Part 3. The Logic of Arguments

Dialectica with Serpent resting
on Aristotle with the Organon.
Chartres Cathedral
West Façade, Right Portal, ca. 1145-1155.

Into the assembly of the gods came Dialectic, a woman whose weapons are complex and knotty utterances....In her left hand she held a snake twined in immense coils; in her right hand a set of formulas, carefully inscribed on wax tablets, which were adorned with the beauty of contrasting colors, was held on the inside by a hidden hook; but since her left hand kept the crafty device of the snake hidden under her cloak, her right was offered to one and all. Then if anyone took one of those formulas, he was soon caught on the hook and dragged toward the poisonous coils of the hidden snake, which presently emerged and after first biting the man relentlessly with the venomous points of its sharp teeth then gripped him in its many coils and compelled him to the intended position. If no one wanted to take any of the formulas, Dialectic confronted them with some other questions; or secretly stirred the snake to creep up on them until its tight embrace strangled those who were caught and compelled them to accept the will of their interrogator.

Martianus Capella, *The Seven Liberal Arts*,
327-329 (310-339 A.D.)

Table of Contents

Part 3. The Logic of Arguments

INTRODUCTION

In Part 1 we studied the properties of terms. These were the words and phrases used to form sentences. We investigated what sorts of entities "in the world" the various parts of speech refer to. We reviewed the accounts of Plato, Aristotle and modern set theory. In Part 2 we studied sentences themselves. We investigated their grammar, and the different the truth conditions appropriate to each type of grammatical sentence. We began with the relatively simple grammar of categorical logic, then discussed the complex sentences of propositional logic, and finished with first-order logic, which contains simple and complex formulas sufficient for formulating most mathematics and science. In Part 3 we shall investigate deductive arguments. Our goal here is to review various attempts to distinguish good arguments from bad, and explain how they differ. Indeed, we have been building to this investigation all along because *logic* is simply another name for the study of good and bad deductive arguments.

As in Part 2, our investigation will be progressive. We will start with arguments formulated in the simple grammar of categorical propositions, move on to those formulated in propositional logic, and finish with those of first-order logic. Overarching this investigation, however, is one of the most interesting features of logic: there is more than one way to define "good and bad argument". First is the straightforward definition: "good arguments" are those that carry one from true premises to a true conclusion. Because this definition uses of the concept of truth, which consists of correspondence to the world, it is called *semantic*. The second definition of "good argument" is formulated in terms of the notion of proof. An argument is good if we can prove that the conclusion follows from the premises. The second approach requires that we be able to lay out how the conclusion follows in steps that are in some sense obvious. Each step then follows because it conforms to some simple logically obvious rule. These rules are obvious because they are formulated in terms of the shapes of formulas in the proof. Because they talk about shape they are said to be *syntactic* or *proof theoretic*. There is yet a third way to distinguish good from bad arguments. It makes use of what is called a *decision procedure*. This is any calculation process that has several important features: it works on any argument, it produce a judgment "yes" is if is a good and "no" if it is bad, it finishes in a finite time, and we can easily apply each step and understand its result. If a decision procedure exists for the "good" and "bad" arguments of a language, its logic is said to be *decidable*.

Our study will be divided according to this threefold division. We shall study arguments first semantically, then proof theoretically, and finally in terms of decidability. Each division will be progressive, moving from categorical, to propositional, and finally to first-order languages. We shall also remark on arguments in set theory, which is a branch of mathematics formulated in a first-order language. As we move through the different approaches to the different languages, you will be introduced to some of the major discoveries of modern logic.

One major result is that the semantic and syntactic approaches to "good" argument turn out to coincide, at least in categorical, propositional and first-order logic. This result, which is called the completeness metatheorem, is very important. It shows that two quite different approaches to the central idea of logic – valid argument – arrive at the same place.

One reason completeness is important is that it cannot be extended to other sciences. The completeness theorem for first-order logic entails that the set of logical truths are the same as those that can be proven as theorems in an axiom system. But the axiomatic method is not always successful. In particular it cannot be used to "capture" the truths of arithmetic. The Austrian logician Kurt Gödel showed in 1931 a famous result know as the *incompleteness theorem for arithmetic*. He proved that in any axiom system for arithmetic there at least one truth of arithmetic that is not a theorem. Since arithmetic is a subtheory of set theory, his result also shows that the axioms of set theory are incomplete. They leave out at least one theorem of set theory. Philosophers and mathematicians are still trying to figure out the implications of incompleteness for mathematics.

Another important result is that the "good" arguments of categorical and propositional logic are decidable, but those of first-order logic and any subject matter formulated in a first-order language – like set theory, arithmetic, and much of mathematics and natural science – are not decidable. Let us now see how this story unfolds.

Kurt Gödel

"If to the Peano axioms we add the logic of *Principia mathematica* .... (with the natural numbers as the individuals) together with the axiom of choice (for all types) we obtain a formal system S, for which the following theorems hold: I. The system S is not complete; that is it contains propositions A (and we can in fact exhibit such propositions) for which neither A nor $\sim A$ is provable and, in particular, it contains (even for decidable properties F of natural numbers) undecidable problems of the simple structure $\exists x F x$ where x ranges over the natural numbers."

Kurt Gödel, "Some Metamathematical Results on Completeness and Consistency," 1930

LECTURE 12. VALIDITY, CONSISTENCY, AND LOGICAL TRUTH

Semantic Entailment

The Semantic Definition of Validity

An argument is just a series of declarative sentences divided into two: its premises $\{P_1,\ldots,P_n,\ldots\}$ and a conclusion Q. It is much more difficult, however, to explain what is it is that makes some arguments "good" or "logical", and others not. Whatever a "good" argument is, languages like English provide many synonymous for talking about them. We can praise the argument from $P_1,\ldots,P_n$ to Q in all the following equivalent ways:

*If $P_1,\ldots,P_n$ are true, then Q **must be** true.*

*Q is a **logical consequence** of $P_1,\ldots,P_n$.*

*It is **necessary** that if $P_1,\ldots,P_n$ then Q.*

*$P_1,\ldots,P_n$ **entail** Q.*

*$P_1,\ldots,P_n$ **imply** Q.*

*The argument from $P_1,\ldots,P_n$ to Q is **valid.***

Valid arguments are the topic of Part 3. We will investigate how this set of arguments is properly defined and look into different ways to explain why some arguments are valid and others not. Validity is also known as *logical entailment, implication*, and *consequence*. It is also said to be logically *necessary*.

To represent the entailment relation, we shall use the symbol $\models$. Using "infix" notation, we place $\models$ between the premises and the conclusion:

$$\{P_1,\ldots,P_n\} \models_L Q$$

is read

$\{P_1,\ldots,P_n\}$ *semantically entails Q in L,*

or

the argument from $P_1,\ldots,P_n$ to Q is (logically) valid in L.

To simplify the notation, we will often delete the set brackets {…} and *L*, and write:

$$P_1,\ldots,P_n \models Q \qquad \text{for} \qquad \{P_1,\ldots,P_n\} \models_L Q\,.$$

Let us start with the definition of validity. The first point to make about the idea is that it is semantic. It concerns "the way signs related to the world." More precisely, it is semantic because it is defined in terms of the concept of truth, which in turn is defined as a correspondence to the world. A valid argument is one such that if its premises are true in a world, so is its conclusion.

Valid arguments, as it were, pass truth from premises to conclusions. That is why it is called entailment. In mediaeval law entailment was a restriction on property transfer between generations: property necessarily passed from the legator to a special category of legatee. In logical entailment truth passes necessarily from premises to a restricted category of sentences.

This entailment is often expressed in ordinary English by the subjunctive mood. If the premises *were* true the conclusion *would be* true. In modern logic the idea suggested by the subjunctive is usually explained in terms of "possible interpretations" or "possible worlds". Consider the argument:

Every M is P
Every S is M
∴*Every S is P*

No matter what sets *S, M* and *P* stand for, no matter what their interpretations, no matter what the world is like, if *Every M is P* and *Every S is M* are true in that interpretation or in that world, then so is *Every S is P*. Accordingly the standard definition of *validity* quantifies over "all interpretations":

$$\{P_1,\ldots,P_n\} \models_L Q \quad \text{iff} \quad \forall\Im\ (\text{if } \Im(P_1)=\text{T} \,\&\ldots\&\, \Im(P_n)=\text{T}) \text{ then } \Im(Q)=\text{T}).$$

Thus, an argument is *invalid* if there is an interpretation that makes the premises true and the conclusions false:

$$\{P_1,\ldots,P_n\} \not\models_L Q \quad \text{iff} \quad \exists\Im\ (\Im(P_1)=\text{T} \,\&\ldots\&\, \Im(P_n)=\text{T} \,\&\, \Im(Q)=\text{F}).$$

Showing Arguments are Valid

By definition, to show that an argument is valid, i.e. to show

$$\{P_1,\ldots,P_n\} \models_L Q\ ,$$

we must show a universally quantified conditional:

$$\forall\Im\ (\text{if } \underline{\Im(P_1)=\text{T} \,\&\ldots\&\, \Im(P_n)=\text{T})} \text{ then } \underline{\underline{\Im(Q)=\text{T}}}).$$

How do you prove a universally quantified conditional? The short answer is that you assume the <u>antecedent</u> and then deduce the $\underline{\underline{\text{consequen}}}$t. That is, for an arbitrary $\Im$, we prove the conditional:

$$\text{If } \underline{\Im(P_1)=\text{T} \,\&\ldots\&\, \Im(P_n)=\text{T}}, \text{ then } \underline{\underline{\Im(Q)=\text{T}}}.$$

(Here we are using single and double underlines to highlight the antecedent and consequent of the conditional to be proved.) How do we prove the conditional? We adopt the strategy called conditional proof. We set up a subproof:

Start of subproof			
	1.	$\underline{\Im(P_1)=\text{T} \,\&\ldots\&\, \Im(P_n)=\text{T}}$	Assumption for conditional proof
	…	…	
	n.	$\underline{\underline{\Im(Q)=\text{T}}}$	
End of Subproof			
n+1	If $\underline{\Im(P_1)=\text{T} \,\&\ldots\&\, \Im(P_n)=\text{T}}$, then $\underline{\underline{\Im(Q)=\text{T}}}$		1-*n*, conditional proof

n+2	$\forall\Im$ (if $\Im(P_1)$=T &…& $\Im(P_n)$=T) then $\Im(Q)$=T).		universal generalization, $\Im$ arbitrary
n+3	$\{P_1,\ldots,P_n\}\vDash_L Q$	n+2,	def of $\vDash$

How do we complete the subproof? Recall that the definition of $\Im$ allows us to calculate for each sentence type its "truth-conditions" in $\Im$. That is, the definition of $\Im$, entails for each *P* an instance of the (T) schema:

(T) $\Im(P)$=T iff $TC_\Im(P)$

such that $TC_\Im(P)$ states in the terminology of set theory that some relation that holds among the $\Im$-values of the smallest referring expressions in *P.* We saw how to determine such truth-conditions for categorical propositions, sentences in propositional logic, and formulas in first-order logic. We apply this technique for the premises and conclusion to determine the list of equivalences:

$\Im(P_1)$=T iff $TC_\Im(P_1)$
…
$\Im(P_n)$=T iff $TC_\Im(P_n)$
$\Im(Q)$=T iff $TC_\Im(Q)$

Now, from the assumption of the subproof we know for each *I*

$\Im(P_i)$=T

We also know

$\Im(P_i)$=T iff $TC_\Im(P_i)$

From these two we may deduce by modus ponens:

$TC_\Im(P_i)$

The combined information in these various truth conditions $TC_\Im(P_1),\ldots,TC_\Im(P_n)$ lays out in detail what must be true in the world when $P_1,\ldots,P_n$ are true. This information will in fact be sufficient to show, by appeal to logic and facts of set theory, that the conditions $TC_{\Im}(Q)$ for making Q true are met. That is we can show from

$$TC_{\Im}(P_1)\&\ldots\&TC_{\Im}(P_n)$$

that

$$TC_{\Im}(Q)$$

But we also know

$$\Im(Q)=T \text{ iff } TC_{\Im}(Q)$$

Hence by *modus ponens* again

$$\Im(Q)=T$$

Hence and the subproof is complete and we have shown by conditional proof the conditional:

$$\underline{TC_{\Im}(P_1)\&\ldots\&TC_{\Im}(P_n)} \rightarrow \underline{\underline{\Im(Q)=T}}$$

We may summarize the strategy for a "validity proof" in a schema. The schema is repeated below. We use underlines and colors to indicate its structure. The overall strategy is to show that a conditional is true: *if* the argument's premises are true, *then* its conclusion is. The technique used to prove the conditional is conditional proof, a rule which requires a subproof. The *if*-part is assumed as the assumption of the subproof, and the *then*-part is deduced as its last line . The subproof then "proves" the conditional. To indicate the structure of the subproof, the *if*-part assumed as the subproof's first line is underlined, and the *then*-part deduced as its last line is double underlined.

Within the subproof, there are various applications of *modus ponens.* The (T) formula for a proposition *P,* which is a biconditional of the form $\Im(P)=T$ iff $TC_{\Im}(P)$, is written as a line of the proof. Using *modus ponens* one side of the biconditional is then shown to be true by showing that the other side is true. To indicate the structure, the side being deduced is colored yellow, and the side previously proven is colored green.

Schema for Proofs of Validity

Metatheorem Proof Schema. $\{P_1,\ldots,P_n\} \models_L Q$

Proof

Start of subproof

1.	$\Im(P_1)$=T &…& $\Im(P_n)$=T	Assumption for conditional proof, $\Im$ arbitrary
2.	$\Im(P_1)$=T	line 1, conjunction
3.	$\Im(P_1)$=T iff $TC_{\Im}(P_1)$	(T) schema entailed by the definition of $\Im$
4.	$TC_{\Im}(P_1)$	*modus ponens* on the previous two lines
…	…	…
3*n*+1.	$\Im(P_n)$=T	line 1, conjunction
3*n*+2.	$\Im(P_n)$=T iff $TC_{\Im}(P_n)$	(T) schema entailed by the definition of $\Im$
3*n*+3.	$TC_{\Im}(P_n)$	*modus ponens* on the previous two lines
3*n*+4.	$TC_{\Im}(P_1)$ &…& $TC_{\Im}(P_n)$,	conjunction of previous TC lines
3*n*+5.	$TC_{\Im}(Q)$	by set theory and logic from the previous line
3*n*+6.	$\Im(Q)$=T iff $TC_{\Im}(Q)$	(T) schema entailed by the definition of $\Im$
2*n*+7.	$\Im(Q)$=T	*modus ponens* on the previous two lines

End of subproof

3*n*+8.	If ($\Im(P_1)$=T &…& $\Im(P_n)$=T) then $\Im(Q)$=T	1 to *n*+5, conditional proof
3*n*+9.	$\forall\Im$(if ($\Im(P_1)$=T &…& $\Im(P_n)$=T) then $\Im(Q)$=T)	*n*+6, universal generalization, $\Im$ arbitrary
3*n*+10.	$\{P_1,\ldots,P_n\} \models_L Q$	*n*+7, definition of $\models$

Although the details of the various truth conditions $TC_{\Im}(P_1),\ldots,TC_{\Im}(P_n)$, and $TC_{\Im}(Q)$ will vary from one argument to anther, if the argument is valid, the facts contained in $TC_{\Im}(P_1),\ldots,TC_{\Im}(P_n)$ will allow us to jump by logic and set theory to those contained in $TC_{\Im}(Q)$. That is, given the facts of $TC_{\Im}(P_1),\ldots,TC_{\Im}(P_n)$, it follows by set theory that the facts of $TC_{\Im}(Q)$ hold. We will justify this step by the simple annotation "by set theory."

Showing Arguments are Invalid

By definition, to show that the argument from $P_1,\ldots,P_n$ to Q is invalid, i.e.

$\{P_1,\ldots,P_n\} \nvDash_L Q,$

we must prove an existentially quantified conjunction:

$\exists\Im\ (\Im(P_1)=T\ \&\ldots\&\ \Im(P_n)=T\ \&\ \Im(Q)=F).$

How do you prove an existentially quantified conjunction that quantifies over interpretations? The short answer is that you construct an interpretation $\Im$ that satisfies the conjunction. Construction here means "define." An interpretation is a set (a set of pairs), and we know that if we can define a set it exists. All we need do is define an interpretation $\Im$ that makes true each of the conjuncts in the conjunction.

Notice that what we must make true are facts about the truth-values of the sentences in the argument. That is we must prove $\Im(P_1)=T\ \&\ldots\&\ \Im(P_n)=T\ \&\ \Im(Q)=F$. (Note that another way to say $\Im(Q)=F$ is $\Im(Q)\neq T$.)

But how we prove facts about truth-values? One way is through truth-conditions. For each of these sentences we know by its (T) schema when it is T or F in $\Im$:

$\Im(P_1)=T$ iff $TC_\Im(P_1)$

…

$\Im(P_n)=T$ iff $TC_\Im(P_n)$

$\Im(Q)=T$ iff $TC_\Im(Q)$

The working backwards, we see that all we need to prove that $P_1,\ldots,P_n$ are all T is to prove that there truth-conditions are satisfied

$\Im(P_1)=T$

…

$\Im(P_n)=T$

If these were true, then given the (T) schemata, it would follows by modus ponens that $P_1,\ldots,P_n$ are all T. Similarly, we know the (T) schema for *Q:*

$\Im(Q)$=T iff $TC_{\Im}(Q)$

Then, to show that $\Im(Q)\neq$T all we would need to know is $\sim TC_{\Im}(Q)$. Given the (T) schema and $\sim TC_{\Im}(Q)$ it would follow by modus tollens that

$\Im(Q)\neq$T

The strategy of the proof then is to define an interpretation $\Im$ that makes all of $TC_{\Im}(P_1),\ldots,TC_{\Im}(P_n)$ true but $TC_{\Im}(Q)$ false. How do we do this? How do we make $TC_{\Im}(P_1),\ldots,TC_{\Im}(P_n)$ true but $TC_{\Im}(Q)$ false? Recall what each of these clauses says. The each assert that $\Im$ assigns sets and set members to objects in the world in a way that makes the sentences in question T or F. All we have to do is make sure that $\Im$ assigs sets and entities to the terms in $P_1,\ldots,P_n$, and Q in the right way. Recall and $\Im$ is just a set of pairs and we can make up a set just by defining it. Therefore we make up $\Im$ by defining the $\Im$-values of the words in $P_1,\ldots,P_n$, and Q it in such a way that $TC_{\Im}(P_1),\ldots,TC_{\Im}(P_n)$ turn our true and $TC_{\Im}(Q)$ false. What values these will be will depend on the grammar of $P_1,\ldots,P_n$, and Q and the referring terms they contain. In hard cases it also requires ingenuity, but the examples we will consider are all fairly easy. We will use Venn diagrams and truth-tables to find the values we need.

We may summarize the strategy for proving invalidity in the following schema:

Schema for an Invalidity Proof

Metatheorem Proof Schema. $\{P_1,\ldots,P_n\} \nvDash_L Q$

Proof. Let $e_1,\ldots,e_m$ be the smallest referring terms that occur in $P_1,\ldots,P_n$, Q. Let us $\Im$ for these expressions be as follows: $\Im(e_1)=\ldots$, …, $\Im(e_m)=\ldots$. [It is here is that ingenuity may be required.]

1.	$TC_\Im(P_1)$ & … &$TC_\Im(P_n)$ & not $TC_\Im(Q)$	By set theory & the def of $\Im(e_1),\ldots,\Im(e_m)$.
2.	$TC_\Im(P_1)$	1, conjunction
3.	$\Im(P_1)$=T iff $TC_\Im(P_1)$	(T) schema entailed by the definition of $\Im$
4.	$\Im(P_1)$=T	2,3 *modus ponens*
…	…	
n+4.	$TC_\Im(P_n)$	1, conjunction
n+5.	$\Im(P_n)$=T iff $TC_\Im(P_n)$	(T) schema entailed by the definition of $\Im$
n+6	$\Im(P_n)$=T	n+4,n+5 *modus ponens*
n+7	not $TC_\Im(Q)$	1, conjunction
n+8.	$\Im(Q)$=T iff $TC_\Im(Q)$	(T) schema entailed by the definition of $\Im$
n+9	not $\Im(Q)$=T	n+7,n+8 *modus tollens*
n+10	$\Im(Q)$=F	n+9, bivalence of $\Im$
n+11.	$\Im(P_1)$=T &…& $\Im(P_n)$=T & $\Im(Q)$=F	4, …,n+6,n+10 conjunction
n+12.	$\exists\Im$ ($\Im(P_1)$=T &…& $\Im(P_n)$=T & $\Im(Q)$=F).	n+11, existential construction
n+13.	$\{P_1,\ldots,P_n\} \nvDash_L Q$	n+13, definition of $\nvDash$

Consistency

The Semantic Definition of Consistency

A second logical concept closely tied to validity is consistency. A set of sentences is consistent if all the sentences contained in it can be true together, and is inconsistent if they cannot.

$\{P_1,\ldots,P_n\}$ is *consistent* iff $\exists\Im(\Im(P_1)=T\&\ldots\&\Im(P_n)=T)$.

$\{P_1,\ldots,P_n\}$ is *inconsistent* iff $\sim\exists\Im(\Im(P_1)=T\&\ldots\&\Im(P_n)=T)$.

To simplify notation we will often leave off the set brackets. There are a variety of synonyms for inconsistency. An inconsistent set is also said to be *contradictory*, *absurd, impossible*, and *unsatisfiable.* A single sentence that is never true is also said to be inconsistent. That is, *P* is inconsistent means $\sim\exists\Im(\Im(P)$=T. (In other words the one element set {*P*} is also inconsistent.) For example, in propositional logic $p_1\wedge \sim p_1$ is inconsistent because there is no $\Im$ such that $\Im(p_1\wedge \sim p_1)$=T. Such sentences are also called *contradictions*.

Like validity, inconsistency is a matter of grammatical form, and indeed the two ideas can be defined in terms of each other. To say that the set $\{P_1,\ldots,P_n ,\sim Q\}$ is inconsistent means that there is not way to make $P_1,\ldots,P_n$ all true and *Q* false, but this is just to say that the argument from $P_1,\ldots,P_n$ to *Q* is valid.

Note also a consequence of the definitions. If a set $\{P_1,\ldots,P_n\}$ logically implies every sentence *Q,* it must be contradictory because it would imply contradictions along with everything else. Conversely, if $\{P_1,\ldots,P_n\}$ implies a contradiction, it implies every sentence whatever because every sentence follows from a contradiction. For example, the argument $p_1\wedge \sim p_1 \models q$ is valid because $\forall\Im$ (if $\Im(p_1\wedge \sim p_1)$=T the $\Im(q)$=T) since its antecedent $\Im(p_1\wedge \sim p_1)$=T is always false. Hence, the following equivalence is true:

$\{P_1,\ldots,P_n\}$ implies a contradiction iff $\{P_1,\ldots,P_n\}$ implies every sentence.

We use this fact below:

Interdefinablity of Consistency and Validity

$\{P_1,\ldots,P_n\} \models_L Q$ iff $\{P_1,\ldots,P_n ,\sim Q\}$ is inconsistent.

$\{P_1,\ldots,P_n$ is consistent iff for no Q, $\{P_1,\ldots,P_n\} \models Q\wedge\sim Q$

iff for some Q, $\{P_1,\ldots,P_n\} \not\models Q$

Logical Truth

The Semantic Definition of Logical Truth

A third important logical idea is logical truth. Some sentences are always true. That they are so is a function of their grammatical form. For example, *Ever S is S,* and $p_1 \vee \sim p_1$ are always true. Again "always" is explained in terms of "all interpretations" or "all worlds":

P is a *logical truth* in L iff $\forall\Im(\Im(P)=T)$

Necessity is a synonym for *logical truth*, and in the special case of the propositional logic, logical truths are called *tautologies*. If the language in question (like propositional logic) contains the connectives $\wedge$ and $\rightarrow$, it is possible to define validity, consistency and logical truth in terms of each other.

Interdefinablity of Validity, Consistency and Logical Truth

Theorem

$\{P_1,\ldots,P_n\} \models Q$	iff	$\{P_1,\ldots,P_n, \sim Q\}$ is inconsistent
	iff	$(P_1\wedge\ldots\wedge P_n)\rightarrow Q$ is a logical truth
$\{P_1,\ldots,P_n\}$ is consistent	iff	for no Q, $\{P_1,\ldots,P_n\} \models Q\wedge\sim Q$
	iff	for some Q, $\{P_1,\ldots,P_n\} \not\models Q$
	iff	$\sim(P_1\wedge\ldots\wedge P_n)$ is not a logical truth

P is a logical truth	iff	for every Q, $Q \models P$
	iff	$\sim P$ is inconsistent

The fact that these ideas are interdefinable makes the task of logical theory simpler. It means that there is really one fundamental idea in logic, valid argument, which can be expressed in different ways. In the lectures that follow we shall investigate the idea in detail. First we shall see what validity is for the three increasingly powerful languages we investigated in Part 2: the syllogistic, propositional logic, and first-order logic. We shall explore both how the semantic definition of validity must adapt itself to the different notions of "interpretation" appropriate to each language, and how the set of valid arguments expands along with the language's expressive power. Later in Part 2 we shall investigate syntactic proofs and decision procedures for valid arguments. These provide techniques independent of semantic that impart to our knowledge of validities a degree of certainty virtually unequaled in any other science.

LECTURE 13. CATEGORICAL LOGIC: VALIDITY

The syllogistic form is one of the most beautiful of the human spirit, and indeed one of the most important. It is a kind of universal mathematics, one that is not sufficiently appreciated. You might even say that it contains an art that leads to infallibility.

Leibniz, *New Essays* IV, xvii, 4.

Categorical Logic

In this lecture we shall investigate our first extended example of a "logical theory". The theory we start with is nice and simple, the logic of categorical propositions that were introduced in Lecture 7. Traditionally this logic is divided into two parts, arguments that have a single premise, called *immediate inferences*, and those with two premises and three terms called *syllogisms*. These two groups will suffice to explain valid arguments composed of categorical propositions with any number of premises.

In the last lecture we defined the two main logical ideas we will be investigating, validity and invalidity. But because categorical logic antedates modern logic – it is literally 2,400 years old – its terminology has some quaint quirks. One is the way to refer to a valid argument with a single premise. If the argument from *P* to *Q* is valid, *Q* is said to be a *subaltern* to *P* . *Q* is also said to stand in the *subalternation* relation to *P*. (This usage persists in British army jargon: a lieutenant is a "subaltern" to a higher officer.)

Another logical idea important in categorical logic is *contrariety*. Two propositions are said to be *contrary* if they are never true together. In ordinary language, *red* and *green* are said to be contrary because nothing can be (completely) red and (completely) green at the same time. We will find that in categorical logic, **A***SP* and **E***SP* cannot both be true. Similarly, propositions that cannot both be false are said to be *subcontraries*. Neither contraries nor subcontraries need have opposite true-values, because two contraries may be simultaneous false, and two subcontraries may be simultaneously true. Propositions that must have opposite truth-values are said to be *contradictories.* Of these terms, *contrary* and *contradictory* are still used in modern logic, but outside categorical logic *subcontrary* has also passed out of usage.

Definition of Logical Concepts

1. $P_1,\ldots,P_n \models_{SL} Q \quad \leftrightarrow \quad \forall\mathfrak{I}\ (\ \mathfrak{I}(P_1)\text{=T}\ \&\ldots\&\ \mathfrak{I}(P_1)\text{=T}) \rightarrow \ \mathfrak{I}(Q)\text{=T})$
2. $P_1,\ldots,P_n \not\models_{SL} Q \quad \leftrightarrow \quad \exists\mathfrak{I}\ (\ \mathfrak{I}(P_1)\text{=T}\ \&\ldots\&\ \mathfrak{I}(P_1)\text{=T}\ \&\ \mathfrak{I}(Q)\text{=F}\ )$
3. *P* and *Q* and are *contradictories* $\leftrightarrow\ \forall\mathfrak{I}\ (\mathfrak{I}(P)\text{=T} \leftrightarrow\ \mathfrak{I}(Q)\text{=F})$
4. *P* and *Q* and are *contraries* $\leftrightarrow\ \sim\exists\mathfrak{I}\ (\mathfrak{I}(P)\text{=T}\ \&\ \mathfrak{I}(Q)\text{=T})$
5. *P* and *Q* and are *subcontraries* $\leftrightarrow\ \sim\exists\mathfrak{I}\ (\mathfrak{I}(P)\text{=F}\ \&\ \mathfrak{I}(Q)\text{=F})$

To prove facts about these logical ideas we will appeal to the truth-conditions of for categorical propositions as proven in Lecture 7. For convenience we summarize these here in the following metatheorem:

Theorem. The following instances of the (T) schema are true:

			Truth-conditions
TC1.	$\mathfrak{I}(\mathbf{A}XY)$=T	$\leftrightarrow$	$\mathfrak{I}(X)\subseteq\mathfrak{I}(Y)$
TC2.	$\mathfrak{I}(\mathbf{E}XY)$=T	$\leftrightarrow$	$\mathfrak{I}(X)\cap\mathfrak{I}(Y)=\varnothing$
TC3.	$\mathfrak{I}(\mathbf{I}XY)$=T	$\leftrightarrow$	$\mathfrak{I}(X)\cap\mathfrak{I}(Y) \neq\varnothing$
TC4.	$\mathfrak{I}(\mathbf{O}XY)$=T	$\leftrightarrow$	$\mathfrak{I}(X)-\mathfrak{I}(Y) \neq\varnothing$

Some of the proofs below will also depend on the Assumption of categorical semantics that terms always refer to a non-empty set:

Theorem. For any term X and any interpretation $\Im$,

$$\Im(X)\neq\varnothing \ \& \ \Im(X)\subseteq U$$

Since we have already proven quite a few basic truths of set theory in Part 1, here we will assume that they are true without further argument. We will, for example, write a fact of set theory down as a line in a proof, or advance to a new line of a proof by substituting a set theoretic identity or equivalence with any further justification than "by set theory".

Among the facts about sets that we have previous proven is:

(1) $A\subseteq B$ iff $\sim(A\cap -B \neq\varnothing)$.

A special case of (1) would be:

(2) $\Im(F)\subseteq\Im(G)$ iff $\sim(\Im(F) \cap -\Im(G) \neq\varnothing)$.

This is one of the set theoretic equivalence we shall use in a proof. Lets us see how. Note that the (T) schemata stated in the definition of $\Im$ for **A** and **O** propositions are:

(3) $\Im(\mathbf{A}XY)$=T iff $\Im(X)\subseteq\Im(Y)$

(4) $\Im(\mathbf{O}XY)$=T iff $\Im(X)-\Im(Y) \neq\varnothing$

It then follows by the substitution of equivalents (3) and (4) into (2) that

(5) $\Im(\mathbf{A}FG)$=T iff $\Im(\mathbf{O}FG)$=F.

Likewise, we proved earlier in set theory that

(6) $A\cup B\neq\varnothing$ iff $\sim(A\cap B=\varnothing)$.

A special case of (6) is

(7) $\Im(F)\cup\Im(G)\neq\varnothing$ iff $\sim(\Im(F)\cap\Im(G)=\varnothing)$.

You will need to cite this equivalence in one of your exercises when you show:

(8) $\Im(\mathbf{E}FG)$=T iff $\Im(\mathbf{I}FG)$=F.

Another elementary fact of set theory is

(9) If $A-B\neq\varnothing$, then $A\neq\varnothing$.

We shall also use this fact in a proof below. In exercises you may want to cite other relevant facts about sets. So long as it is a fact about sets, you may write it down as a line of a proof justifying it with the worlds “by set theory”.

The syntax and semantics of categorical propositions are short and sweet and enables us to prove that arguments are valid and invalid with simple and clear applications of the paradigms sketched in the introductory lecture to Part 3. (If these are not fresh in your mind, go back and review the paradigms for proving an argument is valid or invalid.)

Immediate Inference

We start by showing a simple case of validity that hold among pairs of propositions. Because in these cases inference holds between a pair of propositions without the need of mediating additional premises, the relation is called an *immediate inference*.

Metatheorems

Theorem. **A**$XY \models_{SL}$ **I**XY (An **I**-proposition is subaltern to an **A**-proposition.)

Analysis of the Proof. First note that by the definition of $\models_{SL}$,

A$XY \models_{SL}$ **I**XY means $\forall\Im(\Im(\mathbf{A}XY)=T\rightarrow\Im(\mathbf{I}XY)=T)$.

But $\forall\Im(\Im(\mathbf{A}XY)=T\rightarrow\Im(\mathbf{I}XY)=T)$ is a universally quantified. To prove it, we must first prove the conditional $\Im(\mathbf{A}XY)=T\rightarrow\Im(\mathbf{I}XY)=T$ for an arbitrary $\Im$, and then add the $\forall\Im$ by Universal Generalization. We show the conditional $\Im(\mathbf{A}XY)=T\rightarrow\Im(\mathbf{I}XY)=T$ by Conditional Proof, i.e. by constructing a subproof that starts by assuming the antecedent $\Im(\mathbf{A}XY)=T$ as its first line, and concludes with the consequent $\Im(\mathbf{I}XY)=T$ as its last line. So the general form of the proof will be:

Start of Subproof.

$\Im(\mathbf{A}XY)=T$	Assump. for conditional proof, assuming $\Im$ is arbitrary
…	
$\Im(\mathbf{I}XY)=T$	

End of subproof

$\Im(\mathbf{A}XY)=T\rightarrow\Im(\mathbf{I}XY)=T$	Conditional Proof given the successful subproof
$\forall\Im(\Im(\mathbf{A}XY)=T\rightarrow\Im(\mathbf{I}XY)=T)$.	Previous line by Universal Generalization, $\Im$ arbitrary.

Note that $\Im(\mathbf{A}XY)=T$ and $\Im(\mathbf{I}XY)=T$ are explained by (T) schemata in the definition of $\Im$, which provides for each an equivalent formulation in terms of facts about the sets $\Im(X)$ and $\Im(Y)$:

$\Im(\mathbf{A}XY)=T \leftrightarrow \Im(X)\subseteq\Im(Y)$

$\Im(\mathbf{I}XY)=T \leftrightarrow \Im(X)\cap\Im(Y)\neq\varnothing$

So, the lines $\Im(\mathbf{A}XY)=T$ and $\Im(\mathbf{I}XY)=T$ may be replaced by equivalents $\Im(X)\subseteq\Im(Y)$ and $\Im(X)\cap\Im(Y)\neq\varnothing$. The task then would be to prove $\Im(X)\cap\Im(Y)\neq\varnothing$ from $\Im(X)\subseteq\Im(Y)$. Recall that we also know that $\Im(X)\neq\varnothing$ and $\Im(Y)\neq\varnothing$. It is a fact of set theory that if $\Im(X)\subseteq\Im(Y)$ and $\Im(X)\neq\varnothing$, then $\Im(X)\cap\Im(Y)\neq\varnothing$. Moreover we can now assume any fact from set theory that we need. Hence, by appeal set theory we can go from $\Im(X)\subseteq\Im(Y)$ and $\Im(X)\neq\varnothing$, to $\Im(X)\cap\Im(Y)\neq\varnothing$.

Proof. Start subproof.

1.	$\Im(\mathbf{A}XY)=T$	assumption conditional proof, $\Im$ arbitrary in *SI*
2.	$\Im(\mathbf{A}XY)=T \leftrightarrow \Im(X)\subseteq\Im(Y)$	TC1
3.	$\Im(X)\subseteq\Im(Y)$	1, *modus ponens*
4.	$\Im(X)\neq\varnothing$	definition of $\Im$
5.	$\Im(X)\cap\Im(Y)\neq\varnothing$	3 and 4, set theory (note that we need line 4)
6.	$\Im(\mathbf{I}XY)=T \leftrightarrow \Im(X)\cap\Im(Y)\neq\varnothing$	TC3
7.	$\Im(\mathbf{I}XY)=T$	5 and 6, *modus ponens*

End subproof

8.	$\Im(\mathbf{A}XY)=T\rightarrow\Im(\mathbf{I}XY)=T$	1-4, conditional proof
9.	$\forall\Im(\Im(\mathbf{A}XY)=T\rightarrow\Im(\mathbf{I}XY)=T)$	5, universal generalization, $\Im$ arbitrary
10.	**A**$XY \models_{SL}$ **I**XY	6, definition of $\models$

Theorem. **E**$XY \models_{SL}$ **O**XY

∗**Exercise: E**$XY \models_{SL}$ **O**XY

Theorem. **A**XY and **O**XY are contradictories

Analysis of the Proof. First note that by the definition of contradictories, what me must prove is:

$\forall\Im\ (\Im(\mathbf{A}XY)\text{=T} \leftrightarrow\ \Im(\ \mathbf{O}XY)\text{=F})$

which by the bivalence of $\Im$, means

$\forall\Im\ (\Im(\mathbf{A}XY)\text{=T} \leftrightarrow\ \sim\Im(\ \mathbf{O}XY)\text{=T})$

Since this is universally quantified, we must prove $\Im(P)$=T $\leftrightarrow$ $\Im(Q)$=F for an arbitrary $\Im$. Note that each half of this biconditional asserts something about the way $\Im$ assigns truth-values. But the definition of $\Im$ provides (T) equivalents for these assignments formulated in terms of the sets $\Im(X)$ and $\Im(Y)$:

$\Im(\mathbf{A}XY)\text{=T} \leftrightarrow \Im(X)\subseteq\Im(Y)$

$\Im(\mathbf{O}XY)\text{=T} \leftrightarrow \Im(X)-\Im(Y)\neq\varnothing$

So if we replace the facts about $\Im$'s truth-value assignments by their equivalents in terms of conditions on the sets $\Im(X)$ and $\Im(Y)$, we must prove is:

$\Im(X)\subseteq\Im(Y)\leftrightarrow \sim(\Im(X)-\Im(Y)\neq\varnothing)$

which by double negation is just

$\Im(X)\subseteq\Im(Y)\leftrightarrow \Im(X)-\Im(Y)=\varnothing)$

Now, $A\subseteq B$ iff $A-B=\varnothing$ is a truth of set theory, so it remains a truth when A is replaced by $\Im(X)$ and B by $\Im(Y)$. Since we can now simply write down a truth of set theory if we need it, we can simply write down this truth and justify it by appeal to set theory. (Notice that this entire analysis consists of "working backward".)

