Logic for Liberal Arts Students,

An Introduction

John N. Martin

PROFESSOR OF PHILOSOPHY
UNIVERSITY OF CINCINNATI

Version1/5/2009

Part 1. The Logic of Terms

Version1/5/2009

Dialectica among the Seven Liberal Arts

Nicolà Pisano, Siena Cathedral, 1266

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

Saint Augustine, *On Christian Doctrine*, 32.

Table of Contents

Introduction

Lecture Topics

The purpose of these lectures is to introduce students to the Queen of the Liberal Arts, *ars atrium* (the art of arts), as logic was known in the Middle Ages. Logic is the most universal of sciences because it is presupposed by all the others. You cannot do mathematics or physics or biology unless you can trust your tools of reasoning, which you must learn from logic. Logic in this sense has a history as long as Western learning itself, and in the course of time has made major contributions outside logic proper, above all to philosophy and mathematics. The purpose of these lectures is to introduce you to some of these and make you interested in learning more about them.

The lectures are divided into three parts, corresponding to the traditional division of Aristotle's logic: the properties of terms, the truth-conditions of propositions, and the logic of arguments. Each part begins with a historical introduction and move onto the ideas and, as appropriate, the symbolism of modern logic. Sections preceded with an asterisk (*) contain supplementary material and may be skipped without loss of continuity.

Part 1, *The Properties of Terms*, covers the semantics of singular and general terms, both monadic and relational. As a whole, it is an extended rationale for the introduction in modern logic of entities known as "sets" to serve as the objects that words represent. In this role sets replace the more problematic entities called "properties" and "relations" posited for that purpose in ancient and mediaeval logic. Here students encounter technical notation for the first time. You will learn and prove some basic facts about sets. Equipped with this knowledge, we will then use set theory throughout the rest of the lectures as a tool for explaining logical concepts. Among the important set theoretic tools we shall meet is that of an *inductively defined set.* We will see how for every element in such a set there is what is called a *construction sequence* that shows how the element was added to the set as the result of a finite, step by step, process. Inductive definition will prove very important later in explaining central ideas that have resisted more traditional definitions formulated in terms of necessary and sufficient conditions.

Part 2, *The Truth-Conditions of Propositions*, discusses three “languages,” which are progressively more powerful in the propositions they express. We start with the relatively simple subject-predicate language of Aristotle’s categorical propositions, proceed to the complex sentences formed by “conjunctions” like *and, or, then,* and *not* of what is called *propositional logic*, and finish with the full notation of modern logic, called *first-order logic*, which can express most of the propositions of mathematics and science. Not only does it express simple subject-predicate sentences and their compounds, but it also incorporates relational predicates and complex quantified formulas. Here we shall apply the method of inductive definition to define the set of grammatical sentences, and meet the idea of a *phrase-structure grammar*, an approach which has proven fruitful in both modern logic and linguistics.

After defining *sentence*, we then see how to answer the central issue: what is it for sentences to be true or false? In steps, we shall work through what is known as *the correspondence theory of truth*. We start by explaining it for the simple language of categorical propositions, then extend the account to the propositional logic, and finish by stating the full theory for the syntax of first-order logic. Students will be introduced to the standard theory, which was sketched in the 1930's by Alfred Tarski and provides the basic ideas for many of the important discoveries of modern logic. Again by using the technique of inductive definitions, we will see how it is possible to provide, for any sentence of the language, simple or complex, a formulation of the conditions under which it is true, its *truth-conditions,* that mention only the facts that must obtain among those entities referred to by the sentence's nouns and verbs.

In Part 3, *The Logic of Arguments,* we explore how logic succeeds in what is probably its central scientific goal: to distinguish "good" from "bad" arguments. This is logic proper, the area in which logic's greatest discoveries are found. We shall see that there are two quite different approaches to the notion of "good" argument that, in the end, they yield the same idea: a *semantic* approach in which a good argument is defined as one that leads necessarily from true premises to true conclusions, and a *proof theoretic* or *syntactic* approach in which good arguments are distinguished by the fact that they can be given formally correct proofs.

The semantic definition in terms of truth explains why valid arguments are important tools in science – they provide a technique for extending our knowledge from truth to truth. We shall see how to give a set theoretic proof that an argument is valid or invalid by appealing to the truth-conditions for the argument's sentences explained in Part 2.

The proof theoretic definition of "good" argument explains why logic differs from almost every other science in its extraordinary degree of certainty. Proofs turn out to be nothing more than construction sequences that show that a sentence is a member of an inductively defined set. It is because such a sequence lays open to inspection a finite series of physically visible symbols that we can readily "see" whether each step is correct and can do so with a high degree of reliability. We shall see that in special cases there are even *decision procedures* (algorithms, computer programs) to test for "good" arguments.

It is in this part that you will meet some of the famous discoveries of modern logic. These are interesting and rather unusual in science because they concern the limitations of human knowledge – a topic of immense interest about which very little is known with mathematical precision.

The first major result is called a *completeness theorem*: the logical arguments of all three languages investigated – categorical propositions, propositional and first-order logic – can be axiomatized by an inductive definition.

The second major result is called an *incompleteness theorem*. In the 1930's Kurt Gödel discovered that it is impossible to axiomatized (give an inductive definition for) the truths of arithmetic, or of any sciences (like set theory) that entails arithmetic. This result is extremely important for our understanding for what mathematics is and what methods are appropriate to mathematics as a science. Since the time of Euclid in ancient Greece, the standard "scientific method" of mathematicians has been to "prove" a result in the sense of deducing it as a theorem in an axiom system. Moreover, if anything is mathematics, it is arithmetic. Indeed, arithmetic is the central core of mathematics. It follows that the standard method of mathematicians is incapable of explaining the heart of mathematics. Mathematicians, and philosophers of mathematics, have been scratching their heads about this result since it was discovered.

We shall define decision procedures for the valid arguments of categorical and propositional logic. We will be able to test by a finite calculation whether any argument written in their syntax is valid. However, Jean Herbrand showed in the 1930's what is called an *undecidability theorem*: though the valid arguments of first-order logic can be axiomatized, there can be no general decision procedure that tests whether any argument in first-order syntax is valid.

Several themes run through the lectures concerning the scientific method used in logic. One is ontology. It virtually always happens that along with an explanation there comes an ontology. In order to explain something, you need to posit entities that figure in the explanation. One you have the entities, your job as the "explainer" is to state the "laws" that hold among them. But the entities have to be there in the first place. Where appropriate in the text, it will be pointed out when an explanation is positing an ontology, how this ontology is broken down into various sets and relations, what the "laws" are that the explanation says holds among these entities, and what "theorems" follow from these laws.

A second issue of method particularly important in logic is that of definition. A major theme of the lectures is that traditional definitions – the sort favored by

Aristotle and traditional philosophers (to this day) – that define a set in terms of necessary and sufficient conditions is inadequate for concepts in logic. As a result logicians have developed the novel tool defining a set by “induction”, i.e. by constructing it from some basic elements by some rules for adding new elements to old. One feature of such constructions is that a set’s elements are always added sequentially. It is always possible to lay out the order (the sequence) of entities that preceded the inclusion of any particular member. Inductively defined sets and of construction sequences have proven sufficient for explaining an impressive series of concepts beyond the means of traditional definitions, including the notions of sentence, truth-in-an-interpretation, provable sentence (theorem), and calculable function. We can even explain why we can know logical results with a high degree of certainty – a feature that distinguishes logic from the empirical sciences: because a theorem is added to an axiom system by a construction sequence of physically visible steps, we can easily inspect each and “see” that it produces an element that belongs in the set.

Exercises and Symbolism

Like all subjects, logic has its technical terms, but logic’s are more arcane than most. Part of what a student is expected to learn in an introductory logic course is some familiarity with this symbolism, and how to use in proofs. These lectures will indeed make use of symbols and proofs. They will however be introduced in the service of ideas. As a result, as you meet them, you will start to use them as tools to explain other ideas.

In general to understand a technical idea, it is not enough to simply read through the text introducing it. You must stop and figure it out. Plan enough time to do so. To help you in the process, the notes provide examples and exercises. Your method should be to work through these, step by step. More difficult exercises are preceded with an asterisk (*). If there is something you do not understand, note it down and ask about it. Since technical ideas build upon one another, and each takes some time to figure out, they cannot all be learned at once. Cramming will not work. Plan to work through the notes on a regular basis.

Part 1. The Logic of Terms

LECTURE 1. THE PROBLEM OF UNIVERSALS

Parmenides: The One and the Many

Parmenides, who was writing at the dawn of Greek philosophers (about 485 B.C.), was remarkable for advancing what we would now call logical arguments for some perfectly outrageous propositions. In his poem *On Nature*[1] he defends what is called *monism*, the view that there is only one entity in the world, and that time and change do not exist. It is not an exaggeration to say that logic as distinct branch of study evolved in response to Parmenides' arguments and the arguments of others developed to refute him.

Parmenides' main line of reasoning is very simple. What is, it must be granted, is what is. It follows that what is is not what is not. Another way to translate Parmenides' phrase *what is* is *being.* It follows that what is cannot not be, and hence cannot come into being or perish. Hence there is no becoming or passing away (8.3-21). There is, in other words, no change. Since what is is not what is not and what is different from something that is not that thing, it follows that what is not is not different from anything else. If something is divided then something is different from something else. It follows that what is is not different from anything (8.22-25). It follows that what is is not in a different place from anything else. Hence what is is in one place only (8.26-33). If something changes then it is different in quality, and motion requires a change in place. It follows that what is does not change or move (8.26-33). Hence change and motion are unreal. There is nothing moreover that is not what is (8.36-37). Being then is well rounded the same in all ways and directions, like a sphere (8.43-44).

This argument is remarkable because it presupposes what Aristotle would later call the *Law of Non-Contradiction*: no proposition and its negation can both be true. Let us abbreviave a proposition as *P* and its negation *it is not the case that P* as *not P*. Then, the law says that *P* and *not P* cannot both be true. The reason, then, that what is is not what is not, is that if what is is both what is and what is not, then two contradictory propositions, namely *what is is what is* and *not what is is what is* would both be true, and the law would be violated.

Another way to summarize Parmenides' argument is that it is a reduction to the absurd. Given the law of non-contradiction, it follows that if a proposition *P* leads logically to a contradiction *Q and not Q*, then *P* cannot be true. That is, if *P* logically implies *Q and not Q,* it follows that *not P.* The more formal Latin name for this rule is *reductio ad absurdum.* Parmenides in effect show that what is is what is because if what is were not, a contraction would be true, which is impossible.

Aristotle tells us (*Metaphysics* 1010a2-3) that by *what is* Parmenides means the natural world, the world of everyday experience. Hence contrary to all common sense and our current scientific understanding of the world, Parmenides maintains that only one unchanging timeless entity exists in all of nature. Clearly he must be wrong, but where is the mistake in his reasoning?

---

[1] Parmenides, "On Nature" in Diels, 8.

Over time Greek philosophers came to see that his mistake lies in a misunderstanding of what it takes to make a simple subject-predicate sentence of the form *S is P* and its negation *S is not P* true and false. Following the practice of modern symbolic logic, we shall use lower case italic letters from the beginning of the alphabet (*a,b,c,d,* etc.) to represent proper names. For now, we may understand proper names to be words that stand for individual things. Similarly, we shall use the upper case italics *F, G* and *H* to represent predicates that are common nouns, intransitive verbs, or adjectives that are true of individual things. We will also adopt the modern logic's perverse practice from the perspective of English word order of placing the predicate before the subject. Hence the sentence *Fa* says *a is F.* Finally, we introduce ~, our first logical symbol to represent negation. It may be translated as *it is not the case that* or more simply as *not.* We place it at the beginning (left hand side) of a sentence. For example, the sentence ~*Gb* says *it is not the case that b is G.*

With this notation we may now reformulate the issue. To say that two things are *the same* is to say that there are two proper names, call them *a* and *b,* that stand for different individual things, and that there is some predicate, call it *F*, such that the sentences *Fa* and *Fb* are both true. To say that two things named by *a* and *b* are *different* is to say that there is some predicate, call it *G*, such that the sentences *Ga* and ~*Gb* are both true. The general problem is to discover exactly what the conditions must be for propositions of the form *Fa* and ~*Fa* to be true or false. Once these conditions are spelled out in general, it will turn out that Parmenides was wrong to think that all subject terms stand for one and the same individual and that all predicates are true of just one thing. Given this formulation, the problem posed by Parmenides is sometimes called that of *sameness and difference:* What are the conditions that must obtain for propositions to be true of *a* and *b* such that they are the same in some respect but different in another?

The problem may be formulated in yet another way. Classification consists of putting together in one grouping things that are "the same" in some way, and it consist of "dividing" things that are different into these groups. The problem then may be described as one of *classification* and *division*: under what conditions are things classified as falling together, and under what conditions are they divided into separate groups?

We shall see that philosophers, and sometimes logicians, answer these questions by a appeal to a scientific technique called *reification*. This method explains a phenomenon by attributing it to the workings of some special entity with properties designed to produce just that effect. What caused that lightning flash? The ancient Greeks might blame Zeus venting his angry disposition. What explanation do peasants offer for the fact that a fisherman drowned in the local stream? They might say a pretty girl who live under the water enticed him in.

Scientists too make successful use of the technique. The orbit of Uranus exhibits a deviation in its orbit from that predicted by its mass and that of the inner planets. The cause of the anomaly was hypothesized to be an entity, a planet. Its mass and position were estimated based on Kepler's laws of motion – the planet could not cause the phenomenon unless its had those properties and obeyed those laws. Astronomers looked, and beheld Neptune. Entities called genes obeying precise laws of dominance and recession were posited by Mendel to explain the observed phenotypes of sweet peas, and atoms composed of electrons and protons with the various energy levels of the period table were posited by Mendeleyev to explain the observed reactions among chemicals in the laboratory.

Similarly, philosophers and logicians posit entities to explain sameness and difference. The entity in question, which is called a *universal*, is correlated with the predicate of a true subject-predicate sentence. Philosophers disagree about what exactly a universal is and how it correlates to a predicate, but they do agree on two points. First, the relation between the predicate and a universal is semantic; that is, it is a relation that holds between a sign and the thing the sign signifies. Secondly, they agree that whatever a universal is, it is selective; it embraces some individuals but not others. It is the selectivity of a universal that allows for the explanation, first, of the truth-conditions for a subject predicate sentences and their negation and, subsequently, of sameness and difference. The explanation take the following general form:

Truth-Conditions for Subject-Predicate Sentences and their Negations

- *Fa* is *true* if and only if the universal that *F* signifies embraces the individual that *a* stands for.
- ~*Fa* is *true* if and only if the universal that *F* signifies does not embrace the individual that *a* stands for.

Explanation of Sameness and Difference

- Two individuals are *the same* with respect to a universal *U* if and only if they may be referred to by two different proper names, say *a* and *b*, *U* may be signified by a predicate, say *F*, in such a way that the sentences *Fa* and *Fb* are both true.
- Two individuals are *different* with respect to a universal *U´* if and only if they may be referred to by different proper names, say *a* and *b*, *U´* may be signified by a predicate, say *G*, in such a way that either the sentences *Ga* and ~*Gb* are both true, or the sentences ~*Ga* and *Gb* are both true.

Things that are the same with respect to a universal may be classified together in the same group, and things that fall under different universals may be divided into separate groups.

Parmenides mistake may now be pin-pointed. A subject-predicate sentence consists of a subject that stands for an individual and a predicate that signifies a universal. Parmenides proposition *what is is what is* then is at best elliptical and at worst badly formed. If the left hand occurrence of *what is* is supposed to be a proper name, then it is a rather unusual proper name. It does not stand for normal sorts of individuals but rather for the whole universe. If the second occurrence of *what is* is supposed to be a predicate and signify a universal, it is not clear what that universal that could be. Certainly *what is* in not a normal predicate because it is not a common noun, intransitive verb, or adjective. But whatever that universal is, it should be capable of embracing more than one individual. For example, if we understand this predicate to mean the same as *exists*, then *what is is what is* means roughly *the* *universe exists.* But if this is its meaning, it is a perfectly harmless truth, and certainly does not imply that there is only one thing, or that time and change are unreal.

Sameness and difference are thus, in some sense, "explained" by appeal to "universals." But what kind of explanation is this, and what are universals? In the course of these lectures we shall have a lot to say about what sorts of explanations are appropriate to logic. Indeed, logic as a "science" and the standards of explanation appropriate to it are a central theme of the course. At this point, let us pause in our review of the explanation of sameness and difference to say something about explanation in general. We are about to see a parade of different attempts at an explanation, and we need some criteria for evaluating whether any of them succeed and if so which is the most plausible. That is, let us pause to make some general observations about the scientific standards that must be met by the sort of science that logic aspires to be.

Explanation in Logic

What is a Theory?

An explanation can always be divided into what it is trying to explain, the *phenomenon*, and the explanation. Logic, at least in its modern form, is fortunate in being a mathematical science. As such its explanations take a clear and precise form called a *deductive theory*, by which we generally mean a whole consisting of three parts: axioms (also called or "laws"), definitions, and theorems that follow from these. The "phenomenon to be explained" can then be described in a sentence *P*. It is successfully explained by the theory if this sentence is occurs as part of the theory. *P* could be an axiom or definition of the theory and then its role would be to be to serve as an assumption used to prove a larger group of propositions, but the more normally case is that *P* is a theorem deduced from prior axioms and definitions. The explanation, then, is evaluated not by looking at the truth of *P* alone, but by evaluating the entire theory, because it is theories as a whole that are acceptable or unacceptable, verified or unverified, plausible or implausible. There is, however, no one test for gauging a theory's plausibility. A number of considerations are relevant that must be balanced against one another. Here let us introduce four that are especially important in logic.

Mathematical Rigor

In modern logic we require that any acceptable theory must be mathematically rigorous. We shall learn more about what mathematical rigor means later when we study set theory, which normally provides the language and background assumptions for mathematical theories. The great virtue of mathematical rigor is that when present it allows us to distinguish very clearly between a theory's laws, definitions, and theorems. It allows us to actually provide logical proofs that the theorems follow from the laws and definitions. One of the great strengths of logic since about 1850 is that it is formulated with this sort of rigor.

Parsimony

Another criterion marking a good theory is parsimony, which is also sometimes called mathematical elegance or simplicity. The theory should be states as simply as possible, making minimal use of assumptions and definitions. In practice this goal means we shall posit as few basic sets and relations as necessary, and describe their interactions by means of the shortest set of “laws” possible.

We shall see in some detail shortly how concern over parsimony bedevils one of the famous controversies in the history of logic, “the problem of universals.” This problem consists of explaining what kind of entity in the world accounts for the fact that two things can be “the same”? The problem can be recast as one about the meaning of predicates. What kind of entity, if any, does a predicate stand for in a true subject-predicate sentence? The scientific function of “universals,” if they exist, is to serve as the referent of a predicates in a subject-predicate sentence. After all, such a sentences on the surface seems to assert that there “exists” something named by the predicate that subjects man have in common. But any theory that claims that universals exist is adding to the catalogue of things that exist an extra category. The methodological questions then arise. It is scientifically necessary to postulate this category of entity? Can the phenomenon be explained by a simpler theory that does not posit this extra class of things? The most famous formulation of the ideal of parsimony is due the medieval logician William of Ockham (1285-1348/49) and is called Ockham’s Razor: *entia non sunt multiplicanda praeter necessitatem.* Entities should not be multiplied beyond necessity. (The rule is “razor” because it calls for shaving away a theory’s unnecessary entities.) As we shall see, Ockham used this principle to argue against the existence of universals.

Empirical Confirmation

Theories, especially in mathematics, may describe completely non-physical entities. But the natural sciences by definition describe the natural world, which is a world that is accessible to us through the senses. We may use this fact to evaluate the theory. Let us call a sentence *empirical* if it is written in sensory vocabulary and can be evaluated as true or false by simply checking our sensations. Let us now distinguish between two sorts of empirical propositions. The first group consists of the theory's theorems written in empirical vocabulary, called the theory's *empirical predictions*. The second is the group of sentences that actually record what we sense to be the case. Let us call this second set of sentences the relevant empirical data. In so-called normal science, theorists spin axiom systems and deduce predictions, while experimentalists design experiments to collect data. The important point to notice is that there is no guarantee that the two sets are the same. Only if the theorists are very insightful or lucky will the set of predictions coincide with the data. The better the fit, the better the theory.

In practice the fit is never perfect. Sometimes rival theories predict different data, and the choice between the two over-all theories is not obvious. It also happens that we sometimes come to revise or question "test results" themselves, rejecting some data as "questionable" because they conflict with a theory that is very powerful in other respects.

This criterion of evaluation is relevant to logic because there is a sense in which logic makes empirical predictions. Logic investigates arguments that people actually use specialized contexts like mathematics and science, which in turn derive from argumentation found in ordinary life. It is possible moreover to determine empirically which of these arguments people find plausible. These facts of usage, sometimes called *logical intuitions*, are used to evaluate logical theories. Historically a major goal of modern symbolic logic has been to codify the actual deductive arguments found in use among practicing mathematicians.

Conceptual Adequacy

Theories do invent new words but the vast majority of terms used in a theory existed in scientific or every-day language before the theory was conceived. These older terms had meaning and definitions prior to the theory's formulation. It is required that an acceptable theory by and large respect this earlier usage. It is, after all, only by continuing to use old terms in new theories that the phenomena that interested us in the first place gets explained. Thus though an occasional term may receive a new or more precise definition in a new theory, or occur in new and surprising theorems, most terms that have had a prior usage in science or ordinary language should continue to apply to the things they applied to in earlier usage.

This criterion is important to logic, because logic is rich in the terms that come from ordinary life and the history of philosophy. Among these are *individual, property, truth,* and *"the world".* Much of logic's interest to philosophy, and its difficulty, lies in formulating plausible logical explanations that continue to use these words in ways that fit earlier philosophy.

Let us now apply these criteria to the "explanation" of sameness and difference that we sketched earlier that appeals to universals. Historically, critics came to disagree on issues of clarity and parsimony raised by the positing of universals as explanatory entities. Those who believes that universals exist, who think they are real, are called *realists*. Those who do not, who think predicates are empty names that do not really signify anything, are called *nominalists*. A good part of the history of logic, and some of the more famous controversies in philosophy generally, are concerned with this debate. For reasons of parsimony nominalists like Ockham hold that it better to explain truth-conditions in a way that does not posit universals. Modern logic, on the other hand, is distinctly realistic and posits as universals entities it calls sets. The story of how logic arrives at this conclusion takes up Part 1 of these lectures. We start with one of the most celebrated realistic theory, Plato's theory of Ideas, both because it is well know and was historically influential, and because it well illustrates the form assumed by explanations in logic.

