Logic for Liberal Arts Students

Part 2. The Logic of Propositions

Lai d'Aristote

According to the version of Henri d'Andeli (c. 1250)[1]

Aristotle, tutor and counselor to Alexander the Great, sought to separate the youthful monarch from his paramour--now usually known as Phyllis--who was absorbing all his time and energy, and causing him to neglect his political duties. Reluctantly, Alexander agreed to the separation, but soon revealed the fact to Phyllis. She thereupon contrived a scheme to nullify Aristotle's influence, aiming to regain her lover's attentions. The plan was simple. Early in the morning, when good scholars should be laboring at their books, Phyllis slipped into the garden next to Aristotle's study and, not far from his open window, she softly sang and danced. Her hair was loose, her feet were bare, her belt was off her gown. Aristotle heard her song, and then he turned to look at her: "that made him close his books and cry: 'Oh God !'," it being clear that the deity invoked was Eros. When Phyllis came close enough to the window, Aristotle reached out and seized her firmly. He told her of his ardent wish; she promised to fulfill it, if he would first satisfy a trifling whim of hers. He must pretend to be a horse, get on all fours, wear a saddle, and let her ride around the garden on his back. The besotted Aristotle did exactly what was asked, yielding up an image that approached the essence of burlesque. "In this was grammar betrayed and logic much dumb-founded," remarked the commentator in *Le Livre de Leesce* (c. 1373). Riding on the Master's back, Phyllis loudly sang a song of triumph: Master Silly carries me.

'Love leads on, and so he goes,
by Love's authority'.

The song was a signal for Alexander to look into the garden from his window. "Master, can this be?" he called, going on to question Aristotle's flagrantly quadruped behavior. The old sage answered that there was a lesson to be learned from his example. If a wise philosopher, aged and grey as he, is unable to resist the power of Love, then Alexander, yet youthful and hot blooded,

[1] Ayers Bagley (University of Minnesota), "Study of Love: Aristotle's Fall." (See http://education.umn.edu/EdPA/iconics/Lecture_Hall/aristotle.htm)

must be immeasureably more cautious in exposing himself to such danger. Amused by the sophistical defense, Alexander forgives Aristotle's ridiculous indiscretion and then, presumably, reunites himself with Phyllis. The philosopher would trouble them no more, having lost his credibility.

Table of Contents

Part 2. The Logic of Propositions

INTRODUCTION

The language used in logical reasoning is constructed of various levels. We start with terms; we use terms to construct sentences; and we use sentences to construct arguments. In Part 1 we studied the logic of terms. We studied what terms refer to and concluded that they stand for sets and their elements. In Part 2 we shall study sentences. Our goal is to explain what it is for sentences to be true or false. In Part 3 we shall go on to see how it is we organize sentences into logical arguments.

In Part 1 we have already discussed to some extent what it is for a simple subject-predicate sentence *S is P* to be true. We invoked the so-called correspondence theory of truth, which holds that *S is P* is true in a world if what the element *S* stands for is in the set *P* stands for. But this account is really over-simplified because it ignores the complexities of grammar. The sentences we use in science and every day language are really much more varied than *S is P.* In Part 2 we extend the simple insights of Part 1 to a more complete picture of language. When we finish, we will be able to sketch what it is for a wide variety of sentences to be true or false. Indeed the formal language we will have explained, though not as rich as real English, is adequate for expressing most of mathematics and the natural sciences.

Plato: Discourse as the Interweaving of Nouns and Verbs

The grammar we know from high school was essentially an invention of the ancient Greeks. It began earlier with rudimentary distinctions drawn by philosophers like Plato and Aristotle, who in the course of their philosophical writings on other matters distinguish grammatical ideas like *noun*, *verb* and *sentence*. By the end of the classical period grammarians like Donatus (fl 350 A.D.) and Priscian (fl. 500) were able to give summaries of Greek and Latin grammar that would be familiar to students today. Donatus, for example, in his *Ars minor* distinguishes the eight parts of speech: nouns (which include adjectives), pronouns, verbs, adverbs, participles, conjunctions, prepositions, and interjections. He also summarizes rules for declining nouns and conjugating verbs.

But in its early stages grammar was studied in the service of logic. Plato, for example, is one of the first to discuss the form of the simple subject-predicate sentence. In the following passage from the *Sophist* the Stranger, who is the spokesman for Plato's views in the dialogue, remarks on the fact that a simple assertion requires a noun and a verb:

STRANGER: …[T]here is no expression of action or inaction, or of the existence of existence or non-existence indicated by the sounds, until verbs are mingled with nouns; then the words fit, and the smallest combination of them forms language, and is the simplest and least form of discourse.

THEAETETUS: Again I ask, What do you mean?

STRANGER: When any one says 'A man learns,' should you not call this the simplest and least of sentences?

THEAETETUS: Yes.

STRANGER: Yes, for he now arrives at the point of giving an intimation about something which is, or is becoming, or has become, or will be. And he not only names, but he does something, by connecting verbs with nouns; and therefore we say that he discourses, and to this connexion of words we give the name of discourse.

STRANGER: And as there are some things which fit one another, and other things which do not fit, so there are some vocal signs which do, and others which do not, combine and form discourse.

He goes on to say that sentences of this sort are either *true* or *false* depending on whether they accurately describe "what is." By "what is" here Plato is referring to the reality that the sentence either corresponds to when it is true or does not correspond to when it is false. He explains his own views of what this reality in the theory of Forms. It consists of material objects imitating the Forms. As discussed in an earlier lecture, this metaphysics is not very plausible, but from the perspective of logic, one of its serious weaknesses of the theory of is its lack of attention to grammar. Plato's rudimentary description of sentences fails to say enough about them to allow us to discuss their differences in grammatical structure.

Logical relations are formal. In key cases the fact that a sentence of a particular grammatical form is true or false forces another sentence of a different form to be true or false, depending on the details of their respective forms. Another name for form is grammatical structure. That is, logical relations hold among sentence types because of differences in their grammatical structure. It follows that no account of logical relations is even possible until grammar is sufficiently well developed to distinguish among the different kinds of grammatical sentences. It is Aristotle who first pushed the science of grammar to the point that it can divide sentences into different grammatical kinds that stand in logical relations to one another. Indeed, it is fair to say that in making grammatical distinctions, Aristotle's primary motivation is logical. For him grammar is a necessary prerequisite to logic.

Aristotle's Categorical Propositions

Aristotle develops his account of the grammar and truth-conditions of sentences in the *De Interpretatione.*

He distinguishes between, on the one hand, the impressions imposed on the soul by natural processes of perception. These mental impressions are what strictly speaking form the language with which we think. He contrasts these with the spoken words of audible language. Whereas mental impression are determined by fixed natural processes, the sounds we attach to them in spoken language are a matter of convention. He thus lays the foundation for the later mediaeval theory of mental language.

Nouns and verbs, he says, are the basic words used to make up mental complexes called *propositions.* It is propositions that are either true or false because it is these that either correspond or fail to correspond to "what is." Different propositions do this differently. It is the job of his grammatical theory and the associated account of truth-conditions to explain these differences, and the logical relations among propositions that result. Today we would distinguish between *grammar*, which studies the differences in syntactic forms, from *semantics*, which explains truth-conditions and used them to define logical relations. In outlining Aristotle's ideas here, we shall try to follow this modern division.

Nouns and verbs, he says, are predicated or "said of" things in the world in the sense that they truly apply to some objects but not others. (Nouns and verbs in this broad sense include all the nine different sorts of predicates he distinguishes in the *Categories.*) Like later grammarians of Greek and Latin, he does not distinguish between nouns and adjectives the way we do in English because in the classical languages adjectives have the same form as nouns and can in fact be used as nouns. For example, in Latin nominalized adjective constructions like *the beautiful* (*illud pulcrum*), *all rationals* (*toti rationales*), and *five whites* (*quinque albi*) may all function as nouns. Aristotle recognizes that nouns may be negated by the addition of a negative particle or prefix, as we do in English by attaching the prefix *un-* to an adjective. We use a negated noun (adjective) as a predicate that applies to those things that the un-negated predicate is not. For example, *unjust* is true of those objects that *just* is not true of, and vice versa.

Nouns may also be modified by the words that declare how many of those objects that the subject term is predicated of are actually relevant to the true of a proposition. The role of this sort of modifier, Aristotle says, is to indicate the term's *quantity*. Today in fact we call such modifiers *quantifiers*. Although there are many ways to limit the relevant quantity, Aristotle distinguishes just two. The first consists of modifiers, like *all* and *every* in English, that indicate that *all* the objects the noun is true of are relevant. Nouns modified this way are said to be *universal*. The second sort consists of modifiers, like *some* and *at least one* in English, that indicate that *some* or, more precisely, that *at least one but not necessarily all* the objects the noun is true of are relevant. Nouns modified this way are said to be *particular*. In Greek as in English only common or collective nouns (for example *horse*) can be modified by quantifiers like *every* and *some.* Some common nouns, however, occur in sentences without any explicit modifying quantifier, as in the proposition *birds fly.* Aristotle calls such nouns *indefinite* and, somewhat arbitrarily, understands them to have an implicit quantifier *some* understood*.* That is, indefinite nouns, he says, are a special case of particular nouns. Thus, on his account *birds fly* should be understood to means *some birds fly.* Proper nouns in Greek and English behave differently from common nouns. For one thing, it is ungrammatical to modify them with quantifiers like *every* and *some.* To provide a unified account, however, he treats proper nouns as a special case of common nouns. They are common nouns that just happen to apply to just one object. He calls propositions that contain a proper noun singular. Moreover, he understands proper nouns to be modified by an unstated but implicit universal quantifier. Note that the quantifier in, say, *every Socrates,* does not need to be stated because the noun only applies to one thing and *all* one is one. The quantifier on a proper noun would be redundant. In Aristotle's terminology, the quantity of a proposition's subject determines the quantity of the proposition as a whole. Thus, a proposition with a subject term modified by *every* is universal, and one with a subject modified by *some* is particular.

Verbs are distinguished from nouns by the fact that they have a *tense* indicated special inflected endings (or in English by helping verbs) that indicate the time at which the predication is true. It is either present, past, or future. There are also linking two "nouns" by the verb *to be,* as in *Socrates is wise* or *Socrates was hungry*. This use of *to be* to form a simple nominal propositions came to be called the *copula*, and it too has special forms depending on its tense. According to Aristotle, then, a verb is both similar to and different from a noun: (1) a verb predicates something, a feature it shares with a noun, and (2) it indicates a time of predication, a feature it shares with the copula, but not with a noun.

The simplest proposition, which he calls an *affirmation*, consists of conjoining two predicative words by a “third thing” that indicates the time of predication. The first predicative word is required to be a noun, and is called the *subject term*. The second may be either a noun (including adjectives) or a verb, and is called the *predicate term*. If the subject is a noun and the predicate a verb, then the verb incorporates the time indicator. If the predicate is a noun (or an adjective), then the “third thing” is the verb *to be* functioning as the copula (*De Interpretatione,* 16b24, 19b23). The affirmation is *true* if some or all, as indicated by the subject term’s quantity, of the objects that the subject term is predicated of are also, at the time indicated, among those that the predicate term is predicated of. It is false otherwise. Hence Aristotle accepts the logical law called the *law of excluded middle*: every proposition is either true or false.[2]

Verbs too may be negated. As in English, this may be done in one of two ways. The first is to insert a negative word in front of the verb, as we do in English by using *not*, as in *Socrates is not green*. The second way is to precede the entire affirmation with a negative term, as we do in English with phrase *it is not the case that*, as in *it is not the case that Socrates is green.* This is called a *sentential* negation. The two negations say the same thing. In either form the role of a verbal negation is to affect the proposition as whole, changing it from an affirmation to a *denial*. A denial is true if an affirmation is false, and vice versa. That is, the denial is *true* if it is not the case at the indicated time that the predicate is true of the quantity of subject entities indicated. Denials are said to be *negative,* and whether a proposition is affirmative or negative declares its *quality*.