Proof

11.	$\Im(X)\subseteq\Im(Y)\leftrightarrow \Im(X)-\Im(Y)=\varnothing$	truth of set theory
12.	$\Im(X)\subseteq\Im(Y)\leftrightarrow \sim\Im((X)-\Im(Y)\neq\varnothing)$	1, Double Negation
13.	$\Im(\mathbf{A}XY)\text{=T} \leftrightarrow \Im(X)\subseteq\Im(Y)$	TC1
14.	$\Im(\mathbf{O}XY)\text{=T} \leftrightarrow \Im(X)-\Im(Y)\neq\varnothing$	TC4
15.	$\sim\Im(\mathbf{O}XY)\text{=T} \leftrightarrow \sim\Im(X)-\Im(Y)\neq\varnothing$	4, ~ on both sides
16.	$\Im(\mathbf{O}XY)\text{=F} \leftrightarrow \sim\Im(X)-\Im(Y)\neq\varnothing$	5, bivalence of $\Im$
17.	$\Im(\ \mathbf{A}XY)\text{=T} \leftrightarrow \Im(\ \mathbf{O}XY)\text{=F}$	2, 3, and 6, substitution of equivalents

Theorem. **E***XY* and **I***XY* are contradictories

∗**Exercise.** Prove: **E***XY* and **I***XY* are contradictories

Theorem. **A***XY* and **E***XY* are contraries

Analysis of the Proof. First note that by the definition of contraries, what me must prove is:

$\sim\exists\Im\ (\Im(\mathbf{A}XY)\text{=T}\ \&\ \Im(\mathbf{E}XY)\text{=T})$

One common way to show an negative existential proposition is to assume its opposite and deduce a contradiction. That is, if we can construct a subproof that starts in the first line with the assumption

$\exists\Im\ (\Im(\mathbf{A}XY)\text{=T}\ \&\ \Im(\mathbf{E}XY)\text{=T})$

and then terminates in the last line with a contradiction, then we have proven the opposite of the first line by the rule Reduction to the Absurd. So, we start a subproof with that assumption.
Then by Existential Instantiation we drop the initial $\exists\Im$, giving the interpretation in question the temporary name $\Im$:

$\Im(\mathbf{A}XY)\text{=T}\ \&\ \Im(\mathbf{E}XY)\text{=T}$

Now, each side of this conjunction asserts something about $\Im$ assigning truth-values. But the definition of $\Im$ provides (T) equivalents for these assignments formulated in terms of the sets $\Im(X)$ and $\Im(Y)$:

$\Im(\mathbf{A}XY)\text{=T} \leftrightarrow \Im(X)\subseteq\Im(Y)$

$\Im(\mathbf{E}XY)\text{=T} \leftrightarrow \Im(X)\cap\Im(Y)=\varnothing$

So if we replace the facts about $\Im$'s truth-value assignments by their equivalents in terms of conditions on the sets $\Im(X)$ and $\Im(Y)$, we know is:

$\Im(X)\subseteq\Im(Y)$ & $\Im(X)\cap\Im(Y)=\varnothing$

Recall that we also know that $\Im(X)\neq\varnothing$ and $\Im(Y)\neq\varnothing$. It is a fact of set theory that

$(\Im(X)\subseteq\Im(Y)$ & $\Im(X)\neq\varnothing) \rightarrow \Im(X)\cap\Im(Y)\neq\varnothing$.

Since this is fact of set theory we may simply assert it.
It then follows by modus ponens that

$\Im(X)\cap\Im(Y)\neq\varnothing$.

But since we already know

$\Im(X)\cap\Im(Y)=\varnothing$.

We have deduced a contradiction. We may then terminate the subproof, concluding the negation of its first list by appeal to the rule *Reductio*.

Proof . Start of subproof.

18. $\exists\Im(\Im($ **A**$XY)$=T & $\Im($**E**$XY)$=T)	assump. for *reductio*
19. (**A**$XY)$=T & $\Im($**E**$XY)$=T	1, Existential Instantiation, $\Im$ not arbitrary
20. $\Im($**A**$XY)$=T $\leftrightarrow$ $\Im(X)\subseteq\Im(Y)$	TC1
21. $\Im($**E**$XY)$=T $\leftrightarrow$ $\Im(X)\cap\Im(Y)=\varnothing$	TC2
22. $\Im(X)\subseteq\Im(Y)$ & $\Im(X)\cap\Im(Y)=\varnothing$	2, 3 and 4, substitution of equivalents
23. $\Im(X)\subseteq\Im(Y)$	3, conjunction
24. $\Im(X)\cap\Im(Y)=\varnothing$	3, conjunction
25. $\Im(X)\neq\varnothing$	2, definition of $\Im$
26. $\Im(X)\subseteq\Im(Y)$ & $\Im(X)\neq\varnothing$	4 and 6, conjunction
27. $(\Im(X)\subseteq\Im(Y)$ & $\Im(X)\neq\varnothing)\rightarrow\Im(X)\cap\Im(Y)\neq\varnothing$	set theory
28. $\Im(X)\cap\Im(Y)\neq\varnothing$	7 and 8, *modus ponens*
29. $(\Im(X)\cap\Im(Y)=\varnothing)$ & $\Im(X)\cap\Im(Y)\neq\varnothing$	5 and 9, conjunction

End of subproof.

30. ~$\exists\Im(\Im($ **A**$XY)$=T & $\Im($**E**$XY)$=T) 1-10, *reductio*
31. **I**XY and **O**XY are subcontraries

Theorem. **I**XY and **O**XY are subcontraries

∗**Exercise.** Prove: **I**XY and **O**XY are subcontraries

The Square of Opposition

The theory of immediate inference is traditionally summarized in a famous diagram called the *Square of Opposition*, a teaching tool that has been faithfully reproduced in elementary logic texts since at least the logic of Apuleius (124-170 A.D.) Here is a version that incorporates the grammar, truth-conditions, and metatheorems we have just reviewed. The diagram incorporates the usual assumption of traditional logic that for a universal affirmative to be true its subject term must stand for at least one object. Moreover this assumption is a necessary underpinning of some of the logical relations we have just proved. However, the assumption that a term has a non-empty extension is rejected by modern logic, which leaves open that a set might be empty. Hence, a number of the logical relations claimed to hold in traditional logic are rejected by modern logic. The rejected assumptions and claims are indicated below by red type. The formulations and relations depicted in black type are all accepted in modern logic.

The Square of Opposition

(Diagram of Immediate Inferences)

A: *Every S is P* (A*SP*)

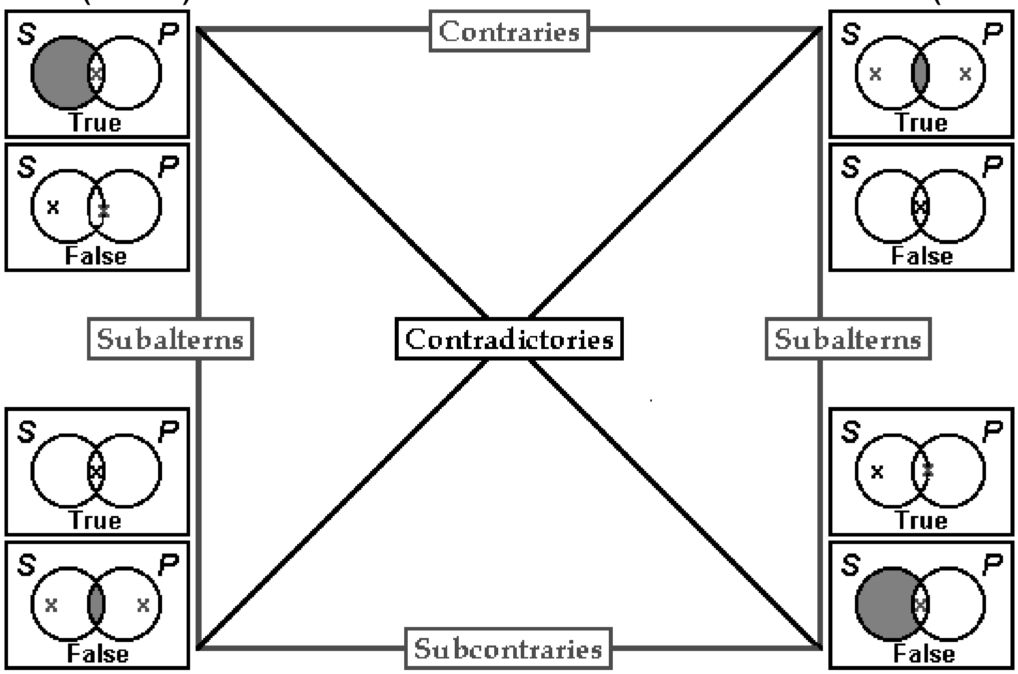
$\Im(\boldsymbol{S}) \neq \emptyset \;\&\; \Im(\boldsymbol{S}) \subseteq \Im(\boldsymbol{P})$

$\exists x Sx \;\&\; \forall x(Sx \rightarrow Px)$

E: *No S is P* (E*SP*)

$\Im(\boldsymbol{S}) \neq \emptyset \;\&\; \Im(\boldsymbol{P}) \neq \emptyset \;\&\; \Im(\boldsymbol{S}) \cap \Im(\boldsymbol{P}) = \emptyset$

$\exists x Sx \;\&\; {\sim}\exists x(Sx \& Px)$



I: *Some S is P* (I*SP*)

$\Im(\boldsymbol{S}) \cap \Im(\boldsymbol{P}) \neq \emptyset$

$\exists x(Sx \& Px)$

O: *Some S is not P* (O*SP*)

$\Im(\boldsymbol{S}) - \Im(\boldsymbol{P}) \neq \emptyset$

$\exists x(Sx \& {\sim}Px)$

For completeness we now list two related metatheorems, which are rather obvious but which we will have reason to refer to later. The first of these follows from the truth conditions for **E** and **O** propositions and the trivial set theoretic facts that $A \cap B = B \cap A$, and $A \cap B = \emptyset \leftrightarrow B \cap A = \emptyset$.

Theorem (Simple Conversion)

1. $\mathbf{E}XY \models_{SL} \mathbf{E}YX$
2. $\mathbf{O}XY \models_{SL} \mathbf{O}YX$

This theorem plus the relations of subalternation on the Square of Opposition entail the next theorem.

Theorem (Conversion *per Accidens*)

1. $\mathbf{A}XY \models_{SL} \mathbf{I}YX$
2. $\mathbf{E}XY \models_{SL} \mathbf{O}YX$

∗The Logic of Empty Terms and Negations

The syntax and the semantics for the extended language that incorporates term negations is summarize below: [1]

Definitions

1. $Trms^+ = Trms \cup \{ \bar{X} \mid X \in Trms\}$
2. $Prop^+ = \{\mathbf{Z}XY \mid \mathbf{Z} \in \{\mathbf{A},\mathbf{E},\mathbf{I},\mathbf{O}\}\ \&\ X \in Trms^+\ \&\ Y \in Trms^+\}$
3. The set SI^+ is the set of all *interpretation* $\Im$ for $SSyn^+$ relative to a domain U that meet these conditions: $\Im$ is a function (set of pairs) that pairs a term in $Trms^+$ to a non-empty subset of U and that pairs a proposition in $Prps^+$ to one of the two truth-values T or F, and is such that, for all terms X and Y,
   a. $\Im(X) \subseteq U\ \&\ \Im(X) \neq \varnothing$
   b. $\Im(\bar{X}) = U - \Im(X)$
   c. $\Im(\mathbf{A}XY)=\mathrm{T} \leftrightarrow \Im(X) \subseteq \Im(Y)$
   d. $\Im(\mathbf{E}XY)=\mathrm{T} \leftrightarrow \Im(X) \cap \Im(Y) = \varnothing$
   e. $\Im(\mathbf{I}XY)=\mathrm{T} \leftrightarrow \Im(X) \cap \Im(Y) \neq \varnothing$
   f. $\Im(\mathbf{O}XY)=\mathrm{T} \leftrightarrow \Im(X) - \Im(Y) \neq \varnothing$
4. SL^+, the *enlarged syllogistic language*, is $< SSyn^+, SI^+>$
5. Let $\Im$ stand for interpretations in SI^+.
   $P_1,\ldots,P_n \models_{SL^+} Q) \leftrightarrow \forall\Im\ ((\Im(P_1)=\mathrm{T}\ \&\ldots\&\ \Im(P_n)=\mathrm{T}) \rightarrow \Im(Q)=\mathrm{T})$

[1] This and later supplementary section may be omitted without loss of continuity.

Obversion

A special set of logical relations among propositions that depend on negative terms were summarized in the mediaeval period as the syllogistic rule called *obversion*.[2] Let a term and its predicate negation be called *opposites.* The rule may be formulate as follows

Obversion. If P and Q are of the same quantity but different quality, and have the same subject term but opposite predicate terms, then P and Q logically entail each other.

By this rule the following logically entail each another:

Every F is G	*No F is* $\bar{G}$
Every F is $\bar{G}$	*No F is G*
Some F is G	*Some F is not* $\bar{G}$
Some F is $\bar{G}$	*Some F is not G*

To summarize these facts in logical notation, it will help to introduce an abbreviation for "mutually entailment":

$P \dashv\vDash_{SL} Q$ means ($P \vDash_{SL} Q$ and $Q \vDash_{SL} P$).

The proofs are straightforward, and since they introduce no new ideas, will be omitted here.

Theorem

$\mathbf{A}\,FG \dashv\vDash_{SL} \mathbf{E}\,F\bar{G}$

$\mathbf{A}\,F\bar{G} \dashv\vDash_{SL} \mathbf{E}FG$

$\mathbf{I}\,FG \dashv\vDash_{SL} \mathbf{O}F\bar{G}$

$\mathbf{I}F\bar{G} \dashv\vDash_{SL} \mathbf{O}FG$

Terms with Empty Extensions

The syntax and semantics of the extended language amended so as to allow that terms have empty extensions is summarized below. We retain the syntax *SSyn*$^+$ that includes negative predicates, but alter the definition of an interpretation so that it may assign the empty set to a term. The new set of interpretation is *SI*$^{+\varnothing}$, the language that uses these interpretations is *SL*$^{+\varnothing}$, and the logical entailment relation for this language is $\models_{SL^{+\varnothing}}$.

Formal Syntax and Semantics

Definitions

1. $Trms^+ = Trms \cup \{ \bar{X} \mid X \in Trms\}$
2. $Prop^+ = \{\mathbf{Z}XY \mid \mathbf{Z} \in \{\mathbf{A},\mathbf{E},\mathbf{I},\mathbf{O}\} \;\&\; X \in Trms^+ \;\&\; Y \in Trms^+\}$
3. The set *SI*$^{+\varnothing}$ is the set of all *possibly empty interpretation* $\Im$ for *SSyn*$^+$ relative to a domain *U* that meet these conditions: $\Im$ is a function (set of pairs) that pairs a term in *Trms*$^+$ to a possibly empty subset of *U* and that pairs a proposition in *Prps*$^+$ to one of the two truth-values T or F, and is such that, for any terms *X* and *Y,*
    a. $\Im(X) \subseteq U$
    b. $\Im(\bar{X}) = U - \Im(X)$
    c. $\Im(\mathbf{A}XY)=\mathrm{T} \leftrightarrow \Im(X)=\varnothing \;\&\; \Im(X) \subseteq \Im(Y)$
    d. $\Im(\mathbf{E}XY)=\mathrm{T} \leftrightarrow \Im(X) \cap \Im(Y)=\varnothing$
    e. $\Im(\mathbf{I}XY)=\mathrm{T} \leftrightarrow (\Im(X) \cap \Im(Y) \neq \varnothing \text{ or } \Im(X)=\varnothing \text{ or } \Im(Y)=\varnothing)$
    f. $\Im(\mathbf{O}XY)=\mathrm{T} \leftrightarrow (\Im(X) - \Im(Y) \neq \varnothing \text{ or } \Im(X)=\varnothing)$
4. *SL*$^{+\varnothing}$, the *enlarged syllogistic language*, is < *SSyn*$^+$, *SI*$^{+\varnothing}$>
5. Let $\Im$ stand for interpretations in *SI*$^{+\varnothing}$.

    $$P_1,\ldots,P_n \models_{SL^{+\varnothing}} Q) \leftrightarrow \forall \Im \, (\Im(P_1)=\mathrm{T} \,\&\ldots\&\, \Im(P_n)=\mathrm{T}) \rightarrow \Im(Q)=\mathrm{T})$$

---

2 The rule was know to Proclus in the 7th century.

It should be noted that allowing terms to be "empty" (i.e. stand for empty sets) complicates the number of ways a universal proposition may be false, and the conditions under which particular propositions may be true.

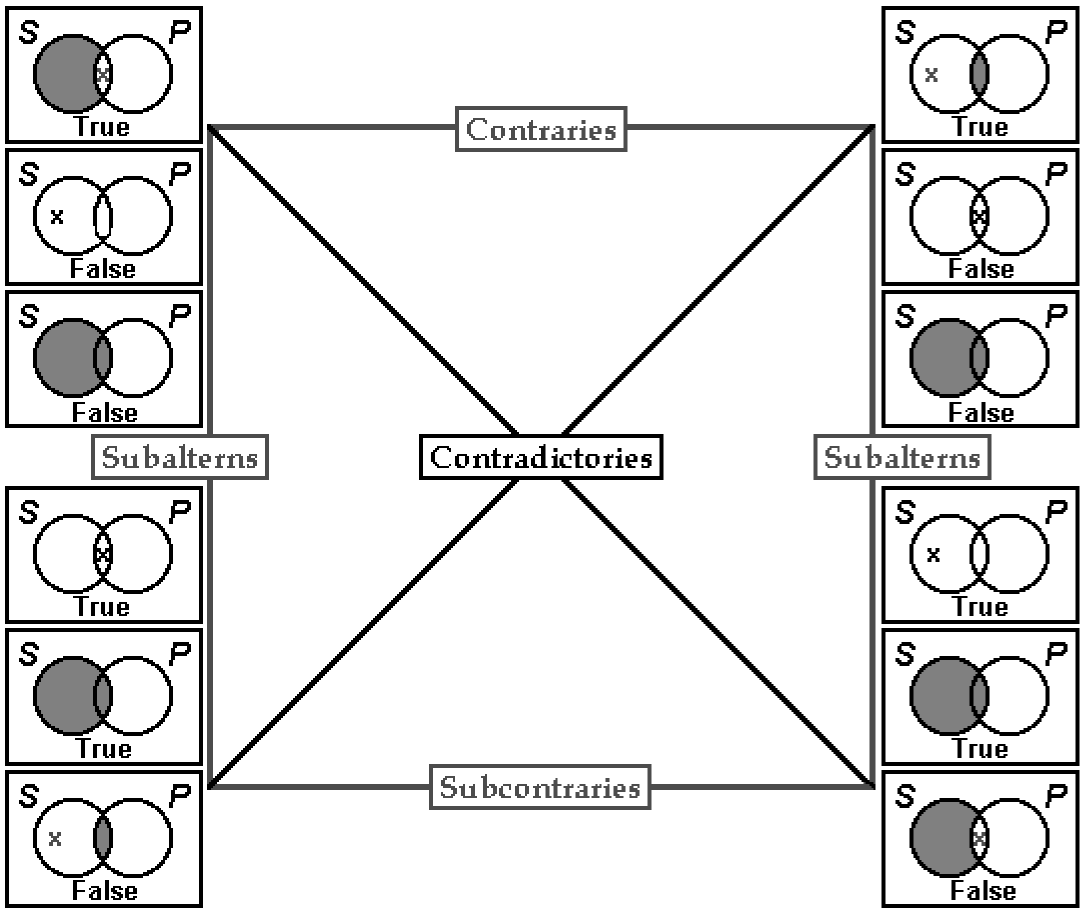
A universal affirmative **A** proposition *Every S is P* is true only if two conditions are met: (1) its subject term must be non-empty and (2) the set it stands for must be a subset of that named by the predicate. If either condition fails the proposition is false. Moreover, its contradictory opposite **I** proposition is *Some S is not P.* This means that whenever the one is true the other is false. It follows that there are now two cases in which the **I** propositions must be true: (1) when the subject term is non-empty and stands for a set that is not a subset of the one named by the predicate, which is the usual case, or (2) when the subject term is empty – this is a new and somewhat odd case. The new case is dictated by two desires: to allow for empty terms, and to retain the relation of contradictoriness across the diagonal of the Square of Opposition.

A similar complication arises for the universal negative *No S is P* and its contradictory *Some S is P.* In the new theory *No S is P* now false in three cases: (1) when the two terms are non-empty and name sets with an empty intersection, which is the normal case, (2) when the subject term is empty, and (3) when the predicate term is empty. Accordingly, its contradictory opposite the **O** proposition *Some S is P* must be true in any of the three cases.

What Aristotle found interesting about predicate negation is that when it is combined with the possibility of empty terms the logic of term negation is different from that of sentence negation. That is, there are two logically distinct forms of negations. Predicate negation is the logically stronger or "more informative" because when applied to a sentence's predicate, a predicate negation entails a sentence negation, but not conversely. This difference appears when the subject term is not truly predicated of anything that exists at the time of the predication, or in modern terms if the subject term stands for the empty set. Aristotle's example of such an empty term is *goatstag.* In *De Interpretatione* X he summarize the facts in a square of opposition, which may be rendered in our terms as follows:

Square of Opposition for Predicate Negations

Every S is P **(A*SP*)** ***Every S is*** $\bar{P}$ **(A*S*$\bar{P}$)**



~***Every S is*** $\bar{P}$ **(~A*S*$\bar{P}$)** ~***Every S is P*** **(~A*SP*)**

Theorems

1. $\mathbf{A}XY \models_{SL+\varnothing} {\sim}\mathbf{A}X\bar{Y}$
2. $\mathbf{A}X\bar{Y} \models_{SL+\varnothing} {\sim}\mathbf{A}XY$
3. $\mathbf{A}XY$ and ${\sim}\mathbf{A}XY$ are contradictories (with respect to $\models_{SL+\varnothing}$)
4. $\mathbf{A}X\bar{Y}$ and ${\sim}\mathbf{A}X\bar{Y}$ are contradictories (with respect to $\models_{SL+\varnothing}$)
5. $\mathbf{A}XY$ and $\mathbf{A}X\bar{Y}$ are contraries (with respect to $\models_{SL+\varnothing}$)
6. ${\sim}\mathbf{A}X\bar{Y}$ and ${\sim}\mathbf{A}XY$ are subcontraries (with respect to $\models_{SL+\varnothing}$)

The Syllogistic

Basic Concepts

In this lecture we shall study what are unquestionably the most well known logical arguments, syllogisms. It is syllogisms that educated people think of when they hear the word “logic”. Etymologically *syllogism* is simply the Greek word for *argument*, but the term was used by Aristotle for a special class of arguments. These are the arguments he singled out as central to scientific reasoning, and it is these that formed the core of his logical studies. In the Middle Ages as part of their regular course of studies university students learned to use syllogisms in the composition of their work, and their professors wrote the details of their scientific research in syllogistic form, regarding syllogisms as an important part of their “scientific method.” The graduation examination at European universities, at some universities as late as the 19th century, included a public disputation in which the student organized his points in syllogisms.

A syllogism is a three-line argument consisting of two premises and a conclusion, each line of which is a categorical proposition. The subject term of the conclusion is called the *minor term* and it must occur in the second premise, which is called the *minor premise*. The predicate of the conclusion is called the *major term*, and it must occur in the fist premise, which is called the *major premise*. (The terminology derives from the fact that in Latin one of the meanings of *major* is *first.* Likewise, *minor* means *second*.) A third term, which is called the *middle term*, must occur in both premises. The major, minor, and middle terms are required to be different.

It follows from these definitions that the three terms can be arranged in only one of four possible arrangements, which are called *figures.* These possibilities are: [3]

Figure 1. The middle term is first in the major premise and second in the minor.
Figure 2. The middle term is second in both the major and the minor premises.
Figure 3. The middle term is first in the major and the minor premises.
Figure 4. The middle term is second in the major premise and first in the minor.

Let *S, P,* and *M* represent terms in the set *Trms.* Then the four patterns are:

1st		2nd		3rd		4th	
	MP		***PM***		***MP***		***PM***
	SM		***SM***		***MS***		***MS***
	SP		***SP***		***SP***		***SP***

These four patterns are used to state a more formal definition for a syllogism. Since the basic theory does not allow for terms to be negated (i.e. to have a *not*- or *un*- prefixed to them), we need only assume in this formal definition the simpler of the syllogistic language defined in Part 2, which does not contain negated terms. There are two stages. We first give a formal definition of a syllogism in each of the four figures, and then combine them. Below we shall use the Greek letters Φ, Χ*,* and Ψ as variables to represent (to stand for) the

[3] Aristotle himself and logicians in the high Middle Ages recognized only three figures, conflating the fourth figure into the first. This combined group was defined as any syllogism in which the position of middle term differed in the two premises. The division of the group into what became know as the first and fourth figures, however, was recognized in ancient logic by the time of Galen (second century A.D.), and became a standard feature of the theory in the late Middle Ages and later. For a discussion of the origins of the fourth figure see Günther Patzig, *Aristotle's Theory of the Syllogism* (Dordrecht: Reidel, 1968). On the irrelevance of the order of the premises in Aristotle own account see Lynn Rose, *Aristotle's Syllogistic* (Springfield, Ill., Thomas, 1968). Theoretically the order of the premises is irrelevant to whether the argument is valid, and in this sense Aristotle's original version, which was followed by the leading mediaevals, is the more elegant.

operators **A**,**E**,**I**, and **O**. When these variables are used below, they may stand for any operator, including the case in which they all stand for the same operator. (That is, even though Φ, X, and Ψ are different variables, they may all stand for the same operator, for example, **A**.) As earlier, we let *S, P,* and *M* represent terms in the set *Trms.* Recall that by definition it is required that the three different terms in a syllogism are all different.

First, we define the notion of a *figure*. The first figure, for example, is the set of all syllogisms of the form $<\Phi MP, \mathrm{X}SM, \Psi SP>$ such that *S, P,* and *M* are all distinct terms. This would include all cases in which Φ, X, and Ψ are replaced by any of the operators **A**,**E**,**I**,**O**. For example, it would include <**A***MP*, **I***SM,* **E***SP*>, <**O***MP*, **E***SM,* **O***SP*>, and <**E***MP*, **E***SM,* **E***SP*>.

Given the four figures it is then a *syllogism* is defined any argument that fits one of the four figures.

Definitions

The set Fig_1 of *Syllogisms of Figure* 1 is:

$$\{<\Phi MP, \mathrm{X}SM, \Psi SP> \mid \{\Phi,\mathrm{X},\Psi\} \subseteq \{\mathbf{A},\mathbf{E},\mathbf{I},\mathbf{O}\} \;\&\; \{S,P,M\} \subseteq Trms \;\&\; S \neq P \;\&\; P \neq M \;\&\; M \neq S\}$$

The set Fig_2 of *Syllogisms of Figure* 2 is:

$$\{<\Phi PM, \mathrm{X}SM, \Psi SP> \mid \{\Phi,\mathrm{X},\Psi\} \subseteq \{\mathbf{A},\mathbf{E},\mathbf{I},\mathbf{O}\} \;\&\; \{S,P,M\} \subseteq Trms \;\&\; S \neq P \;\&\; P \neq M \;\&\; M \neq S\}$$

The set Fig_3 of *Syllogisms of Figure* 3 is:

$$\{<\Phi MP, \mathrm{X}MS, \Psi SP> \mid \{\Phi,\mathrm{X},\Psi\} \subseteq \{\mathbf{A},\mathbf{E},\mathbf{I},\mathbf{O}\} \;\&\; \{S,P,M\} \subseteq Trms \;\&\; S \neq P \;\&\; P \neq M \;\&\; M \neq S\}$$

The set Fig_4 of *Syllogisms of Figure* 4 is:

$$\{<\Phi PM, \mathrm{X}MS, \Psi SP> \mid \{\Phi,\mathrm{X},\Psi\} \subseteq \{\mathbf{A},\mathbf{E},\mathbf{I},\mathbf{O}\} \;\&\; \{S,P,M\} \subseteq Trms \;\&\; S \neq P \;\&\; P \neq M \;\&\; M \neq S\}$$

The set of *Syllogisms* is the union of the set of syllogisms of each figure:

$$Fig_1 \cup Fig_2 \cup Fig_3 \cup Fig_4$$

The list, in order, of the operators used in a syllogism (in its major premise, minor premise, and conclusion) is called the syllogism's *mood.*

Definition

The *mood* of a syllogism <Φ*MP*, Χ*MS,* Ψ*SP*> is the triple <Φ,Χ,Ψ>.

Below we shall drop the pointy brackets and abbreviate <Φ,Χ,Ψ> as ΦΧΨ.

(The word *mood* here is just a translation of the Latin word *modus,* which means manner or fashion. It has nothing to do with emotions.)

What is interesting about moods is that given a syllogism's mood and figure, its form is uniquely determined. For example, using (as is customary) *S* for the minor term, *P* for the major term, and *M* for the middle term, the mood **AEO** in the fourth figure determines the syllogism <**A***PM,***E***MS,***O***SP*>, or in English:

Every P is M
<u>*No M is S*</u>
∴*Some S is not P*

The syllogism with mood **OAI** in the second figure is <**O***PM,***A***SM,***I***SP*>, or:

Some P is not M
<u>*Every M is S*</u>
∴*Some S is P*

There are 64 syllogisms in each figure (three propositions, each of which may have one of four operators, makes 4^3 cases), and hence there are 256 syllogisms in all. Some of these are valid arguments and some not. The remainder of this lecture will concern how we demonstrate which is which.

Proving Syllogisms are Valid

A valid argument is one such that whenever its premise are true its conclusion is true too, and an invalid argument is one such that there is some case in which its premises are true, but its conclusion false. These facts are straightforward consequences of the definition of valid argument, but since we will be using these facts a good deal, it is useful to state the relevant metatheorem for reference. As defined, a syllogism is a triple of propositions $<P,Q,R>$. Saying that the syllogism $<P,Q,R>$ as a whole is valid is just an alternative way of saying that the argument from P,Q to R is valid, or in earlier notation that $P,Q \models_{SL} R$. Recall that the definition says:

$$P,Q \models_{SL} R \leftrightarrow \forall\Im\ ((\Im(P)=\mathrm{T}\ \&\Im(Q)=\mathrm{T}) \rightarrow \Im(R)=\mathrm{T})$$

This follow directly from the definition of valid argument:

$$P_1,\ldots,P_n \models_{SL} Q) \leftrightarrow \forall\Im((\Im(P_1)=\mathrm{T}\&\ldots\&\Im(P_1)=\mathrm{T})\rightarrow \Im(Q)=\mathrm{T})).$$

Proofs of the Validity of Selected Syllogisms

The definition makes clear how to prove that a syllogism $<P,Q,R>$ is valid. You use a conditional proof. You start a subproof by assuming that P and Q are T in an arbitrary $\Im$. Then in the subproof you deduce that R is T in $\Im$. You then conclude the subproof because you have shown the truth of the conditional $(\Im(P)=\mathrm{T}\ \&\Im(Q)=\mathrm{T}) \rightarrow \Im(R)=\mathrm{T}$. Since $\Im$ is arbitrary, you can then generalize to $\forall\Im\ ((\Im(P)=\mathrm{T}\ \&\Im(Q)=\mathrm{T}) \rightarrow \Im(R)=\mathrm{T})$, which by definition means that the syllogism is valid. Consider an example.

Theorem. The syllogism **AAA** in the first figure (called Barbara), i.e. $<\mathbf{A}MP,\mathbf{A}SM,\mathbf{A}SP>$, is valid.

Analysis of the Proof. First note what it is that must be proven: $<\mathbf{A}MP,\mathbf{A}SM,\mathbf{A}SP>$is valid, or in alternative notation

$$\mathbf{A}MP,\mathbf{A}SM \models_{SL} \mathbf{A}SP$$

By the definition of $\models_{SL}$, this means:

$$\forall\Im((\Im(\ \mathbf{A}MP)=\mathrm{T}\ \&\ \Im(\ \mathbf{A}SM)=\mathrm{T})\rightarrow\Im(\ \mathbf{A}SP)=\mathrm{T}).$$

But $\forall\Im((\Im(\ \mathbf{A}MP)=\mathrm{T}\ \&\ \Im(\ \mathbf{A}SM)=\mathrm{T})\rightarrow\Im(\ \mathbf{A}SP)=\mathrm{T})$ is a universally quantified. To prove it, we must first prove the conditional $(\Im(\ \mathbf{A}MP)=\mathrm{T}\ \&\ \Im(\mathbf{A}SM)=\mathrm{T})\rightarrow\Im(\ \mathbf{A}SP)=\mathrm{T}$ for an arbitrary $\Im$, and then add the $\forall\Im$ by Universal Generalization. We show the conditional $(\Im(\mathbf{A}MP)=\mathrm{T}\ \&\ \Im(\ \mathbf{A}SM)=\mathrm{T})\rightarrow\Im(\ \mathbf{A}SP)=\mathrm{T}$ by Conditional Proof, i.e. by constructing a subproof that starts by assuming the antecedent

$\underline{\Im(\mathbf{A}MP)=T \ \& \ \Im(\mathbf{A}SM)=T}$ as its first line, and concludes with the consequent $\underline{\Im(\mathbf{A}SP)=T}$ as its last line. So the general form of the proof will be:

Start of Subproof.

$\underline{\Im(\mathbf{A}MP)=T \ \& \ \Im(\mathbf{A}SM)=T}$ Assump. for conditional proof, assuming $\Im$ is arbitrary

…

$\underline{\Im(\mathbf{A}SP)=T}$

End of subproof

$\underline{(\Im(\mathbf{A}MP)=T \ \& \ \Im(\mathbf{A}SM)=T)} \rightarrow \underline{\Im(\mathbf{A}SP)=T}$ Conditional Proof

$\forall\Im((\Im(\mathbf{A}MP)=T \ \& \ \Im(\mathbf{A}SM)=T) \rightarrow \Im(\mathbf{A}SP)=T)$ Previous line, Universal Gen., $\Im$ arbitrary.

Note that $\Im(\mathbf{A}MP)=T$, $\Im(\mathbf{A}SM)=T)$, and $\Im(\mathbf{A}SP)=T$ are each provided by a (T) schema entailed by the definition of $\Im$. These provide for each an equivalent formulation in terms of facts about the sets $\Im(S)$, $\Im(M)$ and $\Im(M)$:

$\Im(\mathbf{A}MP)=T \leftrightarrow \Im(M)\subseteq\Im(P)$

$\Im(\mathbf{A}SM)=T \leftrightarrow \Im(S)\subseteq\Im(M)$

$\Im(\mathbf{A}SP)=T \leftrightarrow \Im(S)\subseteq\Im(P)$

So, the lines $\Im(\mathbf{A}MP)=T$, $\Im(\mathbf{A}SM)=T)$, and $\Im(\mathbf{A}SP)=T$ may be replaced by equivalents: $\Im(M)\subseteq\Im(P)$, $\Im(S)\subseteq\Im(M)$, and $\Im(S)\subseteq\Im(P)$, respectively. The task then would be to prove $\Im(S)\subseteq\Im(P)$ from $\Im(M)\subseteq\Im(P)$ and $\Im(S)\subseteq\Im(M)$. It is a fact of set theory that if $B\subseteq C$ and $A\subseteq B$, then $A\subseteq C$. Hence, by appeal set theory we can go from $\Im(S)\subseteq\Im(P)$ from $\Im(M)\subseteq\Im(P)$ to $\Im(S)\subseteq\Im(M$.

Proof.

Start of Subproof.

1.	$\underline{\Im(\mathbf{A}MP)=T \ \& \Im(\mathbf{A}SM)=T}$	Assump. for conditional proof. $\Im$ arbitrary
2.	$\Im(\mathbf{A}MP)=T$	1, conjunction
3.	$\Im(\mathbf{A}MP)=T \leftrightarrow \Im(M)\subseteq\Im(P)$	TC1
4.	$\Im(M)\subseteq \Im(P)$	2,4 *modus ponens*
5.	$\Im(\mathbf{A}SM)=T$	1, conjunction
6.	$\Im(\mathbf{A}SM)=T \leftrightarrow \Im(S)\subseteq\Im(M)$	TC1
7.	$\Im(S)\subseteq \Im(M)$	3,5 *modus ponens*
8.	$\Im(S)\subseteq \Im(P)$	6,7 set theory
9.	$\Im(\mathbf{A}SP)=T \leftrightarrow \Im(S)\subseteq\Im(P)$	TC1
10.	$\underline{\Im(\mathbf{A}SP)=T}$	8 ,9 *modus ponens*

End of Subproof.

