I. The Study of Logical Systems

In this course, we'll be studying a number of *logical systems*, also known as *logical theories* or *deductive systems*. Minimally, a logical system consists of four things:

1. A vocabulary of primitive signs used in the language of that system.
2. A list or set of rules governing what strings of signs (called "formulas") are grammatically or syntactically well-formed in the language of that system.
3. A list of axioms, or a subset of the well-formed formulas, considered as basic and unprovable principles taken as true in the system.
4. A specification of what inferences, or inference patterns or rules, are taken as valid in that system.

(3–4 can be done in different ways, semantically or syntactically/deductively, etc.) Because we always start discussing a logical system by discussing the language it uses, it's worth pausing to discuss the notion of using language to study language.

A. Metalanguage and Object Language

The languages of the systems we'll be studying are all symbolic logical languages. They use symbols like "→" and "∨", not found in everyday English. Of course, however, most of our readings, and most of our discussions *about* these languages will be in ordinary English. Whenever one language is used to discuss or study another, we can distinguish between the language being studied, called the ***object language***, and the language in which we conduct the study, called the ***metalanguage***.

In this course, the object languages will be the symbolic languages of first- and higher-order predicate logic, and axiomatic set theory. The metalanguage is English. To be more precise, it is a slightly more technical variant of English than ordinary English. This is because in addition to the symbols of our object language, we'll be adding some technical terms and even quasi-mathematical symbols to ordinary English to make our lives easier.

1. The logic of the metalanguage

Often, we'll be using the metalanguage to prove things about the object language, and proving anything requires logical vocabulary. Luckily, English has handy words like "all", "or", "and", "not", "if", and it allows us to add new words if we want like "iff" for "if and only if". Of course, our object languages also have logical vocabularies, and have signs like "→", "¬", "∨", "∀". But we'd better restrict those signs to the object language unless we want to get ourselves confused.

But we do want our metalanguage to be very clear and precise. For that reason, when we use the word "or", unless suggested otherwise, we mean by this the *inclusive* meaning of "or". Similarly, if we use the phrase "if ... then ..." in this class we always mean the *material conditional* unless stated otherwise. (This makes our metalanguage slightly more precise than ordinary English.) The same sorts of logical inferences that apply in the object language also apply in the metalanguage. So

> If (blah blah blah) then (yadda yadda).
> Blah blah blah.
> Therefore, yadda yadda.

... is a valid inference form. You have to use logic to study logic. There's no getting away from it. However, I'm not going to bother stating all the logical rules that are valid in the metalanguage, since I'd need to do that in the meta-meta-language, and that would just get me started on an infinite regress. However, any process of reasoning used within standard mathematical practice is OK.

2. Metalinguistic variables

Ordinary English doesn't really use variables, but they make our lives a lot easier. Since the metalanguage is usually used in this course to discuss the object language, the variables we use most often in the metalanguage are variables that are used to talk about *all* or *some* expressions of the object language. We don't want to get these variables confused with the variables of the object languages. Since predicate logic uses letters like 'x' and 'y' as variables, typically I use fancy script letters like '$\mathscr{A}$' and '$\mathscr{B}$' in the metalanguage to mean *any object-language expression* of a certain specified type. For example, I might say things like:

> If $\mathscr{A}$ is a sentence of predicate logic, then $\mathscr{A}$ contains no variables not bound by a quantifier.

Notice that, in that statement, the variable '$\mathscr{A}$' is used, not mentioned. The letter '$\mathscr{A}$' is not itself used in predicate logic, and contains no variables bound or free. It's something I use in the metalanguage to mean any sentence of the object language. So $\mathscr{A}$ might be "Fa" or it might be "$(\forall x)(Fx \rightarrow Gx)$", etc.

Variables like this used in the metalanguage are called ***schematic letters***.

Other books use other conventions. Hatcher uses Roman letters 'X' and 'Y' or 'A' and 'B' schematically. Some other books might use Greek letters instead.

II. First-Order Theories: Syntax

Some logical systems are known as *first-order theories*. In order to define this notion, we start by sketching the basic symbols used in the languages of such systems:

Definition: An ***individual constant*** is one of the lowercase letters 'a', 'b', 'c', 'd', or 'e', written with or without a numerical subscript.

Examples: 'a', 'c_3', 'd_{12}', etc.

Definition: An ***individual variable*** is one of the lowercase letters 'x', 'y', or 'z', written with or without a numerical subscript.

Examples: 'x', 'x_1', 'x_{12}' 'y', 'y_2', 'z', 'z_{13}', etc.

Definition: A ***predicate letter*** is one of the uppercase letters from 'A' to 'T', written with a numerical superscript ≥ 1, and with or without a numerical subscript.

Examples: 'A^1', 'R^2', 'H^4', 'F^1_2', 'G^3_4', etc.

- The *super*script indicates how many terms the predicate letter takes to form a statement.
- A predicate letter with a superscript '1' is called a ***monadic*** predicate letter.
- A predicate letter with a superscript '2' is called a ***binary*** or ***dyadic*** predicate letter.
- I leave these superscripts off when it is obvious from context what they must be. E.g., "$R^2(a, b)$" may be written simply "$R(a, b)$".

Definition: A ***function letter*** is one of the lowercase letters from 'f' to 'l', written with a numerical superscript ≥ 1, and with or without a numerical subscript.

Examples: 'f^1', 'g^2', 'h^3_3', etc.

Definition: A ***term*** of a first-order language is defined recursively as follows:

(i) all individual variables are terms;
(ii) all individual constants are terms;
(iii) if f is a function letter with superscript n, and $t_1, \ldots, t_n$ are terms, then $f(t_1, \ldots, t_n)$ is a term;
(iv) nothing that cannot be constructed by repeated applications of the above is a term.

Examples: 'a', 'x', '$f(a)$', '$g(x, f(y))$', etc.

Definition: An ***atomic formula*** is any expression of the form $\mathscr{P}(t_1, \ldots, t_n)$ where $\mathscr{P}$ is a predicate letter with superscript n, and $t_1, \ldots, t_n$ are all terms.

Examples: '$F^1(a)$', '$F^1(f(x))$', '$R^3_4(a, b, c)$', '$H^4(x, b, y, g(a, x))$', etc.

I adopt the convention that if the terms in an atomic formula contain no function letters, the parentheses and comma may be removed.

Examples: 'Fx' is shorthand for '$F^1(x)$', and 'Rab' is shorthand for '$R^2(a, b)$'.

Definition: A ***well-formed formula (wff)*** is recursively defined as follows:

(i) any atomic formula is a wff;
(ii) if $\mathscr{A}$ is a wff, then $\neg\mathscr{A}$ is a wff;
(iii) if $\mathscr{A}$ and $\mathscr{B}$ are wffs, then $(\mathscr{A} \vee \mathscr{B})$ is a wff;

(iv) if $\mathscr{A}$ is a wff and χ is an individual variable, then $((\forall\chi)\mathscr{A})$ is a wff;
(v) nothing that cannot be constructed by repeated applications of the above is a wff.

Other standard logical operators can be included as well, or introduced by definition:

$(\mathscr{A} \land \mathscr{B})$	is defined as	$\neg(\neg\mathscr{A} \lor \neg\mathscr{B})$
$(\mathscr{A} \to \mathscr{B})$	is defined as	$(\neg\mathscr{A} \lor \mathscr{B})$
$(\mathscr{A} \leftrightarrow \mathscr{B})$	is defined as	$((\mathscr{A} \to \mathscr{B}) \land (\mathscr{B} \to \mathscr{A}))$
$((\exists\chi)\mathscr{A})$	is defined as	$\neg((\forall\chi)\neg\mathscr{A})$

Definition: A ***first-order language*** is any logical language that makes use of the above definition of a wff, or modifies it at most by restricting which constants, function letters and predicate letters are utilized (provided that it retains at least one predicate letter). E.g., a language that does not have function letters still counts as a first-order language.

Parentheses conventions

Sometimes when a wff gets really complicated, it's easier to leave off some of the parentheses. Because this leads to ambiguities, we need conventions regarding how to read them. Different books use different conventions. According to standard conventions, we can rank the operators in the order $\neg$, $\exists$, $\forall$, $\lor$, $\land$, $\to$, $\leftrightarrow$. Those earlier on the list should be taken as having narrower scope and those later in the list as having wider scope if possible. For example:

$$Fa \to Fb \lor Fc$$

is an abbreviation of

$$(Fa \to (Fb \lor Fc))$$

Whereas

$$Fa \to Fb \leftrightarrow Fc$$

is an abbreviation of

$$((Fa \to Fb) \leftrightarrow Fc)$$

When the operators are the same, the convention is association to the left, i.e., the leftmost occurrence is taken to have narrow scope. So

$$Fa \to Fb \to Fc$$

is an abbreviation of

$$((Fa \to Fb) \to Fc)$$

Obviously for $\lor$ and $\land$, this last convention is less important, since these operators are associative.

Sometimes parentheses cannot be left off. $Fa \to (Fb \leftrightarrow Fc)$ cannot be written $Fa \to Fb \leftrightarrow Fc$.

Other notations

	My sign	Alternatives
Negation	$\neg$	$\sim$, $-$
Conjunction	$\land$	&, $\bullet$
Disjunction	$\lor$	$+$
Material conditional	$\to$	$\supset$, $\Rightarrow$
Material biconditional	$\leftrightarrow$	$\equiv$, $\Leftrightarrow$
Universal quantifier	$(\forall x)$	$\forall x$, (x), Π_x, $\land_x$
Existential quantifier	$(\exists x)$	$\exists x$, (Ex), Σ_x, $\lor_x$

III. Deduction in First-Order Theories

In addition to a specified list of well-formed formulas, a first-order theory will typically contain a list of axioms, and a list of inference rules. These are divided into two groups. First there are the *logical* axioms and inference rules, which are typically shared in common in all standard first-order theories.

Hatcher formulates them as follows:

Definition: Any instance of the following schemata is a ***logical axiom***, where $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$ are any wffs, χ any variable, and t any term with the specified property:

(1) $((\mathscr{A} \lor \mathscr{A}) \to \mathscr{A})$
(2) $(\mathscr{A} \to (\mathscr{A} \lor \mathscr{B}))$
(3) $((\mathscr{A} \lor \mathscr{B}) \to (\mathscr{B} \lor \mathscr{A}))$
(4) $((\mathscr{A} \to \mathscr{B}) \to ((\mathscr{C} \lor \mathscr{A}) \to (\mathscr{C} \lor \mathscr{B})))$
(5) $(\forall\chi)\mathscr{A}[\chi] \to \mathscr{A}[t]$, provided that no variables in t become bound when placed in the context $\mathscr{A}[t]$.

(6) $(\forall \chi)(\mathscr{B} \to \mathscr{A}[\chi]) \to (\mathscr{B} \to (\forall \chi)\mathscr{A}[\chi])$, provided that χ does not occur free in $\mathscr{B}$.

Notice that even though there are only six axiom schemata, there are infinitely many axioms, since *every* wff of one of the *forms* above counts as an axiom.

Definition: The ***inference rules*** are:

Modus ponens (MP): From $\mathscr{A} \to \mathscr{B}$ and $\mathscr{A}$, infer $\mathscr{B}$.

Universal generalization (UG): From $\mathscr{A}[\chi]$, infer $(\forall \chi)\mathscr{A}[\chi]$.

Many other books give equivalent formulations. E.g., Mendelson has the same inference rules but uses ¬ and → as primitive rather than ¬ and ∨, and schemata (1)–(4) are replaced by:

$\mathscr{A} \to (\mathscr{B} \to \mathscr{A})$

$(\mathscr{A} \to (\mathscr{B} \to \mathscr{C})) \to ((\mathscr{A} \to \mathscr{B}) \to (\mathscr{A} \to \mathscr{C}))$

$(\neg\mathscr{A} \to \neg\mathscr{B}) \to ((\neg\mathscr{A} \to \mathscr{B}) \to \mathscr{A})$

These formulations are completely equivalent; they yield all the same results.

You may instead be used to a completely different approach: one in which there are no axioms, but instead *numerous* inference rules. For example, Hardegree's system SL, which contains the rules DN, →O, ∨I, ∨O, &I, &O, ↔I, ↔O, ∀O, ∃O, ∃I, and the proof techniques CD, ID and UD. Again, the results are the same.

Definition: A ***derivation*** or ***proof*** of a wff $\mathscr{B}$ from a set of premises Γ is an ordered sequence of wffs, $\mathscr{A}_1, \ldots, \mathscr{A}_n$, where $\mathscr{B}$ is $\mathscr{A}_n$, and such that for every $\mathscr{A}_i$, $1 \le i \le n$, (1) $\mathscr{A}_i$ is an axiom of the theory (either logical or proper), (2) $\mathscr{A}_i$ is a member of Γ, or (3) there are previous members of the sequence such that $\mathscr{A}_i$ follows from them by one of the inference rules.

Definition: We use the notation "$\Gamma \vdash_K \mathscr{B}$" to mean that there exists a derivation or proof of $\mathscr{B}$ from Γ in system K. (We leave off the subscript if it is obvious from context what system is meant.)

We write "$\vdash_K \mathscr{B}$" to mean that there exists a proof of $\mathscr{B}$ in system K that does not make use of any premises beyond the axioms and inference rules of K. In such a case, $\mathscr{B}$ is said to be a ***theorem*** of K.

Definition: A ***first-order predicate calculus*** is a first-order theory that does not have any non-logical axioms.

(It might seem at first that there is only one first-order predicate calculus; in fact, however, a distinct first-order predicate calculus exists for every first-order *language*.)

Most likely, what is provable in the natural deduction system you learned in your first logic course is equivalent to the first-order predicate calculus; all of its inference rules can be derived in the above formulation and vice versa. Therefore, I encourage you to make use of the notation and proof structures of whatever natural deduction system for first-order predicate logic you know best, whether it be Hardegree's set of rules, or Copi's, or any other.

You may also shorten steps in proofs that follow from the rules of first-order logic alone by just writing "logic" or "SL" or "FOL [First-Order Logic]", citing the appropriate line numbers. Hatcher will introduce any truth-table tautology by writing "Taut". We're beyond the point where showing each step in a proof is absolutely necessary.

Here is a list of derived rules of any first-order predicate calculus, and the abbreviations I personally am most likely to use.

MT	$\mathscr{A} \to \mathscr{B}, \neg\mathscr{B} \vdash \neg\mathscr{A}$
HS	$\mathscr{A} \to \mathscr{B}, \mathscr{B} \to \mathscr{C} \vdash \mathscr{A} \to \mathscr{C}$
DS	$\mathscr{A} \vee \mathscr{B}, \neg\mathscr{B} \vdash \mathscr{A}$
DS	$\mathscr{A} \vee \mathscr{B}, \neg\mathscr{A} \vdash \mathscr{B}$
Int	$\mathscr{A} \to (\mathscr{B} \to \mathscr{C}) \vdash \mathscr{B} \to (\mathscr{A} \to \mathscr{C})$
Trans	$\mathscr{A} \to \mathscr{B} \vdash \neg\mathscr{B} \to \neg\mathscr{A}$
Trans	$\neg\mathscr{A} \to \neg\mathscr{B} \vdash \mathscr{B} \to \mathscr{A}$
Exp	$\mathscr{A} \to (\mathscr{B} \to \mathscr{C}) \vdash (\mathscr{A} \wedge \mathscr{B}) \to \mathscr{C}$
Exp	$(\mathscr{A} \wedge \mathscr{B}) \to \mathscr{C} \vdash \mathscr{A} \to (\mathscr{B} \to \mathscr{C})$
DN	$\neg\neg\mathscr{A} \vdash \mathscr{A}$
DN	$\mathscr{A} \vdash \neg\neg\mathscr{A}$
FA	$\neg\mathscr{A} \vdash \mathscr{A} \to \mathscr{B}$
TC	$\mathscr{A} \vdash \mathscr{B} \to \mathscr{A}$
TAFC	$\mathscr{A}, \neg\mathscr{B} \vdash \neg(\mathscr{A} \to \mathscr{B})$
TA	$\neg(\mathscr{A} \to \mathscr{B}) \vdash \mathscr{A}$
FC	$\neg(\mathscr{A} \to \mathscr{B}) \vdash \neg\mathscr{B}$
Inev	$\mathscr{A} \to \mathscr{B}, \neg\mathscr{A} \to \mathscr{B} \vdash \mathscr{B}$
Red	$\mathscr{A} \vee \mathscr{A} \vdash \mathscr{A}$
Red	$\mathscr{A} \vdash \mathscr{A} \wedge \mathscr{A}$
Add	$\mathscr{A} \vdash \mathscr{A} \vee \mathscr{B}$
Add	$\mathscr{A} \vdash \mathscr{B} \vee \mathscr{A}$
Com	$\mathscr{A} \vee \mathscr{B} \vdash \mathscr{B} \vee \mathscr{A}$
Com	$\mathscr{A} \wedge \mathscr{B} \vdash \mathscr{B} \wedge \mathscr{A}$
Com	$\mathscr{A} \leftrightarrow \mathscr{B} \vdash \mathscr{B} \leftrightarrow \mathscr{A}$
Assoc	$\mathscr{A} \vee (\mathscr{B} \vee \mathscr{C}) \vdash (\mathscr{A} \vee \mathscr{B}) \vee \mathscr{C}$
Assoc	$\mathscr{A} \wedge (\mathscr{B} \wedge \mathscr{C}) \vdash (\mathscr{A} \wedge \mathscr{B}) \wedge \mathscr{C}$
Assoc	$(\mathscr{A} \vee \mathscr{B}) \vee \mathscr{C} \vdash \mathscr{A} \vee (\mathscr{B} \vee \mathscr{C})$
Assoc	$(\mathscr{A} \wedge \mathscr{B}) \wedge \mathscr{C} \vdash \mathscr{A} \wedge (\mathscr{B} \wedge \mathscr{C})$
Simp	$\mathscr{A} \wedge \mathscr{B} \vdash \mathscr{A}$
Simp	$\mathscr{A} \wedge \mathscr{B} \vdash \mathscr{B}$
Conj	$\mathscr{A}, \mathscr{B} \vdash \mathscr{A} \wedge \mathscr{B}$
DM	$\neg(\mathscr{A} \wedge \mathscr{B}) \vdash \neg\mathscr{A} \vee \neg\mathscr{B}$
DM	$\neg(\mathscr{A} \vee \mathscr{B}) \vdash \neg\mathscr{A} \wedge \neg\mathscr{B}$
DM	$\neg\mathscr{A} \vee \neg\mathscr{B} \vdash \neg(\mathscr{A} \wedge \mathscr{B})$
DM	$\neg\mathscr{A} \wedge \neg\mathscr{B} \vdash \neg(\mathscr{A} \vee \mathscr{B})$
BI	$\mathscr{A} \to \mathscr{B}, \mathscr{B} \to \mathscr{A} \vdash \mathscr{A} \leftrightarrow \mathscr{B}$
BI	$\mathscr{A}, \mathscr{B} \vdash \mathscr{A} \leftrightarrow \mathscr{B}$
BI	$\neg\mathscr{A}, \neg\mathscr{B} \vdash \mathscr{A} \leftrightarrow \mathscr{B}$
BE	$\mathscr{A} \leftrightarrow \mathscr{B} \vdash \mathscr{A} \to \mathscr{B}$
BE	$\mathscr{A} \leftrightarrow \mathscr{B} \vdash \mathscr{B} \to \mathscr{A}$
BMP	$\mathscr{A} \leftrightarrow \mathscr{B}, \mathscr{A} \vdash \mathscr{B}$
BMP	$\mathscr{A} \leftrightarrow \mathscr{B}, \mathscr{B} \vdash \mathscr{A}$
BMT	$\mathscr{A} \leftrightarrow \mathscr{B}, \neg\mathscr{A} \vdash \neg\mathscr{B}$
BMT	$\mathscr{A} \leftrightarrow \mathscr{B}, \neg\mathscr{B} \vdash \neg\mathscr{A}$
UI	$(\forall \chi)\mathscr{A}[\chi] \vdash \mathscr{A}[t]$
EG	$\mathscr{A}[t] \vdash (\exists \chi)\mathscr{A}[\chi]$
CQ	$\neg(\forall \chi)\mathscr{A}[\chi] \vdash (\exists \chi)\neg\mathscr{A}[\chi]$
CQ	$\neg(\exists \chi)\mathscr{A}[\chi] \vdash (\forall \chi)\neg\mathscr{A}[\chi]$
CQ	$(\forall \chi)\neg\mathscr{A}[\chi] \vdash \neg(\exists \chi)\mathscr{A}[\chi]$
CQ	$(\exists \chi)\neg\mathscr{A}[\chi] \vdash \neg(\forall \chi)\mathscr{A}[\chi]$

Definition: The ***deduction theorem*** is a result that holds of any first-order theory in the following form:

> If $\Gamma \cup \{\mathscr{B}\} \vdash \mathscr{A}$, and no UG step is applied to a free variable of $\mathscr{B}$ on a step of the proof dependent upon $\mathscr{B}$, then $\Gamma \vdash \mathscr{B} \to \mathscr{A}$.

This underwrites both the conditional proof technique and the indirect proof technique, provided the restriction on UG is obeyed. It is also used in showing that if one obtains an existential, and another result from an arbitrary instance, one can obtain that result from the existential (EI, or existential out/instantiation).

Hatcher annotates proofs in the following way. E.g., consider this proof of the result:

$$\vdash (\forall x)(Fx \to ((\forall z)(Fz \to Gz) \to (\exists y)Gy))$$

(1)	1. Fx	H[ypothesis]
(2)	2. $(\forall z)(Fz \to Gz)$	H
(2)	3. $Fx \to Gx$	2, $e\forall$ (=UI)
(1, 2)	4. Gx	1,3 MP
(1, 2)	5. $(\exists y)Gy$	4, iE (=EG)
(1)	6. $[2 \to 5]$	2, 5 eH (=DT/CD/CP)
	7. $[1 \to 6]$	1, 6 eH
	8. $(\forall x)[7]$	7, UG

The numbers in parentheses on the left indicate the line number of any premise or hypotheses from which a step is derived. A number in brackets abbreviates the wff found at a given line. Hence "$(\forall x)[7]$" is the same as our result "$(\forall x)(Fx \rightarrow ((\forall z)(Fz \rightarrow Gz) \rightarrow (\exists y)Gy))$".

You may annotate proofs this way if you like. However, you may use whatever method you like, whether it uses boxes, lines, indentations, etc., for subproofs, or a device similar to the above, or uses "$\vdash$" with its relata on every line, etc. I don't care.

Definition: In the metalanguage, "K(Γ)" is used to denote the set of all wffs $\mathscr{B}$ such that $\Gamma \vdash_K \mathscr{B}$.

HOMEWORK 1

1. Prove informally that Γ is a subset of K(Γ) for any first-order system, i.e., that every member of Γ is a member of K(Γ).
2. Prove informally that if Γ is a subset of Δ, then K(Γ) is a subset of K(Δ), for any first-order system.

IV. Formal Semantics and Truth for First-Order Predicate Logic

A sentence or wff of a first-order language is true or false relative to an interpretation or model.

Loosely speaking, an interpretation is just a way of interpreting the language: a specification of (i) what things the variables range over, (ii) what each constant stands for, (iii) what each predicate letter stands for, and (iv) what each function letter stands for.

This can be made more precise by thinking of an interpretation set-theoretically.

Definition: An ***interpretation*** M of a first-order language consists of the following four things:

1. The specification of some non-empty set D to serve as the ***domain of quantification*** for the language
   - This set is the sum total of entities the quantifiers are interpreted to "range over".
   - The domain might include numbers only, or people only, or anything else you might imagine.
   - The domain of quantification is sometimes also known as the ***universe of discourse***.
2. An assignment, for each individual constant in the language, some fixed member of D for which it is taken to stand.
3. An assignment, for each predicate letter with superscript n in the language, a set of n-tuples taken from D.
   - This set can be thought of as the *extension* of the predicate letter under the interpretation.
4. An assignment, for each function letter with superscript n in the language, some n-place operation on D.
   - Set-theoretically, an operation is thought of as a set of ordered-pairs, the first member of which is itself some n-tuple of members of D, and the second member of which is some member of D.
   - This operation can be thought of as representing the mapping for the function represented by the function-letter. The first member of the ordered pair represents the possible arguments, and the second member of the ordered pair represents the value.

In a sense, the four parts of a model fix the meanings of the quantifiers, constants, predicate letters, and function letters, respectively. (Or at the very least, they fix as much of their meanings as is relevant in an extensional logical system such as first-order predicate logic.)

Given an interpretation, we can determine the truth-value of any ***closed wff*** (wff without any free or unbound variables). Roughly speaking, this goes as follows.

First, we consider the set of all possible assignments of a member of the domain as value to the variables. Each indirectly correlates an entity of the domain with each term. (The model fixes the correlated entity for each constant. The assignment provides an entity for each variable. In the case of terms built up from function letters, we simply take the value of the operation associated by the model for the ordered n-tuple built from the entities associated by the assignment with the terms in the argument-places to the function letter.)

A wff is said to be *satisfied* or *unsatisfied* by a given variable assignment.

Definition: The notion of ***satisfaction*** is defined recursively. For a given interpretation M with domain D and variable assignment s:

(i) If $\mathscr{A}$ is an atomic wff $\mathscr{P}(t_1, \ldots, t_n)$, then s satisfies $\mathscr{A}$ iff the ordered n-tuple formed by those entities in the domain D that s correlates with $t_1, \ldots, t_n$ is in the extension of $\mathscr{P}$ under the interpretation.
(ii) s satisfies a wff of the form $\neg\mathscr{A}$ iff s does not satisfy $\mathscr{A}$.
(iii) s satisfies a wff of the form $(\mathscr{A} \vee \mathscr{B})$ iff either s satisfies $\mathscr{A}$ or s satisfies $\mathscr{B}$.
(iv) s satisfies a wff of the form $(\forall \chi)\mathscr{A}$ iff every variable assignment s^* that differs from s at most with regard to what entity of D it correlates with the variable χ satisfies $\mathscr{A}$.

Definition: A wff $\mathscr{A}$ is said to be ***true for the interpretation M*** iff *every* variable assignment one can form from the domain D of M satisfies $\mathscr{A}$.

The notation

$$\vDash_{\mathrm{M}} \mathscr{A}$$

means that $\mathscr{A}$ is true for M. (The subscript on $\vDash$ is necessary here.)

Definition: If M is an interpretation that makes every axiom of a first-order theory K true, then M is called a ***model for*** K.

Definition: A wff $\mathscr{A}$ is said to be ***logically true*** or ***logically valid*** iff $\mathscr{A}$ is true for *every* possible interpretation.

The notation

$$\vDash \mathscr{A}$$

(leaving off any subscript) means that $\mathscr{A}$ is logically valid.

Definition: A wff $\mathscr{A}$ is a ***logical consequence*** of a set of wffs Γ iff in every interpretation, every variable assignment that satisfies every member of Γ also satisifes $\mathscr{A}$.

This is abbreviated $\Gamma \vDash \mathscr{A}$.

V. Some results covered in Mathematical Logic I

1. If M is a model for a given first-order theory K, then every theorem of K is true for M.

2. **Soundness:** Every theorem of a first-order predicate-calculus is logically valid, i.e., if $\vdash \mathscr{A}$ then $\vDash \mathscr{A}$.

3. (Another way of stating the same result.) Every interpretation for a given first-order language is a model for the first-order predicate calculus for that language.

   Definition: A first-order theory K is said to be ***consistent*** if there is no wff $\mathscr{A}$ such that both $\vdash_{\mathrm{K}} \mathscr{A}$ and $\vdash_{\mathrm{K}} \neg\mathscr{A}$.

4. **Consistency:** First-order predicate calculi are consistent, i.e., there is no wff such that both $\vdash \mathscr{A}$ and $\vdash \neg\mathscr{A}$.

5. **Gödel's Completeness Theorem:** Every logically valid wff of a given first-order language is a theorem of the associated first-order predicate calculus, i.e., if $\vDash \mathscr{A}$, then $\vdash \mathscr{A}$.

6. A first-order theory is consistent if and only if it has a model.

7. **The Skolem-Löwenheim Theorem:** If a first-order theory has any sort of model, then it has a denumerable model, i.e., a model with as many elements in the domain as there are natural numbers $\{0, 1, 2, 3, \ldots\}$.

VI. Theories with Identity/Equality

Definition: A first-order theory K is ***a first-order theory with identity [equality]*** iff there is a predicate letter A^2_1 used in the theory such that:

1. $(\forall x)A^2_1xx$ is either an axiom or theorem, and
2. for every wff $\mathscr{B}$ not containing bound occurrences of the variable 'y', the wff:

$$(\forall x)(\forall y)(A^2_1(x, y) \rightarrow (\mathscr{B}[x, x] \rightarrow \mathscr{B}[x, y]))$$

where $\mathscr{B}[x, y]$ is obtained from $\mathscr{B}[x, x]$ by substituting 'y' for zero or more free occurrences of 'x' in $\mathscr{B}[x, x]$, is either an axiom or theorem.

In such theories, the wff $A^2_1(t, u)$ is typically abbreviated as $t = u$, and $\neg A^2_1(t, u)$ is abbreviated as $t \neq u$.

Definition: If instead of a predicate letter A^2_1, K is a theory where there is a complex wff with two free variables $\mathscr{A}[x, y]$ such that similar axioms and theorems obtain, K is said to be ***a theory in which identity is definable***.

HOMEWORK 2

Prove that if K is a theory with identity, then we have the following results for any terms, t, u and v:

(Ref=)	$\vdash_K t = t$
(Sym=)	$t = u \vdash_K u = t$
(Trans=)	$t = u, u = v \vdash_K t = v$
(LL)	$t = u, \mathscr{B}[t, t] \vdash_K \mathscr{B}[t, u]$, provided that neither t nor u contain variables bound in $\mathscr{B}[t, t]$ or $\mathscr{B}[t, u]$.

These likely correspond to the rules you learned for "identity logic" in your intermediate logic courses. Hereafter, feel free to use whatever abbreviations were found there in your proofs.

Semantics for Theories With Identity

Definition: An interpretation M is a ***normal model*** iff the extension it assigns to the predicate A^2_1 is all ordered pairs of the form $\langle o, o\rangle$ taken from the domain D of M.

Definition: A ***first-order predicate calculus with identity*** is a first-order theory whose only proper ("non-logical") axioms are $(\forall x)\, x = x$ and every instance of $(\forall x)(\forall y)(x = y \rightarrow (\mathscr{B} \rightarrow \mathscr{B}^*))$ where $\mathscr{B}^*$ is obtained from from $\mathscr{B}$ by substituting 'y' for zero or more free occurrences of 'x' in places where 'y' does not become bound in the narrow context of $\mathscr{B}^*$.

Some results:

1. **Soundness:** If a wff $\mathscr{A}$ is a theorem of a first-order predicate calculus with identity, then $\mathscr{A}$ is true in all normal models.
2. Every normal model for a given first-order language is a model for the associated predicate calculus with identity.
3. **Consistency:** Every first-order predicate calculus with identity K is consistent.
4. If a first-order theory with identity is consistent, then it has a normal model.
5. **Semantic Completeness:** If $\mathscr{A}$ is a wff of a given first-order language, and $\mathscr{A}$ is true in all normal models, then if K is the associated first-order predicate calculus with identity for that language, $\vdash_K \mathscr{A}$.

VII. Variable-Binding Term Operators (vbtos)

A. Introduction

Definition: A ***first-order language with vbtos*** is just like a first-order language, except adding terms of the form $\nu\chi\, \mathscr{A}[\chi]$, where χ is any variable and $\mathscr{A}[\chi]$ any wff, and ν some new symbol special to the language.

Variable-binding term operators are also called ***subnectives***.

All occurrences of the variable χ occurring in a term of the form $\nu\chi\, \mathscr{A}[\chi]$ are considering *bound*.

We shall limit our discussion to vbtos with one bound variable. We shall also limit our discussion to *extensional* vbtos, i.e., those yielding terms standing for the same object whenever the open wffs to which the vbto is applied are satisfied by the same entities.

Example vbtos:

1. The description operator $\imath x\, \mathscr{A}[x]$

   read: the x such that $\mathscr{A}[x]$
2. Selection/choice operator: $\mu x\, \mathscr{A}[x]$ or $\varepsilon x\, \mathscr{A}[x]$

read: the least/first x such that $\mathscr{A}[x]$, or *an* x such that $\mathscr{A}[x]$

3. Set/class abstraction: $\{x|\mathscr{A}[x]\}$

   read: the set/class of all x such that $\mathscr{A}[x]$

B. Semantics for vbtos

Interpretations for first-order languages with vbtos must also include an additional component:

5. For each vbto, an assignment of a function mapping every subset of D to a member of D.

For example, and intended model for the use of the description operator will assign to it a function mapping every singleton subset of D to its sole member, and will map every empty and non-singleton subset of D to some chosen entity in D (e.g., the number 0).

In determining the satisfaction/truth of a wff containing vbtos, a variable assignment s will assign to a term of the form $\nu\chi\,\mathscr{A}[\chi]$ the value of the function assigned to the vbto by the model for the subset of the domain made up of those entities o for which the assignment s^* just like s except with o assigned to the variable χ satisfies $\mathscr{A}[\chi]$ as argument.

The definitions of truth/satisfaction remain unchanged otherwise.

C. Deduction for vbtos

To preserve completeness, we add the following axiom schemata to the first-order predicate calculus for a given first-order language with vbtos:

V.1 $(\forall x)(\mathscr{A}[x] \leftrightarrow \mathscr{B}[x]) \rightarrow (\mathscr{C}[\nu x\,\mathscr{A}[x]] \leftrightarrow \mathscr{C}[\nu x\,\mathscr{B}[x]])$, provided that no free variable of $\nu x\,\mathscr{A}[x]$ or $\nu x\,\mathscr{B}[x]$ becomes bound in the context $\mathscr{C}[\nu x\,\mathscr{A}[x]]$ or $\mathscr{C}[\nu x\,\mathscr{B}[x]]$.

V.2 $\mathscr{C}[\nu\chi\,\mathscr{A}[\chi]] \leftrightarrow \mathscr{C}[\nu y\,\mathscr{A}[y]]$, provided that no free variable of $\nu\chi\,\mathscr{A}[\chi]$ or $\nu y\,\mathscr{A}[y]$ becomes bound in the context $\mathscr{C}[\nu\chi\,\mathscr{A}[\chi]]$ or $\mathscr{C}[\nu y\,\mathscr{A}[y]]$.

For a first-order predicate calculus *with identity*, we add instead the schemata:

V.1′ $(\forall x)(\mathscr{A}[x] \leftrightarrow \mathscr{B}[x]) \rightarrow \nu x\,\mathscr{A}[x] = \nu x\,\mathscr{B}[x]$

V.2′ $\nu\chi\,\mathscr{A}[\chi] = \nu y\,\mathscr{A}[y]$

HOMEWORK 3

Show that schemata V.1 and V.2 are provable from V.1′ and V.2′ in a theory with identity.

VIII. First-Order Peano Arithmetic

The ***Peano(-Dedekind) axioms*** are the following principles, in English:

1. Zero is a natural number.
2. Every natural number has a successor natural number.
3. Zero is not the successor of any natural number.
4. No two natural numbers have the same successor.
5. Whatever is true of zero and is always true of the successor of any natural number of which it holds, hold of all natural numbers. (The principles of mathematical induction.)

Various versions of “Peano arithmetic” comes from formalizing these axioms in a formal system with an identity predicate, along with some mechanism for capturing recursive functions. In a first order system, this can be done by introducing addition and multiplication functions.

As an example of a first-order system, we shall consider system S, or ***first-order Peano arithmetic***.

The first-order *language* of S is as follows:

1. There is one constant, ‘a’, but instead we write ‘0’.
2. There are three function letters, ‘f^1’, ‘f_1^2’, and ‘f_2^2’, but instead we write ‘′’, ‘+’ and ‘·’ with infix notation.
3. There is one predicate letter ‘A_1^2’, but instead we write ‘=’ (again, with infix notation).

Definition: The ***proper*** or ***non-logical*** axioms of S are:

S1. $(\forall x)\, x = x$
S2. $(\forall x)(\forall y)(x = y \rightarrow y = x)$

S3. $(\forall x)(\forall y)(\forall z)(x = y \land y = z \rightarrow x = z)$
S4. $(\forall x)(\forall y)(x = y \rightarrow x' = y')$
S5. $(\forall x_1)(\forall x_2)(\forall y_1)(\forall y_2)(x_1 = x_2 \land y_1 = y_2 \rightarrow (x_1 + y_1) = (x_2 + y_2) \land (x_1 \cdot y_1) = (x_2 \cdot y_2))$
S6. $(\forall x)(x' \neq 0)$
S7. $(\forall x)(\forall y)(x' = y' \rightarrow x = y)$
S8. $(\forall x)(x + 0 = x)$
S9. $(\forall x)(\forall y)((x + y') = (x + y)')$
S10. $(\forall x)(x \cdot 0 = 0)$
S11. $(\forall x)(\forall y)((x \cdot y') = ((x \cdot y) + x))$
S12. $\mathscr{A}[0] \land (\forall x)(\mathscr{A}[x] \rightarrow \mathscr{A}[x']) \rightarrow (\forall x)\mathscr{A}[x]$

Notice that because S12 is an axiom schema, not a single axiom, system S has infinitely many proper axioms.

Semantics for S

The intended interpretation for S, called the ***standard interpretation***, is as follows:

1. The domain of quantification D is the set of natural numbers $\{0, 1, 2, 3, \ldots\}$.
2. The constant '0' stands for zero.
3. The extension of '=' is the identity relation on the natural numbers, i.e., $\{\langle 0,0\rangle, \langle 1,1\rangle, \langle 2,2\rangle, \ldots\}$.
4. The function signs ''', '+', and '·' are respectively assigned to the successor, addition and multiplication operations.

Many other arithmetical notions can be introduced into S by definition, e.g.:

1	for	$0'$
2	for	$0''$
3	for	$0'''$, etc.
$t < u$	for	$(\exists \chi)(\chi \neq 0 \land t + \chi = u)$

Some Results

1. Because the standard interpretation is a model for S, S is consistent.
2. Every recursive number-theoretic function is representable in S, and every recursively decidable number-theoretic property or relation is expressible in S.
3. **Gödel's First Incompleteness Theorem:** There are closed wffs $\mathscr{A}$ of the language of S such that neither $\mathscr{A}$ nor $\neg\mathscr{A}$ are theorems of S.
   - Hence there are wffs that are true in the standard interpretation that are *not* theorems of S.
   - The same is true for any consistent theory obtained from S by adding a finite number of axioms, or even by adding a recursively decidable infinite number of axioms.
   - The same is true for any consistent first-order theory with recursively decidable lists of wffs and axioms for which the previous result holds, including any such first-order theory in which the axioms of S (or equivalents) are derivable as theorems.
4. **Gödel's Second Incompleteness Theorem:** The consistency of S cannot be proven in S itself (and the same holds for similar systems).
5. There are non-standard models of S, i.e., models not isomorphic to the standard interpretation.

HOMEWORK 4

Prove that $\vdash_S (\forall x)(\forall y)((y' + x) = (y + x)')$. Hint: use S12 with the above as consequent.

IX. System F

A. Frege and Hatcher's System F

Gottlob Frege (1848–1925) was the first logician ever to lay out fully an axiomatic deductive calculus, and is widely regarded as the father of both predicate logic and modern mathematical logic.

Frege was philosophically convinced of a thesis now known as ***logicism***: the thesis that mathematics is, in some sense or another, reducible to logic, or that mathematical truth is a species of logical truth.

Frege, somewhat naïvely, attempted to argue in favor of logicism by inventing a formal system in which he believed all of pure arithmetic could be derived, but whose axioms consisted only of general logical principles. The core predicate logic was laid out in Frege's 1879 *Begriffsschrift*, and expanded into set or class theory in his 1893 *Grundgesetze der Arithmetik*.

Unfortunately, the system of Frege's *Grundgesetze* turned out to be inconsistent.

Hatcher creates a system he calls "System F", after Frege, and claims he presents it "in a form quite close to the original". This isn't even close to being true.

Nevertheless, Hatcher's system does a good job explaining how it is, from a naïve perspective, one might attempt to create a system in which the basic terminology of arithmetic could be defined, and the basic principles of number theory proven from some basic axioms about the nature of classes, sets or (as Frege called them) "extensions of concepts".

B. Syntax of F

We use a first-order language with one vbto; in particular

(i) There are no individual constants in the language;
(ii) There are no function letters in the language;
(iii) There is one predicate letter, but instead of writing $A_1^2(t,u)$, we write $t \in u$. This is read either "t is a member of u", "t is an element of u", or simply "t is in u".
(iv) There is one vbto, forming terms written $\{\chi|\mathscr{A}[\chi]\}$. This is read, "the set [or class] of all χ such that $\mathscr{A}[\chi]$".

It is assumed (for Hatcher's system) that every member of the domain of quantification is a set or class. This allows F to be a system in which *identity is defined*. In particular:

Definition: $t = u$ is defined as $(\forall\chi)(\chi \in t \leftrightarrow \chi \in u)$, where χ is the first variable not occurring free in either t or u.

C. Axioms of F

In addition to the logical axioms, F also includes all instances of the following two schemata:

F1 $(\forall x)(\forall y)(x = y \rightarrow (\mathscr{A}[x,x] \rightarrow \mathscr{A}[x,y]))$, where $\mathscr{A}[x,y]$ is derived from $\mathscr{A}[x,x]$ by replacing one or more free occurrences of 'x' with 'y' without thereby binding 'y'.

F2 $(\forall y)(y \in \{\chi|\mathscr{A}[\chi]\} \leftrightarrow \mathscr{A}[y])$, provided that y is not bound in $\mathscr{A}[y]$.

D. Some set-theoretic notions

Definitions: Where χ is the first variable not occurring free in t or u:

$t \notin u$	for	$\neg t \in u$
$t \neq u$	for	$\neg t = u$
V	for	$\{x|x = x\}$
Λ	for	$\{x|x \neq x\}$
$\{t\}$	for	$\{\chi|\chi = t\}$
$\{t_1,\ldots,t_n\}$	for	$\{\chi|\chi = t_1 \vee \ldots \vee \chi = t_n\}$
$\overline{t}$	for	$\{\chi|\chi \notin t\}$

$(t \cap u)$	for	$\{\chi|\chi \in t \wedge \chi \in u\}$
$(t \cup u)$	for	$\{\chi|\chi \in t \vee \chi \in u\}$
$(t \subseteq u)$	for	$(\forall\chi)(\chi \in t \rightarrow \chi \in u)$
$(t - u)$	for	$\{\chi|\chi \in t \wedge \chi \notin u\}$

From F2, we get two derived rules, where no free variables of t are bound in $\mathscr{A}[t]$:

F2R $\mathscr{A}[t] \vdash_{\mathrm{F}} t \in \{\chi|\mathscr{A}[\chi]\}$
F2R $t \in \{\chi|\mathscr{A}[\chi]\} \vdash_{\mathrm{F}} \mathscr{A}[t]$

Proof: Direct from F2, UI and BMP.

Some theorems:

T1. $\vdash_{\mathrm{F}} (\forall x)\, x = x$

(Hence, F is a theory with identity.)

T2. $\vdash_{\mathrm{F}} (\forall x)\, x \in \mathrm{V}$

Proof:

1. $x = x$ — T1, UI
2. $x \in \{x | x = x\}$ — 1 F2R
3. $x \in \mathrm{V}$ — 2 Df. V
4. $(\forall x)\, x \in \mathrm{V}$ — 3 UG

T2.1 $\vdash_{\mathrm{F}} \mathrm{V} \in \mathrm{V}$

T3. $\vdash_{\mathrm{F}} (\forall x)\, x \notin \Lambda$

T4. $\vdash_{\mathrm{F}} (\forall x)((\forall y)(y \notin x) \rightarrow x = \Lambda)$

T5. $\vdash_{\mathrm{F}} (\forall x)(\forall y)(y \in x \rightarrow ((x - \{y\}) \cup \{y\}) = x)$

HOMEWORK 5

Prove the following, without making use of Russell's paradox or other contradiction:

(a) $\vdash_{\mathrm{F}} (\forall x)(\forall y)(x \cap y = y \cap x)$
(b) $\vdash_{\mathrm{F}} (\forall x)(\forall y)(\forall z)(x \cup (y \cup z) = (x \cup y) \cup z)$
(c) $\vdash_{\mathrm{F}} (\forall x)(\forall y)(x - y = x \cap \overline{y})$

E. Some mathematical notions

On Frege's conception of (cardinal) numbers, a number was taken to be a class of classes all of which had members that could be put in 1–1 correspondence with each other. Zero would be the class of all classes with no members, one would be the class of single-membered classes, two would be the class of all two-membered classes, and so on.

