TRUTH DEFINITIONS AND CONSISTENCY PROOFS

BY

HAO WANG

1. **Introduction.** From investigations by Carnap, Tarski, and others, we know that given a system S, we can construct in some stronger system S' a criterion of soundness (or validity) for S according to which all the theorems of S are sound. In this way we obtain in S' a consistency proof for S. The consistency proof so obtained, which in no case with fairly strong systems could by any stretch of imagination be called constructive, is not of much interest for the purpose of understanding more clearly whether the system S is reliable or whether and why it leads to no contradictions. However, it can be of use in studying the interconnection and relative strength of different systems. For example, if a consistency proof for S can be formalized in S', then, according to Gödel's theorem that such a proof cannot be formalized in S itself, parts of the argument must be such that they can be formalized in S' but not in S. Since S can be a very strong system, there arises the question as to what these arguments could be like. For illustration, the exact form of such arguments will be examined with respect to certain special systems, by applying Tarski's "theory of truth" which provides us with a general method for proving the consistency of a given system S in some stronger system S'. It should be clear that the considerations to be presented in this paper apply to other systems which are stronger than or as strong as the special systems we use below.

Originally the studies reported here were motivated by a desire to look more carefully into the following somewhat puzzling situation.

Let S be a system containing the usual second-order predicate calculus with the usual number theory as its theory of individuals, and S' be a system related to S as an $(n+1)$th order predicate calculus is to an nth except that we do not use variables of the $(n+1)$th type in defining classes of lower types. Tarski's assertions seem to lead us to believe that we can prove the consistency of S in S'. On the other hand, it is known that if S is consistent then S has a model in the domain of natural numbers. But if S has such a model, then, we seem also to be able to argue, S' has a model in S because S contains both natural numbers and their classes. Therefore, we can formalize (so it appears) these arguments in S' and prove within S' that if S is consistent then S' is. If that be the case we shall have a proof of the consistency of S' within S' and therefore, by Gödel's theorem on consistency proofs, S' (and probably also S) will be inconsistent. Indeed, since we need at least a system like S' to develop analysis and since these reasonings do not depend

Received by the editors April 26, 1951 and, in revised form, August 24, 1951.

on peculiar features of the systems under consideration, we shall be driven to the conclusion that practically every system adequate to analysis is inconsistent.

In trying to examine exactly where the above arguments break down, we have found it helpful to formalize more explicitly certain truth definitions and consistency proofs with such definitions. The results of such formalizations as presented below, it is thought, bring out more clearly than usual certain features in the procedures of constructing truth definitions and proving consistency. For example, the use of impredicative classes is dispensable for defining truth but does not seem so for proving consistency; whether the number of axioms of a system to be proved consistent is finite or infinite seems also to imply much difference in formalizing a consistency proof; and the employment of variables of higher types in defining classes of a given type engenders essentially new classes even for systems which contain otherwise already certain impredicative classes.

It turns out that the arguments of a paragraph back break down because of the relativity of number theory to the underlying set theory. As a result, certain intuitively simple reasonings cannot be formalized in even very strong systems. Thus, for systems S and S' related as above, no matter how strong they are, the following results hold for them if they are consistent.

If natural numbers are taken as primitive notions or introduced with the same definitions in both S and S', then (1) for some predicate ϕ in S', we can show that $\phi(0)$, $\phi(1)$, $\cdots$ are all provable in S' but that $\forall m(\phi(m) \supset \phi(m+1))$ is not; (2) for some other predicate ϕ' of S', we can prove $\phi'(0)$ and $\forall m(\phi'(m) \supset \phi'(m+1))$ in S' but not $\forall m \phi'(m)$. These immediately yield new examples of consistent but ω-inconsistent systems. On the other hand, if we choose in S and S' suitable (different) definitions for natural numbers, we can prove in S' that if S is ω-consistent then S' is consistent and also prove in S' the consistency of S, but not the ω-consistency of S. This also shows that although S' contains a truth definition for S, we cannot prove in S' that S must possess a standard or nonpathological model.

In order to separate two different moments of a truth definition, we shall distinguish between a truth definition and a normal truth definition. If we can find in S' a predicate or a class Tr for which we can prove with regard to the sentences of S all the cases of the Tarski truth schema, we say that S' contains a truth definition for S. If in addition we can prove in S' that all theorems of S are true according to the truth definition, then we say that S' contains a normal truth definition for S. This rather natural distinction will be assumed throughout this paper.

We are greatly indebted to Professors Bernays, Quine, and Rosser who have all generously helped us by scrutinizing our earlier proofs, suggesting criticisms, and pointing out fallacies.

2. **A truth definition for Zermelo set theory.** Expressions of Zermelo set

theory are built up from the set variables x_1, x_2, x_3, $\cdots$ and three constants: the sign $|$ for alternative denial (Sheffer's stroke function, disjunctive negation), the sign $\forall$ for general or universal quantification (all-operator), and the sign $\in$ for the membership relation (belonging to). Parentheses for grouping different parts of an expression, although theoretically dispensable, are also employed. A sentence (or well-formed formula) is either of the simple form $y \in z$ or of the complex form $p \mid q$ or $\forall y p$, where we may substitute in place of y and z any variables, and in place of p and q any sentences. From among the sentences some may be selected as theorems. However, since this section is concerned merely with the construction of a truth definition and not a normal truth definition, the selection of theorems is irrelevant for the considerations here. So just let us imagine for the moment that an arbitrary definite set of sentences are taken as theorems of the Zermelo theory.

The problem is to find a suitable system S_1 in which we can find a class (or a predicate) Tr and prove as theorems the special cases of the Tarski truth schema for all the statements (closed sentences, well-formed formulas containing no free variables) of Zermelo theory.

To simplify the structure of the required metasystem, let us assume that the syntax of Zermelo theory (as well as that of every other system we consider) has been arithmetized after the manner of Gödel(1). Then each expression (and in particular, each statement) is represented by a definite number (say)(2) $\mathfrak{m}$. Let $H(\mathfrak{m})$ be the expression represented by the number $\mathfrak{m}$. The problem is to define in a system S_1 a class Tr of natural numbers for which the following holds:

(TS) For each statement $H(\mathfrak{m})$ of Zermelo theory, we can prove in S_1: $H(\mathfrak{m})$ if and only if $\mathfrak{m}$ belongs to Tr.