In sum, propositions possess one of two quantities: they are affirmative or negative. They possess one of two quantities: they are universal or particular. As in the table below, the four types universal affirmative, universal negative, particular affirmative, and particular negative are traditionally labeled respectively A, E, I, and O propositions.

	affirmative	negative
universal	A. *Every man is rational*	E. *No man is rational*
particular	I. *Some man is white*	O. *Some man is not white*
singular	A. *Socrates is mortal*	E. *Socrates is not mortal*
indefinite	I. *Man is just*	O. *Man is not just*

The abbreviations A and E come from the Latin word *affirmo*, meaning *I affirm*; and the E and O from *nego,* which means *I deny.*[3]

Using sentence negation, it is possible to formulate for each universal form a proposition that is its logical equivalent.

Every F is G *It is not the case that some F is not G*

---

[2] In *De Interpretatione* IX Aristotle discusses the unusual case of propositions about the future (i.e. future tense propositions) that are contingent in the sense that they are not necessarily true or false, like *there will be a sea battle tomorrow*. These he suggests may lack a truth-value today.

[3] In the *Port Royal Logic*, Arnauld and Nicole provide a Latin mnemonic poem for students to memorize:

Asserit A, negat E, verum generaliter ambo,
Asserit I, negat O, set particulariter ambo.

[A asserts the truth, E denies it, both do so in a general way. But I asserts it, O denies it, both in a particular way.]

No F is G *It is not the case that some F is G*

Likewise for each negative form, it is possible to formulate a logically equivalent that is the sentence negation of an affirmation.

No F is G *It is not the case that some F is G*
Some F is not G *It is not the case that all F is G*

Syntax and Semantics for the Syllogistic

Let us now use modern logic to make the theory shorter and clearer.[4] We need just a few symbols and definitions. So that the symbolism throughout the lectures matches that of modern logic, we will use upper case letters *F, G,* and *H* to represent Aristotle's predicative terms. Following the terminology of traditional logic, we shall simply call these letters *terms*. These include for him not only common nouns, adjectives, and verbs, but also proper nouns. Though in the *Categories* Aristotle spells out ontological distinctions corresponding to different varieties of predicative terms, in the *De Interpretatione* and his other logical works these are irrelevant. All that is important is that, as a result of natural processes, a term is truly predicated of a specific group of objects at a specific time. As determined by the propositions

[4]It was only after many years that modern logicians, schooled in formal logic and mathematics, turned back to the syllogistic to explore its logical properties from a modern perspective. Early studies, most importantly Łukasiewicz, Jan. *Aristotle's Syllogistic* (Second ed. Oxford: Clarendon Press, 1957, and Sheperdson, J. C. "On the Interpretation of Aristotelian Syllogistic" (*Journal of Symbolic Logic* 21 (1956): 137-47) were followed by the important work of Timothy Smiley and John Corcoran, who brought the syllogistic within the modern framework of natural deduction and set out its core properties. (See Smiley, Timothy. "Syllogism and Quantification." *Journal of Symbolic Logic* 27 (1962), 58-72, "What is a Syllogism?" *Journal of Philosophical Logic* 2 (1973), 136-154, and Corcoran, John. "Completeness of an Ancient Logic." *Journal of Symbolic Logic* 37 (1972), 696-702, and "Aristotle's Natural Deduction System," in J. Corcoran (ed.) *Ancient Logic and its Modern Interpretation* (Dordrecht: Reidel, 1974), 85-131.) Except as noted in footnotes the formal framework presented in these lectures is set forth more fully in the author's papers "Aristotle's Natural Deduction Reconsidered" (*History and Philosophy of Logic* 18 (1997): 1-15) reprinted in John N. Martin, *Themes in Neoplatonic and Aristotelian Logic* (Aldershot: Ashgate, 2004), and "Ecthesis and Existence in the Syllogistic," in the Ashgate volume.

quantifier, only a stated quantity of this group is relevant to determining a proposition's truth. To avoid the long-winded locutions that Aristotle needed when talking about predicates being true of members of this group, modern logicians usually summarize his view in terms of sets. That is, we shall abbreviate his more long-winded view by saying that a predicate stands relative to a time for a set of objects that exist at that time.

To start with, we shall make some simplifying assumptions. First, although we shall shortly allow for genuine sentential negations – those that negate the verb or the entire proposition – we will postpone until later the introduction of term negations like *unjust*, which are prefix to the term itself. Also, although Aristotle allows that terms may at times not truly stand for anything, i.e. that they represent in modern terms an empty set, to start with we will first make the simplifying assumption that all terms stand for non-empty sets.

Since there are only four basic propositional types – A , E, **I**, and O – we shall use bold face versions of these letters themselves to represent the propositional structure. We will write the subject in front of the predicate[5] and prefix both by one of the letters **A** , **E**, **I**, and **O** (called *operators*) indicating the sort of proposition it is. That is, if *X* and *Y* are terms, **A***XY,* **E***XY,* **I***XY,* **O***XY* will be propositions. In this notation,

A*XY*	means	*Every X is Y,*
E*XY*	means	*No X is Y,*
I*XY*	means	*Some X is Y*
O*XY*	means	*Some S is not Y*

[5] Aristotle himself placed the predicate in front of the subject. Thus he preferred to write *Every F is G, which* is the more natural order in Greek as well as English, in the reverse order, as *G is predicated of every F.*

Because Aristotle uses this syntax to study the logic of a special set of arguments called syllogisms, which we shall study in Part 3, the syntax is called that of "the syllogistic."

Formal Syntax

By *the syllogistic syntax SSyn* is meant the pair <*Trms,Prps*> such that

1. The set *Trms* of *terms*:

   $Trms = \{F, G, H, \ldots\}$

2. The set *Prps* of *propositions:*

   $Prps = \{\mathbf{Z}XY \mid \mathbf{Z} \in \{\mathbf{A},\mathbf{E},\mathbf{I},\mathbf{O}\} \;\&\; X \in Trms \;\&\; Y \in Trms\}$

If $\mathbf{Z}XY$ is a proposition, then X is its *subject term* and Y is its *predicate term*. A proposition starting with the operators **A** or **E** is called *universal;* one starting with **I** or **O** is called *particular;* one starting with **A** or **I** is called *affirmative;* and one starting with **E** or **O** is called *negative.*

We introduce sentential negations by means of abbreviate definitions:

Definitions:

$\sim \mathbf{A}XY$	means	$\mathbf{O}XY$
$\sim\mathbf{E}XY$	means	$\mathbf{I}XY$
$\sim\mathbf{I}XY$	means	$\mathbf{E}XY$
$\sim\mathbf{O}XY$	means	$\mathbf{A}XY$

Semantic Intuitions

We will now define what it is to "give an interpretation," which we shall call $\Im$, to the terms and propositions of the syntax in a universe U representing the objects that exist . $\Im$ will assign every predicate F, G, H, etc. a subset of objects in U. These are the objects that the predicate is true of in that universe. We will use the notation $\Im(F)$

for the set that *F* stands for. Further, $\mathfrak{I}$ will give each proposition a truth-value T (for *true*) or F (for *false*) depending on the proposition type and the objects its terms pick out. Using the notation of set theory, it is a simple matter to state the appropriate truth-conditions for **A**, **E**, **I**, and **O** propositions.

We should pause briefly to comment on the truth-conditions for the universal propositions form *Every F is G* and *No F is G.* Traditional logic, prior to the advent of modern symbolic logic in the 19th century, regularly assumed that for these propositions to be true the terms had to stand for at least one existing thing. Quite reasonably from the perspective of common sense, they assumed that there is no point to talking about empty sets. In Aristotelian "science," moreover, it was an assumption that the genera and species that were the subject of scientific study never came into or passed out of existence. In other words, they are never empty. At other times however he does talk about terms – his example is *goatstag* – that are not true of anything. Hence when an "empty term" is used as the subject of a universal proposition, that proposition is traditionally viewed to be false. We should note, however, that modern logic rejects this assumption because it uses the empty set and has explored its properties. (There are true and useful things to be said about it.) In the initial presentation below, however, we shall follow a narrower Aristotelian practice and restrict the language to terms that refer to existing things.

To make the statement of truth-conditions more intuitive, we shall illustrate them by means of what are called Venn diagrams. These were invented by the English logician John Venn (1843-1942). Such a diagram is used to represent a world in which the terms of the language are "interpreted." In modern logic such a world is called a *model* for the set of sentences that are true in that world. In the diagram a rectangle represents the set *U* of all entities that exist in that world. This set is called the *domain* of the model. Circles are drawn within this rectangle each of which is labeled by a term. A circle labeled by a term *T* represents the set of entities named by *T*. A small **x** in a circle represents an individual element (a "thing") that is in that set. If a region of the diagram is shaded, this fact indicates that the corresponding region of the set is "empty," i.e. has no elements. More precisely, the rules governing the interpretation of a Venn diagram may be are stated as follows:

1. The outside rectangle represents the domain *U* of all existing things in a given "world," and the circles represent the sets of objects assigned by an interpretation $\mathfrak{I}$ to the terms of the language.
2. If a region is shaded, then it is empty, containing no objects. That region is an empty set.
3. If a region contains an **x**, then it is not empty, and the **x** represents an object that exists in that region.
4. If a region is neither shaded nor occupied by an **x**, but then the diagram declares that the region is either empty or non-empty, that as far as the diagram is concerned, it is unknown which. It may be either empty or not.
5. If an **x** is drawn on the line between two regions, then the diagram declares that it is known only that one of the two bordering regions is not empty, but that it is not known which it is. That is, the **x** on the line is an object in one of the regions, but as far as the diagram is concerned, it is not known which of the two regions it is in.

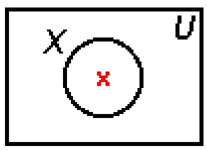
In the diagrams we incorporate the usual assumption of traditional logic that for a universal affirmative to be true its subject term must stand for at least one object. This assumption is rejected, however, by modern logic, and in later lectures we will invent languages in which terms do stand for the empty set. In cases in which Aristotle's logic but not modern logic assumes that there is an entity in a set that entity will be drawn in by use of a small red **x**.

Formal Semantics

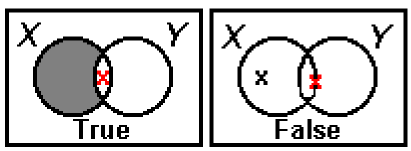
The semantics defines the notion of an interpretation as a "pairing" that assigns to each term a set and to each proposition a truth-value.

Definition. Let *SI* be the set of all *interpretation* $\Im$ for *SSyn* relative to a domain *U* that meet these conditions: $\Im$ is a function (set of pairs) that pairs a term in *Trms* to a non-empty subset of *U* and that pairs a proposition in *Prps* to one of the two truth-values T or F, and is such that, for any terms *X* and *Y*:

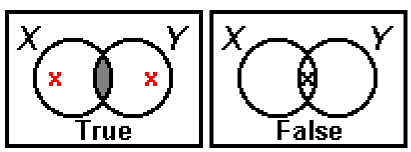
1. $\Im(X)\neq\varnothing$ & $\Im(X)\subseteq U$



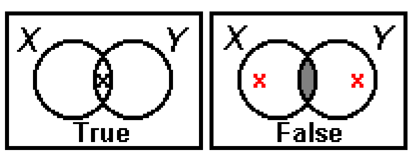
2. $\Im(\mathbf{A}XY)=\mathrm{T} \leftrightarrow \Im(X)\subseteq\Im(Y)$



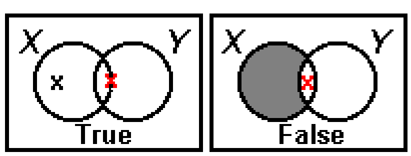
3. $\Im(\mathbf{E}XY)=\mathrm{T} \leftrightarrow \Im(X)\cap\Im(Y)=\varnothing$



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



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



Let us identify the *syllogistic language SL* with <*SSyn,SI*>, i.e. with the “structure” formed by joining the syntax of the language with the set of its interpretations. (This use of pointy-brackets here to define the "language" as a special ordered pair has no special significance other than the fact that it is the custom in modern algebra to describe "structures" as ordered *n*-tuples. The notation is used here because there is a sense in which a language is a "structure," namely one made up set *SSyn* and *SI*.) We shall use the symbol $\Im$ to stand for interpretation in *SI.*

Exercise. Translate each of the following English sentence into A, E, I, or O propositions, with or without predicate negations, using the predicates F and G. First indicate for each example what word or words the letters *F* and *G* represent.