Definitions:

0	for	$\{\Lambda\}$
t'	for	$\{\chi | (\exists y)(y \in \chi \wedge \chi - \{y\} \in t)\}$
N	for	$\{x | (\forall y)(0 \in y \wedge (\forall z)(z \in y \rightarrow z' \in y) \rightarrow x \in y)\}$
1	for	$0'$
2	for	$1'$
3	for	$2'$ and so on

$\mathrm{Fin}(t)$	for	$(\exists \chi)(\chi \in N \wedge t \in \chi)$
$\mathrm{Inf}(t)$	for	$\neg \mathrm{Fin}(t)$

Some theorems:

T6. $\vdash_{\mathrm{F}} 0 \in N$ (=Peano Postulate 1)

Proof:

1.	$0 \in y \wedge (\forall z)(z \in y \rightarrow z' \in y) \rightarrow 0 \in y$	Taut
2.	$(\forall y)(0 \in y \wedge (\forall z)(z \in y \rightarrow z' \in y) \rightarrow 0 \in y)$	1 UG
3.	$0 \in \{x | (\forall y)(0 \in y \wedge (\forall z)(z \in y \rightarrow z' \in y) \rightarrow x \in y)\}$	2 F2R
4.	$0 \in N$	3 Df. N

T7. $\vdash_{\mathrm{F}} (\forall x)\, 0 \neq x'$ (=Peano Postulate 3)

Proof:

(1)	1.	$0 = x'$	Hyp
	2.	$\Lambda = \Lambda$	T1 UI
	3.	$\Lambda \in \{x | x = \Lambda\}$	2 F2R
	4.	$\Lambda \in 0$	3 Dfs. 0, $\{t\}$
(1)	5.	$\Lambda \in x'$	1, 4 LL
(1)	6.	$\Lambda \in \{y | (\exists z)(z \in y \wedge y - \{z\} \in x)\}$	5 Df. $'$
(1)	7.	$(\exists z)(z \in \Lambda \wedge \Lambda - \{z\} \in x)$	6 F2R
(1)	8.	$a \in \Lambda \wedge \Lambda - \{a\} \in x$	7 EI
	9.	$a \notin \Lambda$	T3 UI
(1)	10.	$a \in \Lambda \wedge a \notin \Lambda$	8, 9 SL
	11.	$0 \neq x'$	1–10 RAA
	12.	$(\forall x)\, 0 \neq x'$	11 UG

T8. $\vdash_{\mathrm{F}} (\forall x)(x \in N \rightarrow x' \in N)$ (=Peano Postulate 2)

Proof:

(1)	1. $x \in N$	Hyp
(1)	2. $x \in \{x \mid (\forall y)(0 \in y \wedge (\forall z)(z \in y \to z' \in y) \to x \in y)\}$	1 Df. N
(1)	3. $(\forall y)(0 \in y \wedge (\forall z)(z \in y \to z' \in y) \to x \in y)$	2 F2R
(4)	4. $0 \in y \wedge (\forall z)(z \in y \to z' \in y)$	Hyp
(1)	5. $0 \in y \wedge (\forall z)(z \in y \to z' \in y) \to x \in y$	3 UI
(1,4)	6. $x \in y$	4, 5 MP
(4)	7. $(\forall z)(z \in y \to z' \in y)$	4 Simp
(4)	8. $x \in y \to x' \in y$	7 UI
(1,4)	9. $x' \in y$	6, 8 MP
(1)	10. $0 \in y \wedge (\forall z)(z \in y \to z' \in y) \to x' \in y$	4–9 CP
(1)	11. $(\forall y)(0 \in y \wedge (\forall z)(z \in y \to z' \in y) \to x' \in y)$	10 UG
(1)	12. $x' \in \{x \mid (\forall y)(0 \in y \wedge (\forall z)(z \in y \to z' \in y) \to x \in y)\}$	11 F2R
(1)	13. $x' \in N$	12 Df. N
	14. $x \in N \to x' \in N$	1–13 CP
	15. $(\forall x)(x \in N \to x' \in N)$	14 UG

T9. $\vdash_{\mathrm{F}} (\forall y)(0 \in y \wedge (\forall z)(z \in y \to z' \in y) \to N \subseteq y)$

Proof:

(1)	1. $0 \in y \wedge (\forall z)(z \in y \to z' \in y)$	Hyp
(2)	2. $x \in N$	Hyp
(2)	3. $x \in \{x \mid (\forall y)(0 \in y \wedge (\forall z)(z \in y \to z' \in y) \to x \in y)\}$	2 Df. N
(2)	4. $(\forall y)(0 \in y \wedge (\forall z)(z \in y \to z' \in y) \to x \in y)$	3 F2R
(2)	5. $0 \in y \wedge (\forall z)(z \in y \to z' \in y) \to x \in y$	4 UI
(1,2)	6. $x \in y$	1, 5 MP
(1)	7. $x \in N \to x \in y$	2–6 CP
(1)	8. $(\forall x)(x \in N \to x \in y)$	7 UG
(1)	9. $N \subseteq y$	8 Df. $\subseteq$
	10. $0 \in y \wedge (\forall z)(z \in y \to z' \in y) \to N \subseteq y$	1–9 CP
	11. $(\forall y)(0 \in y \wedge (\forall z)(z \in y \to z' \in y) \to N \subseteq y)$	10 UG

T10. $\vdash_F \mathscr{A}[0] \wedge (\forall x)(\mathscr{A}[x] \rightarrow \mathscr{A}[x']) \rightarrow (\forall x)(x \in N \rightarrow \mathscr{A}[x])$
(=Peano Postulate 5)

Proof:

(1)	1. $\mathscr{A}[0] \wedge (\forall x)(\mathscr{A}[x] \rightarrow \mathscr{A}[x'])$	Hyp
(2)	2. $x \in N$	Hyp
	3. $0 \in \{y \mid \mathscr{A}[y]\} \wedge (\forall z)(z \in \{y \mid \mathscr{A}[y]\} \rightarrow z' \in \{y \mid \mathscr{A}[y]\}) \rightarrow N \subseteq \{y \mid \mathscr{A}[y]\}$	T9 UI
(1)	4. $\mathscr{A}[0]$	1 Simp
(1)	5. $0 \in \{y \mid \mathscr{A}[y]\}$	4 F2R
(6)	6. $z \in \{y \mid \mathscr{A}[y]\}$	Hyp
(6)	7. $\mathscr{A}[z]$	6 F2R
(1)	8. $\mathscr{A}[z] \rightarrow \mathscr{A}[z']$	1 Simp, UI
(1,6)	9. $\mathscr{A}[z']$	7, 8 MP
(1,6)	10. $z' \in \{y \mid \mathscr{A}[y]\}$	9 F2R
(1)	11. $z \in \{y \mid \mathscr{A}[y]\} \rightarrow z' \in \{y \mid \mathscr{A}[y]\}$	6–10 CP
(1)	12. $(\forall z)(z \in \{y \mid \mathscr{A}[y]\} \rightarrow z' \in \{y \mid \mathscr{A}[y]\})$	11 UG
(1)	13. $N \subseteq \{y \mid \mathscr{A}[y]\}$	3, 5, 12 SL
(1)	14. $(\forall x)(x \in N \rightarrow x \in \{y \mid \mathscr{A}[y]\})$	13 Df. $\subseteq$
(1,2)	15. $x \in \{y \mid \mathscr{A}[y]\}$	2, 14 UI, MP
(1,2)	16. $\mathscr{A}[x]$	15 F2R
(1)	17. $x \in N \rightarrow \mathscr{A}[x]$	2–16 CP
(1)	18. $(\forall x)(x \in N \rightarrow \mathscr{A}[x])$	17 UG
	19. $[1 \rightarrow 18]$	1–18 CP

Corollaries (see Hatcher):

T11. $\vdash_F (\forall x)(0 \in x \wedge (\forall y)(y \in x \wedge y \in N \rightarrow y' \in x) \rightarrow N \subseteq x)$

T12. $\vdash_F \mathscr{A}[0] \wedge (\forall x)(x \in N \wedge \mathscr{A}[x] \rightarrow \mathscr{A}[x']) \rightarrow$
$(\forall x)(x \in N \rightarrow \mathscr{A}[x])$

F. To Infinity and Beyond

We now have four of the five Peano postulates. The last one, that no two natural numbers have the same successor, is more difficult to prove.

The reason it is difficult is best understood by considering the numbers as defined by Frege. They are classes of like-membered classes:

$0 = \{\{\}\}$
$1 = \{\{\text{Ernie}\}, \{\text{Bert}\}, \{\text{Elmo}\}, \{\text{Gonzo}\}, \ldots\}$
$2 = \{\{\text{Ernie}, \text{Bert}\}, \{\text{Big Bird}, \text{Snuffy}\}, \{\text{Kermit}, \text{Piggy}\}, \ldots\}$
$3 = \{\{\text{Animal}, \text{Fozzie}, \text{Rowlf}\}, \{\text{Jennifer}, \text{Angelina}, \text{Brad}\}, \ldots\}$
etc.

Notice that if there were only finitely many objects, there would be some natural number n such that:

$$n = \{\mathrm{V}\}$$

Then the *successor* of n would be the empty class, since there is no set which has members we could take away to form V. Hence:

$$n' = \{\}$$
$$n'' = \{\} \text{ and therefore } n' = n''$$

In that case, n and n' have the same successor, but it still holds that $n \neq n'$. Hence, if there were only finitely many objects, the fourth Peano postulate would be false.

In order to prove the fourth Peano postulate, then, we need to *prove* a theorem of infinity: that the universal class is not a member of any natural number. In Frege's logic, this can be done a number of ways. We're going to do it by proving that the following sequence, ω, never ends:

Λ
$\{\Lambda\}$
$\{\Lambda, \{\Lambda\}\}$
$\{\Lambda, \{\Lambda\}, \{\Lambda, \{\Lambda\}\}\}$
$\{\Lambda, \{\Lambda\}, \{\Lambda, \{\Lambda\}\}, \{\Lambda, \{\Lambda\}\{\Lambda, \{\Lambda\}\}\}\}$
and so on, where each member is the set of all before it.

This sequence is also called the ***von Neumann sequence***.

Definition:

ω for $\{x | (\forall y)(\Lambda \in y \wedge (\forall z)(z \in y \rightarrow z \cup \{z\} \in y) \rightarrow x \in y)\}$

Notice the similarity between this and the definition of N.

T13. $\vdash_F (\forall x)(x \in N \rightarrow$
$(\forall y_1)(\forall y_2)(\forall z)(y_1 \in z \wedge y_2 \in z \wedge z - \{y_1\} \in x \rightarrow z - \{y_2\} \in x))$

(The proof of T13 is long and boring; see Hatcher.)

T14. $\vdash_F (\forall x)(x \in N \rightarrow (\forall y)(\forall z)(y \in x \wedge z \in x \wedge y \subseteq z \rightarrow y = z))$

HOMEWORK 6

Prove T14. Use T12 and T13.

T15. $\vdash_F \Lambda \in \omega$

T16. $\vdash_F (\forall x)(\Lambda \neq x \cup \{x\})$

T17. $\vdash_F (\forall x)(x \in \omega \rightarrow x \cup \{x\} \in \omega)$

T18. $\vdash_F (\forall x)(\Lambda \in x \wedge (\forall z)(z \in x \rightarrow z \cup \{z\} \in x) \rightarrow \omega \subseteq x)$

T19. $\vdash_F \mathscr{A}[\Lambda] \wedge (\forall x)(\mathscr{A}[x] \rightarrow \mathscr{A}[x \cup \{x\}]) \rightarrow (\forall x)(x \in \omega \rightarrow \mathscr{A}[x])$

The proofs of T15–T19 are straightforward from the definitions.

T20. $\vdash_F (\forall x)(x \in \omega \rightarrow (\forall y)(y \in x \rightarrow y \subseteq x))$

Proof:

	1. $y \notin \Lambda$	T3 UI
	2. $y \in \Lambda \to y \subseteq \Lambda$	1 FA
	3. $(\forall y)(y \in \Lambda \to y \subseteq \Lambda)$	2 UG
(4)	4. $(\forall y)(y \in x \to y \subseteq x)$	Hyp
(5)	5. $y \in x \cup \{x\}$	Hyp
(5)	6. $y \in \{y \mid y \in x \vee y \in \{x\}\}$	5, Df. $\cup$
(5)	7. $y \in x \vee y \in \{x\}$	6 F2R
	8. $y \in \{x\} \to y = x$	Df. $\{t\}$, F2, UI, SL
	9. $y = x \to y \subseteq x$	Dfs. $=$, $\subseteq$, Logic
	10. $y \in \{x\} \to y \subseteq x$	8, 9 HS
(4)	11. $y \in x \to y \subseteq x$	4 UI
(4,5)	12. $y \subseteq x$	7, 10, 11 SL
(13)	13. $z \in y$	Hyp
(4,5,13)	14. $z \in x$	12, 13 Df. $\subseteq$, UI, MP
(4,5,13)	15. $z \in x \vee z \in \{x\}$	14 Add
(4,5,13)	16. $z \in \{y \mid y \in x \vee y \in \{x\}\}$	15 F2R
(4,5,13)	17. $z \in x \cup \{x\}$	16 Df. $\cup$
(4,5)	18. $z \in y \to z \in x \cup \{x\}$	13–17 CP
(4,5)	19. $(\forall z)(z \in y \to z \in x \cup \{x\})$	18 UG
(4,5)	20. $y \subseteq x \cup \{x\}$	19 Df. $\subseteq$
(4)	21. $y \in x \cup \{x\} \to y \subseteq x \cup \{x\}$	5–20 CP
(4)	22. $(\forall y)(y \in x \cup \{x\} \to y \subseteq x \cup \{x\})$	21 UG
	23. $(\forall y)(y \in x \to y \subseteq x) \to (\forall y)(y \in x \cup \{x\} \to y \subseteq x \cup \{x\})$	4–22 CP
	24. $(\forall x)((\forall y)(y \in x \to y \subseteq x) \to (\forall y)(y \in x \cup \{x\} \to y \subseteq x \cup \{x\}))$	23 UG
	25. $(\forall x)(x \in \omega \to (\forall y)(y \in x \to y \subseteq x))$	3, 24, T19 Conj, MP

T21. $\vdash_{\mathrm{F}} (\forall x)(\Lambda \in x \wedge (\forall z)(z \in x \wedge z \in \omega \to z \cup \{z\} \in x) \to \omega \subseteq x)$

T22. $\vdash_{\mathrm{F}} \mathscr{A}[\Lambda] \wedge (\forall x)(\mathscr{A}[x] \wedge x \in \omega \to \mathscr{A}[x \cup \{x\}]) \to (\forall x)(x \in \omega \to \mathscr{A}[x])$

These are proven similarly to T11 and T12.

T23. $\vdash_{\mathrm{F}} (\forall x)(x \in \omega \to x \notin x)$

Proof:

	1. $\Lambda \notin \Lambda$	T3 UI
(2)	2. $x \notin x \wedge x \in \omega$	Hyp
(3)	3. $x \cup \{x\} \in x \cup \{x\}$	Hyp
(3)	4. $x \cup \{x\} \in \{y \mid y \in x \vee y \in \{x\}\}$	3 Df. $\cup$
(3)	5. $x \cup \{x\} \in x \vee x \cup \{x\} \in \{x\}$	4 F2R
(6)	6. $x \cup \{x\} \in \{x\}$	Hyp
(6)	7. $x \cup \{x\} \in \{y \mid y = x\}$	6 Df. $\{t\}$
(6)	8. $x \cup \{x\} = x$	7 F2R
	9. $x \cup \{x\} \in \{x\} \rightarrow x \cup \{x\} = x$	6–8 CP
(10)	10. $x \cup \{x\} \in x$	Hyp
(2)	11. $(\forall y)(y \in x \rightarrow y \subseteq x)$	2, T20 UI, MP
(2,10)	12. $x \cup \{x\} \subseteq x$	10, 11 UI, MP
	13. $x \subseteq x \cup \{x\}$	Dfs. $\subseteq$, $\cup$, F2R
(2,10)	14. $x \cup \{x\} = x$	12, 13 Dfs. $\subseteq$, $=$
(2)	15. $x \cup \{x\} \in x \rightarrow x \cup \{x\} = x$	10–14 CP
(2,3)	16. $x \cup \{x\} = x$	5, 9, 15 SL
(2,3)	17. $x \in x$	3, 11 LL
(2,3)	18. $x \in x \wedge x \notin x$	2, 17 SL
(2)	19. $x \cup \{x\} \notin x \cup \{x\}$	3–18 RAA
	20. $x \notin x \wedge x \in \omega \rightarrow x \cup \{x\} \notin x \cup \{x\}$	2–19 CP
	21. $(\forall x)(x \notin x \wedge x \in \omega \rightarrow x \cup \{x\} \notin x \cup \{x\})$	20 UG
	22. $(\forall x)(x \in \omega \rightarrow x \notin x)$	1, 21, T22 SL

T24. $\vdash_{\mathrm{F}} (\forall x)(\forall y)(x \in \omega \wedge y \in N \wedge x \in y \rightarrow x \cup \{x\} \in y')$

(T24 is an easy consequence of T23; see Hatcher, p. 94.)

T25. $\vdash_{\mathrm{F}} (\forall x)(x \in N \rightarrow (\exists y)(y \in \omega \wedge y \in x))$

Proof:

	1. $\Lambda \in 0$	Df. 0, Ref= F2R
	2. $\Lambda \in \omega \wedge \Lambda \in 0$	1, T15 Conj
	3. $(\exists y)(y \in \omega \wedge y \in 0)$	2 EG
(4)	4. $x \in N \wedge (\exists y)(y \in \omega \wedge y \in x)$	Hyp
(4)	5. $(\exists y)(y \in \omega \wedge y \in x)$	4 Simp
(4)	6. $a \in \omega \wedge a \in x$	5 EI
	7. $a \in \omega \wedge x \in N \wedge a \in x \rightarrow a \cup \{a\} \in x'$	T24 UI×2

(4)	8. $a \cup \{a\} \in x'$	4, 6, 7 SL
(4)	9. $a \cup \{a\} \in \omega$	6, T17 UI, MP
(4)	10. $a \cup \{a\} \in \omega \wedge a \cup \{a\} \in x'$	8, 9 Conj
(4)	11. $(\exists y)(y \in \omega \wedge y \in x')$	10 EG
	12. $x \in N \wedge (\exists y)(y \in \omega \wedge y \in x) \rightarrow (\exists y)(y \in \omega \wedge y \in x')$	4–11 CP
	13. $(\forall x)(x \in N \wedge (\exists y)(y \in \omega \wedge y \in x) \rightarrow (\exists y)(y \in \omega \wedge y \in x'))$	12 UG
	14. $(\forall x)(x \in N \rightarrow (\exists y)(y \in \omega \wedge y \in x))$	3, 13, T12 SL

T26. $\vdash_F \Lambda \notin N$

T27. $\vdash_F (\forall x)\, x \subseteq \mathrm{V}$

(T26 is an obvious consequence of T25; T27 is obvious, period.)

T28. $(\forall x)\neg(\mathrm{V} \in x \wedge x \in N)$

Proof:

(1)	1. $\mathrm{V} \in x \wedge x \in N$	Hyp
(2)	2. $y \in x'$	Hyp
(2)	3. $y \in \{y \mid (\exists z)(z \in y \wedge y - \{z\} \in x)\}$	2 Df. $'$
(2)	4. $(\exists z)(z \in y \wedge y - \{z\} \in x)$	3 F2R
(2)	5. $a \in y \wedge y - \{a\} \in x$	4 EI
	6. $y - \{a\} \subseteq V$	T27 UI
	7. $x \in N \rightarrow (\forall y)(\forall z)(y \in x \wedge z \in x \wedge y \subseteq z \rightarrow y = z)$	T14 UI
(1)	8. $(\forall y)(\forall z)(y \in x \wedge z \in x \wedge y \subseteq z \rightarrow y = z)$	1, 7 SL
(1)	9. $y - \{a\} \in x \wedge \mathrm{V} \in x \wedge y - \{a\} \subseteq \mathrm{V} \rightarrow y - \{a\} = \mathrm{V}$	8 UI×2
(1,2)	10. $y - \{a\} = \mathrm{V}$	1, 5, 6 SL
(11)	11. $a \in y - \{a\}$	Hyp
(11)	12. $a \in \{x \mid x \in y \wedge x \notin \{a\}\}$	11 Df. $-$
(11)	13. $a \in y \wedge a \notin \{a\}$	12 F2R
	14. $a \in \{a\}$	Ref=, F2R, Def. $\{t\}$
(11)	15. $a \in \{a\} \wedge a \notin \{a\}$	13, 14 SL
	16. $a \notin y - \{a\}$	11–15 RAA
(1,2)	17. $a \notin \mathrm{V}$	10, 16 LL
	18. $a \in \mathrm{V}$	T2 UI
(1,2)	19. $a \in \mathrm{V} \wedge a \notin \mathrm{V}$	17, 18 Conj
(1)	20. $y \notin x'$	2–19 RAA

(1)	21. $(\forall y)\, y \notin x'$	20 UG
(1)	22. $x' = \Lambda$	21, T4 UI, MP
(1)	23. $x' \in N$	1, T8 UI, SL
(1)	24. $\Lambda \in N$	22, 23 LL
	25. $\Lambda \notin N$	T26
(1)	26. $\Lambda \in N \wedge \Lambda \notin N$	24, 25 Conj
	27. $\neg(\mathrm{V} \in x \wedge x \in N)$	1–26 RAA
	28. $(\forall x)\neg(\mathrm{V} \in x \wedge x \in N)$	27 UG

T28.1 $\vdash_{\mathrm{F}} \mathrm{Inf}(\mathrm{V})$

(Obtain from DN on T28.)

T28.2 $(\forall x)(\forall y)(x \in N \wedge y \in x \to (\exists x_1)\, x_1 \notin y)$

Proof:

(1)	1. $x \in N \wedge y \in x$	Hyp
(1)	2. $y \neq \mathrm{V}$	1, T28, F1 logic
	3. $(\forall x_1)\, x_1 \in y \to y = \mathrm{V}$	Df. =, T2 logic
(1)	4. $\neg(\forall x_1)\, x_1 \in y$	2, 3 MT
(1)	5. $(\exists x_1)\, x_1 \notin y$	4 CQ, Df. $\notin$
	6. $x \in N \wedge y \in x \to (\exists x_1)\, x_1 \notin y$	1–5 CP
	7. $(\forall x)(\forall y)(x \in N \wedge y \in x \to (\exists x_1)\, x_1 \notin y)$	6 UG×2

We are finally ready to tackle our last Peano postulate.

T29. $(\forall x)(\forall y)(x \in N \wedge y \in N \wedge x' = y' \to x = y)$
(=Peano Postulate 4)

Proof:

(1)	1. $x \in N \wedge y \in N \wedge x' = y'$	Hyp
(2)	2. $z \in x$	Hyp
(1,2)	3. $(\exists x_1)\, x_1 \notin z$	1, 2, T28.2 UI, SL
(1,2)	4. $a \notin z$	3 EI
(5)	5. $x_1 \in z$	Hyp
(1,2,5)	6. $x_1 \neq a$	4, 5, F1 UI, SL
(1,2,5)	7. $x_1 \notin \{a\}$	F2, Df. $\{t\}$, UI, BMT
(5)	8. $x_1 \in z \vee x_1 \in \{a\}$	5 Add
(5)	9. $x_1 \in \{x \mid x \in z \vee x \in \{a\}\}$	8 F2R
(5)	10. $x_1 \in z \cup \{a\}$	9 Df. $\cup$
(1,2,5)	11. $x_1 \in z \cup \{a\} \wedge x_1 \notin \{a\}$	7, 10 Conj
(1,2,5)	12. $x_1 \in \{x \mid x \in z \cup \{a\} \wedge x \notin \{a\}\}$	11 F2R
(1,2,5)	13. $x_1 \in (z \cup \{a\}) - \{a\}$	12 Df. $-$
(1,2)	14. $x_1 \in z \to x_1 \in (z \cup \{a\}) - \{a\}$	5–12 CP
(15)	15. $x_1 \in (z \cup \{a\}) - \{a\}$	Hyp
(15)	16. $x_1 \in \{x \mid x \in z \cup \{a\} \wedge x \notin \{a\}\}$	15 Df. $-$
(15)	17. $x_1 \in z \cup \{a\} \wedge x_1 \notin \{a\}$	16 F2R
(15)	18. $x_1 \in z \cup \{a\}$	17 Simp
(15)	19. $x_1 \in \{x \mid x \in z \vee x \in \{a\}\}$	18 Df. $\cup$
(15)	20. $x_1 \in z \vee x_1 \in \{a\}$	19 F2R
(15)	21. $x_1 \in z$	17, 20 SL
	22. $x_1 \in (z \cup \{a\}) - \{a\} \to x_1 \in z$	15–21 CP
(1,2)	23. $x_1 \in z \leftrightarrow x_1 \in (z \cup \{a\}) - \{a\}$	14, 22 BI
(1,2)	24. $(\forall x_1)(x_1 \in z \leftrightarrow x_1 \in (z \cup \{a\}) - \{a\})$	23 UG
(1,2)	25. $z = (z \cup \{a\}) - \{a\}$	24 Df. =
(1,2)	26. $(z \cup \{a\}) - \{a\} \in x$	2, 25 LL
	27. $a \in \{a\}$	Ref= F2R
	28. $a \in z \vee a \in \{a\}$	27 Add
	29. $a \in z \cup \{a\}$	28 F2R, Df. $\cup$
(1,2)	30. $a \in z \cup \{a\} \wedge (z \cup \{a\}) - \{a\} \in x$	26, 29 Conj
(1,2)	31. $(\exists y)(y \in z \cup \{a\} \wedge (z \cup \{a\} - \{y\} \in x)$	30 EG
(1,2)	32. $z \cup \{a\} \in \{z \mid (\exists y)(y \in z \wedge z - \{y\} \in x)\}$	31 F2R
(1,2)	33. $z \cup \{a\} \in x'$	32 Df. $'$
(1,2)	34. $z \cup \{a\} \in y'$	1, 33 Simp, LL
(1,2)	35. $z \cup \{a\} \in \{x \mid (\exists x_1)(x_1 \in x \wedge x - \{x_1\} \in y)\}$	34 Df. $-$
(1,2)	36. $(\exists x_1)(x_1 \in z \cup \{a\} \wedge (z \cup \{a\}) - \{x_1\} \in y)$	35 F2R
(1,2)	37. $b \in z \cup \{a\} \wedge (z \cup \{a\}) - \{b\} \in y$	36 EI

(continued...)

	38. $y \in N \to (\forall y_1)(\forall y_2)(\forall z)(y_1 \in z \land y_2 \in z \land z - \{y_1\} \in y \to z - \{y_2\} \in y)$	T13 UI
(1)	39. $(\forall y_1)(\forall y_2)(\forall z)(y_1 \in z \land y_2 \in z \land z - \{y_1\} \in y \to z - \{y_2\} \in y)$	1, 38 SL
(1)	40. $b \in z \cup \{a\} \land a \in z \cup \{a\} \land (z \cup \{a\}) - \{b\} \in y \to (z \cup \{a\}) - \{a\} \in y$	39 UI×3
(1,2)	41. $(z \cup \{a\}) - \{a\} \in y$	29, 37, 40 SL
(1,2)	42. $z \in y$	25, 41 LL
(1)	43. $z \in x \to z \in y$	2–42 CP
(1)	44. $z \in y \to z \in x$	CP just like 2–42
(1)	45. $(\forall z)(z \in x \leftrightarrow z \in y)$	43, 44 BI, UG
(1)	46. $x = y$	45 Df. =
	47. $x \in N \land y \in N \land x' = y' \to x = y$	1–46 CP
	48. $(\forall x)(\forall y)(x \in N \land y \in N \land x' = y' \to x = y)$	47 UG×2

To complete the argument that all of Peano arithmetic can be derived in System F, we should proceed to give definitions of addition and multiplication and derive their recursive properties. However, the point is rather moot in light of the problems with the system.

G. Russell's Paradox

System F is inconsistent due to ***Russell's Paradox***: Some sets are members of themselves. The universal set, V, is a member of V. Some sets are not members of themselves, such as Λ or 0. Consider the set W, consisting of all those sets that are not members of themselves. *Problem:* Is it a member of itself? It is if and only if it is not.

$\vdash_F \{x|x \notin x\} \in \{x|x \notin x\} \land \{x|x \notin x\} \notin \{x|x \notin x\}$

Proof:

1. $(\forall y)(y \in \{x|x \notin x\} \leftrightarrow y \notin y)$	F2
2. $\{x|x \notin x\} \in \{x|x \notin x\} \leftrightarrow \{x|x \notin x\} \notin \{x|x \notin x\}$	1 UI
3. $\{x|x \notin x\} \in \{x|x \notin x\} \to \{x|x \notin x\} \notin \{x|x \notin x\}$	2 BE
4. $\{x|x \notin x\} \notin \{x|x \notin x\} \lor \{x|x \notin x\} \notin \{x|x \notin x\}$	3 Df. →
5. $\{x|x \notin x\} \notin \{x|x \notin x\}$	4 Red
6. $\{x|x \notin x\} \in \{x|x \notin x\}$	2,5 BMP
7. $\{x|x \notin x\} \in \{x|x \notin x\} \land \{x|x \notin x\} \notin \{x|x \notin x\}$	5, 6 Conj

From this it follows that every wff of F is a theorem.

$\vdash_F \mathscr{A}$

Proof:

8. $\{x|x \notin x\} \in \{x|x \notin x\} \lor \mathscr{A}$	6 Add
9. $\mathscr{A}$	5, 8 DS

Moreover, this is not the only contradiction derivable in System F. It also succumbs to Cantor's Paradox, the Burali-Forti Paradox, the Russellian Paradox of Relations in Extension, and others.

Obviously, this poses a major obstacle to the use of system F as a foundation for arithmetic.

The question is: what's wrong with F? Is there a way of modifying it or changing it that preserves its method of constructing arithmetic but without giving rise to the contradictions?

HOMEWORK 7

Do the exercise on p. 96 of Hatcher. I.e., prove the inconsistency of the system like F but without any vbto, and whose axioms are (F1) and (F2′): $(\exists y)(\forall x)(x \in y \leftrightarrow \mathscr{A}[x])$, where $\mathscr{A}[x]$ does not contain 'y' free.

X. The Historical Frege/System GG

A. Introduction

Frege's actual system from the 1893 *Grundgesetze der Arithmetik* (GG) was quite a bit different from Hatcher's System F, and interesting in its own right. Some dramatic differences include:

1. Frege's system was ***second-order***, meaning that it had predicate/function variables as well as individual variables.

2. Frege's system was a ***function calculus*** rather than a predicate calculus. Function calculi do not distinguish between terms and formulas, nor between predicates and function letters. Even connectives are thought of as standing for functions.

   For Frege, a formula was thought to stand for something, just a term does, and in particular, a truth value, either the True or the False. A sentence consisted of a name of a truth value along with an illocutionary force marker he called "the judgment stroke".

   The judgment stroke was written "⊢", but this is not to be confused with a metatheoretic sign used to say something is a theorem.

   If $\mathscr{A}$ is a well-formed expression, then $\vdash \mathscr{A}$ is a ***proposition***, which asserts that $\mathscr{A}$ denotes the True.

B. Primitive Function Signs and their Intended Interpretations

Because it is inconsistent, one cannot develop "models" in the usual sense for this system, but one can list what each sign was intended to mean.

(i) $$-\, \mathscr{A} = \begin{cases} \text{the True, if } \mathscr{A} \text{ is the True;} \\ \text{the False, otherwise.} \end{cases}$$

(ii) $$\top\!\!\!-\, \mathscr{A} = \begin{cases} \text{the False, if } \mathscr{A} \text{ is the True;} \\ \text{the True, otherwise.} \end{cases}$$

Hereafter, we write $\neg\mathscr{A}$ instead.

(iii) $$\begin{array}{l} \ulcorner\, \mathscr{B} \\ \llcorner\, \mathscr{A} \end{array} = \begin{cases} \text{the False, if } \mathscr{A} \text{ is the True and } \mathscr{B} \text{ is not the True;} \\ \text{the True, otherwise.} \end{cases}$$

Hereafter, we write $(\mathscr{A} \rightarrow \mathscr{B})$ instead.

(iv) $$(\mathscr{A} = \mathscr{B}) = \begin{cases} \text{the True, if } \mathscr{A} \text{ and } \mathscr{B} \text{ are the same object;} \\ \text{the False, otherwise.} \end{cases}$$

(v) $$\overset{\chi}{\smile}\, \mathscr{A}[\chi] = \begin{cases} \text{the True, if } \mathscr{A}[\chi] \text{ is the True for every } \chi\text{;} \\ \text{the False, otherwise.} \end{cases}$$

Hereafter, we write $(\forall\chi)\,\mathscr{A}[\chi]$ instead.

(vi) $\dot{\alpha}\,\mathscr{A}[\alpha] =$ the value-range of the function $\mathscr{A}[\ ]$ represents, or its extension/class if $\mathscr{A}[\chi]$ is always a truth-value.

Hereafter, we write $\{\chi|\mathscr{A}[\chi]\}$ instead.

(vii) $$\backslash\mathscr{A} = \begin{cases} \text{the sole object falling under the concept whose} \\ \qquad \text{value-range is } \mathscr{A}\text{, if } \mathscr{A} \text{ is such a value-range;} \\ \mathscr{A} \text{ itself, otherwise.} \end{cases}$$

Other common operators ∧, ∨, ∃, etc., could be defined from these as you would expect.

In addition to the primitive constants, the language also contains individual variables x, y, z, etc., function/predicate variables, F, G, H, etc. (of different polyadicities) and even second level function variables written $M_\beta \ldots \beta \ldots$, so that $M_\beta F(\beta)$ represented a second-level variable with its first-level argument.

A second-level variable $M_\beta \ldots \beta \ldots$ can for most intents and purposes be thought of as a *variable quantifier* whose values are, e.g., the universal quantifier, the existential quantifier, and so on.

C. Axioms and Rules of GG

In modern notation, the axioms ("basic laws") were:

(I) $\vdash x \rightarrow (y \rightarrow x)$
(II) $\vdash (\forall x) F(x) \rightarrow F(y)$
$\vdash (\forall F) M_\beta F(\beta) \rightarrow M_\beta G(\beta)$
(III) $\vdash g(x = y) \rightarrow g((\forall F)(F(x) \rightarrow F(y)))$
(IV) $\vdash \neg(—x = \neg y) \rightarrow (—x = —y)$
(V) $\vdash (\{x|F(x)\} = \{y|G(y)\}) = (\forall x)(F(x) = G(x))$
(VI) $\vdash x = \backslash\{y|y = x\}$

Frege's Basic Law (V) is notorious. In predicate calculi analogues of Frege's system, you will often see it written instead:

$$(\{x|F(x)\} = \{y|G(y)\}) \leftrightarrow (\forall x)(F(x) \leftrightarrow G(x))$$

Given that an identity between names of truth-values works much like a biconditional, this version is very similar, though not *quite* equivalent, to Frege's form above.

Inference rules: (i) MP, (ii) Trans, (iii) HS, (iv) Inev, (v) Int, (vi) Amal, or antecedent amalgamation, e.g., from $\vdash \mathscr{A} \rightarrow (\mathscr{A} \rightarrow \mathscr{B})$ infer $\vdash \mathscr{A} \rightarrow \mathscr{B}$, (vii) Hor, or amalgamation of successive horizontal strokes, (viii) UG, which may be applied also to the consequent of the conditional, provided the antecedent does not contain the variable free, (ix) IC, innocuous change of bound variable, and (x) Repl, or variable instantiation/replacement. Variable instantiation requires replacing all occurrences of a free-variable with some expression of the same type, e.g., individual variable with any "complete" complex expression, or function variables with any "incomplete" complex of the right sort.

Stating the variable instantiation rule precisely is difficult and not worth our bother here, but here are some examples:

$\vdash (\forall x) F(x) \rightarrow F(y)$ to $\vdash (\forall x) F(x) \rightarrow F(\{z|z = z\})$
(Replaced 'y' with '$\{z|z = z\}$'.)
$\vdash (\forall x) F(x) \rightarrow F(y)$ to $\vdash (\forall x)(x = x) \rightarrow (y = y)$
(Replaced '$F(\ )$' with '$(\ ) = (\ )$'.)
$\vdash (\forall F) M_\beta F(\beta) \rightarrow M_\beta G(\beta)$ to $\vdash (\forall F) \neg (\forall x) F(x) \rightarrow \neg (\forall x) G(x)$
(Replaced $M_\beta(\ldots \beta \ldots)$ with $\neg (\forall x)(\ldots x \ldots)$.)

An example proof (Hatcher's F1):

1.	$\vdash g(x = y) \rightarrow g((\forall F)(F(x) \rightarrow F(y)))$	(III)
2.	$\vdash (x = y) \rightarrow (\forall F)(F(x) \rightarrow F(y))$	2 Repl
3.	$\vdash (\forall F) M_\beta F(\beta) \rightarrow M_\beta G(\beta)$	(II)
4.	$\vdash (\forall F)(F(x) \rightarrow F(y)) \rightarrow (G(x) \rightarrow G(y))$	3 Repl
5.	$\vdash (x = y) \rightarrow (G(x) \rightarrow G(y))$	2, 4 HS
6.	$\vdash (x = y) \rightarrow (\mathscr{A}[x, x] \rightarrow \mathscr{A}[x, y])$	5 Repl

Frege gave a slightly more complicated (but more useful) definition, but we can very simply define membership as follows:

$t \in u$ for $(\exists f)(u = \{\chi|f(\chi)\} \wedge f(t))$

From this we can derive an analogue of Hatcher's F2:

$$\vdash (\forall x)(x \in \{y|F(y)\} \leftrightarrow F(x))$$

HOMEWORK 8

Prove that the above is a theorem of GG. (Hint, use (V).)

Some interesting additional definitions:

$t \cong u$ for $(\exists R)((\forall x)(x \in t \rightarrow (\exists y)(y \in u \wedge (\forall z)(Rxz \leftrightarrow z = y))) \wedge (\forall x)(x \in u \rightarrow (\exists y)(y \in t \wedge (\forall z)(Rzx \leftrightarrow z = y))))$

$\#(t) = \{x|t \cong x\}$

Inconsistency: Russell's paradox follows just as for System F.

Paper idea and discussion topic: What led the Serpent into Eden? In other words, what is the source of the inconsistency in systems GG or F?

Is it the very notion of the *extension of a concept*, the supposition that one exists for every wff, or Basic Law (V) or F2?

Or is it the impredicative nature of Frege's logic, that is, that it allows quantifiers for functions to range over functions definable only in terms of quantification over functions, thus creating "indefinitely extensible concepts"?

Consider comparing the views, e.g., of Boolos (*Logic, Logic and Logic*, chaps. 11 and 14) with Dummett (*Frege: Philosophy of Mathematics*, especially last chapter), and/or others.

XI. Type-Theory Generally

"The theory of types" is not one theory, but several. Indeed, there are two broad types of type theory. The first are typed set theories. The second are typed higher-order set theories. Even within each category, there are several varieties. What they have in common are syntactic rules that make distinctions between different kinds of variables (and/or constants), and restrictions to the effect one type of variable (and/or constant) may not meaningfully replace another of a different type.

We begin our examination with so-called *simple theories of types*. The simple theory of types for sets, when interpreted standardly, has distinct kinds of variables for the following domains:

Type 0: individuals (entities)
Type 1: sets whose members are all individuals
Type 2: sets whose members are all sets of type 1
Type 3: sets whose members are all sets of type 2
And so on.

A simple theory of types for a higher-order language would have a similar hierarchy: variables for individuals, variables for properties/relations applicable to individuals, variables for properties/relations applicable to such properties/relations, and so on. We begin our study however, with type-theoretical set theories.

XII. The System ST (Simple Types)

A. Syntax

Definition: A ***variable*** is any of the letters x, y, or z written with a numerical superscript n (where $n \geq 0$), and with or without a numerical subscript. (If a subscript is left off, it should be taken as a 0.)

The subscripts as usual, are simply used to guarantee an infinite supply of variables. The superscripts indicate the ***type*** of a variable.

Examples:

a) "x_1^0" is a variable of type 0 (a variable for individuals).
b) "y_4^1" is a variable of type 1 (a variable for sets of individuals)
c) "x_2^3" is a variable of type 3 (a variable for sets of sets of sets of individuals)

We continue to use the same *vbto* for set abstraction as before. However, a term of the form $\{\chi|\mathscr{A}[\chi]\}$ will be of type $n+1$ where n is the type of the variable χ (as given by its superscript).

A string of symbols of the form "$t \in u$" is a wff only if the type of u is one greater than the type t; otherwise the syntax is the same as those of system F.

Hence "$x^0 \in x^0$" is not well-formed, and so neither is "$\{x^0|x^0 \notin x^0\}$". However "$x^0 \in y^1$" is fine.

We introduce the signs $\rightarrow$, $\leftrightarrow$, $\wedge$, $\notin$, $\exists$, etc. by definition as one would expect.

However, we define identity thusly:

$t = u$ for $(\forall\chi)(t \in \chi \leftrightarrow u \in \chi)$, where χ is the first variable of type $n+1$, where n is the type of t and u, not ocurring free in either t or u.

Note that $t = u$ is only well-formed if t and u have the same type.

We cannot define $=$ as we did for system F, since entities of type 0 are not sets and hence cannot have members, much less the "same members" as each other.

We define $t \neq u$, $t \cup u$, $t \cap u$, $t \subseteq u$, $t - u$, $\{t\}$, $\{t_1, \ldots, t_n\}$ similarly to how they're defined for system F. Notice, again, however that some of these are well-formed only if u and t share the same type.

V^{n+1} is defined as $\{x^n|x^n = x^n\}$.
Notice there is a distinct universal class, of type $n+1$ for every type n.

Λ^{n+1} is defined as $\{x^n|x^n \neq x^n\}$.

$\langle t, u\rangle$ is defined as $\{\{t\}, \{t, u\}\}$.

B. Axiomatics

System ST has logical axioms corresponding to the logical axioms of first-order theories. We note only that for well-formed instances of the axiom schema $(\forall \chi)\mathscr{A}[\chi] \rightarrow \mathscr{A}[t]$, the type t must match that of the variable χ. (Of course, were this not the case, we would not even have a wff.) The inference rules are the same, and all the same derived rules carry over.

The ***proper axioms*** of ST are the following:

ST1. $(\forall \chi^n)(\chi^n \in \{y^n | \mathscr{A}[y^n]\} \leftrightarrow \mathscr{A}[\chi^n])$, provided that χ^n does occur bound in $\mathscr{A}[y^n]$.

ST2. $(\forall x^{n+1})(\forall y^{n+1})((\forall z^n)(z^n \in x^{n+1} \leftrightarrow z^n \in y^{n+1}) \rightarrow x^{n+1} = y^{n+1})$, for every type n.

ST3. $(\exists x^3)((\forall x^0)\langle x^0, x^0\rangle \notin x^3 \wedge (\forall x^0)(\exists y^0)\langle x^0, y^0\rangle \in x^3 \wedge$
$(\forall x^0)(\forall y^0)(\forall z^0)(\langle x^0, y^0\rangle \in x^3 \wedge \langle y^0, z^0\rangle \in x^3 \rightarrow \langle x^0, z^0\rangle \in x^3))$

(The need and import of ST3 will be discussed below.)

C. Basic Results

Derived rule from ST1 (ST1R): Same as F2R, except obeying type-restrictions. Proof the same.

T1. $(\forall x^n)\, x^n = x^n$

Proof:

1. $x^n \in y^{n+1} \leftrightarrow x^n \in y^{n+1}$	Taut
2. $(\forall y^{n+1})(x^n \in y^{n+1} \leftrightarrow x^n \in y^{n+1})$	1 UG
3. $x^n = x^n$	2 Df. =
4. $(\forall x^n)\, x^n = x^n$	3 UG

T2. $(\forall x^n)(\forall y^n)(x^n = y^n \rightarrow (\mathscr{A}[x^n, x^n] \rightarrow \mathscr{A}[x^n, y^n]))$

Proof:

(1)	1. $x^n = y^n$	Hyp
(1)	2. $(\forall z^{n+1})(x^n \in z^{n+1} \leftrightarrow y^n \in z^{n+1})$	1 Df. =
(3)	3, $\mathscr{A}[z^n, x^n]$	Hyp
(3)	4. $x^n \in \{x^n | \mathscr{A}[z^n, x^n]\}$	3 ST1R
(1)	5. $x^n \in \{x^n | \mathscr{A}[z^n, x^n]\} \leftrightarrow y^n \in \{x^n | \mathscr{A}[z^n, x^n]\}$	2 UI
(1,3)	6. $y^n \in \{x^n | \mathscr{A}[z^n, x^n]\}$	4, 5 BMP
(1,3)	7. $\mathscr{A}[z^n, y^n]$	6 ST1R
(1)	8. $\mathscr{A}[z^n, x^n] \rightarrow \mathscr{A}[z^n, y^n]$	3–7 CP
	9. $x^n = y^n \rightarrow (\mathscr{A}[z^n, x^n] \rightarrow \mathscr{A}[z^n, y^n])$	1–8 CP
	10. $x^n = y^n \rightarrow (\mathscr{A}[x^n, x^n] \rightarrow \mathscr{A}[x^n, y^n])$	9 UG, UI
	11. $(\forall x^n)(\forall y^n)(x^n = y^n \rightarrow (\mathscr{A}[x^n, x^n] \rightarrow \mathscr{A}[x^n, y^n]))$	10 UG×2

From T1 and T2 we get, in essence, that ST is a theory with identity, and so we can use all associated derived rules.

The following have proofs that are either obvious or directly parallel to the analogous results in system F:

T3. $\vdash_{ST} (\forall x^n)\, x^n \in V^{n+1}$

T4. $\vdash_{ST} (\forall x^n)\, x^n \notin \Lambda^{n+1}$

T5. $\vdash_{ST} (\forall y^{n+1})((\forall x^n)\, x^n \in y^{n+1} \rightarrow y^{n+1} = V^{n+1})$

T6. $\vdash_{ST} (\forall y^{n+1})((\forall x^n)\, x^n \notin y^{n+1} \rightarrow y^{n+1} = \Lambda^{n+1})$

T7. $\vdash_{ST} (\forall x^{n+1})(\forall y^{n+1})(x^{n+1} \subseteq y^{n+1} \wedge y^{n+1} \subseteq x^{n+1} \rightarrow x^{n+1} = y^{n+1})$

T8. $\vdash_{ST} (\forall x^{n+1})(\forall y^{n+1})(x^{n+1} \subseteq x^{n+1} \cup y^{n+1})$

T9. $\vdash_{ST} (\forall x^{n+1})(\forall y^{n+1})(x^{n+1} \cap y^{n+1} \subseteq x^{n+1})$

T10. $\vdash_{ST} (\forall x^n)(\forall y^n)(x^n \in \{y^n\} \leftrightarrow x^n = y^n)$

T11. $\vdash_{ST} (\forall x^n)(\forall y^n)(\forall z^n)(x^n \in \{y^n, z^n\} \leftrightarrow x^n = y^n \vee x^n = z^n)$

T12. $\vdash_{ST} (\forall x^{n+1})(\forall y^n)(y^n \in x^{n+1} \rightarrow (x^{n+1} - \{y^n\}) \cup \{y^n\} = x^{n+1})$

T13. $\vdash_{ST} (\forall x^{n+1})(\forall y^n)(y^n \notin x^{n+1} \rightarrow (x^{n+1} \cup \{y^n\}) - \{y^n\} = x^{n+1})$

Another important result:

T14. $\vdash_{ST} (\forall x_1^n)(\forall x_2^n)(\forall y_1^n)(\forall y_2^n)(\langle x_1^n, y_1^n\rangle = \langle x_2^n, y_2^n\rangle \leftrightarrow$
$x_1^n = x_2^n \wedge y_1^n = y_2^n)$

Proof of right to left half of biconditional is an obvious consequence of LL. Proof of left to right half of biconditional is straightforward but tedious.