11.	$\underline{(\Im(\mathbf{A}MP)=T \ \& \Im(\mathbf{A}SM)=T)} \rightarrow \underline{\Im(\mathbf{A}SP)=T}$	1-10, conditional proof
12.	$\forall\Im(\Im(\mathbf{A}MP)=T \ \& \Im(\mathbf{A}SM)=T) \rightarrow \Im(\mathbf{A}SP)=T)$	8, universal generalization, $\Im$ arbitrary
13.	$\mathbf{A}MP, \mathbf{A}SM \models_{SL} \mathbf{A}SP$	9, definition of $\models_{SL}$

All proofs that a syllogism is valid follow this form. They differ mainly in the statement of the truth-conditions of the premises (lines 4 and 5) and the conclusion (line 6), and in the citation of relevant set theoretic truths to justifying line 6. (For several of the syllogisms the proof that the truth-condition for the conclusion are satisfied depends in part on the fact that the subject term stands for a non-empty set, and this fact must be cited in the proof. We will see an example in the next lecture.) Hence, the validity of the syllogisms turns on the fact that the categorical propositions have truth-conditions stated in set theory. Given that these conditions are met for the premises, it then follows straight from basic facts about sets, which we learned about in Part 1, that the condition for the conclusion is satisfied. Accordingly proofs that syllogisms are valid are really quite trivial logically.[4]

∗**Exercise.** Prove that he syllogism **EAE** in the first figure (called Celarent), i.e. <**E***MP***,A***SM***,E***SP*>, is valid.

There are in fact 24 valid moods. These are divided by figure and assigned traditional names, which encode information that will be explained in a later lecture. Here it is sufficient to observe that the vowel sequence within the name records the mood of the syllogism. The mood of Celaront, for example, is **EAO**.

[4] In the high Middle Ages, the central curriculum at the great universities in northern Europe, like Paris and Oxford, centered on the *trivium*, the three central subjects of the liberal arts. These consisted of grammar, logic and rhetoric. However, since students already knew grammar before they came to the university, and rhetoric was regarded as of lesser importance, it was logic that formed the core of the trivium and the main part of a student's university education. (To a lesser extent students also studied the *quadrivium*, the four remaining liberal arts – arithmetic, geometry, astronomy, and music – which at that time were all really quasi mathematical subjects.) When the humanists in Italy, where logical studies never took root, rediscovered the literature of classical Greece and Rome, they viewed (obviously quite wrongly) the logic chopping of the trivium at the northern universities as fatuous and silly, or as they put it, as "trivial." Our modern usage of *trivial* derives from this epithet applied to logic by the humanists.

The Names for the Valid Moods

First Figure:			**Third Figure:**		
	AAA	**Barbara**		**AAI**	**Darapti**
M,P	**EAE**	**Celarent**	***M,P***	**EAO**	**Felapton**
S,M	**AII**	**Darii**	***M,S***	**IAI**	**Disamis**
S,P	**EIO**	**Ferio**	***S,P***	**AII**	**Datisi**
	EAO	***Celaront**		**OAO**	**Bocardo**
	AAI	***Barbari**		**EIO**	**Ferison**

Second Figure:			**Fourth Figure:**		
	EAE	**Cesare**		**EIO**	**Fresison**
P,M	**AEE**	**Camestres**	***P,M***	**EAO**	**Fesapo**
S,M	**EIO**	**Festino**	***M,S***	**IAI**	**Dimaris**
S,P	**AOO**	**Baroco**	***S,P***	**AAI**	**Bramantip**
	AEO	***Camestrop**		**AEE**	**Camenes**
	EAO	***Cesaro**		**AEO**	***Camelop**

Those whose names are prefixed with an asterisk are the so-called the *subaltern moods*, because their validity follows from that of an earlier valid syllogism in that figure by an application of the subalternation relation to the conclusion of the earlier syllogism. Consider Celaront. It's validity follows from that of Celarent. Suppose Celarent is valid, i.e. $\mathbf{E}MP,\mathbf{A}SM \models_{SL} \mathbf{E}SP.$ But by the subalternation relation (on the Square of Opposition) we know also that $\mathbf{E}SP \models_{SL} \mathbf{O}SP.$ Hence by the fact that $\models_{SL}$ is transitive, it follows that $\mathbf{E}MP,\mathbf{A}SM \models_{SL} \mathbf{O}SP,$ i.e. that Celaront is valid.

The syllogistic names contain various coded information. Some of this will be relevant later when we investigate how the set of valid syllogisms is axiomatized.

We have seen how to prove, using the (T) schemata entailed by the definition of $\Im$ and facts of set theory, that the listed syllogisms are valid. For centuries, however, students just memorized the list. To help them, mnemonic poems were devised. These date back to the 13th century. Here is one contained in Henry Aldrich (1647–1710), *Artis Logicae Rudimenta*, which was a standard logic textbook at Oxford until the second half of the 19th century:[5]

> *Barbara, Celarent, Darii, Ferioque* prioris:
> *Cesare, Camestres, Festino, Baroco* secundae:
> Tertia, *Darapti, Disamis, Datisi, Felapton,*
> *Bocardo, Ferison,* habet; Quarta insuper addit
> *Bramantip, Camenes, Dimaris, Fesapo, Fresison.*
> Quinque *Subalterni, todidem Generalibus* orti,
> Nomen habent nullum, nec, si bene colligatur, usum.
>
> [*Barbara, Celarent, Darii, Ferioque* are of the First:
> *Cesare, Camestres, Festino, Baroco* are of the Second:
> The Third has *Darapti, Disamis, Datisi, Felapton,*
> *Bocardo, Ferison*; The Fourth adds in addition
> *Bramantip, Camenes, Dimaris, Fesapo, Fresison.*
> Fifth are the *Subalterns,* which all come from the *Universals*,
> They do not have a name, nor, if well connected, a use.]

Exercise. Memorize the poem (just kidding).[6]

Though there are only 24 valid syllogistic moods, they are sufficient for proving any valid argument that can be formulated in the syllogistic syntax, no matter how many premises that argument has, so long as it has only a finite number. Though it would be inappropriate to actually give the proof here, we will state this fact now as a metatheorem because it shows that within the context of the language of categorical propositions, syllogisms have sufficient power to explain all finite arguments.

---

[5] See, Henry Aldrich *Artis Logicae Rudimenta* (Oxford: Hammans, 1862), H. L. Mansel, ed. and commentator, p.84.

[6] When the author started teaching at the University of Cincinnati in the 1970's, he met an elderly alumna who had studied logic here in the 1920's. She had been asked to memorize a similar list.

Theorem. If $P_1,\ldots,P_n \models_{SL} Q$, then there is some finite sequence of valid syllogisms such that (1) the conclusion of the last syllogism is *Q,* and (2) each premise of any syllogism in the sequence is either in $\{P_1,\ldots,P_n\}$ or is the conclusion of a previous syllogism in the sequence.

∗LECTURE 14. CATEGORICAL LOGIC: INVALIDITY

We know by definition that the syllogism <*P,Q,R*> is valid, i.e. that

$$P,Q \models_{SL} R, \text{ if and only if } \forall\Im\ ((\Im(P)=T\ \&\ \Im(Q)=T) \rightarrow \Im(R)=T)$$

From this we know what it means for it to be *in*valid, which we write in notation as $P,Q \not\models_{SL} R$ (here the slash is really a form of negation[7]). It is invalid if there is some interpretation in which the premises are true and the conclusion is false. Since this fact underlies the methods that will be occupying us for the rest of this lecture, let us state this fact formally as a metatheorem.

Theorem

$$P,Q \not\models_{SL} R \leftrightarrow \exists\Im\ (\Im(P)=T\ \&\ \Im(Q)=T\ \&\ \Im(R)=F)$$

Any interpretation that makes an argument's premises true but its conclusion false is called a *counter-example* to the argument. Note that there is an existential quantifier in the last metatheorem. How to we prove an existentially quantified claim? By construction. That is, to show that an argument is invalid, we must construct an interpretation $\Im$ that makes the premises true and the conclusion false. Now, $\Im$ is a set of pairs <*X,A*> such that X is term and *A* is the set that *X* stands for according to $\Im$, and this set is some subset of the universe *U.* To construct $\Im$, therefore, all we have to do is specify its pairs. In the proofs below we shall show that a syllogism is invalid by first specifying a universe *U* and then singling out special subsets of *U* to serve as the referents of the terms

---

[7] That is, $P,Q \not\models_{SL} R$ means ~($P,Q \models_{SL} R$).

in the syllogism. (We will ignore terms of the language that do not occur in the syllogism under discussion because their interpretation is irrelevant to the validity of the syllogism in question.) When we specify sets, we must be sure that *U* and the subsets we specify really exist. To make sure we are dealing only with sets that exist, we will limit our constructions to sets that we have proven exist earlier in this series of lectures. In particular, we will only use small sets of natural numbers, which we have shown to exist in Part 1. These we know exist because we constructed them in set theory by appeal to the Principle of Abstraction.

Traditional Term Rules for Invalid Syllogisms

If there are 256 syllogisms and 24 or them are valid, then 232 must be invalid. We turn now turn to valid syllogisms. We shall divide them into subclasses by reference to a series of syntactic rules devised for weeding out invalid syllogisms. As we shall see, a syllogism is invalid if and only it violates at least one of these rules. Historically, the rules evolved out of some observations Aristotle himself made when he was describing syllogisms for the first in the *Prior Analytics.* In the Middle Ages they were more carefully stated and generalized, and lists of these rule became a regular part of textbooks about the syllogism – and were duly memorized by centuries of undergraduates. In the 16th century Antoine Arnauld (1612–1694) and Pierre Nicole (1625-1695) included a set of rules, which they called axioms, in their influential *Port Royal Logic.* Leibniz (1646 -1716) reformulated Arnauld and Nicole's list in more rigorous way. Though his formulations are quite rigorous from the perspective of modern logic, this and his other logical work was almost entirely unknown until the 20th century.

He proposed seven rules that were intentionally formulated in syntactic terms, and then argued that the rule set taken together was characteristic of the valid moods. [8] We shall use Leibniz's rules here.

The first two rules make use of a technical distinction between what Aristotle called *universal* and *particular* terms. In the Middle Ages these were called *distributed* and *undistributed* terms. Authors sometimes try to make the distinction semantically – it is roughly the difference between common nouns that stand for an entire set and ones that do not – but authors differ a good deal on the details. What is important, however, is that the distinction can be clearly drawn syntactically, and then it can be used to identify invalid syllogisms. Three syntactic features determine whether a term is distributed:

1. whether it occurs in a proposition as its subject or predicate,
2. whether the proposition in which it occurs is universal or particular,
3. whether the proposition in which it occurs is affirmative or negative.

Clearly these features are determined by grammar alone.

Definition of Distribution

1. A subject term is *distributed (universal)* if and only it occurs in a proposition that is universal (**A** or **E**). A subject term is *undistributed (particular)* if and only if the proposition in which it occurs is particular (**I** or **O**).
2. A predicate term is *distributed (universal)* if and only if its proposition is affirmative (**A** or **I**). A predicate term is *undistributed (particular)* if and only if its proposition is negative (**E** or **O**).

[8] See Antoine Arnauld and Pierre Nicole, Book II, *Logic or the Art of Thinking,* Jill Vance Buroker, ed.(Cambridge: Cambridge University Press, 1996). On Leibniz's formally more sophisticated

These distinctions are easily summarized in a table:

	Subject Term *S*	Predicate Term *P*
A. *All S is P*	*Distributed*	*Non- Distributed*
E. *No S is P*	*Distributed*	*Distributed*
I. *Some S is P*	*Non- Distributed*	*Non- Distributed*
O. *Some S is not P*	*Non- Distributed*	*Distributed*

We may now state the term rules.

Leibniz's Term Rules

Rule 1. Undistributed Middle. No valid syllogism has an undistributed middle term.

Rule 2. Distributed Term in the Conclusion. No syllogism is valid that has a term that is distributed in the conclusion but not in the premises.

Rule 3. Affirmative premise. No syllogism is valid that has two negative premises.

Rule 4. Negative Conclusion. No syllogism is valid that has negative conclusion without a negative premise.

Rule 5. Particular Premise. No syllogism is valid that has particular premise and a universal conclusion.

Rule 6. Negative Premise. No syllogism is valid that has a negative premise and an affirmative conclusion.

Rule 7. Universal Premise. No syllogism is valid that does not have at least one universal premise.

Leibniz demonstrated essentially the following metatheorem:

---

extension see Wolfgang Lensen "On Leibniz's Essay *Mathesis rationis,*" *Topoi* 9 (1990), 29-59).

Theorem. [9] A syllogism is valid iff it does not violate any of the rules 1-7.

Examples of Applications of the Rules

Let us now illustrate the rules. It is easy to test by inspecting each of the valid syllogisms that none of them violate any of the rules. We shall therefore now illustrate the converse, that any syllogism that violates a rule is invalid. For each rule we shall cite a syllogism that violates it and then prove by construction that the syllogism is, in fact, invalid.

Rule 1. Undistributed Middle

No valid syllogism has an undistributed middle term.

Let us illustrate by an example. Consider the syllogism **AIO** in the fourth figure, i.e. <**A***PM,***I***MS,***O***SP*>. This syllogism violates the rule because *M* is not distributed. It is also invalid, as the following Venn diagram illustrates. In the

---

[9] The rule test can be generalized to categorical arguments of any finite number of premises. To do so precisely, it is necessary to introduce distinction that allows us to ignore those premises listed but not actually used to justify the conclusion.

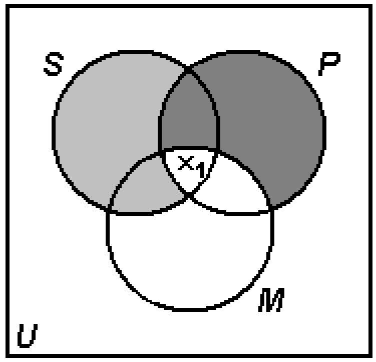
Definition. The argument from $P_1,\ldots,P_n$ to Q is *minimally valid* (briefly, $P_1,\ldots,P_n \models^{min}_{SL} Q$) iff there is some finite sequence of valid syllogisms such that (1) the conclusion of the last syllogism is Q, and (2) each premise of any syllogism in the sequence is either in $\{P_1,\ldots,P_n\}$ or is the conclusion of a previous syllogism in the sequence.

By this definition every proposition listed as a premise in a minimally valid argument is actually used to justify the conclusion by syllogistic reasoning. In order to apply Rule 2 to the general case, we also define a *generalize middle term* as any term that occurs in a premise but not in the conclusion. In the metatheorem below Rule 2 is to understood as applying to generalized terms in this sense.

Theorem. $P_1,\ldots,P_n \models^{min}_{SL} Q$ iff the argument from $P_1,\ldots,P_n$ to Q does not violated any of the rules 1-7.

Though proof is not difficult (an induction on the length of the series of the syllogisms associated with the argument), it requires techniques more advanced that we are using in these lectures.

diagram the major premise is true because the extension of *M* is a non-empty subset of that of *P.* Moreover the minor premise is true because the intersection of the extension of *S* and *M* is non-empty. But the conclusion is false because the region that *S* is true of but *P* is not is empty.



We now illustrate how to prove that the syllogism is invalid.

Theorem. The syllogism **AIO** in the fourth figure, i.e. <**A***PM*,**I***MS*,**O***SP*>, is invalid.

Analysis of the Proof. First note what it is that must be proven: <**A***PM*,**I***MS*,**O***SP*> is invalid, or in alternative notation

$$\mathbf{A}PM,\mathbf{I}MS \nvDash_{SL} \mathbf{O}SP$$

By the definition of $\nvDash_{SL}$, this means:

$$\exists\Im(\Im(\mathbf{A}PM)=T \;\&\; \Im(\mathbf{I}MS)=T)\&\; \Im(\mathbf{O}SP)=F).$$

How do you show an existential proposition? Normally you do so by construction. An interpretation is a set – it is a set of pairs in which each term is paired with a set and each sentence with a truth-value. We know from the Principle of Abstraction that all we need do to insure a set exists is define it. Hence the task is to define an $\Im$ such that $\Im(\mathbf{A}PM)=T$ & $\Im(\mathbf{I}MS)=T)$& $\Im(\mathbf{O}SP)=F$. But (T) schemata entailed by the definition of $\Im$ tells us what these three facts means:

$$\Im(\mathbf{A}PM)=T \text{ iff } \Im(P)\subseteq\Im(M),$$
$$\Im(\mathbf{I}MS)=T \text{ iff } \Im(M)\cap\Im(S)\neq\varnothing,$$
$$\Im(\mathbf{O}SP)=F \text{ iff } \Im(S)-\Im(P)=\varnothing.$$

Hence, what we want to insure is that $\Im(P)\subseteq\Im(M)$, $\Im(M)\cap\Im(S)\neq\varnothing$, and $\Im(S)-\Im(P)=\varnothing$. Taking our clue from the Venn diagram above, we assign sets to $\Im(S)$, $\Im(P)$, $\Im(M)$ to fit the diagram. Let us make up our universe out of numbers because we know that numbers exist (we constructed them earlier). Let $U=\{1\}$, $\Im(S)=\{1\}$, $\Im(P)=\{1\}$, $\Im(M)=\{1\}$. It will then turn out that $\Im(P)\subseteq\Im(M)$, $\Im(M)\cap\Im(S)\neq\varnothing$, and $\Im(S)-\Im(P)=\varnothing$ because, respectively, $\{1\}\subseteq\{1\}$, $\{1\}\cap\{1\}\neq\varnothing$, and $\{1\}-\{1\}=\varnothing$ are all facts of set theory, as you may check for yourself.

Proof. Let us define $\Im$ (as set of pairs) as follows:

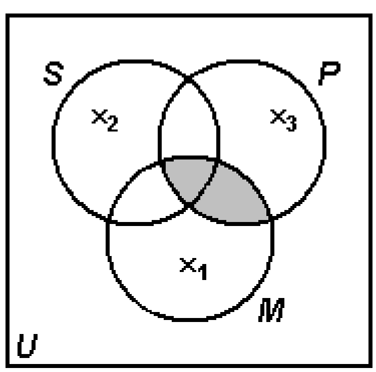
1.	$U=\{1\}$, $\Im(S)=\{1\}$, $\Im(P)=\{1\}$, $\Im(M)=\{1\}$	Set theory
2.	$\{1\}\subseteq\{1\}$	Set theory
3.	$\Im(P)\subseteq\Im(M)$	1 and 2, sub of =

4.	$\Im(\mathbf{A}PM)$=T iff $\Im(P)\subseteq\Im(M)$	TC1
5.	$\Im(\mathbf{A}PM)$=T	4,5 *modus ponens*
6.	$\{1\}\cap\{1\}=\{1\}\neq\varnothing$	Set theory
7.	$\Im(M)\cap\Im(S)\neq\varnothing$	1 and 6, sub of =
8.	$\Im(\mathbf{I}MS)$=T iff $\Im(M)\cap\Im(S)\neq\varnothing$	TC3
9.	$\Im(\mathbf{I}MS)$=T	7,8 *modus ponens*
10.	$\{1\}-\{1\}=\varnothing$	Set theory
11.	$\Im(S)\cap\Im(P)=\varnothing$	1,10 sub of =
12.	$\Im(\mathbf{O}SP)$=F iff $\Im(S)-\Im(P)=\varnothing$.	TC4
13.	$\Im(\mathbf{O}SP)$=F	11,12 *modus ponens*
14.	$\Im(\mathbf{A}PM)$=T & $\Im(\mathbf{I}MS)$=T & $\Im(\mathbf{O}SP)$=F	5, 9, and 13, conjunction
15.	$\exists\Im(\Im(\mathbf{A}PM)$=T & $\Im(\mathbf{I}MS)$=T & $\Im(\mathbf{O}SP)$=F)	14, construction
16.	$\mathbf{A}PM$, $\mathbf{I}MS \nvDash_{SL} \mathbf{O}SP$	15, def of $\nvDash_{SL}$

Rule 2. Distributed Term in the Conclusion

> No valid syllogism has a term that is distributed in the conclusion but not in the premises.

Consider the syllogism **EOA** in the first figure, i.e. <**E***MP*,**O***SM*,**A***SP*>. This syllogism violates the rule because *S* is distributed in the conclusion but not in the minor premise. In the diagram below *No M is P* is true because their intersection is empty. Moreover, *There is an S that is not a M* is true because x_2 is in the right region. However, the conclusion *Every S is P* is false because there is an object that *S* is true of but not *P*, namely x_2.



Theorem. The syllogism **EOA** in the fourth figure, i.e. <**E***MP,***O***SM,***A***SP*>, is invalid.

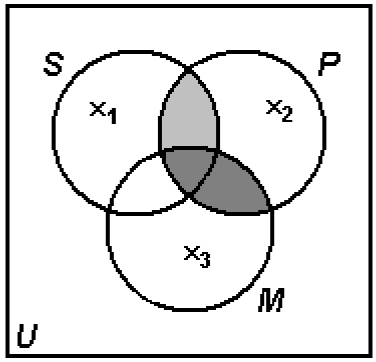
Proof. Let us define $\Im$ (as set of pairs) as follows:

1.	$U=\{1,2,3\}$, $\Im(S)=\{2\}$, $\Im(P)=\{3\}$, $\Im(M)=\{1\}$	Set theory
2.	$\{3\}\cap\{1\}\neq\varnothing$	Set theory
3.	$\Im(P)\cap\Im(M)\neq\varnothing$	1,2 sub of =
4.	$\Im(\mathbf{E}MP)=T$ iff $\Im(P)\cap\Im(M)\neq\varnothing$	TC2
5.	$\Im(\mathbf{E}MP)=T$	3,4 *modus ponens*
6.	$\{1\}-\{2\}=\{1\}\neq\varnothing$	Set Theory
7.	$\Im(M)-\Im(S)\neq\varnothing$	1,6 sub of =
8.	$\Im(\mathbf{O}SM)=T$ iff $\Im(M)-\Im(S)\neq\varnothing$	TC4
9.	$\Im(\mathbf{O}SM)=T$	7,8 *modus ponens*
10.	$\sim(\{2\}\subseteq\{3\})$	Set theory
11.	$\sim(\Im(S)\subseteq\Im(P))$	1,10, sub of =
12.	$\Im(\mathbf{A}SP)=T$ iff $\Im(S)\subseteq\Im(P))$	TC1
13.	$\sim\Im(\mathbf{A}SP)-T$	11,12 *modus tollens*
14.	$\Im(\mathbf{A}SP)=F$	13, bivalence of $\Im$
15.	$\Im(\mathbf{E}PM)=T$ & $\Im(\mathbf{O}SM)=T$ & $\Im(\mathbf{A}SP)=F$	5, 9, and 14, conjunction
16.	$\exists\Im(\Im(\mathbf{E}PM)=T$ & $\Im(\mathbf{O}SM)=T$ & $\Im(\mathbf{A}SP)=F)$	15, construction
17.	$\mathbf{E}PM, \mathbf{O}SM \nvDash_{SL} \mathbf{A}SP$	17, def of $\nvDash_{SL}$

Rule 3. Affirmative premise

Every valid syllogism has at least one affirmative premise.

Consider the syllogism **EOI** in the first figure, i.e. <**E***MP,***O***SM,***I***SP*>. This syllogism violates Rule 3 because both premises are negative In the diagram below *No M is P* is true because the intersection of the extensions of *M,* which is non-empty, and *P* is empty. Moreover, *Some S is not M* is true because x_1 is in the right region. However, the conclusion *Some S is P* is false because the intersection of the extensions of *S* and *P* is empty.



Theorem. The syllogism **EOI** in the fourth figure, i.e. <**E***MP*,**O***SM*,**I***SP*>, is invalid.

Proof. Let us define $\Im$ (as set of pairs) as follows:

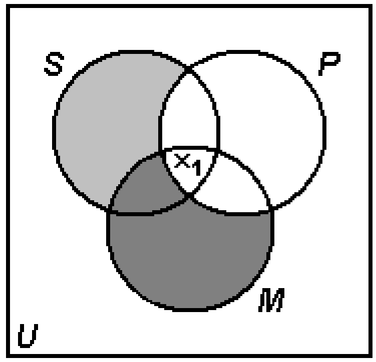
1.	$U=\{1,2,3\}$, $\Im(S)=\{1\}$, $\Im(P)=\{2\}$, $\Im(M)=\{3\}$	Set theory
2.	$\{2\}\cap\{3\}=\varnothing$	Set theory
3.	$\Im(P)\cap\Im(M)=\varnothing$	1,2 sub of =
4.	$\Im(\mathbf{E}MP)=T$ iff $\Im(P)\cap\Im(M)=\varnothing$	TC2
5.	$\Im(\mathbf{E}MP)=T$	3,4 *modus ponens*
6.	$\{3\}-\{1\}=\{3\}\neq\varnothing$	Set theory
7.	$\Im(M)-\Im(S)\neq\varnothing$	1 and 5, sub of =
8.	$\Im(\mathbf{O}SM)=T$ iff $\Im(M)-\Im(S)\neq\varnothing$	TC4
9.	$\Im(\mathbf{O}SM)=T$	7,8 *modus ponens*
10.	$\{1\}\cap\{2\}=\varnothing$	Set theory
11.	$\Im(S)\cap\Im(P)=\varnothing$	1,8 sub of =
12.	$\Im(\mathbf{I}SP)=T$ iff $\sim[\Im(S)\cap\Im(P)=\varnothing]$	TC3
13.	$\sim\Im(\mathbf{I}SP)=T$	1,12 *modus tollens*
14.	$\Im(\mathbf{I}SP)=F$	13, bivalence of $\Im$
15.	$\Im(\mathbf{E}PM)=T$ & $\Im(\mathbf{O}SM)=T$ & $\Im(\mathbf{I}SP)=F$	5, 9, and 14, conjunction
16.	$\exists\Im(\Im(\mathbf{E}PM)=T$ & $\Im(\mathbf{O}SM)=T$ & $\Im(\mathbf{I}SP)=F)$	15, construction
17.	$\mathbf{E}PM, \mathbf{O}SM \nvDash_{SL} \mathbf{I}SP$	16, def of $\nvDash_{SL}$

Rule 4. Negative Conclusion

No valid syllogism has negative conclusion without a negative premise.

Consider the syllogism **AIO** in the first figure, i.e. <**A***MP*,**I***SM*,**O***SP*>. This syllogism violates Rule 4 because the conclusion is negative but both premises are affirmative. In the diagram below *Every M is P* is true because the extension of *M* is a non-empty subset of that of *P.* Moreover, *Some S is M* is true because

x_1 is in the right region. However, the conclusion *Some S is not P* is false because the extension of *S* is outside of that of *P* is empty.



Theorem. The syllogism **AIO** in the fourth figure, i.e. <**A***MP*,**I***SM*,**O***SP*>, is invalid.

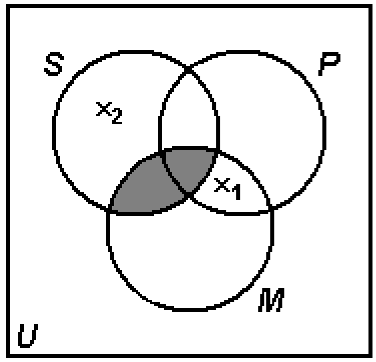
Proof. Let us define $\Im$ (as set of pairs) as follows:

1.	$U=\{1\}, \Im(S)=\{1\}, \Im(P)=\{1\}, \Im(M)=\{1\}$	Set theory
2.	$\{1\}\subseteq\{1\}$	Set theory
3.	$\Im(P)\subseteq\Im(M)$	1 and 2, sub of =
4.	$\Im(\mathbf{A}MP)=T$ iff $\Im(P)\subseteq\Im(M)$	TC1
5.	$\Im(\mathbf{A}MP)=T$	3,4 *modus ponens*
6.	$\{1\}\cap\{1\}=\{1\}\neq\emptyset$	Set theory
7.	$\Im(M)\cap\Im(S)\neq\emptyset$	1,6 sub of =
8.	$\Im(\mathbf{I}SM)=T$ iff $\Im(M)\cap\Im(S)\neq\emptyset$	TC3
9.	$\Im(\mathbf{I}SM)=T$	7,8 *modus ponens*
10.	$\{1\}-\{1\}=\emptyset$	Set theory
11.	$\Im(S)\cap\Im(P)=\emptyset$	1,10 sub of =
12.	$\Im(\mathbf{O}SP)=T$ iff $\sim[\Im(S)\cap\Im(P)=\emptyset]$	TC4
13.	$\sim\Im(\mathbf{O}SP)=T$	11,12 *modus tollens*
14.	$\Im(\mathbf{O}SP)=F$	13, bivalence of $\Im$
15.	$\Im(\mathbf{A}PM)=T$ & $\Im(\mathbf{I}SM)=T$ & $\Im(\mathbf{O}SP)=F$	5, 9, and 13, conjunction
16.	$\exists\Im(\Im(\mathbf{A}PM)=T$ & $\Im(\mathbf{I}SM)=T$ & $\Im(\mathbf{O}SP)=F)$	16, construction
17.	$\mathbf{A}PM, \mathbf{I}SM \not\models_{SL} \mathbf{O}SP$	17, def of $\not\models_{SL}$

Rule 5. Particular Premise

No valid syllogism has particular premise and a universal conclusion.

Consider the syllogism **IEA** in the first figure, i.e. <**I***MP*,**E***SM*,**A***SP*>. This syllogism violates Rule 5 because the conclusion is universal but one of the premises is particular. In the diagram below *No S is M* is true because the intersection of the extension of *M*, which is non-empty, with that of *S* is empty. Moreover, *Some M is P* is true because x_1 is in the right region. However, the conclusion *Every S is P* is false because the extension of *S* outside of that of *P* is non-empty.



Theorem. The syllogism **IEA** in the fourth figure, i.e. <**I***MP*,**E***SM*,**A***SP*>, is invalid.

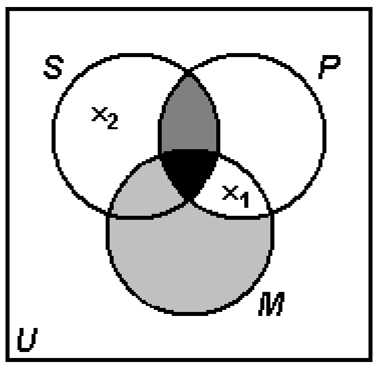
Proof. Let us define $\Im$ (as set of pairs) as follows:

1.	$U=\{1,2\}$, $\Im(S)=\{2\}$, $\Im(P)=\{1\}$, $\Im(M)=\{1\}$	Set theory
2.	$\{1\}\cap\{1\}=\{1\}\neq\varnothing$	Set theory
3.	$\Im(M)\cap\Im(P)=\{1\}$	1,2 sub of =
4.	$\Im(M)\cap\Im(P)\neq\varnothing$	3, set theory
5.	$\Im(\mathbf{I}MP)=T$ iff $\Im(M)\cap\Im(P)\neq\varnothing$	TC3
6.	$\Im(\mathbf{I}MP)=T$	4,5 *modus ponens*
7.	$\{1\}\cap\{2\}=\varnothing$	Set theory
8.	$\Im(M)\cap\Im(S)=\varnothing$	1.7 sub of =
9.	$\Im(\mathbf{E}SM)=T$ iff $\Im(M)\cap\Im(S)=\varnothing$	TC2
10.	$\Im(\mathbf{E}SM)=T$	8,9 *modus ponens*
11.	$\sim(\{2\}\subseteq\{1\})=\varnothing$	Set theory
12.	$\sim(\Im(S)\subseteq\Im(P))$	11, sub of =
13.	$\Im(\mathbf{A}SP)=T$ iff $\Im(S)\subseteq\Im(P)$	TC1
14.	$\sim\Im(\mathbf{A}SP)=T$	12,13 *modus tollens*
15.	$\Im(\mathbf{A}SP)=F$	14, bivalence of $\Im$
16.	$\Im(\mathbf{I}PM)=T$ & $\Im(\mathbf{E}SM)=T$ & $\Im(\mathbf{A}SP)=F$	6, 10, and 14, conjunction
17.	$\exists\Im(\Im(\mathbf{I}PM)=T$ & $\Im(\mathbf{E}SM)=T$ & $\Im(\mathbf{O}SP)=F)$	16, construction
18.	$\mathbf{I}PM$, $\mathbf{E}SM \nvDash_{SL} \mathbf{A}SP$	17, def of $\nvDash_{SL}$

Rule 6. Negative Premise

No valid syllogism with a negative premise, has an affirmative conclusion.

Consider the syllogism **AEI** in the first figure, i.e. <**A***MP,***E***SM,***I***SP*>. This syllogism violates Rule 6 because the conclusion is universal but one of the premises is particular. In the diagram below *Every M is P* is true because the intersection of the extension of *M*, which is non-empty, is a subset of that of *P.* Moreover, *No S is M* is true because the intersection of the extension of *S,* which is non-empty, with that of *M* is empty. However, the conclusion *Some S is P* is false because the intersection of the extensions of *S* and *P* is empty.



Theorem. The syllogism **AEI** in the fourth figure, i.e. <**A***MP,***E***SM,***I***SP*>, is invalid.

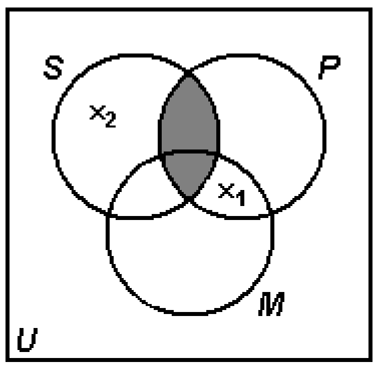
Proof. Let us define $\Im$ (as set of pairs) as follows:

1.	$U=\{1,2\}$, $\Im(S)=\{2\}$, $\Im(P)=\{1\}$, $\Im(M)=\{1\}$	Set theory
2.	$\{1\}\subseteq\{1\}=\{1\}$	Set theory
3.	$\Im(M)\subseteq\Im(P)$	1,2 sub of =
4.	$\Im(\mathbf{A}MP)=T$ iff $\Im(M)\subseteq\Im(P)$	TC1
5.	$\Im(\mathbf{A}MP)=T$	3,4 *modus ponens*
6.	$\{1\}\cap\{2\}=\varnothing$	Set theory
7.	$\Im(M)\cap\Im(S)=\varnothing$	1,6 sub of =
8.	$\Im(\mathbf{E}SM)=T$ iff $\Im(M)\cap\Im(S)=\varnothing$	TC2
9.	$\Im(\mathbf{E}SM)=T$	7,8 *modus ponens*
10.	$\{2\}\cap\{1\}=\varnothing$	Set theory
11.	$\Im(S)\cap\Im(P)=\varnothing$	1,10 sub of =
12.	$\Im(\mathbf{I}SP)=T$ iff $\sim[\Im(S)\cap\Im(P)=\varnothing]$	TC3
13.	$\sim\Im(\mathbf{I}SP)=T$	11,12 *modus tollens*
14.	$\Im(\mathbf{I}SP)=F$	13, bivalence of $\Im$
15.	$\Im(\mathbf{A}PM)=T$ & $\Im(\mathbf{E}SM)=T$ & $\Im(\mathbf{I}SP)=F$	4, 9, and 14, conjunction
16.	$\exists\Im(\Im(\mathbf{A}PM)=T$ & $\Im(\mathbf{E}SM)=T$ & $\Im(\mathbf{I}SP)=F)$	15, construction
17.	$\mathbf{A}PM, \mathbf{E}SM \nvDash_{SL} \mathbf{I}SP$	16, def of $\nvDash_{SL}$

Rule 7. Universal Premise

No valid syllogism with at least one universal premise.

Consider the syllogism **IOI** in the first figure, i.e. <**I***MP*,**O***SM*,**I***SP*>. This syllogism violates Rule 7 because the conclusion is universal but one of the premises is particular. In the diagram below *Every M is P* is true because the intersection of the extension of *M*, which is non-empty, is a subset of that of *P.* Moreover, *No S is M* is true because the intersection of the extension of *S,* which is non-empty, with that of *M* is empty. However, the conclusion *Some S is P* is false because the intersection of the extensions of *S* and *P* is empty.



Theorem. The syllogism **IOI** in the fourth figure, i.e. <**I***MP*,**O***SM*,**I***SP*>, is invalid.

Proof. Let us define $\Im$ (as set of pairs) as follows:

1.	$U=\{1,2\}$, $\Im(S)=\{2\}$, $\Im(P)=\{1\}$, $\Im(M)=\{1\}$	Set theory
2.	$\{1\}\cap\{1\}=\{1\}$	Set theory
3.	$\Im(M)\cap\Im(P)\neq\varnothing$	1,2 sub of =
4.	$\Im(\mathbf{I}MP)$=T iff $\Im(M)\cap\Im(P)\neq\varnothing$	TC3
5.	$\Im(\mathbf{I}MP)$=T	2,3 *modus ponens*
6.	$\{1\}\cap\{2\}=\{1\}\neq\varnothing$	Set theory
7.	$\Im(M)-\Im(S)\neq\varnothing$	1,6 sub of =
8.	$\Im(\mathbf{O}SM)$=T iff $\Im(M)-\Im(S)\neq\varnothing$	TC4
9.	$\Im(\mathbf{O}SM)$=T	7,8 *modus ponens*
10.	$\{2\}\cap\{1\}=\varnothing$	Set theory
11.	$\Im(S)\cap\Im(P)=\varnothing$	1,10 sub of =
12.	$\Im(\mathbf{I}SP)$=T iff $\sim[\Im(S)\cap\Im(P)=\varnothing]$	TC3
13.	$\sim\Im(\mathbf{I}SP)$=T	11,12 *modus tollens*
14.	$\Im(\mathbf{I}SP)$=F	13, bivalence of $\Im$
15.	$\Im(\mathbf{I}PM)$=T & $\Im(\mathbf{O}SM)$=T & $\Im(\mathbf{I}SP)$=F	5, 9, and 14, conjunction

16. $\exists\Im(\Im(\mathbf{I}PM)$=T & $\Im(\mathbf{O}SM)$=T & $\Im(\mathbf{I}SP)$=F) 15, construction
17. **I***PM*, **O***SM* $\nvDash_{SL}$ **I***SP* 17, def of $\nvDash_{SL}$

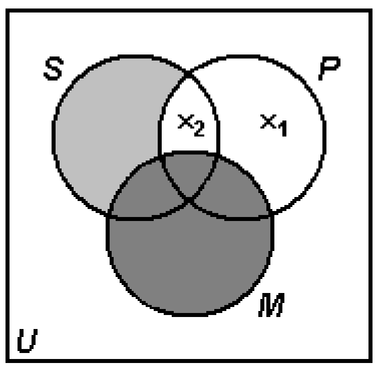
Exercise. The two following syllogisms are invalid. For each,

- name which of the seven syllogistic rules it violates,
- draw a Venn diagram illustrating that its premises are true but its conclusion false in that universe, and
- ∗give a proof like those in the previous examples that the syllogism is invalid:
  1. **AIE** in the fourth figure
  2. **IOA** in the second figure

∗Syllogism with Empty Subject Terms

Let us conclude the topic of invalid syllogisms by discussing those special cases that are valid according to Aristotle's definitions, but which modern logicians would reject as invalid. These are the syllogisms that are valid because Aristotle builds into the truth-conditions for universal propositions a condition that modern logicians reject, namely that their subjects terms that stand for non-empty sets. If unlike Aristotle, but like modern logic, we allowed universal affirmatives and negatives to be true when their subject extensions were empty, then some of the syllogisms that Aristotle accepts as valid would turn out to be invalid. These are named in red in the earlier list of valid syllogisms.

