NOTES ON THE SCIENCE OF LOGIC

Nuel Belnap

University of Pittsburgh

2009

February 25, 2009

Draft for use by friends

Contents

List of exercises

Preface

These notes began many years ago as an auxiliary to the fine text Hunter 1971, which he updated in 2001 (University of California Press). Please bear with inadequacies.

P. Woodruff taught us several of the key ideas around which these notes are organized. Thanks to J. M. Dunn for suggestions, and to W. O'Donahue and B. Basara for collecting errors. G. A. Antonelli and P. Bartha contributed greatly, as did a variety of necessarily anonymous students.

Chapter 1

Preliminaries

1A Introduction

This course assumes you know how to use truth functions and quantifiers as tools; such is part of the *art* of logic. Our principal task here will be to study these very tools; we shall be engaged in part of the *science* of logic.

1A.1 Aim of the course

Our aim is not principally mathematical, but foundational; undoubtedly "logical" is the best word. We are not aiming to hone your *mathematical* intuitions, and this is not a "mathematical" study. When we prove a theorem, we shall therefore not try to do it as swiftly and intuitively as possible, which is the aim of the mathematician, but instead we shall try to keep track of the *principles* involved.

Principles, however, come at many levels, so we must be more specific. In contrast to much of mathematical logic, proofs in these notes will use what most logicians consider only the most elementary of steps. In fact, this presentation of logic is unlike any other of which we know in that (unless clearly labeled) *there is no armwaving* in the following perfectly good sense: Whenever we give a proof, it is possible *for you* to fully formalize this proof using *only* techniques you learned in a one-term course in logic. In particular, (1) no proof relies on geometrical or arithmetical intuition to gain your assent to its conclusion, since we do not presuppose your mastery of either of these topics, and (2) no proof is so complicated or so long that you cannot see how it goes. With regard to (1), if our aim were the mathematical one of helping you to acquire the ability to see that our conclusions were true, then reliance on arithmetic or geometrical intuitions would be reasonable; but instead our principal aim is the *logical* one of seeing *how*, using only quantifier and truth-functional principles, our conclusions follow from the fundamental axioms and definitions that characterize our subject matter. For example, it is "obvious" that every formula contains at most a finite number of atomic symbols—you can "see" that this must be so. These notes will rely on no such "vision," but will instead analyze the principles lying behind such intuitions so that for justifying proofs the intuitions themselves can be dispensed with.

With regard to (2), it is equally important that we organize the work in such a way that you can see how all the quantifier and truth-functional steps are put together—so that you can "survey" our proofs; for unless you can do that, you cannot after all see how our conclusions follow. In short, we bring to proofs in the subject the same standards of rigor and comprehensibility that are enshrined in the logic of which we are speaking. You must be aware, however, that to be rigorous is an ideal, and one that we have not always attained. Our more realistic hope is that we have been sufficiently clear so that with diligence you can see where we have fallen short—as successive groups of students have done.

On the other hand, without intuition conception is blind, and we shall make every effort to give you "pictures" of theorems and proofs to help you see what is going on. Attaining a how-to-carry-out-the proof understanding of these matters is not enough—you *also* need to acquire an arithmetic and geometrical *and philosophical* intuitive understanding.

There is one foundational matter, however, that we do not treat in this course: the justification of definitions of operators by proving the existence of functions satisfying given conditions (see NAL:**7B-7**). In fact what permits us to be rigorous without overwhelming you is that whenever a definition is justified, we feel free to employ it; but the justification itself, which is indeed an important matter for logic, is left for another course.

We are furthermore compelled to say that there are some topics that we do not treat in these notes only because we do not yet know how to do so; for these we have so far been unable to find treatments that both restrict themselves to truth-functions-and-quantifier steps[1] and are also comprehensible.

[1]The adjective "first order" is frequently used as short for "quantifier-and-truth-function-and-identity" in the special case that all quantification is with respect to variables that occupy the places of singular terms, and never with respect to variables standing in place of predicates, operators, or connectives. The sense of rigor encouraged here does not require limitation to the first order—though by and large that is where we shall stay.

Finally, we endorse M. McRobbie's view that logic is at root all about trees; and you will find that these notes are better at attending to a variety of trees than they are at offering a map of the forest. In the References at the end you will find a partly annotated list of books that you might wish to consult. Several of these books can be relied upon for a complementary broader perspective—sometimes at the expense of rigor.

1A.2 Absolute prerequisites and essential background

As we have said, it is assumed that you have mastered truth functions and quantifiers as tools. The elementary part of any branch of logic involves four parts or aspects: grammar, semantics, proof theory, and applications. Hence, to understand truth functions and quantifiers as tools means the following.

1. You should be in command of the *grammar* of truth functions and quantifiers.

2. You should understand at the informal level their *semantics*: truth tables, domains, interpretations of predicate letters and constants (and perhaps operators).

3. You should be in command of some natural deduction technique for constructing *proofs*, including conditional proof (perhaps you know it under another name), reductio ad absurdum (or indirect proof), and the four quantifier rules.

4. Under "*applications*" you should know how to translate between henscratches and English. Fluently.

All of this is covered in the *Notes on the art of logic* (NAL). You may use NAL (1) for review, and (2) to become acquainted with the style of proof we will be using. You do not have to learn how to construct proofs in the style of these notes, however, if you wish to stay with some other system. The same remark goes for notation. You need to learn to read ours, but you do not have to learn to write it.[2]

We will begin by reviewing this material with extreme brevity.

In addition to the absolutely standard items listed above, there are a series of topics that are covered in an ideal one-term logic course such as we are presupposing, but

[2]Although this remark continues to be true, it is also worth observing that as yet no student has both learned this material and not learned to write proofs in the style of NAL.

which are left out of some. We will go over these a little more slowly, but still leave it largely to you to learn this material on your own. We provide a few references to NAL, and to other sections of these notes. The topics are as follows.

1. Logical grammar. Important as background and essential for every logician, but not technically required. See NAL:§1A.
2. Fitch's method of subordinate proofs. See NAL:§2C and NAL:chapter 3.
3. Definitions. Ditto as for logical grammar. See NAL:chapter 7.
4. Easy set theory. *Absolutely required*. See NAL:§9A. We use "easy set theory" or "EST" to reference this material.
5. Elementary theory of functions and relations. Ditto as for elementary set theory; see NAL:§9B.
6. In addition, we explain certain set-theoretical ideas on an as-needed basis in these notes; but it is proper to delay learning these ideas.
7. Elementary arithmetic. Ditto as for elementary set theory. You especially need to be able to use induction on numbers; see NAL:§6B for one way of presenting the material; but the absolutely essential material is covered in §2C.1.

1A.3 About exercises

Preliminary exercises. Anyone who can do all of the exercises in NAL is better off than anyone who cannot do all those exercises.

On the other hand, although making fully efficient use of these notes requires something like the preparation indicated above, still, with a critical exception, it is sufficient to have taken just about any good one-term course in symbolic logic that includes a thorough grounding in *relational* quantifier logic. The exception is this: These notes on occasion use the style of NAL in order to give proofs or describe admissible rules. They do so when it is believed that communication is thereby served. And even though presentation in these notes of proofs and rules in that style is always redundant (proofs and rules are always given in English as well), still, to profit from these theoretically redundant passages, you need to learn the style of proof of NAL (due essentially to Fitch) even if you are already familiar with some other logical system. Exactly which sections of NAL are required for this limited purpose? Answer (as also indicated above): NAL:§2C and NAL:chapter 3.

A word on the exercises in these notes. Throughout these notes you will find appropriate exercises, some scattered, some gathered. These exercises have been kept straightforward, even boring (but not necessarily "easy"). Except for some labeled "optional," their purpose is not to encourage the creative mathematician in you, but rather to assist you in the "elementary mastery" of the central concepts of the science of logic.

1A.4 Some conventions

In order to facilitate your study of this material, we have tried to adhere to certain uniform ways of saying things; these we list here for your convenience of reference.

1A-1 CONVENTION. *(Use language)*

Following a conception of Curry, our *use language* is the language that we—that is, we and you both—are using. We use our use language for many things, among them to discuss our own language and that of others.

1A-2 CONVENTION. *(Policy on numbered statements and sections)*

Numbered statements are numbered within sections (which in turn are numbered—that is, lettered—within chapters), and we always use boldface, e.g. **1A-2**, when referring to one of these numbered statements. In contrast, references to chapters, sections, or subsections are always plain, as e.g. §1A or §2C.1. We use e.g. NAL:§6B and NAL:**9A-3** respectively to refer to a section or to a numbered statement in *Notes on the art of logic*.

Numbered statements are labeled as one of the following:

- CONVENTION. Such an item will state in a relaxed way some agreement as to how we propose to use the language with which we communicate with you—our *use language*, as we said in Convention **1A-1**. You do not in general need to refer to conventions.
- AXIOM. An axiom is a clear-cut, formalizable postulate that is to be *used* in proving something. Though most are stated in (technical) English, *all* are to be thought of as cast in the language of quantifiers and truth functions. Refer to an axiom by name or number.

- DEFINITION. A definition is *also* a clear-cut, formalizable postulate that is to be *used* in proving something. It should also be referred to by name or number. The difference between axioms and definitions is this: We do not call something a definition unless we have reason to believe that, relative to axioms, the definition satisfies the usual criteria of eliminability and non-creativity, which are explained roughly in NAL:**7A-4** and more carefully in NAL:§12A. Most definitions not only permit brevity; they are also revelatory of the structure of the topic under discussion (they show without saying). But from *your* point of view, axioms and definitions are to be treated alike as premisses from which to prove something.

- VARIANT. A variant is a definition that is so trivial that it doesn't even need reference.

- THEOREM. Theorems are statements that follow from axioms and definitions (sometimes via conventions and variants). Unless we explicitly mark the proof of the theorem otherwise, "follow" means: Follows by the elementary quantifier techniques you already know from axioms and definitions explicitly declared in these notes. See **1A-3** for conventions concerning their proofs.

- FACT. A fact is a theorem that isn't so important.

- MINIFACT. A minifact is a minifact.

- PROPOSITION. A proposition is the same as a fact.

- COROLLARY. A corollary is a theorem that can be proved in an easy way from what closely precedes it; for this reason it usually bears no indication of proof.

- LEMMA. A lemma is a theorem that is of little interest on its own account, but that is useful in proving something else.

- LOCAL DEFINITION. These hold only for a while; they are most frequently used to give temporary meaning to a single letter.

- LOCAL CHOICE. Rather like Local definitions, except they are justified by existential statements not usually conveying uniqueness. (Reference to the "axiom of choice" is not intended.)

- LOCAL FACT. A fact that depends on a local definition or choice.

- CONJECTURE. In these notes, all conjectures are true, but they receive this heading because they are stated before we are ready to prove them; of course they cannot be cited in proofs.

1A-3 CONVENTION. *(Convention on proofs)*

We will mark each proof in one of the following ways.

- PROOF. *Trivial.* This one you can do in your head.
- PROOF. *Exercise.* This indicates that every student of this material should be able to reconstruct this proof as a rigorous proof, without armwaving, given such references and indications as are provided, using only axioms, definitions, etc. as are explicitly present in these notes; *and* that it is a worthwhile thing to do so. Usually in such a case the indication "PROOF" is omitted in favor of an explicitly indicated numbered exercise.
- PROOF. *Straightforward.* The difficulty is comparable with "*Exercise*," but it is not so important to spend the time figuring the matter out. Many students will nevertheless wish to take the trouble to do so.
- PROOF. *Tedious.* This indicates that a first order proof is available from previous numbered items, but that the argument might (though consisting of small steps) be long. Only a few students should address these.
- PROOF. Some proofs are just given. These are in general too difficult for most students to find easily, but *every* student should be able to recast these proofs into a rigorous first order argument. (We count it *our* fault if a student who did excellent work in a sound one-term "art of logic" course cannot carry out this task—such is as good a statement as there is of the aspiration of these notes.)
- PROOF. *Omitted.* This indicates that not enough groundwork has been laid for a lucid, rigorous first order proof. No one is responsible for reconstructing these rigorously, though informal indications are sometimes given. (You may prefer to regard these items as so many additional axioms.)

1A-4 CONVENTION. *(Use-language notation)*

We use "if_then_" in the technical passages of these notes in a truth functional sense, equivalent to "not both _ and not _." We also use "_→_" as a connective in our use-language for the *truth functional* if-then connective—sometimes called the "material conditional" or "material implication." Thus, with "S" and "P" in the place of use-language sentences, S→P if and only if (the truth functional sense of) either not S or P.[3]

Similarly, in technical passages "_if and only if_" is truth functional, and we also use "_↔_" for the so-called "material biconditional" or "material equivalence" of our use-language, with the understanding that the output is true just in case the two inputs are alike in truth value.

We do not happen to use any special notation for conjunction or disjunction; but we do intend that "_and_" and "_or_" shall have their standard truth functional readings.

Negation we express by means of standard English or mathematical locutions rather than by means of a symbolic connective.

Occasionally we use "(x)" or one of its cousins, say "∀x," to help express a universal quantification in our use-language, and "∃x" to help express existential quantification. More often, especially as we go along, we use standard "middle English" expressions such as "for all x" and "for some x" (see NAL:**1B-3**).

We use the familiar "=" for identity.

1A-5 CONVENTION. *(Omitted universal quantifiers)*

As is usual, NAL:**9A-2**, we will often drop *outermost* universal quantifiers from what we assert. So if we use a variable without binding it, you are to supply a universal quantifier whose scope is our entire axiom, definition, theorem, etc. (In *proofs* we use variables in another standard way, as "parameters" or "temporary constants"; the present convention does not of course apply to such uses.)

If you do not understand the foregoing two conventions, you are not ready for these notes.

[3]This convention is necessary because we do not believe that "if_then_" in normal English is truth functional. Nor does the locution permit itself to be governed by some *other* rigorous theory. "If_then_" escapes every effort to put it in a formal straighjacket. It is only by means of this convention that we can successfully appeal to the standard canons of truth-functional logic.

Chapter 2

The logic of truth functional connectives

This chapter considers a language analyzed in terms of its truth functional connectives. By §1A.2 you know that the elementary portion of our deliberations should fall into four parts: grammar, proof theory, semantics, and applications—but that we will not be concerned with applications. Before commencing our linguistic work proper, however, we first take up a language-independent topic: truth values and truth functions. We treat the theory of truth values and their functions because, although marvelously simple, it forms the foundation for the *particular* semantic theory we shall be offering. Our treatment of this theory (as well as much of the theory developed further on in this book) presupposes a modicum of understanding of the *general* theory of relations and functions such as is explained in NAL:§9B; we place particular reliance on Cartesian powers and the function-space operator.

We then turn to *grammar*. In this section you should expect an account of the *structure* of the language under consideration in so far as that structure is relevant for semantics and proof theory. The ideas about which we shall be theorizing under the heading of "grammar" will be these: some idea of a sentence, some idea of ways of making new sentences out of old (that will be interpreted truth functionally), and some idea of atomic or uncompounded parts.

Next we shall deal with *semantics*. We shall consider the meaning of the grammatically atomic parts, and how the meaning of the compound sentences derives from these by way of the meaning of the modes of combination permitted by the grammar (in this case, the truth-functional connectives). *Semantics always follows grammar*; so, since our grammar in this chapter is relatively poor, you should expect a matching impoverishment of the notion of meaning we develop. We shall, however, be able to develop a notion of "implication" or "logical consequence" relativized to the grammar under consideration. This notion is the classical *semantic* rendering of the concept of "good argument" in terms of meaning.

After dealing with elementary semantics, we shall be turning to the development of *proof theory*. A negative characterization of proof theory is that in contrast to semantics, it does not rely on meaning. This feature of proof theory is prized by those who think of meaning as a Bad Thing. A positive characterization is this: The fundamental concepts of proof theory are given inductively. That is, first we are given some paradigm cases of "good argument," cases we can immediately recognize, and then we are told how to obtain other cases of good argument out of these in certain standard ways (for example, by using modus ponens). The upshot is a *proof-theoretic* account of "good argument."

Since we develop our semantic notion of "good argument" independently of our proof theoretic notion, we have no obvious guarantee that the two coincide. The principle nonelementary task of this chapter will be to establish this connection: There is perfect agreement between our fundamental semantic account of "good argument" in terms of meanings on the one hand, and our proof theoretic account given inductively on the other. If you think of the semantic concept in terms of meaning as giving a more profound and less ad hoc but (alas) "invisible" account of good argument, and if you think of the proof theoretic concept as giving us a perhaps unmotivated or ad hoc but at least "visible" and hence humanly applicable criterion for good argument, you will appreciate the importance of bringing the twain together. And that is good, because the enterprise is not trivial.

2A Truth values and functions

This section is about truth values and truth functions. We are going to study them because the simple sort of semantic theory we are going to be developing runs most smoothly when we think of sentences as (metaphorically) pointing to or denoting truth values, and when we think of connectives as (metaphorically) expressing truth functions. At this point, however, we are going to study the structure of truth values and functions in isolation from the semantic use to which we shall be putting them.

You are entitled to wonder why we proceed in this way, and the answer is: Because we must. A rigorous semantics of anything like the kind we are eventually going to articulate *always* requires a prior theory about the "subject matter." If you want to describe some linguistic entity as denoting something or other, then if you want your theory about the denoting to be rigorous, you better first have a rigorous theory about the "something or other." Our present case is merely the most simple-minded of illustrations of this philosophical truth.

2A.1 Truth values

Are there any truth values? It isn't important to decide the question when it is taken as a "deep" one; it suffices to observe that it is *useful* to postulate truth values. We call this "practical Platonism." The reason the existence of truth values is not a deep question is this: It is *obvious* how to do without them. Therefore, why not use them with an easy conscience? (How to do without them is largely indicated by the recipe: To say that the-truth-value-of A is T is to say that A is true.)

The theory has just two primitives, "T" and "F," both terms. Of course we are thinking of T as The True, and of F as The False; but such thinkings are not part of the theory.

These axioms and definitions and propositions may be referred to generally as "TF," or individually by name or number.

2A-1 DEFINITION. *(**2**)*

It is convenient to let **2** be the set of truth values:

$$\mathbf{2} = \{\mathrm{T}, \mathrm{F}\};\ \text{that is,}$$
$$\mathbf{x} \in \mathbf{2} \leftrightarrow (\mathbf{x} = \mathrm{T}\ \text{or}\ \mathbf{x} = \mathrm{F}) \qquad (\mathbf{2})$$

2A-2 CONVENTION. *(**x** and **y** over **2**)*

We use boldface "**x**" and "**y**" as variables ranging over **2**.[1] That is, in proofs you are entitled to write "$\mathbf{x} \in \mathbf{2}$" for free.

Exercise 1	*(How big is **2**?)*

How many members are there in **2**?

▷ . . . . ◁

[1] On the blackboard we write these as editors do, with a curly underline.

2A-3 VARIANT. *(Truth value)*

$\mathbf{x}$ is a truth value $\leftrightarrow$ $\mathbf{x} \in \mathbf{2}$; that is, $\leftrightarrow$ $\mathbf{x} = \mathrm{T}$ or $\mathbf{x} = \mathrm{F}$.

The entire theory of truth values hangs from a single ingenious

2A-4 AXIOM. *(*T$\neq$F*)*

(T$\neq$F) (T$\neq$F)

2A-5 COROLLARY. *(*TF *flip-flop)*

For $\mathbf{x} \in \mathbf{2}$,

$\mathbf{x} \neq \mathrm{T} \leftrightarrow \mathbf{x} = \mathrm{F}$.

$\mathbf{x} \neq \mathrm{F} \leftrightarrow \mathbf{x} = \mathrm{T}$.

PROOF. See Exercise 2 just below. ☐

2A-6 COROLLARY. *(Two truth values)*

There are exactly two truth values.

In fact, that is all we know or need to know about the truth values; in particular, there is nothing in the theory of truth values telling us which value wears the White Hat and which the Black. (Of course *we* confer Good Guy status on one of them by calling it "T"; but the theory gives it no intrinsic property on the basis of which it deserves that status.)

Exercise 2 *(Two truth values)*

This exercise is intended to increase your sensitivity to the difference between an intuitive argument and a rigorous proof.

1. *State* Corollary **2A-6** symbolically.

2. Carry out its proof.
3. And reconsider your answer to the question asked by Exercise 1, noting that it was asked *before* the laying down of Axiom **2A-4**.

▷ . . . . ◁

2A.2 Truth functions

"Truth functions" are what we should next like to speak of; with this in mind it would be good to review the appropriate section on relations and functions in NAL.

A truth function is not a piece of language (for example, a truth function is not a "truth-functional connective"); instead, a truth function is a function whose arguments and values are both limited to truth values:

2A-7 DEFINITION. *(Truth function)*

---

f is an n-place *truth function* just in case $f \in (Z \mapsto \mathbf{2})$, where Z is some Cartesian power of **2**, that is, $Z = (\mathbf{2} \times \ldots \times \mathbf{2})$ [n times]. See NAL:**9B-7**.

Recall that our Definition NAL:**9B-7** of "Cartesian power" is partly casual; the same accordingly holds for our definition of a truth function, which we offer just to put our discussion into perspective. For example, it is clear that the following definition yields a truth function, the very truth function corresponding to the classical two valued conditional.

2A-8 DEFINITION. *($\supset^*$)*

$\supset^*$ is that function in $(\mathbf{2}\times\mathbf{2})\mapsto\mathbf{2}$ ($\supset^*$type)

such that for all $\mathbf{x}, \mathbf{y} \in \mathbf{2}$,

$(\mathbf{x} \supset^* \mathbf{y}) = \mathrm{T} \leftrightarrow (\mathbf{x} = \mathrm{T} \rightarrow \mathbf{y} = \mathrm{T}) \leftrightarrow (\mathbf{x} = \mathrm{F}$ or $\mathbf{y} = \mathrm{T})$. ($\supset^*$)

Hence, $(\mathrm{T} \supset^* \mathrm{T}) = (\mathrm{F} \supset^* \mathrm{T}) = (\mathrm{F} \supset^* \mathrm{F}) = \mathrm{T}$; and $(\mathrm{T} \supset^* \mathrm{F}) = \mathrm{F}$. ($\supset^*$)

That is, explicitly, this one time only:

$$\supset^* = \{\langle\langle T, T\rangle, T\rangle, \langle\langle T, F\rangle, F\rangle, \langle\langle F, T\rangle, T\rangle, \langle\langle F, F\rangle, T\rangle\}.$$

The following subproof principles are useful:

2A-9 COROLLARY. *($\supset^*$MP and $\supset^*$CP)*

Provided **x** and **y** are truth values, the following procedures are acceptable:

$$\begin{array}{r|ll} 1 & (\mathbf{x}\supset^*\mathbf{y})=T & \\ 2 & \mathbf{x}=T & \\ & \mathbf{y}=T & 1, 2, \supset^*\text{mp} \end{array}$$

$$\begin{array}{r|l|ll} 1 & & \underline{\mathbf{x}=T} & \text{hyp} \\ \cdot & & \cdot & \\ \cdot & & \cdot & \\ \cdot & & \cdot & \\ k & & \mathbf{y}=T & \\ & \multicolumn{2}{l}{(\mathbf{x}\supset^*\mathbf{y})=T} & 1\text{–}k, \supset^*\text{CP} \end{array}$$

Don't confuse these rules with MP and CP in our use-language.

PROOF. Straightforward, using the first form of ($\supset^*$), **2A-8**: $\supset^*$MP and $\supset^*$CP emerge by MP and CP for our use-language "→" or "if-then." ☐

Exercise 3 *(Exercise on the "type" of $\supset^*$)*

Show that $\mathbf{x}\supset^*\mathbf{y}\in\mathbf{2}$. Use Convention **2A-2** and NAL:**9B-15** with $\supset^*$type of **2A-8**. (An alternative proof could use the very definition of $\supset^*$ instead of $\supset^*$type, but it would not be as illuminating.)

▷ . . . . ◁

The truth function corresponding to negation is given by the following

2A-10 DEFINITION. *($\sim^*$)*

$\sim^*$ is that function in $\mathbf{2}\mapsto\mathbf{2}$ ($\sim^*$type)

such that $\sim^*T=F$ and $\sim^*F=T$. ($\sim^*$)

2A-11 COROLLARY. *($\sim^*$)*

For $\mathbf{x} \in \mathbf{2}$,

$$\sim^* \mathbf{x} = \mathrm{T} \leftrightarrow \mathbf{x} = \mathrm{F} \leftrightarrow \mathbf{x} \neq \mathrm{T} \qquad (\sim^*)$$
$$\sim^* \mathbf{x} = \mathrm{F} \leftrightarrow \mathbf{x} = \mathrm{T} \leftrightarrow \mathbf{x} \neq \mathrm{F} \qquad (\sim^*)$$

Hence, for $\mathbf{x}$, $\mathbf{y} \in \mathbf{2}$,

$$\sim^* \mathbf{x} = \mathbf{y} \leftrightarrow \mathbf{x} \neq \mathbf{y} \qquad (\sim^*)$$
$$\sim^* \mathbf{x} \neq \mathbf{y} \leftrightarrow \mathbf{x} = \mathbf{y} \qquad (\sim^*)$$

Exercise 4 *(Other truth functions)*

We may want other usual connective symbols marked in similar ways to stand for corresponding truth functions. We have in mind especially

$\&^*$ corresponding to “and”
$\vee^*$ corresponding to “or”
$\equiv^*$ corresponding to “if and only if”

The exercise is to give these rigorous definitions, *and* supply subproof rules analogous to **2A-9** for each of them .[2]

▷ . . . . ◁

The following turns out to play an important role in later proceedings.

2A-12 THEOREM. *(Three ways to name* T*)*

For every $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ in $\mathbf{2}$,

[2]Note that in our use-language we use specially marked symbols as names of truth functions. They do not name connectives of our use-language, and they do not name symbols of any kind of the language TF described below, §2B.2.

$$\mathbf{x}\supset^*(\mathbf{y}\supset^*\mathbf{x})=\mathrm{T} \tag{T1}$$

$$(\mathbf{x}\supset^*(\mathbf{y}\supset^*\mathbf{z}))\supset^*((\mathbf{x}\supset^*\mathbf{y})\supset^*(\mathbf{x}\supset^*\mathbf{z}))=\mathrm{T} \tag{T2}$$

$$(\sim^*\mathbf{x}\supset^*\sim^*\mathbf{y})\supset^*(\mathbf{y}\supset^*\mathbf{x})=\mathrm{T} \tag{T3}$$

Note that Theorem **2A-12** does not refer to any language (though of course it cannot be stated without using language).

PROOF. We will prove T1, and leave T2–3 as an exercise. The proof looks best as a subproof.

1	$\quad\vert\ \ \vert\ \underline{\mathbf{x}}=\mathrm{T}$	hyp
2	$\quad\vert\ \ \vert\ \ \vert\ \underline{\mathbf{y}}=\mathrm{T}$	hyp
3	$\quad\vert\ \ \vert\ \ \vert\ \mathbf{x}=\mathrm{T}$	1, reiteration
4	$\quad\vert\ \ \vert\ \mathbf{y}\supset^*\mathbf{x}=\mathrm{T}$	2–3, $\supset^*$ CP
5	$\quad\vert\ \mathbf{x}\supset^*(\mathbf{y}\supset^*\mathbf{x})=\mathrm{T}$	1–4, $\supset^*$ CP

□

Exercise 5 *(Three ways to name* T*)*

Prove T2–3 of **2A-12**. Use **2A-9** to save some "casework."

▷ . . . . ◁

2A.3 Truth functionality of connectives

It is worthwhile to relate the theory of truth functions to the usual concept of truth functionality so as to give intuitive point to the theory; but this section is (a) not rigorous and (b) is not used in what follows.

As background, we need to be sure we understand "truth value of A" and "connective." Given a sentence A of an understood language, *the truth value of* A is generally defined as T if A is true, and as F if A is false. If brevity is required, we can write

$$Val(\mathrm{A})$$

for "the truth value of A." A more notational definition would then be

$$Val(\text{A}) = \mathbf{x} \leftrightarrow [(\text{A is true} \rightarrow \mathbf{x} = \text{T}) \text{ and } (\text{A is false} \rightarrow \mathbf{x} = \text{F})]$$

This explanation of the meaning of Val is acceptable (conservative, NAL:**7A-4**) only when each sentence A is either true or false, but not both.

The best thing to mean by a *connective* is: a mapping from sentences into sentences (NAL:§1A.2). Often the mapping can be visualized as the placing of the input sentences into blanks in some larger context, but that is not here part of the definition.

As a local convenience, we will let Sent be the set of sentences, and where *Conn* is a one-place connective and A is a sentence, we write

Conn(A)

as the result of applying *Conn* to A.

Conn(A) is clearly a sentence. Why? Well, to say that *Conn* is a one-place mapping from sentences into sentences is to say that *Conn* belongs to (Sent$\mapsto$Sent), and since A belongs to Sent, so does *Conn*(A), by MP for function space, NAL:**9B-14**.

Now for the standard definition of truth functionality: A (one-place) connective *Conn* is *truth functional* just in case, for every sentence A and B,

$$Val(\text{A}) = Val(\text{B}) \rightarrow Val(Conn(\text{A})) = Val(Conn(\text{B})).$$

A similar definition is in effect for n-place connectives generally. In words: If the truth values of the inputs are identical, then so are the truth values of the outputs.

We can now state a fact that relates the truth-functionality of connectives to the concept of a truth function.

2A-13 FACT. *(Existence of unique truth function)*

If *Conn* is a one-place truth functional connective, and if (!) there are both true and false sentences, then there is a unique one-place truth function f such that, for every sentence A,

Val(*Conn*(A)) = f(Val(A)).

Similarly for n-place truth functional connectives. Under these circumstances, f is sometimes thought of as the "meaning" of *Conn*.

Exercise 6 *(Exercise on some types)*

1. State the "type" of Val, *Conn*, and f in Fact **2A-13** by stating to which function space each belongs.

2. You should try to "run an example" to convince yourself of Fact **2A-13**. Try it with about five sentences. (1) Define any one-place connective *Conn* you like by a five-entry table that says what the output *Conn*(A) is for each input A, (2) mark any sentence true and another one false, (3) mark the rest of the sentences true or false, being sure that if you mark any two inputs to *Conn* alike, the outputs are also marked alike, and (4) feel compelled to describe the only possible corresponding truth function f by a two-entry table that says into what f maps T, and into what it maps F. Finally, verify for *each* of your five sentences A the equation stating that Val(*Conn*(A)) = f(Val(A)).

3. Consider the "!" of Fact **2A-13**. Show that the marked "if" clause is needed. That is, show that if all sentences are true, then there is more than one one-place truth function f such that Val(*Conn*(A)) = f(Val(A)). Here an intuitive description of the two one-place truth functions and a suitable plausibility-argument will suffice. (There is no need to show that an analogous problem emerges if all sentences are false, but feel free to do so if you wish.)

▷ . . . . ◁

2A-14 REMARK. *(Exercise on some types)*

So far we have in our use-language an "if-then" connective involving the use of "$\rightarrow$," and a name for a function on truth values that is written "$\supset^*$." We also use "$\mapsto$" for the function space operator on sets. When we introduce the language TF, things are going to get worse.

Exercise 7 *(Truth values and functions)*

With respect to the following, observe that apart from the first, these exercises are not concerned with any language.

1. Why is "if_then_" a connective? Why is "_ $\supset^*$ _" an operator?
2. Prove that for all **x** in **2**, ($\sim^*$**x**$\neq$**x**).
3. Prove that for all **x** in **2**, (**x**$\supset^*$**x**)$=$T.
4. Prove that there is an **x** in **2** such that (**x**$\supset^*\sim^*$x)$=$F. (You may find this confusing because it is so easy; after all, what you are to prove is an *existential* statement—not a universal as in the previous exercises—so that all you need do is provide an instance.)
5. Exhibit or describe a function in **2**$\mapsto$**2** other than $\sim^*$.

▷ . . . . ◁

2B Grammar of truth functional logic

The elementary study of any branch of logic divides, as we have said, into four parts: (1) grammar, (2) semantics, (3) proof theory, and (4) applications. Here we take up the grammar of truth functional logic, aiming to outline the grammatical structure of any language endowed with some truth functional connectives.

2B.1 Grammatical choices for truth functional logic

There are not many choices to be made, but there are a few.

Which connectives? First, which connectives should we deal with? Certainly (1) we shall wish to choose a family of connectives that are "adequate" in the sense of being able to express all truth functions (see §2D.5); and (2) with an eye on applications it seems best to choose from among the six truth functional connectives figuring most prominently in (the logicians' picture of) our use-language: falsehood, negation, the conditional, the biconditional, conjunction, and disjunction. It would complicate the technical side of our task, however, to choose *all* of these; (3) we want instead to choose as few connectives as possible. The pair, negation and the conditional, satisfy these three conditions (but not uniquely), and we so choose: We shall study a language with just these connectives.

Grammatical style. Second there is a choice of the style in which we discuss grammar. The truth functional language we are about to study we call "TF"; imagine it spoken on an island far away, so that you are not tempted to confuse it with our own language, i.e., with our use-language. (I will sometimes call its speakers "the islanders.") The question of style is this: Should we tell you what TF looks like? We decide in the negative; we decide to explain as much of the abstract structure of TF as is relevant for the science of logic without in any way giving you a hint as to the appearance of TF. We neither *describe* the appearance of TF (by saying e.g. that one of its sentences is the sixteenth letter of the English alphabet, or that another is formed by intersecting two lines of such and such a length at such and such an angle), nor do we *display* any expression of TF (so that you would be shown rather than told what it looks like).

There is a reason for this. If we were going to be interested much in applications, then we would need to know how the expressions of TF look; we would need to interest ourselves in their communicative possibilities as related to our sensory apparatus, or perhaps as related to the sensory apparatus of the islanders. If for example we wished you to be able to write the expressions of TF, or to know how the islanders manage to do so, we would need to bring you to appreciate the visual features of its expressions. Since, however, we are not interested in applications, we need only introduce so much grammar as required for semantics and proof theory. For these purposes it does not matter what TF looks like, or indeed if it looks like anything at all; all that matters is the structure of its expressions, so far as that structure is relied upon by semantics and proof theory.

For example, semantics requires that there be some "sentences" to which to assign truth values, but it does not care what these sentences are. And proof theory requires that we be able uniquely to recover the antecedent of a conditional (for modus ponens, say), but it does not need to know just how this is done.

So TF is not only "uninterpreted," but even beyond that (we promise to use this word just twice): "unconcretized." For example, you will never be told that a certain "variable" is made by constructing two short lines of the same length which meet at about a ninety degree angle, the lines being at respective angles of 45 and 315 degrees to the vertical.

This choice dictates a subtle but important change in our grammatical jargon when we pass from the art to the science of logic. In NAL:**1A-2** we defined an "elementary functor" as a pattern of words with blanks such that when its blanks are filled (input) with sentences or terms, the result (output) is either a term or a sentence. This gives us an idea of functor that is easy to picture and easy to work with in practice, and is sufficient for the art of logic. It is, however, insufficiently general and insufficiently abstract. We cannot use it, for instance, in application to a language whose physical features are entirely unknown. For this reason, whenever we discuss either elementary functors in general or the special cases constituted by connectives, operators, and predicates, you should expect us to consider *arbitrary* grammatical functions taking terms or sentences as inputs, and yielding terms or sentences as output, with no implication as to the "physical" realization or these grammatical functions.

Primitive vs. defined ideas. There is a third choice. We are in any event going to need to work with four ideas: (1) sentences, which we represent via the set-primitive, Sent, the (2) conditional and (3) negation connectives, and (4) some "atoms"—let us call them "TF-atoms,"and let the set of them be TF-atom—from which all sentences can be constructed via the connectives. It turns out that we could choose either one of "TF-atom" or "Sent" as primitive, and the other as defined; but instead we choose to take both as primitive, because it seems to make the exposition flow more smoothly and to enable more straightforward articulation of the essential conceptual structures.

One language or many? One last choice. The science of logic must inevitably deal with many languages, and hence with many different notions of "sentence"; for this reason, we might have chosen to take "sentence of TF" or "TF-Sent" instead of just "Sent" as our primitive. In these notes, however, since our purposes are elementary, we can confine ourselves to just a single notion of "sentence" or "Sent" so that the longer phrase is unnecessary. The picture to have is this. We find the distant islanders speaking "sentences." In this chapter on truth functional logic we content ourselves with treating only the truth functional structure of these sentences, even though we know full well that there is more structure there to be found. In the next chapter we add a treatment of the quantificational structure of these *very same* sentences.

2B.2 Basic grammar of TF

We describe the grammar of TF by laying down some axioms. The intuitive idea is that each of the "sentences" is built from truth functionally simple sentences called "TF-atoms" by constructions corresponding to the conditional and negation.

Our use-language has, governing the grammar of TF, the following primitives: "Sent," "TF-atom," "$\supset$," "$\sim$." The types of these expressions are given in the following

2B-1 AXIOM. *(Primitives of* TF*)*

Sent is a set; its members are called *sentences*. (Sent type)

TF-atom is a set, and TF-atom $\subseteq$ Sent (TF-atom type)

$\supset$ is a two-place function on Sent:

$$\supset \in ((\text{Sent} \times \text{Sent}) \mapsto \text{Sent}) \qquad (\supset\text{type})$$

$\sim$ is a one-place function on Sent:

$$\sim \in (\text{Sent} \mapsto \text{Sent}) \qquad (\sim\text{type})$$

2B-1 is part of an "inductive definition" of Sent. The clause "TF-atom type" is a "base clause," which gives you some sentences to begin with. "$\supset$type" and "$\sim$type" are "inductive clauses," which tell you how to get new sentences out of old ones. The "inductive definition" is only completed later, with Axiom **2B-17**.

We have now a number of usages in the vicinity of negation.

- In these notes, for negation we use (in our use-language) just "not," or a stroke in the standard negating way as in "$\neq$"; see **1A-4**. On the blackboard, we use in addition "$-$" or "$\sim$" or "$\neg$" for negation in our use-language—but never in these notes.
- We use "_ $-$ _" as a set-operator in our use-language, signifying relative complementation or set difference:**9A-15**.
- We use "$\sim$" as a name in our use-language denoting a connective of TF—so that $\sim$ (but of course not "$\sim$") is a connective of TF (see **2B-1**).
- We use "$\sim^*$" as a name of a truth function, **2A-10**

All this is troublesome if you think about it; but it doesn't cause much difficulty in practice. Something similar happens for the conditional: "$\rightarrow$" helps with our conditional use-connective, there is no useful corresponding set-operator, "$\supset$" names the connective of TF, "$\supset^*$" names the truth function, **2A-8**, and we use "$\mapsto$" for the function space operator on sets, NAL:**9B-14**.

2B-2 CONVENTION. *(Set terms as common nouns)*

We use set *terms* (names of sets) as *common nouns* in middle English constructions (see NAL:Remark **1B-3**), a convention that includes permission to pluralize. These are examples of the general policy:

- A is a TF-atom ↔ A∈TF-atom
- "Every TF-atom" or "all TF-atoms" for "every member of TF-atom"
- "Some TF-atom" for "some member of TF-atom"

Also we sometimes pair a distinct English common noun with a set term, as "sentence" is paired with "Sent" in the clause Sent-type of **2B-1**, so that "every sentence" means "every member of Sent."

2B-3 CONVENTION. *("p" and "q" for* TF-*atoms)*

We use "p"and "q" as variables ranging over TF-atom.

That is, "For all p" means "For all p, if p∈TF-atom then"; and "For some p" means "For some p, p∈TF-atom and." In proofs, the convention means you can always add "p∈TF-atom" for free. We used this same sort of convention above for set letters "X," "Y," etc., NAL:**9A-3**, and we will use the same idea again:

2B-4 CONVENTION. *("A," "B," "C," "D," "E" for sentences)*

We use "A," "B," "C," "D," and "E" as variables ranging over sentences.

Many writers say "formula" instead of "sentence."[3] Later, in chapter 3, we introduce "formula" to encompass both sentences and terms, but of course TF has only sentences.

More grammatical jargon to ease our talk about TF:

2B-5 DEFINITION. *(Conditional and negation)*

A is a *conditional* ↔ A=(B⊃C), some sentences B and C.

[3]Some writers call sentences or formulas "well formed formulas," and then having invoked such a long phrase for such a short idea, introduce "wff" instead for frequent use. But (1) the abbreviation is Not Attractive, and (2) in our context there is in any event no work done by the adjectival phrase "well formed," for we shall never at any point be considering some larger class of "formulas" of which only some are "well formed." In a nutshell: The science of logic does not require "ill formed" formulas, hence it also does not require "well formed" formulas; so it is pointless to pay the price of Talking Ugly by using "wff." (It is recommended that "wff" or its homonym "WFF" be reserved for exclusive use by the justly famous White Flower Farm of Litchfield, CT.)

A is a *negation* $\leftrightarrow$ A = ∼B, some sentence B.

Therefore, (B)(C)[B ⊃ C is a conditional, and ∼B is a negation].

It is useful to add some standard definitions of other connectives (compare Exercise 4).[4]

2B-6 DEFINITION. *(Other connectives)*

A∨B = (∼A ⊃ B)

A&B = ∼(A ⊃ ∼B)

A≡B = (A ⊃ B)&(B⊃A)

⊥(A) = ∼(A⊃A)

NAL introduced ⊥ as a zero-place connective, so that ⊥ can itself be taken as a sentence. Here, for strictly technical reasons, we invoke ⊥ as a one-place connective. No matter, whether you get to ⊥ or ⊥(A), you are in big trouble.

Exercise 8 *(Other connectives)*

Given that *Dom*(∨) = (Sent × Sent), show that

$\vee \in ((\text{Sent} \times \text{Sent}) \mapsto \text{Sent})$.

Be sure to use and refer to definitions of all defined terms; and also be explicit, in this exercise, concerning any use of conventions such as **2B-4**.

▷ . . . . ◁

2B.3 Set theory: finite and infinite

Passing on now to a new topic, we want eventually to be able to say that each sentence is "finitely long"; but sentences do not have length (as far as we know). We

[4] As in all other cases, we are defining symbols of our use-language; TF itself is unchanged.

could assign a metaphorical length to each sentence if we wished, but more important for our purposes is to observe that each sentence has finitely many "parts," i.e., "subformulas." To say this, we shall have to understand "finite" and "subformula"; and since "finite" is a purely set-theoretical notion, we leave grammar for long enough to develop it.

The concept of the finite is important for logic, partly because of the widespread conviction that the human mind can in some sense deal with only finitely many items at once. (We do not endorse this view, which appears to rest on an "identity" (or "similarity") account of intentionality: To know something is to be (or be like) that thing.) This conviction appears most evidently in the nearly unspoken presupposition that each inference to a conclusion must depend on only finitely many premisses (consider $\vdash_{S_{TF}}$fin, **2E-9**).

In any event, we want to give what amounts to a definition of "finite set" for our use. We define the concept not in the shortest way (due to Dedekind[5]), but in a manner which is (a) intuitive and (b) useful for our immediate purposes. What we do is to tell you that the finite sets are those obtainable from the empty set by adding one member at a time. The definition is in effect inductive: The Basis clause sets us under way, the Inductive clause tells us how to continue, and the Closure clause, as always, tells us that "that's all," *and* provides us with a form of inductive proof as in the corollary just below.

2B-7 DEFINITION. *(Finite)*

Basis clause. $\varnothing$ is finite. (Fin$\varnothing$)

Inductive clause. If X is finite, so is $(X\cup\{y\})$. (Fin+1)

Closure clause. "That's all the finite sets: There are no others." Which is to say: Let $\Psi(X)$ be a use-language context, **2B-8**. **Suppose** the following,

> **Basis step.** $\Psi(\varnothing)$.
>
> **Inductive step.** $(X_1)(y)[\Psi(X_1) \rightarrow \Psi(X_1\cup\{y\})]$.
>
> **Then** for all X, X is finite $\rightarrow \Psi(X)$. Ind. on Fin. Sets

In the above definition we have relied on the following

[5] A set X is Dedekind-finite $\leftrightarrow$ there is no one-one function (**9B-13**) whose domain is X and whose range is a proper subset of X. You can use Induction on finite sets, **2B-9**, to show that all finite sets are Dedekind-finite. What about the converse?

2B-8 CONVENTION. *(Ψ for use-contexts)*

We use "Ψ(__)" with a use-language variable in the blank as both denoting and as taking the place of a use-language sentential context involving that variable. Furthermore, any nearby occurrence of "Ψ(__)" with some use-language term instead of the variable in the blank is to name or take the place of a use-language sentence derived from the first by putting the term for all (free) occurrences of the variable.[6]

It is part of the convention that e.g. $\Psi(X)$ may contain a free variable; in this case, that variable shall have the widest sensible scope.

2B-9 COROLLARY. *(Induction on finite sets)*

The following subproof rule is admissible:

.		
.		
.		
$\Psi(\varnothing)$	Conclusion of Basis	$[\varnothing/X]$
| $\Psi(X_1)$	Inductive hyp, flag X_1, y	$[X_1/X]$
| .		
| .		
| .		
| $\Psi(X_1 \cup \{y\})$	Conclusion of ind. step	$[(X_1 \cup \{y\})/X]$
$(X)(X \text{ is finite} \rightarrow \Psi(X))$		Ind. on fin. sets

Examine this rule carefully; it is an excellent paradigm of what a "closure clause" of an inductive definition such as **2B-7** really comes to. Note on the technical side

---

[6]The heaviness of the convention is an example of how troublesome it is to speak with generality about our own use of our own use-language without becoming enmeshed in use-mention difficulties. You will note how seldom the necessity arises. In the meantime, let us explain the phrase "both denoting and taking the place of." When we say "Let $\Psi(X)$ be any use-context," our use-expression "$\Psi(X)$" occupies the place of an English singular term, and hence by the logician's version of English grammar should be taken as denoting something. But when we later say, for instance, "$\Psi(X_1) \rightarrow \Psi(X_1 \cup \{y\})$," our use-expressions are occupying the places of English sentences and hence must be construed as being rather than denoting sentential contexts.

Some logicians think such modes of speech are reprehensible; but we think that they are wrong. What *is* true is that it is difficult (but not impossible) to codify such practices formally; but it is equally true that you will understand every word we say by means of this convention; which, since its purpose is communication, is enough. (Oddly enough, most of these same logicians are happy with "For every x, x is either swift or sorry," even though the first "x" takes the place of a common noun and the second the place of a singular term.)

that it involves three distinct substitutions for "X." Isolating the use-context represented by Ψ(X) and making these or similar substitutions is the key to formulating cogent inductive arguments. Be aware that the eye can be easily confused when Ψ(X) is complex, and do not be slow to make notes to yourself as to the various substitutions involved. In that spirit, we have annotated the three key substitutions for X. You should certainly do likewise.

Here is an abstract example.

2B-10 EXAMPLE. *(Induction on finite sets)*

Suppose you wish to use Induction on finite sets, **2B-9**, to prove something having the form

> For every X, if it is both a G and a finite set then it is an H.

Symbolization makes your work easier:

> (X)[(X is G and X is finite) → X is H]

This form does *not* match the conclusion of **2B-9**; but it is close. First transform it (by quantifier-and-truth-function principles) so that the match is exact:

> (X)[X is finite → (X is G → X is H)].

This is an exact match to the conclusion of **2B-9**, picturing Ψ(X) there as "X is G → X is H" here. Stop to verify this match. Now that Ψ(X) is identified as "X is G → X is H," the form of the required proof is *uniquely* determined:

	·		
	·		
	·		
j	∅ is G → ∅ is H	Conclusion of basis	[∅/X]
j+1	| X_1 is G → X_1 is H	Ind. hyp., flag y, X_1	[X_1/X]
·	| ·		
·	| ·		
·	| ·		
k	| (X_1 ∪ {y}) is G → (X_1 ∪ {y}) is H	Concl. of ind. step	[(X_1 ∪ {y})/X]
k+1	(X)(X is finite → (X is G → X is H))	Ind. on finite sets	

Please study the substitution of $\varnothing$ for X (for the conclusion of the basis), of X_1 for X (for the inductive hypothesis), and of $(X_1 \cup \{y\})$ for X (for the conclusion of the inductive step). These substitutions must be exact in order for you to describe yourself as using Induction on finite sets, **2B-9**.

There is always some "logic" going on inside an inductive proof. In this particular example, at least the following possibilities are easy to see. (1) Step j very likely comes by conditional proof. (2) Step $j+1$ will very likely be used for a modus ponens or a modus tollens. (3) Step k, again, is likely to come by conditional proof. But not necessarily; what *is* essential is that step $k+1$ come by induction on finite sets.

Exercise 9 *(Induction on finite sets)*

Show how you would set up to prove each of the following abstract forms by means of Induction on finite sets. F and G are supposed to be properties of sets. The exercise is to make sure that your proof outline matches the form of Corollary **2B-9**.

1. Every finite set is either F or G.
2. No F is finite.
3. Only Fs are finite.

▷ . . . . ◁

Note that you cannot use induction on finite sets to show "all Fs are finite"—much as it sounds like one of the foregoing examples 1–3 of Exercise 9.

We will need the following facts in order to get on with the study of finite sets.

2B-11 FACT. *(Finiteness of unit sets)*

$\{y\}$ is finite

PROOF. Trivial. □

2B-12 FACT. *(Union of finite sets)*

If X and Y are finite, so is $(X \cup Y)$.

PROOF. Exercise. □

Exercise 10 *(Unions of finite sets)*

Prove **2B-12**, using, if you choose, the following scheme. Prove instead: Y is finite → (X)[X is finite → X∪Y is finite]. Suppose Y is finite; now use **2B-9**, choosing Ψ(X) there as "X∪Y is finite" here. You may refer to and use parts of axioms, definitions, and exercises from NAL:§9A. Here and elsewhere you may *also* use any very elementary fact of set theory not explicitly listed in NAL provided you know how to prove it; e.g., the fact that (X∪Y) = (Y∪X), which comes very quickly from the axiom or rule of extensionality.

▷ . . . . ◁

This is almost all we need to know of the finite. We will later need to know a little bit about the infinite; it seems convenient to develop the matter here. The following is deeper than it looks, but we still call it a "variant."

2B-13 VARIANT. *(Infinite)*

X is infinite ↔ X is not finite

So much for The Concept of the Infinite. One of its properties that we will need to *use* (**3B-19**) is this: If you take a finite set away from an infinite set, you still have lots left. For its proof the following is convenient.

2B-14 LEMMA. *(Infinite less one)*

X is infinite → X − {y} is infinite

PROOF. Suppose X is infinite. If y ∉ X, then X = (X − {y}) by NAL:Exercise 64(1), and thus X − {y} is infinite. If on the other hand y ∈ X, we know by NAL:Exercise 64(2) that X = (X − {y}) ∪ {y}, so that the latter is also infinite. But then contraposing Fin+1 (**2B-7**) guarantees that X − {y} is infinite, as required. □

2B-15 FACT. *(Infinite less finite)*

Y is infinite and X is finite → Y − X is infinite.

PROOF. Prove instead: Y is infinite → (X)[X is finite → Y − X is infinite]. Suppose that Y is infinite. Now use Induction on finite sets, **2B-9**, choosing Ψ(X) there as "Y − X is infinite" here. The base case, that Y − ∅ is infinite, is trivial by easy set theory. Suppose, as hypothesis of induction, that Y − X is infinite; then Lemma **2B-14** implies that (Y − X) − {y} is infinite. But by applying easy definitions from NAL:§9A, this is just (Y − (X ∪ {y}), which accordingly is infinite, as required. □

It may be useful to exhibit this in subproof notation. In interpreting the square-bracket annotations on lines 2, 3, and 5 that indicate substitutions, keep in mind that we have chosen Ψ(X) of **2B-9** as "Y − X is infinite." You should check whether or not we have made a mistake.

1	| | Y is infinite	hyp, flag Y
2	| | Y − ∅ is infinite	Basis, from 1 by NAL:§9A [∅/X]
3	| | | Y − X_1 is infinite	hyp of ind, flag X_1, y [X_1/X]
4	| | | (Y − X_1) − {y} is infinite	3, Lemma **2B-14**
5	| | | Y − (X_1 ∪ {y}) is infinite	4, by NAL:§9A [(X_1 ∪ {y})/X]
6	| | (X)[X is finite → Y − X is infinite]	2, 3–5, Ind. Fin. Sets, i.e., **2B-9**
7	| Y is infinite → (X)(X is finite → (Y − X) is infinite)	1–6, CP

One more

2B-16 FACT. *(Subsets of finite sets)*

Subsets of finite sets are also finite: X is finite and Y ⊆ X → Y is finite. Hence, by contraposition, supersets of infinite sets are infinite.

PROOF. Straightforward. Use Induction on finite sets, **2B-9**, choosing Ψ(X) there as "(Y)[Y ⊆ X → Y is finite]" here. For the base case, use the fact that Y ⊆ ∅ → Y = ∅ (NAL:§9A) and Fin∅, **2B-7**.

Now suppose as inductive hypothesis that (Y)[(Y ⊆ X) → Y is finite], and further suppose that Y′ ⊆ (X ∪ {z}). (Y′ − {z}) ⊆ X by set theory, so Y′ − {z} is finite, by instantiating the inductive hypothesis. So Y′ is finite, by contraposing Infinite less one, **2B-14**, which concludes the inductive step and the proof. □

Note that in contrast to the proof of Infinite less finite, **2B-15**, here the variable "Y" is universally bound *inside* the inductive context $\Psi(X)$. We needed the inductive hypothesis true for all Y (not just for a particular Y chosen outside of the inductive step) in order to be able to instantiate to $Y' - \{z\}$.

Perhaps this point will be clearer if we lay out the proof as a subproof, using "EST" as short for "easy set theory" as presented in NAL:§9A.

1	$\quad\quad$ _ $Y \subseteq \varnothing$	hyp, flag Y
2	$\quad\quad Y = \varnothing$	1, EST
3	$\quad\quad$ Y is finite	2, Fin$\varnothing$, **2B-7**
4	$(Y)[Y \subseteq \varnothing \rightarrow$ Y is fin]	1–3 logic (Basis) $\quad[\varnothing/X]$
5	$\quad$ _ $(Y)[Y \subseteq X \rightarrow$ Y is finite]	hyp of ind, flag X, z $[X_1/X]$
6	$\quad\quad$ _ $Y' \subseteq (X \cup \{z\})$	hyp, flag Y'
7	$\quad\quad (Y' - \{z\}) \subseteq X$	6, EST
8	$\quad\quad Y' - \{z\}$ is finite	5, 7
9	$\quad\quad Y'$ is finite	8, **2B-14**
10	$\quad (Y)[Y \subseteq (X \cup \{z\}) \rightarrow$ Y is finite]	6–9, logic $\quad[(X_1 \cup \{z\})/X]$
11	$(X)[X$ is finite $\rightarrow (Y)[Y \subseteq X \rightarrow$ Y is finite]]	4, 5–9, Ind. fin. sets

You will find it instructive to try (unsuccessfully) to prove this by holding Y fixed instead; i.e., try to prove that Y is infinite $\rightarrow$ (X)[X is finite $\rightarrow$ not $(Y \subseteq X)$] by assuming that Y is infinite, and then trying to show the consequent by Induction on finite sets, **2B-9**, choosing $\Psi(X)$ there as "not $(Y \subseteq X)$" here.

Exercise 11 *(Finitude of Cartesian product)*

Show that if X is finite and Y is finite, then $(X \times Y)$ is finite. For this purpose, you may use some elementary facts about the Cartesian product $(X \times Y)$: $(X \times \varnothing) = (\varnothing \times X) = \varnothing$; $(X \times (Y \cup Z)) = ((X \times Y) \cup (X \times Z))$ and $(X \cup Y) \times Z)) = (X \times Z) \cup (Y \times Z)$; and $(\{x\} \times \{y\}) = \{\langle x, y\rangle\}$.

