Symbolic Logic

Notes on the Interpretation of First-Order Logic

Notes for Symbolic Logic
Fall 2005
John N. Martin

Table of Contents

Dialectica among the Seven Liberal Arts

Nicolà Pisano, Siena Cathedral, 1266

And yet the validity of logical sequences is not a thing devised by men, but is observed and noted by them that they may be able to learn and teach it; for it exists eternally in the reason of things, and has its origin with God. For as the man who narrates the order of events does not himself create that order; and as he who describes the situations of places, or the natures of animals, or roots, or minerals, does not describe the arrangements of man; as he who points out the stars and their movements does not point out anything that he himself or any other man has ordained; in the same way he who says, “when the consequence is false, the antecedent must be true,’ says what is most true; but he does not himself make it so, he only points out that it is so.

Saint Augustine, *On Christian Doctrine*, 32.

LECTURE 1. NAÏVE SET THEORY

The Axioms of Naïve Set Theory

Sets were studied intuitively in the 19th century by Georg Cantor (1845-1918) and later axiomatized by Gottlob Frege (1848-1925). A simplified account designed to highlight the central ideas was provided shortly afterwards by Bertrand Russell.[1] This is the version we shall review here. It is now called *naïve set theory*. It contains just three axioms. The first is an axiom that occurs in every axiom system in mathematics and science. It says simply that every truth of logic may be written down as a theorem in this axiom system. The axiom insures that all the truths discovered in the more basic science of logic can be carried over into the new system. It is the next two axioms that lay out the basic properties of sets themselves. They are written using the "primitive notation" of set theory, $\in$ and =.

Set Identity and The Principle of Extensionality

The symbol = is familiar. In set theory it is intended to stands for identity between sets. This relation is explained in Axiom 2, called the *Principle of Extensionality*, which lays out the "identity conditions" for sets. Philosophers sometime require "identity conditions" as a necessary requirement for an acceptable ontology. They admonish, "No entity without identity." This axiom satisfies that requirement. More precisely, the axiom sets out the conditions under which two names stand for one and

[1] In *Principle of Mathematics*, *op. cit.*

the same set. The stand for the same set if the sets they name have the same members.

Axiom 2. The Principle of Extensionality

$A{=}B \leftrightarrow$ for any x ($x{\in}A \leftrightarrow x{\in}B$)

Simply put, two sets are the same if and only if they have the same members.

(The axiom may be formulated in terms of a name's extension.[2] The *extension* of a set name is simply the set that it names. The axiom says that two names form a true identity sentence exactly when they have the same extension, i.e. exactly when they stand for the same set. It is this formulation in terms of extension that gives the axiom its name.)

Set Membership and the Principle of Abstraction

The Greek letter $\in$ (epsilon) is used to indicate set membership:

$x{\in}A$ is read *x is a member (or element of) A.*

We use $\in$ to classify, to assign entities to sets. In English we accomplish this by using the verb *to be* in one of its various senses. Thus, the following sentences all say the same thing:

Socrates is a human

Socrates is a member of the set of humans

Socrates $\in$ *the set of humans*

(The Greek letter epsilon is used for membership because in Greek verb *to be* (*einai*) begins with $\in$.)

[2] The idea goes back to Antoine Arnauld and Pierre Nicole in the *Port Royal Logic* (1645), Arnauld and Nicole thought it was ideas of indiviuals rather than indivials themselves that "fall under" a general term.

Axiom 3 uses $\in$ to declare the conditions under which a set exists: a set exists if its conditions for membership can be stated in language. Another way to say the same thing is that a set exists if it can be defined. Russell used the term *abstraction* to name the process of defining a set by its membership conditions, and calls this axiom *the Principle of Abstraction.*

But let us be clearer. What is it to state the "membership conditions" of a set? Briefly, it is to formulate a sentence that must be true of all and only the set's members. Let the variable x represent an arbitrary individual. Then, to formulate a condition is simply to write some sentence that must be true of x. For example, *x is red* is a sentence that describes a property of x. The axiom then says that the set of all x such that x is red exists. Again, $2 \leq x$ describes a property of x, The axiom says the set of all x such that $2 \leq x$ exists.

To say this in an axiom we must introduce some notation to represent a sentence that talks about an arbitrary individual. Let v be a variable. We use $P(v)$ to represent a formula P in which v occurs as a free-variable v. Later we shall also talk about formulas that contain two or more free variables. For example, *x loves y* is a formula with two free variables, and *x loves y but hates z* is one containing three. In general we shall represent a formula P with n free variables $v_1,\ldots,v_n$ by the notation $P(v_1,\ldots,v_n)$.

We may now state the Principle of Abstraction informally.

For any formula $P(x)$ the following is an axiom:
there exists an A such that for all x, $x \in A$ if and only if $P(x)$.

The axiom is stated formally as follows:[3]

Axiom 3. Principle of Abstraction

$$\exists A\ \forall x\ (\ x{\in}A \leftrightarrow P(x))$$

Inference Rules

Having now set down the theory's axioms, the next step is to lay down its rules of inference, and then to start deducing theorems. Naïve set theory has just one rule of inference:

Modus (podendo) ponens[4] or **detachment**. Satisfying the antecedent of a conditional proves that the consequent is true.

$$\begin{array}{l} P{\rightarrow}Q \\ \underline{\quad\ \ P} \\ \therefore Q \end{array}$$

In practice however we may use any rule of inference from sentential or first-order logic that we know to be valid, because these are all provable from *modus ponens* with standard logical truths.

Abbreviative Definitions

Set Abstracts

The next step is to introduce the abbreviations used in the theory. We begin with some notation that allows for a more useful formulation of the Principle of Abstraction. The principle assures us that if there is a sentence $P(x)$, we can make up

[3] Strictly speaking this is what logicians call an *axiom schema*, because there are as many axioms of this form as there are differenct open sentences of the form $P(y)$.
[4] [Given a conditional,] the way of positing [the consequent] by positing [the antecedent].

a set A that contains all and only the entities x such that $P(x)$ is true. It is useful to name this set by the notation $\{x \mid P(x)\}$, which is read *the set of all x such that* $P(x)$.

Definition. $\{x \mid P(x)\}$ abbreviates *the one and only A such that* $\forall x\ (x \in A \leftrightarrow P(x))$.

With this notation it is possible to say more directly that any element y is in a set defined by a property if and only if y possess that property:

Theorem 1. $\forall y(\ y \in \{x \mid P(x)\} \leftrightarrow P(y)\ )$.

The proof of theorem 1 is not difficult, but here it will be accepted without proof. A set name of the form $\{x \mid P(x)\}$ is called a *set abstract.*

Defined Relations on Sets

The next set of definitions introduce several usefully defined relations on sets: $\neq$, $\notin$, $\subseteq$, and $\subset$. These are genuine relations on sets, but they are relations that stand in a systematic relation to the primitive relations = and $\in$, and may be introduced by definition.

Abbreviation	**Phrase Abbreviated**	**How to read the notation out loud in English**	**The Abbreviation's Name**
$x \neq y$	$\sim(x=y)$	*x is not identical to y*	***non-identity*** or ***inequality***
$x \notin A$	$\sim(x \in A)$	*x is not an element of set A*	***non-membership***
$A \subseteq B$	$\forall x(x \in A \rightarrow x \in B)$	*A is a subset of B*	***subset***
$A \subset B$	$A \subseteq B \& \sim(A=B)$	*A is a proper subset of B*	***proper subset***

Defined Sets and Operations on Sets

The next set of definitions introduce notation for ways to names sets. First there are the names $\varnothing$ for the empty set (the set with nothing in it) and V for the universal set (the set of everything). Then there is the notation for the set operations: $\cap$ (intersection), $\cup$ (union), – (complementation), and **P** (the power set operation). Intuitively, the intersection of two sets is their overlap, the union of two sets is their combination, and the complement of a set includes everything outside the set, either without restriction (complement) or within a restricted range (relative complement). Lastly there is the abbreviation $\{x_1, \ldots, x_n\}$ that names a set by just listing its members $x_1, \ldots, x_n$.

Abbreviation	**Phrase Abbreviated**	**How to read the notation out loud in English**	**The Abbreviation's Name**
$\varnothing$ or Λ	$\{x \mid x \neq x\}$	*the empty set*	***empty set***
V	$\{x \mid x = x\}$	*the universal set*	***universal set***
$A \cap B$	$\{x \mid x \in A \& x \in B\}$	*the intersection of A and B*	***intersection***
$A \cup B$	$\{x \mid x \in A \vee x \in B\}$	the *union* of *A* and *B*	***union***
$A - B$	$\{x \mid x \in A \& x \notin B\}$	the *relative complement of B in A*	***relative complement***
$-A$	$V - A$	*the complement of A*	***complement***
$\mathbf{P}(A)$	$\{B \mid B \subseteq A\}$	*the set of subsets of A*	***power set***
$\{x_1, \ldots, x_n\}$	$\{y \mid y = x_1 \vee \ldots \vee y = x_n\}$	*the set containing* $x_1, \ldots, x_n$	

In the above abbreviations the particular variable used in not important. Others may be substituted (like *y* for *x* in the definitions above) so long as the variable is new and it replaces every occurrence of the variable being replaced.

Notice that $x \neq y$, $x \notin A$, $A \subseteq B$, and $A \subset B$ are all sentences. Hence they are true or false. On the other hand, $\varnothing$, Λ, V, $A \cap B$, $A \cup B$, $A - B$, $-A$, $\mathbf{P}(A)$, and $\{ x_1, \ldots, x_n \}$ are not sentences (they are not either true or false). They are names of sets. In ordinary grammar there is a huge difference between a name and a sentence. Names stand for entities, sentences combine names with verbs and make assertions about entities that are either true or false. It hard not to spot the difference between names and sentences in English, but it is easy to loose track of which is which in the new notation of set theory. Keep your eyes open.

Statement of the Axiom System

Having stated the axioms, inference rules, and definitions, we are now in a position to prove theorems. We begin by summarizing the axioms, rules and definitions, and the list of theorems we shall prove.

Summary of the System

Axioms

Logical Truth	*Every truth of logic is a theorem.*
Extensionality.	$A{=}B \leftrightarrow \forall x\,(x{\in}A \leftrightarrow x{\in}B)$
Abstraction.	$\exists A\ \forall x\ (\ x{\in}A \leftrightarrow P(x))$

Rules of Inference

Modus ponens
Modus tollens
Disjunctive Syllogism
Hypothetical Syllogism
Conjunction
Addition
Universal Generalization
Universal Instantiation
Construction
Existential Instantiation
Substitution of Logical Equivalents
Reductio ad absurdum
Ex Falso Quodlibet
Conditional Proof
Conditional Proof for Biconditionals
Proof by Cases

Abbreviative Definitions

Abbreviation	Phrase Abbreviated	Abbreviation's Name

$\{x \mid P(x)\}$	*the one and only A such that* $\forall x\ (x \in A \leftrightarrow P(x))$	***set abstract***
$x \neq y$	$\sim(x=y)$	***non-identity***
$x \notin A$	$\sim(x \in A)$	***non-membership***
$A \subseteq B$	$\forall x(x \in A \rightarrow x \in B)$	***subset***
$A \subset B$	$A \subseteq B \& \sim A=B$	***proper subset***
$\varnothing$ or Λ	$\{x \mid x \neq x\}$	***empty set***
V	$\{x \mid x=x\}$	***universal set***
$A \cap B$	$\{x \mid x \in A \& x \in B\}$	***intersection***
$A \cup B$	$\{x \mid x \in A \vee x \in B\}$	***union***
$A-B$	$\{x \mid x \in A \& x \notin B\}$	***relative complement***
$-A$	$V-A$	***complement***
$\mathbf{P}(A)$	$\{B \mid B \subseteq A\}$	***power set***
$\{x_1, \ldots, x_n\}$	$\{y \mid y = x_1 \vee \ldots \vee y = x_n\}$	

Reduction of Relations to Sets

From the early days of logic in ancient Greece, relations have been puzzling. Plato, in addition to the ordinary Forms that make subject-predicate sentences true, posits the Forms called Sameness, Difference, and Identity. Similarly, Aristotle posits a special category for relations. Both doctrines seem to presuppose that relational truths linking two proper names can be explained as some conjunction of simple subject-predicate truths. But in an earlier lecture we saw the problems this analysis faces.

Set theory rises to the challenge. Relational assertions can be represented in set theory without supplementing its ontology or assumptions. Relations in this sense are "reduced to" sets.

To see how this is done, let us review what relations are. Just as in Aristotle's ontology qualities like whiteness or rationality constitute a commonality shared by two substances, relations are what pairs may have in common. The pairs Cain and Able, Castor and Polux, Romulus and Remus all share the fact that they are brothers. Each instance of brotherhood requires that there be two people, or in other words, a pair. This pair is said to stand in the brotherhood relation. Realists go further and claim that relations are actual entities. They do so to "explain" what the pair Cain and Able has in common with the pair Castor an Polux by positing the existence of a relation as a special sort of "universal" that can be instantiated in multiple pairs. Set theory offers a similar account. Instead of positing a new category of entity, however, set theory stipulates that relations are a special sort of set.

To sketch the account we must first explain what a pair is in set theory. Consider the *less than* relation. It holds of many different pairs <*x,y*> such that *x is less than y,* including, for example, the pairs <1,2>, <5,7>, and <36,215>. These all share the feature that the first is less than the second. Notice however that if the pairs are reverse, the relation fails. In the pairs <2,1>, <7,5>, and <215,36> the first is not less than the second. In technical jargon, the less-than relation is asymmetric: if *x* is less than *y*, it is not the case that *y* is less than *x.*

According, logicians say that the order of the pair "makes a difference". We must define a pair so that <*x,y*> is not the same as <*y,x*> except in the unusual case in which *x* and *y* are the same thing. Pairs that obey this rule are said to be *ordered.* A *two-place relation*, which is what we call a relation that holds beteen the elements of a pair, will then be defined as a set of ordered pair.

There are also, however, relations that hold among triples. For example, it takes three things for there to be a case of *between-ness.* Utah is between Nevada and Colorado, Cincinnati is between Dayton and Lexington. These are *three-place* relations. In principle there are also four-place relations, which hold among groups of four things, and likewise for any number you choose. Logicians generalize this fact and allow for relations among ordered groups of any size. An ordered series of *n* elements is called an *n-tuple* and is represented by the notation $<x_1,\ldots,x_n>$. As in the case of two-place relations, the order continues to matter. If x is between y and z, then y cannot be between x and z. An *n-place relation* is then defined as a set of ordered n-tuples.

In languages like English relations are tied to characteristic grammatical forms. For example, two place relations are typically expressed in English by subject-verb-object sentences, like *x loves y,* and *x teaches y.* They are also expressed by sentences that link a subject to an "oblique object" by an intransitive verb and a preposition, as in *x talks to y,* and *x sits under y.* Comparative adjectives also link two *relata*, for example *x is taller than y, x is less than y,* and *x is sillier than y.* Possessive expressions also link two objects, as in *x is the brother of y,* and *x is the creator of y.* All these syntactic forms share the feature that they link two proper noun phrases. Three place relations link three proper noun phrases, as in *x is between y and z, x talked to y about z,* and *x saw y sitting on z.* In general, an open sentence P with n free variables $x_1,\ldots,x_n$,, which is represented by $P(x_1,\ldots,x_n)$, can be used to describe what is shared by a group of ordered n-tuples $<x_1,\ldots,x_n>$.

To define an *n-tuple* $<x_1,\ldots,x_n>$ within set theory, we have to find some definition that makes $<x_1,\ldots,x_n>$ different from $<y_1,\ldots,y_n>$ except in the unusual case in which each x_i is identical to y_i. Be forewarned that the definition usually given is not very intuitive because it does not provide a very natural paraphrase of what we mean by *pair* in English. In the context of the theory, however, it works very well. It allows that an n-tuple's order matters; it allows us to define relations as sets of n-tuples; and it allows us to prove a body of desired theorems about relations. To state the definition efficiently, let us abbreviate the string of quantifiers $\forall x_1 \forall x_2 \ldots \forall x_n$, (which says *for all* $x_1,\ldots,x_n$) by the shorter form $\forall x_1,\ldots,x_n$. We first define ordered-pair, and then using it define ordered n+1-tuple on the assumption that we have defined an ordered n-tuple.

Definitions

$<x,y>$	means	$\{x,\{x,y\}\}$
$<x_1,\ldots,x_{n+1}>$	means	$<<x_1,\ldots,x_n> x_{n+1}>$

We now state without proof the theorem that says that the order makes a difference. (Though not difficult, we do not state the proof for this and several later theorems because the details are irrelevant to the topics in these lectures.)

Theorem 25. $\forall x_1,\ldots,x_n, y_1,\ldots,y_n$

$$<x_1,\ldots,x_n>=<y_1,\ldots,y_n> \leftrightarrow (x_1 = y_1 \ \& \ \ldots \ \& \ x_n = y_n)$$

Let us now group all n-tuples into a set and call this set V^n. [5] Any set of n-tuples then is a subset of V^n. We use this fact to define *n-place relation.*

Definitions

V^n	means	$\{\{<x_1,\ldots,x_n>\mid x_1 \in V \& \ldots \& x_n \in V\ \}$
R is a *n-place relation*	means	$R \subseteq V^n$

Since relations are sets, the principles of extensionality and abstraction apply to them. Two n-place relations are identical if and only if they are made up of the same n-tuples. Similarly, if $P(x_1,\ldots,x_n)$ is a formula with free variables $x_1,\ldots,x_n$, then there is a set R (an n-place relation) such that any n-tuple $<x_1,\ldots,x_n>$ is in R if and only if $P(x_1,\ldots,x_n)$ is true.

Theorem 26. (Extensionality for Relations).

$$(R\subseteq V^n \;\&\; S\subseteq V^n) \rightarrow$$
$$(R=S \leftrightarrow \forall x_1,\ldots,x_n(<x_1,\ldots,x_n>\in R \leftrightarrow <x_1,\ldots,x_n>\in S))$$

Theorem 27. (Abstraction for Relations).

$$\exists R\,\forall\, x_1,\ldots,x_n\,(<x_1,\ldots,x_n>\in R \leftrightarrow P(x_1,\ldots,x_n))$$

As with sets in general, it is possible to refer to relations by set abstracts: $\{<x_1,\ldots,x_n>|\ P(x_1,\ldots,x_n)\}$ is the set of all n-tuples $<x_1,\ldots,x_n>$ such that $P(x_1,\ldots,x_n)$. Abstracts allow us to express the Principle of Abstraction for relations is a simple form:

Theorem 28. $\forall\, y_1,\ldots,y_n\,(<y_1,\ldots,y_n>\in \{<x_1,\ldots,x_n>|\ P(x_1,\ldots,x_n)\} \leftrightarrow P(y_1,\ldots,y_n))$

Properties of Relations and Order

In the last section we accomplished the our main theoretical goal, namely of explaining what relations are within set theory. Here we shall list some of the basic properties of relations that logicians frequently use. Some you will recognize because they have already been introduced informally. In order to make the notation more natural, we shall sometimes rewrite the relational assertion $<x,y>\in R$ in the subjet-veb-object order xRy (so-called *infix*) familiar to English speakers. We shall also say that a two place relation is *a relation on a set A* if all its *relata* are in *A,* i.e.$\forall x,y\,(xRy \rightarrow (x\in A \;\&\; y\in A)$.

[5] The notation derives from the fact that the number of *n*-tuples formed from elements of a set *A* is

Definitions. Properties of Relations. A two-place relation R is said to be:

reflexive iff $\forall x$, xRx

transitive iff $\forall x,y,z$, $((xRy \,\&\, yRz) \rightarrow xRz)$

symmetric iff $\forall x,y$, $(xRy \rightarrow yRx)$

asymmetric iff $\forall x,y$, $(xRy \rightarrow \sim yRx)$

antisymmetric iff $\forall x,y$, $((xRy \,\&\, yRx) \rightarrow x=y)$

connected iff $\forall x,y$, $(xRy \vee yRx)$

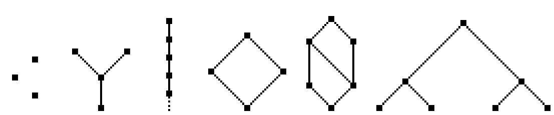
By imposing a relation with these properties on a set its elements may be "ordered".

Definitions. Orderings. A two-place relation R on U is said to be:

partial ordering on U iff R is reflexive, transitive and antisymmetric

total ordering on U iff R is reflexive, transitive and antisymmetric, and connected

A partial order is imposes a minimum amount of structure.



Examples of Partial Orderings

But adding to a partial order the property of connectedness forces the elements to form a line.

Functions

A special sort of relation is one that allows us to identity an entity indirectly by first finding something that it is related to and then using the relation to pinpoint the entity

---

precisely the number of entities in A raised to the power n.

itself. We can find Philip of Macedon, for example, by first finding his son Alexander the Great and then pin-pointing the entity that fathered him. Let us be set-theoretic. Let R be $\{<x,y>|\ x \text{ is fathered by } y\}$. Then $<Alexander, Philip>\in R$. Alexander is the one and only entity paired to Philip in the relation *R.* This is true because *R* uniquely pairs a *relatum* on the left side with one on the right side. More formally, *R* obeys this law:

$$\forall x,y\ \ (<x,y>\in R\ \&\ <x,z>\in R) \to y\text{=z}.$$

In this case *R* is said to be a *function*, and we rewrite $<x,y>\in R$ as $R(x)$=y. Hence in this case $R(x)$ is read *the father of,* and $R(Alexander)=Philip$ is read *the father of Alexander is Philip.* Though *R* is a two-place relation, it is called a *one-place* function, because the notation $R(x)$ has only one variable place.

There are *n*-place functions as well. These are *n*+1-place relations such that if the first *n* members of entities that stand in the relation uniquely pinpoint the $n+1^{\text{th}}$ member.

Definitions

1. An $n+1$ place relation f is called a *n*-place *function* iff
   $\forall\ x_1,\ldots,x_n,y,z,\quad (<x_1,\ldots,x_n,y>\in f\ \&\ <x_1,\ldots,x_n,z>\in f) \to y\text{=z}.$

2. if f is an *n*-place function, we write $<x_1,\ldots,x_n,y>\in f$ as $f(x_1,\ldots,x_n)=y$

Though functions are extremely important in applications of logic to mathematics and we shall see some examples in these lectures, we include them here mainly because they are important to the next topic, the analysis of the notion of a "structure."

Construction and Inductive Definitions

Definitions by Necessary and Sufficient Conditions

We return in this section to the topic of definition. Let us review its history. We saw in the Platonic dialogues that Socrates seeks definitions as answers to *What is*? questions. For example in the Carmedies, he seeks the definition of *temperance*, and in the Republic the definition of *justice*. In the *Euthyphro* when trying to define *piety* Socrates tells Euthyphro that a list of examples will not do. He wants the "general idea,"

> Remember that I did not ask you to give me two or three examples of piety, but to explain the general idea which makes all pious things to be pious. Do you not recollect that there was one idea which made the impious impious, and the pious pious? (6d)

He is alluding to the Platonic Form of Piety. In Plato's theory any true subject-predicate proposition *All F are G* is like a definition because, if true, it describes an immutable fact about the participation of one Platonic Idea in another.

Aristotle and his followers propose a more plausible account. They make a distinction between definitions and other sorts of truths. They contrast conventional agreements to use words to stand for particular concepts, which they call *nominal definitions*, with the necessary natural laws of generic classifications, which they call *real definitions*, and both sorts of definition are contrasted with contingent matters of fact. Real definitions are supposed to observe a fixed form: a species is defined by its genus and its difference. But Aristotelian essentialism is not accepted by modern science.

Definition today is understood, rather, as a part of a scientific theory. Although moder definitions are not understood to stand for Aristotelian forms, they do sometimes look structurally similar to traditional Aristotelian definitions. This is especially true of some abbreviate definitions for set names in sciences that make use of set theory. Consider some examples we have already met:

$\varnothing = \{x|x \neq x\}$

$A \cup B = \{x|x \in A \vee x \in B\}$

These definition fit a general form:

$A = \{x|\ P(x)\}$

In virtue of the Principle of Abstraction, this kind of definition can be recast in an equivalent form as a biconditional:

$\forall x\ (x \in A \leftrightarrow P(x))$

Moreover, it is not unusual for the defining condition to be spelled out even further as a conjunction $P_1(x)$ &…& $P_n(x)$ of conditions. That is, frequently a definition takes this form:

$A = \{x|\ P_1(x)\ \&\ldots\&\ P_n(x)\ \}$

When it does so, it entails the theorem:

$\forall x\ (x \in A \leftrightarrow (P_1(x)\ \&\ldots\&\ P_n(x)))$

Each $P_i\,(x)$, considered as an individual conjunct, is said to be a *necessary condition* for membership in A, and all the conditions together, i.e. the complete conjunction $P_1(x)$ &…& $P_n(x)$, is called the *sufficient condition* for membership in A. One example is the definition of $A \cap B$:

$A \cap B = \{x|\ x \in A\ \&\ldots\&\ x \in B\ \}$

which entails

$\forall x\ (x{\in}A \leftrightarrow\ x{\in}A\ \&\ldots\&\ x{\in}B)$.

Aristotelian real definitions have a similar structure:

$\forall x$ (*x is a Men* $\leftrightarrow$ (*x is rational* & *x is animal*))

The pattern of analysis in terms of necessary and sufficient conditions still has a firm grip on philosophers. Some of the most central claims of epistemology, ethics, and metaphysics are formulated in theses with this structure:

Knowledge is justified true belief
Truth is correspondence with the world.
The good is what maximizes total social utility.
God is the most perfect being.

However, as scientific principles, definitions in terms of necessary and sufficient conditions are problematic.

First of all, in logical theory, which is formulated in set theory, they must be careful to avoid contradictions. As we have seen, the unrestricted axiom of abstraction leads to paradoxes, and it is the application of this very principle that makes definitions by necessary and sufficient conditions possible. Any choice of necessary and sufficient conditions must be crafted to avoid these technical problems.

Secondly, a term can be introduced into a theory by an eliminative definition only if the terms used to formulate the definition (i.e. the terms in the *definiens*) are themselves already part of the theory. It is hard for a philosophical theory or for a logical theories that employs philosophical ideas to met this goal. For example, to define *knowledge* as *justified true belief*, the notion of *truth* must already be part of the theory, either explained by the axioms or by an earlier definition. Likewise a theory that explains *truth* as *corresponds with the world* would need a definition or axioms that explains *world.* No serious mathematical theory comes close to explaining these difficult ideas.

Even purely logical theories have difficulty with definitions of this sort. Conceptually, for example, one might like to define a *logical truth* as a sentence that we can "know" is true from its shape alone. But any such definition would use the word *know*, and we have no satisfactory background theory of knowledge in which to embed it.

Technical difficulties, and difficulty in defining background ideas thus prompt logicians to seek alternatives to the use of necessary and sufficient conditions. It is one such technique that is our topic here. It is definition by construction.

Inductive Definitions and Sets

Instead of defining a set by membership conditions, the technique simply constructs the set. We so in stages. First we specify some initial elements. Next, we lay down some rules for making new elements from old. We then expand the set of initial elements by applying the rules to them. This set is then expanded yet again by applying the rules to its members. The process is repeated, *ad infinitum* if necessary, until no further elements can be added. A set that is constructed in this way is said to be defined by *induction*. (Here the term *induction* has a specialized sense, and has nothing in common with concepts of the same name in statistics or physics.) We summarize the process in the following definition:

Definition. An *inductive system* is any $<E,R,C>$ such that

1. *E* is a set of basic elements;

2. R is a set of relations;
3. C is the set such that[6]
   a. E is a subset of C;
   b. if the elements $x_1,\ldots,x_n$ are in C and bear the relation R to x_{n+1}, then x_{n+1} is in C;
   c. nothing else is in C.

If the set C is defined inductively in this way from a set of basic elements E and s set of rules R, we say that C is defined by *closing E under R.*

It is possible to add more restrictions to that would insure that C will not generate paradoxes.[7] Were we to do so, the set would be genuinely constructive in a strict sense.

In these lectures we have already encountered one important example of an inductively defined set. We used it without remarking on its unusual definition. This is the set of theorems in naïve set theory. Indeed we defined two sets inductively. First we defined the set of *simple theorems*. This was the set that consists of the closure of all instances of logical truths and the axioms of set theory under the non-subproof rules. We then defined the set of *theorems*. This is the closure of the set of simple theorems under the set of all inference rules including the subproof rules. At this point however, it will more instructive to look in detail at two simpler examples of inductive systems.

---

[6] The definition of C can be stated entirely in the notation of set theory. First we define the intersection of a family $\{F_1,\ldots,F_n,\ldots\}$ of sets as $F_1\cap\ldots\cap F_n,\ldots$:

$$\cap\{F_1,\ldots,F_n,\ldots\} = F_1\cap\ldots\cap F_n,\ldots$$

The we define Cas follows:

$$C=\bigcap\{B|\ E\subseteq B\ \&\ (\langle x_1,\ldots,x_n,x_{n+1}\rangle\in R\ \&\ \{x_1,\ldots,x_n,\}\subseteq B)\rightarrow x_{n+1}\in B\}$$

[7] For example, that the basic set or the set of rules the be countable.

Let us consider first an inductive description of “score keeping” as done in a game like cribbage. Let us start “keeping score” by drawing a single vertical line: |. Let us have a rule called *adding one* that consists of drawing a new line | to the right of whatever we apply the rule to. next to it on the right. That is, if we apply the rule to |, we get ||. If we apply it to ||, we get |||. If we apply it to |||, we get ||||, etc. We now define by induction the set of *scores*:

An *scoring system* is any <{ | }*,adding one, scores*> such that

a. { | } is a subset of *scores;*
b. if the elements *x* is in *scores,* then the entity we get by *adding one* to *x* is in *scores*;
c. nothing else is in *scores.*

It follows that *scores* = { |,||,|||,||||,|||||,||||||,|||||||,||||||||,|||||||||,.... }.