D. Development of Mathematics

Almost all the methodology we examined for capturing natural number theory in System F can be carried over to ST. However, notive that we have distinct numbers in distinct types starting with type 2. In type 2, we have the number 2 as the set of all two-membered sets of individuals. In type 3, we have the number 2 as the set of all two-membered sets of sets of individuals, and so on.

Definitions:

0^{n+2} for $\{\Lambda^{n+1}\}$

t' for $\{\chi | (\exists y)(y \in \chi \wedge \chi - \{y\} \in t)\}$, where χ is the first variable one type below t not found free in t, and y is the first variable two types below t not found free in t.

N^{n+3} for $\{x^{n+2} | (\forall y^{n+3})(0^{n+2} \in y^{n+3} \wedge (\forall z^{n+2})(z^{n+2} \in y^{n+3} \rightarrow z^{n+2\prime} \in y^{n+3}) \rightarrow x^{n+2} \in y^{n+3})\}$

Below, unless specified otherwise, we shall use "0" for "0^2" and "N" for "N^3".

The following results are derived exactly as for system F:

T15. $\vdash_{ST} 0 \in N$ (=Peano Postulate 1)

T16. $\vdash_{ST} (\forall x^2)(0 \neq x^{2\prime})$ (=Peano Postulate 3)

T17. $\vdash_{ST} (\forall x^2)(x^2 \in N \rightarrow x^{2\prime} \in N)$ (=Peano Postulate 2)

T18. $\vdash_{ST} (\forall x^3)(0 \in x^3 \wedge (\forall y^2)(y^2 \in x^3 \wedge y^2 \in N \rightarrow y^{2\prime} \in x^3) \rightarrow N \subseteq x^3)$

T19. $\mathscr{A}[0] \wedge (\forall x^2)(x^2 \in N \wedge \mathscr{A}[x^2] \rightarrow \mathscr{A}[x^{2\prime}]) \rightarrow (\forall x^2)(x^2 \in N \rightarrow \mathscr{A}[x^2])$ (=Peano Postulate 5)

The tricky one is, again, the fourth Peano Postulate, which requires that we prove that V^1 is infinite. Here, however, the problem is more difficult because we need to prove an infinity of *individuals*, and not an infinity of sets. Indeed, the method used for System F to establish an infinity will not work, because $x \cup \{x\}$ is not even well-formed in type theory, and so the set ω cannot be defined.

ST3 has been added to the system for the sole purpose of proving an infinity of individuals. The proof is quite complicated.

T20. $\vdash_{ST} \Lambda^2 \notin N$

Proof:

1. $(\exists x^3)((\forall x^0)\langle x^0,x^0\rangle \notin x^3 \wedge (\forall x^0)(\exists y^0)\langle x^0,y^0\rangle \in x^3 \wedge (\forall x^0)(\forall y^0)(\forall z^0)(\langle x^0,y^0\rangle \in x^3 \wedge \langle y^0,z^0\rangle \in x^3 \rightarrow \langle x^0,z^0\rangle \in x^3))$	ST3
2. $(\forall x^0)\langle x^0,x^0\rangle \notin r^3 \wedge (\forall x^0)(\exists y^0)\langle x^0,y^0\rangle \in r^3 \wedge (\forall x^0)(\forall y^0)(\forall z^0)(\langle x^0,y^0\rangle \in r^3 \wedge \langle y^0,z^0\rangle \in r^3 \rightarrow \langle x^0,z^0\rangle \in r^3)$	1 EI
3. $(\forall x^0)\langle x^0,x^0\rangle \notin r^3$	2 Simp
4. $(\forall x^0)(\exists y^0)\langle x^0,y^0\rangle \in r^3$	2 Simp
5. $(\forall x^0)(\forall y^0)(\forall z^0)(\langle x^0,y^0\rangle \in r^3 \wedge \langle y^0,z^0\rangle \in r^3 \rightarrow \langle x^0,z^0\rangle \in r^3)$	2 Simp
6. $\langle x^0,y^0\rangle \in r^3 \wedge \langle y^0,x^0\rangle \in r^3 \rightarrow \langle x^0,x^0\rangle \in r^3$	5 UI
7. $\langle x^0,x^0\rangle \notin r^3$	3 UI
8. $\langle x^0,y^0\rangle \in r^3 \rightarrow \langle y^0,x^0\rangle \notin r^3$	6, 7 SL
9. $(\forall x^0)(\forall y^0)(\langle x^0,y^0\rangle \in r^3 \rightarrow \langle y^0,x^0\rangle \notin r^3)$	8 UG×2

It is convenient to pause to introduce an abbreviation:
C^2 for $\{z^1 | z^1 = \Lambda^1 \vee (\exists y^0)(y^0 \in z^1 \wedge (\forall x^0)(x^0 \in z^1 \wedge x^0 \neq y^0 \rightarrow \langle x^0,y^0\rangle \in r^3))\}$
I now prove inductively that every natural number has a member of C^2 in it.

	10. $\Lambda^1 = \Lambda^1$	T1 UI
	11. $\Lambda^1 = \Lambda^1 \vee (\exists y^0)(y^0 \in \Lambda^1 \wedge (\forall x^0)(x^0 \in \Lambda^1 \wedge x^0 \neq y^0 \rightarrow \langle x^0,y^0\rangle \in r^3))$	10 Add
	12. $\Lambda^1 \in C^2$	11 ST1R, Df. C^2
	13. $\Lambda^1 \in 0$	Df. 0, Ref=, T10, UI, BMP
	14. $\Lambda^1 \in C^2 \wedge \Lambda^1 \in 0$	12, 13 Conj
	15. $(\exists y^1)(y^1 \in C^2 \wedge y^1 \in 0)$	14 EG
(16)	16. $x^2 \in N \wedge (\exists y^1)(y^1 \in C^2 \wedge y^1 \in x^2)$	Hyp
(16)	17. $(\exists y^1)(y^1 \in C^2 \wedge y^1 \in x^2)$	16 Simp
(16)	18. $b^1 \in C^2 \wedge b^1 \in x^2$	17 EI
(16)	19. $b^1 \in \{z^1 \mid z^1 = \Lambda^1 \vee (\exists y^0)(y^0 \in z^1 \wedge (\forall x^0)(x^0 \in z^1 \wedge x^0 \neq y^0 \rightarrow \langle x^0,y^0\rangle \in r^3))\}$	18 Simp, Df. C^2
(16)	20. $b^1 = \Lambda^1 \vee (\exists y^0)(y^0 \in b^1 \wedge (\forall x^0)(x^0 \in b^1 \wedge x^0 \neq y^0 \rightarrow \langle x^0,y^0\rangle \in r^3))$	19 ST1R
(21)	21. $b^1 = \Lambda^1$	Hyp
	22. $z^0 \notin \Lambda^1$	T4 UI
(21)	23. $z^0 \notin b^1$	21, 22 LL
(21)	24. $(b^1 \cup \{z^0\}) - \{z^0\} = b^1$	23, T13 UI, MP
	25. $z^0 \in \{z^0\}$	Ref=, T10 UI, BMP
	26. $z^0 \in b^1 \cup \{z^0\}$	25 Add, ST1R, Df. $\cup$
(16,21)	27. $(b^1 \cup \{z^0\}) - \{z^0\} \in x^2$	18, 24 Simp, LL
(16,21)	28. $z^0 \in b^1 \cup \{z^0\} \wedge (b^1 \cup \{z^0\}) - \{z^0\} \in x^2$	26, 27 Conj
(16,21)	29. $(\exists x^0)(x^0 \in b^1 \cup \{z^0\} \wedge (b^1 \cup \{z^0\}) - \{x^0\} \in x^2)$	29 EG
(16,21)	30. $b^1 \cup \{z^0\} \in x^{2\prime}$	29 ST1R, Def. $'$
(16)	31. $x^{2\prime} \in N$	16, T17 Simp, UI, MP
(32)	32. $x^0 \in b^1 \cup \{z^0\} \wedge x^0 \neq z^0$	Hyp
(32)	33. $x^0 \in b^1 \vee x^0 \in \{z^0\}$	32 Simp, Df. $\cup$, ST1R

(32)	34. $x^0 \notin \{z^0\}$	32, T10 Simp, UI, BMT
(32)	35. $x^0 \in b^1$	33, 34 DS
(21,32)	36. $x^0 \in \Lambda^1$	21, 35 LL
	37. $x^0 \notin \Lambda^1$	T4 UI
(21,32)	38. $x^0 \in \Lambda^1 \wedge x^0 \notin \Lambda^1$	36, 37 Conj
(21)	39. $\neg(x^0 \in b^1 \cup \{z^0\} \wedge x^0 \neq z^0)$	32–38 RAA
(21)	40. $x^0 \in b^1 \cup \{z^0\} \wedge x^0 \neq z^0 \to \langle x^0, z^0 \rangle \in r^3$	39 FA
(21)	41. $(\forall x^0)(x^0 \in b^1 \cup \{z^0\} \wedge x^0 \neq z^0 \to \langle x^0, z^0 \rangle \in r^3)$	40 UG
(21)	42. $z^0 \in b^1 \cup \{z^0\} \wedge (\forall x^0)(x^0 \in b^1 \cup \{z^0\} \wedge x^0 \neq z^0 \to \langle x^0, z^0 \rangle \in r^3)$	26, 41 Conj
(21)	43. $(\exists y^0)(y^0 \in b^1 \cup \{z^0\} \wedge (\forall x^0)(x^0 \in b^1 \cup \{z^0\} \wedge x^0 \neq y^0 \to \langle x^0, y^0 \rangle \in r^3))$	42 EG
(21)	44. $b^1 \cup \{z^0\} = \Lambda^1 \vee (\exists y^0)(y^0 \in b^1 \cup \{z^0\} \wedge (\forall x^0)(x^0 \in b^1 \cup \{z^0\} \wedge x^0 \neq y^0 \to \langle x^0, y^0 \rangle \in r^3))$	43 Add
(21)	45. $b^1 \cup \{z^0\} \in C^2$	44 ST1R, Df. C^2
(16,21)	46. $b^1 \cup \{z^0\} \in C^2 \wedge b^1 \cup \{z^0\} \in x^{2\prime}$	30, 45 Conj
(16,21)	47. $(\exists y^1)(y^1 \in C^2 \wedge y^1 \in x^{2\prime})$	46 EG
(16)	48. $b^1 = \Lambda^1 \to (\exists y^1)(y^1 \in y^1 \in C^2 \wedge y^1 \in x^{2\prime})$	21–47 CP
(49)	49. $(\exists y^0)(y^0 \in b^1 \wedge (\forall x^0)(x^0 \in b^1 \wedge x^0 \neq y^0 \to \langle x^0, y^0 \rangle \in r^3))$	Hyp
(49)	50. $d^0 \in b^1 \wedge (\forall x^0)(x^0 \in b^1 \wedge x^0 \neq d^0 \to \langle x^0, d^0 \rangle \in r^3)$	49 EI
(49)	51. $d^0 \in b^1$	50 Simp
(49)	52. $(\forall x^0)(x^0 \in b^1 \wedge x^0 \neq d^0 \to \langle x^0, d^0 \rangle \in r^3)$	50 Simp
	53. $(\exists y^0)\langle d^0, y^0 \rangle \in r^3$	4 UI
	54. $\langle d^0, e^0 \rangle \in r^3$	53 EI
	55. $e^0 \neq d^0$	3, 54, T2 SL
	56. $\langle e^0, d^0 \rangle \notin r^3$	9, 54 UI, MP
(49)	57. $e^0 \in b^1 \wedge e^0 \neq d^0 \to \langle e^0, d^0 \rangle \in r^3$	52 UI
(49)	58. $e^0 \notin b^1$	55, 56, 57 SL
(49)	59. $(b^1 \cup \{e^0\}) - \{e^0\} = b^1$	58, T13 UI×2, MP
	60. $e^0 \in \{e^0\}$	Ref=, T10 UI, BMP
	61. $e^0 \in b^1 \cup \{e^0\}$	60 Add, ST1R, Df. $\cup$
(16,49)	62. $(b^1 \cup \{e^0\}) - \{e^0\} \in x^2$	18, 59 Simp, LL
(16,49)	63. $e^0 \in b^1 \cup \{e^0\} \wedge (b^1 \cup \{e^0\}) - \{e^0\} \in x^2$	61, 62 Conj
(16,49)	64. $(\exists x^0)(x^0 \in b^1 \cup \{e^0\} \wedge (b^1 \cup \{e^0\}) - \{x^0\} \in x^2)$	63 EG
(16,49)	65. $b^1 \cup \{e^0\} \in x^{2\prime}$	64 ST1R, Df. $'$
(66)	66. $x^0 \in b^1 \cup \{e^0\} \wedge x^0 \neq e^0$	Hyp
(66)	67. $x^0 \in b^1 \vee x^0 \in \{e^0\}$	66 Simp, Df. $\cup$, ST1R
(66)	68. $x^0 \notin \{e^0\}$	66 Simp, T10, UI, BMT
(66)	69. $x^0 \in b^1$	67, 68 DS
(70)	70. $x^0 = d^0$	Hyp

(70)	71. $\langle x^0, e^0\rangle \in r^3$	54, 70 LL
	72. $x^0 = d^0 \to \langle x^0, e^0\rangle \in r^3$	70–71 CP
(73)	73. $x^0 \neq d^0$	Hyp
(49)	74. $x^0 \in b^1 \wedge x^0 \neq d^0 \to \langle x^0, d^0\rangle \in r^3$	52 UI
(49,66,73)	75. $\langle x^0, d^0\rangle \in r^3$	69, 73, 74 Conj, MP
(49,66,73)	76. $\langle x^0, e^0\rangle \in r^3$	5, 54, 75 UI×3, Conj, MP
(49,66)	77. $x^0 \neq d^0 \to \langle x^0, e^0\rangle \in r^3$	73–76 CP
(49,66)	78. $\langle x^0, e^0\rangle \in r^3$	72, 77 Inev
(49)	79. $x^0 \in b^1 \cup \{e^0\} \wedge x^0 \neq e^0 \to \langle x^0, e^0\rangle \in r^3$	66–78 CP
(49)	80. $(\forall x^0)(x^0 \in b^1 \cup \{e^0\} \wedge x^0 \neq e^0 \to \langle x^0, e^0\rangle \in r^3)$	79 UG
(49)	81. $e^0 \in b^1 \cup \{e^0\} \wedge (\forall x^0)(x^0 \in b^1 \cup \{e^0\} \wedge x^0 \neq e^0 \to \langle x^0, e^0\rangle \in r^3)$	61, 80 Conj
(49)	82. $(\exists y^0)(y^0 \in b^1 \cup \{e^0\} \wedge (\forall x^0)(x^0 \in b^1 \cup \{e^0\} \wedge x^0 \neq y^0 \to \langle x^0, y^0\rangle \in r^3))$	81 EG
(49)	83. $b^1 \cup \{e^0\} = \Lambda^1 \vee (\exists y^0)(y^0 \in b^1 \cup \{e^0\} \wedge (\forall x^0)(x^0 \in b^1 \cup \{e^0\} \wedge x^0 \neq y^0 \to \langle x^0, y^0\rangle \in r^3))$	82 Add
(49)	84. $b^1 \cup \{e^0\} \in C^2$	83 ST1R, Df. C^2
(16,49)	85. $b^1 \cup \{e^0\} \in C^2 \wedge b^1 \cup \{e^0\} \in x^{2\prime}$	65, 84 Conj
(16,49)	86. $(\exists y^1)(y^1 \in C^2 \wedge y^1 \in x^{2\prime})$	85 EG
(16)	87. $(\exists y^0)(y^0 \in b^1 \wedge (\forall x^0)(x^0 \in b^1 \wedge x^0 \neq y^0 \to \langle x^0, y^0\rangle \in r^3)) \to (\exists y^1)(y^1 \in C^2 \wedge y^1 \in x^{2\prime})$	49–86 CP
(16)	88. $(\exists y^1)(y^1 \in C^2 \wedge y^1 \in x^{2\prime})$	20, 48, 87 SL
	89. $x^2 \in N \wedge (\exists y^1)(y^1 \in C^2 \wedge y^1 \in x^2) \to (\exists y^1)(y^1 \in C^2 \wedge y^1 \in x^{2\prime})$	16–88 CP
	90. $(\forall x^2)(x^2 \in N \wedge (\exists y^1)(y^1 \in C^2 \wedge y^1 \in x^2) \to (\exists y^1)(y^1 \in C^2 \wedge y^1 \in x^{2\prime}))$	89 UG
	91. $(\forall x^2)(x^2 \in N \to (\exists y^1)(y^1 \in C^2 \wedge y^1 \in x^2))$	15, 90, T19 Conj, MP
	92. $\Lambda^2 \in N \to (\exists y^1)(y^1 \in C^2 \wedge y^1 \in \Lambda^2)$	91 UI
	93. $y^1 \notin \Lambda^2$	T4 UI
	94. $\neg(y^1 \in C^2 \wedge y^1 \in \Lambda^2)$	93 SL
	95. $(\forall y^1)\neg(y^1 \in C^2 \wedge y^1 \in \Lambda^2)$	94 UG
	96. $\neg(\exists y^1)(y^1 \in C^2 \wedge y^1 \in \Lambda^2)$	95 DN, Df. $\exists$
	97. $\Lambda^2 \notin N$	92, 96 MT

To make use of T20 we also need:

T21. $\vdash_{\mathrm{ST}} (\forall x^2)(x^2 \in N \to (\forall x^1)(\forall y^0)(\forall z^0)(y^0 \in x^1 \wedge z^0 \in x^1 \wedge x^1 - \{y^0\} \in x^2 \to x^1 - \{z^0\} \in x^2))$

T22. $\vdash_{\mathrm{ST}} (\forall x^2)(x^2 \in N \to (\forall y^1)(\forall z^1)(y^1 \in x^2 \wedge z^1 \in x^2 \wedge y^1 \subseteq z^1 \to y^1 = z^1))$

(Proofs of these are exactly parallel to those for System F.) From these and T20 we get:

T23. $\vdash_{\mathrm{ST}} (\forall x^2)\neg(\mathrm{V}^1 \in x^2 \wedge x^2 \in N)$

T24. $\vdash_{\mathrm{ST}} (\forall x^2)(\forall y^1)(x^2 \in N \wedge y^1 \in x^2 \to (\exists x^0)\, x^0 \notin y^1)$

T25. $\vdash_{\mathrm{ST}} (\forall x^2)(\forall y^2)(x^2 \in N \wedge y^2 \in N \wedge x^{2\prime} = y^{2\prime} \to x^2 = y^2)$ (=Peano Postulate 4)

Homework 9

Prove T23. (Hint: the proof is very similar to the proof of T28 for System F. Use T20 and T22.)

E. Evaluation

System ST is, as far as we know, consistent. One of the problems with establishing the consistency of any given set-theory is that the typical way to establish the consistency of a theory is to show that the theory has (one or more) models. However, models are themselves usually defined set-theoretically. (E.g., an ordered pair $\langle D, g\rangle$ where D is a domain of quantification and g is an assignment function, both understood as sets.) What theories we can construct models for in the metalanguage depends on what sort of set theory we employ in the metalanguage. Obviously, the proof of consistency will be question begging if we use the very same kind of set-theory in the metalanguage as in the object-language.

Usually, the best we can hope for is a *relative consistency* proof: a proof that models for a given set theory can be constructed in a different set theory. Often this simply takes the form of a translation matrix from one theory to another, along with a proof that if a contradiction was provable in the one theory, a contradiction would also be provable in the other.

System ST is consistent relative to weak Zermelo set theory (System Z). Before discussing System Z, we cannot actually go over the proof. While an absolute proof of the consistency of Z has not been found either, no one (so far as I know) has any serious doubts about it.

Quite a lot of mathematics can be captured in ST. Even more can be captured if we add the axiom of choice (discussed in our next unit).

Still there are philosophical issues regarding System ST:

(1) What is the philosophical motivation for thinking that "$x \in x$" and "$x \notin x$" are *meaningless*? Even if one thinks there is no such set as the set of all sets not members of themselves, is the very question as to its existence nonsense?

A related worry: can the theory itself be stated without violating its own strictures?

(2) Is the system restrictive enough? It still allows for impredicativity, i.e., sets that are defined in terms of quantification over a range that includes itself. Consider, e.g.,

$$E^1 = \{x^0 | (\forall y^1)\, x^0 \in y^1\}$$

Intuitively, this is the set of all individuals that are members of every set. Notive that one of the sets an individual would have to be a member of, in order to be a member of E^1 is E^1 itself. Does this set have well-defined membership conditions? (Notice, however, that it is provable that $E^1 = \Lambda^1$.)

(3) Is the axiom of infinity well-motivated, or *ad hoc*? What if it were dropped?

XIII. Higher-Order Simple Type Theories (HOSTs)

A. Syntax

Definition: A ***type symbol*** is defined recursively as follows: (i) o is a type-symbol; (ii) if $\tau_1, \ldots, \tau_n$ are type symbols, then $(\tau_1, \ldots, \tau_n)$ is a type symbol, (iii) nothing that cannot be constructed from repeated applications of (i) and (ii) is a type symbol.

Examples: (with intended meaning):

o	is the type of *individuals*
(o)	is the type of *properties* applicable to individuals
(o, o)	is the type of *dyadic relations* applicable to individuals
((o))	is the type of *properties applicable to properties* of individuals
(o,(o))	is the type of dyadic relations between individuals and properties of individuals

The use of the words "properties" and "relations" here may be controversial. It may be best to stick to the linguistic level. The symbol "o" is the type-symbol for individual terms; the type symbol "(o)" is the type for *monadic predicates* applied to individual terms, etc.

Definition: A ***variable*** is any lower or uppercase letter between $f, \ldots, z$, written with or without a numerical subscript, and with a type symbol as superscript.

Examples: "$f^{(o)}$", "x_2^o", "$R_1^{(o,o,o)}$".

Definition: A ***constant*** is any lower or uppercase letter between $a, \ldots, e$, written with or without a numerical subscript, and with a type symbol as superscript.

Different higher-order languages may use different sets of constants, or different subsets of the type-symbols.

Definition: A ***well-formed expression (wfe)*** is defined recursively as follows:

(i) A variable or constant is a wfe of the type given by its superscript.
(ii) if $\mathscr{P}$ is a wfe of the type $(\tau_1, \ldots, \tau_n)$, and $\mathscr{A}_1, \ldots, \mathscr{A}_n$ are wfes of types $\tau_1, \ldots, \tau_n$, respectively, then $\mathscr{P}(\mathscr{A}_1, \ldots, \mathscr{A}_n)$ is a type-less wfe;
(iii) if $\mathscr{A}$ and $\mathscr{B}$ are type-less wfes, then $(\mathscr{A} \vee \mathscr{B})$ is a type-less wfe;
(iv) if $\mathscr{A}$ is a type-less wfe, then $\neg\mathscr{A}$ is a type-less wfe;
(v) if $\mathscr{A}$ is a type-less wfe, and χ is a variable, then $(\forall \chi)\,\mathscr{A}$ is a type-less wfe;
(vi) if $\chi_1, \ldots, \chi_n$ are *distinct* variables of types $\tau_1, \ldots, \tau_n$, respectively, and $\mathscr{A}$ is a type-less wfe, then $[\lambda \chi_1 \ldots \chi_n\, \mathscr{A}]$ is a wfe of type $(\tau_1, \ldots, \tau_n)$ (and all occurrences of those variables are considered bound in that context);
(vii) nothing that cannot be constructed from repeated applications of the above is a wfe.

Definitions: A type-less wfe is called a ***well-formed formula (wff)***. A wfe that is not a wff is called a ***term***.

Terms of the form "$[\lambda \chi \ldots \chi \ldots]$" might be read "the property of being an χ such that $\ldots \chi \ldots$". Terms of the form "$[\lambda \chi y \ldots \chi \ldots y \ldots]$" might be read "the relation that holds between χ and y when $\ldots \chi \ldots y \ldots$". Again, however, this phrasing may be controversial.

I use the notation "$[\lambda x \ldots x \ldots]$" where Hatcher would write simply "$\ldots \hat{x} \ldots$". Hatcher's notation is older, and now somewhat outdated, although the two notations are historically connected.

$\wedge$, $\rightarrow$, $\leftrightarrow$, $(\exists \chi)$ are defined as one would expect.

Typically, $t = u$ is defined as $(\forall \chi)(\chi(t) \leftrightarrow \chi(u))$ where t and u are of the same type, τ, and χ is the first variable of type (τ) not occurring free in either t or u.

B. Formulation

Different HOSTs have different axioms. However, all standard theories have as axioms (or theorems) all truth-table tautologies, and the following schemata:

- $(\forall \chi)\,\mathscr{A}[\chi] \rightarrow \mathscr{A}[t]$, where χ is a variable of any type, t any term of the same type, and no free variables of t become bound in the context $\mathscr{A}[t]$.
- $(\forall \chi)(\mathscr{B} \rightarrow \mathscr{A}[\chi]) \rightarrow (\mathscr{B} \rightarrow (\forall \chi)\,\mathscr{A}[\chi])$, where χ is a variable of any type, and $\mathscr{B}$ does not contain χ free.
- $(\forall y_1) \ldots (\forall y_n)([\lambda \chi_1 \ldots \chi_n\, \mathscr{A}[\chi_1, \ldots, \chi_n]](y_1, \ldots, y_n) \leftrightarrow \mathscr{A}[y_1, \ldots, y_n])$, where $\chi_1, \ldots, \chi_n, y_1, \ldots, y_n$ are distinct variables and each χ_i matches y_i in type.

The inference rules are MP and UG (applicable to any type variable).

So in addition to the normal first-order instances of the above, we also have, e.g.
$(\forall F^{(o)})\,(\exists x^o)\,F^{(o)}(x^o) \rightarrow (\exists x^o)[\lambda y^o\; y^o = y^o](x^o)$
$(\forall R^{(o,o)})(F^{(o)}(x^o) \rightarrow R^{(o,o)}(x^o, x^o)) \rightarrow (F^{(o)}(x^o) \rightarrow (\forall R^{(o,o)})\,R^{(o,o)}(x^o, x^o))$
$(\forall G^{(o)})([\lambda F^{(o)}\, F^{(o)}(x^o)](G^{(o)}) \leftrightarrow G^{(o)}(x^o))$

Derived rules: *lambda conversion* (λ-cnv): where no free variable in $t_1, \ldots, t_n$ becomes bound in the context $\mathscr{A}[t_1, \ldots, t_n]$:
$[\lambda \chi_1, \ldots, \chi_n\, \mathscr{A}[\chi_1, \ldots, \chi_n]](t_1, \ldots, t_n) \vdash \mathscr{A}[t_1, \ldots, t_n]$
$\mathscr{A}[t_1, \ldots, t_n] \vdash [\lambda \chi_1, \ldots, \chi_n\, \mathscr{A}[\chi_1, \ldots, \chi_n]](t_1, \ldots, t_n)$

Definition: The ***pure higher-order predicate calculus (HOPC)*** is the HOST whose only axioms are instances of the above schemata in the higher-order language that contains *no* constants.

Definition: Begin counting at 0, and parse a type symbol, adding 1 for every left parenthesis, (, and substracting one for every right parenthesis,

). The highest number reached while parsing in the ***order*** of the type symbol.

Definition: A ***theory of order*** n is just like a HOST, except eliminating from the syntax all variables of order n or greater and all constants of order $n+1$ or greater.

We adopt the following convention: Type symbols may be left off variables in later occurrences in the same wff, provided that the same letter and subscript are not also used in that wff for a variable of a different type.

Many HOSTs have the following as axioms or theorems, including the versions of type theory outlined by Hatcher:

The axiom (principle) of extensionality (Ext):

$$(\forall F^{(\tau_1,\dots,\tau_n)})(\forall G^{(\tau_1,\dots,\tau_n)})((\forall x_1^{\tau_1})\dots(\forall x_n^{\tau_n})(F(x_1^{\tau_1},\dots,x_n^{\tau_n}) \leftrightarrow G(x_1^{\tau_1},\dots,x_n^{\tau_n})) \rightarrow F = G)$$

The axiom (principle) of infinity (Inf):

$$(\exists R^{(o,o)})((\forall x^o)\neg R(x,x) \wedge (\forall x)(\exists y^o)R(x,y) \wedge (\forall x)(\forall y)(\forall z^o)(R(x,y) \wedge R(y,z) \rightarrow R(x,z)))$$

(Ext) forces us to think of the values of the higher-order variables as entities with *extensional* identity conditions, such as sets or classes. We may also think of them as Fregean functions from objects to truth-values, or something similar. However, we may not think of them as Platonic universals or other intensionally-individuated entities. Thus we have as the simplest instance:

$$(\forall F^{(o)})(\forall G^{(o)})((\forall x^o)(Fx \leftrightarrow Gx) \rightarrow F = G)$$

This says that "properties" of individuals are identical when they apply to all and only the same individuals.

With (Ext) we are free to regard the values of variables of type (o) simply as *classes of individuals*, and the variables of type $((o))$ as *classes of classes*, etc. We are then free to regard the notation $[\lambda \chi\, \mathscr{A}[\chi]]$ as a mere variant of $\{\chi|\mathscr{A}[\chi]\}$, and the notation $\chi(y)$ as simply a variant of $y \in \chi$.

With both (Ext) and (Inf) added to HOPC, we can develop Peano arithmetic *precisely* as we did for ST. The principle of lambda conversion is a more general version of ST1, (Ext) is a more general version of ST2, and (Inf) replaces ST3.

(HOPC + (Ext) + (Inf) is, in effect, Hatcher's TT.)

HOMEWORK 10

Prove the following:

(a) $\vdash_{\text{HOPC}} (\forall x^\tau)\, x^\tau = x^\tau$
(b) $\vdash_{\text{HOPC}} (\forall x^\tau)(\forall y^\tau)(x = y \rightarrow (\mathscr{A}[x,x] \rightarrow \mathscr{A}[x,y]))$, where $\mathscr{A}[x,x]$ does not contain y bound.
(c) Explain why it would be unnecessary to take the quantifier $(\forall \chi)$ as a primitive variable binding operator in the syntax of HOSTs if one instead used *constants* of the form $A^{((\tau))}$, with the intended meaning of a property of properties applicable to those properties which hold of everything in their type.

C. Doing without (Ext) or (Inf)

Is it possible to define numbers and capture arithmetic without assuming (Ext)?

Russell used a version of type theory in which class abstracts were introduced via contextual definitions. The general approach could be summarized as follows:

The wff $\mathscr{B}[\{\chi|\mathscr{A}[\chi]\}]$ is an abbreviation for the wff, $(\exists f)((\forall \chi)(f(\chi) \leftrightarrow \mathscr{A}[\chi]) \wedge \mathscr{B}[f])$, where f is the first (predicative) variable of type (τ), where τ is the type of the variable χ.

This approach was of a piece with his theory of descriptions, the view that a wff of the form $\mathscr{B}[\imath\chi\, \mathscr{A}[\chi]]$ is to be regarded as an abbreviation of a wff of the form $(\exists \chi)((\forall y)(\mathscr{A}[y] \leftrightarrow y = \chi) \wedge \mathscr{B}[\chi])$.

Both kinds of contextual definition give rise to scope ambiguities. For example, $\neg M^{((o))}(\{x^o|x^o = x^o\})$ may mean either:
$(\exists F^{(o)})((\forall x^o)(F(x) \leftrightarrow x = x) \wedge \neg M^{((o))}(F^{(o)}))$, or
$\neg(\exists F^{(o)})((\forall x^o)(F(x) \leftrightarrow x = x) \wedge M^{((o))}(F^{(o)}))$

Conventions had to be adopted in *Principia Mathematica* to avoid ambiguities.

It is clear, however, that when the contextual definitions are interpreted with wide-scope, from $\mathscr{B}[\{\chi|\mathscr{A}[\chi]\}]$ and $(\forall\chi)(\mathscr{A}[\chi] \leftrightarrow \mathscr{C}[\chi])$, one can always infer $\mathscr{B}[\{\chi|\mathscr{C}[\chi]\}]$ even without (Ext).

Numbers, etc., would then be defined as classes, as before, but since class symbols are eliminated in context, number signs would be eliminated. The name of a number is, for Russell, an "incomplete symbol": one that contributes to the meaning of a complete sentence without having a meaning on its own. For this reason, Russell calls numbers "logical fictions" or "logical constructions."

Other approaches for avoiding (Ext) may be possible in a richer, extended language, such as one involving modal operators. A weaker, but perhaps more plausible principle may be:
(□Ext) $(\forall F^{(o)})(\forall G^{(o)})(\Box(\forall x^o)(Fx \leftrightarrow Gx) \rightarrow F = G)$

If numbers are construed then as properties of properties, this may be sufficient for the development (at least) of natural number theory.

We might also do without (Inf) in a given HOST if the HOST in question had some other means of establishing an infinity of individuals (or the fourth Peano postulate, which itself can be used to establish an infinity of individuals). Principles regarding other abstract or modal entities (senses, propositions, possibilia) may provide the means for establishing an infinity in another way, though such principles are liable to be philosophically controversial, and have pitfalls of their own.

D. Some Philosophical Issues Revisited

The issues surrounding HOPC and other HOSTs are in effect the same as those for ST, viz.:

1. Is "$\neg F(F)$" really meaningless? Can the theory be stated without violating itself?
2. Is impredicativity a problem? Notice, we allow the abstract "$[\lambda x^o (\forall F^{(o)}) F^{(o)} x^o]$" to define a property of type (o), despite that quantification over all such properties is part of its defining condition.
3. Is there sufficient reason for accepting (Inf)? What if we drop it?
4. What are the entities quantified over by higher-order variables? What are their identity conditions? Are there any other philosophical issues that arise?
5. Does higher-order logic count as *logic*?

XIV. Ramified Type Theory

In order to avoid problems with impredicativity, one may further divide types into various *orders*.

The notion can be introduced by means of some natural language examples.

Consider such properties as *being green*, or *being North of London*. Intuitively, at least these properties are not defined or constituted by quantification over other properties and relations. Let us call these *order one properties*.

Now consider such properties as *having some property in common with Kevin* or *having every property had by every great general*. On a *simple* type theory, these would be regarded as having the same type as *being green*: properties of individuals or type (o). However, in ramified type theory, these are treated as *order two properties*, properties involving quantification over order one properties.

The exact formulation of ramified type theory is very controversial, and there is no universally agreed upon form for it to take. Hatcher's formulation of what he calls System RT is as follows:

Definition: An ***order-type symbol*** is either a symbol of the form $o/0$, or one of the form $(\tau_1, \ldots, \tau_n)/k/(m_1, \ldots, m_n)$, where $(\tau_1, \ldots, \tau_n)$ is a type symbol, $m_1, \ldots, m_n$ are integers, each at least as high as the order of the corresponding type symbol, and k is an integer at least as high as the order of $(\tau_1, \ldots, \tau_n)$ and higher than any of $m_1, \ldots, m_n$.

(The "order" of a type-symbol was defined in the last section.)

In an order-type symbol $(\tau_1, \ldots, \tau_n)/k/(m_1, \ldots, m_n)$, k represents the *order* of the variable, constant or term in question, and $m_1, \ldots, m_n$ represent the greatest possible orders of its arguments.

Examples:

$o/0$	individuals
$(o)/1/(0)$	order-1 properties of individuals
$(o)/2/(0)$	order-2 properties of individuals
$((o))/2/(1)$	order-2 properties of order-1 properties of individuals
$((o))/3/(1)$	order-3 properties of order-1 properties of individuals
$((o), o)/2/(1, 0)$	order-2 relations between properties of individuals and individuals
etc.	

The syntax of RT requires that variables and constants have order-type symbols as superscripts rather than simply type symbols.

If f is a variable or abstract of order-type $(\tau_1, \ldots, \tau_n)/k/(m_1, \ldots, m_n)$, then $f(t_1, \ldots, t_n)$ is well-formed only if $t_1, \ldots, t_n$ are of types $\tau_1, \ldots, \tau_n$ respectively, and have orders of no greater than $m_1, \ldots, m_n$, respectively.

An abstract of the form $[\lambda \chi_1, \ldots, \chi_n \mathscr{A}]$ has order-type $(\tau_1, \ldots, \tau_n)/j + k/(m_1, \ldots, m_n)$, where $\chi_1, \ldots, \chi_n$ are respectively of type $\tau_1, \ldots, \tau_n$ and order $m_1, \ldots, m_n$, and k is the highest order of all variables and constants in $[\lambda \chi_1, \ldots, \chi_n \mathscr{A}]$ or the minimal order possible for the type, and j is either 1 or 0 depending on whether any variables of order k occur bound in $[\lambda \chi_1, \ldots, \chi_n \mathscr{A}]$.

Examples: $[\lambda x^{o/0} F^{(o)/1/(0)}(x^{o/0})]$ has order-type $(o)/1/(0)$, but $[\lambda x^{o/0} (\forall F^{(o)/1/(0)}) F^{(o)/1/(0)}(x^{o/0})]$ has order-type $(o)/2/(0)$.

Definition: A term is said to be ***predicative*** iff its order-type is $o/0$, or its order-type is $(\tau_1, \ldots, \tau_n)/k/(m_1, \ldots, m_n)$, where k is one more than the greatest of $m_1, \ldots, m_n$.

Hence, terms of order-type $(o)/1/(0)$ are predicative; those of order-type $(o)/2/(0)$ are not.

Hatcher's system PT is a system like HOPC, but that excludes from the syntax all but *predicative* constants and variables. Notice that there may still be non-predicative abstracts, as with $[\lambda x^{o/0} (\forall F^{(o)/1/(0)}) F^{(o)/1/(0)}(x^{o/0})]$ since all variables used within it are predicative. PT restricts the axioms involving instantiation and lambda conversion to predicative terms.

The system RT, however, makes use of variables of all orders; instantiation and conversion are allowed for all orders, provided that instantiated terms match the instantiated variable in order, and the appropriate restrictions on well-formed formulas are obeyed. (It is otherwise like HOPC.)

The philosophical upshot in that no term can be introduced that involves quantification over a range that includes itself: perhaps avoiding "vicious circularity" or other philosophical problems. Notice, however, that such restrictions are not needed to rule out Russell's paradox or other direct source of inconsistency.

Problems with Ramified Type-Theory

With order restrictions in place, RT is very weak, and very little mathematics can be captured in it without adding additional axioms. Some problems:

Identity

$t = u$ is normally defined as $(\forall \chi)(\chi(t) \leftrightarrow \chi(u))$, but here the variable χ must be restricted to a given order: most naturally, it would be one above t and u.

However, this would rule out applications of Leibniz's law to wffs involving quantified variables of the same (or higher) type or order.

This could perhaps be gotten around by taking identity as primitive and adding LL as an axiom schema, but this seems less than ideal.

Mathematical Induction

In the approaches to acquiring natural number theory we've looked at, the principle of mathematical induction to supposed to fall out of the definition of natural numbers. In HOPC, we might have given this definition of N: $[\lambda m^{((o))}(\forall F^{(((o)))})(F(0) \wedge (\forall n^{((o))})(F(n) \rightarrow F(n')) \rightarrow F(m))]$. In RT, however, we would have to limit F to a certain order, and the rule of mathematical induction would not apply to formulas involving terms of greater order. However, applications of mathematical induction to non-predicative properties of numbers are widespread and necessary for many applications of number theory.

To get around such problems, ramified type-theories often make use of an ***axiom*** (or axioms) ***of reducibility***, e.g.:

$$(\forall F^{(o)/n/(0)})(\exists G^{(o)/1/(0)})(\forall x^{o/0})(G(x) \leftrightarrow F(x)), \text{ for any } n > 1.$$

This says that for every non-predicative property of individuals, there is a coextensional predicative property of individuals. (Typically there would also be included analogues for higher-types, and relations, etc.)

Ever since Russell and Whitehead suggested something similar in *Principia Mathematica*, the axiom(s) of reducibility have been the source of significant controversy:

1. Are they plausible? Is there logical (or other) justification for supposing them?
2. Do they more or less *undo* the effects of ramification?

(But as against answering "yes" to (2), it is worth noting that they did not accept (Ext), and so the G posited wasn't thought *identical* with F.)

XV. Meta-theory for Higher-Order Logic

We here deal with HOPC and other HOSTs, not with ramified type-theory.

HOPC, with or without (Ext) and (Inf) is, like ST, consistent relative to weak Zermelo set theory (Z).

The precise way to do formal semantics for higher-order logic is controversial, in part because there is disagreement over what entities are quantified over by higher-order predicate variables.

The usual way, however, involves assigning *sets of objects of the domain* (or sets of sets, etc.) as possible values of the predicate variables. (Models constructed in this way presuppose that (Ext) will come out as true.)

A. Standard Semantics

Definition: A ***full model*** M is a specification of the following two things:

1. A non-empty set D to serve as the domain of quantification for variables of type o.

   In a full model, the domain of quantification for other types of variables is determined entirely by D. More precisely, the domain for type of the form (τ) is the powerset of the domain for type τ, and the domain of quantification for type $(\tau_1, \ldots, \tau_n)$ is the powerset of the set of all n-tupes $\langle o_1, \ldots, o_n \rangle$ where each o_i is taken from the domain of quantification for type τ_i.

   We refer to the domain of quantification for type τ as $D(\tau)$.

2. A function assigning, to each constant c of type τ, some member of $D(\tau)$. (We refer to this member of $D(\tau)$ as $(c)^M$ in the metalanguage.)

In a language, such as that of HOPC, without constants, a full model consists of *nothing* but a domain of quantification.

Definition: A ***variable assignment*** is a function assigning to each variable χ some member of $D(\tau)$, where τ is the type of χ.

Each variable assignment s determines a valuation function V^s which assigns to each typeless wfe either 1 or 0 (satisfaction or non-satisfaction), and assigns to each wfe of type τ a member of $D(\tau)$. V^s is defined recursively as follows:

(i) For any constant c, $V^s(c)$ is $(c)^M$.
(ii) For any variable χ of type τ, $V^s(\chi)$ is the member of $D(\tau)$ that s associates with χ.
(iii) For any wfe of the form $\mathscr{P}(t_1, \ldots, t_n)$, $V^s(\mathscr{P}(t_1, \ldots, t_n))$ is 1 if $\langle V^s(t_1), \ldots, V^s(t_n) \rangle$ is a member of $V^s(\mathscr{P})$, and is 0 otherwise.
(iv) For any wfe of the form $\mathscr{A} \vee \mathscr{B}$, $V^s(\mathscr{A} \vee \mathscr{B})$ is 1 if either $V^s(\mathscr{A})$ is 1 or $V^s(\mathscr{B})$ is 1 or both, or is 0 otherwise.

(v) For any wfe of the form $\neg\mathscr{A}$, $V^s(\neg\mathscr{A})$ is 1 if $V^s(\mathscr{A})$ is 0, and is 0 if $V^s(\mathscr{A})$ is 1.

(vi) For any wfe of the form $(\forall\chi)\mathscr{A}$, $V^s((\forall\chi)\mathscr{A})$ is 1 if every variable assignment s^* differing from s at most with regard to what gets assigned to variable χ is such that $V^{s^*}(\mathscr{A})$ is 1, and is 0 otherwise.

(vii) For any wfe of the form $[\lambda\chi_1 \dots \chi_n\,\mathscr{A}]$, $V^s([\lambda\chi_1 \dots \chi_n\,\mathscr{A}])$ is the set of all n-tuples $\langle o_1, \dots, o_n\rangle$ such that $V^{s^*}(\mathscr{A})$ is 1 where s^* is the variable assignment differing from s at most in assigning o_1 to χ_1, …, and o_n to χ_n.

Definition: A variable assignment s is said to ***satisfy*** a wff $\mathscr{A}$ iff $V^s(\mathscr{A}) = 1$.

Definition: A wff $\mathscr{A}$ is said to be ***true for M*** iff every variable assignment for model M satisfies $\mathscr{A}$.
We write this as $\vDash_{\mathrm{M}} \mathscr{A}$.

Definition: A wff $\mathscr{A}$ is said to be ***standardly valid*** iff $\mathscr{A}$ is true for every full model.

Soundness

HOPC + (Ext) is sound, i.e., if $\vdash_{\mathrm{HOPC+(Ext)}} \mathscr{A}$ then $\mathscr{A}$ is standardly valid.

(*Proof sketch:* all the axioms are true in every full model, and the inference rules are truth-preserving.)

B. Incompletess of Standard Semantics

The converse of soundness is not true, i.e., there are standardly valid wffs that are not theorems of HOPC + (Ext). (The same holds for any other recursively axiomatized consistent HOST.)

To see this, consider the language obtained from HOPC by adding a constant a^o, a two-place relation sign $A^{(o,o)}$, and two three-place relation signs $B^{(o,o,o)}$ and $C^{(o,o,o)}$.

- Interpret a^o as standing for the number 0. (Never mind that we have considered defining numbers as properties of properties in HOPC.)
- Interpret $A^{(o,o)}$ as standing for the relation that holds between a number and its successor.
- Interpret $B^{(o,o,o)}$ and $C^{(o,o,o)}$ as standing for the relations that hold between two numbers and their sum, and product, respectively.