Let us consider the case of Felapton: <**E***MP,***A***MS,***O***SP*>. Suppose that the premise **E***MP* (*No M is P*) is true because $\Im(M)\cap\Im(P)=\varnothing$ even when its subject term *M* stands for the empty set, i.e. even when $\Im(M)=\varnothing$. Now consider the set $\Im(P)$. It does not matter what set this is, but for concreteness in the Venn diagram below we set it equal to {1}. Now let us give some more detail about $\Im$ so that we make Felapton's second premise *Ever M is S* also true. Let us make $\Im(S)$={2} and $\Im(P)$={1,2}. Since $\varnothing\subseteq$\{1\}, it follows that $\Im(M)\subseteq\Im(S)$, and hence by the truth-conditions for *Every M is S*, as amended to allow for *M* to be empty, we know $\Im$(**A***MS*)=T. Now, {2}–{1}=∅. Hence, $\Im(S)-\Im(P)=\varnothing$. Thus, by the truth-conditions for *Some S is not P,* the conclusion turns out to be false: $\Im$(**O***SP*)=F. Thus in the interpretation $\Im$, the premises of Felapton are true but its conclusion is false. Thus, if we allow the subject of a true universal affirmative to be non-empty, one of Aristotle's traditionally valid syllogisms turns out invalid.



Felapton however is genuinely valid given Aristotle's assumptions.

Theorem. The syllogism **EAO** in the first figure (Felapton), i.e. <**E***MP*,**A***MS*,**O***SP*>, is valid.

Proof.
Start of Subproof.

1.	$\Im$(**E***MP*)=T & $\Im$(**A***SM*)=T	Assump. for conditional proof. $\Im$ arbitrary
2.	$\Im$(**E***MP*)=T	1, conjunction
3.	$\Im$(**E***MP*)=T iff $\Im(M)\cap\Im(P)=\varnothing$	TC2
4.	$\Im(M)\cap\Im(P)=\varnothing$	2,4 *modus ponens*
5.	$\Im$(**A***MS*)=T	1, conjunction
6.	$\Im$(**A***MS*)=T iff $\Im(M)\subseteq\Im(S)$	TC1
7.	$\Im(M)\subseteq\Im(S)$	3,5 *modus ponens*
8.	$\Im(M)\neq\varnothing$	definition of $\Im$
9.	$\Im(S)-\Im(P)\neq\varnothing$	6, 7 and 8, set theory

10. $\Im(\mathbf{O}SP)=T$ iff $\Im(S)-\Im(P)\neq\varnothing$ TC4
11. $\underline{\Im(\mathbf{O}SP)=T}$ 9,10 *modus ponens*

End of Subproof.

12. $\underline{(\Im(\mathbf{E}MP)=T \;\&\; \Im(\mathbf{A}MS)=T)} \rightarrow \underline{\Im(\mathbf{O}SP)=T}$ 1-11, conditional proof
13. $\forall\Im(\Im(\mathbf{E}MP)=T \;\&\Im(\mathbf{A}MS)=T)\rightarrow \Im(\mathbf{O}SP)=T)$ 12, universal generalization, $\Im$ arbitrary
14. $\mathbf{E}MP,\mathbf{A}MS \models_{SL} \mathbf{O}SP$ 3, definition of $\models_{SL}$

∗**Exercise:** Prove that he syllogism **AAI** in the fourth figure (Bramantip), i.e. <**A***PM,***A***MS,***I***SP*>, is valid.

Summary

In this lecture we have studied the syllogism. Logically syllogisms are as dated as their mediaeval background, but as an exercise in the introduction to logic they provide an excellent example of a little logical system at work.

First they show how logical arguments are a matter of form. To explain what a syllogism is, and which are logical and which are not – which is the goal of logic – we had to do grammar. To philosophers or mathematicians, grammar is of little interest in itself. But before we could identity the valid moods, we had to first distinguish the four syllogistic propositions, major, minor and middle terms, the figures and moods, which are all concepts in grammar.

For a large part of the lecture we rubbed out noses in what it is to “prove” that an argument is valid or invalid. A proof of validity almost always consists of conditional proof in which it is assumed that an arbitrary possible interpretation makes the premises of an argument true, and then it is shown that the conclusion is true by a appeal to (T) schemata, which are entailed by the definition of $\Im$, and to facts of set theory. A proof of invalidity is almost always a construction in which an interpretation is defined for the terms that occur in the argument. You state what sets the terms stand for, and then show by appeal to (T) schemata and set theory that the premises of the argument are true but the conclusion is false.

In later lectures we shall meet more complicated languages. The exposition used for the syllogistic will serve as the model: (1) it starts with the definitions of the syntax including that of sentence, (2) it then defines the notion of an interpretation $\Im$ appropriate to the syntax that explains how $\Im$ is a function that maps expressions to their “meanings”, and (3) it finishes with the definition of validity, and proofs that various arguments are valid or invalid, using proof strategies very similar to those used above.

LECTURE 15. PROPOSITIONAL AND FIRST-ORDER LOGIC: VALIDITY

Propositional Logic

The Truth-Table Test for Validity

We now begin our investigation of valid arguments in modern symbolic logic. In this lecture we start with propositional logic. Recall that the basic concepts of logic are *validity, invalidity, consistency, logical equivalence,* and *logical truth*. It is helpful to repeat here their definitions, relative to a language L:

Definitions

$\{P_1,\ldots,P_n\} \models_L Q$ iff $\forall\Im\ ((\Im(P_1)=\text{T}\&\ldots\&\Im(P_n)=\text{T}) \rightarrow \Im(Q)=\text{T})$

$\{P_1,\ldots,P_n\} \not\models_L Q$ iff $\exists\Im\ (\Im(P_1)=\text{T}\&\ldots\&\Im(P_n)=\text{T}\&\ \Im(Q)=\text{F})$

P is a *logical truth* in L (in symbols $\models_L P$) iff $\forall\Im\ (\Im(P)=\text{T})$

$\{P_1,\ldots,P_n\}$ is *consistent* in L iff $\exists\Im\ (\Im(P_1)=\text{T}\&\ldots\&\Im(P_n)=\text{T})$

(We abbreviate $\{P_1,\ldots,P_n\} \models_L Q$ as $P_1,\ldots,P_n \models_L Q$.) In the propositional logic a logical truth is called a *tautology.*

Identifying logical properties in propositional logic is greatly simplified by the use of truth-tables. Consider validity. First we need a tool. Let us call $(P_1\wedge\ldots\wedge P_n)\rightarrow Q$ *the conditional corresponding to* the argument from $P_1,\ldots,P_n$ to Q. Thus we make up the conditional corresponding to the argument $P_1,\ldots,P_n$ to Q by conjoining all its premises as conjunctions in the conditional's antecedent and using its conclusion as the conditional's consequent. All we need to do to check whether the argument from $P_1,\ldots,P_n$ to Q is valid is do a truth-table for $(P_1\wedge\ldots\wedge P_n)\rightarrow Q$. If $(P_1\wedge\ldots\wedge P_n)\rightarrow Q$ is a tautology, the argument is valid; if it is not a tautology, the argument is invalid. The test works because the circumstances that make an argument valid (the is no case in which $P_1,\ldots,P_n$ are all T and Q is F) are the very circumstances that make $(P_1\wedge\ldots\wedge P_n)\rightarrow Q$ a tautology.

Theorem. $\{P_1,\ldots,P_n\}\models_{PL}Q$ iff $(P_1\wedge\ldots\wedge P_n)\rightarrow Q$ is a tautology.

Proof. The following are all equivalent by definitions:

$\{P_1,\ldots,P_n\}\models_{PL}Q$ iff $\forall\Im$ (if $\Im(P_1)$=T&…&$\Im(P_n)$=T&…) then $\Im(Q)$=T)

iff $\forall\Im$ (if $\Im(P_1\wedge\ldots\wedge P_n)$=T) then $\Im(Q)$=T)

iff $\forall\Im$ (if $\Im(P_1\wedge\ldots\wedge P_n)\rightarrow Q)$=T)

iff $(P_1\wedge\ldots\wedge P_n)\rightarrow Q$ is a tautology

Below we give two examples. The truth-values for the premises and conclusion of the argument are colored blue, and those for the corresponding conditional formed from the argument are colored yellow. In these cases, the yellow values are all T, and therefore the argument is valid.

Examples

Theorem. Disjunctive Syllogism in valid in propositional logic: $\{p_1\vee p_2, {\sim}p_1\}\models p_2$.

Proof. Let us construct a truth-table for the corresponding conditional:

	p_1	p_2	((p_1	$\vee$	p_2)	$\wedge$	~	p_1)	$\rightarrow$	p_2)
$\Im_1$	T	T	T	T	T	F	F	T	T	T
$\Im_2$	T	F	T	T	F	F	F	T	T	F
$\Im_3$	F	T	F	T	T	T	T	F	T	T
$\Im_4$	F	F	F	F	F	F	T	F	T	F

From the truth-table we can summarize the sentence's truth-conditions: for any $\Im$,

$\Im(((p_1\vee p_2)\wedge{\sim}p_1)\rightarrow p_2)$=T iff (($\Im(p_1)$=T and $\Im(p_2)$=T) or

($\Im(p_1)$=T and $\Im(p_2)$=F) or

($\Im(p_1)$=F and $\Im(p_2)$=T) or

($\Im(p_1)$=F and $\Im(p_2)$=F))

That is, $\Im(((p_1 \vee p_2) \wedge \sim p_1) \rightarrow p_2)$=T holds in any $\Im$. Hence, by the previous metatheorem, $\{p_1 \vee p_2, \sim p_1\} \models p_2$.

Theorem. Contraposition is valid in propositional logic: $\{p_1 \rightarrow p_2\} \models \sim p_2 \rightarrow \sim p_1$.

Proof

	p_1	p_2	$((p_1$	$\rightarrow$	$p_2)$	$\rightarrow$	$(\sim$	p_2	$\rightarrow$	$\sim$	$p_1)$
$\Im_1$	T	T	T	T	T	T	F	T	T	F	T
$\Im_2$	T	F	T	F	F	T	T	F	F	F	T
$\Im_3$	F	T	F	T	T	T	F	T	T	T	F
$\Im_4$	F	F	F	T	F	T	T	F	T	T	F

We may summarize these facts as follows: for any $\Im$,

$\Im(((p_1 \rightarrow p_2) \rightarrow (\sim p_2 \rightarrow \sim p_1))$=T iff (($\Im(p_1)$=T and $\Im(p_2)$=T) or
($\Im(p_1)$=T and $\Im(p_2)$=F) or
($\Im(p_1)$=F and $\Im(p_2)$=T) or
($\Im(p_1)$=F and $\Im(p_2)$=F))

That is, $\Im(((p_1 \rightarrow p_2) \rightarrow (\sim p_2 \rightarrow \sim p_1))$=T for any $\Im$. Hence, by the earlier metatheorem, $\{p_1 \rightarrow p_2\} \models \sim p_2 \rightarrow \sim p_1$.

Exercise

Show *modus tollens* is valid in propositional logic: $\{p_1 \rightarrow p_2, \sim p_2\} \models \sim p_1$.

	p_1	p_2	$((p_1$	$\rightarrow$	$p_2)$	$\wedge$	$\sim$	$p_2)$	$\rightarrow$	$\sim$	$p_1)$
$\Im_1$	T	T									
$\Im_2$	T	F									
$\Im_3$	F	T									
$\Im_4$	F	F									

Determine when $\Im(((p_1 \rightarrow p_2) \wedge \sim p_2) \rightarrow \sim p_1)$=T.

Proving Invalidity by Truth-Tables

Essentially the same technique may be used to show an argument is invalid. If the conditional corresponding to an argument is not a tautology, then there is some case in which it is false, i.e. there is an interpretation in which all the premise are true and the conclusion false. If there is one, it is invalid.

Theorem. $\{P_1,\ldots,P_n\} \not\models_L Q$ iff $(P_1 \wedge \ldots \wedge P_n) \rightarrow_L Q$ is not a tautology.

In the example below we show how to use this equivalence.

Theorem. Denying the antecedent is invalid: $\{p_1 \rightarrow p_2, \sim p_1\} \not\models \sim p_1$.

Proof

	p_1	p_2	$((p_1$	$\rightarrow$	$p_2)$	$\wedge$	$\sim$	$p_1)$	$\rightarrow$	$\sim$	$p_2)$
$\Im_1$	T	T	T	T	T	F	F	T	T	F	T
$\Im_2$	T	F	T	F	F	F	F	T	T	T	F
$\Im_3$	F	T	F	T	T	T	T	F	F	F	T
$\Im_4$	F	F	F	F	F	F	T	F	T	T	F

. We may summarize these facts as follows: for any $\Im$,

$\Im(((p_1 \rightarrow p_2) \wedge \sim p_1) \rightarrow \sim p_2))=T$ iff $((\Im(p_1)=T$ and $\Im(p_2)=T)$ or
$(\Im(p_1)=T$ and $\Im(p_2)=F)$ or
$(\Im(p_1)=F$ and $\Im(p_2)=F))$

Also, for any $\Im$,

$\Im(((p_1 \rightarrow p_2) \wedge \sim p_1) \rightarrow \sim p_2))=F$ iff $(\Im(p_1)=F$ and $\Im(p_2)=T)$

Hence define $\Im(p_1)=F$ and $\Im(p_2)=T$. Clearly such an $\Im$ exists (by construction) because we can define it. Then $\Im((p_1 \rightarrow p_2) \wedge \sim p_1) \rightarrow \sim p_2))=F$. Hence $\Im((p_1 \rightarrow p_2)=T$ and $\Im(\sim p_1)=T$ and $\Im(\sim p_2)=F$. Hence, $\exists \Im$, $\Im((p_1 \rightarrow p_2)=T$ and $\Im(\sim p_1)=T$ and $\Im(\sim p_2)=F$.

Hence $\{p_1 \rightarrow p_2, \sim p_1\} \not\models \sim p_2$.

Exercise

Show Affirming the Consequent is invalid in propositional logic: $\{p_1 \rightarrow p_2, p_2\} \not\models p_2$.

	p_1	p_2	$(((p_1$	$\rightarrow$	$p_2)$	$\wedge$	$p_2)$	$\rightarrow$	$p_1)$
$\Im_1$	T	T							
$\Im_2$	T	F							
$\Im_3$	F	T							
$\Im_4$	F	F							

Show the corresponding conditional is invalid and use the truth-table to define an interpretation that makes the premises true but the conclusion false.

Showing Consistency and Inconsistency

Like validity it is easy to test whether a finite sets of sentences $\{P_1,\ldots,P_n\}$ in propositional logic is consistent. If the truth-table for the conjunction $P_1\wedge\ldots\wedge P_n$ of the sentences in the set is T in some interpretation, it is consistent. If it is F in every interpretation, it is inconsistent. In the example below the truth-values of the sentences at issue are highlighted in blue, and the truth-value of their conjunction is in yellow. If the yellow values are all F in all interpretations, then the set of sentences is inconsistent.

Theorem. The set $\{p_1\vee p_2, \sim p_1\wedge\sim p_2\}$ is inconsistent in propositional logic.

Proof

	p_1	p_2	$((p_1$	$\vee$	$p_2)$	$\wedge$	$(\sim$	p_1	$\wedge$	$\sim$	$p_2))$
$\Im_1$	T	T	T	T	T	F	F	T	F	F	T
$\Im_2$	T	F	T	T	F	F	T	F	F	F	T
$\Im_3$	F	T	F	T	T	F	F	T	F	T	F
$\Im_4$	F	F	F	F	F	F	T	F	T	T	F

There is no $\Im$ such that $\Im(p_1\vee p_2)$=T and $\Im(p_1\vee p_2)$=T. Hence $\{p_1\vee p_2, p_1\vee p_2\}$ is inconsistent.

Exercise

Show $\{p_1\rightarrow p_2, \sim(\sim p_1\vee p_2)\}$ is inconsistent in propositional logic:

	p_1	p_2	$((p_1$	$\rightarrow$	$p_2)$	$\wedge$	$\sim$	$(\sim$	p_1	$\vee$	$p_2))$
$\Im_1$	T	T									
$\Im_2$	T	F									
$\Im_3$	F	T									
$\Im_4$	F	F									

First-Order Logic

Validity and Logical Entailment

Arguments in first-order logic are shown to be valid by proving metatheorems that show that in any interpretation in which the premises are true, the conclusion is true. These proofs consist of marshalling three ingredients that we are already familiar with: (1) the definition of validity, (2) the schema for a proof showing that an argument is valid, and (3) the truth-conditions for the premises and conclusion as fixed by the definition of an interpretation. Let's review each briefly.

Definition of Validity. The definition of the logical ideas including validity are the same for first-order logic as they were for the categorical and propositional logic:

Definitions

$P_1,\ldots,P_n \models_L Q$ iff $\forall\mathfrak{I}\ ((\mathfrak{I}(P_1)=T \& \ldots \& \mathfrak{I}(P_n)=T\ ) \rightarrow \mathfrak{I}(Q)=T)$

$P_1,\ldots,P_n \not\models_L Q$ iff $\exists\mathfrak{I}\ (\mathfrak{I}(P_1)=T \& \ldots \& \mathfrak{I}(P_n)=T \& \mathfrak{I}(Q)=F)\ )$

P is a *logical truth* in L (in symbols $\models_L P$) iff $\forall\mathfrak{I}\ (\mathfrak{I}(P)=T)$

$\{ P_1,\ldots,P_n\}$ is consistent in L iff $\exists\mathfrak{I}\ (\mathfrak{I}(P_1)=T \& \ldots \& \mathfrak{I}(P_n)=T)$

Proofs of Validity. In first-order logic we cannot make use of truth-tables to show arguments are valid, but must return to the general proof schema that we used earlier to justify arguments in categorical logic. The schema is repeated below. Recall that the overall strategy of the proof is to show that a conditional is true: *if* the argument's premises are true, *then* its conclusion is. The technique used to prove the conditional is conditional proof, a rule which requires a subproof. The *if*-part is assumed at the assumption of the subproof, and the *then*-part is deduced as its last line . The subproof then "proves" the conditional. To indicate the structure of the subproof, the *if*-part assumed as the subproofs first line is underlined, and the *then*-part deduced as its last line double underlined.

Within the subproof, there are various applications of *modus ponens.* The (T) formula for a proposition *P,* which is a biconditional of the form $\Im(P)$=T iff TC(*P*), is written as a line of the proof. Then using *modus ponens* one side of the biconditional is then shown to be true by showing that the other side is true. To indicate the structure, the side being deduced is colored yellow, and the side previously proven is colored green.

Schema for a Validity Proof

Metatheorem Proof Schema. $\{P_1,\ldots,P_n\} \models_L Q$

Proof

Start of subproof

1.	$\Im(P_1)$=T &…& $\Im(P_n)$=T	Assumption for conditional proof, $\Im$ arbitrary
2.	$\Im(P_1)$=T	line 1, conjunction
3.	$\Im(P_1)$=T iff $TC_{\Im}(P_1)$	(T) schema entailed by the definition of $\Im$
4.	$TC_{\Im}(P_1)$	*modus ponens* on the previous two lines
…	…	…
3n+1.	$\Im(P_n)$=T	line 1, conjunction
3n+2.	$\Im(P_n)$=T iff $TC_{\Im}(P_n)$	(T) schema entailed by the definition of $\Im$
3n+3.	$TC_{\Im}(P_n)$	
3n+4.	$TC_{\Im}(P_1)$ &…&$TC_{\Im}(P_n)$,	conjunction of previous TC lines
3n+5.	$TC_{\Im}(Q)$	by set theory and logic from the previous line
3n+6.	$\Im(Q)$=T iff $TC_{\Im}(Q)$	(T) schema entailed by the definition of $\Im$
2n+7.	$\Im(Q)$=T	*modus ponens* on the previous two lines

End of subproof

3n+8.	If ($\Im(P_1)$=T &…& $\Im(P_n)$=T) then $\Im(Q)$=T	1 to n+5, conditional proof
3n+9.	$\forall\Im$(if ($\Im(P_1)$=T &…& $\Im(P_n)$=T) then $\Im(Q)$=T)	n+6, universal generalization, $\Im$ arbitrary
3n+10.	$\{P_1,\ldots,P_n\} \models_L Q$	n+7, definition of $\models$

Truth-Conditions. The proof schema requires that we be able to plug in the truth-conditions $TC_{\Im}(P_1)$ &…&$TC_{\Im}(P_n)$ of the premises and those $TC_{\Im}(Q)$ of the conclusion. When we studied the semantics of first-order logic in Part 2, we learned what truth-conditions were. Relative to an interpretation $\Im$, the truth-conditions of a formula state what facts must obtain among the objects and sets referred to by the formula's constants and predicates for the formula to be true in $\Im$. We also learned how to calculate the truth-conditions for any formula in first-order logic. We will make use of this knowledge to show that arguments are valid. However, rather that actually recalculating the truth-conditions of formulas we have already studied, we will just summarize the truth-conditions already worked out in Lecture 11. We shall refer back to the list below in later proofs. The list begins by stating the general form of Tarski's T-schema and then lists beneath it various formulas and their truth-conditions that we have previously calculated. Below let *F* and *G* range over one-place predicates and *R* over two-place predicates :

(T) $\Im(P)$=T iff $TC_{\Im}(P)$

Truth-Conditions for *P*

TC0. $\Im(Fc)$=T iff $\Im^{D}(c) \in \Im^{D}(F)$

TC1. $\Im(Fc \wedge Gb)$=T iff $\Im^{D}(c) \in \Im^{D}(F)$ and $\Im^{D}(b) \in \Im^{D}(G)$

TC2. $\Im(Rac \rightarrow Gx)$=T iff $<\Im^{D}(a),\Im^{D}(c)> \notin \Im^{D}(R)$ or $\Im^{D}(x) \in \Im^{D}(G$

TC3. $\Im(\forall xFx)$=T iff for all $d \in D$, $d \in \Im^{D}(F)$

TC4. $\Im(\exists xFx)$=T iff for some $d \in D$, $d \in \Im^{D}(F)$

TC5. $\Im(\forall x\exists yRxy)$=T iff for all $d \in D$, for some $d' \in D$, $<d,d'> \in \Im^{D}(R))$

TC6. $\Im(\exists x\forall yRxy)$=T iff for some $d \in D$, for all $d' \in D$, $<d,d'> \in \Im^{D}(R))$

TC7. $\Im(\forall xRxx)$=T iff for all $d \in D$, $<d,d> \in \Im^{D}(R))$

TC8. $\Im(\forall x(Fx \rightarrow Gx))$=T iff for all $d \in D$, either $d \notin \Im^{D}(F)$ or $d \in \Im^{D}(G)$

TC9. $\Im(\exists x(Fx \wedge Gx))$=T iff for some $d \in D$, $d \in \Im^{D}(F)$ and $d \in \Im^{D}(G)$

TC10. $\Im(\forall x(Fx \wedge Gx))$=T iff for all $d \in D$, $d \in \Im^{D}(F)$ and $d \in \Im^{D}(G)$

TC11. $\Im(\exists x(Fx \rightarrow Gx))$=T iff for some $d \in D$, either $d \in \Im^{D}(F)$ or $d \notin \Im^{D}(G)$

TC12. $\Im(\forall x(Fx \rightarrow \exists yRxy))$=T iff for all $d \in D$, either ($d \notin \Im^{D}(F)$ or for some $d' \in D$ $<d,d'> \in \Im^{D}(R))$

TC13. $\Im(\forall x\exists y(Rxy \rightarrow Ryx))$=T iff for all $d \in D$, for some $d' \in D$, either $<d,d'> \notin \Im^{D}(R))$ or $<d',d> \in \Im^{D}(R))$

TC14. $\Im(\forall x\forall y(Rxy\leftrightarrow Ryx))=T$ iff for all $d\in D$, for all $d'\in D$,
$<d,d'>\in\Im^D(R))$ iff $<d',d>\in\Im^D(R))$

TC15. $\Im(\exists xFx\wedge\exists yGy))=T$ iff for some $d\in D$, $d\in\Im^D(F)$ and
for some $d'\in D$, $d'\in\Im^D(G)$

TC16. $\Im(\forall x(Fx\rightarrow\forall yGy))=T$ iff either for some $d'\in D$, $d'\notin\Im^D(F)$ or
for all $d\in D$, $d\in\Im^D(G)$,

Examples of First-Order Validity Metatheorems

Let us now show that various arguments are valid in first-order logic. We begin with first-order forms of the syllogisms Barbara and Celarent, just to show that they are first-order valid.

Every G is H	$\forall x(Gx\rightarrow Hx)$	*No G is H*	$\sim\exists x(Gx\wedge Hx)$
<u>*Every F is G*</u>	$\underline{\forall x(Fx\rightarrow Gx)}$	<u>*Every F is G*</u>	$\underline{\forall x(Gx\rightarrow Hx)}$
Every F is H	$\forall x(Fx\rightarrow Hx)$	*No F is H*	$\sim\exists x(Gx\wedge Hx)$

These and the other valid syllogistic moods remain valid in first-order logic, though some, like the subaltern mood Barbari below, require explicit existence assumptions that are built into the truth-conditions of categorical propositions:

Every G is H	$\forall x(Gx\rightarrow Hx)$
Every F is G	$\forall x(Fx\rightarrow Gx)$
<u>There exists an *F*</u>	$\underline{\exists xFx}$
Some F is H	$\forall x(Fx\rightarrow Hx)$

More important, however, are arguments that cannot be shown valid in simpler languages, like the syllogistic or propositional logic, but that are valid when formulated with the increased expressive power of first-order syntax. Examples of this sort are listed below, written both in English and in their symbolic form.

Socrates is human	$\underline{Fa}$	*Everything is red*	$\underline{\forall xFx}$
Something is human	$\exists xFx$	*Something is red*	$\exists xFx$

Everything is red and everything is round	$\underline{\forall xFx \wedge \forall yGy}$
Everything is red and round	$\forall x(Fx \wedge Gx)$

Something is red and round	$\underline{\exists x(Fx \wedge Gx)}$
Something is red and something is round	$\exists xFx \wedge \exists yGy$

The relation R is complete	$\underline{\forall x\forall y(Rxy \vee Ryx)}$
The relation R is reflexive	$\forall xRxx$

Somebody loves everybody	$\exists x\forall yLxy$
Love is reciprocal	$\underline{\forall x\forall y(Lxy\leftrightarrow Lyx)}$
Everybody loves somebody	$\forall x\exists yLxy$

Notice that these arguments make use of expressive features of first-order syntax that are not available in the simpler languages of the syllogistic or propositional logic: multiple quantifiers, quantifiers nested inside one another, and relational predicates.

Proofs of the Metatheorems

Theorem (Barbara). $\forall x(Gx\rightarrow Hx), \forall x(Fx\rightarrow Gx) \models \forall x(Fx\rightarrow Hx)$

Proof

Start of subproof

1.	$\underline{\mathfrak{I}(\forall x(Gx\rightarrow Hx))=T \;\&\; \mathfrak{I}(\forall x(Fx\rightarrow Gx))=T}$	Assump for CP, $\mathfrak{I}$ arbitrary
2.	$\mathfrak{I}(\forall x(Gx\rightarrow Hx))=T$	1, conjunction
3.	$\mathfrak{I}(\forall x(Gx\rightarrow Hx))=T$ iff for all $d\in D$, either $d\in\mathfrak{I}^D(G)$ or $d\notin\mathfrak{I}^D(H)$	TC 8
4.	for all $d\in D$, either $d\in\mathfrak{I}^D(G)$ or $d\notin\mathfrak{I}^D(H)$	2,3, *modus ponens*
5.	$\mathfrak{I}(\forall x(Fx\rightarrow Gx))=T$	1, conjunction
6.	$\mathfrak{I}(\forall x(Fx\rightarrow Gx))=T$ iff for all $d\in D$, either $d\in\mathfrak{I}^D(F)$ or $d\notin\mathfrak{I}^D(G)$	TC 8
7.	for all $d\in D$, either $d\in\mathfrak{I}^D(F)$ or $d\notin\mathfrak{I}^D(G)$	5,6, *modus ponens*
8.	for all $d\in D$, either $d\in\mathfrak{I}^D(G)$ or $d\notin\mathfrak{I}^D(H)$, and for all $d\in D$, either $d\in\mathfrak{I}^D(F)$ or $d\notin\mathfrak{I}^D(G)$	4,7, conjunction
9.	for all $d\in D$, either $d\in\mathfrak{I}^D(F)$ or $d\notin\mathfrak{I}^D(H)$	8, by set theory
10.	$\mathfrak{I}(\forall x(Fx\rightarrow Hx))=T$ iff for all $d\in D$, either $d\in\mathfrak{I}^D(F)$ or $d\notin\mathfrak{I}^D(H)$	TC 8
11.	$\underline{\mathfrak{I}(\forall x(Fx\rightarrow Hx))=T}$	9, 10 *modus ponens*

End of subproof

12.	If $\underline{(\mathfrak{I}(\forall x(Gx\rightarrow Hx))=T \;\&\ldots\&\; \mathfrak{I}(\forall x(Fx\rightarrow Gx))=T)}$ then $\underline{\mathfrak{I}(\forall x(Fx\rightarrow Hx))=T}$	1-11, CP
13.	$\forall\mathfrak{I}($ If $(\mathfrak{I}(\forall x(Gx\rightarrow Hx))=T \;\&\ldots\&\; \mathfrak{I}(\forall x(Fx\rightarrow Gx))=T)$ then $\mathfrak{I}(\forall x(Fx\rightarrow Hx))=T)$	12, universal generalization, $\mathfrak{I}$ arbitrary
14.	$\forall x(Gx\rightarrow Hx), \forall x(Fx\rightarrow Gx) \models \forall x(Fx\rightarrow Hx)$	13, definition of $\models$

Theorem (Celarent). $\sim\exists x(Gx\wedge Hx),\ \forall x(Fx\rightarrow Gx) \models \sim\exists x(Fx\wedge Hx)$

∗**Exercise.** Prove Celarent is valid in first-order logic.

Theorem. $Fa \models \exists xFx$

Proof

Start of subproof

1.	$\Im(Fa)$=T & $\Im(\exists xF)$=T	Assump for CP, $\Im$ arbitrary
2.	$\Im(Fa)$=T	1, conjunction
3.	$\Im(Fa)$=T iff $\Im(a)\in\Im^D(F)$	TC 0
4.	$\Im(a)\in\Im^D(F)$	2,3, *modus ponens*
5.	for some $d\in D$, $d\in\Im^D(F)$	4, logic (existential generalization)
6.	$\Im(\exists xFx)$=T iff for some $d\in D$, $d\in\Im^D(F)$	TC 4
7.	$\Im(\exists xFx)$=T	5,6 *modus ponens*

End of subproof

8.	If $\Im(Fa)$=T then $\Im(\exists xFx)$=T	1-7, CP
9.	$\forall\Im$(if $\Im(Fa)$=T then $\Im(\exists xFx)$=T)	8, universal generalization, $\Im$ arbitrary
10.	$Fa \models \exists xFx$	9, definition of $\models$

Theorem. $\forall xFx \models \exists xFx$

∗**Exercise.** Prove the metatheorem $\forall xFx \models \exists xFx$.

Theorem. $\forall xFx\wedge\forall yGy \models \forall x(Fx\wedge Gy)$

Proof

Start of subproof

1.	$\Im(\forall xFx\wedge\forall yGy)$=T	Assump for CP, $\Im$ arbitrary
2.	$\Im(\forall xFx\wedge\forall yGy)$=T	1, conjunction
3.	$\Im(\forall xFx\wedge\forall yGy)$=T iff for all $d\in D$, $d\in\Im^D(F)$ and for all d', $d'\in\Im^D(G)$	TC 10
4.	for all $d\in D$, $d\in\Im^D(F)$ and for all d', $d'\in\Im^D(G)$	2,3, *modus ponens*
5.	for all $d\in D$, $d\in\Im^D(F)$ and $d\in\Im^D(G)$	4, set theory
6.	$\Im(\forall x(Fx\wedge Gy))$=T iff for all $d\in D$, $d\in\Im^D(F)$ and $d\in\Im^D(G)$	TC 10
7.	$\Im(\forall x(Fx\wedge Gy))$=T	5,6 *modus ponens*

End of subproof

8.	If $\Im(\forall xFx\wedge\forall yGy)$=T then $\Im(\forall x(Fx\wedge Gy))$=T	1-7, CP
9.	$\forall\Im$(if $\Im(\forall xFx\wedge\forall yGy)$=T then $\Im(\forall x(Fx\wedge Gy))$=T)	8, universal generalization, $\Im$ arbitrary
10.	$\forall xFx\wedge\forall yGy \models \forall x(Fx\wedge Gy)$	9, definition of $\models$

Theorem. $\exists x(Fx\wedge Gx) \models \exists xFx\wedge\exists yGy$

Theorem. $\forall x\forall y(Rxy\leftrightarrow Ryx) \models \forall xRxx$

Proof

Start of subproof

1.	$\Im(\forall x\forall y(Rxy\leftrightarrow Ryx))$=T	Assump for CP, $\Im$ arbitrary
2.	$\Im(\forall x\forall y(Rxy\leftrightarrow Ryx))$=T	1, conjunction
4.	$\Im(\forall x\forall y(Rxy\leftrightarrow Ryx))$=T iff for all $d\in D$, for all d', $<d, d'>\in\Im^D(R)$ iff $<d',d>\in\Im^D(R)$	TC 13
4.	for all $d\in D$, for all d', $<d, d'>\in\Im^D(R)$ iff $<d',d>\in\Im^D(R)$	2,3, *modus ponens*
5.	for all $d\in D$, $<d,d>\in\Im^D(R)$	7, set theory and logic

6. $\Im(\forall xRxx)$=T iff for all $d\in D$, $<d,d>\in\Im^D(R)$ TC 7
7. $\Im(\forall xRxx)$=T 5,6 *modus ponens*

End of subproof

8. If $\Im(\forall x\forall y(Rxy\leftrightarrow Ryx))$=T then $\Im(\forall xRxx)$=T 1-7, CP
9. $\forall\Im$(if $\Im(\forall x\forall y(Rxy\leftrightarrow Ryx))$=T then $\Im(\forall xRxx)$=T) 8, universal generalization, $\Im$ arbitrary
10. $\forall x\forall y(Rxy\leftrightarrow Ryx)\models\forall xRxx$ 9, definition of $\models$

$\exists x\forall yLxy, \forall x\forall y(Lxy\leftrightarrow Lyx)\models\forall x\exists yLxy$

Proof

Start of subproof

1. $\Im(\exists x\forall yLxy)$=T & $\Im(\forall x\forall y(Lxy\leftrightarrow Lyx))$=T Assump for CP, $\Im$ arbitrary
2. $\Im(\exists x\forall yLxy)$=T 1, conjunction
5. $\Im(\exists x\forall yLxy)$=T iff for some $d\in D$, for all d', $<d, d'>\Im^D(L)$ TC 6
4. for some $d\in D$, for all d', $<d, d'>\Im^D(L)$ 2,3, *modus ponens*
5. $\Im(\forall x\forall y(Lxy\leftrightarrow Lyx))$=T 1, conjunction
7. $\Im(\forall x\forall y(Lxy\leftrightarrow Lyx))$=T iff for all $d\in D$, for all d', $<d, d'>\Im^D(L)$ iff $<d',d>\Im^D(L)$ TC 14
8. for all $d\in D$, for all d', $<d, d'>\Im^D(L)$ iff $<d',d>\Im^D(L)$ 5,7, *modus ponens*
9. for all $d'\in D$, for some d, $<d',d>\in\Im^D(L)$ 8, set theory and logic
10. $\Im(\forall x\exists yLxy)$=T iff for all $d'\in D$, for some d, $<d',d>\in\Im^D(L)$ TC 5
11. $\Im(\forall x\exists yLxy)$=T 9,10 *modus ponens*

End of subproof

12. If $\Im(\exists x\forall yLxy)$=T and $\Im(\forall x\forall y(Lxy\leftrightarrow Lyx))$=T, then $\Im(\forall x\exists yLxy)$=T 1-11, CP
13. $\forall\Im$(if $\Im(\exists x\forall yLxy)$=T and $\Im(\forall x\forall y(Lxy\leftrightarrow Lyx))$=T)=T, then $\Im(\forall xRxx)$=T) 12, universal generalization, $\Im$ arbitrary
14. $\exists x\forall yLxy, \forall x\forall y(Lxy\leftrightarrow Lyx)\models\forall x\exists yLxy$ 13, definition of $\models$

We show by construction that Barbari without an explicit existence assumption is invalid in first-order logic.