Plato's Theory of Ideas

Plato (428 or 427-348 or 347 B.C.) advanced a famous realistic theory that attempted to explain not only sameness and difference – classification and division – but a number of other problems in natural science and philosophy as well. The details matured over his lifetime and are scattered throughout his many works, which take dialogue form. The main points relevant here however are easily summarized. In summarizing his theory I will make use of the standard format used for presenting theories in modern logic, even though doing so will impose some precision on the theory that himself Plato did not provide. This slight historical inaccuracy will make it easier later to compare his theory to those of others.

Plato posits two main categories of entities and a single fundamental relation that holds among them. The first categories is called Ideas (*idea*) or Forms (*eidos*). The second category consists of material objects, which are also called bodies. There are two basic "laws:"

(1) Both Forms and material object possess properties.

(2) Forms and bodies differ in that Forms never change. If a Form ever has a property, it will always have it. A material object, however, is constantly changing. If it ever has a property, it will shortly cease to have it.

The third entity postulated by the theory's ontology is the relation of Participation. Participation is a two-place relation that links a material object or a Form to another Form. That is, a material object may participate in a Form, and one Form may participate in another Form. The central "law" governing Participation is this:

(3) if one entity *E* (which may be either a material object or a Form) participates in a Form *F,* and if *F* has the property *P*, then *E* also has the property *P.*

Plato summarizes this by saying that the participant *copies* or *imitates* the Form. In this sense participation is a synonym for copying or imitation. However, this law interacts with the two previous ones. If the participating entity *E* is material, it only manages to copy a Form briefly or imperfectly because *E* will shortly cease to exhibit any property possessed by the Form. If the participating entity *E* is itself a Form, it will copy the Form it participates in perfectly and forever.[2]

In his dialogue the *Timaeus* Plato even goes so far as to describe an kind of creation myth. In this story, once upon a time, a divine craftsman, the Demiurge, put preexisting inert matter into perpetual motion so that from then on it is constantly changing, first copying ("imitating" or "partaking of") one perfect unchanging Form and then shifting to copy another. The story is not a genuine creation myth, however, because the Demiurge does not at a single instant make the world out of nothing. Indeed, the fact that Plato describes the story as a "myth" suggests to most scholars that he does not believe the Demiurge is real or that there really was an original launching as a historical event. On this reading, the "myth" is a literary device Plato uses to make vivid the relation of matter to Forms.

The basic theory of Forms, matter and participation allows for a simple statement of the truth-conditions for two sorts of subject-predicate sentences: those in which the subject term stands for a material object (or a soul), as in *Fa* , which the reader will recall is the notation of symbolic logic for *a is* F, and general truths of the form *All F are G.* [3]

Plato's Truth-Conditions for Subject-Predicate Sentences

- *Fa* is *true* if and only if the material body (or soul) named by *a* participates in the Form named by F.
- *All F are G* is true if and only if the Form named by *F* participates in that named by *G.*

We may then apply the Theory or Forms to provides the textbook example of "realistic explanation" of sameness and difference in which Forms take the role of universals.

Explanation of Sameness and Difference

- Two material bodies are *the same* with respect to a Form *U* if and only if they may be referred to by two different proper names, say *a* and *b*, *U* may be signified by a predicate, say *F*, in such a way that the sentences *Fa* and *Fb* are then both true.
- Two material bodies are *different* with respect to a Form *U´* if and only if they may be referred to by different proper names, say *a* and *b*, *U´* may be signified

[2] The central theory of Forms as outlined here may be abstracted from the *Phaedo, Republic* IV-VII.

by a predicate, say *G*, in such a way that then either the sentences *Ga* and ~*Gb* are both true, or the sentences ~*Ga* and *Gb* are both true.

The Correspondence Theory of Truth

Plato's is the first of several examples we shall meet of what is known as the *correspondence theory of truth.* Truth is one of the most important ideas in logic. Logical arguments, the main object of study in logic, are defined in terms of truth: an argument *follows logically* (is *valid*) if whenever its premises are true, its conclusion is also true. It follows, then, that to understand validity we must understand truth.

The standard way to explain truth is by a correspondence theory. The general picture is the following. We formulate propositions in language. These propositions have a "form", which consists of the sequential order in which the component words are put together according to rules of grammar. Parallel to the "weaving together" process in grammar, there is a second sort of "weaving together" that takes place in "the world." In the *Sophist* Plato describes a sentence as consisting of a "weaving together" of noun and verb. He also describes the participation relation among Forms as a "weaving together" of Forms. Accordingly, the proposition *All F is G,* which is a weaving together or the words *F* and *G*, is true if it "corresponds" to a weaving together of Forms, i.e. to a process in which the Form named by *F* participates in that named by *G.*

Something to watch as we review the history of logic is how opinions about the grammatical form of subject-predicate sentence matures at the same time as views about truth as correspondence. Later theories reject the details of Plato's theory – for

---

[3] The theory of the intermingling of Forms is most explicitly stated in the *Sophist*.

example, that reality consists of unchanging Forms – but they retain the general notion that truth is correspondence.

Definition and Division, Necessity and Contingency

A second point to make about Plato's theory of truth is that it makes possible scientific classification, or what Plato calls "division" (*diairesis*). Plato assumes that the reality beyond language consists of timeless Forms that stand in fixed participation relations to one another. In the dialogue the *Statesman* he makes clear that he thinks these participation relations form a tree structure, each more general Form serving as a genus that divides into subordinate Forms as subspecies, much as today we divide the biological world into a hierarchy of taxonomic classes.

In the history of logic this theory of division is closely connected to the theory of definition. In modern philosophy it is customary to recognize that there are indeed fixed truths, like the laws of mathematics or technical definitions in science, which are "always" true. These immutable truths are said by philosophers to be *necessary.* A standard definition of a *necessary truth* is one that is always true and could not be otherwise. For example, it is impossible that 2+2 could be other than 4. Hence truths of arithmetic are thought to be necessary. Likewise, it is a matter of definition that a triangle has three sides and that a bachelor is an unmarried male – both favorite examples of philosophers. Plato's theory of Forms, if true, would provide an explanation for such immutable truths because the Forms themselves are supposed to be fixed and unchanging.

However, in modern philosophy we also recognize a second category of truths, those that are not always true but rather change their "truth-value" over time or according to circumstance. Examples include ordinary propositions, both about individuals and about groups, like *Canada is chilly* and *Swans are white.* Climate changes could make Canada warmer, and there could be a black swan, as explores were surprised to find in Australia. Propositions that are not always true or that could be false are said to be *contingent.* Most propositions of daily life and many we meet in science fall in this class. A problem with Plato's theory is that it provides no place for contingent general truths. On his theory any true general proposition of the form *All F is G* is true because it describes immutable facts about Forms and is therefore necessary. Because they are about "ideas", they appear to be more like what we would today call definitions, rather than empirical statements of facts. Something else to watch as the course of lectures develops is how the notions of definition and contingent natural truth mature over time.

Relations

A final logical topic to mention that is addressed by the Theory of Forms is relations. Relations are indicated in English by a variety of phrases that share the general feature that they link more than one proper name. Let us start by considering two-place relations, which are named by phrases that link two proper names. Transitive verbs are examples, as in *Plato teaches Aristotle* and *Socrates loves Xanthippe*. Other examples are provided by verbs that conjoin a subject term obliquely to the object of a preposition: *Plato is sitting next to Aristotle, Theatetus is talking about Socrates.* Comparative adjectives provide a third case: *Theatetus is taller than Plato, Aristotle is stronger than Plato.* Verbal forms like *teaches, loves, is sitting next to, is talking about, is taller than, is stronger than* are called relational predicates and are said to stand for relations, much as common nouns, intransitive verbs, and adjectives are said to stand for properties.

Like a property, a relation can be multiply instantiated. Just as a property can be shared by multiple *individuals*, a two-place relation is shared by multiple *pairs* of individuals. Why is it that the “pair” Socrates and Plato, on the one hand, and the “pair” Aristotle and Theophrastus, on the other, are both the “same”? We can give an answer in ordinary language. It is because they both consist of a cases in which the first is the teacher and the second is his pupil. But what is the “reality” that explains this similarity? Plato is constrained by his metaphysics to explain the commonality by appeal to material objects and Forms. He does so by positing Forms for relations. His examples include Sameness, Difference, and Equality.[4] The relational fact that Socrates and Plato are the same in then “explained” by the fact that they both participate in the Form of Sameness, and the fact that they are different by the fact that they both participate in the Form of Difference. Another way of putting the view is that relational proposition like *Socrates is the same a Plato* means the same as the conjunction of the two subject-predicate propositions *Socrates is the same* and *Plato is the same*, and the relational proposition *Socrates is different from Plato* means the same as the conjunction of the two subject-predicate propositions *Socrates is different* and *Plato is different*. Plato thus initiates a long tradition that was to bedevil logic for two millennia, only to be superceded in modern times, in which unsuccessful attempts were made to represent relational assertions by simpler subject-predicate propositions. We shall see the difficulties such theories face when we discuss more sophisticated versions in next lecture.

Epistemology and Ethics

Much of the attractiveness of Plato's theory is its use it to explain matters beyond logic.[5] Indeed it was issues in ethics and epistemology that were Plato's main concern. For example, he used Forms to explain the central distinction in epistemology between knowledge and opinion. He does so by positing in addition to material objects and Forms an additional category of entity called souls. Souls cycle between an existence in which they are tied to a material body and inhabit the world of matter, and an immaterial state in which they are disassociated from a body and inhabit the world of Forms. Moreover, souls can perceive and note the properties of Forms when in their presence, and remember more or less accurately what these properties are when they later are rejoined to a material body. A soul then knows the proposition *All F are G* if either it is actually perceiving that the Form *F-ness* is participating in *G-ness* or it recalls having perceived it. This is Plato's famous doctrine that knowledge is recollection.[6]

Forms are also used to explain morality, which was the central concern of Socrates, Plato's teacher. Plato holds that if we know what is good we will do it. More precisely he holds that moral virtues are Forms and that if we have knowledge of the Form, we cannot help but act according to that knowledge.[7] This is his celebrated thesis that virtue is knowledge. It follows that the wisest soul is also the most virtuous. To be good, he said, you should study philosophy.

---

[4] Sameness and Difference are discussed most fully in the *Sophist*, and Equality in the *Phaedo*.
[5] For a now classic discussion of the variety of explanatory uses of the Theory of Forms see Harold Cherniss, "The Philosophical Economy of the Theory of Ideas", *American Journal of Philology* 57 (1936), pp. 445-56.
[6] It is laid out in *Meno* and *Phaedrus*.

Because our subject here is logic we will not evaluate Plato's epistemology or ethics, though both knowledge is recollection and virtue is knowledge are highly controversial. We will however criticize the theory on logical grounds.

The Third Man Argument and Parsimony

Let us apply to Plato the general criteria of theory evaluation we reviewed earlier. Consider first mathematical precision. Clearly Plato does not attempt to write his theory as we would today using concepts from mathematics and set theory. It would be unfair to expect him to. On the other hand, his account is precise enough for us to draw out some of its implications. In that sense it does generate a body of sentences that we could call a "theory". It posits basic sets of entities and lays standard properties that are formulated in simple law-like rules. Their statement is clear enough for us to drawn out some of its logical implications and thus to identify at least vaguely a set of theorems. These "laws" and their implications make up the theory we can evaluate.

Is the theory empirically adequate? In this life we cannot approach the Forms empirically because they are not in the sensible world. The best we could do is investigate the way material objects fall in subordination classes. If these correspond to fixed division patterns and correlate to fixed "ideas", then these would be indirect evidence of structure among Forms. Unfortunately, both the evolution of species over time and the difficulty modern biology has encounted in trying to find fixed sets of defining features for species tell against Plato's metaphysics of immutable truths.

---

[7] The role of the Forms in virtuous action is explicit in the *Euthyphro*, *Phaedo* and *Republic* among other dialogues.

Thus, since most of the laws of nature fail to meet the test of immutability required by Plato's theory, the theory fails rather badly as an account of the natural world.

Conceptually it must also be granted that Plato's account put sever strain on the ordinary meanings of terms. In the *Parmenides* Plato himself admits that it is odd to think that there are immutably perfect Forms of Hair-ness and Mud-ness that account for the properties of material hair and mud.

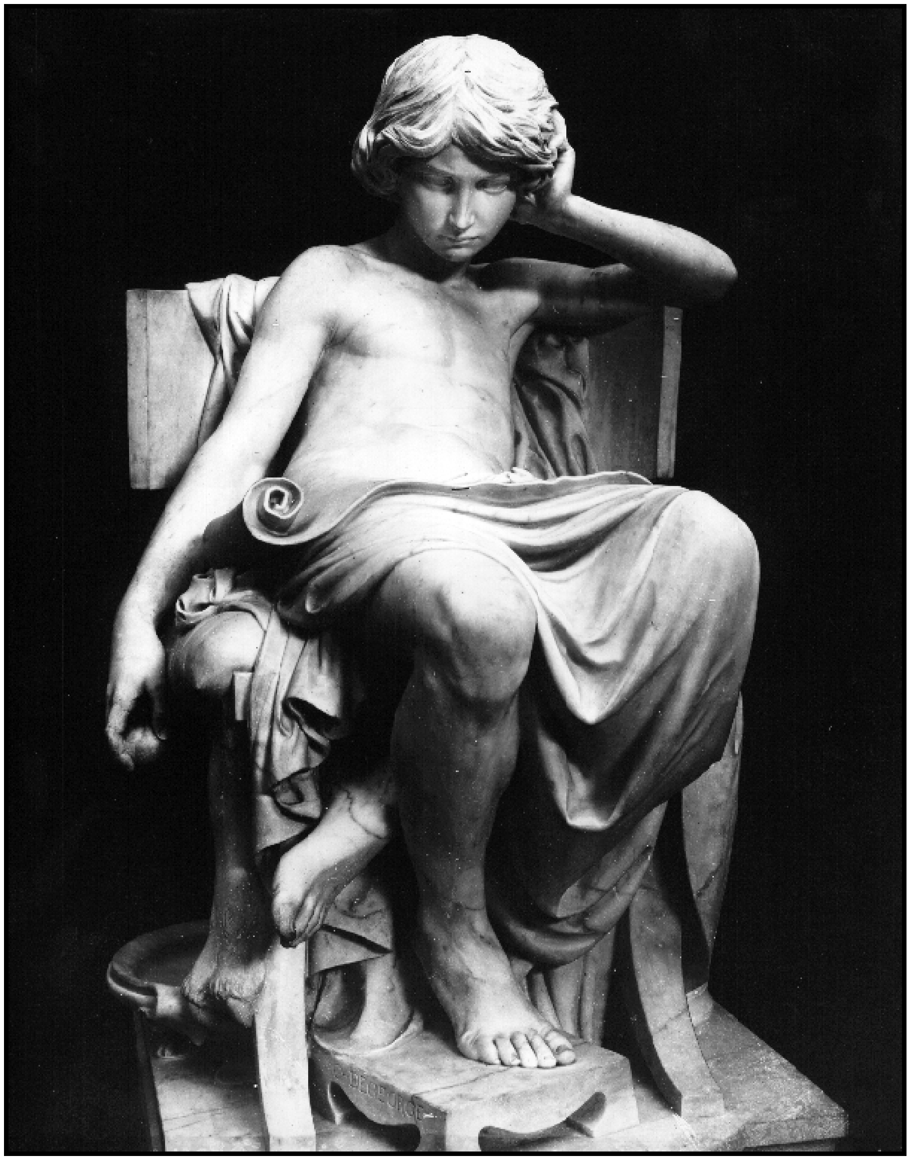
The criterion, however, in terms of which the theory has been most severely criticized is parsimony. Here Plato himself is perhaps his best critic. In his later dialogue the *Parmenides* (123 AB) he advances a famous refutation of his own invention, which is called *the third man argument*. If an entity is *P* because it participates in a Form that is *P*, and if the Form of *F-ness* is *P,* then there must be a form, call it *F´-ness* that *F-ness* participates in such that *F´-ness* is also *P.* But if *F´-ness* is *P,* there must be a further Form, call it *F´´-ness* , that *F´-ness* participates in and is such that it is *P.* The process continues *ad infinitum*. Thus the existence of a single Form entails the existence of an infinite number of Forms in a infinite regress. Plato presents this as a reduction of to the absurd of the theory.

Plato seems to regard the generation of the infinite regress as a serious problem. But the problem may be described in a slightly different way. Notice that the "reason" why a material triangle's participation in the Form Triangularity "explains" the properties of the object is that participation requires that the object resemble the Form. But resemblance means that the two share a property in common. That is, both the material object and the Form Triangularity are triangular. This means that Forms have properties. But if Forms have properties, the problem of why something has a property arises again on the level of Forms. To explain that one thing has a property because it resembles another, its model, that had the property beforehand pushes the explaination of why something has a property onto the model. That is, instead of solving the problem, Plato shifts it to the level of Forms, where it must be solved for Forms themselves. Positing a new level of reality with “models” that themselves possess properties in common merely postpones the problem of explaining what it is to have a common property.

Plato himself stops mentioning the theory of Forms in his last dialogues, a fact that suggests he may have abandoned it himself. Aristotle, Plato’s pupil, parted with his teacher because he found Forms implausible, and philosophers every since have cited Plato as the most egregious example of a realist with a bloated metaphysics. The onus, however, falls on the critics of Plato to provide a better theory, one that explains as much but with a more parsimonious ontology. In the next lecture we see how Aristotle rose to the challenge. In the process he invented logic as a distinct branch of study and contributed an important body of work launching it on its way.

Aristotle's Ontology of Matter and Form

Aristotle (348-322 B.C.), the student of Plato and teacher of Alexander the great, in addition to writing on metaphysics, ethics and natural science, wrote extensively on logic and indeed started it as independent branch of study. His six logical works were edited as a unit after his death. This collection is called the *Organon*, which means "tool" in Greek, because logic was viewed as the preparatory study necessary for more advanced work in philosophy or science. The *Organon* includes the *Categories, De Interpretatione,* and *Prior Analytics*, which concern the logic respectively of terms, propositions, and arguments, and it is this division that provides the format for standard logic texts in the Middle Ages, and for this set of lectures. Three additional worlds, the *Posterior Analytics, Topics,* and *Sophistical Refutations* deal with topics in the theory of definition, scientific method and logical fallacies. It is no exaggeration to say that Aristotle's work formed the core of logical theory until the mid 19th century, when new methods of formal logic were developed to deal with issues raised by advances in mathematics and the natural sciences. The predominance of Aristotelian logic therefore lasted more than two millennia. Views only sketched by Aristotle himself were worked out by subsequent logicians into a remarkably uniform theory, reaching a high point in the Middle Ages. My remarks here will be limited to several themes that have proved important to the development of modern logic. As a result the historical discussion will be somewhat



Degeorge, **Mussé de Luxemburg (© Roger Viollet)**

Aristotle as a Youth

Some relate that when Aristotle went to sleep, a bronze ball was placed in his hand with a vessel under it, in order that, when the ball dropped from his hand into the vessel, he might be waked up by the sound.

Diogenese Laertius, *Lives of Eminent Philosophers*, V. 16

superficial. I hope however that there will be enough detail to interest students in pursuing the topic on their own.

In this lecture we will investigate Aristotle's solution to the problem of universals. In particular we are interested in the theory of truth he proposes for subject-predicate propositions – for it is this that "explains" how one predicate can be true of two subjects or different predicates of the same subject. But to understand the way sentences work, we must first study the words that are used to compose sentences. For Aristotle the words that form subject-predicate sentences fall into two types: proper names, which stand for individuals, and "general terms", which is a broad class that includes common nouns, verbs and adjectives. General terms stand in some sense for what individuals have in common. It is by appeal to the "semantics" of these terms that it is possible to formulate truth-conditions for subject-predicate propositions, and thus explain what it is for individuals to be the same and different.

Like Plato, Aristotle and his tradition advocate a version of the correspondence theory of truth. It presupposes a prior explanation of the grammar of sentences. The grammar specifies the "parts of speech" and how they go together to form subject-predicate combinations. Grammar is supplemented by a rich ontology that divides reality into a series of fundamental entity classes. It is combinations of these entities that form the "facts" that make subject-predicate propositions true and false. The discussion here will be broken down, first into a discussion of the relevant ontology, then of the mechanism whereby the world generates language "in the mind," and finally of language itself, of how it is supposed to reside in the mind and represent truths outside itself.

The theory of terms was developed in the *Categories, Topics,* and *Metaphysics.* Details were elaborated by subsequent writers throughout the ancient period and were summarized by Porphyry (234-305 A.D.) in an influential handbook, the *Isagoge*, that survived the Dark Ages and formed the nucleus for Aristotelian logic in the Middle Ages. This logic, which formed the core of the undergraduate university curriculum in the Middle Ages, was taught widely in standard textbooks, like Peter of Spain's (d. 1277) *Summa Logicales.* In was developed to a high level of sophistication by the masters in the Arts faculties at the great universities of 13th and 14th centuries. Though the quality of research diminished in subsequent centuries, Aristotelian logic continued to be a fixture of university education until superseded by symbolic logic in the 19th and 20th centuries. [8] We begin our account with Aristotle's ontology.

The Categories

Aristotle divides entities into fundamental groupings called *categories* in his introductory book, which is called accordingly the *Categories.* The standard list includes ten groups. First is *substance.* This is defined as including those individuals that can exist in their own right. There are then nine subordinate groups characterized by the fact that they can exist only if they are "in," or as we now say "inhere in," a substance. These are *quantity, quality, relation, place, time, position, state, action* and *passion*.

[8]For an excellent histroy of mediaeval philosophy including Aristotls' metaphysics as it is relevant to logic consult Paul Vincent Spade's *A Survey of Mediaeval Philosophy,* Version 2.0 (August 29, 1985) at Hhttp://pvspade.com/Logic/docs/Survey%202%20Interim.pdfH. For a general account of logic's place in the intellectual life of the Middle Ages see Jacques LeGoff, *Intellectuals in the Middle Ages* (Oxford,

These divisions in ontology are reflected by grammar. Corresponding to the category of substance is the part of speech called proper names, which includes any expression that we use to name an individual, for example, *Socrates, Plato* and *the teacher of Aristotle.* Individual substances are members of genera and species, and corresponding to the genera and species of substances are common or collective nouns, like *man* and *animal.* Corresponding to the categories of non-substances are expressions like verbs, adjectives, and adverbs of quality, quantity, time, place, and manner.