Naturally we can choose the metasystem S_1 in different manners, and the proof of (TS) would become somewhat easier if the metasystem we use is richer or stronger. However, since one of our purposes is to make explicit the material needed for constructing a truth definition, it seems desirable to choose as weak (or simple) a system as is conveniently possible. The system S_1 we choose may be roughly described as of equal strength as a second-order predicate calculus founded on natural numbers. It does not seem possible to use any system which is substantially weaker than this system S_1 (but compare the system S_2 given below).

S_1 contains variables x_1, x_2, $\cdots$ for elements and variables X_1, X_2, $\cdots$ for classes. Sentences are built up from simple sentences of the forms $x_1 \in x_2$ (the element x_1 belongs to the element x_2), etc. and $x_1 \eta X_2$ (the element x_1 belongs to the class X_2), etc. by truth-functional connectives and quantifiers

(1) See Gödel [1] and Hilbert-Bernays [1, vol. 2, §4].

(2) In contrast with the variables m, n, etc., the symbols $\mathfrak{m}$, $\mathfrak{n}$, etc. stand ambiguously for the numerals 0, 1, 2, etc.

for both kinds of variable in the usual manner.

The theorems of S_1 are determined as follows. It contains the ordinary quantification theories (the predicate calculi) for both kinds of variable, as well as proper axioms given below (where $x_1=x_2$ stands for $\forall x_4(x_4\in x_1\equiv x_4\in x_2)$):

Ax1. First axiom of extensionality. $(x_1=x_2 \ \& \ x_1\in x_3)\supset x_2\in x_3$.

Ax2. Existence of the null element. $\exists x_2\forall x_1(-x_1\in x_2)$.

Ax3. Existence of finite elements. $\exists x_2\forall x_1(x_1\in x_2\equiv(x_1\in x_3\vee x_1=x_4))$.

Ax4. Second axiom of extensionality. $(x_1=x_2 \ \& \ x_1\eta X_1)\supset x_2\eta X_1$.

Ax5. The class axiom. For every sentence p of S_1 in which the variable X_1 does not occur, $\exists X_1\forall x_1(x_1\eta X_1\equiv p)$.

From Ax2 and Ax3, we can obtain all finite elements constructed from the null element by taking unit sets and sum sets. Ax5 states that every property or predicate of these elements which are expressible in S_1 determines a class of S_1. This system is closely related to certain standard systems. Thus, on the one hand, it differs from the part consisting of the axioms of groups I–III of Bernay's system([3]) only in that Ax5 takes the place of the somewhat weaker axioms of his group III. On the other hand, it is practically of the same strength as a system proposed by Quine([4]), although Quine uses in place of Ax3 an axiom stating that the sum set of two elements is again an element.

In S_1 we can follow either the definitions used by Bernays or those used by Quine and develop the ordinary number theory while taking certain elements as natural numbers. Thus, we can define with von Neumann([5]):

2.1. The number zero 0 is identified with the null element.

2.2. The successor x_1+1 of a natural number x_1 is identified with the element consisting of all the members of x_1 together with x_1 itself (i.e., the sum set of x_1 and its unit set).

Then we can define the predicate Nn of being a natural number in either of the following two manners([6]).

2.3. $\mathrm{Nn}(x_1)$ if and only if $\forall X_1((0\eta X_1 \ \& \ \forall x_2(x_2\eta X_1\supset(x_2+1)\eta X_1))\supset x_1\eta X_1)$.

2.4. $\mathrm{Nn}(x_1)$ if and only if the following conditions are satisfied:

([3]) See Bernays [1, Part I]. We note incidentally that the axiom Ax1 of S_1 is actually redundant by virtue of Ax4 and Ax5. For similar reasons, as Bernays of course realizes, the middle two of the four axioms on identity listed on Bernays [1, p. 67], as well as the first of the two on p. 68, are derivable from the other axioms on identity with the help of his axioms of group III.

([4]) See Quine [2, p. 140].

([5]) See von Neumann [1].

([6]) The first definition is essentially the same as the last definition on Quine [2, p. 142] (cf. also Quine [1, p. 216]).

The second definition is the definition of finite ordinals on p. 11 of Bernays [1, Part II], while ordinals are defined by the second definition appearing on p. 9, ibid. The possibility of developing number theory in S_1 with 2.4 in place of 2.3 has been explained to us by Professor Bernays in conversation.

(1) $\forall x_2 \forall x_3((x_2 \in x_3 \,\&\, x_3 \in x_1) \supset x_2 \in x_1)$;

(2) $\forall x_2 \forall x_3((x_2 \in x_1 \,\&\, x_3 \in x_1 \,\&\, x_2 \neq x_3) \supset (x_2 \in x_3 \vee x_3 \in x_2))$;

(3) $\forall X_1((\exists x_2(x_2 \eta X_1) \,\&\, \forall x_2(x_2 \eta X_1 \supset x_2 \in x_1)) \supset \exists x_3(x_3 \eta X_1 \,\&\, \forall x_4 - (x_4 \in x_3 \,\&\, x_4 \eta X_1)))$;

(4) $\forall x_2((x_2 \in x_1 \vee x_2 = x_1) \supset (x_2 = 0 \vee \exists x_3(x_2 = x_3 + 1))$.

Recursive functions such as addition and multiplication can be defined in known manner(7). Hence, number theory can be developed in S_1. Variables m, n, k, etc. ranging over natural numbers can be introduced by contextual definitions such as: $\forall m \phi m$ stands for $\forall x_1(\mathrm{Nn}(x_1) \supset \phi x_1)$, etc. The principle of induction can be proved, with the help of Ax5 (the class axiom), for all sentences of S_1:

2.5. If ϕx_1 is any sentence of S_1, then $(\phi 0 \,\&\, \forall n(\phi n \supset \phi(n+1))) \supset \forall m \phi m$.

From the considerations thus far, we see that S_1 is, in the matter of linguistic forms, an extension of the Zermelo set theory, and contains: (1) variables of a higher type; (2) usual notations of general logic; (3) all expressions of the Zermelo theory; (4) the ordinary number theory. Therefore, according to an assertion of Tarski(8), if the axioms of S_1 are sufficiently strong, we can use S_1 as a metalanguage and define in it the concept of truth for Zermelo theory, with merely the help of names and predicates of expressions (or terms from the morphology) of Zermelo theory. However, since S_1 contains number theory and the syntax of Zermelo theory is assumed to have been arithmetized, we can use in S_1 natural numbers and their predicates as names and predicates of expressions of Zermelo theory. Hence, we may expect, with good reasons, that we can obtain in S_1 a truth class (or predicate) for Zermelo theory. The only remaining question is whether the axioms of S_1 are strong enough to enable us to prove the schema (TS). We shall show that the answer is positive. The constructions follow in their general line the methods developed by Tarski and others(9), although the demonstrations, as well as the exact theorems of the metasystem which are needed, are here exhibited more explicitly than by these previous authors.