Induction is thus a simple method for defining quite large sets – *scores* for example is infinite – yet we do so by construction without having to list necessary and sufficient conditions for elements of the set.

The Natural Numbers

Another standard example of a set defined by induction is the set ***N*** of *natural numbers*, which consists of all the positive integers 1,2,3,… plus 0. Let us work through it in some detail because though easy to state, it illustrates the power of inductive definitions. The set is constructed. We start with 0 as the only initial element. We then define the so-called successor relation. The natural numbers then are inductively defined as the set of all entities that can be constructed from 0 by the successor relation.

The entire construction can be done in set theory if 0 and the successor relation are defined in terms of sets. Lets do so here, not because we will be doing any arithmetic, but to illustrate how a real construction of this sort is done in mathematics. Be forewarned. Because we are constraining ourselves to use notions only from set theory, the definitions of 0 and successor will not be very intuitive. But once stated we will be able to show that they work very well. That is, give the definitions and the background theorems of set theory, we can then prove all the theorems of elementary arithmetic. The definitions "work," in other words, by yielding as theorems the right theoretical results.

The basic idea is that the number n is defined as a set that has exactly n things in it. This means that 0 should have nothing in it, i.e. that 0 should be $\varnothing$. It also means that the successor relation should take the number n, which is a set that has n things in it, and make up its successor n+1 (which we shall indicate with the notation $\boldsymbol{S}(x)$) by adding a new element to the set n that was not already in n. What entity should be added to n? The standard trick is just to add the set n itself. This works as a definition of successor, not because it is very intuitive, but because the new entity n is a a genuine entity (it exists because it is a set) and because the set n itself is not an element of n, but it is perfectly possible to make up a new set that all all the original elements from n plus a new element, namely the set n itself. In the notation of set theory, "adding a new entity" is accomplished by taking the union of the original set with a set that has the new entity: i.e. $\boldsymbol{S}(x)= x\cup\{x\}$.

Thus the construction starts by defining 0 as $\varnothing$. Then we define the successor of *x,* indicated by the notation ***S***(*x*), as $x\cup\{x\}$. The set of natural numbers ***N*** is then defined by induction in terms of 0 and ***S.***

Definition. The inductive *system of natural numbers* is <{0}***,S,N***> such that

4. $0=\varnothing$;
5. ***S***(x)= $x\cup\{x\}$;
6. ***N*** is the set such that
   d. {0} is a subset of ***N;***
   e. if $x\in$***N*** and ***S***(x)= *y,* then $x\in$***N*** *;*
   f. nothing else is in ***N.***

These definitions, which at first may seem odd, are justified because they entail just the right theorems – they generate the right "theory". Below some of these standard definitions and theorems are listed, not because we will be using them – you already know elementary arithmetic – but to illustrate how the definition generates the right theory:

- each natural number exists – because it is a set – and is definable, for example,

  $0 = \varnothing$

  $1 = 0+1 = \varnothing+1 = \varnothing\cup\{\varnothing\} = \{\varnothing\}$

  $2 = 1+1 = \{\varnothing\}+1 = \{\varnothing\}\cup\{\{\varnothing\}\} = \{\varnothing,\{\varnothing\}\}$

  $3 = 2+1 = \{\varnothing,\{\varnothing\}\}+1 = \{\varnothing,\{\varnothing\}\}\cup\{\{\varnothing,\{\varnothing\}\}\} = \{\varnothing,\{\varnothing\},\{\varnothing,\{\varnothing\}\}\}$
- there are in infinite number of natural numbers,
- each natural number has exactly as many members as the number suggests:

  0, aka $\varnothing$, has no members,

  1, aka $\{\varnothing\}$ has one member, namely $\varnothing$,

  2 , aka $\{\varnothing,\{\varnothing\}\}$, has two members, namely $\varnothing$ and $\{\varnothing\}$

3, aka {∅,{∅},{∅,{∅}}}, has three members, namely ∅, {∅}, and {∅,{∅}} etc.

- $n \leq m$ is definable because it turns out that $n \leq m \leftrightarrow n \subseteq m \leftrightarrow n \in m$
- the addition function + is definable:
  a. $x + 0 = x$
  b. $x + \boldsymbol{S}(y) = \boldsymbol{S}(x + y)$
- the multiplication function × is definable:
  a. $x \times 0 = 0$
  b. $x \times \boldsymbol{S}(y) = (x \times y) + x$

Indeed, these are the definitions that generate the structure <***N***,≤,+,×,0,1> which most you spent hours working out the details of in high school algebra.[8]

Here in a philosophy class, we are not going to do algebra, but make points about the general nature of scientific explanation. We have here an example in which one "science", the algebra of the natural numbers, is reduced to or subsumed within another "science", set theory. The sentences that were true in algebra then become theorems of set theory because if all the abbreviative definitions in theorems mentioning numbers, + or × were translated out into their defining notation, the resulting formulas would be theorems of set theory. Thus we see the explanatory power of axioms systems like set theory and of techniques like inductive definitions: "counting numbers" are entities that can be explained in a well developed theory (set theory) , the set of "counting numbers" can be given a special sort of definition that was not available to Aristotle or traditional philosophy (an inductive definition), and the

[8] You may for example have learned to work out the equations that are true in an "ordered field" <***N***,≤,+,×,0,1> or "distributive ring," which are structures that obey "laws" that are exemplified and abstracted from the natural numbers.

"counting numbers" literally form a structure with well defined relations and operations ($\leq$, $+$, $\times$) and two special entities (0 and 1).

Proof by Induction

The special form of an inductive definition also introduces a special way to prove things about set. To show all members of an inductive set have a property, all we have to show is two things:

1. that the initial elements have a property,
2. that if we apply a construction rule to something with that property, the result also has that property.

Every element of the set would then have to have the property because every element is either an initial element or results from applying one of the rules to earlier elements of the set.

Theorem. Let $<E,R,C>$ be an inductive system.

If 1. $E \subseteq A$, and

2. for any r in R, if ($<x_1,\ldots,x_n>$ bears r to x_{n+1} & $\{x_1,\ldots,x_n\} \subseteq B$) $\rightarrow$ $x_{n+1} \in A$

then $C \subseteq A$

A proof of this sort is called a proof *by induction*.

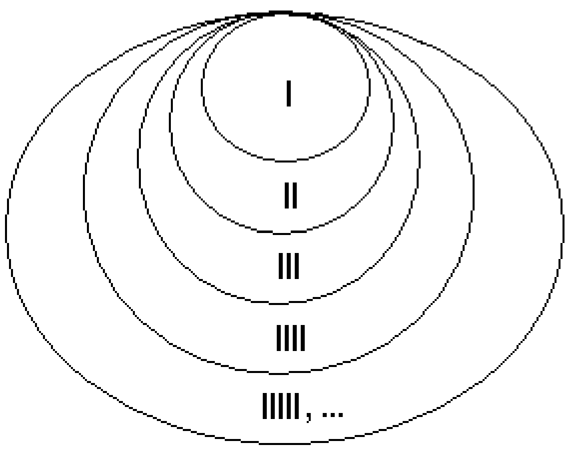
Construction Sequences

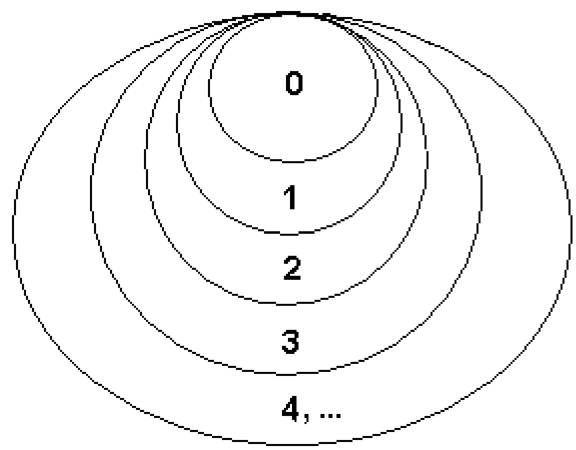
One of the reasons that inductive sets are theoretically attractive is that unlike definitions by abstraction they insure that for every element of the set, there is a finite construction process that places that element in the set. This construction moreover can be set out in what is called a construction sequence.

Definition. If <*E,R,C*> is an inductive system, then *construction sequence for x* relative <*E,R,C*> is a finite series <$y_1,...,y_n$> such that $y_n=x$ and each y_i is either in *E* or bears some relation in *R* to earlier members of the series.

Theorem. If <*E,R,C*> is an inductive system, then
($x \in C$ iff there is a construction sequence for *x* relative to <*E,R,C*>).

The existence of a sequence terminating in *x* is therefore evidence that *x* is in the set. We do not have to show that *x* meets a list of necessary and sufficient conditions. Rather we construct the right sequence. The technique is different. For example <|,||,|||,||||,|||||,||||||,|||||||> is a construction sequence of |||||| and is evidence that it is a member of the set of *scores*. Likewise the fact that <0,1,2,3,4,5,6,> is a construction sequence of 6 show that 6 is a natural numbers.



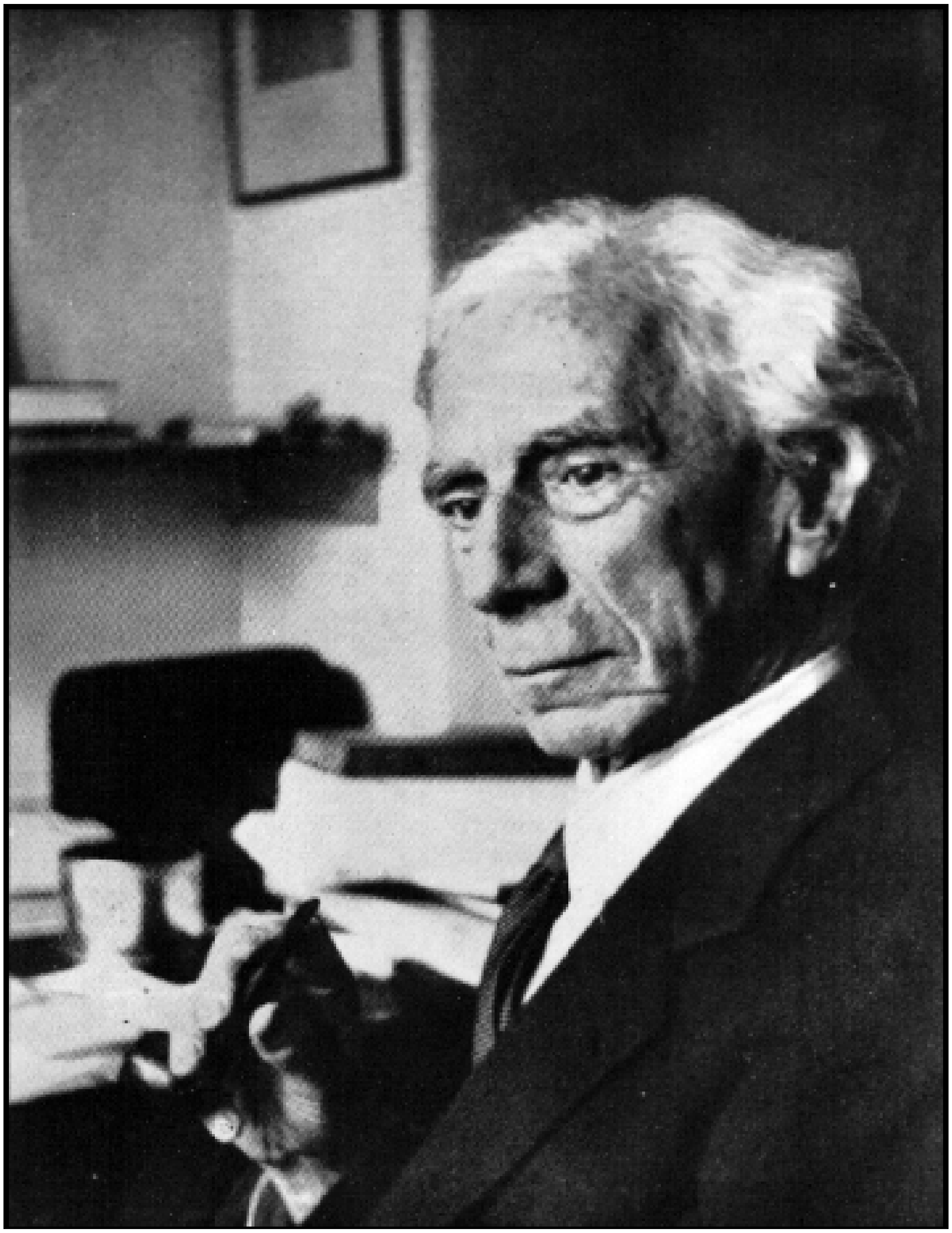


In later lecture we shall met important examples of this device in logic. What for example is a *sentence*? In high school you learned it was something that expresses a “complete thought”, but what is a “thought”? Try finding a mathematically precise theory of thought! In a later lecture we shall define the set of sentences inductively, and show that something is a sentence, not by appeal to thoughts, but by constructing it in a construction sequence from simpler sentences. Similarly, instead of defining a *logical truth* in terms of knowledge as contemplated sketched earlier, we shall show define the set of *logical theorems* inductively and then show that something is a theorem if is the last line in the sort of construction sequence known as a proof.

We have in fact been using this technique in "characterizing" the "truths" of set thoery. Instead of trying to define this set in terms of necessary and sufficient conditions, we defined by reference to "proofs" the set of theorems of naïve set theory. But the notion of proof we used, and indeed the proofs we have been construction to show the theorems of naïve set theory are just construction sequences.

Let us review the sequence of earlier ideas. First we defined the notion of a *simple proof*, and then the notion of *proof*. A *simple proof* is any sequence of formulas that are either (1) truths of logic or instances of the axioms, or (2) follow from earlier lines of the sequence by a non-subproof rule. A *simple theorem* was then defined as any formula that is the last line of some simple proof. Thus, a simple proof is nothing other than a construction sequence for the set of simple theorems. Moreover, a formula is a simple theorem if and only if it is the last line element of a constructions sequence for simple theorems, i.e. the last line of a simple proof. Likewise, a *proof* was defined as any sequence of formulas or simple proofs such that each element of the sequence is either (1) a law of logic, an axiom instance or a simple proof, or (2) follows from earlier elements of the sequence by one of the inference rules. That is, a proof is nothing other than a construction sequence for the set of theorems. Accordingly, a formula is a theorem if and only if it is the last line of a construction sequence for a theorem, i.e. of a proof.

In the course of these lectures we shall see numerous examples of sets that cannot easily be defined by traditional necessary and sufficient conditions, but which are definable inductively and thus allow membership to be fixed by construction sequences. Indeed the applicability of these methods is one of the distinctive features of logic as science.



(© Roger Viollet)

Bertrand Russell, aged 77 years

"Cantor had a proof that there is no greatest cardinal; in applying this proof to the universal class, I was led to the contradiction about classes that are not members of themselves. It soon became clear that this is only one of an infinite class of contradictions. I wrote to Frege, who replied with the utmost gravity that '*die Arithmetik is ins Schwanken geraten*.' At first I hoped that the matter was trivial and could easily be cleared up; but early hopes were succeeded by something very near to despair. Throughout 1903 and 1904, I pursued will-o'-the wisps and made no progress. At last, in the spring of 1905, a different problem, which proved soluble, gave the first glimmer of hope. The problem was that of descriptions, and its solution suggested a new technique."
Bertrand Russell, *My Philosophical Development*, 1943

Axiomatized Set Theory

Russell's paradox and similar contradictions entailed by the axiom presents a serious problem. Indeed there is no greater flaw in a mathematical theory than a contradiction. As Russell recounts, Frege, who used essentially this axiom system to deduce the laws of arithmetic, wrote to him that the discovery raised doubts in his mind about the truth of arithmetic itself.

A number of diagnoses were proposed for the root of the problem. Russell's own account is that the principle errs in allowing sets that are ***un***grounded in the sense that they may form ∈-loops. These are sets that may be a ∈-descendent of themselves, for example a set *x* such that there is some chain $x \in y \in \ldots . \in z \in x$. Here *x* is a member of something that is a member of something in a ∈-hierarchy that eventually leads to a member of *x* itself. In 1910-13, together with Alfred North Whitehead (1861-1947), Russell published *Principia Mathematica*, an important work that revises the axioms so as to proscribe sets that form ∈-loops. It does so by proposing the so-called theory of types in which sets form ranks such that only elements of one rank can enter into sets of the next. With this restriction an element *x* of rank *n* cannot be an element of itself at rank *n*+1. As far as is known, this new system is consistent. It does, however, require additional axioms, including a so-called axiom of reducibility, which requires, without much intuitive plausibility, that the set theoretic relations at higher levels be replicated in the structure of elements at the lowest level. Though the theory is technically successful in entailing the theorems necessary for applications of set theory to mathematics, a more intuitively plausible account is now preferred.[9]

This second explanation of the paradoxes is due to Ernst Zermelo (1871-1953).[10] According to his analysis the problem with the Principle of Abstraction is that it is over generous it the size of the sets it asserts exist. According to the principle, a set of any size may exist so long as it is definable. Indeed, it directly implies that the universal set V exists, and there can be no set bigger than the set of everything. Russell's set $\{x \mid x \notin x\}$ too is "very large". It includes as a subset another very large set, the set of all cardinal numbers, which was shown independently to entail a contradiction (the Burali-Forti paradox). Zermelo proposes a new axiom system that specifies we start with a limited variety of sets, which are "small" enough that we can be fairly sure that they do not entail contradictions. These "starter sets" are limited to the empty set (the empty set axiom) and a set of countably many entities like the positive integers (the axiom of infinity). The system then specifies a restricted number of ways in which new sets may be constructed from those we previously know exist. One method is definability, but definability is restricted. Definable sets exist only if there is another set that we know already exists and either the old set contains the new set as one of its subsets (axiom of separation) or the elements of the old set can be mapped onto the elements of the new set (axiom of replacement). In addition to definability there are several other construction methods: forming a "pair" out of two previously existing sets (the pairing axiom), taking their union (the union axiom),

[9] For an account of the theory of types, which is accessible with even the limited logic of these lectures, see Irving M. Copi, *The Theory of Types* (London: Routledge & Kegan Paul, 1971).
[10] There are many introductions to set theory, but a good account that stresses philosophical issues is Shaughan Levine, *Understanding the Infinite* (Cambridge: Harvard, 1994).

forming a power set (the power set axiom), forming a set by taking a representative from each set in a family of already existing sets (axiom of choice).[11] A far as is known, the system is consistent, and it has be developed within the branch of mathematics known as *axiomatized set theory* to demonstrate a large number of results, interesting to both mathematicians and philosophers.

For these lectures we shall follow the practices of most users of set theory. We shall continue to use the naïve version even though we know that strictly speaking it is contradictory. We will do so, however, with the understanding that we will allow ourselves to talk only about sets that are not "too big" and that we know in principle could be shown to exist in the more precise versions of axiomatized set theory. The upshot for the purposes of this lecture is that set theory can in fact be developed into a plausible, mathematically precise theory, which we may apply, as we have earlier in this lecture, to give a convincing explanation of sameness and difference.

[11] The axioms of Zermelo-Frankle Set Theory, usually called *ZF*, are more precisely stated as follows:

1. *Axiom of Separation.* Let $P(x)$ be an open sentence. $\forall A\exists B\ \forall x(x\in B \leftrightarrow (x\in A \wedge P(x)))$
2. *Union Axiom.* $\forall A\forall B$, $A\cup B$ exists.
3. *Pair Axiom.* $\forall x\forall y$, $<x,y>$ exists.
4. *Power Set Axiom.* $\forall A$, $\mathbf{P}(A)$ exists.
5. *Axiom of Infinity.* An infinite set exists. (Below the set N={0,1,2,3,…} of *natural numbers* is defined. This axiom may be phrased: *N exists.*)
6. *Axiom of Replacement.* $\forall A\ \forall f\ (f\text{“}A$ exists), where $f\text{“}A=\{y|\ \exists x\ \ y=f(x)\}$
7. *Axiom of Choice.* For any family of sets *F,* a choice set of *F* exists, where *C* is a *choice set* of *F* iff for any $A\in F$, there is one and only one element *x* of *A* such that $x\in C$.

Sentential Syntax

Modern Symbolic Notation

The modern treatment of the connectives begins with the introduction of symbolic notation for the representation of logical arguments in mathematics in the mid 19th century. Formulas complex enough to state mathematical propositions required the sentential connectives. Gottlob Frege invented his own symbolization, called the *begriffsschrift* ("concept writing") in his groundbreaking set theoretic axiomatization of arithmetic (1879).[12] The standard modern symbolization began in 19th century studies of arithmetic by Dedekind and Peano[13]. It became regularized in the notation of Bertrand Russell and Alfred North Whitehead in *Principia Mathematica* in the early 20th century, and has evolved little since. A third standard symbolization was invented by Polish logicians in the early 20th century. It is still in use and excels other notation in its simplicity.

Frege's notation was designed for use in an axiom system. A formula starts with a short vertical line, |, indicating that the formula that follows it to the left is a theorem. The formula then continues with a horizontal line. The horizontal is an assertion sign. It indicates that the formula that follows to the left is being asserted as true. Thus every formula in his system stats with the symbol

[12] See Gottlob Frege, *Begriffsschrift, a Formal Langauge, Modeled Upon that of Arithemetic, for Pure Thought*, Jean van Heijnoort, trans., *From Frege to Gödel* (Cambridge: Harvard University Pres, 1967)

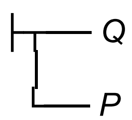
├── which is read *It is a theorem that it is true that ….* To indicate a conjunction of two sentences *P* and *Q,* Frege joins them one to the left of the other connected by a horizontal line:

├── *P* ── *Q*

which is read *It is a theorem that it is true that P and it is true that Q*. To indicate a negation, Frege inserts a short vertical bar from the horizontal prior to a formula. Thus, a formula

├─┬─ *P*

is read, *It is a theorem that it is true that it is not the case that it is true that P*. He indicates the conditional *if P then Q* by subjoining an assertion of the antecedent *P,* namely ── *P,* to that of *Q,* ── *Q,* by a vertical line. Thus



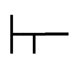
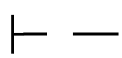
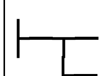
which is read*, It is a theorem that it is true that if it is true that P, then it is true that Q.* Frege has no special notation for disjunction but it may be expressed by means of negation and the conditional because $P\vee Q$ is equivalent to $\sim P\rightarrow Q$.

Polish notation uses letters for connectives: N for negation, K for conjunction (*konjunction* in Polish), A for disjunctions (*alternation* in Polish), C for the conditional, and E for the biconditional (equivalence in Polish). The placement of the connectives differs from standard notation in that a two-place connective is place to the left of the formulas it joins and no parentheses are used. Thus K*PQ*

[13] Guiseppe Peano, *Arithemetices Principia, Nova Method Exposita* (Turin: Fratres Bocca, 1889)

is read *P and Q,* and APNQ is read *P or not Q.* Thus, $(P\&{\sim}(Q\rightarrow R))$ is written K*P*NC*QR.*

Standard notation, which we are using in these lectures, derives from that of Russell and Whitehead. They used the dot • for conjunction, ∨ for disjunction, ⊃ (called *the horseshoe*) for the conditional, and ≡ (called *triple bar*) for the biconditional. This notation is still in use. The ∨ for disjunction comes from the Latin word *vel* which means *or.*[14] Though up to this point we have been using the ampersand & for conjunction[15], from now we shall use for *and* the symbol ∧, which is the more usual symbol in technical logic. It comes from turning ∨ on its head, which makes some sense in that conjunction is the logical "dual" of disjunction.[16]

	negation	conjunction	disjunction	conditional	biconditional
Frege					
Polish	N	K	A	C	E
Russell	~	•	∨	⊃	≡
Modern	¬	∧	∨	→	↔

[14] The more standard word for *or* in Latin is *aut*, which is normally used when there is a contrast between *P* and *Q.* Thus *P aut Q* tends to mean *P or Q but not both.* However, the normal use of *vel* in Latin is to list items that may or may not be mutually disjoint. That is, *P vel Q* trends to mean *P or Q or possibly both,* which is the desired meaning of *or* in logic.
[15] From "and *per se* and". It represents the Latin word *et* which means *and.* It is formed by a combination of the letter e with a cross bar ~ from the letter t: ɇ .
[16] Given DeMorgan's Laws and Double Negation, you can show that if all disjunctions in a formula are replaced by conjunctions, and every formula (atomic or complex) that is negated has its negation removed, and every formula (atomic or complex) that is unnegated has a negation inserted, then the result will be logically equivalent. Such pairs are said to be *duals* to one another, e.g. $\sim(P\vee\sim Q)\wedge R$ is dual to (and hence logically equivalent to) $\sim((\sim P\wedge Q)\vee\sim R)$.

Formation Rules, Generative Grammar, Inductive Sets

In the early days of symbolic logic, logicians merely declared what symbols they would be using for what and set about writing. They did not pause to formulate the rules of grammar for their symbolic languages very carefully. In the 1920's, however, Rudolf Carnap (1891-1970) showed how to state the rules for formal grammar.[17]

Any high school student who has been forced to diagram sentences and then had to argue with his or her teacher about whether their diagram was right – something I remember doing with some irritation – will remember that the rules for diagramming were not very well defined. The reason is that the rules for English grammar are not very well defined. Indeed, the entire field of grammar of the sort you learned in high school – and which is still taught by most English professors – is little more developed than the grammar known by Donatus and Priscian for ancient Greek and Latin. Modern linguists were well aware of this fact and attempted to advance the field in the early decades of the 20th century but without much success. Important advances were made however in the 1950's and 60's with the work of Noam Chomsky, who applied the techniques of generative grammar to natural languages. It is fair to say that Chomsky's revolution in grammar consists in large part of applying to natural languages techniques that were first explored for formal languages by Carnap and subsequent logicians.[18]

In more modern terms what Carnap did was show how the set of grammatical formulas could be defined. His definition is not the traditional sort

---

[17] Rudolf Carnap, *The Logical Syntax of Language* (London: Kegan Paul *et al*., 1937).
[18] See Frederick Newmeyer, *Linguistics in America*, 2nd ed. (Orlando: Academic Press, 1986)

common in philosophy that defines as set in terms of necessary and sufficient conditions. Rather it is constructive. His method consists of first laying down a set of *atomic* expressions and set of *formation rules*. The set of grammatical expressions is then defined as the closure of the atomic expressions by the rules – i.e. it is the set of all formulas that can be constructed from the atomic sentences by the rules.

Before defining the set of sentences we must choose the atomic formulas we shall use. Let us arbitrarily assume these to be $p_1,\ldots,p_n,\ldots$. We well also define the basic formation rules. There will be five of these, one for each connective. The rule for negation will be a 1-place function because it takes a single sentence as input (argument) and produces a negated formula as its output (value). The rules for conjunction, disjunction, conditional, and biconditional are 2-place functions because they take two inputs (a pair of sentences as argument) and produce a complex sentence as their output (value).

Definition. A *sentential syntax* is a structure $<ASen,FR,Sen>$ such that

7. *ASen,* called the set of *atomic sentences, is* a subset of $\{p_1,\ldots,p_n,\ldots\}$;
8. *FR,* called the set of *formation rules,* is a set of functions $\{fr_{\sim}, fr_{\wedge}, fr_{\vee}, fr_{\rightarrow}, fr_{\leftrightarrow}\}$ defined as follows:
   a. $fr_{\sim}(x)=\sim x$
   b. $fr_{\wedge}(x,y)= (x\wedge y)$
   c. $fr_{\vee}(x,y)= (x\wedge y)$
   d. $fr_{\rightarrow}(x,y)= (x\rightarrow y)$
   e. $fr_{\leftrightarrow}(x,y)= (x\leftrightarrow y)$
9. *Sen* is the set such that
   g. ASen is a subset of *Sen;*

h. if the elements *P,* and *Q* are in *Sen,* then $fr_{\sim}(P)$, $fr_{\wedge}(P,Q)$, $fr_{\vee}(P,Q)$, $fr_{\rightarrow}(P,Q)$, $fr_{\leftrightarrow}(P,Q)$ are in *Sen;*

i. nothing else is in *Sen.*

Strictly speaking the formation rules of the two-place connectives $\wedge$, $\vee$, $\rightarrow$ and $\leftrightarrow$ always form a sentence with an outside pair of parentheses, e.g. the rule of $\rightarrow$ produces $(p_3\rightarrow(p_2\rightarrow p_3))$ rather than $p_3\rightarrow(p_2\rightarrow p_3)$. In practice we shall often delete the outer most set to make sentences easier to read.

Grammatical Derivations

A constructive definition of this sort has a number of interesting theoretical properties. Not the least of these is that it succeeds as a definition. Prior to definitions of this sort, there just was no rigorous way to define the set of grammatical sentences. Chomsky and later linguists are working on the hypothesis that some such generative definitions will also work for natural languages.

A second feature of the definition follows from the fact that it is constructive, and therefore that membership in the set is demonstrable by producing a construction sequence. As we saw in Part 1, a set is constructive if and only if there exists, for each element of the set, a construction sequence that shows step by step how the element was added to the set. Accordingly, for each well-formed sentence there is a construction sequence that shows it is so. Linguists call these sequences *grammatical derivations* though they should not be confused with proofs in a logical sense. They do not show that a sentence is *true*, only that it is *grammatical*. Both a sentence and its negation, for example, are grammatical, and hence have construction sequences, but they are not both true, and hence could not both have *proofs* that they are true.