1. *All humans are mortal.*
2. *Some cows fly.*
3. *There are green monkeys.*
4. *There are some monkeys that do not swim.*
5. *There are no pink elephants.*
6. *Students are poor.*

Correspondence Theory of Truth

The main lesson to carry away from this introduction to the simple language of **A**, **E**, *I* and **O** propositions is that it provides a clear model for what a correspondence theory of truth should aspire to. Each of the four proposition types is provided with a semantic rule that states the exact conditions in which is it true, its 'truth-conditions." These conditions lay out "what must be happening in the world" – i.e. the fact it must correspond to – if the proposition is true. These conditions moreover are facts about the sets referred to by the proposition's terms. That is, to say a proposition corresponds to the world means that certain conditions obtain in the world among the referents of the proposition's terms. It is these referents that are the entities that the proposition talks about and it is some relation among them that the proposition asserts holds in fact. In more complex languages that we shall meet in later lectures conditions on the referents of the expressions that make up a proposition will be taken as a mark of a correspondence theory of truth. Let $X_1,\ldots, X_n$ be the referring parts of a sentence *P,* and let $\Im(X_1),\ldots,\Im(X_n)$ be the referents of these terms as stipulated in an interpretation $\Im$. Further let TC(*P*) be a set of conditions $\Im(X_1),\ldots,\Im(X_n)$ that hold among the referents $\Im(X_1),\ldots,\Im(X_n)$. Because $\Im(X_1),\ldots,\Im(X_n)$ are all entities "in the world" these conditions amount to some fact that must obtain in that world. Normally in modern logic these conditions are formulated in the language of set theory. That is, TC(*P*) is a set of conditions formulated in the language of set theory that must hold among the referents $\Im(X_1),\ldots,\Im(X_n)$ of the parts $X_1,\ldots, X_n$ of *P.* For example the TC(**A***XY*), i.e. the condition for $\Im(\mathbf{A}XY)$=T, is $\Im(X)\subseteq\Im(Y)$ as summarized in the rule:

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

Here $\Im(X)\subseteq\Im(Y)$ states a condition on the referents $\Im(X)$ and $\Im(Y)$ of the parts *X* and *Y* of **A***XY* for this proposition to be true. Likewise, TC(**E***XY*), i.e. the condition for $\Im(\mathbf{E}XY)$=T, is $\Im(X)\cap\Im(Y)=\varnothing$, as summarized in the rule:

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

Here $\Im(X)\cap\Im(Y)=\varnothing$ states a condition on the referents $\Im(X)$ and $\Im(Y)$ of the parts *X* and *Y* of **E***XY*. Similarly the clauses in the definition of $\Im(\mathbf{I}XY)$=T and $\Im(\mathbf{O}XY)$=T state the TC(**I***XY*) and TC(**O***XY*), in the rules:

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

and

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

In later chapters, therefore, we shall take as a mark of a correspondence theory of truth that its definition of an interpretation $\Im$ should provide a rule of the form:

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

such that TC(P) states a set a condition formulated in the language of set theory that must hold among the referents $\Im(X_1),\ldots,\Im(X_n)$ of the parts $X_1,\ldots, X_n$ of P. The elements of a correspondence theory as exemplified by the syllogistic may be summarized as follows:

Format of a Correspondence Theory of Truth:

P is true in $\Im$	**iff**	**the world is the way *P* says it is**
$\Im(\mathbf{A}XY)=\mathrm{T}$	iff	$\Im(X)\subseteq\Im(Y)$
$\Im(\mathbf{E}XY)=\mathrm{T}$	iff	$\Im(X)\cap\Im(Y)=\varnothing$
$\Im(\mathbf{I}XY)=\mathrm{T}$	iff	$\Im(X)\cap\Im(Y)\neq\varnothing$
$\Im(\mathbf{O}XY)=\mathrm{T}$	iff	$\Im(X)-\Im(Y)\neq\varnothing$
$\Im(P)=\mathrm{T}$	iff	$\mathrm{TC}(P)$

The truth conditions of *P* in $\Im$
Set theoretic conditions of the referents in $\Im$ of the parts of *P*

Summary

Though limited in its expressive power the language of the syllogistic is interesting for a number of reasons. It provides our first example of a carefully defined syntax. It provides as a set of grammar rules and formally precise definitions of the two parts of speech of its grammar: terms and propositions, which are what sentences are traditionally called in this theory.

It also provides our first example of a rigorously defined notion of an interpretation of a syntax understood as an assignment of "meanings" to the parts of speech. Relative to an interpretation $\Im$, a term is assigned a unique set, and a proposition is assigned a unique truth-value. Set theoretically an interpretation is understood to be a set of pairs. The fact that a term *X* is paired with sets *A* in $\Im$ is expressed as $<X,\mathrm{A}>\in\Im$, or equivalently as $\Im(X)=A$. The fact that the proposition *P* is assigned in $\Im$ the truth-value *V* is expressed as $<P,V>$ or equivalently as $\Im(P)=V$.

The definition of interpretation is also our first example of a rigorously defined correspondence theory of truth. For each sentence in the syntax, the definition of $\Im$ provides a principle of the form

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

in which TC(*P*) expresses the "truth-conditions" of *P* . These are the conditions under which *P* is true in $\Im$. They are written set theory, and they state what must hold in that case among the $\Im$-values of the parts of *P* , i.e. among "the referents of the parts" of *P.*

*Predicate Negation[6]

Extension of the Syntax and Semantics

As defined previously the definitions of the syntax and semantics do not allow for nominal negations of the sort Aristotle describes in which a negative particle is attached directly to an adjective, as in the English *unjust.* We now incorporate this additional feature to the language we have already defined. Let us use the symbolism $\bar{F}$ to represent the nominal negation *un-F.* Semantically $\bar{F}$ stands for the complement of the set that *F* stands for. Formally, when we add negations of this sort to the syntax, we are expanding the set *Trms* to a larger set, which we shall call $Trms^+$. In its turn, the new set of terms generates a larger set $Prop^+$ of propositions, because there is now a new variety of terms that can be used to make propositions that were not there before. The new set of terms and propositions are then used to define an "enlarged" syntax $SSyn^+$ as $< Trms^+, Prop^+ >$. In the semantics we will have to augment the definition of an interpretation so that in addition to the sets it assigns to non-negated terms, it will pair a set with each term negation. The rule we shall add is that the extension of a negated predicate is the complement of its un-negated form: $\Im(\bar{F})=U-\Im^+(F)$. We collect the extended interpretations into a set we shall call SI^+, and define a new "extended" language SL^+, namely $< SSyn^+, SI^+>$. Lastly, since a language's set of valid argument is defined in terms of its interpretations, we must revise its definition too in terms of the new set. Accordingly we defined a new relation

[6] This and subsequent section is not required for an understanding of later lectures. The material here rounds out the semantics of the syllogistic by providing additional distinctions necessary for a full

$\vDash_{SL^+}$ that holds between the premises and conclusion of a valid argument in the expanded language.

Definitions

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

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

Exercise. Translate each of the following English sentence into A, E, I, or O propositions, with or without predicate negations, using the predicates F and G. First indicate for each example, what word or words the letters *F* and *G* represent.

1. Fred is unfriendly.

2. No unprepared student passes.

Terms with Empty Extensions

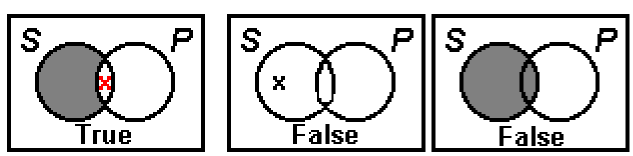
Though we have extended the theory to allow for negative terms, we have yet to allow for the possibility that terms stand for empty sets. Modern logic uses this

comparison syllogistic to modern semantics, especially as its concerns the logic of negation and the existential presuppositions of terms.

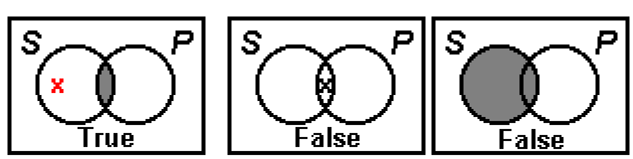
option whenever it talks about the empty set, and Aristotle himself recognized that it was sometimes useful to talk, and form logical arguments, using terms that do not stand for anything. His example is *goatstag.* To incorporate this possibility we retain the syntax $SSyn^{+}$ that includes negative predicates, but alter the definition of an interpretation so that it may assign the empty set to a term. The new set of interpretation will be called $SI^{+\varnothing}$, the language that uses these interpretations will be called $SL^{+\varnothing}$, and the logical entailment relation for this language is called $\models_{SL^{+\varnothing}}$.

Definitions

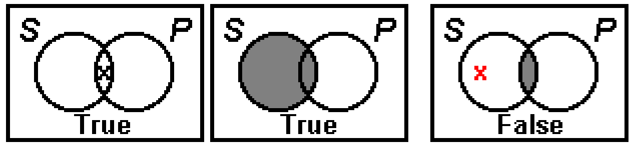
1. $Trms^{+} = Trms \cup \{\ \bar{X} \mid X \in Trms\}$
2. $Prop^{+}=\{\mathbf{Z}XY \mid \mathbf{Z}\in\{\mathbf{A},\mathbf{E},\mathbf{I},\mathbf{O}\}\ \&\ X\in Trms^{+}\ \&\ Y\in Trms^{+}\}$
3. The set $SI^{+\varnothing}$ is the set of all *possibly empty interpretation* $\Im$ for $SSyn^{+}$ relative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



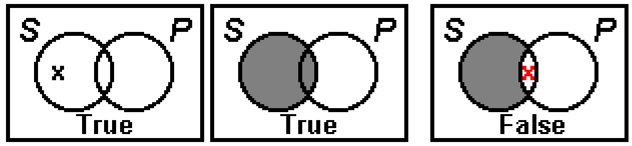
   d. $\Im(\mathbf{E}XY)=\mathrm{T} \leftrightarrow \Im(X)\cap\Im(Y)=\varnothing$



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



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



4. $SL^{+\varnothing}$, the *enlarged syllogistic language*, is $< SSyn^{+}, SI^{+\varnothing} >$

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

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

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

E. Tapissier (© Roger Viollet)

Buridan's Ass

Buridan, the scholastic, used to say that an ass placed at an equal distance between two haystacks, perfectly equal between the two, having no reason to decide for one rather than the other, would die of hunger between the two.

Émile Littré, *Dictionnaire de la langue française*, 1863.

> Between two kinds of food, distant and attractive
> In equal measure, a free man would die of hunger,
> Before he would bring one of them to his teeth.

Dante, *Paradiso* IV

If man does not act from free will, what will happen to him if he finds himself in equilibrium like Buridan's ass?

Spinoza, *Ethics* II, 49.

Ancient and Mediaeval Logic

Simple and Complex Sentences

Aristotle investigated valid arguments that turn on the internal grammar of simple sentences. Such is the categorical logic of the Square of Opposition and of syllogisms. There are however "compound sentences" and "complex sentences," to use the terms given them in high school grammar, and these generate valid arguments that turn on their grammatical forms.