With the arithmetization of the syntax of the Zermelo theory, we can define in S_1 syntactical predicates and functions for it. Let us assume that all the sentences (well-formed formulas) of the Zermelo set theory have been enumerated, and in particular in such a way that when a sentence is a part

(7) See Quine [1, p. 259] and Bernays [1, Part II, p. 11].

(8) See Tarski [1, p. 399, Thesis A]. Instead of (1), Tarski requires that the order or type of the metalanguage should be higher than the language for which the truth definition is constructed. His notion of the order or type of a language seems to concern both the linguistic forms and the proving power of the language system.

(9) See Tarski [1; 2; 3], Hilbert-Bernays [1, vol. 2, p. 324 ff.], Professor Quine's lectures at Harvard University in the Spring of 1948, and, especially, Kemeny [1]. In connection with the formalization of truth definitions, as presented below, we want to thank Professors Bernays and Rosser for extremely helpful criticisms of our earlier attempts, as well as suggestions on ways of making improvements.

of another, the part always precedes the whole. For each $\mathfrak{i}$, let $F(\mathfrak{i})$ be the $\mathfrak{i}$th sentence of Zermelo theory in the assumed enumeration. For example, $F(1)$ may be just $x_1 \in x_1$. We can define in S_1 three predicates $M(m, n, k)$, $T(m, n, k)$, $Q(m, n, k)$, and a function R_m satisfying the following conditions.

2.6. $M(\mathfrak{m}, \mathfrak{n}, \mathfrak{k})$ when and only when $F(\mathfrak{m})$ is $x_{\mathfrak{n}} \in x_{\mathfrak{k}}$.

2.7. $T(\mathfrak{m}, \mathfrak{n}, \mathfrak{k})$ when and only when $F(\mathfrak{m})$ is $(F(\mathfrak{n}) \mid F(\mathfrak{k}))$.

2.8. $Q(\mathfrak{m}, \mathfrak{n}, \mathfrak{k})$ when and only when $F(\mathfrak{m})$ is $\forall x_{\mathfrak{n}} F(\mathfrak{k})$.

2.9. $R_{\mathfrak{m}}$ is the number of logical operators in $F(\mathfrak{m})$.

Then the following elementary properties are easily provable in S_1.

2.10. $(M(m, n, k) \ \& \ M(m, n', k')) \supset (n = n' \ \& \ k = k')$.

2.11. $(T(m, n, k) \ \& \ T(m, n', k')) \supset (n = n' \ \& \ k = k')$.

2.12 $(Q(m, n, k) \ \& \ Q(m, n', k')) \supset (n = n' \ \& \ k = k')$.

2.13. $R_m \leqq 0 \supset \exists n \exists k M(m, n, k)$.

2.14. $(R_m \leqq (j+1) \ \& \ (T(m, n, k) \vee Q(m, i, k))) \supset (R_n \leqq j \ \& \ R_k \leqq j)$.

The definition of truth for Zermelo theory is roughly this[10]: a sentence (with or without free variables) of Zermelo theory is true if and only if it is satisfied by all finite sequences of sets of Zermelo set theory, and a finite sequence g satisfies (a sentence $F(\mathfrak{m})$ represented by the number) $\mathfrak{m}$ when and only when $F'(\mathfrak{m})$, where $F'(\mathfrak{m})$ is the result obtained from $F(\mathfrak{m})$ by substituting simultaneously, for all its free variables $x_{\mathfrak{n}}$ ($\mathfrak{n}$ among 1, 2, $\cdots$), the $\mathfrak{n}$th term of g for $x_{\mathfrak{n}}$[11]. In the particular cases where $F(\mathfrak{m})$ is a statement (of Zermelo theory), it follows that $\mathfrak{m}$ belongs to the class of numbers representing true statements when and only when $F(\mathfrak{m})$.

In Tarski's definitions, infinite sequences are used instead of finite sequences. This is not possible in our approach, because the axioms of S_1 only guarantee that all finite sequences of elements are again elements but not that infinite sequences are also. However, as we have sequences with any arbitrary finite number of terms, we can dispense with infinite sequences altogether in our considerations.

The definition for a sequence is simply this. An element x_1 of S_1 is a finite sequence of elements of S_1 if there exists a number n ($n = 1, 2, \cdots$) such that x_1 is an n-termed sequence; and x_1 is an n-termed sequence when and only when for all m, $m > 0$ and $m \leqq n$, there exists a unique element x_2 of S_1 such that the ordered pair of m and x_2 belongs to x_1. More explicitly, the definition may be stated as follows[12].

---

[10] See Tarski [1, pp. 311–313].

[11] It should be noted that here, as elsewhere in this paper, we are trying to avoid the use of quotation marks and corners (cf. Quine [1, p. 33]). It is hoped that no serious misunderstandings or confusions will result from such a practice.

[12] In this connection we should like to mention an interesting alternative definition of sequences which was introduced by Professor Quine in his lectures already referred to and applies equally well to finite and infinite sequences: $\mathrm{Sq}(x_1)$ when and only when $\forall x_2 \ (x_2 \in x_1 \supset \exists x_3 \exists m (x_2 = (x_3, m) \ \& \ m \neq 0))$. In other words, instead of labelling the terms of the sequences, we label the members of each term and call the sum of all these labelled members

2.15. The ordered pair (x_1, x_2) of two elements x_1 and x_2 of S_1 is the element of S_1 consisting of the unit set of x_1 and the sum set of x_1 and x_2.

2.16. $\mathrm{Sq}(x_1)$ (i.e., x_1 is a finite sequence) when and only when $\exists n(n>0 \ \& \ \forall x_3(\exists x_2((x_2, x_3)\in x_1) \equiv \exists m(m=x_3 \ \& \ m\leqq n)) \ \& \ \forall x_2\forall x_3\forall x_4 (((x_2, x_3)\in x_1 \ \& \ (x_4, x_3)\in x_1)\supset x_2=x_4))$.

In S_1 there exist of course elements which are finite sequences according to this definition. Indeed, for each n and any n elements $x_1, \cdots, x_n$ of S_1, the set consisting of $(x_1, 1)$, $(x_2, 2)$, $\cdots$, (x_n, n) is one such. Let us use the letter g as a variable ranging over those elements of S_1 which are finite sequences:

2.17. $\forall g\phi g$ when and only when $\forall x_1(\mathrm{Sq}(x_1)\supset\phi x_1)$.