Aristotle groups together genera and species terms, on the one hand, and contrasts them with the so-called the subordinate categories, on the other, and explains the difference at first in a way that sounds as odd in Greek as it does in English. He says that genera and species terms are "said of" a substance, but that the subordinate categories terms are "said in" a substance.

He makes clear what he means by "category" by associating with each category a question of the form, *What is S?* This question, and its reply *S is P,* can mean ten different things. Each way corresponds to a different category.

First of all, by *What is S?* we might be asking *What kind of thing is S?* The appropriate answer is *S is P* where *P* is a *substance* term. That is, by *What is S?* we are asking for the genus or species of the subject.

Alternatively, by *What is S?* we might be asking *What qualities does S have?* In English the natural way to say this is not much different from the first question. The appropriate answer would be *S is P* where *P* is an adjective or other expression that

---

Blackwell, 1993), and Hasting Rasdall, *The Universities of the Middle Ages* (Oxford, 1936, ed. F. M. Powicke and A. B. Emden).

names a quality. In English, for example, we might ask this sort of question by saying, *What is S like?* or *What sort of thing is S?* Greek and Latin actually have special interrogative adverbs especially for this purpose. In Latin *Quale S?* literally means *What is S's quality?*

Again, by *What is S?* we might mean *How much is there of S?* This question would prompt an answer *S is P* in which *P* expresses a quantity, yet a third category. An answer *S is P* in which *P* falls in the category of relation is appropriate when *What is S?* means *In what way is S related to things?* If the question means *When is S?* the appropriate answer uses a predicate from the category of time, and if it means *Where is S?* the right answer would use a predicate from the category of place, and so forth for the other subordinate categories. Each type of category predicate corresponds to a fundamentally different way of inquiring about a subject. The difference between the category of substance, on the one hand, and that of the other categories is that substance predicates describe the broader group of substance in which the subject is included. In Aristotle's terminology this group is "said of" the subject. The other categories describe various sorts of predicates that something can have only if it is first a substance. These predicates then are "said in" the subject.

The Predicables

The ontology of the ten categories is complemented with a cross-classification that divides predicates according to how permanent or characteristic they are of their subject. Aristotle sketches the doctrine in the *Topics* and it is detailed by Porphyry. According to this doctrine terms fall into one of five varieties called *predicables*: genus, species, difference, proprium, and accident.

A fundamental feature of any subject is its species and genus. These are part of its very nature. It could not exist if were not in its species and genus. Hence a predicate that describes something's genus or species is definitive of it. It holds by definition. It always holds, and it holds of every member of the species and genus. Species and genus predicates form the first variety of predicable.

The second sort consists of predicates that describe those qualities or other non-substance terms that are true of the subject by definition. The function of a term of this sort is to distinguish one variety of substance from another. Aristotle calls it a difference (*differentia* in Latin). For example, the predicate *rational* names a quality "said in" Socrates that is true of him by definition because the definition of Socrates' species, which is *mankind,* is *rational animal.* Like predicates that stand for genera and species differences are true of their subjects necessarily and by definition, but they are less fundamental because unlike a substance a non-substance has no independent existence. It can only exist "in" a substance.

Aristotelians also believed that there were some qualities (and other non-substance categories) that are "in" a subject necessarily but are not actually part of its definition. Such a predicate holds of all members of a species because it is a law of nature that anything with the species' properties would also exhibit this category trait. For example, Aristotle says humans are *risible*, i.e. capable of laughter. Risibility is not part of the definition of *mankind,* but it law of nature, he thought, that any rational animal must also be able to laugh. Such traits are said in the subject necessarlity but are not true by definition. Necessary non-definitional features are called *propria* (singular *proprium*).

Lastly, there are non-substantial terms that are not necessary. These are either not always true of the subject, or are true of some members of its species but not all. These are called *accidents*. A person's height, color, mood, wealth, location, age, position are accidents because they are not permanent. A subject's definition determines its nature and this in turn restrict the possible accidents the subject might possess. Humans, for example, can sit or stand, both accidental features, but they cannot melt or dissolve, which are accidents of substances like ice cream.

Matter and Form

In the *Metaphysics*, a later work, Aristotle explains more fully the relation that holds among the categories and the predicables. The standard interpretation of his view was that an individual substance is a composite entity made up of two aspects called *matter* and *form*. The form of a species is its definition, i.e. its difference and genus. A statement that defines a species in terms of its genus and difference is called a *real definition* because it records what differentiates the species in reality from other members of the genus. A species form is also called its *nature* and its *essence*.

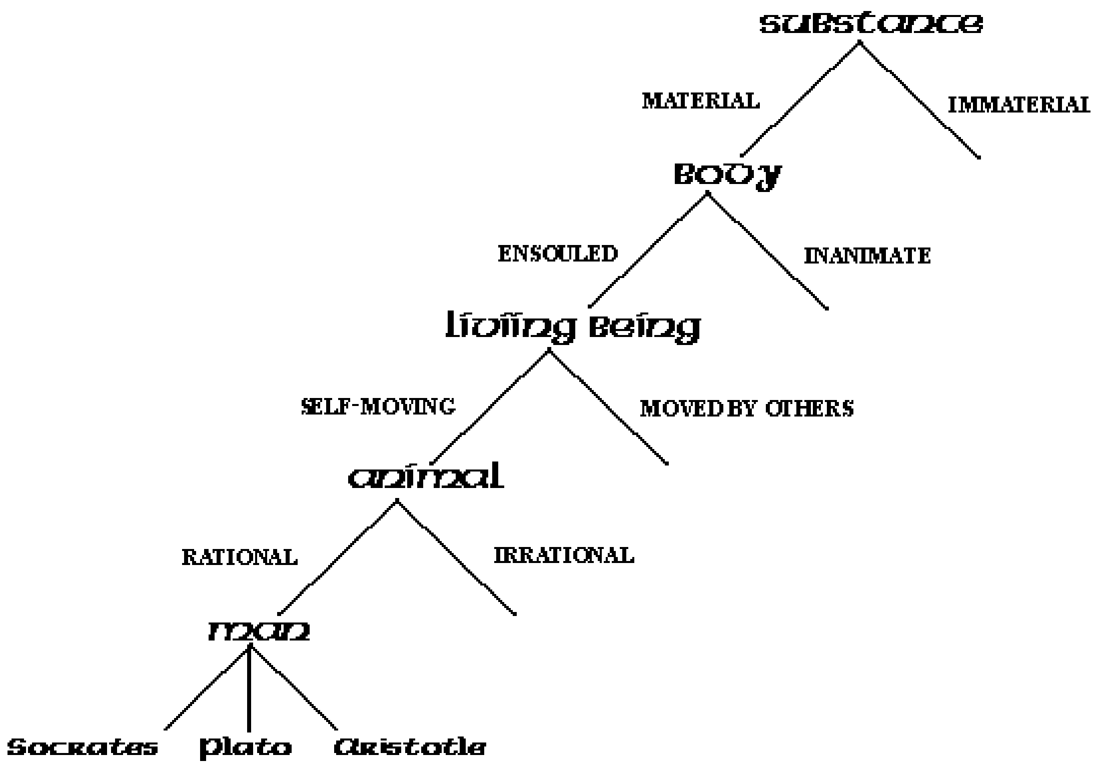
The Standard Form of a Real Definition

$$\text{species} = \overbrace{\text{genus} + \text{difference } (\textit{differentia})}^{\text{Essence, nature, "form"}}$$

$$\underbrace{\underset{\textit{definiendum}}{\textbf{Man}} = \underbrace{\textbf{Animal} + \textbf{Rationality}}_{\textit{definiens}}}_{\textit{definitio per genus et differentiam}}$$

As Aristotle describes genera and species they fall into a tree structure, each genus dividing into a series of species, often two but sometimes more, each species having a difference appropriate to it and to it alone. Porphyry's account of "the tree" became well know.

The Tree of Porphyry



The full classification of the individuals Socrates, Plato, and Aristotle into the hierarchy of **genera** and **species**. The **difference** in terms of which a species is defined within its genus is indicated above the species name **IN SMALL CAPITOLS**. It is possible to read off from the Tree the full **essence** or **form** of an individual: Socrates is a rational, self-moving,ensouled ("animate"), material substance.

Two properties of real definitions must be stressed. First, according to Aristotle's metaphysics they are necessarily true. They record the necessary features of a species that are required by nature. Secondly though real definitions are necessarily true, they are not easy to discover. Scientific classification of the sort reflected in the Tree of Porphyry is the result of difficult scientific research. Moreover, even when discovered, our knowledge of it may still be provisional. Hitting upon a definition of this sort, and even being right, does not automatically bring along with it the certainty that we can attach to other sorts of knowledge, like the truths of mathematics.

Aristotelians believed that, strictly speaking, it is species not individual members of species that are defined. Individuals, after all, come and go and hence cannot possess any property necessarily. Species, on the other hand, were regarded as eternal, having neither a beginning nor end because in Aristotle's view the world is eternal. It had no beginning and will have no end. Moreover the species that now exist have always existed and always will. (Accepting this view posed a difficult for Christian logicians of the Middle Ages who accepted most of his other views, and for later proponents of evolution.) An individual then possesses a nature or definition only derivatively and to the extent that it is a member of a species.

It should also be remarked that Aristotle did not think that all the defining features appropriate to the species appear in the individual of the species at the same time. An oak tree develops acorns as a matter of definition, but only when it is mature. Likewise human babies take a while to become rational. That is, some necessary features are dispositional. They manifest themselves only if the appropriate circumstances obtain.

In Aristotelian terminology a substance that changes is said to be *matter* for the substance at its later state. This terminology is used to explain what is meant by matter and form. A substance is a composite of matter and form in the sense that it is the result of another substance (its matter) changing so as to instantiate the defining features of a species (its form).

Concept Formation: Perception and Abstraction

Aristotle and his followers exploited the ontology to explain how it is we obtain the language by which we think about the world. The account is relevant here because it explains how words “stand for” things. It explains the mechanism of correspondence at work when propositions are true and false. It is the Aristotelian theory of meaning.

First of all, we must say something about the human “mind.” To an Aristotelian, humans are just substances like any other. That is, they are combinations of matter and form. Their knowledge then must be explained using the categories of Aristotelian ontology, i.e. as some sort of inherence of non-substantial categories in a human substance. Since most knowledge is accidental these properties must be accidents of some sort. For an Aristotelian, then, there is no special entity called “the mind” above and beyond the substance that makes up the human being. Even though Aristotle and his followers talk about the soul, and about the soul having properties, strictly speaking, this is a somewhat misleading way of speaking. Mental life is explained as qualities, quantities, etc. being instantiated in the human substance. It became the usual practice to refer to instantiations of this sort as “mental acts.” This terminology too is a bit misleading. It does not mean that there is a special entity, an "action," that exists above and beyond the human being performing the act. It means rather that certain qualities or other accidents, which we regard as mental or intellectual, are *actually* instantiated in the human substance.

What then happens to us when we "see" that the rose is red? Aristotle's story, as recounted in the *De Anima*, goes this way. Redness, a quality, inheres in the rose, a substance. As a general rule, Greek philosophers, including Aristotle and Plato, believe that the way one substance causes another to have a quality Q is that one substance must have Q in the first place, and the causal process consists of passing Q on from the first substance to the second, like a baton in a relay race. Now, the rose passes on the information that it is red. It does so by affecting the sensory medium between itself and the perceiver. In this case the medium, i.e. the physical entities standing between the perceiver and the thing perceived has three parts: (1) particles of air, (2) the sensory organs of the perceiver, which in the case of sight are the eyes, and (3) the bodily parts that stand between the sense organs and the organ of "understanding," which Aristotle thinks is the heart. The rose passes "redness" first to the nearest bit of air; it in turn passes it to the next until it reaches the sense organs; these pass it on in order to the various bodily parts until it reaches the heart. Redness, however, once it leaves the rose is not redness in its ordinary sense. Though it is genuinely a quality and inheres in the intervening substances, it does not make them red. Neither the air itself nor the eyes nor the nerves become red. Rather, redness exists in a reduced or secondary sense. In the Middle Ages redness in this sense was said to be *intentional.* Today we would probably say that what is transferred is a physical change that constitutes "information" that the rose is red.

In any case the causal transfer of the intentional-quality-redness transforms the perceiver so that the quality of intentional redness inheres in him or her. Humans have as part of their nature the feature that when an intentional quality comes to inhere in them they become *aware* or *conscious* of it. This consciousness of the redness is likewise an accidental quality of the perceiver. In a somewhat misleading way, Aristotle and mediaeval philosophers regularly referred to these as qualities inhering in the perceiver's intellective *soul*, but as explained earlier, this way of talking is really just a way of saying that certain qualities, which we happen to regard as intellective or rational, inhere in the perceiving human being.

A special feature of the rose's redness is that it is particular to that rose. If I perceive the redness of a second rose, it might be slightly different. Likewise, I can smell the particular scents of several roses, and see their particular shapes, all of which are slightly different. When I perceive many roses, a large variety of particular sensory qualities will in inhere in me ("in my soul"), and I will be conscious of that fact. Aristotle called this stage of the perceptual process *intuition*.

Recall that it is a feature of Aristotelian ontology that substance can take on only the accidents appropriate to that substance. Nothing can instantiate the accidents appropriate to a member of a species unless it instantiates the formal features of that species. Therefore, I cannot instantiate the *accidents* of various roses unless I also instantiate the *defining features* of the species rose. This does not mean that I myself have to become a real rose, because the various rose features are only in me intentionally. But it does mean that I have to instantiate the species features of a rose intentionally. The upshot of these facts is that my perception of multiple sensory qualities of numerous roses is possible only if I also instantiate intentionally the species form of the rose. Moreover it is part of human nature that when the intentional form of the rose is instantiated in me, I become conscious of the fact. Aristotelians give the name *abstraction* to the process of instantiating consciously an intentional species form that happens as a result of intuiting multiple individual intentional sensory qualities caused by various different members of the species that affect the body's sensory organs. I sense many roses. Their various individual sensory qualities travel in an intentional mode to my sense organs and then to my heart. I become aware of them. But they, as it were, bring along the form of "rose" itself, because I could not instantiate these rose accidents if I did not also instantiate, at least intentionally, the form of rose itself. I become conscious of possessing the form of rose intentionally, and voilà, I have abstracted a "thought" of a rose. Notice that having an abstract thought or idea, then, is ontologically nothing more than a substance, namely the perceiver, instantiating intentionally the formal qualities of the species perceived. Mediaeval logicians use the term *concept* as a name for an intentional form instantiated in the perceiver's soul.

Mental Language

In the latter part of the 13^{th} and beginning of the 14^{th} centuries mediaeval logicians like William of Ockham and John Buridan (1295/1300-1358+) developed the theory of abstraction into a full blown account of thought and language.[9] Language and understanding in their view is a mental phenomenon. By this they mean not that it is a special sort of entity, called the mind or soul, that is above and beyond the human substance – even though they did believe that some aspect of human personality survived death. Rather, mental acts in Aristotelian terms consist of the instantiating of intentional qualities in a human being considered as an Aristotelian substance. The qualities in question are concepts, which are understood as intentional forms abstracted through the perceptual process.

These conceptual instantiations, they suggests, are the "words of the language of thought." Abstracting thoughts is a prerequisite to instantiating further mental acts that constitute constructing, asserting, and believing propositions. My thinking of the proposition *All S is P,* then, is simply the instantiating in me as a thinker (the instantiating of the quality "in my intellective soul") of a quality we regard as mental, i.e. that of forming a proposition. It is not unlike my performing a ballet. My having danced in this particular way is a quality of me as a substance. It is a quality that manifests itself in my acting in a certain way. It is regarded as artistic. Therefore my dancing it is an artistic act. Moreover, it presupposes that I perform other acts in the process, e.g. standing on my toes and jumping in the air. Ontologically these subordinate acts too are simply instantiations of artistic qualities. My thinking of a proposition is a mental act that consists of instantiating mental qualities and that presupposes other mental acts, like concept formation, which is also an instantiation of a mental quality, just as my dancing is an artistic act that is an instance of an artistic quality and presupposes other artistic acts, like dancing *en pointe* and jumping, which are also instances of artistic qualities.

[9] Paul Vincent Spade, *Thoughts, Words and Things: An Introduction to Late Mediaeval Logic and Semantic Theory,* Version 1.1a, at http://pvspade.com/Logic/docs/thoughts1_1a.pdf.

In this way language is both “explained” in the categories of Aristotelian ontology, and – this is the clever part – the relation of thoughts to the world is explained. The medaievals coined the term *signification* to describe the link between a concept and things in the world. A concept *signifies* all possible objects that could have caused it to exist in the perceiver (“in his intellect”) through the process of abstraction. Because causation figures in this definition of signification, philosophers call this a causal theory of reference.

Notice that according to the definition, a “mental word” signifies more than the individual or individuals that actually figured in the causal process that lead to its abstraction in a particular thinker. It signifies all individuals that *could* have caused it. This broad group includes not only the individuals that did cause it, but others that, because they are of the same species and have the same form, could have caused it. This group includes all objects in that species, both that actually exist and that, though they do not exist, could exist. Possible individuals are included because it is true even of possible members that they *could* have caused it – that they *could have* is, after all, exactly what is meant by saying they are “possible” members of the species. Thus, a concept that intentionally instantiates a species form *F* signifies all *possibilia* that fall in the species *F.*

Truth is explained in terms of signification. Common nouns are concepts that, as just explained, signify all possible individuals that could have caused their intentional instantiation in the thinker. Proper nouns, which signify just a single individual, are treated as a special case of common nouns, ones that happen to signify just one possible object, namely the single entity that could have cause the name to be abstracted.

In addition to common and proper nouns, there are further classes of non-individual terms that are associated with the non-substance categories. These are the terms associated with the non-substance categories of quantity, quality, relation, time, place, action, passion, etc. Their signification is more complex. They were said to *signify* all possible substances in which they might inhere, and to *connote* the associated quantity, quality, relation, time, place etc. that is instantiated in all the individuals signified. Such non-substance terms are accordingly called *connotative*.

With these definitions we can now state the truth-conditions for subject-predicate sentences. The brief formula customarily used is that a proposition *All F is G* is true if the subject and predicate "signify for the same." This idea is the same as that found in modern mathematics when we say, using the notation of set theory, that $F \subseteq G$ if and only if $F \cap G = F$. Here *all F are G* is written $F \subseteq G$, and this holds exactly when $F \cap G$ is the same as F. (Here $\subseteq$ indicates the subset relation and $\cap$ the intersection operation.)

Truth-Conditions for Subject-Predicate Sentences

- *All F is G* is true on an occasion of use if and only if all the actual objects signified by *F* and *G* on that occasion are the same as the actual objects signified by *F.*

- *Fa* is true on an occasion of use if and only if all the actual objects signified by *a* and *G* (i.e. all the actual *G's* that are *a* – there will only be one) on that occasion are the same as the actual object signified by *a.*

It follows then that the same predicate would be true of two different actual subjects – that the two subjects would be "the same" – if the predicate was a general enough concept to signify both the actual entities signified by the subject terms. Likewise a second predicate could simultaneously be true of one subject term but not the other if it included the one but not the other in its signification class.

A close inspection of the definition of signification, which is used in the truth-conditions, shows that it seems to assume the existence of universals. Concepts for example are supposed to instantiate intentional qualities, which appear to be universals, and connotative terms appear to connote non-substances, like quantities, qualities, and relations, which also appear to be universals. We shall return to this ontological issue shortly, but first we must comment on relations and on how mental language is linked to speech.

Relations

In Aristotelian ontology relations are one of the non-substance categories of entity that are "said in" a substance. Like qualities and the other non-substance categories they are spoken of in language by means of connotative predicates. The relational fact that Philip is the father of Alexander must accordingly be unpacked as some fact described in subject-predicate sentences that make use of relational predicates as applied to the subject terms *Philip* and *Alexander*. Such sentences would be propositions of the form *Philip is* ____ and *Alexander is* ____ . Moreover, as a group they must contain all the information of the proposition *Philip is the father of Alexander*.

It must be admitted that traditional logic never succeeded in showing how this could be done. Bertrand Russell (1882-1970), a patriarch of formal logic, claims that any such theory ends up forcing upon relations some logical features they do not generally have.[10] Let us abbreviate the assertion *a bears the relation R to b* by *aRb.*

Suppose that some such translation of *aRb* into subject-predicate sentences is possible. Suppose, for example, that any relational proposition *aRb* is really a translation of two subject-predicate sentences with the same predicate *F*. Suppose that *aRb* means the same as *Fa* and *Fb*. It then follows that the relation *R* must be symmetric. Here is the proof:

> Proof: Suppose *aRb.* Then by definition *Fa* and *Fb*. It follows then that *Fb* and *Fa*. Hence by definition *bRa.* Hence, if *R* holds in one direction, it holds in the other. That is, *R* is symmetric.

But many relations, for example *love*, are not in general symmetric.

Another proposal is that *aRb* is really equivalent to two subject-predicate sentences with different predicates, for example to *Fa* and *Gb*, for some *F* and *G*. It then follows that *R* is transitive. Here is the proof:

> Proof. Assume *aRb* and *bRc*. Then by definition *Fa* and *Gb*, and *Fb* and *Gc*. Hence, *Fa* and *Gc*. But this means by definition *aRc*. Hence if *aRb* and *bRc*, it follows that *aRc*. That is, *R* is transitive.

But many relations are not transitive. If *a loves b* and *b loves c,* it is not unlikely that *a hates c.*

One of the more serious attempts to provide subject-predicate accounts of relational propositions was made in the Enlightenment by Leibniz (1646-1716). His account is complex, requiring reformulation of a relational assertion into a long list of simpler ones. It will have to suffice here merely to assure the reader that in the end the attempt is not plausible.[11] One of the successes of modern logic is that, as we shall see, it provides a more successful account of the logic of relational sentences.