Compound sentences of this sort are formed from simpler sentences by the use of "conjunctions." The traditional list of "coordinating conjunctions" that yields "compound" sentences is *and, but, or, for,* and *yet.* Complex sentences are said to be formed by joining a dependent clause to a main clause, both sentences, by adverbial conjunctions like *although, then, since, because,* etc. We met in Lecture 4 various valid inferences that turn on the grammar of compound sentences when we made use of these patterns as inferences rules in the axiomatization of naïve set theory. For example, we saw there that if a complex sentence of the form P&*Q* is true, then the sentence of the form Q&*P* must be true.

For the purpose of constructing logical arguments, it is sufficient to restrict their attention to a small number of "conjunctions." Aristotle recognizes the existence of grammatically complex propositions (*De Interpretatione* V, 17a9) and on some level he was aware that some logical arguments turn on their structure. In the *Prior Analytics* for example he makes systematic use of the argument form called *reductio per impossible*:

$$\frac{(P \,\&\, Q) \rightarrow R}{\therefore (P \,\&\, {\sim}R) \rightarrow {\sim}Q}$$

But he does not remark on the fact that in addition to arguments like syllogisms, which turn on the internal structure of simple propositions, there is another sort of valid argument that depends on the grammar of complex sentences.

Hypothetical Propositions

Later logicians, both ancient and mediaeval, did remark on inferences of this type. In the tradition, the grammatical distinction is made in terms of a contrast between categorical and hypothetical positions. We have already met the four types of categorical propositions. There are also *hypothetical* propositions, defined as those that are formed from two categorical propositions. These were traditionally divided into three kinds: (1) *conjunctions* formed by joining two propositions by *and,* (2) *disjunctions* formed by joining two propositions by *or,* and (3) *conditionals* formed by joining two propositions by using the word if.[7] Because in logic the term *conjunction* is reserved for compounds formed by *and, a* new term is needed for the entire group of connecting words that includes *and, or,* and *if…then*. In modern logic, we also include two other connectives: *negation* and the *biconditional.* Negation is marked in English by the word *not* attached to the verb or by sentence prefixes like *it is not the case that.* The

[7] See for example Peter of Spain, *Summa Logicales*, I, 16. (Francis P. Dinnen, trans., *Peter of Spain: Langauge in Dispute* (Amsterdam: John Benjamins, 1990).

biconditional is expressed in English by the words *if and only if*. As understood in modern logic, connectives form more complex sentences from shorter sentences, which may themselves be either simple or complex. Complex sentences are also called *molecular* and simple sentences are called *atomic.* Molecular sentences can be very complex indeed.

Until the advent of modern logic in the 19th century, however, the logic of the connectives was restricted to that of simple compounds of categorical propositions and was not highly developed. There was little discussion of the fact that connectives can be nested inside one another or that there are logical inferences that hold due to very complex structure. The theory of the connectives, such as it was, consisted mainly in noting some simple logical inference patterns, some of which we have already met, like *modus podendo ponens, modus tollendo tollens*, disjunctive syllogism (*modus tollendo ponens*), and hypothetical syllogism. These were studied – and named – by ancient Stoic logicians.[8] The study of the inference patterns, which were called *consequentia*, consisted mainly of collecting lists of them without much attempt to explain why these rather than other patterns were logically valid.[9]

[8] See I. M. Bocheński, *A History of Formal Logic,* Second ed. (Notre Dame Indiana: University of Notre Dame Press, 1961) (Reprinted by Chelsea Publishing Co.), and Benson Mates, *Stoic Logic* (Berkeley: University of California Press, 1953).

[9] One of the better collections of this sort is John Buridan's *Treatise on Consequences* (early 14th century), which does remark that on the fact that some inference patterns can be shown to be valid on the basis of others. The treatment, however, is not very systematic by modern standards. It does not, for example, provide an axiomatization of the recognized consequences. See Peter King, ed. and trans, *John Buridan's Logic: The Treatise on Supposition and the Treatise on Consequences* (Dordrecht: Reidel, 1985.)

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. Gottleb Frege invented his own symbolization, called the *begriffsschrift* ("concept writing") in his groundbreaking set theoretic axiomatization of arithmetic (1879).[10] The standard modern symbolization began in 19th century studies of arithmetic by Dedekind and Peano[11]. 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 exceeds other notation in its simplicity.

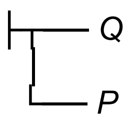
Frege's notation was designed for use in an axiom system. A formula starts with a short vertical line, |, indicating that the formula that follows it to the left is a theorem. The formula then continues with a horizontal line. The horizontal is an assertion sign. It indicates that the formula that follows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

[10] 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)
[11] Guiseppe Peano, *Arithemetices Principia, Nova Method Exposita* (Turin: Fratres Bocca, 1889)

of two sentences *P* and *Q,* Frege joins them one to the left of the other connected by a horizontal line:

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

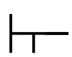
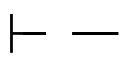
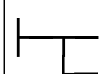
is read, *It is a theorem that it is true that it is not the case that it is true that P*. He indicates the conditional *if P then* Q by subjoining an assertion of the antecedent *P,* namely —— *P,* to that of *Q,* —— *Q,* by a vertical line. Thus



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 placed to the left of the formulas it joins and no parentheses are used. Thus K*PQ* 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.*[12] Though up to this point we have been using the ampersand & for conjunction[13], from now on 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.[14]

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

[12] 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 give a list, without implying that the items in the list are mutually exclusive. That is, *P vel Q* tends to mean *P or Q or possibly both,* which is the more useful meaning of *or* in logic.
[13] 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 .
[14] 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. ~(*P*∨~*Q*)∧*R* is dual to (and hence logically equivalent to) ~((~*P*∧*Q*)∨~*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.[15]

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.[16]

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

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

common in philosophy that defines a 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 a 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 will 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

1. *ASen,* called the set of *atomic sentences, is* a subset of $\{p_1,\ldots,p_n,\ldots\}$;
2. *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. $fr_{\leftrightarrow}(x,y)= (x\leftrightarrow y)$
3. *Sen* is the set such that
    a. ASen is a subset of *Sen;*

b. 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;*

c. 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 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. $((\sim p_4 \vee p_2) \wedge p_4)$
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 proven by producing a grammatical derivation (construction sequence) for each:

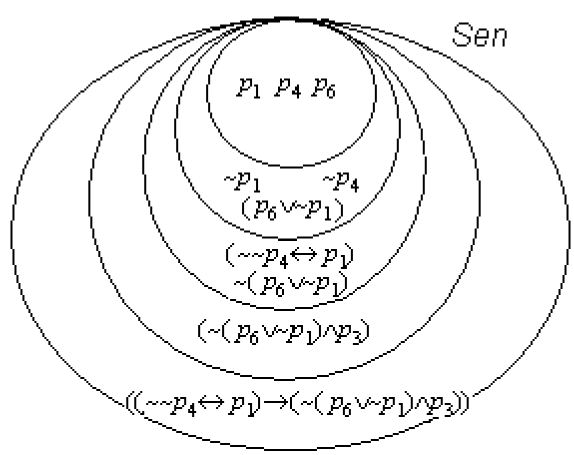
1. p_2	atomic
2. p_4	atomic
3. $\sim p_4$	2, $fr_{\sim}$
4. $(\sim p_4 \vee p_2)$	2 & 1, $fr_{\vee}$
5. $((\sim p_4 \vee p_2) \wedge \sim p_4)$	4 & 3, $fr_{\wedge}$

1. p_2	atomic
2. $\sim p_2$	1, $fr_{\sim}$
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, $fr_{\leftrightarrow}$
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 $\quad T \longrightarrow F$
$\qquad\qquad\quad F \longrightarrow T$

P	$\sim P$
T	F
F	T

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

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

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:

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

Translating English using ~

The symbol ~ reverses the truth-value of a sentence. It corresponds in English to several negative expressions:

1. verbal negations, e.g. *cows do* ***not*** *fly, Socrates in* ***not*** *wealthy*
2. sentence negations, e.g. ***it is not the case that*** *cows fly,* ***it is not true that*** *Socrates is ugly*
3. negative predicate prefixes, e.g. *huskies are* ***un****friendly,* √2 *is* ***ir****rational*.

Of the three, predicate negation is the most tricky. You can only translate *huskies are unfriendly* by ~(*huskies are friendly*) in only contexts in which everything you are talking about is either friendly or unfriendly, Likewise, you can only translate √2 *is irrational* as ~(√2 *is rational*) in contexts in which it assumed that everything is either a rational or an irrational number.

The next set of tables sets out the rules for the two place connectives. In these two truth-values in a given order are paired with a unique truth-value.

Disjunction

The rule for disjunction is:

Disjunction

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

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	$tf_\vee(F,F)=F$

Translating English using $\vee$

The symbol $\vee$ captures the so-called *inclusive disjunction* used asserts that one, or the other, or both of two alternative is true. Very often it fits our use of the word *or.* There are however cases in which we intend the so-called *exclusive disjunction* used to assert the truth of that one or the other of two alternatives, but not both. If you say *I would like milk or lemon in my tea*, you do not mean both – because the milk might curdle. (As we shall check shortly, the way to symbolize *P or Q but not both* is $\sim(P\leftrightarrow Q)$ because $\sim(P\leftrightarrow Q)$ has the right truth-values: $\sim(P\leftrightarrow Q)$ is T if either *P* is T or *Q* is T, but $\sim(P\leftrightarrow Q)$ is F if *P* and *Q* are either both T or both F.)

Conjunction

The next rule is that for conjunction:

Conjunction

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

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:

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

Translating English using $\wedge$

The symbol $\wedge$ captures the idea that both conjuncts are true. We sometimes express this idea in English by the term *and,* but it is only appropriate in cases in which it is only the truth-values of the conjuncts that it conveys. But sometimes in English more information is asserted by *and* than just the truth-value of the parts. Sometimes, for example the temporal order of the conjunctions maters, as in *I filled the pool with water and dove off the high dive.*

If however, we abstract away from any information other than the truth-value of the parts, then $\wedge$ is the appropriate translation. There are many so-called "contrastive" conjunctions and adverbs in English that normally assert that both parts are true, but imply as well that the two stand in some contrasting way to one another. These include *but, yet, moreover, although, however, when,* and many others. If however it is appropriate to abstract away from that fact that the two components contrast with one another, as is normally done when translating into symbolic notation, then it is appropriate to translate these too by $\wedge$.

The Conditional

The next rule is that for the conditional:

The Conditional

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

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:

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

Translating English using $\rightarrow$

The truth-table for $\rightarrow$ captures a rather specialized sense of *if…then,* one that is really not common in English. Normally in ordinary contexts when we say *if then* we intend to impart more information than simply that the truth-value of *P* is F or that of *Q* is T, which is all that is asserted by $P \rightarrow Q$. Normally we intend to say that there is some connection between the fact indicated by *P* and that indicated by *Q* such that there is some kind of law or rule such that the truth-value of *Q* depends on that of *P*. To make this distinction clear, let us distinguish three senses of *if...then.*

The Material Conditional. This is the sense captured by the connective $\rightarrow$. Only truth-values are relevant to calculating the truth of $P \rightarrow Q$. It asserts merely that, as a matter of fact, the component sentences *P* and *Q* have the correct truth-values. Exactly what values *P* and *Q* must have for $P \rightarrow Q$ to be true may be summarized in several different but equivalent ways:

1. If *P* is T, then *Q* is T.[17]
2. Either *P* is false, or *Q* is T.
3. It is not the case both that *P* is T and *Q* is false.

Given that every sentence is T or F, all three of these formulations correctly summarize the truth-table for $P \rightarrow Q$.

The material conditional has two somewhat disconcerting features. The first is that the conditional may be true even though there is no connection between the antecedent and the consequent, e.g. because *Egypt is in Africa* is T and *the drinking age in Ohio is* 21 is T, it follows that this conditional is T: *Egypt is in Africa* $\rightarrow$ *the drinking age in Ohio is* 21. The material conditional also has the feature that a conditional with a false antecedent is automatically true, e.g.

the moon is made of green cheese$\rightarrow$ 3 *is the greatest prime number* is T because its antecedent is false. It is this latter feature that makes the empty set a subset of every set: since $x\in\varnothing$ is F, $x\in\varnothing \rightarrow x\in A$ is T.