The jth term g_j of a finite sequence g is the set correlated with j:

2.18. $x_1\in g_j$ when and only when $\exists x_2((x_2, j)\in g \ \& \ x_1\in x_2)$.

This definition involves a rather undesirable complication which would not arise if infinite sequences were employed instead. Thus, when g has only k terms and j is greater than k, we would want to say that g has no jth term; however, according to this definition, g_j would then be the null set. In such degenerating cases, we have the result that (g_j, j) does not necessarily belong to g. It turns out that this unnatural feature is harmless for our further developments in the sense that it does not affect the definitions and theorems in which we are mainly interested.

Next is the notion of "the sequence $t(g, n, x_2)$ obtained from g by substituting the set x_2 for its nth term":

2.19. $(x_1, m)\in t(g, n, x_2)$ when and only when $(m\neq n \ \& \ (x_1, m)\in g) \vee (m=n \ \& \ x_1=x_2)$.

Using the preliminary notions introduced in 2.6–2.9 and 2.16–2.19, we can now characterize the notion of satisfiability ("gSm", meaning "g satisfies the mth sentence of Zermelo theory") by the following conditions: (1) $M(m, n, k)\supset(gSm\equiv g_n\in g_k)$; (2) $T(m, n, k)\supset(gSm\equiv(gSn \mid gSk))$; (3) $Q(m, n, k)\supset(gSm\equiv\forall x_1(t(g, n, x_1)Sk))$. This recursive characterization can be converted into an explicit definition for "gSm" with known methods and then the concept of truth for Zermelo theory can be defined.

Df1. $G1(g, m)$ when and only when $\exists n\,\exists k(M(m, n, k) \ \& \ g_n\in g_k)$.

Df2. $G2(g, m, X_1)$ when and only when $\exists n\,\exists k(T(m, n, k) \ \& \ ((g, n)\eta X_1 \mid (g, k)\eta X_1))$.

---

the sequence consisting of these terms. Thus, if x_1 is a sequence, the jth term of x_1 is just the set of all x_2 such that $(x_2, j)\in x_1$. When the sets or classes of a system are divided into types so that a set and its members are of different types, this definition has over the ordinary one the advantage of keeping the sequence in the same type as its terms. In certain cases, it seems necessary to use such a definition, replacing the definiens of 2.18 by $(x_1, j)\in g$ and that of 2.19 by $((m\neq n \ \& \ (x_1, m)\in g)\vee(m=n \ \& \ x_1\in x_2))$. When necessary, we shall assume that these definitions have been adopted instead of 2.16, 2.18, and 2.19 given in the text. For example, when we define truth for R in R' and prove 5.8 in the last section, we shall assume such alternative definitions.

Df3. $G3(g, m, X_1)$ when and only when $\exists n \exists k(Q(m, n, k) \& \forall x_1((t(g, n, x_1), k)\eta X_1))$.

Df4. $G(g, m, X_1)$ when and only when $(G1(g, m) \vee G2(g, m, X_1) \vee G3(g, m, X_1))$.

Df5. $G_j(X_1)$ when and only when $\forall g \forall m(R_m \leqq j \supset ((g, m)\eta X_1 \equiv G(g, m, X_1)))$.

Df6. gSm when and only when $\forall X_1(G_{R_m}(X_1) \supset (g, m)\eta X_1)$.

Df7. $m\eta Tr$ when and only when $\forall g(gSm)$.

Df5 amounts essentially to this: if $G_j(X_1)$, then X_1 contains all the ordered pairs (g, m) such that g satisfies m and R_m is no greater than j. In order to prove the conditions (1)–(3) as theorems of S_1, we first show that for each j, there exists in S_1 a class X_1 such that $G_j(X_1)$ and that if $G_{R_m}(X_1)$ then for every g, gSm when and only when (g, m) belongs to X_1.

The next three theorems of S_1 are obvious from the definitions.

2.20. $M(m, n, k) \supset (-G2(g, m, X_1) \& -G3(g, m, X_1))$.

2.21. $T(m, n, k) \supset (-G1(g, m, X_1) \& -G3(g, m, X_1))$.

2.22. $Q(m, n, k) \supset (-G1(g, m, X_1) \& -G2(g, m, X_1))$.

We prove the theorem that for each j there exists in S_1 a class X_1 such that $G_j(X_1)$.

2.30. $\exists X_1(G_0(X_1))$.

Proof. By Df4 (using 2.13 and 2.20), for every X_1, if $R_m \leqq 0$, then $(G1(g, m) \equiv G(g, m, X_1))$. By Ax5, there exists a class X_1 such that for every g and every m, $((g, m)\eta X_1 \equiv G1(g, m))$. Hence, by Df5, $G_0(X_1)$.

2.31. $\exists X_1(G_j(X_1)) \supset \exists X_2(G_{j+1}(X_2))$.

Proof. Let X_1 be a class such that $G_j(X_1)$. Hence, by Df5, $R_m \leqq j \supset ((g, m)\eta X_1 \equiv G(g, m, X_1))$. By Ax5, there exists a class X_2 such that $\forall g \forall m((g, m)\eta X_2 \equiv G(g, m, X_1))$. So, if $R_i \leqq j$, then $((g, i)\eta X_1 \equiv (g, i)\eta X_2)$. Hence, by 2.14 and Df1–Df4, if $R_m \leqq j+1$, then $(G(g, m, X_1) \equiv G(g, m, X_2))$. Therefore, $R_m \leqq j+1 \supset ((g, m)\eta X_2 \equiv G(g, m, X_2))$. Hence, by Df5, 2.31 is proved.

From 2.30 and 2.31, we have immediately:

2.32. For each constant $\mathfrak{j}$, we can prove in S_1: $\exists X_1(G_{\mathfrak{j}}(X_1))$.

Moreover, by applying induction (the consequence 2.5 of Ax5) to the sentence $\exists X_1(G_i(X_1))$, we obtain from 2.30 and 2.31:

2.33. $\forall j \exists X_1(G_j(X_1))$.

We note that in order to prove 2.33, we require that bound large (class) variables be allowed in the defining sentence p of Ax5 for class formation. It will be emphasized later that this is one of the few places where such cases of Ax5 must be applied in our considerations.

Having on hand for each j the existence of some class X_1 such that $G_j(X_1)$, we want now to prove some kind of uniqueness theorem for these classes. As in the definition of G_j we are interested only in the numbers m such that $R_m \leqq j$, we shall prove merely that for all m where $R_m \leqq j$, if $G_j(X_1)$ and $G_j(X_2)$, then $(g, m)\eta X_1$ if and only if $(g, m)\eta X_2$.