Consider then the system HOPA (*Higher-Order Peano Arithmetic*), which has the following non-logical axioms in addition to (Ext):

PA1. $(\forall x^o)(\exists y^o)(\forall z^o)(A(x,z) \leftrightarrow z = y)$
PA2. $(\forall x^o)(\forall w^o)(\exists y^o)(\forall z^o)(B(x,w,z) \leftrightarrow z = y)$
PA3. $(\forall x^o)(\forall w^o)(\exists y^o)(\forall z^o)(C(x,w,z) \leftrightarrow z = y)$
PA4. $(\forall x^o)\neg A(x,a)$
PA5. $(\forall x^o)(\forall y^o)(\forall z^o)(A(x,z) \wedge A(y,z) \rightarrow x = y)$
PA6. $(\forall x^o)B(x,a,x)$
PA7. $(\forall u^o)(\forall v^o)(\forall x^o)(\forall y^o)(\forall z^o)(A(y,u) \wedge B(x,u,v) \wedge B(x,y,z) \rightarrow A(z,v))$
PA8. $(\forall x^o)C(x,a,a)$
PA9. $(\forall u^o)(\forall v^o)(\forall x^o)(\forall y^o)(\forall z^o)(A(y,u) \wedge C(x,u,v) \wedge C(x,y,z) \rightarrow B(z,x,v))$
PA10. $(\forall F^{(o)})(F(a) \wedge (\forall x^o)(\forall y^o)(F(x) \wedge A(x,y) \rightarrow F(y)) \rightarrow (\forall x)F(x))$

These correspond to the axioms of S, except instead of an infinite number of instances of schema for mathematical induction, we have a single axiom involving a bound predicate variable.

Notice that the full model N^* whose domain of quantification is the set of natural numbers, and whose assignment to a, A, B and C are what one would expect, is a model for HOPA. But we have something stronger.

Categoricity of HOPA: Every full model for HOPA is isomorphic to N^* (and hence makes all the same wffs true).

I shall not sketch the proof of this in detail, but notice that each model M of HOPA must have denumerably many entities in its domain of quantification. Correlate 0 with $(a)^{\mathrm{M}}$, and for each object in the domain there is exactly one entity to which it is related by A: its "successor". (Notice that because identity is defined, we need not worry about non-normal models.) We get that $(a)^{\mathrm{M}}$ and its "successors" must exhaust the domain of quantification. Otherwise, there would exist some set in the domain of quantification for type (o) containing a^{M} and all its successors but not every member of the domain—but this is ruled out by the truth of PA10. Model M must place the appropriate n-tuples in the extensions of A, B and C in order to make the remaining axioms true, and hence is isomorphic to N^*.

Obviously, HOPA is at least as strong as first-order Peano Arithmetic. Therefore, it is capable of representing every recursive function and expressing every recursive relation. Moreover, it has a recursive syntax and recursive axiom set. Hence, if it is consistent (which it must be since it had a model), the Gödel and Gödel-Rosser results apply. So, there is a closed wff $\mathcal{G}$, true in the model N^*, but for which it is not the case that $\vdash_{\text{HOPA}} \mathcal{G}$.

Let $\mathcal{PA}$ abbreviate the conjunction of PA1–PA10.

Let $\mathcal{PA}[x^o, f_1^{(o,o)}, f_2^{(o,o,o)}, f_3^{(o,o,o)}]$ be the open wff derived from $\mathcal{PA}$ by replacing every occurrence of the constant a^o with the variable x^o, and ever occurrence of the predicate constants $A^{(o,o)}$, $B^{(o,o,o)}$ and $C^{(o,o,o)}$ with, respectively, the predicate variables $f_1^{(o,o)}$, $f_2^{(o,o,o)}$ and $f_3^{(o,o,o)}$, where these are the first variables of the appropriate types not found in either $\mathcal{PA}$ or $\mathcal{G}$.

Let $\mathcal{G}[x^o, f_1^{(o,o)}, f_2^{(o,o,o)}, f_3^{(o,o,o)}]$ be derived from $\mathcal{G}$ similarly.

Now consider the wff (#):

$$(\forall x^o)(\forall f_1^{(o,o)})(\forall f_2^{(o,o,o)})(\forall f_3^{(o,o,o)})(\mathcal{PA}[x, f_1, f_2, f_3] \to \mathcal{G}[x, f_1, f_2, f_3])$$

This wff is standardly valid but not a theorem of HOPC + (Ext).

Sketch of proof that it is standardly valid: Consider any variable assignment s. If s satisfies $\mathcal{PA}[x, f_1, f_2, f_3]$, it must assign appropriate values to these variables. If the same assignments were made to the constants they replaced in some full model M^* for the language of HOPA, the result would be a model for HOPA. Hence M^* would be isomorphic to N^*, and $\mathcal{G}$ would be true in M^* just as it is for N^*. Since s treats these variables just as M^* treats those constants, it must satisfy $\mathcal{G}[x, f_1, f_2, f_3]$. Hence, every assignment that satisfies the antecedent of (#) also satisfies the consequent, so (#) is true in any full model.

Proof it is not a theorem of HOPC + (Ext): If (#) were a theorem of HOPC + (Ext), it would also be a theorem of HOPA. By UI, we'd obtain $\vdash_{\text{HOPA}} \mathcal{PA} \to \mathcal{G}$, and by Conj and MP, then, $\vdash_{\text{HOPA}} \mathcal{G}$, which is impossible.

Notice the argument would apply just as well the second-order predicate calculus, or any other specific order above first, not just HOPC.

The incompleteness of higher-order logic is sometimes used as an argument in favor of first-order logic. However, the incompleteness is due largely to its expressive power and categoricity: features which could be used to build arguments in favor of higher-order logic over first-order logic.

There is another, weaker, notion of completeness according to which HOPC + (Ext) is complete:

C. Henkin Semantics

Definition: A ***general structure*** is the specification of two things:

1. A function assigning to each type a domain of quantification, for type o some non-empty set D, and for each type of the form $(\tau_1, \ldots, \tau_n)$, some non-empty *subset* of the powerset of the set of all n-tuples taken from $\tau_1, \ldots, \tau_n$.

2. An assignment to each constant (if any) of type τ, some member of $D(\tau)$.

The difference between a general structure and a full model is that the higher-order variables need not be taken as quantifying over *all* sets of entities or n-tuples taken from the type of possible arguments, but may be restricted. So, for example, the variables of type (o) may not range over all possible sets of individuals, but may instead be restricted only to certain subsets of the individuals.

Notice that full models can be conceived as a certain kind of general structure, those in which the domains of quantification do include the full range. All full models are general structures, but not vice versa.

For each general structure, we can consider every possible variable assignment and define our valuation functions V^s exactly as before.

However, with the definitions as stated, we do encounter a problem (or at least an oddity). Consider this statement from the definition of V^s:

> $V^s([\lambda \chi_1 \ldots \chi_n \mathscr{A}])$ is the set of all n-tuples $\langle o_1, \ldots, o_n \rangle$ such that $V^{s^*}(\mathscr{A})$ is 1 for the variable assignment s^* differing from s at most in assigning o_1 to χ_1, … and o_n to χ_n.

Notice, however, that the set defined this way may or may not be included in the domain of quantification for type $(\tau_1, \ldots, \tau_n)$, where $\tau_1, \ldots, \tau_n$ are the types of $\chi_1, \ldots, \chi_n$ respectively.

To rectify this, we restrict our attention to those structures where this problem does not arise:

Definition: A ***general model*** M is a general structure such that for every variable assignment s for M, $V^s(t)$ is a member of $D(\tau)$ for every term t of type τ.

In effect, a general model is a structure in which the axioms of instantiation and λ-conversion hold.

Definition: A wff $\mathscr{A}$ is ***generally valid*** iff it is true for every general model.

We get the following results:

1. *Soundness:* Every wff $\mathscr{A}$ which is a theorem of HOPC + (Ext) is generally valid.

2. *Henkin Completeless:* Every wff $\mathscr{A}$ which is generally valid is a theorem of HOPC + (Ext).

The proof of Henkin completeness is not interestingly different from the standard proof of the completeness of the first-order predicate calculus. Every consistent set of closed wffs can be expanded to become a maximally consistent (and universal) set of closed wffs. This set can be used to construct a denumerable model for the set. Every consistent set of closed wffs has a model. If there a generally valid wff $\mathscr{A}$ for HOPC + (Ext) which was not a theorem, the set $\{\neg\mathscr{B}\}$, where $\mathscr{B}$ was the universal closure of $\mathscr{A}$, would be consistent, and hence would have a general model. But if $\mathscr{A}$ is generally valid, $\neg\mathscr{B}$ cannot have a general model. For more details, see Robbin (pp. 141–42) or the original paper by Henkin (*Journal of Symbolic Logic* 15 (1950): 81–91.)

Notice that (#), while standardly valid, is not generally valid. This is because $\mathcal{PA}$ is not categorical among general models: there are general models of HOPA not isomorphic to N^*. Notice, e.g., that the domain of quantification for a general model need not be exhaused by $(a)^M$ and its successors in order to make PA10 true, provided that the set that contains $(a)^M$ and its successors but not the other entities in the domain of quantification is not included in the domain for type (o). Hence there can be variable assignment in general models that satisfy the antecedent of (#) but not the consequent.

Contemporary Axiomatic Set Theories

We now turn our attention to the systems that can serve as a foundation for arithmetic most often discussed by contemporary mathematicians. They are usually formulated as first-order set theories. We begin with the simplest.

XVI. Zermelo and Zermelo-Fraenkel Set Theories (Systems Z, ZF)

The theory outlined below is a modernized version of the theory of sets outlined in Zermelo's classic 1908 paper, "Untersuchungen über die Grundlagen der Mengenlehre" ("Investigations into the Foundations of Set Theory"), roughly contemporaneous with Russell's "Mathematical Logic as Based on the Theory of Types".

A. Syntax

The language contains:

(i) two primitive constants, but rather than writing 'a' and 'b', we write '0' and 'ω';

(ii) one primitive two-place predicate, but rather than writing $A^2(t,u)$, we write $t \in u$;
(iii) four function constants, three of them one place, one of them two-place, but instead of writing $f^1(t)$, $g^1(t)$, $h^1(t)$ and $f^2(t,u)$, we write $\wp(t)$, $\bigcup(t)$, $\sigma(t)$ and $\{t,u\}$;
(iv) one vbto, written $\{\chi|\mathscr{A}[\chi]\}$.

However, it would be possible to formulte more or less the same theory using only the predicate "∈". (Indeed, you will likely find formulations of Z or ZF in other textbooks like this.) The others are added for convenience.

Definitions: (With many others to follow...)[1]

$t = u$	for	$(\forall\chi)(\chi \in t \leftrightarrow \chi \in u)$
$t \neq u$	for	$\neg t = u$
$t \notin u$	for	$\neg t \in u$
$t \subseteq u$	for	$(\forall\chi)(\chi \in t \rightarrow \chi \in u)$
$t \cup u$	for	$\bigcup(\{t,u\})$
$\{t\}$	for	$\{t,t\}$

B. Formulation of Z

In addition to the axioms and inference rules for the predicate calculus for this first-order language, Z has the following axioms:

Leibniz's Law / Axiom(s) of extensionality[2]
Z1: $(\forall x)(\forall y)(x = y \rightarrow (\mathscr{A}[x,x] \rightarrow \mathscr{A}[x,y]))$, where y does not occur bound in $\mathscr{A}[x,x]$.

Axiom(s) of separation/selection (Aussonderung)
Z2: $(\forall z)(\forall\chi)(\chi \in \{y|y \in z \wedge \mathscr{A}[y]\} \leftrightarrow \chi \in z \wedge \mathscr{A}[\chi])$, where χ, y and z are distinct variables, and χ does not become bound in the context $\mathscr{A}[\chi]$.

[1]From here on out when giving a definition like this, I shall stop explicitly saying that χ is the first variable not occurring free in t and u, and take for granted that you understand the conventions for using schematic letters in definitions.

[2]This is what Hatcher calls it; it is a bit misleading. Zermelo himself took identity as primitive and had something more like (Ext) from type-theory as an axiom, which would more properly be called an axiom of extensionality. However, our Z1 can be seen roughly as saying that Z only deals with extensional contexts.

Null set axiom:
Z3: $0 = \{x|x \in 0 \wedge x \neq x\}$

Power set axiom:
Z4: $(\forall x)(\forall y)(x \in \wp(y) \leftrightarrow x \subseteq y)$

Pairing axiom:
Z5: $(\forall x)(\forall y)(\forall z)(x \in \{y,z\} \leftrightarrow x = y \vee x = z)$

Union / Sum set axiom:
Z6: $(\forall x)(\forall y)(x \in \bigcup(y) \leftrightarrow (\exists z)(z \in y \wedge x \in z))$

Axiom of infinity:
Z8′: $(\exists x)(0 \in x \wedge (\forall y)(y \in x \rightarrow y \cup \{y\} \in x))$
(This is not how Hatcher formulates his ZF.8, though this is closer to its usual formulation. I skip Z7 in numbering to match Hatcher's.)

Modern Zermelo-Fraenkel Set Theory (ZF) is obtained from Z by adding the axioms of *replacement* and *regularity*. More on these below.

If one wished to avoid use of primitive constants and function letters and the vbto $\{\chi|\mathscr{A}[\chi]\}$, one could formulate what amounts to the same using existentially formulated principles such as:

$(\forall z)(\exists y)(\forall\chi)(\chi \in y \leftrightarrow \chi \in z \wedge \mathscr{A}[\chi])$ (separation)
$(\forall y)(\exists z)(\forall x)(x \in z \leftrightarrow x \subseteq y)$ (powerset)

It is convenient, however, to have names for the sets postulated to exist by the axioms.

In some ways the most important axiom of the system Z is the axiom(s) of separation. Rather than postulating a set of x such that $\mathscr{A}[x]$ for any wff $\mathscr{A}$, as F2 of system F does, Z's principle of separation postulates, for any set z we already know to exist, a subset of z of which x is a member iff x is in z and $\mathscr{A}[x]$.

In this way we can divide any set already known to exist into two subsets, one for which $\mathscr{A}[x]$ and another for which $\neg\mathscr{A}[x]$. However, we cannot speak of the set of *all* x absolutely such that $\mathscr{A}[x]$.

The notation $\{x|x \in y \wedge \mathscr{A}[x]\}$ is used for the subset of y which includes the members of it such that $\mathscr{A}[x]$. Other books may use the notation $\{x \in y|\mathscr{A}[x]\}$. Notice that Hatcher's formulation allows $\{y|\mathscr{A}[y]\}$ to be a well-formed term even if the wff in the abstract does not begin with $y \in t$, it is only when it does so begin that anything interesting can be proven using that term.

It is obvious that Z cannot countenance a universal set, or else we could define the Russell set R as $\{z|z \in \mathrm{V} \wedge z \notin z\}$, and derive a paradox. Indeed, we have the following:

$\vdash_{\mathrm{Z}} (\forall x)(\exists y)\, y \notin x$

Proof:

(1)	1. $\neg(\exists y)\, y \notin x$	Hyp
(1)	2. $(\forall y)\, y \in x$	1 CQ, DN
	3. $(\forall x)(\forall y)(y \in \{z|z \in x \wedge z \notin z\} \leftrightarrow y \in x \wedge y \notin y)$	Z2
	4. $\{z|z \in x \wedge z \notin z\} \in \{z|z \in x \wedge z \notin z\} \leftrightarrow \{z|z \in x \wedge z \notin z\} \in x \wedge \{z|z \in x \wedge z \notin z\} \notin \{z|z \in x \wedge z \notin z\}$	3 UI×2
	5. $\{z|z \in x \wedge z \notin z\} \in \{z|z \in x \wedge z \notin z\} \rightarrow \{z|z \in x \wedge z \notin z\} \notin \{z|z \in x \wedge z \notin z\}$	4 SL
	6. $\{z|z \in x \wedge z \notin z\} \notin \{z|z \in x \wedge z \notin z\}$	6 SL
	7. $\{z|z \in x \wedge z \notin z\} \in x \wedge \{z|z \in x \wedge z \notin z\} \notin \{z|z \in x \wedge z \notin z\} \rightarrow \{z|z \in x \wedge z \notin z\} \in \{z|z \in x \wedge z \notin z\}$	4 SL
(1)	8. $\{z|z \in x \wedge z \notin z\} \in x$	2 UI
(1)	9. $\{z|z \in x \wedge z \notin z\} \in \{z|z \in x \wedge z \notin z\}$	6, 7, 8 SL
	10. $(\exists y)\, y \notin x$	1, 6, 9 RAA
	11. $(\forall x)(\exists y)\, y \notin x$	10 UG

Whence in particular: $\vdash_{\mathrm{Z}} (\exists y)\, y \notin \{x|x = x\}$

This is despite that we do have (by a very easy proof): $\vdash_{\mathrm{Z}} (\forall y)\, y = y$.

Obviously, then $\{x|x = x\}$ cannot stand for the set of *all* (and only) self-identical things.

Since the separation axiom only allow us to divide sets already known to exist, we need additional principles guaranteeing the existence of interesting sets. This is the role of the other axioms. They allow us to deduce the existence of sets that are built up from sets already known to exist by means of certain processes (pairing, forming a powerset, etc.). Iterating these processes, we can deduce the the existence of more and more sets.

This is why this system and others like it are sometimes called *iterative set theory*.

Some theorems of Z (some with obvious proofs):

T4. $\vdash_{\mathrm{Z}} (\forall x)((\forall y)(y \notin x) \leftrightarrow x = 0)$
T5. $\vdash_{\mathrm{Z}} (\forall x)(\forall y)(x \subseteq y \wedge y \subseteq x \rightarrow x = y)$
T6. $\vdash_{\mathrm{Z}} (\forall x)\, x \subseteq x$
T7. $\vdash_{\mathrm{Z}} (\forall x)(\forall y)(\forall z)(x \subseteq y \wedge y \subseteq z \rightarrow x \subseteq z)$
T8. $\vdash_{\mathrm{Z}} (\forall x)\, 0 \subseteq x$

Definitions:

$\langle t, u\rangle$	for	$\{\{t\},\{t,u\}\}$
$\bigcap(t)$	for	$\{x \mid x \in \bigcup(t) \wedge (\forall y)(y \in t \rightarrow x \in y)\}$
$t \cap u$	for	$\bigcap(\{t,u\})$
$t - u$	for	$\{x \mid x \in t \wedge x \notin u\}$

T9. $\vdash_Z (\forall x)(\forall y)(x \in \{y\} \leftrightarrow x = y)$
T9a. $\vdash_Z (\forall x_1)(\forall x_2)(\forall y_1)(\forall y_2)(\langle x_1, y_1\rangle = \langle x_2, y_2\rangle \leftrightarrow x_1 = x_2 \wedge y_1 = y_2)$
T10. $\vdash_Z (\forall x)(\forall y)(\forall z)(x \in y \cup z \leftrightarrow x \in y \vee x \in z)$
T11. $\vdash_Z (\forall x)(\forall y)(\forall z)(x \in y \cap z \leftrightarrow x \in y \wedge x \in z)$
T11a. $\vdash_Z (\forall x)(\forall y)(\forall z)(x \in y - z \leftrightarrow x \in y \wedge x \notin z)$

We also get the commutativity and associativity of $\cap$ and $\cup$ and similar results as theorems.

C. Cardinal and Ordinal Numbers in Z(F)

Cardinal numbers

Cardinal numbers derive from the notion of "cardinality," roughly synonymous with the notion of "size". A cardinal number is one that answers the question "how many?" Usually, we apply the notion of cardinality to sets or to properties. We say that two sets *have the same cardinality* or *have the same cardinal number* when there is one-one correspondence between their members.

This, however, does not tell what the cardinal numbers themselves are. Frege defined a cardinal number as a set of sets sharing the same cardinality: so 0 is the set of all zero-membered sets, 1 is the set of all one-membered sets, 2 is the set of all two-membered sets. Russell adopted more or less the same definitions, though, as we've seen, in type theory, this means distinct numbers in distinct types.

This definition could not be used in Z or ZF. The axioms of ZF do not allow us to deduce the existence of the set of all one-membered sets; indeed, it is provable that ZF cannot countenance any such set. (Consider that if "1" represented the set of *all* singletons, since every set is a member of its own singleton, "$\bigcup(1)$" would name the universal set, and we have already seen how from the supposition of a universal set in Z(F) we can derive a contradiction.)

Typically, then in ZF, cardinal numbers are defined as representative sets having a given cardinality. 0 is defined as as a particular set having no members: in this case, there is only one such set: the empty set. 1 is defined as the set whose only member is 0. 2 is defined as the set whose only members are 0 and 1. This gives us the progression:

$$0, \{0\}, \{0, \{0\}\}, \{0, \{0\}, \{0, \{0\}\}\}, \ldots$$

Notice that these sets are the same as the members of the set ω, used to guarantee an infinity even in System F. There we only proved that each natural number has a member of ω in it; here we define the natural numbers *as* members of ω.

Ordinal numbers

Whereas cardinal numbers are associated with sets or properties, ordinal numbers are associated with relations, or their fields. We begin by rehearsing some terminology widely used in discussing relations (here informally):

Definitions:

- The ***domain*** of a relation R is the set of all x for which there is a y such that xRy.
- The ***range*** of a relation R is the set of all x such that there is a y such that yRx.
- The ***field*** of a relation R is the union of its domain and range, i.e., all the relata of the relation.
- Relation R is ***irreflexive*** on set S iff there is no x in S such that xRx.
- Relation R is ***asymmetric*** on set S iff for all x and y in S, if xRy then it is not the case that yRx.
- Relation R is ***transitive*** on set S iff for all x, y and z in S, if xRy and yRz then xRz.
- Relation R is ***connected*** on S iff for all x and y in S, if $x \neq y$ then either xRy or yRx.
- Relation R is a ***partial order*** on set S iff S is subset of (or is) its field, and R is irreflexive, asymmetric and transitive on S.

- Relation R is a ***total order*** on set S iff it is a partial order on S and connected on S.
- Relation R is a ***well order*** on set S iff it is a total order on S and for every non-empty subset T of S, there is a member x of T for which, for all y in T, if y is not x, then xRy.
- Relations R and R' are ***isomorphic*** iff there is a 1–1 function f from the field of R into and onto the field of R' such that xRy iff $f(x)R'f(y)$ for all x and y in the field of R.

The notion of a sequence of series is roughly that of a well-ordered set. In such cases, we can lay out the members of the well-ordered set as follows:

$$e_1 \Rightarrow e_2 \Rightarrow e_3 \Rightarrow e_4 \Rightarrow e_5 \Rightarrow e_6 \Rightarrow e_7 \Rightarrow \ldots$$

Such a series can either be of finite length, or infinite length. Different lengths correspond to different *ordinal numbers*.

Two well orders or well-ordered sets are said to have the same *order-type* or *ordinal number* iff they are isomorphic (or are the fields of isomorphic relations). Again, this does not fix what ordinal numbers themselves are. Naïvely, it would be natural to define an ordinal number as a set of isomorphic well-ordering relations, or a set of isomorphically well-ordered sets. (Frege did discuss ordinal numbers, but this is how Russell defined them.) However, this is not possible in ZF, for reasons similar to those for cardinal numbers.

We shall pursue the strategy of defining ordinal numbers in ZF as certain well-ordered sets having a given order-type: in particular ordinals are defined as sets ordered by the membership ($\in$) relation. Interestingly, the progression

$$0, \{0\}, \{0, \{0\}\}, \{0, \{0\}, \{0, \{0\}\}\}, \ldots$$

is itself well-ordered by the $\in$-relation. (Notice that each one is a member of all subsequent ones.) This simplifies things, as we can identify the finite ordinal numbers with the finite cardinal numbers in ZF. The set ω, of all such numbers, is itself an ordinal number, though not a finite ordinal.

D. Natural Numbers in ZF

The development is made much easier by the following axiom:

The axiom of regularity / foundation:
ZF7: $(\forall x)(x \neq 0 \rightarrow (\exists y)(y \in x \wedge y \cap x = 0))$

(This axiom was not initially included by Zermelo, or even by Fraenkel, but is now standard. It sometimes also considered part of Z.)

The axiom of foundation says that every non-empty set has a member which has no members in common with it. The practical effect of this axiom is to eliminate the possibility of sets being members of themselves, or being members of their members, etc. This is in keeping with the general ZF conception of sets as being "built up" by iterative processes.

(Since all theorems of Z are theorems of ZF, we continue the theorem numbering from before.)

The Irreflexivity of ∈.
T12. $\vdash_{ZF} (\forall x)\, x \notin x$

Proof:

(1)	1. $x \in x$	Hyp
	2. $x \in \{x\}$	T2, T9 UI, BMP
(1)	3. $x \in x \cap \{x\}$	1, 2, T11 UI, BMP
	4. $\{x\} \neq 0$	2, T4 SL, EG, LL
	5. $(\exists y)(y \in \{x\} \wedge y \cap \{x\} = 0)$	4, ZF7 UI, MP
	6. $d \in \{x\} \wedge d \cap \{x\} = 0$	5 EI
	7. $d = x$	6, T9 UI, SL
	8. $x \cap \{x\} = 0$	6, 7 SL, LL
(1)	9. $x \in 0$	3, 8 LL
	10. $x \notin 0$	T3 UI
	11. $x \notin x$	1, 9, 10 RAA
	12. $(\forall x)\, x \notin x$	11 UG

The Asymmetry of ∈.
T13. $\vdash_{ZF} (\forall x)(\forall y)(x \in y \rightarrow y \notin x)$

Proof:

(1)	1. $x \in y$	Hyp	(For hints, see Hatcher, p. 146.)
	2. $x \in \{x, y\}$	Z5, T2 UI, BMP	
	3. $\{x, y\} \neq 0$	2, T4 SL, EG, CQ, LL	
	4. $(\exists z)(z \in \{x, y\} \wedge z \cap \{x, y\} = 0)$	3, ZF7 UI, UG, BMP	
	5. $d \in \{x, y\} \wedge d \cap \{x, y\} = 0$	4 EI	
	6. $d = x \vee d = y$	5, Z5 UI, SL	
(7)	7. $d = y$	Hyp	
(1,7)	8. $x \in d$	1, 7 LL	
(1,7)	9. $x \in d \cap \{x, y\}$	2, 8, T11 UI, SL	
(1,7)	10. $x \in 0$	5, 9 SL, LL	
	11. $x \notin 0$	T3 UI	
(1)	12. $d \neq y$	7, 10, 11 RAA	
(1)	13. $d = x$	6, 12 DS	
(14)	14. $y \in x$	Hyp	
	15. $y \in \{x, y\}$	Z5, T2 UI, SL	
(1,14)	16. $y \in d$	13, 14 LL	
(1,14)	17. $y \in d \cap \{x, y\}$	15, 16, T11 UI, SL	
(1,14)	18. $y \in 0$	5, 17 Simp, LL	
	19. $y \notin 0$	T3 UI	
(1)	20. $y \notin x$	14, 18, 19 RAA	
	21. $x \in y \rightarrow y \notin x$	1–20 CP	
	22. $(\forall x)(\forall y)(x \in y \rightarrow y \notin x)$	21 UG×2	

Definitions:

t'	for	$t \cup \{t\}$
1	for	$0'$
2	for	$1'$
3	for	$2'$ (etc.)
$\{t_1, \ldots t_n, t_{n+1}\}$	for	$\{t_1, \ldots, t_n\} \cup \{t_{n+1}\}$
$\in$-Trans(t)	for	$(\forall z)(z \in t \rightarrow z \subseteq t)$
$\in$-Con(t)	for	$(\forall z)(\forall y)(z \in t \wedge y \in t \wedge z \neq y \rightarrow z \in y \vee y \in z)$
On(t)	for	$\in$-Trans$(t) \wedge \in$-Con(t)

"On(t)" means "t is an ordinal number".

HOMEWORK 11

Prove:

T14. $\vdash_{ZF} (\forall x)(\forall y)(\forall z)(x \in y \wedge y \in z \rightarrow z \notin x \wedge z \neq x)$

Well ordering of ordinals by $\in$.

T15. $\vdash_{ZF} (\forall x)(\forall y)(x \subseteq y \wedge x \neq 0 \wedge \mathrm{On}(y) \rightarrow (\exists z)(z \in x \wedge (\forall x_1)(x_1 \in x \rightarrow x_1 = z \vee z \in x_1)))$

Proof:

(1)	1. $x \subseteq y \wedge x \neq 0 \wedge \mathrm{On}(y)$	Hyp
	2. $x \neq 0 \rightarrow (\exists y)(y \in x \wedge y \cap x = 0)$	ZF7 UI
(1)	3. $(\exists y)(y \in x \wedge y \cap x = 0)$	1, 2 Simp, MP
(1)	4. $c \in x \wedge c \cap x = 0$	3 EI
(5)	5. $x_1 \in x \wedge x_1 \neq c$	Hyp
(1)	6. $\in$-Con(y)	1, Df. On, Simp
(1)	7. $(\forall x)(\forall y)(x \in y \wedge z \in y \wedge x \neq z \rightarrow x \in z \vee z \in x)$	7 Df. $\in$-Con
(1)	8. $c \in y$	1, 4, Df. $\subseteq$ UI, MP
(1,5)	9. $x_1 \in y$	1, 5, Df. $\subseteq$ UI, MP
(1,5)	10. $c \in x_1 \vee x_1 \in c$	5, 7, 8, 9 UI, SL
(11)	11. $x_1 \in c$	Hyp
(5,11)	12. $x_1 \in c \wedge x_1 \in x$	5, 11 SL
(5,11)	13. $x_1 \in c \cap x$	12, T11 UI, BMP
(1,5,11)	14. $x_1 \in 0$	4, 13 Simp, LL
	15. $x_1 \notin 0$	T3 UI
(1,5)	16. $x_1 \notin c$	11, 14, 15 RAA
(1,5)	17. $c \in x_1$	10, 16 DS
(1)	18. $x_1 \in x \wedge x_1 \neq c \rightarrow c \in x_1$	5–17 CP
(1)	19. $x_1 \in x \rightarrow x_1 = c \vee c \in x_1$	18 SL
(1)	20. $(\forall x_1)(x_1 \in x \rightarrow x_1 = c \vee c \in x_1)$	19 UG
(1)	21. $c \in x \wedge (\forall x_1)(x_1 \in x \rightarrow x_1 = c \vee c \in x_1)$	4, 20 SL
(1)	22. $(\exists z)(z \in x \wedge (\forall x_1)(x_1 \in x \rightarrow x_1 = z \vee z \in x_1))$	21 EG
	23. $x \subseteq y \wedge x \neq 0 \wedge \mathrm{On}(y) \rightarrow (\exists z)(z \in x \wedge (\forall x_1)(x_1 \in x \rightarrow x_1 = z \vee z \in x_1))$	1–22 CP
	24. $(\forall x)(\forall y)(x \subseteq y \wedge x \neq 0 \wedge \mathrm{On}(y) \rightarrow (\exists z)(z \in x \wedge (\forall x_1)(x_1 \in x \rightarrow x_1 = z \vee z \in x_1)))$	23 UG×2

T16. $\vdash_{ZF} \mathrm{On}(0)$

T17. $\vdash_{ZF} (\forall x)(\forall y)(\in\text{-Con}(x) \wedge y \subseteq x \rightarrow \in\text{-Con}(y))$

T18. $\vdash_{ZF} (\forall x)(\forall y)(\mathrm{On}(x) \wedge y \in x \rightarrow \mathrm{On}(y))$

There are three kinds of ordinals: zero, successors, and limit ordinals. ***Successors*** are those of the form $n \cup \{n\}$ for some other ordinal n. ***Limits*** are those such as ω, which have infinitely many other ordinals getting "closer and closer" to it.

Definition: We define a ***natural number*** as an ordinal which is either zero or a successor, and all members of which are either zero or a successor. (The ordinal ω', i.e., $\omega \cup \{\omega\}$ is a successor, but not a natural number, since it has a limit ordinal as member.)

Definitions:

$\mathrm{Sc}(t)$ for $(\exists \chi)(\mathrm{On}(\chi) \wedge \chi' = t)$
$\mathrm{Lim}(t)$ for $\mathrm{On}(t) \wedge t \neq 0 \wedge \neg\mathrm{Sc}(t)$
$N(t)$ for $\mathrm{On}(t) \wedge (t = 0 \vee \mathrm{Sc}(t)) \wedge (\forall \chi)(\chi \in t \rightarrow \chi = 0 \vee \mathrm{Sc}(\chi))$

Following Hatcher, from now on, rather than giving full proofs, I shall often only provide sketches.

T19. $\vdash_{ZF} N(0)$ (=Peano postulate 1)

Proof sketch: Obvious.

T20. $\vdash_{ZF} (\forall x)\, 0 \neq x'$ (=Peano postulate 3)

Proof sketch: $x \in x'$ but $x \notin 0$, so $0 = x'$ is impossible.

T21. $\vdash_{ZF} (\forall x)(\mathrm{On}(x) \rightarrow \mathrm{On}(x'))$

Proof sketch: Suppose $\mathrm{On}(x)$. Hence $\in$-Con(x) and $\in$-Trans(x). Consider any $y \in x'$. Either $y \in x$ or $y = x$. If $y \in x$, since $\in$-Trans(x), $y \subseteq x$, and since $x \subseteq x'$, $y \subseteq x'$. If $y = x$, since $x \subseteq x'$, $y \subseteq x'$. Hence $y \in x' \rightarrow y \subseteq x'$. This holds for any y, so $\in$-Trans(x').

Assume $y \in x'$and $z \in x'$ and $y \neq z$. There are three possibilities: either (i) $y \in x$ and $z = x$, or (ii) $y = x$ and $z \in x$, or (iii) $y \in x$ and $z \in x$. (It cannot be that both $y = x$ and $z = x$ since $y \neq z$.) In case (i), $y \in z$, and (ii) $z \in y$. In case (iii), since $\in$-Con(x), we have $y \in z \vee z \in y$. Hence, $y \in x' \wedge z \in x' \wedge y \neq z \rightarrow y \in z \vee z \in y$. This holds for any y and z, and so $\in$-Con(x'), thus $\mathrm{On}(x')$.

T22. $\vdash_{ZF} (\forall x)(N(x) \rightarrow N(x'))$ (=Peano postulate 2)

Proof sketch: Suppose $N(x)$. It follows that $\mathrm{On}(x)$, and so, by T21, $\mathrm{On}(x')$. Obviously, $\mathrm{Sc}(x')$. Consider any member y of x': either $y \in x$, and so is 0 or a successor, or $y = x$, which is either 0 or a successor, since $N(x)$. So $(\forall y)(y \in x' \rightarrow y = 0 \vee \mathrm{Sc}(y))$ and $N(x')$.

T23. $\vdash_{ZF} (\forall x)(\forall y)(x' = y' \rightarrow x = y)$ (=Peano postulate 4)

Proof sketch: Suppose $x' = y'$, and suppose for reductio $x \neq y$. Since $x \in x'$, we get $x \in y'$. This means that either $x \in y$ or $x \in \{y\}$, but since $x \neq y$, it must be that $x \in y$. By a parallel argument, $y \in x$. But this contradicts T13. Hence $x = y$.

HOMEWORK 12

Prove: T24. $\vdash_{ZF} (\forall x)(\forall y)(N(x) \wedge y \in x \rightarrow N(y))$

The principle of mathematical induction (=Peano postulate 5)

T25. $\vdash_{ZF} \mathscr{A}[0] \wedge (\forall x)(N(x) \wedge \mathscr{A}[x] \rightarrow \mathscr{A}[x']) \rightarrow (\forall x)(N(x) \rightarrow \mathscr{A}[x])$

Proof:

(In the proof, L abbreviates $\{y | y \in x' \wedge \neg\mathscr{A}[x]\}$.)

(1)	1. $\mathscr{A}[0] \wedge (\forall x)(N(x) \wedge \mathscr{A}[x] \to \mathscr{A}[x'])$	Hyp
(2)	2. $N(x)$	Hyp
(3)	3. $\neg\mathscr{A}[x]$	Hyp
	4. $x \in x'$	T2, T9, T10 QL
(3)	5. $x \in L$	3, 4, Z2, Df. L QL
(3)	6. $L \neq 0$	5, T4 QL
(3)	7. $(\exists y)(y \in L \wedge y \cap L = 0)$	6, ZF7, QL
(3)	8. $c \in L \wedge c \cap L = 0$	7 EI
(3)	9. $c \in x' \wedge \neg\mathscr{A}[c]$	8, Z2, Df. L, QL
(1,3)	10. $c \neq 0$	1, 9, Z1 QL
(2)	11. $N(x')$	2, T22 QL
(2,3)	12. $N(c)$	9, 11, T24 QL
(2,3)	13. $c = 0 \vee \mathrm{Sc}(c)$	12, Df. N, SL
(1,2,3)	14. $(\exists x)(\mathrm{On}(x) \wedge x' = c)$	10, 13, Df. Sc SL
(1,2,3)	15. $\mathrm{On}(d) \wedge d' = c$	14 EI
(1,2,3)	16. $d \in c$	4, 15 QL, LL
(1,2,3)	17. $N(d)$	12, 16, T24 QL
(1,2,3)	18. $d \notin L$	8, 16, T3, T11 QL
(1,2,3)	19. $d \notin x' \vee \mathscr{A}[d]$	18, Df. L, Z2 QL
(1,2,3)	20. $d \in x'$	9, 11, 16, Dfs. QL
(1,2,3)	21. $\mathscr{A}[d]$	19, 20 SL
(1)	22. $N(d) \wedge \mathscr{A}[d] \to \mathscr{A}[d']$	1 QL
(1,2,3)	23. $\mathscr{A}[d']$	17, 21, 22 SL
(1,2,3)	24. $\neg\mathscr{A}[d']$	9, 15 SL, LL
(1,2)	25. $\mathscr{A}[x]$	3, 23, 24 RAA
(1)	26. $N(x) \to \mathscr{A}[x]$	2–25 CP
(1)	27. $(\forall x)(N(x) \to \mathscr{A}[x])$	26 UG
	28. $\mathscr{A}[0] \wedge (\forall x)(N(x) \wedge \mathscr{A}[x] \to \mathscr{A}[x'])$	
	$\to (\forall x)(N(x) \to \mathscr{A}[x])$	1–27 CP

T26. $(\forall x)(\forall y)(N(x) \wedge y \in x \to y' \in x')$

That the Peano postulates hold in ZF is enough to establish that the domain of quantification for any of its models must be infinite. One can prove, e.g., that $0 \neq 1$, $1 \neq 2$, $2 \neq 3$, etc. However, this does not yet establish the existence of any *set* having an infinite number of members. This is the role of the next axiom, which Hatcher puts in this form:

The axiom of infinity:

ZF8: $(\forall x)(x \in \omega \leftrightarrow N(x))$

Hence, ω is the set of natural numbers.

T27. $\vdash_{ZF} \mathrm{On}(\omega)$

For a sketch of a proof of this, see Hatcher, p. 155.

T28. $\vdash_{ZF} \mathrm{Lim}(\omega)$

Proof sketch: ω is a limit ordinal if it is neither 0 nor a successor. Since ω has 0 as a member, it cannot be 0. Suppose for *reductio* that it is a successor. Then there is some c such that $\omega = c'$. Now $c \in c'$, so $c \in \omega$. This means that c is a natural number, then so is c', i.e., ω, but that would mean that $\omega \in \omega$, contrary to T12. Hence $\neg\mathrm{Sc}(\omega)$, and so $\mathrm{Lim}(\omega)$.

Consider now the ordinal number ω'. It is a successor, but not a natural number. It too has a successor ω'', and so on. However, in order to obtain any further limit ordinals, such as $\omega + \omega$ (a.k.a. $\omega 2$), we need a principle postulating the existence of sets stronger than Z2 alone. Fraenkel suggested the following (and this is the primary difference between Z and ZF):

The axiom of replacement:

ZF9: $(\forall x)((\forall y)(\forall z_1)(\forall z_2)(\mathscr{A}[y, z_1] \wedge \mathscr{A}[y, z_2] \to z_1 = z_2) \to (\exists y)(\forall z)(z \in y \leftrightarrow (\exists y_1)(y_1 \in x \wedge \mathscr{A}[y_1, z])))$,

where z_2 does not become bound in $\mathscr{A}[y, z_2]$.

This is called the axiom of *replacement* because, beginning with any set x, it postulates the existence of a set y, got by replacing members of x with the result of some function on those members. (Above, interpret $\mathscr{A}[t, u]$ as meaning that the function in question assigns u to t as "argument".)

This is a powerful axiom, and indeed, strictly stronger than the axiom of separation. Indeed, the usual form of the axiom of separation, viz:

$$(\forall y)(\exists z)(\forall x)(x \in z \leftrightarrow x \in y \wedge \mathscr{A}[x])$$

can be proven from ZF9 straightaway.

ZF9 completes the axioms of ZF.

E. Relations, Functions and Recursion in ZF

In set theory, relations are often treated as sets of ordered pairs. (I might prefer the "extension of a relation" to "relation", but I am out-numbered.)

Definitions:

(t is relation):
$R(t)$ for $(\forall \chi)(\chi \in t \rightarrow (\exists y)(\exists z)\, \chi = \langle y, z\rangle)$

(t is a function):
$F(t)$ for $R(t) \wedge (\forall \chi)(\forall y)(\forall z)(\langle \chi, y\rangle \in t \wedge \langle \chi, z\rangle \in t \rightarrow y = z)$

(domain of t):
$D(t)$ for $\{\chi | \chi \in \bigcup(\bigcup(t)) \wedge (\exists y)\langle \chi, y\rangle \in t\}$

(range of t):
$I(t)$ for $\{\chi | \chi \in \bigcup(\bigcup(t)) \wedge (\exists y)\langle y, \chi\rangle \in t\}$

(value of function t for u as argument):
$t``u$ for $\bigcup(\{\chi | \chi \in I(t) \wedge \langle u, \chi\rangle \in t\})$

(Cartesian product of t and u):
$t \times u$ for $\{\chi | \chi \in \wp(\wp(t \cup u)) \wedge (\exists y)(\exists z)(y \in t \wedge z \in u \wedge \chi = \langle y, z\rangle)\}$

Indirectly, the axiom of replacement allows us to deduce the existence of a relation, considered as a set of ordered-pairs, corresponding to any "function" $\mathscr{A}[y,z]$, with any set x as domain.

We can now deduce the existence of recursively defined functions. Informally, we say that a function g on the natural numbers is defined by ***simple recursion*** in terms of a value a and a function f when $g(0) = a$ and for all other natural numbers, $g(n+1) = f(g(n))$.

The *principle of simple recursion* asserts the existence of function g as a set of ordered pairs for any function f and value a. In ZF:

T29. $\vdash_{\text{ZF}} (\forall x)(\forall y)(F(x) \wedge I(x) \subseteq D(x) \wedge y \in D(x) \rightarrow$
$(\exists! z)(F(z) \wedge \omega = D(z) \wedge I(z) \subseteq D(x) \wedge z``0 = y \wedge$
$(\forall y_2)(y_2 \in \omega \rightarrow z``y_2' = x``(z``y_2))))$

Proof sketch: Suppose $F(x) \wedge I(x) \subseteq D(x) \wedge y \in D(x)$. Now consider the set of relations $\mathscr{E}$ defined as:

$$\{z | z \in \wp(\omega \times D(x)) \wedge \langle 0, y\rangle \in z \wedge (\forall x_2)(\forall y_2)(\langle x_2, y_2\rangle \in z \rightarrow \langle x_2', x``y_2\rangle \in z)\}$$

This is the set of all relations that contain $\langle 0, y\rangle$ and always the ordered pair $\langle x_2', x``y_2\rangle$ whenever it contains $\langle x_2, y_2\rangle$. Consider now the set of ordered pairs $\bigcap(\mathscr{E})$, i.e., what all members of $\mathscr{E}$ have in common. We can prove that $\bigcap(\mathscr{E})$ represents a function having the desired traits of the consequent (and that it is the only such function). For a fuller sketch that it does, see Hatcher, pp. 161–62.

This proof also gives us a recipe for defining simple recursive functions among the natural numbers. E.g., we *might* introduce the following definitions:

S for $\{x | x \in \omega \times \omega \wedge (\exists y)(x = \langle y, y'\rangle)\}$
$+_t$ for $\bigcap(\{z | z \in \wp(\omega \times \omega) \wedge \langle 0, t\rangle \in z \wedge (\forall x_2)(\forall y_2)(\langle x_2, y_2\rangle \in z \rightarrow \langle x_2', S``y_2\rangle \in z)\})$
$(t + u)$ for $+_t``u$

HOMEWORK 13

Give similar definitions for multiplication. I.e., first define $\cdot_t$ as the set of ordered pairs $\langle n, m\rangle$ for which t times n is m, and then define $t \cdot u$. Hint: use $+_t$.

The downside to such definitions of addition and multiplication is that they are (correctly) defined only for the natural numbers. Nevertheless with them we could obtain analogues to the remaining theorems of system S:

$\vdash_{\text{ZF}} (\forall x)(N(x) \rightarrow x + 0 = x)$
$\vdash_{\text{ZF}} (\forall x)(\forall y)(N(x) \wedge N(y) \rightarrow x + y' = (x + y)')$
$\vdash_{\text{ZF}} (\forall x)(N(x) \rightarrow x \cdot 0 = 0)$
$\vdash_{\text{ZF}} (\forall x)(\forall y)(N(x) \wedge N(y) \rightarrow x \cdot y' = (x \cdot y) + x)$

From this, it follows that ZF can capture all of Peano arithmetic, and hence can represent all recursive functions and express all recursive relations. (Unfortunately, ZF has its own Gödel sentence, and so has undecidable sentences.)

F. Infinite Ordinals

Speaking loosely, if a and b are ordinal numbers, the ordinal number $a + b$ is regarded as the order-type had by a well-ordered sequence got by appending a sequence with ordinal number b after a sequence with ordinal number a.

Since ω is the ordinal number of sequences that could be depicted as follows:

$$\bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \ldots$$

It follows that $\omega + 1$, or ω' is the ordinal number of a series of the following form:

$$\bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \ldots \bullet$$

Here there is a "last" member of the series, but there are infinitely many members of the series preceding the last member.

