather than eliminating it. This is true, but take the name as an article of faith. And yet, you should appreciate the difference between $\lor I$ and $\lor E$. For one, $\lor I$ introduces a disjunction *below* the derivation line, and on the other hand, rule $\lor E$ introduces a disjunction *before* the derivation line, not after the derivation line. In this sense, we can say that $\lor E$ "eliminates" the disjunction below the derivation line.