The material conditional is adequate for most uses of *if…then* in technical work in mathematics and the mathematical sciences. It is for this reason that it is used in symbolic logic. It is, for example, the sense of *if…then* used in the axiomatization of set theory in Part 1, and in the laws of logic used there. For example, the logical laws *modus ponens, modus tollens*, hypothetical syllogism, implication, and transposition are all valid if written in terms of $\rightarrow$. It is for this reason that $\rightarrow$ is adequate for representing *if..then* in contexts in which the primary motivation is to represent the fact that the words *if..then* obey the laws of logic. As a general rule we shall use $\rightarrow$ to translate *if…then* because it is our purpose to represent uses of *if…then* that conform to the rules of logic.

Note that *only if* in English is used to express the converse of *if.* Hence, *if P then Q, Q only if P,* and $P\rightarrow Q$ all say the same thing.

There are two important senses of *if…then* that are not adequately translated by $\rightarrow$.

Logical Entailment. Sometimes if we say *if P then Q* we mean that *Q* "follows logically" from *P,* or that the argument from *P* to Q is logically "valid." For example, we may assert *if* $p_1\wedge p_2$ *then* p_1 to convey the information that the sentence $p_1\wedge p_2$ logically implies the sentence p_1. This sense of *if…then* is called *logical entailment* or *logical implication.* It asserts not only that as a matter of fact

---

[17] Note that the use of *if…then* here is itself a case of the material conditional, but in the

$P \rightarrow Q$, but in addition that $P \rightarrow Q$ *must be* the case. The idea may be formulated in terms of interpretations. No matter what the interpretation (truth-value) given to the atomic formulas in P and Q, it must be the case that $P \rightarrow Q$ is T. This happens due to the internal grammar of *P* and *Q.* Their grammar is such that if the atomic sentences in *P* are assigned values that make *P* true, they also assign values that make *Q* true. In logic we symbolize *logically entails* by $\models$, and define validity in terms of "all interpretations" and $\rightarrow$:

$$P \models Q \quad \text{iff} \quad \forall \Im (P \rightarrow Q \text{ is T in } \Im)$$

It is primarily this sense of *if..then* that was studied by logicians in the Middle Ages in the branch of logic called the theory of consequences, which is the prototype for Part 3 of this lecture series. A true *consequentia* in their terminology was a conditional in which the consequent follows logically from the antecedent.

Causal Entailment and Subjunctive Conditionals. In many cases when we say *if P then Q* we mean *P caused Q,* or *the fact that the P occurred caused it to be the case that the fact that Q occurred.* In English we signal this meaning when we use the subjunctive mood in combination with *if…then: if it were the case that P, it would be the case that Q.* But this sort of conditional has quite a different logic from that of the material conditional.

Normally, for example, it is a logically valid argument to substitute sentences that have the same truth value (that are "materially equivalent") in conditional propositions. That is, the following is a valid argument form:

metalanguage.

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

However, subjunctive conditionals do not obey this pattern. Consider the example of a professor in a constitutional law class who asserts,

if the President signs a bill passed by Congress, it becomes U.S. law.

He says something is true because there is a general causal law that a presidential signature is a sufficient condition for the enactment of a law passed by Congress. But here more than the factual truth-value of the antecedent and consequent of the conditional is at issue. This is shown by the failure of the substitution inference pattern above. Due the particular facts that hold of our current president, the sentences below happen to have the same truth-value:

the President signs a bill

the former governor of Texas with big ears signs a bill.

In other words, the following biconditional is true:

the President signs a bill ↔ *the former governor of Texas with big ears signs a bill.*

If truth-values were all that were at issue, then the following should be true by the substitution inference rule:

if the former governor of Texas with big ears signs a bill passed by Congress, then it becomes U.S. law

But if the constitutional law professor asserted this proposition, he would say something false because former governors, with or without big ears, do not have the authority to make U.S. laws. We see this more clearly perhaps if we formulate the conditional in the subjunctive mood: *if the former governor were to sign a bill, then it would become law U.S.* The subjunctive mood invites us to imagine cases, which it is possible to do, in which former governors with big ears try but fail to sign bills into law.

The Biconditional

The next rule is that for the biconditional:

The Biconditional		
	T,T $\longrightarrow$ T	
	T , F $\longrightarrow$ F	
	F , T $\longrightarrow$ F	
	F , F $\longrightarrow$ T	

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:

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

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

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

Translating English using $\leftrightarrow$

The biconditional is invented by logicians to translate the mathematical jargon *if and only if,* which does not really occur in ordinary English. We use $\leftrightarrow$ when we intend to convey that the material conditional works in both directions, i.e. that $P\rightarrow Q$ and $Q\rightarrow P$. Occasionally we express this idea in English when we say, *P exactly when Q.*

General Remarks on Translating using the Propositional Connectives

Natural languages make use of an abbreviating rule that linguists call *gapping* which shortens a sentence that in logical notation would be represented as a complex conjunctive or disjunctive sentence into a simple sentence with a complex noun or verb phrase. For example, in the sentence pairs below both say the same thing, but the first is an abbreviation of the second:

Jack and Jill went up the hill.
Jack went up the hill and Jill went up the hill.

Jack fell down and broke his crown.
Jack fell down and Jack broke his crown.

Similar abbreviations are made with disjunctions.

Often the grammatical structure of English sentences is exhibited by separating larger parts by punctuation, as in

Jack fell down and broke his crown, and Jill came tumbling after.

This should be translated as $(P\wedge Q)\wedge R$ because the gapping indicates that the first clause abbreviates a conjunction $P\wedge Q$, and the comma plus *and* indicates a major division in which the clause to its left and that to its right are joined by a conjunction.

Here are some general rules for translating using the connectives:

1. Underline all punctuation, like commas and semicolons, and words like *not, and, or,* and *if…then* that usually mean the same as a connective.
2. Replace pronouns by their antecedents. Guess which is which if there is an ambiguity.
3. Expand any gappings to explicit conjunctions or disjunctions of atomic sentences.

4. If a conjunctive or disjunctive series consists of three or more sentences, subdivide these into groups of pairs because $\wedge$ and $\vee$ must join only pairs of sentences. Beware that negations may govern a complex part.
5. Identify the atomic sentences and give each a letter p_1, p_2, p_3, etc. using the same letter for repeated occurrences of the same atomic sentence.
6. If the punctuation does not tell you which of two clauses is the whole and which the part, you have to make a judgment based on what you think the sentence was trying to say.
7. Working from the outside in, place parentheses around the larger parts and separate them using the connective indicated by the connecting words.

Example

When Fido comes home with Jack and barks, Bill won't even notice or if he does he'll throw a shoe at him.

p_1 $\wedge$ p_2 $\rightarrow$ $\sim$ p_3 $\vee$

<u>When</u> **(***Fido comes home with Jack* <u>*and*</u> *[Fido] barks***)** **((***[Bill] will* <u>*[not]*</u> *even notice***)** <u>*or*</u>

p_3 $\rightarrow$ p_4

(<u>*if*</u> *[Bill] does[notice],* *Bill [will] throw a shoe at [Fido]*.**))**

$(p_1 \wedge p_2) \rightarrow (\sim p_3 \vee (p_3 \rightarrow p_4))$

Exercise. Translate each of the following English sentences into the syntax of the propositional logic. Do so in stages:

1. Write down the English sentence in its original form.
2. Rewrite it in an expanded form:
   a. replace pronouns with their antecedents,
   b. expand gapped clauses into conjunctions and disjunctions,
   c. underline connecting words and writing above them (or in the correct place elsewhere) the symbol for the connective it translates,
   d. write above each occurrence of simple sentence a letter (p_1, p_2, p_3, etc.,) indicating it is an atomic sentence,
   e. add parentheses indicating sentence structure as indicated by the punctuation and "sense" of the original.
3. Write the symbolic translation.

Translate the following:

1. *If Jill sees Fido trying to eat Jack's shoe, she will take it away and feed him.*

2. *Although Jack loves Fido, it is Jill, not he, that feeds him*.
3. *Fido wags his tail when he sees Jack, however he doesn't and only barks if he's hungry.*
4. *Fido will eat only if he is very hungry and isn't excited, unless it is Jack who feeds him.*
5. *Fido loves both Jack and Jill, but will obey neither of them.*
6. *Jack will either walk the dog and not feed it, or feed it and not walk it, but when he remembers, he does both.*
7. *Fido isn't lonely when Jack is home, except when Jack isn't paying attention and is reading the paper.*
8. *When the dog eats only when Jack does, Jill won't eat with either of them.*

Summary

In this lecture we have encountered two main ideas. First we saw what it is to define the notion of sentence using constructive methods. The set of sentences in propositional logic is inductively defined from the set of atomic sentences by five formation rules, one for each of the traditional sentential connectives.

Secondly we studied the meaning of the connectives. Each is explained by its characteristic truth-function, displayed in a truth-table. These explain how the truth-value of a whole sentence formed from the connective is calculated from those of its parts. We also confronted the problem of how to translate sentences in English into propositional notation, noted some differences in meaning and some pointers on how to find the best translation.

LECTURE 9. THE INDUCTIVE DEFINITION OF TRUTH

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. If the sentence in question is grammatically simple, then this idea is relatively easy to understand. We have discussed at some length what steps are required for explaining when a simple subject-predicate sentence *S is P* "corresponds to the world." First we lay out a background ontology, next explain how words that fall in one part of speech stand for entities in a particular category in the ontology, and lastly for each type of sentence state its "truth-conditions." This statement lays out the conditions that must hold "in the world "among the entities referred to by the sentence's referring terms for the sentence to be true. In the syllogistic, for example, **A***XY* "corresponds" relative to an interpretation $\Im$ if $\Im$ assigns sets to *X* and *Y*, and the former is a subset of the latter.

When the syntax contains complex sentences, however, there is a problem. How do simple sentences "refer" to anything? There seems to be no definition of truth that simultaneously (1) applies to sentence of all types, simple and complex, (2) provides an analysis of truth as "correspondence," and (3) consists of a set of necessary and sufficient conditions of the sort found in traditional definitions.

In the 1930's the Polish logician Alfred Tarski (1902-1983) provided a solution to the problem.[18] 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."

In the syllogistic the basic parts of a proposition are its subject and predicate terms, and these stand in an intuitively plausible way for "entities in the world." It is in terms of the referents of these basic parts that the truth-conditions for the four categorical propositions are formulated. Moreover, because the terms have referents that we would intuitively recognize as entities that make up "the world," we are ready to grant that these formulations do amount to conditions on how the proposition type "corresponds to the world." In propositional logic, however, the situation is not so intuitive. It is true that the truth-value of a complex proposition is ultimately determined by the truth-values of its *atomic* parts. However the interpretations of these atomic parts consists of truth-values, and truth-values 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 Frege did, however, adopt this somewhat odd turn of phrase. Since it is uncommon 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 known as abstraction, he "abstracts" those features shared by both true simple and complex sentences. The

[18] 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*

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 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 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 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.

Research 4 (1944), 341-375.

By stating conditions on the $\Im$-values of referring parts of *P* , $TC_{\Im}(P)$ states what relations must hold among these values in order for *P* to "correspond to the world." Thus, in propositional logic, $TC_{\Im}(P)$ 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 it is fair to say that $TC_{\Im}(P)$ defined this way do, in an abstract sense, state 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 terms 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: [19]

(T) *P* is true relative to $\Im$ iff $TC_{\Im}(P)$

Consider the example of the syllogistic. For every **A**-proposition **A***XY* it is possible to prove a instance of

$$\mathbf{A}XY \text{ is true relative to } \Im \text{ iff } TC_{\Im}(\mathbf{A}XY),$$

namely,

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

Here $\Im(X)\subseteq\Im(Y)$ counts as the "truth-conditions of **A***XY*, i.e. as $TC_{\Im}(\mathbf{A}XY)$, is because $\Im(X)\subseteq\Im(Y)$ is the condition under which **A***XY* is true and this condition is stated solely in terms of the interpretations of *X* and *Y,* the basic referring terms that occur in **A***XY.*

Likewise for every **E**-proposition **E***XY* it is possible to prove a instance of

E*XY* is true relative to $\Im$ iff $TC_{\Im}(\mathbf{E}XY)$,

namely

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

and similarly for **I**- and **O**-proposition

In propositional logic $TC_{\Im}(P)$ is a statement in the metalanguage that (1) is equivalent to $\Im(P)$=T and (2) 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 a 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.