2.34. $(R_m \leqq 0 \ \& \ G_0(X_1) \ \& \ G_0(X_2)) \supset ((g, m)\eta X_1 \equiv (g, m)\eta X_2)$.

Proof. By Df4 and Df5 (using 2.13 and 2.20), if $(G_0(X_1) \ \& \ G_0(X_2) \ \& \ R_m \leqq 0)$, then $((g, m)\eta X_1 \equiv G1(g, m))$ and $((g, m)\eta X_2 \equiv G1(g, m))$. Hence, the theorem is proved.

2.35. If $\forall X_1 \forall X_2 \forall g \forall m((R_m \leqq j \ \& \ G_j(X_1) \ \& \ G_j(X_2)) \supset ((g, m)\eta X_1 \equiv (g, m)\eta X_2))$, then $\forall g \forall m((R_m \leqq j+1 \ \& \ G_{j+1}(X_3) \ \& \ G_{j+1}(X_4)) \supset ((g, m)\eta X_3 \equiv (g, m)\eta X_4))$.

Proof. Assume that $G_{j+1}(X_3)$ and $G_{j+1}(X_4)$. Then, by Df5, we have also: $G_j(X_3)$ and $G_j(X_4)$. Hence, by hypothesis, if $R_m \leqq j$, then $(g, m)\eta X_3 \equiv (g, m)\eta X_4$. Therefore, by 2.14 and Df4, if $R_m \leqq j+1$, then $G(g, m, X_3) \equiv G(g, m, X_4)$. Hence, by Df5, the theorem is proved.

An immediate consequence of 2.34 and 2.35 is:

2.36. For each constant $\mathfrak{j}$, we can prove in S_1: $(R_m \leqq \mathfrak{j} \ \& \ G_{\mathfrak{j}}(X_1) \ \& \ G_{\mathfrak{j}}(X_2)) \supset ((g, m)\eta X_1 \equiv (g, m)\eta X_2)$.

Again, by applying the induction principle 2.5, we have:

2.37. $(j)((R_{m \leqq} j \ \& \ G_j(X_1) \ \& \ G_j(X_2)) \supset ((g, m)\eta X_1 \equiv (g, m)\eta X_2))$.

This is the second place where we need a case of Ax5 in which the defining sentence for a class contains large bound variables.

The next theorem follows from 2.37.

2.38. $G_{R_m}(X_1) \supset (gSm \equiv (g, m)\eta X_1)$.

Proof. By 2.37, $(G_{R_m}(X_1) \ \& \ (g, m)\eta X_1) \supset \forall X_2(G_{R_m}(X_2) \supset (g, m)\eta X_2)$. Hence, by Df6, the theorem is easily proved.

If we have only 2.36 but not 2.37, then we can prove only:

2.39. For every constant $\mathfrak{m}$, we can prove in S_1: $G_{R\mathfrak{m}}(X_1) \supset (gS\mathfrak{m} \equiv (g, \mathfrak{m})\eta X_1)$.

Now we are ready to prove the characteristic properties of the relation gSm.

2.40. $M(m, n, k) \supset (gSm \equiv g_n \in g_k)$.

Proof. By 2.20 and Df5, $(M(m, n, k) \ \& \ G_0(X_1)) \supset ((g, m)\eta X_1 \equiv G1(g, m))$. Hence, by Df1, $(M(m, n, k) \ \& \ G_0(X_1)) \supset ((g, m)\eta X_1 \equiv \exists n \exists k(M(m, n, k) \ \& \ g_n \in g_k))$. Obviously, $M(m, n, k) \supset (g_n \in g_k \supset \exists n \exists k(M(m, n, k) \ \& \ g_n \in g_k))$. On the other hand, by 2.10, $(M(m, i, j) \ \& \ g_i \in g_j) \supset (M(m, n, k) \supset g_n \in g_k)$. Hence, $(M(m, n, k) \ \& \ G_0(X_1)) \supset ((g, m)\eta X_1 \equiv g_n \in g_k)$. Therefore, by 2.38, $(M(m, n, k) \ \& \ G_0(X_1)) \supset (gSm \equiv g_n \in g_k)$. By 2.30, the theorem is proved.

2.41. $T(m, n, k) \supset (gSm \equiv (gSn \mid gSk))$.

Proof. By 2.21 and Df5, $(T(m, n, k) \ \& \ G_{R_m}(X_1)) \supset ((g, m)\eta X_1 \equiv G_2(g, m, X_1))$. Hence, by 2.11 and Df2, $(T(m, n, k) \ \& \ G_{R_m}(X_1)) \supset ((g, m)\eta X_1 \equiv ((g, n)\eta X_1 \mid (g, k)\eta X_1))$. By 2.38, $(T(m, n, k) \ \& \ G_{R_m}(X_1)) \supset (gSm \equiv (gSn \mid gSk))$. Therefore, by 2.33, the theorem is proved.

2.42. $Q(m, n, k) \supset (gSm \equiv \forall x_1(t(g, n, x_1)Sk))$.

Proof. By 2.22 and Df5, $(Q(m, n, k) \ \& \ G_{R_m}(X_1)) \supset ((g, m)\eta X_1 \equiv G3(g, m, X_1))$. Hence, by 2.12 and Df3, $(Q(m, n, k) \ \& \ G_{R_m}(X_1)) \supset ((g, m)\eta X_1 \equiv \forall x_1((t(g, n, x_1), k)\eta X_1))$. Therefore, by 2.38, $(Q(m, n, k) \ \& \ G_{R_m}(X_1))$

$\supset(gSm \equiv \forall x_1(t(g, n, x_1)Sk))$. By 2.33, the theorem is proved.

From 2.40–2.42 we can derive the following important metatheorem about S_1.

2.43. Let g be a variable not occurring in the sentence $F(\mathfrak{m})$ of Zermelo theory (and, a fortiori, of S_1) which contains free occurrences of the variables $x_i, \cdots, x_j$ and of no others. If $F'(\mathfrak{m})$ is the sentence obtained from $F(\mathfrak{m})$ by substituting the set expressions $g_i, \cdots, g_j$ (as defined in 2.18) respectively for these variables $x_i, \cdots, x_j$, then we can prove in S_1: $gS\mathfrak{m} \equiv F'(\mathfrak{m})$.

Proof. We prove 2.43 by making induction on the number of logical operators in $F(\mathfrak{m})$.