$\omega + 2$ or ω'' is the ordinal number of a series of the following form:

$$\bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \ldots \bullet \Rightarrow \bullet$$

In ZF, we define ordinals as exemplar well-ordered sets having a certain order type. Thus ω is the set containing the members 0, 1, 2, 3, 4, ...
ω' is the set containing 0, 1, 2, 3, 4, ... , ω.
ω'' is the set set containing 0, 1, 2, 3, 4, ... ω, ω'.
We can similarly define ω''', ω'''', etc.

The ordinal number $\omega + \omega$ or $\omega 2$ is the ordinal number of well-ordered series having the form:

$$\bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \ldots \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \bullet \Rightarrow \ldots$$

We cannot obtain $\omega + \omega$ using the definition of + given in the previous section: it is defined only over natural numbers. Nevertheless we can use the axiom of replacement to prove the existence of such ordinals.

Consider the wff below, abbreviated $M(y,z)$:

$$N(y) \land F(z) \land D(z) = y \land z``0 = \omega \land (\forall y_2)(y_2' \in y \rightarrow z``y_2' = (z``y_2)')$$

Intuitively, this says that y is a natural number, and z is a function whose domain is y (remember that a natural number is the set of all natural numbers less than it), and whose value for 0 is ω and whose value for 1 is ω', and whose value for 2 is ω'' and so on. We can prove that such a z exists for every natural number:

$$\vdash_{\mathrm{ZF}} (\forall y)(y \in \omega \rightarrow (\exists z)\, M(y,z))$$

We can also establish that this wff is functional, i.e.:

$$\vdash_{\mathrm{ZF}} (\forall y)(\forall z_1)(\forall z_2)(M(y,z_1) \land M(y,z_2) \rightarrow z_1 = z_2)$$

It follows by the axiom of replacement that:

$$\vdash_{\mathrm{ZF}} (\exists x)(\forall z)(z \in x \leftrightarrow (\exists y)(y \in \omega \land M(y,z)))$$

Call the set postulated by this theorem m. Then m is the set of all such functions from natural numbers to corresponding successors of ω. The sum set $\bigcup(m)$ is thus the set of ordered pairs of the form $\langle 0,\omega\rangle$, $\langle 1,\omega'\rangle$, $\langle 2,\omega''\rangle$, and so on. Its range, $I(\bigcup(m))$ is the set of all successors of ω, and hence $\omega 2$ or $\omega + \omega$ would be the well-ordered set $\omega \cup I(\bigcup(m))$.

(This does not allow us to define $\omega 2$ outright in ZF, since "m" is a dummy constant arrived at from EI, not an actual constant of the language. But clearly such large ordinals are included in the domain of quantification for the language.)

By similar means, we could obtain $\omega + \omega + \omega$, and further, ω^2, or $\omega + \omega + \ldots$, the ordinal number corresponding to a well-ordered series of the form:

$$\bullet \Rightarrow \bullet \Rightarrow \ldots \bullet \Rightarrow \bullet \Rightarrow \ldots \bullet \Rightarrow \bullet \Rightarrow \ldots \bullet \Rightarrow \bullet \Rightarrow \ldots \bullet \Rightarrow \bullet \Rightarrow \ldots\ \ldots$$

(I.e., a denumerable number of denumerable sequences following after one another.)

G. Cardinal Numbers and the Axiom of Choice

Cardinal numbers measure the *size* of a set. We say that sets have the same size, or same cardinality, when there is a 1–1 function having one set as domain and the other as range.

Definitions:

$F^1(t)$	for	$F(t) \wedge (\forall \chi)(\forall y)(\forall z)(\langle \chi, z\rangle \in t \wedge \langle y, z\rangle \in t \rightarrow \chi = y)$
$t \cong u$	for	$(\exists \chi)(F^1(\chi) \wedge t = D(\chi) \wedge u = I(\chi))$
$t \ncong u$	for	$\neg t \cong u$

It then follows (by proofs similar to those for F):

(Ref$\cong$)	$\vdash_{ZF} (\forall x)\, x \cong x$
(Sym$\cong$)	$\vdash_{ZF} (\forall x)(\forall y)(x \cong y \rightarrow y \cong x)$
(Trans$\cong$)	$\vdash_{ZF} (\forall x)(\forall y)(\forall z)(x \cong y \wedge y \cong z \rightarrow x \cong z)$

Hatcher writes "t Sm u" instead of "$t \cong u$".

HOMEWORK 14

Informally sketch a proof showing that: $\vdash_{ZF} \omega \cong \omega'$

Since the Frege-Russell definition of cardinal numbers cannot be used in ZF, we use representative sets having a given cardinality. In particular, cardinals are identified with particular ordinal numbers having a given cardinality. Because multiple ordinals (such as ω and ω') can have the same cardinality, we must select only one such number as *the* cardinal number for that cardinality. It is convenient to pick the *least* ordinal (as ordered by membership).

Definition:

Card(t) for On(t) $\wedge (\forall \chi)(\chi \in t \rightarrow \chi \ncong t)$

We can also define greater-than and less-than relations among sets in virtue of their cardinality as follows:

Definitions:

$t \preceq u$	for	$(\exists \chi)(\chi \subseteq u \wedge t \cong \chi)$
$t \prec u$	for	$t \preceq u \wedge \neg u \preceq t$

We can also define what is for a set to be finite or infinite:

Fin(t)	for	$(\exists \chi)(\chi \in \omega \wedge t \cong \chi)$
Inf(t)	for	$\neg$Fin(t)

One potential downfall of this way of defining cardinals is that we do not yet have any guarantee that every set is of the same cardinality as an ordinal. A proof of this result, viz.:

$\vdash_{ZFC} (\forall x)(\exists y)(\mathrm{On}(y) \wedge x \cong y)$

requires the ***axiom of choice***. The usual formulation of this axiom is as follows:

(AC) $(\forall x)(\exists y)(F(y) \wedge (\forall z)(z \subseteq x \wedge z \neq 0 \rightarrow y``z \in z))$

This states that for every set x, there is a function y that has as value, for any nonempty subset z of x as argument, some member of that subset. This gives us a way of "selecting" a member from every subset.

Hatcher instead formulates the axiom in a more convenient way as follows:

ZFC10: $(\forall x)(x \neq 0 \rightarrow \sigma(x) \in x)$

Definition: ***ZFC*** is the system obtained from ZF by adding ZFC10 as an axiom.

Here "σ" is taken as a primitive function constant, which returns, for *any* set as argument some member of that set. Since "σ" is a primitive function constant, it is not defined as a set of ordered pairs, and indeed, no such set as the set of all ordered pairs $\langle z, \sigma(z)\rangle$ can be proven to exist, although by means of the axiom of replacement, for any set x, one can establish the existence of a set of ordered pairs of this form where $z \subseteq x$, i.e., one can establish (AC) above.

ZFC10 has been shown to be independent of the axioms of ZF, and indeed, both (AC) and its negation are consistent with ZF1–ZF9.

Very rough sketch of proof of:
$\vdash_{ZFC} (\forall x)(\exists y)(\mathrm{On}(y) \wedge x \cong y)$

Let x be any set. For every ordinal t we can define a function f^t from t to members of x as follows:

$$\bigcap(\{z | z \in \wp(t \times x) \wedge (\forall y)(y \in t \rightarrow \langle y, \sigma(x - \{x_2 | x_2 \in x \wedge (\exists y_2)(y_2 \in y \wedge \langle y_2, x_2\rangle \in z)\})\rangle \in z)\})$$

I.e., what is in common between all sets of ordered pairs sharing the following:

$$\langle 0, \sigma(x)\rangle$$
$$\langle 1, \sigma(x - \{\sigma(x)\})\rangle$$
$$\langle 2, \sigma(x - \{\sigma(x), \sigma(x - \{\sigma(x)\})\})\rangle$$
etc.

Call $\sigma(x)$, "x:0", and $\sigma(x - \{\sigma(x)\})$, "$x$:1", and $\sigma(x - \{\sigma(x), \sigma(x - \{\sigma(x)\})\})$, "$x$:2", and so on.

If t is greater than ω, then this intersection will also contain, e.g.:

$$\langle \omega, \sigma(x - \{x{:}0, x{:}1, x{:}2, x{:}3, \dots\})\rangle$$

Now consider the wff below, abbreviated $O(y, z)$:

$$(\mathrm{On}(z) \wedge f^{z'}``(z) = y \wedge (\forall z_2)(z_2 \in z \to f^{z'}``(z_2) \neq y)) \vee (z = 0 \wedge (\forall z_2)(\mathrm{On}(z_2) \to f^{z_2'}``(z_2) \neq y))$$

This says that z is the least ordinal such that $x{:}z = y$ or $z = 0$ if there is no such ordinal. This wff is functional, i.e., for a given member y of x it cannot hold for more than one z, and hence, by the axiom of replacement, there is a set s such that:

$$(\forall z)(z \in s \leftrightarrow (\exists y)(y \in x \wedge O(y, z)))$$

This set s is thus the set of all ordinals which are the least ordinals corresponding to a given member of x in the ordering generated by the choice function σ.

Consider now the set $\bigcup(s)$: the set of all members of any ordinal in s. Notice that since ordinals only have other ordinals as members, $\bigcup(s)$ is the set of all ordinals "less" than any member of s. In fact, $\bigcup(s)$ is itself an ordinal, at least as great as any ordinal in s. Hence $\bigcup(s)'$ is greater than any ordinal in s.

It follows that $x \subseteq I(f^{\bigcup(s)'})$. For if not, then notice that $x - I(f^{\bigcup(s)'})$ would be non-empty. Since $f^{\bigcup(s)''}``\bigcup(s)' = \sigma(x - I(f^{\bigcup(s)'}))$, it would follow that $O(\sigma(x - I(f^{\bigcup(s)'})), \bigcup(s)')$, and hence that $\bigcup(s)'$ was a member of s, which is impossible. Basically, this says that the ordinal $\bigcup(s)'$ is big enough to "order" the members of x.

Now consider the set $\{y | y \in \bigcup(s)'' \wedge x \subseteq I(f^y)\}$. This is the set of all ordinals big enough to order x. This set is non-empty, since $\bigcup(s)'$ is a member. By the axiom of regularity, there is an ordinal b, which is a member of this set and which contains no members in common with this set. (In effect, b the lowest such ordinal.) Hence $\mathrm{On}(b)$ and $x \subseteq I(f^b)$.

Now consider the function f^b. By the definition of f^t, the domain of f^b must be b, and the range of f^b must be a subset of x. Since $x \subseteq I(f^b)$ and $I(f^b) \subseteq x$, it follows that $x = I(f^b)$.

We now show that f^b is a 1–1 function. Suppose that $\langle y_1, x_1\rangle \in f^b$ and $\langle y_2, x_1\rangle \in f^b$. Since y_1 and y_2 are members of b, which is an ordinal, they are ordinals themselves. Since ordinals are well-ordered by $\in$, we have either $y_2 \in y_1$ or $y_1 \in y_2$ or $y_1 = y_2$. Suppose that $y_2 \in y_1$. By the definition of f^b, since $\langle y_1, x_1\rangle \in f^b$, it must be that $x_1 = \sigma(x - \{x_2 | x_2 \in x \wedge (\exists y_2)(y_2 \in y_1 \wedge \langle y_2, x_2\rangle \in f^b)\})$. But notice we have $y_2 \in y_1 \wedge \langle y_2, x_1\rangle \in f^b$, so $(\exists y_2)(y_2 \in y_1 \wedge \langle y_2, x_1\rangle \in f^b)$ by EG. But this means that $x_1 \in \{x_2 | x_2 \in x \wedge (\exists y_2)(y_2 \in y_1 \wedge \langle y_2, x_2\rangle \in f^b)\}$ and hence $x_1 \notin x - \{x_2 | x_2 \in x \wedge (\exists y_2)(y_2 \in y_1 \wedge \langle y_2, x_2\rangle \in f^b)\}$. By ZFC10, it follows that $x - \{x_2 | x_2 \in x \wedge (\exists y_2)(y_2 \in y_1 \wedge \langle y_2, x_2\rangle \in f^b)\}$ must be the empty set. This is only possible only if x is the empty set, or $x \subseteq \{x_2 | x_2 \in x \wedge (\exists y_2)(y_2 \in y_1 \wedge \langle y_2, x_2\rangle \in f^b)\}$. But x cannot be the empty set, since $x_1 \in x$. Moreover, it cannot be that $x \subseteq \{x_2 | x_2 \in x \wedge (\exists y_2)(y_2 \in y_1 \wedge \langle y_2, x_2\rangle \in f^b)\}$, since this would mean that $y_1 = b$, which is impossible, since $y_1 \in b$. By an exactly parallel argument, it follows that it cannot be that $y_1 \in y_2$. So $y_1 = y_2$, and $F^1(f^b)$.

Then $F^1(f^b) \wedge D(f^b) = b \wedge I(f^b) = x$. Thus $b \cong x$. Hence by (Sym$\cong$), $x \cong b$. So $(\exists y)(\mathrm{On}(y) \wedge x \cong y)$. This holds for any set x, and so $\vdash_{\mathrm{ZFC}} (\forall x)(\exists y)(\mathrm{On}(y) \wedge x \cong y)$, QED.

HOMEWORK 15

Prove: $\vdash_{\mathrm{ZFC}} (\forall x)(\exists y)(\mathrm{Card}(y) \wedge x \cong y)$

Hint: apply the axiom of regularity to the set of all ordinals equinumerous with x less than or equal to (i.e., in the successor of) the ordinal given by the theorem proven above.

Definitions:

(These match the definitions given informally on p. 40.)

$$
\begin{array}{rll}
t\text{-Irr}(u) & \text{for} & R(t) \wedge (\forall \chi)(\chi \in u \to \langle \chi, \chi \rangle \notin t) \\
t\text{-Trans}(u) & \text{for} & R(t) \wedge (\forall \chi)(\forall y)(\forall z)(\chi \in u \wedge y \in u \wedge z \in u \to \\
 & & \qquad (\langle \chi, y \rangle \in t \wedge \langle y, z \rangle \in t \to \langle \chi, z \rangle \in t)) \\
t\text{-Part}(u) & \text{for} & t\text{-Irr}(u) \wedge t\text{-Trans}(u) \\
t\text{-Con}(u) & \text{for} & R(t) \wedge (\forall \chi)(\forall y)(\chi \in u \wedge y \in u \wedge \chi \neq y \to \\
 & & \qquad \langle \chi, y \rangle \in t \vee \langle y, \chi \rangle \in t) \\
t\text{-Tot}(u) & \text{for} & t\text{-Irr}(u) \wedge t\text{-Trans}(u) \wedge t\text{-Con}(u) \\
t\text{-We}(u) & \text{for} & t\text{-Tot}(u) \wedge (\forall \chi)(\chi \subseteq u \wedge \chi \neq 0 \to \\
 & & \qquad (\exists y)(y \in \chi \wedge (\forall z)(z \in \chi \wedge z \neq y \to \langle y, z \rangle \in t)))
\end{array}
$$

Consequences of the above results

1) *The trichotomy principle:*

(TP) $\vdash_{\text{ZFC}} (\forall x)(\forall y)(x \preceq y \vee y \preceq x)$

2) *The well ordering principle:*

(WO) $\vdash_{\text{ZFC}} (\forall x)(\exists y)\, y\text{-We}(x)$

In fact, both the well-ordering principle and the trichotomy principle are *equivalent* to the axiom of choice, in the sense that (AC) would folow from either of them, if added as an axiom to ZF. I.e.:

$$\vdash_{\text{ZF}} \text{(AC)} \longleftrightarrow \text{(TP)}$$
$$\vdash_{\text{ZF}} \text{(AC)} \longleftrightarrow \text{(WO)}$$

Notice that these results hold in weaker ZF, not just in ZFC.

Other equivalent principles include:

3) *The multiplicative axiom:*

(MP) $\vdash_{\text{ZFC}} (\forall x)((\forall y)(y \in x \to y \neq 0) \wedge (\forall y)(\forall z)(y \in x \wedge z \in x \wedge y \neq z \to y \cap z = 0) \to (\exists y)(\forall z)(z \in x \to (\exists! x_1)(x_1 \in y \cap z)))$

I.e., for every set of disjoint, non-empty sets, there is a set having exactly one member from each.

4) *Zorn's lemma:*

(ZL) $\vdash_{\text{ZFC}} (\forall x)(\forall y)((y\text{-Part}(x) \wedge (\forall z)(z \subseteq x \wedge y\text{-Tot}(z) \to (\exists x_1)(x_1 \in x \wedge (\forall z_1)(z_1 \in z \to z_1 = x_1 \vee \langle z_1, x_1 \rangle \in y)))) \to (\exists x_1)(x_1 \in x \wedge (\forall z_1)(z_1 \in x \to \langle x_1, z_1 \rangle \notin y)))$

I.e., every partially ordered set x which is such that every totally ordered subset has an "upper bound" has a "maximal" element.

(The theorem proven above, that every set is equinumerous with some ordinal, is also equivalent to (AC) in ZF.)

Additional consequences of the axiom of choice

5) Every infinite set has a denumerable subset:

$\vdash_{\text{ZFC}} (\forall x)(\text{Inf}(x) \to \omega \preceq x)$

6) Every infinite set is equinumerous with one of its proper subsets (i.e., is "Dedekind-infinite"):

$\vdash_{\text{ZFC}} (\forall x)(\text{Inf}(x) \to (\exists y)(y \subseteq x \wedge y \neq x \wedge x \cong y))$

It has been proven by Kurt Gödel that ZFC is consistent if ZF consistent.

H. Cantor's Theorem and the Continuum Hypothesis

Cantor proved the result:

$$(\forall x)\, x \ncong \wp(x)$$

Informal sketch: Suppose $x \cong \wp(x)$. Each member of x can be put in 1–1 correlation with a subset of x. Some members of x will be members of the subset with which they're correlated. Some will not. Consider the subset of x of all members of x not members of that to which they're correlated. Call it w. Some member y of x must be correlated with w. Is y a member of w? It is just in case it is not. Contradiction. Hence the assumption that $x \cong \wp(x)$ must be mistaken.

Although quite long, this proof can be given in ZF. It follows that there are different sizes of infinite sets, since, e.g.:

$$\omega \ncong \wp(\omega)$$

The cardinality of $\wp(\omega)$ is sometimes called the cardinality of the ***continuum*** since it can be proven that there are this many real numbers, or points on a geometric line. It is also called c, $\beth_1$, or $2^{\aleph_0}$. It represents the number of different subsets of natural numbers there are. (Each subset is the result of making $\aleph_0$ many yes/or or in-or-out choices: hence there are $2^{\aleph_0}$ possible choice combinations.)

Cantor conjectured that there are no sets with cardinalities *in between* $\aleph_0$ and $2^{\aleph_0}$, or that $\aleph_1 = 2^{\aleph_0}$. This has come to be known as the ***Continuum Hypothesis***. Since, in ZF, ω is our representative ordinal with cardinality $\aleph_0$, we can represent the hypothesis as follows:

(CH) $\neg(\exists x)(\omega \prec x \wedge x \prec \wp(\omega))$

It has been shown (by Paul Cohen) that the above is independent of the axioms of ZF and ZFC. However, it has also been shown (by Kurt Gödel) that it is not inconsistent with the axioms of ZF or ZFC. Neither it nor its negation can be proven in ZF(C).

The same holds for the ***Generalized Continuum Hypothesis***, which suggests that there is never any sets of infinite cardinality in between that of an infinite set and its powerset:

(GCH) $(\forall y)(\mathrm{Inf}(y) \rightarrow \neg(\exists x)(y \prec x \wedge x \prec \wp(y)))$

These hypotheses are controversial, but given their consistency, they are sometimes suggested as additional axioms for set theory.

It is perhaps worth noting that (AC) is provable in ZF + (GCH).

XVII. von Neumann/Bernays/Gödel (NBG) and Mostowski/Kelley/Morse (MKM) Set Theories

Rivaling ZF in popularity is a theory of sets originally formulated by John von Neumann (1925) and further specified and examined by Paul Bernays (1937) and Kurt Gödel (1940).

The chief hallmark of NBG is its distinction between *sets* and *proper classes*. Sets are collections of objects, postulated to exist when built up by iterative processes, much like the sets of ZF. Classes are postulated to exist as the extensions of certain concepts applicable to sets. Some classes are sets, some are not. In particular, a class is a set when it a member of a class. A *proper class* is a class that is not a set. While a proper class can have members, it is not itself a member of any classes.

Generally, proper classes are very large collections, such as the universal class (of which all sets are members), and the class of all ordinals numbers, etc.

A. Syntax

The syntax of NBG is the same as that of ZF, except that we shall use uppercase letters W, X, Y and Z (with or without numerical subscripts) for individual variables. (This does not preclude NBG from being a first-order system, since these are officially the only sort of variable, and the difference in notation is trivial.)

We use $\mathscr{X}$, $\mathscr{Y}$, $\mathscr{Z}$ in the metalanguage schematically for such variables.

Definitions:

$M(t)$ for $(\exists \mathscr{X})\, t \in \mathscr{X}$
$Pr(t)$ for $\neg M(t)$

$M(t)$ means "t is a set" (from the German word *Menge*). "$Pr(t)$" means "t is a proper class".

Lowercase letters are introduced as abbreviations for quantification *restricted to sets*.

Definitions:

Where χ is a lowercase letter used as variable (x, y, z, etc.) written with or without a numerical subscript, and $\mathscr{X}$ is the first (uppercase) variable not in $\mathscr{A}[\chi]$:

$(\forall \chi)\, \mathscr{A}[\chi]$ for $(\forall \mathscr{X})(M(\mathscr{X}) \rightarrow \mathscr{A}[\mathscr{X}])$
$(\exists \chi)\, \mathscr{A}[\chi]$ for $(\exists \mathscr{X})(M(\mathscr{X}) \wedge \mathscr{A}[\mathscr{X}])$
$\{\chi | \mathscr{A}[\chi]\}$ for $\{\mathscr{X} | \mathscr{A}[\mathscr{X}]\}$

Definition: A wff $\mathscr{A}$ is said to be ***predicative*** iff all quantifiers in $\mathscr{A}$ are restricted to quantification over sets.

For the most part, it is harmless to treat the restricted quantifiers as normal quantifiers for a different type of object, with, e.g., rules like UI, UG, EG, etc., so long as in proofs, the different types of variables are kept straight. Notice however that one needs $M(t)$ to instantiate a variable of the form $(\forall x)\ldots x\ldots$ to t.

B. Formulation

The proper axioms of NBG are the following:

NBG0. $(\forall x)(x \in \{\chi|\mathscr{A}[\chi]\} \leftrightarrow \mathscr{A}[x])$, where $\mathscr{A}[\chi]$ is any predicative wff.

NBG1. $(\forall X)(\forall Y)(X = Y \rightarrow (\mathscr{A}[X,X] \rightarrow \mathscr{A}[X,Y]))$, where Y does not occur bound in $\mathscr{A}[X,X]$.

NBG2. $(\forall x)M(\{y|y \in x \wedge \mathscr{A}[y]\})$, where $\mathscr{A}[y]$ is predicative.

NBG3. $M(0) \wedge (\forall x)\, x \notin 0$

NBG4. $(\forall x)(M(\wp(x)) \wedge (\forall y)(y \in \wp(x) \leftrightarrow y \subseteq x))$

NBG5. $(\forall x)(\forall y)(M(\{x,y\}) \wedge (\forall z)(z \in \{x,y\} \leftrightarrow z = x \vee z = y))$

NBG6. $(\forall x)(M(\bigcup(x)) \wedge (\forall y)(y \in \bigcup(x) \leftrightarrow (\exists z)(z \in x \wedge y \in z)))$

NBG7. $(\forall X)(X \neq 0 \rightarrow (\exists y)(y \in X \wedge y \cap X = 0))$

NBG8. $M(\omega) \wedge (\forall x)(x \in \omega \leftrightarrow N(x))$

NBG9. $(\forall x)((\forall y)(\forall z_1)(\forall z_2)(\mathscr{A}[y,z_1] \wedge \mathscr{A}[y,z_2] \rightarrow z_1 = z_2) \rightarrow (\exists y)(\forall z)(z \in y \leftrightarrow (\exists y_1)(y_1 \in x \wedge \mathscr{A}[y_1,z])))$, where $\mathscr{A}[y,z_1]$ is predicative and z_2 does not become bound in $\mathscr{A}[y,z_2]$.

Notice that the axioms NBG1–9 correspond to ZF1–9 save that the postulate the existence of *sets* whose members are *sets* having certain conditions. NBG0 postulates the existence of classes. It appears superficially similar to the problematic axiom schema of System F, but notice that even without the restriction to predicative wffs, NBG0 really reads:

$$(\forall X)(M(X) \rightarrow (X \in \{\chi|\mathscr{A}[\chi]\} \leftrightarrow \mathscr{A}[X]))$$

HOMEWORK 16

Prove: $\vdash_{\text{NBG}} Pr(\{y|y \notin y\})$

Although, as formulated above, NBG0, NBG1, NBG2 and NBG9 are schemata, it is possible to formulate NBG using a finite number of axioms. This is easiest if we abandon the use of the vbto $\{\chi|\mathscr{A}[\chi]\}$ and the constant function signs $\wp$, $\bigcup$, etc.

Formulated this way, the following are called "the axioms of class existence":

$(\exists X)(\forall y)(\forall z)(\langle y,z\rangle \in X \leftrightarrow y \in z)$
$(\forall X)(\forall Y)(\exists Z)(\forall x)(x \in Z \leftrightarrow x \in X \wedge x \in Y)$
$(\forall X)(\exists Y)(\forall z)(z \in Y \leftrightarrow z \notin X)$
$(\forall X)(\exists Y)(\forall x)(x \in Y \leftrightarrow (\exists y)(\langle x,y\rangle \in X))$
$(\forall X)(\exists Y)(\forall x)(\forall y)(\langle x,y\rangle \in Y \leftrightarrow x \in X)$
$(\forall X)(\exists Y)(\forall x)(\forall y)(\forall z)(\langle x,y,z\rangle \in Y \leftrightarrow \langle y,z,x\rangle \in X)$
$(\forall X)(\exists Y)(\forall x)(\forall y)(\forall z)(\langle x,y,z\rangle \in Y \leftrightarrow \langle x,z,y\rangle \in X)$

All instances of the following version of NBG0 are then derivable as theorems:

$(\exists Y)(\forall x)(x \in Y \leftrightarrow \mathscr{A}[x])$, for any predicative $\mathscr{A}[x]$.

The other axioms are formulated as follows:

LL:
$(\forall X)(\forall Y)(X = Y \rightarrow (\forall Z)(X \in Z \rightarrow Y \in Z))$
Pairing:
$(\forall x)(\forall y)(\exists z)(\forall w)(w \in z \leftrightarrow w = x \vee w = y)$
Null set:
$(\exists x)(\forall y)\, y \notin x$
Sum set:
$(\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow (\exists w)(z \in w \wedge w \in y))$
Power set:
$(\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow z \subseteq x)$
Separation:
$(\forall x)(\forall X)(\exists y)(\forall z)(z \in y \leftrightarrow z \in x \wedge z \in X)$
Infinity:

$(\exists x)((\exists y)(y \in x \wedge (\forall z) z \notin y) \wedge$
$(\forall y)(y \in x \rightarrow (\exists z)(z \in x \wedge (\forall w)(w \in z \leftrightarrow w \in y \vee w = y))))$
Replacement:
$(\forall X)(F(X) \rightarrow (\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow (\exists w)(\langle w,z\rangle \in X \wedge w \in x)))$
Regularity:
$(\forall X)(X \neq 0 \rightarrow (\exists y)(y \in X \wedge y \cap X = 0))$

This marks a contrast from ZF, which cannot be axiomatized with a finite number of axioms.

Mostowski-Kelley-Morse set theory (***System MKM***) is just like NBG, except removing the "predicativity" restrictions on NBG0, NBG2 and NBG9.

C. Development of Mathematics

It is possible to define individual Frege-Russell numbers in NBG: e.g., one might define 1 as follows:

1 for $\{x|(\exists y) x = \{y\}\}$

Here "1" would name the class of all one-membered sets. The problem with such definitions is that this is a proper class, not a set. Hence, it is not itself a member of anything, and it would then become impossible to have a set containing all the natural numbers, etc. Consequently, mathematics is typically developed in NBG just as it is in ZF. E.g., we define x' as $x \cup \{x\}$, and define $\mathrm{On}(x)$ and $N(x)$ as before.

Unlike ZF, in NBG there is a *class* of all ordinal numbers, viz., $\{x|\mathrm{On}(x)\}$. However, it is a proper class.

Precisely the same sets can be proven to exist in ZF as compared to NBG. The notation for proper classes simplifies certain proofs, and allows an easier or more economical way to state certain axioms or results, as with the single axiom of replacement.

In fact, the relationship between ZF and NBG can be more precisely characterized as follows:

(1) While ZF is not strictly speaking a subtheory of NBG (since it is not strictly the case that every theorem of ZF is a theorem of NBG), for every theorem $\mathscr{A}$ of ZF there is a theorem $\mathscr{A}^*$ of NBG obtained by replacing all the quantifiers of the ZF formula with restricted quantifiers for sets.

   Proof sketch: For every proper axiom of ZF, $\mathscr{A}$, either $\mathscr{A}^*$ is an axiom of NBG or any easy consequence of the axioms of NBG. The two systems have the same inference rules.

   It follows that if ZF is inconsistent, then NBG is inconsistent as well. (And hence that if NBG is consistent, so is ZF.)

(2) Similarly if $\mathscr{A}$ is a predicative wff of the language of NBG, and $\vdash_{\text{NBG}} \mathscr{A}$, then the corresponding wff, $\mathscr{A}^\dagger$ of ZF obtained by replacing all the restricted quantifiers of $\mathscr{A}$ with regular quantifiers is such that $\vdash_{\text{ZF}} \mathscr{A}^\dagger$.

   It follows from the above that if NBG is inconsistent, then ZF is inconsistent as well. (Notice that if NBG is inconsistent, then every wff of NBG is a theorem, including both some predicative wff and its negation. The corresponding wff and its negation would then both be theorems of ZF.) It follows that if ZF is consistent, then so is NBG.

(3) *Result:* ZF is consistent iff NBG is consistent.

NBG is a proper subtheory of MKM, and MKM is strictly stronger, since the consistency of NBG can be proven in MKM.

XVIII. Set Theory with Urelements

As we have formulated them, systems Z, ZF, NBG and MKM do not allow entities other than classes (or sets) in their domains of quantification. According to the definition of identity

$t = u$ for $(\forall \chi)(\chi \in t \leftrightarrow \chi \in u)$

all entities that have no members or elements are identical to the empty set.

However it does not take much to alter the systems in order to make them consistent with having non-classes, typically called ***urelements*** (or sometimes, ***individuals***) within the domains of quantification.

A. The Systems ZU/ZFU

We begin by adding to our language a new monadic predicate letter 'K'. "$K(t)$" is to mean "t is a set or class". (In Z or ZF there is no distinction made between sets and classes.)

Definition:

$U(t)$ for $\neg K(t)$

(We could have taken 'U' as primitive, and defined 'K' in terms of it; it would not make a significant difference to the system.)

Mimicking Hatcher's introduction of the constant 'ω', we will introduce a new constant 'u' for the set of all urelements.

We also modify the syntax and use boldface nonitalicized letters **w**, **x**, **y** and **z**—with or without subscripts—for our variables. As with NBG, this does not prevent the language from being first-order. I use italicized boldface letters ***x***, etc., schematically for such variables.

Non-boldface letters (w, x, y and z) are introduced for restrictive quantification over classes/sets, as in NBG. In particular:

Definitions:

Where χ is a non-bold variable and $\boldsymbol{x}$ is the first bold-face variable not in $\mathscr{A}[\chi]$:

$$(\forall\chi)\,\mathscr{A}[\chi] \quad\text{for}\quad (\forall\boldsymbol{x})(K(\boldsymbol{x})\rightarrow\mathscr{A}[\boldsymbol{x}])$$
$$(\exists\chi)\,\mathscr{A}[\chi] \quad\text{for}\quad (\exists\boldsymbol{x})(K(\boldsymbol{x})\wedge\mathscr{A}[\boldsymbol{x}])$$
$$\{\chi|\chi\in t\wedge\mathscr{A}[\chi]\} \quad\text{for}\quad \{\boldsymbol{x}|\boldsymbol{x}\in t\wedge\mathscr{A}[\boldsymbol{x}]\wedge K(\boldsymbol{x})\}$$

For reasons that should be clear from the above, the normal definition of identity is not appropriate. One may choose either to take the identity relation as a primitive 2-place predicate, and build the system upon the *predicate calculus with identity* rather than the predicate calculus simpliciter. Alternatively, one may adopt the following rather counterintuitive definitions:

$$t=u \quad\text{for}\quad (K(t)\wedge K(u)\wedge(\forall\boldsymbol{x})(\boldsymbol{x}\in t\leftrightarrow\boldsymbol{x}\in u))\vee (U(t)\wedge U(u)\wedge(\forall\boldsymbol{x})(t\in\boldsymbol{x}\leftrightarrow u\in\boldsymbol{x}))$$
$$t\subseteq u \quad\text{for}\quad K(t)\wedge K(u)\wedge(\forall\boldsymbol{x})(\boldsymbol{x}\in t\rightarrow\boldsymbol{x}\in u)$$

The axioms of ZU are the following:

ZU0a. $(\forall\mathbf{x})(\forall\mathbf{y})(\mathbf{x}\in\mathbf{y}\rightarrow K(\mathbf{y}))$
ZU0b. $K(u)\wedge(\forall\mathbf{x})(U(\mathbf{x})\leftrightarrow\mathbf{x}\in u)$
ZU1. $(\forall\mathbf{x})(\forall\mathbf{y})(\mathbf{x}=\mathbf{y}\rightarrow(\mathscr{A}[\mathbf{x},\mathbf{x}]\rightarrow\mathscr{A}[\mathbf{x},\mathbf{y}]))$, where $\mathbf{y}$ does not occur bound in $\mathscr{A}[\mathbf{x},\mathbf{x}]$.
ZU2. $(\forall\chi)(\forall\boldsymbol{x})(\boldsymbol{x}\in\{\boldsymbol{y}|\boldsymbol{y}\in\chi\wedge\mathscr{A}[\boldsymbol{y}]\}\leftrightarrow\boldsymbol{x}\in\chi\wedge\mathscr{A}[\boldsymbol{x}])$, where $\boldsymbol{x}$ does not become bound in the context $\mathscr{A}[\boldsymbol{x}]$.
ZU3. $K(0)\wedge 0=\{\mathbf{x}|\mathbf{x}\in 0\wedge\mathbf{x}\neq\mathbf{x}\}$
ZU4. $(\forall\mathbf{x})(\forall y)(\mathbf{x}\in\wp(y)\leftrightarrow\mathbf{x}\subseteq y)$
ZU5. $(\forall\mathbf{x})(\forall\mathbf{y})(\forall\mathbf{z})(\mathbf{x}\in\{\mathbf{y},\mathbf{z}\}\leftrightarrow\mathbf{x}=\mathbf{y}\vee\mathbf{x}=\mathbf{z})$
ZU6. $(\forall\mathbf{x})(\forall y)(\mathbf{x}\in\bigcup(y)\leftrightarrow(\exists\mathbf{z})(\mathbf{z}\in y\wedge\mathbf{x}\in\mathbf{z}))$
ZU7. $(\forall x)(x\neq 0\rightarrow(\exists\mathbf{y})(\mathbf{y}\in x\wedge\neg(\exists\mathbf{z})(\mathbf{z}\in\mathbf{y}\wedge\mathbf{z}\in x)))$
ZU8. $(\forall\mathbf{x})(\mathbf{x}\in\omega\leftrightarrow N(\mathbf{x}))$[3]

ZFU and ZFCU would also add these:

ZFU9. $(\forall x)((\forall\mathbf{y})(\forall\mathbf{z}_1)(\forall\mathbf{z}_2)(\mathscr{A}[\mathbf{y},\mathbf{z}_1]\wedge\mathscr{A}[\mathbf{y},\mathbf{z}_2]\rightarrow\mathbf{z}_1=\mathbf{z}_2)\rightarrow (\exists y)(\forall\mathbf{z})(\mathbf{z}\in y\leftrightarrow(\exists\mathbf{y}_1)(\mathbf{y}_1\in x\wedge\mathscr{A}[\mathbf{y}_1,\mathbf{z}])))$, where $\mathbf{z}_2$ does not become bound in $\mathscr{A}[y,\mathbf{z}_2]$.
ZFCU10. $(\forall x)(x\neq 0\rightarrow\sigma(x)\in x)$

Metatheoretic results:
Z is consistent iff ZU is consistent.
ZF is consistent iff ZFU is consistent.
ZFC is consistent iff ZFCU is consistent.

To obtain a system NBGU (for NBG with urelements), it would be necessary to adopt the following definitions (for *set* and *element*, respectively):

$M(t)$ for $K(t)\wedge(\exists\boldsymbol{x})\,t\in\boldsymbol{x}$
$E(t)$ for $M(t)\vee U(t)$

Uppercase letters could then be used as restricted quantifiers for all classes, and lowercase for sets only, so that $(\forall X)\,\mathscr{A}[X]$ would mean $(\forall\mathbf{x})(K(\mathbf{x})\rightarrow\mathscr{A}[\mathbf{x}])$ and $(\forall x)\,\mathscr{A}[x]$ would mean $(\forall\mathbf{x})(M(\mathbf{x})\rightarrow\mathscr{A}[\mathbf{x}])$, etc.

[3] The definition of $\mathrm{On}(t)$ would also require $K(t)$, and so $N(t)$ would also require $K(t)$.

The axioms of NBG would then need to be modified like those of ZF were modified for ZFU.

HOMEWORK 17

Prove: $\vdash_{ZU} (\forall \mathbf{x})(\forall \mathbf{y})K(\{\mathbf{x},\mathbf{y}\})$

XIX. Relative Consistency

We have several times cited results to the effect that one system is consistent *if* another is. This is established by a relative consistency proof.

Relative consistency proofs can take two forms. One is to show that a model can be constructed for a given theory using the sets that can be proven to exist in another theory. Since a theory is consistent if it has a model, this shows that the theory in question is consistent if the theory in which the model is constructed is. This method is slightly more controversial, however, since the proof that every system that has a model is consistent makes use of a fair amount of set theory that may itself be contested.

Another more direct way towards establishing relative consistency is to give an argument that any proof of an inconsistency in one system could be mirrored to produce an inconsistency in another. We give an example of such a proof below.

The Consistency of ST relative to System ZF

Definition: Let $\wp^n(t)$ represent n occurrences of $\wp$ applied to t, so that, e.g., $\wp^3(t)$ is $\wp(\wp(\wp(t)))$. (Limiting case, $\wp^0(t)$ is just t itself.)

Lemma 1: For every n, we have $\vdash_{ZF} \wp^n(t) \in \wp^{n+1}(t)$.
Proof: from repeated applications of T6 and Z4.

We now sketch the following ***translation scheme*** between the languages of ST and ZF:

(i) Begin with a formula $\mathscr{A}$ of ST.

(ii) Replace each occurrence of $(\forall \chi^n)\,\mathscr{B}[\chi^n]$ with $(\forall y)(y \in \wp^n(\omega) \rightarrow \mathscr{B}[y])$, where y is the first variable of ZF's language not yet used in the translation.

(iii) Replace each occurrence of $\{\chi^n|\mathscr{B}[\chi^n]\}$ with $\{y|y \in \wp^n(\omega) \wedge \mathscr{B}[y]\}$, where y is, again, the first variable of ZF's language not yet used in the translation.

(iv) for each free variable χ^n of $\mathscr{A}$, replace all free occurrences of χ^n with y (the first variable of ZF's language not yet used) and prefix the result with $y \in \wp^n(\omega) \rightarrow \ldots$.

If $\mathscr{A}$ is a wff of ST, let $(\mathscr{A})^Z$ be the wff of ZF that results from the steps above. Let $\mathscr{A}^\#$ be the result of only applying steps (i)–(iii) and replacing each typed variable with the type free variable that would replace it in step (iv).

Example: If $\mathscr{A}$ is $(\forall x^1)\, x^1 \in y^2$, then $\mathscr{A}^\#$ is $(\forall x)(x \in \wp(\omega) \rightarrow x \in y)$ and $(\mathscr{A})^Z$ is $y \in \wp(\wp(\omega)) \rightarrow (\forall x)(x \in \wp(\omega) \rightarrow x \in y)$

Lemma 2: If $\mathscr{A}$ is an axiom of ST, then $\vdash_{ZF} (\mathscr{A})^Z$.

Proof:

Suppose $\mathscr{A}$ is an axiom of ST. There are the following cases to consider.

1). $\mathscr{A}$ is a truth-table tautology; then $(\mathscr{A})^Z$ is also a truth-table tautology, whence $\vdash_{ZF} (\mathscr{A})^Z$.

2). $\mathscr{A}$ takes the form $(\forall \chi^n)\,\mathscr{B}[\chi^n] \rightarrow \mathscr{B}[t^n]$ where t^n is free for χ^n in $\mathscr{B}[\chi^n]$. We have the following as an axiom of ZF: $(\forall y)(y \in \wp^n(\omega) \rightarrow \mathscr{B}^\#[y]) \rightarrow ((t^n)^\# \in \wp^n(\omega) \rightarrow \mathscr{B}^\#[(t^n)^\#])$. By SL, (*) $(t^n)^\# \in \wp^n(\omega) \rightarrow ((\forall y)(y \in \wp^n(\omega) \rightarrow \mathscr{B}^\#[y]) \rightarrow \mathscr{B}^\#[(t^n)^\#])$. Because t^n is a term of ST, it is either a variable or a term of the form $\{z^{n-1}|\mathscr{C}[z^{n-1}]\}$ where $n \geq 1$. If t^n is a variable then it is free in $\mathscr{A}$, and $(\mathscr{A})^Z$ follows from (*) by at most ST (TC) alone. If t^n is a term of the form $\{z^{n-1}|\mathscr{C}[z^{n-1}]\}$, then $(t^n)^\#$ takes the form $\{z|z \in \wp^{n-1}(\omega) \wedge \mathscr{C}^\#[z]\}$. It then follows by Z2 that $\vdash_{ZF} (t^n)^\# \subseteq \wp^{n-1}(\omega)$, and hence, by Z4 that $\vdash_{ZF} (t^n)^\# \in \wp^n(\omega)$. By this and (*) we get $\vdash_{ZF} (\forall y)(y \in \wp^n(\omega) \rightarrow \mathscr{B}^\#[y]) \rightarrow \mathscr{B}^\#[(t^n)^\#]$. From this, $\vdash_{ZF} (\mathscr{A})^Z$ follows by at most SL (TC) alone.

3). $\mathscr{A}$ takes the form $(\forall \chi^n)(\mathscr{C} \rightarrow \mathscr{B}[\chi^n]) \rightarrow (\mathscr{C} \rightarrow (\forall \chi^n)\,\mathscr{B}[\chi^n])$, where $\mathscr{C}$ does not contain χ^n free. By a simple proof, we have (*) $\vdash_{ZF} (\forall y)(y \in \wp^n(\omega) \to (\mathscr{C}^\# \to \mathscr{B}^\#[y])) \to (\forall y)(\mathscr{C}^\# \to (y \in \wp^n(\omega) \to \mathscr{B}^\#[y]))$. We also have an axiom of Z: $\vdash_{ZF} (\forall y)(\mathscr{C}^\# \to (y \in \wp^n(\omega) \to \mathscr{B}^\#[y])) \to (\mathscr{C}^\# \to (\forall y)(y \in \wp^n(\omega) \to \mathscr{B}^\#[y]))$. By HS, $\vdash_{ZF} (\forall y)(y \in \wp^n(\omega) \to (\mathscr{C}^\# \to \mathscr{B}^\#[y])) \to (\mathscr{C}^\# \to (\forall y)(y \in \wp^n(\omega) \to \mathscr{B}^\#[y]))$. By SL, we get $\vdash_{ZF} (\mathscr{A})^Z$.

4). $\mathscr{A}$ is an instance of ST1, $(\forall \chi^n)(\chi^n \in \{z^n | \mathscr{B}[z^n]\} \leftrightarrow \mathscr{B}[\chi^n])$. We have $\vdash_{ZF} (\forall y)(y \in \wp^n(\omega) \to (y \in \{z | z \in \wp^n(\omega) \land \mathscr{B}^\#[z]\} \leftrightarrow \mathscr{B}^\#[y]))$ as an easy consequence of Z2. By SL, $\vdash_{ZF} (\mathscr{A})^Z$.