▷ . . . . ◁

2B.4 More basic grammar

After this excursus, let us return to the axiomatic development of TF. The first axiom, **2B-1**, gave the type of the four primitives of the grammar of TF. This "type information" compressed, however, two important facts. First, since TF-atom $\subseteq$ Sent, you can tell that every TF-atom is a sentence, which therefore gives you a way to *start* identifying sentences. It gives you a *basis* for the sentences, and is therefore often called a "base clause." Second, since

$$\sim \in (\text{Sent} \mapsto \text{Sent})$$

you know that whenever you have a sentence, A, you can also identify ∼A as a sentence, and proceed in a like manner *inductively*. This clause, and the similar clause for ⊃, are therefore often called "inductive clauses." Base clauses start you off and inductive clauses keep you going. A base clause can always be taken to have the form of a membership or subset statement, and an inductive clause can always be taken to have the form of a function space statement (though these forms are often well concealed). Both base and inductive clauses always give *sufficient* conditions. Turning now to the next axiom, it says in effect that all sentences can be constructed from TF-atoms by means of the conditional and negation connectives. It says that the basis and inductive clauses not only give you *a* way to identify sentences, but the *only* way, where your logical ear should pick up "only" as a signal of a *necessary* condition.

2B-17 AXIOM. *(Axiom of induction on sentences for* TF*)*

There are no sentences other than those constructed from TF-atoms by means of ⊃ and ∼; that is, let Ψ(A) be any use-context; **Suppose**

Basis step. For all p, Ψ(p).

Inductive step. For all A, B, Ψ(A) and Ψ(B) → (Ψ(A ⊃ B) and Ψ(∼A)).

Then (A)(A ∈ Sent → Ψ(A)). (Ind. on sentences for TF)

Axiom **2B-17** licenses "Induction on sentences": If every TF-atom has a property, and if A ⊃ B and ∼A have it whenever A and B do, then all sentences have it. Hence,

2B-18 FACT. *(Induction on sentences for* TF*)*

The following subproof rule is legal, where Ψ(A) is any use-context (context of our use-language) involving the variable "A":

flag p, for Basis		[p/A]
.		
.		
.		
Ψ(p)	[This is the conclusion of the Basis step]	
$\Psi(A_1)$		$[A_1/A]$
$\Psi(A_2)$	Inductive hyp, flag A_1, A_2	$[A_2/A]$
.		
.		
.		
$\Psi(\sim A_1)$	[These two are the conclusions	$[\sim A/A]$
$\Psi(A_1 \supset A_2)$	of the Inductive step]	$[(A_1 \supset A_2)/A]$
(A)(A $\in$ Sent $\rightarrow \Psi$(A))	Induction on sentences for TF	

Take notice of the five substitutions indicated on the right. In the light of Convention **2B-4**, the conclusion could have been stated as just

$$(A)\Psi(A)$$

Think of it in that way: Then you will see that Induction on sentences is a way of reasoning to a universal statement, a way that is an alternative to universal generalization. That is, if (A)Ψ(A) is a desired conclusion, we now have two different things to try: universal generalization and Induction on sentences.

Conventions **2B-3** and **2B-4** also allow us to insert "p $\in$ TF-atom" in the Basis step, or either "$A_1 \in$ Sent" or "$A_2 \in$ Sent" in the inductive step, if needed.

Note that Induction on sentences requires two premisses, each of which is itself a subproof. The first is the "Basis" or "base step" of Induction on sentences, and answers directly to the Basis step of the axiom. And indirectly it answers to what we have informally described as the "basis clause" in our axiomatic account of what it is to be a sentence, **2B-1** on p. 22. The second constitutes the "inductive step" or "induction step" of Induction on sentences, and answers directly to the "Inductive step" of the axiom. And indirectly it answers to what we have informally described as the "inductive clauses" of **2B-1**.

Drawing the conclusion corresponds directly to the "conclusion" of the axiom. Note that we are using the convention **2B-8** according to which Ψ(p), $\Psi(A_1)$, $\Psi(A_2)$, $\Psi(\sim A_1)$, and $\Psi(A_1 \supset A_2)$ are all obtained from Ψ(A) by substitution for the variable A. (As always, it is the *quantified* letter that is substituted for.) Be *sure* to choose flag-letters p, A_1, and A_2 that do not occur in $\Psi(A)$—either free or bound.

At this point, before the introduction of semantic ideas, there isn't much we want to say about all sentences, so that we have little use for Induction on sentences; but here is something useful:

2B-19 FACT. *(Sentence cases for* TF*)*

Every sentence is either a TF-atom, a conditional, or a negation. That is, (A)[A$\in$ Sent $\rightarrow$ (A$\in$TF-atom or A is a conditional or A is a negation)].

PROOF. This is an easy corollary of Induction on sentences, choosing $\Psi(A)$ there as what follows the arrow here. In fact, it is so easy you may well become confused, so that we provide some details.

First we establish the base step (of **2B-17** or **2B-18**); but p$\in$TF-atom by **2B-3** (this is a rare example of an explicit reference to one of our Conventions), so trivially either p$\in$TF-atom or p is a conditional or p is a negation.

Now suppose as hypothesis of the inductive step (of either **2B-17** or **2B-18**) that A_1 is either a TF-atom or a conditional or a negation, and that A_2 is either a TF-atom or a conditional or a negation. We need to show that $A_1 \supset A_2$ and $\sim A_1$ is each either a TF-atom or a conditional or a negation, and it is easy to see that we may use Definition **2B-5** on p. 23 to do so—without, oddly enough, appealing to the hypothesis of the inductive step.

This completes the proof. □

Even though the matter is trivial, it seems best to offer a subproof version.

1	flag p, for Basis	
2	p is a TF-atom	Convention**2B-3**
3	p is a TF-atom, or conditional, or negation	2
4	A_1 is a TF-atom, or conditional, or negation	
5	A_2 is a TF-atom, or conditional, or negation	Ind. hyps, flag A_1, A_2
6	$A_1 \supset A_2$ is a conditional	**2B-5**
7	$\sim A_1$ is a negation	**2B-5**
8	$A_1 \supset A_2$ is a TF-atom, or cond., or neg.	6
9	$\sim A_1$ is a TF-atom, or cond., or neg.	7
10	(A)(A$\in$Sent $\rightarrow$ A is a TF-atom, or cond., or neg.)	Ind. on sent., **2B-18**

The following two axioms are an essential part of the structure of TF; but we will not be using them explicitly. Instead they are included in order to justify certain "inductive definitions" over the sentences. They correspond directly to certain of the Peano axioms, NAL:§6B or §2C.1.

Their intuitive content is that TF is nowhere (grammatically) ambiguous: Each sentence of TF has a unique grammatical construction (decomposition; structure; form). For example, there is nothing in TF like the English "A only if B only if C," which has two distinct grammatical constructions depending on whether the first or the second "only if" is taken as major.[7]

2B-20 AXIOM. *(Distinctness for* TF*)*

Nothing is both a TF-atom and a negation, or both a TF-atom and a conditional, or both a negation and a conditional.

A deft combination of Axiom **2B-20** with Fact **2B-19** leads to

2B-21 COROLLARY. *(*TF*-atom equivalence)*

A is a TF-atom $\leftrightarrow$ A is neither a conditional nor a negation.

We could in fact have *defined* "TF-atom" in this way. But it is even more important to see that the whole significance of the "atomicity" of TF-atoms is caught in this equivalence: TF-atoms are "atomic" *only* in that they are neither conditionals nor negations, only from the point of view of this one mode of analysis. It is not just that we should not think of TF-atoms as necessarily "short"; it is furthermore true that some of them may even involve the conditional and negation in their construction—just as long as they are not themselves conditionals or negations. The point is this: TF-atoms are to be those sentences such that *their* truth values fix the truth values of *all* sentences reachable by the conditional and negation constructions.

[7]But, you may say, what about "A $\supset$ B $\supset$ C"? Isn't that ambiguous? Answer: Yes, but irrelevant. For (a) the expression "A $\supset$ B $\supset$ C" is not an expression of TF (but only of our use-language), and (b) the expression "A $\supset$ B $\supset$ C" does not succeed in denoting a sentence of TF (just because it is ambiguous in our use-language). You may continue, "but isn't A $\supset$ B $\supset$ C itself ambiguous?" Answer: We reject your question as ambiguous on the grounds that it uses language (use-language) that is ambiguous. (Confusing—but only if you think about it.)

2B-22 AXIOM. *(One-one for* TF*)*

$(A_1 \supset B_1) = (A_2 \supset B_2) \rightarrow (A_1 = A_2 \text{ and } B_1 = B_2)$

$\sim A_1 = \sim A_2 \rightarrow A_1 = A_2$

This axiom says that the whole uniquely determines the parts; e.g., if you have a conditional, the antecedent and consequent are each uniquely recoverable.

2B.5 More grammar: subformulas

In the upcoming grammatical discussion, watch for the promised deployment of the concept of the finite.

Sentences do not have "parts" in any physical sense; but they do have parts in a usefully metaphorical sense, namely, their subformulas.[8] We define "Subf(A)"—"the set of subformulas of A"—inductively; this by first explaining what Subf means for TF-atoms, and then explaining what it means for larger sentences in terms of what it means for smaller sentences. This procedure is justified by the axioms governing sentences, including the Distinctness and One-one axioms. The exact course of this justification, important as it is, is beyond the scope of these notes. (It is nevertheless useful to notice that there is no "closure clause" appropriate in the inductive definition of an *operator*. Once we have enough clauses to tell us what the appropriate value is for each member of the domain of definition, why, then we have enough clauses. Of course this only works because the predicate for the domain of definition is itself associated with a "closure clause.")

2B-23 DEFINITION. *(*Subf *)*

Subf(A)—the set of subformulas of A—is characterized as follows.

$\text{Subf} \in (\text{Sent} \mapsto \mathcal{P}(\text{Sent}))$ (Subftype)

$\text{Subf}(p) = \{p\}$ (Subfatom)

$\text{Subf}(A \supset B) = \text{Subf}(A) \cup \text{Subf}(B) \cup \{A \supset B\}$ (Subf⊃)

$\text{Subf}(\sim A) = \text{Subf}(A) \cup \{\sim A\}$ (Subf∼)

[8] Since we have systematically used "sentence" instead of "formula," perhaps we should here say "subsentence"; but the phrase "subformula" is so absolutely standard in the literature that it seems best to keep it.

(The powerset operation $\mathcal{P}$ is introduced in NAL:**9B-20**.) Because the difference in technique between inductive definitions of operators on the one hand and predicates on the other is confusing, permit us to repeat the observation that immediately preceded this definition: In an inductive definition of an *operator* (as opposed to a predicate), no closure clause is appropriate. Once we have stated the domain of definition of the operator, and then have stated what the value is of applying the operator to each thing in the domain of definition, there is nothing more to say. Perhaps it helps to note that inductive definitions of operators are always "piggy-backed" on antecedently given inductively generated sets (here the set of sentences). There is a closure clause for that underlying set, which is the domain of definition of the operator, but none for the operator itself.

2B-24 FACT. *(Finitude of sentences)*

Subf(A) is finite.

Exercise 12 *(Finitely many subformulas)*

Prove **2B-24**. Use Induction on sentences, **2B-18**, using facts about finiteness (§2B.3).

▷ . . . . ◁

It is especially important that each sentence is constructed from a finite group of TF-atoms; this obvious fact is connected with the finitude of truth tables.

2B-25 COROLLARY. *(Finitude of* TF-*atoms)*

Subf(A)∩TF-atom is finite.

PROOF. Trivial, using Fact **2B-16** on p. 30. □

2B-26 VARIANT. *(Subformula)*

A *is a subformula of* B ↔ A *is a part of* B ↔ B *contains* A ↔ A subf B ↔ A∈ Subf(B).

2B-27 FACT. *(Subformula partially orders Sent)*

The subformula relation expressed by "A subf B" is reflexive, transitive, and antisymmetric on the sentences, and accordingly partially orders them.

Sometimes we want a notion that stands to "subformula" as "proper subset" stands to "subset."

2B-28 DEFINITION. *(Proper subformula)*

A *is a proper subformula of* B $\leftrightarrow$ A *propsubf* B $\leftrightarrow$ (A subf B and A$\neq$B).

2C Yet more grammar: substitution

We are going to define substitution of a sentence for (every occurrence of) a TF-atom in a sentence. This definition isn't used for a while, but it becomes important later, and it finds its place here because it is a purely grammatical idea. The definition is inductive in form.

2C-1 DEFINITION. *(Substitution for* TF*-atoms)*

[B/p](A)—read "B for p in A"—is the result of putting B for all occurrences of p in A (even though we do not know what an "occurrence" is). That is,

$$[B/p](A) \in \text{Sent} \qquad ([]\text{A-type})$$

$$[B/p](p) = B \qquad ([]\text{A-atom1})$$

$$[B/p](q) = q \text{ if } q \neq p \qquad ([]\text{A-atom2})$$

$$[B/p]((C \supset D)) = ([B/p](C) \supset [B/p](D)) \qquad ([]\text{A}\supset)$$

$$[B/p]((\sim C)) = \sim([B/p](C)) \qquad ([]\text{A}\sim)$$

Observe that []A-atom1 could have been stated as

$$[B/p](q) = B \text{ if } q = p.$$

That longer version emphasizes that between them []A-atom1 and []A-atom2 cover an important excluded middle: When the thing into which you substitute (named inside the round parentheses) is a TF-atom, it either is or is not identical to the TF-atom for which you are substituting (named after the slash). Numerous arguments involving substitution will need to rely on this case structure, as you will see.

Sometimes [B/p] taken by itself is called "the substitution prefix," and is considered as denoting a function in the space (Sent$\mapsto$Sent). That's a good perspective to keep in mind: You hand [B/p] a sentence, and it hands you back another.

Exercise 13 *(Substitution for* TF-*atoms)*

The following are intuitively true according to our intuitive concept of substitution for TF-atoms. Test the success of our effort to catch this concept in a rigorous definition by giving them rigorous proofs—using, of course, our definition. (Recall that p and q are TF-atoms by **2B-3**; do *not* assume that they are distinct):

1. [p/q](p) = p
2. [q/p](p) = q
3. [p/q]((p ⊃ q)) = (p ⊃ p)
4. [∼p/p]((∼p ⊃ p)) = (∼∼p ⊃ ∼p)

▷ . . . . ◁

Next on the grammatical agenda is a statement as to how many TF-atoms there are in the language TF; but such a statement must be cast in language that is rendered intelligible by containing only language we understand. Part of what we want to say is that TF-atom is nonempty—there are some TF-atoms; and we already know what this means, NAL:**9A-11**. But we also want to say that the TF-atoms are "countable"; and for this we must lay some groundwork.

2C.1 Arithmetic: N and N$^+$

A set of special usefulness is the set N of non-negative integers 0, 1, … . It is usual in set theory to define 0, 1, …, as certain sets, but we leave that development for another course. Instead, we are going to assume that "0" and "1" are primitive, as well as ordinary addition "+" and multiplication "×"; we shall simply lay down for these primitives the usual Peano axioms—except that we use "$n+1$" instead of Peano's successor operator. For convenience we divide the axioms into three groups.

2C-2 AXIOM. *(Peano axioms, types)*

N is a set	(Ntype)
$0 \in N$	(0type)
$1 \in N$	(1type)
$+ \in ((N \times N) \mapsto N)$	(+type)
$\times \in ((N \times N) \mapsto N)$	(×type

2C-3 CONVENTION. *("k," "m," "n" over N)*

We use "k," "m," "n" as variables ranging over N.

2C-4 AXIOM. *(Peano axiom, induction on N)*

Every member of N can be reached by successive additions of 1 to 0; that is, where $\Psi(n)$ is any use-context, **Suppose**

Basis step. $\Psi(0)$.

Inductive step. $(n)(\Psi(n) \rightarrow \Psi(n+1))$.

Then $(n)(n \in N \rightarrow \Psi(n))$. (Induction on N)

Exercise 14 *(Induction on N as a subproof rule)*

Formulate a subproof rule that carries the content of Axiom **2C-4**.

▷ . . . . ◁

We will need just a few facts of arithmetic. First the remaining basic Peano facts, transformed into our setting, and labeled as "axioms" because for us they are starting points.

2C-5 AXIOM. *(Peano axioms, more)*

$n+1 \neq 0$

$m+1=n+1 \rightarrow m=n$

$m+0=m$

$m+(n+1)=(m+n)+1$

$m \times 0=0$

$m \times (n+1)=(m \times n)+m$

Much later (**2F-14**) we will use the less-than-or-equals concept, and somewhat sooner some cousins, and so we define them all now.

2C-6 DEFINITION. *(Less-than-or-equals; less-than; predecessors-of)*

$m \leqslant n \leftrightarrow$ for some $k \in N$, $m+k=n$. $(\leqslant)$

$m < n \leftrightarrow m \leqslant n$ but $m \neq n$. $(<)$

Variant: m *precedes* (or *is a predecessor of*) $n \leftrightarrow m < n$.

For each n, *predecessors-of*(n) is that set such that for all k, $k \in$ predecessors-of(n) $\leftrightarrow k < n$.

Variant: *the*-n-*initial* = predecessors-of(n).

X is a *finite initial of* N $\leftrightarrow$ X = predecessors-of(n) for some $n \in N$.

It is easy to see intuitively that since 0 has 0 predecessors, 1 has 1 predecessor, etc., predecessors-of(n) has always n members. Thus predecessors-of(n) is a convenient set with a known number of known elements. One speaks of finite initials when one wants to generalize over this situation.

2C-7 FACT. *(Useful arithmetical facts)*

$m \leqslant n$ or $n \leqslant m$

$(m+1)+k=m+(k+1)$

PROOF. Tedious. Almost unbelievably, however, these facts are the ones directly useful in our logical enterprise. □

Even later (§3E.3), we will rely on the following

2C-8 FACT. *(One-one-ness of doubling function)*

$(m+m)=(n+n) \rightarrow m=n$.

PROOF. Tedious. □

Sometimes it is considerably more convenient to begin counting with 1.

2C-9 DEFINITION. *(N^+)*

$N^+=(N-\{0\})$

So N^+ starts with 1.

A principal use of N is as a counting device.

2C-10 DEFINITION. *(Countable)*

A set X is *countable* ↔ either it is empty or if there is a function f in $(N \mapsto X)$ with *Rng*(f) = X.

The promised f can be thought of as counting X. That *Rng*(f) = X guarantees that everything in X is eventually counted; in fact, if X is finite, something in X is bound to be counted many, many times over!

It is worth observing that countable sets can be either finite or infinite; for the immediate application we have in mind it just doesn't matter.

The following gives the most usual notation for counting the countable.

2C-11 CONVENTION. *(Subscripts)*

When x is a function in $(N \mapsto X)$ (or in $(N' \mapsto X)$ for N' a subset of N), and when $n \in N$ (or $n \in N'$), we may write the argument as a subscript:

$$x_n = x(n)$$

The picture to have is that x itself is a sequence: $\langle x_0, x_1, x_2, \ldots\rangle$. From the picture you can see that there is no bar to some items being counted twice.

The following is a useful combination of **2C-10** and **2C-11**.

2C-12 COROLLARY. *(Subscripting the countable)*

Suppose X is a countable nonempty set. Then there is a function x in $(N \mapsto X)$ such that

$$(y)[y \in X \leftrightarrow \exists n(n \in N \text{ and } y = x_n)]$$

This is so convenient because we can obtain the effect of using "x_n" as a variable over X by using "n" as a variable over N.

2C-13 VARIANT. *(Enumeration)*

Just in case x has the property of Corollary **2C-12**, then it is said to be an *enumeration of*—or to *enumerate*—X.

There is also a concept of counting or enumerating "without repetition," but we do not happen to require that concept for our immediate purposes.

2C.2 More grammar—countable sentences

Back to grammar: Armed with the concept of countability, we add an axiom that caps off our account of the grammar of TF by saying how many TF-atoms there are.

2C-14 AXIOM. *(How many* TF-*atoms)*

TF-atom is nonempty and countable (see NAL:**9A-11** and **2C-10**).

Note that we tell you something about *how many* TF-atoms there are, but not (so to speak) *what* they are. According to the axiom there might be either finitely or infinitely many TF-atoms; nothing[9] we say here depends on which, so that there is no point in saying which—and some point (generality) in not.

In contrast, when we come to quantifiers it will turn out to be logically important to have infinitely many individual variables in their own right, but here this axiom is used only to guarantee the truth of the following fact, which in turn is used only much later.

2C-15 FACT. *(Countable sentences)*

Sent is nonempty and countable; or more idiomatically, there are countably many sentences.

That is, by **2C-10**, there is a function B in (N$\mapsto$Sent) such that for every sentence C, there is an n in N such that $C = B_n$ (every sentence is counted). Note that in this context, B is a function and B_n is a sentence. This is confusing only if you think about it.

PROOF. Omitted. As a matter of fact, Axiom **2C-14** collaborates with the Distinctness axiom **2B-20** to guarantee *infinitely* many sentences in all; but we do not need this information. The proof is omitted because there is just a lot of set-theoretical argle-bargle involved in showing that whenever a countable set (such as TF-atom) is generated by a finite number of finite operations (such as $\supset$ and $\sim$), the result (in this case, sentence) is bound to be countable. Pictorially, you can see what is going on if you can see that if a rectangular or square array has countable rows and also countable columns, then it has countable items in all—as Cantor taught us, one can start in the upper left hand corner and count through the entire array on the "lower left/upper right" diagonals, counting up one (from lower left to upper right) and then down the next (from upper right to lower left). With this picture firmly in mind, proceed to make more pictures as follows. In row 0, list all the TF-atoms. In each odd row, list negations of all preceding items (they will be countable because the preceding items "clearly" form a rectangular array). In each even row, list conditionals formed in all possible ways from all preceding items (the preceding items are countable as before, and the conditionals formed from them are countable because they can be arranged in a square array with antecedents as row headings and consequents as column headings). "Clearly" this Big Array exhausts the sentences, and has countable rows and columns, so that the sentences are themselves countable. □

[9]Well, hardly anything. The expressive powers of TF can be limited by too few TF-atoms; this subtlety is discussed just after the statement of Theorem **2D-38**.

Exercise 15 *(Grammar of* TF*)*

1. Assume that A and B are sentences; prove that the following are sentences: $\sim$(A $\supset$ B) and $\sim$A $\supset$ B.
2. Prove that for all A, A is a subformula of A. (Use TF-sentence cases, **2B-19**.)
3. Prove that A is a subformula of A $\supset$ A. (Use universal generalization and the previous exercise.)
4. Prove that for all A, A $\supset$ A $\neq$ A. (Use Induction on sentences.) Also prove that $\sim$A $\neq$ A.
5. Optional. Prove that for all A, B, ($\sim$B subf A) $\rightarrow$ (B subf A). Do this most briefly (?) by proving instead that it holds for all B, A—with quantifiers in reversed order. For this latter, use universal generalization for B and Induction on sentences for A.
6. Optional. Take as a fact that (B $\supset$ C) subf A implies (B subf A) and (C subf A). Using this, now show that subf is transitive: (A subf B) and (B subf C) $\rightarrow$ (A subf C). Try this by arranging the variables in the order A, B, C; and then using universal generalization on A and B, and Induction on sentences on C.
7. Optional. Show that subf is antisymmetric: ((A subf B) and (B subf A)) $\rightarrow$ A = B. This may require using Induction on sentences on both A and B.
8. Show by example that the following does not hold for all A, B, C: [C/p](([B/p](A))) = [B/p](([C/p](A))). To figure this out, try picturing A = ...p..., B = _p_, etc.
9. Draw a "persuasive picture" representing the Big Array of the omitted proof of Fact **2C-15**.

▷ . . . . ◁

2D Elementary semantics of TF

As we have said, for each branch of logic (for each language), semantics and proof theory are themselves mutually independent, while each depends on grammar. This section treats elementary semantics, the next elementary proof theory.

2D.1 Valuations and interpretations for TF

The picture to have in mind in the course of working through the following material is: a single row of a single truth table, for example

	p	q	r	p	r	$p \supset r$	$\sim(p \supset r)$	$\sim p$	$\sim p \supset r$
	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.
⇒	T	F	T	T	T	T	F	F	T
	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.

Consider just the part to the *right* of the vertical line: Each displayed sentence is given a truth value in every row. The first thing we do is to record this in a definition, at the same time building in the idealization that in a row of a truth table a value is given not only to some few sentences in which we may happen to be interested, but to all sentences whatsoever. This idealization of the right hand portion of a row of truth table is called a "sentential valuation," a mapping from the sentences into the truth values:

2D-1 DEFINITION. *(Sentential valuation)*

sentential valuation $= (\text{Sent} \mapsto \mathbf{2})$.

There are two remarks to be made about this definition, one negative and one positive. The negative one is that we have so far extracted very little from the picture of a row of a truth table—the concept of a sentential valuation puts *no* constraints whatsoever on how truth values are to be assigned to sentences; e. g., p and $\sim$p might both be given the value T by some sentential valuation. On the positive side, however, we do know something: Each sentential valuation gives each sentence precisely one of the values T and F. There are no "truth value gaps," and there are no truth values other than T and F.[10] This property is called "bivalence" (two valuedness):

[10] Many-valued logic begins here—by permitting additional truth values.

2D-2 FACT. *(Bivalence of sentential valuations)*

If A is a sentence and f is a sentential valuation, then

$$f(A)=T \leftrightarrow f(A)\neq F, \text{ and}$$
$$f(A)=F \leftrightarrow f(A)\neq T.$$

PROOF. Trivial; but here it is anyhow. Suppose that A is a sentence and that f is a sentential valuation, **2D-1**. Since $f \in (\text{Sent}\mapsto \mathbf{2})$, $f(A)\in \mathbf{2}$ by MP for function space, NAL:**9B-15**. But $\mathbf{2}=\{T, F\}$, **2A-1**, so the Fact follows by **2A-5**. □

Remark. Bivalence, you will note, has nothing whatsoever to do with truth functionality; one can have a many-valued but still truth functional logic, or a bivalent logic whose connectives are not truth functional.

Returning to our picture of a row of a truth table, p. 46, it is obvious that only some of the sentential valuations represent really acceptable truth table calculations—those, namely, that correctly derive the value of $A\supset B$ from the values of A and B, and the value of $\sim A$ from the value of A. Because they respect the grammar of TF, we will label as "TF-valuations" those sentential valuations that treat ⊃ and ∼ properly.

2D-3 DEFINITION. *(*TF*-valuation)*

$f\in$ TF-valuation ↔

f is a sentential valuation, **2D-1**	(TF-valtype)
$f(A\supset B)=f(A) \supset^{*} f(B)$	(TF-val⊃)
$f(\sim A)=\sim^{*}(f(A))$	(TF-val∼

The last two clauses, TF-val⊃ and TF-val∼, display with maximum clarity Frege's principle that the truth value of each compound "is a function of" the truth values of the parts—for these clauses tell you explicitly *which* function (which truth function, **2A-7**) to employ. They are special cases of a much more general (and less precise) principle of formal semantics that we must equally ascribe to Frege: Let the value of the whole be a function of the value of the parts. It is to highlight this feature of the semantics of TF that we chose to define TF-valuations as in **2D-3**, instead of using clauses like those of **2D-13** below, even though the latter are often more useful in proving things about the semantics of TF.

So far we have served notice in our definitions that a row of a truth table gives each sentence a value T or F, and that it does so in accordance with standard truth tables for $\supset$ and $\sim$. There is also the feature of truth tables that they have "reference columns," arbitrary assignments to those parts which are not (from the point of view of TF) complex—these are given on the *left* of the vertical line in the picture on p. 46. This leads us to a central idea of semantics, that of an "interpretation." We use the word as follows. Envision a language as having some "atomic" parts whose "meaning" can vary, and some grammatical ways of making larger items out of smaller items. The job of an interpretation (in any branch of logic) is to give us enough information about these "atomic" parts to fix a value for each of the larger items.

What does this mean in the special case TF? What do we intuitively need to know about a sentence in order to compute its value? Clearly, the truth value of each of its TF-atoms. This motivates the following

2D-4 DEFINITION. *(TF-interpretation)*

TF-interpretation = (TF-atom $\mapsto$ **2**)

That is, **i** is a TF-interpretation just in case **i** is a function defined on all TF-atoms that takes values in **2**. A TF-interpretation assigns a truth value to each TF-atom.

Picture: A TF-interpretation is the same as the "reference column" portion of a row of a truth table—except that there is a reference column for every TF-atom, not just for a few. Later, when we turn our attention to other languages, we will need to distinguish other sorts of interpretations—but not now.

2D-5 CONVENTION. *("**i**" for TF-interpretations)*

We use "**i**," "$\mathbf{i}_1$," etc., as ranging over TF-interpretation.

How many TF-interpretations are there? Answer: many.

2D-6 FACT. *(Existence of TF-interpretations)*

Let G be any set of sentences. Then there is a TF-interpretation **i** such that

$\mathbf{i}(p) = T \leftrightarrow p \in (\text{TF-atom} \cap G)$ and
$\mathbf{i}(p) = F \leftrightarrow p \in (\text{TF-atom} - G)$.

PROOF. Consult NAL:**9B-19** concerning characteristic functions, choosing X there as TF-interpretation here and Y there as G here. □

So far we have introduced a couple of phrases, "sentential valuation" and "TF-valuation," for the right side of a row of a truth table, and the phrase "TF-interpretation" for the left side; we next want to make good on what is obvious to everyone: The TF-interpretation (of the TF-atoms) is enough to fix the TF-valuation (of all sentences). In fact this is precisely what one learns in learning truth tables: One learns a method of fixing the value of each sentence when the value of each TF-atom is known. In the jargon of set theory, what we have here is the concept of "the TF-valuation determined by **i**," where **i** is a TF-interpretation. For notation it is convenient to write "$Val_{\mathbf{i}}$" for "the TF-valuation determined by **i**." The idea is that the TF-*valuation* $Val_{\mathbf{i}}$ shall exactly agree with the TF-*interpretation* **i** on all those items on which both are defined—namely, the TF-atoms—just as in the picture on p. 46, p and r have the same values on the right of the vertical ("valuation" part of picture) as they do in the "reference columns" on the left ("interpretation" part of picture); and that otherwise $Val_{\mathbf{i}}$ shall be a TF-valuation, **2D-3**. That is,

2D-7 DEFINITION. *($Val_{\mathbf{i}}$)*

For each TF-interpretation **i**, $Val_{\mathbf{i}}$ is the (unique) TF-valuation such that for every TF-atom p, $Val_{\mathbf{i}}(p) = \mathbf{i}(p)$. That is,

$$Val_{\mathbf{i}} \in (\text{Sent} \mapsto \mathbf{2}) \qquad (Val_{\mathbf{i}}\text{type})$$

$$Val_{\mathbf{i}}(p) = \mathbf{i}(p) \qquad (Val_{\mathbf{i}}\text{atom})$$

$$Val_{\mathbf{i}}(A \supset B) = Val_{\mathbf{i}}(A) \supset^{*} Val_{\mathbf{i}}(B) \qquad (Val_{\mathbf{i}} \supset)$$

$$Val_{\mathbf{i}}(\sim A) = \sim^{*}(Val_{\mathbf{i}}(A)) \qquad (Val_{\mathbf{i}} \sim$$

This is an inductive definition of "$Val_{\mathbf{i}}$," the justification of which is beyond our scope. But note that for each argument **i**, "$Val_{\mathbf{i}}$" for more complex sentences is explained in terms of itself for simpler sentences, until we get to TF-atoms—when "$Val_{\mathbf{i}}$" disappears in favor of just "**i**." We could therefore prove (but we won't) that in fact *there is* a unique function satisfying the conditions laid down; and that is quite enough to ensure that the definition satisfies the criteria for good definitions (eliminability and conservativeness). The uniqueness side is actually only tedious; but the existence side, which requires the Distinctness and One-one axioms **2B-20** and **2B-22**, is of a degree of complexity we wish to avoid. It nevertheless seems useful to point out in a more informal way the roles of the Distinctness and One-one axioms in semantics. Let $\mathbf{i}$(p)=T. Suppose Distinctness, **2B-20**, failed, so (say) ∼p=p. Then it would be easy to use **2D-7** to show that $Val_{\mathbf{i}}$(∼p) equals T *and* that it equals F, contradicting Axiom **2A-4**; thus, Definition **2D-7** of $Val_{\mathbf{i}}$ would be nonconservative in the worst way. Similarly, suppose the One-one axiom **2B-22** failed; say ∼p=∼q while p≠q. Let $\mathbf{i}$(p)=T and $\mathbf{i}$(q)=F to produce a contradiction with Axiom **2A-4** via Definition **2D-7**.

Note that we could have let Val ∈ TF-interp ↦ (Sent ↦ **2**). In that case, its argument, $\mathbf{i}$, appears as a subscript.

The individual clauses of Definition **2D-7**, all important, are named at its right. The first, $Val_{\mathbf{i}}$type, tells us what kind of function $Val_{\mathbf{i}}$ is. It has the following consequences, which are mostly used silently.

2D-8 COROLLARY. *($Val_{\mathbf{i}}$(A))*

$Val_{\mathbf{i}}$(A)∈**2** ($Val_{\mathbf{i}}$type)

2D-9 COROLLARY. *(Bivalence)*

$Val_{\mathbf{i}}$(A)=T ↔ ($Val_{\mathbf{i}}$(A)≠F)

$Val_{\mathbf{i}}$(A)=F ↔ ($Val_{\mathbf{i}}$(A)≠T

The second clause of **2D-7**, namely, ($Val_{\mathbf{i}}$atom), says that the value of a TF-atom, p, is determined by the TF-interpretation $\mathbf{i}$. This clause corresponds to the following in truth table calculations: When you need a truth value for p, look to the reference column for p.[11]

It is easy to confuse a "TF-interpretation" $\mathbf{i}$ and a "TF-valuation" $Val_{\mathbf{i}}$; how are they different? The most important difference lies in their respective domains of definition, for

---

[11]An analog of the clause "$Val_{\mathbf{i}}$(p)=$\mathbf{i}$(p)" of **2D-7** is expressed in the following wonderfully mysterious way by Church 1956, p. 73: "A form which consists of a variable **a** standing alone has the value t for the value t of **a**, and the value f for the value f of **a**."

$\mathbf{i} \in (\text{TF-atom} \mapsto \mathbf{2})$, while

$Val_{\mathbf{i}} \in (\text{Sent} \mapsto \mathbf{2})$.

That is, $\mathbf{i}$ is defined on *only* TF-atoms, while $Val_{\mathbf{i}}$ is defined on those, *and* on complex sentences as well. In particular, "$\mathbf{i}(A \supset B)$" is quite senseless. In words: A TF-interpretation $\mathbf{i}$ interprets (or values) only the TF-atoms, while a TF-valuation $Val_{\mathbf{i}}$ values (or interprets) all sentences.[12]

Even though they are set-theoretically different, TF-interpretations and TF-valuations are closely linked, and in fact stand in a natural one-one correspondence:

2D-10 FACT. *(*TF-*interpretations and* TF-*valuations)*

For each TF-valuation f there is a unique TF-interpretation $\mathbf{i}$ such that f is in fact the TF-valuation determined by $\mathbf{i}$: $f = Val_{\mathbf{i}}$. In other words, $Val_{\mathbf{i}} = Val_{\mathbf{i}'} \rightarrow \mathbf{i} = \mathbf{i}'$.

For each TF-interpretation $\mathbf{i}$ there is a unique TF-valuation f such that $\mathbf{i}$ and f agree on all TF-atoms: $\mathbf{i}(p) = f(p)$, all TF-atoms p. Namely, set $f = Val_{\mathbf{i}}$.

PROOF. Omitted. If you can see the picture, that is enough. □

Because of this intimate link between TF-interpretations and TF-valuations, we could have chosen either as fundamental and dispensed with the other; but it seems more illuminating to carry both concepts.

2D.2 Truth and models for TF

This section deals chiefly with variants on the concept of a valuation; the only part that must not be skipped is Fact **2D-13** and the following Exercise 16.

Sometimes we want truth as a predicate; or, since truth is relative to a TF-interpretation ("true on a row of a truth table"), as a two place predicate:

2D-11 VARIANT. *(True on)*

A is true on $\mathbf{i} \leftrightarrow Val_{\mathbf{i}}(A) = T$

[12] The words "interpretation" and "valuation" are jargon to be memorized; surely the distinction is not marked in everyday English.

We have thus first postulated a primitive, "T," that we think of as naming *Truth*, and then we have defined a relational truth predicate in terms of it. That such a definitional sequence is possible is philosophically interesting; but it is bad philosophy either to suppose that our particular ordering of concepts is inevitable or to suppose that it is obfuscating.

We have a choice when it comes to defining "A is false on **i**"; do we mean it isn't true (doesn't have the value T), or do we mean it has the value F? In three valued logic, or logic with truth value gaps, the choice would make a real difference, but here, of course, bivalence (**2D-9**) tells us that it doesn't matter; nevertheless we must choose.

2D-12 VARIANT. *(False on)*

A is false on $\mathbf{i} \leftrightarrow Val_{\mathbf{i}}(\mathrm{A}) \neq \mathrm{T}$

The variant "true on" suggests the following obvious variations on the clauses ($Val_{\mathbf{i}}\supset$) and ($Val_{\mathbf{i}}\sim$) of **2D-7**; they are frequently *useful*, and so we give them names.

2D-13 FACT. *(Truth and connectives)*

$$Val_{\mathbf{i}}(\mathrm{A}\supset\mathrm{B})=\mathrm{T} \leftrightarrow [Val_{\mathbf{i}}(\mathrm{A})=\mathrm{T} \rightarrow Val_{\mathbf{i}}(\mathrm{B})=\mathrm{T}] \qquad (\mathrm{Tr}\supset)$$

$$Val_{\mathbf{i}}(\sim\mathrm{A})=\mathrm{T} \leftrightarrow Val_{\mathbf{i}}(\mathrm{A})\neq\mathrm{T} \qquad (\mathrm{Tr}\sim$$

The clause Tr⊃ is to be read: $Val_{\mathbf{i}}(\mathrm{A}\supset\mathrm{B})=\mathrm{T}$ just in case, if $Val_{\mathbf{i}}(\mathrm{A})=\mathrm{T}$ then $Val_{\mathbf{i}}(\mathrm{B})=\mathrm{T}$. It is more usual to say, instead, that $Val_{\mathbf{i}}(\mathrm{A}\supset\mathrm{B})=\mathrm{T}$ just in case either $Val_{\mathbf{i}}(\mathrm{A})=\mathrm{F}$ or $Val_{\mathbf{i}}(\mathrm{B})=\mathrm{T}$. We put the matter as we do for two reasons. In the first place, the "either-or" statement frequently suggests proofs by case argument, whereas the statement Tr⊃ suggests proofs by modus ponens and conditional proof; the latter are generally shorter and more elegant. In the second place, avoiding the use of the "if then" construction in English to read the "if then" construction of TF seems to us somehow disingenuous *if* one believes that the English "if then" is truth functional. We do not believe this; but (as we have stated, **1A-4**) in technical passages in these notes we are so using "if then"; so that it would be philosophically misleading not to extract all the good we can out of this decision.

PROOF. Straightforward ☐

Exercise 16 *($Val_{\mathbf{i}}$⊃MP and $Val_{\mathbf{i}}$⊃CP)*

Formulate natural deduction "rule" versions of (Tr⊃), which are to be called "$\text{Tr}_{\mathbf{i}}$⊃MP" and "$\text{Tr}_{\mathbf{i}}$⊃CP," in analogy with **2A-9** on p. 14. These principles may be referenced and used.

▷ . . . . ◁

The following two definitions illustrate that sometimes we like to use the same word in two different contexts with somewhat different (though related) meanings. The first merely gives a *converse* of "true on."

2D-14 VARIANT. *(Model of a sentence)*

---

i *is a model of* A ↔ A is true on **i** ↔ $Val_{\mathbf{i}}$(A) = T

The second lets "model of" apply to *sets* of sentences,

2D-15 DEFINITION. *(Model of a set of sentences)*

If G is a set of sentences (i.e., if G ⊆ Sent), then **i** *is a model of* G ↔ (A)(A ∈ G → $Val_{\mathbf{i}}$(A) = T).

That is, a model of a set makes every member come out true.

An interesting side excursion is provided by the following

2D-16 DEFINITION. *(Model set)*

M is the unique function such that

M ∈ (Sent ↦ $\mathcal{P}$(TF-interpretation)) (Mtype)

i ∈ M(A) ↔ **i** is a model of A (M)

Recall that $\mathcal{P}$(TF-interpretation) is the set of all subsets of TF-interpretation, NAL:**9B-20**. Thus M(A), "the model set of A," is the set of TF-interpretations that make A true. The following fact exploits this definition to make a connection between set operators and connectives.

2D-17 FACT. *(Connectives and operators on model sets)*

$M(A\&B) = M(A) \cap M(B)$

$M(A \vee B) = M(A) \cup M(B)$

$M(\sim A) = -M(A)$

$M(A \supset B) = -M(A) \cup M(B$

where complementation of sets is relative to TF-interpretation:
$-X = (\text{TF-interpretation} - X)$.

PROOF. Straightforward, using Extensionality, NAL:**9A-7**. □

2D.3 Two two-interpretation theorems and one one-interpretation theorem

This section contains three fundamental theorems concerning TF-valuations $Val_{\mathbf{i}}$ taken, as it happens, one or two at a time. Only in the following section do we go on to the concepts defined by quantifying over TF-valuations.

The first theorem tells us that our extravagance in permitting each TF-interpretation to interpret every one of the perhaps infinitely many TF-atoms is not so extravagant after all in the sense that all the extra "information" does no work. In fact, the value of each sentence is wholly determined by the interpretation of the TF-atoms that are part of it—the values of all the rest are irrelevant. This, then, justifies the practice of writing down reference columns for only those TF-atoms actually occurring in the sentences to be truth-tabled

The name of the theorem derives from the fact that what it says is that the value of each sentence is determined "locally" (by its own TF-atoms) instead of "globally" (by perhaps all TF-atoms). One of the reasons the Local determination theorem for TF is of such primary importance is that it touches on the "learnability of language" theme: It is thought (probably without warrant) that a language is learnable, or more learnable, if the entire semantic value of a piece of it derives from parts of that piece.

2D-18 THEOREM. *(Local determination theorem for* TF*)*

If TF-interpretations $\mathbf{i}$ and $\mathbf{i}'$ agree on every TF-atom subformula of A, then the associated TF-valuations $Val_{\mathbf{i}}$ and $Val_{\mathbf{i}'}$ agree on A:

$$(\mathrm{q})(\mathrm{q}\in\mathrm{Subf}(\mathrm{A}) \rightarrow \mathbf{i}(\mathrm{q})=\mathbf{i}'(\mathrm{q})) \rightarrow Val_{\mathbf{i}}(\mathrm{A})=Val_{\mathbf{i}'}(\mathrm{A})$$

PROOF. By Induction on sentences, **2B-18**. First consider the base case $\mathrm{A}=\mathrm{p}$. Since $\mathrm{p}\in\mathrm{Subf}(\mathrm{p})$, **2B-23** on p. 36, hence $\mathbf{i}(\mathrm{p})=\mathbf{i}'(\mathrm{p})$ by the hypothesis of the theorem, so $Val_{\mathbf{i}}(\mathrm{p})=Val_{\mathbf{i}'}(\mathrm{p})$ by ($Val_{\mathbf{i}}$atom), **2D-7** on p. 49.

Now consider $\mathrm{A}=\mathrm{A}_1\supset\mathrm{A}_2$, with the theorem supposed true for A_1 and A_2. Suppose $\mathbf{i}$ and $\mathbf{i}'$ agree on every TF-atom subformula of $\mathrm{A}_1\supset\mathrm{A}_2$. Then (since subformulas of A_1 or of A_2 are—by subf$\supset$ of **2B-23**—subformulas of $\mathrm{A}_1\supset\mathrm{A}_2$) they must agree on every TF-atom subformula of each of A_1 and A_2, so that by the hypothesis of induction, $Val_{\mathbf{i}}(\mathrm{A}_1)=Val_{\mathbf{i}'}(\mathrm{A}_1)$ and $Val_{\mathbf{i}}(\mathrm{A}_2)=Val_{\mathbf{i}'}(\mathrm{A}_2)$. So

$$Val_{\mathbf{i}}(\mathrm{A}_1\supset\mathrm{A}_2) = Val_{\mathbf{i}'}(\mathrm{A}_1\supset\mathrm{A}_2)$$

as desired, by ($Val_{\mathbf{i}}\supset$), **2D-7**, twice.

The case for $\mathrm{A}=\sim\mathrm{A}_1$ is similar. □

A "subproof" version of this proof might go as follows. We write "($\mathbf{i}=\mathbf{i}'$ on A)" for "for all q, if q subf A, then $\mathbf{i}(\mathrm{q})=\mathbf{i}'(\mathrm{q})$."

1	$\mathrm{p}\in$TF-atom	hyp, flag p, for basis
2	$(\mathbf{i}=\mathbf{i}'$ on p)	hyp for CP
3	p subf p	1, Subfatom, **2B-23**
4	$\mathbf{i}(\mathrm{p})=\mathbf{i}'(\mathrm{p})$	2, 3
5	$Val_{\mathbf{i}}(\mathrm{p})=Val_{\mathbf{i}'}(\mathrm{p})$	4, ($Val_{\mathbf{i}}$atom, **2D-7**)
6	$(\mathbf{i}=\mathbf{i}'$ on p)$\rightarrow Val_{\mathbf{i}}(\mathrm{p})=Val_{\mathbf{i}'}(\mathrm{p})$	CP [p/A]
7	$(\mathbf{i}=\mathbf{i}'$ on $\mathrm{A}_1) \rightarrow Val_{\mathbf{i}}(\mathrm{A}_1)=Val_{\mathbf{i}'}(\mathrm{A}_1)$	hyp of ind., flag A_1 [A_1/A]
8	$(\mathbf{i}=\mathbf{i}'$ on $\mathrm{A}_2)\rightarrow Val_{\mathbf{i}}(\mathrm{A}_2) = Val_{\mathbf{i}'}(\mathrm{A}_2)$	hyp of ind., flag A_2 [A_2/A]
9	$(\mathbf{i}=\mathbf{i}'$ on $(\mathrm{A}_1\supset\mathrm{A}_2))$	hyp for CP
10	p subf A_1	hyp, flag p, for UGC
11	p subf $\mathrm{A}_1\supset\mathrm{A}_2$	10, Subf$\supset$, **2B-23**
12	$\mathbf{i}(\mathrm{p}) = \mathbf{i}'(\mathrm{p})$	9, 11
13	$(\mathbf{i}=\mathbf{i}'$ on $\mathrm{A}_1)$	UGC
14	$Val_{\mathbf{i}}(\mathrm{A}_1)=Val_{\mathbf{i}'}(\mathrm{A}_1)$	7, 13
15	$Val_{\mathbf{i}}(\mathrm{A}_2)=Val_{\mathbf{i}'}(\mathrm{A}_2)$	like 14
16	$Val_{\mathbf{i}}(\mathrm{A}_1) \supset^* Val_{\mathbf{i}}(\mathrm{A}_2)=Val_{\mathbf{i}'}(\mathrm{A}_1)\supset^* Val_{\mathbf{i}'}(\mathrm{A}_2)$	14, 15
17	$Val_{\mathbf{i}}(\mathrm{A}_1\supset\mathrm{A}_2) = Val_{\mathbf{i}'}(\mathrm{A}_1\supset\mathrm{A}_2)$	16, $Val_{\mathbf{i}}\supset$, **2D-7**
18	$(\mathbf{i}=\mathbf{i}'$ on$(\mathrm{A}_1\supset\mathrm{A}_2)) \rightarrow Val_{\mathbf{i}}(\mathrm{A}_1\supset\mathrm{A}_2) = Val_{\mathbf{i}'}(\mathrm{A}_1\supset\mathrm{A}_2)$	CP [$(\mathrm{A}_1\supset\mathrm{A}_2)$/A]
19	$(\mathbf{i}=\mathbf{i}'$ on $\sim\mathrm{A}_1)\rightarrow Val_{\mathbf{i}}(\sim\mathrm{A}_1) = Val_{\mathbf{i}'}(\sim\mathrm{A}_1)$	like 18 [$\sim\mathrm{A}_1$/A]
20	$(\mathrm{A})[(\mathbf{i}=\mathbf{i}'$ on A$) \rightarrow Val_{\mathbf{i}}(\mathrm{A})=Val_{\mathbf{i}'}(\mathrm{A})]$	1–6, 7–19, Ind. on sent

By **2B-4**, "(A)" here stands in for "(A)(A is a sentence →" or "(A)(A∈Sent →." See the further discussion after **2B-18** on p. 32.

The next theorem shows how the grammatical operation of substitution of a sentence for a TF-atom is related to semantic values. That is, if we regard a certain sentence as having arisen as a result of the substitution operation, then the theorem gives us a way of figuring out the semantic value of that result. In more detail, the upcoming theorem—the "semantics of substitution theorem"—says that we can calculate the value of a sentence ...B... containing a substituted sentence B in two stages: First calculate the value of B, and then calculate the value of ...p..., that is, the value of the sentence into which B was substituted, where p has the value given to B. Somewhat more precisely, let ...p... be a sentence with TF-atom p, and let ...B... result by putting B for all p. Given TF-interpretation **i**, let **i**′ be the (unique) TF-interpretation that is just like **i**, except that $\mathbf{i}'(\text{p}) = Val_{\mathbf{i}}(\text{B})$. Then $Val_{\mathbf{i}}(\ldots\text{B}\ldots) = Val_{\mathbf{i}'}(\ldots\text{p}\ldots)$.

It is worth observing, just to keep us on our toes, that although the substitution operation [B/p](A) has *three* arguments, the truth value of the result depends only on truth value calculations for A and for B, and does not depend on knowing what value was assigned to p. But observations aside, it is clear that our statement is not sufficiently precise to permit a proof-without-arm-waving; first, the dots won't do, and second, we need a sharper characterization of **i**′. The first problem is easily solved, since we already have a notation for substitution (**2C-1**): If C = ...p..., then ...B... is just [B/p](C). The second problem is solved by introducing some notation corresponding to the idea that we start with an interpretation **i**, and then shift it to **i**′ by giving p the truth value $Val_{\mathbf{i}}(\text{B})$. More generally, with **x** a truth value, we will define [**x**/p](**i**) as that TF-interpretation just like **i** except that the truth value **x** is given to p. Note that the following definition is formulated so as to give the notation [**x**/p] a separate meaning as a mapping from TF-interpretations into TF-interpretations.

2D-19 DEFINITION. *(Interpretation shift for* TF*)*

Where **x** is a truth value and p is a TF-atom,

$[\mathbf{x}/\text{p}] \in (\text{TF-interpretation} \mapsto \text{TF-interpretation})$ ([]**i**-type)

$([\mathbf{x}/\text{p}](\mathbf{i}))\text{p} = \mathbf{x}$ ([]**i**1)

$([\mathbf{x}/\text{p}](\mathbf{i}))\text{q} = \mathbf{i}_{(}\text{q})$ if $\text{q} \neq \text{p}$ ([]**i**2

This is inevitably hard to read. Observe especially that [**x**/p](**i**) is itself a function, which we know because it is said in ([]**i**-type) to be a member of TF-interpretation, and *all* members of TF-interpretation are functions. Accordingly, it makes metaphysical sense to apply [**x**/p](**i**) to an appropriate argument, namely, a TF-atom. Note further that a result such as ([**x**/p](**i**))q of applying such a function to such an argument has to be a truth value. Consequently, you can meaningfully put the expression, ([**x**/p](**i**))q, wherever you can meaningfully put any name of a truth value.

Furthermore, we must be able to use [**x**/p](**i**) with complex as well as simple names for truth values in place of "**x**"; the prime example is the $\mathbf{i}'$ of the motivating discussion preceding Definition **2D-19**, which will be [$Val_{\mathbf{i}}$(B)/p](**i**); note that this survives type-checking because $Val_{\mathbf{i}}(\mathrm{B}) \in \mathbf{2}$.

Recall that TF-interpretation = (TF-atom $\mapsto$ **2**). So [**x**/p] $\in$ ((TF-atom $\mapsto$ **2**) $\mapsto$ (TF-atom $\mapsto$ **2**)). Now think about [_/_] as itself a function. Then its place amid function spaces would have to be as follows: ((**2** $\times$ TF-atom) $\mapsto$ ((TF-atom $\mapsto$ **2**) $\mapsto$ (TF-atom $\mapsto$ **2**))). If you think about the matter in this way, you can come to see that the idea of the interpretation-shift operator is entirely general, having nothing special to do with TF-interpretations, TF-atoms or truth values. It is just a matter of shifting from one function to another by fixing one of the arguments and leaving all of the others alone.

It is well to keep in mind that we have also used the notation [_/_] for substitution. But although the substitution notation is related to the interpretation-shift notation—and in important ways—it is *not* an instance of the same general idea. In particular, [B/p] when considered as a function by itself maps sentences into sentences (it belongs to (Sent $\mapsto$ Sent)); it does not shift from one function to another. We keep informally straight by referring to the properties of substitution (**2C-1**) with "[]A," and to the properties of interpretation shift (**2D-19**) with "[]**i**."

In addition to having in mind some "algebraic combinations," you also want to be able to visualize the shifted interpretation [**x**/p](**i**). The picture is easier than the notation. Recall that we are speaking of functions **i** and [**x**/p](**i**), and that we are in full control of a function when we know what it does to each member of its domain of definition. So Picture **2D-20** tells it all:

2D-20 PICTURE. *(Interpretation shift for* TF*)*

TF-atoms	**i**	[**x**/p](**i**)
q	T	T
r	F	F
p	T	**x**
s	T	T

The left column pictures the domain of the two functions (we are pretending that TF-atom = {q, r, p, s}, all of which are distinct), while the middle column displays the assumed values of **i** (e.g., **i**(q) = T) and the right column displays the values of [**x**/p](**i**). You can literally *see* that [**x**/p](**i**) has exactly the same pattern as **i**, except at the argument p, and there it delivers the value **x**—whatever that value may be.

Here, then, is the

2D-21 THEOREM. *(Semantics of substitution theorem for* TF*)*

$Val_{\mathbf{i}}([B/p](A)) = Val_{[Val_{\mathbf{i}}(B)/p](\mathbf{i})}(A)$

Note that sentence substitution, **2C-1**, is used on the left and TF-interpretation shift, **2D-19**, on the right. The very point of the theorem is to relate the two concepts carried by these "square bracket" notations.

Exercise 17 *(Proof of semantics of substitution theorem)*

Provide a proof. Choose arbitrary **i**, B, and p, and then use Induction on sentences, **2B-18**, choosing Ψ(A) there as the theorem here. Be sure in the base case to choose a letter distinct from "p," say "q," so that you will be looking at

$$Val_{\mathbf{i}}([B/p](q)) = Val_{[Val_{\mathbf{i}}(B)/p](\mathbf{i})}(q)$$

and may then take up the two cases, p = q and p ≠ q. Similarly, in the inductive step do not choose letters that will become confused with the already chosen "B." Otherwise it is just a matter of appealing to the definitions of the terminology used in the statement of the theorem: **2C-1**, **2D-7**, and **2D-19**.

▷ . . . . ◁

Bizarre or unmemorable is doubtless how Theorem **2D-21** appears before you have given it the deep attention it deserves. You may come to increase your admiration for the complexities it summarizes when you see for yourself the smooth role it plays in the proof to follow.