Spoken Language; Real and Nominal Definitions

The mediaevals held that we thought using concepts. Indeed, we all think using the same concepts no matter what language we happen to actually speak because the concepts are caused by the perceptual-abstractive mechanism that depends on what species there are in the world, and has nothing to do with our culture or what the spoken sounds of our language are. Spoken language, on the other hand, was believed to be a matter of convention. A given culture makes arbitrary conventions about what sounds to use to talk about the concepts it finds provided by nature "in the mind." We might have decided to use the "word" *cheval* like the French to name the concept horse, or *pferd* like the Germans, but we in fact decided to use *horse*. Ontologically, my saying out loud the sound "horse" is the instantiation in me of a physical quality, the property of making a certain noise. It is quite different from my instantiation of the intentional form of the species horse. Nevertheless, I and all others who speak English have learned to use the first of these acts to represent the

[10] Bertrand Russell, *Principles of Mathematics* [1902] (N.Y.: Norton), section 212.
[11] See Massimo Mugnai, *Leibniz' Theory of Relations* (1992)

second. The rule that states that we should attach the utterance of a certain sort to a given concept is called a *nominal definition*.

Nominal definitions then are arbitrary matters of convention. As such, we can be fairly certain that they are right. It is ourselves, after all, that by an act of the will decide what sound to use to stand for what concept. As a result there is a kind of automatic certainty attached to verbal conventions of this sort that is not common to most of our knowledge about the world. The arbitrariness and certainty of nominal definitions make them very different from what the mediaevals called real definitions, which are the rules that lay out the genus and difference of a species. Real definitions are necessary truths of nature that are hard to discover, and even when we do, we may not be certain we are right.

The Realism Nominalism Debate

As we have seen, mediaeval logicians made use of Aristotle's ontology. Some of the best of them had serious questions about universals. Concepts are instances of intentional qualities, and connotative terms connote non-substance universals, like quantities, qualities, and relations.

But many had serious doubts. These doubts centered on the precise formulation of what a universal is. Boetius (d. 525 or 526 A.D.) clarified the mediaeval understanding: a *universal*, he says, is an entity that enters wholly and completely into a composite to form a unit and does so with more than one individual at the same time.[12] William of Ockham, one of the great champions of nominalism, notes that if Socrates were to be annihilated so presumably would all the parts that make him up, including the universal humanity if it is wholly and completely “in” Socrates. God, Ockham observes, is omnipotent. Thus, he could certainly annihilate Socrates and all his parts and at the same time spare Plato. But if God annihilates Socrates together with his parts, he annihilates the universal humanity, which is wholly and completely “in” Socrates. But the universal is in Plato as well. Hence a key “part” of Plato would cease to exist, and so too would presumably Plato himself. Hence contrary to the supposition of omnipotence, God cannot both annihilate Socrates and not annihilate Plato. Hence the theory of universals entails that God’s power is limited and that he is not omnipotent. Ockham regarded this argument as a reduction to the absurd of the theory of universals.

What ultimately turned logicians from Aristotelian logic was not parsimony but other issues. We have already seen that it does not explicate relational assertions very well. In the 19th century logical difficulties in mathematics compelled logically minded mathematicians and mathematically minded logicians to jointly attempt to develop logical theories that were capable of expressing the complex mathematical formulas of the period. A much more complex grammar was required than that of the simple subject-predicate propositions we have been discussing. Because the new logic was to be used in mathematics it also had to meet the requirements of mathematical rigor.

The theory of evolution and advances in chemistry and physics had also made clear that Aristotle’s ontology of fixed necessary species divisions was implausible. With the rejection of his basic ontology of form and matter, with its distinction between

---

[12] On Boethius on universals consult Paul Vincent Spade, "Boethius against Universals: The Arguments in the *Second Commentary on Porphyry*", http://pvspade.com/Logic/docs/boethius.pdf

necessary and accidental properties, his theory of perception, abstraction and mental language also fell. Indeed the increasing scientific rigor of the period made clear how little we knew of the mind and how we think, or of language and its basic mechanisms. The view called *psychologism*, that logic concerns the forms of thought, was rejected in favor of a more objective scientific project: the identification and explication of the logically valid arguments. On this view, there are logical facts about what arguments are valid, just as there are truths about what formulas of algebra are true, or what laws of physics. It is the logician's task to find and explain them. The theory of truth remains central because a valid argument necessarily takes one from true premises to a true conclusion. But the focus of study turned from the psychology to the study directly of grammatical forms and how they correspond to the world. In later lectures we shall study this new grammar.

But any theory of correspondence even within a more complex grammar was still going to face the task of explaining its simpler sentences, including subject-predicate sentences. They were not going to disappear simply because we also wanted to write longer formulas too. Moreover, as long as there are subject-predicate sentences there will be a problem of sameness and difference. The logicians had to face the problem afresh but without the help of Aristotle's ontology.

By a happy coincidence, set theory was being developed by mathematicians like Georg Cantor, who was interested in it as a tool to understanding problems in number theory and the theory of the infinite. It proved to be an ideal tool for the logical projects of the day. One of its key contributions to logic is that it provides an ontology suitable for explaining the problem of universals.

Let us turn then to the theory of sameness and difference as seen from the perspective of set theory. Working through the basic ideas of set theory will also provide some seat-of-the-pants experience for students new to logic, so that they can experience what it is to think an issue through logically and construct a simple proof. This intuitive experience is really a prerequisite for appreciating the power of the symbolic techniques for constructing proofs that we shall meet in later lectures.

Lecture 3. Naïve Set Theory

The Motivation for Set Theory

Set theory is regarded as superior to Aristotle's theory of form and matter in two ways. First of all, it is not committed to Aristotelian essentialism, the view now rejected by natural science that species possess necessary defining features. Set theory explains sets, which are not part of the physical world like atoms or volcanoes, but are rather abstract mathematical entities. Accordingly set theory itself makes no claim whatever about the natural world explained by the empirical sciences. Thus, set theory is not committed to Aristotle's account of perception and abstraction, or to the mediaeval doctrine of mental language, both of which are accounts of natural processes. As we shall see in later lectures, sets are used in empirical theories, including scientific accounts of language, but the naturalistic theories are applications of set theory, much as physics makes use of applications of geometry or economics of calculus. Set theory as a theory of abstract entities, as a "mathematical" theory in this sense, is prior to and independent of such applications. It stands on its own as an account of what sets are.

Secondly, set theory is superior to the Aristotelian account in that it can actually explain its fundamental ideas in a way the Aristotelians could not. The mediaeval account explains sameness and difference, language and thought, etc. by appeal to the catalogue of entities listed in the Aristotelian categories. But it does not define these in any serious way, nor does it state precisely the "laws" that they obey.

Though we were able in the last lecture to sketch their relations, the picture has many obscurities.

What, for example is a difference? Is it an entity from one of the non-substance categories, like a quality, so that it is "said in" a substance? If so, is form then reducible to the list of the necessary differences that characterize a substance's genera and species? If so, how does form differ from genera and species, which are "said of," not "said in," a substance?

The matter of a substance is a temporally prior substance that changes so as to become the new substance. More precisely, the matter of a substance consists of the substantial parts that come together to form the new substance. But how deep does the part-whole relation go? It was a standard doctrine that the part-whole relation "bottoms out" at a fundamental level of substantial parts, which was called *prime matter.* This was the basic "stuff" out of which the smallest substances are composed. But prime matter is a mysterious entity because it does not fall in any of the categories.

Obscurities of this sort derive from the fact that the theory is not very well worked out. Set theory, on the other hand, is mathematically rigorous, and states its claims with admirable clarity. The contrast can be made in terms of the basic elements of an explanatory theory, which were set out in Lecture 1. As described there, a theory has three parts: a list of its basic ontology, which consists of a classification of the entities the theory assumes to exist, a list of relations that hold among these entities, and a statement of the "laws" that explain these relations. Aristotle provides the lists, but does an imperfect job of providing the laws.

Set theory does a better job. Its ontology is simple. It presumes only that sets and their elements exist. It also posits a minimal list of two basic relations: identity and set membership. But it is in the statement of its "laws" that the theory is strikingly superior. They take the form of an axiom system.

Early forms of axiom systems have been employed in mathematics since Euclid's *Elements* of the 2^{nd} century B.C. Their modern format was developed with the advent of symbolic logic in the 19^{th} century. An *axiom system* has three parts: a set of *axioms*, a set of *inference rules*, and a set of *theorems* that are deduced from the axioms by careful proofs. The system is usually supplemented with a set of *definitions* that allow for the abbreviation of longer expressions by more familiar shorter ones. In this lecture and later we shall state and prove theorems in a simple axiom system for naïve set theory, so that the reader may learn something about set theory and at the same time gain first-hand experience with an axiom system.

Abbreviative and Implicit Definitions

Before we start, however, it is important to draw attention to two rather different ways in which an axiom system "explains" basic words or concepts.

Abbreviative Definitions

One way is by a direct definition. As we shall see, some terms are explained by being given an explicit definition in terms of other more basic expressions of the theory. In the context of an axiom system definitions of this sort are abbreviations. The axioms are stated first in terms of the basic vocabulary of the theory. None of these terms receive an explicit definition. They are said to be *primitive,* and the undefined language of the axioms is called *primitive notation.* Given the axioms and rules, it is then possible to deduce theorems. Sometimes these theorems get rather long. It is often possible and informative to abbreviate some of this longer notation by shorter expressions. The declarations describing such abbreviations are called *abbreviative definitions*. The definition has two parts: the term to be used as the abbreviation (the *definiendum*) and the longer expression it abbreviates (the *definiens*).

Officially, from the point of view of logical deductions, any abbreviated item can be eliminated. It is simply replaced by the longer expression it stands for. In this way, by the progressive elimination of all abbreviations, any formula can be translated back into one written entirely in primitive notation. In this sense the abbreviated terms are not really part of the basic axiom system.

Abbreviative definitions, nevertheless, play an explanatory role. Usually the term used as an abbreviation has a prior usage in ordinary language or in science. At the same time, the words that are being abbreviated often have a previous usage. That is, the term being defined and the terms used to define it are ones we already have opinions about, including views about what they mean. When this is true, it is required that the abbreviation conforms to or clarifies this past usage. It would not do, for example, to abbreviate the expression *three sided plane figure* by the term *duck* because past usage does not use the term *duck* to mean anything like a triangle. Accordingly, the scientific community that is responsible for developing the axiom system for a body of science is constrained to build into it only definitions that are plausible. Often such definitions are also quite illuminating as well because, given the definition, it is then possible to prove theorems using an abbreviation that shows how the abbreviated idea relates to the other ideas in the theory. In the axiom system below a number of ideas are given abbreviative definitions, for example *subset.* You should cast a skeptical look at these definitions to see if they do a good job of capturing what you think the abbreviated terms mean in ordinary language. We shall then use these abbreviations to prove theorems, and discover how the defined terms relate to one another.

Implicit Definitions

In addition to explicit definitions, there is another way in which an axiom system "explains" its terms. It is this technique that is used to explain those terms that occur in the axioms. Because they are in the axioms, they are part of the primitive notation and do not have explicit definitions. These terms, rather, provide the basic vocabulary that makes abbreviative definitions possible. But if a term is being used to write a basic assumption, how is it that the term itself is to be "explained"? The usual answer is that the axioms themselves in a sense "explain" the terms used to write them. The rationale goes this way. The axioms and rules together determine all the theorems that can be proven from them. That is, they determine all the truths that can be stated using the primitive terms. What more could you want in the way of explaining a term? It remains true that the term must be used in a way that fits ordinary usage. If the axiom system uses a familiar term in one of the axioms, then it cannot use it in a way that violates the ordinary meaning of the term. The axiom *every duck is divisible by* 2 is an unacceptable axiom because in the ordinary language *ducks* does not apply to the sort of thing that is divisible by a number. Axioms that are consistent with ordinary language, on the other hand, can be quite informative about the ideas they assume. The axioms are therefore said to provide an *implicit definition* of the terms they contain. Newton's theory of mechanics is axiomatized by his four laws of motion, which presuppose various terms as part of their undefined vocabulary. They are said, for example, to implicitly define the notions of Newtonian time and space because they use variables that range over distances and durations. In the axiom system for set theory that we shall sketch in this lecture, the symbols $\in$ for set membership and = for set identity will be used in the axioms and be implicitly defined by them. It is in this sense then that the theory will explain its fundamental ideas.

The Axioms of Naïve Set Theory

Sets were studied intuitively in the 19th century by Georg Cantor (1845-1918) and later axiomatized by Gottlieb Frege (1848-1925). A simplified account designed to highlight the central ideas was provided shortly afterwards by Bertrand Russell.[13] This is the version we shall review here. It is now called *naïve set theory*. It contains just three axioms. The first is an axiom that occurs in every axiom system in mathematics and science. It says simply that every truth of logic may be written down as a theorem in this axiom system. The axiom insures that all the truths discovered in the more basic science of logic can be carried over into the new system. It is the next two axioms that lay out the basic properties of sets themselves. They are written using the "primitive notation" of set theory, $\in$ and =. Strictly speaking, it is the axioms

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

themselves that are supposed to explain (”implicitly define”) what these symbols mean, but let us start by translating them into English.

Set Identity and The Principle of Extensionality

The symbol = is familiar. It stands for identity between sets. Axiom 2, called the *Principle of Extensionality*, lays out the identity conditions for sets. Philosophers sometime require “identity conditions” as a necessary requirement for an acceptable ontology. They admonish, “No entity without identity.” This axiom satisfies that requirement. More precisely, the axiom sets out the conditions under which two names stand for one and the same set. The stand for the same set if the sets they name have the same members.

♦Axiom 2. The Principle of Extensionality

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

Simply put, two sets are the same if and only if they have the same members. (The ♦ is used to indicate axioms, definitions, rules, theorems etc. that are important in the sense that they are presupposed in the later lecture material.)

The axiom may be formulated in terms of a name’s extension.[14] The *extension* of a set name is simply the set that it names. The axiom says that two names form a true identity sentence exactly when they have the same extension, i.e. exactly when they stand for the same set. It is this formulation in terms of extension that gives the axiom its name.

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

The axiom is obviously true, but a related principle for Platonic Forms and Aristotelian universals is false. It is not true, for example, that two Forms are identical if the same bodies imitate them, or that two Aristotelian qualities are identical if they are "said in" the same substance. The very same things could be both red and square but that would not mean that the Ideas or qualities of redness and squareness would be the same, because understanding the one would not entail understanding the other. Unlike sets, Forms and qualities need not be identical when they have the same extensions. It is not at all clear in fact how to state identity conditions for Forms or qualities. This imprecision is a major weakness in the theories of Plato and Aristotle because it prevents them from achieving the ideal of clear mathematical formulation.

Set Membership and the Principle of Abstraction

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

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

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

Socrates is a human

Socrates is a member of the set of humans

Socrates $\in$ *the set of humans*

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

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

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

To say this in an axiom we must introduce some notation to represent a sentence that talks about *x*.

First we must explain what a variable is. The letters we shall use as variables are *x,y* or *z.* They function as a pronoun that stands for sets or for the element contained in a set. In ordinary grammar, in order to know what a pronoun stands for, its antecedent must be fixed by the sentence in which it occurs. Similarly, if a variable is used in a formula in which it does not have a fixed antecedent, we cannot know what set or element it is supposed to stand for. A variable without a fixed antecedent is said to be *free.* For example, in the sentence *it is red* and the formula *x is red,* it is not possible to determine an antecedent for the subject term. Thus, in *x is red* we say that the variable x is "free."

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

We may now state the Principle of Abstraction.

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

It is useful later when constructing proofs to state the axiom more succinctly. To do so we introduce some abbreviating notation. First is the standard notation for the quantifier expressions *all* and *some*:

for all x	is abbreviated by	$\forall x$
for some x	is abbreviated by	$\exists x$

We will also use symbols for *not, and, or, if…then,* and *if and only if,* which are called the *sentential connectives*. We have already met ~ for negation. I this lecture we shall also use the ampersand & for conjunction (*and*), the symbol $\vee$ for disjunction (*or*), $\rightarrow$ for the conditional (*if…then*), and $\leftrightarrow$ for the so-called biconditional (*if and only if*):

P or Q	is abbreviated as	$P\vee Q$.
if P then Q	is abbreviated as	$P\rightarrow Q$
P if and only if Q	is abbreviated as	*P iff Q,* or as $P\leftrightarrow Q$

The axiom may then be stated:[15]

♦Axiom 3. Principle of Abstraction

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

Having now set down the theory's axioms, the next step is to lay down its rules of inference, and then to start deducing theorems. We will postpone this task, however, until the next lecture, and jump ahead to the abbreviative definitions of the theory. In this way we will be able to discuss in the same lecture both the primitive and defined ideas of set theory. When we do move on to proving theorems, we will then use the defined terms to abbreviate longer formulas that occur in the derivations.

Abbreviative Definitions

We shall now introduce the abbreviations used in the theory. There are two important things to recall about abbreviative definitions. First, each such defined term is in principle eliminable from a formula by replacing it with the longer formula that it abbreviates. Second, the definition is suppose to conform to our "pre-analytic" usage. That is, according to ordinary language or earlier scientific usage the term doing the abbreviating should mean the same as the paraphrase that abbreviates.

Set Abstracts

We begin with some notation that allows for a more useful formulation of the Principle of Abstraction. The principle assures us that if there is a sentence $P(x)$, we can make up a set A that contains all and only the entities x such that $P(x)$ is true. It

[15] Strictly speaking this is what logicians call an *axiom schema*, because there are as many axioms of this form as there are different open sentences of the form $P(y)$.

is useful to name this set by the notation $\{x \mid P(x)\}$, which is read *the set of all x such that P*(*x*).

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

It is now possible to say more directly that any element *y* is in a set defined by a property if and only if *y* possess that property:

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

The proof of theorem 1 is not difficult, but since it requires steps of logic that have not yet been introduced, the theorem will have to be accepted for now on faith. A set name of the form $\{x \mid P(x)\}$ is called a *set abstract.*

Defined Relations on Sets

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

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

Defined Sets and Operations on Sets

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

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

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

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

The Problem of Universals

Let us conclude this introduction to sets by returning to the problem of sameness and difference that motivated the theories of Plato and Aristotle. Given set theory we can explain when two things are the same: two things are of the "same kind" when they are elements of the same set.

Truth-Conditions for Subject-Predicate Sentences and their Negations

- *Fa* is *true* if and only if the set that *F* names has as an element the individual that *a* stands for.
- ~*Fa* is *true* if and only if the set that *F* names does not have as an element the individual that *a* stands for.

Explanation of Sameness and Difference

- Two individuals are *the same* with respect to a set *U* if and only if they may be referred to by two different proper names, say *a* and *b*, *U* may be named by a predicate, say *F*, and the sentences *Fa* and *Fb* are both true.
- Two individuals are *different* with respect to a set *U´* if and only if they may be referred to by different proper names, say *a* and *b*, *U´* may be named by a

> predicate, say *G*, and either the sentences *Ga* and ~*Gb* are both true, or the sentences ~*Ga* and *Gb* are both true.

That is, sets function as universals. The events in the world that underlie our attributions of sameness and difference are grounded in facts that are part of the ontology of sets. Parmenides' puzzles of how the world could be coherently described so as to allow for sameness and difference are solved.

Lecture 4. Set Theory as an Axiom System

Proofs

This lecture is an introduction to two notion: a proof and an axiom system. It is fair to say that proofs are what logic studies, and it is axiom systems that display proofs in action in the clearest manner. But it is hard to study something you are not familiar with to some extent already. Accordingly a standard way to prepare for more serious study of logic is to have some hands-on experience at doing proofs. Here we shall use proofs to develop an elementary axiom system, naïve set theory, as based on the axioms of abstraction and extensionality discussed in the previous lecture.

How does one learn to do proofs? Historically, mathematicians have been employing proofs since ancient Greece. In the nineteenth century, however, it is fair to say that mathematicians became more careful in the way they did proofs. This care was partly motivated by the difficulty of the problems they were addressing and partly because they were learning how to formulate proofs more carefully. It was at this period that symbolic logic in its modern form was born. One of its main motivations was to represent even more precisely the proof methods that were part of mathematical practice.

Mathematicians did not then, and do not now, learn to construct proofs by taking logic courses, sad to say. Rather, they learn how to do so in math courses, by imitating what they see in math texts and by copying their professors. This process is highly imperfect. Neither textbooks nor math professors do a very good job of explaining what a proof is or how to go about constructing one. To an amazing extent, proof learning by mathematicians is actually non-verbal, a matter of aping professorial behavioral.

The existence of this proof behavior is important to logic as a "science." It is the proofs found in mathematics that logicians try to understand. These proofs are "the data to be explained," much as the movement of the planets or experimental results in the chemistry laboratory provide the data to be explained by the laws of physics or chemistry.

For these reasons it is a good idea for logic students to begin by trying to do proofs even before they understand very well what they are doing. We shall engage in this sort of baptism by fire in this lecture. We shall do so by proving theorems in naïve set theory. We start by explaining the rules of the game. A *proof* is a series of sentences, each of which is an axiom, a previously proven theorem, or a sentence that follows from earlier lines by a "rule of inference." A *rule of inference* tells you when you are permitted to write down a new line. Any line that you can prove as the last line of a proof is called a *theorem*. (We shall give more technically accurate definitions later.)

Inference rules all follow a certain pattern. They tell you the shape of the line you are permitted to write down, and they do so by describing the shape that must be met by earlier lines in the proof. By the "shape" of a sentence – often called its *form* – is meant its grammatical structure. If there are earlier lines of a certain form, the rule says, then you can write down a new line of another form. The new line is then "justified" by citing both these earlier lines and the rule used to derive it. Writing down the justification next to each line is called *annotating* the proof. In the previous lecture we have already laid down the axioms and abbreviative definitions will be using. It remains for us to explain the rules of inference. Here, to make proof construction easy, we will employ a generous set of rules, and rules that are fairly obvious and easy to use. They are also chosen because they are the ones that logicians most often use themselves when they have to construct proofs.

Inference Rules

The following rules explain when it is permitted to write a new line of a proof by citing as reasons previously proven lines. Read these rules carefully. First make sure you understand each symbol. Next make sure you understand what the rule says, i.e. what sort of sentence is said to follow logically from what other sort of sentence. Finally, make sure the rule "clicks." That is, try to convince yourself that the rule is right, that the inference it says follows logically really does. These rules are supposed to be obvious. If one is not, then you simply have not understood what it says. Keep trying to figure out what is says. If all else fails, ask somebody. There is really no point trying to use the rules until you understand them.

Below the rules are explained first in English and then by means of a display that is suppose to help you focus on the forms of the formulas involved. The display has the form:

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

The pyramid of three dots ∴ below the bar means *therefore.* The display means that a formula of the form *R* may be written down as a new line of the proof if there are earlier lines of the proof, one of form *P* and another of the form *Q.* Note that though *P* is written above *Q* in the display, it does not matter which was proven first so long as both occurs as lines earlier in the proof. In the examples below the various formulas will be complex sentences. Pay attention to their grammatical form. A rule only works when applied to formulas of the right shape. Rules marked with ♦ are used frequently and those with ♦♦ very frequently.