Theorem (Barbari without its existence assumption). $\forall x(Gx\rightarrow Hx)$, $\forall x(Fx\rightarrow Gx)\not\models\exists x(Fx\wedge Hx)$

Proof

1. $\Im(F)=\varnothing$ & $\Im(G)=\varnothing$ & $\Im(H)=\varnothing$ Def of $\Im$ (principle of abstraction)
2. $\Im(G)=\varnothing$ & $\Im(H)=\varnothing$ 1, conjunction
3. $\Im(G)\subseteq\Im(H)$ 2, set theory
4. for all $d\in D$, if $d\in\Im^D(G)$ then $d\in\Im^D(H)$ 3, def of $\subseteq$
5. for all $d\in D$, either $d\in\Im^D(G)$ or $d\notin\Im^D(H)$ 4, logic
6. $\Im(\forall x(Gx\rightarrow Hx))$=T iff for all $d\in D$, either $d\in\Im^D(G)$ or $d\notin\Im^D(H)$ TC 8
7. $\Im(\forall x(Gx\rightarrow Hx))$=T 5,6 *modus ponens*
8. $\Im(F)=\varnothing$ & $\Im(G)=\varnothing$ 1, conjunction
9. $\Im(F)\subseteq\Im(G)$ 8, set theory
10. for all $d\in D$, if $d\in\Im^D(F)$ then $d\in\Im^D(G)$ 9, def of $\subseteq$
11. for all $d\in D$, either $d\in\Im^D(F)$ or $d\notin\Im^D(G)$ 10, logic
12. $\Im(\forall x(Fx\rightarrow Gx))$=T iff for all $d\in D$, either $d\in\Im^D(F)$ or $d\notin\Im^D(G)$ TC 8
13. $\Im(\forall x(Fx\rightarrow Gx))$=T 11,12 *modus ponens*
14. $\Im(F)=\varnothing$ & $\Im(H)=\varnothing$ 1, conjunction
15. for all $d\in D$, $d\notin\Im^D(F)$ or $d\notin\Im^D(H)$ 14, set theory and logic
16. $\Im(\exists x(Fx\wedge Hx))$=T iff for some $d\in D$, $d\in\Im^D(F)$ and $d\in\Im^D(H)$ TC 8

17.	$\Im(\exists x(Fx\wedge Hx))\neq$T iff for all $d\in D$, $d\notin\Im^{D}(F)$ or $d\notin\Im^{D}(H)$	16, logic
18.	$\Im(\exists x(Fx\wedge Hx))\neq$T	15, 17 *modus ponens*
19.	$\Im(\forall x(Gx\rightarrow Hx))$=T & $\Im(\forall x(Fx\rightarrow Gx))$=T& $\Im(\exists xFx)$=T and $\Im(\exists x(Fx\wedge Hx))$=F)	7,13,18 conjunction
20.	$\exists\Im(\Im(\forall x(Gx\rightarrow Hx))$=T & $\Im(\forall x(Fx\rightarrow Gx))$=T& $\Im(\exists xFx)$=T and $\Im(\exists x(Fx\wedge Hx))$=F)	19, construction
21.	$\forall x(Gx\rightarrow Hx)$, $\forall x(Fx\rightarrow Gx)$, $\exists xFx \nvDash \exists x(Fx\wedge Hx)$	19, definition of $\vDash$

With the explicit assumption, however, Barbari is valid, as are the other traditional subaltern moods (Celaront, Camestrop, Cesaro, Camelop) and as well as Darapti, Felapton, Fesapo, and Bramantip**.**

Theorem (Barbari in First-Order Logic). $\forall x(Gx\rightarrow Hx)$, $\forall x(Fx\rightarrow Gx)$, $\exists xFx \vDash \exists x(Fx\wedge Hx)$

Proof

Start of subproof

1.	$\Im(\forall x(Gx\rightarrow Hx))$=T & $\Im(\forall x(Fx\rightarrow Gx))$=T & $\Im(\exists xFx)$=T	Assump for CP, $\Im$ arbitrary
2.	$\Im(\forall x(Gx\rightarrow Hx))$=T	1, conjunction
3.	$\Im(\forall x(Gx\rightarrow Hx))$=T iff for all $d\in D$, either $d\in\Im^{D}(G)$ or $d\notin\Im^{D}(H)$	TC 8
4.	for all $d\in D$, either $d\in\Im^{D}(G)$ or $d\notin\Im^{D}(H)$	2,3, *modus ponens*
5.	$\Im(\forall x(Fx\rightarrow Gx))$=T	1, conjunction
6.	$\Im(\forall x(Fx\rightarrow Gx))$=T iff for all $d\in D$, either $d\in\Im^{D}(F)$ or $d\notin\Im^{D}(G)$	TC 8
7.	for all $d\in D$, either $d\in\Im^{D}(F)$ or $d\notin\Im^{D}(G)$	5,6, *modus ponens*
8.	$\Im(\exists x(Fx)$=T	1, conjunction
9.	$\Im(\exists x(Fx)$=T iff for some $d\in D$, $d\in\Im^{D}(F)$	TC 4
10.	for some $d\in D$, $d\in\Im^{D}(F)$	8,9, *modus ponens*
11.	for all $d\in D$, either $d\in\Im^{D}(G)$ or $d\notin\Im^{D}(H)$, and for all $d\in D$, either $d\in\Im^{D}(F)$ or $d\notin\Im^{D}(G)$ and for some $d\in D$, $d\in\Im^{D}(F)$	4,7,10 conjunction
12.	for some $d\in D$, $d\in\Im^{D}(F)$ and $d\in\Im^{D}(H)$	11, by set theory and logic
13.	$\Im(\exists x(Fx\wedge Hx))$=T iff for some $d\in D$, $d\in\Im^{D}(F)$ and $d\in\Im^{D}(H)$	TC 9
14.	$\Im(\exists x(Fx\wedge Hx))$=T	12, 13 *modus ponens*

End of subproof

15.	If $(\Im(\forall x(Gx\rightarrow Hx))$=T & $\Im(\forall x(Fx\rightarrow Gx))$=T& $\Im(\exists xFx)$=T) then $\Im(\exists x(Fx\wedge Hx))$=T	1-14, CP
16.	$\forall\Im$(If $(\Im(\forall x(Gx\rightarrow Hx))$=T & $\Im(\forall x(Fx\rightarrow Gx))$=T& $\Im(\exists xFx)$=T) then $\Im(\exists x(Fx\wedge Hx))$=T)	15, universal generalization, $\Im$ arbitrary
17.	$\forall x(Gx\rightarrow Hx)$, $\forall x(Fx\rightarrow Gx)$, $\exists xFx \vDash \exists x(Fx\wedge Hx)$	16, definition of $\vDash$

$*$**Exercise.** Prove the metatheorems in the semantics for first-order logic:

1. $\forall xFx \vDash \exists xFx.$

2. $\forall x(Bx\rightarrow\exists y(Gx\wedge Lxy))$, $\sim\exists x(Gx \wedge Lcx) \vDash \sim Bc$

3. $\forall x\forall y\forall z((Lxy \wedge Lyz))\rightarrow Lxz)$, $\forall x\sim Lxx \vDash \sim\exists x\exists y(Lxy \wedge Lyx)$

4. $\exists x(Gx \wedge Hx)$, $\exists x(Fx \wedge\sim Gx) \nvDash \forall x(Fx\rightarrow Hx)$

LECTURE 16. PROOF THEORY

The Axiomatic Method

Validity, it has now been said many times over, is defined semantically, in terms of truth – an argument is *valid* if and only if in any interpretation in which the premises are true, the conclusion is true also. But this is not the only "definition" of valid arguments. Notice that the semantic definition has the form of a traditional "if and only if" definition in philosophy: it lays out the necessary and sufficient conditions for an argument to be valid. There is however another way to "characterize" the valid arguments, and sometimes it is put forward as a definition: the argument from $P_1,\ldots,P_n$ to Q is "logical" if a a formally correct *proof* can be constructed that has $P_1,\ldots,P_n$ as premises and Q as its conclusion. This idea of logical argument is so basic that logic is sometimes even defined as the study of proofs. In this and subsequent lectures we will be exploring what proofs are and how they are related to what we have been calling valid arguments. We begin with the history of the idea.

The notion of *proof* has be know in various forms since ancient Greece, and it has had profound influence on the history of philosophy generally because it has been used as a model for the ideal form of knowledge. Proofs impress philosophers. They do so because of their connection to knowledge, perhaps the most basic philosophical idea. The very word *philosophy* means *lover of wisdom,* and *wisdom* is a synonym of *knowledge.* *Knowledge*, in turn, has a standard definition in philosophy. Since Plato first formulated the definition in the *Theatetus, knowledge* has been defined as *justified true belief*. Let S be a person. Then, the standard definition is:

S knows that *P* if and only if

1. *P* is true;
2. *S* believes that *P*;
3. *S* is well justified in believing that *P.*

The important part of this definition for logic is the phrase “well justified”.

What is “good justification” in the case of knowledge? That is philosopher’s question *par excellence.* There is one example, however, that thinkers have always seen as a paradigm: logical proof. If logical proofs do not justify our claims to knowledge, nothing does. More precisely, if a proof of *P* is adequate justification for knowledge of *P,* then two things are required: (1) the proof must have premises that are known to be true, and (2) the steps of the proof must transparently take us from something true to something true. If the premises are known not only to be true, but are certainty, necessary, or self-evident, so much the better.

The paradigm of justification by proof became dominant in Western philosophy and in natural science as it emerged from philosophy because the first truly successful “science” was mathematics, especially geometry. The propositions of these disciplines are particularly suitable for justification by proofs. Some basic mathematical propositions do seem to be self-evident, and others can be proven from them in simple logical steps in elegant proofs. Euclid’s *Elements* (4^{th} century B.C.) formulates the truths of plane geometry in just this way, deducing an extensive set of “theorems” from five “postulates:”

Euclid’s Postulates for Plane Geometry

(*Elements*, 300 B.C.)

1. Any two points are contained in some line.

2. Any finite line is contained in some line not contained in any other line.
3. Any point and any line segment beginning with that point determine a circle with the point as its center and the line as its radius.
4. All right angles are equal.
5. (Euclid's original version.) If a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, then the two straight lines, if produced indefinitely, meet on that side on which the angles are less than the two right angles.

Much of this geometric lore was known to Plato and Aristotle, and its format affected their views on knowledge generally. Plato thought of geometry as a model of human knowledge. As portrayed in his dialogue the *Meno,* he believed that the process of uncovering truths by geometrical reasoning "reminded" a person of a prior direct experience of mathematical Forms encountered in a life before birth. Aristotle, who is much more important for the history of logic, explained the process of proof systematically, and indeed launched logic on its way as a separate science. In the *Topics* he used this logic to formulate his notion of the scientific method. Science should consist of deductions using syllogistic logic from premises that consist of necessary definitions, which classify natural objects into genera and species.

Throughout the mediaeval period philosophers continued to advocate methods similar to Aristotle's, to the degree that they often formulated their work as a series of syllogistic arguments. Likewise, mathematicians and physicist subsequent to Euclid – like Archimedes, Ptolemy, Copernicus, Galileo, and Newton – followed Euclid in formulating their work as "axiom systems." Newton mechanics for, example, begins with three formulas familiar from high school physics:[10]

Newton's Three Laws of Motion

(*Mathematical Principles of Natural Philosophy*, 1686)

1. Every body continues in its state of rest, or of uniform motion in a right line, unless it is compelled to change that state by forces impressed upon it.
2. The change of motion is proportional to the motive force impressed; and is made in the direction of the right line in which that force is impressed.
3. To every action there is always an equal reaction: or, the mutual actions of two bodies upon each other are always equal, and directed to contrary parts.

Occasionally even philosophers formulated their theories in axiomatic form. In the fifth century Proclus' does so in the *Elements of Theology,* which is set forth as a series of axioms and remarks on them. In the 17th century Descartes declares in *Discourse on Method*, a manifesto setting forth the "rationalist" scientific method, that all science must be deduced from self-evident first-principles. He tried to do so in various works of philosophy and natural science, for example, his well knows *Meditations* that tries to establish the basis for our knowledge of ourselves and the world. Other philosophers adopted his program. Spinoza, for example, lays out his philosophical system as a series of theorems deduced, he claims, from seven axioms, which – it must be said – are famous for their obscurity:

[10] These laws are in Newton's original formulations. You may recall them as (1) a body at rest tends to remain at rest, and a body in motion tends to remain in motion, until acted upon by an external force, (2) $f = ma$, and (3) for every action, there is an equal and opposite reaction.

Spinoza's Philosophical Axioms

(*Ethics*, 1670)

1. Everything which is, is either in itself or in another.
2. That which cannot be conceived through another must be conceived through itself.
3. From a given determinate cause an effect necessarily follows; and, on the other hand, if no determinate cause is given, it is impossible that an effect can follow.
4. The knowledge of an effect depends upon and involves the knowledge of the cause.
5. Those things which have nothing mutually in common with one another cannot through one another be mutually understood, that is to say, the conception of the one does not involve the conception of the other.
6. A true idea must agree with that which is the idea.
7. The essence of that thing which can be conceived as not existing does not involve existence.

Leibniz advocated formulating all science is a special universal language and deducing all the truths of science from self-evident propositions asserting conceptual identities, like *Every S is S*, using his adaptations of syllogistic reasoning.

In the 18th century practicing scientists and philosophers under the lead of the British Empiricists – most prominently, Locke, Berkeley and Hume – rejected the rationalist project because, as it became increasingly clear, scientific discoveries were best justified not by proofs from self-evident axioms but by highly fallible generalizations from observations and experiments. Logical proofs, however, were still scientifically important because they remained, as they still do, the primary means of justification in mathematics.

In 1790, Kant advanced a systematic philosophy that attempted to explain the special role of logic. Its laws, he claimed, express the very forms of thought. The underlying nature of reality dictates that when we perceive something, we organize the world in accordance with the rules of logic and mathematics. By logic he means syllogisms and by math he means Euclid.

Among philosophers and mathematicians Kant's views were extremely influential. They came into conflict, however, with a niggling doubt that had been troubling specialists in geometry since ancient times. Since ancient Greece Euclid's axioms have been regarded as self-evident. Indeed, Kant claimed we know them *a priori,* i.e. without an appeal to sense experience. The fifth postulate, however, was not as obvious as the other five. If you actually count its worlds and parse its grammar, you will see that it is significantly longer and more complicated than the other four. For this reason alone it is less likely to express an instantly obvious truth. Indeed, even in ancient times it was viewed as different. In the fifth century, for example, Proclus claimed he was able to prove it from the other four. Though not apparent until much later, his proof contains an error by assuming in a subtle way the very postulate he is trying to prove.[11] But the postulate remained troubling and attempts to put in on sounder footing continued without much success until the early 19th century.

Non-Euclidean Geometry

In the 19th century it is was shown that in fact the fifth postulate does not follow from the other four. [12] An important step in doing so was the reformulation of the fifth

[11] See Glenn R. Morrow, trans., *Proclus, A Commentary on the First Book of Euclid's Elements* (Princeton: Princeton University Press, 1970), ll. 371.10-373.2, pp. liv-lv, 290-291.

[12] There are introductions at all levels to non-Euclidean geometry. One that is non-technical is Philip J. Davis and Reuben Hersh, *The Mathematical Experience* (Boston: Houghton Mifflin, 1981). On the relation of non-Euclidean geometry to logic, see Howard DeLong, *A Profile of Mathematical Logic*

postulate in a shorter but equivalent manner. John Playfair (1748-1819) proposed a revised version (also know in antiquity) stated in terms of parallel lines.

Playfair's Version of Euclid's Fifth Postulate

Given a line and a point not on that line, there is exactly one line through that point parallel to the given line.

Karl Friedrich Gauss (1777-1855), Johann Bolyai (1802-1860), and Nikilai Lobachevski (1793-1856) independently discovered that a consistent geometry would result by retaining Euclid's first four axioms but replacing Payfair's fifth postulate by one inconsistent with it. The replacement specifies that through a point not on a line *l* there is more than one line parallel to *l*.

Postulate 5 in Lobachevskian Geometry

Given a line and a point not on that line, there are at least two distinct lines through that point parallel to the given line.

This system, not surprisingly, has some novel theorems. For example, the sum of the angles formed by a line bisecting two parallels is in general less than that of two right angles. This geometry, unlike Euclid's, also has the property that the measure of the least angle formed by the intersection of a parallel to the perpendicular of a line varies directly with the distance of the intersection from the line.

(Reading, MA: Addison-Wesley, 1970). On non-Euclidean geometry as a logistic system see Raymond L. Wilder, *Introduction to Foundations of Mathematics*, 2nd ed. (N.Y.: Wiley, 1967).

Soon after this discovery a third variety of geometry was discovered by Bernhard Rienmann (1826-1866), who observed that the angles of a triangle may be greater than that of two right angles and that a line and a point might well determine no parallel.

Postulate 5 in Riemanninan Geometry

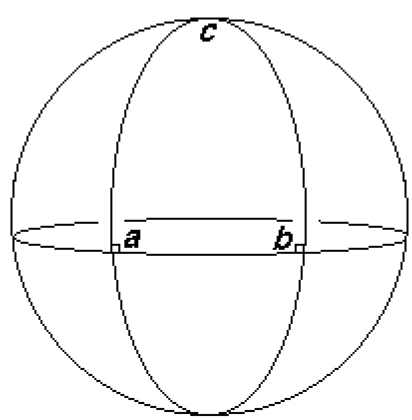
> Given a line and a point not on that line, there is no line through that point parallel to the given line.

It is easy to show that Rienmann's geometry is consistent. According to the earlier definition a set $\{P_1,\ldots,P_n\}$ is *consistent* if and only if there is an interpretation $\Im$ such that $\Im(P_1)$=T &…& $\Im(P_n)$=T. To use alternative terminology, we show $\{P_1,\ldots,P_n\}$ is consistent by constructing a "model" in which $P_1,\ldots,P_n$ are simultaneously true. In the case of geometry we construct a model – defined an interpretation $\Im$ – for the terms *point* and *line* that occur in the language of geometry in such a way that the five axioms are jointly true in $\Im$. For the domain of the model let us employ the points that make up the surface of a sphere. By *point* in $\Im$ let us mean any object in the domain, i.e. any point on the sphere's surface. By a *line* let us mean any great circles on the sphere's surface, that is, any set of points that satisfies the equation for a circle and which as its center the center of the sphere itself.

It is easy to see that in this interpretation the first four postulates are true. Any two points on the sphere's surface fall on some great circle, confirming postulate 1. Any finite arch of a great circle is contained in some great circle, which is itself not a portion of another circle, satisfying postulate 2. Any arch of a great circle from a point on the surface determines a circle on the sphere with that arch as its radius, verifying postulate 3. Finally, all right angles are equal, as required by postulate 4.

It is also true that Rienmann's postulate 5 is true. For consider any great circle on the sphere and a point on the surface that is not on the circle. Now imagine a second a great circle passing through that point. This circle will interact the original circle, and hence is not parallel to it. Hence a "line" and a point determine no parallel.

It is also easy to see that the sum of the angles in a triangle are in general greater than that of two right angles. For example, consider the equator of the sphere given in the figure below. Consider in addition the sphere's "north pole" point *c*. Clearly *c* is not on the equator. Hence any great circle that passes through *c* will also intersect the equator. Indeed, any such circle will be a line of longitude of the sphere, forming a right angle with the equator. Now consider two points *a* and *b* on the equator, and the lines of longitude passing through them. Notice also that ∠*cab* and ∠*cba* are right angles to the equator. Hence the sum of the angles of the triangle *cab* will be greater than that of two right angles.



Theorem. Rienmannian Geometry is consistent.

This discussion shows that Proclus and others were wrong in thinking that the first four postulates are sufficient for determining a Euclidean world. The fifth postulate is required as well.

It is hard to overstate the shock that resulted from the discovery of non-Euclidean geometry. No longer was geometry a paradigm of *a priori* knowledge. Indeed, which geometry was the right one, i.e. which was true in the actual world, became an open question, and evidently a matter to be resolved by empirical observation. Gauss and others started actually measuring the sum of the angles of large terrestrial triangles to see it they equaled 90 degrees. He found that within the margin of error of his measuring devices Euclid's geometry seemed to be confirmed. In the 20th century, however, it was a version of Rienmann's geometry that was incorporated into Einstein's theory of relativity. Geometry, in short, lost its status as *a priori*, knowledge, and doubt was sown about the rest of mathematics. If geometry was not known *a prior*, then perhaps other branches of mathematics were also empirical.

Axiom Systems and Proofs

Throughout the debate, however, the security of the logical used to deduce the various steps in mathematical proof remained constant. It was the status of the *premises* that changed, not that of the logic used to deduce new steps from old. In those parts of the natural sciences open to mathematical description in which proofs were applicable, there was never any doubt about the proof steps themselves. If the premise of an argument were true and the steps of the proof fit the “rules of logic”, then the conclusion was true too. Thus, though the status of the premises, and hence of the conclusion, of logical arguments in science came to be downgraded from *a priori* or self-evident to merely true, the certainty of the logical inferences themselves was not questioned. If a step follows logically, we can be *sure* it does.

The proof rules, however, did come under increased scrutiny. Through the 18th century proofs were often informal and sometimes quite obscure. In the 19th century, in part because of the need for the greater precision to distinguish the different forms of geometry, proofs became more rigorous and explicit. By the mid 19th century the standards of mathematical proof were essentially what they are today. Some mathematicians, however, took their interest further. They proposed precise symbolic languages for the formulation of mathematical propositions and codified restricted sets of inference rules that were generally accepted by the mathematical community. In the 19th and early 20th centuries logical systems of this sort were designed by Frege, and by Russell and Whitehead to deduce the truths of arithmetic from axioms of set theory and logic. We studied a simplified version of Frege’s system in Part 1 under the name *naïve set theory.*

The framework the used is called an *axiom* or *logistic system.* It has three parts: (1) axioms, (2) inference rules and (3) theorems deduced from the axioms by the inference rules. Let us say something about each, and about how they are related.

First of all, the axioms and rules are used to define the set of theorems in an inductive definition. The axioms constitute the “stater elements” of the set of theorems, and new elements are added by the iterative application of the inference rules. The set of theorems is defined as the closure of the axioms under the inference rules. Since the set of theorems is inductively defined, each theorem *P* has a constructions sequence. This is a series $Q_1,\ldots,Q_n$ such that $P=Q_n$, and each Q_i is either a starter element (i.e. an axiom), or is constructed (i.e. derive) from earlier elements of the series. In an axiom system, another name for a construction sequence is *a proof.*

The second important feature of an axiom system is that axioms and inference rules are “syntactic” or “formal”. Lets see how logicians explain what syntax means. Syntax begins by stipulating a set Σ of *signs.* A sign is understood to be a physical object of a sort that can be easily perceived to be distinct from other signs and that can be used to represent another entity. In general, a concept is *syntactic* if it is defined in terms of the easily perceptible physical characteristics of signs (marks or sounds). Signs are usually understood to be written (*marks*) or spoken (*sounds*).

An important example of a syntactic idea is the operation called *concatenation*. This is the process of “writing” one mark to the left of another, or of “saying” one word right after another. “Writing” and “saying” of this kind are supposed to be physical operations that are easy to perceive. The series of marks or sounds, moreover, is supposed to form a perceptual whole easily distinguished from its parts, yet one that displays in its physical shape the very parts from which it was formed. Normally, when you see a series of marks or hear a “stream” of words, you can pick out the individual signs used to compose the whole. Logicians use the symbol $^\frown$ to represent the concatenation operation: $x^\frown y$ is the name of the result of concatenating the sign *y* after the sign *x*. When *x, y,* and *z* are written marks, we can delete the $^\frown$ symbol because, $x^\frown y$ is just another name for *xy*, $x^\frown y^\frown z$ is *xyz* , etc.

Any rule that applies to signs and is defined by means of concatenation is automatically syntactic because it forms an easily perceptible physical whole that exhibits its parts. The formation rules of grammar that we have met in Part 2 are concatenation rules. For example, in categorical logic to form the proposition **A***SP* proposition, we concatenate the marks **A,** *S* and *P,* and in propositional logic to form the sentence $(p_3\wedge{\sim}p_5)$, we first concatenate ~ with p_5 to get ${\sim}p_5$. We then concatenate (,p_3,$\wedge$, ${\sim}p_5$, and) to get $(p_3\wedge{\sim}p_5)$.

The axioms and inference rules of an axiom system are required to be definable syntactically. Normally they are defined by concatenation. For example, later in defining an axiom system for propositional logic, we shall say:

Any sentence of the form $(P\rightarrow(Q\rightarrow P))$ is an axiom.

This is a shorthand way of saying something that when more fully stated, makes use of the idea of concatenation:

> the set of all formula that result from concatenating (with *P* with $\rightarrow$ with (with *Q* with $\rightarrow$ with *P* with) is a subset of the set of axioms

This statement can be said both more precisely and briefly in the notation of set theory. Let *Sen* be the set of sentences and *Ax* the set of axioms. The reformulation is then:

$$\{\, (^\frown P^\frown\rightarrow^\frown(^\frown Q^\frown\rightarrow^\frown P)^\frown) : P\in Sen \;\&\; Q\in Sen\} \subseteq Ax$$

Because this set is defined in terms of concatenation it is syntactic. We can easily "see" whether a sentence has the form ($P\rightarrow(Q\rightarrow P)$), and if it does, we can tell that it is made up of the parts P, Q, and ($Q\rightarrow P$).

As an example of a syntactic inference rule defined in terms of concatenation consider *modus ponens.* We shall use this rule shortly to axiomatize propositional logic. For illustration let us formulate the rule in two different but equivalent ways, both of which use concatenation. The first formulation stresses that *modus ponens* is, a construction rule used in the inductive definition of theorem, and second emphasizes that it is an "inference rule" used in justifying the steps of a proof.

From ($^\frown P^\frown\rightarrow^\frown Q^\frown$) and P construct Q.
From ($^\frown P^\frown\rightarrow^\frown Q^\frown$) and P deduce Q.

Because a proof is simply another name for a construction sequence for a theorem, the two formulations really describe the same operation in different words. The important point, however, is that because *modus ponens* is defined in terms of concatenation, it is syntactic.

We are ready to state the definition of an axiom system. Let *Sen* be the set of formulas of a formal language.

Axiom System

Definition

An *axiom system AS* for *Sen* is a structure <*Ax, PR, Th*> such that

1. *Ax* (the *axioms* of *AS*) is a syntactically defined subset of *Sen,*
2. *PR* is a syntactically defined set of relations on *Sen* (the *inference rules* of *AS*), and
3. *Th* (the set of *theorems* of *AS*) is the inductively defined set such that
   a. $Ax\subseteq Th$;

b. for any $x_1,\ldots,x_n,x_{n+1}$ in *Sen,* $x_1,\ldots,x_n$, and any *n*-place relation *R* in *PR,* if $x_1,\ldots,x_n$ are in *Th* and $< x_1,\ldots,x_n,x_{n+1} > \in R$, then x_{n+1} is in *Th;*

c. nothing else is in *Th.*

It is customary to use the symbol $\vdash$ to indicate that a sentence is a theorem.

Definitions

1. A *proof* for a theorem *P* of *AS* is any construction sequence for *P* showing it is an element of *Ax.*
2. $\vdash P$ is an abbreviation for *P is a theorem* or $P \in Th$

Derivation

An axiom system can also be used to construct formally correct arguments. Intuitively an argument is formally correct if its conclusion "follows from" its premises by appeal to the axioms, theorems and inference rules of the system. We are said to "derive" a conclusion *Q* from a set of premises $P_1,\ldots,P_n$ in an axiom system if we can "prove" *Q* by adding $P_1,\ldots, P_n$ to the axioms of the system and then deducing *Q* from the combined set of axioms and premises.

Definition

Relative to an axiom system *AS,* we say $\{P_1,\ldots,P_n\}$ is *derivable* from *Q* (abbreviated $\{P_1,\ldots,P_n\} \vdash Q$) iff there is a finite sequence (called a *derivation*) such that

a. *Q* is its last element*;* and
b. each element in the sequence is either in $\{P_1,\ldots,P_n\}$ or *Ax* or follows from earlier elements in the sequence by some rule in *PR.*

Another way to say that $\{P_1,\ldots,P_n\}$ is derivable from *Q* would be to say that *Q* would be a theorem if we added $\{P_1,\ldots,P_n\}$ to the axioms, as the following theorem confirms:

Theorem

1. $\{P_1,\ldots,P_n\} \vdash Q$ iff $AS' = \langle Ax \cup \{P_1,\ldots,P_n\}, PR, Th' \rangle$ is an axiom system and $Q \in Th'$.
2. A derivation of Q from $P_1,\ldots,P_n$ is a proof of Q in AS'.

Again, if the axioms and inference rules of the system are well chosen it will turn out that all and only valid arguments of the language are derivable in the axiom system.

Certainty

At this point we should pause to explain why syntactic concepts are especially important to the scientific method of logic. (Note that *formal* is a synonym for *syntactic* because syntax is a matter of form.) As we have seen, built into the definition of a syntactic property is the fact that it is a crudely perceivable physical property of marks or sounds. The important of such properties lies in the fact that they are *epistemically transparent*. It is easy to perceive *with a high degree if certainty* whether something has them or not. It is easy, for example, to perceive by looking at the physical shapes of three consecutive lines of a proof $P \to Q$, *P,* and Q that they fit the form of *modus ponens* .

A feature of logical arguments that cries out for explanation is that if the premises are true then *we call tell with an extremely high degree of certainty* that the conclusion is also true. What accounts for this epistemological sureness? It is due to the epistemic transparency of syntactic ideas. In the case of *modus ponens* we are certain because the sentences in question fit a readily perceivable syntactic pattern, something that is easy to judge by the simple perception of the physical properties of the signs being used.

On the other hand, though syntactic ideas of grammar and proof theory can be spotted with a high degree of certainly, they might be totally uninteresting. For example, let us call a written word of English *odd* if its normal spelling contains an odd number of letters. We can tell then with a high degree of certainty that *cow* is odd but *goat* is not odd. But this concept is utterly uninteresting. Who cares whether a word is odd or not in this sense? Likewise, it may be easy to perceive whether a sequence of lines in an axiom system is a proof, and hence to know with a high degree of certainty whether a formula is a theorem of an axiom system, but that fact may be utterly uninteresting. Let us now explain what makes axiom systems interesting.

Soundness and Completeness

What makes axiom systems interesting is that their theorems happen to be logical truths and their inference rules happen to be valid arguments. Notice that axioms of a system are simple human creations. Some logician decides that these sentences are to be axioms and those operations are to be inference rules. There is no guarantee that the axioms are logical truths or the rules valid. For example, consider the system with $p_1 \wedge \sim p_1$ as its only axiom, together with one rule: from P deduce $\sim P$. That system's axioms and theorems are inconsistent, and its rule is invalid. A logician who proposed it would be fired. What makes a good axiom system interesting is that the axioms and rules have been chose with special care. The axioms, which are defined syntactically, are chosen because they are logical truths, and the rules, also defined syntactically, are chose because they are valid. Moreover, in the ideal case, there are enough axioms and rules so that no logical truth is left out of the set of theorems and every valid argument is derivable. In that happy coincidence, the axiom system is very interesting indeed. In that case the set of theorems, defined syntactically, coincides with the set of logical truths, defined semantically; and the set of derivable arguments, defined syntactically, coincides with the set of valid arguments, defined semantically. Because every logical truth and every valid argument is captured by the axiom system, the system is said to be *complete*. Conversely, because every theorem and derivable argument is logically true or valid, it is said do be *sound*. Let $P_1,\ldots,P_n$, Q be formulas of a formal language.

Definitions

An axioms system is *statement sound* iff every theorem is a logical truth, and is *argument sound* iff

for any $P_1,\ldots,P_n$, Q, if $\{P_1,\ldots,P_n\} \vdash Q$ then $\{P_1,\ldots,P_n\} \models Q$.

An axioms system is *statement complete* iff every logical truth is a theorem, and is *argument complete* iff

for any $P_1,\ldots,P_n$, Q, if $\{P_1,\ldots,P_n\} \models Q$ then $\{P_1,\ldots,P_n\} \vdash Q$.

Sound and complete theories are extremely interesting scientifically because they characterize the same thing in two ways. First, in the semantic theory they define the set of logical truths and valid arguments. These are scientifically interesting because they identity an important set of truths common to all the sciences and the set of logical arguments that can be used to advance knowledge from things already known to things yet to be discovered. Secondly, in the proof theory (the axiom system) they impart to the logical truths and valid arguments a degree of certainty unrivaled in other branches of science. Soundness and completeness theorems show that the key ideas of logic – logical truth and valid argument – are simultaneously scientifically interesting and epistemically accessible.

In the following lectures we shall review a series of sound and complete systems for the languages we studied in Part 2: categorical, the propositional, and first-order logic. We will begin with categorical logic both because it it provides a good example due to its simplicity and because it was very important in history of pre-20th century philosophy. We then investigate the two languages of modern logic – propositional logic, which is simpler, and first-order logic, which is expressively powerful.

Summary

In this lecture we have meet the notion of an axiom system. In such a system a set of theorems is has a definition that is both inductive and syntactic. It is defined inductively as the closure of a set of axioms under a set of inference rules. Because it is inductive, every theorem has a construction sequence, also know in the case of an axiom system as a proof. Because the set of axioms and inference rules are defined syntactically they are epistemically transparent. If a formula has a proof, we can know with a high degree of certainty that it is in the set of theorems.

We also learned that there is no guarantee that the set of theorems or derivable arguments captured in an axiom system is interesting. Whether they are depends on whether the set of theorems happens to correspond to the set of logical truths of the language and whether the arguments derivable in the axiom system happen to coincide with the set of valid arguments of the language. If every theorem is logically true and every derivable argument is valid, the system is sound. If every logical truth is a theorem and every valid argument is derivable, it is complete. If a system is sound and complete, it simultaneously explains why the set of logical truths and valid arguments are interesting and knowable with certainty. The semantic theory, which defines truth in an interpretation, logical truth, and valid argument, provides the analyses showing that the these sets are scientifically interesting because they say something about the world and provide reasoning patterns for the advancement of knowledge. The proof theory shows why the logical truths and derivable argument are knowable with certainty because it defines them in finite epistemically transparent syntactic ideas.

LECTURE 17. REDUCTIONS TO THE PERFECT SYLLOGISMS

Syllogistic Reduction

In this lecture we shall have some fun. We will meet our first example of an axiom system – that part is serious – but at the same time we will expose some occult lore from the Middle Ages. In the *Prior Analytics* Aristotle not only described the valid syllogisms. He also sketched what was really an early version of an axiom system. All the valid syllogisms, he suggested, were is some fundamental way "reducible" by logical rules to the "perfect syllogisms" of the first figure, Barbara and Celarent:

Barbara	Celarent
Every M is P	*No M is P*
<u>*Every S is M*</u>	<u>*Every S is M*</u>
∴ *Every S is P*	∴ *No S is P*

Barbara is perfect (i.e. complete) in the sense that it is logically transparent. What could be more logically evident than the transitivity of $\subseteq$? Celarent in turn may be seen to be almost as perfect as Barbara because in a sense it is a version of Barbara. If we reformulate *No X is Y* contrapositively as *Every Y is non-X,* Celarent becomes a case of Barbara:

Every M is non-P
<u>*Every S is M*</u>
∴ *Every S is non-P*

In this lecture we shall see how to "reduce" all the valid syllogisms to Barbara and Celarent. Viewed in reverse a syllogism's reduction is what we would call today a proof. So we will also be seeing an early form of axiom system in which all the valid syllogisms are proven from Barbara and Celarent by proof rules.

Aristotle's insight was to define in a rudimentary way what we call today an inductive set. This is *RSyl,* the set of " *reducible syllogisms"*. As in any inductive definition, we begin by stipulating the set of "starter" elements. This is the set *PSyl* of *perfect syllogisms,* defined as all syllgisms that have the form of Barbara and Celarent. These are the axioms of the system. Construction is accomplished by means of a set of rules, which we shall call *RR,* the set of *reduction rules* The rules in RR are early versions of what we now call inference rules because by strictly formal (i.e. syntactic) manipulations of "marks on a page", they generate valid syllogisms from other valid syllogism. There are four rules. We shall discuss each informally and then state its precise syntactic definition.

Transposition

The first rule, called *transposition of the premises* (*transpositio praemissarum*), says that the order of the premises in a syllogism may be reversed.

The Greek word for *transposition* is *metathesis*, so the rule is referred to by the abbreviated code letter *m.* This rule is easy to define precisely. Single and double underlinings highlight which premise is switched where:

Transposition (*m).* From <*P*,*Q*,*R*> form <*Q*,*P*,*R*>

Simple Conversion

The second rule says that the order of subject and predicate in **E** or **I**-statements is irrelevant. Whether two sets are empty or not – the fact that is at issue in **E** and **I** statements – does not depend on the order in which the two sets are named. The irrelevance of term order is immediately obvious from the Venn

Diagrams for **E** and **I** statements. In general, changing the order of the a proposition's subject and predicate is called *conversion.* The particular sort of order inversion **E** and **I** statements is called *simple* conversion (*conversio* <u>*s*</u>*implex* in Latin), and the code letter for this rule is *s.* It is called *simple* because, unlike the next rule which is more complex, this rule changes only the order of the terms. In the definition below colors indicate which term is switched where

Simple Conversion (*s*)

From $< \mathbf{E}XY,Q,R>$ form $< \mathbf{E}YX,Q,R>$
From $< P,\mathbf{E}XY,R>$ form $< P,\mathbf{E}YX,R>$
From $< P,Q,\mathbf{E}XY>$ form $< P,Q,\mathbf{E}YX>$
From $< \mathbf{I}XY,Q,R>$ form $< \mathbf{I}YX,Q,R>$
From $< P,\mathbf{I}XY,R>$ form $< P,\mathbf{I}YX,R>$
From $< P,Q,\mathbf{I}XY>$ form $< P,Q,\mathbf{I}YX>$

Accidental Conversion

When the argument from *P* to *Q* is valid, i.e. when $P \models Q$, let us say that *P* is *stronger than Q.* (It is stronger because, in general, it entails more.) The third rule exploits the fact that a stronger proposition may replace a weaker one as the premise of a valid argument, and as the conclusion a weaker may replace a stronger. Consider in particular the two stronger-weaker pairs:

1. $\mathbf{A}XYP \models \mathbf{I}YX$ and
2. $\mathbf{E}XY \models \mathbf{O}YX$,

In these examples, the conclusion (weaker) can be seen to follow from the premise (stronger) as the result of two logical steps: a conversion of terms in the **I** and **O** propositions, and the subalternation entailment of an **I** from an **A,** and an **O** from an **E** proposition. Let us now form a rule by applying the earlier fact about stronger-weaker replacement to these pairs in particular:

> You may replace a premise of a syllogism by a proposition that is stronger, or a conclusion by one that is weaker, according to the entailments 1 or 2.