A principal corollary of the semantics of substitution theorem for TF is the validation of familiar "replacement rules" such as the rule of Double Negation, from ...A... to infer ...∼∼A..., and conversely. In fact the matter is so important in applications that we promote the corollary to a full

2D-22 THEOREM. *(Semantics of replacement theorem for* TF*)*

$Val_{\mathbf{i}}$(A)$=$$Val_{\mathbf{i}}$(B) $\rightarrow$ $Val_{\mathbf{i}}$([A/p](C))$=$$Val_{\mathbf{i}}$([B/p](C))

For example, since $Val_{\mathbf{i}}$(A)$=$$Val_{\mathbf{i}}$(∼∼A), we can conclude that

$Val_{\mathbf{i}}$(...A...)$=$$Val_{\mathbf{i}}$(...∼∼A...)

by taking C as ...p... .

PROOF. Straightforward, using **2D-21**; and endlessly tedious without it. □

Observe that only one interpretation is mentioned in the statement of the theorem, whereas the proof we suggest makes essential reference to more than one interpretation. What do you make of that?

Evidently a version concerning the value of triple-bar sentences is also possible, when the triple bar is taken as a connective of TF, as in Definition **2B-6**.

Exercise 18 *(Basic semantic definitions for* TF*)*

Feel free to cite previous exercises as well as other items.

1. Type-checking. Use $Val_{\mathbf{i}}$type, **2D-7**, and modus ponens for function space, NAL:**9B-15**, to show that $Val_{\mathbf{i}}$atom, $Val_{\mathbf{i}}\supset$, and $Val_{\mathbf{i}}\sim$ of **2D-7** "make sense" in that the two sides of the equation anyhow belong to the same basic set. (You cannot of course show that these equations are "true" by these means. And on the other hand, you can always show that there is a set to which two entities both belong; so to make the exercise interesting, you must use judgment in interpreting the word "basic.")

2. More type-checking. All but one of the following do not make sense:

   (a) $\mathbf{i}(\sim p) = \sim^{*}\mathbf{i}(p)$

   (b) $Val_{\mathbf{i}}(\sim^{*} p)$

   (c) $p \supset p = T$

   (d) $Val_{\mathbf{i}}(A) = T$

   Explain. For example, even using extra parentheses, $([T/q](\mathbf{i}))(\sim p) = q$ does not make sense, because $[T/q](\mathbf{i})$ is a TF-interpretation by []**i**-type, **2D-19**, so that because TF-interpretation = (TF-atom ↦ **2**) by **2D-4**, to make sense its argument ~p would need to be a TF-atom. It isn't, by Distinctness, **2B-20**. And there is another problem: Even if the argument of $[T/q](\mathbf{i})$ were a TF-atom, its value when applied to that argument would be in **2**, whereas the right side of the equation denotes a TF-atom, not a truth value. (In fact we do not have a guarantee that no TF-atom is a truth value, and hence that the two sides of the equation denote the same thing, but in type-checking we look at the other side of the coin: It is the absence of a guarantee that the entities denoted by the two sides of an equation belong to the same basic type that leads us to sense the likelihood of a conceptual error.)

3. Suppose that p, q, and r are distinct TF-atoms. Describe or display a TF-interpretation that gives T to p and q, and gives F to r. Be sure that what you describe really is a TF-interpretation, **2D-4**. (But you don't have to *prove* that it is—we haven't developed techniques for this.)

4. Prove that $\mathbf{i}(p) = T$ implies $Val_{\mathbf{i}}(A \supset p) = T$.

5. Prove that $Val_{\mathbf{i}}(A \supset A) = T$. Use a previous exercise, Exercise 7.

6. Prove that for each **i**, every sentence is either true or false, but not both. Be *sure* to use the definitions rather than proceeding intuitively.

7. Prove that A is false on **i** ↔ ~A is true on **i**.

8. Prove that if A and A ⊃ B are true on **i**, so is B.

9. Learn the proof of Theorem **2D-18**

10. Important: Give a prose explanation of the point and content of each of the three fundamental theorems of §2D.3: **2D-18**, **2D-21**, and **2D-22**. You may wish to present them as not only true of our TF language, but as *desirable* features of a language.

11. Optional. Prove **2D-17**.
12. Formulate a triple-bar version of **2D-22** as mentioned just after that theorem, assuming that the triple bar is a connective of TF.

▷ . . . . ◁

2D.4 Quantifying over TF-interpretations

We want to give a sharp definition of an appropriate concept of logical consequence to the extent that it pertains to TF. We are going to be thinking of this relation—"tautological implication" is a usual name—as holding between a set of sentences (the premisses) on the one hand and a single sentence (the conclusion) on the other. Since not only here but elsewhere we shall be concerning ourselves with sets of sentences, it is convenient to introduce the following

2D-23 CONVENTION. *("G" and "H" for sets of sentences)*

We use "G" and "H" (sometimes with marks) as variables ranging over $\mathcal{P}$(Sent) (see NAL:**9B-20**). That is, "G ⊆ Sent" and "H ⊆ Sent" are for free.

It is customary to use the double turnstile, "⊨," for the relation of logical consequence; we subscript it with a "TF" to remind us that tautological implication has competitors as an analysis of "good argument," and that in any event even classically tautological implication does not capture the full intent of logical consequence (think of quantifiers). Other than that, you will notice that the double turnstile just signals truth-preservation: Whenever all the premisses are true, so is the conclusion—with the metaphorical quantifier "whenever" cashed out in terms of quantification over TF-interpretations (equals rows of a truth table).

2D-24 DEFINITION. *(Tautological implication: $\vDash_{TF}$)*

$\vDash_{TF} \subseteq (\mathcal{P}(\text{sentences}) \times \text{sentences})$ ($\vDash_{TF}$type)

G $\vDash_{TF}$ A ↔ for all TF-interpretations $\mathbf{i}$, if (B)(B ∈ G → $Val_{\mathbf{i}}$(B) = T) then $Val_{\mathbf{i}}$(A) = T

In words: $\vDash_{TF}$ is a relation between sets of sentences and (single) sentences such that G $\vDash_{TF}$ A ↔ every TF-interpretation that makes every member of G true also makes A true. Carefully observe that whereas the scope of the quantifier "for all TF-interpretations **i**" is the entire conditional, in contrast the scope of the quantifier "(B)" is restricted to the antecedent of the body of the definiens.

2D-25 VARIANT. *(Variants of $\vDash_{\mathrm{TF}}$)*

As variants of "G $\vDash_{\mathrm{TF}}$ A" we can use "G tautologically implies A" or "A is a semantic consequence of G" or "G implies A in TF" or "G implies A" or "the argument from G to A is truth functionally valid" (or just "valid").

These variants raise the possibility of a philosophical error. Observe that the Definition **2D-24** is itself purely mathematical, a matter of formal semantics. It is an *application* to hook up this mathematical concept with the intuitive and (note) normative notion of a truth functionally valid argument, an application which may or may not be correct—it is a matter of philosophical discussion, not mathematics. It is therefore an error to suppose that the philosophical discussion is brought to a close by the definition or its variants.

As a matter of fact we doubt the final correctness of $\vDash_{\mathrm{TF}}$ as an account of valid truth functional argument, but this is not the place for a minority report. We are certain $\vDash_{\mathrm{TF}}$ is an interesting and sensible preliminary *candidate* for an account of valid truth functional argument, and in the rest of these notes we shall proceed accordingly.

It should be clear from **2D-10** that "$\vDash_{\mathrm{TF}}$" could have been defined by quantification over TF-valuations instead of TF-interpretations:

2D-26 COROLLARY. *($\vDash_{\mathrm{TF}}$ via* TF-*valuations)*

G $\vDash_{\mathrm{TF}}$ A $\leftrightarrow$ for all TF-valuations f, if f(B) = T for all B in G, then f(A) = T.

The following convention permits a somewhat simpler notation.

2D-27 CONVENTION. *(Dropping "$\cup$" and "$\{\}$" and "$\varnothing$")*

On the left of "$\vDash_{\mathrm{TF}}$," it is often convenient to use a comma in place of the union sign "$\cup$," to drop curly brackets around sentence-terms, and to drop the sign, "$\varnothing$," for the empty set. Thus

$$\mathrm{G, A} \vDash_{\mathrm{TF}} \mathrm{B} \leftrightarrow \mathrm{G} \cup \{\mathrm{A}\} \vDash_{\mathrm{TF}} \mathrm{B}$$

$$\mathrm{A} \vDash_{\mathrm{TF}} \mathrm{B} \leftrightarrow \{\mathrm{A}\} \vDash_{\mathrm{TF}} \mathrm{B}$$

$$\mathrm{A, B, C, D} \vDash_{\mathrm{TF}} \mathrm{E} \leftrightarrow \{\mathrm{A, B, C, D}\} \vDash_{\mathrm{TF}} \mathrm{E}$$

$$\vDash_{\mathrm{TF}} \mathrm{A} \leftrightarrow \varnothing \vDash_{\mathrm{TF}} \mathrm{A}$$

This convention applies *only* on the left of turnstiles; as a matter of practice it turns out to be confusing to try to use it elsewhere.

Exercise 19 *(Variants of what?)*

Of what would the following be variants: "A is a tautology"; "A tautologically implies B"; "A is tautologically equivalent to B"; "G is truth functionally inconsistent"?

▷ . . . . ◁

There are three important "structural" properties of $\vDash_{TF}$, where by calling them "structural" we mean that they describe the concept of logical consequence itself, without mention of connectives or other grammatical features of TF.

2D-28 FACT. *(Structural $\vDash_{TF}$ facts)*

Identity. $A \in G \rightarrow G \vDash_{TF} A$. ($\vDash_{TF}$id)

Weakening. $(G \subseteq H$ and $G \vDash_{TF} A) \rightarrow H \vDash_{TF} A$. ($\vDash_{TF}$weak)

Cut. $(G \vDash_{TF} A$ and $G, A \vDash_{TF} B) \rightarrow G \vDash_{TF} B$. ($\vDash_{TF}$cut)

These facts, though easy to prove, as we shall soon see, are important.

- *Identity* tells us that if your conclusion is among your premisses, you cannot go wrong: Circular arguments are paradigmatically valid, even if to stutter is not to reason. Consideration of this principle reminds us that the classical logician does not presume to tell us how to reason correctly, but only how to avoid making a fresh mistake ("safety first," as Alan Ross Anderson used to say).

- *Weakening* says that an argument remains valid if you add extra premisses. The confusing thing is this: By "strengthening" your set of premisses, you "weaken" your $\vDash_{TF}$ statement. To say that C follows from A and B together is weaker than to say that C follows from A alone.

- *Cut* spells out that if one of your premisses, A, follows from the rest, G, then the premiss A is redundant and may be cut out

It is worth recording for future reference (from **3C-21**) that Identity, Weakening, and Cut depend only on the form of Definition **2D-24**, and not at all on either the set, TF-interpretation (**2D-4**), or the definition **2D-7** of the TF-valuation $Val_{\mathbf{i}}$ determined by $\mathbf{i}$. (Recall that a sentential valuation is any mapping whatsoever from the sentences into **2**.)

2D-29 LEMMA. *(Identity, Weakening, and Cut abstractly)*

Let Val be an *arbitrary* subset of the sentential valuations (Sent$\mapsto$**2**), **2D-1**, not necessarily TF-valuations. Suppose that $\vDash$ is a relation on ($\mathcal{P}$(Sent)$\times$Sent) such that G$\vDash$A if and only if, for all functions f$\in Val$, if (B)(B$\in$G $\rightarrow$ f(B)$=$T) then f(A)$=$T. Then Identity, Weakening, and Cut hold for$\vDash$.

PROOF. We are given that G$\vDash$A $\leftrightarrow$ for all f in Val, (B)(B$\in$G $\rightarrow$ f(B)$=$T) $\rightarrow$ f(A)$=$T; call this "the equivalence for$\vDash$," and its right side "the definiens."

Identity. Suppose A$\in$G, and choose f$\in Val$. Then the antecedent of the body of the definiens trivially leads to its consequent.

Weakening. Suppose that *(a)* G is a subset of H and that *(b)* G$\vDash$A; choose f in Val, and suppose f(B)$=$T for every member B of H. Hence also f(B)$=$T for every member B of G by *(a)*, so f(A)$=$T by *(b)*. The conclusion follows by the equivalence for $\vDash$.

Cut. See below, Exercise 20. □

Exercise 20 *(Cut abstractly)*

Provide the proof for Cut. Be careful about your quantifiers, especially the scope of "(B)" in the definiens. Expect to use definitions governing the symbols for union and for unit set, and to use the logic of identity.

▷ . . . . ◁

Returning now to the proof of Fact **2D-28**, it is enough to observe that Definition **2D-24** has the right form to render it a corollary of **2D-29**, characterizing Val there by the condition, f$\in Val$ $\leftrightarrow$ for some TF-interpretation $\mathbf{i}$, f$=Val_{\mathbf{i}}$.

There is a fourth structural fact that in contrast does depend on both what counts as a TF-interpretation and the details of the definition of the TF-valuation $Val_{\mathbf{i}}$, a fact that we record here without proving; indeed, its proof is of substantial difficulty, and we postpone it. (You therefore cannot at this point use it in proving later theorems.)

2D-30 CONJECTURE. *(Finiteness for* $\vDash_{TF}$*)*

Finiteness. If $G \vDash_{TF} A$, then there is a finite subset G_0 of G such that $G_0 \vDash_{TF} A$. ($\vDash_{TF}$fin)

This is a deep and important principle: If A is a tautological consequence of a whole infinite bag of premisses, then it is a tautological consequence of some finite subgroup of them. It follows that we do not have to countenance truth-functional rules with infinitely many premisses, for any such proposed rule can without loss be replaced by another with only finitely many premisses.

You should pay special attention to the four structural principles just stated; there are no others to learn, as all structural principles follow from these—this is a Fact (advanced type).

Next come several facts that show how each connective interacts with $\vDash_{TF}$, when it is on the right, and when it is on the left. Even though we shall not need to refer to these later, they are of fundamental importance in understanding the role of the various connectives in inference.

2D-31 FACT. *(Connectives and* $\vDash_{TF}$*)*

$G \vDash_{TF} A \supset B \leftrightarrow G, A \vDash_{TF} B$ ($\vDash_{TF}\supset$)

$G, A \supset B \vDash_{TF} C \leftrightarrow (G, \sim A \vDash_{TF} C$ and $G, B \vDash_{TF} C)$ ($\supset\vDash_{TF}$)

$G \vDash_{TF} \sim A \leftrightarrow G, A \vDash_{TF} \sim(B \supset B)$ ($\vDash_{TF}\sim$)

$G, \sim A \vDash_{TF} \sim(B \supset B) \leftrightarrow G \vDash_{TF} A$ ($\sim\vDash_{TF}$)

$G \vDash_{TF} A\&B \leftrightarrow (G \vDash_{TF} A$ and $G \vDash_{TF} B)$ ($\vDash_{TF}\&$)

$G, A\&B \vDash_{TF} C \leftrightarrow G, A, B \vDash_{TF} C$ ($\&\vDash_{TF}$)

$G \vDash_{TF} A \vee B \leftrightarrow G, \sim A \vDash_{TF} B$ ($\vDash_{TF}\vee$)

$G, A \vee B \vDash_{TF} C \leftrightarrow (G, A \vDash_{TF} C$ and $G, B \vDash_{TF} C)$ ($\vee\vDash_{TF}$)

We are using $\sim(B \supset B)$ to play in TF something of the role that "$\perp$" plays in our use-language, as was more explicit in the definition of $\perp(B)$, **2B-6**.

PROOF. Straightforward, using the definition of $Val_{\mathbf{i}}$, **2D-7**. Most people find that it is easier to show the two sides equivalent by showing their negations equivalent □

Here are some more facts you should know.

2D-32 COROLLARY. *(More $\vDash_{TF}$ facts)*

---

1. $A \vDash_{TF} A$
2. $(G \vDash_{TF} A$ and $A \vDash_{TF} B) \rightarrow G \vDash_{TF} B$
3. $(G \vDash_{TF} A \supset B$ and $G \vDash_{TF} A) \rightarrow G \vDash_{TF} B$
4. $\vDash_{TF} A \rightarrow G \vDash_{TF} A$
5. $A\&B \vDash_{TF} A$; $A\&B \vDash_{TF} B$
6. $A \vDash_{TF} A \vee B$; $B \vDash_{TF} A \vee B$
7. $\vDash_{TF} A \vee {\sim}A$

PROOF. Straightforward, without directly using the definition of $Val_{\mathbf{i}}$, or even of $\vDash_{TF}$, but instead relying on the already established Facts **2D-28** and **2D-31** □

The following are easy to prove, so easy that we make a point of emphasizing their importance for some of our later work.

2D-33 THEOREM. *(Three tautologies)*

---

1. $G \vDash_{TF} A \supset (B \supset A)$
2. $G \vDash_{TF} (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$
3. $G \vDash_{TF} ({\sim}A \supset {\sim}B) \supset (B \supset A)$

PROOF. We give two separate proofs. First proof, using the definition of $\vDash_{TF}$ and $Val_{\mathbf{i}}$. We illustrate with 1. Choose $\mathbf{i}$, and assume all members of G are true on $\mathbf{i}$ (this assumption will in fact not be used). By two uses of ($Val_{\mathbf{i}}\supset$), **2D-7**,

$$Val_{\mathrm{i}}(\mathrm{A} \supset (\mathrm{B} \supset \mathrm{A})) = Val_{\mathrm{i}}(\mathrm{A}) \supset^* (Val_{\mathrm{i}}(\mathrm{B}) \supset^* Val_{\mathrm{i}}(\mathrm{A})).$$

By Val_{i}type, **2D-7**, Val_{i}(A) and Val_{i}(B) are in **2** (**2D-8**). So by T1 of Theorem **2A-12** on p. 15, $Val_{\mathrm{i}}(\mathrm{A} \supset (\mathrm{B} \supset \mathrm{A})) = \mathrm{T}$, which was to be shown. The proofs for 2 and 3 are exactly similar.

Second proof, using **2D-28** and **2D-31**. Again we consider only (1). G, A, B $\vDash_{\mathrm{TF}}$ A by $\vDash_{\mathrm{TF}}$id (**2D-28**) and easy set theory; so (1) by $\vDash_{\mathrm{TF}}\supset$ (**2D-31**) twice. □

It is important, on the one hand, to notice that the three previous statements concerning the connection between $\vDash_{\mathrm{TF}}$ and connectives would *not* hold were $\vDash_{\mathrm{TF}}$ defined in terms of arbitrary sentential valuations; in contrast to Lemma **2D-29**, the restriction to TF-valuations is essential.

Exercise 21 *(Three tautologies and arbitrary sentential valuations)*

Show this informally; it's easy. That is, for each statement in Theorem **2D-33** (or for all three together, your choice) describe a sentential valuation and a set G (the empty set will do) such that on that sentential valuation, each of the displayed formulas would turn out to be false.

▷ . . . . ◁

On the other hand, we discover in **3C-21** on p. 171 that it is also important to notice that these three statements *would* continue to hold were $\vDash_{\mathrm{TF}}$ defined in terms of any subset of the TF-valuations whatsoever; that $\vDash_{\mathrm{TF}}$ refers to *all* TF-valuations (via all TF-interpretations) is not relevant to these properties:

2D-34 FACT. *(Connectives and $\vDash_{\mathrm{TF}}$ in an abstract setting)*

Let Val be an *arbitrary* subset of the TF-valuations, **2D-3**. Suppose that $\vDash$ is a subset of ($\mathcal{P}$(Sent) × Sent) such that G $\vDash$ A if and only if, for all functions f ∈ Val, if (B)(B ∈ G → f(B) = T) then f(A) = T. Then analogs of all of **2D-31**, **2D-32**, and **2D-33** hold for $\vDash$.

PROOF. Tedious. The guiding principle is that proof of each of these statements involves consideration of only one TF-interpretation (hence one TF-valuation) at a time. For example, consider modus ponens, Corollary **2D-32**(3). Make the same beginning as in the proof of **2D-29**. Now suppose G $\vDash$ A$\supset$B and G $\vDash$ A, and choose an arbitrary member f of Val making *(a)* every member of G true; we need to show *(b)* f(B)=T. By *(a)* and two uses of the equivalence for $\vDash$, f(A$\supset$B)= T and f(A)=T; but then *(b)* as required, because f is known to be a TF-valuation, **2D-3** □

It is "intuitively obvious" that tautologyhood is preserved under substitution: Every substitution-instance of a tautology is again a tautology; in fact, many logical systems utilize the "rule of substitution" as a device with which to generate their "theorems." It is therefore good that we justify this procedure.

2D-35 FACT. *(Preservation of tautologyhood by substitution)*

$\vDash_{\text{TF}}$ A $\rightarrow \vDash_{\text{TF}}$ [B/p](A)

PROOF. Since the fact is semantic and employs the concept of substitution in its statement, it is not surprising that the semantics of substitution theorem **2D-21** will do the work required. Begin by assuming $Val_{\mathbf{i}}$(A)=T for all TF-interpretations **i**. Choose a TF-interpretation **i**′; we need to show that $Val_{\mathbf{i}'}$([B/p](A))=T. Use the notation defined in **2D-19** to choose **i** in the assumption as the TF-interpretation [$Val_{\mathbf{i}'}$(B)/p]($\mathbf{i}'$), whence

$$Val_{[Val_{\mathbf{i}'}(\text{B})/\text{p}](\mathbf{i}')}(\text{A})=\text{T}.$$

But by the semantics of substitution theorem **2D-21**, this implies that

$$Val_{\mathbf{i}'}([\text{B/p}](\text{A}))=\text{T},$$

which was to be shown □

Exercise 22 *(More semantics for* TF*)*

1. Learn the two proofs of **2D-33**.

2. Prove Fact **2D-28**. Be sure to use the definitions. But you will note that you need only the definition of "$\vDash_{\mathrm{TF}}$," not of "Val_{i}."

3. Prove $\vDash_{\mathrm{TF}}\supset$, for example, of Fact **2D-31**. You will need the definitions of both "$\vDash_{\mathrm{TF}}$" and "Val_{i}."

4. Prove (6), for example, of Corollary **2D-32**. Previous facts should suffice, without recourse to definitions.

5. Optional. To think. What about a proof of **2D-35** that does *not* rely on the semantics of substitution theorem **2D-21**?

▷ . . . . ◁

2D.5 Expressive powers of TF

Any time one is considering a certain grammar with an attendant semantics, the question of "expressive power" arises: What can be expressed by the grammatical-semantic means at hand? Often this question-type is of considerable philosophical interest, such as the question of whether the language of set theory is expressively adequate for all of mathematics. Here, however, we raise a much less interesting case of this question-type, with an eye on setting a sharp question with a sharp answer as an intellectual foundation for more difficult cases.

What is it possible to say with the means provided by TF? In thinking about this question, it is good to begin by asking what the sentences of TF express. One might answer that they express "truth functions," or one might instead answer "truth tables." The intuitive difference between the two concepts lies in what each takes as inputs:

A truth function has as its inputs *appropriate sequences of truth values*, and tells you for each input what the output truth value is to be.

A truth table has as its input *assignments of truth values to the* TF*-atoms* (a "row," as we might say), and tells you for each input what the output truth value is to be.

You can see from the above that there is this sharp difference: The idea of a truth function is entirely *language-independent*, whereas the idea of a truth table makes essential appeal to the notion of a TF-atom, which is linguistic. For example, what TF-atoms there are determine what truth tables there are, but not what truth functions there are.

We choose to discuss the expressive powers of TF in terms of truth tables instead of truth functions.[13] Such choices always carry costs as well as benefits on both sides; in this case, the treatment in terms of truth tables is noticeably simpler but not as deep as would be a similar discussion in terms of the language-free concept of truth function.

First we enter several definitions that we will need for our discussion, beginning with a sharp definition of "TF-truth-table," where we add "TF-" to remind us of the aforementioned linguistic dependence. After giving these definitions, we comment on each.

2D-36 DEFINITION. *(*TF-*truth-table, based on, expresses)*

TF-truth-table = (TF-interpretation $\mapsto$ **2**) (TF-truth-table)

t ranges over TF-truth-tables.

For $\mathbf{i}_1$, $\mathbf{i}_2 \in$ TF-interpretation, and X a set of TF-atoms, $\mathbf{i}_1$ *agrees with* $\mathbf{i}_2$ *on* X $\leftrightarrow$ for each TF-atom, p, if p $\in$ X then $\mathbf{i}_1(\mathrm{p}) = \mathbf{i}_2(\mathrm{p})$. (Agrees with on)

For *t* a TF-truth-table and X a set of TF-atoms, *t is based on* X (or X *is a basis for t*) $\leftrightarrow$ for all TF-interpretations $\mathbf{i}_1$ and $\mathbf{i}_2$, if $\mathbf{i}_1$ and $\mathbf{i}_2$ agree on X, then $t\mathbf{i}_1 = t\mathbf{i}_2$. (Based on)

For sentence A and TF-truth-table *t*, A *expresses t* $\leftrightarrow$ for each TF-interpretation $\mathbf{i}$, $Val_{\mathbf{i}}(\mathrm{A}) = t\mathbf{i}$. (Expresses)

The definition of "TF-truth-table" evidently correctly catches the idea we put intuitively above. As for "agrees on," this is exactly the same idea that played a critical role in the local determination theorem, **2D-18**; here we are only making it official. To say that a truth table is "based on" a set X of TF-atoms is to say that it is only the values of the TF-atoms in X that matter; the values of others are irrelevant. For example, if a TF-truth-table is based on the empty set, then that TF-truth-table is a constant function, delivering either always T or always F. We are going to be interested in TF-truth-tables that have *finite* bases, where "finite" comes from **2B-7**. Finally, the "expresses" relation is obviously of immediate intuitive significance:

[13] N. Tennant pointed out to us how much simpler the development can be if carried out in those terms.

A sentence expresses a TF-truth-table if sentence and table always deliver the same truth value for the same TF-interpretation.

As an illustration at this point of the difference between truth functions and truth tables, note that whereas on the one hand, to establish the sense in which the two sentences $p \supset q$ and $q \supset p$ "express" different truth *functions* is going to require a method for keeping track of the *order* of the TF-atoms p and q, on the other hand we can see immediately and without paying attention to order that the two sentences express different TF-truth-*tables* (provided $p \neq q$).

There are two central theorems about the expressive powers of TF, both of which involve the concept of finitude, **2B-7**, as well as those concepts defined just above.

2D-37 THEOREM. *(Expressive limits of* TF*)*

If *t* is a TF-truth-table expressed by a sentence, A, of TF, then *t* is based on a *finite* set of TF-atoms.

Which is to say this: Even though there are such things as TF-truth-tables that are properly infinite in the sense that they are not based on any finite set of TF-atoms, no such table can be expressed by any sentence of TF.[14]

PROOF. This theorem is a corollary of facts that we have already established. The local determination theorem, **2D-18**, tells us that any TF-truth-table expressed by A is based on the set of TF-atoms that are subformulas of A (i.e., is based on $\text{Subf}(A) \cap \text{TF-atom}$); and Finitude of TF-atoms, **2B-25**, tells us that this set is finite, which concludes the proof.

□

The other theorem is the converse of the first.

2D-38 THEOREM. *(Truth-tabular expressive completeness of* TF*)*

Every TF-truth-table *t* based on a *finite* set of TF-atoms is expressed by some sentence, A, of TF.

[14] There is no reason that we should not make up a theory of expression of a truth table by means of an *infinite* set of sentences of TF; but here we decline the invitation.

One subtlety needs to be noted: In spite of the apparent strength of our statement of truth-tabular expressive completeness, the fact that the notion of TF-truth-table is language-dependent renders the interest of this sort of expressive completeness equally dependent on language. For instance, TF can in the stated sense be expressively complete with respect to TF-truth-tables even if there is only one TF-atom! Reason: If there are very few TF-atoms then there are very few TF-truth-tables that need to be expressed for completeness to hold, so that completeness is easy. When, however, there are or might be infinitely many TF-atoms (recalling that **2C-14** does not say one way or the other), the theorem is powerful. To prove it we shall need to introduce the following concept of truth-table shift.

2D-39 DEFINITION. *(Truth-table shift)*

Let $[\mathbf{x}/\mathrm{p}] \in (\text{TF-interpretation} \mapsto \text{TF-interpretation})$ shift TF-interpretations in according with **2D-19**. Let t be a TF-truth-table.

$t \circ [\mathbf{x}/\mathrm{p}]$ is a TF-truth-table (Truth-table-shift type)

$(t \circ [\mathbf{x}/\mathrm{p}])\mathbf{i} = t([\mathbf{x}/\mathrm{p}](\mathbf{i}))$. (Truth-table shift)

The beauty of $t \circ [\mathbf{x}/\mathrm{p}]$ is that it is guaranteed to be a TF-truth-table whose value for any argument $\mathbf{i}$ is *not* dependent on the value that $\mathbf{i}$ gives to p.

It is worth observing that the notation is consonant with that for "composition of functions" as in NAL:Exercise 66. The general idea is that when you have two functions f and g with matching function spaces ($\mathrm{f} \in (\mathrm{Y} \mapsto \mathrm{Z})$, $\mathrm{g} \in (\mathrm{X} \mapsto \mathrm{Y})$), then their composition $\mathrm{f} \circ \mathrm{g}$ makes sense as a function in $(\mathrm{X} \mapsto \mathrm{Z})$ if we define $(\mathrm{f} \circ \mathrm{g})\mathrm{x} = \mathrm{f}(\mathrm{gx})$ for every $\mathrm{x} \in \mathrm{X}$. Truth-table shift, as you can see from its definition, is just a special case.

In the proof of Theorem **2D-38** we shall need the following small facts.

2D-40 MINIFACT. *(Facts for expressive completeness)*

---

1. For all TF-interpretations $\mathbf{i}_1$ and $\mathbf{i}_2$, $\mathbf{i}_1$ and $\mathbf{i}_2$ agree on $\varnothing$. (This is vacuously true: $\mathbf{i}_1$ and $\mathbf{i}_2$ agree on everything in $\varnothing$ because there isn't anything in $\varnothing$.)
2. If a TF-truth-table t is based on $\varnothing$, then for every $\mathbf{i}_1$ and $\mathbf{i}_2$, $t\mathbf{i}_1 = t\mathbf{i}_2$.
3. Let X be a set of TF-atoms, let p be a TF-atom, and let $\mathbf{x}$ be a truth value. If a TF-truth-table t is based on $\mathrm{X} \cup \{\mathrm{p}\}$, then the shifted TF-truth-table $t \circ [\mathbf{x}/\mathrm{p}]$ is based on X.

4. Let $\mathbf{i}$ be a TF-interpretation and $\mathbf{x}$ a truth value. If $\mathbf{i}(p) = \mathbf{x}$, then $[\mathbf{x}/p](\mathbf{i}) = \mathbf{i}$. (If the value of p is already $\mathbf{x}$, then fixing it to $\mathbf{x}$ does not change things.)

5. Suppose A expresses the shifted TF-truth-table $t \circ [\mathbf{x}/p]$. Then for any TF-interpretation $\mathbf{i}$, if $\mathbf{i}(p) = \mathbf{x}$, then $Val_{\mathbf{i}}(A) = t\mathbf{i}$.

PROOF. Straightforward—though unfamiliar notation may confuse you for a while, especially as to what is function and what is argument. We will not in fact be needing either **2D-40**(1) or **2D-40**(4) separately; we enter the former only to smooth the proof of **2D-40**(2), and the latter only for **2D-40**(5). Here is a sketch of the proof of **2D-40**(3).

Suppose *(a)* t is based on $X \cup \{p\}$; we need to show that *(z)* $t \circ [\mathbf{x}/p]$ is based on X. For *(z)*, let *(b)* $\mathbf{i}_1$ and $\mathbf{i}_2$ agree on X; we need to show that *(y)* $(t \circ [\mathbf{x}/p])\mathbf{i}_1 = (t \circ [\mathbf{x}/p])\mathbf{i}_2$, which by Definition **2D-39** requires showing that *(x)* $t([\mathbf{x}/p](\mathbf{i}_1)) = t([\mathbf{x}/p](\mathbf{i}_2))$. But *(x)* will follow from *(a)* once we establish that *(w)* $[\mathbf{x}/p](\mathbf{i}_1)$ and $[\mathbf{x}/p](\mathbf{i}_2)$ agree on $X \cup \{p\}$, so suppose that *(c)* $q \in (X \cup \{p\}$. In the only (easy but) nonobvious move in this proof, we infer from *(c)* that *(d)* either ($q \in X$ and $q \neq p$) or $q = p$. In the former case of *(d)* we can use *(b)* and $\mathbf{i}$[]2, **2D-19**, and in the latter just $\mathbf{i}$[]1, **2D-19**, to conclude that $([\mathbf{x}/p](\mathbf{i}_1))q = ([\mathbf{x}/p](\mathbf{i}_2))q$, which is what is required for *(w)* □

We break for some familiarization exercises.

Exercise 23 *(TF-truth-values)*

1. Type-check the least transparent among the definitions and equations so far given in this section, using MP for function space, NAL:**9B-15**. (Consult Exercise 18 for some discussion of type-checking.)

2. Imagine that there are just four TF-atoms, p, q, r, and s. Draw a picture of a TF-truth-table based on {p, q}, recalling that the picture of a TF-interpretation, $\mathbf{i}$, is the reference-columns part of a row.

3. Draw a picture of a TF-truth-table that is *not* based on any finite set of TF-atoms. Also describe your table in words.

4. Draw a picture of a TF-truth-table being expressed by some sentence of your choice.

5. Draw a picture of a TF-truth-table t and its two "shifts," $t\circ$[T/p] and $t\circ$[F/p].

6. Think about the proofs of the various items under Minifact **2D-40**.

▷ . . . . ◁

We return now to the proof of Theorem **2D-38**.

PROOF. Induction on finite sets, **2B-9**, provides the outer structure of the proof, and accordingly we rearrange what we wish to prove into the following equivalent form: For every finite set X, if X$\subseteq$TF-atom then for every TF-truth-table t, if t is based on X then there is a sentence A that expresses t. (The important tinker consists in placing t as a *bound* variable inside the scope of the induction.)

For the base case, when X is the empty set, we are supposing that *(a)* t is based on $\varnothing$. There are two cases: Either $t\mathbf{i}=$T for some TF-interpretation $\mathbf{i}$, or not. The second is easiest, so we carry out only the first. Let *(b)* $t\mathbf{i}=$T. We are going to show that t is the constant function T. To show that t is the constant function T, let $\mathbf{i}'$ be an arbitrary TF-interpretation; but then *(c)* $t\mathbf{i}'=t\mathbf{i}$ from *(a)* by **2D-40**(2), so that $t\mathbf{i}'$ =T from *(b)* and *(c)*, as required. Now we need to rely on the existence of at least one TF-atom, **2C-14**, say p, so that we can be sure that p$\supset$p is "some sentence." It is a matter of brief reference to Tr$\supset$, **2D-13**, to conclude that this sentence expresses the constant function T, which is now our desired conclusion.

For the inductive step, we suppose as inductive hypothesis *(a)* that if X$\subseteq$TF-atom, then for every t, if t is based on X, then some A expresses t. As "step hypothesis" suppose (for arbitrary p) that *(b)* (X$\cup\{$p$\}$)$\subseteq$TF-atom, and (for arbitrary t) that *(c)* t is based on X$\cup\{$p$\}$. We need to show that some sentence expresses t. In order to find such a sentence, we first observe that *(d)* each of the two shifted TF-truth-tables $t\circ$[T/p] and $t\circ$[F/p] are based on X; this follows from *(c)* and **2D-40**(3). Hence, by the inductive hypothesis *(a)*, it must be that *(e)* there are sentences A_1 and A_2 such that A_1 expresses $t\circ$[T/p] and A_2 expresses $t\circ$[F/p], and we choose one of each. We now for mere convenience *(f)* let Ch[15] be the sentence (p$\supset A_1$)&($\sim$p$\supset A_2$), and we propose to show that Ch expresses the TF-truth-table t. To show this requires that $Val_{\mathbf{i}}$(Ch)$=t\mathbf{i}$ for arbitrary $\mathbf{i}$, so choose $\mathbf{i}$. There are now two cases: $\mathbf{i}$(p)=T or $\mathbf{i}$(p)=F. They are exactly parallel, and so we undertake only the former. Suppose, then, that *(g)* $\mathbf{i}$(p)=T. Then *(h)* $Val_{\mathbf{i}}(A_1)=t\mathbf{i}$ from *(e)* and *(g)* by **2D-40**(5). But it is easy to see from *(g)* and Tr$\supset$ and Tr$\sim$, **2D-13**, and of course from the local definition *(f)* of Ch, that *(i)* $Val_{\mathbf{i}}$(Ch)$=Val_{\mathbf{i}}(A_1)$, so that we may conclude from *(h)* and *(i)* that Ch expresses t, as desired. □

[15]In honor of Church, to whom this argument is due.

The key step in understanding this proof lies in becoming clear that we may choose the expressing sentence Ch *after* we have A_1 and A_2 in hand, but that we must choose it *before* we are given the cases $\mathbf{i}(p)=T$ or $\mathbf{i}(p)=F$.

The above theorems about the expressive powers of TF are made more interesting by seeing what would happen if we tinkered with TF in various ways. For example, suppose we consider a language TF′ whose *only* connective is ≡, with a valuation clause such that $Val_{\mathbf{i}}(A\equiv B)=T$ just in case $Val_{\mathbf{i}}(A)=Val_{\mathbf{i}}(B)$. Then it is easy to show that TF′ is *not* expressively complete with respect to truth tables, at least as long as there is at least one TF′-atom, because, where p is some TF′-atom, no sentence of TF′ can express either of the truth tables t_1 or t_2, where t_1 is that truth table such that for every TF′-interpretation $\mathbf{i}$, $t_1\mathbf{i}=T$ iff $\mathbf{i}(p)=F$ (that is, t_1 corresponds to the negation of p), and where t_2 is that truth table such that for every TF′-interpretation $\mathbf{i}$, $t_2\mathbf{i}=F$ (that is, t_2 corresponds to constant falsity). We could show this failure of expressive completeness by induction on sentences of TF′, in analogy with **2B-18**.[16] In the base case we would choose an arbitrary TF′-atom q, and consider two cases, $q=p$ and $q\neq p$, showing that in neither case does q express either t_1 or t_2. Then under the inductive hypothesis that neither A_1 nor A_2 expresses either t_1 or t_2, we would show that $A_1\equiv A_2$ equally fails to express either t_1 or t_2, which would complete the argument that no sentence of TF′ expresses either of the truth tables t_1 or t_2 described above.

There are many similar facts; for a good survey consult Hunter 1971, §21.

Exercise 24 *(Grammar and semantics)*

1. Define consistency of a set of sentences. Describe its chief structural properties (of which it has very few), and its chief properties involving reference to connectives.

2. Invent a three valued logic to be called "TNF" ("N" for "neutral"):

   (a) Offer a theory of the truth values, **3**.

   (b) Use the same grammar (TF-atoms, ⊃, ∼) as §2B.2, but add conjunction and disjunction connectives & and ∨.

[16] At this point we would certainly wish we had a general theory of grammar instead of a theory about just the grammar of TF.

(c) Provide a semantics—at least a definition of "$Val_{\mathfrak{i}}$" and "$\vDash_{\text{TNF}}$." You will do best to think about *alternative* possibilities for $\vDash_{\text{TNF}}$, and (if you are a *philosophical* logician) what they might mean.

(d) Establish a few facts.

▷ . . . . ◁

2E Elementary proof theory of S_{TF}

So far we have a rigorous account of part of the grammar and semantics of the language TF. Now we add some proof theory, giving us a formal theory we call "S_{TF}."

To put proof theory in perspective, remind yourself of the Logician's Task: to separate the good arguments from the bad. And recall that semantics in fact accomplishes that task at the level of meaning, for our intention was that the relation $\vDash_{\text{TF}}$ of logical consequence should represent our account of "good" arguments: An argument from G to A is "good" just in case $G \vDash_{\text{TF}} A$, at least insofar as its goodness is claimed to depend on the structure conferred by the grammar of TF. But although $\vDash_{\text{TF}}$ gives us an account, definition, or perhaps analysis of good argument, from its form we cannot immediately extract a humanly usable *criterion* of good argument. Its form, that is, does not inevitably enable us to recognize a good argument when we see one, or a bad one either. It is the job of applied proof theory to provide such a criterion, if it can.

Just for local purposes, and without the aim of introducing permanent jargon, it is useful to mark with Terminology that criteria (for a given feature) are of three kinds. (1) "Decision criteria" permit humans to settle the question as to whether a candidate does or does not have the feature. (2) "Certification criteria" permit humans to be sure a candidate has the feature, if in fact it does (but may give no information concerning candidates lacking the feature). (3) "Decertification criteria" are just certification criteria for the negative or complement of a feature: They permit humans to be sure a candidate does not have the feature, if in fact it does not (but may give no information concerning candidates having the feature). It is clear that if we have both certification and a decertification criteria for a feature, then we have a decision criterion for it, and conversely.

In the most fortunate cases, applied proof theory will provide a decision criterion for good arguments, or equivalently, both a certification criterion and a decertification criterion. It is possible to have such a pair for a suitable concrete version of

TF, but in fact we will be dealing only with half the story: a criterion for certifying the good arguments as good.

In less fortunate cases, the logician cannot supply both certification and decertification criteria for the good arguments, but can supply only a certification criterion for the good arguments; such is the case with quantification theory, as discussed in the next chapter. In still less fortunate cases, the logician is unable even to supply adequate certification criteria for good arguments, as Gödel established in showing that no formal system can precisely capture good arguments in arithmetic.[17]

It is because of the charge of applied proof theory that it try (if it can) to produce a humanly usable certification criterion of good argument that abstract proof theory (the only sort that we are developing) is essentially *inductive* in character: One begins by certifying a few very elementary good arguments, and then one shows how to certify others in terms of these in a simple step-by-step fashion. Observe the contrast with the semantic account as given in **2D-24**.

This account of proof theory marks not just one but two contrasts with semantics: The discipline of semantics gives us an "essential" account of the concept of good argument (in terms of meanings), but an account that treats this property as "invisible," while proof theory gives us an "accidental" account of the concept of good argument (via inductively generated lists), but an account that renders good-argumenthood "visible."

There is another contrast worth keeping in mind. In semantics one typically begins with an understanding of the meaning of the connectives (or other grammatical modes of combination) and works from this relatively "atomic" level towards an understanding of the somewhat more global idea of good argument; one tries to understand the "wholes" (good arguments) in terms of their "parts" (sentences and their ingredient connectives, etc.). In proof theory everything is reversed: One begins with certain arguments given as good, and can then try to understand the ingredient sentences and connectives in terms of their roles; one tries to understand the "parts" in terms of the "wholes." For example, the part-to-whole logician observes that *modus ponens* is valid *because* of the truth table meaning of the horseshoe, whereas the whole-to-part logician takes *modus ponens* as a given expression of the islander's way of inferential life, and thinks of the meaning of the horseshoe as determined by its role in this and other patterns of inference.

[17] Some people use such considerations as an argument in favor of the thesis that logic stops with the quantifiers. The argument seems to us based on a confusion concerning the Logician's Task: Though the task indeed defines the core of logic, it would be perverse to build into the definition of the logician that he is bound to succeed.

In the favorable cases we may hope that the two approaches mesh, as in fact they do in the case of TF.

One final remark: Note that what we do here *does* depend on the grammar of TF but does *not* depend on the semantics of TF (though of course it is motivated by those semantics).

2E.1 Proof theoretical choices

Just as there are alternative ways to develop grammar and semantics, so it is with proof theory. We make the rationale of our choices as clear as we can.

Theoremhood or derivability? Should we take "theoremhood" or "derivability" as the central idea of proof theory? Some logicians take theoremhood as central, thus posing the chief task of proof theory as the sorting out of the logical truths, because, they say, the question as to whether B follows truth functionally from some premisses A_1, …, A_n is just a matter of whether or not the sentence $A_1 \supset (A_2 \supset \ldots (A_n \supset B)\ldots)$ is a tautology. We instead take derivability as central, for the following reasons. First, it is a matter of substantive philosophical debate whether or not one's views of tautologyhood automatically settle one's views of truth functional logical consequence; we do not think that they do.[18] Second, it is useful to leave room for the possibility of infinitely many premisses, even if in these notes we shall not be wanting to explore that possibility. Third and most important, no one really cares what the logical truths are; instead, the practically important question is in fact what the valid inferences are; so this matter should be made central in proof theory.

Elementary vs nonelementary concepts of derivability. The fundamental idea of proof theory, then, is "derivability"; and what makes an account of derivability *proof theoretical* instead of semantical is that it shall be inductive in character: Some relations of derivability shall be taken as fundamental, and others shall be generated from these by some fixed patterns. One has an *elementary* concept of derivability if the relation is explainable in terms of a *conclusion* following by means of some *axioms* and *inferences* from some *premisses*. An example of a

[18]The matter is subtle; for some introductory considerations, see *Entailment: the logic of relevance and necessity* by A. R. Anderson and N. D. Belnap, Princeton University Press, 1975. As an example, we accept that $\sim A \supset (A \supset B)$ is a tautology, but reject that arbitrary B follows from $\sim A$ and A.

nonelementary concept of derivability is that underlying Fitch's method of subproofs, and we treat such an example in §2G.2 below. But this will be by the way; on grounds of pedagogical simplicity, we decide to restrict attention to elementary concepts of derivability.

Axioms. What is an "axiom" of a formal system? Heuristically, it is a starting point; it is a sentence you are permitted to write down without reference to previous items. Technically the important point is that the expression (of our use-language) "axiom of S" for a formal system S must be defined: What counts as an axiom of S is whatever we *say* counts as an axiom of S. One does not, however, count S as a *formal system* unless there is some kind of effective decision procedure that will mechanically decide whether a candidate sentence is or is not an axiom of S. This is an important matter for logic, and the idea should stay at the back of your mind; but as it turns out, we do not involve ourselves technically with such matters in these notes.[19]

We note that the separation into (1) "premisses" of a derivation and (2) "axioms" of the formal system in which the derivation is carried out is not essential technically, but it certainly is a motivational convenience. You are to picture the axioms of a formal system as "permanent," and as "always available," while the premisses are changeable. But that is only a picture; if you study the definitions below, you will see that premisses and axioms are given quite the same technical role (which is not so in other contexts or studies).

Inferences, rules, and modus ponens. An elementary concept of derivability needs axioms to get started, and "inferences" to keep going. What is an inference? We suppose it is some kind of "transition" from some premisses to a conclusion; but all that abstract logic needs from this intuitive concept is the division itself into premisses and conclusion:

2E-1 DEFINITION. *(Inference and rule)*

An *inference* is an ordered pair $\langle G, A \rangle$, with $G \subseteq$ Sent and $A \in$ Sent; G is the set of its *premisses* and A is its *conclusion*.

A *rule* is a set of inferences.

[19] See Church 1956 for a classic discussion of the issues.

We have at the same time defined "rule": From the point of view of the science of logic, a rule is a set of inferences—which in turn may be called the "instances" of the rule.

As a matter of fact, although these ideas of proof theory lie behind much of what we do, we shall not be working with them, for we decide to outfit all our concepts of derivability with but a single rule, a single kind of inference; namely, modus ponens, the instances all of which have the form

$$\langle\{A, A\supset B\}, B\rangle.$$

We make this choice for pedagogical simplicity: It permits us to invoke just modus ponens instead of the abstract concepts of inference and rule.

Modus ponens has two special features of which you should be aware. In the first place, each instance has finitely many premisses (namely, two); and in the second place, offered a candidate inference, it is a mechanical matter to determine whether or not one has an instance of modus ponens. Both of these features are normally required of the inferences and rules of a "formal system"; and the former figures technically in our results below. See Church 1956 for additional discussion.

Direct or indirect account of derivability? It is most usual first to give an account of a "derivation" from axioms by inferences sanctioned by rules, and then to define "derivability" indirectly in terms of "derivation." In order to secure a substantial technical simplification, we choose to define "derivability" directly instead. The matter is explained in more detail below, after elaboration of the notion of an axiom.

2E.2 Basic proof-theoretical definitions for S_{TF}

We shall call the formal theory we are developing "S_{TF}"; and first we highlight the notion of "axiom" by defining the set AxS_{TF} of axioms of S_{TF}.

2E-2 DEFINITION. *(*AxS_{TF}*)*

AxS_{TF} is that subset of Sent (AxS_{TF}type)
such that for any sentence D, $D\in AxS_{TF} \leftrightarrow$ there are sentences A, B, C such that D is identical to one of the following:

$S_{TF}1$. $A \supset (B \supset A)$ ($S_{TF}1$)

$S_{TF}2$. $(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$ ($S_{TF}2$)

$S_{TF}3$. $(\sim A \supset \sim B) \supset (B \supset A)$ ($S_{TF}3$)

This definition will hardly come as a surprise to those who have considered Theorem **2A-12** (Three ways to name T) and Theorem **2D-33** (Three tautologies), though of course the motivation goes in reverse: Having, in honor of the great Polish logician, Łukasiewicz, decided to choose $S_{TF}1$-$S_{TF}3$ as axioms of our formal theory S_{TF}, those otherwise undistinguished facts assume the prominence of Theorems.

2E-3 VARIANT. *(Axiom of* S_{TF}*)*

A is an axiom of $S_{TF} \leftrightarrow A \in AxS_{TF}$.

Note that what is being defined is "axiom of S_{TF}." Obviously each axiom of S_{TF} is a sentence of S_{TF}, *not* an expression of our use-language. Sometimes the use-expressions displayed above are called "axiom schemata," and the axioms are said to be "instances" of them. This is a useful mode of speech even though we don't know what our axioms look like.[20]

The standard definition of "derivability" is in terms of a "derivation": A is derivable from a set of sentences G just in case *there is* a derivation of A from G. (The definition is existential.) The standard definition of a derivation from G, in turn, is as a finite sequence of sentences, each of which is either an axiom, a member of G, or follows from predecessors by means of the rules—in our case modus ponens is the only rule. The picture to have is this. When an islander arises in the morning, he says some things before breakfast, which he thinks of as premisses. He is silent during breakfast, and after breakfast he adheres to the following rule: He only utters axioms, or perhaps repeats something he said before breakfast, or something B after he has already uttered C and $C \supset B$. All this time, until sunset, he thinks of himself as in the process of constructing a "derivation." And he thinks of the last thing he says, just as the sun goes down, as the "conclusion" of his "derivation," and as "derived from" his pre-breakfast premisses. So:

[20]We hope that this discussion is enough to prevent any possible confusion between the concept of the axioms of S_{TF} and the concept of *our* use-language axioms, but we do not suppose that it is enough to promote understanding of the second concept. It is a universal truth and a great irony that we can generally understand concepts referring to the islanders' language much more readily than we can understand concepts referring to our own linguistic practices.

2E-4 DEFINITION. *(S_{TF} derivation)*

A S_{TF}-*derivation of* A *from* G is a finite sequence of sentences, the last of which is A, and such that each member B satisfies one of the following three conditions:

1. $B \in G$;
2. B is an axiom of S_{TF};
3. for some C, there are predecessors C and $C \supset B$ in the sequence.

One can then say that A is S_{TF}-derivable from G if there is a S_{TF}-derivation of A from G. And one could in this way introduce the notation "$G \vdash_{S_{TF}} A$." (Note that, as advertised, the axioms and the "premisses" G are treated just alike.)

This is the most usual definition, and it is an intelligible account that you should keep firmly in mind; but note that it relies on some unexplained terminology: "finite sequence," "last," "member of a sequence," "predecessor." To use this definition on a quantifiers-and-truth-functions basis (as opposed to relying on geometrical intuition) would therefore require us to explain all those words, which indeed could be done; but you would discover that the development would be complicated, and unnecessarily so.

An interesting alternative, not at first as intuitive, but more manageable, is to give a direct inductive definition of derivability that bypasses any talk of derivations. That means we shall have a base clause, an inductive clause, and a closure clause, *directly* for derivability. As usual, the base clause tells us how to get started, in this case with axioms or premisses: The base clause tells us that any axiom or any premiss is derivable from any set G of premisses. The inductive clause tells us how to continue, using the familiar rule of modus ponens: If A and $A \supset B$ are derivable from a set G, so is B. And the closure clause tells us that "that's all," i.e., that is all the derivability-relations sanctioned by S_{TF}. We put these together in the following

2E-5 DEFINITION. *($\vdash_{S_{TF}}$)*

$\vdash_{S_{TF}} \subseteq (\mathcal{P}(\text{Sent}) \times \text{Sent})$ ($\vdash_{S_{TF}}$type)

Base clause. If $A \in G$, or if A is a S_{TF}-axiom, then $G \vdash_{S_{TF}} A$. In other words, with names for future reference, the following hold:

(A)[A ∈ G → G $\vdash_{S_{TF}}$ A] ($\vdash_{S_{TF}}$id)

(A)[A is an axiom of S_{TF} → G $\vdash_{S_{TF}}$ A] ($\vdash_{S_{TF}}$AxS_{TF}

Inductive clause.

(A)(B)[(G $\vdash_{S_{TF}}$ A ⊃ B and G $\vdash_{S_{TF}}$ A) → G $\vdash_{S_{TF}}$ B] ($\vdash_{S_{TF}}$MP)

Closure clause. We want to say that these are the *only* ways that A can be a S_{TF}-consequence of G, the *only* ways that one can have G $\vdash_{S_{TF}}$ A. That is, let Ψ(A) be any use-context. **Suppose** the following.

Basis step 1. A_1 ∈ G → Ψ(A_1), all A_1.
Basis step 2. A_1 ∈ AxS_{TF} → Ψ(A_1), all A_1.
Inductive step. (Ψ(A_1) and Ψ(A_1 ⊃ A_2)) → Ψ(A_2), all A_1, A_2.

Then for all sentences A, G $\vdash_{S_{TF}}$ A → Ψ(A). ($\vdash_{S_{TF}}$Induction)

Note that the "**Then**" part of the closure clause says that all S_{TF}-consequences of G have the property expressed by Ψ(A).

We will be using it chiefly in the following form.

2E-6 COROLLARY. *(Induction on* S_{TF}*-consequence)*

The following subproof rule is acceptable:

$A_1 \in G$	hyp for part I of Basis, flag A_1	
.		
.		
.		
$\Psi(A_1)$		$[A_1/A]$
$A_1 \in AxS_{TF}$	hyp for part II of Basis, flag A_1	
.		
.		
.		
$\Psi(A_1)$		$[A_1/A]$
$\Psi(A_1)$		$[A_1/A]$
$\Psi(A_1 \supset A_2)$	Ind. hyp, flag A_1, A_2	$[(A_1 \supset A_2)/A]$
.		
.		
.		
$\Psi(A_2)$		$[A_2/A]$
$(A)[G \vdash_{S_{TF}} A \to \Psi(A)]$	$\vdash_{S_{TF}}$ Induction	

Sometimes it is more convenient to gather the two parts of the Basis into a single subproof with a disjunctive hypothesis, as in the proof of **2E-15** below; it clearly makes no logical difference, and we will do so silently.

We install the analog of Convention **2D-27** for "$\vdash_{S_{TF}}$":

2E-7 CONVENTION. *(Dropping $\cup$ and $\{\}$ and $\varnothing$)*

Same as **2D-27**, with single instead of double turnstile.

2E-8 VARIANT. *(Variants of $\vdash_{S_{TF}}$)*

G S_{TF}-*implies* A $\leftrightarrow$ A *is a* S_{TF}-*consequence of* G $\leftrightarrow$ A is S_{TF}-*derivable from* G $\leftrightarrow$ $G \vdash_{S_{TF}} A$

A *is a theorem of* S_{TF} $\leftrightarrow$ A *is* S_{TF}-*provable* $\leftrightarrow$ $\vdash_{S_{TF}} A$

Before leaving, let us go back to the island. We may think of our statements "$G \vdash_{S_{TF}} A$" as giving a *relational* account of what the islanders are up to, perhaps in terms of that to which they take themselves to be *committed*. Thus, let G represent what some islander says before breakfast. Then Identity tells us he is committed to all of that, AxS_{TF} tells us that he is committed to all the axioms of S_{TF} regardless of what he says before breakfast, and *modus ponens* tells us about how some commitments lead inevitably to other commitments. The closure clause then limits his commitments to just those.