Logical Truths

♦♦**The Axiom of Logical Truth**. The first axiom of set theory, and indeed the first axiom of any science other than logic, asserts that any truth of logic is a truth of the science in question. The axiom captures the idea that logic is the most general of sciences and is presupposed in any other scientific endeavor. The axiom is true because, as we shall see in Part 2, the truths of logic are necessary in the sense that they are true "in any possible world". They are true, in other words, no matter what the facts of the particular science in question. Whatever the "special axioms" of the subject in question, these axioms must be compatible with the truths of logic. In the Middle Ages philosophers recognized the fundamental nature of logic in their recognition that even God is bound by its laws. As a practical matter it is difficult to use this axiom unless one already knows what the truths of logic are. In the proofs we shall be doing this prior knowledge is not a problem because we shall only be using three or four truths of logic which are themselves extremely obvious, like *P or not P, Everything is self-identical,* and *Every A is A.*

In Part 3 of the course we shall see how to replace this one rather question begging axiom by a short group of logical formulas that are obviously logical truths and that in turn serve as the axioms for all the other truths of logic. But for simplicity here we adopt the expedient of simply declaring that if we know something is a truth of logic, we can write it down as a line in a proof.[16] In practice, logicians and mathematicians also permit you to introduce as a line in a proof any other very simple truth of mathematics like simple facts of arithmetic (e.g. 2+2=4 or *Every positive integer is odd or even*), but we shall not be doing so here.

In this lecture we shall make use of just three truths of logic:

$P\vee\sim P$

$\forall x(x=x)$, and

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

[16] W.V.O. Quine adopts exactly the same expedient for the truths of propositional logic in his axiomatization of set theory. See the very first axiom (*100) in his *Mathematical Logic* (Cambridge: Harvard University Press, Revised Edition 1951).

Sentential Logic Rules

The next set of rules explains legitimate inferences that turn on the form of what are called the sentential connectives ~, &, ∨, →, and ↔.

♦♦***Modus (podendo) ponens***[17] (called by mathematicians **detachment**). Satisfying the antecedent of a conditional proves that the consequent is true.

$P \rightarrow Q$ $\underline{\quad P}$ $\therefore Q$	Note that because $P \leftrightarrow Q$ entails $P \rightarrow Q$ we shall also call the rule to the right *modus ponens*:	$P \leftrightarrow Q$ $\underline{\quad P}$ $\therefore Q$

♦***Modus (tollendo) tollens***[18]***.*** Refuting the consequent of a conditional proves that the antecedent is false.

$P \rightarrow Q$ $\underline{\sim Q\quad}$ $\therefore \sim P$	Note that because $P \leftrightarrow Q$ entails $P \rightarrow Q$ we shall also call the rule to the right *modus tollens*:	$P \leftrightarrow Q$ $\underline{\quad \sim Q}$ $\therefore \sim P$

Disjunctive Syllogism (*Modus tollendo ponens*[19]). If one horn of a dilemma is refuted (or one case of those cases that are possible), then one of the remaining alternatives must be true.

$P \vee Q$ $\underline{\sim P}$ $\therefore Q$	$P \vee Q$ $\underline{\sim Q}$ $\therefore P$

♦**Hypothetical Syllogism.** Conditionals are transitive.

$P \rightarrow Q$
$\underline{Q \rightarrow R}$
$\therefore P \rightarrow R$

♦♦**Conjunction.** A conjunction may be broken down and either half written as a new line, or two previous lines may be put together on a single line if joined by &.

Conjunction Elimination	$\underline{P \& Q}$ $\therefore P$	$\underline{P \& Q}$ $\therefore Q$	Conjunction Introduction	P $\underline{Q}$ $\therefore P \& Q$

We shall call both rules simply **Conjunction.**

[17] [Given a conditional,] the way of positing [the consequent] by positing [the antecedent].
[18] [Given a conditional,] the way of taking away [the antecdent] by taking away [the cosequent].
[19] [Given a disjunction,] the way of positing [one disjunct] by taking away [the other disjunct].

Addition or **Disjunction Introduction.** To prove a disjunction $P \vee Q$ (to "introduce" a line containing $P \vee Q$ into a proof) it suffices to prove either the disjunct P or the disjunct Q individually.

$$\frac{P}{\therefore P \vee Q} \qquad\qquad \frac{Q}{\therefore P \vee Q}$$

Quantifier Rules

The next set of rules explain the legitimate inferences that turn on the so-called quantifiers *all* and *some,* i.e. that depend of the symbols $\forall$ and $\exists$.

♦♦**Universal Generalization.** If you know some fact $P(x)$ about an individual x, and you know that this x "typical of everything" or is a so-called "arbitrary individual", then you can generalize from this one case to the entire universe and write down $\forall x P(x)$. To do so, you cannot have assumed anything about x that you do not know is true of everything in the universe. Normally the only way a free variable for arbitrary individuals enters a proof is by the rule universal instantiation below.

$$\frac{P(x)}{\therefore \forall x\, P(x)} \qquad\qquad \frac{P(y)}{\therefore\ \forall x P(x)}$$ here x and y must be arbitrary

♦♦**Universal Instantiation.** If you know a universal proposition $\forall x P(x)$, you can always deduce an "instance" $P(x)$ or $P(y)$, for any free variable you wish. When you do this, the instantiated free variable, in this case x or y, counts as an "arbitrary individual" and can be used later in an application of universal generalization above. Here are two forms, both valid:

$$\frac{\forall x P(x)}{\therefore\ P(x)} \qquad\qquad \frac{\forall x P(x)}{\therefore\ P(y)}$$ (here x and y count as arbitrary)

♦♦**Proof of an Existential Proposition by Construction.** To prove a proposition of existential form $\exists x P(x)$, you must find an example of some c such that $P(c)$ is true. Normally "finding" here consists of defining the individual if, for example, it is a set, or otherwise appealing to facts of mathematics that assure you the right entity exists. (Another way to prove $\exists x P(x)$ is by the rule reductio, i.e. assume $\sim\exists x P(x)$ and then deduce a contradiction. See the reductio rule below for the proper form.)

$$\frac{P(c)}{\therefore\ \exists x P(x)}$$ (here you may know $P(c)$ because of the definition of c or from math)

♦♦**Existential Instantiation.** If you know an existential proposition $\exists xP(x)$, you can "give a "name" to this x, which may be any free variable that has not occurred in any previous line of the proof, for example $P(x)$ or $P(y)$ if x or y does not occur earlier as a free variable.

$\underline{\exists xP(x)}$	$\underline{\exists xP(x)}$	
$\therefore P(x)$	$\therefore P(y)$	(here x and y must be new to the proof and are ***not*** to be considered arbitrary.)

Substitution Rule

♦♦**Substitution**. If $a=b$ is a logical truth or a previous line and a and b are proper names or variables, then one may be substituted for the other in one line of the proof to deduce a new line. Likewise, if a formula of the form $P\leftrightarrow Q$ is a logical truth or a previous line in a proof, and if P and Q contain the same variables (in this case P and Q are said to be *logically equivalent*), then if the formula P occurs in an earlier line, a new line may be introduced that consists of replacing one or more occurrences of P in the earlier line with a expression of the form Q. Likewise, any occurrence of a formula of the form Q in an earlier line may be replaced in a new line by one of the form P.

To state the rule more precisely, we need a bit of notation. We say that E_1 is in R by writing $R(E_1)$, and then that E_2 replaces E_1 one or more times in R by writing $R(E_1)$. That is, let $R(E_1)$ refer to a formula R that contains one or more occurrences of E_1, and let $R(E_2)$ be like R except that it contains one of more occurrences of E_2 were R contains E_1. Now we can state the rule, first for the substitutivity of identities and then for the substitution of sentences that are logically equivalent:

$\underline{R(t)}$	$\underline{R(s)}$	if $t=s$ is a logical truth or a previous line
$\therefore R(s)$	$\therefore R(t)$	with the same free variables[20].
$\underline{R(P)}$	$\underline{R(Q)}$	if $P\leftrightarrow Q$ is a logical truth or a previous line
$\therefore R(Q)$	$\therefore R(P)$	and P and Q have the same free variables

For this rule we list several logical truths that are particularly useful:

[20] Here s and t are proper names (which are called *constants* in formal logic), or free variables, or abstracts. If they are abstracts $\{x \mid \ldots x\ldots\}$. , they contain a sentence $\ldots x\ldots$ that may itself contain variables other than x and some of these might be "free." For example, in . $\{x \mid x \textit{ loves } y\}$ the variable y is free because it has no antecedent that fixes its referent. In $\forall x(\ldots x\ldots)$, $\exists x(\ldots x\ldots)$,, and $\{x \mid \ldots x\ldots\}$, the occurrence of x in $\ldots x\ldots$ is said to be *bound* to the earlier x in $\forall x$, $\exists x$, or the abstract notation $\{x \mid \ \}$ which fixes the referent of the second x. The technical definition is that variable x is *free* if and only if it does not occur as part of some expression $\forall x(\ldots x\ldots)$, $\exists x(\ldots x\ldots)$,, and $\{x \mid \ldots x\ldots\}$. Free variables, thus, function as pronouns that lack a reference because they lack an antecedent. Binding a variable by a quantifier or an abstract of the same variable provides that antecedent. Since free variables are functioning as pronouns, in substitution the free variables must remain the same to make sure the same objects are being named in both the original expression and the one that replaces it.

Association: $((P\&Q)\&R) \leftrightarrow (P\&(Q\&R))$
$((P \vee Q) \vee R) \leftrightarrow (P \vee (Q \vee R))$

Commutation: $(P\&Q) \leftrightarrow (Q\&P)$
$(P \vee Q) \leftrightarrow (Q \vee P)$
$(P \leftrightarrow Q) \leftrightarrow (Q \leftrightarrow P)$

DeMorgan's Laws: $\sim(P\&Q) \leftrightarrow (\sim P \vee \sim Q)$

DeMorgan's Laws: $\sim(P\&Q) \leftrightarrow (\sim P \vee \sim Q)$
$\sim(P \vee Q) \leftrightarrow (\sim P\&\sim Q)$

♦**Double Negation :** $\sim\sim P \leftrightarrow P$

Implication : $(P \to Q) \leftrightarrow (\sim P \vee Q)$
$\sim(P \to Q) \leftrightarrow (P \,\&\, \sim Q)$

Contraposition : $(P \to Q) \leftrightarrow (\sim Q \to \sim P)$

Tautology : $(P\&P) \leftrightarrow (P \vee P) \leftrightarrow P$

♦**The Biconditional :** $(P \leftrightarrow Q) \leftrightarrow ((P \to Q)\& (Q \to P))$
$(P \leftrightarrow Q) \leftrightarrow ((P\&Q) \vee (\sim P\&\sim Q))$

Quantifier Negations : $\sim\forall x P(x) \leftrightarrow \exists x \sim P(x)$
$\sim\exists x P(x) \leftrightarrow \forall x \sim P(x)$
$\sim\forall x(P(x) \to Q(x)) \leftrightarrow \exists x(P(x)\&\sim Q(x))$
$\sim\exists x(P(x)\&Q(x)) \leftrightarrow \forall x(P(x) \to \sim Q(x))$

Subproof Rules

The next rules says that if a proposition of a certain form can be proven from other propositions of a certain form, then you can write down a new line of a certain form. They are explained in displays of the form:

If $\dfrac{P}{\therefore Q}$ then R

These means that if somewhere earlier in the proof you have actually written down another mini-proof, called a *subproof*, that starts with *P* as an assumption and finishes with *Q,* then you can write down a new line of the form *R.*

There are two rules of this form, one says that if you can prove a contradiction from a formula, its opposite is true. The second says that if on the assumption of *P* you can prove *Q*, then the conditional $P \rightarrow Q$ must be true. But showing that a formula leads to a contradiction, or that *P* leads to *Q,* means that you need to produce a subproof showing the steps. You add this subproof to your overall proof, and then apply the subproof rule to deduce the new line. In what follows those lines that constitute a subproof will be offset to the right several spaces to distinguish the subproof from the proof proper.

Note that when proving a line within a subproof, it is permitted to cite lines proven earlier in the proof, including those that occur prior to the subproof itself. These can be cited because they have already been proven. On the other hand, once the subproof is concluded, its content must be "sealed off." You cannot later cite lines that occur within an earlier subproof because its lines were only conditionally true, depending on the special *ad hoc* assumption that started the subproof. The examples below will show how these restrictions work in practice.

♦***Reductio ad absurdum***
If you can prove from an assumption P a contradiction $Q\&{\sim}Q$, or indeed any other proposition that you know to be false, you may conclude that P is false, i.e. you can conclude ${\sim}P$. The rule has two forms, depending on whether the proposition you are showing is absurd is itself negated or not:

If $\dfrac{P}{\therefore Q\&{\sim}Q}$ then ${\sim}P$

If $\dfrac{{\sim}P}{\therefore Q\&{\sim}Q}$ then P

Ex Falso Quodlibet (from a falsity anything follows). If you show a contradiction, you can write down anything you want as a following line.

$$\frac{P\&{\sim}P}{\therefore Q}$$

♦♦**Conditional Proof** If on the assumption of P you can prove Q, then you may conclude $P\rightarrow Q$.

If $\dfrac{P}{\therefore Q}$ then $P\rightarrow Q$

Conditional Proof for Biconditionals. If on the assumption of P you can prove Q, and on the assumption of Q you can prove P, then you may conclude $P\leftrightarrow Q$.

If $\dfrac{P}{\therefore Q}$ and $\dfrac{Q}{\therefore P}$, then $P\leftrightarrow Q$

Proof by Cases (**Disjunction Elimination**). If R is provable from each of the horns of a dilemma $P\vee Q$ (i.e. if R is provable from case 1, namely P, and R is also provable from case 2, namely Q), then R is provable from the disjunction $P\vee Q$, which lays out the possible cases.

If $\dfrac{P}{\therefore R}$ and $\dfrac{Q}{\therefore R}$, then $\dfrac{P\vee Q}{\therefore R}$

Statement of the Axiom System

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

Summary of the System

Axioms

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

Rules of Inference

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

Abbreviative Definitions

Abbreviation	Phrase Abbreviated	Abbreviation's Name
$\{x \mid P(x)\}$	*the one and only A such that* $\forall x\ (\ x\in A \leftrightarrow P(x))$	***set abstract***[21]
$x\neq y$	$\sim(x=y)$	***non-identity***
$x\notin A$	$\sim(x\in A)$	***non-membership***
$A\subseteq B$	$\forall x(x\in A\rightarrow x\in B)$	***subset***
$A\subset B$	$A\subseteq B\&\sim A=B$	***proper subset***

[21] See the later note on Theorem 1 for a more precise statement of this definition.

$\varnothing$ or Λ	$\{x \mid x \neq x\}$	***empty set***
V	$\{x \mid x = x\}$	***universal set***
$A \cap B$	$\{x \mid x \in A \& x \in B\}$	***intersection***
$A \cup B$	$\{x \mid x \in A \vee x \in B\}$	***union***
$A - B$	$\{x \mid x \in A \& x \notin B\}$	***relative complement***
$-A$	$V - A$	***complement***
$\mathbf{P}(A)$	$\{B \mid B \subseteq A\}$	***power set***
$\{x_1, \ldots, x_n\}$	$\{y \mid y = x_1 \vee \ldots \vee y = x_n\}$	

Definition of Proof

We can now also formally define a proof. Since a proof can contain subproofs, to avoid a circular definition, we must first define the notion of a simple proof that does not make use of a subproof.

> A *simple proof* is any finite sequence of lines each of which is an instance of an axiom or follows from earlier lines of the sequence by one of the non-subproof rules.
>
> A *simple theorem* is any formula that occurs as the final line of some simple proof.

Given these two definitions, it follows that – and we shall demonstrate this in a later lecture –any simple theorem is (1) either a logical truth or an instance of an axiom of naïve set theory or (2) follows from a another simple theorem by one of the non-subproof rules.

> A *proof* is a finite sequence of lines such that each is an axiom instance or a simple proof or follows from earlier lines or sequences of lines by one of the proof rules.
>
> A *theorem* is any formula that occurs as the final line of some proof.

As before, given these two definitions, it follows that any theorem is (1) either a logical truth or an instance of an axiom of naïve set theory or (2) follows from a another simple theorem by one of the inference rules, including the subproof rules.

To make proofs shorter, when we want to prove a theorem that follows in part from a theorem we have already proven, rather than repeating that theorem's proof in the proof under construction, we shall adopt the convention of merely writing down the theorem as a line or of citing it if it states an equivalence that would justify a useful substitution. We now list for reference the theorems that we shall prove in the next section.

Theorems

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

Lecture 5. Proofs of the Theorems

Below we prove theorems of naïve set theory and do so by producing examples of proofs. Some proofs have completed annotations. You should work through these annotated examples to see why each line follows. Look at the lines it depends on and see how the rule applied yields the new line.

To follow proofs, it helps tremendously to print out the axioms, and definitions, and then have them before your eyes. Print the full versions of the rules that display the shapes of the required forms. This way you can check that the lines have the relevant shapes.

Some proofs are only partly annotated. The relevant earlier lines are cited but not the rule used. You are supposed to find the rule yourself. A few theorems are left for you to prove entirely on your own, with annotations. A couple of these are not easy.

Note that the rules reductio and conditional proof require "subproofs," i.e. mini proofs within longer proofs. Below the line of such subproofs are offset several spaces to the right so that it is easy to see that they are subproofs. In a few cases there are subproofs within subproofs, and these are set off even further to the right, etc.

Theorem 1 (Principle of Abstraction, useful form). $\forall y(\ y\in\{x \mid P(x)\} \leftrightarrow P(y)\ )$.

The proof is not difficult, but even with our rich set of rules it is relatively complicated, too much so to make it worthwhile reproducing here. The reader will have to take it on faith that is follows using the inference rules alone from axioms 1 and 2. [22]

---

[22] For the technically minded, the complete proof is given below. Students new to logic should be able to follow the proof after working through the later parts of Part 1 in these lectures. Before stating the proof, however, the definition of {x| $P(x)$} must be stated more accurately. To define {x| $P(x)$} by eliminative definition, it must be possible to eliminate the notation wherever it occurs. It occurs only in four possible positions, either to the left or right of the predicates $\in$ and =. That is, we must define 1-4 below:

1. $y\in\{x|\ P(x)\}$
2. $\{x|\ P(x)\}\in y$
3. $\{x|\ P(x)\}=y$
4. $y=\{x|\ P(x)\}$

Consider the first. In this case $y\in\{x|\ P(x)\}$ says that there is a set, call it A, that meet three conditions:

(i) something z is in A iff $P(z)$}. In symbols:
$\forall z(z\in A \leftrightarrow P(z))$

(ii) A is the only such set or, in other words, if any B is such, it must be A. In symbols:
$\forall B(\forall z(z\in B \leftrightarrow P(z))\rightarrow B=A)$

(iii) y has the defining property of A. In symbols:
$P(y)$

Hence, the formal definition of 1 draws together all three conditions:

$y\in\{x|\ P(x)\}$ $\quad \exists A(\forall z(z\in A \leftrightarrow P(z))\ \&\ \forall B(\forall z(z\in B \leftrightarrow P(z))\rightarrow B=A)\ \&\ P(y))$

The case is similar for 2. Here $\{x|\ P(x)\}\in y$ says there is a set A that (i) $\forall z(z\in A \leftrightarrow P(z))$, (ii) A is the only such set, and (iii) $A\in y$. In symbols the definition reads:

$\{x|\ P(x)\}\in y$ $\quad \exists A(\forall z(z\in A \leftrightarrow P(z))\ \&\ \forall B(\forall z(z\in B \leftrightarrow P(z))\rightarrow B=A)\ \&\ A\in y)$

The definitions of 3 and 4 are similar:

$y=\{x|\ P(x)\}$ $\quad \exists A(\forall z(z\in A \leftrightarrow P(z))\ \&\ \forall B(\forall z(z\in B \leftrightarrow P(z))\rightarrow B=A)\ \&\ y=A)$

$\{x|\ P(x)\}=y$ $\quad \exists A(\forall z(z\in A \leftrightarrow P(z))\ \&\ \forall B(\forall z(z\in B \leftrightarrow P(z))\rightarrow B=A)\ \&\ A=y)$

Given these definitions it is possible to state and prove Theorem 1.

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

Proof.

Start of Subproof 1.

1. $\exists A(\forall z(z\in A \leftrightarrow P(z))\ \&\ \forall B(\forall z(z\in B \leftrightarrow P(z))\rightarrow B=A)\ \&\ P(y))$ Assumption for CP, for arbitrary y
2. $\forall z(z\in A \leftrightarrow P(z))\ \&\ \forall B(\forall z(z\in B \leftrightarrow P(z))\rightarrow B=A)\ \&\ P(y))$ 1, Existential Instantiation
3. $P(y)$ 2, Conjunction

End of Subproof 1.

Start of Subproof 2.

4. $P(y)$ Assumption for CP, for arbitrary y
5. $\exists A\forall z(z\in A \leftrightarrow P(z))$ Axiom 3, Abstraction
6. $\forall z(z\in A \leftrightarrow P(z))$ 5, Existential Instantiation

Start of Subproof 2a.

7. $\forall z(z\in B \leftrightarrow P(z))$ Assumption for CP, arbitrary B
8. $y\in A \leftrightarrow P(y)$ 6, Universal Instantiation
9. $y\in B \leftrightarrow P(y)$ 7, Universal Instantiation
10. $y\in B\leftrightarrow y\in A$ 8 and 9, Hypothetical Syllogism
11. $\forall y(y\in B\leftrightarrow y\in A)$ 10, Universal Generalization
12. $B=A$ 11, Axiom 1, Extensionality

End of Subproof 2a.

13. $\forall z(z\in B \leftrightarrow P(z))\rightarrow B=A$ 7-12, Conditional Proof
14. $\forall B(\forall z(z\in B \leftrightarrow P(z))\rightarrow B=A)$ 13, Universal Generalization

Theorem 1, however, is the useful version of the Principle of Abstraction. It may be applied directly to each of the defined set names and operations to yield for each an elementary theorem that makes it relatively easy to prove more complex facts about these sets.