[19] The sense of *iff* in (T) is $\leftrightarrow$, which is equivalent to $\rightarrow$ in both directions. Since $\rightarrow$ is the material conditional

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 $<P,V>\in\Im$ 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 a 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 $<P,\mathrm{T}>\in\Im$ and $<Q,\mathrm{F}>\in\Im$, we know we should put $<P\wedge Q,\mathrm{F}>\in\Im$ 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

, (T) is sometimes called Tarski's *material adequacy condition* .

or F but not both. Then we can formulate this rule in several equivalent ways, getting shorter each time:

Negation Rule

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

As the truth-table for $\vee$ stipulates, if P is paired with the truth-value V in $\mathfrak{I}$ 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. Again, we can formulate this rule in several equivalent ways, getting shorter each time.

Disjunction Rule

1. 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 $<P\vee Q,\mathrm{F}>\in\mathfrak{I}$

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

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

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

Interpreting Conjunctions

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 $P \wedge Q$ 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 $<P,T>\in\Im$ and $<Q,T>\in\Im$, then $<P\leftrightarrow Q,T>\in\Im$
   If $<P,T>\in\Im$ and $<Q,F>\in\Im$, then $<P\leftrightarrow Q,F>\in\Im$
   If $<P,F>\in\Im$ and $<Q,T>\in\Im$, then $<P\leftrightarrow Q,F>\in\Im$
   If $<P,F>\in\Im$ and $<Q,T>\in\Im$, then $<P\leftrightarrow Q,T>\in\Im$

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.

An atomic interpretation is any mapping *Atomic*-$\Im$ from atomic sentences to the truth-values (i.e. set of pairs $<p_i,V>$) such that,

1. for any atomic p_i (in *ASen*) there is some *V* truth-value *V* (T or F) such that $<p_i,V>$ is in *Atomic*-$\Im$; and
2. p_i is not paired with more than one value in *Atomic*-$\Im$ (i.e. 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,…,k places respectively, then the function h defined as follows is an j+…+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, ~

$tf_{\leftrightarrow}(tf_{\wedge}(\Im(p),tf_{\sim}(tf_{\vee}(\Im(q),\Im(r)))), \Im(s))$ 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 that 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) $\Im(P)$=T iff $TC_{\Im}(P)$

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.*

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

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

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

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

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 bit of 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 even a very complex sentence can be expressed in terms of the truth-values of its atomic sentences. This method will it allow us to figure out very easily the truth-conditions $TC_{\Im}(P)$, for any sentence *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		$< p_2,\mathrm{T}>$	$< p_2,\mathrm{T}>$	$< p_2,\mathrm{F}>$	$< p_2,\mathrm{F}>$
2. p_4		$< p_4,\mathrm{T}>$	$< p_4,\mathrm{F}>$	$< p_4,\mathrm{T}>$	$< p_4,\mathrm{F}>$
3. $\sim p_4$		$< \sim p_4 \mathrm{f}>$	$<\sim p_4,\mathrm{T}>$	$<\sim p_4,\mathrm{F}>$	$< \sim p_4,\mathrm{T}>$
4. $(\sim p_4 \vee p_2)$		$< \sim p_4 \vee p_2,\mathrm{T}>$	$< \sim p_4 \vee p_2,\mathrm{T}>$	$< \sim p_4 \vee p_2,\mathrm{F}>$	$< \sim p_4 \vee p_2,\mathrm{T}>$
5. $((\sim p_4 \vee p_2) \wedge p_4)$		$<((\sim p_4 \vee p_2) \wedge p_4),\mathrm{T}>$	$<((\sim p_4 \vee p_2) \wedge p_4),\mathrm{F}>$	$<((\sim p_4 \vee p_2) \wedge p_4),\mathrm{F}>$	$<((\sim p_4 \vee p_2) \wedge p_4),\mathrm{F}>$

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 $(\sim p_4 \vee p_2) \wedge p_4$:

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

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-conditions 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 are two atomic sentence, and therefore $2^2=4$ 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 $\sim(p_1 \vee \sim p_3)$:

$TC_{\Im}\ (\sim(p_3 \vee \sim p_3))=F$ iff [($\Im(p_1)$=T and $\Im(p_3)$=T) or

($\Im(p_1)$=F and $\Im(p_3)$=T) or

($\Im(p_1)$=F and $\Im(p_3)$=F)]

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-conditions 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 3.

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

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 emphasize 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 correspond 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 10. FIRST-ORDER LOGIC

Expressive Power

Simple and Complex Sentences in a Single Syntax

In Part 2 of 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 to 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 sentences, which had no internal grammatical parts and could only be said to have a "referent" in the sense that they had 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 words 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 built 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 syllogistic's 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*].

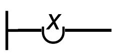
Notice, for example, that in the Principle of Abstraction there is a universal quantifier nested inside an existential quantifier, and that Peano's second postulate begins with two universal quantifiers. These propositions cannot be formulated in syllogistic syntax in a way that allows the deduction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

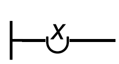
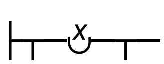
Secondly, he introduced special symbols, called predicates, which stand for sets and relations. For these he used upper case letters, as we continue to do today: *F,G,H,…* Predicates that stand for sets are followed by a single symbol naming an individual and are called *one-place predicates.* For example, *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.

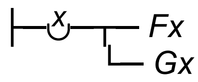
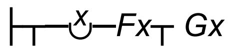
Thirdly, he also introduced symbolization for the universal and existential quantifiers, and for their accompanying variables. Though Frege used lowercase gothic letters for variables, we shall follow the modern practice of using lowercase letters from the end of the alphabet: *u,v,w,x,y,* and *z.* For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



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	 *Fx*	ΠxFx	$(x)Fx$	$\forall xFx$
for some x, Fx	 *Fx*	ΣxFx	$(\exists x)Fx$	$\exists xFx$

for all x, if Fx then Gx		Π*x*C*FxGx*	(*x*)(*Fx*⊃*Gx*)	∀*x*(*Fx*→*Gx*)
for some, x Fx and 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 the following conventions.

<u>Singular Terms.</u> 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 on 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. $Trms = Cns \cup Vbls$.

<u>Predicates.</u> Predicates are represented by P^n_m, and by upper case letters *F,…,M*, with and without subscripts and superscripts. A super-script indicates the predicate's *degree*, i.e. the number of singular terms that follow 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 set. 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 it is a convention that the first two-place predicate stands for the identity relation among elements in the domain. For this purpose we use the symbol **=** (in bold type), which is the traditional symbol for identity. 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.

<u>Formulas.</u> 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 of 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. $\exists z\forall x(\ z=x\rightarrow\exists yHzy)$
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. $\sim Fx$	1, $fr_{\sim}$
4. $(\sim Fx\vee Gxb)$	2 & 3, $fr_{\vee}$
5. $((\sim Fx\vee Gxb)\wedge\sim Fx)$	4 & 3, $fr_{\wedge}$
6. $\exists x((\sim Fx\vee Gxb)\wedge\sim Fx)$	5, $fr_{\exists}$

1. Hzy	atomic
2. $\exists yHzy$	1, $fr_{\exists}$
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. $\sim Fxc$	1, $fr_{\sim}$
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 $\Im$ in which quantified formulas are true or false. We shall use Venn diagrams for this purpose. The universe of entities that exist relative to $\Im$, 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 $\Im$. 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 $\Im$. 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 $\Im$, 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 $\mathfrak{I}$.

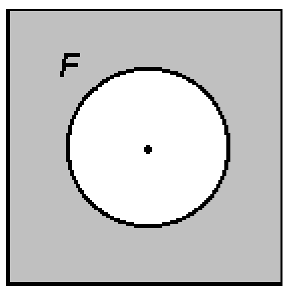
Unrestricted Quantifiers

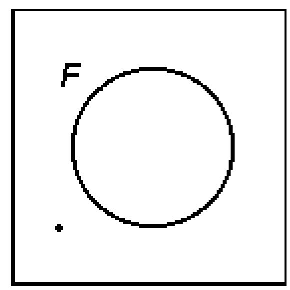
Let us begin with the simple use of the universal and existential quantifiers to say (1) that everything in the universe falls in the class named by *F,* and (2) that at least one thing falls in that class:

True

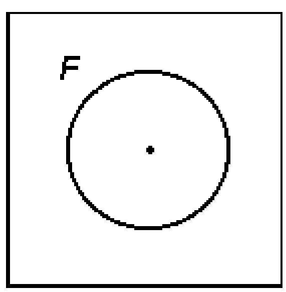
False

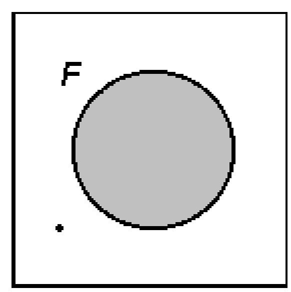
Everything is F
$\forall xFx$





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

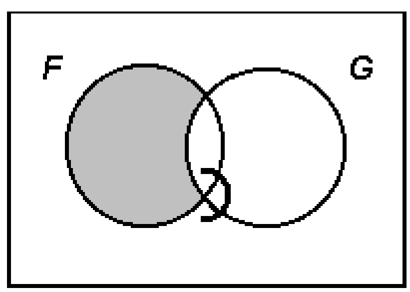
In modern notation the universal affirmative **A**-proposition *Every F is G* is reformulated as a conditional and symbolized using $\rightarrow$:

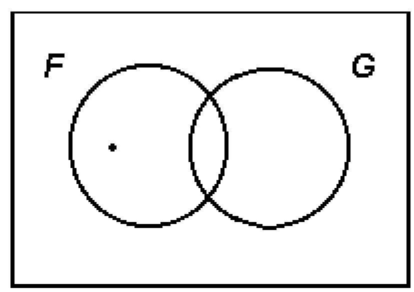
Every F is G *For all x, if Fx then Gx* $\forall x(Fx \rightarrow Gx)$

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
$\forall x(Fx \rightarrow Gx)$

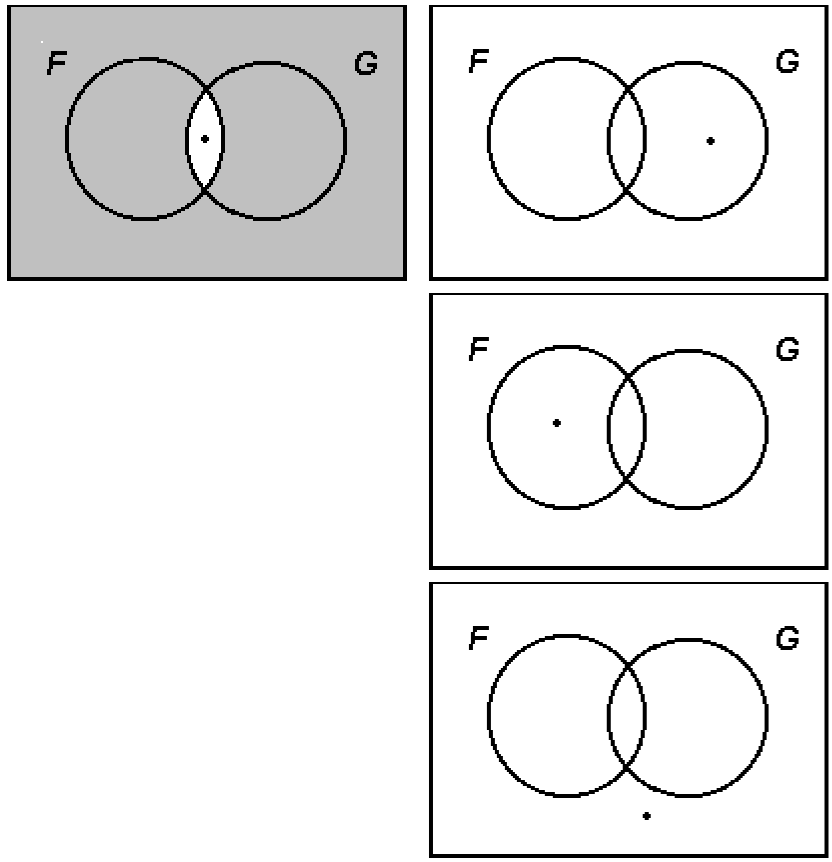




(Here the bold line crossing the lines separating the three subregions is an entity "on the line." It indicates that there is at least one entity in the domain without declaring which subregion it is in.)