Let us consider some examples. Recall that like proofs, a construction sequence is a series such that each element is either a basic element, which in sentential grammar is an atomic sentence, or is produced from an earlier elements of the series by one of the generative rules, which in sentential grammar are the formation rules. We shall display a grammatical construction in the style used by linguists as a list of lines going down the page. We shall also annotate the construction by writing next to each line how it was obtained, either from the set of atomic sentences or by the application of a formation rule to earlier lines.

Grammatical Metatheorem. The following are in *Sen:*

1. $((\sim p_4 \vee p_2) \wedge p_1)$
2. $\sim(p_2 \vee \sim p_2)$
3. $((\sim\sim p_4 \leftrightarrow p_1) \rightarrow (\sim(p_6 \vee \sim p_1) \wedge p_3))$
4. $((p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2)))$

The theorem is prove by producing a grammatical derivation (construction sequence) for each:

1. p_2	atomic
2. p_4	atomic
3. $\sim p_4$	2, $fr_{\sim}$
4. $(\sim p_4 \vee p_2)$	2 & 3, $fr_{\vee}$
5. $((\sim p_4 \vee p_2) \wedge \sim p_4)$	4 & 2, $fr_{\wedge}$

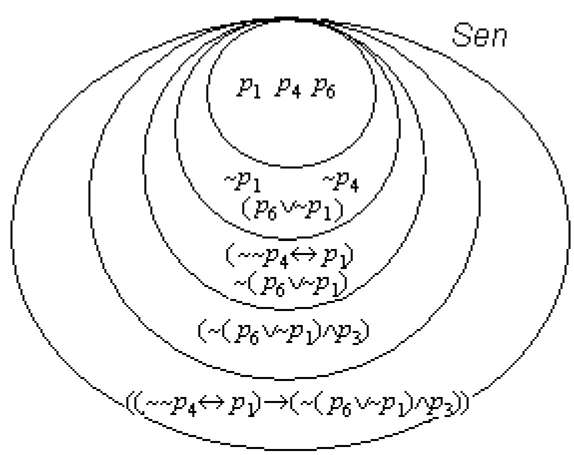
1. p_2	atomic
2. $\sim p_2$	2, $fr_{\sim}$
3. $(p_2 \vee \sim p_2)$	1 & 2, $fr_{\vee}$
4. $\sim(p_2 \vee \sim p_2)$	3, $fr_{\sim}$

1. p_1	atomic
2. p_3	atomic
3. p_4	atomic
4. p_6	atomic
5. $\sim p_1$	1, $fr_{\sim}$
6. $\sim p_4$	3, $fr_{\sim}$
7. $\sim\sim p_4$	6, $fr_{\sim}$
8. $(\sim\sim p_4 \leftrightarrow p_1)$	7 & 1, $fr_{\rightarrow}$
9. $(p_6 \vee \sim p_1)$	4 & 1, $fr_{\vee}$
10. $\sim(p_6 \vee \sim p_1)$	9, $fr_{\sim}$
11. $(\sim(p_6 \vee \sim p_1) \wedge p_3)$	10 & 2, $fr_{\wedge}$
12. $((\sim\sim p_4 \leftrightarrow p_1) \rightarrow (\sim(p_6 \vee \sim p_1) \wedge p_3))$	8 & 11, $fr_{\rightarrow}$

1. p_1	atomic
2. p_2	atomic
3. $(p_1 \leftrightarrow p_2)$	1 & 2, $fr_{\leftrightarrow}$
4. $(p_1 \wedge p_2)$	1 & 2, $fr_{\wedge}$
5. $\sim p_1$	1, $fr_{\sim}$
6. $\sim p_2$	2, $fr_{\sim}$
7. $(\sim p_1 \wedge \sim p_2)$	5 & 6, $fr_{\wedge}$
8. $((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2))$	4 & 7, $fr_{\vee}$
9. $((p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2)))$	3 & 8, $fr_{\leftrightarrow}$

The additions to *Sen* as stipulated by the third construction sequence may be

illustrated as follows:



Exercise. Provide grammatical derivations (construction sequences) showing that the following are in *Sen:*

1. $\sim(p_1 \leftrightarrow \sim p_1)$
2. $((p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \rightarrow p_2) \wedge (p_2 \rightarrow p_1)))$
3. $(p_1 \rightarrow (p_1 \vee (p_2 \wedge \sim p_2)))$

Having explained in some detail the grammar of the connectives, it is now time to talk about what they mean.

Truth-Functionality

Truth-Tables for the Connectives

The first observation to make about the meaning of the connectives is that they are *truth-functional* in a precise sense: given the truth-value of the parts of a sentence formed by a connective, there is a rule corresponding to that connective that determines uniquely the truth-value of the whole. These rules are customarily stated in what are called the *truth-tables* for the connectives:

P	$\sim P$
T	F
	T

P	Q	$P \wedge Q$	$P \vee Q$	$P \rightarrow Q$	$P \leftrightarrow Q$
T	T	T	T	T	T
T	F	F	T	F	F
F	T	F	T	T	F
F	F	F	F	T	T

Negation

The first table sets out the rule for negation, where $\longrightarrow$ (the "long arrow") means "is paired with" (this is not the material conditional $\rightarrow$):

Negation $\quad \begin{array}{l} T \longrightarrow F \\ F \longrightarrow T \end{array}$

P	$\sim P$
T	F
F	T

In set theory this "rule" is understood as a set of pairs:

$tf_{\sim} = \{\langle T,F\rangle, \langle F,T\rangle\}$

Note that this is a one-place function since each initial value is uniquely paired with a second value, as the above diagram illustrates. Hence we can write:

$\langle T,F\rangle \in tf_{\sim}$ as $tf_{\sim}(T)=F$
$\langle F,T\rangle \in tf_{\sim}$ as $tf_{\sim}(F)=T$

Disjunction

The rule for disjunction is:

Disjunction $\quad \begin{array}{l} T,T \longrightarrow T \\ T,F \longrightarrow T \\ F,T \longrightarrow T \\ F,F \longrightarrow F \end{array}$

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

In set theory the "rule" is a set of triples:

$tf_{\vee} = \{\langle T,T,T\rangle, \langle T,F,T\rangle, \langle F,T,T\rangle, \langle F,F,F\rangle\}$

Note that this is a two-place function since each initial pair of values is uniquely paired with a third value. Hence we can write:

$\langle T,T,T\rangle \in tf_{\vee}$ as $tf_{\vee}(T,T)=T$
$\langle T,F,T\rangle \in tf_{\vee}$ as $tf_{\vee}(T,F)=T$
$\langle F,T,T\rangle \in tf_{\vee}$ as $tf_{\vee}(F,T)=T$
$\langle F,F,F\rangle \in tf_{\vee}$ as $tf_{\vee}(F,F)=F$

Conjunction

The next rule is that for conjunction:

Conjunction

$$\begin{aligned} T,T &\longrightarrow T \\ T,F &\longrightarrow F \\ F,T &\longrightarrow F \\ F,F &\longrightarrow F \end{aligned}$$

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

In set theory the “rule” too is really a set of triples:

$$tf_{\wedge} = \{<T,T,T>,<T,F,F>,<F,T,F>,<F,F,F>\}$$

Since this is a two-place function, we can write:

$<T,T,T> \in tf_{\wedge}$	as	$tf_{\wedge}(T,T)=T$
$<T,F,F> \in tf_{\wedge}$	as	$tf_{\wedge}(T,F)=F$
$<F,T,F> \in tf_{\wedge}$	as	$tf_{\wedge}(F,T)=F$
$<F,F,F> \in tf_{\wedge}$	as	$tf_{\wedge}(F,F)=F$

The Conditional

The next rule is that for the conditional:

The Conditional

$$\begin{aligned} T,T &\longrightarrow T \\ T,F &\longrightarrow F \\ F,T &\longrightarrow T \\ F,F &\longrightarrow T \end{aligned}$$

P	Q	$P \rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

In set theory the “rule” too is really a set of triples:

$$tf_{\rightarrow} = \{<T,T,T>,<T,F,F>,<F,T,T>,<F,F,T>\}$$

Since this is a two-place function, we can write:

$<T,T,T> \in tf_{\rightarrow}$	as	$tf_{\rightarrow}(T,T)=T$
$<T,F,F> \in tf_{\rightarrow}$	as	$tf_{\rightarrow}(T,F)=F$
$<F,T,T> \in tf_{\rightarrow}$	as	$tf_{\rightarrow}(F,T)=T$
$<F,F,T> \in tf_{\rightarrow}$	as	$tf_{\rightarrow}(F,F)=T$

The Biconditional

The next rule is that for the biconditional:

The Biconditional

$$\begin{array}{lcl} T,T & \longrightarrow & T \\ T,F & \longrightarrow & F \\ F,T & \longrightarrow & F \\ F,F & \longrightarrow & T \end{array}$$

P	Q	$P \leftrightarrow Q$
T	T	T
T	F	F
F	T	F
F	F	T

In set theory the "rule" is a set of triples:

$$tf_{\to} = \{\langle T,T,T\rangle,\langle T,F,F\rangle,\langle F,T,F\rangle,\langle F,F,T\rangle\}$$

Since this is a two-place function, we can write:

$\langle T,T,T\rangle \in tf_{\to}$	as	$tf_{\to}(T,T)=T$
$\langle T,F,F\rangle \in tf_{\to}$	as	$tf_{\to}(T,F)=F$
$\langle F,T,F\rangle \in tf_{\to}$	as	$tf_{\to}(F,T)=F$
$\langle F,F,T\rangle \in tf_{\to}$	as	$tf_{\to}(F,F)=T$

Sentential Semantics

Tarski's Correspondence Theory for Complex Grammars

The standard definition of truth is that a sentence is true if it corresponds to the world. For the time being let us divide this task into two parts: explaining truth for atomic (simple) sentences and explaining it for complex (molecular) sentences. We shall concentrate in this section of the second task: defining truth as correspondence for molecular sentences. Here we shall assume that every atomic sentence is either true or false in a world. Our task then, for the moment, is to then explain what it would be for molecular sentences to be true or false.

In the 1930's the Polish logician Alfred Tarski (1902-1983) provided a solution to the problem.[19] He rejects the requirement of a traditional definition by necessary and sufficient conditions. Instead he defines truth inductively. Moreover, he does so in such a way that there is a sense in which even complex sentences can be said to "correspond to the world".

[19] For an account of Tarski's theory in his own words see Alfred Tarski, "Truth and Proof", *Scientific American* 194 (1968), 63-77, and "The Semantic Conception of Truth", *Philosophy and Phenomenological Research* 4 (1944), 341-375.

In the propositional logic the truth-value of a complex proposition is ultimately determined by the truth-values of its *atomic* parts. Moreover we are assuming that the interpretation of an atomic part consist of a truth-value. Truth-values, however, are not the sort of thing that we would normally count as entities “in the world”. It is odd to say, for example, that a sentence “refers” to the value T or F. In the 19th century, however, the logician Gottlob Frege did exactly that. He recommended that we regard sentences as “standing for” truth-values. Since it is odd to think of truth-values as entities in the world, it is odd to say that we can explain how a sentence “corresponds to the world” by defining the truth-value of a whole sentence in terms of the truth-values of its parts. But, following in Frege’s footsteps, this is just what Tarski does. Applying the mathematical method know as abstraction, he “abstracts” those features shared by both true simple and complex sentences. The method presumes that it is these common features that contain “the core” of the correspondence. The resulting commonality is then judged to capture the central idea of correspondence. What is it that true simple and complex sentences share?

Let us call the *basic referring parts* of a sentence (in an abstract sense) its grammatically simple parts that in a given an interpretation stand for something in the world. In the syllogistic these are the subject predicate terms of a categorical proposition because it is these that are the basic expressions from which more complex ones are formed and because they are the terms that are given an interpretation “in the world”. In the propositional logic, a sentence’s basic referring parts are its atomic sentences because it is from these that the sentence is constructed, and it is these that constitute the atomic parts that are in an interpretation have a truth-value. The “essence” then that Tarski takes to be indicative of a correspondence theory, in an abstract sense, is the general rule that the "referents of an expression's atomic parts determine that of the whole".

He makes this precise in terms of *truth-conditions.* By the *truth-conditions* of *P* relative to $\mathfrak{I}$, which we shall abbreviate as $TC_{\mathfrak{I}}(P)$, we shall mean the conditions that must hold in the world among the various entities assigned by $\mathfrak{I}$ to the referring parts of *P.* Thus, apart from necessary mathematical concepts, the only entities that $TC_{\mathfrak{I}}(P)$ talks about are the entities that $\mathfrak{I}$ assigns to the atomic or basic terms in *P.* These "assignments" are entities in the world but, possibly, only in a rather abstract sense. In the syllogistic they are relatively normal denizens of "the world", namely the subsets of the universe of existing things. In propositional logic, however, atomic sentences "refer" to truth-values, which can be called "entities in the world" only in a rather abstract sense.

By stating conditions on the $\mathfrak{I}$-values of referring parts of *P* , $TC_{\mathfrak{I}}(P)$ states what relations must hold among these values in order for *P* to "correspond to the world". Thus, in propositional logic, $TC_{\mathfrak{I}}(P)$ will accordingly state what must hold among the truth-values of the atomic parts of *P* in order for *P* to be true. The important conceptual point here is that its is fair to say that $TC_{\mathfrak{I}}(P)$ defined this way do, in an abstract sense, states what it is for *P* to "correspond to the world". Since what is common to the definition of truth for both simple and complex sentences is the fact that their truth is explained in terms of conditions on the interpretation-values of their parts, it is this feature that is abstracted as the content of the idea "corresponds to the world."

Tarski summarizes his view in a simple way by proposing a criterion that he says must be met by any theory that calls itself a genuine correspondence theory of truth. Every correspondence theory, he says, should entail, for every sentence in the language, a statement that it is true relative to an interpretation if and only if its truth-conditions hold under that interpretation. More precisely, let $TC_{\Im}(P)$ be a sentence in the metalanguage formulated only in mathematical that states some condition on the $\Im$-values of the atomic or basic expressions of *P.* Tarksi's criterion for an acceptable correspondence theory, then, is that it should entail, for every sentence of the syntax, a metatheorem of the form: [20]

In the propositional logic $TC_{\Im}(P)$ may be define a statement in the metalanguage such that

a. $TC_{\Im}(P)$ is equivalent to $\Im(P)$=T and

b. $TC_{\Im}(P)$ is formulated only in terms of conditions on the $\Im$-values of the atomic parts of *P.*

We shall see below that we can in fact prove an instance of (T) in this sense for the sort of theory advocated by Tarski.

The Strategy for an Inductive Definition

To state the inductive definition of "an interpretation" for the propositional logic, Tarski's strategy is to use the truth-functions for the connectives. The method understands an interpretation $\Im$ to be two-place relation in the set theoretic sense, i.e. an interpretation is a set of pairs $<P,V>$, the first element of which is a sentence *P* and the second element is the truth-value *V* that the interpretation assigns to *P* in $\Im$. It is assumed that every interpretation is two-valued (*bivalent*) in the sense that *V* must be either T or F. Moreover, in the set theoretic sense an interpretation is a function, i.e. it is a relation that assigns only one truth-value to each sentence. Thus we may rewrite the fact that

[20] The sense of *iff* in (T) is $\leftrightarrow$, which is equivalent to $\rightarrow$ in both directions. Since $\rightarrow$ is the material conditional , (T) is sometimes called Tarski's *material adequacy condition* .

$\langle P,V\rangle \in \Im$ in functional notation: $\Im(P)=V$. Accordingly, $\Im(P)=V$ means that the sentence P has the value V in the interpretation $\Im$.

To define any set inductively, we first stipulate a set of basic elements, and then define a set of construction rules. To define the particular set $\Im$ inductively, we must stipulate a basic set of sentence truth-value pairs. In this case we form the basic set by taking each atomic sentence and forming a pair by joining the sentence with a truth-value. This pair will declare the truth-value of that atomic sentence in $\Im$.

Next we define a set of rules that makes new elements of $\Im$ from old. These rules will make new sentence truth-value pairs from others. The key idea is to use truth-tables. If we know what truth-values $\Im$ assigns to the parts of sentence formed by a connective, we can use the connective's truth-function to calculate the truth-value that $\Im$ should assign to the whole sentence. For example, if $\langle P,\mathrm{T}\rangle \in \Im$ and $\langle Q,\mathrm{F}\rangle \in \Im$, we know we should put $\langle P\wedge Q,\mathrm{F}\rangle \in \Im$ because $tf_{\wedge}(\mathrm{T},\mathrm{F})=\mathrm{F}$. That is, if $\Im$ assigns T to P but F to Q, we know it should assign F to $P\wedge Q$ because the truth-table $tf_{\wedge}$ tells us a conjunction with a false conjunct should be false.

Interpreting Negations

If P is paired with the truth-value V to $\Im$, we add the pair consisting of $\sim P$ and the opposite truth-value. Let us assume that an interpretation is bivalent, i.e. assigns either T or F but not both. Then we can formulate this rule is several equivalent ways, getting shorter each time:

Negation Rule

1. If $\langle P,\mathrm{T}\rangle \in \Im$ then $\langle \sim P,\mathrm{F}\rangle \in \Im$

If $<P,\mathrm{F}>\in\mathfrak{I}$ then $<{\sim}P,\mathrm{T}>\in\mathfrak{I}$

2. If $<P,V>\in\mathfrak{I}$ then $<P,tf_{\sim}(V)>\in\mathfrak{I}$

3. $\mathfrak{I}({\sim}P)=tf_{\sim}(\mathfrak{I}(P))$

All three formulations say the same thing. They each describe the same rule for adding a pair to $\mathfrak{I}$ that consist of a negated sentence and its truth-value. We use a similar method for the other connectives.

Interpreting Disjunctions

If P is paired with the truth-value V to $\mathfrak{I}$, and Q with the truth-value V' in $\mathfrak{I}$, then we add the pair consisting of $P\vee Q$ and the value T if either V or V' is T, otherwise we add the pair consisting of $P\vee Q$ and F, as the truth-table for $\vee$ stipulates. Again, we can formulate this rule in several equivalent ways, getting shorter each time.

Disjunction Rule

a. If $<P,\mathrm{T}>\in\mathfrak{I}$ and $<Q,\mathrm{T}>\in\mathfrak{I}$, then $<P\vee Q,\mathrm{T}>\in\mathfrak{I}$
If $<P,\mathrm{T}>\in\mathfrak{I}$ and $<Q,\mathrm{F}>\in\mathfrak{I}$, then $<P\vee Q,\mathrm{T}>\in\mathfrak{I}$
If $<P,\mathrm{F}>\in\mathfrak{I}$ and $<Q,\mathrm{T}>\in\mathfrak{I}$, then $<P\vee Q,\mathrm{T}>\in\mathfrak{I}$
If $<P,\mathrm{F}>\in\mathfrak{I}$ and $<Q,\mathrm{F}>\in\mathfrak{I}$, then $<P\vee Q,\mathrm{F}>\in\mathfrak{I}$

b. If $<P,V>\in\mathfrak{I}$ and $<Q, V'>\in\mathfrak{I}$, then $<P\vee Q,tf_{\vee}(V,V')>\in\mathfrak{I}$

c. $\mathfrak{I}(P\vee Q)=tf_{\vee}(\mathfrak{I}(P), \mathfrak{I}(Q))$

Again, all three of these say the same thing. They each describe the same rule for adding a pair to $\mathfrak{I}$ that consist of a disjunction and its truth-value.

Interpreting Conjunctions

If P is paired with the truth-value V to $\mathfrak{I}$, and Q with the truth-value V' in $\mathfrak{I}$, then we add the pair consisting of $P\wedge Q$ and T if both V and V' are T, otherwise we add the pair consisting of $P \wedge Q$ and F, as the truth-table for $\wedge$ declares. We formulate this rule in three ways, getting shorter each time.

Conjunction Rule

1. If $<P,\mathrm{T}> \in \Im$ and $<Q,\mathrm{T}> \in \Im$, then $<P \wedge Q,\mathrm{T}> \in \Im$
   If $<P,\mathrm{T}> \in \Im$ and $<Q,\mathrm{F}> \in \Im$, then $<P \wedge Q,\mathrm{F}> \in \Im$
   If $<P,\mathrm{F}> \in \Im$ and $<Q,\mathrm{T}> \in \Im$, then $<P \wedge Q,\mathrm{F}> \in \Im$
   If $<P,\mathrm{F}> \in \Im$ and $<Q,\mathrm{F}> \in \Im$, then $<P \wedge Q,\mathrm{F}> \in \Im$

2. If $<P,V> \in \Im$ and $<Q, V'> \in \Im$, then $<P \wedge Q, tf_{\wedge}(V,V')> \in \Im$

3. $\Im(P \wedge Q) = tf_{\wedge}(\Im(P), \Im(Q))$

Interpreting the Conditional

If P is paired with the truth-value V to $\Im$, and Q with the truth-value V' in $\Im$, then we add the pair consisting of $P \rightarrow Q$ and T if V is T or V' is F, and we add $P \rightarrow Q$ with F if V is T and V' is F, as the truth-table for $\rightarrow$ dictates. We formulate this rule in three ways, getting shorter each time.

The Rule for the Conditional

1. If $<P,\mathrm{T}> \in \Im$ and $<Q,\mathrm{T}> \in \Im$, then $<P \rightarrow Q,\mathrm{T}> \in \Im$
   If $<P,\mathrm{T}> \in \Im$ and $<Q,\mathrm{F}> \in \Im$, then $<P \rightarrow Q,\mathrm{F}> \in \Im$
   If $<P,\mathrm{F}> \in \Im$ and $<Q,\mathrm{T}> \in \Im$, then $<P \rightarrow Q,\mathrm{T}> \in \Im$
   If $<P,\mathrm{F}> \in \Im$ and $<Q,\mathrm{F}> \in \Im$, then $<P \rightarrow Q,\mathrm{T}> \in \Im$

2. If $<P,V> \in \Im$ and $<Q, V'> \in \Im$, then $<P \rightarrow Q, tf_{\rightarrow}(V,V')> \in \Im$

2. $\Im(P \rightarrow Q) = tf_{\rightarrow}(\Im(P), \Im(Q))$

Interpreting the Biconditional

If P with the truth-value V to $\Im$, and Q with the truth-value $V'$is in $\Im$, we add the pair $P \leftrightarrow Q$ with T if V and V' are the same, and we add $P \leftrightarrow Q$ with F if V and V' are different, as the truth-table for $\leftrightarrow$ declares. We formulate this rule in three ways, getting shorter each time.

The Rule for the Biconditional

1. If $<P,\mathrm{T}>\in\Im$ and $<Q,\mathrm{T}>\in\Im$, then $<P\leftrightarrow Q,\mathrm{T}>\in\Im$
   If $<P,\mathrm{T}>\in\Im$ and $<Q,\mathrm{F}>\in\Im$, then $<P\leftrightarrow Q,\mathrm{F}>\in\Im$
   If $<P,\mathrm{F}>\in\Im$ and $<Q,\mathrm{T}>\in\Im$, then $<P\leftrightarrow Q,\mathrm{F}>\in\Im$
   If $<P,\mathrm{F}>\in\Im$ and $<Q,\mathrm{T}>\in\Im$, then $<P\leftrightarrow Q,\mathrm{T}>\in\Im$

2. If $<P,V>\in\Im$ and $<Q, V'>\in\Im$, then $<P\leftrightarrow Q,tf_{\leftrightarrow}(V,V')>\in\Im$

3. $\Im(P\leftrightarrow Q)=tf_{\leftrightarrow}(\Im(P), \Im(Q))$

The Inductive Definition of Interpretation

We can now define the set of sentential interpretations $\Im$ for a sentential syntax $<ASen,FR,Sen>$ as follows. Let V be either T or F. First we define a "basic set". This is the set of interpretation-value pairs limited to atomic sentences. By a *basic set Atomic*-$\Im$ we mean some functional pairing of atomic sentences with the truth-values T and F.

Atomic-$\Im$ is a set such that for any p_i in *ASen,*

1. there is some V, such that $<p_i,V>$ is in *Atomic*-$\Im$ and
2. p_i is not paired with more than one value . That is, if $<p_i,V>\in$*Atomic*-$\Im$ and $<p_i,V'>\in$*Atomic*-$\Im$, then $V=V'$.

Note that if there are n atomic sentences, there are 2^n basic sets *Atomic*-$\Im$.

The interpretation $\Im$ *relative to Atomic*-$\Im$ is the set of pairs defined inductively as follows:

1. *Atomic*-$\Im\subseteq\Im$ (i.e. if $<p_i,V>\in$*Atomic*-$\Im$, then $<p_i,V>\in\Im$)
2. Construction Steps:
   a. If $<P,V>\in\Im$ then $<\sim P,tf_{\sim}(V)>\in\Im$

b. If $<P,V>\in\Im$ and $<Q, V'>\in\Im$, then $<P\vee Q,tf_{\vee}(V,V')>\in\Im$

c. If $<P,V>\in\Im$ and $<Q, V'>\in\Im$, then $<P\wedge Q,tf_{\wedge}(V,V')>\in\Im$

d. If $<P,V>\in\Im$ and $<Q, V'>\in\Im$, then $<P\rightarrow Q,tf_{\rightarrow}(V,V')>\in\Im$

e. If $<P,V>\in\Im$ and $<Q, V'>\in\Im$, then $<P\leftrightarrow Q,tf_{\leftrightarrow}(V,V')>\in\Im$

3. Nothing else is in $\Im$.

In alternative notation, $\Im$ *defined relative to Atomic*-$\Im$ is the set such that:

1. *Atomic*-$\Im \subseteq \Im$
2. Construction Steps:

   a. $\Im(\sim P)=tf_{\sim}(\Im(P))$

   b. $\Im(P\vee Q)=tf_{\vee}(\Im(P), \Im(Q))$

   c. $\Im(P\wedge Q)=tf_{\wedge}(\Im(P), \Im(Q))$

   d. $\Im(P\rightarrow Q)=tf_{\rightarrow}(\Im(P), \Im(Q))$

   e. $\Im(P\leftrightarrow Q)=tf_{\leftrightarrow}(\Im(P), \Im(Q))$
3. Nothing else is in $\Im$.

We shall let *SenIntrp* be the set of all sentential interpretations $\Im$ defined relative to any *basic set Atomic*-$\Im$, and let $\Im$ stand for interpretations in *SenIntrp.* We define a *sentential language L* as the pair <*SenSyn*, *SenIntrp* >.

Truth-Conditions

General Truth-Functions

The language of propositional logic possesses a number of interesting semantic properties as a result of its inductive definition of “truth in an interpretation.” These turn on the fact that the truth-value of a whole sentence can be calculated from the values of its immediate parts by the use of the basic truth-function for the connectives $tf_{\sim}$, $tf_{\vee}$, $tf_{\wedge}$, $tf_{\rightarrow}$, and $tf_{\leftrightarrow}$. This idea is stated more precisely in the following metatheorem.

Metatheorem.

a. $\Im(\sim P)$=T iff $tf_{\sim}(\Im(P))$ =T
b. $\Im(P\vee Q)$ =T iff $tf_{\vee}(\Im(P), \Im(Q))$ =T
c. $\Im(P\wedge Q)$ =T iff $tf_{\wedge}(\Im(P), \Im(Q))$ =T
d. $\Im(P\rightarrow Q)$ =T iff $tf_{\rightarrow}(\Im(P), \Im(Q))$ =T
e. $\Im(P\leftrightarrow Q)$ =T iff $tf_{\leftrightarrow}(\Im(P), \Im(Q)$ =T

The theorem is an immediate consequence of the previous definition of $\Im$ and the fact that $\Im$ is two-valued. Below we shall call the term on the right of the identity sign in the metatheorem the *truth-functional analysis* of the term on the left.

The calculation process, moreover, may be generalized. Not only is the truth-value of a sentence calculable from those of its immediate parts, it is calculable from the value of its *atomic* sentences.

This property is a bit more complicated to state. To do so we must first define the general notion of a truth-function as one defined in terms of the basic truth-functions. The idea is that if you can apply the functions $tf_{\sim}$, $tf_{\vee}$, $tf_{\wedge}$, $tf_{\rightarrow}$, or $tf_{\leftrightarrow}$ to truth-values to get an new truth-value, then you can keep applying these function to the results so as to get yet further values. A “general truth-function” is any result of repeated applications of the basic functions $tf_{\sim}$, $tf_{\vee}$, $tf_{\wedge}$, $tf_{\rightarrow}$, and $tf_{\leftrightarrow}$. For example, the function *h* defined below is a general truth-function:

$$h(\mathrm{w,x,y,z}) = tf_{\leftrightarrow}(tf_{\wedge}(w,tf_{\sim}(tf_{\vee}(x,y)),z)$$

Here h is a general truth-function because it is defined by repeated applications of $tf_{\sim}$, $tf_{\vee}$, $tf_{\wedge}$, $tf_{\rightarrow}$, and $tf_{\leftrightarrow}$. As we shall see shortly, we may use h to calculate the truth-value of the sentence $(p\wedge\sim(q\vee r)))\leftrightarrow s$ if we know the values of its atomic parts p, q, r, and s. The obvious way to define a "general truth-function" is by induction:

Definition

1. Any of the basic truth-functions $tf_{\sim}$, $tf_{\vee}$, $tf_{\wedge}$, $tf_{\rightarrow}$, and $tf_{\leftrightarrow}$ is a *truth-function*
2. If f, g_1, …,g_n are *truth-functions* of $n, j,\ldots,k$ places respectively, then the function h defined as follows is an $j+\ldots+k$-place *truth-function*:

   $$h(x_1,\ldots,x_{j+\ldots+k})=f(g_1(x_1,\ldots,x_j), \ldots,g_n(x_1,\ldots,x_k))$$
3. Nothing else is a *truth-function*.