2E.3 Structural properties of derivability

The chief structural facts for "$\vdash_{S_{TF}}$" repeat those for "$\vDash_{TF}$," but their proofs are entirely different. They are "structural" because they do not involve any connectives or other features of the grammar of TF. These facts are important, and will be used again and again.

2E-9 FACT. *($\vdash_{S_{TF}}$ facts)*

Identity. $A \in G \to G \vdash_{S_{TF}} A$ ($\vdash_{S_{TF}}$id)

Weakening. $G \subseteq H \to [G \vdash_{S_{TF}} A \to H \vdash_{S_{TF}} A]$ ($\vdash_{S_{TF}}$weak)

Cut. $(G \vdash_{S_{TF}} A$ and $G, A \vdash_{S_{TF}} B) \to G \vdash_{S_{TF}} B$ ($\vdash_{S_{TF}}$cut)

Also the following variant is useful (and evidently follows by $\vdash_{S_{TF}}$weak): $(G \vdash_{S_{TF}} A$ and $H, A \vdash_{S_{TF}} B) \to G, H \vdash_{S_{TF}} B$

Finiteness. $G \vdash_{S_{TF}} A \to$ there is a finite subset G_0 of G such that $G_0 \vdash_{S_{TF}} A$; that is, $\exists G_0(G_0$ is finite and $G_0 \subseteq G$ and $G_0 \vdash_{S_{TF}} A)$. ($\vdash_{S_{TF}}$fin)

Exercise 25 *(Structural rules)*

You are to prove **2E-9**. We discuss each part in turn.

Identity. Trivial.

Weakening. First restate in the form: $G \subseteq H \to (A)[G \vdash_{S_{TF}} A \to H \vdash_{S_{TF}} A]$. Next assume the antecedent, $G \subseteq H$; then prove the consequent by Induction on S_{TF}-consequence, **2E-6**, choosing $\Psi(A)$ there as "$H \vdash_{S_{TF}} A$" here.

Cut. Restate as $G \vdash_{S_{TF}} A \to [G, A \vdash_{S_{TF}} B \to G \vdash_{S_{TF}} B]$. Next, to facilitate using Induction on S_{TF}-consequence correctly, relabel as: $H \vdash_{S_{TF}} B \to [H, B \vdash_{S_{TF}} A$

$\rightarrow$ H $\vdash_{S_{TF}}$ A]. Also express one of the implicit universal quantifiers explicitly at a convenient place: H $\vdash_{S_{TF}}$ B $\rightarrow$ (A)[H, B $\vdash_{S_{TF}}$ A $\rightarrow$ H $\vdash_{S_{TF}}$ A]. Assume the antecedent; now prove the consequent by Induction on S_{TF}-consequence, **2E-6**, choosing G there as (H$\cup$\{B\}) here and Ψ(A) there as "H $\vdash_{S_{TF}}$ A" here. (With regard to "(H$\cup$\{B\})," recall that by **2D-27**, the union sign and the curly brackets can be dropped *only* on the left of the turnstiles.)

Finiteness. Prove by Induction on S_{TF}-consequence, choosing Ψ(A) there as the consequent here. Use facts from §2B.3. Be careful when treating the inductive case $\vdash_{S_{TF}}$MP that you do not fallaciously assume that the finite sets promised for A_1 and $A_1 \supset A_2$ are the same; instead, generate from these separate sets a single finite set that will do the job for A_2.

▷ . . . . ◁

Exercise 26 *(Subtractive cut)*

One can also argue for $\vdash_{S_{TF}}$cut by proving that G$-$\{B\} $\vdash_{S_{TF}}$ B $\rightarrow$ (A)[G $\vdash_{S_{TF}}$ A $\rightarrow$ G$-$\{B\} $\vdash_{S_{TF}}$ A]. Do so.

▷ . . . . ◁

At this point it is incumbent upon you to compare the four $\vdash_{S_{TF}}$ facts just proved with their analogs stated but not entirely proved as Lemma **2D-29** and Conjecture **2D-30**. In particular you should note that the semantic version of Finiteness was Too Hard to Prove, while the proof-theoretical version just above is Easy. But almost equally illuminating is that Weakening and Cut here require careful inductive argument, while their semantic analogs require no more than keeping your quantifiers straight.

There is an additional structural turnstile-fact, mysterious in itself, that proves useful later. We preface it with a variant and a couple of definitions—at least one of the latter being itself mysterious until later.

2E-10 VARIANT. *(E-free)*

G *is* E-*free in* S_{TF} $\leftrightarrow$ G $\nvdash_{S_{TF}}$ E.

Observe that we use "$\nvdash_{S_{TF}}$" to deny a consequence statement, in analogy with "$\neq$," so that for G to be E-free is for E not to be a S_{TF}-consequence of G.

2E-11 DEFINITION. *(Maximal E-free)*

G* is *maximal* E-*free in* S_{TF} ↔ G* is E-free in S_{TF} but no proper superset of G* is E-free in S_{TF}. That is,

$$G^* \nvdash_{S_{TF}} E \qquad (MEF_{S_{TF}}1)$$

$$(H)(G^* \subset H \rightarrow H \vdash_{S_{TF}} E) \qquad (MEF_{S_{TF}}2)$$

The two parts always go together. Note that the hypothesis of ($MEF_{S_{TF}}2$) is that G* is a *proper* subset, NAL:**9A-10**, of H. So G* is maximal E-free if G* is itself E-free, while no "bigger" set is. (The use of "maximal" for such an idea is common in applications of set theory.) In still other words, G* itself doesn't yield E, but every "bigger" set does.

Three remarks. (1) Note the names for the two parts of the definition; we use them later. (2) From now on when you see "G*," you will know you are in the presence of a set that either is or is hoped to be maximal E-free. (3) In order to avoid confusion, we use "E" (instead of "A" or some earlier letter) in order to pick out a sentence that G* is said to be free of. But these uses of "G*" and "E" are heuristic only; logically "G*" and "E" are just use-language variables like any others.

2E-12 DEFINITION. *(S_{TF}-closure)*

G *is closed under* $\vdash_{S_{TF}}$ ↔ (A)(G $\vdash_{S_{TF}}$ A → A∈G).

In words: Every S_{TF}-consequence of G already belongs to G, so that S_{TF}-consequence does not lead outside of G.

This is an interesting idea in itself, but we won't be making much use of it except by the way. Such a set is sometimes called a "S_{TF} *theory*" since it is like the collection of all the theorems (consequences) of some axiom set.

Exercise 27 *(A closure property)*

Optional. You may wish to divert yourself by proving that if G is the set of all S_{TF}-consequences of some set G_0 (i.e., if (A)[A∈G ↔ $G_0 \vdash_{S_{TF}}$ A]), then G itself is closed under $\vdash_{S_{TF}}$. (Use Induction on S_{TF}-consequence, **2E-6**.)

▷ . . . . ◁

Here is an easy but useful fact about maximal E-free sets and S_{TF}-closure.

2E-13 FACT. *(Maximality/*S_{TF}*-closure)*

G* is maximal E-free in S_{TF} $\rightarrow$ (A)(G* $\vdash_{S_{TF}}$ A $\rightarrow$ A$\in$G*).

That is, we can usually tell the difference between a set and its S_{TF}-consequences; but not in the case of maximal E-free sets: They contain all and only their S_{TF}-consequences; they are closed under $\vdash_{S_{TF}}$, **2E-12**. Maximal E-free sets, which you will learn to prize, are "rare." For instance, evidently any maximal E-free set has to be infinite.

Exercise 28 *(Maximality/*S_{TF}*-closure)*

Prove Fact **2E-13**. Even though the consequent of the Fact has the right form for proof by Induction on S_{TF}-consequence, it is *not* required. Try reductio, and use $\vdash_{S_{TF}}$cut, **2E-9**.

▷ . . . . ◁

The properties Identity, Weakening, Cut, Finiteness, and maximality/closure (which follows from Cut alone) hold under quite general conditions; the only aspect of this generality that is used in these notes, however, is that the properties hold not only for S_{TF}, but also for any "axiomatic extension" of S_{TF}. We delay proof of this until needed, §3D.2 below.

2E.4 Derivability and connectives

In the end we want to show that the single and double turnstiles have exactly the same properties, and indeed that G $\vDash_{TF}$ A $\leftrightarrow$ G $\vdash_{S_{TF}}$ A. We now know that they have much the same *structural* properties (**2D-28** and **2E-9**); what about the way they relate to connectives? This is a long story; we make the following modest beginning, but a beginning sufficiently important to give it a S_{TF}-number.

2E-14 FACT. *(*S_{TF}4*)*

G $\vdash_{S_{TF}}$ A $\supset$ A (S$_{TF}$4)

PROOF. Here is a slow and painful articulation of how it has to go in our framework, which is not as efficient for establishing such facts as are some other frameworks.[21]

1	$A \supset [(A \supset A) \supset A] \in AxS_{TF}$	$S_{TF}1$, **2E-2**
2	$G \vdash_{S_{TF}} A \supset [(A \supset A) \supset A]$	1, $\vdash_{S_{TF}} AxS_{TF}$, **2E-5**
3	$G \vdash_{S_{TF}} A \supset [(A \supset A) \supset A] \supset [A \supset (A \supset A)] \supset (A \supset A)$	$S_{TF}2$, $\vdash_{S_{TF}} AxS_{TF}$
4	$G \vdash_{S_{TF}} [A \supset (A \supset A)] \supset (A \supset A)$	2, 3 $\vdash_{S_{TF}}$MP, **2E-5**
5	$G \vdash_{S_{TF}} A \supset (A \supset A)$	$S_{TF}1$, $\vdash_{S_{TF}} AxS_{TF}$
6	$G \vdash_{S_{TF}} A \supset A$	4, 5 $\vdash_{S_{TF}}$MP

□

The following important theorems relate $\vdash_{S_{TF}}$ to $\supset$ in the same way that $\vDash_{TF}$ is related to $\supset$, **2D-31**. The name is standard in the literature; in contrast to the previous theorem, its proof in our framework is both smooth and explanatory.

2E-15 THEOREM. *(Deduction theorem)*

$$G, A \vdash_{S_{TF}} B \rightarrow G \vdash_{S_{TF}} A \supset B$$

PROOF. To prepare, relabel: $H, B \vdash_{S_{TF}} A \rightarrow H \vdash_{S_{TF}} B \supset A$. Now use Induction on S_{TF}-consequence, **2E-6**, choosing G there as $(H \cup \{B\})$ here and $\Psi(A)$ there as "$H \vdash_{S_{TF}} B \supset A$" here. Suppose for the Basis that either $A_1 \in (H \cup \{B\})$ or $A_1 \in AxS_{TF}$, i.e., that *(a)* $A_1 = B$ or *(b)* ($A_1 \in H$ or $A_1 \in AxS_{TF}$); we need to show that $H \vdash_{S_{TF}} B \supset A_1$. In case *(a)* the conclusion of the Basis comes by $S_{TF}4$ (**2E-14**), since $B = A_1$. In case *(b)*, we obtain $H \vdash_{S_{TF}} A_1$ by either $\vdash_{S_{TF}}$id or $\vdash_{S_{TF}} AxS_{TF}$, and further note that $H \vdash_{S_{TF}} A_1 \supset (B \supset A_1)$ by $\vdash_{S_{TF}} AxS_{TF}$, **2E-5** ($S_{TF}1$, **2E-2**); so $H \vdash_{S_{TF}} B \supset A_1$ by $\vdash_{S_{TF}}$MP.

Now suppose for the Inductive hypothesis that

$$H \vdash_{S_{TF}} B \supset A_1 \text{ and } H \vdash_{S_{TF}} B \supset (A_1 \supset A_2)$$

we need to show that $H \vdash_{S_{TF}} B \supset A_2$. But

$$H \vdash_{S_{TF}} (B \supset (A_1 \supset A_2)) \supset ((B \supset A_1) \supset (B \supset A_2))$$

[21] No framework is equally efficient for all tasks; why should anyone think otherwise? Why, that is, should anyone think that some one framework is logically privileged? Perhaps because of the understandable but regrettable professional tendency on the part of both mathematicians and philosophers to neglect keeping in mind that logic is at bottom a practical discipline. Thus "mere efficiency" is sometimes imagined not to matter, a formulation that ironically forgets that mattering itself is a matter of practical (or esthetic) purposes.

by $\vdash_{S_{TF}}$AxS$_{TF}$ (S$_{TF}$2, **2E-2**), so H $\vdash_{S_{TF}}$ B$\supset$A$_2$ by two uses of $\vdash_{S_{TF}}$MP □

Exercise 29 *(Subtractive deduction theorem)*

The Deduction theorem comes to this as well: (A)[G $\vdash_{S_{TF}}$ A $\rightarrow$ G $-$ {B} $\vdash_{S_{TF}}$ B$\supset$A].

1. Show this "subtractive deduction theorem" equivalent to **2E-15**, using only easy set theory and logic (especially: Do not use induction; but you may need a bit of **2E-9**).
2. Identify what here should be chosen as Ψ(A) for an alternative proof by Induction on S$_{TF}$ consequence, **2E-6**.
3. Optional. Complete the inductive proof.

▷ . . . . ◁

The converse of the Deduction theorem is called just that:

2E-16 FACT. *(Converse of deduction theorem)*

G $\vdash_{S_{TF}}$ A$\supset$B $\rightarrow$ G, A $\vdash_{S_{TF}}$ B.

PROOF. Straightforward. Suppose G $\vdash_{S_{TF}}$ A$\supset$B. But then G, A $\vdash_{S_{TF}}$ A$\supset$B by $\vdash_{S_{TF}}$weak; and also of course G, A $\vdash_{S_{TF}}$ A by $\vdash_{S_{TF}}$id. So the desired conclusion comes by $\vdash_{S_{TF}}$MP □

The Deduction theorem and its converse also hold for every axiomatic extension of S$_{TF}$, as noted in §3D.2 below.

We will certainly want analogs of the other conditional and negation properties listed in **2D-31**. The Deduction theorem and its converse are repeated with another name, just to have everything we want in the same place, and to highlight the analogy with **2D-31**.

2E-17 CONJECTURE. *(Connectives and* $\vdash_{S_{TF}}$*)*

$G \vdash_{S_{TF}} A \supset B \leftrightarrow G, A \vdash_{S_{TF}} B$ ($\vdash_{S_{TF}}\supset$)

$G, A \supset B \vdash_{S_{TF}} C \leftrightarrow (G, \sim A \vdash_{S_{TF}} C$ and $G, B \vdash_{S_{TF}} C)$ ($\supset\vdash_{S_{TF}}$)

$G \vdash_{S_{TF}} \sim A \leftrightarrow G, A \vdash_{S_{TF}} \sim(B \supset B)$ ($\vdash_{S_{TF}}\sim$)

$G, \sim A \vdash_{S_{TF}} \sim(B \supset B) \leftrightarrow G \vdash_{S_{TF}} A$ ($\sim\vdash_{S_{TF}}$)

We postpone proof of this conjecture until **2G-2**; instead we record a lemma and two small facts that depend on the lemma. The facts, though humble, are going to be useful later, so that we give them S_{TF} numbers; the lemma (though indeed revelatory of the structure of S_{TF}) we use just for the facts. (You will see that our basic plan is to prove just enough about $\vdash_{S_{TF}}$ to carry out the program of the next section.)

2E-18 LEMMA. *(Some* S_{TF} *consequences)*

---

a. $G, \sim A \vdash_{S_{TF}} \sim B \rightarrow G, B \vdash_{S_{TF}} A$

b. $A, \sim A \vdash_{S_{TF}} B$

c. $\sim\sim A \vdash_{S_{TF}} A$

d. $A \vdash_{S_{TF}} \sim\sim A$

PROOF. *Ad* a. This is, in context, a form of $S_{TF}3$, **2E-2**. Suppose (1) the antecedent of (a), and apply the Deduction theorem, **2E-15**, to (1) in order to obtain (2) $G \vdash_{S_{TF}} \sim A \supset \sim B$. Also (3) $G, \vdash_{S_{TF}} (\sim A \supset \sim B) \supset (B \supset A)$ is an instance of $\vdash_{S_{TF}}$AxS$_{TF}$, **2E-5** ($S_{TF}3$, **2E-2**). So (4) $G \vdash_{S_{TF}} B \supset A$ from (2) and (3) by $\vdash_{S_{TF}}$MP. Obtain the consequent of (a) from (4) by the Converse of the deduction theorem, **2E-16**.

Ad b. We have $\sim A, \sim B \vdash_{S_{TF}} \sim A$ by Identity, **2E-9**, and so obtain (b) by applying (a) to this. It is amazing how little trouble it is to show that whatever "follows from" an arbitrary contradiction.

Ad c. By (b) we have that $\sim\sim A, \sim A \vdash_{S_{TF}} \sim\sim\sim A$. Then (c) comes from this by (a) and some easy set theory.

Ad d. By (c) we know that $\sim\sim\sim A \vdash_{S_{TF}} \sim A$. Now (d) follows by applying (a). □

2E-19 FACT. *(*$S_{TF}5$*,* $S_{TF}6$*)*

$S_{TF}5.\ {\sim}A, A \vdash_{S_{TF}} B$ ($S_{TF}5$)

$S_{TF}6.\ A \supset B, {\sim}A \supset B \vdash_{S_{TF}} B$ ($S_{TF}6$)

PROOF. $S_{TF}5$ is just **2E-18**b. For proof of $S_{TF}6$ we first define $H = \{A \supset B, {\sim}A \supset B\}$, just to unclutter the presentation; we are accordingly trying to show that $H \vdash_{S_{TF}} B$, as on line 9 below. References to Identity, Weakening, and Cut (including its variant form) are to **2E-9**; the Converse of the deduction theorem is **2E-16**. Letter references are to **2E-18**.

1	$H, A \vdash_{S_{TF}} B$	Def. of H, Ident., Conv. of Ded. Th.
2	$H, {\sim}{\sim}A \vdash_{S_{TF}} {\sim}{\sim}B$	1, c, d, Cut, Cut
3	$H, {\sim}B \vdash_{S_{TF}} {\sim}A$	2, a
4	$H, {\sim}A \vdash_{S_{TF}} B$	Def. of H, Ident., Conv. of Ded. Th.
5	$H, {\sim}A \vdash_{S_{TF}} {\sim}{\sim}B$	4, d, Cut
6	$H, {\sim}B \vdash_{S_{TF}} A$	5, a
7	$A, {\sim}A \vdash_{S_{TF}} {\sim}(A \supset B)$	$S_{TF}5$
8	$H, {\sim}B \vdash_{S_{TF}} {\sim}(A \supset B)$	3, 6, 7, Cut, Cut
9	$H \vdash_{S_{TF}} B$	8, a, Easy set theory

□

Leaving for a moment our basic plan of preparing for the next section, at this point we turn to the two most important proof theoretical theorems involving substitution; the first one says that S_{TF}-theoremhood (see **2E-8**) is closed under the substitution operation, **2C-1**. (Closure under substitution holds equally—and more generally—for S_{TF}-consequence, but since we cannot state the latter fact without a definition that we do not yet need, we let the former suffice):[22]

2E-20 THEOREM. *(Closure of* S_{TF}*-theoremhood under substitution)*

$\vdash_{S_{TF}} A \rightarrow \vdash_{S_{TF}} [B/p](A)$.

This theorem has a certain importance since many formal systems like S_{TF} have a substitution rule as primitive, in addition to modus ponens. These formal systems typically replace our infinite set of axioms with just three, choosing one instance each of $S_{TF}1$, $S_{TF}2$, and $S_{TF}3$. Thus, S_{TF} is "stronger" than these systems in its axioms, and "weaker" in its primitive rules. Theorem **2E-20** suggests that the weakness is only apparent, since after all S_{TF} can obtain the effect of substitution.

[22]The more general notion, of substitution throughout a set, will be defined when it is finally needed, in §3F below.

PROOF. Tedious. By Induction on S_{TF}-consequence, choosing G there as ∅ here and Ψ(A) there as "$\vdash_{S_{TF}}$ [B/p](A)" here; using of course Definition **2C-1**, especially []A⊃ (over and over) □

You should compare the above with the semantic Fact **2D-35**.

There is another important theorem involving the substitution notation, not to be confused with the foregoing:

2E-21 THEOREM. *(Replacement theorem for* S_{TF}*)*

(G, B $\vdash_{S_{TF}}$ C and G, C $\vdash_{S_{TF}}$ B) → (A)(G, [B/p](A) $\vdash_{S_{TF}}$ [C/p](A) and G, [C/p](A) $\vdash_{S_{TF}}$ [B/p](A)).

Think of A as a sentence with a single occurrence of p: A = ...p... . Then the theorem says that when in the context of G, B and C are interdeducible, then so are ...B... and ...C.... You should see that ...B... and ...C... are just alike except that B has been "replaced" by C—which gives the theorem its name.

The theorem is often stated in terms of the triple bar connective of TF:

G $\vdash_{S_{TF}}$ B ≡ C → G $\vdash_{S_{TF}}$ ([B/p](A) ≡ [C/p](A)).

Or sometimes it is stated in the form:

($\vdash_{S_{TF}}$ B ≡ C and $\vdash_{S_{TF}}$ [B/p](A)) → $\vdash_{S_{TF}}$ [C/p](A).

Given usual proof-theoretical postulates for the triple bar, these are corollaries of the form given above.

The Replacement theorem is not needed for consistency or completeness, **2G-1**, and given the latter, the proof theoretical Replacement theorem **2E-21** is an easy corollary of the *Semantic* replacement theorem **2D-22**, with which you should compare it. But the proof theoretical version can also be given a direct proof here using only elementary techniques:

PROOF. Tedious. Assume the antecedent. Prove the consequent by Induction on sentences, **2B-18**, choosing Ψ(A) there as the scope of the quantifier "(A)" here, and use the following features of S_{TF}, the proofs of which are omitted: G, A $\vdash_{S_{TF}}$ B → G, ∼B $\vdash_{S_{TF}}$ ∼A; (G, A $\vdash_{S_{TF}}$ B and G, C $\vdash_{S_{TF}}$ D) → G, B ⊃ C $\vdash_{S_{TF}}$ A ⊃ D □

It is worthwhile noting that the two chief theorems relating S_{TF}-consequence to substitution require quite different methods of proof: The proof of the Closure of S_{TF}-theoremhood under substitution, **2E-20**, relies on Induction on S_{TF} *consequence*, **2E-6**, while the proof of the Replacement theorem for S_{TF}, **2E-21**, uses Induction on *sentences*, **2B-18**. You will understand these matters much more thoroughly if you spend some time trying (and failing) to prove each of these by the other (wrong) method.

We close this section by listing *precisely* the properties of $\vdash_{S_{TF}}$ on which we shall be relying in showing that $\vDash_{TF}$ and $\vdash_{S_{TF}}$ coincide. All parts have already been established, as indicated. Be sure you know what these properties are, and how to prove them. The first four are of course the defining properties of $\vdash_{S_{TF}}$, **2E-5**.

2E-22 THEOREM. *(Properties of* $\vdash_{S_{TF}}$*)*

---

1. $\vdash_{S_{TF}}$id, **2E-5**, **2E-9**
2. $\vdash_{S_{TF}}$AxS_{TF}, **2E-5**
3. $\vdash_{S_{TF}}$MP, **2E-5**
4. $\vdash_{S_{TF}}$Induction, or Induction on S_{TF}-consequence, **2E-5**
5. $\vdash_{S_{TF}}$weak, **2E-9**
6. $\vdash_{S_{TF}}$cut, **2E-9**
7. $\vdash_{S_{TF}}$fin, **2E-9**
8. Maximality/S_{TF}-closure, **2E-13**
9. Deduction theorem, **2E-15**
10. $\vdash_{S_{TF}}$$S_{TF}$5, **2E-19**
11. $\vdash_{S_{TF}}$$S_{TF}$6, **2E-19**

Exercise 30 *(Elementary proof theory of* TF*)*

The chief thing is to understand and be able to prove all parts of **2E-22**.

1. Flesh out the proofs of the various parts of Fact **2E-9**.

2. Flesh out the proof of Theorem **2E-15**. You may wish to compare this proof with other proofs of the "Deduction theorem" in other books; our direct inductive definition of derivability in **2E-5** permits a proof without "auxiliary" constructions, without geometrical intuition, and without appeal to arithmetic.

3. Optional. Prove Conjecture **2E-17**.

4. Optional. Define uniform substitution of B for p in a whole set G by completing the following:

   $$\mathrm{A} \in [\mathrm{B/p}](\mathrm{G}) \leftrightarrow$$

   Then use this notation to state that $\mathrm{S_{TF}}$-consequence is closed under uniform substitution. (You are not expressly forbidden to prove this generalization of Fact **2E-20**.)

5. Optional. Give a short, *non*-inductive proof of the $\vdash_{\mathrm{S_{TF}}}$cut part of **2E-9**, using the Deduction theorem, **2E-15**. (Since this proof uses properties of the connective $\supset$ in order to establish $\vdash_{\mathrm{S_{TF}}}$cut, it is inferior to the one outlined in these notes.)

▷ . . . . ◁

2F Consistency and completeness of $\mathrm{S_{TF}}$

There are some profoundly important *connections* between the semantics and the proof theory we have hitherto developed separately. It is easy to say what we want to show:

$$\mathrm{G} \vdash_{\mathrm{S_{TF}}} \mathrm{A} \leftrightarrow \mathrm{G} \vDash_{\mathrm{TF}} \mathrm{A}.$$

But it will not be easy to show what we want to say; in fact, the entire section will be devoted to this enterprise.

2F.1 Consistency of S_{TF}

Consider for a moment the title of this section. "Consistency" is used in many ways: The term can be applied to a single sentence or to a set of sentences, in either a semantic or a proof theoretic sense; but here we are invoking a sense spanning both proof theory and semantics, since we wish to speak of the semantic consistency of a formal (proof theoretic) system. In particular, what we are after is the "semantic consistency" of S_{TF}. To say a formal system is semantically consistent is to say that it is consistent with (harmonizes with) its intended semantic interpretation. It is to say that the formal system "doesn't do anything wrong." In the case at issue, you should think of the matter in this way. On the one hand, S_{TF} inductively certifies certain inferences as acceptable by permitting us to say that $G \vdash_{S_{TF}} A$. In other words, "$G \vdash_{S_{TF}} A$" says that A follows from G *by the lights of* S_{TF}; that the inference from G to A has S_{TF}'s inductive stamp of approval. When we learn that $G \vdash_{S_{TF}} A$, we have found a route in the order of knowledge to the claim that G yields A. On the other hand, "$G \vDash_{TF} A$" means to tell us that A *really does* follow from A, that the inference really is (in itself) a good one, in the order of being.

Now the "semantic consistency" (or "soundness") of S_{TF} amounts to this: *only* the really good inferences are certified as such by S_{TF}.[23] Anything that S_{TF} says is O.K. in fact is O.K. That is: $G \vdash_{S_{TF}} A \rightarrow G \vDash_{TF} A$. Such Knowledge as we claim agrees with Being. So half of our job is to prove

2F-1 THEOREM. *(Semantic consistency of* S_{TF}*)*

$$G \vdash_{S_{TF}} A \rightarrow G \vDash_{TF} A$$

In words: If there is a S_{TF}-derivation of A from G, then in fact whenever all members of G are true, so is A. If we believe that $\vDash_{TF}$ gives the correct account of valid inference, then the theorem tells us that $\vdash_{S_{TF}}$ does not "say too much." Happily, this is easy to prove.

Exercise 31 *(Semantic consistency of* S_{TF}*)*

Prove Theorem **2F-1**. You can see from its form (even without tinkering) that we should be able to use Induction on S_{TF}-consequence, **2E-6**, to prove it, choosing

[23] The term "soundness" is nowadays much more common than "semantic consistency." Both, however, are jargon, and so we choose as a matter of light preference.

Ψ(A) there as "G$\vDash_{TF}$ A" here. Of use will be Corollary **2D-32** and Theorem **2D-33**, as well as Fact **2D-28**.

▷ . . . . ◁

It is important to observe that although this theorem is a mathematical fact, the degree of its *interest* is not. In particular, how much point the theorem has depends partly on the degree to which the definition of "$\vDash_{TF}$" really does catch a sensible notion of "good inference"; and this is indeed a proper matter for philosophical (not mathematical) debate, as we indicated in the discussion following Definition **2D-24** of "$\vDash_{TF}$."

2F.2 Completeness of S_{TF}

"Completeness" also has many senses in logic; the "semantic completeness" of a formal system is (we think always) the converse of its semantic consistency. Just as consistency says that *only* really good inferences are O.K.'d by S_{TF}, so completeness says that *all* really good inferences are authorized by S_{TF}. If in fact the inference from G to A is a good one, then S_{TF} will definitely provide a way for you to derive A from G. S_{TF} doesn't leave out any good inferences; it is "complete." Hence, the semantic completeness of S_{TF} amounts to this: G $\vDash_{TF}$ A $\rightarrow$ G $\vdash_{S_{TF}}$ A.

In contrast to consistency, completeness is a "hard" theorem. It is by no means "self-proving"; in particular, it requires auxiliary concepts and constructions, which is to say, concepts and constructions not automatically suggested by unpacking the definitions involved in the language of the statement of the theorem. There are many proofs of it; the version we give is one based on that of Leon Henkin. It is not easy in conception at all, but (a) it is beautiful, and (b) understanding it now makes things much easier when we come to quantifiers.

We want to approach the proof from the outside in, by first giving you the general structure, and by then filling in the details. And first we want to call your attention to the four auxiliary ideas the proof invokes:

1. The idea of a "maximal E-free set." We gave you this above as Definition **2E-11**. It is a notion from proof-theory, but it is a notion that links easily with semantics.

2. The idea of a "truth-like set." We define this below, Definition **2F-2**, without any reliance whatsoever on either semantics or proof theory; but it is, perhaps, the central idea mediating those two disciplines, the keystone of the whole proof.

3. The idea of a TF-valuation, $Val_{\mathbf{i}}$, "agreeing with a set, G*." Sometimes such a $Val_{\mathbf{i}}$ is called a "canonical valuation." This semantic idea is critical for our proof that proof-theoretical consequence, $\vdash_{S_{TF}}$, is co-extensive with semantic consequence, $\vDash_{TF}$.

4. The idea of a TF-interpretation, **i**, "agreeing with a set, G*." Sometimes such a TF-interpretation is called a "canonical interpretation." Think of an agreeing TF-intepretation as a kind of "base case" for an agreeing TF-valuation.

Here are the official definitions of the three of the four ideas above that are new.

2F-2 DEFINITION. *(Truth-like set for* TF*)*

G* is a *truth-like set for* TF (we will drop the "TF" unless specially needed) if and only if G* $\subseteq$ Sent, and both of the following hold:

$$\sim A \in G^* \leftrightarrow A \notin G^* \qquad (TL\sim)$$

$$A \supset B \in G^* \leftrightarrow [A \in G^* \rightarrow B \in G^*] \qquad (TL\supset)$$

We choose the name "truth-like set" because a truth-like set works exactly like "$Val_{\mathbf{i}}(A) = T$," **2D-7**, or its variant "A is true on **i**," **2D-11**, with respect to connectives. We have named the clauses of the definition to emphasize this (compare Fact **2D-13** on p. 52). We use "G*" because, as it turns out, maximal E-free sets are truth-like sets (see remark 2 after Definition **2E-11** on p. 87).

2F-3 DEFINITION. *(Agreement of* **i** *and* G**)*

Given any set of sentences, G*, a TF-interpretation, **i**, *agrees with* G* $\leftrightarrow$ **i** is the characteristic function of G* relative to the domain, TF-atom, NAL:**9B-19**.[24] Namely,

[24] We have (just) rewritten this definition in NAL to conform to present usage. When consulting an unrevised version of NAL, on p. 232, Definition **9B-19**, delete "$X_0 \subseteq X$" and change the third line to "$f(y) = T \leftrightarrow y \in (X \cap X_0)$."

$\mathbf{i} \in$ TF-interpretation; i.e., $\mathbf{i} \in$ (TF-atom $\mapsto$ **2**), **2D-4**.

For every TF-atom, p, $\mathbf{i}(p) = T \leftrightarrow p \in G^*$.

We use "G*" because the idea of a TF-interpretation agreeing with a set, G*, is (meaningful but) not interesting except for maximal E-free (hence truth) sets.

2F-4 DEFINITION. *(Agreement of* $Val_{\mathbf{i}}$ *and* G*)

Given any set of sentences, G*, and TF-interpretation, **i**, a TF-valuation, $Val_{\mathbf{i}}$, *agrees with* G* $\leftrightarrow$

For every sentence, A, $Val_{\mathbf{i}}(A) = T \leftrightarrow A \in G^*$.

Note that Definition **2F-3** states the agreement of G* with a TF-interpretation (of just TF-atoms), whereas Definition **2F-4** concerns agreement of G* with a TF-valuation (of all sentences). Warning: The existence of an agreeing TF-interpretation is cheap, but the existence of an agreeing TF-valuation is not.

We now have enough conceptual machinery to outline for you a Henkin style version of the completeness theorem for S_{TF}. We are going to do this by first stating a minifact and three lemmas without proof or explanation (even as to their names). Then we will show you in a fourth lemma how the minifact and the three lemmas work together to yield the theorem we want, so the only explanation you are going to get of them just now is this: They work.[25]

2F-5 MINIFACT. *(Existence of an agreeing* TF*-interpretation)*

For each set G*, there is a TF-interpretation, **i**, such that **i** agrees with G*.

2F-6 LEMMA. *(Lindenbaum's lemma)*

Every set E-free in S_{TF}, **2E-10**, can be extended to a set maximal E-free in S_{TF}, **2E-11**. That is, if $G \nvdash_{S_{TF}} E$, then there is a G* such that

[25] Some philosophers of mathematics think this is the only kind of mathematical understanding there is. Such a view seems to us partial.

$G \subseteq G^*$ (LL1)

$G^* \nvdash_{S_{TF}} E$ ($MEF_{S_{TF}}1$)

$(H)(G^* \subset H \rightarrow H \vdash_{S_{TF}} E)$ ($MEF_{S_{TF}}2$

The "$MEF_{S_{TF}}$" names come from **2E-11**.

2F-7 LEMMA. *(Maximality/truth-like-set lemma)*

If a set of sentences G* is maximal E-free in S_{TF}, **2E-11**, then G* is a truth-like set, **2F-2**.

Just as maximal E-free sets are tied to truth-like sets, so the latter are bound to agreeing TF-valuations:

2F-8 LEMMA. *(Truth-like set/agreeing interpretation/agreeing valuation lemma)*

If G* is a truth-like set, and if the TF-interpretation, **i**, agrees with G* (**2F-3**), then the TF-valuation, $Val_{\mathbf{i}}$, agrees with G*.

Now we turn to the fourth lemma, which shows how the minifact and the other three lemmas suffice for the completeness of S_{TF}. This one we prove without delay.

2F-9 LEMMA. *(Course of proof of* S_{TF} *completeness)*

Suppose Lindenbaum's lemma (**2F-6**), the maximality/truth-like-set lemma (**2F-7**), the existence of an agreeing TF-interpretation minifact (**2F-5**), the and the truth-like set/agreeing TF-interpretation/agreeing TF-valuation lemma (**2F-8**). Then S_{TF} is complete: $(G \vDash_{TF} E) \rightarrow (G \vdash_{S_{TF}} E)$.

PROOF. Suppose the four hypotheses of the lemma. Suppose, for a proof by contraposition, that $G \nvdash_{S_{TF}} E$, that is, that G is E-free. Let G* be the set promised by Lindenbaum's lemma. Clearly G* is a truth-like set, by the maximality/truth-like-set lemma (**2F-7**). By the Existence of an agreeing TF-interpretation, **2F-5**, choose **i** as a TF-interpretation agreeing with G*, so that $Val_{\mathbf{i}}$ by the truth-like set/agreeing TF-interpretation/agreeing TF-valuation lemma. That is, for all sentences A, $Val_{\mathbf{i}}(A) = T \leftrightarrow A \in G^*$. This implies that $Val_{\mathbf{i}}(A) = T$ for all $A \in G$, since $G \subseteq G^*$ by (LL1) of Lindenbaum's lemma. It also implies that $Val_i(E) \neq T$, since $E \notin G^*$ (for if $E \in G^*$, we would have $G^* \vdash_{S_{TF}} E$ by $\vdash_{S_{TF}}$id, which would contradict $MEF_{S_{TF}}1$ of Lindenbaum's lemma). So $G \nvDash_{TF} E$ by the definition of $\vDash_{TF}$, **2D-24** on p. 61. Hence, by contraposition,

$$(G \vDash_{TF} E) \rightarrow (G \vdash_{S_{TF}} E).$$

□

Thus the completeness of S_{TF} follows from the minifact and the three lemmas. Let us verify these four propositions in reverse order so that we may pass from the easiest to the most difficult.

The proof of Minifact **2F-5** requires only the observation that there is a characteristic function for each Y relative to each domain, X, as recorded pedantically in NAL:**9B-19**. For the present application, choose Y there as G* here, and the domain, X, there as TF-atom here.

We set as an exercise the proof of the truth-like set/agreeingTF-interpretation/agreeing TF-valuation lemma, **2F-8**:

Exercise 32 *(Truth-like set/agreeing* TF*-interpretation/agreeing* TF*-valuation lemma)*

Prove Lemma **2F-8**. Use Induction on sentences, **2B-18** on p. 32. This one is longish, but transparent because the definitions of "agreeing TF-interpretation," **2F-3**, and "Val_i," **2D-7** on p. 49, give us the base case for the Induction on sentences, and the clauses for the definitions of "Val_i," **2D-7**, and of "truth-like set," **2F-2**, match perfectly to yield the inductive step.

▷ . . . . ◁

Next we turn to a proof of the maximality/truth-like-set lemma, **2F-7**. We rely heavily on the properties stated in Theorem **2E-22**.

PROOF. Suppose G* is maximal E-free, so that $MEF_{S_{TF}}1$ and $MEF_{S_{TF}}2$ of **2E-11** hold. We need to show TL$\sim$ and TL$\supset$ of **2F-2**.

First show TL$\sim$ left-to-right. Suppose $\sim A \in G^*$, and then suppose $A \in G^*$ for reductio; hence $\sim A, A \subseteq G^*$. But $\sim A, A \vdash_{S_{TF}} E$ by $S_{TF}5$, **2E-19**, so $G^* \vdash_{S_{TF}} E$ by $\vdash_{S_{TF}}$weak. This contradicts $MEF_{S_{TF}}1$.

Next, for TL$\sim$ right-to-left, suppose A $\notin$ G*, and then suppose $\sim$A $\notin$ G* for reductio. Obviously G* $\subset$ (G*$\cup$\{A\}) and G* $\subset$ (G*$\cup$\{$\sim$A\}), so by MEF$_{S_{TF}}$2, G*, A $\vdash_{S_{TF}}$ E and G*, $\sim$A $\vdash_{S_{TF}}$ E. So by the Deduction theorem, **2E-15**, G* $\vdash_{S_{TF}}$ A$\supset$E and G* $\vdash_{S_{TF}}$ $\sim$A$\supset$E. Now S_{TF}6 (**2E-19**) gives A$\supset$E, $\sim$A$\supset$E $\vdash_{S_{TF}}$ E; so G* $\vdash_{S_{TF}}$ E by $\vdash_{S_{TF}}$cut; which contradicts MEF$_{S_{TF}}$1.

Next, to show TL$\supset$ left-to-right, suppose A$\supset$B$\in$G*, and further suppose the antecedent of the right side of TL$\supset$, A$\in$G*. Hence G* $\vdash_{S_{TF}}$ A$\supset$B and G* $\vdash_{S_{TF}}$ A by $\vdash_{S_{TF}}$id. These, by $\vdash_{S_{TF}}$MP (**2E-5**), give G* $\vdash_{S_{TF}}$ B. So B$\in$G* (which is the consequent of the right side of TL$\supset$) by maximality/S_{TF}-closure, **2E-13**.

Lastly, for TL$\supset$ right-to-left, suppose that A$\in$G* $\rightarrow$ B$\in$G*. In preparation for an lapse from elegance, we prepare for an ugly detour through negation by equivalently supposing that either A $\notin$ G* or B$\in$G*. Suppose first the left disjunct holds: A $\notin$ G*. Hence $\sim$A$\in$G* by TL$\sim$, established above, so G* $\vdash_{S_{TF}}$ $\sim$A by $\vdash_{S_{TF}}$id. So by S_{TF}5 and $\vdash_{S_{TF}}$cut, G*, A $\vdash_{S_{TF}}$ B. Hence G* $\vdash_{S_{TF}}$ A$\supset$B by the Deduction theorem. The second disjunct, B$\in$G*, also gives G* $\vdash_{S_{TF}}$ A$\supset$B, using $\vdash_{S_{TF}}$id, S_{TF}1, and $\vdash_{S_{TF}}$MP, so that G* $\vdash_{S_{TF}}$ A$\supset$B in either case. But then A$\supset$B$\in$G* by maximality/S_{TF}-closure □

At this point all we have left to prove is Lindenbaum's lemma, **2F-6**. This is the part of the proof that uses ideas most distant from those with which we started (the two turnstiles). But first some exercises.

Exercise 33 *(Consistency and completeness of* S_{TF}*)*

The chief thing to master at this point is the proof of consistency and completeness of S_{TF}, both forest and such trees as have been encountered. But here are some other things to do.

1. Give an example of a sentence E such that no G is E-free. Explain.
2. Give an example of a G that is not E-free for any E. Explain.
3. This exercise is a bit tedious, but clarifying: Show that every truth-like set is closed under S_{TF}-consequence: If G* is a truth-like set, then for all A, if G* $\vdash_{S_{TF}}$ A then A$\in$G*. Use Induction on S_{TF}-consequence. The most tedious but most revealing part of the argument will be the portion of the base case that shows that every member of AxS_{TF} is a member of G*.

4. Use the previous exercise in order to help show that every truth-like set, G*, is maximal E-free for some E. (You may find it natural to start with the fact that there are some sentences, and let A be such a sentence, and consider whether A is in or out of the truth-like set G*.)

5. For your choice of E, give an example of an E-free set that is not maximal E-free, and explain.

6. Figure out the appropriate clauses for the definition of "truth-like set" (**2F-2**) when conjunction and disjunction and the biconditional are part of the grammar.

7. Optional: Figure out an argument for the TL⊃ part of the maximality/truth-like-set lemma that will avoid the inelegant detour through negation, at the cost of adding a new family of members to AxS_{TF}, namely, a cousin, $(A \supset E) \supset [((A \supset B) \supset E) \supset E]$, of Peirce's Law. (Peirce's Law itself is $((A \supset B) \supset A) \supset A$.)

▷ . . . . ◁

2F.3 Quick proof of Lindenbaum's lemma

Lindenbaum's lemma, **2F-6**, is mostly a set theoretical fact. It is easy to prove if you are sufficiently relaxed about what counts as a proof, and you will find many; here's another.

PROOF. Line up all the sentences. Starting with the E-free set G to be extended, add each sentence in its turn just in case its addition to what you already have does *not* lead to E; let G* be the result of this process. Clearly G* satisfies the "maximality" part of the lemma, since if we failed to put a sentence in G* it was because it led to E. And it is only a little harder to see that G* itself is also E-free. Suppose it isn't; then by $\vdash_{S_{TF}}$fin, $G_0 \vdash_{S_{TF}} E$ for some *finite* subset G_0 of G*. G_0 must consist entirely of sentences either in G, or added along the way. But G is by hypothesis E-free, and so is each result of addition; so G_0 must also be E-free; which is a contradiction □

If you found this proof convincing, you have good intuitions. If you didn't, draw yourself some pictures. But in either event, were you to spell out all the steps, substituting rigorous language for pictorial language ("line up all the sentences," "add each sentence in turn," "the result of this process," etc.), you would be into some set theory you hadn't expected.

Here it is; then we will re-present the argument for Lindenbaum's lemma, filling in a few steps.

2F.4 Set theory: chains, unions, and finite subsets

Most of the set theory we need revolves around the concept of the finite, as treated in §2B.3; you may wish to review that section.

We are also going to need a fact relating (possibly infinite) unions, chains, and finiteness. First we enter the appropriate definition. Keep in mind that Z is two levels up: a set of sets—of whatever; but our applications will be to sets of sets of sentences. (We will informally reserve "Z" for this level.)

2F-10 DEFINITION. *(Chain)*

A family Z of sets is a *chain* ↔ every pair of its members is comparable by the subset relation: $(X_1)(X_2)[(X_1 \in Z \text{ and } X_2 \in Z) \rightarrow (X_1 \subseteq X_2 \text{ or } X_2 \subseteq X_1)]$.

There are two useful pictures of chains of sets. The first arranges them horizontally, in a chain, with the subset relation forming the links: $\ldots X_1 \subseteq \ldots \subseteq X_2 \ldots$. The second thinks of Euler diagrams, and thus arranges the members of Z in a giant array of concentric circles; with the second picture in mind, chains of sets are sometimes called "nests."

We need the notation for taking the union of an entire family of sets (not just two, or a few) that was provided in NAL:Axiom **9B-21**, here repeated: $\bigcup Z$, the *union* of a family of sets Z, is the set of members of members of Z: $\bigcup Z$ is a set, and

$$x \in \bigcup Z \leftrightarrow \exists Y (x \in Y \text{ and } Y \in Z)$$

That is, $\bigcup Z$ lumps together all the members of members of Z. Pictorially, if

$$Z = \{X, X', \ldots Y, Y', \ldots\}$$

then

$$\bigcup Z = (X \cup X' \cup \ldots \cup Y \cup Y' \cup \ldots)$$

For example,

$$\bigcup\{\{1,2\},\{1,4\},\{4,5,6\}\}=\{1,2,1,4,4,5,6\}=\{1,2,4,5,6\}$$

So $\bigcup$ is the Great Eraser, causing the deletion of lots of curly brackets. Or suppose you draw Z as a Big Circle with its members (each a set) as Little Circles inside it, and with some dots inside the little circles to represent their members. Then you obtain a picture of $\bigcup$Z by erasing the Little Circles, leaving only their dots inside the Big Circle. So again $\bigcup$ is the Great Eraser.

Notice that $\bigcup$Z is "down" a level from Z, and is a set on the "same level" as the *members* of Z. Obviously if $Z=\{X_1, X_2\}$, then $\bigcup Z=(X_1\cup X_2)$.

Now for the fact mentioned above.

2F-11 FACT. *(Finite subsets of unions of nonempty chains)*

If Z is a nonempty chain of sets, then every finite subset of $\bigcup$Z is a subset of a member of Z. That is, (Z is nonempty and Z is a chain and X is finite and $X\subseteq \bigcup Z) \rightarrow \exists Y(X\subseteq Y$ and $Y\in Z)$.

For "nonempty" see NAL:**9A-11**.

This is one of those facts that seems intelligible only *after* one has seen a proof; before that it just sounds like doubletalk. (Some people "see" the fact when it is stated in terms of "nests" instead of "chains"; you might try this.)

Exercise 34 *(Finite subsets of unions of nonempty chains)*

1. Prove Fact **2F-11**. First rearrange, in preparation for an induction on finite sets, to say that if Z is nonempty, and is a chain, then for every X, if X is finite, then: If X is a subset of $\bigcup$Z then X is a subset of some member of Z. Now begin by choosing a nonempty chain Z. Show that X is finite $\rightarrow$ [$X\subseteq\bigcup Z \rightarrow \exists Y(X\subseteq Y$ and $Y\in Z)$] by Induction on finite sets, **2B-9**, choosing $\Psi(X)$ there as the bracketed part here.
2. Optional but strongly suggested. Show by three informal examples that the Fact fails if you omit any one of the three conjuncts in the antecedent of its rearranged form, i. e., any one of (1) Z is nonempty, (2) Z is a chain, and (3) X is finite.

▷ . . . . ◁

2F.5 Slow proof of Lindenbaum's lemma

So much for the set theoretical background. Now for a somewhat more articulated proof of Lindenbaum's lemma **2F-6**.

Suppose as *hypothesis of the lemma* **2F-6** that $G \nvdash_{S_{TF}} E$.

By **2C-15** the sentences are countable,[26] so we may rely on Corollary **2C-12** to choose an enumeration B of the sentences such that each sentence is B_n for some $n \in N$:

2F-12 LOCAL CHOICE. *(Enumeration of sentences)*

$A \in Sent \leftrightarrow \exists n(n \in N \text{ and } A = B_n)$

Relative to G and to E and to the enumeration B of **2F-12**, inductively define a sequence $G_0, \ldots,$ of sets of sentences as follows.

2F-13 LOCAL DEFINITION. *(G_n)*

For each $n \in N$, $G_n \subseteq Sent$. ($G_n$type)

Basis clause. $G_0 = G$.

Inductive clause.

$G_{n+1} = G_n$ *if*: $G_n, B_n \vdash_{S_{TF}} E$.

$G_{n+1} = (G_n \cup \{B_n\})$ *if*: $G_n, B_n \nvdash_{S_{TF}} E$.

Note that as usual the Basis tells us how to get started, while the Inductive clause tells us how to continue. And note that the meaningfulness of "G_n" depends on the previous Local choice **2F-12** of an enumeration of the sentences.

Let us pause to establish two facts about this sequence. (The technical legitimacy of the definition itself—even its grammar—is something we omit.)

2F-14 LOCAL FACT. *(G_n increasing)*

---

[26] There are alternative proofs that do not require the sentences to be countable; they instead rely on the axiom of choice as discussed in §3E.3, or an equivalent as laid out in §2F.6 immediately below.

1. $G_n \subseteq G_{n+1}$.
2. $m \leqslant n \rightarrow G_m \subseteq G_n$.

That is, later sets are supersets of earlier sets.

Exercise 35 *(Proof that* G_n *is increasing)*

1. Prove **2F-14**(1) first.
2. Then rely on **2C-6** to restate **2F-14**(2) as "$k \in N \rightarrow (m)(n)[m+k=n \rightarrow G_m \subseteq G_n]$," and use Induction on N, **2C-4**, with "k" the inductive variable. By this we mean that you are to restate **2C-4** with "k" in place of "n"; then choose $\Psi(k)$ of your restatement as what is here to the right of the main arrow.

▷ . . . . ◁

2F-15 LOCAL FACT. *(*G_n *E-free)*

For each $n \in N$, G_n is E-free in S_{TF}; that is, $G_n \nvdash_{S_{TF}} E$.

Exercise 36 *(*G_n *E-free)*

Prove **2F-15**. Use Induction on N, **2C-4**. The hypothesis of Lindenbaum's lemma together with the Basis of **2F-13** guarantees this for G_0; and the Inductive clause of **2F-13** clearly suffices to show that G_{n+1} is E-free in S_{TF} if G_n is so.

▷ . . . . ◁

Next let us fix our attention on the family of all the G_n:

2F-16 LOCAL DEFINITION. *(Z is the set of all the* G_n*)*

Z = the set of all the G_n; that is,

$$H \in Z \leftrightarrow H = G_n, \text{ some } n \in N.$$

(The standard axioms of set theory imply the existence of a set satisfying the given condition.)

We are going to show that the G* promised by Lindenbaum's lemma can be chosen as $\bigcup Z$. The following prepare the way.

2F-17 LOCAL FACT. *(Lemma for* LL1*)*

$G_n \subseteq \bigcup Z$, all $n \in N$.

2F-18 LOCAL FACT. *(Z a nonempty chain)*

Z is a nonempty chain.

To complete the proof of Lindenbaum's lemma, we need to show that $\bigcup Z$ satisfies LL1, $MEF_{S_{TF}}1$, and $MEF_{S_{TF}}2$ of **2E-11**.

2F-19 LOCAL FACT. *(*LL1 *for* $\bigcup Z$*)*

$G \subseteq \bigcup Z$.

Exercise 37 *(Some facts about* Z*)*

Prove Local facts **2F-17**, **2F-18**, and **2F-19**. For **2F-18**, combine **2F-14** with **2C-7**.

▷ . . . . ◁

2F-20 LOCAL FACT. *(*$MEF_{S_{TF}}$ *1 for* $\bigcup Z$*)*

$\bigcup Z$ is E-free in S_{TF}.

PROOF. Suppose not; that is, suppose that $\bigcup Z \vdash_{S_{TF}} E$. Then by $\vdash_{S_{TF}}$fin, **2E-9**, let H be a finite subset of $\bigcup Z$ such that $H \vdash_{S_{TF}} E$. But Z is a nonempty chain (Local fact **2F-18**); so by the set theoretical Fact **2F-11**, $H \subseteq G_n$, some $n \in N$. So $G_n \vdash_{S_{TF}} E$ by $\vdash_{S_{TF}}$weak, **2E-9**; which contradicts that all the G_n are E-free (Local fact **2F-15**) □

2F-21 LOCAL FACT. *(MEF$_{S_{TF}}$ 2 for $\bigcup$Z)*

No proper superset of $\bigcup Z$ is E-free.

PROOF. Choose a proper superset H of $\bigcup Z$, and let $C \notin \bigcup Z$ but $C \in H$. C must by Local choice **2F-12** be enumerated somewhere, so let $C = B_n$: $B_n \notin \bigcup Z$ and $B_n \in H$.

Since $B_n \notin \bigcup Z$, Local fact **2F-17** guarantees $B_n \notin G_{n+1}$; so obviously $G_{n+1} \neq (G_n \cup \{B_n\})$. Hence, by the Inductive clause of **2F-13**, $G_n, B_n \vdash_{S_{TF}} E$. But $G_n \subseteq \bigcup Z$ (by **2F-17**) $\subseteq H$ (by hypothesis); so, since $B_n \in H$, $(G_n \cup \{B_n\}) \subseteq H$.

So $\vdash_{S_{TF}}$weak gives $H \vdash_{S_{TF}} E$, as was to be shown. □

Thus, since $\bigcup Z$ is a superset of G (**2F-19**) and is maximal E-free (**2F-20** and **2F-21**), the proof of Lindenbaum's lemma **2F-6** is complete.

Exercise 38 *(Lindenbaum lemma)*

Provide a proof for each item that is needed in order to prove Lindenbaum's lemma. Let Exercise 37 stand as part of your work. Be sure that you have proved each item, whether or not there is a proof given in these notes. Be sure to indicate how the pieces fit together. (In the case of items for which a proof is given herein, try to restate the matter in your own words—or symbols.) ▷ . . . . ◁

2F.6 Lindenbaum and Zorn

Earlier we said that Lindenbaum's lemma is largely a fact of set theory, and we also said that it could be proved without the assumption of the countability of the sentences, provided something further is postulated in the way of set theory. Here we sketch an alternative proof of Lindenbaum's lemma that makes good on both of these claims.

"Zorn's lemma" is a famous provable equivalent of the even more famous axiom of choice (§3E.3), but, even though it is historically a "lemma," since we are working on logic rather than set theory, we just enter it as an axiom for this section.

2F-22 AXIOM. *(Zorn's lemma)*

Let Z* be any set of sets. Suppose

Z* is nonempty.

The union of every nonempty chain of sets in Z* is itself in Z*; that is, for every Z, [(Z⊆Z* and Z is nonempty and Z is a chain → ⋃Z∈Z*)].

Then Z* has a maximal member, i.e., a member no proper superset of which is itself a member of Z*: There is a G* such that G*∈Z* and (H)[G* ⊂ H → H ∉ Z*].