Case 1. $F(\mathfrak{m})$ contains no logical operators. We may assume that $F(\mathfrak{m})$ is the sentence $(x_\mathfrak{n} \in x_\mathfrak{f})$. Therefore, since we can develop number theory in S_1, we can prove: $M(\mathfrak{m}, \mathfrak{n}, \mathfrak{f})$. Hence, 2.43 follows from 2.40 because $F'(\mathfrak{m})$ is $(g_\mathfrak{n} \in g_\mathfrak{f})$.

Case 2. Assume 2.43 holds true for all cases where $F(\mathfrak{m})$ contains no more than s logical operators. We want to prove it for the cases where $F(\mathfrak{m})$ contains $s+1$ logical operators.

Case 2a. $F(\mathfrak{m})$ is of the form $(p \mid q)$. We may assume that the sentence is $(F(\mathfrak{n}) \mid F(\mathfrak{f}))$. Therefore, we can prove in S_1: $T(\mathfrak{m}, \mathfrak{n}, \mathfrak{f})$. Let $F'(\mathfrak{n})$ and $F'(\mathfrak{f})$ be related to $F(\mathfrak{n})$ and $F(\mathfrak{f})$ in the same manner as $F'(\mathfrak{m})$ is to $F(\mathfrak{m})$. By 2.41, $gS\mathfrak{m} \equiv (gS\mathfrak{n} \mid gS\mathfrak{f})$. Hence, by induction hypothesis, $gS\mathfrak{m} \equiv (F'(\mathfrak{n}) \mid F'(\mathfrak{f}))$. Hence, 2.43 is proved, because $(F'(\mathfrak{n}) \mid F'(\mathfrak{f}))$ is $F'(\mathfrak{m})$.

Case 2b. $F(\mathfrak{m})$ is of the form $\forall x_\mathfrak{n} p$. We may assume that $F(\mathfrak{m})$ is the sentence $\forall x_\mathfrak{n} F(\mathfrak{f})$. Therefore, we can prove in S_1: $Q(\mathfrak{m}, \mathfrak{n}, \mathfrak{f})$. Let $F'(\mathfrak{f})$ be related to $F(\mathfrak{f})$ as $F'(\mathfrak{m})$ is to $F(\mathfrak{m})$ except that free occurrences of $x_\mathfrak{n}$ in $F(\mathfrak{f})$ are not replaced by those of $g_\mathfrak{n}$ in $F'(\mathfrak{f})$. By 2.42, $gS\mathfrak{m} \equiv \forall x_\mathfrak{n}(t(g, \mathfrak{n}, x_\mathfrak{n})S\mathfrak{f})$. Therefore, by induction hypothesis, $gS\mathfrak{m} \equiv \forall x_\mathfrak{n} F'(\mathfrak{f})$. Hence, 2.43 is proved, because $F'(\mathfrak{m})$ is $\forall x_\mathfrak{n} F'(\mathfrak{f})$.

From this the truth schema (TS) for Zermelo theory can be proved directly.

2.44. If $F(\mathfrak{m})$ is a closed sentence (statement) of Zermelo theory, then we can prove in S_1: $\mathfrak{m}\eta Tr \equiv F(\mathfrak{m})$.

Proof. By 2.43, if $F(\mathfrak{m})$ is a closed sentence, then we can prove in S_1: $gS\mathfrak{m} \equiv F(\mathfrak{m})$. Hence, by Df7, $\mathfrak{m}\eta Tr \equiv F(\mathfrak{m})$.

Hence, we reach the main theorem of this section.

THEOREM I. *In S_1 we can construct a truth definition for the Zermelo set theory.*

It may be worthwhile to emphasize here again that in this section we are merely concerned with the construction of a truth definition which need not be also a normal one; in other words, we do not assert that according to the truth definition given above, we can prove in S_1 that all the theorems of Zermelo theory are true. In order that in a system S we be able to prove

such an assertion, it would be necessary to require that S contain S_1, as well as theorems which answer to those of the Zermelo set theory under consideration. This problem will be studied more carefully in a later section.

Although, when compared with standard axiomatic systems for set theory, S_1 should be considered a very weak system, it already contains the impredicative feature (through Ax5) which separates typical set theories from more elementary disciplines. It is therefore of interest to note in this connection that we can also obtain a truth definition (but not a normal one) for Zermelo theory in a system which does not contain impredicative classes.

Let S_2 be the system which contains the same linguistic forms as S_1 and is obtained from S_1 by substituting for Ax5 the following axioms([13]).

Ax6. For every sentence p of S_2 in which neither X_1 nor any bound large variables occur, $\exists X_1 \forall x_1 (x_1 \eta X_1 \equiv p)$.

Ax7. $\exists x_3 \forall x_1 (x_1 \in x_3 \equiv (x_1 \in x_2 \ \& \ x_1 \eta X_2))$.

We want to prove that in S_2 we can also obtain a truth definition for the Zermelo theory. Since S_2 has the same notation as S_1, all the definitions in S_1 are also definitions in S_2, except that it would be more correct to write $Tr(m)$ in place of $m \eta Tr$ in Df7, because there is no truth class in S_2 but only a predicate. However, this point is not important for our purpose.

Our problem is to prove a metatheorem for S_2 which answers to 2.44 for S_1. In the first place, when Ax5 is replaced by Ax6, the theorems 2.33 and 2.38 are no longer demonstrable, and we can only prove the metatheorems 2.32 and 2.39. Consequently, in place of the theorems 2.40–2.42, we can prove in S_2 only the following metatheorems:

2.45. $M(\mathfrak{m}, \mathfrak{n}, \mathfrak{f}) \supset (g S \mathfrak{m} \equiv g_{\mathfrak{n}} \in g_{\mathfrak{f}})$.

2.46. $T(\mathfrak{m}, \mathfrak{n}, \mathfrak{f}) \supset (g S \mathfrak{m} \equiv (g S \mathfrak{n} \mid g S \mathfrak{f}))$.

2.47. $Q(\mathfrak{m}, \mathfrak{n}, \mathfrak{f}) \supset (g S \mathfrak{m} \equiv \forall x_1 (t(g, \mathfrak{n}, x_1) S \mathfrak{f})$.

However, an examination of the proofs for 2.43 and 2.44 should make it clear that 2.45–2.47 are already adequate to the derivation of 2.43 and 2.44. Therefore, it would seem that we can define truth for Zermelo theory in S_2 even without applying Ax7.