We will now describe a general method for defining the general truth-function that may be used to evaluate a sentence, simple or complex. We find the function by progressive applications of the clauses of the definition of $\Im$, first to the sentence as a whole then to each smaller part until we reach its atomic sentences. Let us find the function appropriate to evaluating $\Im((p\wedge\sim(q\vee r)))\leftrightarrow s)$. We do so in the following steps, applying the clauses in the definition of $\Im$ annotated to the right.

$\Im((p\wedge\sim(q\vee r)))\leftrightarrow s)$	=	$tf_{\leftrightarrow}(\Im(p\wedge\sim(q\vee r)), \Im(s))$	clause e, $\leftrightarrow$
		$tf_{\leftrightarrow}(tf_{\wedge}(\Im(p),\Im(\sim(q\vee r))), \Im(s))$	clause b, $\wedge$
		$tf_{\leftrightarrow}(tf_{\wedge}(\Im(p),tf_{\sim}(\Im(q\vee r))), \Im(s))$	clause a, $\sim$
		$tf_{\leftrightarrow}(tf_{\wedge}(\Im(p),tf_{\sim}(tf_{\vee}(\Im(q),\Im(r)))), \Im(s))$	clause c, $\vee$

We now generalize this method to every sentence in the following metatheorem. Let us use the notation $P[Q_1,\ldots,Q_n]$ to refer to the sentence P that has as its atomic parts in left to right order the sentences $Q_1,\ldots,Q_n$.

Metatheorem. For any sentence $P[Q_1,\ldots,Q_n]$ there is some n-place truth-function f such that for any $\Im$,

$$\Im(P)=f(\Im(Q_1),\ldots,\Im(Q_n))$$

Proof. Using the previous metatheorem, we define a procedure that consists of writing down the page a series of terms that stand for a truth-values. The first line will $\Im(P[Q_1,\ldots,Q_n])$. The last line will be a term of the form $f(\Im(Q_1),\ldots,\Im(Q_n))$ for an n-place truth-function f. Moreover, the procedure is designed so that if t_n is the term on line n and t_{n+1} is the term on line n+1, then by the previous metatheorem it will be true that $t_n= t_{n+1}$. Hence, each term in the list will be identical to the next one in the list. It will then follow that the first term in the series is identical to the last, i.e. that $\Im(P[Q_1,\ldots,Q_n])= f(\Im(Q_1),\ldots,\Im(Q_n))$

Procedure for sentence $P[Q_1,\ldots,Q_n]$). Complete each step below as directed, starting with step 1.

1. Write down the term $\Im(P)$ as line 1. Make line 1 the current line. Go to the next rule.
2. In the current line if every whole sentence that occurs in the line is atomic, stop. If there are some occurrences of a whole sentence that are non-atomic go to the next step.
3. If the current line n contains an occurrence of a whole sentence Q that is non-atomic, write a new line n+1 which is like line n except that every such occurrence of Q is replaced by its truth-functional analysis (as defined in the last metatheorem). Make line n+1 the current line and go to step 2.

There will be only a finite number of applications of rule 3 because each truth-functional analysis is formulated in terms of the parts of the formula that it analyses. Since the construction sequence for any formula is finite, there can therefore be only a finite number of applications of step 3. Hence at some point step 2 must apply, and the procedure stops. Moreover, since step 2 applies, the last line contains some truth-function $f(\Im(Q_1),\ldots,\Im(Q_n))$ of the values of the atomic parts $Q_1,\ldots,Q_n$ of P. It is also clear from the earlier metatheorem that each term in the list is a truth-functional analysis of the one above it. Hence the first and the last are identical: $\Im(P[Q_1,\ldots,Q_n])= f(\Im(Q_1),\ldots,\Im(Q_n))$. End of Proof.

An important corollary of this theorem, which we will not pause to prove here, is that sentences with the same truth-value may be substituted for one another in longer sentences. That is, if two sentences have the same truth-value, one may be substituted for the other in a longer sentence without altering the truth-value of the longer sentence. To state this more precisely, let us use the notation $P[Q/R]$ to stand for the result of replacing some of the occurrences of Q in P by R.

Corollary. For any $\Im$, if $\Im(Q)=\Im(R)$, then $\Im(P[Q/R])=\Im(P)$

As we shall see in Part 3, it is this corollary underlies the validity of the logical rule called the substitution of material equivalents, which we remarked earlier failed for subjunctive conditionals:

$$\begin{array}{l} P \\ \underline{Q \leftrightarrow R} \\ \therefore\ P[Q/R] \end{array}$$

As we shall now see, the truth-functionality metatheorem also shows that the theory of truth defined using the inductive method meets Tarski's criterion for a correspondence theory.

Satisfaction of Tarski's Adequacy Condition

Given the inductive nature of the definition of an interpretation, it is possible to show that Tarski's condition (T) for a correspondence theory of truth is satisfied for every sentence. Let us illustrate how. Recall that the goal is to produce for any sentence *P* a metatheorem of the form:

(T) $\mathfrak{I}(P)$=T iff $TC_{\mathfrak{I}}(P)$

where $TC_{\mathfrak{I}}(P)$ states only facts about the interpretation of the atomic parts of *P* relative to $\mathfrak{I}$.

Metatheorem (Tarski's T Principle): for any *P* in *Sen.*

$$\mathfrak{I}(P)=\text{T iff } TC_{\mathfrak{I}}(P)$$

Proof. According to Tarski a statement of the truth-conditions of *P,* in symbols $TC_{\mathfrak{I}}(P)$, should be a metalinguistic statement that is equivalent to $\mathfrak{I}(P)$=T but is formulated by in terms that mention only the $\mathfrak{I}$-values of the atomic parts of *P.* Now consider the truth-function for *P,* such that (as shown by the previous metatheorem): $\mathfrak{I}(P)=f(\mathfrak{I}(Q_1),\ldots,\mathfrak{I}(Q_n))$. The proposition

$$f(\mathfrak{I}(Q_1),\ldots,\mathfrak{I}(Q_n))=\text{T}$$

meets Tarski's conditions for $TC_{\mathfrak{I}}(P)$ because it is formulated in terms that mention the $\mathfrak{I}$-values of the atomic parts of *P*. *Moreover,* it is equivalent to $\mathfrak{I}(P)$=T because we have proven:

(1) $\mathfrak{I}(P)=f(\mathfrak{I}(Q_1),\ldots,\mathfrak{I}(Q_n))$.

Since *f* is a function that assigns either T or F, (1) may be rewritten in an equivalent form as:

(2) $\mathfrak{I}(P)$=T iff $f(\mathfrak{I}(Q_1),\ldots,\mathfrak{I}(Q_n))$=T

But (2) states that $\mathfrak{I}(P)$=T is equivalent to $f(\mathfrak{I}(Q_1),\ldots,\mathfrak{I}(Q_n))$=T. Hence $f(\mathfrak{I}(Q_1),\ldots,\mathfrak{I}(Q_n))$=T is $TC_{\mathfrak{I}}(P)$.) End of proof.

The ability to prove instances of the (T) principle is of considerable theoretical interest because it shows that the notion of "truth in an interpretation" as defined inductively does in fact meet Tarski's minimal condition for being a correspondence notion of truth. It does so even though sentences mirror "the world" only in the abstract sense that they have truth-values.

The ability to prove instances of the (T) principle is also of practical value in allowing us to show arguments are valid. First, let us rephrase the results of an earlier metatheorem in simpler language that eliminates the difficult to read notation that refers to truth-functions.

Metatheorem. For any interpretaion $\Im$,

1. $\Im(\sim P)$=T iff $\Im(P)\neq$T
2. $\Im(P\wedge Q)$ =T iff, $\Im(P)$ =T and $\Im(Q))$ =T
3. $\Im(P\vee Q)$ =T iff, $\Im(P)$ =T or $\Im(Q))$ =T
4. $\Im(P\rightarrow Q)$ =T iff, $\Im(P)\neq$T or $\Im(Q))$ =T
5. $\Im(P\leftrightarrow Q)$ =T iff, either $\Im(P)$=T and $\Im(Q)$=T , or $\Im(P)\neq$T and $\Im(Q)\neq$T

   [or equivalently, $\Im(P)=\Im(Q)$]

Proof. Note first the following facts that hold given the definitions of the truth-functions:

1. $tf_{\sim}(x)$ =T iff x=F
2. $tf_{\wedge}(x,y)$ =T iff, x=T and y=T
3. $tf_{\vee}(x,y)$ =T iff, x=T or y=T
4. $tf_{\rightarrow}(x,y)$ =T iff, either x=F or y=T
5. $tf_{\leftrightarrow}(x,y)$ =T iff $x=y$

The metatheorem above then follows from the previous metatheorem by substituting into its biconditionals the the equivalences above.

This latest metatheory shows us how to explain when a sentence is true in terms of the truth-values of its parts. Let us turn now to an even easier way to calculate how the truth-value of a very complex sentence can be expressed in terms of the truth-values of its atomic sentences, allows for us to figure out very easily for any sentence *P* is truth-conditions $TC_{\Im}(P)$.

Calculating Sentence Values by Truth-Tables

There is a standard procedure for calculating the truth-value of a whole sentence from those of its atomic parts, called the *truth-table method*. It is easy to describe and use. First construct the construction sequence for a sentence *P*. If *P* contains *n* atomic sentences $Q_1,\ldots,Q_n$, there are 2^n possible interpretations $\Im_1,\ldots,\Im_{2^n}$ that assign truth-values T or F to $Q_1,\ldots,Q_n$. Parallel to the steps in the construction sequence for P, start 2^n new construction sequences, one for each $\Im_1,\ldots,\Im_{2^n}$, as follows. Next to line of each atomic formula Q_i in the construction sequence of P, write under in the column for $\Im_j$ the truth-value $\Im_j(Q_i)$ that $\Im_j$ assigns to Q_i. Proceed to complete the construction sequence for $\Im_j$ by using the construction rules for the definition of $\Im_j$, writing next to a part R of P the value $\Im(R)$. Below we highlight the fact that $\Im_j$ is a set of pairs by using the ordered pair notation $<P,V>\in\Im$ instead of $\Im(P)=V$.

Once the series of interpretation constructions parallel to *P's* grammatical derivation is produced, it is easy to see the information they contain. IN particular the last element in each sequence states the assignment in that interpretation of the truth-value of the sentence P as a whole.

As the examples below show, however, actually writing out the series of parallel sequence takes up lots of paper. It is customary to summarize the process in what is called *the truth-table for P.* This is a two-dimensional table constructed as follows:

- Write the sentence P to be evaluated across the top of a page.
- Under it draw and label a series of rows, one for each interpretation $\Im_i$ of the atomic sentence in P. If P contains n atomic sentences, there will be 2^n rows.
- Draw a series of columns, one under each atomic sentence and under each occurrence of a connective in P.
- In the row for interpretation $\Im_i$ enter in the column under each atomic sentence p_j the truth-values that $\Im_i$ assigns to p_j and under each occurrence of a connective the truth-values that $\Im_i$ assigns to part of P formed by that connective. Progress from the smaller to larger parts of P.

With very little practice it is possible to construct such truth-tables directly without first producing the construction sequences for the sentence and its interpretations.

Examples of Truth-functional Computation and Truth-Tables

For each of the following sentences, which were earlier provided with construction sequences showing their membership in *Sen*, we provide a parallel series of construction sequences, one for each interpretation. We then summarize this information in a traditional truth-table for the sentence.

1. $((\sim p_4 \vee p_2) \wedge p_4)$
2. $\sim(p_3 \vee \sim p_3)$
3. $\sim(p_1 \vee \sim p_3)$
4. $((p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2)))$

Example 1. $((\sim p_4 \vee p_2) \wedge p_4)$

There are two atomic sentences and therefore 2^2=4 possible interpretations.

		$\mathfrak{I}_1$	$\mathfrak{I}_2$	$\mathfrak{I}_3$	$\mathfrak{I}_4$
1. p_2		$< p_2,T>$	$< p_2,T>$	$< p_2,F>$	$< p_2,F>$
2. p_4		$< p_4,T>$	$< p_4,F>$	$< p_4,T>$	$< p_4,F>$
3. $\sim p_4$		$< \sim p_4 f>$	$<\sim p_4,T>$	$<\sim p_4,F>$	$< \sim p_4,T>$
4. $(\sim p_4 \vee p_2)$		$< \sim p_4 \vee p_2,T>$	$< \sim p_4 \vee p_2,T>$	$< \sim p_4 \vee p_2,F>$	$< \sim p_4 \vee p_2,T>$
5. $((\sim p_4 \vee p_2) \wedge p_4)$		$<((\sim p_4 \vee p_2) \wedge p_4),T>$	$<((\sim p_4 \vee p_2) \wedge p_4),F>$	$<((\sim p_4 \vee p_2) \wedge p_4),T>$	$<((\sim p_4 \vee p_2) \wedge p_4),T>$

The truth-table for $(\sim p_4 \vee p_2) \wedge p_4$:

	p_2	p_4	$((\sim$	p_4	$\vee$	$p_2)$	$\wedge$	$p_4)$
$\mathfrak{I}_1$	T	T	F	T	T	T	T	T
$\mathfrak{I}_2$	T	F	T	F	T	T	F	F
$\mathfrak{I}_3$	F	T	F	T	F	F	F	T
$\mathfrak{I}_4$	F	F	T	F	F	F	F	F

From this table we can read off the truth-conditions of $(\sim p_4 \vee p_2) \wedge p_4$:

$$TC_{\mathfrak{I}}((\sim p_4 \vee p_2) \wedge p_4) = T \text{ iff } (\ \mathfrak{I}(p_4)=T \text{ and } \mathfrak{I}(p_2)=T\ )$$

Example 2. $\sim(p_3 \vee \sim p_3)$

There is one atomic sentence, and therefore $2^1=2$ possible interpretations.

		$\Im_1$	$\Im_2$
1. p_3		< p_3,T>	< p_3,F>
2. $\sim p_3$		< $\sim p_3$,F>	< $\sim p_3$,T>
3. $(p_3 \vee \sim p_3)$		< $(p_3 \vee \sim p_3)$,T	<$(p_3 \vee \sim p_3)$,T>
4. $\sim(p_3 \vee \sim p_3)$		< $\sim(p_3 \vee \sim p_3)$,F>	< $\sim(p_3 \vee \sim p_3)$,F>

The truth-table for $\sim(p_3 \vee \sim p_3)$:

	p_3	$\sim($	p_3	$\vee$	$\sim$	$p_3)$
$\Im_1$	T	F	T	T	F	T
$\Im_2$	F	F	F	T	T	F

Hence we can read off the truth-coditions for $\sim(p_3 \vee \sim p_3)$:

$$TC_{\Im}\,(\sim(p_3 \vee \sim p_3))=T \text{ iff } (\Im(p_3)=T \text{ and } \Im(p_3)=F).$$

That is, $TC_{\Im}\,(\sim(p_3 \vee \sim p_3))=T$ is never true.

Example 3. $\sim(p_1 \vee \sim p_3)$

There is one atomic sentence, and therefore $2^1=2$ possible interpretations.

	$\Im_1$	$\Im_2$	$\Im_3$	$\Im_4$
1. p_1	< p_1,T>	< p_1,T>	< p_1,F>	< p_1,F>
2. p_3	< p_3,T>	< p_3,F>	< p_3,T>	< p_3,F>
3. $\sim p_3$	< $\sim p_3$,F>	< $\sim p_3$,T>	< $\sim p_3$,F>	< $\sim p_3$,T>
4. $(p_1 \wedge \sim p_3)$	< $(p_1 \wedge \sim p_3)$,F>	< $(p_1 \wedge \sim p_3)$,T>	< $(p_1 \wedge \sim p_3)$,F>	<$(p_1 \vee \sim p_3)$,F>
5. $\sim(p_1 \wedge \sim p_3)$	< $\sim(p_1 \wedge \sim p_3)$,T>	<$\sim(p_1 \wedge \sim p_3)$,F>	<$\sim(p_1 \wedge \sim p_3)$,T>	< $\sim(p_1 \vee \sim p_3)$,T>

The truth-table for $\sim(p_1 \vee \sim p_3)$:

	p_1	p_3	$\sim($	p_1	$\wedge$	$\sim$	$p_3)$
$\Im_1$	T	T	T	T	F	F	T
$\Im_2$	T	F	F	T	T	T	F
$\Im_3$	F	T	T	F	F	F	T
$\Im_4$	F	F	T	F	F	T	F

Hence we can read off the truth-coditions for $\sim(p_1 \vee \sim p_3)$:

$$\begin{aligned} TC_{\Im}\,(\sim(p_3 \vee \sim p_3))=F \quad \text{iff} \quad & [(\Im(p_1)=T \text{ and } \Im(p_3)=T) \text{ or} \\ & (\Im(p_1)=F \text{ and } \Im(p_3)=T) \text{ or} \\ & (\Im(p_1)=F \text{ and } \Im(p_3)=F)\ ] \end{aligned}$$

Example 4. $((p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2)))$

There are two atomic sentences, and therefore $2^2=4$ possible interpretations.

		$\mathfrak{I}_1$	$\mathfrak{I}_2$	$\mathfrak{I}_3$	$\mathfrak{I}_4$
1. p_1		$< p_1,T>$	$< p_1,T>$	$< p_1,F>$	$< p_1,F>$
2. p_2		$< p_2,T>$	$< p_2,F>$	$< p_2,T>$	$< p_2,F>$
3. $(p_1 \leftrightarrow p_2)$		$< p_1 \leftrightarrow p_2,T>$	$< p_1 \leftrightarrow p_2,F>$	$< p_1 \leftrightarrow p_2,F>$	$< p_1 \leftrightarrow p_2,T>$
4. $(p_1 \wedge p_2)$		$< p_1 \wedge p_2,T>$	$< p_1 \wedge p_2,F>$	$< p_1 \wedge p_2,F>$	$< p_1 \wedge p_2,F>$
5. $\sim p_1$		$<\sim p_1,F>$	$<\sim p_1,F>$	$<\sim p_1,T>$	$<\sim p_1,T>$
6. $\sim p_2$		$<\sim p_2,F>$	$<\sim p_2,T>$	$<\sim p_2,F>$	$<\sim p_2,T>$
7. $(\sim p_1 \wedge \sim p_2)$		$<\sim p_1 \wedge \sim p_2,F>$	$<\sim p_1 \wedge \sim p_2,F>$	$<\sim p_1 \wedge \sim p_2,F>$	$<\sim p_1 \wedge \sim p_2,T>$
8. $((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2))$		$<((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2)),T>$	$<((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2)),F>$	$<((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2)),F>$	$<((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2)),T>$
9. $((p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2)))$		$<((p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2))),T>$	$<((p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2))),T>$	$<((p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2))),T>$	$<((p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2))),T>$

The truth-table for $(p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2))$:

	p_1	p_2	$((p_1$	$\leftrightarrow$	$p_2)$	$\leftrightarrow$	$((p_1$	$\wedge$	$p_2)$	$\vee$	$(\sim$	p_1	$\wedge$	$\sim$	$p_2)))$
$\mathfrak{I}_1$	T	T	T	T	T	T	T	T	T	T	F	T	F	F	T
$\mathfrak{I}_2$	T	F	T	F	F	T	T	F	F	F	F	T	F	T	F
$\mathfrak{I}_3$	F	T	F	F	T	T	F	F	T	F	T	F	F	F	T
$\mathfrak{I}_4$	F	F	F	T	F	T	F	F	F	T	T	F	F	T	F

Hence we can read off the truth-coditions for $(p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2))$:

$TC_{\mathfrak{I}}\ ((p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2)))=T$ iff $((\ \mathfrak{I}(p_1)=T$ and $\mathfrak{I}(p_2)=T\ )$ or
$(\ \mathfrak{I}(p_1)=T$ and $\mathfrak{I}(p_2)=F\ )$ or
$(\ \mathfrak{I}(p_1)=F$ and $\mathfrak{I}(p_2)=T\ )$ or
$(\ \mathfrak{I}(p_1)=F$ and $\mathfrak{I}(p_2)=F\ )\ )$

That is, $TC_{\mathfrak{I}}\ ((p_1 \leftrightarrow p_2) \leftrightarrow ((p_1 \wedge p_2) \vee (\sim p_1 \wedge \sim p_2)))=T$ holds no matter what.

Notice that example 1 is true in some interpretations and false in others. Such sentences are said to be *contingent.* Example 2 is false in every interpretation. Such sentences are said to be *contradictory* or *inconsistent.* Example 4 is true in every interpretation. Sentences of propositional logic that are always true are called *tautologies.*

Exercise. Analyze the following sentences P like the previous example:

(a) for all possible interpretations of the sentence's atomic parts, provide a construction sequence that is parallel to the sentence's grammatical derivation,

(b) summarize the information from the construction sequences in a traditional truth-table for the sentence,

(c) summarize the truth-conditions $TC_{\mathfrak{I}}(P)$ for P.

1. $\sim(p_1\leftrightarrow\sim p_1)$ [two possible interpretations]
2. $\sim\sim(p_1\vee\sim p_1)$ [two possible interpretations]
3. $\sim(p_1\leftrightarrow\sim p_2)$ [four possible interpretations]
4. $(((p_1\rightarrow p_2)\wedge\sim p_2))\rightarrow\sim p_1)$ [four possible interpretations]
5. $(((p_1\rightarrow p_2)\wedge p_2))\rightarrow p_1)$ [four possible interpretations]
6. $((p_1\rightarrow p_2)\leftrightarrow(\sim p_2\rightarrow\sim p_1))$ [four possible interpretations]
7. $((p_1\leftrightarrow p_2)\leftrightarrow((p_1\rightarrow p_2)\wedge(p_2\rightarrow p_1)))$ [four possible interpretations]

Exercise. For the sentences below construct their truth-table only, without first producing the construction sequences for the sentence itself and its interpretations.

1. $(p_1\rightarrow(p_1\vee (p_2\wedge\sim p_2)))$ [four possible interpretations]
2. $((p_1\leftrightarrow p_2)\leftrightarrow ((p_1\rightarrow p_2)\wedge(p_2\rightarrow p_1)))$ [four possible interpretations]
3. $(\sim(p_1\wedge p_2)\leftrightarrow (\sim p_1\vee\sim p_2))$ [four possible interpretations]
4. $((p_1\wedge(p_2\vee p_2))\rightarrow ((p_1\wedge p_2)\vee (p_1\wedge p_3)))$ [eight possible interpretations]

We complete this introduction to the semantics of propositional logic by defining several important logical ideas, which we shall investigate more fully in Part.

The Definition of Logical Concepts

We complete the semantic theory by defining the key concepts of logic, which will be the main topic of Part 3: valid argument, and consistency.

To represent a valid argument we will continue to use the notation

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

which is read "the argument from the set of premises $P_1,\ldots,P_n$ to conclusion Q is *valid*."

Definitions

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

P is a *tautology* (in symbols $\models_L P$) iff $\forall\Im\ (\ \Im(P)=T)$

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

In this notation, we group the sentences $P_1,\ldots,P_n$ into the set $\{P_1,\ldots,P_n\}$ to emphasis the fact that the order of the sentences does not matter when the issue is whether they are the premises of a logically valid argument or as a group are jointly consistent. In practice, however, we often omit the {…} notation and write $\{P_1,\ldots,P_n\} \models_L Q$ simply as $P_1,\ldots,P_n$,… $\models_L Q$, which is easier to read. However, this notation should be understood as imposing no definite order on the sentences $P_1,\ldots,P_n$.

Summary

The material in this lecture is of great theoretical importance in logic. We saw how to define a correspondence theory of truth for a sentential grammar with simple and complex propositions that stand for truth-values. This is a theoretical challenge for two reasons.

First of all it is not clear how to make sense of the notion of truth as "correspondence with the world" in cases in which what is supposed to corresponds to the world are the simple and complex sentence of the propositional logic. These stand for truth-values, but it is odd to think of truth-values as entities that make up "the world". We saw how Alfred Tarski suggests a solution by proposing his T principle as a criterion for any theory claiming to be a genuine "correspondence theory of truth." It is a fair abstraction of "correspondence" because it fits the clear cases like the syllogistic, in which the parts of sentences genuinely do refer to things in the world and true sentences genuinely do impose some condition on the structure of these entities. But it also fits the propositional logic. In both, the truth-value of the whole is determined by the values of the expression's atomic parts. Thus he proposes that a genuine correspondence theory is marked by the fact that every sentence is such that its truth is a function of the "referents" of its parts, where "referent" is understood in an abstract way, one broad enough to include truth-values. It is this idea that is captured in his requirement that correspondence theory of truth must entail an instance of the T schema for each sentence.

(T) $\mathfrak{I}(P)$=T iff $TC_{\mathfrak{I}}(P)$

where $TC_{\mathfrak{I}}(P)$ spells out the truth-conditions of P in terms of the "referents" of its atomic parts.

Tarksi also solves the difficulty of how to define truth without recourse to a traditional definition in terms of necessary and sufficient conditions. Truth is one of those ideas that it hard to define in terms of necessary and sufficient conditions. Tarski's solution is to employ the method of inductive definition invented by logicians to deal with difficult ideas of this sort. He shows how to state an inductive definition for each "interpretation" of the syntax. He does so by understanding an interpretation to be a set of pairs. His task then is to define this set of pairs inductively. As in any inductive definition, he first defines a set of "basic" pairs. These are pairs that assign a unique truth-value to each atomic sentence. He then defines a series of rules designed to add new pairs to the set, one rule for each of the sentential connectives. Each rule tells, for a given connective, how to add a sentence truth-value pair given the sentence truth-value pairs of its immediate parts. In this way, every sentence is paired with one and only one truth-value in a given interpretation.

The finale of the discussion is the proof that Tarski's inductive definition of interpretation actually meets his T criterion for a correspondence theory of truth. Thus, the difficult idea of truth correspondence is shown to be well-defined for propositional logic and in a way that insures it qualifies in an abstract sense as a correspondence theory of truth.

LECTURE 3. FIRST-ORDER LOGIC

Expressive Power

Simple and Complex Sentences in a Single Syntax

In Part 2 these lectures the topic has been the "logic of propositions", by which we mean the grammar and semantics of sentences. In the syllogistic we investigated the syntax and semantics of subject-predicate sentences. In the propositional logic we did the same for complex sentences formed by the connectives from unanalyzed atomic sentences. In this lecture we investigate how to combine both in one language. Syntactically, atomic sentences will themselves have grammatical parts, made of up parts of speech similar to the nouns and verbs of traditional grammar. Putting together these atomic sentences by means of the connectives of the propositional logic, we will then be able to form a myriad of complex forms, all those that are possible by repeated applications of the formation rules for the connectives. Semantically, we will be able to combine the versions of the correspondence theory of truth developed for the syllogistic and the propositional logic. The notion of correspondence appropriate to atomic sentences will be quite intuitive, as it is for **A**, **E**, **I** and **O** propositions in the syllogistic, because the parts of speech into which atomic sentences divide do "refer" to entities "in the world" in an intuitively plausible way in terms of which it is possible in state truth-conditions for the sentence as a whole. We will be able to extend this correspondence theory to molecular sentences as well by making use of the correspondence theory in the sense proposed by Tarski , which is suitable for grammars with complex sentences. For every sentence *P,* simple and complex, the theory will entail a metalinguistic principle that will spell out when the sentence is true in terms of its truth-conditions:

$$\Im(P)=\mathrm{T} \text{ iff } \mathrm{TC}_{\Im}(P).$$

Here $\mathrm{TC}_{\Im}(P)$ will state the conditions that must obtain among the referring parts of *P* in order for *P* to be true in $\Im$.

In the language we will be developing in this lecture, however, the conditions stated in $\mathrm{TC}_{\Im}(P)$ will be less abstract and more intuitive than those in propositional logic. Recall that in the propositional logic the basic parts of *P* were atomic sentence, which had no internal grammatical parts and could only be said to have a "referent" in the sense that they had a truth-values. Truth-values, however, can be called "entities in the world" only in a very abstract sense. In the richer syntax we are about to explore, on the other hand, the atomic parts of *P* are word much more like the nouns and verbs of traditional grammar. They will "stand for" sets and the elements of sets, which are entities that it is much more intuitively plausible to think of as constituting "the world". Thus the truth-conditions of every sentence P, simple or complex, will be formulated in terms of conditions on the sets and set members represented by the simple words that go into the formation of *P.*

We will not however simply combine the syllogistic with propositional logic. We could for example simply say that the set of atomic sentences for the syntax was the set of syllogistic propositions. Let us see what such a syntax would be like and what its limitations would be.

The Limitations of the Syllogistic and Propositional Logic

Logicians in the Middle Ages in fact did work with a combination of syllogistic and hypothetical propositions. Their understanding of what they were doing is somewhat different from that of modern logic because they did not think of themselves as inventing a new or restricted syntax with formal rules of grammar as we do now. Rather they thought of themselves as describing carefully a subset of the grammatical sentences of Latin. In their view there were large parts of natural language about which they had little to say, but which were just as real as the propositions they did study. The concentrated on simple forms of **A**, **E**, **I** and **O** propositions, and short hypothetical propositions formed with them by conjunctions and disjunctions.