True False

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



Particular Affirmatives

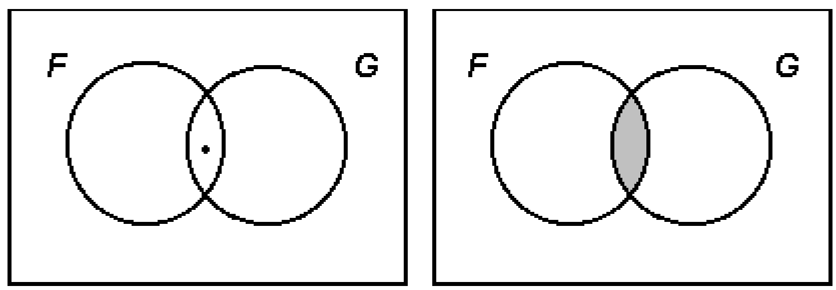
In modern notation the particular affirmative **I**-proposition *Some F is G* is reformulated as a conjunction and symbolized using $\wedge$:

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

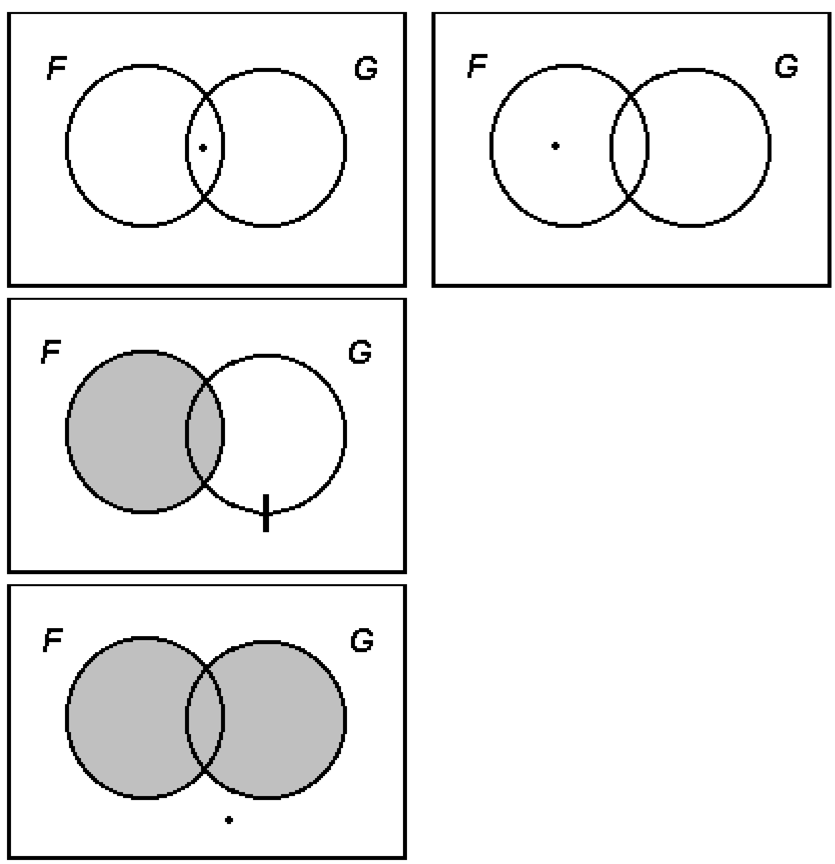
It is important to see how this differs from the conditional $\exists x(Fx \rightarrow Gx)$. As the diagrams below show, the latter asserts a rather odd claim. Given the truth-table for $\rightarrow$, this conditional is true in three cases: (1) when both *Fx* and *Gx* are true, (2) when *Fx* is false 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, ∃*x*(*Fx*→*Gx*) is an inappropriate translation of *some F are G.* We use rather ∃*x*(*Fx*∧*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
∃*x*(*Fx*∧*Gx*)



Something is such that if it is F then it is G
∃*x*(*Fx*→*Gx*)



Distribution of the Quantifiers over Connectives

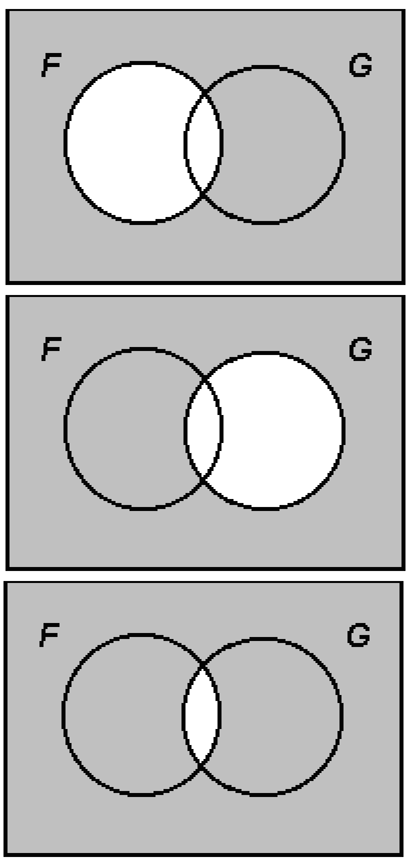
In some cases the distribution of a quantifier makes a difference in meaning. Though the following pairs are equivalent:

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

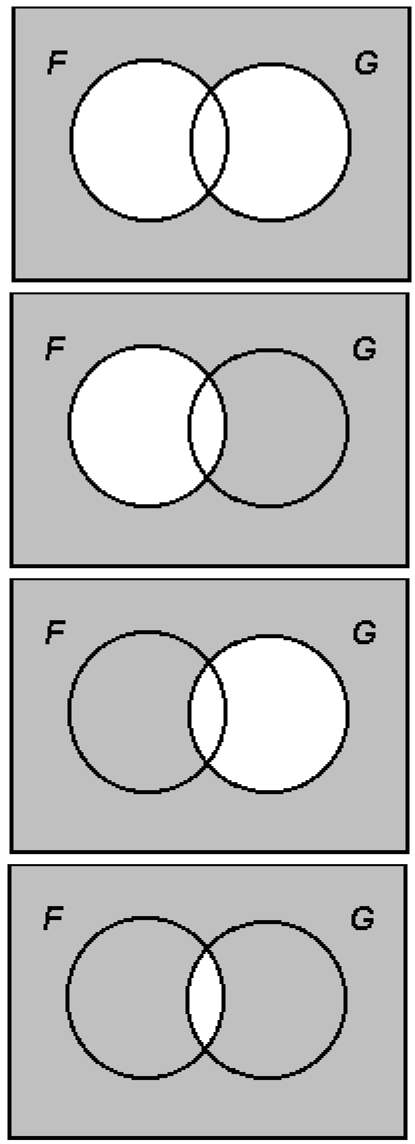
$\exists xFx \vee \forall xGx \qquad \exists x(Fx \vee Gx)$

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

$\forall xFx \vee \forall xGx$

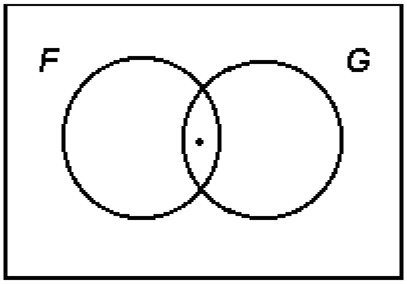


$\forall x(Fx \vee Gx)$

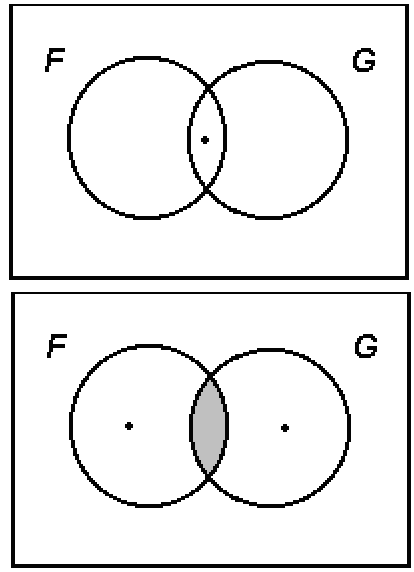


Likewise $\exists x(Fx \wedge Gx)$ entails but is not entailed by $\exists xFx \wedge \exists xGx$:

$\exists x(Fx \wedge Gx)$



$\exists xFx \wedge \exists xGx$

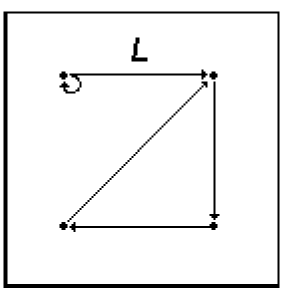


Embedded Quantifiers

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

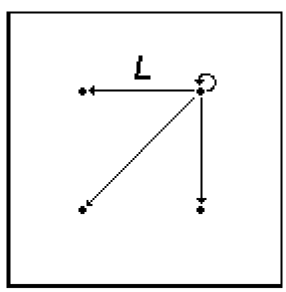
Everybody loves somebody or other.

$\forall x \exists y Lxy$



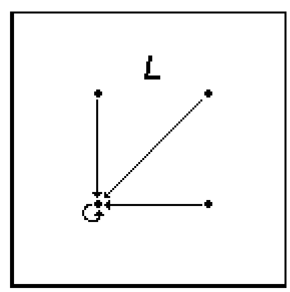
Some body loves everybody.

$\exists x \forall y Lxy$



Everybody loves some one person.

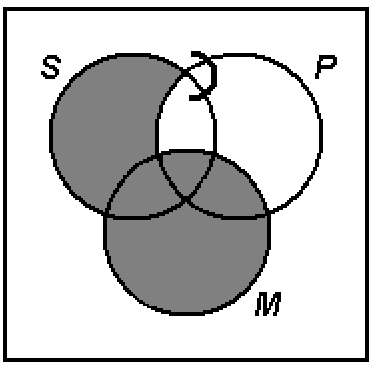
$\forall x \exists y(Lxy \wedge \forall z(Lxz \rightarrow z=y))$



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
Every M is S	$\forall x(Mx \rightarrow Sx)$	true
Some S is not P	$\exists x(Sx \wedge \sim Px)$	false



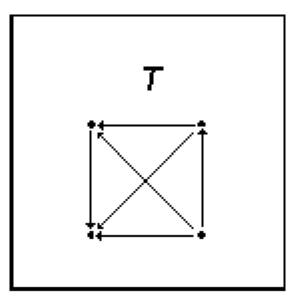
Properties of Relations

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:

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

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.