In the early proofs, the justification of each step is followed by a commentary in a box that explains more fully what rule is being used and why. These explanations are designed to introduce you to the late proofs in which these sorts of explanations are "understood" without being stated. Read them closely to so that you do see what rule is being used. Keep a print out of the rules in front of you so that you can confirm that the right shapes are being applied.

Theorem 2. $\forall x\,(x\in\varnothing \leftrightarrow x\neq x)$

Proof.

1. $\forall y(y\in\{x|\ x\neq x\} \leftrightarrow y\neq y)$ — Theorem 1 (Principle of Abstraction), substituting *y* for *x.*

> Theorem 1 is $\forall y(\ y\in\{x \mid P(x)\} \leftrightarrow P(y))$.
> Let $P(x)$ be $x\neq x$.
> Let $P(y)$ be $y\neq y$.
> Hence, an instance of Theorem 1 is $\forall y(y\in\{x|\ x\neq x\} \leftrightarrow y\neq y)$

2. $\forall y(y\in\varnothing \leftrightarrow y\neq y)$ — definition of $\varnothing$

> The def of $\varnothing$ is $\{x|\ x\neq x\}$.
> Substitute $\{x|\ x\neq x\}$ for $\varnothing$ in line 1.

3. $\forall x\,(x\in\varnothing \leftrightarrow x\neq x)$ — 2, substituting *x* for *y.*

---

15. $\forall z(z\in A \leftrightarrow P(z))\ \&\ \forall B(\forall z(z\in B \leftrightarrow P(z))\rightarrow B=A)\ \&\ P(y)$ — 3, 6, and 13, Conjunction

End of Subproof 2.

16. $(\forall z(z\in A \leftrightarrow P(z))\ \&\ \forall B(\forall z(z\in B \leftrightarrow P(z))\rightarrow B=A)\ \&\ P(y)) \leftrightarrow P(y)$ — Subproof 1 and 2, CP
17. $\forall y(\forall z(z\in A \leftrightarrow P(z))\ \&\ \forall B(\forall z(z\in B \leftrightarrow P(z))\rightarrow B=A)\ \&\ P(y)) \leftrightarrow P(y))$ — 16, Universal Generalization
18. $\forall y(y\in\{x \mid P(x)\} \leftrightarrow P(y))$ — 17, Definition of $y\in\{x \mid P(x)\}$

Line 3 is like line 2 except that y is replaced by x. You can always uniformly replace an old variable by one not already in a formula.

The proofs of theorems 3-9 are virtually the same, using the definitions of the appropriate set names.

Theorem 3. $\forall x\,(x\in \mathrm{V} \leftrightarrow x=x)$
Proof. **Exercise.** Annotate the proof.

1. $\forall y(y\in\{x|\ x=x\} \leftrightarrow y=y)$
2. $\forall y(y\in \mathrm{V} \leftrightarrow y=y)$
3. $\forall x\,(x\in \mathrm{V} \leftrightarrow x=x)$

Theorem 4. $\forall x\,(x\in A\cap B \leftrightarrow (x\in A \& x\in B)\,)$
Proof.

1. $\forall y(y\in\{x|\ x\in A \& x\in B\ \} \leftrightarrow (y\in A \& y\in B))$ Principle of Abstraction, with y for x.

Theorem 1 is $\forall y(\ y\in\{x\ |\ P(x)\} \leftrightarrow P(y))$.
Let $P(x)$ be $x\in A \& x\in B$.
Let $P(y)$ be $y\in A \& y\in B$.
Hence, an instance of Theorem 1 is
$\forall y(y\in\{x|\ x\in A \& x\in B\ \} \leftrightarrow (y\in A \& y\in B))$

2. $\forall y(y\in\ A\cap B \leftrightarrow (y\in A \& y\in B))$ definition of $A\cap B$

The def of $A\cap B$ is $\{x|\ x\in A \& x\in B\}$.
Substitute $\{x|\ x\in A \& x\in B\}$. for $A\cap B$ in line 1.

3. $\forall x\,(x\in\ A\cap B \leftrightarrow (x\in A \& x\in B))$ 2, substituting y for x.

Line 3 is like line 2 except that y is replaced by x. You can always uniformly replace an old variable by one not already in a formula.

Theorem 5. $\forall x\,(x\in A\cup B \leftrightarrow (x\in A \vee x\in B)\,)$
Proof. **Exercise.** Annotate the proof.

1. $\forall y(y\in\{x|\ x\in A \vee x\in B\ \} \leftrightarrow (y\in A \vee y\in B))$
2. $\forall y(y\in\ A\cup B \leftrightarrow (y\in A \vee y\in B))$
3. $\forall x\,(x\in\ A\cup B \leftrightarrow (x\in A \vee x\in B))$

Theorem 6. $\forall x\,(x\in A-B \leftrightarrow (x\in A \& x\notin B)\,)$ This proof is like that of Theorem 4
Proof.

1. $\forall y(y\in\{x|\ x\in A \& x\notin B\ \} \leftrightarrow (y\in A \& y\notin B))$ Principle of Abstraction, with y for x.
2. $\forall y(y\in\ A-B \leftrightarrow (y\in A \& y\notin B))$ definition of $A-B$
3. $\forall x\,(x\in\ A-B \leftrightarrow (x\in A \& x\notin B))$ 2, substituting y for x.

Theorem 7. $\forall x\ (x\in -A \leftrightarrow x\notin A)$

Proof. **Exercise.** Annotate line 2-4 of the proof.

1. $\forall y((y\in \mathrm{V}\&{\sim}y\in A) \leftrightarrow {\sim}y\in A)$ Truth of Logic (Axiom 1, Theorem 3)
2. $\forall y((y\in \mathrm{V}\&y\notin A) \leftrightarrow y\notin A)$
3. $\forall y(y\in \mathrm{V}-A \leftrightarrow y\notin A)$
4. $\forall x\ (x\in -A \leftrightarrow x\notin A)$

Theorem 8. $\forall B\ (B\in \mathbf{P}(A) \leftrightarrow B\subseteq A)$

This proof is like that of Theorem 4

Proof.

1. $\forall B(B\in\{x|\ x\subseteq A\} \leftrightarrow B\subseteq A)$ Principle of Abstraction, with B for x.
2. $\forall B(B\in \mathbf{P}(A) \leftrightarrow B\subseteq A)$ definition of $A-B$

Theorem 9. $\forall y\ (y\in\{x_1, \ldots, x_n\} \leftrightarrow (y = x_1 \vee \ldots \vee y = x_n))$

Proof. **Exercise.** Construct the proof.

Theorem 10. $--A=A$

Proof.

1. $\forall x(x\in A\leftrightarrow x\in A)$ logical truth

One of the logical truths listed earlier is:
$\forall x(P(x) \leftrightarrow P(x))$
Let $P(x)$ be $x\in A$.
Hence a particular instance of this logical truth is:
$\forall x(x\in A\leftrightarrow x\in A)$
You are always permitted to write down a logical truth as a line of a proof.

2. $\forall x({\sim}{\sim}(x\in A)\leftrightarrow x\in A)$ 1, double negation

One of the logical equivalences listed under the Substitutions Rule is Double Negation:
${\sim}{\sim}P \leftrightarrow P$
Let P be $x\in A$.
Hence, an instance of Double Negation is:
${\sim}{\sim}(x\in A)\leftrightarrow x\in A$
The substitution rule says you can always replace a formula by a logical equivalent. So, in this step we replace the first occurrence of $x\in A$ in line 1 with its equivalent ${\sim}{\sim}(x\in A)$ and thus obtain line 2.

3. $\forall x({\sim}(x\notin A)\leftrightarrow x\in A)$ 2, definition of $\notin$

The definition of $x\notin A$ is ${\sim}(x\in A)$. That is, $x\notin A$ is an abbreviation for ${\sim}(x\in A)$. Note that the second of these occurs in line 2. You can always replace an expression by its abbreviation or replace an abbreviation by the expression it abbreviates. Here we replace ${\sim}(x\in A)$ in line 2 by its abbreviation $x\notin A$ to obtain line 3.

4. $\forall x({\sim}(x\in -A)\leftrightarrow x\in A)$ 3, theorem 7

Theorem 7 tells you how to replace the occurrence of the complement sign $-$ by an equivalent in terms of negation, and vice versa. The theorem is:
$\forall x\,(x\in -A \leftrightarrow x\notin A)$.
Hence we replace $x\notin A$ in line 3 by its equivalent $x\in -A$ to obtain line 4.

5. $\forall x(x\notin -A\leftrightarrow x\in A)$ — 4, definition of $\notin$

Like line 3.

6. $\forall x(x\in --A\leftrightarrow x\in A)$ — 5, theorem 7

Like line 4.

7. $--A=A$ — 6, Principle of Extensionality

The Principle of Extensionality tells you when two sets are identical:
$A=B \leftrightarrow \forall x\,(x\in A \leftrightarrow x\in B)$
Here A and B can be any sets.
Let A be $--A$ and let B be A. Then an instance of the principle is:
$--A=A \leftrightarrow \forall x\,(x\in --A \leftrightarrow x\in A)$
But line 6 is one half of this equivalence. Hence, by the Substitution of Equivalents rule, we replace line 6 by its equivalent to obtain line 7.

Theorem 11. $A\subseteq A$
Proof. **Exercise.** Construct the proof.

Theorem 12. $\forall x((x\in A\ \&\ A\subseteq B)\rightarrow x\in B)$

General Proof Strategy. Notice that this theorem is a universally quantified conditional. Such theorems are often proven by a regular strategy: assume the antecedent (the "if" part) of the conditional is true for an arbitrary object, and then deduce that the conditional's consequent (the "then" part). The rule Conditional Proof sets out how to do this. Start a subproof (indent right) and assume the antecedent. Deduce lines until you reach the consequent. At that point end the subproof (move left) and write down a new line stating the conditional, justified by the subproof lines and the words "Conditional Proof." If you have treated the object as genuinely arbitrary -- you have not assumed anything about it that is not true of everything – then you may follow this line with one that encloses the conditional in a universal quantifier, justified by citing the Universal Generalization rule.

Proof.
Start of Subproof 1. (a conditional proof)

Here you declare that you are starting a subproof, this case in order to prove a conditional. The next line will be the antecedent of the conditional to be proved, and the last line of the subproof will be its consequent.

1. $x \in A$ & $A \subseteq B$ — assumption for conditional proof, with x arbitrary

> This is the antecedent of the conditional to be proved. You say "x is arbitrary" because, if from this point on you do not assume anything about x except facts that are true of everything, then you will later be able to apply the rule Universal Generalization to this x, converting a fact about it into a generalization that talks about "everything".

2. $x \in A$ — 1, conjunction

> The Conjunction inference rule is:
>
> $$\frac{P\&Q}{\therefore P} \qquad \frac{P\&Q}{\therefore Q} \qquad \frac{P \quad Q}{\therefore P\&Q}$$
>
> Let P be $x \in A$ and let Q *be* $A \subseteq B$
> Then an instance of the rule is
>
> $$\frac{x \in A \,\&\, A \subseteq B}{x \in A} \qquad \frac{x \in A \,\&\, A \subseteq B}{A \subseteq B}$$
>
> We apply the first of these rule to line 1 to get line 2.

3. $A \subseteq B$ — 1, conjunction

> As in line 2 above an instance of Conjunction rule is
>
> $$\frac{x \in A \,\&\, A \subseteq B}{x \in A} \qquad \frac{x \in A \,\&\, A \subseteq B}{A \subseteq B}$$
>
> We apply the second of these to line 1 to get line 3.

4. $\forall x(x \in A \rightarrow x \in B)$ — 3, definition of $\subseteq$

> The definition of $A \subseteq B$ is $\forall x(x \in A \rightarrow x \in B)$. That is, $A \subseteq B$ is an abbreviation for $\forall x(x \in A \rightarrow x \in B)$. Note that the first of these occurs in line 3. You can always replace an expression by its abbreviation or replace an abbreviation by the expression it abbreviates. Here we replace $A \subseteq B$ in line 3 by the expression it abbreviates $\forall x(x \in A \rightarrow x \in B)$ to obtain line 4.

5. $x \in A \rightarrow x \in B$ — 4, universal instantiation, x arbitrary

> The Universal Instantiation inference rule is:
>
> $$\frac{\forall x P(x)}{\therefore P(x)} \qquad \frac{\forall x P(x)}{\therefore P(y)}$$ for arbitrary x and y.
>
> Let $\forall x P(x)$ be $\forall x(x \in A \rightarrow x \in B)$
> Then an instance of the rule is
>
> $$\frac{\forall x(x \in A \rightarrow x \in B)}{\therefore x \in A \rightarrow x \in B}$$
>
> We apply this rule to line 4 to get line 5.

6. $x \in B$ — 2 and 5, *modus ponens*

> The *Modus Ponens* inferece rule is:

$P \to Q$
$\underline{P}$
$\therefore Q$
Let P be $x \in A$, Q be $x \in B$, and $P \to Q$ be $x \in A \to x \in B$
Then an instance of the rule is
$x \in A \to x \in A$
$\underline{x \in A}$
$\therefore x \in B$
We apply the first of the rule to lines 2 and 5, to get line 6.

End of subproof

Here you declare that you are ending a subproof, this case in order to prove a conditional formed by taking as its antecedent the assumption of the subproof (line 1) and as its consequent the last line of the subproof (line 6). The next line will be the conditional itself.

7. $(x \in A \ \&\ A \subseteq B) \to x \in B$ 1-6, conditional proof

Here you move to the left, and declare that you have proven the conditional determined by the first and last line of the subproof that immediately precedes this line. You cite *all* the line of the subproof as part of the justification. From this point on you cannot draw on any line from the subproof to prove further lines (the subproof is "sealed off") because its lines depend on the extra assumption that starts the subproof (line 1) which you are no longer entitled to assume. This assumption is said to have been "discharged" by moving it to the antecedent of this conditional.

8. $\forall x((x \in A \ \&\ A \subseteq B) \to x \in B)$ 7, universal generalization, x arbitrary

The Universal Generalization is:
$\underline{\forall x P(x)}$ $\underline{P(y)}$
$\therefore \forall x P(x)$ $\therefore \forall x P(x)$ for arbitrary x and y.
Let $P(x)$ be $(x \in A \ \&\ A \subseteq B) \to x \in B$
Then an instance of the rule is
$\underline{(x \in A \ \&\ A \subset B) \to x \in B}$
$\therefore \forall x((x \in A \ \&\ A \subseteq B) \to x \in B)$
We know that x ix in fact arbitrary because we have not assumed anything about it that is not true of everything. Hence, we apply this rule to line 7 to get line 8.

Theorem 13. $A \cap A = A = A \cup A$
This really two propositions, $A \cap A = A$ and $A = A \cup A$, which we shall call theorems 13a and 13b respectively. Each requires its own proof.
Theorem 13a. $A \cap A = A$
Proof.
1. $\forall x(x \in A \leftrightarrow x \in A)$ logical truth

2. $\forall x((x\in A \& x\in A) \leftrightarrow x\in A)$	1, tautology
3. $\forall x(x\in A\cap A \leftrightarrow x\in A)$	2, theorem 4
4. $A\cap A=A$	3, theorem 1

Theorem 13b. $A=A\cup A$.

Proof. **Exercise.** Annotate the proof.

1. $\forall x(x\in A \leftrightarrow x\in A)$
2. $\forall x(x\in A \leftrightarrow (x\in A\vee x\in A))$
3. $\forall x(x\in A \leftrightarrow x\in A\cup A)$
4. $A=A\cup A$

Theorem 14. $A=B \leftrightarrow (A\subseteq B \ \& \ B\subseteq A)$

Proof.

Start of Subproof 1. (a conditional proof)	
1. $A=B$	assumption for conditional proof
2. $\forall x(x\in A \leftrightarrow x\in B)$	1, Principle of Extensionality
3. $x\in A \leftrightarrow x\in B$	2, an instance
4. $(x\in A \rightarrow x\in B) \ \& \ (x\in B \rightarrow x\in A)$	3, biconditional
5. $x\in A \rightarrow x\in B$	4 , conjunction
6. $\forall x(x\in A \rightarrow x\in B)$	5, generalization since x is arbitrary
7. $A\subseteq B$	6, definition of $\subseteq$
8. $x\in B \rightarrow x\in A$	4, conjunction
9. $\forall x(x\in B \rightarrow x\in A)$	7, generalization since x is arbitrary
10. $B\subseteq A$	9, definition of $\subseteq$
11. $A\subseteq B \ \& \ B\subseteq A$	7 and 10
End of Subproof 1	
12. $A=B \rightarrow (A\subseteq B \ \& \ B\subseteq A)$	1-11, conditional proof
Start of Subproof 2. (a conditional proof)	
13. $A\subseteq B \ \& \ B\subseteq A$	assumption for a conditional proof
14. $A\subseteq B$	13, conjunction
15. $\forall x(x\in A \rightarrow x\in B)$	14, definition of $\subseteq$
16. $x\in A \rightarrow x\in B$	instance of 15
17. $B\subseteq A$	13, conjunction
18. $\forall x(x\in B \rightarrow x\in A)$	17, definition of $\subseteq$
19. $x\in B \rightarrow x\in A$	instance of 18
20. $(x\in A \rightarrow x\in B) \ \& \ (x\in B \rightarrow x\in A)$	16 and 19
21. $x\in A \leftrightarrow x\in B$	20, biconditional
22. $\forall x \ (x\in A \leftrightarrow x\in B)$	21, universal generalization, x arbitrary
23. $A=B$	22, Principle of Extensionality
End of Subproof 2.	
24. $(A\subseteq B \ \& \ B\subseteq A) \rightarrow A=B$	13-23, conditional proof
25. $A=B \leftrightarrow (A\subseteq B \ \& \ B\subseteq A)$	12 and 24, biconditional

Theorem 15. $A\cap B\subseteq A\subseteq A\cup B$

This is really two theorems.

Theorem 15a. $A \cap B \subseteq A$
Proof.
Start of Subproof 1 (a conditional proof)

1. $x \in A \cap B$	assumption for conditional proof
2. $x \in A$ & $x \in B$	1, theorem 4
3. $x \in A$	2, conjunction

End of Subproof 1

4. $x \in A \cap B \rightarrow x \in A$	1-3, conditional proof
5. $\forall x(x \in A \cap B \rightarrow x \in A)$	4, universal generalization, *x* arbitrary
6. $A \cap B \subseteq A$	5, definition of $\subseteq$

Theorem 15b, $A \subseteq A \cup B$
Proof.
Start of Subproof 1. (a conditional proof)

1. $x \in A$	assumption for conditional proof
2. $x \in A \vee x \in B$	1, addition
3. $x \in A \cup B$	2, theorem 5

End of Subproof 1.

4. $x \in A \rightarrow x \in A \cup B$	1-4, conditional proof
5. $\forall x(x \in A \rightarrow x \in A \cup B)$	4, universal generalization, *x* arbitrary
6. $A \subseteq A \cup B$	5, definition of $\subseteq$

Theorem 16. $\varnothing \subseteq A \subseteq \mathrm{V}$
This is two theorems.
16a. $\varnothing \subseteq A$
Proof. **Exercise.** Annotate the proof.

1. $\forall x(x=x)$
2. $x=x$
3. $\sim\sim(x=x)$
4. $\sim\sim(x=x) \vee x \in A$
5. $\sim(x=x) \rightarrow x \in A$
6. $x \neq x \rightarrow x \in A$
7. $\forall x(x \neq x \rightarrow x \in A)$
8. $\forall x(x \in \varnothing \rightarrow x \in A)$
9. $\varnothing \subseteq A$

16b. $A \subseteq \mathrm{V}$
Proof. **Exercise.** Construct the proof.

Theorem 17. $-(A\cup B)=-A\cap -B$
Proof.

1. $\forall x(\ {\sim}(x\in A \vee x\in B) \leftrightarrow {\sim}(x\in A \vee x\in B)\ )$ — logical truth
2. $\forall x({\sim}(x\in A \vee x\in B) \leftrightarrow({\sim}(x\in A)\ \&\ {\sim}(x\in B)))$ — 1, DeMorgan's Laws
3. $\forall x({\sim}(x\in A \vee x\in B) \leftrightarrow (x\notin A\ \&\ x\notin B))$ — 2, definition of $\notin$
4. $\forall x({\sim}(x\in A \vee x\in B) \leftrightarrow (x\in -A\ \&\ x\in -B))$ — 3, theorem 7
5. $\forall x({\sim}(x\in (A\cup B))\leftrightarrow (x\in -A\ \&\ x\in -B))$ — 4, theorem 5
6. $\forall x({\sim}(x\in (A\cup B))\leftrightarrow x\in -A\cap -B)$ — 5, theorem 4
7. $\forall x(x\notin (A\cup B)\leftrightarrow x\in -A\cap -B)$ — 6, definition of $\notin$
8. $\forall x(x\in -(A\cup B)\leftrightarrow x\in -A\cap -B)$ — 7, theorem 7
9. $-(A\cup B)=-A\cap -B$ — 8, Principle of Extensionality

Theorem 18. $-(A\cap B)=-A\cup -B$
Proof. **Exercise.** Construct the proof.

Theorem 19. $A\subseteq B \leftrightarrow -B\subseteq -A$
Proof. Part 1.
Start of Subproof 1. (a conditional proof)

1. $A\subseteq B$ — assumption for conditional proof
2. $\forall x(x\in A \rightarrow x\in B)$ — 1, definition of $\subseteq$
3. $\forall x({\sim}(x\in B) \rightarrow {\sim}(x\in A))$ — 2, contraposition
4. $\forall x(x\notin B \rightarrow x\notin A)$ — 3, definition of $\notin$
5. $\forall x(x\in -B \rightarrow x\in -A)$ — 4, theorem 7
6. $-B\subseteq -A$ — 5, definition of $\subseteq$

End of Subproof 1.
7. $A\subseteq B \rightarrow -B\subseteq -A$ — 1-6, conditional proof
Start of Subproof 2. (a conditional proof)

8. $-B\subseteq -A$ — assumption for conditional proof
9. $\forall x(x\in -B \rightarrow x\in -A)$ — 8, definition of $\subseteq$
10. $\forall x(x\notin B \rightarrow x\notin A)$ — 9, theorem 7
11. $\forall x({\sim}(x\in B) \rightarrow {\sim}(x\in A))$ — 10, definition of $\notin$
12. $\forall x(x\in A \rightarrow x\in B)$ — 11, contraposition
13. $A\subseteq B$ — 12, definition of $\subseteq$

End of Subproof 2.
14. $-B\subseteq -A \rightarrow A\subseteq B$ — 9-13, conditional proof
15. $A\subseteq B \leftrightarrow -B\subseteq -A$ — 7 and 14, biconditional

Theorem 20. $A\subseteq B\leftrightarrow\sim(A\cap-B\neq\varnothing)$ (*All A are B* iff *it is not that case that some are A are not B.)*

Proof. **Exercise.** Annotate the missing annotations.