They did discuss several more complex forms of the basic **A**, **E**, **I** and **O** proposition types. For example, as sketched in the supplementary section of Lecture 7, they studied predicate negations. Following the lead of Aristotle in the *Prior Analytics,* they also investigated the logic of propositions in which the verb or sentence as a whole was modified by the adverbs *necessarily* and *possibly*, as in:

Every man is necessarily rational
Possibly some man is just.

Known as *modal logic*, this field today is an important part of advanced work in logic. They studied “exceptive” quantifiers like *only* and *except,* as in the sentences:

> *Only birds fly.*
> *All birds except ostriches fly.*

To some extent they also studied the logic of **A**, **E**, **I** and **O** propositions in which the subject or predicate term is a grammatically complex noun or verb phrase, or is formed by a conjunction or disjunctions of nouns and verb, or by relative clauses, as in:

> *Every cat and dog is an animal.*
> *No dog is either a fish or a bird.*
> *Every man who laughs is happy.*

What mediaeval logicians have to say about the logic of such propositions is interesting and, in some instances, helpful in modern logic. We will not pursue it further because of serious limitations build into grammars based on the four syllogistic forms. Even their refined versions are inadequate for the purpose for which modern logic was invented: expressing the argument forms used in mathematics and the mathematical sciences. There are a number of ways in which the expressive power of syllogistic syntax is limited. Here I will mention three. It cannot adequately express propositions about the empty set or relations, nor does it have the power to express multiple or embedded quantifiers.

We have seen that traditional logic builds into the truth-conditions of **A**-propositions the assumption that the subject term stands for a non-empty set. In mathematics, however, it is often important to say that a set, or region of a set, is empty.

We have also already seen how it is difficult to express relational properties using only nouns and verbs that stand for simple sets. Despite ingenious tries, traditional logicians never solved the problem how to talk about relations using just **A**, **E**, **I** and **O** propositions.

More importantly perhaps is the syllogistics inability to express multiple quantifiers or to nest quantifiers inside one another. Consider, for example, the task that Frege set for himself. He invented a syntax with several general goals in mind. First he wanted to be able to express the axioms of set theory, which we formulated in Part 1 as follows:

Abstraction. There is some *A* such that for every *x,* x is in *A* if and only if *P*[*x*]
Extensionality. For every *A* and *B, A=B* if and only if for every *x, x* is in *A* if and only if *x* is in *B.*

He also wanted to prove as theorems the five basic postulates of the natural numbers as studied by Dedekind and Peano:

1. 0 is a natural number.
2. For every natural number *n* and every entity *x,* if (*x* stands in the successor relation to *n*) then (*x* is a natural number).
3. 0 stands in the successor relation to no natural number.
4. For every natural number *n* and *m,* if (*x* stands in the successor relation to *n, y* stands in the successor relation to *m*, and *n=m)* then *x=y*.
5. If (0 is in *A*) and if [(for every natural number *n* and for every entity *m* such that, if [(*m* stands in the successor relation to *n,* and *n* is in *A*) then *m* is in *A*)] then [every natural number is in *A*].

Notice, for example, that in the Principle of Abstraction there is a universal quantifier nested inside an existential quantifier, and that Peano's second postulate begins with two universal quantifiers. These propositions cannot be formulated in syllogistic syntax in a way that allows the deduction of their simple mathematical consequences.

New Notation: Constants, Predicates and the Quantifiers

Frege invented a new syntax. It incorporates the features of the syllogistic and propositional logic, but it also has a great deal of expressive power these simpler languages lack. You have in fact been introduced to this language in the lecture on set theory. His syntax contains three key innovations. The first is a new part of speech used to stand for the individuals that are members of sets. In traditional grammar this role is filled by proper nouns, demonstratives like *this* and *that,* and singular noun phrases that begin with *the* like *the tallest man in New York.* Expressions that stand for individuals are called *constants.* For these Frege used lower case Greek letters, but we shall follow the modern practice of using the lower case letters: *a,b,c,d,e,f,g,h.*

Secondly, he introduced special symbols, called predicates, which stand for sets and relations. For these he used upper case letters, as we continue to do today: *F,G,H,…* Predicates that stand for sets are followed by a single symbol naming an individual, and are called *one-place predicates.* For example, *Fc* says that *c* is in *F.* Predicates that name a two-place relation are followed by two symbols for individuals. For example *Gcb* say *c stands in the relation G to b.* Predicates that stand for a three-place relation are followed by symbols for three individuals. For example, *Habc* say that *the individuals a, b, and c stand (in that order) in the H relation to one another.* Likewise, a predicate followed by *n* names for individuals is called an *n-place predicate* and stands for an *n*-place relation.

Thirdly, he also introduced symbolization for the universal and existential quantifiers, and for their accompanying variables. Though Frege used lowercase gothic letters for variables, we shall follow the modern practice of using lowercase letters from the end of the alphabet: *u,v,w,x,y,* and *z.* For the universal quantifier *for all x* he uses:

$\vdash\!\overset{x}{\smile}\!\!-\!\!-$

He represents an existential quantifier by means of the universal because *for some x* means the same as *it is not the case that for all x it is not the case that.*

In later logic the notation was simplified, along with its intended reading. In the notation of Russell and Whitehead (1910) the universal quantifier *for all* is (x), in Polish logic it is Πx (the letter Π come from "product" in arithmetic, and *panta,* which means *everything* in Greek), and in modern notation is $\forall x$. The existential quantifier *for some x* is ($\exists x$) in the notation of Russell and Whitehead, Σx in Polish notation (analogous to arithmetical "sum"), and is $\exists x$ in modern notation.

	Frege	Polish Notation	Russell	Modern
for all x , Fx	$\vdash\!\overset{x}{\smile}\!\!-\!\!-\ Fx$	ΠxFx	$(x)Fx$	$\forall xFx$
for some x, Fx	$\vdash\!\!\top\!\!-\!\overset{x}{\smile}\!\!-\!\!\top\!\!-\ Fx$	ΣxFx	$(\exists x)Fx$	$\exists xFx$

for all x, if Fx then Gx	├─$\overset{x}{\smile}$─┬─ *Fx* └─ *Gx*	Π*x*C*FxGx*	(*x*)(*Fx*⊃*Gx*)	∀*x*(*Fx*→*Gx*)
for some, x Fx and Gx	├┬─$\overset{x}{\smile}$─*Fx*┬ *Gx*	Σ*x*K*FxGx*	(∃*x*)(*Fx*•*Gx*)	∃*x*(*Fx*∧*Gx*)

The new syntax is called *first-order logic* because it allows quantification over individuals, which are the lowest "order" in the hierarch of sets that consists of the series: individuals, set of individuals, set of sets of individuals, etc. With this introduction we are now ready to state the formation rules for the new grammar precisely.

Syntax for First-order Logic

Definition of Well-Formed Formula

In preparation for stating the precise definitions of the grammar, let us adopt he following conventions.

Singular Terms. Constants, which are the equivalents in formal grammar of proper names because they stand for individuals, will be represented by the letter *c,* with and without subscripts, and by other lower case letters from *a* to *t*. The set of all constants is *Cns*. It may or may not be infinite depending o the syntax we happen to be using. In addition to constants there are also variables, represented by lower case letters *w* through *z,* with and without subscripts, that also stand for individuals. They function like pronouns because when they are used with a quantifier as their antecedent their referent is determined by that antecedent. The set of variables is *Vbls*. For technical reasons that will not concern us here it is always assumed to be infinitely large. The set of constants and variables is combined in the set *Trms* of *(singular) terms,* i.e. $Trms=Cns \cup Vbls$.

Predicates. Predicates are represented by P^n_m, and by upper case letters $F,\ldots,M$, with and without subscripts and superscripts. A super-script indicates t predicate's *degree*, i.e. the number of singular terms that follows it when it forms an atomic formula. A predicate's degree also determines what type of set or relation it stands for. For example, the predicate P^1_m, with superscript 1, is the *m*-th one-place predicate. It forms an atomic formula when it is followed by a single constant or variable, and it stands for a sets. The predicate P^n_m, with superscript *n*, is the *m*-th *n*-place predicate. It forms an atomic formula when it is followed by *n* constants or variables, and it stands for an *n*-place relation. In first-order logic the first two-place predicate is usually stands for the identity relation among elements in the domain. For this purpose we shall use the symbol **=** (in bold type). Since it is a two-place predicate, strictly speaking, it should form an atomic formula by writing two singular terms to its right, e.g. **=***ab.* We will rewrite this, however, in the usual order of English: *a***=***b.* The formula *a***=***b* will be true in an interpretation if and only if in that interpretation the two terms *a* and *b* stand for the same individual.

Formulas. The definition of *formula* is inductive. As in the inductive definition of *sentence* for the propositional logic, the definition presupposes a basic set of formulas, the so-called *atomic formulas,* and a set of construction rules. The set of formulas is then defined as all those that can be constructed from the basic elements by the rules. An atomic formula is defined as any sequence of symbols that consists of an *n*-place predicate followed by *n* singular terms (constants or variables). The construction rules, or as they are called in grammar the *formation rules*, include all those of the propositional logic ($fr_{\sim}$, $fr_{\wedge}$, $fr_{\vee}$, $fr_{\rightarrow}$, $fr_{\leftrightarrow}$), as well as two new rules for quantified formulas: $fr_{\forall}$ and $fr_{\exists}$. The former takes a formula *P* and a variable *x* and forms a new formula $\forall xP$. The latter takes a formula *P* and a variable *x* and forms a new formula $\exists xP$. The set of formulas is then the closure the set of atomic formulas under these rules.

Definition. A *first-order syntax FOSyn* is a structure *<Cns, Vbls, Prds, AFor, FR, For*> such that

1. *Cns* is a subset of $\{c_1,\ldots,c_n,\ldots\}$
2. *Vbsl* is $\{v_1,\ldots,v_n,\ldots\}$. Let $Trms = Cns \cup Vbls$
3. *Prds* is a subset of $\{P^1_1,\ldots,P^1_m;\ldots; P^n_1,\ldots, P^n_m,\ldots;\ldots\}$ such that P^2_1 is **=**.
   (here P^n_m is the *m*-th *n*-place predicate and **=** is the 1st 2-place predicate).
4. *AFor,* called the set of *atomic formulas,* is $\{P^n_m t_1,\ldots,t_n \mid P^n_m \in Prds \;\&\; t_1 \in Trms \;\&\; \ldots \&\; t_n \in Trms\}$
5. *FR,* called the set of *formation rules,* is the set of functions $\{fr_{\sim}, fr_{\wedge}, fr_{\vee}, fr_{\rightarrow}, fr_{\leftrightarrow}, fr_{\forall}, fr_{\exists}\}$ defined as follows:
   a. $fr_{\sim}(x) = \sim x$
   b. $fr_{\wedge}(x,y) = (x \wedge y)$
   c. $fr_{\vee}(x,y) = (x \vee y)$
   d. $fr_{\rightarrow}(x,y) = (x \rightarrow y)$
   e. $fr_{\leftrightarrow}(x,y) = (x \leftrightarrow y)$
   f. $fr_{\forall}(x,y) = \forall xy$
   g. $fr_{\exists}(x,y) = \exists xy$

6. *For* is defined inductively as follows:
   a. *AFor* is a subset of *For;*
   b. if the elements *P,* and *Q* are in *For* and *v* is in *Vbls,* then $fr_{\sim}(P)$, $fr_{\wedge}(P,Q)$, $fr_{\vee}(P,Q)$, $fr_{\rightarrow}(P,Q)$, $fr_{\leftrightarrow}(P,Q)$, $fr_{\forall}(v,P)$, $fr_{\exists}(v,P)$are in *For;*
   c. nothing else is in *For.*

We shall say that a variable *x* is *free* in a formula *P* is it is not part of some formula $\forall xQ$ or $\exists xQ$ in *P.* If the formula is not free, it is *bound.* We reserve the term *sentence* for formulas that have no free variables.

As in propositional logic, since the set of formulas is constructed by an inductive definition, there is a construction sequence, a so-called *grammatical derivation*, showing that it is in the set. The proof of the following metatheorem provides some examples of grammatical derivations in first-order syntax.

Grammatical Metatheorem. The following are in *Sen:*

1. $\exists x((\sim Fx \vee Gxb) \wedge \sim Fx)$
2. $\exists z \forall x(\ Gzx \rightarrow \exists yHzy)$
3. $\forall x((Fxb \vee \exists yGyx) \rightarrow \sim Fxc)$

The theorem is prove by producing a grammatical derivation (construction sequence) for each:

1. Fx	atomic
2. Gxb	atomic
3. $\sim Fx$	2, $fr_{\sim}$
4. $(\sim Fx \vee Gxb)$	2 & 3, $fr_{\vee}$
5. $((\sim Fx \vee Gxb) \wedge \sim Fx)$	4 & 2, $fr_{\wedge}$
6. $\exists x((\sim Fx \vee Gxb) \wedge \sim Fx)$	5, $fr_{\exists}$

1. Hzy	atomic
2. $\exists yHzy$	2, $fr_{\exists}$
3. $z{=}x$	atomic
4. $(z{=}x \rightarrow \exists yHzy)$	2 &3, $fr_{\rightarrow}$
5. $\forall x(\ z{=}x \rightarrow \exists yHzy)$	4, $fr_{\forall}$
6. $\exists z \forall x(\ z{=}x \rightarrow \exists yHzy)$	5, $fr_{\exists}$

1. Fxc	atomic
2. Gyx	atomic
3. $\exists yGyx$	2, $fr_{\exists}$
4. $\sim Fxc$	2, $fr_{\sim}$
5. Fxb	atomic
6. $(Fxb \vee \exists yGyx)$	3 & 5, $fr_{\vee}$
7. $((Fxb \vee \exists yGxb) \rightarrow \sim Fxc)$	6 & 4, $fr_{\rightarrow}$
8. $\forall x((Fxb \vee \exists yGyx) \rightarrow \sim Fxc)$	7, $fr_{\forall}$

Exercises. Construct a grammatical derivation for each of the following showing that they are elements of *For:*

$\forall x\forall y\forall z((Hxy \wedge Hyz)\rightarrow Hxz)$

$\forall x\forall y((x{=}y \wedge Fx)\rightarrow Fy)$

$\sim\exists yFy\rightarrow\forall x(\sim Hx\vee\sim Fx)$

Informal Semantics

Quantifiers and Models

Perhaps the best way to develop a sense of the meaning of the quantifiers is to construct “models” for an interpretation $\mathfrak{I}$ in which quantified formulas are true or false. We shall use Venn diagrams for this purpose. The universe of entities that exist relative to $\mathfrak{I}$, called the model’s *domain*, is represented by the surrounding rectangle. A circle represents a subset of the domain. If the set is labeled by a one-place predicate then that set is the predicates extension in $\mathfrak{I}$. A dot (rather than an x) is used to represent an entity in the domain, and if it labeled by a constant, it is the referent of that constant in $\mathfrak{I}$. To indicate that there is an entity in one of several regions without declaring which a short bold line will be drawn across the line or lines separating these regions. The fact that the domain is non-empty will sometimes be represented by such “on the line” entities. Note, however, that a subset of the domain, even those named by a predicate in $\mathfrak{I}$, may be empty and totally shaded.

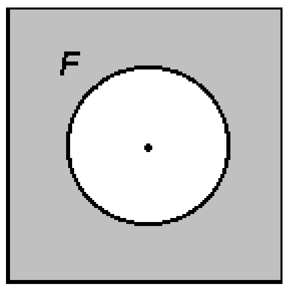
It is not easy to represent relations in a Venn diagram, but we shall do so by means of arrow diagrams. An arrow from one entity to another, possibly even to itself, represents the fact that the entity at the arrow’s source bears the relation to the target entity. Arrows for different relations will be drawn in different colors. Some will be labeled by the relational predicate that stands for them in $\mathfrak{I}$.

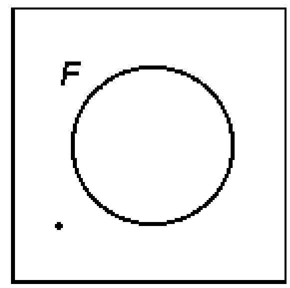
Unrestricted Quantifiers

Let us begin with the simple use of the universal and existential quantifiers to say (1) that everything in the universe falls in the class named by *F,* and (2) that at least one thing falls in that class:

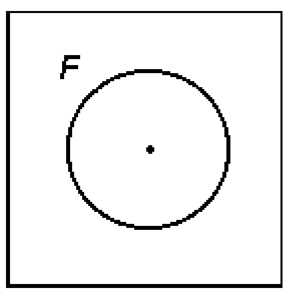
True | False

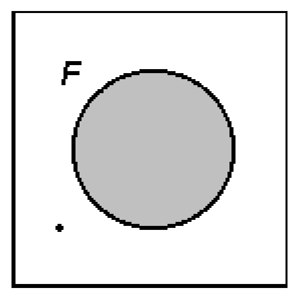
Everything is F
$\forall xFx$





Something is F
$\exists xFx$





Notice in the first case that because it is assumed that the domain *D* is non-empty if *everything is F* is true, then there is at least one entity in the extension of *F.*

Universal Affirmatives

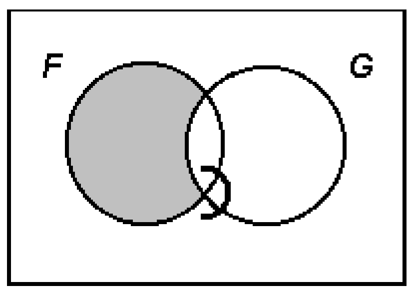
In modern notation the universal affirmative **A**-proposition *Every F is G* is reformulated as a conditional and symbolized using $\rightarrow$:

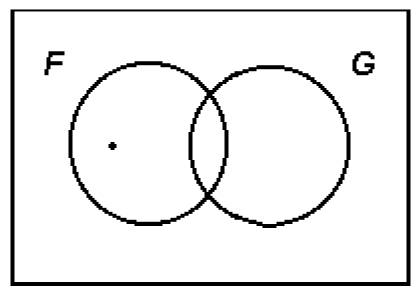
Every F is G *For all x, if Fx then Gx* $\forall x(Fx \rightarrow Gx)$

It is important to see how this differs from the conjunction *for all x, Fx and Gx.* As the diagrams below show, the latter asserts the very strong claim that both *F* and *G* are true of everything in the world . It is hard to find even one predicate true of everything there is, much less two. It is quite common, in contrast, to have cases in which one set is a subset of another, which is what the **A**-proposition asserts.

True False

Every F is G
$\forall x(Fx \rightarrow Gx)$

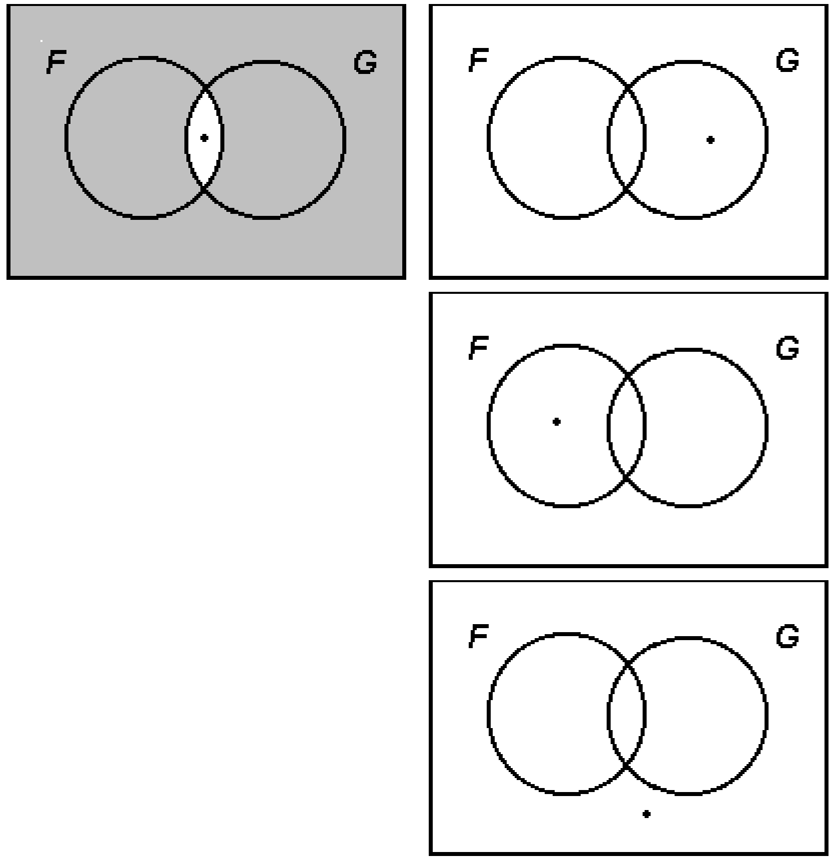




(Here the bold line crossing the lines separating the three subregions is an entity "on the line". It indicates that there is at least one entity in the domain without declaring which subregion it is in.)

True False

Everything is both F and G
$\forall x(Fx \wedge Gx)$



Particular Affirmatives

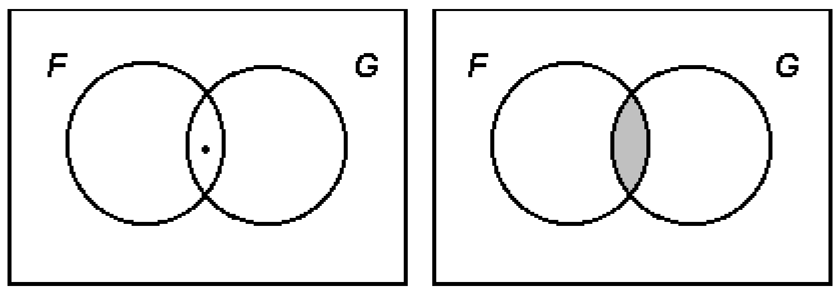
In modern notation the particular affirmative **I**-proposition *Some F is G* is reformulated as a conjunction and symbolized using $\wedge$:

Some F is G *For some x, Fx and Gx* $\exists x(Fx \wedge Gx)$

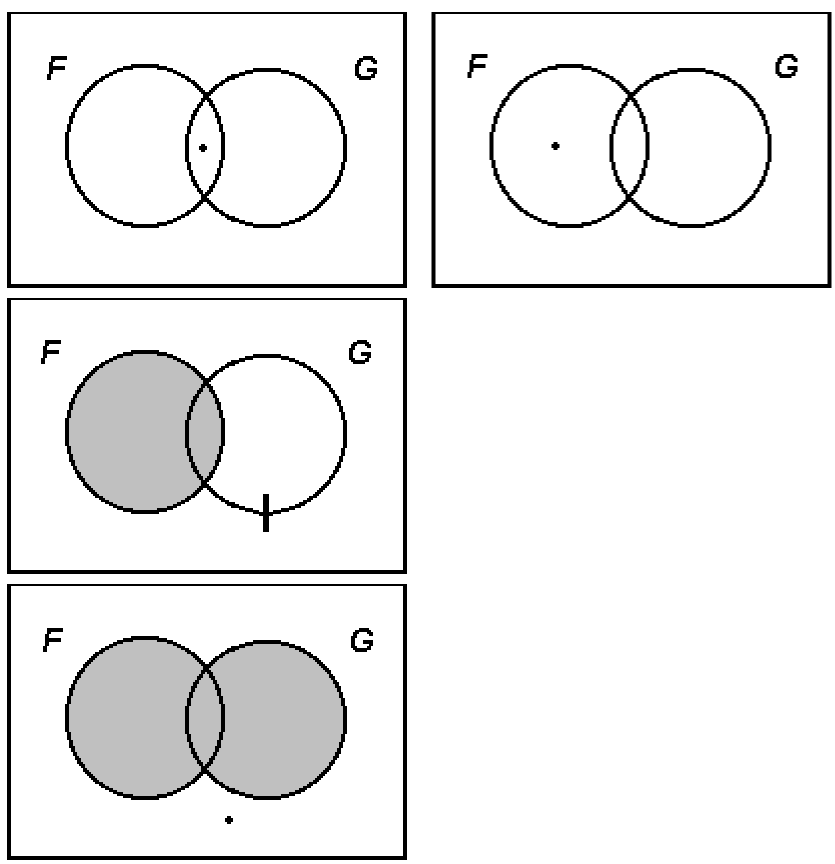
It is important to see how this differs from the conditional $\exists x(Fx \rightarrow Gx)$. As the diagrams below show, the latter asserts a rather odd claim. Given the truth-table for $\rightarrow$, this conditional is true in three cases: (1) when both *Fx* and *Gx* are true, (2) when *Fx* is false and *Gx* is true, and (3) when both *Fx* and *Gx* are false. Clearly, when we say *some F are G,* we do not want our claim to be true if there are no *F’*s, as would be the case in (2) and (3). Hence, $\exists x(Fx \rightarrow Gx)$ is an inappropriate translation of *some F are G.* We use rather $\exists x(Fx \wedge Gx)$, which is true in the right circumstances, *viz.* when there is an object of which both *F* and *G* are true.

True False

Some F is G
$\exists x(Fx \wedge Gx)$



Something is such that if it is F then it is G
$\exists x(Fx \rightarrow Gx)$



Distribution of the Quantifiers over Connectives

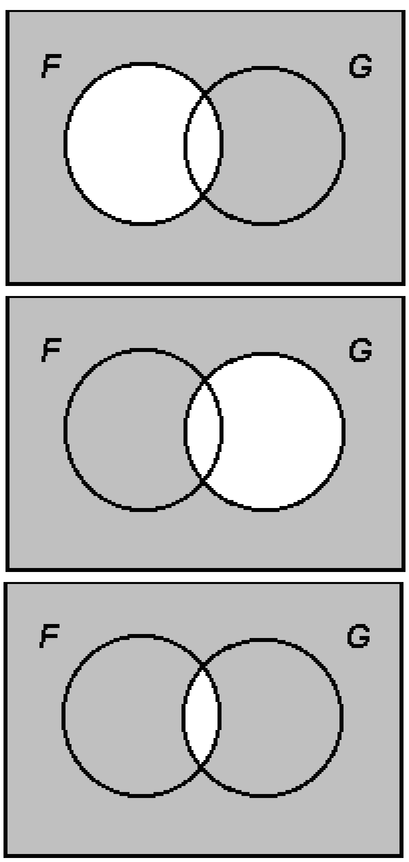
In some cases the distribution of a quantifier makes a difference in meaning. Though the following pairs are equivalent:

$\forall xFx \wedge \forall xGx \qquad \forall x(Fx \wedge Gx)$

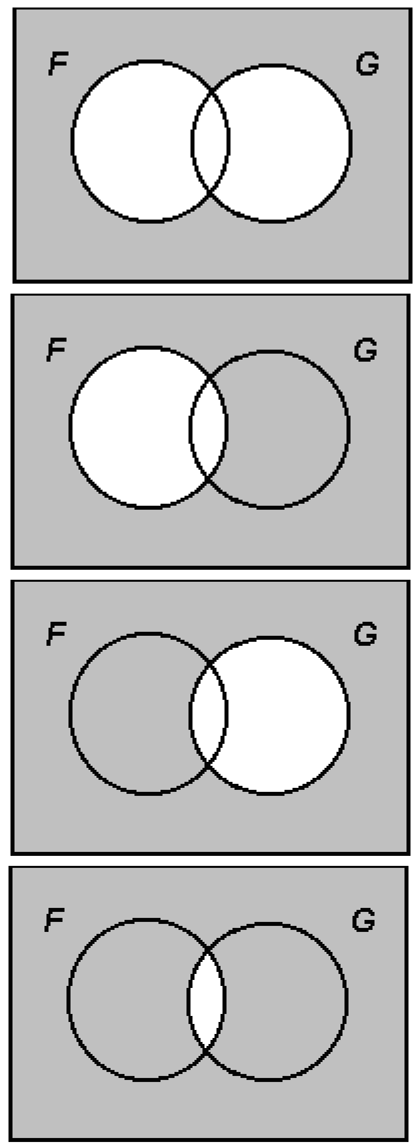
$\exists xFx \vee \forall xGx \qquad \exists x(Fx \vee Gx)$

However $\forall xFx \vee \forall xGx$ entails but is not entailed by $\forall x(Fx \vee Gx)$:

$\forall xFx \vee \forall xGx$

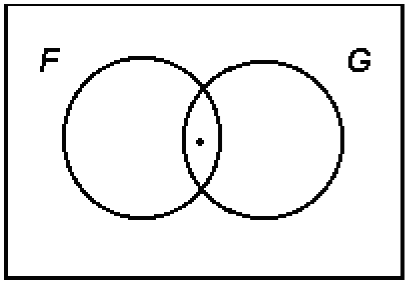


$\forall x(Fx \vee Gx)$

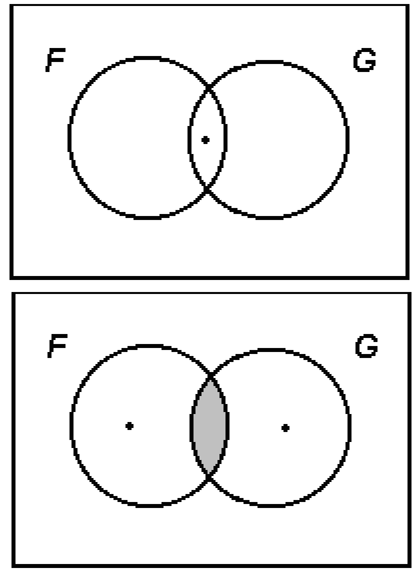


Likewise $\exists x(Fx \wedge Gx)$ entails but is not entailed by $\exists x Fx \wedge \exists x Gx$:

$\exists x(Fx \wedge Gx)$



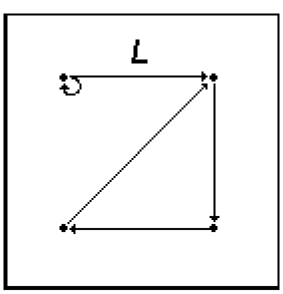
$\exists x Fx \wedge \exists x Gx$

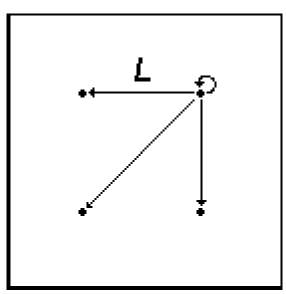


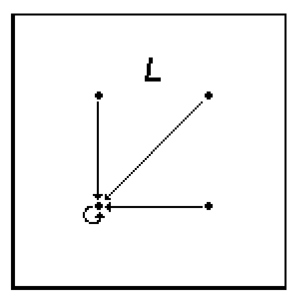
Embedded Quantifiers

The examples below illustrate the affect on meaning of embedding one quantifier within the scope of another.