The rule, which we will restate more precisely below, is called *accidental conversion (conversio per accidens)*; it is abbreviated by the letter *p.* As in simple conversion colors indicate which term is switched where

Accidental Conversion (*p*)

From < **I**XY,Q,R> form < **A**YX,Q,R>
From < P,**I**XY,R> form < P,**A**YX,R>
From < P,Q,**A**XY> form < P,Q,**I**YX>
From < **O**XY,Q,R> form < **E**YX,Q,R>
From < P,**O**XY,R> form < P,**E**YX,R>
From < P,Q,**E**XY> form < P,Q,**O**YX>

Reduction to the Impossible

Recall that in categorical logic negations are limited to:

~**A**XY	is	**O**XY
~**E**XY	is	**I**XY
~**I**XY	is	**E**XY
~**O**XY	is	**A**XY

Recall also that in propositional logic the following biconditionals are tautologies and that the sentences that flank the $\leftrightarrow$ are logical equivalent.

$$[(P\wedge Q)\rightarrow R] \leftrightarrow [(P\wedge \sim R)\rightarrow \sim Q]$$
$$[(P\wedge Q)\rightarrow R] \leftrightarrow [(\sim R\wedge Q)\rightarrow \sim P]$$

The last rule says that you may apply these equivalences to the limited set of negations present in categorical logic. The rule, stated below, is called *reduction to the impossible* (*reductio per impossibilem* or *reductio per contradictionem*). Its abbreviation letter is *c.* As in transposition single and double underlinings highlight which premise is switched where:

Reduction to the Impossible (*c*)

From $\langle P,\underline{Q},\underline{\underline{R}}\rangle$ form $\langle P,\sim\underline{\underline{R}},\sim\underline{Q}\rangle$

From $\langle \underline{P},Q,\underline{\underline{R}}\rangle$ form $\langle \sim\underline{\underline{R}},Q,\sim\underline{P}\rangle$

The Reduction System

We are now ready to state the inductive definition of the set *PSyl* of provable syllogisms. Let <*P,Q,R*>, range over syllogisms.

Definition. The *Aristotelian reduction system ASyl* for the categorical logic is the proof system < *RSyl, RR, PSyl* > such that

1. *PSyl* (the *perfect* syllogisms) is the set of all syllogisms of the form <**A***YZ*, **A***XY,* **A***XZ*> or <**E***YZ*, **A***XY,* **E***XZ*>
1. *RR* is the set of rules:

Transposition (*m)*

From $\langle \underline{P},\underline{\underline{Q}},R\rangle$ form $\langle \underline{\underline{Q}},\underline{P},R\rangle$

Simple Conversion (*s*)

From < **E***XY,Q,R*> form < **E***YX,Q,R*>
From < *P,***E***XY,R*> form < *P,***E***YX,R*>
From < *P,Q,***E***XY*> form < *P,Q,***E***YX*>
From < **I***XY,Q,R*> form < **I***YX,Q,R*>
From < *P,***I***XY,R*> form < *P,***I***YX,R*>
From < *P,Q,***I***XY*> form < *P,Q,***I***YX*>

Accidental Conversion (*p*)

From < **I***XY,Q,R*> form < **A***YX,Q,R*>
From < *P,***I***XY,R*> form < *P,***A***YX,R*>
From < *P,Q,***A***XY*> form < *P,Q,***I***YX*>
From < **O***XY,Q,R*> form < **E***YX,Q,R*>
From < *P,***O***XY,R*> form < *P,***E***YX,R*>
From < *P,Q,***E***XY*> form < *P,Q,***O***YX*>

Reduction to the Impossible (*c*)

From $\langle P,\underline{Q},\underline{\underline{R}}\rangle$ form $\langle P,\sim\underline{\underline{R}},\sim\underline{Q}\rangle$

From $\langle \underline{P},Q,\underline{\underline{R}}\rangle$ form $\langle \sim\underline{\underline{R}},Q,\sim\underline{P}\rangle$

2. *RSyl* is defined inductively as follows:
   i. $PSyl \subseteq RSyl$
   ii. if $\langle P,Q,R\rangle \in RSyl$, and is formed $\langle P',Q',R'\rangle$ from $\langle P,Q,R\rangle$ by the rule *m, s, p,* or *c*, then $\langle P',Q',R'\rangle \in RSyl$;
   iii. nothing else is in *RSyl.*

The set *RSyl* of reducible syllogisms is an excellent example of an inductive system in proof theory. First of all, the definitions of the basic syllogisms, any Barbara or Celarent, are syntactic because they are defined by their form. Moreover, whether a syllogism meets the right syntactic form to count as Barbara or Celarent, i.e. whether it is a member of the "starter set" used in the inductive definition, is readily perceptible – it is "epistemically transparent." Likewise the four rules are formal. They consist of manipulations of symbol strings in prescribed ways, all of which are also "epistemically transparent." We apply these rules to the starter set to construct the inductive set *RSyl*. Membership in this set will in turn be transparent since, being inductive, every element of the set will have a construction sequence. This sequence will consist of a finite series of syllogisms. The syllogism at any stage will either be a member of the starter set, a fact which is transparent, or it will be constructed, transparently, from an earlier syllogism in the series by one of the rules. These construction sequences viewed in one direction are proofs that the last syllogism "follows from" Barbara or Celarent. Viewed in the other direction they are reductions to Barbara or Celarent.

Let us now see some examples. In the examples below rather that reducing every syllogism all the way back to Barbara or Celarent, we will reduce them only to other syllogisms that we have already shown can in turn be reduced to Barbara or Celarent.

Examples of Reductions: Proofs

Let us begin by showing that the other syllogisms of the first figure may be proven from (reduced to) Barbara and Celarent. The following are construction sequences showing that the syllogism on the last line is in the set *PSyl.*

1. <**E**YZ,**A**XZ,**E**XY> Celarent
2. <**E**YZ,**I**XY,**O**XZ> 1, *c* (Ferio)

1. <**A**ZX,**A**YZ,**E**XY> (Barbara)
2. <**I**XY,**A**YZ,**I**ZX> 1, *c*
3. <**I**XY,**A**YZ,**I**XZ> 2, *s*
4. <**A**YZ,**I**XY,**I**XZ> 1, *m* (Darii)

The two remain first figure syllogisms, Celaront and Barbari, are *subaltern moods* and follow from Barbara and Celarent respectively.

1. <**A**YZ,**A**XY,**A**XZ> (Barbara)
2. <**A**YZ,**A**XY,**I**ZX> 1, *p*
3. <**A**YZ,**A**XY,**I**XZ> 2, *s* (Babari)

1. <**E**YZ,**A**XY,**E**XZ> Celarent
2. <**E**YZ,**A**XY,**E**ZX> 1, *s*
3. <**E**YZ,**A**XY,**O**XZ> 2, *p* (Celaront)

The fun starts as we turn to the valid moods of figures two, three, and four because hidden in the names of the syllogisms, which were first devised in the Middle Ages for use in mnemonic poems, there are codes that tell how they can be reduced to the valid moods of the first figures. Recall the poem cited earlier, which divides the valid moods into the four figures and their residual subaltern moods:

> *Barbara, Celarent, Darii, Ferioque* are of the First:
> *Cesare, Camestres, Festino, Baroco* are of the Second:
> The Third has *Darapti, Disamis, Datisi, Felapton,*
> *Bocardo, Ferison*; The Fourth adds in addition
> *Bramantip, Camenes, Dimaris, Fesapo, Fresison.*
> Fifth are the *Subalterns,* which all come from the *Universals*,
> They do not have a name, nor, if well connected, a use.

In the 13th century Peter of Spain, a master in the Arts faculty at the University of Paris, explained the codes in his *Summa logicales*, a textbook that became a standard for many centuries.[13] As Peter explains, the vowels in a syllogism's name record its mood:

> It is important to know that the vowels *A, E, I* and *O* stand for the four types of propositions. The vowel *A* stands for a universal affirmative; *E* for a universal negative; I for a particular affirmative; and *O* for a particular negative.

In addition, built into the other letters are instructions on how to prove the syllogism. The syllogisms all have as their initial letters *B, C, D,* or *F*. The first four syllogisms of the first figure are special because it is from these that the valid syllogisms in later figures are all deduced, each from a syllogism in the first figure having the same initial letter. For example, Disamis in the third figure is proven from Darii in the first figure, indicated by the fact that they both begin with *D*.

The four consonants *m*, *s*, *p*, and *c* when they follow a vowel also have special meaning. Each names one of the four inference rules. Working backward from the syllogism to be proved, you can arrive at the right first figure syllogism by applying

---

[13] The quotations here and below are from Peter of Spain (later Pope John XXI), *Summa logicales,* also know as the *Tractatus,* Normal Kretzman and Eleonore Stump, eds., *Cambridge Translations of Medieval Philosophy Texts* (Cambridge: Cambridge University Press, 1988). The names and their order in Peter's 13th century poem are slightly different. Peter, like Aristotle and most logicians in the middle ages, views the fourth as part of the first figure, defining the first figure as all the valid syllogisms

these rules. The rule letter indicates that the syllogism was derived from an earlier one in which the rule is applied to the line (premise or conclusion) represented by the vowel that immediately precedes the rule letter. The other consonants in the names, like *b*, *l*, *t*, *n*, and *r* no special significance and may be ignored. Peter's own explanation is:

> Again, there are three syllables in each word (if there is any more, it is superfluous, except for *M*, as will be clear later). The first three syllables stand for the major proposition of the syllogism; the second stands for the minor; and the third for the conclusion. For example, the first word – Barbara – has three syllables, in each of which *A* is used; the three occurrences of *A* signify that the first mood of the first figure consists of two universal affirmative premises resulting in a universal affirmative conclusion. The vowels used in the other worlds should also be understood in this way.)
>
> Again, it is important to know that the first four words [i.e. names] of the first verse [refer to the first figure] and all the other subsequent words [i.e. names] begin with these consonants: *B, C, D,* and *F*. In this way we are given to understand that all the moods that a word beginning with *B* stand for should be reduced to the first mood of the first figure; all the moods signified by a word beginning with *C*, to the second mood of the first figure; *D*, to the third mood; *F*, to the fourth.
>
> Again, where *S* is used in these words, it signifies that the proposition that the immediately preceding vowel stands for requires simple conversion. And *P* signifies that the proposition requires conversion *per accidens*. And where *M* is used, it signifies that the premise requires transposition. (Transposition is making the major premise the minor, and vice versa.) And where *C* is used [after a vowel] it signifies that the mood that word stands for should be proved by reduction *per impossibile.*

Here are some examples:

1. <**A** *YX*,**I** *XY*,**I** *XZ*> Darii
2. <**A** *YX*,**I** *XY*,**I** *ZX*> 1, *s*
3. <**A** *YX*,**I** *YX*,**I** *ZX*> 2, *s*
4. <**I** *YZ*,**A** *YX*,**I** *XZ*> 3, *m* (Disamis)

1. <**E** *YZ*,**I** *XY*,**O** *XZ*> Ferio
2. <**E** *YZ*,**A** *YX*,**O** *XZ*> 1, *p* (Felapton)

---

in which the middle term is the subject of one premise but the predicate of the other). The poem cited in the text follows the more modern tradition, which distinguishes the fourth from the first figure.

Only two syllogisms, Bocardo and Baroco, contain *c* in their names. They form a trio of syllogisms with Barbara, all of which are interdeducible in one line by Reduction to the Impossible. Consider the derivation of Baroco from Barbara:

1. $<\mathbf{A}ZY,\underline{\mathbf{A}XZ},\underline{\underline{\mathbf{A}XY}}>$ Barbara
2. $<\mathbf{A}ZY,\underline{\underline{\mathbf{O}XY}},\underline{\mathbf{O}XZ}>$ 1, *c* (Baroco)

By somewhat tedious work we could in fact prove all 24 valid syllogisms. By doing so would show that our rule set is *complete* in the sense explained in the last lecture: every valid syllogism has a proof. It is also easy to prove (by the techniques of Lecture 13) the converse, that the system is sound: every syllogism that has a proof is valid.

Theorem: Soundness and Completeness. The set of valid syllogisms is the same as the set of reducible syllogisms.

Technically, the soundness and completeness theorem for this system is not very interesting because the inductive set itself is finite, and indeed quite small. There are only 24 valid moods. Rather than define the set inductively we could just list its members. I have presented the theory as one of inductive sets, however, because it then provides a simple and clear example of a “proof system” that was historically important. From the viewpoint of technical logic it may be noted that there are generalizations of Aristotle’s system that are not mathematically trivial. If we allow arguments in categorical syntax to have any number of premises, the set of valid arguments is infinite. This infinite set can be inductively defined using Aristotle’s rules together with some simple rules that capture immediate inferences, and the resulting system shown to be sound and complete.[14] It is true that the bare **A**, **E**, **I** and **O** syntax of categorical logic is expressively limited from the perspective of first-order logic, which can express most of mathematics. But by stringing syllogisms together, it practicing logicians and philosophers were able to compose syllogistic arguments of numerous premises. Traditional logic also investigated hypothetical propositions (conjunctions, disjunctions and conditionals) and other syntactic forms that we have not had time to discuss in these lectures, like complex noun and verb phrases, relative clauses, tenses and modal verb modifiers. Generally speaking, the language of pre-19th century logic was complex and subtle, an impressive tool for the science and philosophy of its day, and an interesting subject for a student of philosophy or the history of ideas.[15]

Exercises

Prove form (reduce to) a first figure syllogism with the same initial letter the following:

1. Datisi
2. Camestrop
3. Bocardo
4. Fesapo

[14] See John N. Martin, "Aristotle's Natural Deduction Reconsidered," *op. cit.,* and references therein. For a more complete citation of the relevant literature see note 4 of Part 2.
[15] See I. M. Bocheński, *A History of Formal Logic,* Second ed. (Notre Dame Indiana: University of Notre Dame Press, 1961) (Reprinted by Chelsea Publishing Co.); Alexander Broadie, *Introduction to Medieval Logic* (Oxford:Clarendon Press, 1987).

LECTURE 18. PROPOSITIONAL AND FIRST-ORDER LOGIC: PROOF THEORY[16]

Introduction

In this lecture we shall provide an inductive definition for the set of theorems of propositional and first-order logic. This will be a "textbook example" of a fully rigorous modern axiom system. First a warning. Proof theory is not everybody's cup of tea. It takes a particularly obsessive mind to like doing the minutely careful and often quite complicated symbolic manipulation necessary to work out a proof system work. Our interest here is wholly theoretically. Our goal is to see what an axiom system looks like and to gain an appreciation for why they are important. We will not be mastering the actual derivation of theorems in these systems. Logicians themselves hardly actually use these systems in their daily work. What is interesting is that they exist in principle.

We will first sketch basic proof theory the propositional logic. We do so in two stages. We begin by axiomatizing its logical truths – i.e. the tautologies. These will be defined inductively as the set of theorems that follow from a set of four kinds of axioms by the single inference rule *modus ponens.* This is the way proof theory was done in the early decades of the 20th century. The system we will define is utterly impractical but extremely elegant theoretically. Given the axiom system, we can then define the "logically acceptable" arguments as derivation relative to the system, in the manner sketched in a previous lecture: Q is *derivable* from $P_1,\ldots,P_n$ if Q would be a theorem if

$P_1,\ldots,P_n$ were added to the axioms. The resulting system is both statement and argument sound and complete.

We will then go on to a more interesting proof theory for the propositional logic. In this second approach the set of “acceptable arguments” is given a direct definition that is both inductive and syntactic. This second method is important for two reasons, one practical and one theoretical. Because the rule set is comprehensive, it is actually quite easy to use. It is the system logicians most often employ if they have to state a formal proof – and for this reason it is the system drilled into student in advance logic courses. But since our purpose here is not learning how to do proofs – that is for another course – we are interested in the system for the second, theoretical reason. Unlike an axiom system, which defines a set of sentences (the theorems), this second system defines a relation, a set of ordered pairs. This is the relation that holds between the premises and conclusion of a formally correct argument. This second system is argument sound and complete.

After having sketched the proof theory for propositional logic in two ways, we extended it to first-order logic. Again, our interest is mainly theoretical. What you should be noticing is how the key ideas– logical truth and “acceptable argument” – are being explained by definitions that are at once inductive and syntactic.

---

[16] For a review of the technical history in this and subsequent lectures see John Martin and John Franco, “A History of Satisfiability”, in Armin Biere, Hans van Maaren, and Toby Walsh, eds., *Handbook of Satisfiability,* IOS Press, to appear, 2008.

Proof Theory for Propositional Logic

Substitution

Before defining the axioms we must say something about substitution. Below we will specify four sets of axioms. Each axiom set is defined by reference to a specific sentence form called a *sentence schema*. One of the forms we shall use is $P\rightarrow(Q\rightarrow P)$. An axiom will be any sentence that "has the same form" as this sentence schema. But what do we mean by "has the same form"? This idea is explained by substitution. Any way of substituting sentences for p_1 and p_2 in $p_1\rightarrow(p_2\rightarrow p_1)$ is an axiom. This happens if we can alter the construction sequence of by putting in place of the individual sentences p_1 and p_2 the either construction sequences of the formulas replacing them. The resulting construction sequence is longer than the original but it produces a formula that has the same form as $p_1\rightarrow(p_2\rightarrow p_1)$ but with p_1 and p_2 replaced by longer formulas. Consider the example below. It consists of four construction sequences. The first is for $p_1\rightarrow(p_2\rightarrow p_1)$, the second is for $p_{25}\wedge p_6$, the third is for $\sim\sim p_6$, and the fourth the construction tree for $(p_{25}\wedge p_6)\rightarrow((\sim\sim p_6\rightarrow(p_{25}\wedge p_6))$, which results from replacing every occurrence of p_1 by the construction tree for $p_{25}\wedge p_6$, and every occurrence of p_2 by the construction tree for $\sim\sim p_6$ in the construction tree for $p_1\rightarrow(p_2\rightarrow p_1)$. It follows that $(p_{25}\wedge p_6)\rightarrow((\sim\sim p_6\rightarrow(p_{25}\wedge p_6))$ is a substitution instance of $p_1\rightarrow(p_2\rightarrow p_1)$.

			p_6
			p_{25}
			$p_{25}\wedge p_6$
			p_6
p_1			$\sim p_6$
p_2	p_6	p_6	$\sim\sim p_6$
$(p_2\rightarrow p_1)$	p_{25}	$\sim p_6$	$\sim\sim p_6\rightarrow(p_{25}\wedge p_6)$
$p_1\rightarrow(p_2\rightarrow p_1)$	$p_{25}\wedge p_6$	$\sim\sim p_6$	$(p_{25}\wedge p_6)\rightarrow((\sim\sim p_6\rightarrow(p_{25}\wedge p_6))$

We make this idea precise in the next definition. Let *CS(P)* be a construction sequence for *P.*

Definition

A sentence *Q* is *a substitution instance* of *P* if and only if there is some construction sequence *CS(Q)* of *Q* formed from some construction sequence *CS(P)* of *P* by replacing some atomic sentences $R_1,\ldots,R_n$ of *P* in *CS(P)* by (possibly molecular) sentences $S_1,\ldots,S_n$, and inserting some construction sequences $CS(S_1),\ldots, CS(S_n)$ into the new sequence prior to occurrences of $S_1,\ldots,S_n$.

Łukasiewicz's Axiom System

The *Łukasiewicz's axiom system for propositional logic* is <*AxPL, PR, ThPL*> such that

1. *AxPL* the set that contains all and only sentences that are substitution instances of one of the following: [17]
   - Axiom Schema 1. $p_1\rightarrow(p_2\rightarrow p_1)$
   - Axiom Schema 2. $(p_1\rightarrow(p_2\rightarrow p_3))\rightarrow((p_1\rightarrow p_2)\rightarrow(p_1\rightarrow p_3))$
   - Axiom Schema 3. $(\sim p_1\rightarrow\sim p_2)\rightarrow(p_2\rightarrow p_1)$
2. *PR* contains only the rule *modus ponens:* from two sentences of the form $\underline{P}$ and $\underline{P}\rightarrow\underline{\underline{Q}}$ the sentence of the form $\underline{\underline{Q}}$ follows.
3. *AxPL* is defined by induction as follows
   a. *AxPL*$\subseteq$ *ThPL*.

[17] Frege axiomatized the propositional logic using five axioms in 1879. This reduced set is due to Jan Łukasiewicz. See Jan Łukasiewicz and Alfred Tarski, "Untersuchungen über den Aussagenkalkül," in *Comptes Rendus des Séances de la Société des Sciences et des Lettres de Varsovie*, 23:III (Warsaw 1930), pp. 30-50. For fuller presentation see Alonzo Church, *Introduction to Logic* Chapter II, §27 (Princeton: Princeton University Press, 1956).

b. If P and Q are in *ThPL* and R is constructed ("derived") from P and Q by *modus ponens*, then R is in *ThPL*.
c. Nothing else is in *ThPL.*

It is customary to abbreviate the fact that P is a theorem, i.e that $P \in ThPL$, by the turnstyle notation $\vdash P$. Colors are added to aid the eye in seeing multiple occurrences in a single line of the same sentence. As in used of *modus ponens* in earlier lectures, underlinings are added here to aid the eye in spotting the relevant antecedent and consequent of the conditional used in the rule.

Examples of Proofs

To make it easier to read proofs, we shall abbreviate the names of atomic sentences, which employ subscripts, by single letters. Let *p, q, and r* abbreviate respectively p_1, p_2, and p_3. In the following proofs colors will be used to indicate that a sentence is a instance of the formula of the same color in an axiom, and underlining will be used to indicate the antecedent and consequents of conditionals used in an application of *modus pones.*

Theorem. $\vdash p \rightarrow p$

1. $p \rightarrow ((p \rightarrow q) \rightarrow p)$ (*Axiom Schema* 1)
2. $(p \rightarrow ((p \rightarrow p) \rightarrow p)) \rightarrow ((p \rightarrow (p \rightarrow p)) \rightarrow (p \rightarrow p))$ (*Axiom Schema* 2)
3. $(p \rightarrow (p \rightarrow p)) \rightarrow (p \rightarrow p)$ 1 & 2, *modus ponens*
4. $p \rightarrow (p \rightarrow p)$ (*Axiom Schema* 1)
5. $p \rightarrow p$ 3 & 4, *modus ponens*

Theorem. $\vdash {\sim}p \rightarrow (p \rightarrow q)$

1. $({\sim}q \rightarrow {\sim}p) \rightarrow (p \rightarrow q)$ (*Axiom Schema* 3)
2. $(({\sim}q \rightarrow {\sim}p) \rightarrow (p \rightarrow q)) \rightarrow ({\sim}p \rightarrow (({\sim}q \rightarrow {\sim}p) \rightarrow (p \rightarrow q)))$ (*Axiom Schema* 1)
3. ${\sim}p \rightarrow (({\sim}q \rightarrow {\sim}p) \rightarrow (p \rightarrow q))$ 1 & 2, *modus ponens*

4. $(\sim p\rightarrow((\sim q\rightarrow\sim p)\rightarrow(p\rightarrow q)))\rightarrow((\sim p\rightarrow(\sim q\rightarrow\sim p))\rightarrow(\sim p\rightarrow(p\rightarrow q)))$ (*Axiom Schema* 2)
5. $(\sim p\rightarrow(\sim q\rightarrow\sim p))\rightarrow(\sim p\rightarrow(p\rightarrow q))$ 3 & 4, *modus ponens*
6. $\sim p\rightarrow(\sim q\rightarrow\sim p)$ (*Axiom Schema* 1)
7. $\sim p\rightarrow(p\rightarrow q)$ 5& 6, *modus ponens*

Defining "Derivability" in the Axiom System

As explained in the previous lecture, it is possible to define by reference to the axiom system the notion of a formally correct proof, called a *derivation,* of the conclusion Q from the premise set $\{P_1,\ldots,P_n\}$. The conclusion follows from the premises if by adding the premises to the axiom set we could then prove the conclusion as a theorem in the augmented axiom system. Let $<AxPL\cup\{P_1,\ldots,P_n\}, PR, ThPL'>$ be this axiom system.

Definition. Q is (*syntactically*) *derivable* from $\{P_1,\ldots,P_n\}$ (abbreviated $P_1,\ldots,P_n \vdash Q$) iff $Q\in ThPL'$ $<AxPL\cup\{P_1,\ldots,P_n\}, PR, ThPL'>$ is an axiom system and $Q\in ThPL'$.

Since the syntax contains the material conditional $\rightarrow$ and the system has the rule *modus ponens*, there is a way to relate the theorems of the system to the derivable arguments. We do so by showing that the conditional $(P_1\rightarrow(P_2\rightarrow(\ldots\rightarrow P_n)))\rightarrow Q$ is a theorem whenever there is a derivation of Q form $\{P_1,\ldots,P_n\}$.

Theorem (The Deduction Theorem). There is a derivation of Q from $\{P_1,\ldots,P_n\}$ iff $(P_1\rightarrow(P_2\rightarrow(\ldots\rightarrow P_n)))\rightarrow Q$ is a theorem.

Proof Sketch. If $(P_1\rightarrow(P_2\rightarrow(\ldots\rightarrow P_n)))\rightarrow Q$ is a theorem, then we can construct a proof exhibiting the derivation of Q form $\{P_1,\ldots,P_n\}$ as follows:

…		
m.	$(P_1\rightarrow(P_2\rightarrow(\ldots\rightarrow P_n)))\rightarrow Q$	(previously prove this as a theorem)
m+1. P_1		assumption
m+2. P_2		assumption
…		…

m+n.	P_n	assumption
m+n+1.	$P_2\rightarrow(\ldots\rightarrow P_n)))\rightarrow Q$	*m*+1 & *m+n*, *modus ponens*
…		…
m+n+(*n*-1).	$(\ldots\rightarrow P_n)))\rightarrow Q$	*m*+(*n*-1) & *m+n*+(*n*-2), *modus ponens*
m+n+n.	Q	*m+n* & *m+n*+(*n*-1), *modus ponens*

Conversely, if there is a derivation of Q form $\{P_1,\ldots,P_n\}$ it is possible to convert it to a proof of $(P_1\rightarrow(P_2\rightarrow(\ldots\rightarrow P_n)))\rightarrow Q$, though we shall not do so here.

Hence, any axiomatization of tautologies suffices for a "syntactic explanation" of derivability as well.

Soundness and Completeness

The similarity in design of the symbol $\vdash$ to that of $\models$ is intentional. Though the two have different definitions – $\{P_1,\ldots,P_n\}\models Q$ is defined semantically ("for any $\Im$, if $P_1,\ldots,P_n$ are T in $\Im$, then Q is T in $\Im$") and $P_1,\ldots,P_n \vdash Q$ is defined syntactically ("$(P_1\rightarrow(P_2\rightarrow(\ldots\rightarrow P_n)))\rightarrow Q$ is a theorem of the axiom system") – the two relations are intended to be the same. If the axiom system is well designed – if it is sound and complete – then the two relations are in fact identical. Indeed, the whole point of the axiom system is to define $\vdash$ so that it will turn out to be the same as $\models$. In the case of Łukaisiewicz's axiom system this goal is achieved, as the following metatheorem states:

Theorem. Statement Soundness and Completeness of Łukaisiewicz's Axioms.

1. Statement Soundness and Completeness. P is a theorem iff P is a tautology.
2. Finite Argument Soundness and Completeness. $P_1,\ldots,P_n\models Q$ iff $P_1,\ldots,P_n \vdash Q$.
3. Argument Soundness and Completeness

$$X \models Q \text{ iff, for some subset } \{P_1,\ldots,P_n\} \text{ of } X,\ P_1,\ldots,P_n\vdash Q$$

Though the theorem is not difficult to prove, that is a task for a more technical discussion. Notice that part 3 says that the result remains true even if the premise set is allowed to be infinitely large.

The theorem is important because it shows that the two relations $\vdash$ (syntactic derivability) and $\models$ (logical entailment) are the same. Recall what it means according to naïve set theory as set forth in Part 1 for two relations to be the same. First of all, in set theory the fact that $\models$ and $\vdash$ are two-place relations means that they are sets of pairs $<X,P>$. In this case X is a set of premises and P is a sentence. Hence, the fact that $P_1,\ldots,P_n$ logically entails Q, which we write $P_1,\ldots,P_n \models Q$, could equally well be written in set theoretic notation as $<X,P> \in \models$. Similarly the fact that Q is derivable from $P_1,\ldots,P_n$, which we write as $P_1,\ldots,P_n \vdash Q$, could be written as $<\{P_1,\ldots,P_n\},Q> \in \vdash$. The previous theorem therefore could equally well be stated in ordered-pair notation:

For any $<\{P_1,\ldots,P_n\},Q>$, $<\{P_1,\ldots,P_n\},Q> \in \vdash$ iff $<\{P_1,\ldots,P_n\},Q> \in \models$.

But this statement is exactly what is required by the Principle of Extensionality for the identity of the two sets of pairs $\models$ and $\vdash$.

Theoretically the completeness theorem is a major result. It shows that two rather different approaches to validity – the semantic and a proof theoretic– characterize the same concept.

Exercise. Add the annotation to the following proof indicating for each line (1) the axiom schema it instantiates or (2) the prior lines it follows from by *modus ponens*.

Theorem. $\vdash (q\rightarrow r)\rightarrow((p\rightarrow q)\rightarrow(p\rightarrow r))$

1. $((p\rightarrow(q\rightarrow r))\rightarrow((p\rightarrow q)\rightarrow(p\rightarrow r)))\rightarrow((q\rightarrow r)\rightarrow((p\rightarrow(q\rightarrow r))\rightarrow((p\rightarrow q)\rightarrow(p\rightarrow r))))$
2. $(p\rightarrow(q\rightarrow r))\rightarrow((p\rightarrow q)\rightarrow(p\rightarrow r))$
3. $(q\rightarrow r)\rightarrow((p\rightarrow(q\rightarrow r))\rightarrow((p\rightarrow q)\rightarrow(p\rightarrow r)))$
4. $((q\rightarrow r)\rightarrow((p\rightarrow(q\rightarrow r))\rightarrow((p\rightarrow q)\rightarrow(p\rightarrow r))))\rightarrow(((q\rightarrow r)\rightarrow(p\rightarrow(q\rightarrow r)))\rightarrow((q\rightarrow r)\rightarrow((p\rightarrow q)\rightarrow(p\rightarrow r))))$
5. $((q\rightarrow r)\rightarrow(p\rightarrow(q\rightarrow r)))\rightarrow((q\rightarrow r)\rightarrow((p\rightarrow q)\rightarrow(p\rightarrow r)))$
6. $(q\rightarrow r)\rightarrow(p\rightarrow(q\rightarrow r))$
7. $(q\rightarrow r)\rightarrow((p\rightarrow q)\rightarrow(p\rightarrow r))$

∗Natural Deduction Proof Theory for Propositional Logic

Gentzen and Natural Deduction

Once you've seen one inductive set, you've seen them all. They differ in detail but have the same form. Below we offer a second way to capture the semantic entailment relation by a coextensive, i.e. identical, relation defined by induction in purely syntactic, i.e. proof theoretic, terms. In this approach it is not the tautologies that is axiomatized, but the set of valid arguments itself.

In this definition what is defined inductively is not a set of sentences, but a set of ordered pairs <*X,P*>. We have already seen a similar inductive definition when we defined the set of syllogisms reducible to Barbara and Celarent. Recall that a syllogism is a triple <*P,Q,R*> of three categorical propositions. It follows the set of syllogisms reducible to Barbara and Celarent is a set of triples. It was this set that defined inductively.

There is a theoretical reason in favor of directly defining the set of "logically acceptable" arguments directly rather than doing so by the indirect method of first defining "theorem" in an axiom system and then defining the notion of "derivation" in terms of it. The reason is that the main subject matter of logic is validity, not logical truth. Indeed, logical truth is really just a special case of validity. It is easy to show, for example, that *P* is a logical truth if and only if $\varnothing \models P$. That is, a logical truth is the "degenerate case" of a proposition that is true – that "follows" – "no matter what".

The system we shall use to define derivability directly is due to Gerhard Gentzen (1909-1945)[18], called *natural deduction.* It gets its name in part due to the fact that it is relatively easy to construct proofs using its rules.

Motivation: Intuitionistic Logic

The rule set we are about to explore also has a special theoretical interest for philosophers of logic because in a sense it provides a "theory of meaning" for the logical connectives. As you will see, for each connective there will be two rules, a so-called "introduction rule" that tells you how to add the connective to a new step of a proof, and a so-called "elimination rule" that tells you how to deduce a new line of the proof without that deletes from the proof the connective from proven line. Advocates of the system say that the rule set therefore explains "how the connectives are used" in logic. After all, they say, there is nothing more to logic than proofs, and therefore knowing how to use connectives in logic means nothing more than knowing how to add and subtract them from proofs. Moreover, some philosophers, like Ludwig Wittgenstein (1889-1951) in his *Philosophical Investigations*, have argued that the proper way to explain a word's meaning is to explain its use. Meaning, in short, is use. It would follow then that if the rules explain the use of the connectives, they explain its meaning.

[18] See Gerhard Gentzen, "Untersuchungen über der logische Schliessen", *Mathematische Zeitschrift* 39 (1934-35), 176-210, 405-431, and D. Prawitz, "Natural Deduction" (Stickholm: Almqvist and Wiksell,

This line of argument is especially attractive to logicians who would like to explain meaning, but who have serious doubts about set theory and therefore have serious doubts about the semantic theory we have been setting out in these lectures which makes extensive use of sets. Naïve set theory, after all, harbors contradictions and even modern axiomatic set theory cannot be proven consistent. They reason that semantic theory that makes use of sets is then equally dubious. Logicians who question set theory in this way are called *intuitionists* or *constructivists*. They also usually question several other features of traditional logic, especially the law excluded middle (*P*∧~*P*) and indirect proof (i.e. proving *P* by showing ~*P* is absurd). Suffice it to say that this is an important and interesting minority opinion, which we will not be able to investigate further here.[19] We shall continue to make use of sets in semantics and shall continue to use all the traditional logical rules.

The Inductive Strategy

What we are going to inductively define a set of *arguments.* Arguments have two parts: a set *X* of premises and a conclusion *P*. The argument from *X* to *P* is represented by the ordered pair <*X,P*>. In natural deduction theory the argument <*X,P*> is called a *deduction.* The set to be defined is call the set of "acceptable deduction," and it will be a set of ordered pairs. It will be defined using only syntactic ideas, but as before our intention is that when finished, this set will turn out to be the same as the set of valid arguments defined semantically.

1965). An excellent introduction to first-order logic using natural deduction is Neil Tennant, *Natural Logic* (Edinburgh: Edinburgh University Press, 1978).

[19] Students interested in pursuing the subject further may consult Grigori Mints, *A Short Introduction to Intuitionistic Logic* (New York: Kluwer, 2000).

Since the definition of an acceptable deduction is inductive, it begins with a set of starter elements. These will be a group of completely trivial arguments, which are called *basic deductions* (in the set *BD*). These are arguments in which the conclusion simply repeats one of the premises. They are trivial because, in a sense, they wear their validity on their sleeves. For example, the argument from the premise set {*P,Q*} to the conclusion *P* is basic because the conclusion *P* is in the premise set {*P,Q*}. Such arguments are obviously valid, because if all the sentences in the premise set are true, so is the sentence repeated as the conclusion.

The construction rules make up new arguments from old. There are two rules for each connective. There are also two rules for a new "connective" represented by the symbol ⊥, and a rule called *thinning* that adds extra premises to an argument. The new symbol ⊥ is called *the contradiction sign.* It is intended to represent a contradiction. It does not matter what this contradiction is so long as is a sentence that is false in every interpretation. We could, for example, simply define ⊥ as $p_1 \wedge \sim p_1$. Most of the rules all have the same form:

> From argument <*X,P*> construct the argument <*Y,Q*>.

An intuitive way to reformulate this rule would be:

> If by assuming the premises in *X* we can prove *P,* then by assuming the premises in *Y* we can prove *Q.*

Consider the a version of double negation:

> If by assuming the premises in *X* we can prove ~~*P,* then from the same premises, namely those in *X,* we can prove *P.*

The rule is written more simply:

> From <*X,*~~*P*> construct <*X,P*>.

Several of the rules, however, need two input deductions to make up a new deduction, and one rule needs three. Consider a version of *modus ponens*. It builds on two deductions:

> If by assuming the premises in *X* we can prove *P* and by assuming the premises in *Y* we can prove $P \rightarrow Q$, then by assuming the combined set of premises $X \cup Y$ we can prove *Q.*

The rule is written more simply:

> From $<X,P>$ and $<Y,P \rightarrow Q>$ construct $<X \cup Y,Q>$.

Or consider the following rules written first informally and then more precisely:

Reduction to the Absurd

> If by assuming (as background assumptions) the premises in *X* and assuming for the sake of argument *P,* we can prove the contradiction sign $\perp$, then on the basis of the background assumptions in *X* alone, we know $\sim P$.
>
> From $<X,\perp>$ construct $<X-\{P\},\sim P>$.