Such would be true if we could develop number theory in S_2 without using Ax7. As a matter of fact, there seems to be no way of doing so, although it is known that number theory can be developed in S_2 with the help of Ax7([14]).

Consequently, we have the next theorem:

THEOREM II. *In S_2 we can construct a truth definition for the Zermelo theory.*

([13]) If we have merely Ax6 instead of Ax5, we have to require that the predicate Nn of being a natural number be defined by a "constitutive expression" (cf., e.g., Bernays [1, Part II, p. 12]), while the bound class variable in condition [3] of 2.4 contradicts this. If we assume also Ax7, the class variable in the condition can be replaced by a set variable; see Bernays [1, Part II, top of p. 9].

([14]) See Bernays [1, Part II].

3. **Remarks on the construction of truth definitions in general**[15]. In order to exhibit the procedure of constructing truth definitions more explicitly, we studied in the previous section only two special systems. Here we indicate how similar considerations are applicable to other formal systems as well.

In the Zermelo theory we can consider the predicate $\in$ as an operator for generating sentences from variables, and the logical operators $|$ and $\forall$ as operators for generating new sentences from given ones. Let L be an arbitrary system which contains one kind of variable just as Zermelo theory, but contains in addition to (or instead of) $\in$, other predicates $P_1, \cdots, P_i$, in addition to (or instead of one or both of) $|$ and $\forall$, other operators $O_1, \cdots, O_j$. Here again, we can suppose that the theorems of L have been selected in an arbitrary but definite way.

Let L^* be the system which is like S_1 except that it contains, besides the sentences of S_1, also the sentences generated by the predicates $P_1, \cdots, P_i$ and the logical operators $O_1, \cdots, O_j$. Let the axioms of L^* be like those of S_1 except that the domain of the sentences p in Ax5 be extended accordingly. Then we can construct in L^* a truth definition for L just as we did in S_1 for Zermelo theory. The necessary changes are few and simple. Thus, instead of Df1, we define in similar manner $G1(g, m), \cdots, Gi(g, m)$ for the predicates $P_1, \cdots, P_i$ of L, and instead of Df2–Df3, we define $G(i+1)(g, m, X_1), \cdots, G(i+j)(g, m, X_1)$ for the logical operators $O_1, \cdots, O_j$. The definition Df4 for $G(g, m, X_1)$ is then modified as an alternation of the $(i+j)$ clauses $G1, \cdots, G(i+j)$ thus defined.

After these changes in the definitions, it would be a routine matter to modify the proofs in the preceding section and demonstrate two theorems in L^* which are notationally the same as 2.33 and 2.38. From these two theorems, theorems answering to 2.40–2.42 can be proved with analogous proofs. Hence, we have:

THEOREM III. *If L is any system with one kind of variable and L^* contains S_1, as well as all sentences of L, we can construct in L^* a truth definition for L.*

In the above proof, we have assumed that the number of predicates and logical operators is finite. The theorem also holds if in L either the number of predicates or that of logical operators or both are denumerably infinite. In such a case, some further modifications in the procedure are necessary, for otherwise we would need for the definition of $G(g, m, X_1)$ an infinite alternation of clauses. We can proceed in the following manner. Define, as in Df1–Df3, $G1$, $G2$, $\cdots$ for all the predicates and logical operators. Let $F(\mathfrak{m})$ be the $\mathfrak{m}$th sentence of L, $R_{\mathfrak{m}}$ be the number of logical operators in $F(\mathfrak{m})$, and

[15] Later sections are independent of the material contained in this section, which can be omitted by a reader interested only in the few conclusions regarding the relativity of number theory and induction, to be presented in the last part.

$N_{\mathfrak{m}}$ be the greatest $\mathfrak{k}$ such that the $\mathfrak{k}$th predicate occurs in $F(\mathfrak{m})$. Instead of $G(g, m, X_1)$ as defined in Df4, define $G(g, m, X_1, 1)$, $G(g, m, X_1, 2)$, $\cdots$ as (finite) alternations of terms drawn from $G1$, $G2$, $\cdots$ in such a way that for each $\mathfrak{i}$, the clauses $G1$, $G2$, $\cdots$ for all the predicates and logical operators occurring in some $F(\mathfrak{m})$ for which $N_{\mathfrak{m}} \leqq \mathfrak{i}$ and $R_{\mathfrak{m}} \leqq \mathfrak{i}-1$ occur as alternation terms in $G(g, m, X_1, \mathfrak{i})$. Instead of definitions Df5–Df7, we use definitions of the following forms ($\mathfrak{j}$, $\mathfrak{m}=1, 2, \cdots$):

Df5′. $G_{\mathfrak{j}}(X_1)$ when and only when $\forall g \forall_m((R_m \leqq \mathfrak{i}-1 \;\&\; N_m \leqq \mathfrak{j}) \supset ((g, m)\eta X_1 \equiv G(g, m, X_1, \mathfrak{j})))$.

Df6′. $gS\mathfrak{m}$ when and only when $\forall X_1(G_{R\mathfrak{m}+\mathfrak{N}\mathfrak{m}}(X_1) \supset (g, \mathfrak{m})\eta X_1)$.

Df7′. $\mathfrak{m}\eta Tr$ when and only when $\forall g(gS\mathfrak{m})$.

With these definitions we can prove for L^* metatheorems answering to 2.32 and 2.39 (although not the theorems corresponding to 2.33 and 2.38), and therefore those answering to 2.45–2.47. However, these are precisely what we need in order to prove the metatheorems for L and L^* which answer to 2.43 and 2.44 for Zermelo theory and S_1. Hence, in these cases, we can also give in L^* a truth definition for L. In other words, Theorem III holds true no matter whether L contains finitely or (denumerably) infinitely many predicates and logical operators.

We insert here a few remarks on function symbols and constant names. We have thus far been assuming that systems are so formulated that function symbols and constant names do not occur among their primitive notations. Such an approach seems to be in accord with Tarski's procedure. And it is partly justified by the assertion([16]) that when we use sufficiently many predicates, constant names and function symbols are theoretically dispensable. Nevertheless, it must be admitted, the alternative procedure of including among the primitive notation of a system constant names and function symbols from which terms are generated seems to be more intuitive and tends to clarify matters in many connections. However, if we consider a system formulated with terms among its primitive notation, the construction of a truth definition for it would have to differ considerably from what is described in the preceding section. Moreover, for such systems it would often appear possible to construct truth definitions with more direct and intuitive methods([17]). Since considerations regarding such possibilities would lead us far afield, we shall continue to assume that no function symbols or constant names occur in the primitive notations of the systems which we study.