In order to use Zorn's lemma, we need a set of sets to begin with. Since our aim is to prove Lindenbaum's lemma, **2F-6**, we suppose its antecedent, that G $\nvdash_{S_{TF}}$ E, and enter a

2F-23 LOCAL DEFINITION. *(*$Z^{E, G}$ = *set of* E*-free supersets of* G*)*

H∈$Z^{E, G}$ ↔ G⊆H and H $\nvdash_{S_{TF}}$ E.

Now we may use Zorn's lemma with $Z^{E, G}$ for Z*; to do so, we must show that the hypotheses of Zorn's lemma hold for $Z^{E, G}$. First we must establish that $Z^{E, G}$ is nonempty; this is easy, because that is precisely the antecedent of Lindenbaum's lemma that we have given ourselves. Second and last, we must be sure that the union of every nonempty chain of sets in $Z^{E, G}$ is itself a member of $Z^{E, G}$ (see the statement of Zorn's lemma, **2F-22**, for a spelling out of this). The proof of this second hypothesis for Zorn's lemma closely follows the proof that we already gave for **2F-20**, except that there we had to work to prove the chain condition, while here we are given it. What is crucial in both proofs is the use of $\vdash_{S_{TF}}$fin from **2E-9**, and the use of the set theoretical fact **2F-11**. With the two hypotheses of Zorn's lemma established, we may use Zorn's lemma to conclude to the existence of a maximal member, say G*, of $Z^{E, G}$. That is, by **2F-23**, and the spelling out of "maximal member" in the conclusion of Zorn's lemma, we conclude that LL1, $MEF_{S_{TF}}1$ and $MEF_{S_{TF}}2$ hold for the G* promised by the conclusion of Zorn's lemma. This completes the proof that Lindenbaum's lemma is indeed true.

This proof is less intuitive than the one of the previous section, but it has its own merits. First, the Local definition **2F-23** on which it relies is much easier to justify rigorously than is the Local Definition **2F-13** of G_n used in that other proof—we have swept much under the rug by "omitting" the justification of **2F-13**. Second, this proof separates to a much larger degree the purely set theoretical from the proof-theoretical considerations (only $\vdash_{S_{TF}}$fin and $\vdash_{S_{TF}}$weak come from proof theory). Third, this proof does not invoke concepts of arithmetic as auxiliaries, either generally for the idea of countability or in detail for the facts in section §2C.1. Somehow it is better if the completeness of S_{TF} is not made to seem to depend on the fact that all numbers m and n are such that either $m \leqslant n$ or $n \leqslant m$.

Exercise 39 *(Zorn and Lindenbaum)*

Optional. The foregoing proof of Lindenbaum's lemma is compressed and allusive; recast it rigorously.

▷ . . . . ◁

2G Consistency/completeness of S_{TF} and its corollaries

Upshot:

2G-1 THEOREM. *(Consistency/completeness of* S_{TF}*)*

$G \vdash_{S_{TF}} A \leftrightarrow G \vDash_{TF} A$

PROOF. Theorem **2F-1** gives consistency. Completeness is an immediate corollary of Lindenbaum's lemma, **2F-6**, (the proof of which was given in §2F.5 via the maximality/truth-like-set lemma (**2F-7**), the truth-like set/canonical TF-interpretation lemma (**2F-8**), and the Course of proof of S_{TF} completeness (**2F-9**). □

2G-2 COROLLARY. *(Connectives and* $\vdash_{S_{TF}}$ *corollary)*

The Conjecture **2E-17** relating connectives and $\vdash_{S_{TF}}$ is established, since all these properties transfer from Fact **2D-31** relating connectives and $\vDash_{TF}$.

Of course the only *new* information here relates to $\sim$, since the Deduction theorem and its converse were used in obtaining the very theorem (**2G-1**) underlying the transfer.

2G.1 Finiteness and compactness

2G-3 COROLLARY. *(Finiteness property for* $\vDash_{TF}$*)*

The conjecture $\vDash_{TF}$fin (**2D-30**) is established by transference from $\vdash_{S_{TF}}$fin, **2E-9**.

This corollary is of considerable interest; note that unlike $\vdash_{S_{TF}}$fin, it is not at all "built in" to the definition of $\vDash_{TF}$ that finitely many premisses always suffice. Nor does it follow from the local determination lemma, **2D-18**, working together with the Finitude of TF-atoms, **2B-25**, as you might suppose. Instead, it really depends on there being "enough" interpretations, as the following contrived example shows.

Define "TF-interpretation$_{(p)}$" as a certain subset of TF-interpretation, **2D-4**, namely, as the set of all TF-interpretations **i** such that if $\mathbf{i}(q) = T$ for every TF-atom q other than p, then $\mathbf{i}(p) = T$ as well. This definition does not alter the truth of the Local determination lemma, **2D-18**. Now define $\vDash_{(p)}$ like $\vDash_{TF}$ (**2D-24**), except quantify over TF-interpretation$_{(p)}$ instead of over TF-interpretation. Clearly

$$(\text{TF-atom} - \{p\}) \vDash_{(p)} p,$$

but this statement surely does not hold for *any* proper subset of (TF-atom $-$ {p}), much less for any finite subset.

There is an alternative and elegant proof of $\vDash_{TF}$fin that does not detour through $\vdash_{S_{TF}}$. The idea is to define

$$G \vDash_{fin} A \leftrightarrow \text{there is a finite subset } G_0 \text{ of } G \text{ such that } G_0 \vDash_{TF} A.$$

Then show that each of the parts of Theorem **2E-22** except Induction on S_{TF}-consequence holds for $\vDash_{fin}$. Since these were the only properties of $\vdash_{S_{TF}}$ we used for completeness, precisely the same argument will then establish

$$G \vDash_{TF} A \rightarrow G \vDash_{fin} A$$

which is just $\vDash_{TF}$fin.

What is nice about this alternative proof is that it gives a purely semantic proof of a purely semantic theorem, with no detour through proof theory. Such ideas have led to the contemporary flowering of the discipline of "model theory," the aim of which is to plumb the depths of semantic ideas without admitting the existence of proof theory.

Exercise 40 *($\vDash_{fin}$)*

Optional. Show that $G \vDash_{TF} A \rightarrow G \vDash_{fin} A$ by the method discussed above; or at least attack some parts of the problem; e.g., prove analogs of $\vdash_{S_{TF}}$weak and $\vdash_{S_{TF}}$cut, **2E-9**, for $\vDash_{fin}$.

▷ . . . . ◁

Finally, note that $\vDash_{TF}$fin, **2D-30**, is equivalent to what is often called "*compactness*": If every finite subset of G is (semantically) consistent, then so is G itself.

Exercise 41 *(Compactness)*

First convince yourself (by a proof) of the equivalence of compactness and $\vDash_{TF}$fin, **2D-30**. For this purpose, observe—if you haven't already done so—that to say that a set of sentences G is consistent is to say that for every A, G is (A&∼A)-free.

▷ . . . . ◁

Then reflect that the property is by no means obvious; for the premiss says that for each finite subset there is a TF-interpretation (perhaps different for each choice of finite subset) that makes the subset true, while the conclusion says there is a single TF-interpretation that makes the whole set true. (The term "compactness," incidentally, arises out of associations with topology.)

2G.2 Transfer to subproofs

We began §1A by saying that we were going to study the tools you already were using; perhaps you therefore felt a bit cheated when we took S_{TF} as our proof theory, instead of Fitch subproofs as developed in *Notes on the Art of Logic*. We want here partly to repair the damage done to your expectations; these sections, however, are not required for subsequent developments.

First a name: We use "Fi_{TF}" for the upcoming result of our redescription of Fitch's system. And now reflect: What does a step in a Fitch subproof of "mean"? Clearly writing it down doesn't mean it's *true*, for we are allowed to write down anything we like, as long as we "protect" it by signifying it as a hypothesis. If you think along these lines for very long, you will soon conclude that the whole geometry of the Fitch proof plays a role in determining what each step means, and that in fact writing down a step "means" that the sentence you write down follows from—from what? From the sequence of all "available"—in the technical sense—hypotheses.

That is, the statement made, in our use-language, by each step is that it follows from the hypotheses under which it is written.

This suggests another notation for the same statement, namely the turnstile notation. So let us enter the following definition, at least for starters: Where S is a (finite) sequence, $S \vdash^{Fi}_{TF} A \leftrightarrow$ there is a Fitch subproof containing a certain step such that (1) A is the sentence recorded at that step, and (2) S is the sequence of all hypotheses "available" at that step, the hypotheses to be listed in order, from the outside to the inside.

Note the switch from sets to sequences on the left of the turnstile; this is because of the nature of the Fitch system itself. Both the rule of reiteration, and the rule CP, refer not to an arbitrary available hypothesis, but to a hypothesis identified as "innermost." So the notation must reflect that feature of the situation.

What we have given is something like our first definition of "$\vdash_{S_{TF}}$" in terms of derivations, **2E-4**; let us immediately turn to giving the proper inductive definition of $\vdash^{Fi}_{TF}$ analogous to Definition **2E-5**. The idea is that we have just five rules: hyp, reit, MP, CP, and RAA. (Since $\perp$ is not in the vocabulary of TF, we are thinking of RAA in the form that obtains A from a subproof with hypothesis $\sim$A and containing two steps B and $\sim$B.) What we want to do is to cast these five rules into turnstile form.

Exercise 42 *(Transfer to subproofs I)*

Try it yourself.

▷ . . . . ◁

To prepare the way, we are going to present a small Fitch proof, in fact a proof of $S_{TF}3$, and then show how each line can be translated into the $\vdash^{Fi}_{TF}$ notation. Please note the form of RAA used.

1	| $\underline{\sim A \supset \sim B}$	hyp
2	| | $\underline{B}$	hyp
3	| | $\sim A \supset \sim B$	1, reit
4	| | | $\underline{\sim A}$	hyp
5	| | | $\sim A \supset \sim B$	3, reit
6	| | | $\sim B$	4, 5 MP
7	| | | B	2, reit
8	| | A	4–(6, 7) RAA
9	| $B \supset A$	2–8, CP
10	$(\sim A \supset \sim B) \supset (B \supset A)$	1–9, CP

This is the translation, where the sentence itself is on the right of the turnstile, while listed to the left is the sequence of all the available hypotheses, in order.

1	$\sim A \supset \sim B$	$\vdash^{Fi}_{TF} \sim A \supset \sim B$
2	$\sim A \supset \sim B, B$	$\vdash^{Fi}_{TF} B$
3	$\sim A \supset \sim B, B$	$\vdash^{Fi}_{TF} \sim A \supset \sim B$
4	$\sim A \supset \sim B, B, \sim A$	$\vdash^{Fi}_{TF} \sim A$
5	$\sim A \supset \sim B, B, \sim A$	$\vdash^{Fi}_{TF} \sim A \supset \sim B$
6	$\sim A \supset \sim B, B, \sim A$	$\vdash^{Fi}_{TF} \sim B$
7	$\sim A \supset \sim B, B, \sim A$	$\vdash^{Fi}_{TF} B$
8	$\sim A \supset \sim B, B$	$\vdash^{Fi}_{TF} A$
9	$\sim A \supset \sim B$	$\vdash^{Fi}_{TF} B \supset A$
10		$\vdash^{Fi}_{TF} (\sim A \supset \sim B) \supset (B \supset A)$

So now you should be able to transcribe the Fitch rules into turnstile form.

Exercise 43 *(Transfer to subproofs II)*

Try again. To see how, annotate the above 10 lines, using the original subproof as a guide.

▷ . . . . ◁

We are going to give these rules as part of an inductive definition of "$\vdash^{Fi}_{TF}$"; but first, because what appears on the left of the turnstile must name a sequence rather than a set, we need a little conceptualization of sequences.

2G.3 Set theory: sequences, stacks, and powers

We are going to need some small theories of (finite) sequences. There are several standard ways to be rigorous about this topic. Since each is more convenient for some purposes and less so for others (even though any one can be made to do all the work if required), we will, in spite of a certain cost in tedium, pursue two.

One standard technique is to define a "sequence" as a function whose domain is constituted by some standard finite set, say by a set consisting of all the predecessors of some number n (knowing as we do that 0 has 0 predecessors, 1 has 1 predecessor, etc.); in fact, let us reserve "finite sequence" for the creature of this technique (in context we sometimes omit "finite"). With this technique a typical sequence of length three might look like

$$\{\langle 0, a\rangle, \langle 1, b\rangle, \langle 2, b\rangle\}.$$

(See NAL:**9B-8** for the conception of a function as a set of ordered pairs.) It was with this technique in mind that we earlier defined "predecessors-of(n)," **2C-6**, as the set of all predecessors of n, so that we may now enter the following

2G-4 DEFINITION. *(Finite sequence; X^n)*

- x *is an* n-*ary sequence of* X $\leftrightarrow$ x $\in$ (predecessors-of(n) $\mapsto$ X).
- "*is a finite sequence of* X," "*is an* n-*ary sequence*," and "*is a finite sequence*" are defined by appropriate existential quantification from the above.
- Variant: X^n = (predecessors-of(n) $\mapsto$ X).

We use the concept of sequence indicated above when we want a handy name for the length of the sequence ("n" will do), or when we want handy names for each member of the sequence indifferently (x_k is via **2C-11** "the k-th" member of an n-ary sequence x when $0 \leqslant k < x$).[27]

A second method is to define sequences as resulting by repeated use of the ordered pair construction, **9B-1**, so that a typical sequence of length three might look like

[27] The generalization to "a set of X indexed by I," i.e. a member y of (I $\mapsto$ X), is useful when the order of the chosen elements of X has no conceptual point, or when we don't much care how many there are. We can still use y_i, for $i \in I$, as a name for an arbitrary indexed item.

$\langle\langle a, b\rangle, b\rangle$

These two methods are near-enough equivalent, but a slightly modified version of the second proves more transparent when we want to keep our attention away from irrelevant features of the numbers in favor of concentrating on the inductive structure of the sequence itself.

The picture of a "sequence" created in this way by successive use of the ordered pair construction leads one to see that such "sequences" generally have the form $\langle S, a\rangle$, where S is an old "sequence" and where a is the last item added to make the sequence at hand. In this sense, the pairing mode of construction gives immediate access to only the last item, so that especially in computer science such a structure is called a "stack." We shall follow this usage. It is exactly the concept we want in order to model the sequence of hypotheses in a Fitch proof. The image to have is of the "push down stack" of dishes in the cafeteria line: To gain access to a lower plate, one must pass one-by-one through each intermediate plate.

The modification is this. Just as there is the 0-ary sequence, so it is convenient to provide an "empty" stack, one, so to speak, with no plates; we use "$\widehat{\varnothing}$" for the empty stack. The idea of the empty stack is perhaps even less intuitive than the number zero, that measures nothing, or the empty set, that members nothing; but though it stacks nothing, the empty stack has the same kind of utility in stack theory that zero has in number theory, or the empty set in set theory. The only thing we need to know about it is that it really is empty:[28]

2G-5 AXIOM. *(Empty stack:* $\widehat{\varnothing}$*)*

$\langle x, y\rangle \neq \widehat{\varnothing}$.

With the "reality" of the empty stack in mind it turns out to be useful to insist that every stack begin with the empty stack. Therefore, instead of picturing a typical stack of length three as e.g. $\langle\langle a, b\rangle, b\rangle$, you should visualize such a stack as follows:

$\langle\langle\langle\widehat{\varnothing}, a\rangle, b\rangle, b\rangle$.

You can imagine the empty stack, $\widehat{\varnothing}$, as the spring-loaded mechanism on which the cafeteria plates all sit—if you like.

[28]We could define the empty stack as some entity we already know, since we hardly care what it is (except that it should never result from the stacking (pairing) operation), but we leave such reductions to conceptually indolent Ockamists.

2G-6 DEFINITION. *(*stacks-of(X)*)*

Let X be a set; we define "stacks-of(X)" inductively.

Basis clause. $\widehat{\varnothing} \in$ stacks-of(X).

Inductive clause. If $S \in$ stacks-of(X) and $x \in X$, then $\langle S, x \rangle \in$ stacks-of(X).

Closure clause. (Induction on stacks-of(X)). Let $\Psi(S)$ be any use-context. **Suppose**

> **Basis step.** $\Psi(\widehat{\varnothing})$.
>
> **Inductive step.** $\Psi(S_1)$ implies $\Psi(\langle S_1, x\rangle$, all S_1 in stacks-of(X) and x in X.

Then (S)[S ∈ stacks-of(X) → Ψ(S)].

The definition for stacks-of(X) in words comes to this: The empty stack is a stack of X, and so is the result $\langle S, x \rangle$ of adding one more member x of X to any stack S of X. The inductive part of the definition gives us, as expected, a way of proving something about all stacks of X, so that nothing unexpected is a stack of X.

2G-7 VARIANT. *(Stack variants)*

S *is a stack of* X ↔ S ∈ stacks-of(X).

The empty stack $= \widehat{\varnothing}$.

S *is a nonempty stack of* X ↔ S is a stack of X other than the empty stack.

A *stack* is a stack of X for some set X.

Stacks give rise to a concept of "Cartesian power," that is, the result of repeatedly taking a "product" of the same thing.

2G-8 DEFINITION. *(Cartesian power)*

Basis clause. $\{\widehat{\varnothing}\}$ is in Cartesian-powers-of(X).

Inductive clause. If Z is in Cartesian-powers-of(X), then $Z \times X$ is in Cartesian-powers-of(X).

Closure clause. Let Ψ(Z) be any use-context. **Suppose** this context is satisfied by whatever is admitted by the Basis clause (this is the **Basis step**), and this context is preserved by the inductive clause (this is the **Inductive step**). **Then** (Z)[Z ∈ Cartesian-powers-of(X) → Ψ(Z)].[29]

When one wishes a notation for the collection of members of Cartesian-powers-of(X) resulting from n uses of the inductive clause, it is common to use X^n, with the understanding that $X^0 = \{\widehat{\varnothing}\}$, and so forth. To do so might be confusing here, however, since in Definition **2G-4** we introduced X^n as (predecessors-of(n) ↦ X).

Exercise 44 *(Cartesian-powers-of)*

The above definition of Cartesian-powers-of(X) is compressed; spell it out, taking special care in your formulation of the closure clause to avoid the part of the terminology used above that is either metaphorical ("preserved") or unexplained ("satisfied"). See other such definitions for guidance.

▷ . . . . ◁

The last thing we want to do is to define what it is to be a "member" of a sequence or stack. The intuitive idea is clear; we need only enter the rigorous definitions. To avoid too much terminology, we will use "members-of" for members of both stacks and sequences, leaving it to context to disambiguate in case some unfortunate candidate falls under both cases.

2G-9 DEFINITION. *(*members-of(S)*)*

Case 1. S is an n-ary sequence of X, **2G-4**. Then x ∈ members-of(S) ↔ for some k in the predecessors-of(n), $x = S_k$.

Case 2. S is a stack of X, **2G-6**. In this case we proceed inductively.

> **Basis clause.** members-of($\widehat{\varnothing}$) = ∅;
>
> **Inductive clause.** members-of(⟨S, x⟩) = (members-of(S) ∪ {x}).

[29] You can see that we are becoming tired of expressing ourselves by means of rigorous repetition. You, however, should not become weary in this way.

2G.4 Back to subproofs

Of the two ways outlined in §2G.3 for conceptualizing sequences, in order to represent the sequence of available hypotheses in a Fitch proof, it is conceptually best to invoke the idea of a stack. The positive reasons are two. (1) In Fitch proofs we do add hypotheses one at a time, and always at the end. (2) The empty sequence of hypotheses really does play a special role (in connection with the idea of categorical proofs). The negative reasons are also conceptually important. (1) We do not need or want access to an arbitrary hypothesis; only the last-entered hypothesis plays a role in the rules. (2) We do not care how many hypotheses have been laid down. First a

2G-10 CONVENTION. *("S" for stacks of sentences)*

Here we sometimes use "stack" with the special meaning of "stacks-of(Sent)." We use "S," etc., as ranging over stacks of sentences (including the empty stack). We suppress corners "$\langle\rangle$" and "$\widehat{\varnothing}$" on the left of "$\vdash^{\text{Fi}}_{\text{TF}}$."

Now, finally, we are prepared for the central definition of proof-theoretical consequence in a sense answering to the method of subproofs.

2G-11 DEFINITION. *($\vdash^{\text{Fi}}_{\text{TF}}$)*

This is an inductive definition, with clause 1 the **basis**, and clauses2 through 5 **inductive**.

$\vdash^{\text{Fi}}_{\text{TF}}$ is that subset of (stacks-of(Sent) $\times$ Sent) ($\vdash^{\text{Fi}}_{\text{TF}}$type)
such that

1. S, A $\vdash^{\text{Fi}}_{\text{TF}}$ A ($\vdash^{\text{Fi}}_{\text{TF}}$hyp)
2. S $\vdash^{\text{Fi}}_{\text{TF}}$ A $\rightarrow$ S, B $\vdash^{\text{Fi}}_{\text{TF}}$ A ($\vdash^{\text{Fi}}_{\text{TF}}$reit)
3. S $\vdash^{\text{Fi}}_{\text{TF}}$ A$\supset$B and S $\vdash^{\text{Fi}}_{\text{TF}}$ A $\rightarrow$ S $\vdash^{\text{Fi}}_{\text{TF}}$ B ($\vdash^{\text{Fi}}_{\text{TF}}$MP)
4. S, A $\vdash^{\text{Fi}}_{\text{TF}}$ B $\rightarrow$ S $\vdash^{\text{Fi}}_{\text{TF}}$ A$\supset$B ($\vdash^{\text{Fi}}_{\text{TF}}$CP)
5. S, $\sim$A $\vdash^{\text{Fi}}_{\text{TF}}$ B and S, $\sim$A $\vdash^{\text{Fi}}_{\text{TF}}$ $\sim$B $\rightarrow$ S $\vdash^{\text{Fi}}_{\text{TF}}$ A ($\vdash^{\text{Fi}}_{\text{TF}}$RAA)

Closure: For any relational use-context Ψ(S, A), **Suppose** (1) as **Basis step** and (2)–(5) as **Inductive steps**.

1. $\Psi(\langle S_1, A_1\rangle, A_1)$
2. $\Psi(S_1, A_1) \rightarrow \Psi(\langle S_1, B\rangle, A_1)$
3. $\Psi(S_1, A_1 \supset A_2)$ and $\Psi(S_1, A_1) \rightarrow \Psi(S_1, A_2)$
4. $\Psi(\langle S_1, A_1\rangle, A_2) \rightarrow \Psi(S_1, A_1 \supset A_2)$
5. $\Psi(\langle S_1, {\sim}A_1\rangle, A_2)$ and $\Psi(\langle S_1, {\sim}A_1\rangle, {\sim}A_2) \rightarrow \Psi(S_1, A_1)$

Then $(S)(A)[S \vdash^{Fi}_{TF} A \rightarrow \Psi(S, A)]$. (Induction on Fi_{TF}-consequence)

This is the first occasion on which we have had a *relational* use-context Ψ. Reason: Unlike the inductive clause (3) for modus ponens, inductive clauses (2), (4), and (5) involve *different* stacks on the two sides of "$\rightarrow$."

2G-12 FACT. *($\vdash^{Fi}_{TF}$ and Fitch proofs)*

$S \vdash^{Fi}_{TF} A \leftrightarrow$ there is a Fitch proof of A with available hypotheses S (in order).

PROOF. Omitted. Without a rigorous definition of "Fitch proof," "available," and "in order," a rigorous proof is not possible. So we hope you are satisfied with this: It's "obvious," since the turnstile rules **2G-11** were designed to reflect precisely the Fitch system. Or you may wish to join us in taking the truth of Fact **2G-12** as a condition of adequacy on anyone's (rigorous) account of the nature of Fitch proofs
□

Now what about consistency and completeness? To state the claim rigorously, we need to be able to refer to the set of "members of" a stack S, for it is only *sets* of sentences that are related by the semantic relation $\vDash_{TF}$. We can do this by using "members-of(S)" as defined in **2G-9**; a wanted theorem is then that if $S \vdash^{Fi}_{TF} A$, then the set of members of S tautologically implies A:

2G-13 THEOREM. *(Consistency of* Fi_{TF}*)*

$S \vdash^{Fi}_{TF} A \rightarrow$ members-of(S) $\vDash_{TF}$ A.

PROOF. By Induction on Fi_{TF}-consequence; that is, using the closure clause of the definition of $\vdash^{Fi}_{TF}$, **2G-11**. Choose the relational use-context $\Psi(S, A)$ there as "members-of(S) $\vDash_{TF}$ A" here. □

Exercise 45 *(Consistency of* $\mathrm{Fi}_{\mathrm{TF}}$*)*

Prove Theorem **2G-13**.

▷ . . . . ◁

It seems awkward to prove the converse, members-of(S) $\vDash_{\mathrm{TF}}$ A $\to$ S $\vdash^{\mathrm{Fi}}_{\mathrm{TF}}$ A, directly, or even to prove that members-of(S) $\vdash_{\mathrm{S_{TF}}}$ A $\to$ S $\vdash^{\mathrm{Fi}}_{\mathrm{TF}}$ A (which, given previous results, would certainly suffice). The shortest path to completeness seems to require the following, where we recall that $\widehat{\varnothing}$ is the empty stack as characterized in **2G-5**.

2G-14 LEMMA. ($\mathrm{Fi}_{\mathrm{TF}}$ *contains* $\mathrm{S_{TF}}$ *as to theorems)*

$\vdash_{\mathrm{S_{TF}}}$ A $\to$ $\vdash^{\mathrm{Fi}}_{\mathrm{TF}}$ A. That is, $\varnothing \vdash_{\mathrm{S_{TF}}}$ A $\to$ $\widehat{\varnothing} \vdash^{\mathrm{Fi}}_{\mathrm{TF}}$ A.

PROOF. By Induction on $\mathrm{S_{TF}}$-consequence, **2E-6**. Choose G there as $\varnothing$ here and Ψ(A) there as "$\widehat{\varnothing} \vdash^{\mathrm{Fi}}_{\mathrm{TF}}$ A" here. □

2G-15 THEOREM. *(Completeness of* $\mathrm{Fi}_{\mathrm{TF}}$*)*

---

members-of(S) $\vDash_{\mathrm{TF}}$ A $\to$ S $\vdash^{\mathrm{Fi}}_{\mathrm{TF}}$ A.

PROOF. By Induction on stacks-of(sentences), **2G-6**. **Basis step.** We need to show that members-of($\widehat{\varnothing}$) $\vDash_{\mathrm{TF}}$ A $\to$ $\widehat{\varnothing} \vdash^{\mathrm{Fi}}_{\mathrm{TF}}$ A; but this comes from the completeness of $\mathrm{S_{TF}}$, **2G-1**, and the previous Lemma **2G-14**.

Inductive step. Suppose as Inductive hypothesis that members-of(S) $\vDash_{\mathrm{TF}}$ A $\to$ S $\vdash^{\mathrm{Fi}}_{\mathrm{TF}}$ A, all A. Show that members-of($\langle$S, B$\rangle$) $\vDash_{\mathrm{TF}}$ A $\to$ S, B $\vdash^{\mathrm{Fi}}_{\mathrm{TF}}$ A, all A. Now choose $\mathrm{A_1}$ and suppose members-of($\langle$S, B$\rangle$) $\vDash_{\mathrm{TF}}$ $\mathrm{A_1}$. So (members-of(S) $\cup$ {B}) $\vDash_{\mathrm{TF}}$ $\mathrm{A_1}$, by **2G-9**, so members-of(S) $\vDash_{\mathrm{TF}}$ B $\supset$ $\mathrm{A_1}$ by properties of $\vDash_{\mathrm{TF}}$; so S $\vdash^{\mathrm{Fi}}_{\mathrm{TF}}$ B $\supset$ $\mathrm{A_1}$ by the Inductive hypothesis, putting B $\supset$ $\mathrm{A_1}$ for A. Now S, B $\vdash^{\mathrm{Fi}}_{\mathrm{TF}}$ $\mathrm{A_1}$ by elementary properties of $\vdash^{\mathrm{Fi}}_{\mathrm{TF}}$. □

This proof used Induction on stacks-of(X), i.e., the closure clause of **2G-6**. The idea is that if anything is true of the empty stack, and if it is true of $\langle$S, A$\rangle$ whenever it is true of S, then it must be true of every stack of X; for clearly every such can be built by adding one member at a time. Compare Induction on finite sets, **2B-9**.

Exercise 46 *(Transfer to subproofs)*

Some exercises are wanted here. Please supply.

▷ . . . . ◁

Chapter 3

The first order logic of extensional predicates, operators, and quantifiers

By a “first order” logic or language or theory, etc., is meant a logic (etc.) that goes beyond connectives to add quantifiers, but which restricts quantification to individuals. Such a language is more complicated than TF in two independent ways. First, there is more than one basic grammatical category (at least both terms, which name things, and sentences, which have truth values) instead of only one (just sentences); and second, there are variables and ways to bind them. We discuss the grammar, semantics, and proof theory of one such language, and we establish by way of a consistency and completeness theorem that there is a perfect match between semantics and proof theory.

3A Grammatical choices for first order logic

In this section we survey some of the possible features of first order languages, making various choices as we go along, and explaining as clearly as we can our reason for each choice.

3A.1 Variables and constants

Individual variables Every such language contains an alphabet of individual variables, needed at least for quantifiers and for cross-references to them. How many should there be? The first answer arises out of consideration of applications: Because we want to be able to treat quantification statements of arbitrary degree of nesting of quantifiers, we will need infinitely many individual variables. The second answer is this: Without infinitely many variables, it would appear that the proof theory presented below is incomplete; we will call attention to the place at which the assumption of infinitely many variables appears inescapable.

It is a hallmark of first order languages that there are no other bindable variables.

Individual constants. Our aim is to idealize and theorize about various of our use-language practices. In arithmetic there are symbols like "0" and "1," and in set theory "∅," and in the theory of truth values there are "T" and "F," all of which are used to name some individual or other. These are called "individual constants"; they typically appear both in our axioms and definitions (or assumptions under any name) and in our conclusions.

How many should there be? In any one application there are likely to be only a finite number of these, even if we make infinitely many assumptions about them, though we can certainly imagine having a constant for (say) each integer. In any of these applications there may also be no individual constants at all, for some theories do not rely on naming any individual whatsoever. One might therefore make no assumption about their number; but wait.

Individual parameters. In addition and in contrast, there are symbols like "a" and "b" that we use only as "temporary names" in the course of a proof. They do not appear in either our assumptions or our conclusions. They are used only in connection with arguing from existential generalizations or arguing to universal generalizations (by whatever rules). Such symbols are often called "individual parameters" to keep them distinct both from individual variables on the one hand, which are bindable by quantifiers, and individual constants on the other, which in applications name particular entities. Or to look at the matter from the opposite point of view, individual parameters are a kind of cross between individual variables and individual constants.

How many individual parameters should there be? If we speak from their function, it is certain we shall need infinitely many, for we cannot tell how complicated our proofs will have to be and therefore how many parameters will be needed for existential instantiation or universal generalization. And we will see precisely where this assumption needs to be used in order to be sure our proof theory is strong enough (see proof of Local Fact **3E-18**).

Even so, the question as to how many parameters there should be is not simple. Some logicians in fact use the individual variables themselves to play the role of the individual parameters, so that there are, officially, *none* of the latter. Technically there could not possibly be any objection, for after all there is already a need for infinitely many variables, so that surely a few are available for existential instantiation and the like. In applications, however, and even technically, the practice leads to complications, for it requires talk about just when a variable can be instantiated, talk that is at the very best: messy.

Other logicians lump together the individual constants and the individual parameters, calling them all just "individual constants"; for even though constants and parameters function somewhat differently in applications, they are from a certain point of view treated much alike in the technical parts of grammar, semantics, and proof theory.

If we were interested more in applications here, we would keep the individual parameters separate from the constants for the sake of clarity; but in fact since we are not treating applications except by the way, we will join these later logicians in lumping together the parameters and the constants.

Accordingly, even though we may envisage that only finitely many constants will be needed to do service as "real" constants like "T" and "0," still we will need infinitely many constants for the following reason: We cannot tell how complicated our proofs will have to be and therefore how many constants will be needed to serve as "parameters" for existential instantiation or universal generalization. (To forestall confusion, we remark that we will not need to say which constants we use in which ways, since that is a matter of application on which nothing technical depends.)

Atomic terms. The "atomic terms" are just the individual variables and the individual constants and the individual parameters. Below we discuss ways of making new terms out of old; all such ways must commence with the atomic terms.

3A.2 Predicates and operators

Predicates. Predicates correspond to our use-language predicates such as "_ is finite" or "_ is a subformula of _."

We remind you that by a "predicate" we mean a function that takes singular terms as inputs and delivers sentences as outputs (NAL:Definition **1A-3** as amended here on p. 20. We shall keep to this scheme in what follows; predicates of a language we describe will map its terms into sentences. Our use-language will posit terms naming such predicates; especially since the language we discuss is "unconcretized," these terms will not name "symbols" of any kind, but instead certain grammatical functions. Such a choice is not common; but it is a good one, and you will very likely not even notice the difference.

A predicate is *logical* if our logic (grammar, semantics, proof theory) makes special assumptions about it—typically, enough assumptions so as to fix its "meaning" uniquely. The contrast is with "nonlogical" predicates; for each of these, semantics and proof theory admit their existence, and give them properties shared by all predicates of the same n-arity, but do not single out any nonlogical predicate for special mention.

Logical constants from the theoretical point of view can be distinguished from the nonlogical by the following rules of thumb. In *grammar* logical constants are ordinarily enumerated retail, and each is given a separate grammatical clause; e.g., the clauses for $\supset$ and $\sim$ in the grammar of TF. In contrast, the TF-atoms were awarded a wholesale grammatical clause. In *proof theory* the logical constants are specifically mentioned by name in the statement of the axioms or rules; nonlogical constants seldom are. This test, however, is like the others in being seldom sure; for example, a formulation might give a name "p" to some special TF-atom, say p, and choose p$\supset$p as an axiom, obtaining others (e.g. q$\supset$q for q$\neq$p) by substitution. That would indeed give p a special role, but would not in most minds qualify it as a logical constant. In *semantics* the "meaning" of the logical constants is typically given by certain of the semantic clauses—think again of $\supset$ and $\sim$. In contrast, the nonlogical constants are given their "meaning" via a variable interpretation. One might wish to say that the "meaning" of the logical constants is "the same" on every interpretation, or is anyhow independent of the interpretation; thus, $\supset$ "means" $\supset^*$, regardless of the interpretation **i**, while what a TF-atom p "means" depends on **i**.

In practice the *only* predicate about which logicians customarily make special assumptions is that corresponding to our use-language notion of identity: It is a deep and difficult question to ascertain how much that practice is a matter of historical accident and how much philosophically justified. In any event, we want to have it both ways; for a while we will not be singling out identity at all, so that if it is thought of as present it will be treated just like any other two-place predicate, but in §3G we will treat it as a logical constant. The reason for postponing its consideration is not that it much complicates the grammar, semantics, or proof theory; it doesn't. But its presence does in fact complicate our chief theorem—completeness—connecting proof theory and semantics, so that the postponement seems justified.

How many predicates? Because of envisaging a diversity of applications, we are open about how many predicates there are of each n-arity; but it has to be insisted that there is at least one of some n-arity or other, for otherwise we might wind up with no sentences at all—which would be a pity.

Operators. An "operator" is a function from terms into terms, corresponding to such use-language locutions as "_+_," "_∪_," and (when we are discussing TF) "_⊃_." Some first order languages include them, while some do without on the practical ground that they don't need them, and others do without on the theoretical grounds that they aren't necessary and complicate the technical development. It is the theoretical point we wish to take up.

The sense in which operators are dispensable is this. Suppose you thought you needed a two-place *operator* "(_+_)" with which to do arithmetic; then we point out that the three place *predicate* "(_+_)=_" will do all the work required, without doing anything extra, at least assuming identity, =, is present in the language with the operator. The chief technical point is this: Given identity, (…t…) is equivalent to $\exists$x(t=x & (…x…)), so that in particular the following are equivalent:

(…(t+u)…), and

$\exists$x([(t+u)=x] & (…x…)).

As you can see, on the right the operator appears *only* to the left of the identity, so that its separate role can be *fully* played by the *predicate* "(_+_)=_"; for example, we could introduce a *primitive* three place predicate to play that role.

There are some subsidiary technical questions in the vicinity (e.g., what axioms would we need in our proof theory to ensure that the operator and predicate versions were in fact equivalent?), but we do not pursue the matter, having made plausible the technical dispensability of operators. The technical payoff, which encourages many logicians to dispense with the dispensable, is that the grammar and semantics and (to a slight extent) the proof theory can be simplified by ignoring operators.

We shall nevertheless include operators, for these reasons. On the one hand, if we are careful the extra complications can be minimized. On the other, operators are so extremely important in our use-language that it seems a pity not to deal with them in our formal theory so as to illustrate how they can be treated in grammar, semantics, and proof theory.

There is another reason whose range of application quite outdistances our enterprise: It is bad philosophy to cater to those who would encourage us to dispense with the dispensable. The philosophical point is that in interesting cases *each thing* is dispensable, but always at some cost. The story never varies, though numerous careers (perhaps yours?) will ride on the pizzazz that emerges with the non-argument from "can-dispense-with" to "should-dispense-with."

How many operators should we have? It doesn't matter, so that we shall permit any number (including zero), and of any n-arity.

As is the case of a predicate, an operator would be "logical" if its role and "meaning" were fixed by semantics and proof theory. In practice, first order logic has employed *no* logical operators; we suppose the identity operator ($fx = x$) is the only plausible candidate. (There are many more candidates at higher orders, for instance, functional application itself.)

3A.3 Sentences, connectives, and quantifiers

Atomic sentences. Just as there are "atomic" terms that are not analyzed by the means at our disposal, so there are always "atomic" sentences, namely, any result of applying an n-ary predicate to n terms, that are not analyzable by any of our sentential modes of combination. But observe that the terms themselves can be of arbitrary complexity, so that atomic sentences can in *that* sense be complex, *as well as* being complex in the sense of arising by way of predication (shall we note that "atomic" sentences nevertheless have "parts"?).

Some languages also include some totally unanalyzed (non-predicational) atomic sentences. For a slight gain in simplicity, with some loss in perspicuity, we decline to discuss any such "wholly atomic" (that is, wholly unanalyzed) sentences.

Connectives. As with TF, we need standard ways of making new sentences out of old: connectives. As with predicates and operators, connectives can be logical or nonlogical according as to whether or not the proof theory and semantics do or do not make some special assumptions about them. In practice it is uncommon for a first order logic to treat *any* nonlogical connectives; usually the connectives are drawn up in a small, very finite list, each with its own grammatical clause, its own semantic clause, and its own axioms or rules (think again of $\supset$ and $\sim$ in TF). We shall follow the crowd.

Another way in which we move with the majority is this: In spite of the fact that many of the most interesting connectives of our use-language are non-truth-functional, such connectives will not be represented in the language we discuss; there will be nothing to correspond to "that _ is necessary" or "Hermione believes that _." Reason: only pedagogical simplicity;[1] in particular, there is no implied thought that such connectives are "dispensable" in the way that operators are.

Upshot: The only connectives will be $\supset$ and $\sim$, just as for TF; and indeed the very same connectives characterized by **2B-1**.

Elementary functors. Connectives and predicates and operators are alike in that they all take terms or sentences as arguments (inputs), and produce terms or sentences as values (outputs), though they differ as to how many and what kind of inputs, and as to what kind of output. When we wish to focus on their likeness, we will call them all *elementary functors*, an uncommon name. Some of grammar and semantics treats all elementary functors alike, which is the point of introducing a term whose lack of euphony would otherwise cause pause.

This usage of "elementary functor" differs in a subtle but important way from that of NAL:**1A-2**. There we identified an "elementary functor" as a pattern of words with blanks (etc.), which is easy to understand. Here, however, we need to be more abstract for the reasons given on p. 20.

Quantifiers. Grammatically a quantifier like $\forall$ takes two arguments: a variable in its first position and a sentence in its second. Or more helpfully, if you supply $\forall$ with a variable v, then $\forall$v is a one-place connective: $\forall v \in (\text{Sent} \mapsto \text{Sent})$, so that $\forall vA \in \text{Sent}$. But the semantic and proof-theoretic behavior of $\forall$v is so different from that of other connectives that we don't make much of this (true) grammatical fact, except for purposes of comparison with other variable-binding functors just below.[2]

Many logicians take both $\forall$ and $\exists$ as primitive. Occasionally one finds a two-place variable-binding connective, like the "formal implication" "(_ $\supset_v$ _)" of *Principia Mathematica*, equivalent to "$\forall$v(_ $\supset$ _)." Nothing else is much used;

[1] Whitehead: Seek simplicity; and distrust it.

[2] We continue occasional use of "(x)" (or even "$\forall$x") and "$\exists$x" as use-language quantifiers when perspicuity so dictates.

again the question as to why is a difficult one. We will content ourselves with just ∀, simply to have less to talk about and because it is so obvious how to define "∃" and anything else we might want.

You should note that even though ∀v is a connective, we do not count it as among the *elementary functors*; the reason is that both grammar and semantics go more smoothly if it is conceded that ∀v acts very unlike routine connectives that do not bind variables.

Atoms. We collect the bits with which we begin under the heading of "atoms": atomic terms, elementary functors, and the universal quantifier.

Sentences, terms, and formulas. In considering TF, we identified the formulas with the sentences (just below Convention **2B-4**), but the most striking grammatical fact about first order logic is that it has *two* categories of formulas: terms as well as sentences. The explanations of these sets are about what you might expect: We obtain terms from the atomic ones by means of the operators; atomic sentences come by applying predicates to terms; and we obtain more sentences by means of connectives and quantifiers. Then, finally, the formulas are characterized as the terms and sentences taken together.

3A.4 Other variable-binding functors

We are not going to put any other variable-binding functors into the language we describe, but here we offer some notes on what is being left out.

Russell's definite descriptions. "ιv(…v…)" is to be read: "the sole v such that …v… ."[3] When in fact there is such a v, and a unique one (the "existence and uniqueness conditions") there is no problem, but if one tries to be rigorous instead of just wafflous, there is an enormous problem about what to do with definite descriptions when the existence and uniqueness clauses required for their sensible use are not fulfilled. And this is a problem to which there is no simple solution. Russell's own solution declared a sentence containing such a term *false* if the conditions were not met, but as Russell knew, this "solution" leads to a further problem of "scope": Given say "∼(…ιv(…v…)…)," *which* sentence is false?—the one inside or the one beginning with the negation? All solutions to this scope problem

[3] We are using a plain ι, whereas Russell turned his upside down.

are in some way ugly, sometimes even at the grammatical level (e.g., Russell introduced explicit scope-markers). Hence, even though definite descriptions abound in our use-language, and even though it would be illuminating to treat them, we don't.

Hilbert's μ and ϵ functors. Hilbert, in the context of arithmetic, used "μv(…v…)" for "the least number v such that …v…." This functor has exactly the same problems as do Russell's definite descriptions.[4] Partly to avoid these problems, Hilbert introduced his famous epsilon functor, so that ϵv(…v…) is a term when (…v…) is a sentence. This term does not represent anything in quite ordinary English, but we like to give an approximate reading of "ϵv(…v…)" as "the paradigm v such that …v…," e. g. "the paradigm horse"; where that would be the thing which is a horse if anything is—the sole v such that: If $\exists$v′(v′ is a horse) then v is a horse. The nice thing is that anyhow *existence* is guaranteed, since it is a Q-logical-truth that $\exists$v($\exists$v′Fv′ $\supset$ Fv).

Exercise 47 *(Hilbert's epsilon)*

Prove this.

▷ . . . . ◁

There remains a *uniqueness* problem; normally there will be more than one paradigm horse, which means our suggested reading of "*the* paradigm horse" is not wholly justified. But let that go. Hilbert showed a number of interesting things about the epsilon operator, the only one of which we want to mention here is this: The existential quantifier can be defined in terms of it:

$$\exists\text{v}(\ldots\text{v}\ldots) \leftrightarrow (\ldots(\epsilon\text{v}(\ldots\text{v}\ldots)\ldots)$$

We don't want to go into much in the way of formal details, but the relevant proof theoretical fact is this: Postulation of only the "epsilon axiom," $\vdash$ [(…t…) $\supset$ (…ϵv(…v…)…)], together with the expected closure of logical truths under substitution, is enough for (…ϵv(…v…)…) to have all the properties of $\exists$v(…v…).

[4]Hilbert had a splendidly interesting solution which, unlike the program of any other logician of which we happen to know, mixed grammar and proof theory: "μv(…v…)" was counted as grammatical only after one had provided a proof of the appropriate existence and uniqueness conditions. (Since Hilbert was a Formalist, he didn't have to provide a semantics, as we Platonists do; and we are unsure what he would have done in this department.)

Lambda abstraction. Church introduced lambda abstraction; tailored to our own purposes here, given a sentence A or a term t, λvA is a predicate and λvt is an operator.[5] Hence, (λvA)t is a sentence—it is equivalent to the result of substituting t for v in A—and (λvt)u is a term—equivalent to substituting u for v in t. Adding this facility still does not take us out of the first order, for the variable-binding is first order; but if we begin to postulate that λvA (say) is itself accessible to quantification, or indeed that predicate positions generally are so accessible, then we should be leaving first order logic.

It is not a bad thing to leave first order logic, and it is a good thing to have a theory of lambda abstraction; but because we do not see much enlightenment in adding lambda abstraction in the context of first order logic, perhaps especially because it would be so simple to do so, we omit it.

Set abstraction. More used in practice than any of the above ways of variable binding is set abstraction: $\{$v: Av$\}$ is supposed to be the set of entities b such that Ab. The difference between this and first-order lambda abstraction is that the set abstract is itself a term with full status; in particular, a set abstract can stand on the left of $\in$, and falls with the range of the (first order) quantifiers used by set theory. Russell's paradox (for instance) using the set abstract $\{$x: x$\notin$x$\}$ shows that something must be done to preserve consistency, for excluded middle tells us that either $\{$x: x$\notin$x$\}\in\{$x: x$\notin$x$\}$ or not, and both alternatives seem to lead to contradiction. Set theory avoids this disaster by permitting the Church principle,

for every term t, t$\in\{$v: Av$\}\leftrightarrow$At,

only with the additional premiss that $\{$v: Av$\}$ is known to be a set. And indeed a chief business of "practical" set theory might be described as giving an account of which set abstracts really name sets. In any event, the language of which we speak does not have set theory as a part; and although we *use* set theory in these deliberations, we do not happen to employ set abstraction.

[5]These are the first cases we have had in which the output of a grammatical process has been some kind of functor (grammatical function) instead of a term or sentence. We should point out that most logicians do not or cannot distinguish between treating a use-expression such as "λvA" as denoting on the one hand a predicate, and on the other a term; in the latter case, one would have to supply the language being theorized about with a separate functor, the predication-functor, that maps a "predicate term" together with an "individual term" into a sentence. The reason most logicians cannot or will not distinguish these two is this: If we *exhibit* the language we are discussing and say no more (as is usual), then we will not be able to tell the difference. For example, given only the exhibited sequence "Fa," we cannot tell if it arose by applying to "a" the one-place predicate functor corresponding to "F_," or if it arose instead by applying to "F" and "a" the two-place predication functor corresponding to "_ _."

3A.5 Summary

1. Individual variables and individual constants make up the atomic terms.
2. And we have predicates, operators, and connectives as the elementary functors; with the result of predications being atomic sentences.
3. And there are quantifiers.
4. Atoms are constituted by the atomic terms, predicates, operators, connectives, and the universal quantifier.
5. Terms are built from the atomic terms by operators, and sentences from the atomic sentences by connectives and the quantifier.
6. (Identity is not yet treated as logical.)

3B Grammar of Q

We call our quantificational language "Q." After setting down its basic grammatical principles, we deal with substitution and some related matters, most of which are (a) uninteresting in themselves but (b) essential for our semantic and proof-theoretic deliberations.

3B.1 Fundamental grammatical concepts of Q

First we set down the fundamental axioms and definitions suggested by the discussion of §3A, with due attention to such properties as distinctness and uniqueness in analogy with §2B.2. There is, alas, an inevitable surfeit of clauses.

3B-1 AXIOM. *(*Q *primitives: their types, and how many)*

We retain "Sent," "TF-atom," "$\supset$," and "$\sim$" governed by Axiom **2B-1** and all the other axioms so far postulated. (Note: At this point we should explicitly cross-reference those axioms. Instead we enter a request that someone else do so and send us the list.) Our new primitives are "variable," "constant," "term," "n-operator," "n-predicate," and "$\forall$."

Term is a set.

Variable $\subseteq$ term. Variable is countable and infinite.

Constant $\subseteq$ term. Constant is countable and infinite.

For each $n \in N^+$, **2C-9**, n-operator $\subseteq$ (term$^n \mapsto$ term) (see **2G-4** for "term^n").

For each $n \in N^+$, n-operator is countable. (Observe that term$^n \mapsto$ term is in general *not* countable.)

For each $n \in N^+$, n-predicate $\subseteq$ (term$^n \mapsto$ Sent).

For each $n \in N^+$, n-predicate is countable.

For some $n \in N^+$, there is at least one n-predicate.

The universal quantifier: $\forall \in$ (variable $\mapsto$ (Sent $\mapsto$ Sent)).

What follows is a set of "collecting definitions," useful when we want to consider all of some population uniformly.[6]

3B-2 DEFINITION. *(Collecting definitions for* Q*)*

---

atomic term = (variable $\cup$ constant)

t is a *complex term* $\leftrightarrow$ for some n, for some f, for some $t_1, \ldots, t_n$, $t = ft_1 \ldots t_n$, where f is an n-operator and for each k ($1 \leqslant k \leqslant n$), t_k is a term.

A is a *predication* (indeed, an n-*ary* predication) $\leftrightarrow$ for some n, for some F, for some $t_1, \ldots, t_n$, $A = Ft_1 \ldots t_n$, where F is an n-predicate and for each number k (where $1 \leqslant k \leqslant n$), t_k is a term.

A is a *universal quantification* $\leftrightarrow$ for some v and B, $A = \forall vB$.

$f \in$ *operator* $\leftrightarrow$ for some $n \in N^+$, f is an n-operator. Also, n-*place operators* are n-operators.

$F \in$ *predicate* $\leftrightarrow$ for some $n \in N^+$, F is an n-predicate. Also, n-*place predicates* are n-predicates.

Q-*atom* = (atomic term $\cup$ operator $\cup$ predicate $\cup \{\supset, \sim, \forall\}$)

[6]Here we begin to use "triple dot" notation for the first time in a definition used as a basis for proofs. This notation appeals to a combination of geometrical and arithmetical intuitions and by so much does not meet our standards of rigor. In a subsequent version of these notes, the triple dot notation should be replaced either by an absolutely rigorous account of its use (which is certainly possible) or by another notation involving reference to finite sequences (§2G.3) or stacks (**2G-6**); in the meantime, you are likely to find the triple dot notation easy to understand on the intuitive level, and Convention **3B-6** spells out the greatest part of the story of what it comes to.

Non-logical Q-*atom* = (atomic term ∪ operator ∪ predicate)

Logical Q-*atom* = ⊃, ∼, ∀

An *elementary functor* is either one of ∼, or ⊃, or is an operator or a predicate. (Shortly we introduce a convention according to which we use "Φ" as ranging over elementary functors.)

formula = (term ∪ Sent)

3B-3 CONVENTION. *(Reserved letters for Q)*

We reserve letters (sometimes marked) for restricted variables as follows.

n, k	N^+ as well as N, depending on context.
v	variable
c	constant
a, b	atomic term
t, u	term
f	operator
F	predicate
A, B	sentence
C, D, E	also reserved for sentences.
O	formula (i. e., either a term or a sentence)
Φ	elementary functor

This convention is heavily used in what follows; it is best to memorize it. The chart of Figure 3.1 may help organize the various pieces of the grammar of Q for you. (If it doesn't help, skip it. If it does, pay special attention to the uses of the letters "t," "a," "c," "v.") You will see that "Q-atom" represents a cross-classification of what has already been broken down in a different way.

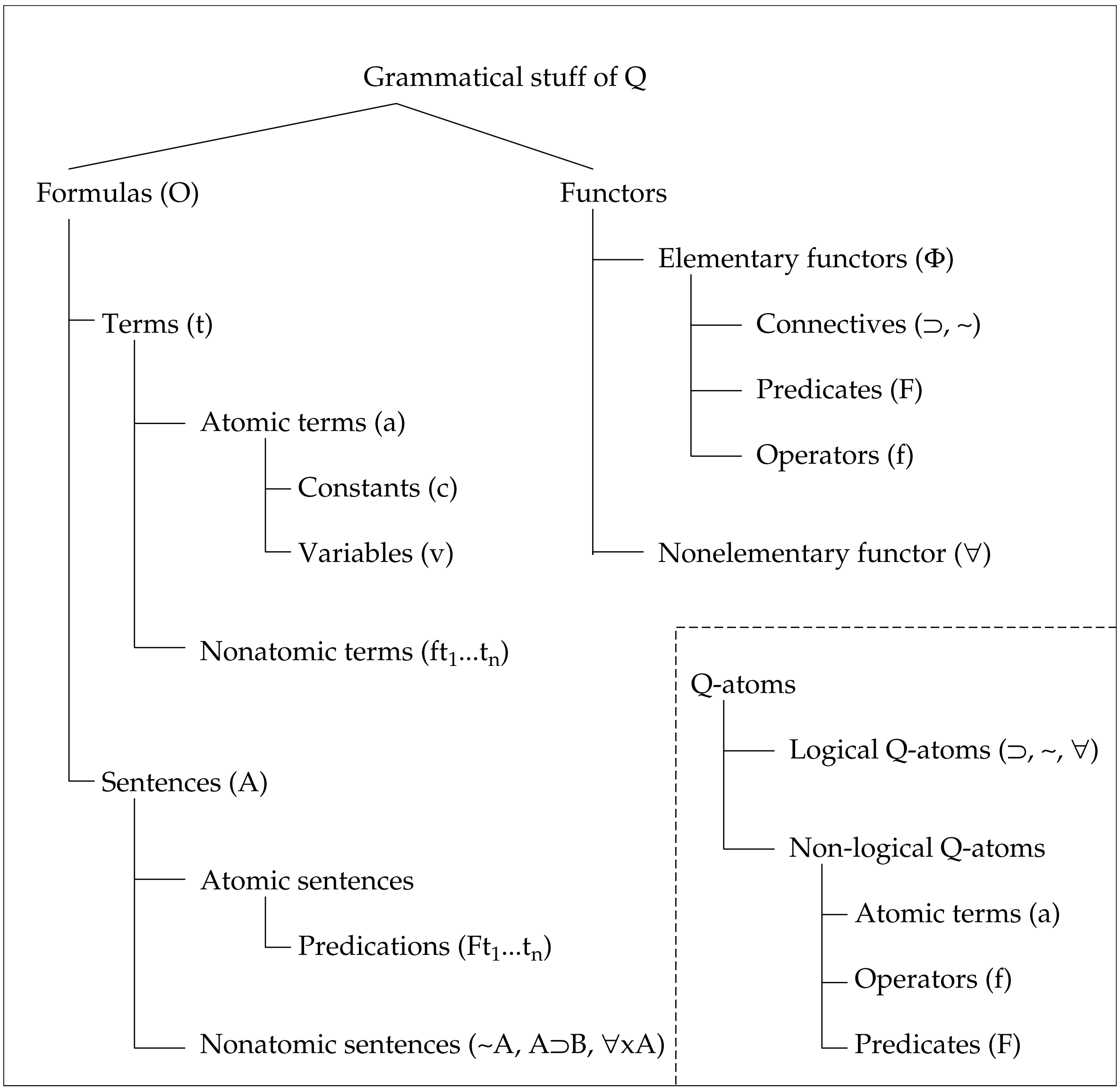

Figure 3.1: Organizing the grammar of Q

The following confers the appropriate inductive structure on the terms and sentences.