5). $\mathscr{A}$ is an instance of ST2, i.e., $(\forall x^{n+1})(\forall y^{n+1})((\forall z^n)(z^n \in x^{n+1} \leftrightarrow z^n \in y^{n+1}) \to x^{n+1} = y^{n+1})$. Given ST's definition of identity $(\mathscr{A})^Z$ is $(\forall x)(x \in \wp^{n+1}(\omega) \to (\forall y)(y \in \wp^{n+1}(\omega) \to ((\forall z)(z \in \wp^n(\omega) \to (z \in x \leftrightarrow z \in y)) \to (\forall z)(z \in \wp^{n+2}(\omega) \to (x \in z \leftrightarrow y \in z)))))$. This can be proven in ZF as follows:

(1)	1. $x \in \wp^{n+1}(\omega)$	Hyp
(1)	2. $x \subseteq \wp^n(\omega)$	1, Z4 QL
(3)	3. $y \in \wp^{n+1}(\omega)$	Hyp
(3)	4. $y \subseteq \wp^n(\omega)$	3, Z4 QL
(5)	5. $(\forall z)(z \in \wp^n(\omega) \to (z \in x \leftrightarrow z \in y))$	Hyp
(6)	6. $z \in x$	Hyp
(1,6)	7. $z \in \wp^n(\omega)$	2, 6, Df. $\subseteq$ QL
(1,5,6)	8. $z \in y$	5, 6, 7 QL
(1,5)	9. $z \in x \to z \in y$	6–8 CP
(3,5)	10. $z \in y \to z \in x$	CP like 6–8
(1,3,5)	11. $(\forall z)(z \in x \leftrightarrow z \in y)$	9, 10 BI, UG
(1,3,5)	12. $x = y$	11 Df. =
	13. $z \in \wp^{n+2}(\omega) \to (x \in z \leftrightarrow x \in z)$	Taut
(1,3,5)	14. $z \in \wp^{n+2}(\omega) \to (x \in z \leftrightarrow y \in z)$	12, 13 LL
(1,3,5)	15. $(\forall z)[14]$	14 UG
(1,3)	16. $[5] \to [15]$	5–15 CP
(1)	17. $[3] \to [16]$	3–16 CP
(1)	18. $(\forall y)[17]$	17 UG
	19. $[1] \to [18]$	1–18 CP
	20. $(\forall x)[19]$	19 UG

6). $\mathscr{A}$ is ST3, i.e., $(\exists x^3)((\forall x^0)\langle x^0, x^0\rangle \notin x^3 \land (\forall x^0)(\exists y^0)\langle x^0, y^0\rangle \in x^3 \land (\forall x^0)(\forall y^0)(\forall z^0)(\langle x^0, y^0\rangle \in x^3 \land \langle y^0, z^0\rangle \in x^3 \to \langle x^0, z^0\rangle \in x^3))$. Then $(\mathscr{A})^Z$ is $(\exists x)(x \in \wp(\wp(\wp(\omega))) \land (\forall y)(y \in \omega \to \langle y, y\rangle \notin x) \land (\forall z)(z \in \omega \to (\exists x_1)(x_1 \in \omega \land \langle z, x_1\rangle \in x)) \land (\forall y_1)(y_1 \in \omega \to (\forall z_1)(z_1 \in \omega \to (\forall x_2)(x_2 \in \omega \to (\langle y_1, z_1\rangle \in x \land \langle z_1, x_2\rangle \in x \to \langle y_1, x_2\rangle \in x)))))$. Consider the relation E, defined as follows: $\{x | x \in \wp(\wp(\omega)) \land (\exists y)(\exists z)(x = \langle y, z\rangle \land y \in z)\}$, i.e., the less-than relation among members of ω. We have all of the following:

$\vdash_{ZF} E \in \wp(\wp(\wp(\omega)))$
$\vdash_{ZF} (\forall y)(y \in \omega \to \langle y, y\rangle \notin E)$
$\vdash_{ZF} (\forall z)(z \in \omega \to (\exists x_1)(x_1 \in \omega \land \langle z, x_1\rangle \in E))$
$\vdash_{ZF} (\forall y_1)(y_1 \in \omega \to (\forall z_1)(z_1 \in \omega \to (\forall x_2)(x_2 \in \omega \to$
$(\langle y_1, z_1\rangle \in E \land \langle z_1, x_2\rangle \in E \to \langle y_1, x_2\rangle \in E))))$

Hence, we get $\vdash_{ZF} (\mathscr{A})^Z$ by Conj and EG.

Lemma 3: For any wff $\mathscr{A}$, if $\vdash_{ST} \mathscr{A}$, then $\vdash_{ZF} (\mathscr{A})^Z$.

Proof:

Assume $\vdash_{ST} \mathscr{A}$. By definition, there is an ordered sequence of wffs $\mathscr{B}_1, \ldots, \mathscr{B}_n$, where $\mathscr{B}_n$ is $\mathscr{A}$ and for each $\mathscr{B}_i$, where $1 \le i \le n$, either (a) $\mathscr{B}_i$ is an axiom, (b) $\mathscr{B}_i$ follows from previous members by MP, or (c) $\mathscr{B}_i$ follows from previous members by UG. We shall prove by strong/complete induction on its line number that for each such $\mathscr{B}_i$, we have $\vdash_{ZF} (\mathscr{B}_i)^Z$. (We can assume that we already have $\vdash_{ZF} (\mathscr{B}_j)^Z$ for all $1 \le j < i$.) There are three cases to consider.

(a) $\mathscr{B}_i$ is an axiom of ST. By Lemma 2, we have $\vdash_{ZF} (\mathscr{B}_i)^Z$.

(b) $\mathscr{B}_i$ follows by previous members of the sequence by MP. Hence, there are previous members of the sequence $\mathscr{B}_j$ and $\mathscr{B}_k$ where $\mathscr{B}_k$ takes the form $(\mathscr{B}_j \to \mathscr{B}_i)$. By the inductive hypothesis, we have $\vdash_{ZF} (\mathscr{B}_j)^Z$ and $\vdash_{ZF} (\mathscr{B}_k)^Z$, i.e., $\vdash_{ZF} y_1 \in \wp^{m_1}(\omega) \to (y_2 \in \wp^{m_2}(\omega) \to \ldots (\mathscr{B}_j)^\#)$, where y_1, y_2, etc., are the free variables of $(\mathscr{B}_j)^\#$, and $\vdash_{ZF} y_1 \in \wp^{m_1}(\omega) \to (y_2 \in \wp^{m_2}(\omega) \to \ldots (z_1 \in \wp^{q_1}(\omega) \to (z_2 \in \wp^{q_2}(\omega) \to \ldots ((\mathscr{B}_j)^\# \to (\mathscr{B}_i)^\#))))$, where z_1, z_2 are the free variables of $(\mathscr{B}_i)^\#$ not in $(\mathscr{B}_j)^\#$. By SL, $\vdash_{ZF} y_1 \in \wp^{m_1}(\omega) \to (y_2 \in \wp^{m_2}(\omega) \to \ldots (z_1 \in \wp^{q_1}(\omega) \to (z_2 \in \wp^{q_2}(\omega) \to \ldots (\mathscr{B}_i)^\#)))$. Also, by SL, one can pull to the front those variables in the list y_1, y_2 that do not occur free in $(\mathscr{B}_i)^\#$. By UG, and UI, one can instantiate each such variable to obtain an antecedent either of the form $\wp^{m_i-1}(\omega) \in \wp^{m_i}(\omega)$ or of the form $0 \in \omega$, and discharge them by Lemma 1 (or T19 and ZF8) and MP, resulting in $\vdash_{ZF} (\mathscr{B}_i)^Z$.

(c) $\mathscr{B}_i$ follows by previous members of the sequence by UG. hence, there is some previous member of the sequence $\mathscr{B}_j$ and $\mathscr{B}_i$ takes the form $(\forall x^n)\, \mathscr{B}_j$. By the inductive hypothesis, we have $\vdash_{ZF} y_1 \in \wp^{m_1}(\omega) \rightarrow (y_2 \in \wp^{m_2}(\omega) \rightarrow \ldots (\mathscr{B}_j)^\#)$, where y_1, y_2, etc., are the free variables of $(\mathscr{B}_j)^\#$. Let z be the variable we use to replace x^n. Most likely, it occurs in the list y_1, y_2, but even if not, we can push it in (or add it then push it in) to the end, to get $\vdash_{ZF} y_1 \in \wp^{m_1}(\omega) \rightarrow (y_2 \in \wp^{m_2}(\omega) \rightarrow \ldots (z \in \wp^n(\omega) \rightarrow (\mathscr{B}_j)^\#))$. By QL, $\vdash_{ZF} y_1 \in \wp^{m_1}(\omega) \rightarrow (y_2 \in \wp^{m_2}(\omega) \rightarrow \ldots (\forall z)(z \in \wp^n(\omega) \rightarrow (\mathscr{B}_j)^\#))$, which is $\vdash_{ZF} (\mathscr{B}_i)^Z$.

Hence, for all such $\mathscr{B}_i$, $\vdash_{ZF} (\mathscr{B}_i)^Z$, and since $\mathscr{A}$ is $\mathscr{B}_n$, we have $\vdash_{ZF} (\mathscr{A})^Z$. Hence, if $\vdash_{ST} \mathscr{A}$ then $\vdash_{ZF} (\mathscr{A})^Z$.

Finally, we are ready to establish our main result: *if ZF is consistent, then ST is consistent.*

We shall prove the transposition: *If ST is inconsistent, then ZF is inconsistent.* Suppose that ST is inconsistent. From any contradiction in ST, everything follows. Hence there are closed wffs $\mathscr{A}$ and $\neg\mathscr{A}$ such that $\vdash_{ST} \mathscr{A}$ and $\vdash_{ST} \neg\mathscr{A}$. By Lemma 3, it follows that $\vdash_{ZF} (\mathscr{A})^Z$ and $\vdash_{ZF} (\neg\mathscr{A})^Z$. Since $\mathscr{A}$ is closed, $(\neg\mathscr{A})^Z$ is $\neg(\mathscr{A})^Z$, and so $\vdash_{ZF} \neg(\mathscr{A})^Z$, and ZF is inconsistent as well.

More or less the same line of proof can be used to show:

- that ST is relative to NBG (or even weaker system Z, since nothing in the above uses the axiom of replacement); and
- that HOPC (the higher-order predicate calculus) is consistent relative to ZF.

Other relative consistency results (e.g., that ZF is consistent relative to NBG) can be carried out using similar "translation" schemes.

XX. Quine's System NF

We now turn our attention to less common approaches to the foundations of mathematics, with very different philosophical motivations. Two of the most influential were developed by W. V. Quine.

A. Background

The system of Quine's "New Foundations" was meant as a compromise between the limitations of type-theory and the naiveté of unrestricted set abstraction as in system F.

In system ST, a formula is only regarded as well-formed when the type of any term preceding the sign "∈" is always one less than the type of the term following the sign. Hence the following is well-formed:

$$x^1 \in y^2 \wedge y^2 \in z^3$$

Whereas the following is not:

$$x^1 \in y^2 \wedge y^2 \in x^1$$

In *Principia Mathematica*, Whitehead and Russell often employed a device known as ***typical ambiguity*** when stating general results and giving proofs. The upshot was that they would leave off type indices, leaving it to the reader to determine the types of the terms as being any types consistent with the rules governing what formulas are well-formed. Thus instead of writing the single formula:

$$(\forall x^0)(\exists y^1)(\forall z^0)(z^0 \in y^1 \leftrightarrow z^0 = x^0)$$

Or the explicitly schematic:

$$(\forall x^n)(\exists y^{n+1})(\forall z^n)(z^n \in y^{n+1} \leftrightarrow z^n = x^n)$$

They would write simply:

$$(\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow z = x)$$

It would then be understood that the reader could understand the variables to be of any type consistent with it being a wff, so that the reader could take the above as any of:

$$(\forall x^0)(\exists y^1)(\forall z^0)(z^0 \in y^1 \leftrightarrow z^0 = x^0)$$
$$(\forall x^1)(\exists y^2)(\forall z^1)(z^1 \in y^2 \leftrightarrow z^1 = x^1)$$
$$(\forall x^2)(\exists y^3)(\forall z^2)(z^2 \in y^3 \leftrightarrow z^2 = x^2), \text{ etc.}$$

This leads to the question as to whether it might be possible to have a system like type-theory but in which variables simply *do not have types*, but in which a formula would only be well-formed if it were *possible* to place superscripts on all the terms such that a number placed for a term preceding "∈" were always one less than the number placed on the term following it.

That is, consider a first-order language with only one type of variable, with a single two-place predicate, ∈, and no further predicates, constants or function-letters or vbtos. Then consider the following definition:

Definition: A formula $\mathscr{A}$ is ***stratified*** iff there is a function F from the variables of $\mathscr{A}$ to natural numbers such that for any portion of $\mathscr{A}$ of the form $\chi \in y$, it holds that $F(\chi) + 1 = F(y)$.

In other words, a formula is stratified iff it would be possible to regard it as the "typically ambiguous" representation of a possible wff of ST. Thus, e.g., "$(\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow z = x)$" is stratified in virtue of the function assigning 0 to both x and z, and 1 to y. However $y \in y$ or $x \notin x$ are unstratified.

Notice that the distinction between stratified and unstratified formulas applies to formulas of an *untyped* language. (Roughly, a wff $\mathscr{A}$ of an untyped language is stratified iff it could be *translated into* the typed language of ST.)

One suggestion (not taken up by Quine) is to restrict well-formed formulas to *stratified formulas* and attempt to carry out mathematics in a type-free language, but in which it is be impossible to even to speak of the existence of, e.g., a Russell class, since, e.g., $(\exists y)(\forall x)(x \in y \leftrightarrow x \notin x)$ is unstratified.

Quine takes things a step further. One question that arises is whether type-restrictions should really be restrictions on *meaning* or what is well-formed. Even if one regards sets as divided into different types, so that a set is never a member of itself, but only ever a member of sets the next type-up, does this mean that "$x \in x$" should be *meaningless* or syntactically *ill-formed*? (Indeed, one cannot even state that no set can be a member of itself, since, $(\forall x)\, x \notin x$ is not well-formed!)

Quine proposes a system not in which stratification is adopted as a criterion for well-formedness, but in which stratification is used as a criterion for determining what sets or classes exist. That is, he takes an underlying idea behind the theory of types and modifies it from a theory restricting the meaningfulness of certain formulas to a theory of *ontological economy*, or a theory restricting what sets are postulated to exist.

(In particular, $\{x|\mathscr{A}[\chi]\}$ is postulated to exist only when $\mathscr{A}[x]$ or an equivalent wff is *stratified*.)

The result is a type-free first-order system whose variables are unrestricted in scope: ranging over everything. Moreover, unstratified formulas are allowed in the language: "$x \in x$", "$y \notin y$" are nevertheless wffs, even though Quine does not postulate the existence of sets of all and only those things satisfying such formulas.

B. Syntax of NF

As Quine himself formulated NF, it had only a single binary connective ("∈"), and no additional predicates, constants, function letters or vbtos.

Quine adopted Russell's theory of descriptions, wherepon a quasi-term $\imath x\, \mathscr{A}[\chi]$, read "the χ such that $\mathscr{A}[\chi]$" would be used as part of a contextual definition:

$\mathscr{B}[\imath x\, \mathscr{A}[\chi]]$ is an abbreviation for:

$$(\exists \chi)((\forall y)(\mathscr{A}[y] \leftrightarrow y = \chi) \land \mathscr{B}[\chi])$$

I.e., "one and only one thing satisfies $\mathscr{A}[\chi]$ and it also satisfies $\mathscr{B}$".

Those of you familiar with Russell's theory of descriptions know that such "contextual definitions" give rise to scope ambiguities. For example, does $\neg F(\imath x\, Gx)$ mean $(\exists x)((\forall y)(Gy \leftrightarrow y = x) \wedge \neg Fx)$ or $\neg(\exists x)((\forall y)(Gy \leftrightarrow y = x) \wedge Fx)$?

For Quine, since the only place where terms may occur is flanking the sign $\in$, he gives explicit definitions for atomic wffs containing description "quasi-terms" as follows:

$\imath\chi\, \mathscr{A}[\chi] \in z$	for	$(\exists\chi)((\forall y)(\mathscr{A}[y] \leftrightarrow y = \chi) \wedge \chi \in z)$
$z \in \imath\chi\, \mathscr{A}[\chi]$	for	$(\exists\chi)((\forall y)(\mathscr{A}[y] \leftrightarrow y = \chi) \wedge z \in \chi)$
$\imath\chi\, \mathscr{A}[\chi] \in \imath z\, \mathscr{B}[z]$	for	$(\exists\chi)((\forall y)(\mathscr{A}[y] \leftrightarrow y = \chi) \wedge$ $(\exists z)((\forall y)(\mathscr{B}[y] \leftrightarrow y = z) \wedge \chi \in z))$

Such definitions are applied to atomic parts of a wff, so descriptions always have narrow scope. Quine then introduces class abstracts by means of descriptions:

$\{\chi|\mathscr{A}[\chi]\}$ abbreviates $\imath y(\forall\chi)(\chi \in y \leftrightarrow \mathscr{A}[\chi])$

In order to sidestep such complications regarding descriptions, Hatcher formulates NF using the vbto $\{\chi|\mathscr{A}[\chi]\}$ as primitive. He must then give a definition of *stratification* according to which the function must assign a number to every term of a wff, and the number assigned to a term of the form $\{\chi|\mathscr{A}[\chi]\}$ must be one more than the number assigned to the variable χ. Then, in any portion of the form $t \in u$ in any stratified wff $\mathscr{A}$, the number assigned to u must be one more than that assigned to t.

So long as one deals only with terms of the form $\{\chi|\mathscr{A}[\chi]\}$ when $\mathscr{A}[\chi]$ is stratified, the differences between Hatcher's practice and Quine's are trivial.

Additional definitions can be added showing a remarkable similarity to those for System F:

Definitions:

$t \notin u$	for	$\neg t \in u$
$t = u$	for	$(\forall\chi)(\chi \in t \leftrightarrow \chi \in u)$
$t \neq u$	for	$\neg t = u$
$t \subseteq u$	for	$(\forall\chi)(\chi \in t \rightarrow \chi \in u)$
$\{t\}$	for	$\{\chi|\chi = t\}$
$\{t, u\}$	for	$\{\chi|\chi = t \vee \chi = u\}$
$\langle t, u\rangle$	for	$\{\{t\}, \{t, u\}\}$
$\overline{t}$	for	$\{\chi|\chi \notin t\}$
$(t \cap u)$	for	$\{\chi|\chi \in t \wedge \chi \in u\}$
$(t \cup u)$	for	$\{\chi|\chi \in t \vee \chi \in u\}$
$(t - u)$	for	$\{\chi|\chi \in t \wedge \chi \notin u\}$
$\wp(t)$	for	$\{\chi|\chi \subseteq t\}$
V	for	$\{x|x = x\}$
Λ	for	$\{x|x \neq x\}$

We say that a term $\{\chi|\mathscr{A}[\chi]\}$ is stratified when $\mathscr{A}[\chi]$ is stratified. All of the abstracts used above are stratified if t and u are, though notice that when embedded in a larger wff, the function must be able to assign the same value to t and u in any wff in which $t = u$, $t \neq u$, $t \subseteq u$, $\{t, u\}$, $\langle t, u\rangle$, $t \cap u$, $t \cup u$, or $t - u$ appear.

The definition of identity here rules out urelements, or forces us (as Quine deems "harmless") to identify an urelement with its own singleton. (In fact, it may not be so harmless!)

C. Axiomatization

Mostly due their difference as to whether or not to take the vbto $\{\chi|\mathscr{A}[\chi]\}$ as primitive, Quine and Hatcher give slightly different axiomatizations. The differences are unimportant, as we shall now prove.

Hatcher's Formulation

NF has the following proper axiom schemata:

NF1. $(\forall x)(\forall y)(x = y \rightarrow (\mathscr{A}[x, x] \rightarrow \mathscr{A}[x, y]))$, where y does not occur bound in $\mathscr{A}[x, x]$.

NF2. $(\forall\chi)(\chi \in \{y|\mathscr{A}[y]\} \leftrightarrow \mathscr{A}[\chi])$, provided that $\mathscr{A}[y]$ is stratified, and χ does not become bound when placed in the context $\mathscr{A}[\chi]$.

NF2 appears superficially the same as F2, except for the limitation to stratified formulas. In fact, the limit of F2 to stratifed formulas is *the only difference between NF and System F*.

Quine's Formulation

Instead of NF2 as above, Quine adopts the existential posit:

NF2′. $(\exists y)(\forall x)(x \in y \leftrightarrow \mathscr{A}[x])$, where $\mathscr{A}[x]$ is stratified and does not contain y free.

However, given his understanding of $\{\chi|\mathscr{A}[\chi]\}$ as a definite description, from NF2′ one can derive Hatcher's NF2 as follows:

Assume throughout that $\mathscr{A}[y]$ is stratified.

(1)	1. $\chi \in \{y \mid \mathscr{A}[y]\}$	Hyp
(1)	2. $\chi \in \iota z\,(\forall y)(y \in z \leftrightarrow \mathscr{A}[y])$	1 Df. $\{y \mid \mathscr{A}[y]\}$
(1)	3. $(\exists z)((\forall w)((\forall y)(y \in w \leftrightarrow \mathscr{A}[y]) \leftrightarrow w = z) \wedge \chi \in z)$	2 Df. ιz
(1)	4. $(\forall w)((\forall y)(y \in w \leftrightarrow \mathscr{A}[y]) \leftrightarrow w = b) \wedge \chi \in b$	3 EI
(1)	5. $(\forall y)(y \in b \leftrightarrow \mathscr{A}[y]) \leftrightarrow b = b$	4 Simp, UI
	6. $b = b$	Ref=
(1)	7. $(\forall y)(y \in b \leftrightarrow \mathscr{A}[y])$	5, 6 BMP
(1)	8. $\chi \in b \leftrightarrow \mathscr{A}[\chi]$	7 UI
(1)	9. $\mathscr{A}[\chi]$	4, 8 SL
	10. $\chi \in \{y \mid \mathscr{A}[y]\} \rightarrow \mathscr{A}[\chi]$	1–9 CP
(11)	11. $\mathscr{A}[\chi]$	Hyp
	12. $(\exists y)(\forall x)(x \in y \leftrightarrow \mathscr{A}[x])$	NF2′
	13. $(\forall x)(x \in c \leftrightarrow \mathscr{A}[x])$	12 EI
(11)	14. $\chi \in c$	11, 13 UI, BMP
(15)	15. $(\forall y)(y \in w \leftrightarrow \mathscr{A}[y])$	Hyp
	16. $y \in c \leftrightarrow \mathscr{A}[y]$	13 UI
(15)	17. $y \in w \leftrightarrow \mathscr{A}[y]$	15 UI
(15)	18. $y \in w \leftrightarrow y \in c$	16, 17 SL
(15)	19. $w = c$	18, UG, Df. =
	20. $(\forall y)(y \in w \leftrightarrow \mathscr{A}[y]) \rightarrow w = c$	15–19 CP
(21)	21. $w = c$	Hyp
(21)	22. $(\forall x)(x \in w \leftrightarrow \mathscr{A}[x])$	13, 21 LL
(21)	23. $(\forall y)(y \in w \leftrightarrow \mathscr{A}[y])$	22 UI, UG
	24. $w = c \rightarrow (\forall y)(y \in w \leftrightarrow \mathscr{A}[y])$	21–23 CP
	25. $(\forall y)(y \in w \leftrightarrow \mathscr{A}[y]) \leftrightarrow w = c$	20, 24 BI
	26. $(\forall w)((\forall y)(y \in w \leftrightarrow \mathscr{A}[y]) \leftrightarrow w = c)$	25 UG
(11)	27. $(\forall w)((\forall y)(y \in w \leftrightarrow \mathscr{A}[y]) \leftrightarrow w = c) \wedge \chi \in c$	14, 26 Conj
(11)	28. $(\exists z)((\forall w)((\forall y)(y \in w \leftrightarrow \mathscr{A}[y]) \leftrightarrow w = z) \wedge \chi \in z)$	27 EG
(11)	29. $\chi \in \{y \mid \mathscr{A}[y]\}$	28 Dfs. $\{y \mid \mathscr{A}[y]\}$, ιz
	30. $\mathscr{A}[\chi] \rightarrow \chi \in \{y \mid \mathscr{A}[y]\}$	11–29 CP
	31. $(\forall \chi)(\chi \in \{y \mid \mathscr{A}[y]\} \leftrightarrow \mathscr{A}[\chi])$	10, 30 BI, UG

We also have the following derived rules on Quine's formulation:

UI*: $(\forall \chi)\, \mathscr{B}[\chi] \vdash_{\text{NF}} \mathscr{B}[\{y|\mathscr{A}[y]\}]$, provided that $\mathscr{A}[y]$ is stratified and $\{y|\mathscr{A}[y]\}$ contains no free variables that become bound in the context $\mathscr{B}[\{y|\mathscr{A}[y]\}]$.

EG*: $\mathscr{B}[\{y|\mathscr{A}[y]\}] \vdash_{\text{NF}} (\exists \chi)\, \mathscr{B}[\chi]$, provided that $\mathscr{A}[y]$ is stratified and $\{y|\mathscr{A}[y]\}$ contains no free variables that become bound in the context $\mathscr{B}[\{y|\mathscr{A}[y]\}]$.

It is therefore harmless to treat $\{y|\mathscr{A}[y]\}$ as though it were a standard term, *provided that* it is stratified. Thus in what follows we shall not worry about the differences between Quine's and Hatcher's formulations.

Notice that if $\mathscr{A}[y]$ is not stratified, and is not equivalent to any stratified wff, then, there is no proof that the term $\{y|\mathscr{A}[y]\}$ is a "proper description". If there no such unique class matching this description, any atomic wff or atomic part of a wff containing it is always false.

For instance, $x \notin x$ is not stratified, and so any atomic wff containing $\{x|x \notin x\}$ is false, and indeed, in Quine's formulation demonstrably so.
$\vdash_{\text{NF}} \{x|x \notin x\} \notin \{x|x \notin x\}$

From the above no contradiction follows, as we do not have an instance of NF2 of the form:

$$(\forall y)(y \in \{x|x \notin x\} \leftrightarrow y \notin y)$$

D. Development of Mathematics

Many theorems of NF are proven exactly as they are for system F. Hence we shall not usually bother to give the details.

T0. $\vdash_{\text{NF}} (\forall x)\, x = x$
T1. $\vdash_{\text{NF}} (\forall x)\, x \in \text{V}$
T1a. $\vdash_{\text{NF}} \text{V} \in \text{V}$

Notice that T1a marks a point of departure of NF from both ZF and ST. Some sets are members of themselves. (Given Quine's intention to identify individuals as their singletons, they too would be members of themselves.)

T2. $\vdash_{\text{NF}} (\forall x)\, x \subseteq \text{V}$
T3. $\vdash_{\text{NF}} \wp(\text{V}) = \text{V}$
T4. $\vdash_{\text{NF}} (\forall x)\, x \notin \Lambda$

Some mathematical definitions are below. Again, they are similar to those for System F. Notice that the barrier present to using the Frege-Russell definition of numbers as classes of like-membered classes does not exist for NF. (Notice, moreover, that we *cannot* use the von Neumann numbers, since $x \cup \{x\}$ is unstratified.)

We can even define $Nc(t)$ for "the cardinal number of t," i.e., the number to which t belongs.

Definitions:

0	for	$\{\Lambda\}$
t'	for	$\{\chi \mid (\exists y)(y \in \chi \wedge \chi - \{y\} \in t)\}$
N	for	$\{x \mid (\forall y)(0 \in y \wedge (\forall z)(z \in y \rightarrow z' \in y) \rightarrow x \in y)\}$
Fin	for	$\{x \mid (\exists y)(y \in N \wedge x \in y)\}$
Inf	for	$\overline{\text{Fin}}$
R	for	$\{x \mid (\forall y)(y \in x \rightarrow (\exists z_1)(\exists z_2)(y = \langle z_1, z_2\rangle))\}$
F	for	$\{x \mid x \in R \wedge (\forall y)(\forall z_1)(\forall z_2)(\langle y, z_1\rangle \in x \wedge \langle y, z_2\rangle \in x \rightarrow z_1 = z_2)\}$
F^1	for	$\{x \mid x \in F \wedge (\forall y)(\forall z_1)(\forall z_2)(\langle z_1, y\rangle \in x \wedge \langle z_2, y\rangle \in x \rightarrow z_1 = z_2)\}$
$D(t)$	for	$\{\chi \mid (\exists y)\langle \chi, y\rangle \in t\}$
$I(t)$	for	$\{y \mid (\exists \chi)\langle \chi, y\rangle \in t\}$
$t \cong u$	for	$(\exists \chi)(\chi \in F^1 \wedge D(\chi) = t \wedge I(\chi) = u)$
$Nc(t)$	for	$\{\chi \mid \chi \cong t\}$
$(t+u)$	for	$\{\chi \mid (\exists y)(\exists z)(\chi = y \cup z \wedge y \in t \wedge z \in u \wedge y \cap z = \Lambda)\}$
1	for	$0'$
2	for	$1'$
3	for	$2'$, etc.

T5. $\vdash_{\text{NF}} 0 \in N$ (=Peano postulate 1)
T6. $\vdash_{\text{NF}} (\forall x)(x \in N \rightarrow x' \in N)$ (=Peano postulate 2)
T7. $\vdash_{\text{NF}} (\forall x)\, 0 \neq x'$ (=Peano postulate 3)
T8. $\vdash_{\text{NF}} (\forall x)(0 \in x \wedge (\forall y)(y \in x \rightarrow y' \in x) \rightarrow N \subseteq x)$
T9. $\vdash_{\text{NF}} \mathscr{A}[0] \wedge (\forall x)(x \in N \wedge \mathscr{A}[x] \rightarrow \mathscr{A}[x']) \rightarrow (\forall x)(x \in N \rightarrow \mathscr{A}[x])$, where $\mathscr{A}[x]$ is any stratified wff.

The above theorem is a weaker version of the fifth Peano postulate. Its weaker because it limits the application of mathematical induction to stratified formulas. This is often seen as a defect of the system. (Fixing this defect was one of Quine's motivations for creating System ML, which we shall discuss soon.)

The remaining Peano postulate, viz.:

$$(\forall x)(\forall y)(x \in N \wedge y \in N \wedge x' = y' \rightarrow x = y)$$

is a theorem of NF. However, its proof is very complicated, and indeed follows from what many see as another defect of system NF: that the negation of the axiom of choice is a theorem. As we explained in our discussions of this result in F and ST, this Peano postulate requires that there be an infinite number of objects, or that the universal set V is not a member of any natural number.

Notice that we cannot establish an infinite set using the sort of proof we used for System F, in which we proved that the set ω has an infinite number of members. The definition given for ω cannot be given since $x \cup \{x\}$ is not stratified.

In his original article, Quine notes that $\Lambda, \{\Lambda\}, \{\{\Lambda\}\}, \{\{\{\Lambda\}\}\}, \ldots$ can each be shown to be different from each other. However, we cannot define a set that contains all and only members of this series. It is tempting to attempt to define such a set using the abstract:

$$\{x \mid (\forall y)(\Lambda \in y \wedge (\forall z)(z \in y \rightarrow \{z\} \in y) \rightarrow x \in y)\}$$

But this term is not stratified either.

We shall discuss how it is that the existence of an infinite set can be established for NF a bit later.

Some other results:

T10. $\vdash_{\text{NF}} (\forall x_1)(\forall x_2)(\forall y_1)(\forall y_2)(\langle x_1, y_1\rangle = \langle x_2, y_2\rangle \leftrightarrow x_1 = x_2 \wedge y_1 = y_2)$
T11. $\vdash_{\text{NF}} (\forall x)\, x \cong x$
T12. $\vdash_{\text{NF}} (\forall x)(\forall y)(x \cong y \rightarrow y \cong x)$
T13. $\vdash_{\text{NF}} (\forall x)(\forall y)(\forall z)(x \cong y \wedge y \cong z \rightarrow x \cong z)$

From which we derive Hume's Law in the form:

(HL) $\vdash_{\text{NF}} (\forall x)(\forall y)(Nc(x) = Nc(y) \leftrightarrow x \cong y)$

Definitions:

$t \ncong u$	for	$\neg t \cong u$
$\wp_u(t)$	for	$\{\chi \mid (\exists y)(y \in t \wedge \chi = \{y\})\}$

$\wp_u(t)$ is the set of all singletons (or "unit sets") of members of t. So if t is $\{0, 1, 2\}$, then $\wp_u(t)$ is $\{\{0\}, \{1\}, \{2\}\}$. ($\wp_u(t)$ is stratified if t is.)

T14. $\vdash_{\text{NF}} 1 = \wp_u(\text{V})$

HOMEWORK 18

Prove T14.

T15. $\vdash_{\text{NF}} (\forall x)\, x' = x + 1$

E. Cantorian and Non-Cantorian Sets in NF

Cantor's theorem states that no set is equinumerous with its own power set, i.e.:

$$(\forall x)\, x \not\cong \wp(x)$$

This theorem is usually proven by a *reductio as absurdum*. Suppose for any set a there is a 1–1 function f with a as domain and $\wp(a)$ as range. Let $e_1, e_2, e_3, \ldots$, etc. be the members of a. We can depict the mapping between a and its powerset as something such as the following:

Members of a		Members of $\wp(a)$
e_1	$\Longleftrightarrow$	$\{e_1, e_2\}$
e_2	$\Longleftrightarrow$	$\{e_1\}$
e_3	$\Longleftrightarrow$	$\{e_3\}$
e_4	$\Longleftrightarrow$	$\{e_1, e_2, e_3\}$
e_5	$\Longleftrightarrow$	$\{\ \}$
$\vdots$		$\vdots$

Notice that in any such purported mapping, some members of a will be mapped to a subset of a that contains it, and others won't be. Here, e_1 and e_3 are members of their corresponding subsets, but the others are not. Now, let w be the subset of a which contains all members of a not in the subset of a onto which f maps them, i.e.:

$$w \text{ for } \{x | x \in a \wedge (\forall y)(\langle x, y\rangle \in f \rightarrow x \notin y)\}$$

Now, w is a subset of a, and hence a member of the powerset of a, and therefore *should* be in the range of f. However, this is not possible, since there would have to be some member of a, viz., e_w of a for which w is the value of f for e_w as argument. But then one would get a contradiction since $e_w \in w \leftrightarrow e_w \notin w$. Hence, there cannot be such a function as f.

This proof works for systems such as F, ZF and NBG. However, it does not work for NF, since the definition given above for w is not stratified. Recall that in order for a wff containing "$\langle x, y\rangle$" to be stratified, it must be possible to assign the same number to "x" and "y". This is not possible with w since it contains "$x \notin y$".

Hence one cannot prove that $e_w \in w \leftrightarrow e_w \notin w$ or even that w is a member of $\wp(a)$. The usual proof of Cantor's theorem is not possible in NF.

In fact, the negation of Cantor's theorem is a theorem of NF. Notice that $\vdash_{\text{NF}} \text{V} \cong \text{V}$ by T11, and by T3, $\vdash_{\text{NF}} \wp(\text{V}) = \text{V}$. By LL, $\vdash_{\text{NF}} \text{V} \cong \wp(\text{V})$.

However, by a proof very similar to the one just given, it is possible to establish that $\wp_u(x)$ is never equinumerous with $\wp(x)$.

We simply consider any such mapping as:

Members of $\wp_u(a)$		Members of $\wp(a)$
$\{e_1\}$	$\Longleftrightarrow$	$\{e_1, e_2\}$
$\{e_2\}$	$\Longleftrightarrow$	$\{e_1\}$
$\{e_3\}$	$\Longleftrightarrow$	$\{e_3\}$
$\{e_4\}$	$\Longleftrightarrow$	$\{e_1, e_2, e_3\}$
$\{e_5\}$	$\Longleftrightarrow$	$\{\ \}$
$\vdots$		$\vdots$

We then define the subset of a that must be left out of the mapping as follows:

$$w^* \text{ for } \{x | x \in a \wedge (\forall y)(\langle \{x\}, y\rangle \in f \rightarrow x \notin y)\}$$

Here, the definition of w^* is stratified. It follows that $w^* \subseteq a$, and so that $w \in \wp(a)$. Therefore, there must be some e_{w^*} in a such that $\langle \{e_{w^*}\}, w^*\rangle \in f$. It then follows that $e_{w^*} \in w^* \leftrightarrow e_{w^*} \notin w^*$.

The conclusion of this line of reasoning gives us:

T16. $\vdash_{\text{NF}} (\forall x)\, \wp_u(x) \not\cong \wp(x)$
T16a. $\vdash_{\text{NF}} \wp_u(\text{V}) \not\cong \wp(\text{V})$

However, in virtue of T3 (and T11), we get the following *startling* result:

T17. $\vdash_{\text{NF}} \text{V} \not\cong \wp_u(\text{V})$

There are not equally many members of V as there are singletons sets taken from members of V. Sets that are not equinumerous with their own members' singletons are called ***non-Cantorian sets***.

The presence of non-Cantorian sets in NF is often regarded as its oddest feature. Notice, that the seemingly "obvious" proof that every set is Cantorian is blocked in NF, since $\langle x, \{x\}\rangle$ is not stratified.

F. The Failure of the Axiom of Choice in NF

Another startling result is that the Axiom of Choice is inconsistent with NF, i.e., one can prove its negation in NF:

$$\vdash_{\mathrm{NF}} \neg(\forall x)(\exists y)(F(y) \wedge (\forall z)(z \subseteq x \wedge z \neq \Lambda \rightarrow (\exists z_1)(\langle z, z_1\rangle \in y \wedge z_1 \in z)))$$

The actual proof is somewhat complicated. However, suppose (AC) held in NF. Apply UI on x to V. Then there would be a function y such that for all non-empty sets z, the value of y for z would be some member of z. If we consider the restriction of y to $\wp_u(\mathrm{V})$, we have a function whose domain is $\wp_u(\mathrm{V})$ and whose value for every singleton is a member of that singleton. The function in question would be a 1–1 function, since no two singletons share a member. Hence there would be a 1–1 function whose domain is $\wp_u(\mathrm{V})$ and whose range is V and it would follow that $\wp_u(\mathrm{V}) \cong \mathrm{V}$, contradicting T17 (and T12).

It follows from the above that the various equivalents of the Axiom of Choice (the multiplicative axiom, the trichotomy principle, the well-ordering principle, etc.) are also all disprovable in NF. Since the axiom of choice (and its equivalents) are widely accepted by practicing mathematicians, this is another reason for the relative unpopularity of NF.

Another issue is its consistency, which through most of NF's existence was considered an unsolved question. There is, however, currently a paper under review (by Randall Holmes) alleging to establish NF's consistency relative to traditional set theory.

Two more consequences of the failure of the axiom of choice are worth mentioning.

1. It can be proven that the axiom of choice holds for finite sets. It follows that the universal set is infinite, from which one can establish the fourth Peano postulate without postulating a special axiom of infinity (as was done by some early proponents of NF).
2. Since the generalized continuum hypothesis implies the axiom of choice, the generalized continuum hypothesis is also disprovable in NF.

The oddities of NF are sufficient to establish that at any rate it is less than an ideal system for the foundations of mathematics; indeed, Quine himself eventually stopped championing it as his sole preferred theory.

Nevertheless, the general method of finding some restriction on the comprehension principle of naïve set theory (F2) that leaves the system both intuitive, easy to work with, and allows for the derivation of large portions of mathematics remains an attractive one.

XXI. Variants of NF (ML, NFU)

A. Quine's ML

Perhaps the most influential variation on NF was the system ML, also proposed by Quine, three years after proposing NF, in his book *Mathematical Logic* (from which ML gets its name). The main motivation for ML was to obtain an unrestricted principle of mathematical induction, and a straightforward proof of infinite sets. (The proof of infinity for NF that makes use of the failure of the axiom of choice had not yet been found when Quine proposed ML.) The original version of ML was found to be inconsistent, but it was shown that the inconsistency was simply due to a slip in presentation. We here examine the revised version of ML from the second edition of Quine's book.

ML stands to NF much the way that NBG stands to ZF. ML makes a distinction between sets and proper classes. Every wff "comprehends" a class, but only stratified wffs "comprehend" sets.

ML is a first-order theory with only a single binary predicate, "∈". I find it convenient, as we did for NBG, to use capital letters for the variables, and use lowercase letters for restricted quantification over sets. The definition of being a set is the same as NBG.

Definitions:

$M(t)$ for $(\exists \mathscr{X})\, t \in \mathscr{X}$

Where χ is a lowercase variable, and $\mathscr{X}$ is the first uppercase variable that does not occur in $\mathscr{A}[\chi]$:
$(\forall \chi)\, \mathscr{A}[\chi]$ for $(\forall \mathscr{X})(M(\mathscr{X}) \to \mathscr{A}[\mathscr{X}])$
$(\exists \chi)\, \mathscr{A}[\chi]$ for $(\exists \mathscr{X})(M(\mathscr{X}) \land \mathscr{A}[\mathscr{X}])$

Definition: We say that a wff is ***predicative*** if and only if all its quantifiers are restricted quantifiers.

We define a wff as ***stratified*** just as we did for NF.

The proper axiom schemata of ML are the following:

ML1. $(\forall X)(\forall Y)(X = Y \to (\mathscr{A}[X,X] \to \mathscr{A}[X,Y]))$, where Y does not occur bound in $\mathscr{A}[X,X]$.

ML2. $(\exists Z)(\forall x)(x \in Z \leftrightarrow \mathscr{A}[x])$, where $\mathscr{A}[x]$ does not contain Z free.

ML3. $(\forall \mathscr{X}_1)\ldots(\forall \mathscr{X}_n)(M(\mathscr{X}_1) \land \ldots \land M(\mathscr{X}_n) \to (\exists Z)(M(Z) \land (\forall X)(X \in Z \leftrightarrow M(X) \land \mathscr{A}[X])))$,
where $\mathscr{A}[X]$ is predicative and stratified, and does not contain Z free, and $\mathscr{X}_1, \ldots, \mathscr{X}_n$ are the free variables of $\mathscr{A}[X]$ besides X.

ML2 postulates the existence of *classes*. ML3 postulates *sets* corresponding to predicative, stratified formulas.

Using contextual definitions similar to those in NF, we obtain:

$\vdash_{\text{ML}} (\forall x)(x \in \{y|\mathscr{A}[y]\} \leftrightarrow \mathscr{A}[x])$, provided that x does not become bound in the context $\mathscr{A}[x]$.

This appears to be the same as F2, but once again, notice that the above uses a restricted quantifier for sets. We also get:

$\vdash_{\text{ML}} M(\{y|\mathscr{A}[y]\}) \land (\forall x)(x \in \{y|\mathscr{A}[y]\} \leftrightarrow \mathscr{A}[x])$, provided that x does not become bound in the context $\mathscr{A}[x]$, and $\mathscr{A}[y]$ contains no free variables besides y, and $\mathscr{A}[y]$ is predicative and stratified.

There are important differences between the set/proper class distinction in ML and the similar distinction in NBG. In NBG, whether or not a class is a set has to do with its *size*. In ML, the distinction has to do with its defining membership conditions. Indeed, in ML, the universal class turns out to be a set, and indeed, is a member of itself.

HOMEWORK 19

Prove: $\vdash_{\text{ML}} \{y|y = y\} \in \{y|y = y\}$

Definition:

N for $\{x|(\forall Y)(0 \in Y \land (\forall z)(z \in Y \to z' \in Y) \to x \in Y)\}$

From this, one gets the full principle of mathematical induction, for any wff $\mathscr{A}[x]$:

$\vdash_{\text{ML}} \mathscr{A}[0] \land (\forall x)(x \in N \land \mathscr{A}[x] \to \mathscr{A}[x']) \to (\forall x)(x \in N \to \mathscr{A}[x])$

Hence, all of Peano arithmetic is provable in ML.

Despite the greater mathematical strength of ML, it has been proven that ML and NF are equiconsistent: one is consistent if and only if the other is.

B. Jensen's System NFU

Ernest Specker proved that NF is consistent relative to system sometimes known as *simple type-theory with complete typical ambiguity*, which is obtained from ST (without the axiom of infinity), by adding an an inference to the effect that whatever holds in one type holds in any other. However, since this theory has not been shown consistent even relative to ST, this does not put NF on more solid footing.

Another variant of NF, called NFU, however, has been shown consistent relative to ST. (And since ST is consistent relative to ZF, this shows that NFU is also consistent relative to ZF.) NFU is a relatively natural system that one might suggest anyway as a modification of NF suited to allow for the existence of urelements (or urelements *not* identified with their own singletons.)

The syntax of NFU differs from that of NF in taking both membership and identity as primitive. (Thus identity is not defined in terms of having all the same members.) The system is built on the *predicate calculus with* *identity* (and thus has the reflexivity of identity and LL as axioms). In addition to these, it adds axioms of extensionality (for sets) and class abstraction:

NFU1. $(\forall x)(\forall y)((\exists z) z \in x \wedge (\forall z)(z \in x \leftrightarrow z \in y) \rightarrow x = y)$

NFU2. $(\forall \chi)(\chi \in \{y|\mathscr{A}[y]\} \leftrightarrow \mathscr{A}[\chi])$, provided that $\mathscr{A}[y]$ is stratified, and χ does not become bound when placed in the context $\mathscr{A}[\chi]$.

(NFU2 is the same as NF2.)

Unlike NF, one cannot prove the negation of the axiom of choice in NFU: nor even can one establish the existence of any infinite sets. (Unfortunately, then, one also cannot prove the fourth Peano postulate without adding some additional axioms.)

Nevertheless, the consistency of NFU to ST may show that Quine's equating individuals with their own singletons might not have been so "harmless".

XXII. Cocchiarella's Higher-Order Variants

A. Background

Nino Cocchiarella has attempted what he calls a "reconstruction" of logicism based on a higher-order system employing a notion of stratification at least partly inspired by Quine.

The other part of motivation comes from the work of Frege (and very early Russell). According to the way Frege thinks language works, some expressions refer to objects, and others to concepts. Names like "Plato", "Berlin" and descriptions, e.g., "the largest star in the Big Dipper", refer to objects. The remainder of a sentence, when a referring expression such as a name is removed, refers to a *concept*.

____ is a horse.
____ orbits the Sun.

Frege thought that both such expressions, and *what they refer to* (concepts) are in some sense "incomplete" or "unsaturated".

The above are examples of sentences from which a name has been removed. Such expressions refer to what Frege calls *first-level concepts*, or concepts applicable to objects. If one removes the name of a first-level concept from a sentence, one is left with the name of a *second-level concept*.

E.g., if we remove "is a horse" from these sentences:

> Something is such that it is a horse.
> If Socrates is a horse, then Plato is a horse.

We obtain:

> Something is such that . . . (it) . . . .
> If . . . (Socrates). . . , then . . . (Plato). . . .

Depending on which name of a first-level function completes the gap, one is left with something true or something false. Second-level concepts are applicable (or not) to first-level concepts. This forms the beginning of a *hierarchy* of levels of concepts akin to what is found in higher-order simple type-theory.

Frege regarded it as impossible for a concept to be predicated of itself. The reason of course is that different kinds of concepts exhibit different kinds of incompleteness. A concept never has the right sort of incompleteness in order to "complete" (or as Frege says, "saturate" itself.)