Start of Subproof 1. (a conditional proof)

1. $A\subseteq B$ — assumption for conditional proof

Start of Subproof 1a. (a reductio)

2. $A\cap-B\neq\varnothing$ — assumption for reductio

> General Strategy. The goal in this part of the proof is to show $\sim(A\cap-B\neq\varnothing)$. The strategy is to assume its opposite $A\cap-B=\varnothing$ and deduce a contradiction. It does not matter what this contradiction is so long as it is of the general form $P\&\sim P$. If in a subproof you can deduce a contradiction from a formula, then you may terminate the subproof and write down a new line affirming its negation (opposite), and justifying the line by the subproof and the rule *Reductio ad Absurdum*. In this case the subproof is quite long. We do not reach a contradiction until line 28 and then the contradiction is quite long: $((x\in A\&\sim(x\in B))\leftrightarrow\sim(x=x))\&\sim((x\in A\&\sim(x\in B))\leftrightarrow\sim(x=x))$
> You can confirm that this a genuine contradiction of the form $P\&\sim P$ by letting
> P be $(x\in A\&\sim(x\in B))\leftrightarrow\sim(x=x))$ and
> $\sim P$ be $\sim((x\in A\&\sim(x\in B))\leftrightarrow\sim(x=x))$.

3. $\forall x(x\in A\rightarrow x\in B)$
4. $x\in A\rightarrow x\in B$
5. $\sim(A\cap-B=\varnothing)$
6. $\sim\forall x(x\in A\cap-B\leftrightarrow x\in\varnothing)$
7. $\sim\forall x((x\in A\&\ x\in-B)\leftrightarrow x\in\varnothing)$
8. $\sim\forall x((x\in A\&\ x\notin B)\leftrightarrow x\in\varnothing)$
9. $\sim\forall x((x\in A\&\sim(x\in B))\leftrightarrow x\in\varnothing)$
10. $\sim\forall x((x\in A\&\sim(x\in B))\leftrightarrow x\neq x)$
11. $\sim\forall x((x\in A\&\sim(x\in B))\leftrightarrow\sim(x=x))$
12. $\forall x(x=x)$
13. $x=x$
14. $\exists x\sim((x\in A\&\sim(x\in B))\leftrightarrow\sim(x=x))$ — 11, *what rule?*
15. $\sim((x\in A\&\sim(x\in B))\leftrightarrow\sim(x=x))$ — 12, existential instantiation, x is ***not*** arbitrary

Start of Subproof 1ai. (a conditional proof)

16. $x\in A\&\sim(x\in B)$ — assumption for conditional proof
17. $x\in A$ — *What line? What rule?*
18. $x\in B$
19. $\sim(x\in B)$
20. $x\in B\ \&\sim(x\in B)$
21. $\sim(x=x)$

End of Subproof 1ai.

22. $(x\in A\&\sim(x\in B))\rightarrow\sim(x=x)$ — 1-22, conditional proof

Start of Subproof 1b. (a conditional proof)

23. $\sim(x=x)$ — assumption for conditional proof
24. $x=x\ \&\sim(x=x)$
25. $x\in A\&\sim(x\in B)$

End of Subproof 1b.

26. $\sim(x=x)\rightarrow(x\in A\&\sim(x\in B))$ 23-25, conditional proof
27. $(x\in A\&\sim(x\in B))\leftrightarrow\sim(x=x)$
28. $((x\in A\&\sim(x\in B))\leftrightarrow\sim(x=x))\&\sim((x\in A\&\sim(x\in B))\leftrightarrow\sim(x=x))$
End Subproof 1a.
29. $\sim(A\cap-B\neq\varnothing)$ 2-28, reductio
(*note* 28 *is a contradiction*)

End of Subproof 1.
30. $A\subseteq B\rightarrow\sim(A\cap-B\neq\varnothing)$
Start of Subproof 2. (a conditional proof)
31. $\sim(A\cap-B\neq\varnothing)$
Start of Subproof 2a. (a conditional proof)
32. $x\in A$ assumption for conditional proof
33. $\sim\sim(A\cap-B=\varnothing)$
34. $A\cap-B=\varnothing$
35. $\forall x(x\in A\cap-B\leftrightarrow x\in\varnothing)$
36. $\forall x((x\in A\& x\in-B)\leftrightarrow x\in\varnothing)$
37. $\forall x((x\in A\& x\notin B)\leftrightarrow x\in\varnothing)$
38. $\forall x((x\in A\&\sim(x\in B))\leftrightarrow x\in\varnothing)$
39. $\forall x((x\in A\&\sim(x\in B))\leftrightarrow x\neq x)$
40. $\forall x((x\in A\&\sim(x\in B))\leftrightarrow\sim(x=x))$
41. $(x\in A\&\sim(x\in B))\leftrightarrow\sim(x=x))$
42. $(x\in A\&\sim(x\in B))\rightarrow\sim(x=x))$
Start of Subproof 2ai. (a reductio)
43. $\sim(x\in B)$ assumption for reductio
44. $x\in A\&\sim(x\in B)$
45. $\sim(x=x)$
46. $\forall x(x=x)$
47. $x=x$
48. $x=x\&\sim(x=x)$
End of Subproof 2ai.
49. $x\in B$ 43-58, reductio
End of Subproof 2a.
50. $x\in A\rightarrow x\in B$
51. $\forall x(x\in A\rightarrow x\in B)$
52. $A\subseteq B$
End of Subproof 2.
53. $\sim(A\cap-B\neq\varnothing)\leftrightarrow A\subseteq B$

Theorem 21. $\exists x(x\in A\cap B)\leftrightarrow(A\cap B\neq\varnothing)$ (*Some A are B* iff *it is not that case that no A are B.)*
Proof *__Exercise.__ Construct the proof.

Theorem 22. $A\in\mathbf{P}(A)$
Proof.

1. $A \subseteq A$	theorem 11
2. $A \in \mathbf{P}(A)$	1, theorem 8

Theorem 23. $\varnothing \in \mathbf{P}(A)$

Proof

***Exercise.** Annotate the proof

1. $\varnothing \subseteq A$
2. $\varnothing \in \mathbf{P}(A)$

Russell's Paradox

It is now time for us to critically evaluate set theory on logical grounds. Clearly it is mathematically precise. Moreover, it does a fair job of remaining true to the earlier usage of terms, and its empirical strength is testified to by its successful use in the formulations of numerous theories in the natural sciences. It must be admitted, however, that set theory's ontology is bloated. There are lots of sets. Nominalists, accordingly, are skeptical of set theory.

They have good reason. Despite all we have said, the naïve version is demonstrably incoherent. At the turn of the 20th century Bertrand Russell discovered that the axioms entail the contradiction that bears his name.[23] Russell's contradiction is called a paradox because it is seems unavoidable because it is entailed by the axioms that appear to be simple and true. The proof is one line long.

Russell's Paradox

Theorem 24. The Principle of Abstraction is false.

Proof.

1. $\forall x\ (x \in A \leftrightarrow x \notin x)$	Principle of Abstraction
2. $A \in A \leftrightarrow A \notin A.$	1, universal instantiation

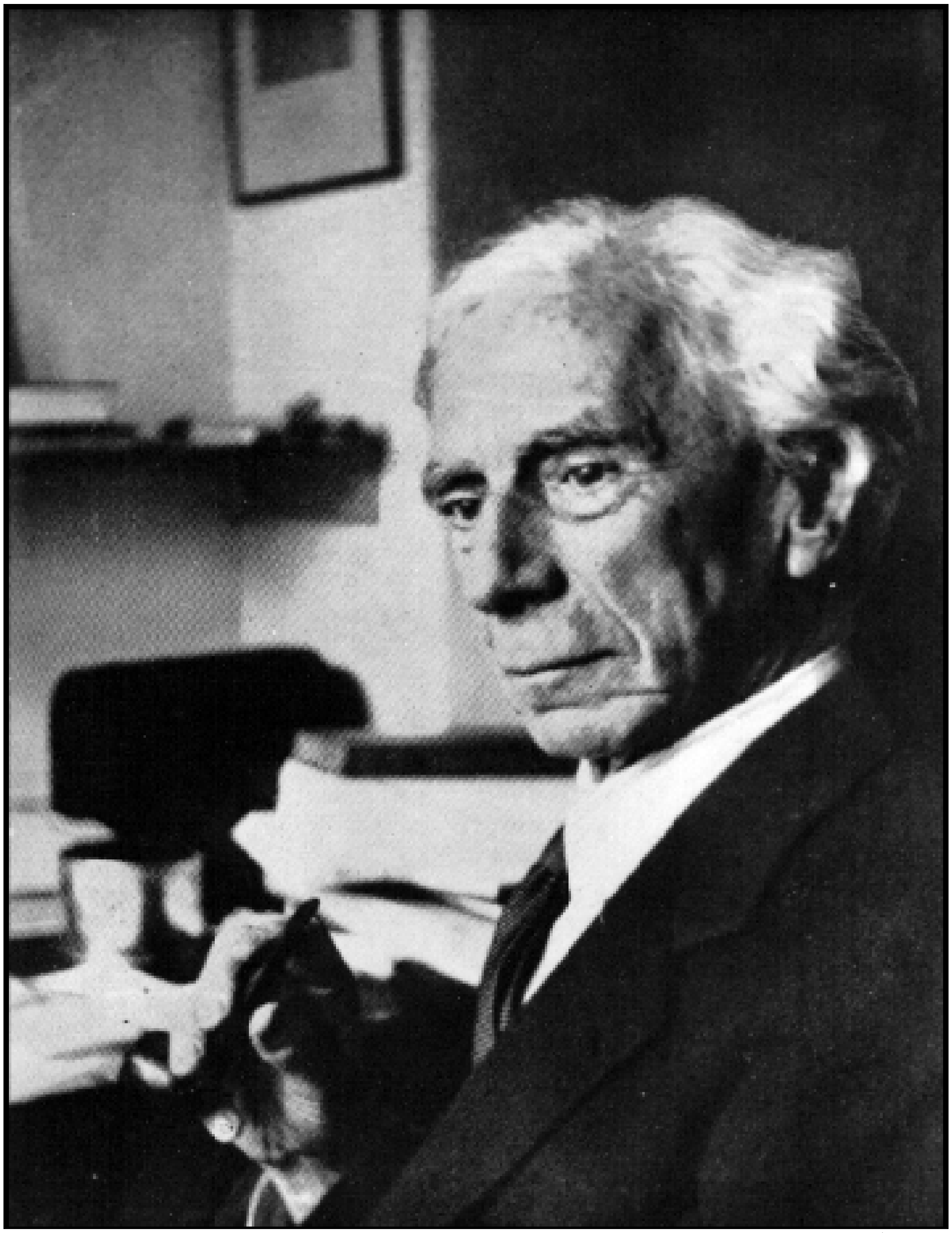
[23] See his account in the *Principles of Mathematics*, *op. cit.*

Line 2 here is contradictory in the sense that it is a logical impossibility. It is not possible for a proposition *P* to be such that it is true if and only if it is false. (It is a straightforward matter to complete the proof so that it concludes with an explicit contradiction of the form $P\&{\sim}P$.)[24] The proof establishes beyond any doubt that the Principle of Abstraction, a seemingly self-evident axiom, is false.

---

[24] The proof continues:

3. $A \in A \rightarrow A \notin A$	3, biconditional
Start subproof 1	
4. $A \in A$	assumption for reductio
5. $A \notin A$	2 and 4, modus ponens
6. $A \in A$ & $A \notin A$	4 and 5 conjunction
End subproof 1	
7. $A \notin A$	4-6, reductio
9. $A \notin A \rightarrow A \in A$	3, biconditional
Start subproof 2	
10. $A \notin A$	assumption for reductio
11. $A \in A$	9 and 10, modus ponens
12. $A \in A$ & $A \notin A$	10 and 11, conjunction
End subproof 2	
13. $A \in A$	10-12, reductio
14. $A \in A$ & $A \notin A$	7 and 13, conjunction



Bertrand Russell, aged 77 years

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

Axiomatized Set Theory

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

A number of diagnoses were proposed for the root of the problem. Russell's own account is that the principle errs in allowing sets that are ***un***grounded in the sense that they may form $\in$-loops. These are sets that may be a $\in$-descendent of themselves, for example a set x such that there is some chain $x\in y\in \ldots\in z\in x$. Here x is a member of something that is a member of something in a $\in$-hierarchy that eventually leads to a member of x itself. In 1910-13, together with Alfred North Whitehead (1861-1947), Russell published *Principia Mathematica*, an important work that revises the axioms so as to proscribe sets that form $\in$-loops. It does so by proposing the so-called theory of types in which sets form ranks such that only elements of one rank can enter into sets of the next. With this restriction an element x of rank n cannot be an element of itself at rank n+1. As far as is known, this new system is consistent. It does, however, require additional axioms, including a so-called axiom of reducibility, which requires, without much intuitive plausibility, that the set theoretic relations at higher levels be replicated in the structure of elements at the lowest level. Though the theory is technically successful in entailing the theorems necessary for applications of set theory to mathematics, a more intuitively plausible account is now preferred.[25]

This second explanation of the paradoxes is due to Ernst Zermelo (1871-1953).[26] According to his analysis the problem with the Principle of Abstraction is that it is over-generous in the size of the sets it asserts exist. According to the principle, a set of any size may exist so long as it is definable. Indeed, it directly implies that the universal set V exists and that there can be no set bigger than the set of everything. Russell's set $\{x \mid x \notin x\}$ too is "very large." It includes as a subset another very large set, the set of all cardinal numbers, which was shown independently to entail a contradiction (the Burali-Forti paradox). Zermelo proposes a new axiom system that specifies we start with a limited variety of sets, which are "small" enough that we can be fairly sure that they do not entail contradictions. These "starter sets" are limited to the empty set (the empty set axiom) and a set of countably many entities like the positive integers (the axiom of infinity). The system then specifies a restricted number of ways in which new sets may be constructed from those we previously know exist. One method is definability, but definability is restricted. Definable sets exist only if there is another set that we know already exists and either the old set contains the new set as one of its subsets (axiom of separation) or the elements of the old set can be mapped onto the elements of the new set (axiom of replacement). In addition to definability there are several other construction methods: forming a "pair" out of two previously existing sets (the pairing axiom), taking their union (the union axiom), forming a power set (the power set axiom), forming a set by taking a representative from each set in

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

a family of already existing sets (axiom of choice).[27] A far as is known, the system is consistent, and it has be developed within the branch of mathematics known as *axiomatized set theory* to demonstrate a large number of results, interesting to both mathematicians and philosophers.

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

---

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

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

Reduction of Relations to Sets

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

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

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

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

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

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

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

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

Definitions

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

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

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

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

Let us now group all n-tuples into a set and call this set V^n. [28] Any set of n-tuples then is a subset of V^n. We use this fact to define *n-place relation.*

Definitions

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

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

Theorem 26. (Extensionality for Relations).

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

Theorem 27. (Abstraction for Relations).

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

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

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

With these definitions we achieved a major goal. We have shown how relations are reducible to sets. Explaining sameness and difference for relational pairs is then a straightforward application of the same account given for sameness and difference for individuals.

Truth-Conditions for Relational Sentences and their Negations

- *Rab* is *true* if and only if the set of pairs that *R* names has as an element the ordered pair $<x,y>$ formed by the individual that *a* stands for and the individual that *b* stands for.
- ~*Rab* is *true* if and only if the set of pairs that *R* names does not have as an element the ordered pair $<x,y>$ formed by the individual that *a* stands for and the individual that *b* stands for.

[28] The notation derives from the fact that the number of n-tuples formed from elements of a set *A* is precisely the number of entities in *A* raised to the power *n.*

Explanation of Sameness and Difference of Relational Pairs

- Two pairs of individuals $<x,y>$ and $<z,w>$ are *the same* with respect to a set *U* if and only if they may be referred to by two pairs of different proper names, say *a* for *x* and *b* for *y,* and *c* for *z* and *d* for *w*, *U* may be named by a predicate, say *R*, and the sentences *Rab* and *Rcd* are both true.
- Two pairs of individuals $<x,y>$ and $<z,w>$ are *different* with respect to a set *U´* if and only if they may be referred to by two pairs of different proper names, say *a* for *x* and *b* for *y,* and *c* for *z* and *d* for *w*, *U´* may be named by a predicate, say *R*, and either the sentences *Rab* and ~*Rcd* are both true, or the sentences ~*Rab* and *Rcd* are both true.

Properties of Relations and Order

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

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

reflexive iff $\forall x,\ xRx$

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

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

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

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

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

In the last lecture we reviewed Russell's criticism of definitions of relations in terms of conjunctions of one-place predicates, namely that they end up attributing to relations properties they do not generally possess. The arguments sketched there may now be formulated as theorems in set theory.

Theorem 29. If $R=\{<x,y>|\ Fx \,\&\, Fy\}$ then R is symmetric
Proof.
Start of subproof.

1. xRy	assumption for conditional proof, x arbitrary
2. $<x,y>\in R$	1, infix notation
3. $<x,y>\in\{<z,w>|\ Fz \,\&\, Fw\}$	2, definition of R, with change of variables
4. $Fx \,\&\, Fy$	3, abstraction
5. $Fy \,\&\, Fx$	4, commutation
6. $<x,y>\in\{<z,w>|\ Fz \,\&\, Fw\}$	5, abstraction
7. $<x,y>\in R$	6, definition of R
8. xRy	7, infix notation

End of Subproof.

7. $xRy \rightarrow yRx$	1-8, conditional proof
8. $\forall x\ (xRy \rightarrow yRx)$	7, universal generalization, x arbitrary

Theorem 30. If $R=\{<x,y>|\ Fx \,\&\, Gy\}$ then R is transitive.
Proof.

***Exercise.** Construct the proof.

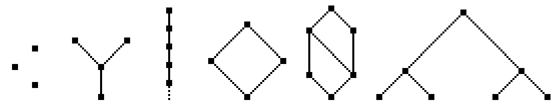
By imposing a relation with these properties on a set its elements may be "ordered".

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

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

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

A partial order is imposes a minimum amount of structure.



Examples of Partial Orderings

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

Functions

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

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

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

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

Definitions

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

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

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

*Abstract Structures

We all have a good intuitive idea of a "structure." Examples include buildings, governmental institutions, ecologies, and polyhedral. We have used the term structure to describe what scientific explanations "explain. They describe an ontology of sets and relations and by laws that state how they are put together. Plato claimed the cosmos was a structure of Forms, matter, souls organized by participation and recollection. Aristotle claimed it was a structure formed by the categories put together by the "said in" and "said of" relations. Set theory posits a universe of sets organized by the identity and set membership relations.

The general properties of structures are studied in the branch of mathematics known as *abstract* or *universal algebra*, a field that was started by Alfred North Whitehead in the fourth volume of *Principia Mathematica.*

The raw intuition behind the mathematical definition of a structure is that of an architect's blueprint. The blueprint succeeds in describing a building by first listing its various materials and then using a diagram to describe the relations that must obtain among the "building blocks" in the finished structure. In algebra a structure is defined in a similar way. First a list of sets $A_1,...,A_k$ is given. This may be viewed as listing the building blocks and dividing them into various kinds or classes. Next are listed the relations $R_1,...,R_i$ and functions $f_1,...,f_m$ that hold among these materials. (Recall that functions are just a sub-variety of relations.) In mathematics it is also customary to list some specific individual building blocks $O_1,...,O_m$ that have special importance in the structure. The entire structure is they summarized as an ordered n-tuple:

$$< A_1,...,A_k,R_1,...,R_i,f_1,...,f_m,O_1,...,O_m >.$$

Definition. An *abstract structure* is any $<A_1,...,A_k,R_1,...,R_l,f_1,...,f_m,O_1,...,O_n>$ such that:

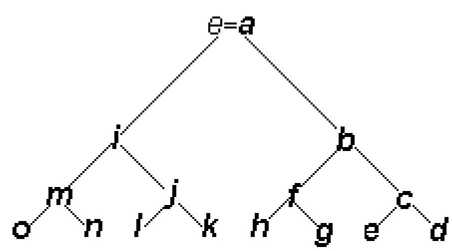
for each i=1…k, A_i is a set,
for each i=1…l, R_i is a relation of elements in $A_1,\cup..\cup A_k$,
for each i=1…m, f_i is a function of elements in $A_1,\cup..\cup A_k$, and
for each i=1…n, $O_i \in A_1,\cup..\cup A_k$.

It is also common to investigate families of structures with similar properties, and to assign a name to families obeying certain ”laws”. Families are defined as those whose structural relations and functions obey these laws. Sometimes these restrictions are referred to as the "axioms" of the structure-type. As an example, let us define the notion of a “tree.” We encountered informal versions of trees in the ontologies of Plato and Aristotle, but one of the weakness of their theories is that they lack the sort of precision that explains exactly what structure was being presupposed. Using set theory we can now make the idea clear.

Definition. A *tree* is any structure $<U,\leq,e>$ meeting these conditions:

1. $<U,\leq>$ is a partial ordering,
2. 1 is the unique maximal element e in U, i.e.
   a. $1\in \underline{U}$
   b. $\forall x\ (x\in U\rightarrow x\leq e)$
   c. $\forall y\ (\forall x\ (x\in U \rightarrow x\leq y)\rightarrow y= e)$
3. Every element of U is the last member of a branch that starts with 1, i.e. $\forall x(x\in U \rightarrow \exists\ y_n,\ldots,y_1\ (\{y_n,\ldots,y_1\}\subseteq U\ \&\ y_n=x\ \&\ y_1= e\ \&\ \forall i\ (i\in\{\ e,\ldots,n\}\rightarrow (\ y_{i+1}\leq y_i\ \&\ \forall z(y_{i+1}\leq z\ \&\ z\leq y_i)\rightarrow(z= y_{i+1} \vee z= y_{i+1}\ ))))$

As an example, consider the structure $<(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d},\boldsymbol{e},\boldsymbol{f},\boldsymbol{g},\boldsymbol{h},\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{l},\boldsymbol{m},\boldsymbol{n},\boldsymbol{o},\},\leq,\boldsymbol{a}>$ in which the relation $\leq$ is defined as follows: $x\leq y$ iff x is connected by a descending line to y in the diagram below:



It is easy to check that the structure meet the conditions (1)-(3) qualifying it as a tree. It root node e is ***a***, and $<\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d},>$ is one of its maximal branches, the one which starts with ***a*** and ends with ***d***.