Conditional Proof

> If by assuming (as background assumptions) the premises in *X* and assuming for the sake of argument *P,* we can prove *Q,* then on the basis of the background assumptions in *X* alone, we know $P \rightarrow Q$.
>
> From $<X,P>$ and $<Y,P \rightarrow Q>$ construct $< X-\{Q\},Q \rightarrow P>$.

Genzen's Natural Deduction System

The *Gentzen' natural deduction system for propositional logic* is $<BD, PR, DPL>$ such that

1. The set *BD* of *basic deductions for propositional logic* is set of all pairs $<X,P>$ such that $X \subseteq Sen$ and $P \in X$.
2. The set *PR* of *natural deduction rules for propositional logic* is the set containing the rules:

$\bot$ Rules:

Introduction. From $\langle X,P\rangle$ and $\langle Y,\sim P\rangle$ construct $\langle X\cup Y, \bot\rangle$. (This "explains" the meaning of $\bot$.)

Elimination. From $\langle X,\bot\rangle$ construct $\langle X-\{\sim P\},P\rangle$. (A version of *Ex Falso Quodlibet*)

$\sim$ Rules:

Introduction. From $\langle X,\bot\rangle$ construct $\langle X-\{P\},\sim P\rangle$. (Reduction to the Absurd)

Elimination. From $\langle X,\sim\sim P\rangle$ construct $\langle X,P\rangle$. (Double Negation)

$\wedge$ Rules:

Introduction. From $\langle X,P\rangle$ and $\langle Y,Q\rangle$, construct $\langle X\cup Y,P\wedge Q\rangle$.

Elimination. From $\langle X,P\wedge Q\rangle$. construct $\langle X,P\rangle$.

Elimination. From $\langle X,P\wedge Q\rangle$. construct $\langle X,Q\rangle$.

$\vee$ Rules:

Introduction. From $\langle X,P\rangle$ construct $\langle X,P\vee Q\rangle$. (Addition, to the right side)

Introduction. From $\langle X,Q\rangle$ construct $\langle X,P\vee Q\rangle$. (Addition, to the left side)

Elimination. From $\langle X,P\vee Q\rangle$, $\langle Y,R\rangle$ and $\langle Z,R\rangle$, construct $\langle X\cup(Y-\{P\})\cup(Z-\{Q\}), R\rangle$. (Argument from cases)

$\rightarrow$ Rules:

Introduction. From $\langle X,P\rangle$ construct $\langle X-\{Q\},Q\rightarrow P\rangle$. (Conditional Proof)

Elimination. From $\langle X,P\rangle$ and $\langle Y,P\rightarrow Q\rangle$ construct $\langle X\cup Y,Q\rangle$. (*Modus Ponens*)

Thinning. From $\langle X,P\rangle$ construct $\langle X\cup Y,P\rangle$. (You can always add more premises.)

3. The set *DPL* of *natural deductions for propositional logic* set of all pairs such that
   a. $BD\subseteq DPL$
   b. If $\langle X,P\rangle$ is constructed by one of the rules a-k in *PR* from elements of *DPL,* then $\langle X,P\rangle$ is in *DPL.*
   c. Nothing else is in *DPL.*

It is customary to write the fact that the deduction $\langle\{P_1\ldots,P_n\},Q\rangle$ is "acceptable", i.e. that $\langle\{P_1\ldots,P_n\},Q\rangle\in DPL$, in turnstule notation as $P_1\ldots,P_n\vdash Q$. Likewise, it turns out that when *P* is a tautology it can be proven from the empty set, i.e. *P* is a tautology iff $\langle\varnothing,P\rangle\in DPL$. This too is customarily written in turnstyle notation, as $\vdash P$.

Examples of Theorems

Theorem. $P\vdash\sim\sim P$

Proof

1. $\langle\{P\}, P\rangle$ bd
2. $\langle\{\sim P\}, \sim P\rangle$ bd
3. $\langle\{P, \sim P\}, \bot\rangle$ 1,2 $\bot+$
4. $\langle\{P\}, \sim\sim P\rangle$ 3 $\sim+$

Theorem. $P,\sim Q\vdash\sim(P\rightarrow Q)$

Proof

1. $\langle\{P, P\rightarrow Q\}, P\rangle$ bd
2. $\langle\{P, P\rightarrow Q\}, P\rightarrow Q\rangle$ bd
3. $\langle\{P, P\rightarrow Q\}, Q\rangle$ 1,2 $\rightarrow-$
4. $\langle\{\sim Q\}, \sim Q\rangle$ bd

5. $< \{P, P\to Q, \sim Q\} , \bot >$ 3,4 $\bot$+
6. $< \{P,\sim Q\} , \sim(P\to Q) >$ 5 ~+

Theorem. $\sim Q \vdash \sim(P\wedge Q)$

Proof

1. $< \{\sim Q, P\wedge Q \} , P\wedge Q >$ bd
2. $< \{\sim Q, P\wedge Q \} , Q >$ 1 $\wedge$-
3. $< \{\sim Q, P\wedge Q \} , \sim Q >$ bd
4. $< \{\sim Q, P\wedge Q\} , \bot >$ 2,3 $\bot$+
5. $< \{\sim Q\} , \sim(P\wedge Q) >$ 4 ~+

∗**Exercise.** Annotate the following proof.

Theorem. $\vdash P\vee\sim P$

Proof

1. $< \{P,\sim(P\vee\sim P)\} , \sim(P\vee\sim P) >$
2. $< \{P,\sim(P\vee\sim P)\} , P >$
3. $< \{P,\sim(P\vee\sim P)\} , P\vee\sim P >$
4. $< \{P,\sim(P\vee\sim P)\} , \bot >$
5. $< \{\sim(P\vee\sim P)\}, \sim P >$
6. $< \{\sim(P\vee\sim P)\}, P\vee\sim P >$
7. $< \{\sim(P\vee\sim P)\}, \sim(P\vee\sim P) >$
8. $< \{\sim(P\vee\sim P)\} , \bot >$
9. $< \varnothing, \sim\sim(P\vee\sim P) >$
10. $< \varnothing, P\vee\sim P) >$

∗**Exercise.** Construct natural deduction proofs of the following

1. $P\to Q,\sim Q \vdash \sim P$
2. $R\to\sim P, Q\to\sim R, P\vee Q \vdash \sim R$

As promised, the set of deductions can be shown to be coextensional with the valid arguments of propositional logic:

Theorem. Soundness and Completeness.

1. Statement Soundness and Completeness. P is a tautology iff $\vdash P$
2. Finite Argument Soundness and Completeness. $P_1,\ldots,P_n \vDash Q$ iff $P_1,\ldots,P_n \vdash Q$.
3. Argument Soundness and Completeness

   $X \vDash Q$ iff, for some subset $\{P_1,\ldots,P_n\}$ of X, $P_1,\ldots,P_n \vdash Q$.

According we have seen two somewhat different ways to "capture" the valid arguments of propositional logic. Both are inductive definitions that make use only of epistemically transparent syntactic ideas. As a result, we are able to explain why philosophers have always thought that the arguments of logic carry with them a variety of certainty unique to the subject matter. The same proof theoretic techniques used thus far in this lecture to characterize the validity relation in propositional logic can be extended to capture validity in first-order logic.

Proof Theory for First-Order Logic

The Axiom System of Russell and Whitehead for FOL

We first extend Łukasiewicz' Axiom System to first-order logic by adding axioms due to Russell and Whitehead for the quantifiers. The trick is to capture the logic of universal instantiation and generalization. Quite cleverly they do so in three axioms. First we must extend the notion of substitution to first-order syntax.

Definition

A formula *Q* is *a substitution instance* of *P* if and only (1) all variables free in *P* are free in *Q* and (2) if there is some construction sequence *CS(Q)* of *Q* formed from some construction sequence *CS(P)* of *P* by replacing some atomic formulas $R_1,\ldots,R_n$ of *P* in *CS(P)* by (possibly molecular) formulas $S_1,\ldots,S_n$, and inserting some construction sequences $CS(S_1),\ldots, CS(S_n)$ in the new sequence prior to occurrences of $S_1,\ldots,S_n$.

The axioms are really axiom schemata – the represent any formula that fits their form. Moreover, as stipulated the formulas may contain free variables. If *P* is a formula of first-order syntax that contains the free variables $x_1,\ldots, x_n$ let us call $\forall x_1,\ldots, x_nP$ a *universal closure* of *P.* In the new axiom system not only are instances of a schema to count as axioms but so are their universal closures. We now define the new set of axioms, that incorporates Russell and Whitehead's quantifier axioms and two axioms for identity (the laws of self-identity and substitution.)

The *Russell and Whitehead axiom system for first-order logic* is <*AxFOL,PR, ThFOL*> such that

1. *AxFOL* is the set that contains all and only the instances and closures of formulas that are substitution instances of one of the following:[20]
   Axiom Schema 1. $p_1\rightarrow(p_2\rightarrow p_1)$
   Axiom Schema 2. $(p_1\rightarrow(p_2\rightarrow p_3))\rightarrow((p_1\rightarrow p_2)\rightarrow(p_1\rightarrow p_3))$
   Axiom Schema 3. $(\sim p_1\rightarrow\sim p_2)\rightarrow(p_2\rightarrow p_1)$
   Axiom Schema 4. $\forall x(p_1\rightarrow p_2)\rightarrow\forall xp_1\rightarrow\forall p_2$
   Axiom Schema 5. $\forall xp_1\rightarrow p_1$
   Axiom Schema 6. $p_1\rightarrow\forall xp_1$ if the formula replacing p_1 contains no free x
   Axiom Schema 7. $x=x$
   Axiom Schema 8. $x=y\rightarrow(p_1[x]\rightarrow p_2[y])$ if the formula replacing $p_2[y]$ contains some free y where that replacing $p_1[x]$ contains free x.
2. *PR* contains just the rule *modus ponens.*
3. The set *ThFOL* is defined inductively as follows:
   a. $AxFOL\subseteq ThFOL$.
   b. If *P* and *Q* are in *ThFOL* and *R* follows from *P* and *Q* by *modus ponens*, then *R* is in *ThFOL*.
   c. Nothing else is in *ThFOL.*

As before, it is customary to abbreviate the fact that *P* is a theorem, i.e that $P\in ThFOL$, by the turnstyle notation $\vdash P$. We also define the notion of derivation as before. Let $<AxFOL\cup\{P_1,\ldots,P_n\}, PR, ThFOL'>$ be the axiom system formed by adding $P_1,\ldots,P_n$ to the axiom set *AxFOL.*

[20] The first-order axioms are due to Russell and Whitehead, *9 of *Principia Mathematica,* vol 1. (Cambridge: Cambridge University Press, First edition 1910, Second Edition 1927). They employ a longer set of axioms for the propositional logic. Here we substitute Łukaisiewicz's shorter set developed later. Proofs of completeness of first-order logic under suitable axiom systems date back at least to G¨odel in 1929.

Definition. Q is (*syntactically*) *derivable* from $\{P_1,\ldots,P_n\}$ (abbreviated $P_1,\ldots,P_n \vdash Q$) iff $Q \in ThPL'$ $<AxPL \cup \{P_1,\ldots,P_n\}, PR, ThPL'>$ is an axiom system and $Q \in ThPL'$.

Soundness and Completeness

The system is sound and complete.

Theorem. Soundness and Completeness.

4. Statement Soundness and Completeness. P is a tautology iff $\vdash P$
5. Finite Argument Soundness and Completeness. $P_1,\ldots,P_n \vDash Q$ iff $P_1,\ldots,P_n \vdash Q$.
6. Argument Soundness and Completeness

$$X \vDash Q \text{ iff, for some subset } \{P_1,\ldots,P_n\} \text{ of } X,\ P_1,\ldots,P_n \vdash Q.$$

An Example of a Theorem

Proofs in first-order logic are generally more complex than those for propositional logic, but as an example we give a simple proof that identity is symmetric. Given that the axioms hold for the closures of a formula, it follows that if P is a theorem, then any universal quantification of P is a theorem:

Theorem. [21] If P is a theorem, then $\forall xP$ is a theorem.

We will make use of this fact in proofs of the example.

Theorem. $\vdash \forall x \forall y(x=y \rightarrow y=x)$

1.	$x=x \rightarrow (x=y \rightarrow x=x)$	Axiom Schema 1
2.	$x=x$	Axiom Schema 7
3.	$x=y \rightarrow x=x$	1 & 2, *modus ponens*
4.	$x=y \rightarrow (x=x \rightarrow y=x)$	Axiom Schema 8
5.	$(x=y \rightarrow (x=x \rightarrow y=x)) \rightarrow ((x=y \rightarrow x=x) \rightarrow (x=y \rightarrow y=x))$	Axiom Schema 2
6.	$(x=y \rightarrow x=x) \rightarrow (x=y \rightarrow y=x)$	4 & 5, *modus ponens*
7.	$x=y \rightarrow y=x$	3 & 6, *modus ponens*
8.	$\forall y(x=y \rightarrow y=x)$	7, previous metatheorem

[21] Strictly speaking this theorem requires a proof, which we will forgo here. See W. V. Quine *Mathematical Logic, op. cit.,* theorem *115.

9. $\forall x\forall y(x=y\rightarrow y=x)$ 8, previous metatheorem

$\ast$A Gentzen Natural Deduction System for FOL

We now extend the natural deduction system defined earlier for the propositional logic by adding introduction and elimination rules for the universal and existential quantifiers, and for the identity predicate. The rules for the quantifiers – once their notation is deciphered – are quite natural. They spell out the ideas behind the quantifier instantiation and generalization rules that we first met in Part 1 doing proofs in naïve set theory. The rules for identity are again a version of the law of self-identity and of the substitution of identity.

Definitions

The *Gentzen' natural deduction system for first-order logic* is $<BD, PR, DFOL>$ such that

1. The set *BD* of *basic deductions for propositional logic* is set of all pairs $<X,P>$ such that $X\subseteq For$ and $P\in X$.
2. The set *PR* of *natural deduction rules for propositional logic* is the set containing the rules:

$\perp$ Rules:
- Introduction. From $<X,P>$ and $<Y,\sim P>$ construct $<X\cup Y, \perp>$. (This "explains" the meaning of $\perp$)
- Elimination. From $<X,\perp>$ construct $<X-\{\sim P\},P>$. (A version of *Ex Falso Quodlibet*)

$\sim$ Rules:
- Introduction. From $<X,\perp>$ construct $<X-\{P\},\sim P>$. (Reduction to the Absurd)
- Elimination. From $<X,\sim\sim P>$ construct $<X,P>$. (Double Negation)

$\wedge$ Rules:
- Introduction. From $<X,P>$ and $<Y,Q>$, construct $<X\cup Y,P\wedge Q>$.
- Elimination. From $<X,P\wedge Q>$. construct $<X,P>$.
- Elimination. From $<X,P\wedge Q>$. construct $<X,Q>$.

$\vee$ Rules:
- Introduction. From $<X,P>$ construct $<X,P\vee Q>$. (Addition, to the right side)
- Introduction. From $<X,Q>$ construct $<X,P\vee Q>$. (Addition, to the left side)
- Elimination. From $<X,P\vee Q>$, $<Y,R>$ and $<Z,R>$, construct $<X\cup(Y-\{P\})\cup(Z-\{Q\}), R>$. (Argument from cases)

$\rightarrow$ Rules:
- Introduction. From $<X,P>$ construct $<X-\{Q\},Q\rightarrow P>$. (Conditional Proof)
- Elimination. From $<X,P>$ and $<Y,P\rightarrow Q>$ construct $<X\cup Y,Q>$. (*Modus Ponens*)

$\forall$ Rules:
- Introduction. From X,P$[t/v]>$ construct $<X,\forall vP>$. (Universal Generalization)

Elimination. From $<X,\forall vP>$ construct $<X,P[t/v]>$ where v is not free in any $P\in X$ (Universal Instantiation)

$\exists$ Rules:

Introduction. From $<X,P[t//v]>$ construct $<X,\exists v'P>$. (Proof by Construction)

Elimination. From $<X,\exists v'P>$ & $<Y\cup\{P[t/v]\},Q>\in$ DFOL construct $<X\cup Y,Q>$. (if t is not free in $X,Y,$ $\exists v'P$ or Q) (Existential Instantiation)

= Rules:

Introduction. From $<X,P>$ construct $<X, t=t>$. (Law of Self-Identity)

Elimination. From $<X,P>$ & $<Y, t=t'>$ construct $<X\cup Y,P[t'//t]>$. (Substitution of Identity.)

Thinning. From $<X,P>$ construct $<X\cup Y,P>$. (You can always add more premises.)

3. The set *DFOL* the set of all pairs defined inductively as follows:
   a. $BDFOL\subseteq DFOL$
   b. If is $<X,P>$ follows from some rule in *PR* from some deductions in *DFOL*, then $<X,P>\in$ DFOL
   c. Nothing else is in *DFOL*.

∗**Exercise.** Construct a proof of the following In Gentzen's natural deduction system:

$$\forall x(Fx\rightarrow Gx),\exists xFx\vdash\exists xGx$$

Soundness and Completeness

The set of deductions is co-extensional with the valid arguments of first-order logic:

Theorem. Soundness and Completeness.

7. Statement Soundness and Completeness. P is a tautology iff $\vdash P$
8. Finite Argument Soundness and Completeness. $P_1,\ldots,P_n\vDash Q$ iff $P_1,\ldots,P_n\vdash Q$.
9. Argument Soundness and Completeness

$$X\vDash Q \text{ iff, for some subset } \{P_1,\ldots,P_n\} \text{ of } X,\ P_1,\ldots,P_n\vdash Q.$$

LECTURE 19. TRUTH-TABLES, ALGORITHMS AND DECIDABILITY

Decidability

Knowing that something has a proof is not the same has actually having one. In the case of categorical and propositional logic it is possible to design procedures that will let you know that an argument is valid without actually having a proof. Such a device is called a decision procedure. One of the more interesting facts we know about first-order logic is that there is in principle no decision procedure for its valid arguments. The only way to know that an argument is valid it is necessary to actually find a proof.

Let us define more carefully what a decision procedure is. For any set *A* we can define what is called its "characteristic function." This is the function *f* that assigns to an entity *x* the value 1 if it is in *A* and 0 if it is not in *A.*

Definition. The *characteristic function* for a set *A* is the function *f* defined as follows:

$\forall x\ (x \in A \text{ iff } f(x)=1, \text{ and } x \notin A \text{ iff } f(x)=0)$

It is a trivial fact of set theory that every set has its characteristic function, because we have just defined it. Some characteristic functions, however, are very special epistemologically. The are functions the values of which can be calculated by what is called an *effective processes.* We can compute in a finite amount of time, and can know the result in an epistemically transparent way. Long division is a classic example. There is a process *g* defined on pairs of natural numbers such that given a devisor *n* (grater than 0) and a dividend *m* it will in a finite number of epistemically transparent steps produce a quotient-remainder pair $<q,r>$ such that

$m=(qn+r)$.

As a result we can define a calculable characteristic function f_{Pr} for the set *Pr* of prime numbers:

$f_{Pr}(m)=1$ iff $1 \leq m$ and the greatest *n* such that $\exists q\ (m=(qn+0)$ is $m)$;

$f_{Pr}(m)=0$ otherwise.

A calculable characteristic function is called a *decision procedure*, and a set that has a decision procedure is called *decidable*.

A number of the most asic idea of logic are decidable, and it was their importance that led logicians to investigate the notion of decidable set and effective process in a general way. For just ideas we have met in these lectures, there are decision tests for well-formed formula, and for the validity of syllogisms, strings of syllogisms, and arguments in propositional logic. There are also decision procedures for testing whether a sentence is a tautology. Equally interesting was the discovery that there is in principle no decision procedure for testing the validity of arguments in first-order logic. Before sketching these results, however, let us investigate a bit more thoroughly exactly what a decision procedure is.

The notion of decision procedure is closely related to calculation, one of the most basic ideas in mathematics. One of the most interesting stories in the history of is how logicians in the first half of the 20th century marshaled their efforts to explain calculation. As a by product they gave birth to computer science, which is founded on calculation, and the dubious machines it studies.

Effective Process

In some sense we all know what arithmetical calculation is. As children we all spend years learning the techniques of adding, subtracting, multiplying and dividing. These and other calculation procedures have some common features. They apply to any number whatever, no matter how large. The process also proceeds in steps that are themselves simple and relatively foolproof. When multiplying, for example, we begin with the *multiplicandum* and *multiplicans*, and by a short series of prescribed steps we arrive at their product. Such a calculation procedure is called an *effective process* or an *algorithm*.

If a philosopher was asked to define the set of a mathematical calculations in terms of their necessary and sufficient conditions in the sort of *iff* definition philosophers favor, the definition would make appeal to the concept of knowledge. An calculation is a finite number of processes of a special sort. It is one that permits you to *know* at each stage what your are starting with, when you are done, what the result is, and what you should do next. You also know when the entire process stops and what the final result is. In mathematics calculations, moreover, the knowledge you possess at each stage carries with it a special kind of certainty. But as a "definition" the account we have just sketched has a serious defect. It explains the obscure, namely mathematical calculation, by the equally obscure, certain knowledge. This is so because of the sad truth that epistemology, the branch of philosophy that studies knowledge, is far from a settled and uncontroversial field. Mathematical calculation is in fact a good example of a difficult idea that resists a frontal definition of the traditional sort in terms of its necessary and sufficient conditions.

Skolem and Gödel on Recursive Functions

Thoralf Skolem (1887-1963) essentially invented the modern idea of an inductive definition to solve this problem. He defined the set of calculable arithmetic operations by construction introducing the general method that we now know as a definition by induction. Subsequently Gödel build upon Skolem's proposal. He formulated it more precisely, and expanded it somewhat to include some calculations that Skolem's original definition omitted. The set Gödel defined is called today the *recursive functions*.[22] These were meant to include all and only the functions on the natural numbers that a mathematician would recognize as a "calculation" or, in other words, the result of applying an "effective process."[23]

Gödel's definition fits the form of the inductive definition we have met before. It starts with some basic "starter" elements. For this purpose Skolem and Gödel single out three varieties of operations that they recognized were intuitively obvious calculations:

- the successor operation s, i.e. $s(x)=x+1$
- the constant functions k_c for each natural number c. Here k_c is defined as that function that assigns to every value x the value c, i.e. $k_c(x)=c$
- the index functions i_n, for each natural number n. Here i_n is defined as that function that assigns to any series $x_1,\ldots,x_n,\ldots,x_m$ its *n-th* element. That is, $i_n(x_1,\ldots,x_n,\ldots,x_m)=x_n$.

The successor function is included because it is obvious that adding one is a calculable operation. Likewise, any constant function is calculable because, for a fixed value c, anybody can set x to equal c. Lastly any index function is calculable

[22] Skolem's paper was written in 1919. Thoralf Skolem, "Begründung der elementaren Arithmetik durch die rekurrienrende Denkweise ohne Anwendung scheinbarer Veränderlichen mit undendlichen Ausdehnungsbereich," *Skrifter utgit av Videnskapsselskapet i Kristiania, I. Mathematicknaturvidenskabelig klass,* (1923) No. 6, pp 1-38. Skolem's definition omits the minimization rule and the smaller set of functions he defined is call the set of *primitive recursive functions.* Kurt, Gödel, "On Formally Undecidable Propositions of *Principia Mathmatica* and Related Systems I."[1931]. Reprinted in Jean van Heijenoort, *From Frege to Gödel* (Cambridge: Harvard University Press, 1967). See Martin Davis (ed.), *The Undecidable* (Raven: New York, 1965).

[23] In his Princeton lectures of 1934, G¨odel, attributing the idea of general recursive functions to a suggestion of Herbrand, did not commit himself to whether all effective functions are characterized by his definition. In 1936, Church [57] and Turing [268] independently proposed a definition of the effectively computable functions. It's equivalence with G¨odel's definition was proved in 1943 by Kleene [166]. Emil Post (1897-1954), Andrei Markov (1903-1979), and others. confirmed G¨odel's work by providing alternative analyses of computable function that are also provably coextensive with his.

because it is easy to calculate that the *n-th* element of a series. Just start counting from x_1 and stop at the *n-th* item.

Next three "construction" rules are identified for making new recursive functions from old. These rules are chosen because they have the property that it is transparently clear that if the rules operate on calculable functions, it produces a new calculable function:

- Composition
- Recursion
- Minimization

$f(x_1,...,x_n) = h(g_1(x_1,...,x_n),...,g_m(x_1,...,x_n))$The fist rule, called *composition*, makes up a new calculation f from old by first applying some of the old functions $g_1,...,g_m$ to a set of inputs $x_1,...,x_n$ to calculate some tentative outputs $g_1(x_1,...,x_n),...,g_m(x_1,...,x_n)$. Next these tentative outputs $g_1(x_1,...,x_n),...,g_m(x_1,...,x_n)$ are taken as inputs of yet another old function *h.* That is we calculate $h(g_1(x_1,...,x_n),...,g_m(x_1,...,x_n))$. This final output is then declared the value of the new function f for he original inputs. That is, $f(x_1,...,x_n) = h(g_1(x_1,...,x_n),...,g_m(x_1,...,x_n))$. The new calculation f is clearly the result of applying already defined calculable functions. Its results from a "heaping up" of old ones.

The second rule is called *recursion*. These methods too requires that we start with a previously defined "old" calculable function g and then defines a new one *f.* f is define in stages. It first $f(0)$ defined. Then f applied to 0+1 is defined by appeal to g and $f(0)$*:* $f(0+1)=g(f(0))$ Then using the fact that f is defined for 0+1 and g, f is defined for 0+1+1, etc. That is, $f(n+1)=g(f(n))$ At any stage clearly the new calculation is again a heaping up of old ones.

The last rule simply requires that we be able to calculate which of a finite list of natural numbers is the minimum. It defines a new calculable function from an old by simplifying the old function by eliminating all arguments mapped onto 0 except the least.

There is no special reason for choosing exactly these three sets of initial elements and rule other than the fact that they work. That is, (1) the initial elements are clearly effective processes and the rules clearly generate effective processes when applied to effective processes, and (2) the functions definable from them by induction are also effective processes. Just to see what the inductive definition looks like – we will not be actually working with these technical ideas – let us state the formal definitions. (Note that in the "general form" of the definitions below introduce additional input places – variables – that are not there in the inductive "simple" version. These extra "parameters" do not make the calculation any less calculable and allow the form capture even more calculations.)

The Basic Recursive Functions

The Successor Function s	For any natural number x:	$s(x) = x+1$
Constant Functions k_c	For a constant c:	$k_c(x) = c$
Index Functions i_n	For the n-th position:	$i_n(x_1,\ldots,x_n,\ldots,x_m)=x_n$

The Composition Rule

Simple Version. If h and g are 1-place recursive functions, then the following defines a 1-place recursive function f:

$$f(x) = h(g(x))$$

General Version. If h is an m-place recursive function and $g_1,\ldots,g_m$ are all n-place recursive functions, then the following defines an n-place recursive function f:

$$f(x_1,\ldots,x_n) = h(g_1(x_1,\ldots,x_n),\ldots,g_m(x_1,\ldots,x_n))$$

The Recursion Rule

Simple Version. If c is a constant and g is a 2-place recursive function, then the following defines a 1-place recursive function f:

$$f(0) = c$$

$$f(x+1) = g(x,f(x))$$

General Version. If g is an n-place recursive function and h is an n+2 place recursive function, then the following defines an n+1 place recursive function f:

$$f(0,y_1,...,y_n) = \mathrm{g}(y_1,...,y_n);$$

$$f(x+1,y_1,...,y_n) = h(x,f(x,y_1,...,y_n),y_1,...,y_n)$$

Minimization Rule

Simple Version. Suppose g is a 2-place recursive function defined on natural numbers such that for any y there is a least one x such that $g(y,\mathrm{x})$=0. Then, we define $f(y)$ to be the least such x:

$$f(x) = \text{the least } y \text{ such that } g(y,x)=0$$

General Version. Suppose g is an n+1-place recursive function defined on natural numbers such that for any x there are $y_1,\ldots,y_n$ such that $g(y_1,\ldots,y_n,x)$=0. Then, we define $f(y_1,\ldots,y_n)$ to be the least such x:

$$f(y_1,\ldots,y_n) = \text{the least } x \text{ such that } g(y_1,\ldots,y_n,x)=0$$

Definition of Recursive Function

The set *RF* of *recursive functions* is the set defined by induction from the set of basic elements $\{s,k_c,i_n\}$ (for all constants c and positive integers n) and the rule set {Composition, Recursion, Minimization}:

1. (Basis Clause.) $\{s,k_c,i_n\} \subseteq RF$
2. (Inductive Clause.) if R is in {Recursion, Composition, Minimization} and $f_1,...,f_m$ are all in *RF* and g is definable from $f_1,...,f_m$ by R, then g is in *RF*.
3. (Closure Clause.) Nothing else is in *RF*.

Notice that by this definition a function is recursive simply by virtue of being a member of the set *RF*. A function may be in that set and we not know it. It may be in the set while we only possess a poor definition of it, one that is not calculable and which is not sufficient for showing it meets the defining conditions for being in *RF*. The way to show it is in *RF* is to show it is either a basic function or is one definable from basic functions by means of recursion or composition.

Examples of RF's

Calculation Operations Defined by the Recursion Rule. The two-pace *addition operation* + is defined by recursion in terms of the one-place successor operation *s*:

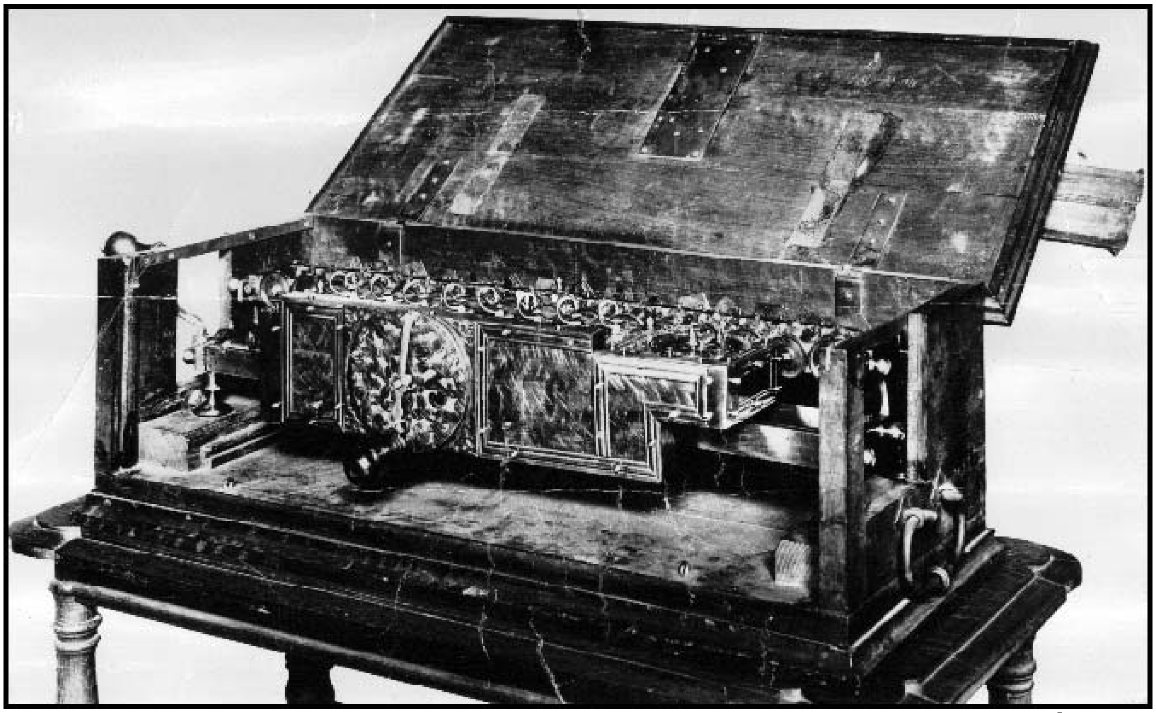
$x+0=x$

$x+s(y)=s(x+y)$

The 2-place *multiplication operation* x is defined by recursion in terms of the two-place addition operation +:

$x\times 0=0$

$x\times s(y)=(x\times y)+x$



Leibniz' Calculating Machine

"Thomas Hobbes, everywhere a profound examiner of principles, rightly stated that everything done by our mind is a *computation*, by which is to be understood the addition of a sum or the subtraction of a difference."

Leibniz, *De arte combinatoria* 63, 1666

<u>Calculation Operations Defined by the Composition Rule.</u> Given the constants *a* and *b*, we construct the linear function *h* with slope *a* and *y*-intersect *b*, namely y=*ax*+*b* as follows. (Note that we can *calculate y* from *a, b,* and *x*.)

Let *n* be the number named by the constant *c*. Then k_c is the constant function pairing any *x* with *n*. Further, let i_1 be the index function that assigns any *x* to itself. Note that both k_c and i_1 are basic recursive functions. Let *a* and *b* be constants (i.e. numerals). We define three functions *f*, *g*, and *h* by composition.

$f(x)=k_a(x) \times i_1(x)$ (In traditional notation: $f(x)=a \times x$)
$g(x)=k_b(x)+i_1(x)$ (In traditional notation: $g(x)=x+b$)
$h(x)=g(f(x))$ (In traditional notation: $h(x)=ax+b$.

Thus, by composition we have defined *h*, the equation for a line with slope *a* and y-intersect *b.*

Implications of Gödel's Definition for the Methodology of Logic

Historically, the inductive characterization of the recursive functions was an important contribution to the methodology of logic. It was one of the first uses of inductive definitions that was clearly seen to be a new way to approach the task of "defining" an important idea. In this case the idea is that of an effectively calculable function. The fruitfulness of the inductive method was shown by the fact that shortly after Gödel definition in the 1930's the logicians Alan Turing (1912-1954) and Alonzo Church (1903-1995) using different techniques proposed their own inductive definitions of the effectively calculable function. These were quickly were shown define exactly the same set of functions as Gödel's.[24] It was possible to prove mathematically that Gödel's set of recursive functions were the same as the sets defined by Church and Turing because all three have precise mathematically clear inductive definitions. Emil Post (1897-1954) later found yet a fourth coextensive characterization. Church explicitly claimed that these new inductive methods provided "definitions" of *effectively calculable function.*

Post, however, pointed out that Church's claim that the new method provided a new "definition" was something that could not itself be proven in mathematics itslef.[25] It cannot be proven because *effectively calculable function* as traditionally defined does not have a precise mathematical definition. Rather, if the traditional idea can be said to have a definition at all, it would be a definition of the traditional sort, formulated in terms of necessary and sufficient conditions. Since it is this definition that is at issue, let us make an attempt to state it, though it has to be understood that any such attempt is only an approximation of rather imprecise mathematical usage:

The Traditional Definition of Effectively Calculable Function

> A function is *effectively calculable* if and only if its value can be calculated in a finite number of steps each of which produces an outcome known with a high degree of certainty (i.e. that is "epistemically transparent"), and which terminate in a final result known with a high degree of certainty.

---

[24] Alonzo Church "An Unsolvable Problem of Elementary Arithmetic", *American Journal of Mathematics,* 2nd ser. 58 (1936), 345-363; and A.M. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem," *Proceedings of the London Mathematical Society*, ser. 2, 42 (1936-37), 230-265.

[25] Emil L. Post, "Finite Combinatory Processes – Formulation 1", *Journal of Symbolic Logic* 1 (1936), 103-105.

We cannot use this definition in a mathematical proof because it makes use of the mathematically unexplained idea of "epistemic transparency".

Logicians agreed that Gödel's basic recursive functions were genuine examples of effectively calculable functions, and that his three method of generating new calculable functions from old would only generate calculable functions from calculable functions. The generally opinion that every recursive function as defined by Gödel is in fact an effectively calculable function as traditionally understood may be summarized in a simple universal law:

Every recursive function is an effectively calculable function.

This was half of Church's claim. Post pointed out however that the converse is not obvious. Moreover since *effectively calculable function* lacks a clear mathematical definition, it cannot be proven. As a result, the converse proposition is now viewed as a substantive claim about the analysis of concepts, and it bears Church's name:

Church's Thesis

Every effectively calculable function is recursive.

It is now generally accepted that the thesis is true.

Equally important for logic are the methodological implications of the technique. The success of the inductive analysis of *effectively decidable function* demonstrated the power of the inductive method for analysis. Since being recognized as an alternative approach to definition, it has become a standard tool for defining sets and proving facts about them. As traditional axioms systems show, inductive sets have been used since the beginning of mathematics without being clearly recognized as different. In these lectures, we have used them to define the natural numbers, the set of well-formed sentences and formulas in formal grammars, various interpretation of propositional and first-order logic, the sets of theorems and acceptable deductions of propositional and first-order logic, and just now the set of calculable functions.

Decidable Sets in Logic

Decidability is a concept we can apply to logic. As we shall now see, it is possible to define decision procedures to test whether arguments written in the syntax of categorical or propositional logic are valid. It turns out, however, to be impossible to do so for first-order logic more generally.

Before taking up these interesting issues in logic, we must pause to make a general remark about the concept of a decidable set. It must not be confused with an inductively definable set. In general, being decidability is not the same as having an inductive definition. Consider the case of the prime numbers. As we saw earlier, there is a decision procedure for the prime numbers. On the other hand, mathematicians do not know a method for listing all the prime numbers, and do not know how to define the set of primes inductively. Thus, some decidable sets are not inductive. The converse is also true. There are some sets that are inductively definable but are not decidable. An important example is the set of valid arguments in first-order logic. We have already seen two ways to define this set inductively: in an axiom system and in a natural deduction system. But though there are decision procedures for the set of valid arguments in categorical and propositional logic, there is no such procedure for the set of valid arguments of first-order logic. It is an inductive set that is not decidable.