Everybody loves somebody or other.	*Some body loves everybody.*	*Everybody loves some one person.*
$\forall x \exists y Lxy$	$\exists x \forall y Lxy$	$\forall x \exists y(Lxy \wedge \forall z(Lxz \rightarrow z=y))$



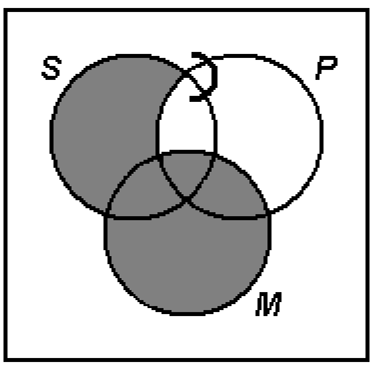




Syllogisms in Modern Notation

Using the informal methods of Venn diagrams, let us illustrate how syllogisms presumed to be valid in Aristotelian and mediaeval logic are invalid in first-order logic given their normal translation. Consider Felapton (**EAO** in the third figure). We construct a model in which its premises are true and its conclusion false:

No M is P	$\sim\exists x(Mx \wedge Px)$	true
<u>*Every M is S*</u>	$\underline{\forall x(Mx \rightarrow Sx)}$	true
Some S is not P	$\exists x(Sx \wedge \sim Px)$	false



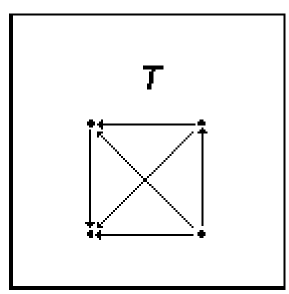
Properties of Relations

Lastly, let us illustrate how to diagram some of the properties of relational predicates.

Taller-than is transitive
$\forall x \forall y \forall z((Txy \wedge Tyz) \rightarrow Txz)$

Taller-than is anti-symmetric

$\forall x \forall y((Txy \rightarrow \sim Tyx)$



Exercises

1. Construct a Venn diagram in which the sentences below are all true together:

$\forall x(Fx \rightarrow Gxy)$
$\exists x(Gx \wedge Hx)$
$\sim\exists x(Fx \wedge Hx)$

2. Construct Venn diagram in which $\forall x(Fx \rightarrow \exists y(Lxy))$ is true but $\exists y \forall x(Fx \rightarrow Lxy)$ false.

3. Symbolize in the notation of first-order logic the syllogism Bramantip (**AAI** in the fourth figure). Construct a Venn diagram showing that in modern notation it is invalid because in the diagram the premises are true but the conclusion is false.

4. Construct an arrow diagram in which the relation *same size as,* represented by the letter *S,* is reflexive, transitive and symmetric.

Intuitions about the Truth-Conditions of Each Formula Types

Atomic Formulas

An interpretation $\Im$ is defined relative to a universe *D*, *the domain* of the interpretation, which represents all the entities that "exist" according to that interpretation. The task of $\Im$ is to assign referents to variables, constants, predicates and formulas. We shall understand $\Im$ to be relation that pairs expression to their "interpretations" relative to *D.* A constant *c* or variable *x* will stand for an individual in the universe *D*. That is,

For any constant *c*, $\Im(c) \in D$.
For any variable *x*, $\Im(\text{x}) \in D$, and

A one-place predicate will stand for a subset of *D,* and an *n*-place predicate (for $n \geq 2$) will stand for an *n*-place relation among members of *D.* That is,

For any P^1_m, $\Im(P^1_m) \subseteq D$.
For any P^n_m such that $n \geq 2$, $\Im(P^n_m)$ is a set of *n*-tuples of elements of *D*

A formula *P* will stand for a truth-value.

For any *P*, $\Im(P)$ is T or F.

Let us consider atomic formulas first. We no longer have as atomic formulas the **A**, **E**, **I**, and **O** propositions of the syllogistic, but rather formulas made up of constants and variables that refer to individuals in the domain *D,* and of predicates that talk about set and relations among these individuals.

An atomic formula made out of a one-place predicate (that refers to a set) will be true if the individual named by its constant or variable are in that set. Let t_i be a constant or variable (i.e. a member of *Trms*). Then, $P^1_m\ t_i$ is an atomic formula that says the individual named by t_i is in the set named by P^1_m:

$\Im(P^1_m\ t_i)$=T iff $\Im(t_i)\in\Im(P^1_m)$.

An atomic formula made of a n-place predicate, which that stands in interpretation $\Im$ for an n-place relation, is true if its n constants or variables pick out entities in $\Im$ that stand in the relation named by the predicate in $\Im$. Let $t_1,\ldots,t_n$ be constants or variables (i.e. members of *Trms*). Then, $P^1_m\ t_1$ is an atomic formula that says the individual named in $\Im$ by t_1 is in the set named in $\Im$ by P^1_m:

$\Im(P^1_m\ t_1)$=T iff $\Im(t_1)\in\Im(P^1_m)$, and

An atomic formula made out of an n-place predicate, which refers in an interpretation $\Im$ to a n-place relation, will be true if the individuals named in $\Im$ by its constants or variables are, in the order indicated, stand in the relation named in $\Im$ by the predicate. Let $t_1,\ldots,t_n$ be constants or variables (i.e. members of *Trms*). Then, $P^n_m\ t_1,\ldots,t_n$ is an atomic formula that says that the individuals named in $\Im$ by $t_1,\ldots,t_n$ stand (in that order) in the relation named in $\Im$ by P^n_m. Now, an n-place relation is a set of n-tuples. Thus, to say that the individuals $\Im(t_1), \ldots, \Im(t_n)$ in that order stand in the relation named by P^n_m may be said more briefly as $<\Im(t_1), \ldots, \Im(t_n)>\ \Im(P^n_m)$. That is,

$\Im(P^n_m\ t_1,\ldots,t_n)$=T iff $<\Im(t_1), \ldots, \Im(t_n)>\in\Im(P^n_m)$

Molecular Formulas: The Connectives

Let us now consider molecular formulas. As in the propositional logic, we shall continue to use the truth-functions $tf_{\sim}$, $tf_{\wedge}$, $tf_{\vee}$, $tf_{\rightarrow}$, and $tf_{\leftrightarrow}$ for the connectives $\sim$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$ (described in their truth-tables) to explain how $\Im$ assigns truth-values to the formulas made up from them:

a. $\Im(\sim P)$=T iff $tf_{\sim}(\Im(P))$ =T
b. $\Im(P\wedge Q)$ =T iff $tf_{\wedge}(\Im(P), \Im(Q))$ =T
c. $\Im(P\vee Q)$ =T iff $tf_{\vee}(\Im(P), \Im(Q))$ =T
d. $\Im(P\rightarrow Q)$ =T iff $tf_{\rightarrow}(\Im(P), \Im(Q))$ =T
e. $\Im(P\leftrightarrow Q)$ =T iff $tf_{\leftrightarrow}(\Im(P), \Im(Q)$ =T

Quantified Formulas

We have one further step: explaining the truth-values of quantified formulas. We must explain when $\Im(\forall xP)$=T and $\Im(\exists xP)$=T. Universally and existential quantified expressions talks about “everything” or “something’, but explaining how they do so precisely is a bit tricky. The easiest way to do so is to look at the formula that the quantifier is attached to. The formula $\forall xFx$, for example, attaches the quantifier $\forall x$ to *Fx.* It says that the open formula *Fx* is true of everything in the universe. One way to say this is that no matter what *x* stands for, it will be true to say *Fx.* Likewise $\exists xFx$ is true if there is at least one thing in the universe that *x* could stand for that would make *Fx* true.

To make this idea precise, let us use the notation $\Im_{[x\rightarrow d]}$ to represent an interpretation that is like $\Im$ in what it assigns to all expressions other than *x* but that reassigns to *x* the entity *d.* That is, $\Im_{[x\rightarrow d]}$ provides a notation for the interpretation that makes x stand for d but otherwise keep all the other assignments the same as those of $\Im$.

Suppose, for example, that the domain has thirty seven different members, i.e. D={d_1, d_2,…, d_{37}}. Then, there will be thirty-seven different ways to change what x stands for in $\Im$, one reassignment for each entity in the domain There will be: $\Im_{[x\to d_1]}$, $\Im_{[x\to d_2]}$,…, $\Im_{[x\to d_{37}]}$. Suppose that in all thirty-seven P is true, i.e. that $\Im_{[x\to d_1]}(P)$=T, $\Im_{[x\to d_2]}(P)$=T, ,…, $\Im_{[x\to d_{37}]}(P)$=T. That would mean, that no matter how the referent of x varied, the formula P is true. Suppose, for example the P is Fx and that $\Im(F)$= {d_1, d_2,…, d_{37}}. That is, in $\Im$ the predicate F stands for the entire domain D. Then it should be the case that $\forall xFx$ is true. Lets see how to express this using the notation $\Im_{[x\to d]}$ to make x stand one at a time for each entity in the domain. Notice first that

(1) $d_1\in\Im(F)$, $d_2\in\Im(F)$,…, $d_{37}\in\Im(F)$,

But by definition of $\Im_{[x\to d]}$, we know that $\Im_{[x\to d]}(x)=d$ because the whole point of $\Im_{[x\to d]}$ is that it reassign x that it stands for d. Hence, we rename d_1, d_2,…, d_{37} in (1) and obtain:

(2) $\Im_{[x\to d_1]}(x)\in\Im(F)$, $\Im_{[x\to d_2]}(x)\in\Im(F)$,…, $\Im_{[x\to d_{37}]}(x)\in\Im(F)$,

But since d_1, d_2,…, d_{37} consist of everything in the domain D , we may summarize (2) as:

(3) for any $d\in D$, $\Im_{[x\to d]}(Fx)$=T.

In this way (3) summarizes the fact that no matter how we vary the referent of x over the domain while at the same time keeping the referents of expressions other than x fixed as specified by $\Im$, the open sentence Fx is true. Thus, (3) is equivalent to:

(4) $\Im(\forall xFx)$=T

and we may use (3) as the "truth-conditions" for $\forall xFx$:

$$\Im(\forall xFx)=\text{T iff for any } d\in D, \Im_{[x\to d]}(Fx)=\text{T}$$

Thus we have a way to state the "truth-conditions" for an arbitrary universally quantified formula $\forall xP$:

$$\Im(\forall xP)=\text{T} \leftrightarrow \text{for any } d\in D, \Im_{[x\to d]}(P)=\text{T}$$

Note that the phrase

for any $d\in D$, $\Im_{[x\to d]}(P)$=T

qualifies as the "the truth-conditions" of $\forall xP$, i.e. as $TC_{\Im}(\forall xP)$, because it is formulated only in terms of the $\Im$-values of the parts of *P*.[21] Hence we have an instance of Tarski's T principle:

$\Im(\forall xP)=\text{T} \leftrightarrow TC_{\Im}(\forall xP)$.

The interpretation of existential quantified formulas is similar: $\exists xP$ *true* if there is at least one way to assign a referent to *x* that makes *P* true:

$\Im(\exists xP)=\text{T} \leftrightarrow$ for some $d\in D$, $\Im_{[x\to d]}(P)$=T

We are now read to define the notion of interpretation inductively. Let us now put these various pieces together and state the general definition for interpretation.

The Inductive Definition of Interpretation

Introduction

As in the propositional logic the definition of an interpretation will be inductive. We first specify a "starter set" and then close this set under some construction rules. The

[21] The induction here is actually on the values of the parts of the formula in all interpretations, as explained shortly.
.

starter set here will be a set of pairs that assign values to the atomic formulas of the syntax. There will then be a set of rules, one for each connective and one for each of the two quantifiers. These rules add a complex formula and its truth-value to a given interpretation given that its parts with its truth-values have already be added to this and other interpretations. We have seen what these rules should be in the discussion we have just completed on the truth-conditions of the various formula types. These are all combined in the definition below. Let us state the definition and then make some comments about it.

Formal Definitions

First we specify a given first-order syntax *FOSyn*= <*Cns, Vbls, Prds, AFor, FR, For*>. Next specify a non-empty set *D* to serve as a domain. Next we define a *basic interpretation* $\Im^D$ relative to *D* as a set of pairs that assigns a entity in *D* to each constant and variable, a set or relation on *D* to each predicate in *Prds,* and a truth-value T or F to each atomic formula in *AFor* as follows

1. For any variable x_n, $\Im^D(x_n)\in D$, and
2. For any constant c_n, $\Im^D(c_n)\in D$
3. For any m, $\Im(P^1_m)\subseteq D$ and
4. For any n and m, $\Im(P^n_m)$ is a set of n-tuples of elements of D
5. $\Im(\mathbf{=})$ is the identity relation on members of D
6. For any m, $\Im(P^1_m\, t_1)$=T iff $\Im(t_1)\in\Im(P^1_m)$, and
7. For any n and m, $\Im(P^n_m\, t_1,\ldots,t_n)$=T iff $<\Im(t_1), \ldots, \Im(t_n)>\in\Im(P^n_m)$
8. For any n and m, $\Im(t_n\mathbf{=}t_m)$=T iff $\Im(t_n)=\Im(t_m)$

We now define the notion of an interpretation inductively in terms of a basic deduction and the series of rules as described earlier:

A first order interpretation relative to basic interpretation $\Im^D$ relative to *D* is a function $\Im$ such that ($\Im$ extends $\Im^D$ as follows):

1. $\Im^D \subseteq \Im$ (i.e. if $<P^n_m\ t_1,\ldots,t_n\ V> \in \Im^D$, then $<P^n_m\ t_1,\ldots,t_n,V> \in \Im$)
2. Construction Steps:
   b. if $tf_{\sim}(\Im(P))$ =T, then $\Im(\sim P)$=T;
      $tf_{\sim}(\Im(P))$ = F otherwise;
   c. if $tf_{\wedge}(\Im(P), \Im(Q))$ =T, then $\Im(P\wedge Q)$ =T;
      $tf_{\wedge}(\Im(P), \Im(Q))$ = F otherwise;
   d. if $tf_{\vee}(\Im(P), \Im(Q))$ =T, then $\Im(P\vee Q)$ =T;
      $tf_{\vee}(\Im(P), \Im(Q))$ = F otherwise;
   e. if $tf_{\rightarrow}(\Im(P), \Im(Q))$ =T, then $\Im(P\rightarrow Q)$ =T;
      $tf_{\rightarrow}(\Im(P), \Im(Q))$ = F otherwise;
   f. if $tf_{\leftrightarrow}(\Im(P), \Im(Q)$ =T, then $\Im(P\leftrightarrow Q)$ =T;
      $tf_{\leftrightarrow}(\Im(P), \Im(Q)$ = F otherwise;
   g. if for any $d\in D$, $\Im_{[x\rightarrow d]}(P)$=T, then $\Im(\forall xP)$=T;
      $\Im(\forall xP)$ = F otherwise;
   h. if for some $d\in D$, $\Im_{[x\rightarrow d]}(P)$=T, then $\Im(\exists xP)$=T;
      $\Im(\exists xP)$= F otherwise.
3. Nothing else is in $\Im$.

We shall let *ForIntrp* be the set of all first-order interpretations $\Im$ defined relative to any *basic interpretation* $\Im^D$, and let $\Im$ stand for interpretations in *ForIntrp.* We define a *first-order language L* as the pair <*FOSyn*, *ForIntrp* >.

Simultaneous Induction and Impossibility of Truth-Tables

Strictly speaking, though "for any $d\in D$, $\Im_{[x\rightarrow d]}(P)$=T" does explain $\Im(\forall xP)$=T, it does not do so in terms of just the $\Im$-values of the immediate parts of $\forall xP$. This is so because "for any $d\in D$, $\Im_{[x\rightarrow d]}(P)$=T" does not talk merely about what $\Im$ assigns to *P,* it also refers to what the various interpretations $\Im_{[x\rightarrow d]}$ assign to *P.* That is, whether a pair <∀*xFx*> is added to $\Im$ will be determined not just by whether <*Fx,*T> is in $\Im$, but on whether <*Fx,*T> is in every $\Im_{[x\rightarrow d]}$. More generally, for any formula *P,* before a pair <*P,V*> is added to $\Im$, it is assumed that for any part *Q* of *P* and any interpretation $\Im'$ whatever, the value *V* of *Q* in $\Im'$ is determined. The definition of $\Im$ remains well-defined, nevertheless.

First the values of the atomic formulas are simultaneously fixed in every $\Im^D$ all at once. Thus the "starter set" for each interpretation $\Im^D$ is fixed. Let us say that atomic formulas are of length 1. These atomic valuations (of formulas of length 1) are then used to determine the values of formulas made up of them, both in $\Im$ and in all other interpretations. For example, the values in $\Im$ and every other interpretation by now be determined for the formulas *P* that have atomic formulas as their immediate parts. The value of $\Im(P)$ can be determined, even if *P* starts with a universal quantifier (if say *P* is ∀xFx) because the values of its atomic part (in this case *Fx*) is already determined not only in $\Im$ but in all other interpretations, including all $\Im_{[x\rightarrow d]}$. In this way all formulas having atomic formulas as their immediate parts get their values fixed for all interpretations at the same time. Let us say a formula is of length 2 if it is either atomic (of length 1) or made up of atomic formulas. As we have just seen, the values in all interpretations of all formulas of length 2 have be fixed. Let us now consider all formulas that are made up of formulas of length 2 or less (i.e. all formulas made up of atomic formulas or of formulas that have atomic formulas as their immediate parts). These we shall say are of length 3. As we have seen all formula of length 3 have immediate parts that already have their interpretations fixed in all interpretations. We can then apply this knowledge of the values of the parts to determine those of the whole, even though the rule fixing the value in $\Im$ may require information about the values of the parts in other interpretations. If we say, generally, that a formula is of length n+1 if it is made of formulas whose parts are of length n or less, we see that when the value of a formula of length n+1 in $\Im$ is defined, all the values of its parts, which are of length n or less, have been predefined, not only in $\Im$ but in all other interpretations as well. In this way the value of formulas in all interpretations is determined in stages corresponding to the stages of construction of each formula. The set of interpretations is said to be defined by *simultaneous induction*.

Tough every $\Im$ is well defined by the process of simultaneous induction, the method lacks an important feature of ordinary definition by induction. It is not longer the case that every element of $\Im$ has a construction sequence. This happens because the information needed to put a pair, say $<\forall xFx>$, into $\Im$ might be infinite but a construction sequence by definition is finite. For example, to put $<\forall xFx>$ in $\Im$ we must have already put $<Fx, \mathrm{T}>$ in all $\Im_{[x\rightarrow d]}$, and there might be an infinite number of these because there might be an infinite number of entities in the domain *D.* We simply could not list put all these prior pairs in a finite construction sequence that ended with $<\forall xFx>$.

We can now see that first-order semantics does not allow us to lay out a finite truth-table displaying how the value of a formula is calculated from those of its parts. There are two reasons there could be no such table. First of all, if there are infinite number of entities, as there are if we include numbers among the things that exist, there are an infinite number of interpretation. But there cannot be an infinite number of lines in a truth-table. Moreover, the "line" laying out the information needed to "calculate" the value of a quantified formula, say $\forall xFx$, might also be infinitely long because it would need to list the values of its immediate part, in this case *Fx,* in other interpretations $\Im_{[x\rightarrow d]}$, of which there may be an infinite number. But a truth-table cannot have a line that is infinitely long. We will see in Part 3 that this difference between first-order and propositional logic is profound. We will be able to used truth-tables as decision procedure to test arguments in propositional logic for their validity, but we shall also see that there is in principle no such test for arguments in first-order logic.

Tarski's Adequacy Condition

The definition of interpretation satisfies Tarski's condition (T) for counting as a correspondence theory of truth. It does so for formulas formed by the truth-functional connectives because the truth-conditions are the same as in propositional logic. The only new case are formulas formed by the quantifiers. We show that the (T) principle is satisfied in the following metatheorem.

Metatheorem. For any formula *P,*

$$\Im(P)=\mathrm{T} \leftrightarrow \mathrm{TC}_{\Im}(P).$$

Proof. Given the definition of $\Im$ it follows that it is two-valued. Given the definition and the fact that it is two-valued it follows that the truth-value of a molecular formula is equivalent to a statement that specifies truth-conditions in some interpretation for its immediate parts and that the truth-value of an atomic formula is equivalent to a statement that specifies conditions on the $\Im$-values of the predicate and terms that occur in the formula, as follows:

1. $\Im(P^1_m\, t_1)$=T iff $\Im(t_1)\in\Im(P^1_m)$,
2. $\Im(P^n_m\, t_1,\ldots,t_n)$=T iff $<\Im(t_1), \ldots, \Im(t_n)>\in\Im(P^n_m)$

3. $\Im(t_n=t_m)$=T iff $\Im(t_n)=\Im(t_m)$
4. $\Im(\sim P)$=T iff $\Im(P)\neq$T
5. $\Im(P\wedge Q)$ =T iff, $\Im(P)$ =T and $\Im(Q)$) =T
6. $\Im(P\vee Q)$ =T iff, $\Im(P)$ =T or $\Im(Q)$) =T
7. $\Im(P\rightarrow Q)$ =T iff, $\Im(P)\neq$T or $\Im(Q)$) =T
8. $\Im(P\leftrightarrow Q)$ =T iff, either $\Im(P)$=T and $\Im(Q)$=T , or $\Im(P)\neq$T and $\Im(Q)\neq$T
9. $\Im(\forall xP)$=T iff, for any $d\in D$, $\Im_{[x\rightarrow d]}(P)$=T
10. $\Im(\exists xP)$=T iff, for some $d\in D$, $\Im_{[x\rightarrow d]}(P)$=T

Given that each formula is has a finite grammatical derivation, it follows that by a finite number of applications of the substitution of equivalents as specified in 1-2 above, a statement $\Im(P)$=T can be transformed into an equivalent that mentions only the interpretations of the predicates and terms that occur in the atomic formulas in *P.* Since this statement is equivalent to $\Im(P)$=T and is formulated only in terms of the interpretations of its terms and predicates is qualifies as $TC_\Im(P)$. Hence $\Im(P)$=T iff $TC_\Im(P)$. End of proof.

Calculating Truth-Values Using Truth-Conditions

The Technique

Since an interpretation $\Im$ does not have a simple inductive definition, it is no longer the case as it is in propositional logic that there is a finite construction sequence for every assignment pair in $\Im$. As a result, it is not possible to calculate by the truth-table method the truth-value of a whole formula from those of its atomic parts. Another technique is needed for determining when a formula is true in $\Im$. We describe one that makes use of a formula's truth-conditions as set forth in instances of Tarski's principle:

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

This principle tells us that all we need do to show that $\Im(P)$=T is prove that the conditions $TC_{\Im}(P)$ are satisfied.

Below we give examples of how to calculate the truth conditions by reference to the equivalences proven earlier:

E1. $\Im(P^1_m\, t_1)$=T iff $\Im(t_1)\in\Im(P^1_m)$,

E2. $\Im(P^n_m\, t_1,\ldots,t_n)$=T iff $<\Im(t_1), \ldots, \Im(t_n)>\in\Im(P^n_m)$

E3. $\Im(t_n\mathbf{=}t_m)$=T iff $\Im(t_n)=\Im(t_m)$

E4. $\Im(\sim P)$=T iff $\Im(P)\neq$T

E5. $\Im(P\wedge Q)$ =T iff, $\Im(P)$ =T and $\Im(Q))$ =T

E6. $\Im(P\vee Q)$ =T iff, $\Im(P)$ =T or $\Im(Q))$ =T

E7. $\Im(P\rightarrow Q)$ =T iff, $\Im(P)\neq$T or $\Im(Q))$ =T

E8. $\Im(P\leftrightarrow Q)$ =T iff, either $\Im(P)$=T and $\Im(Q)$=T , or $\Im(P)\neq$T and $\Im(Q)\neq$T

E9. $\Im(\forall xP)$=T iff, for any $d\in D$, $\Im_{[x\rightarrow d]}(P)$=T

E10. $\Im(\exists xP)$=T iff, for some $d\in D$, $\Im_{[x\rightarrow d]}(P)$=T

Below, for various examples of *P,* we work out the truth conditions for *P* in $\Im$, that is we work out $\Im(P)$=T iff TC(P). We do so by applying the equivalences E1-E10 above, one after anther, to the progressively smaller parts of *P,* whatever they are.

Since E1-E10 they are already proven (indeed, since they follow from the definition of $\Im$ by logic and set theory, they are theorems of naïve set theory), we can simply write an one of them down as true in any proof we are constructing. Moreover since E1-E10 are biconditionals, we can substitute one side for the other. In sum, the way we will deduce $\Im(P)$=T iff $TC_{\Im}(P)$ is by writing down relevant cases of E1-E10, and then make substitutions based on the equivalences they provide. Each line of the proof will either be a direct instance of E1-E10, or will result from an earlier line by an E1-E10 substitution. If we proceed in this way, it will follow, as Tarski required, that each instance of $\mathfrak{I}(P)$=T iff TC(P) is "a theorem of set theory that follows from the definition of $\mathfrak{I}$".

Examples

Before stating the examples, it will help to remark on notation. Recall that $\mathfrak{I}^{D}{}_{[x\to d]}$ is that interpretation like $\mathfrak{I}^{D}$ except that it assigns d to x. That is, $\mathfrak{I}^{D}{}_{[x\to d]}$ pairs x with d. This fact is written in functional notation as $\mathfrak{I}^{D}{}_{[x\to d]}(x)=d$. Likewise $\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}$ is that interpretation like $\mathfrak{I}^{D}{}_{[x\to d]}$ except that it assigns d' to y. Hence, in functional notation $\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(y)=d'$, but it also it remains the case that $\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(x)=d$. Below, to aid the eyes to see these identities, terms that name the same object have the same color. Thus,

> $\mathfrak{I}^{D}{}_{[x\to d]}(x)$ and $\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(x)$ in red are alternative notation for d, and
> $\mathfrak{I}^{D}{}_{[y\to d']}(y)$ and $\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(y)$ in blue are alternative notation for d'.

These will be substituted one for another as instances of the substitution of identity.