3B-4 AXIOM. *(Axiom of induction on terms, and sentences, for* Q*)*

Induction on terms for Q. Let $\Psi(t)$ be any use context, **2B-8**. **Suppose**

> **Basis step.** $\Psi(a)$ for every atomic term, **3B-2**.
>
> **Inductive step.** Provided that f is an n-operator: $\Psi(t_k)$ for all k $(1 \leqslant k \leqslant n)$ $\rightarrow \Psi(ft_1 \ldots t_n)$.

Then $\Psi(t)$ for all terms t.

Induction on sentences for Q. Let $\Psi(A)$ be any use-context. **Suppose**

> **Basis step.** $\Psi(Ft_1 \ldots t_n)$ for any n-ary predication $Ft_1 \ldots t_n$.
>
> **Inductive steps.** $\Psi(A)$ and $\Psi(B) \rightarrow (\Psi({\sim}A)$ and $\Psi(A \supset B)$ and $\Psi(\forall vA)$.

Then $\Psi(A)$ for all sentences A.

Exercise 48 *(Subproof versions of ind. on terms and sentences)*

Formulate subproof versions of Induction on terms, and of Induction on sentences, for Q, using as models previous cases in which we have turned axioms or definitions into subproof rules; e.g., **2B-18** and **2B-9**. Be thoroughly explicit about "flagging" variables.

▷ . . . . ◁

An easy consequence of Induction on terms for Q and Induction on sentences for Q, analagous to **2B-19**, is

3B-5 COROLLARY. *(Term and sentence cases for* Q*)*

> Every term is either a variable or a constant or a complex term.
>
> Every sentence is either a predication, a conditional, a negation, or a universal quantification.

Before continuing, we introduce a tiresome convention permitting us to exploit the concept of an elementary functor.

3B-6 CONVENTION. *(Φ for elementary functors)*

"Φ" is used as a variable ranging over elementary functors. Furthermore, since these functors have different n-arities, it is useful to employ a variety of conventions involving the use of triple dots signifying a run from 1 through n—that is, a quantification of some sort over this range. These conventions pretty much "go without saying," and we recommend that you read the remainder of this convention and then forget it, reading and writing triple dots in a relaxed fashion.

In the first place, in any context in which we write something like "$\Phi O_1 \ldots O_n$," it is assumed that the statement is to be prefaced with something like "n and the arguments $O_1, \ldots, O_n$ are appropriate to Φ"; that is, "if Φ is ⊃, then $n = 2$ and the arguments are sentences; if Φ is ∼, then $n = 1$ and the argument is a sentence; if Φ is an operator or a predicate, then it is an n-operator or an n-predicate and the arguments are terms." Also in the preface is a statement to the effect that O is a function whose domain is the set of numbers between 1 and n, inclusive.

Recall from **9B-10** that "$\Phi(O_1, \ldots, O_n)$" is a use-language variant of "$\Phi O_1 \ldots O_n$," both expressions signaling the result of applying the function Φ to its arguments.

As for more specific indications concerning triple dots, let the following suffice as indications of how they always represent a quantification of some sort. Let n* be (here) that subset of N^+ such that for all k, $k \in n^* \leftrightarrow (1 \leqslant k \text{ and } k \leqslant n)$.[7]

- Let Ψ(k) be any sentential use-context. Then: Ψ(1) and … and Ψ(n) ↔ for all $k \in n^*$, Ψ(k).
- Let O be a function in $(n^* \mapsto X)$. Then: $O_1 \ldots O_n$ [variant: $O_1, \ldots, O_n$] is that member of X^n such that for all $k \in n^*$, its k-th member is O_k. In other somewhat more mysterious words, $O_1 \ldots O_n = O$.

[7] Compare predecessors-of(n), **2C-6**, which is based on N instead of N^+, **2C-9**; both have n well-known "standard" members. This duplication of function ought to be eliminated in a subsequent version of these Notes, but in the meantime there may be some confusion; for instance, we defined X^n as predecessors-of(n)$\mapsto$X, but here we are instead thinking of X^n as $n^* \mapsto X$, trusting to context to disambiguate.

- Let $\Psi(k)$ be a use-language set-term. Then: $x \in (\Psi(1) \cup \ldots \cup \Psi(n)) \leftrightarrow$ for some $k \in n^*$, $x \in \Psi(k)$.

We repeat: Don't bother to learn these; they are included only to insure that in fact we could improve our degree of rigor, were we called upon to do so, but they are really too clumsy to be of much help.

This convention permits us to state a principle that we shall often use when we wish to prove something about all formulas (terms and sentences, **3B-2**).

3B-7 FACT. *(Induction on formulas of* Q*)*

Let $\Psi(O)$ be any use-language context. **Suppose**

> **Basis step.** $\Psi(a)$, for any atomic term a.
>
> **Inductive step.** $\Psi(O_1)$ and … and $\Psi(O_n) \rightarrow \Psi(\Phi O_1 \ldots O_n)$, for every elementary functor, Φ, etc.
>
> **Inductive step.** For any sentence A and variable v, $\Psi(A) \rightarrow \Psi(\forall v A)$.

Then for all formulas O, $\Psi(O)$.

Sometimes this fact is not helpful, since anyhow Convention **3B-6** has to be employed to break up the Φ-clause into its numerous parts; but when later we introduce definitions that themselves use Convention **3B-6**, the Fact does indeed prove useful. Note that it highlights the role of the universal quantifier, which has almost always to be treated specially.

PROOF. Tedious, using the definition of "formula," **3B-2**, and (separate) inductions on terms and sentences for Q, **3B-4**. Also of course the definition of "atomic term" and the Convention **3B-6** governing "Φ" will have roles to play. □

The following subproof rule is accordingly admissible:

3B-8 COROLLARY. *(Induction on formulas of Q)*

---

	flag a, for Basis	
.		
.		
.		
$\Psi(a)$	Basis	[a/O]
$\Psi(O_1)$ and…and $\Psi(O_n)$	hyp. ind., flag O, n, Φ	$[O_1/O], \ldots, [O_n/O]$
.		
.		
.		
$\Psi(\Phi O_1 \ldots O_n)$		$[\Phi(O_1 \ldots O_n)/O]$
$\Psi(A)$	hyp. ind., flag v, A	[A/O]
.		
.		
.		
$\Psi(\forall vA)$		$[\forall vA/O]$
$(O)(O$ is a formula $\rightarrow \Psi(O))$	Ind. form. Q	

If we ever want to speak of existentially quantified sentences, we can use the following

3B-9 DEFINITION. *(Existential quantifier)*

$\exists vA = \sim\forall v\sim A$

The following axiom combines the elements of "distinctness" and "one-one-ness" of **2B-20** and **2B-22** to give a general statement (in addition to those previous statements) that Q is not ambiguous.

3B-10 AXIOM. *(Distinctness and one-one for* Q*)*

Everything is distinct from everything else, and all modes of construction are one-one:

> No constant is a variable: $c \neq v$.
>
> More generally, all "kinds" of Q-atoms are distinct from one another. That is, all of the following sets have empty intersections when taken in pairs: constant; variable; n-operator, each n; n-predicate, each n; $\{\supset\}$; $\{\sim\}$; $\{\forall\}$.

No complex term is an atomic term: $ft_1 \ldots t_n \neq a$.

$ft_1 \ldots t_n = f't'_1 \ldots t'_{n'} \rightarrow f = f'$, $n = n'$, and $t_k = t'_k$, all k $(1 \leqslant k \leqslant n)$.

No term is a sentence: $t \neq A$.

No predication is a universal quantification or a conditional or a negation.

No universal quantification is a predication or a conditional or a negation.

$Ft_1 \ldots t_n = F't'_1 \ldots t'_{n'} \rightarrow F = F'$, $n = n'$, and $t_k = t'_k$, all k $(1 \leqslant k \leqslant n)$.

$\forall vA = \forall v'A' \rightarrow (v = v'$ and $A = A')$

We need to set down

3B-11 FACT. *(How many terms and sentences and formulas of* Q*)*

There are countably many terms, sentences, and formulas of Q.

There are in fact also infinitely many of each; but we record only what is needed.

PROOF. Omitted; see the remarks in the "proof" of Fact **2C-15**. □

You may ask how the grammar we have so far developed for Q is related to that previously developed for TF. The answer is that because we are speaking of the *very same* sentences, the entire structure we have so far developed remains intact; nothing is subtracted. We do, however, have *more* information than before about sentences and TF-atoms; concerning the latter, the following seems worth reporting.

3B-12 FACT. *(*TF*-atom equivalence for Q)*

A is a TF-atom ↔ A is a predication or a universal quantification.

PROOF. Straightforward. □

This equivalence highlights once again that the essential idea of a TF-atom is *truth functional* atomicity—for in Q it is just the predications and universal quantifications that are not the result of truth functional compounding (although universal quantifications may "contain" such compounding). It is a great convenience to have this concept available for the discussion of Q.[8]

[8]This convenience would not have been available to us if we had not been abstract about the TF-atoms; if, for instance, we had displayed some particular series of symbols, postulating that series

3B.2 Substitution in Q

In quantification theory one must deal somehow with substitution, for the idea is intimately associated with the notion of a bound variable; in proof theory, for example, we could not even state the rules for the quantifiers without a concept of substitution. And substitution is a bore, because so heavily notational. In the *Notes on the Art of Logic* we mostly avoided the notational difficulties by relying on you to intuit the significance of the "Av" and "At" notation, the latter signifying the result of putting (i.e., substituting) t for free v throughout the former; and indeed this procedure is entirely adequate when the aim is to be able to *use* logic (applications). But now, when we want to *prove* something about logic, and *without relying on geometrical intuitions*, we had better have an explicit notation, explicitly defined. We do not know of any way to be both rigorous and to avoid some henscratches, but we shall endeavor to introduce only as much notation as is useful.

Sometimes one wants to substitute for variables, and sometimes for individual constants. We can gain some simplicity, however, just as we did in designing our use-logic, by never substituting for bound variables. We remark that this is a matter of tactics, not strategy: Others make a different choice. On the other hand, certain considerations that one might not expect lead us to need not only the concept of retail substitution of a single term for a single atomic term, but also of wholesale substitutions for all atomic terms simultaneously. Treatment of the wholesale concept can, however, be postponed until needed, §3F.

We will want a short and workable notation for one-at-a-time retail substitution; where O is …a…, [t/a](O) will be …t…, *except* that bound occurrences of variables are always immune to substitution. (Compare our account of substitution for TF-atoms, **2C-1**.)

3B-13 DEFINITION. *(Term substitution for Q)*

For any formula O, term t, and atomic term a:

if O is a term then [t/a](O) is a term, and
if O is a sentence then [t/a](O) is a sentence. ([]O-type)

[t/a](a) = t ([]O-atom1)

---

as the TF-atoms. Such is one of the many virtues of abstractness, as Plato well knew, though not so many others have been able to keep it continually in mind. For instance, we neglected this virtue in the case of Q by setting down that all atomic sentences are predications, by so much interfering with later easy consideration of modes of combination not available in Q.

$[t/a](b) = b$, if $b \neq a$ ([]O-atom2)

$[t/a]((\Phi O_1 \ldots O_n)) = \Phi([t/a](O_1), \ldots, [t/a](O_n))$ ([]OΦ)

$[t/v]((\forall vA)) = \forall vA$ ([]O∀1)

$[t/a]((\forall vA)) = \forall v([t/a](A))$, if $a \neq v$ ([]O∀2)

Also, where G is a set of formulas, $O \in [t/a](G) \leftrightarrow$
for some $O' \in G$, $(O = [t/a](O'))$. ([]G)

The "atom" clauses are the **Basis clauses**, and the next three are **Inductive clauses**. Read "[t/a](O)" as "t for free a in O" (noting that clause []O∀1 in effect forbids substitution for unfree v). And [t/a](G) is clearly the result of simultaneously putting t for free a throughout G.

Since this is the first time we have used the "Φ" notation to say in one breath what would otherwise have to be whistled at length, let us make clear what the clause []OΦ of **3B-13** concerning elementary functors comes to.

3B-14 COROLLARY. *(Term substitution for Q for elementary functors)*

$[t/a]((\sim A)) = \sim([t/a](A))$ ([]O∼)

$[t/a]((A \supset B)) = ([t/a](A) \supset [t/a](B))$ ([]O⊃)

$[t/a]((ft_1 \ldots t_n)) = f([t/a](t_1), \ldots, [t/a](t_n))$ ([]Of)

$[t/a]((Ft_1 \ldots t_n)) = F([t/a](t_1), \ldots, [t/a](t_n))$ ([]OF)

This will help you see what is going on; but normally we will treat all of these four cases alike, and hence use the original "Φ" statement instead of the corollary.

3B.3 Occurrence, freedom, closed, and open

By this time you must be clear as to why among Famous Logicians there are Semanticists and Proof Theorists—but no Grammarians. Grammar is essential for the disciplines on which it depends, and even *interesting* for natural languages; but for formal languages, it seems to be essentially Ho Hum. Still, there is some more to go. We cannot proceed further without knowing when one of our Q-atoms (atomic terms or elementary functors or ∀—see **3B-2**) "occurs" in a formula; and for variables we need to know when they "occur free." We will approach these ideas by defining for each formula O a pair of sets, occur(O)—the set of Q-atoms that occur in O, and free-occur(O)—the set of Q-atoms that occur free in O. (The purpose of defining these concepts via sets is that we sometimes wish to prove something about how large they are. The purpose of the ugly notation is to keep it short. Other purposes will, as usual, be served below by a variant.) To avoid repetition, we will use the "parenthetical" style to define "occur(O)" and "free-occur(O)" simultaneously. All the clauses are intended to hold for both, except those separately given for ∀. (Of course it will turn out that "free" is redundant except for variables.)

3B-15 DEFINITION. *(Occurs and occurs free)*

---

occur, free-occur ∈ (formulas ↦ $\mathcal{P}$(Q-atom)) ((free-)occurtype)

They are defined inductively as follows, with the first line as the **Basis clause** and the next three lines as **Inductive clauses**.

(free-)occur(a) = {a} ((free-)occuratom)

(free-)occur($\Phi O_1 \ldots O_n$) =
{Φ} ∪ (free-)occur(O_1) ∪ … ∪ (free-)occur(O_n) ((free-)occurΦ)

occur(∀vA) = {∀} ∪ {v} ∪ occur(A) (occur∀)

free-occur(∀vA) = ({∀} ∪ free-occur(A)) − {v} (free-occur∀)

Also, x ∈ occur(G) ↔ for some A, A ∈ G and x ∈ occur(A); and similarly for free-occur(G).

Let us hasten to supply a more well-favored mode of speech:

3B-16 VARIANT. *(Occurs in and occurs free in)*

---

x *occurs in* O ↔ x ∈ occur(O); and x *occurs free in* O ↔ x ∈ free-occur(O).

Ditto for "x *occurs (or occurs free) in* G."

Observe that these notions were defined without reifying the concept of "occurrences," which is good, because although the pictures of "many occurrences of the same variable" are easy even if not beautiful, the rigorous theory[9] is not easy even if beautiful.

[9] Presented at its best in Quine's *Mathematical logic*.

It is convenient to have notation for the set of constants *not* occurring in a formula or set of formulas. (It is an accident of our later needs that we keep track of nonoccurring *constants* only.)

3B-17 DEFINITION. *(Nonoccur)*

---

nonoccur(O), nonoccur(G) ⊆ constant (nonoccurtype)

nonoccur(O) = constant − occur(O), **3B-16**.

nonoccur(G) = constant − occur(G), **3B-16**.

§3A noted that we want no finite bound on the number of variables, or constants, because we cannot predict in advance the degree of complexity we shall need in order to say what must be said, or prove what must be proved. The following two facts reflect this matter in a form most useful for our later work.

3B-18 FACT. *(Finiteness of occur(O))*

For each formula O, occur(O) is finite.

PROOF. Omitted. We can use Induction on formulas, **3B-8**, and the definition of "occur," **3B-15**; and of course the properties of "finite," §2B.3; but this version of these notes has not developed the apparatus needed to be rigorous about the finitude of the members $t_1 \ldots t_n$ of $term^n$. ☐

3B-19 FACT. *(Existence of nonoccurring variable, and constant)*

For each formula, there is a variable, and a constant, that does not occur in it.

PROOF. Axiom **3B-1** says that the set of variables, and the set of constants, is infinite, so that by **2B-15** and the fact that ∅ is finite (Fin∅, **2B-7**), there must be a variable, and a constant, not in occur(O) by **3B-18**. ☐

The following minifacts are used later (in the proof of an important theorem in §3D). They are of no independent interest. You should, however, pay attention as you go through them. The manner of their statement places essential reliance on Convention **3B-3**, which should be consulted.

3B-20 MINIFACT. *(Substitution, occurrence, and freedom)*

1. **Vacuous substitution.** If a does not occur free in O, [c/a](O) = O.
2. **Inverse substitution.** If v does not occur (at all) in O, then [c/v](([v/c](O))) = O.
3. **Composition of substitution.** If v′ does not occur (at all) in O then [c/v](O) = [c/v′](([v′/v](O)))]
4. **Disappearing substituendum.** c does not occur in [v/c](O).
5. **Appearing substituens.** If v occurs free in O then c occurs in [c/v](O).
6. **Persisting side constants.** If c occurs in [v′/v](O) then c occurs in O.
7. **Unfreedom and ∀.** v does not occur free in ∀v′([v′/v](A)).
8. **Nothing from nothing.** If b occurs (free) in [t/a](O), then either b occurs (free) in O or b occurs (free) in t.

Exercise 49 *(On some minifacts)*

Prove one or two of the minifacts **3B-20**.

Also, for one or two of the listed minifacts that you choose not to prove, provide an example indicating your understanding of the matter. Thus, for (2), choose an O, v, and c rendering the hypothesis of the minifact true, compute the intermediate values, and generally convince yourself that the *definitions* (and not merely intuition) lead to this result. Further, and equally important, provide an example (choice of O, v, c) in which the consequent of the same minifact is *false* because its antecedent is false.

▷ .... ◁

Evidently we can say what it is for a formula to be "closed" or "open," a matter of some interest:

3B-21 DEFINITION. *(Closed and open)*

A formula is *closed* or *open* according as no variable occurs free in it, or some variable occurs free in it.

A *set* of formulas is *closed* if all of its members are closed, and is otherwise *open*.

One reason this concept is important to us is that in the *proof theory* to be presented below, substitution will be restricted to closed terms. Another is that in *semantics*, what values variables receive will be irrelevant for the values of closed formulas. A third is that in *applications*, closed sentences and terms represent ordinary English sentences and terms (with determinate truth values or denotations).

The following is an easy corollary of **3B-20**(8):

3B-22 COROLLARY. *(Closed and substitution)*

If O and t are closed, then so is [t/a](O).

Also, $\Phi O_1 \ldots O_n$ is closed ↔ each O_i is closed.

The following isn't needed until much later (for the proof of **3E-11**). It gives us a method for proving that some property holds of all closed terms and formulas without detouring through any open ones.

3B-23 LEMMA. *(Induction on closed terms, and sentences)*

Induction on closed terms. Let Ψ(t) be a use-context. **Suppose** the following.

> **Basis step.** For all constants c, Ψ(c).
>
> **Inductive step.** For all closed terms $t_1, \ldots, t_n$, if $\Psi(t_1), \ldots, \Psi(t_n)$, then $\Psi(ft_1 \ldots t_n)$.

Then for all closed terms t, Ψ(t).

Induction on closed sentences.

Let Ψ(A) be a use-context. **Suppose** the following.

> **Basis step.** For all closed terms $t_1, \ldots, t_n$, $\Psi(Ft_1 \ldots t_n)$.
>
> **Inductive step.** For all closed A and B, if Ψ(A) and Ψ(B) then Ψ(∼A) and Ψ(A ⊃ B).
>
> **Inductive step.** If ∀vA is closed, and if for all closed terms t, Ψ([t/v](A)), then Ψ(∀vA).

Then for all closed sentences A, Ψ(A).

PROOF. It is straightforward to establish Induction on closed terms by using Induction on terms, **3B-4**, choosing "if t is closed then Ψ(t)" here as the use-context for that induction. You will of course need to appeal to the definitions of "closed," **3B-21**, and of "occurs free," **3B-15**.

In contrast, the proof of Induction on closed sentences cannot be carried out in that straightforward way, as you can convince yourself by analyzing the case for the quantifier. We omit development of the means to carry it out, and ask you to observe that this is yet another case in which surface appearances of similarity (of degree of "straightforwardness") disappear upon a more rigorous inspection. □

Exercise 50 *(Induction on closed terms, sentences, and formulas)*

1. Formulate Induction on closed terms, and Induction on closed sentences, as subproof rules.
2. Formulate a statement that could properly be called "Induction on closed formulas"—a kind of combination of Induction on (all) formulas, **3B-7** and Induction on closed terms, and sentences, **3B-23**.

▷ . . . . ◁

(The following is not part of the main line of development.)

Although in fact we will be dealing with the substitution of only *closed* terms in our serious work below, it seems best to define the wider conditions under which substitution "makes sense."[10] Intuitively we are trying to lay out the conditions under which no variable free in the substituted term, t, is "caught" by a quantifier of the formula, O, into which the substitution (for some atomic term b) is made; under these circumstances one says that "t is free for b in O."

3B-24 DEFINITION. *(Free for)*

[10]The careful reader will observe occasions on which we ourselves use (in our use language) the wider form of substitution, that is, the one permitting substitution of open terms. Even though it is more complicated to theorize about (and so we don't), it certainly is convenient to use (and so we do).

Basis clause. t is free for b in a (always). (freeforatom)

Inductive clause. t is free for b in $\Phi O_1 \ldots O_n \leftrightarrow$ for all k ($1 \leqslant k \leqslant n$), t is free for b in O_k. (freeforΦ)

Inductive clause. t is free for b in $\forall$vA $\leftrightarrow$ [(v does not occur free in t or b does not occur free in $\forall$vA) and t is free for b in A]. (freefor$\forall$)

This inductive definition needs no closure clause, because all of the clauses are biconditionals, so that the necessary conditions that are usually provided by a separate closure clause are here provided piece by piece as we go up the line. The definition is in any event almost impenetrably dense. Here is a non-inductive version that somewhat more closely follows the preceding English. A term t is free for b in O $\leftrightarrow$ for no v and B does the following happen: $\forall$vB is a subformula of O and b occurs free in $\forall$vB and v occurs free in t. Complete rigor, however, does not quite entitle us here now to this second version, because we have not defined "subformula" for Q, so that it would contain some unexplained terminology.

One uses this definition in speaking of principles governing the use of the universal quantifier. For example, the inference from $\forall$vA to [t/v](A) is acceptable *provided* t is free for v in A (but not in general otherwise). The complexity of the final clause of the definition of "t is free for b in O" is the principal reason that it is easier to confine substitution to closed terms, accepting the inference from $\forall$vA to [t/v](A) only when t is closed (noting that a closed term, t, is automatically free for any v in any A).

Exercise 51 *(Grammar of* Q*)*

1. Optional (but easy). Prove Induction on closed terms, **3B-23**.

2. Optional (and difficult). Prove Induction on closed sentences, **3B-23**.

3. Optional (tedious at best). First provide an inductive definition of "subformula" for Q, corresponding closely to **2B-23**. Then prove that the two definitions of "t is free for b in O" are equivalent.

▷ . . . . ◁

3C Elementary semantics of Q

This section articulates the basic semantic concepts and facts concerning the language Q, in analogy with our consideration of the semantics of TF in §2D. The matter here is more complicated, however, as one should expect from the fact that we are dealing with a language that, with the help of its quantifiers, can in some not altogether metaphorical sense say infinitely many things at once: As you already know very well, understanding truth tables hardly suffices for understanding "it is false that there are three integers, and there is an integral power greater than two, such that the sum of raising each of the first two numbers to that power is the same as raising the third to that power"; that is, Fermat's Last Theorem, which denies that anywhere in the vast infinity of integers there are x, y, z, and n (n larger than 2) such that

$$(x^n + y^n) = z^n.$$

3C.1 Semantic choices

In §3A we discussed various *grammatical* choices to be made in designing a first order language; here we want to enter a few remarks on *semantic* choices, keeping the grammar fixed as already given in §3B.

Domain-and-values or substitutional? Is ∀vA to be taken as true (1) when A is true for every value of v in the domain, or (2) when every substitution instance [t/v](A) is true? The former is the "domain-and-values" interpretation of the quantifiers; the latter is the "substitutional" interpretation. Both are legitimate (clear, understandable, interesting, useful), but they are distinct except in those cases in which (a) everything in the domain has a name and (b) every term denotes something in the domain.

Without even considering the at least quasi-factual question of whether quantificational phrases in English are domain-and-values or substitutional, we choose to treat here only the domain-and-values interpretation of the quantifier. The reason is simply that at the very least it is much more useful in applications, because we so often want to treat cases where not everything has a name and secure the effect of "referring" to these nameless entities indirectly by means of our quantifiers; e.g., "between each pair of points on a line there is another (whether named or not!)." So even though the substitutional interpretation is in some respects technically simpler, we choose not to treat it.

Domain of quantification fixed or variable? Should we treat the variables of Q as having a fixed range (that is, as ranging over what the islanders think of as "everything," specified once and for all), or should we think instead of their range as varying with the application we have in mind? The cash value of this choice comes in our definition of validity: Do we refer "validity" to some one pre-chosen domain, or do we make the choice of domain part of what varies, so that "validity" requires reference to all such choices? Either choice is defensible, although the former seems philosophically shaky to the extent that it too closely binds the question of good argument to the question of what there is. It would also reduce the convenience of Q as applied, because we often want to use quantifiers with a restricted range (e.g. the nonnegative integers N), and also, such a restriction is often necessary in interpreting homely English expressions (e.g. "something has gone awry"). Nor is it a good choice technically, as with great frequency we want to cite special domains, often very small, in order to make some theoretical point. Upshot: We shall think of the domain of quantification of Q as varying just as much as our interpretation of its "nonlogical" constants. We will represent this technically by attaching the domain to the universal quantifier; i.e., if $\mathbf{j}$ is an interpretation, $\mathbf{j}(\forall) =$ D is the domain. The point of this representation is entirely convenience narrowly conceived, and is not the only sensible one,[11] but it correctly suggests that domains go with quantifiers.

There are a couple of closely related choices to be made. It turns out to be essential in our framework that each domain $\mathbf{j}(\forall)$ be a *set* so that we can perform certain set-theoretical manipulations upon it; this is a way of putting Tarski's point: If you wish to give a rigorous and consistent semantic theory of Q, then the variables of your use-language are going to have to have an "essentially richer" range than the variables of Q. (That will be true in this particular case because no one set encompasses everything, especially not all sets.) The detailed need for the requirement that $\mathbf{j}(\forall)$ be a set, however, will not be visible in the development to follow, for it would only appear in the justification of Definition **3C-7** below—which as usual we do not give.

The second matter is this: Given as above that each domain must be a set, should the family of all domains that serve as possible ranges of the variables of Q be itself a set? Equivalently, should all the various domains be subsets of some one super-domain? Here we have a choice, and it is conceptually better not to assume that all the domains have a common superset, for that is at least something like

[11] In higher order logics, domains are generally attached to abstract entities called "types," each of which has a different style of terms as well as a different domain associated with it. (It should be noted that this representation is a change from previous versions of NSL, so that there may well be change-created errors.)

assuming that there is a single, fixed domain. It follows that although we shall be able to make sense out of the predicate, "**j** is an interpretation for Q," there will be no set of all Q-interpretations (any more than there is a set of all sets). Technically, however, it turns out (after the fact, as it were) that it would not make any difference in so far as the concept of validity goes: The same arguments turn out valid whether or not we assume that there is a super-domain—as long as we assume the super-domain is large enough.[12]

Permit empty domain? Should we, as a matter of logic, insist that the domain of quantification be nonempty? There cannot be much philosophical reason for doing so, for whether in a particular application there is something rather than nothing can seldom be a matter of logic. To involve a semantics permitting possibly empty domains of quantification is part of what is meant by "free logic," i.e., logic free of existential commitments. Study of free logics is philosophically useful, and it may even happen on occasion that such a logic has application.

We nevertheless make the classical choice of considering only nonempty domains, for several reasons: First, it is technically simpler; second, in the vast majority of applications we can be as certain as we are of anything that the domain of quantification is nonempty; third, the choice simplifies—a little—the proof theory.

Nondenoting terms? We also make the classical choice of requiring every term, as a matter of logic, to denote something in the domain of quantification. Considering terms that do not denote something in the domain (or do not denote at all) is another part of what is meant by "free logic"; and again the matter is technically more complicated—all sorts of semantic decisions would have to be made concerning the treatment of such nondenoting terms; for example, how to treat sentences containing them. Thus we shall not be able to apply our logic to cases involving terms like "Peter Rabbit" or operators like "the mass of _" when grammar permits putting terms like "fourteen" in the blank, with the understanding that the resultant "the mass of fourteen" denotes nothing at all.

The technical expression of this decision is this: Where D is the domain, each atomic term will have a value in D, and each n-operator will be interpreted as a function in $(D^n \mapsto D)$—that is, as a function that delivers a value in D given n arguments in D. (See **2G-4** for the notation D^n.) It follows as the day the night that all terms will have values in D.

[12]Suppose someone believes that there are exactly six things all together in the world. Such a "sixist" might want an alternative account of validity. Puzzle about this if you wish, but it is not going to help you understand logic as a tool.

Nonextensional operators and predicates? The foregoing decision conceals another: We shall not be able to treat "nonextensional" operators such as, perhaps, "John's reason for undertaking _"; because for this operator (some think) one can have identical inputs without identical outputs: The act of John's running yesterday was identical to John's act of killing himself by overexertion (so some say), but his reason for undertaking the one was not the same as his reason for the other (so most say).

For exactly the same reasons, we shall not be able to treat nonextensional predicates like (some readings of) "John believes that _ is good on potatoes" (salt, yes; NaCl, no).[13] Thus we fully fix the interpretation of an n-predicate by any recipe that tells us, for each n-ary sequence of members of D, whether or not the predicate truly applies.

In the case both of operators and predicates, the decision is partly based on considerations of technical simplicity, for there is no doubt we narrow the range of applications of our theory by leaving out nonextensional functors. But another reason is to avoid philosophical argle-bargle at this point, for a (dwindling) number of philosophers easily grow apoplectic when faced with discussions that treat nonextensional operators or predicates as sufficiently respectable to warrant theorizing about.

Semantic styles. The foregoing represent substantive choices; here we mention some further choices of a stylistic nature.

Interpret the logical Q-atoms? It is not customary to "interpret" the logical Q-atoms, $\supset$, $\sim$, or $\forall$, but instead to convey their "meanings" through semantic clauses such as our definition of "$Val_{\mathbf{i}}$," **2D-7**. (You will recall that a "TF-interpretation," **2D-4**, interprets only the TF-atoms. The connectives $\supset$ and $\sim$ were certainly given "meaning," but not through what we called a "TF-interpretation.") We choose, however, to "interpret" $\supset$ and $\sim$ along with all the nonlogical atoms, while still treating $\forall$ via a semantic clause.

Where $\mathbf{j}$ is an interpretation, $\mathbf{j}(\supset) = \supset^*$ and $\mathbf{j}(\sim) = \sim^*$, **2A-8**, **2A-10**. These interpretations will be "constant," as of course is appropriate for logical constants. The reason we "interpret" these logical constants is to emphasize and utilize the similarity between the ways in which (1) the values of complex sentences grow out of the values of their ingredient sentences, (2) the values of complex terms grow out of the value of their ingredient terms, and (3) the values of predications grow out

[13] John is wrong; we recommend a diet free of table-salt.

of the values of their ingredient terms. It is, however, not convenient to treat ∀ in this way.

Denotation and truth separate? Frequently semantics is made a matter of "reference and truth," separately, the terms having reference and the sentences having truth. We choose to lump these together in the concept of "semantic value," the terms having entities in the domain as their values, the sentences having truth values. The chief reason is technical: to avoid a great deal of repetition. But also this style of presentation emphasizes the profound idea common to the semantics of terms and sentences: The values (of whatever kind) of the wholes are functions of the values (of whatever kind) of the parts.

Satisfaction or values to variables? Tarski, the first to get these matters right, "defined truth in terms of satisfaction." What "satisfaction" in effect did was to assign values to the variables; but in an indirect way. Tarski defined as his semantic workhorse, "the (infinite) sequence s satisfies A," the idea being that the n-th member of s should be thought of as assigned as value to the n-th variable. We don't know about Tarski himself, but certainly some later followers of his have spoken as if the satisfaction relation somehow gave a direct relation between the sentence A and "the world" as represented by the sequence s—a sequence being a supposedly nonlinguistic sort of animal. But that is wrong: The sequence really *is* linguistic, since the *only* meaning its order has is relative to the order of the variables—a definitely linguistic matter of language, not of the world.

In order to avoid the false comfort of bad metaphysics, it is therefore better to avoid basing semantics on a relation between A and a sequence; instead, the relation should be between A and a direct assignment of values to the variables. Technically, it is the difference between relating A to a function in (N↦D) (Tarski's sequence) and relating A to a function in (variables↦D) (direct assignments of values to variables). Because the former (1) anyhow requires a separate ordering of the variables—a one-one function in (N↦variables)—and (2) conceals the fundamentally linguistic aspect of the matter by an only apparently language-free device, the latter is to be preferred on both technical and philosophical grounds.

Separate interpretation of variables? We already agreed in the course of designing the grammar of Q to lump together the constants and the parameters, calling them all "constants," whence it follows as the night the day that they will be treated alike semantically. It is however still open to us to give separate semantic treatment to (nonlogical) constants on the one hand and variables on the other; and there is substantial conceptual point in so doing, most particularly this: Values of variables are an *auxiliary* concept, since the values of closed formulas—the only ones in which we are really interested in any application—evidently do not depend on giving values to variables. In other words, every closed formula has a definite semantic value once a domain is chosen and an interpretation is given to each non-logical constant, even if no interpretation is given to the variables at all; the role of the interpretation of the variables is only secondary, coming into play in the course of defining the semantic values of closed formulas in terms of the values of open ones, but not itself influencing the values of the closed formulas. We have in mind that the value of a closed sentence like ∀vA does not depend on any value given to v, but is defined in terms of the semantic values of A, which *does* depend on values of v. (It is quite possible to avoid looking at A with its free v, instead defining a value for ∀vA in terms of some well-chosen substitution instance [c/v](A), c a constant. But this alternative, too, involves its own complications, both technical and philosophical.)

There are also certain technical considerations arising from the set-theoretical form of our central semantic definition, **3C-7**, which make separating off assignments to variables a good thing; for it is only these which *must* vary in the course of the definition; all other parts of the interpretation *can* be held fixed.

For what *we* have to do, however, it is technically somewhat simpler to lump together the assignments to variables with the interpretation of the constants; and we so choose. We can get back most of our cake (having eaten it) through an application of an appropriate "Local determination lemma" such as **3C-12** on p. 165, to the effect that if a formula is in fact closed, then its value does not depend on the value of any variable. And that seems good enough.

Interpretation of predicates. It is a matter of *substance* that the interpretation of an n-predicate is, as said above, fixed by a recipe telling us those n-ary sequences to which it truly applies, and we have already decided this matter. It is, however, a matter of *style* how we do this. The most usual way is to interpret a predicate as a subset of the set D^n of all n-ary sequences of members of D, taking the items *in* the chosen subset to be those to which the n-predicate truly applies. A substantively equivalent choice is to interpret an n-predicate as a function in $(D^n \mapsto \mathbf{2})$, with the n-ary sequences that are mapped into T being those to which the n-predicate truly applies.

We choose this latter recipe for two sorts of reasons. On the philosophical side, it exhibits predicates in the role of "propositional functions," except that since we are being so thoroughly extensional, individuals are being mapped into truth values instead of into full blooded propositions. Further, it emphasizes the semantic similarity between operators and predicates and connectives (where the notation X^n is defined in **2G-4**):

($D^n \mapsto D$) for operators

($D^n \mapsto \mathbf{2}$) for predicates

($\mathbf{2}^n \mapsto \mathbf{2}$) for connectives)

And it highlights the Fregean idea that the meaning of a predicate needs "completion"—by its arguments. And it suggests generalizations to 3-valued logic, etc. Furthermore, on the technical side we can exploit the formal similarity among the sorts of meanings we give operators, predicates, and connectives, as displayed above, in order to give brief and uniform treatments of some matters that would otherwise be case-ridden. All of this outweighs, in our judgment and for present purposes, the consideration that members of ($D^n \mapsto \mathbf{2}$) are from a certain set-theoretical point of view more complex than members of $\mathcal{P}(D^n)$.

Atomic terms as 0-ary operators? Sometimes atomic terms are treated as operators, 0-ary ones (no arguments). This has technical point in smoothing the treatment of certain matters, but because our aims are limited, it would not be likely to improve either conceptual clarity or technical facility for us. Accordingly, we do not introduce the idea of a 0-ary operator at all.

Summary.

- Quantifiers receive a domain-and-values interpretation, not substitutional.
- Domains vary (they must be specified as part of the interpretation), and must be nonempty. We do not suppose that there is a set of all possible domains.
- All terms must denote something in the domain.
- Operators and predicates are all treated extensionally.

The following are matters of semantic style.

- The meanings of $\supset$ and $\sim$ are fixed by the same function that gives meaning to the other atoms.
- Instead of separating the treatment of terms (reference, or denotation) and sentences (truth values), all formulas will be treated together (values).
- The valuing of variables and constants will not be separated.
- Predicates are interpreted by functions from n-ary sequences of individuals into **2**.

3C.2 Basic semantic definitions for Q

An interpretation normally gives us enough information to fix the value of every *closed* formula, but by choices recorded in the previous section, we want a concept of an interpretation that fixes the value of *every* formula, open or closed. Let **j** be such an interpretation. What information must **j** supply?

1. A domain. It is important to be aware that the truth value of a sentence can be changed by changing *only* the domain; e. g., the truth of "$3 \leqslant 2 \leftrightarrow \sim\forall z \sim(3 + z = 2)$" depends on whether or not the domain for the universal quantifier is restricted to *non*negative numbers. We obtain this information artificially by applying **j** to ∀: **j**(∀) will be the domain of quantification.

2. An interpretation (relative to the domain) for each nonlogical atom—in the case of Q, for each atomic term (variable or constant), each operator, and each predicate. **j** will deliver this information for us.

3. The above information wholly suffices. In order to exploit the concept of "elementary functor," however, for both the conceptual and technical reasons given above, we also define **j** for the logical elementary functors ⊃ and ∼. But because they are logical, we won't give **j** any choice in how they are to be interpreted: $\mathbf{j}(\supset) = \supset^*$ and $\mathbf{j}(\sim) = \sim^*$, for all interpretations **j**.

Each atom must of course receive the right type of interpretation. It is illuminating to say what this type is by using a function space symbol "$\rightarrowtail$" instead of "$\mapsto$": the former bears a sense such that $f \in (X \rightarrowtail Y)$ does *not* imply that f is defined *only* on X, so that the use of "$\rightarrowtail$" carries less information than use of "$\mapsto$," NAL:**9B-14**. We must be technically circumspect, however, because if we tried to define $X \rightarrowtail Y$ generally, it might turn out not be a set; for this reason we relativize our use of this generalized function space operator:

3C-1 DEFINITION. *(Generalized function space operator: $\rightarrowtail_Z$)*

$f \in (X \rightarrowtail_Z Y) \leftrightarrow$ f is a function, $X \subseteq Dom(f)$, $Dom(f) \subseteq Z$, and for all x, $x \in X \rightarrow fx \in Y$.

Hence one may infer from $f \in (X \rightarrowtail_Z Y)$ and $a \in X$ to $fa \in Y$ (we call this "MP for $\rightarrowtail$" or "modus ponens for function space").

We may now define what it is for **j** to be a Q-interpretation as follows:

3C-2 DEFINITION. *(Q-interpretation)*

$\mathbf{j}$ is a Q-interpretation ↔ $\mathbf{j}$ is a function, and

> *Dom*($\mathbf{j}$) = ({∀, ⊃, ∼} ∪ atomic term ∪ operator ∪ predicate).
>
> *Local definition*: D = $\mathbf{j}$(∀).
>
> D is a nonempty set.
>
> $\mathbf{j}(\supset) = \supset^*$.
>
> $\mathbf{j}(\sim) = \sim^*$.
>
> *Local definition*: $(X \rightarrowtail Y) = (X \rightarrowtail_{Dom(\mathbf{j})} Y)$.
>
> $\mathbf{j} \in$ (atomic term $\rightarrowtail$ D).
>
> $\mathbf{j} \in$ (n-operator $\rightarrowtail$ $(D^n \mapsto D)$), each $n \in N^+$.
>
> $\mathbf{j} \in$ (n-predicate $\rightarrowtail$ $(D^n \mapsto \mathbf{2})$), each $n \in N^+$.

Observe the significance of our use of "$\rightarrowtail$" and "$\mapsto$," defined in NAL:**9B-14**; and recall "D^n" from **2G-4**.

3C-3 CONVENTION. *("$\mathbf{j}$" for Q-interpretations)*

"$\mathbf{j}$" is reserved for Q-interpretations. Furthermore, in any context using "$\mathbf{j}$," we will let D = $\mathbf{j}$(∀)—the domain—by local definition. (We will still on occasion use "D" for a sentence.[14]) Don't confuse the domain, *Dom*($\mathbf{j}$), of the function, $\mathbf{j}$, with the domain-of-quantification supplied by $\mathbf{j}$(∀). We reduce confusion by always using "D" to refer to "$\mathbf{j}$(∀)."

Definition **2D-21** introduced in the context of the logic of truth functional connectives a way of speaking of shifts in TF-interpretations, and we shall want something analogous for shifts in Q-interpretations. But the analog here is vastly more important than its mate there, for, as you will observe, we must introduce the concept even *before* we can explain the meaning of complex sentences, since we shall be invoking the shift of Q-interpretations in that very explanation. As it turns out, we shall want to consider shifts in the values given to atomic terms, but not to other Q-atoms; hence the following

[14]Grover once noted that printers did not have logicians in mind when they offered their none too numerous fonts.

3C-4 DEFINITION. *(Interpretation shift for* Q*)*

Let **j** be a Q-interpretation and let d ∈ D. Let a be an atomic term. Then [d/a](**j**) is defined as that Q-interpretation that is exactly like **j** except for the argument a; and it gives a the value d. That is,

[d/a](**j**) is a Q-interpretation ([]**j**-type)

([d/a](**j**))a = d ([]**j**1)

([d/a](**j**))x = **j**(x), if x ∈ *Dom*(j) but x ≠ a ([]**j**2)

It is hard to read "[d/a](**j**)" in English; perhaps "**j** but with d at a" will do, as short for "the Q-interpretation that is just like **j** except that it gives a the value d." And though we use a general variable "x" in []**j**2 , in applications we'll nearly always be considering an atomic term b; for nothing else ever shifts by means of this notation:

3C-5 MINIFACT. *(* [d/a](**j**) *for nonterms)*

([d/a](**j**))x = **j**(x) if x is not an atomic term but is in *Dom*(j). In particular, neither the domain of quantification nor the interpretations of the operators or predicates shift.

PROOF. Trivial, using []**j**2 of **3C-4** and Distinctness, **3B-10**. □

Exercise 52 *(Calculate* [d/a](**j**)*)*

Let D = {1, …, 8}. Let the atomic terms be x_1, …, x_4, all distinct. Suppose **j**(x_n) = 2n. First draw a picture of **j** as it acts on the atomic terms. Then of [3/x_2](**j**). Then of [2/x_3](([3/x_2](**j**))). (See Picture **2D-20** on p. 57.)

▷ . . . . ◁

The following properties of interpretation shift are needed later.

3C-6 FACT. *(Properties of interpretation shift for* Q*)*

Let **j** be a Q-interpretation, let d, d′ ∈ D, and let a, b ∈ atomic term.

$$[d'/a]([d/a](\mathbf{j})) = [d'/a](\mathbf{j}) \qquad ([]\mathbf{j}3)$$

$$a \neq b \rightarrow [d'/b]([d/a](\mathbf{j})) = [d/a]([d'/b](\mathbf{j})) \qquad ([]\mathbf{j}4)$$

That is, by []**j**3 , if you twice shift the value of a *single* atomic term, only the last shift counts; and if by []**j**4 you shift the values of *distinct* atomic terms, then it doesn't matter in which order you make the shifts.

PROOF. Tedious. One shows functions identical by showing that they give the same value for each argument in their domain—and have the same domain (NAL:**9B-18**). []**j**-type of **3C-4** gives us the part about the same domains; in addition, there are two cases for []**j**3 (the argument is or is not a) and three cases for []**j**4 (the argument is a or b or something else). Use []**j**1 and []**j**2 of **3C-4** to calculate results in these various cases. ☐

Exercise 53 *(Shifting interpretations)*

Lay out the proofs of []**j**3 and []**j**4 (so as not to embarrass yourself when teaching this material).

▷ . . . . ◁

Next comes *the* central semantic definition, the definition that tells us how the values of complex formulas—terms and sentences—depend on the values of the nonlogical atoms—and the domain. Keep in mind that the values of all terms will be members of D, while values of sentences will be truth values.

3C-7 DEFINITION. *($Val_{\mathbf{j}}$)*

Let **j** be an interpretation for Q, and recall that D = **j**(∀), **3C-3**. Then $Val_{\mathbf{j}}$ is that function such that

$$Dom(Val_{\mathbf{j}}) = \text{formula} \qquad (Val_{\mathbf{j}}\text{type})$$

Local definition: $(X \rightarrowtail Y) = (X \rightarrowtail_{Dom(\mathbf{j})} Y)$.

$$Val_{\mathbf{j}} \in (\text{term} \rightarrowtail D) \qquad (Val_{\mathbf{j}}\text{type})$$

$$Val_{\mathbf{j}} \in (\text{sentence} \rightarrowtail \mathbf{2}) \qquad (Val_{\mathbf{j}}\text{type})$$

$$Val_{\mathbf{j}}(a) = \mathbf{j}(a) \qquad (Val_{\mathbf{j}}\text{atom})$$

$$Val_{\mathbf{j}}(\Phi O_1 \ldots O_n) = (\mathbf{j}(\Phi))(Val_{\mathbf{j}}(O_1), \ldots, Val_{\mathbf{j}}(O_n)) \qquad (Val_{\mathbf{j}}\Phi)$$

$$Val_{\mathbf{j}}(\forall vA) = T \leftrightarrow (d)(d \in D \rightarrow Val_{[d/v](\mathbf{j})}(A) = T) \qquad (Val_{\mathbf{j}}\forall)$$

The "atom" clause is the **Basis clause**, and the remainder are **Inductive clauses.** Note that the clause ($Val_{\mathbf{j}}\forall$) for the universal quantifier is not given as an identity having the form $Val_{\mathbf{j}}(\forall vA)=?$; such would be possible, but not worth the extra henscratches. We mean that the quantifier clause could be "improved" to an identity by using abstraction, p. 133, and by introducing a function $\&^*$ from *sets* of truth values to truth values; i. e., let $\&^* \in (\mathcal{P}(\mathbf{2}) \leftrightarrow \mathbf{2})$ be defined by the condition that $\&^*X=T \leftrightarrow F \notin X$. Then: $Val_{\mathbf{j}}(\forall vA)=\&^*\{Val_{[d/v](\mathbf{j})}(A)\text{: } d \in D\}$.

Definition **3C-7** packs in a lot, especially through the omnibus clause ($Val_{\mathbf{j}}\Phi$); we sort it out in the following.

3C-8 FACT. *($Val_{\mathbf{j}}$ facts)*

$Val_{\mathbf{j}}(ft_1 \ldots t_n)=(\mathbf{j}(f))(Val_{\mathbf{j}}(t_1), \ldots, Val_{\mathbf{j}}(t_n))$ ($Val_{\mathbf{j}}$func)

$Val_{\mathbf{j}}(Ft_1 \ldots t_n)=(\mathbf{j}(F))(Val_{\mathbf{j}}(t_1), \ldots, Val_{\mathbf{j}}(t_n))$ ($Val_{\mathbf{j}}$pred)

$Val_{\mathbf{j}}(Ft_1 \ldots t_n)=T \leftrightarrow (\mathbf{j}(F))(Val_{\mathbf{j}}(t_1), \ldots, Val_{\mathbf{j}}(t_n))=T$ ($Val_{\mathbf{j}}$pred)

$Val_{\mathbf{j}}(A \supset B)=Val_{\mathbf{j}}(A) \supset^* Val_{\mathbf{j}}(B)$ ($Val_{\mathbf{j}}\supset$)

$Val_{\mathbf{j}}(A \supset B)=T \leftrightarrow (Val_{\mathbf{j}}(A)=T \rightarrow Val_{\mathbf{j}}(B)=T)$ ($Val_{\mathbf{j}}\supset$)

$Val_{\mathbf{j}}(\sim A)=\sim^* Val_{\mathbf{j}}(A)$ ($Val_{\mathbf{j}}\sim$)

$Val_{\mathbf{j}}(\sim A)=T \leftrightarrow Val_{\mathbf{j}}(A) \neq T \leftrightarrow Val_{\mathbf{j}}(A)=F$ ($Val_{\mathbf{j}}\sim$)

$Val_{\mathbf{j}}(\forall vA) \neq T \leftrightarrow Val_{\mathbf{j}}(\forall vA)=F \leftrightarrow \exists d(d \in D \text{ and } Val_{[d/v](\mathbf{j})}(A)=F)$ ($Val_{\mathbf{j}}\forall F$)

You will want to consult ($Val_{\mathbf{j}}\Phi$) of **3C-7** when considering the relations between $Val_{\mathbf{j}}$ and other concepts also using "Φ" in their definitions—that is, in considering the relation of $Val_{\mathbf{j}}$ to concepts in which all elementary functors are treated alike. But you will want to consult instead the various separate clauses of **3C-8** in those cases in which $Val_{\mathbf{j}}$ is being related to special concepts about terms, or about predications, or about truth functionally complex sentences. One test of whether or not Definition **3C-7** is satisfactory is to put it back into something like English.

- $Val_{\mathbf{j}}$func. The value of a complex term is obtained by applying the function associated with its operator to the values of the arguments.
- $Val_{\mathbf{j}}$pred. A predication is true just in case the relation associated with its predicate is truly applicable to the values of its arguments.
- Conditionals and negations are as before.

- $Val_{\mathbf{j}}\forall$ and $Val_{\mathbf{j}}\forall$F. A universal quantification is true on a given interpretation if its body is true on every interpretation that can be obtained from the given interpretation by varying the value of its variable; and it is false if its body is false on some interpretation so obtainable.

Another test of whether or not Definition **3C-7** is satisfactory is whether or not it coheres with our intuitive judgments about the truth and falsity of various sentences. The following exercise is designed to help convince you that the test is passed.

Exercise 54 *($Val_{\mathbf{j}}$ facts)*

Let **j** be a Q-interpretation having the following properties.

1. The domain is $\{1, 2, 3, 4\}$: $\mathbf{j}(\forall) = \mathrm{D} = \{1, 2, 3, 4\}$.
2. The atom a has the value 3: $\mathbf{j}(\mathrm{a}) = 3$.
3. The one-place predicate G is true of just 1 and 2: (d)[d $\in$ D $\rightarrow$ (($\mathbf{j}$(G)d) = T $\leftrightarrow$ d = 1 or d = 2)]. That is, $\mathbf{j}$(G)1 = T, $\mathbf{j}$(G)2 = T, $\mathbf{j}$(G)3 = F, $\mathbf{j}$(G)4 = F.

We are going to list some statements that should be true; your job is (1) to decide informally (perhaps pictorially) that each is true, and then (2) to use **3C-7** and **3C-8** (and whatever else you need) to prove these statements. Your most difficult job will be to see how to apply []**j**1 -[]**j**4 (**3C-4**, **3C-6**) in the context of a variety of other notation. Do not proceed without assuring yourself that you can do this.

1. $Val_{\mathbf{j}}$(Ga $\supset$ ∼∀vGv) = T.
2. $Val_{\mathbf{j}}$(∀v(Gv $\supset$ ∀vGv)) = F.
3. $Val_{\mathbf{j}}$(∼Gb $\supset$ ∼∀vGv) = T.

▷ . . . . ◁

If you have worked through this exercise, you will also have convinced yourself that the definitions we give for "$Val_{\mathbf{j}}$" are *not* useful for computing the truth values of individual sentences. For that job excessive rigor is out of place; we are better off winging it. The value of the rigor of Definition **3C-7** lies instead in enabling statement and proof of *deep* properties of our semantic concepts. The analogy with truth tables is perfect: To find the value of a sentence, lay out a truth table, using geometrical intuitions to the fullest; but use rigorous definitions if you want to prove the Finiteness property for semantic consequence, **2G-3**. It all depends. A principal consequence of your success in carrying out this exercise is that now you will know what it *means* to deduce "A is true" from a theory of "truth conditions," a theme that from time to time rightly comes into philosophical fashion. (Sometimes, however, the theme is played in discordant cacophony with ill-composed epistemological ditties. "Logic is logic; that's all I say.")

3C.3 Fundamental properties of $Val_{\mathbf{j}}$

Before defining semantic consequence by quantifying over Q-interpretations, let us establish some fundamental properties of valuations $Val_{\mathbf{j}}$ taken one or two at a time. These will mostly be analogs of theorems in §2D.3, but first comes an easy fact relating $Val_{\mathbf{j}}$ to previous concepts.[15]

3C-9 FACT. *($Val_{\mathbf{j}}$ and* TF-*valuations)*

For each Q-interpretation **j**, $Val_{\mathbf{j}}$ restricted to sentences is a TF-valuation, **2D-3** on p. 47, hence a sentential valuation, **2D-1** on p. 46.

Before proceeding to serious work, we lay down two trivial facts which will help clarify some proofs. The chief reason the first has point is this: While *identity* (a use-language predicate requiring use-language terms) figures in most of the clauses in the definition of $Val_{\mathbf{j}}$, and in many of the theorems about $Val_{\mathbf{j}}$, the clause $Val_{\mathbf{j}}\forall$, **3C-7**, instead relies on *equivalence* (a use-language connective requiring use-language sentences); so that it is good to have a way of passing back and forth.

3C-10 MINIFACT. *(Identity and equivalence)*

$Val_{\mathbf{j}_1}(\text{A}) = Val_{\mathbf{j}_2}(\text{B}) \leftrightarrow (Val_{\mathbf{j}_1}(\text{A}) = \text{T} \leftrightarrow Val_{\mathbf{j}_2}(\text{B}) = \text{T})$.

PROOF. Straightforward. You will see that the only property of $Val_{\mathbf{j}}$ required is $Val_{\mathbf{j}}$type: $Val_{\mathbf{j}_1}(\text{A})$ and $Val_{\mathbf{j}_2}(\text{B})$ are in **2** as defined in **2A-1**. That is, this minifact is really just an instance of the following: For **x**, **y**

in **2**, $\mathbf{x} = \mathbf{y} \leftrightarrow (\mathbf{x} = \text{T} \leftrightarrow \mathbf{y} = \text{T})$. □

[15] The fact may be "easy," but its statement involves some words, "restricted to," that we have not defined. We mark and then pass over this lapse from rigor because in fact we do not use this fact for any purpose beyond incidental illumination; instead we use Definition **3C-15** below.

The next gives us one way to show that the values of two universal quantifications are identical. It relies on and is also an analog to the use-language principle taking us from "$(x)(\Psi x \leftrightarrow \Psi' x)$" to "$(x)\Psi x \leftrightarrow (x)\Psi' x$."