Construction and Inductive Definitions

Definitions by Necessary and Sufficient Conditions

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

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

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

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

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

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

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

These definition fit a general form:

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

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

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

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

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

When it does so, it entails the theorem:

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

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

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

which entails

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

Aristotelian real definitions have a similar structure:

∀*x* (*x is a man* ↔ (*x is rational* & *x is an animal*))

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

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

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

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

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

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

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

Inductive Definitions and Sets

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

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

1. E is a set of basic elements;

2. *R* is a set of relations;
3. *C* is the set such that[29]
   a. *E* is a subset of *C;*
   b. if the elements $x_1,\ldots,x_n$ are in *C* and bear the relation *R* to x_{n+1}, then x_{n+1} is in *C;*
   c. nothing else is in *C.*

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

It is possible to add more restrictions that would insure that *C* will not generate paradoxes.[30] Were we to do so, the set would be genuinely constructive in a strict sense.

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

---

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

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

The we define *C*as follows:

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

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

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

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

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

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

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

The Natural Numbers

Another standard example of a set defined by induction is the set **Error! Bookmark not defined.*N*** of *natural numbers*, which consists of all the positive integers 1,2,3,… plus 0. Let us work through it in some detail because though easy to state, it illustrates the power of inductive definitions. The set is constructed. We start with 0 as the only initial element. We then define the so-called successor relation. The natural numbers then are inductively defined as the set of all entities that can be constructed from 0 by the successor relation.

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

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

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

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

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

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

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

0				$= \varnothing$
$1 =$	$0+1 =$	$\varnothing+1$	$= \varnothing\cup\{\varnothing\}$	$= \{\varnothing\}$
$2 =$	$1+1 =$	$\{\varnothing\}+1$	$= \{\varnothing\}\cup\{\{\varnothing\}\}$	$= \{\varnothing,\{\varnothing\}\}$
$3 =$	$2+1 =$	$\{\varnothing,\{\varnothing\}\}+1$	$= \{\varnothing,\{\varnothing\}\}\cup\{\{\varnothing,\{\varnothing\}\}\}$	$= \{\varnothing,\{\varnothing\},\{\varnothing,\{\varnothing\}\}\}$

- there are in infinite number of natural numbers,
- each natural number has exactly as many members as the number suggests:

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

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

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

3, aka $\{\varnothing,\{\varnothing\},\{\varnothing,\{\varnothing\}\}\}$, has three members, namely $\varnothing$, $\{\varnothing\}$, and $\{\varnothing,\{\varnothing\}\}$ etc.

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

Indeed, these are the definitions that generate the structure $<\boldsymbol{N},\leq,+,\times,0,1>$ which most of you spent hours working out the details of in high school algebra.[31]

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

[31] You may for example have learned to work out the equations that are true in an "ordered field" $<\boldsymbol{N},\leq,+,\times,0,1>$ or "distributive ring," which are structures that obey "laws" that are exemplified and abstracted from the natural numbers.

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

Proof by Induction

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

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

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

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

If 1. $E\subseteq A$, and

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

then $C\subseteq A$

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

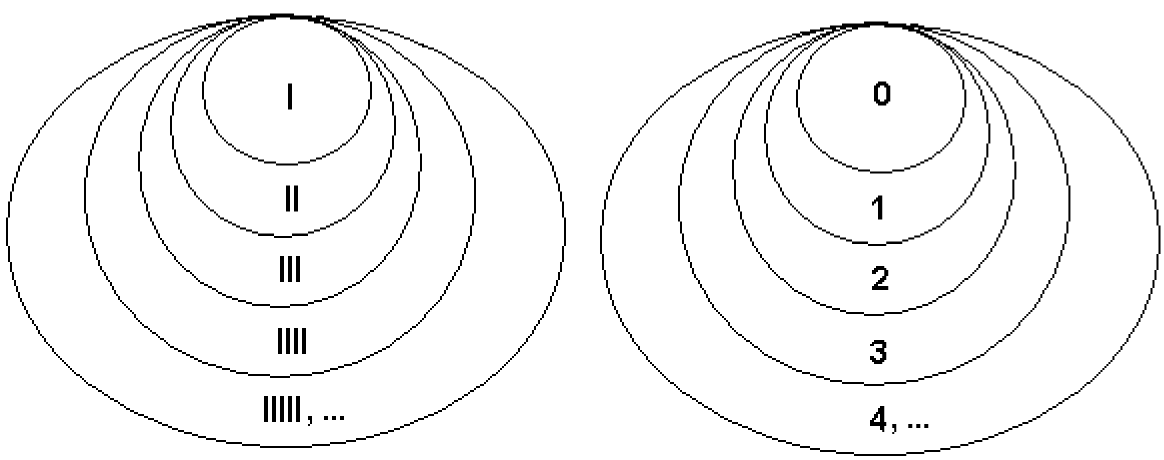
Construction Sequences

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

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

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

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



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

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

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

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

APPENDIX I. INFERENCE RULES

♦♦**Logical Truth**. Any logical truth is a theorem.

$P \vee {\sim}P$

$\forall x(x{=}x)$, and

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

Sentential Logic Rules

♦♦***Modus ponens***

$P \to Q$

$\underline{\quad P}$

$\therefore Q$

♦***Modus tollens***

$P \to Q$

$\underline{{\sim}Q}$

$\therefore {\sim}P$

Disjunctive Syllogism

$P \vee Q \qquad P \vee Q$

$\underline{{\sim}P} \qquad \underline{{\sim}Q}$

$\therefore Q \qquad \therefore P$

♦**Hypothetical Syllogism**

$P \to Q$

$\underline{Q \to R}$

$\therefore P \to R$

♦♦**Conjunction.**

$\underline{P\&Q} \qquad \underline{P\&Q} \qquad P$

$\therefore P \qquad\quad Q \qquad \underline{Q}$

$\therefore P\&Q$

Addition

$\underline{\quad P \quad} \qquad \underline{\quad Q}$

$\therefore P \vee Q \qquad \therefore P \vee Q$

Quantifier Rules

♦♦**Universal Generalization**

$\underline{P(x)} \qquad \underline{P(y)} \qquad$ *x, y* arbitrary

$\therefore \forall x\, P(x \quad \therefore\ \forall x P(x)$

♦♦**Universal Instantiation**

$\underline{\forall x P(x)} \qquad \underline{\forall x P(x)}$

$\therefore\ P(x) \qquad \therefore\ P(y) \qquad$ *x, y* arbitrary

♦♦**Existential Construction**

$\underline{P(c)}$

$\therefore\ \exists x P(x)$

♦♦**Existential Instantiation**

$\underline{\exists x P(x)} \qquad \underline{\exists x P(x)}$

$\therefore\ P(x) \qquad \therefore\ P(y) \qquad$ *x,y* new and ***not*** arbitrary

Substitution Rule

If $P \leftrightarrow Q$ or $s{=}t$ is proven (with the same free variables), *P* may be substituted for *Q*, or *s* for *t*, or *vice versa*, in any subsequent line. **Examples:**

Association

$((P\&Q)\&R) \leftrightarrow (P\&(Q\&R))$

$((P \vee Q) \vee R) \leftrightarrow (P \vee (Q \vee R))$

Commutation

$(P\&Q) \leftrightarrow (Q\&P)$

$(P \vee Q) \leftrightarrow (Q \vee P)$

$(P \leftrightarrow Q) \leftrightarrow (Q \leftrightarrow P)$

DeMorgan's Laws

${\sim}(P\&Q) \leftrightarrow ({\sim}P \vee {\sim}Q)$

${\sim}(P \vee Q) \leftrightarrow ({\sim}P\&{\sim}Q)$

♦**Double Negation**

${\sim}{\sim}P \leftrightarrow P$

♦**Implication**

$(P \to Q) \leftrightarrow ({\sim}P \vee Q)$

${\sim}(P \to Q) \leftrightarrow (P \,\&\, {\sim}Q)$

Contraposition

$(P \to Q) \leftrightarrow ({\sim}Q \to {\sim}P)$

Tautology

$(P\&P) \leftrightarrow (P \vee P) \leftrightarrow P$

The Biconditional

$(P \leftrightarrow Q) \leftrightarrow ((P \to Q)\& (Q \to P))$

$(P \leftrightarrow Q) \leftrightarrow ((P\&Q) \vee ({\sim}P\&{\sim}Q))$

Quantifier Negations

${\sim}\forall x P(x)$	$\leftrightarrow$	$\exists x {\sim}P(x)$
${\sim}\exists x P(x)$	$\leftrightarrow$	$\forall x {\sim}P(x)$
${\sim}\forall x(P(x) \to Q(x))$	$\leftrightarrow$	$\exists x(P(x)\&{\sim}Q(x))$
${\sim}\exists x(P(x)\&Q(x))$	$\leftrightarrow$	$\forall x(P(x) \to {\sim}Q(x))$

Subproof Rules

♦***Reductio ad absurdum***

If $\underline{\quad P}$ then ${\sim}P$

$\therefore Q\&{\sim}Q$

If $\underline{\quad {\sim}P}$ then P

$\therefore Q\&{\sim}Q$

Ex Falso Quodlibet

$\underline{P\&{\sim}P}$

$\therefore Q$

Proof by Cases

If $\underline{P}$ and $\underline{Q}$ then $\underline{P \vee Q}$

$\therefore R \qquad \therefore R \qquad \therefore R$

♦♦**Conditional Proof**

If $\underline{P}$ then $P \to Q$

$\therefore Q$

Conditional Proof for Biconditionals.

If $\underline{P}$ and $\underline{Q}$, then $P \leftrightarrow Q$

$\therefore Q \qquad \therefore P$

Appendix II. Naïve Set Theory

Logical Symbols

a,b,c	constants (proper names)
x,y,z	variables (pronouns)
F,G,H	predicates (commn nouns, intrans.verbs, adjectives)
$\sim$, $\neg$	negation (*not*)
&, $\wedge$	conjunction (*and*)
$\vee$	disjunction (*or*)
$\rightarrow$	the conditional (*if…then*)
$\leftrightarrow$	the biconditional (*if & only if*)
$\forall$	univ. quant. (*for all*)
$\exists$	exist. quant. (*for some*)
$\in$	membership (*is a member of*)
=	identity

Axioms

♦Logical Truth. *Every truth of logic is a theorem.*

♦Extensionality. $A=B \leftrightarrow \forall x\,(x\in A \leftrightarrow x\in B)$

♦Abstraction. $\exists A\ \forall x\ (\,x\in A \leftrightarrow P(x))$

Abbreviations

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

Theorems

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

♦2. $\forall x\,(x\in\varnothing \leftrightarrow x\neq x)$

3. $\forall x\,(x\in V \leftrightarrow x=x)$

♦4. $\forall x\,(x\in A\cap B \leftrightarrow (x\in A\&x\in B)\,)$

♦5. $\forall x\,(x\in A\cup B \leftrightarrow (x\in A\vee x\in B)\,)$

♦6. $\forall x\,(x\in A-B \leftrightarrow (x\in A\&x\notin B)\,)$

♦7. $\forall x\,(x\in -A \leftrightarrow x\notin A)$

8. $\forall B\,(B\in \mathbf{P}(A) \leftrightarrow B\subseteq A)$

9. $\forall y\,(y\in\{\,x_1, \ldots, x_n\,\} \leftrightarrow (y = x_1 \vee \ldots \vee y = x_n)\,)$

10. $--A=A$

11. $A\subseteq A$

12. $\forall x((x\in A\ \&\ A\subseteq B)\rightarrow x\in B)$

13. $A\cap A=A=A\cup A$

14. $A=B \leftrightarrow (A\subseteq B\ \&\ B\subseteq A)$

15. $A\cap B\subseteq A\subseteq A\cup B$

16. $\varnothing\subseteq A\subseteq V$

17. $-(A\cup B)=-A\cap -B$

18. $-(A\cap B)=-A\cup -B$

19. $A\subseteq B \leftrightarrow -B\subseteq -A$

20. $A\subseteq B \leftrightarrow \sim(A\cap -B \neq\varnothing)$

21. $\exists x(x\in A\cap B) \leftrightarrow \sim(A\cap B\neq\varnothing)$

22. $A\in \mathbf{P}(A)$

23. $\varnothing\in \mathbf{P}(A)$

24. $\sim\exists A\ \forall x\ (\,x\in A \leftrightarrow P(x))$

25. $\forall x_1,\ldots,x_n,\ (\langle x_1,\ldots,x_n\rangle=\langle y_1,\ldots,y_n\rangle \leftrightarrow \forall\ x_1,\ldots,x_n\forall\ y_1,\ldots,y_n\ (x_i = y_i))$

26. $(R\subseteq V^n\ \&\ S\subseteq V^n\ ) \rightarrow(R=S \leftrightarrow \forall x_1,\ldots,x_n(\langle x_1,\ldots,x_n\rangle\in A \leftrightarrow \langle x_1,\ldots,x_n\rangle\in B))$

27. $\exists A\ \forall\ x_1,\ldots,x_n\ (\langle x_1,\ldots,x_n\rangle\in A \leftrightarrow P(x_1,\ldots,x_n))$

Lecture 5

Consult the theorem *immediately prior* to each of these in the notes for examples of a similar proof. The strategies are very similar, and use similar, if not the same rules.

If the exercise asks you to annotate a proof you should write to its right of each line of the proof the axiom or theorem of which it is an instance (e.g."Axiom of Abstraction" or "Theroem 5") or the numbers of previous line from which this line is deduced along with rule of logic used (e.g "1,2 *modus ponens*"). See the examples in the text.

Print out this summary and write your answers on the print out.

Annotate the proofs:

Theorem 3. $\forall x\,(x\in \mathrm{V} \leftrightarrow x=x)$
Proof.

1. $\forall y(y\in\{x|\ x=x\} \leftrightarrow y=y)$
2. $\forall y(y\in \mathrm{V} \leftrightarrow y=y)$
3. $\forall x\,(x\in \mathrm{V} \leftrightarrow x=x)$

Theorem 5. $\forall x\,(x\in A\cup B \leftrightarrow (x\in A\vee x\in B)\,)$
Proof.

1. $\forall y(y\in\{x|\ x\in A\vee x\in B\,\} \leftrightarrow (y\in A\vee y\in B))$
2. $\forall y(y\in\ A\cup B \leftrightarrow (y\in A\vee y\in B))$
3. $\forall x\,(x\in\ A\cup B \leftrightarrow (x\in A\vee x\in B))$

Theorem 7. $\forall x\,(x\in -A \leftrightarrow x\notin A)$
Proof.

1. $\forall y(y\in \mathrm{V}\&{\sim}y\in A\ ) \leftrightarrow {\sim}y\in A)$
2. $\forall y(y\in \mathrm{V}\&y\notin A\ ) \leftrightarrow y\notin A)$
3. $\forall y(y\in \mathrm{V}-A \leftrightarrow y\notin A)$
4. $\forall x\,(x\in -A \leftrightarrow x\notin A)$

Exercise. Annotate line 2-4 of the proof.
Truth of Logic (Axiom 1 and Theorem 3)

Construct the proofs:

Theorem 9. $\forall y\,(y\in\{\,x_1,\ \ldots,x_n\,\} \leftrightarrow (y = x_1 \vee\ \ldots \vee\ y = x_n)\,)$
Proof.

Theorem 11. $A \subseteq A$
Proof.

Annotate the proofs:

Theorem 13b. $A = A \cup A$.
Proof.
1. $\forall x(x \in A \leftrightarrow x \in A)$
2. $\forall x(x \in A) \leftrightarrow (x \in A \vee x \in A))$
3. $\forall x(x \in A \leftrightarrow x \in A \cup A)$
4. $A = A \cup A$

Theorem 16a. $\varnothing \subseteq A \subseteq \mathrm{V}$
Proof.
1. $\forall x(x = x)$
2. $x = x$
3. $\sim\sim(x = x)$
4. $\sim\sim(x = x) \vee x \in A$
5. $\sim(x = x) \rightarrow x \in A$
6. $x \neq x \rightarrow x \in A$
7. $\forall x(x \neq x \rightarrow x \in A)$
8. $\forall x(x \in \varnothing \rightarrow x \in A)$
9. $\varnothing \subseteq A$

Construct the proofs:

Theorem 16b. $A \subseteq \mathrm{V}$
Proof.

Theorem 18. $-(A \cap B) = -A \cup -B$
Proof.

More Challenging Exercises

****Complete the annotation:**
Theorem 20. $A \subseteq B \leftrightarrow \sim(A \cap -B \neq \varnothing)$ (*All A are B* iff *it is not that case that some are A are not B.)*
Proof.
Start of Subproof 1. (a conditional proof)

1. $A \subseteq B$	assumption for conditional proof
Start of Subproof 1a. (a reductio)	
2. $A \cap -B \neq \varnothing$	assumption for reductio
3. $\forall x(x \in A \rightarrow x \in B)$	
4. $x \in A \rightarrow x \in B$	
5. $\sim(A \cap -B = \varnothing)$	
6. $\sim\forall x(x \in A \cap -B \leftrightarrow x \in \varnothing)$	
7. $\sim\forall x((x \in A \& x \in -B) \leftrightarrow x \in \varnothing)$	
8. $\sim\forall x((x \in A \& x \notin B) \leftrightarrow x \in \varnothing)$	
9. $\sim\forall x((x \in A \& \sim(x \in B)) \leftrightarrow x \in \varnothing)$	
10. $\sim\forall x((x \in A \& \sim(x \in B)) \leftrightarrow x \neq x)$	
11. $\sim\forall x((x \in A \& \sim(x \in B)) \leftrightarrow \sim(x = x))$	
12. $\forall x(x = x)$	
13. $x = x$	
14. $\exists x\sim((x \in A \& \sim(x \in B)) \leftrightarrow \sim(x = x))$	11, *what rule?*
15. $\sim((x \in A \& \sim(x \in B)) \leftrightarrow \sim(x = x))$	14, existential instantiation,

x is ***not*** arbitrary

Start of Subproof 1ai. (a conditional proof)

16. $x\in A\&\ {\sim}(x\in B)$ — assumption for conditional proof
17. $x\in A$ — *What line? What rule?*
18. $x\in B$
19. ${\sim}(x\in B)$
20. $x\in B\ \&\ {\sim}(x\in B)$
21. ${\sim}(x=x)$

End of Subproof 1ai.

22. $(x\in A\&\ {\sim}(x\in B))\rightarrow{\sim}(x=x)$ — 1-22, conditional proof

Start of Subproof 1b. (a conditional proof)

23. ${\sim}(x=x)$ — assumption for conditional proof
24. $x=x\ \&{\sim}(x=x)$
25. $x\in A\&\ {\sim}(x\in B)$

End of Subproof 1b.

26. ${\sim}(x=x)\rightarrow(\ x\in A\&\ {\sim}(x\in B))$ — 23-25, conditional proof
27. $(x\in A\&\ {\sim}(x\in B))\leftrightarrow{\sim}(x=x)$
28. $((x\in A\&\ {\sim}(x\in B))\leftrightarrow{\sim}(x=x))\&{\sim}\ ((x\in A\&\ {\sim}(x\in B))\leftrightarrow{\sim}(x=x))$

End Subproof 1a.

29. ${\sim}(A\cap-B\neq\varnothing)$ — 2-28, reductio (*note* 28 *is a contradiction*)

End of Subproof 1.

30. $A\subseteq B\rightarrow{\sim}(A\cap-B\neq\varnothing)$

Start of Subproof 2. (a conditional proof)

31. ${\sim}(A\cap-B\neq\varnothing)$

Start of Subproof 2a. (a conditional proof)

32. $x\in A$ — assumption for conditional proof
33. ${\sim}{\sim}(A\cap-B=\varnothing)$
34. $A\cap-B=\varnothing$
35. $\forall x(x\in A\cap-B\leftrightarrow x\in\varnothing)$
36. $\forall x((x\in A\&\ x\in-B)\leftrightarrow x\in\varnothing)$
37. $\forall x((x\in A\&\ x\notin B)\leftrightarrow x\in\varnothing)$
38. $\forall x((x\in A\&\ {\sim}(x\in B))\leftrightarrow x\in\varnothing)$
39. $\forall x((x\in A\&\ {\sim}(x\in B))\leftrightarrow x\neq x)$
40. $\forall x((x\in A\&\ {\sim}(x\in B))\leftrightarrow{\sim}(x=x))$
41. $(x\in A\&\ {\sim}(x\in B))\leftrightarrow{\sim}(x=x))$
42. $(x\in A\&\ {\sim}(x\in B))\rightarrow{\sim}(x=x))$

Start of Subproof 2ai. (a reductio)

43. ${\sim}(x\in B)$ — assumption for reductio
44. $x\in A\&\ {\sim}(x\in B)$
45. ${\sim}(x=x)$
46. $\forall x(x=x)$
47. $x=x$
48. $x=x\ \&{\sim}(x=x)$

End of Subproof 2ai.

49. $x \in B$ 43-58, reductio

End of Subproof 2a.

50. $x \in A \rightarrow x \in B$

51. $\forall x(x \in A \rightarrow x \in B)$

52. $A \subseteq B$

End of Subproof 2.

53. $\sim(A \cap -B \neq \varnothing) \leftrightarrow A \subseteq B$

Lecture 6

***Construct the proof:**

Theorem 21. $\exists x(x \in A \cap B) \leftrightarrow (A \cap B \neq \varnothing)$ (*Some A are B* iff *it is not that case that no A are B.)*

Proof

***Annotate the proof:**

Theorem 23. $\varnothing \in \mathbf{P}(A)$

Proof

1. $\varnothing \subseteq A$
2. $\varnothing \in \mathbf{P}(A)$

***Construct the proof:**

Theorem 30. If $R=\{<x,y>|\ Fx\ \&Gy\}$ then R is transitive.

Proof.

.

REVIEW QUESTIONS

1. State the problem of sameness and difference, first in general terms and then by formulating it in terms of the truth-conditions for subject-predicate propositions.

2. Explain how the problem is purportedly solved by:

    a. Plato's Theory of Forms,

    b. Aristotle's theory of the categories and predicables,

    c. Set Theory.

3. Critically evaluate the three explanations (a-c) as follows. Explain in what ways set theory is better than the accounts of Plato and Aristotle. But even set theory has problems. Sketch why its naïve version is contradictory, and explain in general terms how we deal with this fact.