"____ is a horse" must be completed by a name, not something such as "____ is a horse" to yield a complete, grammatical sentence. "If . . . (Plato). . . , then . . . (Socrates). . . " can be completed by "____ is a horse" to form something grammatical, but not by itself, etc.

This hierarchy of levels of concepts was reflected in the grammar of Frege's logical language. Different styles of variables were used for different levels, and it would be nonsense to instantiate a variable of one type to an expression of another, or place a variable of one type in its own argument spot, e.g., "$F(F)$" or "$\neg F(F)$". Recall that for Frege, the expression for the first-level concept is not simply "F", but "$F(\ )$", revealing the incompleteness or unsaturation of the concept.

The use of different styles of variables, and different "shape" expressions for different logical types in some ways sets Frege's higher-order logic apart from modern renditions that make use of lambda abstracts. If the second level concept written in everyday English as "Something is such that … (it)… ." is written as "$(\exists x)\ \ldots x \ldots$", it is quite clear that it does not "fit" in its own blank space. It is not *physically possible* to violate the type-restriction.

However, if it is written instead "$[\lambda F\,(\exists x)\,Fx]$", it is at least physically possible to write, "$[\lambda F\,(\exists x)\,Fx]([\lambda F\,(\exists x)\,Fx])$", and one would need special grammatical rules (involving, e.g., type indices) to exclude such a formula.

Frege went so far in analyzing ordinary language as to suggest that an expression such as "the concept horse", because it is not an incomplete or gappy expression, and because it can fill the blank spot in an expression such as "____ is a horse", or "____ orbits the Sun" must not, despite appearance, refer to a concept, but to an object.

Thus Frege thought that for every concept, there was an object that could "go proxy for it", i.e., would be referred to by a nominalized predicate derived from a predicate expression for the concept. Thus, "the concept horse", "Humanity", "Kindness", etc., refer to what Cocchiarella calls "concept-correlates". There is also evidence, as Cocchiarella points out, that Frege identified these "concept-correlates" with the extension or value-range of a concept. Frege further claimed that the extension or *value-range* of a concept "had its being" in the concept: that they were in some sense, the same entity in a metaphysical sense.

For a given concept expression $F(\ )$, one refers to its value-range as "$\dot{\alpha}(F(\alpha))$". In my formulation of Frege's GG, I changed this notation to $\{x|Fx\}$, in keeping with the usual reading of Frege's extensions of concepts as *classes*. Cocchiarella, however, suggests that Frege's logic of value-ranges can instead be thought of as a logic of *nominalized predicates*, i.e., one in which expressions for concepts can occur in both predicate and subject positions. In effect, he is suggesting that $F(\dot{\alpha}(F(\alpha)))$ is really just a variation on $F(F)$.

The effect then of including value-ranges is more or less to *undo* the effects of having distinct levels. Every first-level concept corresponds to an object: its concept-correlate (or value-range), etc. Every second-level concept can be represented by a first-level concept which applies to an object just in case that object is the correlate of a concept to which the second-level concept applies. Cocchiarella calls this Frege's "***double-correlation thesis***":

$$(\forall M)(\exists G)(\forall F)(M_\beta F(\beta) \longleftrightarrow G(\dot{\alpha}(F(\alpha))))$$

By repeating this reasoning, in effect, all concepts can be "reduced in level" all the way down.

Of course, the introduction of concept-correlates as objects corresponding to *any arbitrary concept allows* one in effect to apply concepts to themselves, and thus obtain Russell's paradox in the form gotten by considering whether or not the concept whose value range is:

$$\dot{\epsilon}((\exists F)(\epsilon = \dot{\alpha}(F(\alpha)) \wedge \neg F(\epsilon)))$$

applies to that very value-range.

Cocchiarella suggests a *reconstruction* of Frege's logicism in which *some concepts* are thought to have objects corresponding to them, others not. To determine which ones do, and which ones don't, Cocchiarella modifies Quine's notion of *stratification*. Here it is used not to restrict which sets are postulated to exist, but which concepts can be "converted" into objects.

Rather than employing Frege's notation "$\dot{\epsilon}(\ldots\epsilon\ldots)$", or the set theoretic notation "$\{x|\ldots x\ldots\}$", Cocchiarella employs what is now the usual notation for forming complex predicates, viz. "$[\lambda x \ldots x \ldots]$". However, such an expression, if *homogeneously stratified* (defined below) may occur either in a predicate *or* in a subject position, so that we may allow some instances of:

$$[\lambda x\ \mathscr{A}[x]]([\lambda x\ \mathscr{A}[x]])$$

This is supposed to be analogous to Frege's allowance of such constructions as "the concept horse is a horse", in which, a concept is predicated of its own concept-correlate object.

Cocchiarella's systems are formulated in the context of a second-order logic (higher orders are not necessary given the possibility of *reducing level*). Allowing suitably stratified λ-abstracts as valid substituends of both individual and predicate variables in part undoes the usual type-distinctions employed within a second, but this is no more dangerous (suggests Cocchiarella) than the effect of removing the type-restrictions of ST in favor of the stratification restrictions employed in system such as NF.

B. The System of Homogeneous Simple Types (λ-HST*): Syntax

Definition: An ***individual variable*** is any lowercase letter from 'u' to 'z' with or without a numerical subscript.

Definition: A ***predicate variable*** is any uppercase letter from 'A' to 'T', with a numerical superscript ≥ 1 (indicating how many terms it is applied to), and with or without a numerical subscript.

Well-formed expressions come in two varieties: ***well-formed formulas*** (which have no "kind")[4] and ***terms*** (which have a kind depending on how many arguments they can take):

Definition: A ***well-formed expression (wfe)*** is defined recursively as follows:

(i) individual variables are well-formed expressions of kind 0; predicate variables are well-formed expressions both of kind 0 and of kind n given by their superscripts;
(ii) if t and u are wfes of kind 0, then $(t = u)$ is a wfe without a kind;
(iii) if $\mathscr{P}$ is a wfe of kind n, and $t_1, \ldots, t_n$ are all wfes of kind 0, then $\mathscr{P}(t_1, \ldots, t_n)$ is a wfe without a kind.
(iv) If $\mathscr{A}$ is a wfe without a kind, and $\chi_1, \ldots, \chi_n$ are *distinct* individual variables, and $[\lambda \chi_1 \ldots \chi_n \mathscr{A}]$ is homogeneously stratified (defined below), then $[\lambda \chi_1 \ldots \chi_n \mathscr{A}]$ is a wfe both of kind n and of kind 0; (Limiting case: if $\mathscr{A}$ is a wfe without a kind, and $[\lambda\ \mathscr{A}]$ is homogeneously stratified, then $[\lambda\ \mathscr{A}]$ is a wfe of kind 0.)
(v) If $\mathscr{A}$ is a wfe without a kind, then $\neg\mathscr{A}$ is a wfe without a kind;
(vi) If $\mathscr{A}$ and $\mathscr{B}$ are each wfes without a kind, then $(\mathscr{A} \vee \mathscr{B})$ is a wfe without a kind;
(vii) If $\mathscr{A}$ is a wfe without a kind, and χ is either an individual or predicate variable, then $(\forall \chi)\mathscr{A}$ is a wfe without a kind.

Definition: A formula or string of symbols $\mathscr{A}$ made up of the elements of the language of λ-HST* is said to be ***homogeneously stratified*** iff there is a function f from the terms (or would-be terms) t making up $\mathscr{A}$ to the natural numbers such that:

(i) if $t = u$ occurs in $\mathscr{A}$, then $f(t) = f(u)$;
(ii) If $\mathscr{P}(t_1, \ldots, t_n)$ occurs in $\mathscr{A}$, then $f(t_1) = f(t_2) = \ldots = f(t_n)$, and $f(\mathscr{P}) = f(t_1) + 1$.
(iii) If $[\lambda \chi_1 \ldots \chi_n \mathscr{B}]$ occurs in $\mathscr{A}$ then $f(\chi_1) = f(\chi_2) = \ldots = f(\chi_n)$, and $f([\lambda \chi_1 \ldots \chi_n \mathscr{B}]) = f(\chi_1) + 1$.

We can see from the above, that e.g.:

$$[\lambda x\,(\exists F)(x = F \wedge \neg F(x))]$$

is not homogeneously stratified, and so does not count as a well-formed λ-abstract in our syntax.

On the other hand, the abstract:

$$[\lambda x\,(\exists F)(\exists y)(x = F \wedge \neg F(y))]$$

is homogeneously stratified, in virtue of the assignment of 0 to y, 1 to both x and F, and 2 to the whole abstract.

If a λ-abstract occurs without binding any variables, it is meant as a name of a *proposition*, so that, e.g., $[\lambda\ (\forall x)\, x = x]$ represents the proposition that everything is self-identical.

λ-HST*: Formulation

Cocchiarella lays out the system using a slightly different axiomatization of the underlying logic. However, I shall not delve into the details. It suffices to note that it is possible to instantiate individual variables to any term of kind 0, including lambda abstracts (if homogeneously stratified), and to instantiate predicate variables to any term of the appropriate kind, including lambda abstracts (again, if homogeneously stratified).

[4] Instead of "kind", Cocchiarella speaks of "types"; his type 1 is my "no kind" and his type $n+2$ is my kind $n+1$.

The system includes the following axiom schemata, in addition to the standard laws for identity logic (Ref=, LL):

(λ-Conv*)
$(\forall y_1)\ldots(\forall y_n)([\lambda x_1 \ldots x_n\, \mathscr{A}[x_1,\ldots,x_n]](y_1,\ldots,y_n) \leftrightarrow \mathscr{A}[y_1,\ldots,y_n])$, where $x_1,\ldots x_n, y_1,\ldots,y_n$ are distinct individual variables (and in order to be well-formed, the λ-abstract needs to be homogeneously stratified).

(Id*)
$(\forall F^n)([\lambda x_1 \ldots x_n\, F^n(x_1,\ldots,x_n)] = F^n)$

Although not an axiom schema of λ-HST* as such, Cocchiarella considers adding the following in certain expansions of the system:

(Ext*):
$$(\forall x_1)\ldots(\forall x_n)(\mathscr{A}[x_1,\ldots,x_n] \leftrightarrow \mathscr{B}[x_1,\ldots,x_n]) \rightarrow [\lambda x_1 \ldots x_n\, \mathscr{A}(x_1,\ldots,x_n)] = [\lambda x_1 \ldots x_n\, \mathscr{B}(x_1,\ldots,x_n)]$$

This constrains us to take concept-correlates as *extensionally* individuated. This is in keeping with Frege's understanding of concepts as *functions* from objects to truth-values.

HOMEWORK 20

Prove the following version of Basic Law V in Cocchiarella's λ-HST* + (Ext*):

$$\vdash_{\lambda\text{-HST*+(Ext*)}} (\forall F)(\forall G)(F = G \leftrightarrow (\forall x)(Fx \leftrightarrow Gx))$$

The system λ-HST* + (Ext*) is roughly as powerful as (and is equiconsistent with) Jensen's NFU. To obtain something equally powerful as Quine's NF, one must also add the assumption that every entity is a concept-correlate:

(Q*) (for "Quine's Thesis")
$(\forall x)(\exists F)\, x = F$

Since Cocchiarella sees no reason to accept (Q*) (and indeed, alleges it is not something Frege would accept), Cocchiarella believes that λ-HST* + (Ext*) is the better reconstruction of Frege's logicism.

As a reconstruction of other views about the identity conditions of concepts/concept-correlates, Cocchiarella also discusses a modal version of λ-HST* (called $\Box\lambda$-HST*) obtained by adding the modal operator $\Box$ (for necessity) to the language, supplementing it with the axioms and rules of S5 modal logic and weakening (Ext*) to the following:

($\Box$Ext*)
$$(\Box(\forall x_1)\ldots(\forall x_n)(\mathscr{A}[x_1,\ldots,x_n] \leftrightarrow \mathscr{B}[x_1,\ldots,x_n])) \rightarrow [\lambda x_1 \ldots x_n\, \mathscr{A}(x_1,\ldots,x_n)] = [\lambda x_1 \ldots x_n\, \mathscr{B}(x_1,\ldots,x_n)]$$

I.e., concept-correlates are identical when they're *necessarily* coextensive. To obtain a large portion of mathematics, one must also assume that every concept is coextensive with a *rigid* concept (one that either necessarily holds or necessarily doesn't hold of a given object or objects):

Definitions:

Rigid_n	for	$[\lambda x\, (\exists F^n)(x = F^n \wedge (\forall y_1)\ldots(\forall y_n)(\Box F(y_1,\ldots,y_n) \vee \Box\neg F(y_1,\ldots,y_n)))]$
Cls	for	$[\lambda x\, (\exists F^1)(x = F^1 \wedge \text{Rigid}_1(F^1))]$

(PR)
$$(\forall F^n)(\exists G^n)(\text{Rigid}_n(G^n) \wedge (\forall x_1)\ldots(\forall x_n)(F^n(x_1,\ldots,x_n) \leftrightarrow G^n(x_1,\ldots,x_n)))$$

Adding ($\Box$Ext*) and (PR) to $\Box\lambda$-HST* results in a rival (more "Russellian", or so Cocchiarella claims) reconstruction of logicism. This system too is equiconsistent with Jensen's NFU.

Since λ-HST* + (Ext*) is only as strong as NFU, it has certain limitations as a system for the foundations of mathematics. In particular, no theorem of infinity is provable within it (blocking the derivation of the 4th Peano postulate), and mathematical induction is limited to stratified formulas (where natural numbers are defined as follows:)

Definitions:

0	for	$[\lambda x\, (\exists F)(x = F \wedge (\forall y)\neg Fy)]$
S	for	$[\lambda xy\, (\exists F)(\exists G)(x = F \wedge y = G \wedge (\forall H)(G(H) \leftrightarrow (\exists z)(Hz \wedge F([\lambda w\, Hw \wedge w \neq z]))))]$
N	for	$[\lambda x\, (\forall F)(F(0) \wedge (\forall y)(\forall z)(Fy \wedge Syz \rightarrow Fz) \rightarrow Fx)]$

To obtain something strong one must add either an axiom of infinity, or (Q*), etc.

C. HST^*_λ: Formulation

To obtain a stronger principle of mathematical induction, Cocchiarella also considers a system that stands to λ-HST* the way ML stands to NF. Cocchiarella calls the system HST^*_λ.

We shall not discuss in full detail the exact formulation of HST^*_λ, in part because it does not employ a standard logical core, but instead a "free logic" where not all well-formed terms are referential. Those of you familiar with Hardegree's free description logic (from his Intermediate Logic course) may get the idea.

All λ-abstracts are considered well-formed terms, whether or not homogeneously stratified, and all are considered valid substituends of higher-order quantifiers (e.g., "$(\forall F)\ldots$", etc.). However, only some such abstracts are considered valid substituends of the individual variables. I.e., in order to instantiate a quantifier of the form "$(\forall x)\ldots x\ldots$" to a Lambda abstract (or to existentially generalize from it), one needs a result of the form:

$$(\exists y)([\lambda \chi_1 \ldots \chi_n\, \mathscr{A}[\chi_1,\ldots,\chi_n]] = y)$$

One begins only with the assumption:

(∃/HSCP*)
$(\forall f_1)\ldots(\forall f_n)((\exists y)(y = f_1) \wedge \ldots\ (\exists y)(y = f_n) \rightarrow$
$(\exists y)([\lambda \chi_1 \ldots \chi_n\, \mathscr{A}[\chi_1,\ldots,\chi_n]] = y))$,
where f_1, …, f_n are the free predicate variables of $[\lambda \chi_1 \ldots \chi_n\, \mathscr{A}[\chi_1,\ldots,\chi_n]]$, and y is the first individual variable not occurring therein, and $\mathscr{A}[\chi_1,\ldots,\chi_n]$ is *homogeneously stratified* and *bound to individuals* (see below).

Definition: A wff is ***bound to individuals*** when all predicate quantifiers $(\forall f)\ldots$ in it occur in a context of the form $(\forall f)((\exists \chi)\,\chi = f \rightarrow \ldots)$. (The notion of being bound to individuals is analogous to the ML notion of being predicative.)

Whence we get, e.g.:

$$(\exists y)[\lambda x\, x = x] = y$$

However, we do not get:

(a) $(\exists y)[\lambda x\,(\forall F)(x = F \rightarrow \neg F x)] = y$
since the abstract in question is not homogeneously stratified.
We do, however, get:
(b) $(\exists G)[\lambda x\,(\forall F)(x = F \rightarrow \neg F x)] = G$
Without (a), however, the individual variables in (λ-conv*) cannot be instantiated to the abstract $[\lambda x\,(\forall F)(x = F \rightarrow \neg F x)]$, thus preventing a contradiction.

The system HST^*_λ + (Ext*) + (Q*) is equiconsistent with, and roughly equally strong as, Quine's ML. Of course, because Cocchiarella doubts (Q*), he prefers the weaker HST^*_λ + (Ext*), which is roughly as strong as MLU, which would be obtained from ML by allowing for urelements, in the same way that NF is modified to become NFU.

While the definitions given for 0 and S are not bound to individuals, it is possible to prove that they are equivalent to definitions using formulas that are bound to individuals. The same cannot easily be said of N:

N for $[\lambda x\,(\forall F)(F(0) \wedge (\forall y)(\forall z)(F y \wedge S y z \rightarrow F z) \rightarrow F x)]$

Recall that the reason for advanting HST^*_λ instead of the simpler λ-HST* is to obtain an unrestricted principle of mathematical induction, applicable to both stratified and unstratified formulas. To maintain this, we need to keep the higher-order quantifier in F not restricted to individuals.

Hence, Cocchiarella considers adding as an additional axiom the following assumption:

$$(\exists x)\, N = x$$

Thus making N an individual. (This is not actually needed, however, to obtain any results about finite numbers.)

However, it is worth noting that, without (Q*), no theorem of infinity in forthcoming in HST^*_λ + (Ext*). However, Cocchiarella regards an axiom of infinity, when applied to *concept-correlates* rather than sets or classes—or as in ST, to urelements or individuals—to have some plausibility even as a *logical truth*.

Indeed, as he notes, there would need to be an infinite number of concept-correlates if all we assume is the following:

(Inf*) $(\forall F^n)(\forall G^m)F^n \neq G^m$, where n and m are distinct natural numbers.

In other words, concept-correlates corresponding to properties are never identical to concept-correlates corresponding to two-place relations, and concept-correlates corresponding to to two-place relations are never identical to concept-correlates corresponding to three-place relations, and so on. Since at least one concept-correlate (as an individual) can be shown to exist for every kind, this means that there must be an infinite number of concept-correlates.

XXIII. Hume's Law, Frege's Theorem and Frege Arithmetic

A. Review

In our first unit, we examined a method using higher-order logic as a foundations of mathematics in which numbers were defined as concepts applicable to concepts: so 0 would be the concept a concept has if it applies to nothing, 1 would be a concept a number has if it applies to exactly one thing, and so forth. There were (at least) two pitfalls to this approach. One was that in order to get the right identity conditions for numbers, we needed to assume that concepts themselves have extensional identity conditions. I.e., we had to add (Ext) as an axiom.

Another downfall was that if concepts are *only* considered as values of higher-order variables, one cannot use them to establish the existence of any objects. Hence, in order to establish the fourth Peano postulate, we had simply to *assume* the existence of infinitely many objects or individuals, i.e., assume (Inf).

Cocchiarella's systems have similar drawbacks. They require some postulation similar to (Ext)—e.g., Cocchiarella's (Ext*) or (□Ext*)—and although, since concept-correlates are taken as objects or individuals, and so one can establish the existence of some objects or individuals without taking additional axioms, one still needs some further assumption in order to derive an infinity of them.

In our first unit, we also looked at HOPA, or higher-order Peano arithmetic, in which numbers were treated not as concepts, but as objects. However, there the basic principles of numbers were simply taken as axioms.

We shall now look at other second- (or higher-) order systems for the foundations of mathematics, in which numbers are treated as objects or individuals, but in which one does not assume an axiom of extensionality, or take the Peano postulates (or anything similar) as primitive, but instead attempts to derive them from some more basic—possibly logical or at least "analytic"— assumptions.

B. Second-order Logic

We begin by sketching the syntax of the pure second-order predicate calculus. It is more or less just a simplification of HOPC taken by just eliminating all variables with order greater than 1. Thus the remaining variables are of types are o (objects or individuals, (o) (properties of objects or individuals), (o,o) (two-place relations between objects and individuals), (o,o,o) (three-place relations between individuals), and so on for any n-place relation between individuals. It is convenient to replace these type symbols with speaking of variables of kind n, where kind 0 is the kind of individuals, kind 1 is the kind of properties, kind 2 is the kind of two-place relations, and so on, as we did for Cocchiarella's systems. The difference, of course, is that we no longer include predicate variables and lambda-abstracts also to be terms of kind 0.

As before, we use lowercase letters as individual variables, and uppercase letters with superscript n for predicate variables of kind n.

Definition: A well-formed expression is defined as follows:

(i) Individual variables are well-formed expressions of kind 0; predicate variables are well-formed expressions of kind n given by their superscripts;

(ii) if $\mathscr{P}$ is a wfe of kind n, and $t_1, \ldots, t_n$, are n-many wfes of kind 0, then $\mathscr{P}(t_1, \ldots, t_n)$ is a wfe without a kind;

(iii) if $\mathscr{A}$ is a wfe without a kind, and $\chi_1, \ldots, \chi_n$ are n-many *distinct* individual variables, then $[\lambda \chi_1 \ldots \chi_n \mathscr{A}]$ is a wfe of kind n;

(iv) if $\mathscr{A}$ and $\mathscr{B}$ are each a wfe without a kind, then $(\mathscr{A} \vee \mathscr{B})$ is a wfe without a kind;
(v) if $\mathscr{A}$ is a wfe without a kind, then $\neg\mathscr{A}$ is a wfe without a kind;
(vi) if $\mathscr{A}$ is a wfe without a kind, and χ is either an individual or predicate variable, then $(\forall \chi)\, \mathscr{A}[\chi]$ is a wfe without a kind.

Definitions: As before, a wfe without a kind is called a ***well-formed formula (wff)***; a wfe with a kind is called a ***term***.

It is typical to employ the following definition of identity for wfes of kind 0:

$t = u$ for $(\forall f^1)(f^1(t) \leftrightarrow f^1(u))$

In addition to normal logical axioms, allowing the instantiation of quantifiers with variables of kind n to terms of kind n, etc.), we also have a principle of lambda conversion (λ-conv):

$$(\forall y_1)\ldots(\forall y_n)([\lambda \chi_1 \ldots \chi_n\, \mathscr{A}[\chi_1, \ldots, \chi_n]](y_1, \ldots, y_n) \leftrightarrow \mathscr{A}[y_1, \ldots, y_n]),$$

where $\chi_1, \ldots, \chi_n, y_1, \ldots, y_n$ are distinct individual variables.

An easy consequence of λ-conv, and existential generalization for variables of kind n is the following, called the *comprehension principle*:

(CP)
$(\exists F^n)(\forall x_1)\ldots(\forall x_n)(F(x_1, \ldots, x_n) \leftrightarrow \mathscr{A}[x_1, \ldots, x_n])$, where $\mathscr{A}[x_1, \ldots, x_n]$ is any wff not containing F^n free.

Second-order predicate logic is sometimes formulated instead by not including λ-abstracts, but instead taking (CP) as an axiom instead. The differences between such formulations and ours are usually trivial.

In what follows we consider certain small expansions of the above system sufficient for arithmetic.

C. Frege Arithmetic

One obtains what amounts to Frege's system GG from the above either by allowing all Lambda abstracts as terms of kind 0—as done by Cocchiarella, but regardless of whether or not they are stratified or predicative—or by adding a functor that maps every concept to its "extension" or "concept-correlate". That is, we add to the syntax the following stipulation:

(vii) If $\mathscr{P}$ is a wfe of kind 1, then $\mathrm{Ext}(\mathscr{P})$ is a wfe of kind 0.

One can then add what amounts to Frege's Basic Law V as follows:

(BLV) $(\forall F)(\forall G)(\mathrm{Ext}(F) = \mathrm{Ext}(G) \leftrightarrow (\forall x)(Fx \leftrightarrow Gx))$

$\mathrm{Ext}([\lambda x\, \mathscr{A}[x]])$ would then be a variant of the notations $\{x|\mathscr{A}[x]\}$ or $\dot{\epsilon}(\mathscr{A}[\epsilon])$.

Unfortunately, with the following definition:

$t \in u$ for $(\exists f^1)(u = \mathrm{Ext}(f) \wedge f(t))$

BLV leads to the unrestricted theorem schema:

$(\forall x)(x \in \mathrm{Ext}([\lambda y\, \mathscr{A}[y]]) \leftrightarrow \mathscr{A}[x])$

And hence to Russell's paradox.

However, the main use of BLV for Frege's treatment of numbers was in its definition of the number belonging to the concept F as the extension of the concept *extension of a concept equinumerous with the concept F*. I.e.:

Definitions:

Where f and g are terms of kind 1, r is an appropriate variable of kind 2, and χ, y and z are variables of kind 0:

$f \cong g$ for

$$(\exists r)((\forall \chi)(\forall y)(\forall z)(r(\chi, y) \wedge r(\chi, z) \rightarrow y = z) \wedge (\forall \chi)(\forall y)(\forall z)(r(y, \chi) \wedge r(z, \chi) \rightarrow y = z) \wedge (\forall \chi)(f(\chi) \rightarrow (\exists y)(g(y) \wedge r(\chi, y))) \wedge (\forall \chi)(g(\chi) \rightarrow (\exists y)(f(y) \wedge r(y, \chi))))$$

(Note: other authors write F Eq G, $\mathrm{Eq}_x(Fx, Gx)$, $F \approx G$ instead of $F \cong G$.)

$\#(f)$ for $\mathrm{Ext}([\lambda \chi\, (\exists g)(\chi = \mathrm{Ext}(g) \wedge f \cong g)])$

Note: in a formulation of second-order logic without λ-abstracts, one needs to use instead of the simple notation "$\#(f)$", a vbto, usually written

"$\#_x : \mathscr{A}[x]$" or "$N_x : \mathscr{A}[x]$" for "the number of x such that $\mathscr{A}[x]$". With λ-abstracts, one can write simply "$\#([\lambda x\, \mathscr{A}[x]])$" for the same thing.

These definitions allowed Frege to obtain an important principle often known as *Hume's Law* or *Hume's Principle*:

(HL) $(\forall F)(\forall G)(\#(F) = \#(G) \leftrightarrow F \cong G)$

After obtaining (HL), Frege derives the basic properties of numbers, i.e., the Peano postulates, from Hume's Law alone, and never returns to Basic Law V.

It has been suggested (especially by Crispin Wright) that Frege's derivation of the basic properties of numbers from (HL) should be regarded as a substantial achievement, even if his proof of (HL) relied on the dubious (BLV).

This suggests that instead of defining numbers as extensions, we may take numbers as primitive, i.e., instead of:

(vii) If $\mathscr{P}$ is a wfe of kind 1, then Ext($\mathscr{P}$) is a wfe of kind 0.

We add the stipulation:

(vii) If $\mathscr{P}$ is a wfe of kind 1, then $\#(\mathscr{P})$ is a wfe of kind 0.

Thus taking the notation $\#(\mathscr{P})$ as primitive, rather than defining it as Frege did.

Frege Arithmetic (FA) is the system obtained by adding (HL) (with #( ) taken as primitive) as an axiom to the second-order predicate calculus.

It turns out that FA is consistent iff PA2 is. (PA2 is second-order Peano Arithmetic—the system like HOPA except restricting the language to 2nd order formulas.)

D. Frege's Theorem

Frege's theorem is the result that all of Peano Arithmetic can derived in second-order logic from HL alone (i.e., in system FA).

Definitions:

0	for	$\#([\lambda x\, x \neq x])$
$t\ P\ u$	for	$(\exists f)(\exists z)(u = \#(f) \wedge f(z) \wedge t = \#([\lambda y\, f(y) \wedge y \neq z]))$
$N(t)$	for	$(\forall f)(f(0) \wedge (\forall z)(\forall y)(f(z) \wedge z\ P\ y \rightarrow f(y)) \rightarrow f(t))$

"$t\ P\ u$" can be read "t precedes u" or "u is the successor or t".

Results:

(Ref=) $\vdash_{FA} (\forall x)\, x = x$

Proof: UG on a tautology.

(LL) $\vdash_{FA} (\forall x)(\forall y)(x = y \rightarrow (\mathscr{A}[x,x] \rightarrow \mathscr{A}[x,y]))$, where y is not bound in $\mathscr{A}[x,x]$.

Proof: Same as in HOPC.

(Ref≅) $\vdash_{FA} (\forall F)\, F \cong F$
(Sym≅) $\vdash_{FA} (\forall F)(\forall G)(F \cong G \rightarrow G \cong F)$
(Trans≅) $\vdash_{FA} (\forall F)(\forall G)(\forall H)(F \cong G \wedge G \cong H \rightarrow F \cong H)$

Proof sketch: Identity pairs Fs one-to-one with the Fs. If a relation R pairs F*s* up with Gs, then its converse pairs the Gs up with Fs. If R pairs up the Fs and the Gs, and R' pairs up the Gs and the Hs, then the product of R and R' pairs the Fs up with Hs.

I also add:

(Ext≅) $\vdash_{FA} (\forall F)(\forall G)((\forall x)(Fx \leftrightarrow Gx) \rightarrow F \cong G)$

Proof: Easy result of the definition of ≅ and (Ref≅).

(PP1) $\vdash_{FA} N(0)$

Proof: UG on a tautology.

(Zero) $\vdash_{FA} (\forall F)(\#(F) = 0 \rightarrow \neg(\exists x)Fx)$

Proof:

(1)	1. $\#(F)=0$	Hyp
(2)	2. $(\exists x)Fx$	Hyp
(2)	3. Fa	2 EI
(1)	4. $\#(F)=\#([\lambda x\, x\neq x])$	1 Df. 0
	5. $\#(F)=\#([\lambda x\, x\neq x]) \leftrightarrow F\cong[\lambda x\, x\neq x]$	(HL) UI×2
(1)	6. $F\cong[\lambda x\, x\neq x]$	4, 5 BMP
(1)	7. $(\exists R)((\forall y)(\forall z)(\forall w)(Ryz\wedge Ryw\to z=w)\wedge$ $(\forall y)(\forall z)(\forall w)(Rzy\ \ \&\ \ Rwy\to z=w)\wedge$ $(\forall y)(Fy\to(\exists z)([\lambda x\, x\neq x](z)\wedge Ryz))\wedge$ $(\forall y)([\lambda x\, x\neq x](y)\to(\exists z)(Fz\wedge Rzy)))$	6 Df. $\cong$
(1)	8. $(\forall y)(Fy\to(\exists z)([\lambda x\, x\neq x](z)\wedge Ayz))$	7 EI, Simp
(1,2)	9. $(\exists z)([\lambda x\, x\neq x](z)\wedge Aaz)$	3, 8 UI, MP
(1,2)	10. $[\lambda x\, x\neq x](b)\wedge Aab$	9 EI
(1,2)	11. $b\neq b$	10 Simp, λ-conv
	12. $b=b$	Ref= UI
(1)	13. $\neg(\exists x)Fx$	2, 11, 12 RAA
	14. $\vdash_{\text{FA}} (\forall F)(\#(F)=0\to\neg(\exists x)Fx)$	1–13 CP, UG

(PP3) $\vdash_{\text{FA}} (\forall x)\neg x\ P\ 0$

Proof:

(1)	1. $x\ P\ 0$	Hyp
(1)	2. $(\exists F)(\exists y)(0=\#(F)\wedge Fy\wedge x=\#([\lambda z\, Fz\wedge z\neq y]))$	1 Df. P
(1)	3. $0=\#(A)\wedge Aa\wedge x=\#([\lambda z\, Az\wedge z\neq a])$	2 EI×2
(1)	4. $0=\#(A)$	3 Simp
(1)	5. $\neg(\exists x)Ax$	1, (Zero) QL
(1)	6. $(\exists x)Ax$	3 Simp, EG
	7. $(\forall x)\neg x\ P\ 0$	1, 6, 7 RAA, UG

($\cong$–) $\vdash_{\text{FA}} (\forall F)(\forall G)(\forall x)(\forall y)(Fx\wedge Gy\wedge F\cong G\to$ $[\lambda w\, Fw\wedge w\neq x]\cong[\lambda w\, Gw\wedge w\neq y])$

Proof sketch: Assume $Fx\wedge Gy\wedge F\cong G$. We then have a relation R that constitutes a 1–1 correspondence between F and G. There is some unique a such that Rxa, and some unique b such that Rby. The relation $[\lambda wv(Rwv\vee(w=b\wedge v=a))\wedge w\neq x\wedge v\neq y]$ is a 1–1 correspondence between $[\lambda w\, Fw\wedge w\neq x]$ and $[\lambda w\, Gw\wedge w\neq y]$. To see this, notice that there are two cases to consider. Either $(x=b\wedge y=a)$ or $x\neq b\wedge y\neq a)$. In the first case, then the relation thus defined is precisely the same as R except not holding between x and y, so all Fs other than x are mapped to one and only one G, as before. If $(x\neq b\wedge y\neq a)$, then the relation defined above is similar to R except that it does not hold between x and anything, or between anything and y, but it does hold between b and a. Adding these as relata preserves its 1–1 status, since a is the only thing to which b is now related, and b is the only thing that it relates to a.

($\cong$+) $\vdash_{\text{FA}} (\forall F)(\forall G)(\forall x)(\forall y)(Fx\wedge Gy\wedge$ $[\lambda w\, Fw\wedge w\neq x]\cong[\lambda w\, Gw\wedge w\neq y]\to F\cong G)$

HOMEWORK 21

Briefly sketch a proof of ($\cong$+).

(PP4) $\vdash_{\text{FA}} (\forall x)(\forall y)(\forall z)(y\ P\ x\wedge z\ P\ x\to y=z)$

Proof:

(1)	1. $y\ P\ x\wedge z\ P\ x$	Hyp
(1)	2. $(\exists F)(\exists z)(x=\#(F)\wedge Fz\wedge$ $y=\#([\lambda w\, Fw\wedge w\neq z]))$	1 Simp, Df. P
(1)	3. $(\exists F)(\exists y)(x=\#(F)\wedge Fy\wedge$ $z=\#([\lambda w\, Fw\wedge w\neq y]))$	1 Simp, Df. P
(1)	4. $x=\#(A)\wedge Aa\wedge y=\#([\lambda w\, Aw\wedge w\neq a])$	2 EI×2
(1)	5. $x=\#(B)\wedge Bb\wedge z=\#([\lambda w\, Bw\wedge w\neq b])$	3 EI×2
(1)	6. $\#(A)=\#(B)$	4, 5 Simp, LL
(1)	7. $A\cong B$	6, (HL) QL
(1)	8. $[\lambda w\, Aw\wedge w\neq a]\cong[\lambda w\, Bw\wedge w\neq b]$	4, 5, 7, ($\cong$–) QL
(1)	9. $\#([\lambda w\, Aw\wedge w\neq a])=\#([\lambda w\, Bw\wedge w\neq b])$	8, (HL) QL
(1)	10. $y=\#([\lambda w\, Bw\wedge w\neq b])$	4, 9 Simp, LL
(1)	11. $y=z$	5, 10 Simp, LL
	12. $(\forall x)(\forall y)(\forall z)(y\ P\ x\wedge z\ P\ x\to y=z)$	1–11 CP, UG×3

(PP2a) $\vdash_{\text{FA}} (\forall x)(\forall y)(\forall z)(x\ P\ y\wedge x\ P\ z\to y=z)$

Proof: Very similar to proof of (PP4), but with ($\cong$+) instead of ($\cong$–).

(PP2b) $\vdash_{\text{FA}} (\forall x)(\forall y)(N(x)\wedge x\ P\ y\to N(y))$

Proof:

(1)	1. $N(x) \wedge x\ P\ y$	Hyp
(2)	2. $F(0) \wedge (\forall x)(\forall z)(Fx \wedge x\ P\ z \rightarrow Fz)$	Hyp
(1)	3. $(\forall F)(F(0) \wedge (\forall y)(\forall z)(Fy \wedge y\ P\ z \rightarrow Fz) \rightarrow Fx)$	1 Simp, Df. N
(2)	4. $(\forall y)(\forall z)(Fy \wedge y\ P\ z \rightarrow Fz)$	2 Simp, QL
(1,2)	5. Fx	2, 3, 4 QL
(1)	6. $Fx \wedge x\ P\ y \rightarrow Fy$	2 QL
(1,2)	7. Fy	1, 5, 6 SL
(1)	8. $F(0) \wedge (\forall x)(\forall z)(Fx \wedge x\ P\ z \rightarrow Fz) \rightarrow Fy$	2–7 CP
(1)	9. $N(y)$	8 UG, Df. N
	10. $(\forall x)(\forall y)(N(x) \wedge x\ P\ y \rightarrow N(y))$	1–9 CP, UG×2

(PP5a) $\vdash_{\text{FA}} \mathscr{A}[0] \wedge (\forall x)(\forall y)(\mathscr{A}[x] \wedge x\ P\ y \rightarrow \mathscr{A}[y]) \rightarrow (\forall x)(N(x) \rightarrow \mathscr{A}[x])$

Proof:

(1)	1. $\mathscr{A}[0] \wedge (\forall x)(\forall y)(\mathscr{A}[x] \wedge x\ P\ y \rightarrow \mathscr{A}[y])$	Hyp
(2)	2. $N(x)$	Hyp
(2)	3. $(\forall F)(F(0) \wedge (\forall y)(\forall z)(Fy \wedge y\ P\ z \rightarrow Fz) \rightarrow Fx)$	2 Df. N
(2)	4. $[\lambda w\, \mathscr{A}[w]](0) \wedge (\forall y)(\forall z)([\lambda w\, \mathscr{A}[w]](y) \wedge y\ P\ z \rightarrow [\lambda w\, \mathscr{A}[w]](z)) \rightarrow [\lambda w\, \mathscr{A}[w]](x)$	3 UI
(1)	5. $\mathscr{A}[0]$	1 Simp
(1)	6. $[\lambda w\, \mathscr{A}[w]](0)$	5 λ-conv
(1)	7. $\mathscr{A}[y] \wedge y\ P\ z \rightarrow \mathscr{A}[z]$	1 Simp, UI
(8)	8. $[\lambda w\, \mathscr{A}[w]](y) \wedge y\ P\ z$	Hyp
(8)	9. $\mathscr{A}[y] \wedge y\ P\ z$	8 SL, λ-conv
(1,8)	10. $\mathscr{A}[z]$	7, 9 MP
(1,8)	11. $[\lambda w\, \mathscr{A}[w]](z)$	10 λ-conv
(1)	12. $[\lambda w\, \mathscr{A}[w]](y) \wedge y\ P\ z \rightarrow [\lambda w\, \mathscr{A}[w]](z)$	8–11 CP
(1,2)	13. $[\lambda w\, \mathscr{A}[w]](x)$	4, 6, 12 QL
(1,2)	14. $\mathscr{A}[x]$	13 λ-conv
(1)	15. $(\forall x)(N(x) \rightarrow \mathscr{A}[x])$	2–14 CP, UG
	16. [1] → [15]	1–15 CP

(PP5) $\vdash_{\text{FA}} \mathscr{A}[0] \wedge (\forall x)(\forall y)(N(x) \wedge \mathscr{A}[x] \wedge x\ P\ y \rightarrow \mathscr{A}[y]) \rightarrow (\forall x)(N(x) \rightarrow \mathscr{A}[x])$

Proof:

(1)	1. $\mathscr{A}[0] \wedge (\forall x)(\forall y)(N(x) \wedge \mathscr{A}[x] \wedge x\ P\ y \rightarrow \mathscr{A}[y])$	Hyp
(1)	2. $N(0) \wedge \mathscr{A}[0]$	1, (PP1) SL
(3)	3. $N(x) \wedge \mathscr{A}[x] \wedge x\ P\ y$	Hyp
(1,3)	4. $\mathscr{A}[y]$	1, 3 QL
(3)	5. $N(x) \wedge x\ P\ y$	3 SL
(3)	6. $N(y)$	5, (PP2b) QL
(1,3)	7. $N(y) \wedge \mathscr{A}[y]$	4, 6 Conj
(1)	8. $(\forall x)(\forall y)(N(x) \wedge \mathscr{A}[x] \wedge x\ P\ y \rightarrow N(y) \wedge \mathscr{A}[y])$	3–7 CP, UG×2
(1)	9. $N(0) \wedge \mathscr{A}[0] \wedge (\forall x)(\forall y)(N(x) \wedge \mathscr{A}[x] \wedge x\ P\ y \rightarrow N(y) \wedge \mathscr{A}[y]) \rightarrow (\forall x)(N(x) \rightarrow N(x) \wedge \mathscr{A}[x])$	(PP5a)
(1)	10. $(\forall x)(N(x) \rightarrow N(x) \wedge \mathscr{A}[x])$	2, 8, 9 SL
(1)	11. $(\forall x)(N(x) \rightarrow \mathscr{A}[x])$	10 QL
	12. $\mathscr{A}[0] \wedge (\forall x)(\forall y)(N(x) \wedge \mathscr{A}[x] \wedge x\ P\ y \rightarrow \mathscr{A}[y]) \rightarrow (\forall x)(N(x) \rightarrow \mathscr{A}[x])$	1–11 CP

So far we have proven four of the five Peano postulates. In English, the second Peano postulate states that *every natural number has a unique successor which is also a natural number*.

So far we have proven two lemmas, (PP2a) and (PP2b), which state that no natural number has more than one successor, and that any successor of a natural number is itself a natural number. What we have not yet established is that every natural number *has* a successor.

Our strategy for showing this will be to show that the collection of elements in the natural number series leading up to and including a given natural number n always has the successor of n as its number. I.e., there is 1 natural number up to an including 0, 2 natural numbers up to and including 1, 3 natural numbers up to an including 2, and so on. We shall prove this for all natural numbers inductively, starting with 0.

Definitions:

$t <_{\text{N}} u$	for	$(\forall f)((\forall \chi)(\forall y)(\chi\ P\ y \wedge (\chi = t \vee f(\chi)) \rightarrow f(y)) \rightarrow f(u))$
$t \leq_{\text{N}} u$	for	$t <_{\text{N}} u \vee t = u$
$t \not<_{\text{N}} u$	for	$\neg t <_{\text{N}} u$
$@(t)$	for	$t\ P\ \#([\lambda z\, z \leq_{\text{N}} t])$

($@(t)$ says that there are $t+1$ natural numbers up to and including t.)

($\not<0$) $\vdash_{\mathrm{FA}} (\forall x)\, x \not<_{\mathrm{N}} 0$

Proof:

(1)	1. $x <_{\mathrm{N}} 0$	Hyp
(1)	2. $(\forall F)((\forall y)(\forall z)(y\ P\ z \land (y = x \lor Fy) \to Fz) \to F(0))$	1 Df. $<_{\mathrm{N}}$
(1)	3. $(\forall y)(\forall z)(y\ P\ z \land (y = x \lor [\lambda w\, w \neq 0](y)) \to$ $[\lambda w\, w \neq 0](z)) \to [\lambda w\, w \neq 0](0)$	2 UI
(4)	4. $y\ P\ z \land (y = x \lor [\lambda w\, w \neq 0](y))$	Hyp
	5. $\neg y\ P\ 0$	(PP3) UI
(4)	6. $z \neq 0$	4, 5, (LL) QL
(4)	7. $[\lambda w\, w \neq 0](z)$	6 λ-conv
	8. $(\forall y)(\forall z)(y\ P\ z \land (y = x \lor [\lambda w\, w \neq 0](y)) \to$ $[\lambda w\, w \neq 0](z))$	4–7 CP, UG×2
(1)	9. $[\lambda w\, w \neq 0](0)$	3, 8 MP
(1)	10. $0 \neq 0$	9 λ-conv
	11. $0 = 0$	Ref=
	12. $(\forall x)\, x \not<_{\mathrm{N}} 0$	1, 10, 11 RAA, UG

($P<$) $\vdash_{\mathrm{FA}} (\forall x)(\forall y)(x\ P\ y \to x <_{\mathrm{N}} y)$

Proof:

(1)	1. $x\ P\ y$	Hyp
(2)	2. $(\forall z)(\forall w)(z\ P\ w \land (z = x \lor Fz) \to Fw)$	Hyp
(2)	3. $x\ P\ y \land (x = x \lor Fx) \to Fy$	2 UI×2
	4. $x = x$	Ref=
	5. $x = x \lor Fx$	4 Add
(1,2)	6. Fy	1, 3, 5 SL
(1)	7. $(\forall F)((\forall z)(\forall w)(z\ P\ w \land (z = x \lor Fz) \to Fw) \to Fy)$	2–7 CP, UG
(1)	8. $x <_{\mathrm{N}} y$	7 Df. $<_{\mathrm{N}}$
	9. $(\forall x)(\forall y)(x\ P\ y \to x <_{\mathrm{N}} y)$	1–8 CP, UG×2

(Trans$<$) $\vdash_{\mathrm{FA}} (\forall x)(\forall y)(\forall z)(x <_{\mathrm{N}} y \land y <_{\mathrm{N}} z \to x <_{\mathrm{N}} z)$

Proof:

(1)	1. $x <_{\mathrm{N}} y \wedge y <_{\mathrm{N}} z$	Hyp
(1)	2. $(\forall F)((\forall z)(\forall w)(z\ P\ w \wedge (z = x \vee Fz) \to Fw) \to Fy)$	1 Simp, Df. $<_{\mathrm{N}}$
(1)	3. $(\forall F)((\forall x)(\forall w)(x\ P\ w \wedge (x = y \vee Fx) \to Fw) \to Fz)$	1 Simp, Df. $<_{\mathrm{N}}$
(4)	4. $(\forall y)(\forall w)(y\ P\ w \wedge (y = x \vee Fy) \to Fw)$	Hyp
(1)	5. $(\forall z)(\forall w)(z\ P\ w \wedge (z = x \vee Fz) \to Fw) \to Fy$	2 UI
(4)	6. $(\forall z)(\forall w)(z\ P\ w \wedge (z = x \vee Fz) \to Fw)$	4 QL
(1,4)	7. Fy	5, 6 MP
(1)	8. $(\forall x)(\forall w)(x\ P\ w \wedge (x = y \vee Fx) \to Fw) \to Fz$	3 UI
(9)	9. $v\ P\ w \wedge (v = y \vee Fv)$	Hyp
(9)	10. $v\ P\ w$	9 Simp
(9)	11. $v = y \vee Fv$	9 Simp
(1,4)	12. $v = y \to Fv$	7, (LL) QL
(9)	13. $v \neq y \to Fv$	11 SL
(1,4,9)	14. Fv	12, 13 SL
(1,4,9)	15. $v = x \vee Fv$	14 Add
(1,4,9)	16. $v\ P\ w \wedge (v = x \vee Fv)$	10, 15 Conj
(1,4,9)	17. Fw	6, 16 QL
(1,4)	18. $v\ P\ w \wedge (v = y \vee Fv) \to Fw$	9–17 CP
(1,4)	19. $(\forall x)(\forall w)(x\ P\ w \wedge (x = y \vee Fx) \to Fw)$	18 QL
(1,4)	20. Fz	8, 19 MP
(1)	21. $(\forall y)(\forall w)(y\ P\ w \wedge (y = x \vee Fy) \to Fw) \to Fz$	4–20 CP
(1)	22. $x <_{\mathrm{N}} z$	21 UG, Df. $<_{\mathrm{N}}$
	23. $(\forall x)(\forall y)(\forall z)(x <_{\mathrm{N}} y \wedge y <_{\mathrm{N}} z \to x <_{\mathrm{N}} z)$	1–22 CP, UG×3

(Ref≤) $\vdash_{\mathrm{FA}} (\forall x)\, x \leq_{\mathrm{N}} x$

Proof: obvious result of (Ref=).