3C-11 MINIFACT. *(A distribution for $\forall$)*

Let $\mathbf{j}_1$ and $\mathbf{j}_2$ be Q-interpretations with the same domain; that is, let $\mathbf{j}_1(\forall)=\mathbf{j}_2(\forall)$. If $Val_{[d/v](\mathbf{j}_1)}(A)=Val_{[d/v](\mathbf{j}_2)}(B)$ for all $d\in D=\mathbf{j}_1(\forall)=\mathbf{j}_2(\forall)$, then $Val_{\mathbf{j}_1}(\forall vA)=Val_{\mathbf{j}_2}(\forall vB)$.

Exercise 55 *(A distribution for $\forall$)*

Prove **3C-11**. Use Minifact **3C-10** as well as $Val_{\mathbf{j}}\forall$, **3C-7**.

As part of the exercise, also convince yourself that the converse is in fact false.

▷ . . . . ◁

Now for an important

3C-12 THEOREM. *(Local determination theorem for* Q*)*

For any Q-interpretations $\mathbf{j}_1$ and $\mathbf{j}_2$, if they agree on the particular argument, $\forall$ (that is, if $\mathbf{j}_1(\forall)=\mathbf{j}_2(\forall)$, so that $\mathbf{j}_1$ and $\mathbf{j}_2$ have the same domains), and if they also agree on every Q-atom occurring free in O, then: $Val_{\mathbf{j}_1}(O)=Val_{\mathbf{j}_2}(O)$.

Note the word "free": Values of nonfree variables are irrelevant. This is the principal "new" import of this theorem over its predecessor, the local determination theorem for TF, **2D-18**. Otherwise the theorem expresses one of the fundamental notions of semantics in the Fregean tradition: Fixing the values of the atoms that actually occur in a formula fixes the values of the formulas.

PROOF. By Induction on formulas, **3B-8**, choosing $\Psi(O)$ there as the entire theorem here, including the quantifications using "$\mathbf{j}_1$" and "$\mathbf{j}_2$" The cases 1 and 2 of **3B-8**, which cater to atomic terms and elementary functors respectively, proceed much like their analogs in the proof of **2D-18**; there remains case 3 for $\forall$.

Suppose the theorem (complete with quantifications using "$\mathbf{j}_1$" and "$\mathbf{j}_2$"[16]) holds for A; we need to show that it holds for ∀vA. Choose $\mathbf{j}_1$ and $\mathbf{j}_2$ to satisfy the hypothesis of the theorem for ∀vA; hence, $\mathbf{j}_1(\forall)=\mathbf{j}_2(\forall)$, so that by **3C-11**, it clearly suffices to show that for all d∈D,

$$Val_{[d/v](\mathbf{j}_1)}(A)=Val_{[d/v](\mathbf{j}_2)}(A).$$

But this last statement would be guaranteed by the hypothesis of induction *provided* the two interpretations $[d/v](\mathbf{j}_1)$ and $[d/v](\mathbf{j}_2)$ agree on all Q-atoms free in A. We know by hypothesis that $\mathbf{j}_1$ and $\mathbf{j}_2$ agree on all Q-atoms free in ∀vA, hence, by free-occur∀, **3B-15**, also on all Q-atoms free in A, with the possible exception of v. But certainly $[d/v](\mathbf{j}_1)$ and $[d/v](\mathbf{j}_2)$ agree also on v, by []**j**, **3C-4**. □

Since **j** and [d/a](**j**) clearly agree on ∀, and on all Q-atoms except possibly a, the following is an obvious

3C-13 COROLLARY. *(Corollary to the local determination theorem for* Q*)*

If a is not free in O, then $Val_{\mathbf{j}}(O)=Val_{[d/a](\mathbf{j})}(O)$.

The following is far more important for Q than its analog **2D-21** was for TF, just because the whole idea of substitution is more important in the presence of bound variables. For example, one of the standard principles of quantifier logic is instantiation, from ∀vA to infer [t/v](A), t closed; it is hardly possible to settle the question of whether or not this rule is valid without an understanding of the semantic properties of the operation of substitution.

3C-14 THEOREM. *(Semantics of substitution theorem for Q)*

Let t be a *closed* term. For all Q-interpretations **j**,

$$Val_{\mathbf{j}}([t/a](O))=Val_{[Val_{\mathbf{j}}(t)/a](\mathbf{j})}(O).$$

[16] This is the crucial difference between the form of this proof and the form of our proof for the analogous theorem for TF, **2D-18**. There the quantifications could be fixed (flagged) *outside* the induction, whereas here we need the quantifications expressed *inside* the induction.

That is, you can calculate the value of a result [t/a](O) defined by **3B-13** in two stages: First calculate $Val_{\mathbf{j}}$(t); then give a this value, while leaving all other Q-atoms alone, and calculate the value of O. This two-step calculation is guaranteed to yield the correct value for [t/a](O), *provided* t is closed. (By **3C-4**, the complex notation "[$Val_{\mathbf{j}}$(t)/a]($\mathbf{j}$)" names the shifted Q-interpretation that is just like $\mathbf{j}$, except assigning $Val_{\mathbf{j}}$(t) to a. Compare the discussion surrounding the Semantics of substitution lemma for TF, **2D-21**.)

In the upcoming proof, watch for the hypothesis that t is closed.

PROOF. By Induction on formulas, **3B-8**, choosing Ψ(O) there as the part of the theorem here beginning "for all Q-interpretations $\mathbf{j}$." Cases 1 and 2 for atomic terms and elementary functors are wholly straightforward, using the definitions **3B-13** and **3C-7** and **3C-4** of the key concepts occurring in the theorem: [t/a](O), $Val_{\mathbf{j}}$, and [d/a]($\mathbf{j}$). We take up case 3 for ∀, where we need to consider $Val_{\mathbf{j}}$([t/a]((∀vA))), and treat first the more complicated subcase when a$\neq$v. Let d be an arbitrary member of the domain of quantification, $\mathbf{j}$(∀). The inductive hypothesis gives us the theorem for A and all $\mathbf{j}$, hence in particular for the instance [d/v]($\mathbf{j}$):

$$Val_{[d/v](\mathbf{j})}([t/a](A)) = Val_{[Val_{[d/v](\mathbf{j})}(t)/a]([d/v](\mathbf{j}))}(A)$$

Don't think about what this means; but do verify that it comes from the form displayed in Theorem **3C-14** by mindless substitution of [d/v]($\mathbf{j}$) for every occurrence of $\mathbf{j}$. There are now two operations we can perform on the right side. In the first place, since t is closed, v cannot occur free in t, so that by the Corollary to the local determination theorem, **3C-13**, we may replace "$Val_{[d/v](\mathbf{j})}$(t)" by the simpler "$Val_{\mathbf{j}}$(t)." This yields

$$Val_{[d/v](\mathbf{j})}([t/a](A)) = Val_{[Val_{\mathbf{j}}(t)/a]([d/v](\mathbf{j}))}(A)$$

In the second place, since a$\neq$v, we can use []$\mathbf{j}$4 , **3C-6**, to change the order of the shiftings. We will then have

$$Val_{[d/v](\mathbf{j})}([t/a](A)) = Val_{[d/v](([Val_{\mathbf{j}}(t)/a](\mathbf{j}))}(A).$$

Hence

$$Val_{\mathbf{j}}(\forall v([t/a](A)) = Val_{[Val_{\mathbf{j}}(t)/a](\mathbf{j})}(\forall vA)$$

by **3C-11**, together with **3C-5** to obtain its hypothesis. But $a \neq v$ by the hypothesis of the subcase, so that $\forall v([t/a](A)) = [t/a]((\forall vA))$ by []O∀2, **3B-13**; hence, as desired,

$$Val_{\mathbf{j}}([t/a](\forall vA)) = Val_{[Val_{\mathbf{j}}(t)/a](\mathbf{j})}(\forall vA).$$

The subcase with $a = v$ is simpler: The left side of the desired display just above reduces to $Val_{\mathbf{j}}(\forall vA)$ by []O∀1, **3B-13**; and the right side reduces to precisely the same thing by the Corollary to the local determination theorem, **3C-13**, given that by free-occur∀, **3B-15**, v does not occur free in ∀vA. □

Exercise 56 *(Semantics of substitution and "free for")*

1. Optional and tedious, but illuminating. Show that Theorem **3C-14** continues to hold when the hypothesis that t is closed is replaced by the weaker hypothesis that t is free for a in O in the sense of Definition **3B-24**.
2. Equally optional. That this is so is half of the verification that Definition **3B-24** is "correct"; the other half: Show that for every way in which t can fail to be free for a in O, there is a case in which the theorem fails.

▷ . . . . ◁

Just as the *grammars* of TF and Q are tied together by treating the very same sentences, so we can tie the *semantics* of TF and Q together by recovering from each Q-interpretation $\mathbf{j}$ the TF-interpretation—call it "$\mathbf{i}(\mathbf{j})$"—associated with it.[17]

3C-15 DEFINITION. *($\mathbf{i}(\mathbf{j})$)*

Let $\mathbf{j}$ be a Q-interpretation. Then $\mathbf{i}(\mathbf{j})$ is the TF-interpretation that agrees with $Val_{\mathbf{j}}$ on the TF-atoms:

$$\mathbf{i}(\mathbf{j}) \in (\text{TF-atom} \mapsto \mathbf{2}). \qquad (\mathbf{i}(\mathbf{j})\text{type})$$

$$\mathbf{i}(\mathbf{j})(A) = Val_{\mathbf{j}}(A) \text{ for every TF-atom A} \qquad (\mathbf{i}(\mathbf{j})\text{TF-atom}).$$

[17] The unhappy notation is blessedly with us for only a flickering moment.

You may find this definition as confusing as we do because $\mathbf{i}(\mathbf{j})$ is a TF-interpretation while $\mathbf{j}$ is a Q-interpretation; so take time to appreciate the following

3C-16 FACT. *($\mathbf{j}$ and $\mathbf{i}(\mathbf{j})$)*

Let $\mathbf{j}$ be a Q-interpretation, and let $\mathbf{i}(\mathbf{j})$ be defined by **3C-15**. Where "$Val_{\mathbf{i}(\mathbf{j})}$" is defined as denoting a TF-valuation in accordance with **2D-7**, and "$Val_{\mathbf{j}}$" is defined by **3C-7**,

$$Val_{\mathbf{i}(\mathbf{j})}(\mathrm{A}) = Val_{\mathbf{j}}(\mathrm{A}).$$

PROOF. Exercise **3C-16**. You can use Induction on sentences for TF, **2B-18**; choose Ψ(A) there as the equation here. The base and inductive cases are straightforward, provided you keep in mind that you will need **2D-7** for TF-valuations for the left side and Definition **3C-7** for Q-valuations for the right. Of course you will need to appeal to Definition **3C-15** of "$\mathbf{i}(\mathbf{j})$" for the base case. ☐

It is worth observing that although $\mathbf{i}(\mathbf{j})$ is always a TF-interpretation, it is not true that "$\mathbf{i}(\mathbf{j})$" ranges over *all* TF-interpretations as $\mathbf{j}$ varies over all Q-interpretations. For example, there is a TF-interpretation $\mathbf{i}'$ such that $\mathbf{i}'(\forall \mathrm{vFv}) = \mathrm{T}$ and $\mathbf{i}'(\mathrm{Fc}) = \mathrm{F}$. This TF-interpretation, however, cannot be an $\mathbf{i}(\mathbf{j})$ (there is no Q-interpretation $\mathbf{j}$ such that $\mathbf{i}(\mathbf{j})$ has these properties). Put it this way: Whatever is true for all values of "$\mathbf{i}'$" is true for all values of "$\mathbf{i}(\mathbf{j})$"—but not conversely.

3C-17 REMARK. *(Isomorphism)*

This is a good place for a discussion of isomorphism.

3C.4 Semantic consequence for Q

As for TF, a crucial semantic concept for Q is that of semantic consequence, and the definition is just the same as **2D-24**, except that it refers to Q-interpretations instead of TF-interpretations:

3C-18 DEFINITION. *($\vDash_{\mathrm{Q}}$)*

$\vDash_{\mathrm{Q}} \subseteq (\mathcal{P}(\text{sentence}) \times \text{sentence})$. ($\vDash_{\mathrm{Q}}$type)

$\mathrm{G} \vDash_{\mathrm{Q}} \mathrm{A} \leftrightarrow (\mathbf{j})(\mathbf{j}$ is a Q-interpretation $\rightarrow ((\mathrm{B})(\mathrm{B} \in \mathrm{G} \rightarrow Val_{\mathbf{j}}(\mathrm{B}) = \mathrm{T}) \rightarrow Val_{\mathbf{j}}(\mathrm{A}) = \mathrm{T}))$.

The words are just the same: $G \vDash_{Q} A$ just in case every interpretation that makes every member of G true also makes A true; but the sense of "interpretation" is much more refined.

What do we know about semantic consequence for Q? In the first place, $\mathbf{i}(\mathbf{j})$ of **3C-15** allows us to take over for Q any *one* statement $G \vDash_{TF} A$:

3C-19 FACT. *($\vDash_{TF}$ and $\vDash_{Q}$)*

$G \vDash_{TF} A \rightarrow G \vDash_{Q} A$.

The converse of course fails.

PROOF. For contraposition, assume the consequent is false, so that some Q-interpretation $\mathbf{j}$ is such that $Val_{\mathbf{j}}(B) = T$ for all $B \in G$, and $Val_{\mathbf{j}}(A) = F$. But then by **3C-16**, $Val_{\mathbf{i}(\mathbf{j})}(B) = T$ for all $B \in G$, and $Val_{\mathbf{i}(\mathbf{j})}(A) = F$ Let us note explicitly that by **3C-15**, $\mathbf{i}(\mathbf{j})$ is a TF-interpretation; hence, by the Definition **2D-24** of "$\vDash_{TF}$," the antecedent is false as well. □

Given the consistency of S_{TF}, **2G-1**, the following is an obvious

3C-20 COROLLARY. *($\vdash_{S_{TF}}$ and $\vDash_{Q}$)*

$G \vdash_{S_{TF}} A \rightarrow G \vDash_{Q} A$.

In particular, we shall have $\vDash_{Q} A$ whenever A is a tautology, or a theorem of S_{TF}.

We can also obtain certain other facts about $\vDash_{Q}$ in strict analogy to facts about $\vDash_{TF}$, but these we shall have not because of a relation between Q-interpretations and TF-interpretations, but instead because of the formal similarity between the definitions **3C-18** and **2D-24** of $\vDash_{Q}$ and $\vDash_{TF}$; namely, we can have the analog of any fact that does not depend on how $Val_{\mathbf{j}}$ is defined (so we shall have $\vDash_{Q}$id, $\vDash_{Q}$weak, and $\vDash_{Q}$cut in analogy to $\vDash_{TF}$id, $\vDash_{TF}$weak, and $\vDash_{TF}$cut, **2D-28**); and the analog to any fact that depends only on features of the definition of "$Val_{\mathbf{j}}$" that it shares with that of "$Val_{\mathbf{i}}$" (so we shall have $\vDash_{Q}$MP, as well as all the other facts reported in Fact **2D-31** and Corollary **2D-32**):

3C-21 FACT. *($\vDash_Q$ facts)*

All of the easy structural facts about $\vDash_{TF}$ are also true of $\vDash_Q$, and so are the facts relating $\vDash_{TF}$ to connectives. In particular, we have $\vDash_Q$id, $\vDash_Q$weak, $\vDash_Q$cut, $\vDash_Q$MP, and $\vDash_Q$ for the three tautologies; that is, we have Q-analogs of **2D-28**, **2D-31**, **2D-32**, and **2D-33**.

We also have analogs for $\vDash_Q$ of the rest of the facts reported in the four places just cited.

PROOF. One-turnstile statements, such as the Theorem **2D-33** yielding three tautologies, come by **3C-19**. The remainder, which relate turnstile statements, can be made to depend on earlier work as follows (but we are not quite rigorous so that this part of the proof counts as "omitted"). Characterize Val by: f ∈ Val ↔ *Dom*(f) = Sent, and for some Q-interpretation **j**, f(A) = $Val_{\mathbf{j}}$(A), all A. By **3C-9**, Val is a set of TF-valuations, hence of sentential valuations; so that clearly the relation $\vDash_Q$ satisfies the hypotheses of Identity, Weakening, and Cut in an abstract setting, **2D-29**, and Connectives and $\vDash_{TF}$ in an abstract setting, **2D-34**. The conclusion is immediate. □

We are now in a position to verify the following, which corresponds to Three tautologies, **2D-33**.

3C-22 THEOREM. *(Six quantificational truths, and some more)*

---

1. $\vDash_Q$ A ⊃ (B ⊃ A)
2. $\vDash_Q$ (A ⊃ (B ⊃ C)) ⊃ ((A ⊃ B) ⊃ (A ⊃ C))
3. $\vDash_Q$ (∼A ⊃ ∼B) ⊃ (B ⊃ A)
4. $\vDash_Q$ (∀vA ⊃ [t/v](A)), if t is closed (**3B-21**)
5. $\vDash_Q$ (A ⊃ ∀vA), if v does not occur free in A (**3B-15**)
6. $\vDash_Q$ ∀v(A ⊃ B) ⊃ (∀vA ⊃ ∀vB)
7. $\vDash_Q$ [c/v](A) → $\vDash_Q$ ∀vA, if c does not occur in A (**3B-15**)

According to Convention **3B-3**, t in 4 is any term, while c in 7 must be a constant.

It is easy to see that "G" could be systematically added on the left of the turnstiles in 1-6 by $\vDash_Q$weak, **3C-21**; if "G" were present (twice) in 7, we should have to add that c also fails to occur in any member of G.

PROOF. These are of fundamental importance. Dealing with them rigorously will quickly disabuse you of the common notion that the presence in our use-language of quantificational devices makes it a *trivial* matter to prove the legitimacy, in a mathematically defined sense, of quantificational principles in the language we are theorizing about.

1-3. Facts **2D-33** and **3C-19**.

4. Let t be closed. Choose a Q-interpretation **j**. Suppose that $Val_{\mathbf{j}}$(∀vA) = T, so that (d)(d ∈ D → $Val_{[d/v]\mathbf{j}}$(A) = T) by **3C-7**. Choose d as $Val_{\mathbf{j}}$(t), which belongs to D by $Val_{\mathbf{j}}$type, **3C-7**, so that $Val_{[Val_{\mathbf{j}}(t)/v]\mathbf{j}}$(A) = T. So $Val_{\mathbf{j}}$([t/v](A)) = T by the hypothesis that t is closed and the semantics of substitution theorem **3C-14**.

5. Use the Corollary to the local determination theorem, **3C-13**.

6. Only the Definition **3C-7** of $Val_{\mathbf{j}}$ is required.

7. See below. □

Exercise 57	*(Validity of some more quantificational truths)*

Prove part 7 of Theorem **3C-22**. It is in fact more difficult than the other parts, requiring a wise choice of instantiation for the universal quantifier over Q-interpretations used in the definition (**3C-18**) of $\vDash_Q$.

▷ . . . . ◁

3D Elementary proof theory of S_Q

This section on the proof theory of the quantifier language we are describing can be short because we have treated nearly all the complexities in somewhat simpler settings elsewhere; you will not, of course, be misled into supposing that the topic itself is without complexity.

3D.1 Proof theoretical choices for S_Q

Just as for grammar and for semantics, there are choices to be made concerning the proof apparatus—we call it "S_Q"—of our quantificational logic.

Axioms or rules? Just as for S_{TF}, we want to do as much work as possible with axioms rather than rules, partly, we suppose, for elegance, and partly because it makes the technical work in some respects simpler: As for S_{TF}, there will be only one inductive clause in the definition of proof-theoretical consequence, namely, modus ponens. Designing S_Q this way renders it an "axiomatic extension" of S_{TF}, so that we know automatically that numerous properties of S_{TF} *must* hold for S_Q, without further worry.

The price we pay is this: It is hard work to show that we can still get the effect of universal generalization.

Another price is this: Our axioms will themselves be given inductively.

Completeness for all formulas, or only closed ones? In both grammar and semantics we dealt in an evenhanded way with all formulas, both open and closed. In our development of proof theory, however, it is simpler to aim for an inductive characterization of the valid inferences from G to A that will be adequate only under the assumption that G and A are closed. That is, although we catch *some* of the valid consequences among open formulas in our inductive net, we will not try to catch them all (we will be incomplete), but we *will* aim to encompass all the valid inferences involving only closed sentences. For example, the inference from $\forall$vFv to Fv is evidently sanctioned as valid by our semantic account of consequence, that is, $\forall$vFv $\vDash_Q$ Fv; but our proof theory will not place its stamp of approval on the inference from $\forall$vFv to Fv. The reason is technical: If we lower our sights, trying to authorize with our proof theory only semantically correct inferences among closed sentences, then it is not necessary to worry about instantiating quantified variables with terms containing free variables that might become bound (recall the complexity of Definition **3B-24**). And although this is without doubt a limitation, it is not terribly harmful in spite of the fact that quite generally, and as recently as part 4 of the proof of **3C-22**, we ourselves instantiate with variables of quantification. It is relevant that in our use-logic we could have always and only used closed sentences: We could have arranged things so that neither any of our premisses, nor any conclusion, nor even any intermediate step ever contained a free occurrence of a letter used (elsewhere) as a bound variable. We could have done this by never using any letter both as a bound use-variable and as a letter (individual parameter) for instantiation, thus following advice frequently given to beginners in logic. Still, we repeat that restricting our completeness claim to arguments involving only closed sentences is indeed a limitation, a limitation we accept for the sake of technical simplicity.

Only sets of premisses economical as to constants? Call a set of premisses G "economical as to constants" if it omits (that is, fails to contain) infinitely many constants. Such a premiss-set is called "economical" because it considers that in the future it might need some of those unused constants for use with universal generalization. As you might guess, it is technically difficult to deal with premiss-sets that are not economical as to constants, just because of not knowing what to do when faced with a desire to argue by some analog of universal generalization. Even so, it is better not to limit ourselves to economical sets of premisses. It is true that you might not even have noticed the limitation if we had laid it on, since you have been accustomed to think only of *finite* premiss-sets, which are of course economical: Since there are infinitely many constants in all, any finite premiss-set *must* omit infinitely many of them. But even though it might seem that we could confine attention to economical sets by considering only finite sets, it is not so: Infinite premiss-sets are important. As a paradigm example, consider one of our Closure clauses: As we have framed them, they each involve infinitely many use-sentences, one for each use-context Ψ(A). Furthermore, the same example carries us even further, for the Closure clauses are not only infinite, but decidedly uneconomical: Assuming our use-language is fixed (we don't suppose it really is, but make the assumption), "every use-context Ψ(A)" must run through all the use-constants.

For reason of their importance in applications, then, it is good for us *not* to exclude premiss-sets that are not economical as to constants; we must face the problems their wastefulness creates. We shall, however, segregate the technical problems caused by uneconomical premiss-sets into §3F.

3D.2 Basic proof theoretic definitions for S_Q

As before, §2E.2, we first define what it is to be an axiom.

3D-1 DEFINITION. *(Ax*S_Q*)*

Basis clause. If A, B, and C are sentences of Q and if v is a variable, then the following, S_Q1–6, are members of AxS_Q:

$A \supset (B \supset A)$ (S_Q1)

$((A \supset (B \supset C)) \supset ((A \supset B) \supset A \supset C))$ (S_Q2)

$(\sim A \supset \sim B) \supset (B \supset A)$ (S_Q3)

$(\forall vA \supset [t/v](A))$, if t is closed ($S_Q4$)

$A \supset \forall vA$, if v does not occur free in A (S_Q5)

$(\forall v(A \supset B) \supset (\forall vA \supset \forall vB))$ (S_Q6)

Inductive clause. Let c not occur in A. $[c/v](A) \in AxS_Q \rightarrow \forall vA \in AxS_Q$. ($S_Q7$)

Closure clause. That's all the members of AxS_Q: Let $\Psi(A)$ be a use-context. **Suppose**

> **Basis step.** $\Psi(A)$ when A is an axiom by one of S_Q1-6.
>
> **Inductive step.** For all c and v, (c does not occur in A and $\Psi([c/v](A)) \rightarrow \Psi(\forall vA)$).

Then for all A, $A \in AxS_Q \rightarrow \Psi(A)$. (Induction on AxS_Q)

Observe that the set of axioms of S_Q is inductively defined, which puts it in high contrast to the set of axioms of S_{TF}.

Exercise 58 *(Closure clause for* AxS_Q*)*

Represent the closure clause for AxS_Q as a subproof rule.

▷ . . . . ◁

Next we define $\vdash_{S_Q}$. The definition follows precisely that for $\vdash_{S_{TF}}$, **2E-5**.

3D-2 DEFINITION. *(*$\vdash_{S_Q}$*)*

$\vdash_{S_Q}$ is a relation between sets of sentences of Q and sentences. ($\vdash_{S_Q}$type)

Base clause.

$A \in G \to G \vdash_{S_Q} A$ $(\vdash_{S_Q}id)$

$A \in AxS_Q \to G \vdash_{S_Q} A$ $(\vdash_{S_Q}AxS_Q)$

Inductive clause.

$(G \vdash_{S_Q} A \supset B$ and $G \vdash_{S_Q} A) \to G \vdash_{S_Q} B$ $(\vdash_{S_Q}MP)$

Closure clause. That's all. **Suppose**

Basis step. ($A_1 \in G$ or A_1 is an axiom of S_Q) $\to \Psi(A_1)$, all A_1.

Inductive step. ($\Psi(A_1)$ and $\Psi(A_1 \supset A_2)$) imply $\Psi(A_2)$, all A_1, A_2.

Then $G \vdash_{S_Q} A \to \Psi(A)$, all A. $(\vdash_{S_Q}$Ind, or Induction on S_Q consequence)

We may take over single turnstile statements straightway in proof theory, just as we did in semantics, **3C-19**.

3D-3 FACT. *(S_{TF} and S_Q)*

$G \vdash_{S_{TF}} A \to G \vdash_{S_Q} A$

PROOF. Straightforward, by Induction on S_{TF}-consequence, **2E-6**. □

We next wish to show that many relations between S_{TF}-turnstile statements also hold between S_Q-turnstile statements. Much—but not all—of what is needed for completeness is given therein. To this end, we point out that S_Q is an "axiomatic extension" of S_{TF}, and prove some properties about all such.

3D-4 DEFINITION. *(Axiomatic extension)*

Suppose $\vdash_S$ and $\vdash_{S'}$ are subsets of ($\mathcal{P}$(Sent) × Sent). Then S' is an *axiomatic extension* of S if and only if there is a set H (the new axioms of S') such that for all sets G and sentences A,

$$G \vdash_{S'} A \leftrightarrow (G \cup H) \vdash_S A.$$

3D-5 FACT. *(S_Q is an axiomatic extension of S_{TF})*

This fact is the same as its name: S_Q is an axiomatic extension of S_{TF}.

PROOF. Choose H as all the axioms of S_Q, that is, let $H = AxS_Q$, **3D-1**, (you may leave out those that are also axioms of S_{TF} if you wish; it doesn't matter); we need to show $G \vdash_{S_Q} A \leftrightarrow G \cup H \vdash_{S_{TF}} A$. This is perhaps obvious by inspection; but the thing can be verified by using Induction on S_Q-consequence, **3D-2**, from left to right, and Induction on S_{TF}-consequence, **2E-6**, from right to left. □

The reason we are interested in axiomatic extensions is this:

3D-6 FACT. *(Preservation of properties under axiomatic extension)*

All of the properties listed in Theorem **2E-22** except "Induction on" are preserved under axiomatic extension: If they hold for S, then they hold for any axiomatic extension S′ of S. (For $\vdash_{S'}$fin, we must assume that $\vdash_S$weak holds as well as $\vdash_S$fin.)

The proof is too tedious to bear: The fact has to be verified for each property.

Exercise 59 *(Verification of **3D-6**)*

Carry out the verification at least for $\vdash_S$fin.

▷ . . . . ◁

3D-7 THEOREM. *(Properties of S_Q as an axiomatic extension of S_{TF})*

Analogs of all the properties listed in Theorem **2E-22** except "Induction on" hold for $\vdash_{S_Q}$.

PROOF. Trivial, given the previous two facts, **3D-5** and **3D-6**. □

3D.3 Universal generalization in S_Q

This section contains the only really *particular* results about the proof theory of quantifier logic that we take up, namely, results leading to the one rule of our quantifier logic that is most delicately bound up with bound variables: universal generalization.

How can we represent the rule of universal generalization in S_Q? Of course we know that it is *not* the case that whenever [c/v](A) follows from a set of premisses G, so does ∀vA; for G might contain some special information involving c which forbids our interpreting it as "flagged." There is one case, however, in which we can be sure that c can be thought of as "arbitrarily chosen" and hence fit for universal generalization: When it does not occur at all in G, or in A.

We noted just after Theorem **3C-22** the *semantic* fact that the absence of c from G and A guarantees that if $G \vDash_Q$ [c/v](A) then $G \vDash_Q$ ∀vA; we should therefore expect that the analog holds for $\vdash_{S_Q}$, as an appropriate *proof theoretical* representation of the rule of universal generalization. In fact, given the to-be-proved interchangeability of $\vDash_Q$ and $\vdash_{S_Q}$, you can see that the absence of the analog would imperil the completeness (or perhaps the consistency) of S_Q.

Of course we could have decided to help out S_Q by simply taking the analog as primitive; we chose, however, not to do so for powerful reasons given above in §3D.1: We would no longer have an axiomatic extension of S_{TF}, so that *other* aspects of our investigations which are now simple would have instead been complex.

But now we must pay the price: It is (we think) unavoidably tedious to show that our apparently weak system S_Q is strong enough to permit a proof of "universal generalization," ($\vdash_{S_Q}$∀, **3D-8**). The tedium arises from having to rely on properties of substitution; we present them as best we can. If you are not interested, just skip to Theorem **3D-10**.

3D-8 THEOREM. *($\vdash_{S_Q}$∀)*

$G \vdash_{S_Q}$ [c/v](A) → $G \vdash_{S_Q}$ ∀vA, provided that c does not occur in either G or A.

PROOF. Let us work from the outside in. We want to prove the theorem by Induction on S_Q consequence, **3D-2**; but as you can see, what we wish to prove does not exhibit this form on its surface. Instead of making the necessary qualifications implicitly, we reformulate what needs to be proved as follows.[18] (In the following,

[18] This proof replaces a clumsier previous version. Although we have tinkered with presentation in the service of pedagogy, it is in every essential due to M. Kremer.

bold parentheses are just variant parentheses of convenience. And parenthesized letters such as *(a)* give us an easy way to refer to assumptions, intermediate results, or things-to-be-proved.)

> *(z)* $G \vdash_{S_Q} A \rightarrow$ for every v, c, and B, **(**(A = [c/v](B) and c occurs in neither G nor B) $\rightarrow G \vdash_{S_Q} \forall vB$**)**.

We aim, then, to show *(z)*, which will suffice. We show *(z)* by Induction on S_Q consequence, **3D-2**, as announced. We use Ψ(A) for the entire quantified consequent of *(z)*, as follows:

> Ψ(A): For every v, c, and B, ((A = [c/v](B) and c occurs in neither G nor B) $\rightarrow G \vdash_{S_Q} \forall vB$).

The premisses for deriving *(z)* by Induction on S_Q consequence, **3D-2**, are these:

> **Basis step (abbreviated).** For every A_1, (either $A_1 \in G$ or $A_1 \in AxS_Q$) $\rightarrow$ Ψ(A).
>
> **Inductive step (abbreviated).** For every A_1 and A_2, **(**($\Psi(A_1)$ and $\Psi(A_1 \supset A_2)$) $\rightarrow \Psi(A_2)$**)**.

So it suffices to show *(x)* and *(y)* below, which are the unabbreviated forms of the base and inductive steps. (We reletter so as to avoid confusion of bound variables.)

> **Basis step.** *(x)* For every A_1, ($A_1 \in G$ or $A_1 \in AxS_Q$) $\rightarrow$ for all v, c, and B_1, **(**(A_1 = [c/v](B_1) and c occurs in neither G nor B_1) $\rightarrow G \vdash_{S_Q} \forall vB_1$**)**.
>
> **Inductive step.** *(y)* For each A_1 and A_2, suppose both of the following.
>
> > For all v, c, and B_1, ((A_1 = [c/v](B_1) and c occurs in neither G nor B_1) $\rightarrow G \vdash_{S_{TF}} \forall vB_1$).
> >
> > For all v, c, and C, **(**($A_1 \supset A_2$ = [c/v](C) and c occurs in neither G nor C) $\rightarrow G \vdash_{S_Q} \forall vC$**)**.
>
> **Then** for all v, c, B_2, **(**(A_2 = [c/v](B_2) and c occurs in neither G nor B_2) $\rightarrow$ $G \vdash_{S_Q} \forall vB_2$**)**.

So evidently to show *(z)* it suffices to show *(x)* and *(y)*. First we show the base step *(x)*. Choose A_1. Suppose that *(a)* either $A_1 \in G$ or $A_1 \in AxS_Q$. We need to show the consequent of *(x)*. Choose v, c, and B_1, and suppose *(b)* $A_1 = [c/v](B_1)$ and that *(c)* c does not occur in G and *(d)* c does not occur in B_1. We need to show that *(v)* G $\vdash_{S_Q} \forall vB_1$. There are obviously two cases.

Suppose the first disjunct, *(a_1)* $A_1 \in G$. We show *(v)*. In this case the key observation is that v is not free in B_1, which we obtain as follows. By *(c)* and *(a_1)*, c does not occur in A_1. So by *(b)*, c does not occur in $[c/v](B_1)$. So by **3B-20**(5), *(e)* v is not free in B_1. This has two interesting consequences. First, *(e)* means by **3B-20**(1) that the substitution $[c/v](B_1)$ is vacuous, i.e., $[c/v](B_1) = B_1$, so that by *(b)*, $A_1 = B_1$. But then $B_1 \in G$ by *(a_1)*, so that *(f)* G $\vdash_{S_Q} B_1$ by $\vdash_{S_Q}$id, **3D-2**. Second, *(e)* means that $B_1 \supset \forall vB_1$ is an axiom of S_Q (S_Q5, **3D-1**). Therefore *(g)* G $\vdash_{S_Q} B_1 \supset \forall vB_1$ by $\vdash_{S_Q}$AxS$_Q$, **3D-2**. The desired *(v)* now follows from *(f)* and *(g)* by $\vdash_{S_Q}$MP, **3D-2**.

Next suppose the second disjunct, *(a_2)* $A_1 \in$ AxQs. We show *(v)*. The argument here is considerably shorter. We know from *(b)* and *(a_2)* that *(e)* $[c/v](B_1) \in AxS_Q$. But *(e)* together with *(d)* permits us to use S_Q7, **3D-1**, to infer that $\forall vB_1 \in AxS_Q$. So *(v)* by $\vdash_{S_Q}$AxS$_Q$, **3D-2**.

This completes the argument for the base step *(x)*. In order to use Induction on S_Q consequence, **3D-2**, we must also show the inductive step *(y)*. Choose A_1 and A_2, and suppose the two antecedents of *(y)*:

> *(a_1)* For all v, c, and B_1, (($A_1 = [c/v](B_1)$ and c occurs in neither G nor B_1) $\rightarrow$ G $\vdash_{S_{TF}} \forall vB_1$).
>
> *(a_2)* For all v, c, and C, ($A_1 \supset A_2 = [c/v](C)$ and c occurs in neither G nor C) $\rightarrow$ G $\vdash_{S_Q} \forall vC$.

It is enough to show the consequent of *(y)*, namely,

> *(w)* for all v, c, B_2, (($A_2 = [c/v](B_2)$ and c occurs in neither G nor B_2) $\rightarrow$ G $\vdash_{S_Q} \forall vB_2$).

So choose v, c, and B_2, and assume *(b)* $A_2 = [c/v](B_2)$, *(c)* c does not occur in G, and *(d)* c does not occur in B_2. It suffices to show *(u)* G $\vdash_{S_Q} \forall vB_2$.

Making a key choice, let v′ be a variable not occurring in $A_1 \supset B_2$ (**3B-19**), so that by occurΦ, **3B-15**, *(e)* v′ does not occur in A_1 and *(f)* v′ does not occur in B_2.[19]

---

[19]We don't care much about the formula $A_1 \supset B_2$; it is introduced just as a convenient way to apply **3B-19** in order to obtain a variable that occurs in neither of A_1 nor B_2. It's a good thing there is one, for otherwise this proof would not go through.

We now instantiate *(a_1)* as follows: v′ for v, c for c, and (carefully) [v′/c](A_1) for B_1. This yields the following.

> *(a_1′)* (A_1 = [c/v′](([v′/c](A_1))) and c occurs in neither G nor [v′/c](A_1)) → G $\vdash_{S_{TF}}$ ∀v′[v′/c](A_1).

We establish the three conjuncts of the antecedent as follows. That *(g)* A_1 = [c/v′](([v′/c](A_1))) follows from *(e)* and **3B-20**(2). The second conjunct is just *(c)*; and **3B-20**(4) promises that *(h)* c does not occur in [v′/c](A_1). So *(i)* G $\vdash_{S_{TF}}$ ∀v′[v′/c](A_1). We save *(i)* for a bit.

Next we instantiate *(a_2)* with v′ for v, c for c, and (even more carefully) ([v′/c](A_1) ⊃ [v′/v](B_2)) for C, giving

> *(a_2′)* (A_1 ⊃ A_2 = [c/v′](([v′/c](A_1) ⊃ [v′/v](B_2))) and c occurs in neither G nor ([v′/c](A_1) ⊃ [v′/v](B_2))) → G $\vdash_{S_Q}$ ∀v′([v′/c](A_1) ⊃ [v′/v](B_2)).

We establish the antecedent of *(a_2′)* as follows. First the critical identity, starting from the complex (right) side.

[c/v′](([v′/c](A_1) ⊃ [v′/v](B_2)))

= [c/v′]([v′/c](A_1)) ⊃ [c/v′]([v′/v](B_2))	O[]Φ, **3B-13**
= A_1 ⊃ [c/v′]([v′/v](B_2))	*(g)*
= A_1 ⊃ [c/v](B_2)	*(f)*, **3B-20**(3)
= A_1 ⊃ A_2	*(b)*

Second, that c does not occur in G is again just *(c)*. And last, that c does not occur in ([v′/c](A_1) ⊃ [v′/v](B_2)) comes from applying occurΦ (**3B-15**) to *(h)* and the result of using *(d)* with **3B-20**(6). So *(j)* G $\vdash_{S_Q}$ ∀v′([v′/c](A_1) ⊃ [v′/v](B_2)). Putting *(j)* together with the saved *(i)* suggests what to do next: Derive

> *(k)* G $\vdash_{S_Q}$ ∀v′[v′/v](B_2)

by S_Q6 (**3D-1**) together with $\vdash_{S_{TF}}$AxS$_Q$ and $\vdash_{S_Q}$MP of **3D-2**.

We are close, but evidently there is a gap between *(k)* and the desired goal,

> *(u)* G $\vdash_{S_Q}$ ∀vB_2.

We may bridge this last gap—more accurately, a chasm filled with brackets, parentheses, primes, and slashes—by appeal to Lemma **3D-9** just below, for the content of that lemma is precisely that *(f)* and *(k)* imply *(u)*. □

Here is the gap-bridging lemma.[20]

3D-9 LEMMA. *(Special change of bound variable)*

If v' does not occur in A and $G \vdash_{S_Q} \forall v'[v'/v](A)$, then $G \vdash_{S_Q} \forall vA$.

PROOF. Assume *(a)* v' does not occur in A and that *(b)* $G \vdash_{S_Q} \forall v'[v'/v](A)$. We need to show that *(z)* $G \vdash_{S_Q} \forall vA$. Fact **3B-19** justifies choosing c so that *(c)* c does not occur in A. Then *(d)* $[c/v']([v'/v](A)) = [c/v](A)$ by **3B-20**(3). Also v does not occur free in $\forall v'[v'/v](A)$ by **3B-20**(7), so *(e)* $\forall v'[v'/v](A) = [c/v]((\forall v'[v'/v](A)))$ by **3B-20**(1). Now

$[c/v]((\forall v'[v'/v](A) \supset A))$

	$= [c/v]((\forall v'[v'/v](A))) \supset [c/v](A)$	$[]O\Phi$, **3B-13**
	$= \forall v'[v'/v](A) \supset [c/v']([v'/v](A))$	*(d)* and *(e)*
(f)	$\in AxS_Q$.	S_Q4, **3D-1**

Furthermore, *(c)* says that c does not occur in A, and **3B-20**(6) adds that it also fails to occur in $[v'/v](A)$, so *(g)* c coes not occur in $\forall v'([v'/v](A) \supset A)$, by occur$\Phi$ and occur$\forall$ of **3B-15**. So by *(f)* and *(g)* and S_Q7, **3D-1**, $\forall v \forall v'([v'/v](A) \supset A) \in AxS_Q$; which implies that

(h) $G \vdash_{S_Q} \forall v \forall v'([v'/v](A) \supset A)$

by $\vdash_{S_Q} AxS_Q$, **3D-2**. Also

(i) $G \vdash_{S_Q} \forall v(\forall v'[v'/v](A) \supset A) \supset (\forall v \forall v'[v'/v](A) \supset \forall vA)$

[20]**Exercise**. Find a proof of Theorem **3D-8** that does not rely on a separate lemma; and decide whether or not the new proof is an improvement. In this context someone might think to add the content of Lemma **3D-9** as an axiom in the form (S_Q6') $\forall v'[v'/v](A) \supset \forall vA$, provided v' does not occur in A. That would certainly close the gap in a hurry. Of course the self-same proof of the lemma would still be required as a proof that the newly added (S_Q6') is redundant, thus illustrating the precept, "never demolish a bridge once crossed."

by S_Q6 and $\vdash_{S_Q}$AxS$_Q$. So

(j) $G \vdash_{S_Q} \forall v \forall v'[v'/v](A) \supset \forall vA$

from *(h)* and *(i)* by $\vdash_{S_Q}$MP. Save this. Penultimately, by **3B-20**(7), v is not free in $\forall v'[v'/v](A)$. So

$G \vdash_{S_Q} \forall v'[v'/v](A) \supset \forall v \forall v'[v'/v](A)$

by S_Q5 and $\vdash_{S_Q}$AxS$_Q$. We are now in a position to use *(b)* with $\vdash_{S_Q}$MP to obtain

$G \vdash_{S_Q} \forall v \forall v'[v'/v](A)$.

And finally *(z)* comes from this and *(j)* by another $\vdash_{S_Q}$MP. □

Having established that "universal generalization" holds for $\vdash_{S_Q}$, we summarize the chief proof-theoretical facts concerning the universal quantifier. These will all be required in the course of proving that S_Q is complete. You will recognize $\forall\vdash_{S_Q}$ as universal quantifier instantiation; and you can see that $\sim\forall\vdash_{S_Q}$ is a close cousin of existential instantiation.

3D-10 THEOREM. *($\vdash_{S_Q}$ and $\forall$)*

$\forall vA \vdash_{S_Q} [t/v](A)$, if t is closed ($\forall\vdash_{S_Q}$)

$G \vdash_{S_Q} [c/v](A) \rightarrow G \vdash_{S_Q} \forall vA$ if c does not occur in either G or A. ($\vdash_{S_Q}\forall$)

G, $\sim([c/v](A))$, $\sim\forall vA \vdash_{S_Q} B \rightarrow$ G, $\sim\forall vA \vdash_{S_Q} B$, if c does not occur in G, A, or B. ($\sim\forall\vdash_{S_Q}$)

PROOF. Show $\forall\vdash_{S_Q}$ by S_Q4 and obvious properties of $\vdash_{S_Q}$.

We already proved $\vdash_{S_Q}\forall$ as Theorem **3D-8**.

Proof of $\sim\forall\vdash_{S_Q}$ is easy, given $\vdash_{S_Q}\forall$ and **3D-3**. □

In parallel with **2E-22**, we summarize the properties of S_Q required for completeness and consistency. You will note that the level of difficulty in proving these properties ranges from Absolutely Trivial (e.g., (1) is a conjunct of a definition) to Beyond the Pale of Tedium (e.g., (13), for remembering the details of the proof of which only these notes are reasonably held responsible).

3D-11 THEOREM. *(Some properties of $\vdash_{S_Q}$)*

1. $\vdash_{S_Q}$id, **3D-2**.
2. $\vdash_{S_Q}$AxS$_Q$, **3D-2**.
3. $\vdash_{S_Q}$MP, **3D-2**.
4. Induction on S$_Q$-consequence, **3D-2**.
5. $\vdash_{S_Q}$weak, **3D-7**.
6. $\vdash_{S_Q}$cut, **3D-7**.
7. $\vdash_{S_Q}$fin, **3D-7**.
8. Maximality/S$_Q$-closure (which follows from $\vdash_{S_Q}$cut alone, **3D-7**). (In fact we shall need a slightly revised version of this, **3E-3**; but its proof is the same as the proof of **2E-13**.)
9. Deduction theorem, **3D-7**.
10. $\vdash_{S_Q}$S$_{TF}$1-6, **3D-7**.
11. $\vdash_{S_Q}$S$_Q$1-7, **3D-1** with 2 above.
12. $\forall\vdash_{S_Q}$, **3D-10**.
13. $\vdash_{S_Q}\forall$, **3D-8**.
14. $\sim\forall\vdash_{S_Q}$, **3D-10**.

PROOF. You should already be thoroughly clear as to how each of these is established. □

Exercise 60 *(Elementary proof theory of S$_Q$)*

1. Be able to prove all of the properties of $\vdash_{S_Q}$ listed in **3D-11**, except for the too tedious 13. Some are just parts of definitions.
2. Optional (and an opportunity to make a contribution): Find an even less tedious proof of $\vdash_{S_Q}\forall$, **3D-8**.

3. Prove $\sim\forall\vdash_{S_Q}$.

4. Prove a syllogism, say

$$\forall v(Fv \supset Gv), \forall v(Gv \supset Hv) \vdash_{S_Q} \forall v(Fv \supset Hv).$$

   By all means use any fact already established about S_Q (especially **3D-3**, **3D-7**, and **2G-1**), for this and also for succeeding problems. (You should be able to do this problem in a few steps; otherwise you are on the wrong track.)

5. Prove that the De Morgan challenge to Aristotelian logic (since horses are animals, a tail of a horse is a tail of an animal) is admitted by S_Q:

$$\forall v(Fv \supset Gv) \vdash_{S_Q} \forall v'(\exists v(Fv\&Rv'v) \supset \exists v(Gv\&Rv'v)).$$

   (This should take no more than eight or ten steps, provided you are on track.)

6. Prove at least one other "Introductory logic" exercise as a S_Q-consequence.

7. Optional. Prove that according to S_Q, there is a paradigm horse:

$$\vdash_{S_Q} \exists v(\exists vFv \supset Fv).$$

   Or prove that S_Q says that there is a woman such that if she is infertile then the whole human race will die out: $\vdash_{S_Q} \exists v(Fv \supset \forall vFv)$.

▷ . . . . ◁

3E Consistency and completeness of S_Q

Having presented the elements of the grammar, semantics, and proof theory of quantifiers, we turn to the chief theorem: consistency and completeness, on the model of §2F. In fact we aim to treat only those features of the situation that are caused by the presence of the quantifiers themselves.

3E.1 Consistency of S_Q

3E-1 THEOREM. *(Consistency theorem for* S_Q*)*

$G \vdash_{S_Q} A \rightarrow G \vDash_Q A$

PROOF. We need to prepare with a lemma: If $A \in AxS_Q$ then $\models_Q A$. This is proved by Induction on AxS_Q, **3D-1**, using Six logical truths and some more, **3C-22**, which were, of course, designed for this very purpose.

The lemma establishes that every axiom of S_Q is a correct semantic consequence of the empty set. To show that every consequence statement authorized by S_Q is semantically correct, proceed by Induction on S_Q consequence, **3D-2**, choosing $\Psi(A)$ there as "$G \models_Q A$" here. The preceding lemma will be helpful in part of the base case, and you will need some $\models_Q$ facts, **3C-21**. □

3E.2 Completeness of S_Q

The completeness proof comes by analogy with our proof for S_{TF}, with such new auxiliary ideas and lemmas as are required. You may wish to glance over the beginning of §2F.2. We begin with an overview of the ideas.

1. Limitation to closed sentences. It is boring but necessary continually to repeat this limitation of the claim to completeness that we aim to establish. We observe that this is not really an "auxiliary" idea, since the limitation is part of what in §3D.1 we chose to mean by "completeness."

2. Worries as forecast in §3D.1 about economy as to constants: Reduction of the problem of completeness for uneconomical premiss-sets to the problem for those that are economical.

3. Maximal E-free. The idea must be adjusted to S_Q and to closed sets of sentences.

4. $\sim\forall$-complete. A set is $\sim\forall$-complete if whenever it contains $\sim\forall vA$, it also contains $\sim[t/v](A)$ for some closed term t. You may well think of $\sim[t/v](A)$ as a "counterexample" to $\forall vA$, hence as a witness to the truth of $\sim\forall vA$.

5. Truth-like set. Adjust TL$\sim$ and TL$\supset$ to closed sentences, and add a clause TL$\forall$ saying that closed $\forall vA$ is in the set just in case all its instances $[t/v](A)$ (with t closed) are also in the set.

6. Canonical interpretation. This concept needs to be substantially revised to suit the new idea of interpretation, chiefly because terms as well as sentences must be interpreted.

Now for the definitions. The first is identical to **2E-11** except for the reference to S_Q and the restriction to closed sets—recall from **3B-21** that for a *set* of sentences to be closed is for each of its members to be closed.

3E-2 DEFINITION. *(Maximal closed and E-free for* S_Q*)*

G* is *maximal closed and* E-*free in* S_Q ↔ G* is closed and E-free in S_Q but no proper superset of G* is closed and E-free in S_Q:

G* is closed and G* $\nvdash_{S_Q}$ E. $(MCEF_{S_Q}1)$

(H)[(H is closed and G* ⊂ H) → H $\vdash_{S_Q}$ E]. $(MCEF_{S_Q}2)$

We state the closure property required, its proof being in analogy to that of **2E-13**:

3E-3 FACT. *(Maximal closed/*S_Q*-closure)*

G* is maximal closed and E-free in S_Q → G* is closed under S_Q-consequence for closed A; that is, for all closed A, G* $\vdash_{S_Q}$ A → A ∈ G*.

3E-4 DEFINITION. *(∼∀-complete in* Q*)*

G* is ∼∀*-complete in* Q ↔ for every sentence A and for every variable v, if ∼∀vA ∈ G* then for some closed term t, ∼[t/v](A) ∈ G*.

3E-5 DEFINITION. *(Truth-like set for* Q*)*

G* *is a truth-like set for* Q ↔ G* is closed and satisfies TL∼ and TL⊃ of **2F-2** restricted to closed sentences together with a clause TL∀ for ∀:

G* is closed. (TLclosed)

∼A ∈ G* ↔ A ∉ G*, all closed A. (TL∼)

(A ⊃ B) ∈ G* ↔ (A ∈ G* → B ∈ G*), all closed A, B. (TL⊃)

For closed ∀vA: ∀vA ∈ G* ↔ (for all closed t, [t/v](A) ∈ G*). (TL∀)

TL∀ is sometimes given as a truth definition for sentences ∀vA; that is, such a sentence is said to be true on a Q-interpretation when all of its instances are. This is the "substitution interpretation" of the quantifiers: You can convince yourself in various ways that it is *not* the same as our domain-and-values interpretation. Discussion of the relations between these two interpretations has led to some philosophy, some good, some bad.

The definition of "canonical Q-interpretation" must reflect the fact that we have considerably beefed up the notion of "interpretation" for Q.

One idea will be the same: Some set G* will mastermind what we count as true. But in addition, we need a *domain*, and a value for each of our *terms* (as well as sentences). What shall we choose? The fundamental idea is due in effect to Gödel: Be massively self-referential. Choose the domain as the terms themselves, and fix things so that each term has itself as its value—$Val_{\mathbf{j}}(t)=t$ (total autonymy)! This aim suggests the following, which we adopt.

3E-6 DEFINITION. *(Canonical* Q*-interpretation)*

---

j is a *canonical* Q*-interpretation determined by* G* ↔ **j** is a Q-interpretation, **3C-2**, and

$x \in \mathbf{j}(\forall) \leftrightarrow$ x is a closed term; let $D=\mathbf{j}(\forall)$.

$\mathbf{j}(c)=c$.

$(\mathbf{j}(f))(t_1, \ldots, t_n)=ft_1\ldots t_n$, all closed terms t_k, $1\leqslant k\leqslant n$.

$(\mathbf{j}(F))(t_1, \ldots, t_n)=T \leftrightarrow Ft_1\ldots t_n \in G^*$, all closed terms t_k, $1\leqslant k\leqslant n$.

If you keep in mind that the domain consists of terms, you may save your sanity. In the clause for constants, the argument for **j** is a term, and the value (named on the right) is an entity (accidently also a term). In the clauses for operators and predicates, names of entities occur on the left, and names of terms on the right (which of course come to the same).[21]

[21] We might add that because of earlier choices, we could nearly have said simply: $\mathbf{j}(f)=f$. Had we made (for example) "f" range over "operator *symbols*," as is usual, this would not have been possible, for then given $\mathbf{j}(f)=f$, "$\mathbf{j}(f)$" would also have denoted a symbol— say, the symbol "TF." But then "$(\mathbf{j}(f))t$" would have been a senseless string of symbols, just as senseless as ""P"t," which (in our use-language) means nothing at all. But, in contrast, our earlier choice was to use (for example) "f" as a variable ranging over (not symbols but) *operators*; and operators are functions of a certain sort, namely grammatical functions: $f\in(\text{term}\mapsto\text{term})$ when f is one-place. Hence, if we

It is an important part of the definition that we say that **j** is a Q-interpretation; for that clause automatically determines **j**(⊃) and **j**(∼), which are the same for every Q-interpretation. Also we know from this clause that **j**(v) is in the domain for every variable v, since this is so for every Q-interpretation; but we are not told what this value is; for this reason, we must take seriously that **j** is *a* canonical Q-interpretation and not *the* canonical Q-interpretation determined by G*. (Because we shall be dealing only with closed sentences, we can be sure that Q-interpretations of variables don't matter—by the local determination theorem for Q, **3C-12**.)

As before (**2F-5**) we need a small

3E-7 FACT. *(Existence of canonical* Q*-interpretation)*

For each set G*, there is an **j** such that **j** is a canonical Q-interpretation determined by G*.

PROOF. Omitted, as for **2F-5**. □

So much for definitions; next a few words about the problem raised by a G which wastefully runs through all the constants. You will recall from **3B-1** that we assumed infinitely many individual constants with the idea that they would be used as auxiliaries in proofs using the rule of universal generalization or the like. As it happens, we must be sure that there are infinitely many available in the context of evaluating any consequence-statement, so there is a problem if G and A use up (say) all the constants—there are none left with which to instantiate. One thing we could do is to add some new ones; technically this would lead us to start comparing languages with different vocabularies, which is possible and interesting, but complicated. Instead, we show that we already have enough resources in the language as it is. We do this not by pretending that G and A don't use up all the constants when they do, but by showing that in place of considering the given G and A, we can find a closely related H and A′ (in fact obtained by wholesale substitution) which (a) omit infinitely many constants, as desired, and (b) have the same consequence-relation as do G and A. We will have to show this both semantically and proof-theoretically.

---

let **j**(f) = f, both "**j**(f)" and "f" would denote (grammatical) functions, and functions of nearly the same type (i.e., (term ↦ term); so that we could conclude by logic, without further intervention, that ((**j**(f))t = ft. (One of the purposes of this note is to illustrate the contorted forms into which we must twist ourselves when we wish to both use and mention language simultaneously.) But for accuracy, observe that **j**(f), though very like f, is in fact restricted to *closed* terms, whereas f is not.