*LECTURE 11. FORMAL SEMANTICS FOR FIRST-ORDER LOGIC

Intuitions about the Truth-Conditions of Each Formula Type

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 a relation that pairs each expression to its "interpretation" 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*, $\Im(\mathrm{x}) \in D$, 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

* This lecture and its exercises (marked by an asterisks *) may be omitted without loss of continuity.

and variables that refer to individuals in the domain *D,* and of predicates that talk about sets 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 is 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 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, 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 $\Im(t_1)$, …, $\Im(t_n)$ 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 ~, ∧, ∨, →, and ↔ (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 existentially quantified expressions talk 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. $D=\{d_1, d_2,\ldots, d_{37}\}$. 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 $\Im(F)= \{d_1, d_2,\ldots, d_{37}\}$. 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 reassigns x to stand 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 $\forall xFx$:

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

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

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

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 $\mathrm{TC}_{\Im}(\forall xP)$, because it is formulated only in terms of the $\Im$-values of the parts of *P.*[20] Hence we have an instance of Tarski's T principle:

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

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:

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

We are now ready 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

[20] 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. $\Im(=)$ 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, $\Im(t_n{=}t_m)$=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 $\Im(P)$ =T, then $\Im(\sim P)$=T;
      $\Im(P)$ = F otherwise;
   c. if $\Im(P)$ =T and $\Im(Q)$) =T, then $\Im(P \wedge Q)$ =T;
      $\Im(P)$, $\Im(Q)$ = F otherwise;
   d. if $\Im(P)$=T or $\Im(Q)$) =T, then $\Im(P \vee Q)$ =T;
      $\Im(P \vee Q)$) = F otherwise;
   e. if $\Im(P)$=F or $\Im(Q)$) =T, then $\Im(P \rightarrow Q)$ =T;
      $\Im(P \rightarrow Q)$ = F otherwise;
   f. if $\Im(P)$=T & $\Im(Q)$ =T, or if $\Im(P)$=F & $\Im(Q)$ =F, then $\Im(P \leftrightarrow Q)$ =T;
      $\Im(P \leftrightarrow Q)$ = 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 <$\forall xFx$> 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 both in $\Im$ and in all other interpretations of the formulas that are made up of these atomic parts. Suppose, for example *P* is the universally quantified formula ∀xFx. Now, by hypothesis, the value of its atomic part *Fx* has already been determined for all interpretations. That is, it is determined what value *Fx* has not only in $\Im$ but in all other interpretations, including the value that *Fx* has in $\Im_{[x\rightarrow d]}$, for any *d* whatever in *D.* If $\Im_{[x\rightarrow d]}(Fx)$=T for all *d* in the domain, then *Fx* is T in $\Im$ no matter what *x* stands for, and hence we know that $\Im$ should assign T to $\forall xFx$.

In this way all formulas having atomic formulas (of length 1) 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 interpretations of the parts to determine the value in $\Im$ of the whole formula, even though the rule fixing the value in $\Im$ may require information about the values of the parts in interpretations other than $\Im$. 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 all formulas in all interpretations are determined in stages corresponding to the stages of construction of each formula. The set of interpretations is said to be defined by *simultaneous induction*.

Though every $\Im$ is well defined by the process of simultaneous induction, the method lacks an important feature of ordinary definition by induction. It is no 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 $\Im_{[x\to d]}$, 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 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 use 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 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. $\Im(t_n{=}t_m)$=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 any 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 $\Im(P)$=T iff TC(P) is "a theorem of set theory that follows from the definition of $\Im$."

Examples

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

$\Im^{D}{}_{[x\rightarrow d]}(x)$ and $\Im^{D}{}_{[x\rightarrow d,\ y\rightarrow d']}(x)$ in red are alternative notation for d, and
$\Im^{D}{}_{[y\rightarrow d']}(y)$ and $\Im^{D}{}_{[x\rightarrow d,\ y\rightarrow d']}(y)$ 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 biconditionals of the following form:

$$\text{(T)}\qquad \Im(P)=\text{T} \qquad \text{iff} \qquad \underbrace{\text{TC}_{\Im}(P)}_{\substack{\textbf{Truth-Conditions for } \boldsymbol{P}\\ \text{Conditions that must hold in the world among}\\ \text{the entities referred to by the smallest parts}\\ \text{of speech in } P}}$$

Example 1. $Fc \wedge Gb$

1. $\Im(Fc \wedge Gb)$=T	iff	$\Im(Fc)$=T and $Gb)$=T	E5
	iff	$\Im^D(c) \in \Im^D(F)$ and $\Im^D(b) \in \Im^D(G)$	1, sub of eq E1 & E2

Example 2. $Rac \rightarrow Gx$

1. $\Im(Rac \rightarrow Gx)$=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 $d \in D$, $\Im^D{}_{[x\rightarrow d]}(Fx)$=T	E9
2.	iff	for all $d \in D$, $\Im^D{}_{[x\rightarrow d]}(x) \in \Im^D(F)$	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 $d \in D$, $\Im^D{}_{[x\rightarrow d]}(Fx)$=T	E10
2.	iff	for some $d \in D$, $\Im^D{}_{[x\rightarrow d]}(x) \in \Im^D(F)$	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 $d \in D$, $\Im^D{}_{[x\rightarrow d]}(\exists yRxy)$=T	E9
2.	iff	iff for all $d \in D$, for some $d' \in D$, $\Im^D{}_{[x\rightarrow d,\ y\rightarrow d']}(Rxy)$=T	1, sub eq E10
3.	iff	for all $d \in D$, for some $d' \in D$, $<\Im^D{}_{[x\rightarrow d,\ y\rightarrow d']}(x),\Im^D{}_{[x\rightarrow d,\ y\rightarrow d']}(y)> \in \Im^D(R))$	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 $d \in D$, $\Im^D{}_{[x\rightarrow d]}(\forall yRxy)$=T	E10
2.	iff	iff for some $d \in D$, for all $d' \in D$, $\Im^D{}_{[x\rightarrow d,\ y\rightarrow d']}(Rxy)$=T	1, sub eq E19
3.	iff	for some $d \in D$, for all $d' \in D$, $<\Im^D{}_{[x\rightarrow d,\ y\rightarrow d']}(x),\Im^D{}_{[x\rightarrow d,\ y\rightarrow d']}(y)> \in \Im^D(R))$	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 $d\in D$, $\langle\Im^{D}_{[x\to d,\ y\to d']}(x),\Im^{D}_{[x\to d,\ y\to d']}(x)\rangle\in\Im^{D}(R))$	1, sub eq E2
3.	iff	for all $d\in D$, $\langle d,d\rangle\in\Im^{D}(R))$	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 $\Im^{D}_{[x\to d]}(Fx)\neq$T or $\Im^{D}_{[x\to d]}(Gx))$=T	1, sub of eq E7
3.	iff	for all $d\in D$, either $\Im^{D}_{[x\to d]}(x)\notin\Im^{D}(F)$ or $\Im^{D}_{[x\to d]}(x)\in\Im^{D}(G)$	2, sub of eq E1
4.	iff	for all $d\in D$, either $d\notin\Im^{D}(F)$ or $d\in\Im^{D}(G)$	3, sub of =

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

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

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

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

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

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

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

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

4.	iff	for all $d\in D$, either $\Im^{D}_{[x\to d]}(x)\notin\Im^{D}(F)$ or for some $d'\in D$, $<\Im^{D}_{[x\to d,\ y\to d']}(x),\Im^{D}_{[x\to d,\ y\to d']}(y)>\in\Im^{D}(R))$
5.	iff	for all $d\in D$, either $d\notin\Im^{D}(F)$ or for some $d'\in D$ $<d,d'>\in\Im^{D}(R)$

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

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

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

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

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

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

1. $\Im(\exists x(Fx\to Gx))$=T	iff	for some $d\in D$, $\Im^{D}_{[x\to d]}(Fx\to Gx)$=T

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

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. 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 $\Im(\forall x(Fx \to Gx))$=T.

Proof:

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

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

(Note that the last line follows from the line before by substitution of identities because, given the definition of $\Im_{[x\to d]}$, $\Im_{[x\to d]}(x)=d$.)

Hence, $TC_{\Im}(\forall x(Fx \to Gx))$ 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. $\Im(\forall x(Fx \rightarrow Gx))$=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 $d \in D$, $\Im_{[x \rightarrow d]}(Fx \wedge Gx)$=T
	iff	for some $d \in D$, $\Im_{[x \rightarrow d]}(Fx)$=T and $_{x \rightarrow d]}(Gx)$=T
	iff	for some $d \in D$, $\Im_{[x \rightarrow d]}(x) \in \Im(F)$ and $\Im_{[x \rightarrow d]}(x) \in \Im(G)$
	iff	for some $d \in D$, $d \in \Im(F)$ and $d \in \Im(G)$

(Note that the last line follows from the line before by substitution of identities because, given the definition of $\Im_{[x \rightarrow d]}$, $\Im_{[x \rightarrow d]}(x)=d$.)

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

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 expressions would mirror ontological relations that hold among entities in the world.

Truth-Conditions for *P*

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

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 fact "in the world" that its 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 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 $\forall\mathfrak{I}\,(\,\mathfrak{I}(P_1)=\mathrm{T}\&\ldots\&\mathfrak{I}(P_n)=\mathrm{T}\&\ldots) \rightarrow \mathfrak{I}(Q)=\mathrm{T})$

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

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

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.

Lecture 7

Exercise. Translate each of the following English sentence into **A**, **E**, **I**, or **O** propositions, with or without predicate negations, using the predicates *F* and *G*. Precede each translation with a declaration of what word or words the letters *F* and *G* represent. (For example, *F = human, G =mortal*).

1. *All humans are mortal.*
2. *Some cows fly.*
3. *There are green monkeys.*
4. *There are some monkeys that do not swim.*
5. *There are no pink elephants.*
6. *Students are poor.*

Lecture 8

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. Translate each of the following English sentences into the syntax of the propositional logic. Do so in stages:

1. Write down the English sentence in its original form.
2. Rewrite it in an expanded form:
   a. replace pronouns with their antecedents,
   b. expand gapped clauses into conjunctions and disjunctions,
   c. underline connecting words and writing above them (or in the correct place elsewhere) the symbol for the connective it translates,
   d. write above each occurrence of simple sentence a letter (p_1, p_2, p_3, etc.) indicating it is an atomic sentence,
   e. add parentheses indicating sentence structure as indicated by the punctuation and "sense" of the original.
3. Write the symbolic translation.

Exercise. Translate the following:

1. *If Jill sees Fido trying to eat Jack's shoe, she will take it away and feed him.*
2. *Although Jack loves Fido, it is Jill, not he, that feeds him.*
3. *Fido wags his tail when he sees Jack, however he doesn't and only barks if he's hungry.*
4. *Fido will eat only if he is very hungry and isn't excited, unless it is Jack who feeds him.*
5. *Fido loves both Jack and Jill, but will obey neither of them.*
6. *Jack will either walk the dog and not feed it, or feed it not walk it, but when he remembers, he does both.*
7. *Fido isn't lonely when Jack is home, except when Jack isn't paying attention and is reading the paper.*
8. *When the dog eats only when Jack does, Jill won't eat with either of them.*

Lecture 9

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]

Lecture 10

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.

Lecture 11

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

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

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

*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 $TC_{\Im}(\exists x(Gx \wedge {\sim}Fx))$.

REVIEW QUESTIONS

1. **Inductive sets.** In Part 2 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 **A**, **E**, **I**, and **O** propositions of categorical logic,

   b. the sentences of propositional logic, [*c. the formulas of first-order logic}.

It is easier to explain the (T) rule for "atomic sentences," i.e. the **A**, **E**, **I**, and **O** propositions of categorical logic and the formulas *Fa* and *Rab* of first-order logic, because these sentences contain words that refer to sets and their elements and say what has to hold among these sets and elements for the sentence to be true. It is harder, however, to explain what the truth-conditions for *P* should be if *P* is a complex sentence, which are (made up of the connectives $\sim$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$, or of quantifiers $\forall$ and $\exists$). In what sense does the (T) rule apply to complex sentences and formulas? For simplicity you may ignore the quantifiers $\forall$ and $\exists$ and limit your answer to the case of complex sentences made up from the connectives $\sim$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$.