$(<P)$ $(\forall x)(\forall y)(x <_{\mathrm{N}} y \to (\exists u)(x \leq_{\mathrm{N}} u \wedge u\ P\ y))$

Proof:

(1)	1. $x <_{\mathrm{N}} y$	Hyp
(1)	2. $(\forall F)((\forall z)(\forall w)(z\ P\ w \wedge (z = x \vee Fz) \to Fw) \to Fy)$	1 Df. $<_{\mathrm{N}}$
(1)	3. $(\forall z)(\forall w)(z\ P\ w \wedge (z = x \vee [\lambda v\, (\exists u)(x \leq_{\mathrm{N}} u\ \&\ u\ P\ v)](z)) \to [\lambda v\, (\exists u)(x \leq_{\mathrm{N}} u\ \&\ u\ P\ v)](w)) \to [\lambda v\, (\exists u)(x \leq_{\mathrm{N}} u\ \&\ u\ P\ v)](y)$	2 UI
(4)	4. $z\ P\ w \wedge (z = x \vee [\lambda v\, (\exists u)(x \leq_{\mathrm{N}} u\ \&\ u\ P\ v)](z))$	Hyp
(4)	5. $z\ P\ w$	4 Simp
(4)	6. $z = x \vee [\lambda v\, (\exists u)(x \leq_{\mathrm{N}} u\ \&\ u\ P\ v)](z)$	4 Simp
(7)	7. $z = x$	Hyp
	8. $x \leq_{\mathrm{N}} x$	Ref≤ UI

(7)	9. $x \leq_{\mathrm{N}} z$	7, 8 LL
(4,7)	10. $(\exists u)(x \leq_{\mathrm{N}} u \wedge u\ P\ w)$	5, 9 QL
(4)	11. $z = x \rightarrow (\exists u)(x \leq_{\mathrm{N}} u \wedge u\ P\ w)$	7–10 CP
(12)	12. $[\lambda v\,(\exists u)(x \leq_{\mathrm{N}} u \wedge u\ P\ v](z)$	Hyp
(12)	13. $(\exists u)(x \leq_{\mathrm{N}} u \wedge u\ P\ z)$	12 λ-conv
(12)	14. $x \leq_{\mathrm{N}} a \wedge a\ P\ z$	13 EI
(12)	15. $a\ P\ z$	14 Simp
(12)	16. $a <_{\mathrm{N}} z$	15, $(P<)$ QL
(12)	17. $x <_{\mathrm{N}} a \vee x = a$	14 Simp, Df. $\leq_{\mathrm{N}}$
(12)	18. $x = a \rightarrow x <_{\mathrm{N}} z$	16, (LL) QL
	19. $x <_{\mathrm{N}} a \wedge a <_{\mathrm{N}} z \rightarrow x <_{\mathrm{N}} z$	(Trans<) QL
(12)	20. $x <_{\mathrm{N}} a \rightarrow x <_{\mathrm{N}} z$	16, 19 SL
(12)	21. $x <_{\mathrm{N}} z$	17, 18, 20 CD
(12)	22. $x \leq_{\mathrm{N}} z$	21 Add, Df. $\leq_{\mathrm{N}}$
(4,12)	23. $(\exists u)(x \leq_{\mathrm{N}} u \wedge u\ P\ w)$	5, 22 QL
(4)	24. $[\lambda v\,(\exists u)(x \leq_{\mathrm{N}} y \wedge u\ P\ v)](z) \rightarrow (\exists u)(x \leq_{\mathrm{N}} u \wedge u\ P\ w)$	12–23 CP
(4)	25. $(\exists u)(x \leq_{\mathrm{N}} u \wedge u\ P\ w)$	6, 11, 24 CD
(4)	26. $[\lambda v\,(\exists u)(x \leq_{\mathrm{N}} u \wedge u\ P\ v)](w)$	25 λ-conv
	27. $(\forall x)\,(\forall w)(z\ P\ w \wedge (z = x \vee [\lambda v\,(\exists u)(x \leq_{\mathrm{N}} u\ \ \&\ \ u\ P\ v)](z)) \rightarrow [\lambda v\,(\exists u)(x \leq_{\mathrm{N}} u \wedge u\ P\ v)](w))$	4–26 CP, UG×2
(1)	28. $[\lambda v\,(\exists u)(x \leq_{\mathrm{N}} u \wedge u\ P\ v)](y)$	3, 27 MP
(1)	29. $(\exists u)(x \leq_{\mathrm{N}} u \wedge u\ P\ y)$	28 λ-conv
	30. $(\forall x)\,(\forall y)(x <_{\mathrm{N}} y \rightarrow (\exists u)(x \leq_{\mathrm{N}} u \wedge u\ P\ y))$	1–29 CP, UG×2

(Irref<) $\vdash_{\mathrm{FA}} (\forall x)(N(x) \rightarrow x \not<_{\mathrm{N}} x)$

Proof:

	1. $0 \not<_{\mathrm{N}} 0$	$(\not<0)$, UI
(2)	2. $x \not<_{\mathrm{N}} x \wedge x\ P\ y$	Hyp
(3)	3. $y <_{\mathrm{N}} y$	Hyp
(3)	4. $(\exists u)(y \leq_{\mathrm{N}} u \wedge u\ P\ y)$	3, $(<P)$ QL
(3)	5. $y \leq_{\mathrm{N}} b \wedge b\ P\ y$	4 EI
(2,3)	6. $x = b$	2, 5, (PP4) QL
(2,3)	7. $y \leq_{\mathrm{N}} x$	5, 6 Simp, LL
(2,3)	8. $y <_{\mathrm{N}} x \vee y = x$	7 Df. $\leq_{\mathrm{N}}$
(3)	9. $y = x \rightarrow x <_{\mathrm{N}} x$	3, (LL) QL
(2,3)	10. $y \neq x$	2, 9 Simp, MT
(2,3)	11. $y <_{\mathrm{N}} x$	8, 10 DS
(2)	12. $x <_{\mathrm{N}} y$	2, $(P<)$ QL

(2,3)	13. $x <_{\mathrm{N}} x$	11, 12, (Trans<) QL
(2)	14. $x \not<_{\mathrm{N}} x$	2 Simp
(2)	15. $y \not<_{\mathrm{N}} y$	3, 13, 14 RAA
	16. $(\forall x)(\forall y)(x \not<_{\mathrm{N}} x \wedge x\ P\ y \to y \not<_{\mathrm{N}} y)$	2–15 CP, UG×2
	17. $(\forall x)(N(x) \to x \not<_{\mathrm{N}} x)$	1, 16, (PP5a) SL

(<Succ≤) $\vdash_{\mathrm{FA}} (\forall x)(\forall y)(x\ P\ y \to (\forall z)(z <_{\mathrm{N}} y \leftrightarrow z \leq_{\mathrm{N}} x))$

HOMEWORK 22

Prove (<Succ≤). Hint: use (P<), (<P), (PP4) and (Trans<).

(≤0) $\vdash_{\mathrm{FA}} (\forall x)(x \leq_{\mathrm{N}} 0 \leftrightarrow x = 0)$

Proof: easy result of (Ref≤), (≮0) and Df. $\leq_{\mathrm{N}}$.

(@Base) $\vdash_{\mathrm{FA}} @(0)$

Proof uses (≤0), (Ext≅) and (HL).

(@Ind) $\vdash_{\mathrm{FA}} (\forall x)(\forall y)(N(x) \wedge x\ P\ y \wedge @(x) \to @(y))$

Proof:

(1)	1. $N(x) \wedge x\ P\ y \wedge @(x)$	Hyp
(1)	2. $x\ P\ \#([\lambda v\, v \leq_{\mathrm{N}} x])$	1 Simp, Df. @
(1)	3. $(\forall z)(z <_{\mathrm{N}} y \leftrightarrow z \leq_{\mathrm{N}} x)$	1, (<Succ≤) QL
(4)	4. $[\lambda v\, v \leq_{\mathrm{N}} x](w)$	Hyp
(4)	5. $w \leq_{\mathrm{N}} x$	4 λ-conv
(1,4)	6. $w <_{\mathrm{N}} y$	3, 5 QL
(1)	7. $N(y)$	1, (PP2b) QL
(1)	8. $y \not<_{\mathrm{N}} y$	7, (Irref<) QL
(1,4)	9. $w = y \to y <_{\mathrm{N}} y$	6, (LL) QL
(1,4)	10. $w \neq y$	8, 9 MT
(1,4)	11. $w \leq_{\mathrm{N}} y$	6 Add, Df. $\leq_{\mathrm{N}}$
(1,4)	12. $[\lambda v\, v \leq_{\mathrm{N}} y](w)$	11 λ-conv
(1,4)	13. $[\lambda v\, v \leq_{\mathrm{N}} y](w) \wedge w \neq y$	10, 12 Conj
(1,4)	14. $[\lambda z[\lambda v\, v \leq_{\mathrm{N}} y](z) \wedge z \neq y](w)$	13 λ-conv
(1)	15. $[\lambda v\, v \leq_{\mathrm{N}} x](w) \to [\lambda z[\lambda v\, v \leq_{\mathrm{N}} y](z) \wedge z \neq y](w)$	4–14 CP
(16)	16. $[\lambda z[\lambda v\, v \leq_{\mathrm{N}} y](z) \wedge z \neq y](w)$	Hyp
(16)	17. $w \leq_{\mathrm{N}} y \wedge w \neq y$	16 λ-conv, SL
(16)	18. $w <_{\mathrm{N}} y$	17 Df. $\leq_{\mathrm{N}}$, SL

(1,16)	19. $w \leq_{\mathrm{N}} x$	3, 18 QL
(1,16)	20. $[\lambda v\, v \leq_{\mathrm{N}} x](w)$	19 λ-conv
(1)	21. $[\lambda z[\lambda v\, v \leq_{\mathrm{N}} y](z) \wedge z \neq y](w) \rightarrow [\lambda v\, v \leq_{\mathrm{N}} x](w)$	16–20 CP
(1)	22. $(\forall w)([\lambda v\, v \leq_{\mathrm{N}} x](w) \leftrightarrow [\lambda z[\lambda v\, v \leq_{\mathrm{N}} y](z) \wedge z \neq y](w))$	15, 21 BI, UG
(1)	23. $[\lambda v\, v \leq_{\mathrm{N}} x] \cong [\lambda z[\lambda v\, v \leq_{\mathrm{N}} y](z) \wedge z \neq y]$	22, (Ext$\cong$) QL
(1)	24. $\#([\lambda v\, v \leq_{\mathrm{N}} x]) = \#([\lambda z[\lambda v\, v \leq_{\mathrm{N}} y](z) \wedge z \neq y])$	23, (HL) QL
(1)	25. $x\ P\ \#([\lambda z[\lambda v\, v \leq_{\mathrm{N}} y](z) \wedge z \neq y])$	2, 24 LL
(1)	26. $y = \#([\lambda z[\lambda v\, v \leq_{\mathrm{N}} y](z) \wedge z \neq y])$	1, 25, (PP2a) QL
	27. $\#([\lambda v\, v \leq_{\mathrm{N}} y]) = \#([\lambda v\, v \leq_{\mathrm{N}} y])$	Ref=
	28. $[\lambda v\, v \leq_{\mathrm{N}} y](y)$	(Ref$\leq$) UI, λ-conv
(1)	29. $\#([\lambda v\, v \leq_{\mathrm{N}} y]) = \#([\lambda v\, v \leq_{\mathrm{N}} y]) \wedge [\lambda v\, v \leq_{\mathrm{N}} y](y) \wedge y = \#([\lambda z[\lambda v\, v \leq_{\mathrm{N}} y](z) \wedge z \neq y])$	26, 27, 28 Conj
(1)	30. $(\exists F)(\exists x)(\#([\lambda v\, v \leq_{\mathrm{N}} y]) = \#(F) \wedge Fx \wedge y = \#([\lambda z\, Fz \wedge z \neq x]))$	39 EG×2
(1)	31. $y\ P\ \#([\lambda v\, v \leq_{\mathrm{N}} y])$	30 Df. P
(1)	32. $@(y)$	31 Df. @
	33. $(\forall x)(\forall y)(N(x) \wedge x\ P\ y \wedge @(x) \rightarrow @(y))$	1–32 CP, UG×2

(PP2c) $\vdash_{\mathrm{FA}} (\forall x)(N(x) \rightarrow @(x))$

Proof: direct from (@Base), (@Ind) and (PP5).

Finally we establish:

(PP2) $\vdash_{\mathrm{FA}} (\forall x)(N(x) \rightarrow (\exists! y)(x\ P\ y \wedge N(y)))$

Proof sketch: Suppose $N(x)$. By (PP2c), x has a successor, viz., $\#([\lambda v\, v \leq_{\mathrm{N}} x])$. By (PP2b), this successor is a natural number. By (PP2a), this successor is unique.

This establishes Frege's theorem.

E. Philosophical Reflections on Frege's Theorem

The fact that all of Peano arithmetic (and hence almost all of ordinary mathematics) is derivable in second-order logic with only a single premise is very substantial. It naturally prompts questions regarding the nature of (HL): is it plausible? What is its metaphysical and epistemological status? Is it metaphysically innocent? Can it be regarded as a logical or at least analytical truth?

In many ways (HL) seems *constitutive* of our concept of number: an analytic truth regarding what we mean by "the number of . . . ". It seems in some ways to explicate what we mean by a number: a number is a *thing* that equinumerous collections have in common. And since it explicates numbers in terms of the relation $\cong$, itself defined only using logical constants and quantifiers, it adds support for thinking that there is some sense in which arithmetic reduces to logic. If (HL) can be known *a priori*, this helps firm up the epistemology of mathematics.

Principles such as (HL) are sometimes called definitions by ***abstraction***: a number is regarded as the *common thing* all equinumerous collections share, abstracting away their differences.

Frege himself *rejected* taking (HL) as a basic principle, since it does not give an outright *definition* of terms of the form "$\#(F)$" but only fixes the truth conditions for identity statements formed with two terms of this form. It does not tell us what numbers *themselves* are, and does not tell us, e.g., whether or not Julius Caesar is a number. This is in part why Frege thought it necessary to *define* $\#(\ )$ in terms of his theory of extensions, and obtain (HL) as a theorem starting with (BLV).

There are other objections to definitions in abstraction in general. (BLV) is a kind of definition by abstraction, just replacing equinumerosity with coextensionality. Similar principles, such as the supposition that order-type(R) = order-type(S) iff R and S are isomorphic relations, is also inconsistent. This casts suspicion on taking such definitions at face value.

On the other hand, (HL) is known to be consistent (relative to PA2).

XXIV. Boolos's (New V) and Others

Hume's law, Basic Law V, and other "definitions by abstraction" in second-order logic take the form:

$$(\forall F)(\forall G)(\%(F) = \%(G) \longleftrightarrow F \text{ eq } G)$$

where "%( )" is some operator (like "Ext( )" or "#( )"), which when applied to a predicate variable or λ-abstract, forms a term, and "eq" is short for some definable *equivalence relation* between F and G: i.e., a reflexive, symmetric and transitive relation between properties (e.g., such as being *coextensive* or *equinumerous*).

One of the nice things about Basic Law V (had it been consistent) is that it made any *further* axioms of this form unnecessary. "$\%(F)$" could be defined as $\text{Ext}([\lambda x\, (\exists G)(x = \text{Ext}(G) \land F \text{ eq } G)])$, and then above abstraction principle would simply follow by Basic Law V, and the definition of %( ), just as Frege supposed for defining numbers.

Of course, when combined with unrestricted property comprehension (CP), Law V leads to contradictions, as do other "abstraction principles" of roughly the same form as the above, e.g.:

Definition: $t \approx u$ (t is isomorphic to u) for

$$\begin{aligned}(\exists r)((&\forall x)(\forall y)(\forall z)(r(x,y) \land r(x,z) \to y = z) \land\\ &(\forall x)(\forall y)(\forall z)(r(y,x) \land r(z,x) \to y = z) \land\\ &(\forall x)(\forall y)(t(x,y) \to (\exists z)(\exists w)(r(x,z) \land r(y,w) \land u(z,w))) \land\\ &(\forall x)(\forall y)(u(x,y) \to (\exists z)(\exists w)(r(z,x) \land r(w,y) \land t(z,w))))\end{aligned}$$

Consider the abstraction principle:

$$(\forall R)(\forall S)(\S(R) = \S(S) \longleftrightarrow R \approx S)$$

This abstraction principle leads to the Burali-Forti paradox.

Nevertheless, one might be tempted to look for abstraction principles with a wide array of uses, like Basic Law V, but without leading to a contradiction. One such principle that has received significant attention is George Boolos's "New V". It begins with certain definitions:

Definitions:

Big(f)	for	$(\exists g)((\forall x)(g(x) \to f(x)) \land g \cong [\lambda x\, x = x])$
Small(f)	for	$\neg\text{Big}(f)$

A "small" concept is one that applies to fewer things than there are things. A "big" concept has as many things as there are things. (Note that a concept does not have to appy to everything to be "big": it suffices for it to apply to an infinite subset of things that can be put in 1–1 correspondence with everything.)

Definition:

$f \mathrel{\hat{=}} g$ for $(\text{Small}(f) \lor \text{Small}(g)) \to (\forall x)(f(x) \longleftrightarrow g(x))$
(or equivalently, $(\text{Big}(f) \land \text{Big}(g)) \lor (\forall x)(f(x) \longleftrightarrow g(x))$.)

Boolos then suggests taking the following as an axiom:

(New V) $(\forall F)(\forall G)(*(F) = *(G) \longleftrightarrow F \mathrel{\hat{=}} G)$

"$*(F)$" is read "the subtension of F". (New V) holds that "small" concepts have the same subtension when and only when they are coextensive, and that all "big" concepts have the same subtension. Boolos thinks this represents a natural way to represent the "limitation of size approach" to the philosophy of sets. Indeed, with the definitions:

Definitions:

$t \in u$ for $(\exists f)(u = *(f) \wedge f(t))$
$\{x|\mathscr{A}[x]\}$ for $*([\lambda x\, \mathscr{A}[x]])$

We get:

$$\vdash_{\text{FN}} \text{Small}([\lambda x\, \mathscr{A}[x]]) \rightarrow (\forall y)(y \in \{x|\mathscr{A}[x]\} \leftrightarrow \mathscr{A}[y])$$

HOMEWORK 23

Prove the theorem above.

It turns that FN—the system that adds only (New V) to the axioms of second-order logic—is strong enough of a set theory to obtain the Peano postulates (though it requires using the von Neumann definition of numbers rather than the Frege-Russell definition: notice that property of being a singleton is "big", and so the Frege-Russell 1 is the subtension of a "big" concept.)

(New V) is consistent. To see this we need only consider a standard model whose domain of quantification is the set of natural numbers. "Small" concepts would those that apply to finitely many. There are well-known ways of "coding" finite sets of numbers using numbers. E.g., representing $\{n_1, n_2, \ldots, n_m\}$ by raising the first m primes to the powers of $n_1, n_2, \ldots, n_m$ respectively and multiplying them together. Take $*(F)$ to be the result of this procedure applied to set of natural numbers satisfying $F(x)$ if finitely many satisfy it, to be 0 if none do, and to be 5 otherwise. (New V) is true on this model.

Notice that (New V) can be seen as taking the form:

$$(\forall F)(\forall G)(*(F) = *(G) \leftrightarrow (\text{Bad}(F) \wedge \text{Bad}(G)) \vee (\forall x)(Fx \leftrightarrow Gx))$$

It is currently in vogue to explore definitions of "badness" (other than bigness) to use here corresponding to different philosophical positions on what sorts of properties can be seen as defining sets. E.g., one might define "Bad(f)" in terms of Dummett's notion of *indefinite extensibility*.

XXV. Paraconsistent Foundations

A. Paraconsistent Logic

In classical bivalent truth-functional logic, any instance of the following schema is a tautology:

$$\mathscr{A} \wedge \neg\mathscr{A} \rightarrow \mathscr{B}$$

Because the $\mathscr{A}$ and $\mathscr{B}$ could be anything, this means that any contradiction would *explode*, leading to every wff (and its negation) being a theorem.

A ***paraconsistent logic*** in a logic in which the inference $\mathscr{A}, \neg\mathscr{A} \vdash \mathscr{B}$ is not allowed (or the corresponding semantic claim, $\mathscr{A}, \neg\mathscr{A} \vDash \mathscr{B}$ not considered valid).

Paraconsistent logic must at least in some sense be *weaker* than classical logic, disallowing certain inferences allowed by classical logic. Consider the standard "proof" of an arbitrary result $\mathscr{B}$ from $\mathscr{A}$ and $\neg\mathscr{A}$:

1. $\mathscr{A}$	
2. $\neg\mathscr{A}$	
3. $\mathscr{A} \vee \mathscr{B}$	1 Add
4. $\mathscr{B}$	2, 3 DS

Hence, any paraconsistent logic must block either the rule of Addition or the rule of Disjunctive Syllogism (or, I suppose, both). Most forms of paraconsistent logic discard the validity of Disjunctive Syllogism.

Consider Priest's system LP (the Logic of Paradox). Every sentence of the language can be true, false, or *both*. A negation is true just in case what it negates is false, and is false just in case what it negates is true. Hence, negations of statements that are both are also both. A disjunction is true if either side is true, and false if both sides are false; hence a disjunction may be both if both sides are both, or if one side is false and the other is both.

$\mathscr{A}$	$\neg\mathscr{A}$
T	F
B	B
F	T

$\mathscr{A}$	$\mathscr{B}$	$(\mathscr{A} \vee \mathscr{B})$
T	T	T
T	B	T
T	F	T
B	T	T
B	B	B
B	F	B
F	T	T
F	B	B
F	F	F

We then say that a statement is a ***logical truth*** iff it is necessarily true or both true and false, and hence, impossibly false only. We say that an inference is invalid if it is possible for the premises to be true or both while the conclusion is false-only. Thus, we can see that disjunctive syllogism is invalid:

$(\mathscr{A}$	$\vee$	$\mathscr{B}),$	$\neg$	$\mathscr{A}$	$\mathscr{B}$	
T	T	T	F	T	T	
T	T	B	F	T	B	
T	T	F	F	T	F	
B	T	T	B	B	T	
B	B	B	B	B	B	
B	B	F	B	B	F	(**)
F	T	T	T	F	T	
F	B	B	T	F	B	
F	F	F	T	F	F	

At the line marked (**) we see that both premises are at least true, but the conclusion is merely false. Hence, DS must be rejected in LP.

It is also easy to see that the inference from $\mathscr{A}$ and $\neg\mathscr{A}$ to $\mathscr{B}$ in general must be rejected because of the possibility that $\mathscr{A}$ is both true and false, but $\mathscr{B}$ is false only.

B. Paraconsistent Set Theory: Weak or Strong?

One of the most intriguing possibilities that paraconsistent logic offers is that of accepting a (seemingly) naïve set theory, that is a set theory that accepts in full generality the unrestricted class abstraction principle from Hatcher's system F:

$$(\forall x)(x \in \{y|\mathscr{A}[y]\} \leftrightarrow \mathscr{A}[x])$$

(...along with either a definition of identity with a Leibniz's Law rule or axiom, or identity taken as primitive, with the system built on identity logic, with an extensionality axiom.)

Most likely such a theory would be inconsistent, in virtue of having a contradiction, such as the contradiction from Russell's paradox, as a theorem.

$$\{y|y \notin y\} \in \{y|y \notin y\} \wedge \{y|y \notin y\} \notin \{y|y \notin y\}$$

Such inconsistency, however, need not *trivialize* the system in the sense that every wff ends up as a theorem, or even that "undesired" set-theoretic or mathematical absurdities such as that $0 = 1$ or $\Lambda = V$ are provable.

Indeed, it is possible to prove that, e.g., FLP, the system just like system F from our first unit but with LP substituted for the core propositional logic, is non-trivial despite being inconsistent.

Unfortunately, FLP is likely too weak to be of much use as a foundational system for mathematics. To see this, notice that it misses some completely fundamental basic inferences. Indeed, it does not even allow *modus ponens*. Suppose we define $\mathscr{A} \supset \mathscr{B}$ as $\neg\mathscr{A} \vee \mathscr{B}$, as usual. (I here use $\supset$ rather than $\rightarrow$ for reasons I hope become clear later on.) In that case, the inference $\mathscr{A}, \mathscr{A} \supset \mathscr{B} \vdash \mathscr{B}$ is just the combination of DN and DS, and is invalid for the same reason that DS is invalid. If $\mathscr{A}$ is both true and false, then $\neg\mathscr{A}$ is also both true and false, and thus, $\neg\mathscr{A} \vee \mathscr{B}$, i.e., $\mathscr{A} \supset \mathscr{B}$, is true even when $\mathscr{B}$ is false-only.

To capture much of mathematics, a *detachable* conditional is needed. Hence, paraconsistent set theory usually adopts a stronger logical core than LP by taking as primitive a non-truth-functional ("relevant") conditional $\rightarrow$. This conditional is not taken as defined in terms of $\vee$, but as a distinct primitive connective. (We may, as before, define $\mathscr{A} \wedge \mathscr{B}$ as $\neg(\neg\mathscr{A} \vee \neg\mathscr{B})$; we define $\mathscr{A} \leftrightarrow \mathscr{B}$ as $(\mathscr{A} \rightarrow \mathscr{B}) \wedge (\mathscr{B} \rightarrow \mathscr{A})$, rather than, e.g., $(\mathscr{A} \supset \mathscr{B}) \wedge (\mathscr{B} \supset \mathscr{A})$ or $(\mathscr{A} \wedge \mathscr{B}) \vee (\neg\mathscr{A} \wedge \neg\mathscr{B})$, so it too is non-truth-functional.

But it turns out that the logic of this new conditional $\rightarrow$ must itself be rather weak to avoid triviality when a set theory is added. Assuming it allows MP, it cannot, for example, validate either of these intuitively (and classically) valid schemata:

$$\text{(Contraction)}\quad (\mathscr{A} \rightarrow (\mathscr{A} \rightarrow \mathscr{B})) \rightarrow (\mathscr{A} \rightarrow \mathscr{B})$$
$$\text{(MP-theorem)}\quad ((\mathscr{A} \rightarrow \mathscr{B}) \wedge \mathscr{A}) \rightarrow \mathscr{B}$$

The reason for these is that, when combined with set theory, a set-theoretic version of Curry's paradox threatens to lead to triviality. The usual form of Curry's paradox is an argument in favor of any arbitrary conclusion $\mathscr{B}$, e.g., "Santa Claus exists". Take the sentence "If this sentence is true, then $\mathscr{B}$." Assume it's true. To do so is to assume that if it is true, then $\mathscr{B}$. We're assuming it's true, so under that assumption, by modus ponens, $\mathscr{B}$. Discharging our assumption, if that sentence is true, $\mathscr{B}$. But that's what it says. Hence, it is true. By modus ponens on itself, $\mathscr{B}$.

The set theoretic version defines a set, M, of all those sets whose self-membership implies $\mathscr{B}$. That is, let M abbreviate $\{y|y \in y \rightarrow \mathscr{B}\}$. By naïve class abstraction:

$$M \in \{y|y \in y \rightarrow \mathscr{B}\} \leftrightarrow M \in M \rightarrow \mathscr{B}$$

In other words:

$$M \in M \leftrightarrow M \in M \rightarrow \mathscr{B}$$

Now, using (Contraction), we can prove $\mathscr{B}$:

1. $M \in M \leftrightarrow M \in M \rightarrow \mathscr{B}$	(see above)
2. $M \in M \rightarrow (M \in M \rightarrow \mathscr{B})$	1 BE
3. $(M \in M \rightarrow \mathscr{B}) \rightarrow M \in M$	1 BE
4. $(M \in M \rightarrow (M \in M \rightarrow \mathscr{B})) \rightarrow (M \in M \rightarrow \mathscr{B})$	(contraction)
5. $M \in M \rightarrow \mathscr{B}$	2, 4 MP
6. $M \in M$	3, 5 MP
7. $\mathscr{B}$	5, 6 MP

Hence, (Contraction) must be rejected for this conditional. So must (MP-theorem) via a similar argument, though the *rule* form of MP may be maintained.

Personally, this makes me question whether the abstraction principle of these systems deserves its moniker "naïve", since it disallows those instances that would make use of a resonably strong conditional, and hence one should not pretend as if it has no "restrictions", but this is a difficult matter.

C. An Example Paraconsistent Set Theory

As an example of a system that tries to thread the needle between "too strong" and "too weak", we examine the system of Weber (2010), built on the work of Brady and Sylvan (Routley). The system is built upon a core paraconsistent logic called TLQ. The system uses $\wedge$ and $\neg$ as primitive truth-functions, and defines $\vee$ using $\wedge$. (Nothing turns on this, however, and $\vee$ could have been taken as primitive instead.) It also adopts a relevant conditional $\rightarrow$ which is not defined in terms of other truth-functions; $\leftrightarrow$ is defined in terms of it, as above. Identity is taken as primitive; existential quantification is defined in terms of universal quantification in the normal way. The core logic has the following axiom schemata:

I $\mathscr{A} \rightarrow \mathscr{A}$
IIa $\mathscr{A} \wedge \mathscr{B} \rightarrow \mathscr{A}$
IIb $\mathscr{A} \wedge \mathscr{B} \rightarrow \mathscr{B}$
III $\mathscr{A} \wedge (\mathscr{B} \vee \mathscr{C}) \rightarrow (\mathscr{A} \wedge \mathscr{B}) \vee (\mathscr{A} \wedge \mathscr{C})$

IV $(\mathscr{A} \rightarrow \mathscr{B}) \wedge (\mathscr{B} \rightarrow \mathscr{C}) \rightarrow (\mathscr{A} \rightarrow \mathscr{C})$
V $(\mathscr{A} \rightarrow \mathscr{B}) \wedge (\mathscr{A} \rightarrow \mathscr{C}) \rightarrow (\mathscr{A} \rightarrow \mathscr{B} \wedge \mathscr{C})$
VI $(\mathscr{A} \rightarrow \neg\mathscr{B}) \rightarrow (\mathscr{B} \rightarrow \neg\mathscr{A})$
VII $\neg\neg\mathscr{A} \rightarrow \mathscr{A}$
VIII $(\mathscr{A} \rightarrow \mathscr{B}) \rightarrow \neg(\mathscr{A} \wedge \neg\mathscr{B})$
IXa $(\mathscr{A} \rightarrow \mathscr{B}) \rightarrow ((\mathscr{B} \rightarrow \mathscr{C}) \rightarrow (\mathscr{A} \rightarrow \mathscr{C}))$
IXb $(\mathscr{A} \rightarrow \mathscr{B}) \rightarrow ((\mathscr{C} \rightarrow \mathscr{A}) \rightarrow (\mathscr{C} \rightarrow \mathscr{B}))$
X $(\forall\chi)\mathscr{A}[\chi] \rightarrow \mathscr{A}[u]$, where no free variables of u become bound in $\mathscr{A}[u]$.
XI $(\forall\chi)(\mathscr{A} \rightarrow \mathscr{B}) \rightarrow (\mathscr{A} \rightarrow (\forall\chi)\mathscr{B})$, where $\mathscr{A}$ does not contain χ free.
XII $(\forall\chi)(\mathscr{A} \vee \mathscr{B}) \rightarrow (\mathscr{A} \vee (\forall\chi)\mathscr{B})$, where $\mathscr{A}$ does not contain χ free.

And we have the following primitive inference rules:

(Conj) From $\mathscr{A}$ and $\mathscr{B}$, infer $\mathscr{A} \wedge \mathscr{B}$.
(MP) From $\mathscr{A}$ and $\mathscr{A} \to \mathscr{B}$ infer $\mathscr{B}$.
(Cxt) From $\mathscr{A} \to \mathscr{B}$ and $\mathscr{C} \to \mathscr{D}$ infer $(\mathscr{B} \to \mathscr{C}) \to (\mathscr{A} \to \mathscr{D})$.
(UG) From $\mathscr{A}$ infer $(\forall \chi)\mathscr{A}$.
(Sub=) From $u = v$ infer $\mathscr{A}[u] \to \mathscr{A}[v]$.

Some interesting results:

(DN-in) $\vdash \mathscr{A} \to \neg\neg\mathscr{A}$
(EM) $\vdash \mathscr{A} \vee \neg\mathscr{A}$
(NC) $\vdash \neg(\mathscr{A} \wedge \neg\mathscr{A})$
(Inev) $\vdash (\mathscr{A} \to \neg\mathscr{A}) \to \neg\mathscr{A}$
$\vdash (\neg\mathscr{A} \to \mathscr{A}) \to \mathscr{A}$
(MT) $\mathscr{A} \to \mathscr{B}, \neg\mathscr{B} \vdash \neg\mathscr{A}$

(Note: it may be surprising that the theory has (NC). However, in this system all contradictions are false—it just so happens that some are true as well.)

To be able to adequately express "all *A*s are *B*s", the system is strengthened from the above by adding a constant proposition t which intuitively is taken to represent the conjunction of all theorems. Along with this, a rule:

(t) From $\mathscr{A}$ infer $t \to \mathscr{A}$.
From $t \to \mathscr{A}$ infer $\mathscr{A}$.

This allows us to define another conditional, $\mathscr{A} \mapsto \mathscr{B}$, as $\mathscr{A} \wedge t \to \mathscr{B}$, which is in between $\supset$ and $\to$ in strength. "All *A*s are *B*s" is then sometimes taken to mean $(\forall x)(\mathscr{A}[\chi] \mapsto \mathscr{B}[x])$.

We then add two axioms for set theory, including *extensionality*:

$$(\forall x)(\forall y)((\forall z)(z \in x \leftrightarrow z \in y) \leftrightarrow x = y)$$

And a naïve principle of class abstraction. Weber wants this in a form even stronger than what we found in System F. In particular, he wants to validate the schema:

$$(\exists x)(\forall y)(y \in x \leftrightarrow \mathscr{A}[y])$$

even when $\mathscr{A}[y]$ contains x free! This allows wildly circularly defined sets, such as one defined as having exactly its own non-members as members:

$$(\exists x)(\forall y)(y \in x \leftrightarrow y \notin x)$$

To capture circularly defined classes using a vbto, one can employ double-binding vbto $\{xy|\mathscr{A}[x,y]\}$, which means "the set identical to y whose members are all x such that ...." We then adopt the schema:

$$(\forall z)(z \in \{xy|\mathscr{A}[x,y]\} \leftrightarrow \mathscr{A}[z, \{xy|\mathscr{A}[x,y]\}])$$

To capture the silly set whose members are all its non-members, we then have the instance:

$$(\forall z)(z \in \{xy|x \notin y\} \leftrightarrow z \notin \{xy|x \notin y\})$$

We can abbreviate $\{xy|\mathscr{A}[x]\}$ when $\mathscr{A}[x]$ does not contain y free as $\{x|\mathscr{A}[x]\}$, and then as a simplification we have the usual naïve principle:

$$(\forall z)(z \in \{x|\mathscr{A}[x]\} \leftrightarrow \mathscr{A}[z])$$

Letting R abbreviate $\{x|x \notin x\}$ we have:

$$\vdash R \in R \leftrightarrow R \notin R$$

which, via (Inev), gives both $\vdash R \in R$ and $\vdash R \notin R$, so the system is inconsistent. Of course, this does not mean it is trivial.

Weird result: since there is something which is a member of R, but not a member of R, viz., R itself, R is non-coextensional with itself, hence by extensionality, $\vdash R \neq R$. (It also holds that $\vdash R = R$.)

This weird result means that it is undesirable to define the empty set Λ as $\{x|x \neq x\}$, since then the empty set would have members. So Weber uses the definition $\{x|(\forall y)x \in y\}$ instead. Similarly, V is defined not as $\{x|x = x\}$ but as $\{x|(\exists y)x \in y\}$.

Some points about this system

- It is not known whether or not the system is trivial, though the very similar theory obtained by replacing axiom VIII with the weaker (EM) is known to be non-trivial.

- The deduction theorem does not hold, so one cannot in general transform a proof of $\mathscr{B}$ from the assumption that $\mathscr{A}$ into a proof of $\mathscr{A} \rightarrow \mathscr{B}$.
- Weber proves that for every axiom of ZFC, except for the axiom of foundation/regularity, a corresponding axiom can be proven in his system.
- Using something similar to the von Neumann series of ordinals (finite and infinite), versions of the Peano axioms can be proven.
- Notice, however, that this *not* mean that the system can recapture all of Peano arithmetic (or ZFC minus foundation); since the core logic is weakened in certain respects, capturing the axioms does not mean capturing the theorems. It is still an open question how much of classical mathematics is captured as a "subtheory".
- A set On of all ordinals can be shown to exist; by the argument of the Burali-Forti paradox, it is both a member of itself and not, and both identical to itself and not, much like R.
- A strange argument can be given showing the universe can be well-ordered, and hence that the axiom of choice is true.
- Cantor's theorem, and thereby, Cantor's paradox, can be proven. The universe is both the largest possible size, and also smaller than the size of its own powerset. (It is therefore bigger than itself.)
- The continuum hypothesis (and hence the generalized continuum hypothesis as well) is at least disprovable; certain other instances of the GCH are both provable and disprovable.

XXVI. Plural Logic

A. Basics

Notations for plural logic vary; we shall follow Linnebo's *Stanford Encyclopedia* article.

Add to first-order logic plural variables xx, yy, zz, xx_1, yy_1, $\ldots$, etc., and constants aa, bb, etc.

We introduce a sign $\prec$, allowing a singular term on the left and plural term on the right.

$$x \prec yy$$

means x is one of the things yy, or x is among yy.

Some use a sign $\preccurlyeq$ instead, and then typically allow either side to be plural or singular, where $\ldots \preccurlyeq$ ___ means that every individual in $\ldots$ is identical to one individual in ___.

The Geach-Kaplan sentence, "some critics admire only one another":

$$(\exists xx)((\forall y)(y \prec xx \rightarrow \text{Critic}(y)) \wedge \\ (\forall y)(\forall z)(y \prec xx \wedge \text{Admires}(y,z) \rightarrow z \prec xx))$$

Linnebo sketches two very simple theories, PFO and PFO+. The difference has to do with other predicates in addition to $\prec$. PFO+ allows other predicates to take plural terms; PFO does not. Both share the axiom schemata:

$$(\exists x)\,\mathscr{A}[x] \rightarrow (\exists xx)(\forall y)(y \prec xx \leftrightarrow \mathscr{A}[y]) \qquad \text{(Comprehension)}$$

$$(\forall xx)(\exists y)\, y \prec xx \qquad \text{(Nonemptiness)}$$

$$(\forall xx)(\forall yy)[(\forall z)(z \prec xx \leftrightarrow z \prec yy) \rightarrow (\mathscr{A}[xx] \leftrightarrow \mathscr{A}[yy])] \qquad \text{(Extensionality)}$$

B. Interpreting Second-Order Logic/Type-theory

It is somewhat trivial to reinterpret *monadic* second-order logic in plural logic. Take any formula for the form $\mathscr{P}(t)$ and rewrite it as $t \prec \chi\chi$, and replace each quantified statement with $(\exists \mathscr{P})\ldots$ with $(\exists \chi\chi)\ldots \vee (\ldots^*)$, where $\ldots^*$ results from our original second order $\ldots$ by replacing $\mathscr{P}(t)$ with $t \neq t$. (This is to accommodate null predicates, whereas there is no null plurality.)

Only so much work can be done in monadic second-order logic. Second-order logic with quantification over relations is much more useful.

Two approaches to this:

(1) Introduce a primitive pairing functor π, where:

$$(\forall x)(\forall y)(\forall x_1)(\forall y_1)(\pi(x,y) = \pi(x_1,y_1) \rightarrow x = x_1 \wedge y = y_1)$$

Then we can replace talk of (binary) relations with pluralities of pairs.

(2) Introduce pluplurals or superplurals, i.e., pluralities of plurals, in a hierarchy:

$$x \prec^1 xx$$
$$xx \prec^2 xxx$$
$$xxx \prec^3 xxxx$$

The pair (x, y) can be then be treated as the plurality of x and x and y, similar to the Wiener-Kuratowski definition of $\langle x, y\rangle$ as $\{\{x\}, \{x, y\}\}$.

Such a theory can interpret the simple-theory of types.

However, arguably this involves treating a plurality as a single thing, i.e., as one among a plurality of plurals, which is controversial, or at least it is controversial whether or not it preserves the alleged ontological "innocence" of plural logic.

C. Boccuni's PG

We can also build hybrid systems, which employ plural logic alongside higher-order logic, set theory, or something else.

Let us take as an example Boccuni's system PG, *plural Grundgesetze*, which some unimportant changes in notation here.

The system uses the logical constants $\rightarrow$, $\neg$, $=$, $\exists$, $\prec$, a vbto $\{\chi|\mathscr{A}[\chi]\}$ for "extensions", and three types of variables, individual, $x, y, z, \ldots$, monadic predicate, $F, G, H, \ldots$, and plural, xx, yy, zz.

Rather than use polyadic predicates or predicate varaibles, the system introduces ordered pairs (x, y)—see below—and $F(x, y)$ is taken as predicating F of this pair.

To guard against paradox, the vbto $\{\chi|\mathscr{A}[\chi]\}$ is only allowed when the formula $\mathscr{A}[\chi]$ contains neither bound predicate variables nor free plural variables.

There are three axiom schemata:

$$(\exists xx)(\forall y)(y \prec xx \leftrightarrow \mathscr{A}[y]) \qquad \text{(PLC)}$$

where $\mathscr{A}[y]$ does not contain xx free.

$$(\exists F)(\forall x)(Fx \leftrightarrow \mathscr{A}[x]) \qquad \text{(PRC)}$$

where $\mathscr{A}[x]$ does not contain F free plural or bound predicate variables.

$$\{\chi|\mathscr{A}[\chi]\} = \{\chi|\mathscr{B}[\chi]\} \leftrightarrow (\forall\chi)(\mathscr{A}[\chi] \leftrightarrow \mathscr{B}[\chi]) \qquad \text{(V)}$$

Note that (PLC) allows empty pluralities.

(PRC) is limited to so-called "predicative" instances. Along with the restrictions on the vbto, this keeps the system consistent. Note that if $\in$ is defined:

$$t \in u \text{ for } (\exists F)(u = \{x|Fx\} \land Ft)$$

Then any use of it is not allowed in the vbto, as it contains a bound predicate variable. The rival definition:

$$t \in u \text{ for } (\exists xx)(u = \{y|y \prec xx\} \land t \prec xx)$$

is not allowed, as it puts a free plural variable in the vbto. Thus, there is no "Russell class".

The system can interpret Peano arithmetic using the Zermelo sequence for its "numbers". Some definitions:

$\{t\}$	for	$\{\chi|\chi = t\}$ — this is also the *successor* notion
$\{t, u\}$	for	$\{\chi|\chi = t \lor \chi = u\}$
(t, u)	for	$\{\{t\}, \{t, u\}\}$
0	for	$\{x|x \neq x\}$
1	for	$\{0\}$
2	for	$\{1\}$, etc.
$\text{Ind}(tt)$	for	$0 \prec tt \land (\forall x)(x \prec tt \rightarrow \{x\} \prec tt)$
$\mathbb{N}t$	for	$(\forall\chi\chi)(\text{Ind}(\chi\chi) \rightarrow t \prec \chi\chi)$

The Peano "axioms" are then derivable in the following forms:

$$\mathbb{N}0 \tag{PA1}$$

$$(\forall x)(\mathbb{N}x \rightarrow \mathbb{N}\{x\}) \tag{PA2}$$

$$(\forall y)\{y\} \neq 0 \tag{PA3}$$

$$(\forall x)(\forall y)(\{x\} = \{y\} \rightarrow x = y) \tag{PA4}$$

$$(\forall xx)[\mathrm{Ind}(xx) \rightarrow (\forall y)(\mathbb{N}y \rightarrow y \prec xx)] \tag{PA5}$$

HOMEWORK 24

Prove (PA3). Hint: Use (V) and a reductio.

Boccuni gives the usual arguments for the philosophical soundness, or innocence, of plural quantification and predicative second-order logic, but gives a rather more obscure argument in favor of thinking of extensions simply as freely chosen objects, different for non-coextensive formulas.