We shall now state the five required lemmas, the fifth of which outlines the course of the proof.

3E-8 LEMMA. *(Reduction to economical sets lemma)*

Consider closed G and A. There are closed H and A′ such that there are infinitely many constants that occur neither in A′ nor in any member of H—which makes H economical as to constants—and such that the semantic and proof-theoretic consequence questions for G and A have the same answers as those for H and A′:

> H and A′ are closed.
>
> nonoccur(H∪{A′}) is infinite.
>
> $G \vDash_Q A \leftrightarrow H \vDash_Q A'$.
>
> $G \vdash_{S_Q} A \leftrightarrow H \vdash_{S_Q} A'$.

The proof of this lemma is wholly deferred until §3F.

3E-9 LEMMA. *(Lindenbaum's lemma for* S_Q*)*

Under a certain limiting assumption, every set closed and E-free in S_Q can be extended to a set maximal closed and E-free in S_Q that is also ∼∀-complete. Namely, suppose each of the following.

> G and E are closed.
>
> G is economical as to constants, so that nonoccur(G∪{E}) is infinite.
>
> $G \nvdash S_Q$ E.

Then there is a G* as follows.

> G* is closed and G⊆G*. (LL1)
>
> $G^* \nvdash S_Q$ E. ($MCEF_{S_Q}1$)
>
> (H)[(H is closed and G* ⊂ H) → $H \vdash_{S_Q}$ E]. ($MCEF_{S_Q}2$)
>
> G* is ∼∀-complete in Q. (∼∀-completeness)

3E-10 LEMMA. *(Maximality and ∼∀-completeness / truth-like-set lemma for Q)*

If a set of sentences G* is both maximal closed and E-free in S_Q, **3E-2**, and ∼∀-complete, **3E-4**, and if E is closed, then G* is a truth-like set for Q, **3E-5**. (As an accident, the natural hypothesis that E is closed is redundant.)

3E-11 LEMMA. *(Truth-like set/canonical* Q*-interpretation lemma for* Q*)*

If G* is a truth-like set for Q, **3E-5**, and if **j** is a canonical Q-interpretation determined by G*, **3E-6**, then

$Val_{\mathbf{j}}(t) = t$, all closed terms t. (Autonymy sublemma)

$Val_{\mathbf{j}}(A) = T \leftrightarrow A \in G^*$, all closed sentences A.[22]

The limitation to closed sentences is real: We know nothing about values of open terms or of open sentences on canonical Q-interpretations.

Now as before, **2F-9**, we give the course of the proof of the completeness of S_Q in a lemma.

3E-12 LEMMA. *(Course of proof of completeness of* S_Q*)*

Suppose the previous four lemmas, **3E-8** through **3E-11**. Then S_Q is complete: Where G and E are closed, $G \vDash_Q E \rightarrow G \vdash_{S_Q} E$.

PROOF. Exercise **3E-12**. Be careful about subscripts on turnstiles. Be prepared to use Theorem **3D-11**. The proof of Lemma **2F-9** can serve as a partial guide. □

Exercise 61 *(S_Q not completely complete)*

Optional. Perhaps it is worth observing that S_Q is *not* "completely complete" in the sense that the limitation to closed sentences is essential. For example, ∀vFv has Fv as a semantic consequence, but $\forall vFv \nvdash_{S_Q} Fv$. Prove this.

▷ . . . . ◁

[22]If in the spirit of Frege you let your use-sentences occupy the places of terms because you take seriously that—like the sentences of S_Q—they denote truth values, then you might put this clause this way: $Val_{\mathbf{j}}(A) = (A \in G^*)$.

There are four lemmas left to prove, **3E-8** through **3E-11**. Let us take them in reverse order, deferring the first to the next section, and here turning first to the truth-like set/canonical Q-interpretation lemma for Q, **3E-11**.

PROOF. The idea of the proof of **3E-11** is very like that of the proof of the corresponding truth-like set/canonical TF-interpretation lemma, **3E-11**, adjusting for the complications raised by the enriched powers of Q. Though it might seem best to try Induction on terms and sentences, **3B-4**, or Induction on formulas, **3B-8**, these strategies will not work because the clause for ∀vA would require us to look at the possibly *open* sentence A, whereas the hypothesis of induction would give us information only for *closed* sentences. But the lemma succumbs to Induction on *closed* terms, followed by Induction on *closed* sentences, **3B-23**. The Semantics of substitution lemma for Q, **3C-14**, is invaluable. The hardest part of dealing with this lemma is keeping straight about the Gödelian self-referential idea which casts terms in a double role as both grammatical entities having values and also as entities in the domain that are those very values. □

Next comes the Maximality and ∼∀-completeness/truth-like-set lemma for Q, **3E-10**.

PROOF. The proof of TL∼ and TL⊃ is just as before, in our proof of the maximality/truth-like-set lemma for S_{TF}, **2F-7**, except that where that proof uses **2E-13**, this one relies on **3E-3**. The added part for TL∀ is left as an exercise (recall that at this point TL∼ has already been established, and that Theorem **3D-11** is available). □

Next in line is Lindenbaum's lemma for S_Q, **3E-9**. We give a short form, as in §2F.3.

PROOF. Assume we are dealing with only closed sentences and that G and E together omit infinitely many constants; and that G $\nvdash S_Q$ E. Line up all the sentences. Starting with the E-free set G to be extended, add each closed sentence in its turn just in case its addition to what you have so far in hand does *not* lead to E; and *also* check whether it has the form ∼∀vA. If you are adding it, then *also* choose some constant c that occurs neither in E nor in any member of G nor in any sentence you have so far added nor in A. (There will always be one to choose, since G and E omit infinitely many constants, and since you have added at any stage only finitely many sentences, each of which can contain only finitely many constants.) Now add ∼[c/v](A) as well.

Let G* be the result of this process. The same argument as before shows that G* must be maximal closed and E-free in S_Q, except that we must specially verify that the addition of an instance of ∼∀vA does not upset E-freedom, using $\sim\forall\vdash_{S_Q}$, **3D-10**. And obviously G* is ∼∀-complete in Q, since we deliberately fixed things that way. □

Exercise 62 *(Slow proof of Lindenbaum's lemma for* S_Q*)*

Optional. Give a slow version of this proof, as in §2F.5. The key elements will be (1) the analog to Local definition **2F-13** of G_n and (2) managing the choice of constant with which to instantiate—the second problem being much more difficult than the first. Do this exercise without peeking at §3E.4.

▷ . . . . ◁

Most of the extra work required for extending Lindenbaum's lemma to S_Q revolves around choosing that magic constant. It seems best to take this part slowly, beginning with a small set-theoretical excursus.

3E.3 Set theory: axiom of choice

In the course of establishing Lindenbaum's lemma for S_Q, we shall be given a series of nonempty sets of constants (namely, at each stage, the set of those constants not yet occurring at that stage) and asked to pick one from each member of the series. How we pick doesn't matter, as long as each chosen constant is a member of the given set in the series from which it is supposed to be picked.

Were only one pick from a single nonempty set at stake, our "choice" would be just a matter of existential instantiation; and if there were just a fixed and finite number of picks to be made, then a series of existential instantiations would do it. Furthermore, there is not even trouble about an infinite number of picks of members from nonempty sets whenever we are given a *rule* for picking; for example, there is no trouble picking "the first" from each nonempty subset of N; plainly that is not really "picking" in the sense of "choosing," but rather what has been known as a "Hobson's choice" since 1712, when a Cambridge innkeeper allowed his customers to "choose" any horse they liked—as long as it was the one next the stable door.

In the case, however, when we must be prepared to make infinitely many choices and have no guidance as to how to do so, the question arises as to whether *there* *is* such a system of choices. Well, of course there is—intuition tells us so. Clearly infinitely many separate picks can always be reduced to a single pick—namely, they can be reduced to the single pick of a single rule for making the separate picks. It is this intuition that is enshrined in the so-called axiom of choice: Given any family of sets, there is a function, f, which when applied to any nonempty set of the family yields an element of that very set: $f(X) \in X$. The function f is sometimes called a "choice function" because it "chooses" a member from each X.

In the general case, and though as we think intuitively sound, the axiom of choice is (a) independent of many standard sets of "other" axioms for set theory, and (b) a matter of worry or concern or something to many persons interested in the foundations of mathematics. Fortunately, then, we have postulated enough axioms to avoid its use in the matter at hand, for there is a special case in which it is available: If all the nonempty subsets from which we are interested in picking a member are subsets of some one *countable* set, then we can have the effect of the axiom of choice without separate postulation:

3E-13 LEMMA. *(Countable choice)*

(Even without the axiom of choice) if a set Y is countable, then there is a function f in $(\mathcal{P}(Y) \mapsto Y)$ such that for every nonempty subset X of Y, $f(X) \in X$.

PROOF. Omitted. The idea is that we let f pick the *first* member of X—the sense of "first" coming from an enumeration of Y. □

The application we have in mind is to the constants. By Axiom **3B-1**, constant is countable; so by Countable choice, there is a function that will pick for us a constant from each nonempty set of constants. In fact let us choose such a function now, to hold good for the rest of this work, calling it "Chosenconstant":

3E-14 LOCAL CHOICE. *(Chosenconstant)*

Chosenconstant $\in ((\mathcal{P}(\text{constant})) \mapsto \text{constant})$ (Chosenconstanttype)

(X $\subseteq$ constant and X is nonempty) $\rightarrow$ Chosenconstant(X) $\in$ X

This choice is justified by **3B-1**—the part that says that the constants are countable—and **3E-13**. It amounts (as do all uses of Local choice) to a single existential instantiation. Note that in each "useful" application of **3E-14**, its antecedent must be verified: We cannot know that Chosenconstant(X) belongs to X unless we know that X is a nonempty set of constants. We might, however, want to use the notation "Chosenconstant(X)" in advance of knowing that X is not empty, so we let it be defined even in this "useless" case. Chosenconstant(∅) is therefore uselessly some constant, we know not which.

There is in the vicinity a further set theoretical fact that later (**3F-11**) proves useful; since its proof can rely on Countable choice, **3E-13**, it is convenient to locate it here. We may approach the matter in the following way. In a note in §2B.3 on p. 25, we gave Dedekind's definition of a finite set as one that cannot be mapped one-one onto a proper subset of itself. Hence, by negation, a set is "Dedekind-infinite" if it *can* be mapped one-one onto some proper subset of itself, that is, onto a subset that omits at least one member of the original set. For example, the function that takes n into $n+1$ is one-one and omits 0 from its range, and consequently shows N to be Dedekind-infinite.

It turns out via the axiom of choice that each infinite set X is not only Dedekind-infinite, but has an even stronger property: It can be mapped one-one onto a "highly proper" subset of itself, i. e., onto a subset that omits not just at least one, but an entire infinite number of members of X. For example, the function that carries n into $n+n$ is one-one (**2C-8**) and omits all the infinitely many odd numbers from its range, which shows that N is not only Dedekind-infinite, but can even be mapped one-one onto a "highly proper" subset of itself. (The key idea here goes back to Galileo, who observed that sometimes the whole is *not* greater than the part.)

Since the axiom of choice is available for countable sets without extra cost, **3E-13**, we may record what we want as the following

3E-15 FACT. *(Countable whole into part)*

Let X be countable and infinite. Then there is a one-one function f in $X \mapsto X$ such that $(X - Rng(f))$ is infinite.

PROOF. Omitted. The picture, however, is satisfactory: First enumerate $X = x_0, \ldots,$ and then run through the enumeration one stage at a time. At each stage, n, if an f-value has already been assigned to x_n, pass on. Otherwise, use Countable choice, **3E-13**, twice: once to pick out an "f-reserved" member of X (from among those not yet assigned as f-values, and not yet f-reserved) that will never be assigned as f-value to any member of X, and a second time to pick out an f-value for x_n (from among those members of X not yet assigned as f-value, and not yet f-reserved). "Clearly" this procedure generates an f as required. □

3E.4 Slow proof of the Lindenbaum lemma for S_Q

We are ready for a more articulated proof of Lindenbaum's lemma for S_Q, **3E-9**.

Suppose as hypothesis of the lemma that G and E are closed, that $G \nvdash S_Q$ E, and that G and E omit infinitely many constants: nonoccur(G∪{E}) is infinite.

In repetition of **2F-12**, we have

3E-16 LOCAL CHOICE. *(Enumeration of sentences)*

$A \in \text{Sent} \leftrightarrow \exists n(n \in N \text{ and } A = B_n)$

This choice is justified by Fact **3B-11**, which declares the sentences to be countable.

Relative to G and E and the enumeration B of **3E-16**, define a sequence G_0, …, of sets of sentences inductively as follows.

3E-17 LOCAL DEFINITION. *(G_n for S_Q)*

For each $n \in N$, G_n is that subset of Sent such that: ($G_n$type)

Basis clause. $G_0 = G$.

Inductive clauses. (Three main cases, the third with numerous subcases.)

1. Suppose that either B_n is not closed, or that $G_n, B_n \vdash_{S_Q} E$. Then:

   $G_{n+1} = G_n$.

2. Suppose B_n is closed, that $G_n, B_n \nvdash S_Q$ E, and that for all v and A, ($B_n \neq$ ∼∀vA). Then:

   $G_{n+1} = (G_n \cup \{B_n\})$.

3. Suppose B_n is closed, that $G_n, B_n \nvdash_{S_Q} E$, and that $B_n =$ ∼∀vA. First let

   $c_n = \text{Chosenconstant}(\text{nonoccur}(G_n \cup \{E\} \cup \{B_n\}))$;

   that is, c_n is the chosen constant occurring neither in G_n nor in E nor in B_n. Then:

   $G_{n+1} = (G_n \cup \{B_n\} \cup \{\sim[c_n/v](A)\})$.

Omitted are considerations justifying this definition, most salient among which is the observation that the antecedents of the inductive cases (counting 3 as a separate case for each v and A) are exclusive and exhaustive: Exactly one of them holds. In particular, the antecedent of 3 holds (if at all) for at most one v and A, by One-oneness (last clause of **3B-10**), so that the threat of inconsistent directions for constructing G_{n+1} is thereby shown unreal.

We need the following facts about the sequence of G_n:

3E-18 LOCAL FACT. *(Facts about* G_n*)*

---

1. G_n is closed.
2. G_n is increasing: $m \leqslant n \rightarrow G_m \subseteq G_n$.
3. Sublemma. Economy is maintained in the sense that there are always infinitely many constants left with which to instantiate: nonoccur($G_n \cup \{E\} \cup B_n$) is infinite.
4. Sublemma. c_n, as defined in **3E-17**, occurs neither in E, nor in B_n, nor in any member of G_n.
5. Each G_n is E-free.

Exercise 63 *(Facts about* G_n*)*

1. Prove 1–5 of **3E-18**. See **2F-14** and **2F-15** for a little help. Most of these are established by Induction on N, **2C-4**. The ones marked "sublemma" are useful in proving later ones: We want 3 only for 4, and 4 only for 5. Furthermore, the restriction of Lindenbaum's lemma **3E-9** to

   economical sets has its sole point *precisely* here, in obtaining the base case for the Induction on N that underlies the proof of 3.
2. Explain to yourself why the "economy" hypothesis of Lindenbaum's lemma **3E-9** for S_Q is required in the following sense: Without it the lemma is false. That is, find a set G and a sentence E that are closed and are such that G $\nvdash S_Q$ E, but where there is no G* satisfying all parts of the conclusion of the lemma. Hint: Let G contain Ft for every closed term t, and let $E = \forall v F v$. And consider the fate of $\sim\forall v F v$, given that TL$\sim$, **3E-5**, is known to follow for G*.

▷ . . . . ◁

The development continues in analogy with §2F.5. Here insert an exact copy of Local definition **2F-16** and Local facts **2F-17** and **2F-18**; but now the reference is to *this* locality. To complete the proof, we need to show the following.

3E-19 LOCAL FACT. *(Properties of $\bigcup$Z)*

---

1. $\bigcup$Z is closed.
2. LL1 for $\bigcup$Z: G$\subseteq\bigcup$Z.
3. $MCEF_{S_Q}1$ for $\bigcup$Z: $\bigcup$Z is E-free in S_Q.
4. $MCEF_{S_Q}2$ for $\bigcup$Z: No closed proper superset of $\bigcup$Z is E-free in S_Q.
5. $\sim\forall$-completeness for $\bigcup$Z: If $\sim\forall$vA$\in\bigcup$Z, then for some closed Q-term t, $\sim$[t/v](A)$\in\bigcup$Z.

Exercise 64 *(Properties of $\bigcup$Z)*

. Prove Local fact **3E-19** Consult §2F.5 for guidance, especially the proofs of **2F-19** and **2F-20** and **2F-21**.

▷ . . . . ◁

Given **3E-19**, the proof of Lindenbaum's lemma **3E-9** for S_Q is complete.

Exercise 65 *(S_Q is consistent and complete)*

There is just one thing to do: Prove S_Q consistent and complete, taking the Reduction to economical sets lemma **3E-8** as given.

▷ . . . . ◁

3F Wholesale substitution and economical sets

Here, as in the order of heaven, the first is last: It is finally time to turn our attention to the proof that the problem of consequence for arbitrary closed G and A can be reduced to the case in which G is economical; see **3E-8** for an exact statement. The matter can be handled intuitively in a few picturesque sentences. Enumerate the constants without repetition: $c_0, \ldots$. Form H and A′ by substituting, say, c_{n+n} for c_n. This "clearly" omits infinitely many constants from H and A′, since none of the odd-numbered ones will occur therein; and because the substitution is one-one, **2C-8**, it amounts to nothing more than "re-lettering," which can obviously upset neither semantic nor proof-theoretic consequence. In fact, were $H \vDash_Q A'$ to fail in virtue of some interpretation $\mathbf{j}_2$, we could show that $G \vDash_Q A$ fails as well by simply changing $\mathbf{j}_2$ to $\mathbf{j}$ by letting $Val_{\mathbf{j}}(c_n) = \mathbf{j}(c_{n+n})$; and vice versa. And any proof from G to A can "clearly" be transformed into a proof from H to A′ by systematic substitution of c_{n+n} for c_n throughout the entire proof; and vice versa. Q.E.D.

The most standard solution to the problem of uneconomical sets of premisses is even more swift than that of the preceding paragraph: It suggests that if we don't have enough constants (for purposes of instantiation), then we should add some more. The idea is that if the demand for constants goes up, then by all means increase the supply. This intuition is absolutely sound: It is "clear" that the questions as to whether A follows from G, proof-theoretically and semantically, receive the same answers if one adds some new constants to the language, thus rendering G economical after the addition of those constants. It is not so clear just what conceptual apparatus is required to establish this fact in a rigorous fashion; in particular, the cost of relativizing our many semantic and proof-theoretic concepts to languages with different stocks of constants needs to be measured. These notes embody the belief that when one is compelled to be truth-function-and-quantifier rigorous, it is easier to deal with one language than with many; but if rigor at the level of truth functions and quantifiers doesn't matter, as if often doesn't, then it doesn't matter.

If you are satisfied with either of these accounts, you should skip to the very end of this section, Theorem **3F-21**. Otherwise there is quite a lot to go through.

We shall proceed by introducing a concept of wholesale substitution, and in three successive sections studying its grammar, semantics, and proof theory.

3F.1 Grammar of wholesale substitution

"Term-assignment," the first concept we define to treat these matters, is a grammatical analog to the semantic notion of an "interpretation": Closed terms (instead of entities in the domain) are assigned wholesale to the constants (no assignments are made to variables). In fact the analogy to which we here refer is deep, but we let it go by. It is to be noted that the concept is introduced with a particular purpose in mind, and that it is only in virtue of that purpose that (1) term-assignments are made to constants only and (2) only closed terms are assigned. These limitations make our work simpler.

3F-1 DEFINITION. *(Term-assignment)*

Term-assignment = (constant $\mapsto$ closed term).

That is, an assignment of terms (this is a variant) is a function that assigns to each constant some (perhaps complex) closed term, so that if r is an assignment of terms and c is a constant, r(c) is a closed term.

3F-2 CONVENTION. *("r" for term-assignments)*

We reserve "r" for term-assignments.

Just as we sometimes want to shift interpretations, as in **2D-19** and **3C-4**, so we sometimes want to shift term-assignments:

3F-3 DEFINITION. *(Term-assignment shift for* Q*)*

Where r is a term-assignment, t is a closed term, and c is a constant,

[t/c](r) is a term-assignment ([]r-type)

([t/c](r))c = t ([]r1)

$([t/c](r))(c') = r(c')$ if $c' \neq c$ ([]r2)

That is, [t/c](r) is that assignment of terms that is just like r except for giving c the value t.

We are going to define the grammatical concept, "the substitution determined by r," in analogy to the semantic concept, "the valuation determined by **j**," **3C-7**; we will rely on r for a base clause and give the inductive clauses in a natural way.

3F-4 DEFINITION. *(S_r: wholesale substitution for* Q*)*

For any term-assignment r, S_r is a function whose domain is the formulas, and such that:

$S_r(t)$ is a term ($S_r$type)

$S_r(A)$ is a sentence ($S_r$type)

$S_r(c) = r(c)$ ($S_r$const)

$S_r(v) = v$ ($S_r$var)

$S_r(\Phi O_1 \ldots O_n) = \Phi(S_r(O_1), \ldots, S_r(O_n))$ ($S_r\Phi$)

$S_r(\forall vA) = \forall v S_r(A)$ ($S_r\forall$)

Also, $A \in S_r(G) \leftrightarrow$ for some B in G, $A = S_r(B)$. (S_rG)

Here the ($S_r$const) and ($S_r$var) are **Basis clauses**, while ($S_r\Phi$) and ($S_r\forall$) are **Inductive.** The picture to have is straightforward: If r assigns closed terms to the constants, then $S_r(O)$ is just those same substitutions in O. $S_r(O)$, then, is the result of substituting according to the directions of r for all constants in O. Read "$S_r(O)$" as something clumsy like "the result of making the substitutions r in O." (It should be noted that "wholesale substitution" is not a true generalization of "retail substitution," since the notation of **3B-13** for the retail concept, [t/a](O), permits substitution (1) for variables as well as for constants, and (2) of open as well as of closed terms.)

Because a wholesale substitution S_r is so much like a valuation Val_j, it is not surprising that analogs of the local determination and semantics of substitution theorems, **3C-12** and **3C-14**, are forthcoming.

3F-5 FACT. *(Local determination for* S_r*)*

If $r(c) = r'(c)$ for every c occurring in O, then $S_r(O) = S_{r'}(O)$.

PROOF. Straightforward, by Induction on formulas, **3B-8**, using the definitions of the ingredient notations: **3B-15** and **3F-4**. □

3F-6 COROLLARY. *(Corollary to local determination for* S_r*)*

If c does not occur in O, then $S_{[t/c](r)}(O) = S_r(O)$.

The analog to the semantics of substitution theorem relates wholesale and retail substitution, and also term-assignment shift. There are separate clauses for variables and constants because r and accordingly S_r do not touch variables.

3F-7 LEMMA. *(Wholesale and retail substitution)*

Assume t is closed.

$$S_r([t/v](O)) = [S_r(t)/v]((S_r(O))).$$

$$S_r([t/c](O)) = S_{[S.r(t)/c](r)}(O).$$

Note that the left side of each equation represents a retail substitution followed by a wholesale substitution.

PROOF. Tedious. □

By the time we finish with substitution, or vice versa, we will all be clear on why Curry invented combinatory logic in order to do without it. In the meantime, we need to call attention to the family of term-assignments that "rearrange" constants: They map constants into other constants (never into complex terms), and they always map distinct constants into distinct constants. Such term-assignments are intuitively important because one can guess that any such "rearrangement" or "permutation" will preserve not only semantic consequence and proof-theoretical consequence, but also their negations; for we can all see that which constants are employed is a matter of no consequence.

3F-8 DEFINITION. *(Constant-permuting term-assignment)*

---

r is a *constant-permuting term-assignment* provided it is a term-assignment that assigns only constants to constants, and is one-one:

r is a term-assignment (cptype)

r(c) is a constant (cpconst)

$r(c) = r(c') \rightarrow c = c'$ (cponeone)

Exercise 66 *(Three term-assignments)*

Draw pictures of three term-assignments, one that is constant-permuting and two that are not because violating exactly one of cpconst or cponeone of **3F-8**.

▷ . . . . ◁

Sometimes we want to consider what happens when we reverse the effect of a constant-permuting term-assignment; in such circumstances, we will use "r*" for the function that undoes r—it's "left inverse":

3F-9 DEFINITION. *(r*)*

For r a constant-permuting term-assignment, r* is the "left inverse" of r, namely, that term-assignment (not necessarily constant-permuting) such that

r* is a term-assignment (r*type)

r*(r(c)) = c, all constants c (r**Rng*(r))

r*(c) = c, all c not in *Rng*(r), **9B-4** (r*Other)

Exercise 67 *(r* and r)*

Draw a picture illustrating an r and its left inverse r* such that r is constant-permuting but r* is not. Explain why this shows that r* is a *left* inverse of r (i.e., r*(r(c)) = c) without being also a *right* inverse of r (i.e., r(r*(c)) = c). You will need to picture infinitely many constants.

▷ . . . . ◁

We will later need the following, which simply makes good on the claim that r* undoes r.

3F-10 FACT. *(S_{r*} undoes S_r)*

If r is a constant-permuting term-assignment, **3F-8**, and r* is its left inverse according to **3F-9**, then

$S_{r*}(S_r(O)) = O$.

PROOF. Straightforward, by Induction on formulas, **3B-8**, choosing Ψ(O) there as the equation here. First the basis, which has two cases. $S_{r*}(S_r(c))$ equals $S_{r*}(r(c))$ by $S_r$const, **3F-4**, and hence equals r*(r(c)) by $S_r$const again, since r(c) is a constant because r is by hypothesis constant-permuting, so that we may appeal to cpconst, **3F-8**; and so it equals c by the definition of "r*," **3F-9**. The second base case requires us to verify that $S_{r*}(S_r(v)) = v$; consult $S_r$var of **3F-4**. Inductive step: The argument for elementary functors Φ goes through by $S_r\Phi$ twice, and the argument for ∀ is as easy, using $S_r\forall$. □

We need the fact that there is a way of permuting *all* the constants into just *some* of them, which is of course at the heart of the reduction to sets economical as to constants.

3F-11 FACT. *(All constants into some)*

There is a constant-permuting term-assignment r whose range omits infinitely many constants: (constant − *Rng*(r)) is infinite.

PROOF. Because the constants are countable and infinite, **3B-1**, this is an immediate corollary of Fact **3E-15**. □

3F.2 Semantics of wholesale substitution

We must be sure that the condition of Fact **3F-11** suffices for the reduction to economical sets. The theorem we want to obtain is that if G and A on the one hand are related to H and A′ on the other by a *constant-permuting* term-assignment, then $G \vDash_Q A \leftrightarrow H \vDash_Q A'$. Since Fact **3F-11** promises at least one term-assignment whose range omits infinitely many constants, this will suffice.

The basic fact is a wholesale version of the semantics of substitution theorem, **3C-14**; to state it, we need to define a wholesale shift in **j**:

3F-12 DEFINITION. *([r]**j**)*

---

[r]**j** is that Q-interpretation that agrees with **j** on other than constants, but gives each constant c the value that **j** gives r(c):

$[r]\mathbf{j}$ is a Q-interpretation. ($[r]\mathbf{j}$-type)

$([r]\mathbf{j})c = \mathbf{j}(r(c))$. ($[r]\mathbf{j}$-const)

$([r]\mathbf{j})x = \mathbf{j}(x)$ for $x \notin$ constant. ($[r]\mathbf{j}$-other)

The following is then the right analog of **3C-14**:

3F-13 LEMMA. *(Semantics of wholesale substitution lemma for* Q*)*

$Val_{\mathbf{j}}(S_r(O)) = Q_{[r]\mathbf{j}}(O)$.

PROOF. Tedious, and too much like that of **3C-14** to bear. Use Definition **3F-4** of "S_r." □

The following, which is half of what we need, is now easy:

3F-14 FACT. *(Preservation of* Q*-consequence under wholesale substitution)*

$G \vDash_Q A \rightarrow S_r(G) \vDash_Q S_r(A)$.

The same is of course true of retail substitution.

PROOF. For contraposition, let $\mathbf{j}$ witness the falsehood of the consequent according to **3C-18**: $Val_{\mathbf{j}}(S_r(B)) = T$ for all $B \in G$, and $Val_{\mathbf{j}}(S_r(A)) = F$. By the Semantics of wholesale substitution lemma, **3F-13**, $Q_{[r]\mathbf{j}}(B) = T$ for all $B \in G$ and $Q_{[r]\mathbf{j}}(A) = F$; so the antecedent is false as well by **3C-18**. □

The converse is not in general true, so this fact is not by itself enough. But the converse *does* hold if r is constant-permuting, **3F-8**, so that in this special case we can promote the implication to an equivalence:

3F-15 FACT. *(Invariance of* Q*-consequence under permutation of constants)*

Suppose that r is a constant-permuting term-assignment, **3F-8**; then

$$G \vDash_Q A \leftrightarrow S_r(G) \vDash_Q S_r(A).$$

PROOF. We already have from left to right; for the other way, suppose

$$S_r(G) \vDash_Q S_r(A).$$

Then consider the left inverse r* of r as given by Definition **3F-9**. By the previous Fact **3F-14**,

$$S_{r^*}(S_r(G)) \vDash_Q S_{r^*}(S_r(A)).$$

But then by Fact **3F-10**, which says that S_{r^*} undoes S_r, $G \vDash_Q A$ as required. □

3F.3 Proof theory of wholesale substitution

As outlined above in §3F.2, the easiest way to solve the problems raised by premiss-sets that are not economical as to constants is to establish that when given such a wasteful set and some sentence A, there is another set and another sentence such that (a) the other set *is* economical as to constants, and (b) both the question of semantic consequence, $\vDash_Q$, and the question of proof-theoretic consequence, $\vdash_{S_Q}$, yield the same answers in the two cases. Now we take up the proof-theoretic part.

The chief theorem is that S_Q-consequence is closed under wholesale substitution; this will have what we want as a corollary. For the theorem, we require two lemmas.

3F-16 LEMMA. *(S_r preserves being closed)*

If O is closed, so is $S_r(O)$.

PROOF. Straightforward, by Induction on formulas, **3B-8**, choosing Ψ(O) there as the entire lemma here. Or appeal to Exercise 50, if you have carried it out, in this case choosing just the consequent of the lemma for Ψ(O). □

3F-17 LEMMA. *(Closure of AxS_Q under wholesale substitution)*

If $A \in AxS_Q$, **3D-1**, then for all term-assignments r, $S_r(A) \in AxS_Q$.

PROOF. Recall that AxS_Q is defined inductively; so we shall proceed by Induction on AxS_Q, **3D-1**, choosing $\Psi(A)$ there as "$(r)(S_r(A) \in AxS_Q)$" here.

The argument for A an axiom having one of the forms S_Q1, S_Q2, S_Q3, S_Q5, or S_Q6 requires only a series of applications of $S_r\Phi$ and $S_r\forall$, **3F-4**, to show that $S_r(A)$ is an axiom having the same form as A. When A has the form S_Q4, show that $S_r(A)$ does as well by using Lemma **3F-7**.

Suppose now for the inductive step that (1) c does not occur in A and that (2) $(r)(S_r([c/v](A)) \in AxS_Q)$. We need to show that $S_r(\forall vA) \in AxS_Q$. Choose (3) c' not in $S_r(A)$ by **3B-19**, and choose r of (2) as $[c'/c](r)$, so that $S_{[c'/c](r)}([c/v](A)) \in AxS_Q$. Hence

$$(4)\ [(S_{[c'/c](r)}(A))/v](S_{[c'/c](r)}(c)) \in AxS_Q.$$

by Lemma **3F-7**.

Now $S_{[c'/c](r)}(A) = S_r(A)$ by (1) and Fact **3F-6**; and $S_{[c'/c](r)}(c) = c'$ by $S_r$const of **3F-4** and []r1 of **3F-3**; so (4) yields

$$[c'/v]((S_r(A))) \in AxS_Q,$$

whence $\forall v(S_r(A)) \in AxS_Q$ by (3) and the Inductive clause of **3D-1**. So, finally, $S_r(\forall vA) \in AxS_Q$ by $S_r\forall$ of **3F-4**. □

3F-18 THEOREM. *(Preservation of* S_Q*-consequence under wholesale substitution)*

Where r is a term-assignment, **3F-1**, and S_r is the substitution it induces, **3F-4**,

$$G \vdash_{S_Q} A \rightarrow S_r(G) \vdash_{S_Q} S_r(A).$$

PROOF. By Induction on S_Q-consequence, **3D-2**, choosing $\Psi(A)$ there as "$S_r(G) \vdash_{S_Q} S_r(A)$" here. If $A \in G$, then $S_r(A) \in S_r(G)$ by S_rG of **3F-4**, so $\vdash_{S_Q}$id, **3D-2**, gives us $S_r(G) \vdash_{S_Q} S_r(A)$. If $A \in AxS_Q$, Lemma **3F-17**, just proved, guarantees that $S_r(A) \in AxS_Q$, so $S_r(G) \vdash_{S_Q} S_r(A)$ by $\vdash_{S_Q}AxS_Q$, **3D-2**. Suppose inductively that $S_r(G) \vdash_{S_Q} S_r(A)$ and that $S_r(G) \vdash_{S_Q} S_r(A \supset B)$; then $S_r(G) \vdash_{S_Q} (S_r(A) \supset S_r(B))$ by $S_r\Phi$, so $S_r(G) \vdash_{S_Q} S_r(B)$ by $\vdash_{S_Q}$MP, **3D-2**. □

This theorem is indeed important in itself; essential, however, for subsequent use is the following

3F-19 COROLLARY. *(Invariance of* S_Q*-consequence under permutation of constants)*

If r is a constant-permuting term-assignment, **3F-8**, then

$$G \vdash_{S_Q} A \leftrightarrow S_r(G) \vdash_{S_Q} S_r(A).$$

PROOF. Like the proof of **3F-15**. □

3F.4 Reducing consequence questions to economical sets

used.

3F-20 FACT. *(Rng(r) and nonoccurrence in* S_r*)*

Let r be a constant-permuting term-assignment, **3F-8**. If c is not a member of *Rng*(r), NAL:**9B-12**, then c does not occur in S_r(O)—hence not in S_r(G).

PROOF. Take a detour through the contrapositive: Use Induction on formulas, **3B-8**, choosing Ψ(O) there as "c occurs in S_r(O) → c ∈ *Rng*(r))" here.

Basis step. Suppose c occurs in S_r(a). Since $S_r(v) = $v, a cannot be a variable, so must be a constant. So S_r(a) = r(a) by $S_r$const, **3F-4**; which therefore is a constant by **3F-8**. Since the only atom in which c occurs is c itself (occuratom, **3B-15**), c = r(a), and is hence in the range of r by the definition of "*Rng*," NAL:**9B-12**.

Inductive step. The clauses S_rΦ and occurΦ, and S_r∀ and occur∀ (**3F-4**, **3B-15**) match to enable the inductive step. □

We may now turn to a proof of the Reduction to economical sets lemma, **3E-8**.

PROOF. By **3F-11** there is a constant-permuting term-assignment r whose range omits infinitely many constants. Given G and A, choose the H and A′ required by the lemma as S_r(G) and S_r(A). By the invariance under permutation of constants of Q-consequence, **3F-15**, and of S_Q-consequence, **3F-19**, G $\vDash_Q$ A ↔ H $\vDash_Q$ A′, and G

$\vdash_{S_Q}$ A $\leftrightarrow$ H $\vdash_{S_Q}$ A′. It only remains to observe that by **3F-16** H and A′ are closed, and that by **3F-20**, nonoccur(H∪{A′}) is a superset of (constant − *Rng*(r)). So since the latter is promised as infinite, the former must be as well, by contraposition of **2B-16**. □

Since the Reduction to economical sets lemma **3E-8**, just proven, was the only missing piece in our argument for the completeness of S_Q, we are finally ready for the wrap-up:

3F-21 THEOREM. *(Consistency/completeness of* S_Q*)*

G $\vDash_Q$ A $\leftrightarrow$ G $\vdash_{S_Q}$ A.

PROOF. Trivial, given **3E-1** and **3E-7** through **3E-12**. □

Whew!

Exercise 68 *(Things to do)*

Explain the Löwenheim-Skolem theorem, the Craig interpolation theorem, Beth's definability theorem, and second order logic.

▷ . . . . ◁

3G Identity

We discuss the grammar, semantics, proof theory, and consistency and completeness of the first order functional calculus with identity, here called "Q=" for grammar and semantics, and "$S_{Q=}$" for proof theory.

The picture to have is this. The islanders speak a language. In Chapter 2, we theorized about their language insofar as what we wanted to say depended only on the structure conferred on that language by the truth-functional connectives ⊃ and ∼. In the first six sections of Chapter 3, we enriched our theory (while their language stayed just the same) by taking into account the additional structure due to the presence of predicates, operators, and quantifiers. Now, in §3G of Chapter 3, we enrich our theory still further (while their language continues unaltered) by taking into account the special properties of a certain one of their predicates, the identity predicate.

3G.1 Grammar of Q=

We shall use "=" to name the identity predicate of Q=, while continuing to use "_ = _" (as always) as the basis of the identity predicate of our use-language.[23]

With this understanding, we can just declare the entire grammar of Q= to be identical to the grammar of Q as defined in §3B, with the addition of a single

3G-1 AXIOM. *(Grammar of =)*

= is a 2-predicate (a two-place predicate), **3B-1** on p. 134.

Since (1) it may be hard to distinguish "=" and "=" and (2) we don't want to introduce some unfamiliar symbol in either role, to heighten clarity we adopt the following

3G-2 CONVENTION. *(Square brackets for Q= identities)*

We provide with square brackets all use-terms denoting identity sentences of Q=:

[t=u] = =(t, u).

That is, [t=u] is the result of applying the 2-predicate, =, to the two arguments t and u.[24]

Clearly this axiom alters in no way any results we have previously established; nor does it lessen their interest, for it is nearly conservative with respect to our theory of the grammar of Q, adding in the language of that theory only the statement that there is at least one 2-place predicate. In particular, **3B-1** made no assumptions inconsistent with having = as a predicate. We may therefore refer to any previous grammatical result with confidence.

[23]That is, the identity predicate of our use-language is that function from pairs of use-language terms into use-language sentences such that given any pair of the former as arguments, the following sentence is value: the first term followed by a double-bar symbol followed by the second term. More or less: This recipe ignores parentheses, just as we do.

[24]Thus, (1) "[t=u]" *is* a term of our use-language; (2) "[t=u]" *names* a sentence of Q=; (3) [t=u] *is* a sentence of Q=; (4) [t=u] is not a name of anything; (5) "(t = u)" is a sentence of our use-language; (6) "(t = u)" is not a name of anything; and (7) "(t = u) is (or isn't, or names or doesn't name) a sentence (of any language)" is ungrammatical, having a use-sentence where a use-term is wanted. This discussion, like all use-mention discussions, is confusing; but our practice is not.

3G.2 Semantics of Q=

The key definition of "interpretation" is nearly the same as **3C-2**, except that the predicate = is now treated as a "logical" constant, like the connectives ⊃ and ∼, so that every interpretation gives the predicate, =, the same "meaning"—to the extent that its meaning is independent of choice of domain.

3G-3 DEFINITION. *(Q=-interpretation)*

---

j is a Q=-interpretation ↔ **j** is a Q-interpretation, **3C-2**, satisfying one further condition: For all d, d′ ∈ D,

$$(Val_{\mathbf{j}}(=))(\mathrm{d}, \mathrm{d}') = \mathrm{T} \leftrightarrow \mathrm{d} = \mathrm{d}'.$$

Obviously

3G-4 FACT. *($Val_{\mathbf{j}}$ for identities)*

$Val_{\mathbf{j}}([\mathrm{t{=}u}]) = \mathrm{T} \leftrightarrow (Val_{\mathbf{j}}(\mathrm{t}) = Val_{\mathbf{j}}(\mathrm{u}))$.

If one understands that all reference to interpretations is to be taken as reference to Q=-interpretations instead of to Q-interpretations, then we may take over wholesale the semantic development of §3C. In particular, the definition of "$\vDash_{Q=}$" is just like that of "$\vDash_{Q}$" except for reference to Q=-interpretations. One item that needs (easy) checking is that the result [d/a](**j**) of an interpretation shift, **3C-4**, is always a Q= interpretation if **j** is.

3G.3 Proof theory of $S_{Q=}$

There are two new primitive axioms, ready to be thrown together for S_Q7 (**3D-1**) to generalize upon in order to yield the set $AxS_{Q=}$ of axioms of $S_{Q=}$. Everything else is then the same, so that we may take $\vdash_{S_{Q=}}$ as defined. The axioms:

[t=t], all closed terms t. $(S_{Q=}1)$

[t=u] ⊃ (([t/v](A) ⊃ [u/v](A)), all closed t, u; all v. $(S_{Q=}2)$

Compare $S_{Q=}2$ with the Replacement theorem for S_{TF}, **2E-21**, and the surrounding discussion. It tends to be at first confusing to use $S_{Q=}2$ in applications, since although its description in our use-language mentions the variable v, its instances do not contain (free) v—the variable disappears via substitution.

$S_{Q=}2$ is unexpectedly powerful; in particular, it leads easily to the following

3G-5 FACT. *(Symmetry, transitivity, and replacement for* $S_{Q=}$*)*

Where t and u are closed and v is any variable,

G, [t=u] $\vdash_{S_{Q=}}$ [u=t]. ($\vdash_{S_{Q=}}$Symm)

G, [t_1=t_2], [u=t_1] $\vdash_{S_{Q=}}$ [u=t_2]. ($\vdash_{S_{Q=}}$Trans)

G, [t_1=t_2] $\vdash_{S_{Q=}}$ [[t_1/v](u) = [t_2/v](u)]. ($\vdash_{S_{Q=}}$Replace)

Exercise 69 *(Symmetry, transitivity, and replacement for* $S_{Q=}$*)*

Prove **3G-5**. Since substitution notation is used in the statement of $S_{Q=}2$, you will probably need to rely on its definition, **3B-13**. For $\vdash_{S_{Q=}}$Symm, choose A in $S_{Q=}2$ as [v=t]. For $\vdash_{S_{Q=}}$Trans, choose A in $S_{Q=}2$ as [t_1=v]. You may find that choosing A in $S_{Q=}2$ for the proof of Replacement is little delicate.

▷ . . . . ◁

3G.4 Consistency of $S_{Q=}$

3G-6 THEOREM. *(Semantic consistency of* $S_{Q=}$*)*

$S_{Q=}$ is semantically consistent: G $\vdash_{S_{Q=}}$ A → G $\vDash_{Q=}$ A.

PROOF. All the other axioms of $S_{Q=}$ (we mean those that are also axioms of S_Q) are all right because (a) every Q=-interpretation is still a Q-interpretation, and (b) the Q=-interpretations are closed under interpretation shift, **3C-4**, needed in stating the semantic values for universally quantified statements, **3C-7**. While these facts do not render the semantic consistency of $S_{Q=}$ an in-your-head corollary, we nevertheless recommend taking as sufficient unto the day our previous verification that those other axioms are valid for all Q-interpretations.

The two new axioms for $S_{Q=}$ need special checking; use the Semantic substitution theorem, **3C-14**, for $S_{Q=}$—as you should expect, since the axiom uses the concept of substitution in its statement. One has to check that S_Q7 preserves validity in its new meaning; and that modus ponens does as well. □

3G.5 Completeness of $S_{Q=}$

We try to make clear what is *new* when identity is present.

There is no change in our concentration on closed sets and sentences. There is no change in the way the reduction to economical premiss sets is handled.

There is no change in Lindenbaum's lemma, except that it now refers to $S_{Q=}$ instead of S_Q.

The definition of "truth-like set" is adapted by adding a clause TL=, which has two parts:

For all closed terms t, [t=t] ∈ G*. (TL=1)

For all closed terms t and u, and for all A with at most v free: [t=u] ∈ G* → ([t/v](A) ∈ G* ↔ [u/v](A) ∈ G*). (TL=2)

These correspond in an obvious way to the two axioms $S_{Q=}1$ and $S_{Q=}2$, so that it is easy to see that if G* is maximal closed and E-free with respect to $S_{Q=}$, then TL=1 and TL=2 will hold for G*; so that the maximality/truth-like-set lemma for $S_{Q=}$ will go through. (We note that one could combine the two parts into a single statement: [t=u] ∈ G* ↔ for all sentences A with at most the variable v free, [t/v](A) ∈ G* ↔ [u/v](A) ∈ G*. This matches Leibniz's definition of truth for identity statements, but with the important difference that quantification over properties is replaced by quantification over open sentences. And of course we are not talking about truth, but membership in G*.)

The definition of a canonical interpretation determined by G* needs to be modified; the problem to be overcome is this. When we come to the truth-like set/canonical interpretation lemma, we must have $Val_{\mathbf{j}}(A) = T$ ↔ A ∈ G* for *all* closed A; so that as a special case we shall need a clause as follows:

$$Val_{\mathbf{j}}([t{=}u]) = T \leftrightarrow [t{=}u] \in G^*.$$

But then putting this requirement together with Fact **3G-4** implies that it is essential that

Val_{j}(t)$=$$Val_{j}$(u) $\leftrightarrow$ [t=u]$\in$G*;

so that this had better be built into our definition of a canonical interpretation determined by G*. In particular, it will no longer do to have each term denote itself (Val_{j}(t)$=$t, as in the Autonymy sublemma, **3E-11**), since quite clearly we shall have cases of [t=u]$\in$G* where t and u are distinct terms (t$\neq$u). Instead, in contrast to the situation without identity, we shall need to make what each term denotes, and not only the truth value of sentences, depend on G* itself, so as to make the above equivalence come out the way we need it. That is, we need terms to denote the same entity just in case the identity that says that they do is a member of G*.

There are two standard procedures here. As background, we introduce the important ideas of an "equivalence relation" and of an "equivalence class." By an *equivalence relation* over a set, X, we mean a relation that is reflexive, symmetrical, and transitive on that set. Given an equivalence relation over X, the *equivalence class*[25] *of* x (for x$\in$X) is the set of all y in X to which x is equivalent (by the given equivalence relation). You should see that if G* is closed under Q=-consequence (as it will be), then "[t=u]$\in$G*" describes an equivalence relation on the set of terms. With these concepts, we can say that given any one equivalence class, we want all of its members to denote the same thing.

The two standard procedures for securing this kind of result are (1) to invoke the equivalence class itself, and (2) to invoke some representative member of the equivalence class. According to procedure (1), we would let each term in an equivalence class denote that equivalence class itself; evidently on this plan two terms will denote the same thing just in case they fall in the same equivalence class. According to procedure (2), which is an equally happy choice, we choose from each equivalence class a "representative" of that equivalence class, and let everything in the equivalence class denote that representative. Since, however, there is nothing special to distinguish any one member of an equivalence class from another, we will have to do a little axiom-of-choice type work in order to proceed. One realization of the method of representatives begins by enumerating the closed terms. For t closed, define t* as the first term u in that enumeration such that [t=u]$\in$G*. Then, in effect, fix Val_{j} so that for each closed term t, Val_{j}(t)$=$t*; and show that everything comes out right. Such is the plan; and it works. Our very slightly different realization of the method of representatives is to rely on the countability of the terms in a somewhat more abstract fashion, as we did in picking a Chosenconstant

[25] No one ever says "equivalence set," but always "equivalence class." The reason lies in sound, not sense.

(**3E-14**) in our proof of Lindenbaum's lemma for S_Q. Here are a few of the details, utilizing as much as possible previous work.

3G-7 LOCAL DEFINITION. *(*Equivalence-class(t)*)*

For each term t, Equivalence-class(t) is that subset of terms such that

$$u \in \text{Equivalence-class(t)} \leftrightarrow [t{=}u] \in G^*.$$

Next we need a choice, justified by Countable choice, **3E-13**, and the fact that the formulas are countable, **3B-11**:

3G-8 LOCAL CHOICE. *(*Chosen-formula*)*

Chosen-formula $\in (\mathcal{P}(\text{formula}) \mapsto \text{formula})$; and for all nonempty subsets X of formula, Chosen-formula(X) $\in$ X.

Observe that in applications we must be sure that X is nonempty. Now we can define "t*":

3G-9 LOCAL DEFINITION. *(*t**)*

$t^* = \text{Chosen-formula(Equivalence-class(t))}$.

The important properties of the star operation are given by the following

3G-10 FACT. *(Properties of* t**)*

Suppose TL=1 and TL=2 for G*. Then for any closed terms t and u, the following hold.

$[t{=}t^*], [t^*{=}t] \in G^*$.

$[t^*{=}u^*] \in G^* \leftrightarrow (t^* = u^*)$.

$[t^*/v](A) \in G^* \leftrightarrow [t/v](A) \in G^*$, where A is any sentence with at most v free.

$(ft_1{}^* \ldots t_n{}^*)^* = (ft_1 \ldots t_n)^*$.

PROOF. Tedious. □

Canonical Q=-interpretations can now be defined as follows (the definition is "local," **1A-2**, because it depends on Local definition **3G-9**.)

3G-11 LOCAL DEFINITION. *(Canonical* Q=*-interpretation)*

---

j is a *canonical* Q=*-interpretation determined by* G* if and only if

> **j** is a Q=-interpretation, **3G-3**.
>
> **j**(∀) is the set of star-terms: $x \in \mathbf{j}(\forall) \leftrightarrow$ for some closed term t, $x = t^*$.
>
> $\mathbf{j}(c) = c^*$.
>
> $(\mathbf{j}(f))(t_1^*, ,\ldots, t_n^*) = (ft_1^* \ldots t_n^*)^*$, all $t_k^* \in D$, $1 \leqslant k \leqslant n$.
>
> $(\mathbf{j}(F))(t_1^*, \ldots, t_n^*) = T \leftrightarrow Ft_1^* \ldots t_n^* \in G^*$, all F except =, all $t_k^* \in D$, $1 \leqslant k \leqslant n$.

The verification of an appropriate truth-like set/canonical interpretation lemma for Q= is now a routine matter of Induction on closed terms, and sentences, **3B-23**.

Which is enough to complete completeness for $S_{Q=}$:

3G-12 THEOREM. *($S_{Q=}$ is complete)*

For closed G and A,

$$G \vDash_{Q=} A \rightarrow G \vdash_{S_{Q=}} A.$$

Exercise 70 *(Consistency/completeness of* $S_{Q=}$*)*

1. Fill in the details of consistency for $S_{Q=}$.
2. Fill in the details of completeness for $S_{Q=}$; that is, prove **3G-12**.

▷ . . . . ◁

Bibliography

Bocheňski, J. 1956, *Formale Logik*, K. Alber, Freiburg. Translated and edited by I. Thomas (*A history of formal logic*), second edition, New York (Chelsea Publishing Company), 1970.

Boolos, G. S., Burgess, J. P. and Jeffrey, R. C. 2007, *Computability and logic*, 5th edn, Cambridge University Press, Cambridge.

Carnap, R. 1937, *The logical syntax of language*, International library of psychology, philosophy and scientific method, Harcourt, Brace, New York. Translation of *Logische Syntax der Sprache* by A. S. Smeaton (Countess von Zeppelin).

Church, A. 1936, "A note on the Entscheidungsproblem", *The journal of symbolic logic* **1**, pp. 40–41, 101–102.

Church, A. 1956, *Introduction to mathematical logic*, Princeton University Press, Princeton, NJ.

Curry, H. 1963, *Foundations of mathematical logic*, McGraw-Hill, Inc., New York.

Curry, H. and Feys, R. 1958, *Combinatory logic*, Vol. I, North-Holland Publishing Company, Amsterdam.

Fitch, F. B. 1952, *Symbolic logic*, The Ronald Press, New York. Source of the method of subproofs as we use it. Fitch was my teacher.

Frege, G. 1964, *The basic laws of arithmetic; exposition of the system*, University of California Press, Berkeley. Translated and edited, with an introduction, by M. Furth. The translation is of "the introductory portions of the first volume and an epilogue appended to the second.".

Gentzen, G. 1934, “Untersuchungen Über das Logische Schliessen”, *Mathematische Zeitschrift* **39**, 176–210, 405–431. English translation (Investigations into logical deduction) by M. E. Szabo in *The collected papers of Gerhard Gentzen*, ed. M. E. Szabo, Amsterdam (North-Holland Publishing Co.), 1969, pp. 68–131.

Gentzen, G. 1936, “Die Widerspruchsfreiheit der reinen Zahlentheorie”, *Mathematische Annalen* **112**, 493–565. English translation (The consistency of elementary number theory) by M. E. Szabo in *The collected papers of Gerhard Gentzen*, ed. M. E. Szabo, Amsterdam (North-Holland Publishing Co.), 1969, pp. 132–213.

Gödel, K. 1930, “Die Vollständigkeit der Axiome des logischen Funktionkalküls”, *Monatschefte für Mathematik und Physik* .

Gödel, K. 1931, “Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I”, *Monatschefte für Mathematik und Physik* **38**, 173–198. English translation (On formally undecidable propositions of *Principia Mathematica* and related systems I) by B. Meltzer, with an introduction by R. B. Braithwaite, (Oliver and Boyd) 1962; and J. van Heijenoort in van Heijenoort 1967, pp. 596–616.

Hunter, G. 1971, *Metalogic: an introduction to the metatheory of standard first order logic*, University of California Press, Berkeley.

Jaśkowski, S. 1934, “On the rules of supposition in formal logic”, *Studia logica* **1**, 5–32. Reprinted in *Polish logic, 1920–1939*, ed. S. McCall, Oxford (Clarendon Press), 1967, pp. 232–258.

Kneale, W. and Kneale, M. 1962, *The development of logic*, The Clarendon Press, Oxford. Reprinted from corrected sheets of the first edition, 1966.

Kneebone, G. T. 1963, *Mathematical logic and the foundations of mathematics*, D. van Nostrand Publishing Company, Princeton, NJ, New York, Toronto, London. A good survey.

Leśniewski, S. 1981, *Collected works*, Vol. 118 of *Synthese library*, D. Reidel Published Company, Boston. ed. S. J. Surma, J. T. J. Srzednicki, and D. I. Barnett.

Myhill, J. 1952, “Some philosophical implications of mathematical logic”, *The review of metaphysics* **6**, 165–198.

Post, E. L. 1921, "Introduction to a general theory of elementary propositions", *American Journal of Mathematics* **43**, 163–185. Reprinted in van Heijenoort 1967, pp. 264–283.

Robinson, R. 1950, *Definition*, The Clarendon Press, Oxford.

Runes, D. D. 1962, *Dictionary of philosophy*, Littlefield, Adams, Totowa, NJ.

Russell, B. 1902, "Letter to Frege", *in* van Heijenoort 1967, pp. 124–125.

Shapiro, S. 1992, *Foundations without foundationalism: a case for second order logic*, Oxford University Press, Oxford.

Suppes, P. 1957, *Introduction to logic*, D. van Nostrand Publishing Company, Princeton, NJ, New York, Toronto, London.

Tarski, A. 1941, *Introduction to logic and to the methodology of deductive sciences*, enlarged and revised edition edn, Oxford University Press, New York. Translated by Olaf Helmer from a work that appeared in Polish in 1936 and then in German translation in 1937.

Tarski, A. 1956, *Logic, semantics, metamathematics*, The Clarendon Press, London and Oxford. A collection, including the famous paper on the definition of truth, a paper every logician needs to read.

Thomason, R. 1970, *Symbolic logic, an introduction*, The Macmillan Company, New York. A fine advanced undergraduate text.

van Heijenoort, J., ed. 1967, *From Frege to Gödel: A source book in mathematical logic, 1879–1931*, Harvard University Press, Cambridge, MA.