Below we work out are Biconditionals of the following form:

Truth-Conditions for *P*
Conditions that must hold in the world among the entities referred to by the smallest parts of speech in P

(T) $\mathfrak{I}(P)$=T iff $TC_{\mathfrak{I}}(P)$

Example 1. $Fc \wedge Gb$

1. $\mathfrak{I}(Fc \wedge Gb)$=T	iff $\mathfrak{I}(Fc)$=T and Gb)=T	E5
	iff $\mathfrak{I}^{D}(c)\in\mathfrak{I}^{D}(F)$ and $\mathfrak{I}^{D}(b)\in\mathfrak{I}^{D}(G)$	1, sub of eq E1

Example 2. $Rac \to Gx$

1. $\Im(Rac \to Gx)$=T	iff $\Im(Rac)\neq$T or $\Im(Gx)$=T	E7
2.	iff $<\Im^D(a),\Im^D(c)>\notin\Im^D(R)$ or $\Im^D(x)\in\Im^D(G)$	1, sub of eq E1

Example 3. $\forall xFx$

1. $\Im(\forall xFx)$=T	iff for all $d\in D$, $\Im^D_{[x\to d]}(Fx)$=T	E9
2.	iff for all $d\in D$, $\Im^D_{[x\to d]}(x)\in\Im^D(F)$	1, sub of eq E1
3.	iff for all $d\in D$, $d\in\Im^D(F)$	2, sub of =

Example 4. $\exists xFx$

1. $\Im(\exists xFx)$=T	iff for some $d\in D$, $\Im^D_{[x\to d]}(Fx)$=T	E10
2.	iff for some $d\in D$, $\Im^D_{[x\to d]}(x)\in\Im^D(F)$	1, sub of eq E1
3.	iff for some $d\in D$, $d\in\Im^D(F)$	2, sub of =

Example 5. $\forall x\exists yRxy$

1. $\Im(\forall x\exists yRxy)$=T	iff for all $d\in D$, $\Im^D_{[x\to d]}(\exists yRxy)$=T	E9
2.	iff iff for all $d\in D$, for some $d'\in D$, $\Im^D_{[x\to d,\ y\to d']}(Rxy)$=T	1, sub eq E10
3.	iff for all $d\in D$, for some $d'\in D$, $<\Im^D_{[x\to d,\ y\to d']}(x),\Im^D_{[x\to d,\ y\to d']}(y)>\in\Im^D(R))$	2, sub eq E1
4.	iff for all $d\in D$, for some $d'\in D$, $<d,d'>\in\Im^D(R))$	3, sub of =

Example 6. $\exists x\forall yRxy$

1. $\Im(\exists x\forall yRxy)$=T	iff for some $d\in D$, $\Im^D_{[x\to d]}(\forall yRxy)$=T	E9
2.	iff iff for some $d\in D$, for all $d'\in D$, $\Im^D_{[x\to d,\ y\to d']}(Rxy)$=T	1, sub eq E10
3.	iff for some $d\in D$, for all $d'\in D$, $<\Im^D_{[x\to d,\ y\to d']}(x),\Im^D_{[x\to d,\ y\to d']}(y)>\in\Im^D(R))$	2, sub eq E1
4.	iff for some $d\in D$, for all $d'\in D$, $<d,d'>\in\Im^D(R))$	3, sub of =

Example 7. $\forall xRxx$

1. $\Im(\forall xRxx)$=T	iff for all $d\in D$, $\Im^D_{[x\to d]}(Rxx)$=T	E9
2.	iff for all $d\in D$, $<\Im^D_{[x\to d,\ y\to d']}(x),\Im^D_{[x\to d,\ y\to d']}(x)>\in\Im^D(R))$	1, sub eq E1
3.	iff for all $d\in D$, $<d,d>\in\Im^D(R))$	3, sub of =

Example 8. $\forall x(Fx\to Gx)$

1. $\Im(\forall x(Fx\to Gx))$=T	iff for all $d\in D$, $\Im^D_{[x\to d]}(Fx\to Gx)$=T	E9
2.	iff for all $d\in D$, either $\Im^D_{[x\to d]}(Fx)$=T or $\Im^D_{[x\to d]}(Gx))\neq$T	1, sub of eq E7

3. iff for all $d \in D$,
either $\mathfrak{I}^{D}_{[x\to d]}(x) \in \mathfrak{I}^{D}(F)$ or $\mathfrak{I}^{D}_{[x\to d]}(x) \notin \mathfrak{I}^{D}(G)$ — 2, sub of eq E1

4. iff for all $d \in D$, either $d \in \mathfrak{I}^{D}(F)$ or $d \notin \mathfrak{I}^{D}(G)$ — 3, sub of =

Example 9. $\exists x(Fx \wedge Gx)$

1. $\mathfrak{I}(\exists x(Fx \wedge Gx))=T$ iff for some $d \in D$, $\mathfrak{I}^{D}_{[x\to d]}(Fx \wedge Gx)=T$ — E10
2. iff for some $d \in D$, $\mathfrak{I}^{D}_{[x\to d]}(Fx)=T$ and $\mathfrak{I}^{D}_{[x\to d]}(Gx))=T$ — 1, sub of eq E5
3. iff for some $d \in D$, $\mathfrak{I}^{D}_{[x\to d]}(x) \in \mathfrak{I}^{D}(F)$ and $\mathfrak{I}^{D}_{[x\to d]}(x) \in \mathfrak{I}^{D}(G)$ — 2, sub of eq E1
4. iff for some $d \in D$, $d \in \mathfrak{I}^{D}(F)$ and $d \in \mathfrak{I}^{D}(G)$ — 3, sub of =

Example 10. $\forall x(Fx \wedge Gx)$

1. $\mathfrak{I}(\forall x(Fx \wedge Gx))=T$ iff for all $d \in D$, $\mathfrak{I}^{D}_{[x\to d]}(Fx \wedge Gx)=T$
2. iff for all $d \in D$, $\mathfrak{I}^{D}_{[x\to d]}(Fx)=T$ and $\mathfrak{I}^{D}_{[x\to d]}(Gx))=T$
3. iff for all $d \in D$,
$\mathfrak{I}^{D}_{[x\to d]}(x) \in \mathfrak{I}^{D}(F)$ and $\mathfrak{I}^{D}_{[x\to d]}(x) \in \mathfrak{I}^{D}(G)$
4. iff for all $d \in D$, $d \in \mathfrak{I}^{D}(F)$ and $d \in \mathfrak{I}^{D}(G)$

Example 11. $\exists x(Fx \to Gx)$

1. $\mathfrak{I}(\exists x(Fx \to Gx))=T$ iff for some $d \in D$, $\mathfrak{I}^{D}_{[x\to d]}(Fx \to Gx)=T$
2. iff for some $d \in D$, either $\mathfrak{I}^{D}_{[x\to d]}(Fx)=T$ or $\mathfrak{I}^{D}_{[x\to d]}(Gx)) \neq T$ — 1
3. iff for some $d \in D$, either $\mathfrak{I}^{D}_{[x\to d]}(x) \in \mathfrak{I}^{D}(F)$ or $\mathfrak{I}^{D}_{[x\to d]}(x) \notin \mathfrak{I}^{D}(G)$
4. iff for some $d \in D$, either $d \in \mathfrak{I}^{D}(F)$ or $d \notin \mathfrak{I}^{D}(G)$

Example 12. $\forall x(Fx \to \exists yRxy)$

1. $\mathfrak{I}(\forall x(Fx \to \exists yRxy))=T$ iff for all $d \in D$, $\mathfrak{I}^{D}_{[x\to d]}(Fx \to \exists y\, Rxy)=T$
2. iff for all $d \in D$, either $(\mathfrak{I}^{D}_{[x\to d]}(Fx) \neq T$ or $(\mathfrak{I}^{D}_{[x\to d]}(\exists yRxy)=T$
3. iff for all $d \in D$, either $(\mathfrak{I}^{D}_{[x\to d]}(x) \notin \mathfrak{I}^{D}(F)$ or for some
$d' \in D$, $(\mathfrak{I}^{D}_{[x\to d,\, y\to d']}(Rxy)=T$
4. iff for all $d \in D$, either $(\mathfrak{I}^{D}_{[x\to d]}(x) \notin \mathfrak{I}^{D}(F)$ or for some
$d' \in D$, $<\mathfrak{I}^{D}_{[x\to d,\, y\to d']}(x), \mathfrak{I}^{D}_{[x\to d,\, y\to d']}(y)> \in \mathfrak{I}^{D}(R))$
5. iff for all $d \in D$, either $(d \notin \mathfrak{I}^{D}(F)$ or for some
$d' \in D$ $<d,d'> \in \mathfrak{I}^{D}(R))$

Example 13. $\forall x \exists y(Rxy \to Ryx)$

1. $\mathfrak{I}(\forall x \exists y(Rxy \to Ryx))=T$ iff for all $d \in D$, $\mathfrak{I}^{D}_{[x\to d]}\exists y(Rxy \to Ryx)=T$
2. iff for all $d \in D$, for some $d' \in D$, $\mathfrak{I}^{D}_{[x\to d,\, y\to d']}((Rxy \to Ryx)=T$
3. iff for all $d \in D$, for some $d' \in D$,
either $\mathfrak{I}^{D}_{[x\to d,\, y\to d']}(Rxy)=T$ or $\mathfrak{I}^{D}_{[x\to d,\, y\to d']}(Rxy) \neq T$
4. iff for all $d \in D$, for some $d' \in D$,

either $<\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(x),\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(y)>\in\mathfrak{I}^{D}(R))$ iff $<\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(y),\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(x)>\notin\mathfrak{I}^{D}(R))$

5. iff for all $d\in D$, for some $d'\in D$, either $<d,d'>\in\mathfrak{I}^{D}(R))$ or $<d',d>\notin\mathfrak{I}^{D}(R))$

Example 14. $\forall x\forall y(Rxy\leftrightarrow Ryx)$

1. $\mathfrak{I}(\forall x\forall y(Rxy\leftrightarrow Ryx))$=T iff for all $d\in D$, $\mathfrak{I}^{D}{}_{[x\to d]}\forall y(Rxy\leftrightarrow Ryx)$=T
2. iff for all $d\in D$, for all $d'\in D$, $\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}((Rxy\leftrightarrow Ryx)$=T
3. iff for all $d\in D$, for all $d'\in D$, $\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(Rxy)$=T iff $\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(Rxy)$=T
4. iff for all $d\in D$, for all $d'\in D$, $<\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(x),\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(y)>\in\mathfrak{I}^{D}(R))$ iff $<\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(y),\mathfrak{I}^{D}{}_{[x\to d,\ y\to d']}(x)>\in\mathfrak{I}^{D}(R))$
5. iff for all $d\in D$, for all $d'\in D$, $<d,d'>\in\mathfrak{I}^{D}(R))$ iff $<d',d>\in\mathfrak{I}^{D}(R))$

Exercises.

*1. Annotate each line of the Example 10 and 11, repeated below, citing either the equivalence E1-E10 that it instantiates, or the number of previous line and the equivalence E1-E10 from which it is derived by the substitution of equivalents, or the numbers of the previous line from which it is derived by the substitution of identity.

Example 10. $\forall x(Fx\wedge Gx)$

1. $\mathfrak{I}(\forall x(Fx\wedge Gx))$=T iff for all $d\in D$, $\mathfrak{I}^{D}{}_{[x\to d]}(Fx\wedge Gx)$=T
2. iff for all $d\in D$, $\mathfrak{I}^{D}{}_{[x\to d]}(Fx)$=T and $\mathfrak{I}^{D}{}_{[x\to d]}(Gx))$=T
3. iff for all $d\in D$, $\mathfrak{I}^{D}{}_{[x\to d]}(x)\in\mathfrak{I}^{D}(F)$ and $\mathfrak{I}^{D}{}_{[x\to d]}(x)\in\mathfrak{I}^{D}(G)$
4. iff for all $d\in D$, $d\in\mathfrak{I}^{D}(F)$ and $d\in\mathfrak{I}^{D}(G)$

Example 11. $\exists x(Fx\to Gx)$

1. $\mathfrak{I}(\exists x(Fx\to Gx))$=T iff for some $d\in D$, $\mathfrak{I}^{D}{}_{[x\to d]}(Fx\to Gx)$=T
2. iff for some $d\in D$, either $\mathfrak{I}^{D}{}_{[x\to d]}(Fx)$=T or $\mathfrak{I}^{D}{}_{[x\to d]}(Gx))\neq$T 1
3. iff for some $d\in D$, either $\mathfrak{I}^{D}{}_{[x\to d]}(x)\in\mathfrak{I}^{D}(F)$ or $\mathfrak{I}^{D}{}_{[x\to d]}(x)\notin\mathfrak{I}^{D}(G)$
4. iff for some $d\in D$, either $d\in\mathfrak{I}^{D}(F)$ or $d\notin\mathfrak{I}^{D}(G)$

2. Work out the truth-conditions with annotation for the two new examples, call them examples 15 and 16:

Example 15.

1. $\mathfrak{I}(\exists xFx\wedge\exists yGy))$=T iff
2.
3.
4.

5.
Example 16.
1. $\Im(\forall xFx\rightarrow\forall yGy))$=T iff
2.
3.
4.
5.

If we first calculate out the truth-conditions of P in $\Im$, i.e. $TC_{\Im}(P)$, and we also know enough facts about $\Im$ itself, then we can often prove that P is true in $\Im$, i.e. that $\Im(P)$=T. We The following metatheorems provide examples. In each we first state some facts about $\Im$. We then calculate out $TC_{\Im}(P)$ for a particular formula *P.* These together provide enough information that we are then able, given the truths of set theory and logic, to deduce that $\Im(P)$=T.

Metatheorem. If D={1,2,3} $\Im(F)$={1}, $\Im$(G)={1,2,}, then $\Im(\forall x(Fx\rightarrow Gx))$=T.

Proof:
First we calculate $TC_{\Im}(\forall x(Fx\rightarrow Gx))$ by successive applications of the earlier metatheorem:

$\Im(\forall x(Fx\rightarrow Gx)$=T   iff   for any $d\in D$, $\Im_{[x\rightarrow d]}(Fx\rightarrow Gx)$=T
    iff   for any $d\in D$, if $\Im_{[x\rightarrow d]}(Fx)$=T then $_{x\rightarrow d]}(Gx)$=T
    iff   for any $d\in D$, if $\Im_{[x\rightarrow d]}(x)\in\Im(G)$ then $\Im_{[x\rightarrow d]}(x)\in\Im(G)$
    iff   for any $d\in D$, if $d\in\Im(G)$ then $d\in\Im(G)$

(Note that the last line follows from the line before by substitution of identities because, given the definition of $\Im_{[x\rightarrow d]}$, $\Im_{[x\rightarrow d]}(x)=d$.)
Hence, $TC_{\Im}(\forall x(Fx\rightarrow Gx))$ is:

for any $d\in D$, if $d\in\Im(G)$ then $d\in\Im(G)$

That is,

0. $\Im(\forall x(Fx\rightarrow Gx))$=T   iff   for any $d\in D$, if $d\in\Im(G)$ then $d\in\Im(G)$

Hence it suffices to prove: for any $d\in D$, if $d\in\Im(G)$ then $d\in\Im(G)$. We do so as follows:

1. Let D={1,2,3} $\Im(F)$={1}, $\Im$(G)={1,2,}. and let d be an arbitrary member of *D.*
Start subproof for conditional proof.

2. $d\in\Im(F)$	Assumption for conditional proof
3. $d\in\{1\}$	1 and 4, sub of =
4. $\{1\}\subseteq\{1,2\}$	set theory
5. $d\in\{1,2\}$	5 and 6, set theory
6. $d\in\Im(G)$	1 and 5, sub of =

End of subproof

7. If $d \in \Im(F)$ then $d \in \Im(G)$	2-6, conditional proof
8. for any $d \in D$, If $d \in \Im(F)$ then $d \in \Im(G)$	7, universal generalization, d arbitrary
9. $\Im(\forall x(Fx \to Gx))$=T	0 and 8, sub of equivalents

Metatheorem. If D={1,2,3} $\Im(F)$={1}, $\Im$(G)={1,2,}, then $\Im(\exists x(Fx \wedge Gx))$=T.

Proof:

First we calculate $TC_{\Im}(\exists x(Fx \wedge Gx))$ by successive applications of the earlier metatheorem:

$\Im(\exists x(Fx \wedge Gx)$=T	iff	for some $d \in D$, $\Im_{[x \to d]}(Fx \wedge Gx)$=T
	iff	for some $d \in D$, $\Im_{[x \to d]}(Fx)$=T and $_{x \to d]}(Gx)$=T
	iff	for some $d \in D$, $\Im_{[x \to d]}(x) \in \Im(G)$ and $\Im_{[x \to d]}(x) \in \Im(G)$
	iff	for some $d \in D$, $d \in \Im(G)$ and $d \in \Im(G)$

(Note that the last line follows from the line before by substitution of identities because, given the definition of $\Im_{[x \to d]}$, $\Im_{[x \to d]}(x)=d$.)
Hence, $TC_{\Im}(\exists x(Fx \wedge Gx))$ is:

for some $d \in D$, $d \in \Im(G)$ and $d \in \Im(G)$

That is,

0. $\Im(\exists x(Fx \wedge Gx))$=T iff for some $d \in D$, $d \in \Im(G)$ and $d \in \Im(G)$

Hence it suffices to prove: for some $d \in D$, $d \in \Im(G)$ and $d \in \Im(G)$. We do so by existential generalization from the details of definition of $\Im$.

1.	Let D={1,2,3} $\Im(F)$={1}, $\Im$(G)={1,2,}.	Given
2.	$1 \in$ {1} and $1 \in$ {1,2,}.	Set theory
3.	for some d, $d \in$ {1} and $d \in$ {1,2,}.	2, existential generalization
4.	$\Im(\exists x(Fx \wedge Gx))$=T	0 and 3, sub of equivalents

Exercise: Prove that if D={1,2,3} $\Im(F)$={1}, $\Im$(G)={2,3}, then

1. $\Im(\forall x(Fx \vee Gx))$=T,
2. $\Im(\exists x(Gx \wedge {\sim}Fx))$=T.

Prove (1) by first calculating $TC_{\Im}(\forall x(Fx \vee Gx))$ by progressive applications of the earlier metatheorem, as in the previous example. Prove (2) by first calculating $TC_{\Im}(\exists x(Gx \wedge {\sim}Fx))$.

The Correspondence Theory of Truth for First-Order Logic

We are now at a point from which it is possible to drive home exactly how the definition of $\Im$ amounts to a correspondence theory of truth. The vague idea behind the correspondence theory is that a sentence is true if it corresponds to the world. The problem with the theory is that it is a tall order to lay out a plausible theory of “the world” and of a account of what correspondence is. Traditionally, philosophers understood the task of explaining ”the world” as requiring no less that a theory of ontology understood as providing a breakdown of all entities that exist into their fundamental categories and of the basic relations that hold among them. It was then part of the standard account that the correspondence between language and the world would consist in the fact that grammar mirror ontological structure. The various parts of speech in grammar would be distinguished by the fact that each is used to name or refer to a characteristic category of entity in the world. Moreover, the grammatical structures that link one part of speech to another to form longer expresses would mirror ontological relations that hold among entities in the world.

Truth-Conditions for *P*

$$\text{(T)} \qquad \Im(P)\text{=T} \qquad \text{iff} \qquad \overbrace{\mathrm{TC}_{\Im}(P)}^{\text{Conditions that must hold in the world among the entities referred to by the smallest parts of speech in } P}$$

Plato. Nouns and verbs both stand for Forms, and the subject-predicate structure of a true sentence corresponds to the relational fact in “the world” that one form inheres in another.

S is P is true	iff	the Form named by *S* imitates the Form named by *P*

Aristotle. The various parts of speech stand for the various categories of being, and a true **A**-proposition *Every S is P* corresponds to the relational fact "in the world" that "what is said" by the predicate *P* is "in" (i.e. inheres in) or is "of" (i.e. is a genus of) the substance referred to by the subject *S*.

Every S is P is true	iff	*P* is "said of" or "said in" *S* (i.e. the accident or genus/species named by *P* inheres in or includes that named by *S*)

The Syllogistic, Modern Version. Terms stand for non-empty sets, and a true categorical proposition corresponds to a characteristic relational fact "in the world" that holds among the sets referred to by the terms.

$\Im(\mathbf{A}SP)$=T	iff	$\Im(S)\subseteq\Im(P)$ where $\Im(S)\neq\varnothing$
$\Im(\mathbf{E}SP)$=T	iff	$\Im(S)\cap\Im(P)=\varnothing$, etc.

Propositional Logic. Atomic sentences, which in the examples below are p_1 and p_2, stand for truth-values. A true molecular sentence corresponds to a the fact "in the world" that is atomic parts "name" truth-values in the particular combination stipulated by the grammatical structure of the sentence's connectives.

$\Im(\sim p_1)$=T	iff	$\Im(p_1)$=F
$\Im(p_1\wedge p_2)$=T	iff	$\Im(p_1)$=T and $\Im(p_2)$=T
$\Im(p_1\vee p_2)$=T	iff	$\Im(p_1)$=T or $\Im(p_2)$=T or both
$\Im(p_1\rightarrow p_2)$=T	iff	either $\Im(p_1)$=F or $\Im(p_2)$=T, etc.

First-Order Logic. In an interpretation $\Im$, constants and variables stand for entities in the domain *D,* one-place predicates stand for subsets of *D,* and *n*-place predicates stand for *n*-place relations (sets of *n*-tuples) of entities in *D.* Below, $\Im^D(a)$, $\Im^D(b)$, $\Im^D(c)$, $\Im^D(x)$, $\Im^D(y)$, $\Im^D{}_{[x\rightarrow d]}(x)$, $\Im^D{}_{[y\rightarrow d']}(y)$ are members of *D*; $\Im^D(F)$ and $\Im^D(F)$ are subsets of D; and $\Im^D(R)$ is a set of pairs of (a two-place relation on) elements in D. A true formula corresponds to a relation that hold among the entities named by these basics parts of speech in combinations determined by the grammatical structure of the entire formula. Below are the examples of proven earlier of a P and its truth-conditions $TC_{\Im}(P)$. Note how the truth-conditions of P specify a "fact in the world" by referring only to the smallest referring expressions in *P.*

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

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

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

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

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

The Definition of Logical Concepts

We give a foretaste of the ideas in Part 3 and at the same time complete the standard set of definitions that constitutes the semantic theory of first-order logic by defining three logical concepts. These are the ideas of a "good argument", "necessary truth", and "consistent set".

The definitions and notation for validity and consistency are the same as that for sentential logic. Instead of calling a necessary truth a *tautology* as we do in propositional logic, it is the custom to call it a *truth of logic*.

Definitions

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

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

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

We simplify the notation $\{P_1,\ldots,P_n\} \models_L Q$ to $P_1,\ldots,P_n \models_L Q$, which is easier to read.

Summary

In this lecture we incorporated into the single language of *first-order logic* both the syllogistic's simple sentences and propositional logic's complex sentences. The language we developed goes well beyond both, however, in its expressive capacity because it allows both relational predicates, and multiple and nested quantification over complex sentential parts.

We stated a rigorous inductive definition of *well-formed formula*, which includes both open formulas with free variables and sentences without free variables. We also saw how to interpret quantified formulas intuitively in models described by Venn Diagrams.

We also continued the important theoretical work of the last lecture by extending the correspondence theory of truth to the new language. The inductive definition of interpretation incorporates the intuitive aspects of the definition of truth for the syllogistic in that it assigns predicates to sets and constants to elements of sets, both entities it is plausible to say make up "the world." The correspondence theory is extended to molecular sentences using the framework proposed by Tarski. The full definition of interpretation provides for each sentence, simple and complex, a set of truth-conditions that state conditions for the sentence's being true in terms of conditions that must hold "in the world" among the sets and their elements that are referred to by the sentences constants, variables, and predicates.

Finally, working out the truth-conditions of a formula provides a technique for showing that it is true in a given interpretation. Break down the formula's conditions into facts that must hold among the $\mathfrak{I}$-values of the terms and predicates that occur in its atomic parts. Then show that these fact obtain by appeal to the interpretation's definition, which spells out values it assigns to these terms and predicates.

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 $\forall\Im\ ((\Im(P_1)=T\&\ldots\&\Im(P_n)=T) \rightarrow \Im(Q)=T)$

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

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

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

(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	$((p_1$	$\vee$	$p_2)$	$\wedge$	$\sim$	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: $\{p_1 \rightarrow p_2, \sim p_1\} \not\models \sim p_2$.

Proof

	p_1	p_2	$((p_1$	$\rightarrow$	$p_2)$	$\wedge$	$\sim$	$p_1\,)$	$\rightarrow$	$\sim$	$p_2)$
$\mathfrak{I}_1$	T	T	T	T	T	F	F	T	T	F	T
$\mathfrak{I}_2$	T	F	T	F	F	F	F	T	T	T	F
$\mathfrak{I}_3$	F	T	F	T	T	T	T	F	F	F	T
$\mathfrak{I}_4$	F	F	F	F	F	F	T	F	T	T	F

. We may summarize these facts as follows: for any $\mathfrak{I}$,

$\mathfrak{I}(((p_1 \rightarrow p_2) \wedge \sim p_1\,) \rightarrow \sim p_2))$=T iff (($\mathfrak{I}(p_1)$=T and $\mathfrak{I}(p_2)$=T) or
($\mathfrak{I}(p_1)$=T and $\mathfrak{I}(p_2)$=F) or
($\mathfrak{I}(p_1)$=F and $\mathfrak{I}(p_2)$=F))

Also, for any $\mathfrak{I}$,

$\mathfrak{I}(((p_1 \rightarrow p_2) \wedge \sim p_1\,) \rightarrow \sim p_2))$=F iff ($\mathfrak{I}(p_1)$=F and $\mathfrak{I}(p_2)$=T)

Hence define $\mathfrak{I}(p_1)$=F and $\mathfrak{I}(p_2)$=T. Clearly such an $\mathfrak{I}$ exists (by construction) because we can define it. Then $\mathfrak{I}((p_1 \rightarrow p_2) \wedge \sim p_1\,) \rightarrow \sim p_2))$=F. Hence $\mathfrak{I}((p_1 \rightarrow p_2)$=T and $\mathfrak{I}(\sim p_1\,)$=T and $\mathfrak{I}(\sim p_2)$=F. Hence, $\exists \mathfrak{I}$, $\mathfrak{I}((p_1 \rightarrow p_2)$=T and $\mathfrak{I}(\sim p_1\,)$=T and $\mathfrak{I}(\sim p_2)$=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	$(((p_1$	$\rightarrow$	$p_2\,)$	$\wedge$	$p_2\,)$	$\rightarrow$	$p_1)$
$\mathfrak{I}_1$	T	T							
$\mathfrak{I}_2$	T	F							
$\mathfrak{I}_3$	F	T							
$\mathfrak{I}_4$	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	$\wedge$	$\sim$	$p_2))$
$\mathfrak{I}_1$	T	T	T	T	T	F	F	T	F	F	T
$\mathfrak{I}_2$	T	F	T	T	F	F	T	F	F	F	T
$\mathfrak{I}_3$	F	T	F	T	T	F	F	T	F	T	F
$\mathfrak{I}_4$	F	F	F	F	F	F	T	F	T	T	F

There is no $\mathfrak{I}$ such that $\mathfrak{I}(p_1\vee p_2)$=T and $\mathfrak{I}(p_1\vee p_2)$=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	$\vee$	$p_2))$
$\mathfrak{I}_1$	T	T									
$\mathfrak{I}_2$	T	F									
$\mathfrak{I}_3$	F	T									
$\mathfrak{I}_4$	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 $\forall\Im$ (($\Im(P_1)$=T&…&$\Im(P_n)$=T) $\rightarrow$ $\Im(Q)$=T)

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

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

{ $P_1,\ldots,P_n$} is consistent in L iff $\exists\Im$ ($\Im(P_1)$=T&…&$\Im(P_n)$=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 <u>underlined</u>, and the *then*-part deduced as its last line <u>double underlined</u>.

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.	$\underline{\Im(P_1)=T \ \&\ldots\&\ \Im(P_n)=T}$	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.	$TC_\Im(P_1)\ \&\ldots\&TC_\Im(P_n)$,	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.	$\underline{\Im(Q)=T}$	*modus ponens* on the previous two lines

End of subproof

$3n$+8. <u>If ($\Im(P_1)$=T &…& $\Im(P_n)$=T)</u> then <u>$\Im(Q)$=T</u>	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 $\models$

<u>Truth-Conditions.</u> The proof schema requires that we be able to plug in the truth-conditions $TC_\Im(P_1)\ \&\ldots\&TC_\Im(P_n)$ 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	$\underbrace{\mathrm{TC}_{\Im}(P)}_{\textbf{Truth-Conditions for } P}$
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 $d\in\Im^D(F)$ or $d\notin\Im^D(G)$
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 $<d,d'>\in\Im^D(R))$ or $<d',d>\notin\Im^D(R))$
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.

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

<u>*Everything is red and everything is round*</u> $\underline{\forall xFx \wedge \forall yGy}$

Everything is red and round	$\forall x(Fx \wedge Gx)$
<u>*Something is red and round*</u>	$\underline{\exists x(Fx \wedge Gx)}$
Something is red and something is round	$\exists xFx \wedge \exists yGy$

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

Somebody loves everybody	$\exists x\forall yLxy$
<u>*Love is reciprocal*</u>	$\underline{\forall x\forall y(Lxy \leftrightarrow Lyx)}$
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.	$\underline{\Im(\forall x(Gx \rightarrow Hx))=T \ \& \ \Im(\forall x(Fx \rightarrow Gx))=T}$	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.	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 $d \notin \Im^D(G)$	4,7, conjunction
9.	for all $d \in D$, either $d \in \Im^D(F)$ or $d \notin \Im^D(H)$	8, by set theory
10.	$\Im(\forall x(Fx \rightarrow Hx))=T$ iff for all $d \in D$, either $d \in \Im^D(F)$ or $d \notin \Im^D(H)$	TC 8
11.	$\underline{\Im(\forall x(Fx \rightarrow Hx))=T}$	9, 10 *modus ponens*

End of subproof

12.	If $\underline{(\Im(\forall x(Gx \rightarrow Hx))=T \ \&\ldots\& \ \Im(\forall x(Fx \rightarrow Gx))=T)}$ then $\underline{\Im(\forall x(Fx \rightarrow Hx))=T}$	1-11, CP
13.	$\forall\Im$(If $(\Im(\forall x(Gx \rightarrow Hx))=T \ \&\ldots\& \ \Im(\forall x(Fx \rightarrow Gx))=T)$ then $\Im(\forall x(Fx \rightarrow Hx))=T$)	12, universal generalization, $\Im$ 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.	$\underline{\Im(Fa)\text{=T} \;\&\; \Im(\exists xF)\text{=T}}$	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.	$\underline{\Im(\exists xFx)\text{=T}}$	5,6 *modus ponens*

End of subproof

8.	If $\underline{\Im(Fa)\text{=T}}$ then $\underline{\Im(\exists xFx)\text{=T}}$	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.	$\underline{\Im(\forall xFx\wedge\forall yGy)\text{=T}}$	Assump for CP, $\Im$ arbitrary
2.	$\Im(\forall xFx\wedge\forall yGy)$=T	1, conjunction
1.	$\Im(\forall xFx\wedge\forall yGy)$=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.	$\underline{\Im(\forall x(Fx\wedge Gy))\text{=T}}$	5,6 *modus ponens*

End of subproof

8.	If $\underline{\Im(\forall xFx\wedge\forall yGy)\text{=T}}$ then $\underline{\Im(\forall x(Fx\wedge Gy))\text{=T}}$	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.	$\underline{\Im(\forall x\forall y(Rxy\leftrightarrow Ryx))\text{=T}}$	Assump for CP, $\Im$ arbitrary
2.	$\Im(\forall x\forall y(Rxy\leftrightarrow Ryx))$=T	1, conjunction
2.	$\Im(\forall x\forall y(Rxy\leftrightarrow Ryx))$=T iff for all $d\in D$, for all d', $<d, d'>\Im^D(R)$ iff $<d',d>\Im^D(R)$	TC 13
4.	for all $d\in D$, for all d', $<d, d'>\Im^D(R)$ iff $<d',d>\Im^D(R)$	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.	$\underline{\Im(\forall xRxx)\text{=T}}$	5,6 *modus ponens*

End of subproof

8.	If $\underline{\Im(\forall x\forall y(Rxy\leftrightarrow Ryx))\text{=T}}$ then $\underline{\Im(\forall xRxx)\text{=T}}$	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
3.	$\Im(\exists x\forall yLxy)$=T iff for some $d\in D$, for all d', $<d, d'>\in\Im^D(L)$	TC 6
4.	for some $d\in D$, for all d', $<d, d'>\in\Im^D(L)$	2,3, *modus ponens*
5.	$\Im(\forall x\forall y(Lxy\leftrightarrow Lyx))$=T	1, conjunction
1.	$\Im(\forall x\forall y(Lxy\leftrightarrow Lyx))$=T iff for all $d\in D$, for all d', $<d, d'>\in\Im^D(L)$ iff $<d',d>\in\Im^D(L)$	TC 14
8.	for all $d\in D$, for all d', $<d, d'>\in\Im^D(L)$ iff $<d',d>\in\Im^D(L)$	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. $\forall x(Gx\rightarrow Hx)$, $\forall x(Fx\rightarrow Gx)$, $\exists xFx \not\models \exists x(Fx\wedge Hx)$ 19, definition of $\models$

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). $\forall x(Gx\rightarrow Hx), \forall x(Fx\rightarrow Gx), \exists xFx \models \exists x(Fx\wedge Hx)$

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. $\forall x(Gx\rightarrow Hx), \forall x(Fx\rightarrow Gx), \exists xFx \models \exists x(Fx\wedge Hx)$ — 16, definition of $\models$

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 $\forall\Im\ ((\Im(P_1)=T \& \ldots \& \Im(P_n)=T) \rightarrow \Im(Q)=T)$

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

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

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

(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	$((p_1$	$\vee$	$p_2)$	$\wedge$	$\sim$	$p_1)$	$\rightarrow$	$p_2)$
$\Im_1$	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. $\{P_1,\ldots,P_n\} \nvDash_L Q$ 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: $\{p_1\rightarrow p_2, \sim p_1\} \nvDash \sim p_2$.