Syntactic Tests for Valid Syllogisms

If the issue is deciding whether any of the 256 syllogistic moods is valid, there is a trivial test: just look to see whether the syllogism fits any of the listed 24 valid mood paradigms. If it does it is valid, if not it is invalid. This test is syntactic and epistemically transparent because it is possible to determine easily by inspection of its physical shape whether a syllogism fits one of the 24 valid forms. The test moreover can be generalize to an evaluation of any finitely valid categorical argument $P_1,\ldots,P_n \models_{SL} Q$. Since there are only a finite number of premises, there are only a finite (though possibly very large) number of finite sequences of syllogisms that would lead from $P_1,\ldots,P_n$ to Q, i.e. there are only a finite number of finite syllogism sequences such that in each sequence the conclusion of the last syllogism is Q, and every premise of every syllogism in the sequence is either in $\{P_1,\ldots,P_n\}$ or is the conclusion of a previous syllogism in the sequence. We can test each syllogism in each series. If there is a series such that all its syllogisms are valid, then the argument being tested is valid. If there is no such series, it is invalid.

A more interesting and useful test makes use of the traditional term rules. As we saw earlier, a syllogism or, more generally, a minimal categorical argument of any number of premises, is valid iff it does not violate a term rule. A syllogism is valid iff it does not violate a term true. Moreover, whether an argument violates a rule is an epistemically transparent syntactic matter. Hence, we can test in a finite amount of time by epistemically transparent methods whether a syllogism violates any term rule.

The traditional term rule test is, accordingly, a decision procedure for the validity of categorical arguments.[26]

More generally, we also saw that a minimal categorical argument of any finite number of premises is valid iff it does not violate a term rule. It follows that the categorical argument from $P_1,\ldots,P_n$ to Q is valid iff, there is some finite sequence of syllogisms such that (1) no syllogism in the sequence violates a term rule, (2) Q is the conclusion of the last syllogism in the sequence, and (3) every premise of every syllogism in the sequence is either in $\{P_1,\ldots,P_n\}$ or is the conclusion of a previous syllogism in the sequence. Accordingly, there is a finite test of whether there exists a finite series of syllogisms leading from $P_1,\ldots,P_n$ to Q none of which violates any term rule.

The Truth-Table Test for Tautologies and Validities

As an example of a decision procedure the truth-table method for propositional logic is more interesting. Indeed, the truth-table test for validity is a textbook example of a decision procedure. It is easy to sketch the test informally.

We define a function *f* as follows:

Let $<\{P_1,\ldots,P_n\},Q>$ be an "argument " in propositional logic (i.e. a set of premises and a conclusion). Construct the truth-table for $(P_1,\ldots,P_n)\rightarrow Q$.

[26] Wolfgang Lensen makes the claim (in "On Leibniz's Essay *Mathesis rationis," Topoi* 9 (1990), 29-59) that the set of term rules as formulated by Leibniz constitutes a sound and complete axiom set for the valid arguments of a categorical syntax that he attributes to Leibniz. As should be clear from the presentation here, the term rules do not function as inference rules, or more generally as construction rules for an inductively defined set of acceptable syllogisms, whether these be understood as theorems in a traditional axiom system or as deductions as in a natural deduction system, or as any other sort of entity in an inductively defined set. Though in the *Port Royal Logic* Arnauld and Nicole refer to the rules as axioms, and Leibniz follows this usage, they are not axioms in the modern sense. They do not play a role in an inductive definition. From the perspective of modern logic, the function of the rules is different. They formulate necessary and sufficient syntactic criteria for an argument's acceptability. It is therefore more accurate to say that they define a decision procedure rather than an axiom system.

- If in the truth-table $(P_1,\ldots,P_n)\rightarrow Q$ is assigned T in every case, then make *f* assign to the argument from $P_1,\ldots,P_n$) to *Q* the value 1 , i.e.

  $f(<\{P_1,\ldots,P_n\},Q>=1$;
- If in the truth-table $(P_1,\ldots,P_n)\rightarrow Q$ is sometimes assigned F, then make *f* assign to the argument from $P_1,\ldots,P_n$) to *Q* the value 0 , i.e.

  $f(<\{P_1,\ldots,P_n\},Q>=0$.

Though we shall not do so here, it is a relatively simple process to represent the sentences and arguments of propositional logic by natural numbers so that each sentence or argument is represented by a number. It is then possible to define a recursive function on these representatives that produces 1 if the argument is valid according to the truth-table test and 0 if it is not. That is, it is not hard to show that the truth-table test determines a recursive function on the natural numbers in Gödel's sense.

The Undecidability of First-Order Logic

Thus, both categorical and propositional logic is decidable. However, this is not so for every logical system. One of the more interesting logical discoveries in the early 20th century[27], is the following metatheorem:

[27] Using techniques devised by Skolem, Jaques Herbrand (1908-1932)showed that the quantified formula *P* is satisfiable if and only if a specific set of its truth-functional instantiations, each essentially a formula in sentential logic, is satisfiable. Thus, satisfiability of P reduces to an issue of testing by truth-tables the satisfiability of a potentially infinite set S of sentential formulas. Herbrand showed that, for any first-order formula *P*, there is a decision function f such that f(*P*) = 1 if *P* is unsatisfiable because, eventually, one of the truth-functions in S will come out false in a truth-table test, but f(*P*) may be undefined when *P* is satisfiable because the truth-table testing of the infinite set S may never terminate.
The foundational paper which leads to this result is Jacques Herbrand, "Sur la Théorie de la Démonstration," *Comptes Rendus des Séances de la Sociéte des Sciences et des Lettres de Varsovie, Classe III* 24 (1931): 12-56.

Theorem. There is in principle no decision procedure for the valid arguments of first-order logic.

Like Gödel's theorem, which proves the incompleteness of any axiom system of arithmetic, this theorem demonstrates a limitation on human knowledge. Epistemology is a difficult subject about which there is very little known with certainty. It is therefore remarkable that we can show with mathematical rigor that a certain kind of knowledge is impossible. One thing we can now about knowledge with certainty is that there is no finite effectively calculable epistemically transparent method for testing in general whether an argument is valid in first-order logic.

SUMMARY

In Part 3 we have discussed three sorts of syntax of increasing complexity: categorical, propositional and first-order logic. For these we have discussed two sets important to logic: their logical truths and their valid arguments. For each of these sets we have investigated whether it was inductively definable, whether the resulting axiom or natural deduction system was complete, and whether the set was decidable. The conclusions may be summarized in tabular form:

	Set of:	Inductive Definition	Complete	Decidable
Categorical Logic	Syllogisms	**Yes:** Reduction System	**Yes:** A syllogism is valid iff it is reducible.	**Yes:** Leibniz' Term Rule Test
	Finite Arguments	**Yes:** Generalized Reduction System	**Yes:** A generalized argument is valid iff it is reducible.	**Yes:** Generalized Term Rule Test
Propositional Logic	Tautologies	**Yes:** Łukaisiewicz' Axiom System	**Yes:** *P* is a tautology iff *P* is a theorem.	**Yes:** Truth-Table Test
	Valid Arguments	**Yes:** Gentzen's Natural Deduction System	Yes: $P_1,\ldots,P_n \models Q$ iff $P_1,\ldots,P_n \vdash Q$	**Yes:** Truth-Table Test
First-Order Logic	Logical Truths	**Yes:** Russell and Whitehead's Axiom System	**Yes:** *P* is a logical truth iff *P* is a theorem.	**No:** There is no effectively calculable characteristic function.
	Valid Arguments	**Yes:** Gentzen's Natural Deduction System	**Yes**: $P_1,\ldots,P_n \models Q$ iff $P_1,\ldots,P_n \vdash Q$	**No:** There is no effectively calculable characteristic function.
Arithmetic	Truths of Arithmetic	**Yes:** Axiom systems of Russell, Whitehead and others	**No:** In any axiom system there is at least one truth of arithmetic that is not a theorem.	**No:** There is no effectively calculable characteristic function.

The first general conclusion we can draw concerns conceptual identity. The notions of inductive, complete and decidable sets of logical truths and arguments are not the same. Some inductive sets are decidable and others not, and conversely.

An axiom or natural deduction system is a "proof theory." It is so because it provides an inductive definition in syntactic terms, and each element of an inductive set has a construction sequence. When defined syntactically, these construction sequences are proofs. It is possible to see with a high degree of certainty that any step in the sequence is either an initial element (axiom or basic deduction) or is constructed ("follow from") earlier elements of the sequence by a construction rule. In this way the theory explains the peculiar kind of certainty we have about the application of logical rules. They are obvious because they are questions about the fit of simple syntactic rules.

But not every axiom or natural deduction system is successful simply because it is well defined syntactically. To be successful in its scientific purpose the system must also be interesting in the sense of saying something about the world. In technical terms, it must be sound and complete. That is, the system is successful only if its theorem set contains all and only logical truths and its acceptable derivations correspond exactly to the set of valid arguments. If it is sound and complete, then we can be sure that its laws and proofs are not only correct applications of clear syntactical rules but that they also way something about the world that is necessarily true and allow us to reason in ways that are totally dependable.

We also know by the criterion of completeness that axiomatic method does not always work. The truths of arithmetic cannot be completely axiomatized in an inductively defined set.

In some cases as in the syllogistic and propositional logic we can also define decision procedures for testing whether a sentence is a tautology or an argument is valid, though we cannot do so for first-order logic generally. As a rough generalization, it seems that whether a subject matter is complete or decidable is a function of the expressive complexity of its language and the richness of its assumptions. The very simplest logical theories represented by categorical and propositional are not only sound and complete but are also decidable. But their syntax is uncomplicated and they make only the most minimal assumptions about the nature of reality, as reflected in the short list of their extremely formal axioms and rules. First-order logic is much richer in its expressive capacity, adequate for the formulation of most mathematics and science. Its axiom and rule set is also more detailed imposing more demands on the structure of "the world." Though it is sound and complete, it is not decidable. Arithmetic, set theory, and natural sciences that assume the truths of arithmetic or set theory, though they are written in a first-order syntax, imposes much more specific assumptions about the world. As a result they are neither decidable nor complete. Unfortunately, to discover the truths of math and science it is not enough to do logic.

Truth-Conditions of Selected Formulas

Categorical Propositions

TC1. $\Im(\mathbf{A}XY)=\mathrm{T} \quad \leftrightarrow \quad \Im(X)\subseteq\Im(Y)$

TC2. $\Im(\mathbf{E}XY)=\mathrm{T} \quad \leftrightarrow \quad \Im(X)\cap\Im(Y)=\varnothing$

TC3. $\Im(\mathbf{I}XY)=\mathrm{T} \quad \leftrightarrow \quad \Im(X)\cap\Im(Y)\neq\varnothing$

TC4. $\Im(\mathbf{O}XY)=\mathrm{T} \quad \leftrightarrow \quad \Im(X)-\Im(Y)\neq\varnothing$

First-Order Formulas

TC0. $\Im(Fc)=\mathrm{T}$ iff $\Im^{D}(c)\in\Im^{D}(F)$

TC1. $\Im(Fc\wedge Gb)=\mathrm{T}$ iff $\Im^{D}(c)\in\Im^{D}(F)$ and $\Im^{D}(b)\in\Im^{D}(G)$

TC2. $\Im(Rac\rightarrow Gx)=\mathrm{T}$ iff $<\Im^{D}(a),\Im^{D}(c)>\notin\Im^{D}(R)$ or $\Im^{D}(x)\in\Im^{D}(G$

TC3. $\Im(\forall xFx)=\mathrm{T}$ iff for all $d\in D$, $d\in\Im^{D}(F)$

TC4. $\Im(\exists xFx)=\mathrm{T}$ iff for some $d\in D$, $d\in\Im^{D}(F)$

TC5. $\Im(\forall x\exists yRxy)=\mathrm{T}$ iff for all $d\in D$, for some $d'\in D$, $<d,d'>\in\Im^{D}(R))$

TC6. $\Im(\exists x\forall yRxy)=\mathrm{T}$ iff for some $d\in D$, for all $d'\in D$, $<d,d'>\in\Im^{D}(R))$

TC7. $\Im(\forall xRxx)=\mathrm{T}$ iff for all $d\in D$, $<d,d>\in\Im^{D}(R))$

TC8. $\Im(\forall x(Fx\rightarrow Gx))=\mathrm{T}$ iff for all $d\in D$, either $d\notin\in\Im^{D}(F)$ or $d\in\Im^{D}(G)$

TC9. $\Im(\exists x(Fx\wedge Gx))=\mathrm{T}$ iff for some $d\in D$, $d\in\Im^{D}(F)$ and $d\in\Im^{D}(G)$

TC10. $\Im(\forall x(Fx\wedge Gx))=\mathrm{T}$ iff for all $d\in D$, $d\in\Im^{D}(F)$ and $d\in\Im^{D}(G)$

TC11. $\Im(\exists x(Fx\rightarrow Gx))=\mathrm{T}$ iff for some $d\in D$, either $d\notin\Im^{D}(F)$ or $d\in\Im^{D}(G)$

TC12. $\Im(\forall x(Fx\rightarrow\exists yRxy))=\mathrm{T}$ iff for all $d\in D$, either ($d\notin\Im^{D}(F)$ or for some $d'\in D <d,d'>\in\Im^{D}(R))$

TC13. $\Im(\forall x\exists y(Rxy\rightarrow Ryx))$=T iff for all $d\in D$, for some $d'\in D$,
either $<d,d'>\notin\Im^{D}(R))$ or $<d',d>\in\Im^{D}(R))$

TC14. $\Im(\forall x\forall y(Rxy\leftrightarrow Ryx))$=T iff for all $d\in D$, for all $d'\in D$,
$<d,d'>\in\Im^{D}(R))$ iff $<d',d>\in\Im^{D}(R))$

TC15. $\Im(\exists xFx\wedge\exists yGy))$=T iff for some $d\in D$, $d\in\Im^{D}(F)$ and
for some $d'\in D$, $d'\in\Im^{D}(G)$

TC16. $\Im(\forall x(Fx\rightarrow\forall yGy))$=T iff either for all $d\in D$, $d\notin\Im^{D}(F)$, or
for some $d'\in D$, $d'\in\Im^{D}(G)$

The Names for the Valid Syllogistic Moods

First Figure:			Third Figure:		
	AAA	Barbara		AAI	Darapti
M,P	EAE	Celarent	*M,P*	EAO	Felapton
S,M	AII	Darii	*M,S*	IAI	Disamis
S,P	EIO	Ferio	*S,P*	AII	Datisi
	EAO	*Celaront		OAO	Bocardo
	AAI	*Barbari		EIO	Ferison

Second Figure:			Fourth Figure:		
	EAE	Cesare		EIO	Fresison
P,M	AEE	Camestres	*P,M*	EAO	Fesapo
S,M	EIO	Festino	*M,S*	IAI	Dimaris
S,P	AOO	Baroco	*S,P*	AAI	Bramantip
	AEO	*Camestrop		AEE	Camenes
	EAO	*Cesaro		AEO	*Camelop

Red type indicates that their validity depends on the assumption that universal propositions must stand for a non-empty subject to be true.
An asterisk indicates a *subaltern mood.*

Mnemonic Poem for the Valid Moods and Reduction

Henry Aldrich (1647–1710), *Artis Logicae Rudimenta*

Barbara, Celarent, Darii, Ferioque prioris:
Cesare, Camestres, Festino, Baroco secundae:
Tertia, *Darapti, Disamis, Datisi, Felapton,*
Bocardo, Ferison, habet; Quarta insuper addit
Bramantip, Camenes, Dimaris, Fesapo, Fresison.
Quinque *Subalterni,* todidem *Generalibus* orti,
Nomen habent nullum, nec, si bene colligatur, usum.

[*Barbara, Celarent, Darii, Ferioque* are of the First:
Cesare, Camestres, Festino, Baroco are of the Second:
The Third has *Darapti, Disamis, Datisi, Felapton,*
Bocardo, Ferison; The Fourth adds in addition
Bramantip, Camenes, Dimaris, Fesapo, Fresison.
Fifth are the *Subalterns,* which all come from the *Universals*,
They do not have a name, nor, if well connected, a use.]

Square of Opposition

The Square of Opposition
(Diagram of Immediate Inferences)

A: *Every S is P* (A*SP*)

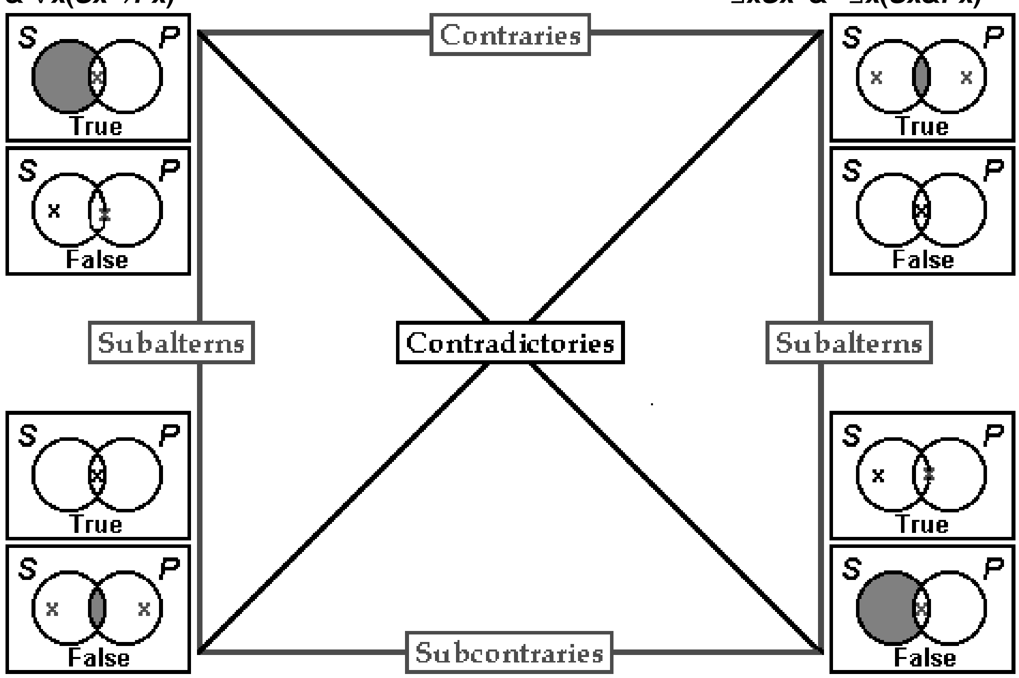
$\Im(S) \neq \emptyset \;\&\; \Im(S) \subseteq \Im(P)$

$\exists x Sx \;\&\; \forall x(Sx \rightarrow Px)$

E: *No S is P* (E*SP*)

$\Im(S) \neq \emptyset \;\&\; \Im(P) \neq \emptyset \;\&\; \Im(S) \cap \Im(P) = \emptyset$

$\exists x Sx \;\&\; \sim\exists x(Sx \& Px)$



I: *Some S is P* (I*SP*)

$\Im(S) \cap \Im(P) \neq \emptyset$

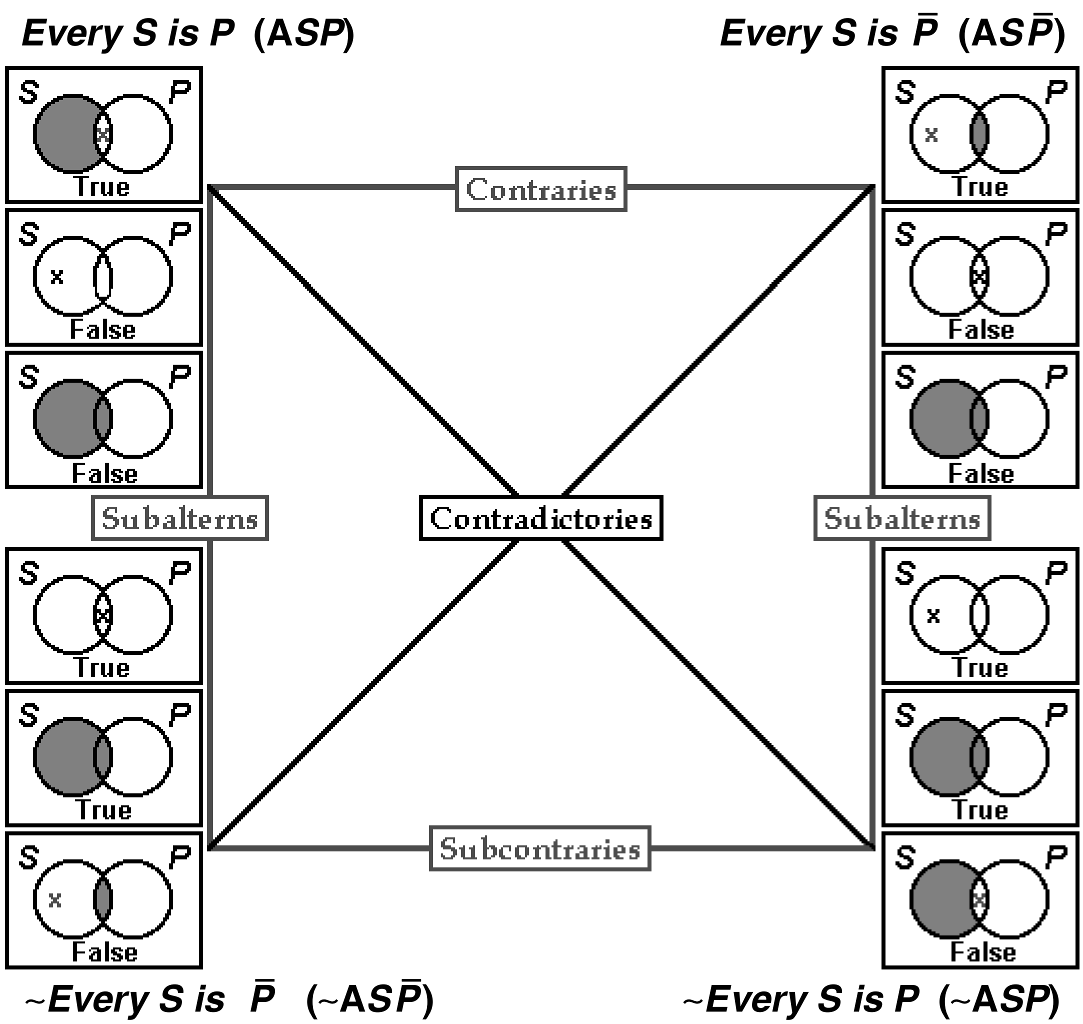
$\exists x(Sx \& Px)$

O: *Some S is not P* (O*SP*)

$\Im(S) - \Im(P) \neq \emptyset$

$\exists x(Sx \& \sim Px)$

Square of Opposition for Predicate Negations



Term Rules for Testing Whether a Mood is Valid

Rule 1. Undistributed Middle. No valid syllogism has an undistributed middle term.

Rule 2. Distributed Term in the Conclusion. No syllogism is valid that has a term that is distributed in the conclusion but not in the premises.

Rule 3. Affirmative premise. No syllogism is valid that has two negative premises.

Rule 4. Negative Conclusion. No syllogism is valid that has negative conclusion without a negative premise.

Rule 5. Particular Premise. No syllogism is valid that has particular premise and a universal conclusion.

Rule 6. Negative Premise. No syllogism is valid that has a negative premise and an affirmative conclusion.

Rule 7. Universal Premise. No syllogism is valid that does not have at least one universal premise.

General Metatheorems on Logical Relations

Theorem

$\{P_1,\ldots,P_n\} \vDash Q$	iff	$\{P_1,\ldots,P_n, \sim Q\}$ is inconsistent
	iff	$(P_1 \wedge \ldots \wedge P_n) \rightarrow Q$ is a logical truth
$\{P_1,\ldots,P_n\}$ is consistent	iff	for no Q, $\{P_1,\ldots,P_n\} \vDash Q \wedge \sim Q$
	iff	for some Q, $\{P_1,\ldots,P_n\} \nvDash Q$
	iff	$\sim(P_1 \wedge \ldots \wedge P_n)$ is not a logical truth
P is a logical truth	iff	for every Q, $Q \vDash P$
	iff	$\sim P$ is inconsistent

Metatheorems on Categorical Logic

Theorem. The following instances of the (T) schema are true:

TC1. $\Im(\mathbf{A}XY)=\mathrm{T} \quad \leftrightarrow \quad \Im(X)\subseteq\Im(Y)$

TC2. $\Im(\mathbf{E}XY)=\mathrm{T} \quad \leftrightarrow \quad \Im(X)\cap\Im(Y)=\varnothing$

TC3. $\Im(\mathbf{I}XY)=\mathrm{T} \quad \leftrightarrow \quad \Im(X)\cap\Im(Y) \neq\varnothing$

TC4. $\Im(\mathbf{O}XY)=\mathrm{T} \quad \leftrightarrow \quad \Im(X)-\Im(Y) \neq\varnothing$

Theorem. For any term X and any interpretation $\Im$,

$$\Im(X)\neq\varnothing \ \& \ \Im(X)\subseteq U$$

Theorem. $\mathbf{A}XY \vDash_{SL} \mathbf{I}XY$

Theorem. $\mathbf{E}XY \vDash_{SL} \mathbf{O}XY$

Theorem. $\mathbf{A}XY$ and $\mathbf{O}XY$ are contradictories

Theorem. $\mathbf{E}XY$ and $\mathbf{I}XY$ are contradictories

Theorem. $\mathbf{A}XY$ and $\mathbf{E}XY$ are contraries

Theorem (Simple Conversion)

1. $\mathbf{E}XY \vDash_{SL} \mathbf{E}YX$
2. $\mathbf{O}XY \vDash_{SL} \mathbf{O}YX$

Theorem (Conversion *per Accidens*)

1. $\mathbf{A}XY \vDash_{SL} \mathbf{I}YX$
2. $\mathbf{E}XY \vDash_{SL} \mathbf{O}YX$

Theorems

$\mathbf{A}FG \;\dashv\vDash_{SL}\; \mathbf{E}F\bar{G}$

$\mathbf{A}F\bar{G} \;\dashv\vDash_{SL}\; \mathbf{E}FG$

$\mathbf{I}FG \;\dashv\vDash_{SL}\; \mathbf{O}F\bar{G}$

$\mathbf{I}F\bar{G} \;\dashv\vDash_{SL}\; \mathbf{O}FG$

Theorems

1. $\mathbf{A}XY \vDash_{SL+\varnothing} \sim\mathbf{A}X\bar{Y}$
2. $\mathbf{A}X\bar{Y} \vDash_{SL+\varnothing} \sim\mathbf{A}XY$
3. $\mathbf{A}XY$ and $\sim\mathbf{A}XY$ are contradictories (with respect to $\vDash_{SL+\varnothing}$)
4. $\mathbf{A}X\bar{Y}$ and $\sim\mathbf{A}X\bar{Y}$ are contradictories (with respect to $\vDash_{SL+\varnothing}$)
5. $\mathbf{A}XY$ and $\mathbf{A}X\bar{Y}$ are contraries (with respect to $\vDash_{SL+\varnothing}$)
6. $\sim\mathbf{A}X\bar{Y}$ and $\sim\mathbf{A}XY$ are subcontraries (with respect to $\vDash_{SL+\varnothing}$)

Theorem. The syllogism **AAA** in the first figure (called Barbara), i.e. <**A***MP*,**A***SM*,**A***SP*>, is valid.

Theorem. If $P_1,\ldots,P_n \vDash_{SL} Q$, then there is some finite sequence of valid syllogisms such that (1) the conclusion of the last syllogism is Q, and (2) each premise of any syllogism in the sequence is either in $\{P_1,\ldots,P_n\}$ or is the conclusion of a previous syllogism in the sequence.

Theorem. The syllogism **AIO** in the fourth figure, i.e. <**A***PM,***I***MS,***O***SP*>, is invalid.

Theorem. The syllogism **EOA** in the fourth figure, i.e. <**E***MP,***O***SM,***A***SP*>, is invalid.

Theorem. The syllogism **EOI** in the fourth figure, i.e. <**E***MP,***O***SM,***I***SP*>, is invalid.

Theorem. The syllogism **AIO** in the fourth figure, i.e. <**A***MP,***I***SM,***O***SP*>, is invalid.

Theorem. The syllogism **IEA** in the fourth figure, i.e. <**I***MP,***E***SM,***A***SP*>, is invalid.

Theorem. The syllogism **AEI** in the fourth figure, i.e. <**A***MP,***E***SM,***I***SP*>, is invalid.

Theorem. The syllogism **IOI** in the fourth figure, i.e. <**I***MP,***O***SM,***I***SP*>, is invalid.

Theorem. The syllogism **EAO** in the first figure, i.e. <**E***MP,***A***MS,***O***SP*>, which is called Felapton, is valid.

Metatheorems on Propositional Logic

Theorem. $\{P_1,\ldots,P_n\} \models_{PL} Q$ iff $(P_1\wedge\ldots\wedge P_n)\rightarrow Q$ is a tautology.

Theorem. Disjunctive Syllogism in valid in propositional logic: $\{p_1\vee p_2, \sim p_1\} \models p_2$.

Theorem. Contraposition is valid in propositional logic: $\{p_1\rightarrow p_2\} \models \sim p_2 \rightarrow \sim p_1$.

Theorem. $\{P_1,\ldots,P_n\} \not\models_L Q$ iff $(P_1\wedge\ldots\wedge P_n)\rightarrow_L Q$ is not a tautology.

Theorem. Denying the antecedent is invalid: $\{p_1\rightarrow p_2, \sim p_1\} \not\models \sim p_2$.

Theorem. The set $\{p_1\vee p_2, p_1\vee p_2\}$ is inconsistent in propositional logic.

Theorem. The set $\{p_1\vee p_2, \sim p_1\wedge\sim p_2\}$ is inconsistent in propositional logic.

Metatheorems on First-Order Logic

Theorem (Barbara). $\forall x(Gx\rightarrow Hx), \forall x(Fx\rightarrow Gx) \models \forall x(Fx\rightarrow Hx)$

Theorem (Celarent). $\sim\exists x(Gx\wedge Hx), \forall x(Gx\rightarrow Hx) \models \sim\exists x(Gx\wedge Hx)$

Theorem. $Fa \models \exists xFx$

Theorem. $\forall xFx \models \exists xFx$

Theorem. $\forall xFx \wedge \forall yGy \models \forall x(Fx \wedge Gy)$

Theorem. $\exists x(Fx \wedge Gx) \models \exists xFx \wedge \exists yGy$

Theorem. $\forall x \forall y(Rxy \leftrightarrow Ryx) \models \forall xRxx$

Theorem. $\exists x \forall yLxy, \forall x \forall y(Lxy \leftrightarrow Lyx) \models \forall x \exists yLxy$

Theorem (Barbari). $\forall x(Gx \rightarrow Hx), \forall x(Fx \rightarrow Gx) \not\models \exists x(Fx \wedge Hx)$

Theorem. $\forall x(Gx \rightarrow Hx), \forall x(Fx \rightarrow Gx), \exists xFx \models \exists x(Fx \wedge Hx)$

SUMMARY OF EXERCISES

Lecture 13

∗**Exercise:** In the standard semantics for the syllogistic prove: **E***XY* $\models_{SL}$ **O***XY*

∗**Exercise.** Prove: **E***XY* and **I***XY* are contradictories

∗**Exercise.** Prove: **I***XY* and **O***XY* are subcontraries

∗**Exercise.** Prove that the syllogism **EAE** in the first figure (called Celarent), i.e. <**E***MP*,**A***SM*,**E***SP*>, is valid.

Lecture 14

Exercise. The two following syllogisms are invalid. For each,

- name which of the seven syllogistic rules it violates,
- draw a Venn diagram illustrating that its premises are true but its conclusion false in that universe, and
- ∗give a proof like those in the previous examples that the syllogism is invalid:
    1. **AIE** in the fourth figure
    2. **IOA** in the second figure

∗**Exercise:** Prove that he syllogism **AAI** in the fourth figure (Bramantip), i.e. <**A***PM*,**A***MS*,**I***SP*>, is valid.

Lecture 15

Exercise

Show *modus tollens* is valid in propositional logic: $\{p_1 \rightarrow p_2, \sim p_2\} \models \sim p_1$.

	p_1	p_2	$((p_1$	$\rightarrow$	$p_2)$	$\wedge$	$\sim$	$p_2)$	$\rightarrow$	$\sim$	$p_1)$
$\Im_1$	T	T									
$\Im_2$	T	F									
$\Im_3$	F	T									
$\Im_4$	F	F									

Determine when $\Im(((p_1 \rightarrow p_2) \wedge \sim p_2) \rightarrow \sim p_1)$=T.

Exercise

Show Affirming the Consequent is invalid in propositional logic: $\{p_1 \rightarrow p_2, p_2\} \not\models p_1$.

	p_1	p_2	$(((p_1$	$\rightarrow$	$p_2)$	$\wedge$	$p_2)$	$\rightarrow$	$p_1)$
$\Im_1$	T	T							
$\Im_2$	T	F							
$\Im_3$	F	T							
$\Im_4$	F	F							

Exercise
Show $\{p_1 \rightarrow p_2, \sim(\sim p_1 \vee p_2)\}$ is inconsistent in propositional logic:

	p_1	p_2	$((p_1$	$\rightarrow$	$p_2)$	$\wedge$	$\sim$	$(\sim$	p_1	$\vee$	$p_2))$
$\Im_1$	T	T									
$\Im_2$	T	F									
$\Im_3$	F	T									
$\Im_4$	F	F									

$*$**Exercise.** Prove Celarent is valid in first-order logic:

$$\sim\exists x(Gx \wedge Hx), \forall x(Fx \rightarrow Gx) \models \sim\exists x(Fx \wedge Hx)$$

$*$**Exercise.** Prove the metatheorems in the semantics for first-order logic:

1. $\forall xFx \models \exists xFx$.
2. $\forall x(Bx \rightarrow \exists y(Gx \wedge Lxy)), \sim\exists x(Gx \wedge Lcx) \models \sim Bc$
3. $\forall x \forall y \forall z((Lxy \wedge Lyz)) \rightarrow Lxz), \forall x \sim Lxx \models \sim\exists x \exists y(Lxy \wedge Lyx)$
4. $\exists x(Gx \wedge Hx), \exists x(Fx \wedge \sim Gx) \not\models \forall x(Fx \rightarrow Hx)$

Lecture 17

Exercises

Using the traditional reduction rules, prove form (i.e. reduce to) a first figure syllogism with the same initial letter the following syllogisms:

1. Datisi
2. Camestrop
3. Bocardo
4. Fesapo

Lecture 18

Exercise. Add the annotation to the following proof indicating for each line (1) the axiom schema it instantiates or (2) the prior lines it follows from by *modus ponens*.

Theorem. $\vdash (q\to r)\to((p\to q)\to(p\to r))$

1. $((p\to(q\to r))\to((p\to q)\to(p\to r)))\to((q\to r)\to((p\to(q\to r))\to((p\to q)\to(p\to r))))$
2. $(p\to(q\to r))\to((p\to q)\to(p\to r))$
3. $(q\to r)\to((p\to(q\to r))\to((p\to q)\to(p\to r)))$
4. $((q\to r)\to((p\to(q\to r))\to((p\to q)\to(p\to r))))\to(((q\to r)\to(p\to(q\to r)))\to((q\to r)\to((p\to q)\to(p\to r))))$
5. $((q\to r)\to(p\to(q\to r)))\to((q\to r)\to((p\to q)\to(p\to r)))$
6. $(q\to r)\to(p\to(q\to r))$
7. $(q\to r)\to((p\to q)\to(p\to r))$

∗**Exercise.** Annotate the following proof.

Theorem. $\vdash P\vee\sim P$

Proof

11. $\langle \{P,\sim(P\vee\sim P)\}, \sim(P\vee P)\rangle$
12. $\langle \{P,\sim(P\vee\sim P)\}, P\rangle$
13. $\langle \{P,\sim(P\vee\sim P)\}, P\vee\sim P\rangle$
14. $\langle \{P,\sim(P\vee\sim P)\}, \bot\rangle$
15. $\langle \{\sim(P\vee\sim P)\}, \sim P\rangle$
16. $\langle \{\sim(P\vee\sim P)\}, P\vee\sim P\rangle$
17. $\langle \{\sim(P\vee\sim P)\}, \sim(P\vee\sim P)\rangle$
18. $\langle \{\sim(P\vee\sim P)\}, \bot\rangle$
19. $\langle \varnothing, \sim\sim(P\vee\sim P)\rangle$
20. $\langle \varnothing, P\vee\sim P)\rangle$

∗**Exercise.** Construct natural deduction proofs of the following

3. $P\to Q, \sim Q \vdash \sim P$

4. $R\to\sim P, Q\to\sim R, P\vee Q \vdash \sim R$

5. $\forall x(Fx\to Gx), \exists xFx \vdash \exists xGx$

REVIEW QUESTIONS

1. Explain how the special certainty characteristic of the knowledge we have of logical relations can be explained by two factors:
   a. the fact that the set of theorems in proof theory has an inductive definition, and thus each element (theorem) the has a construction sequence (proof), and
   b. the basic elements (the axioms) and construction rules (the rules of inference) are defined in terms of the syntactic properties of signs, and therefore they are "epistemically transparent."

2. Give an example of showing that it is possible to prove from the axioms of set theory and the definition of $\mathfrak{I}$ for a given formal language, a metatheorem stating that a given sentence of (say) the propositional logic is a logical truth. (You could also do this for a metatheorem stating that a syllogism is valid in categorical logic or that a formula is a logical truth in first-order logic).

3. Give an example of showing that it is possible to produce a construction sequence (proof) for a given sentence of (say) propositional logic showing it is a member of the inductively set defined set of theorems of propositional logic. (You could also do this: produce a construction sequence (reduction) showing a syllogism is a member of the inductively defined set of acceptable syllogisms, or a construction sequence (proof) that a formula of first-order logic is a member of the inductive defined set of theorems of first-order logic.)

4. The set of logical truths of propositional logic defined in terms of truth in all interpretations $\mathfrak{I}$ has a different definition from the set of theorems of propositional logic defined as the closure of the axioms of propositional logic under the rule *modus ponens*. Nevertheless the two sets are the same. Explain why this fact is interesting.