Proof

	p_1	p_2	$((p_1$	$\rightarrow$	$p_2)$	$\wedge$	$\sim$	$p_1)$	$\rightarrow$	$\sim$	$p_2)$
$\mathfrak{I}_1$	T	T	T	T	T	F	F	T	T	F	T
$\mathfrak{I}_2$	T	F	T	F	F	F	F	T	T	T	F
$\mathfrak{I}_3$	F	T	F	T	T	T	T	F	F	F	T
$\mathfrak{I}_4$	F	F	F	F	F	F	T	F	T	T	F

. We may summarize these facts as follows: for any $\mathfrak{I}$,

$\mathfrak{I}(((p_1\rightarrow p_2)\wedge\sim p_1)\rightarrow\sim p_2))$=T iff (($\mathfrak{I}(p_1)$=T and $\mathfrak{I}(p_2)$=T) or
($\mathfrak{I}(p_1)$=T and $\mathfrak{I}(p_2)$=F) or
($\mathfrak{I}(p_1)$=F and $\mathfrak{I}(p_2)$=F))

Also, for any $\mathfrak{I}$,

$\mathfrak{I}(((p_1\rightarrow p_2)\wedge\sim p_1)\rightarrow\sim p_2))$=F iff ($\mathfrak{I}(p_1)$=F and $\mathfrak{I}(p_2)$=T)

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

Hence $\{p_1\rightarrow p_2, \sim p_1\} \nvDash \sim p_2$.

Exercise

Show Affirming the Consequent is invalid in propositional logic: $\{p_1\rightarrow p_2, p_2\} \nvDash p_2$.

	p_1	p_2	$(((p_1$	$\rightarrow$	$p_2)$	$\wedge$	$p_2)$	$\rightarrow$	$p_1)$
$\mathfrak{I}_1$	T	T							
$\mathfrak{I}_2$	T	F							
$\mathfrak{I}_3$	F	T							
$\mathfrak{I}_4$	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

$P_1,\ldots,P_n \vDash_L Q$ iff $\forall\Im\ ((\Im(P_1)=T\&\ldots\&\Im(P_n)=T\ ) \rightarrow \Im(Q)=T)$

$P_1,\ldots,P_n \nvDash_L Q$ iff $\exists\Im\ (\Im(P_1)=T\&\ldots\&\Im(P_n)=T\&\ \Im(Q)=F)\ )$

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

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

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

End of subproof

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

<u>Truth-Conditions.</u> The proof schema requires that we be able to plug in the truth-conditions $TC_\Im(P_1)\ \&\ldots\&TC_\Im(P_n)$ 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 $d \in \Im^D(F)$ or $d \notin \Im^D(G)$

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 $<d,d'> \in \Im^D(R))$ or $<d',d> \notin \Im^D(R))$

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*	$\sim\exists x(Gx\wedge Hx)$
Every F is G	$\underline{\forall x(Fx\rightarrow Gx)}$	*Every F is G*	$\underline{\forall x(Gx\rightarrow Hx)}$
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	$\forall x(Fx\rightarrow Gx)$
There exists an *F*	$\underline{\exists xFx}$
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	$\exists x(Fx \wedge Gx)$
Something is red and something is round	$\exists xFx \wedge \exists yGy$

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

Somebody loves everybody	$\exists x\forall yLxy$
Love is reciprocal	$\forall x\forall y(Lxy\leftrightarrow Lyx)$
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.	$\Im(\forall x(Gx\rightarrow Hx))$=T & $\Im(\forall x(Fx\rightarrow Gx))$=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.	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 $d\notin\Im^D(G)$	4,7, conjunction
9.	for all $d\in D$, either $d\in\Im^D(F)$ or $d\notin\Im^D(H)$	8, by set theory
10.	$\Im(\forall x(Fx\rightarrow Hx))$=T iff for all $d\in D$, either $d\in\Im^D(F)$ or $d\notin\Im^D(H)$	TC 8
11.	$\Im(\forall x(Fx\rightarrow Hx))$=T	9, 10 *modus ponens*

End of subproof

12.	If ($\Im(\forall x(Gx\rightarrow Hx))$=T &…& $\Im(\forall x(Fx\rightarrow Gx))$=T) then $\Im(\forall x(Fx\rightarrow Hx))$=T	1-11, CP
13.	$\forall\Im$(If ($\Im(\forall x(Gx\rightarrow Hx))$=T &…& $\Im(\forall x(Fx\rightarrow Gx))$=T) then $\Im(\forall x(Fx\rightarrow Hx))$=T)	12, universal generalization, $\Im$ 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. $Fa \vDash \exists xFx$

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.	$Fa \vDash \exists xFx$	9, definition of $\vDash$

Theorem. $\forall xFx \vDash \exists xFx$

***Exercise.** Prove the metatheorem $\forall xFx \vDash \exists xFx$.

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

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
4.	$\Im(\forall xFx\wedge\forall yGy)$=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.	$\forall xFx\wedge\forall yGy \vDash \forall x(Fx\wedge Gy)$	9, definition of $\vDash$

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

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

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
5.	$\Im(\forall x\forall y(Rxy\leftrightarrow Ryx))$=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.	$\forall x\forall y(Rxy\leftrightarrow Ryx) \vDash \forall xRxx$	9, definition of $\vDash$

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

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
6.	$\Im(\exists x\forall yLxy)$=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
2.	$\Im(\forall x\forall y(Lxy\leftrightarrow Lyx))$=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

22.	$\Im(F)=\varnothing$ & $\Im(G)=\varnothing$ & $\Im(H)=\varnothing$	Def of $\Im$ (principle of abstraction)
23.	$\Im(G)=\varnothing$ & $\Im(H)=\varnothing$	1, conjunction
24.	$\Im(G)\subseteq\Im(H)$	2, set theory
25.	for all $d\in D$, if $d\in\Im^D(G)$ then $d\in\Im^D(H)$	3, def of $\subseteq$
26.	for all $d\in D$, either $d\in\Im^D(G)$ or $d\notin\Im^D(H)$	4, logic
27.	$\Im(\forall x(Gx\rightarrow Hx))$=T iff for all $d\in D$, either $d\in\Im^D(G)$ or $d\notin\Im^D(H)$	TC 8
28.	$\Im(\forall x(Gx\rightarrow Hx))$=T	5,6 *modus ponens*
29.	$\Im(F)=\varnothing$ & $\Im(G)=\varnothing$	1, conjunction
30.	$\Im(F)\subseteq\Im(G)$	8, set theory
31.	for all $d\in D$, if $d\in\Im^D(F)$ then $d\in\Im^D(G)$	9, def of $\subseteq$
32.	for all $d\in D$, either $d\in\Im^D(F)$ or $d\notin\Im^D(G)$	10, logic
33.	$\Im(\forall x(Fx\rightarrow Gx))$=T iff for all $d\in D$, either $d\in\Im^D(F)$ or $d\notin\Im^D(G)$	TC 8
34.	$\Im(\forall x(Fx\rightarrow Gx))$=T	11,12 *modus ponens*
35.	$\Im(F)=\varnothing$ & $\Im(H)=\varnothing$	1, conjunction
36.	for all $d\in D$, $d\notin\Im^D(F)$ or $d\notin\Im^D(H)$	14, set theory and logic
37.	$\Im(\exists x(Fx\wedge Hx))$=T iff for some $d\in D$, $d\in\Im^D(F)$ and $d\in\Im^D(H)$	TC 8
38.	$\Im(\exists x(Fx\wedge Hx))\neq$T iff for all $d\in D$, $d\notin\Im^D(F)$ or $d\notin\Im^D(H)$	16, logic
39.	$\Im(\exists x(Fx\wedge Hx))\neq$T	15, 17 *modus ponens*

40. $\Im(\forall x(Gx\rightarrow Hx))$=T & $\Im(\forall x(Fx\rightarrow Gx))$=T& $\Im(\exists xFx)$=T and $\Im(\exists x(Fx\wedge Hx))$=F) 7,13,18 conjunction
41. $\exists\Im(\Im(\forall x(Gx\rightarrow Hx))$=T & $\Im(\forall x(Fx\rightarrow Gx))$=T& $\Im(\exists xFx)$=T and $\Im(\exists x(Fx\wedge Hx))$=F) 19, construction
42. $\forall x(Gx\rightarrow Hx)$, $\forall x(Fx\rightarrow Gx)$, $\exists xFx \not\models \exists x(Fx\wedge Hx)$ 19, definition of $\models$

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). $\forall x(Gx\rightarrow Hx)$, $\forall x(Fx\rightarrow Gx)$, $\exists xFx \models \exists x(Fx\wedge Hx)$

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. $\forall x(Gx\rightarrow Hx)$, $\forall x(Fx\rightarrow Gx)$, $\exists xFx \models \exists x(Fx\wedge Hx)$ 16, definition of $\models$

LECTURE 6. PROPOSITIONAL AND FIRST-ORDER LOGIC: PROOF THEORY

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.

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: [22]

   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*.
   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

[22] 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).

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:

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

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)[23], 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

[23] 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, 1965). An excellent introduction to first-order logic using natural deduction is Neil Tennant, *Natural Logic* (Edinburgh: Edinburgh University Press, 1978).

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.

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 ($P \wedge \sim P$) and indirect proof (i.e. proving P by showing $\sim P$ 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.[24] We shall continue to make use of sets in semantics and shall continue to use all the traditional logical rules.

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

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.

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 $\perp$, and a rule called *thinning* that adds extra premises to an argument. The new symbol $\perp$ 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 $\perp$ 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 $<X,\sim\sim P>$ 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 *~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 $<X \cup Y,Q>$.

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:

$\perp$ 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 $<X,P \rightarrow Q>$ construct $<X \cup Y,P>$. (*Modus Ponens*)

Thinning. From $<X,P>$ construct $<X \cup Y,P>$. (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 $<X,P>$ is constructed by one of the rules a-k in *PR* from elements of *DPL*, then $<X,P>$ is in *DPL*.
   c. Nothing else is in *DPL*.

It is customary to write the fact that the deduction $<\{P_1\ldots,P_n\},Q>$ is "acceptable", i.e. that $<\{P_1\ldots,P_n\},Q> \in DPL$, 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 $<\varnothing,P> \in DPL$. 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.	$\langle \{P, P\rightarrow Q, \sim Q\}, \bot \rangle$	3,4 $\bot$+
6.	$\langle \{P,\sim Q\}, \sim(P\rightarrow Q) \rangle$	5 $\sim$+

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

Proof

1.	$\langle \{\sim Q, P\land Q\}, P\land Q \rangle$	bd
2.	$\langle \{\sim Q, P\land Q\}, Q \rangle$	1 $\land$-
3.	$\langle \{\sim Q, P\land Q\}, \sim Q \rangle$	bd
4.	$\langle \{\sim Q, P\land Q\}, \bot \rangle$	2,3 $\bot$+
5.	$\langle \{\sim Q\}, \sim(P\land Q) \rangle$	4 $\sim$+

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

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

Proof

1. $\langle \{P,\sim(P\lor P)\}, \sim(P\lor P) \rangle$
2. $\langle \{P,\sim(P\lor P)\}, P \rangle$
3. $\langle \{P,\sim(P\lor P)\}, P\lor\sim P \rangle$
4. $\langle \{P,\sim(P\lor P)\}, \bot \rangle$
5. $\langle \{\sim(P\lor P)\}, \sim P \rangle$
6. $\langle \{\sim(P\lor P)\}, P\lor\sim P \rangle$
7. $\langle \{\sim(P\lor P)\}, \sim(P\lor\sim P) \rangle$
8. $\langle \{\sim(P\lor P)\}, \bot \rangle$
9. $\langle \varnothing, \sim\sim(P\lor\sim P) \rangle$
10. $\langle \varnothing, P\lor\sim P)$

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

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

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:[25]
    1. $p_1\to(p_2\to p_1)$
    2. $(p_1\to(p_2\to p_3))\to((p_1\to p_2)\to(p_1\to p_3))$
    3. $(\sim p_1\to\sim p_2)\to(p_2\to p_1)$
    4. $\forall x(p_1\to p_2)\to\forall xp_1\to\forall\ p_2$
    5. $\forall xp_1\to p_1$
    6. $p_1\to\forall xp_1$ if the formula replacing p_1 contains no free x
    7. $x=x$
    8. $x=y\to(p_1[x]\to p_2[y])$ 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.*

[25] 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.

Definition. Q is (*syntactically*) *derivable* from $\{P_1,\ldots,P_n\}$ (abbreviated $P_1,\ldots,P_n \vdash Q$) iff $Q \in ThPL'$ $\langle AxPL \cup \{P_1,\ldots,P_n\}, PR, ThPL' \rangle$ 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.[26] If P is a theorem, then $\forall xP$ is a theorem.

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

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

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

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

∗*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:

⊥ Rules:
- Introduction. From <*X*,*P*> and <*Y*,~*P*> construct <*X*, ⊥>. (This "explains" the meaning of ⊥)
- Elimination. From <*X*,⊥> construct <*X*–{~*P*},*P*>. (A version of *Ex Falso Quodlibet*)

~ Rules:
- Introduction. From <*X*,⊥> construct <*X*–{*P*},~*P*>. (Reduction to the Absurd)
- Elimination. From <*X*,~~*P*> construct <*X*,*P*>. (Double Negation)

∧ Rules:
- Introduction. From <*X*,*P*> and <*Y*,*Q*>, construct <*X*∪*Y*,*P*∧*Q*>.
- Elimination. From <*X*,*P*∧*Q*>. construct <*X*,*P*>.
- Elimination. From <*X*,*P*∧*Q*>. construct <*X*,*Q*>.

∨ Rules:
- Introduction. From <*X*,*P*> construct <*X*,*P*∨*Q*>. (Addition, to the right side)
- Introduction. From <*X*,*Q*> construct <*X*,*P*∨*Q*>. (Addition, to the left side)
- Elimination. From <*X*,*P*∨*Q*>, <*Y*,*R*> and <*Z*,*R*>, construct <*X*∪(*Y*–{*P*})∪(*Z*–{Q}), *R*>. (Argument from cases)

→ Rules:
- Introduction. From <*X*,*P*> construct <*X*–{Q},*Q*→*P*>. (Conditional Proof)
- Elimination. From <*X*,*P*> and <*X*,*P*→*Q*> construct <*X*∪*Y*,*P*>. (*Modus Ponens*)

∀ Rules:
- Introduction. From *X*,P[*t*/*v*]> construct <*X*,∀*vP*>. (Universal Generalization)
- Elimination. From <*X*,∀*vP*> construct <*X*,*P*[*t*/*v*]> where *v* is not free in any *P*∈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,$P[t/v]$>. (if t is not free in X,Y, $\exists v$ ′P or Q) (Existential Instantiatiion)

= 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*.

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. $P_1,\ldots,P_n \models Q$ iff $P_1,\ldots,P_n \vdash Q$.
9. Argument Soundness and Completeness

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

Exercise. Provide grammatical derivations (construction sequences) showing that the following are in *Sen:*

1. $\sim(p_1\leftrightarrow\sim p_1)$
2. $((p_1\leftrightarrow p_2)\leftrightarrow((p_1\rightarrow p_2)\wedge(p_2\rightarrow p_1)))$
3. $(p_1\rightarrow(p_1\vee(p_2\wedge\sim p_2)))$

Exercise. Analyze the following sentences *P* like the previous example:

(a) for all possible interpretations of the sentence's atomic parts, provide a construction sequence that is parallel to the sentence's grammatical derivation,

(b) summarize the information from the construction sequences in a traditional truth-table for the sentence,

(c) summarize the truth-conditions $TC_{\mathfrak{I}}(P)$ for *P.*

8. $\sim(p_1\leftrightarrow\sim p_1)$ [two possible interpretations]
9. $\sim\sim(p_1\vee\sim p_1)$ [two possible interpretations]
10. $\sim(p_1\leftrightarrow\sim p_2)$ [four possible interpretations]
11. $(((p_1\rightarrow p_2)\wedge\sim p_2))\rightarrow\sim p_1)$ [four possible interpretations]
12. $(((p_1\rightarrow p_2)\wedge p_2))\rightarrow p_1)$ [four possible interpretations]
13. $((p_1\rightarrow p_2)\leftrightarrow(\sim p_2\rightarrow\sim p_1))$ [four possible interpretations]
14. $((p_1\leftrightarrow p_2)\leftrightarrow((p_1\rightarrow p_2)\wedge(p_2\rightarrow p_1)))$ [four possible interpretations]

Exercise. For the sentences below construct their truth-table only, without first producing the construction sequences for the sentence itself and its interpretations.

1. $(p_1\rightarrow(p_1\vee(p_2\wedge\sim p_2)))$ [four possible interpretations]
2. $((p_1\leftrightarrow p_2)\leftrightarrow((p_1\rightarrow p_2)\wedge(p_2\rightarrow p_1)))$ [four possible interpretations]
3. $(\sim(p_1\wedge p_2)\leftrightarrow(\sim p_1\vee\sim p_2))$ [four possible interpretations]
4. $((p_1\wedge(p_2\vee p_2))\rightarrow((p_1\wedge p_2)\vee(p_1\wedge p_3)))$ [eight possible interpretations]

Exercises. Construct a grammatical derivation for each of the following showing that they are elements of *For:*

1. $\forall x\forall y\forall z((Hxy\wedge Hyz)\rightarrow Hxz)$
2. $\forall x\forall y((x{=}y\wedge Fx)\rightarrow Fy)$
3. $\sim\exists yFy\rightarrow\forall x(\sim Hx\vee\sim Fx)$

Exercises

1. Construct a Venn diagram showing that the sentences are all true:
   a. $\forall x(Fx \rightarrow Gxy)$
   b. $\exists x(Gx \wedge Hx)$
   c. $\sim\exists x(Fx \wedge Hx)$
2. Construct Venn diagram showing that $\forall x(Fx \rightarrow \exists y(Lxy))$ can be true but $\exists y \forall x(Fx \rightarrow Lxy))$ false.
3. Symbolize in the notation of first-order logic the syllogism Bramantip (**AAI** in the fourth figure). Construct a Venn diagram showing that in modern notation it is invalid because in the diagram the premises are true but the conclusion is false.
4. Construct an arrow diagram in which the relation *same size as,* represented by the letter *S,* is reflexive, transitive and symmetric.

Exercises

1. Annotate each line of the Example 10 and 11, repeated below, citing either the equivalence E1-E10 that it instantiates, or the number of previous line and the equivalence E1-E10 from which it is derived by the substitution of equivalents, or the numbers of the previous line from which it is derived by the substitution of identity.

Example 10. $\forall x(Fx \wedge Gx)$

1. $\Im(\forall x(Fx \wedge Gx))$=T	iff	for all $d \in D$, $\Im^{D}_{[x\rightarrow d]}(Fx \wedge Gx)$=T
2.	iff	for all $d \in D$, $\Im^{D}_{[x\rightarrow d]}(Fx)$=T and $\Im^{D}_{[x\rightarrow d]}(Gx))$=T
3.	iff	for all $d \in D$, $\Im^{D}_{[x\rightarrow d]}(x) \in \Im^{D}(F)$ and $\Im^{D}_{[x\rightarrow d]}(x) \in \Im^{D}(G)$
4.	iff	for all $d \in D$, $d \in \Im^{D}(F)$ and $d \in \Im^{D}(G)$

Example 11. $\exists x(Fx \rightarrow Gx)$

1. $\Im(\exists x(Fx \rightarrow Gx))$=T	iff	for some $d \in D$, $\Im^{D}_{[x\rightarrow d]}(Fx \rightarrow Gx)$=T
2.	iff	for some $d \in D$, either $\Im^{D}_{[x\rightarrow d]}(Fx)$=T or $\Im^{D}_{[x\rightarrow d]}(Gx)) \neq$T 1
3.	iff	for some $d \in D$, either $\Im^{D}_{[x\rightarrow d]}(x) \in \Im^{D}(F)$ or $\Im^{D}_{[x\rightarrow d]}(x) \notin \Im^{D}(G)$
4.	iff	for some $d \in D$, either $d \in \Im^{D}(F)$ or $d \notin \Im^{D}(G)$

2. Work out the truth-conditions with annotation for the two new examples, call them examples 15 and 16:

Example 15

1. $\Im(\exists xFx \wedge \exists yGy))$=T iff
2.
3.
4.

5.
Example 16
1. $\Im(\forall xFx \to \forall yGy)$=T iff
2.
3.
4.
5.

Exercise. Prove that if D={1,2,3} $\Im(F)$={1}, $\Im$(G)={2,3}, then

1. $\Im(\forall x(Fx \vee Gx))$=T
2. $\Im(\exists x(Gx \wedge \sim Fx))$=T

Prove 1 by first calculating $TC_{\Im}(\forall x(Fx \vee Gx))$ by progressive applications of the earlier metatheorem, as in the previous example. Prove 2 by first calculating $TC_{\Im}(\exists x(Gx \wedge \sim Fx))$.

Exercise

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

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

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

Exercise

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

	p_1	p_2	$(((p_1$	$\to$	$p_2)$	$\wedge$	$p_2)$	$\to$	$p_1)$
$\Im_1$	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	$\vee$	$p_2))$
$\mathfrak{I}_1$	T	T									
$\mathfrak{I}_2$	T	F									
$\mathfrak{I}_3$	F	T									
$\mathfrak{I}_4$	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 metatheorem in the semantics for first-order logic:

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

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

Exercise. Annotate the following proof.

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

Proof

11. $< \{P, \sim(P \vee P)\}, \sim(P \vee P) >$
12. $< \{P, \sim(P \vee P)\}, P >$
13. $< \{P, \sim(P \vee P)\}, P \vee \sim P >$
14. $< \{P, \sim(P \vee P)\}, \perp >$
15. $< \{\sim(P \vee P)\}, \sim P >$
16. $< \{\sim(P \vee P)\}, P \vee \sim P >$
17. $< \{\sim(P \vee P)\}, \sim(P \vee \sim P) >$
18. $< \{\sim(P \vee P)\}, \perp >$
19. $< \varnothing, \sim\sim(P \vee \sim P) >$
20. $< \varnothing, P \vee \sim P >$

Review Questions

1. Inductive sets. We have met several sets that have inductive definitions: the set of sentences in propositional logic, the set of formulas in first-order logic, and each interpretation in propositional and first-order logic. (Recall that an interpretation is a set because it is a set of pairs, i.e. a two place relation that pairs an expression with its referent or its truth-value.) See if you can identify for each of these sets the *basic elements* used to start building the set, and *the construction rules* used to add new members from those already in the set. Also, see if you can explain what a *construction sequence* is for each of these sets and what they are used for.

2. Truth. In Part 2 we have seen how the correspondence theory of truth is applied of a variety of sentences, simple and complex. The format use for the definition is to define of each sentence type P a truth-condition rule, called a (T) rule, for the form:

   $\Im(P)$=T iff ____________

   Here the ___________ is filled with the truth-conditions of P, briefly summarized as TC(P). These conditions spell out what has to be true in the world of $\Im$ for P to be true. Be able to discuss the rule as it applies to

   a. the sentences of propositional logic,
   b. the formulas of first-order logic.

It is not easy to explain what the truth-conditions for *P* should be if *P* is a complex sentence, which are (made up of the connectives ~, ∧, ∨, →, and ↔, or of quantifiers ∀ and ∃). In what sense does the (T) rule apply to complex sentences and formulas? For simplicity you may ignore the quantifiers ∀ and ∃ and limit your answer to the case of complex sentences made up from the connectives ~, ∧, ∨, →, and ↔.

3. 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."
4. Give an example of showing that it is possible to prove from the axioms of set theory and the definition of $\Im$ 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).
5. 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.)

6. 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.

APPENDIX I. NAÏVE SET THEORY

Axioms

Logical Truth.	*Every truth of logic is a theorem.*
Extensionality.	$A=B \leftrightarrow \forall x\,(x\in A \leftrightarrow x\in B)$
Abstraction.	$\exists A\ \forall x\,(x\in A \leftrightarrow P(x))$

Abbreviations

$x\neq y$	$\sim(x=y)$
$x\notin A$	$\sim(x\in A)$
$A\subseteq B$	$\forall x(x\in A\rightarrow x\in B)$
$A\subset B$	$A\subseteq B\,\&\sim A=B$
$\varnothing$ or Λ	$\{x \mid x\neq x\}$
V	$\{x \mid x=x\}$
$A\cap B$	$\{x \mid x\in A\,\&\,x\in B\}$
$A\cup B$	$\{x \mid x\in A\vee x\in B\}$
$A-B$	$\{x \mid x\in A\,\&\,x\notin B\}$
$-A$	$\mathrm{V}-A$
$\mathbf{P}(A)$	$\{B \mid B\subseteq A\}$
$\{x_1, \ldots, x_n\}$	$\{y \mid y = x_1 \vee \ldots \vee y = x_n\}$

Theorems

1. $\forall y(\ y\in\{x \mid P(x)\} \leftrightarrow P(y))$
2. $\forall x\,(x\in\varnothing \leftrightarrow x\neq x)$
3. $\forall x\,(x\in \mathrm{V} \leftrightarrow x=x)$
4. $\forall x\,(x\in A\cap B \leftrightarrow (x\in A\,\&\,x\in B))$
5. $\forall x\,(x\in A\cup B \leftrightarrow (x\in A\vee x\in B))$
6. $\forall x\,(x\in A-B \leftrightarrow (x\in A\,\&\,x\notin B))$
7. $\forall x\,(x\in -A \leftrightarrow x\notin A)$
8. $\forall B\,(B\in \mathbf{P}(A) \leftrightarrow B\subseteq A)$
9. $\forall y\,(y\in\{ x_1, \ldots, x_n \} \leftrightarrow (y = x_1 \vee \ldots \vee y = x_n))$
10. $--A=A$
11. $A\subseteq A$
12. $\forall x((x\in A \ \&\ A\subseteq B)\rightarrow x\in B)$
13. $A\cap A=A=A\cup A$
14. $A=B \leftrightarrow (A\subseteq B \ \&\ B\subseteq A)$
15. $A\cap B\subseteq A\subseteq A\cup B$
16. $\varnothing\subseteq A\subseteq \mathrm{V}$
17. $-(A\cup B)=-A\cap -B$
18. $-(A\cap B)=-A\cup -B$
19. $A\subseteq B \leftrightarrow -B\subseteq -A$
20. $A\subseteq B \leftrightarrow \sim(A\cap -B \neq\varnothing)$
21. $\exists x(x\in A\cap B) \leftrightarrow \sim(A\cap B\neq\varnothing)$
22. $A\in\mathbf{P}(A)$
23. $\varnothing\in\mathbf{P}(A)$
24. $\sim\exists A\ \forall x\,(x\in A \leftrightarrow P(x))$
25. $\forall x_1,\ldots,x_n, (<x_1,\ldots,x_n>=<y_1,\ldots,y_n> \leftrightarrow \forall\ x_1,\ldots,x_n\forall\ y_1,\ldots,y_n\ (x_i = y_i))$
26. $(R\subseteq \mathrm{V}^n \ \&\ S\subseteq \mathrm{V}^n) \rightarrow (R=S \leftrightarrow \forall x_1,\ldots,x_n(<x_1,\ldots,x_n>\in A \leftrightarrow <x_1,\ldots,x_n>\in B))$
27. $\exists A\ \forall\ x_1,\ldots,x_n\ (<x_1,\ldots,x_n>\in A \leftrightarrow P(x_1,\ldots,x_n))$