So far we have restricted ourselves to systems each containing only one

([16]) See Quine [1, p. 149]; Hilbert-Bernays [1, vol. 1, p. 460].

([17]) If we compare the truth definition for number theory in Hilbert-Bernays, vol. 2, with the truth definition for Zermelo theory elaborated in the present paper, it seems fair to consider the former more straightforward and intuitively simpler, involving a more direct recursion. It is not very easy to determine the exact conditions which a system must satisfy in order to possess such a truth definition. See also the observations in the paragraph between parentheses on p. 63 of Tarski [3].

kind of variable in its primitive notation. Let us now consider the case where a system contains many kinds of variable.

The simplest way of handling such a case seems to be the following: consider instead of the given system with many kinds of variable (a many-sorted system) an "equivalent" system with one kind (a one-sorted system). Since we know this is always possible([18]), we seem to be able to avoid altogether the question of defining truth for many-sorted theories.

However, since in certain cases it is more natural to use many kinds of variable, it is desirable to consider directly how we can construct a truth definition for a given many-sorted theory. In this connection, Tarski has given indications as to how we should proceed([19]). We state merely the general conditions which a metasystem should satisfy.

Given a one-sorted system, what do we need in a metasystem in order that it be adequate to defining truth for the given system? If we examine the construction of truth definition for Zermelo theory, we see that the following things are needed: (1) General logic (quantification theory and theory of identity) for each kind of variable in the metasystem; (2) ordinary number theory together with variables m, n, $\cdots$ ranging over natural numbers; (3) sufficient resources for defining all finite sequences of entities of the given system and for introducing a variable g which ranges over such sequences; (4) existence of a class of ordered pairs of g (a finite sequence) and m (a natural number), corresponding to each sentence of the metasystem which contains g and m as free variables.

Let K be an arbitrary many-sorted theory. In order that a system K^* be adequate as a metasystem in which we can define truth for K, K^* must contain materials similar to those listed under (1)–(4). The only important necessary alteration is with regard to item (3). To satisfy a sentence of K with n free variables, we need a suitable sequence with n terms such that for each k between 1 and n, the kth term is an object falling within the range of values of the kth free variable in the given sentence. Hence, it is necessary that, besides the things listed under (1), (2), and (4), such finite sequences g of arbitrary terms from K be obtainable in K^*. Roughly speaking, if K^* contains number theory, we need only classes which take as members all entities of K, because finite sequences and ordered pairs can usually be defined with the help of natural numbers. Thus, for example, if K is the predicate calculus of the ith order founded on natural numbers, then it is sufficient to use the predicate calculus of the $(i+1)$th order as K^*([20]).

([18]) See Schmidt [1] and a forthcoming paper *Logic of many-sorted theories* in J. Symbolic Logic. We are here interested in ordinary systems which contain for each kind of their variables the ordinary complete quantification theory.

([19]) See Tarski [1].

([20]) See Tarski [2, p. 110], the first three sentences of §8. More detailed investigations have been made in Kemeny [1] which contains also the treatment of a highly interesting case where the system to be studied contains infinitely many kinds of variables.

4. Consistency proofs via truth definitions. It seems to be widely believed that once we have in L^* a truth definition for L, it is then a routine matter to formalize in L^* a consistency proof for L. Probably partly on account of this belief, details of such consistency proofs are usually not supplied. However, as we remarked before, in order to prove in L^* the consistency of L through a truth definition, we also have to prove in L^* that all theorems of L are true by the definition; and such a proof not only calls for strong axioms in L^* but usually also involves a number of complications. We consider in this section a few special cases of such consistency proofs.

Consider first a weak and simple system S_3 of set theory. The linguistic forms of S_3 are as given in §2 for the Zermelo theory. The theorems of S_3 are specified in the following manner. The proper axioms of S_3 are just the axioms Ax1–Ax3 of the system S_1. Moreover, all axioms of the quantification theory as given below are also theorems of S_3:

Q1. $p \supset (q \supset p)$.

Q2. $(p \supset (q \supset r)) \supset ((p \supset q) \supset (p \supset r))$.

Q3. $(-p \supset -q) \supset (q \supset p)$.

Q4. $\forall x(\phi x \supset \psi x) \supset (\forall x \phi x \supset \forall x \psi x)$.

Q5. If x is not free in p, $p \supset \forall x p$.

Q6. $\forall x \phi x \supset \phi y$.

The only rule of inference is modus ponens for closed sentences.

Q7. If p and $p \supset q$ are theorems, then q is also.

It should be noted that, following Quine([21]), we do not allow free variables to occur in theorems. Thus, when a sentence is listed as an axiom or a theorem, we mean actually that a closure of the given sentence (i.e., a closed sentence obtained from the given sentence by prefixing distinct general quantifiers for all its free variables) is an axiom or a theorem. Moreover, we also follow Quine in calling $\forall x p$ a vacuous quantification when x is not free in p.

By Theorems I and II, S_1 and S_2 each contains a truth definition for S_3. Moreover, all the theorems of S_3 are also theorems of S_1 and S_2. We now show that the consistency of S_3 can be formally proved in S_1. In all probability no consistency proof for S_3 can be formalized in S_2, although we possess no proof of the impossibility.

Since S_1 contains number theory, syntactical notions for S_3 can be defined in S_1 through an arithmetization of the syntax of S_3. In particular, we assume that the following notions have been defined in S_1: $\supset(m, n, k)$ ($F(m)$ being $(F(n) \supset F(k))$), $\exists(m, n, k)$ ($F(m)$ being $\exists x_n F(k)$), $\equiv(m, n, k)$ ($F(m)$ being $F(n) \equiv F(k)$), $\vee(m, n, k)$ ($F(m)$ being $(F(n) \vee F(k))$), inf (m, n, k) ($F(m)$, $F(n)$, $F(k)$ being all closed, and $\supset(m, n, k)$), ax$fr(m)$ (a closure of $F(m)$ is an axiom of S_3 and none of the initial general quantifiers of $F(m)$ whose ranges extend to the end of $F(m)$ is vacuous), ax(m) ($F(m)$ is an axiom of S_3 which, incidentally, must be a closed sentence), pr(n, m) (the nth proof of S_3 is a

([21]) See Quine [1, chap. 2].

proof for $F(m)$, an arbitrary enumeration of the proofs of S_3 having been assumed), thm(m) ($F(m)$ is a theorem of S_3: $\exists n(\mathrm{pr}(n, m))$), neg (m) (the number k such that $F(k)$ is the negation of $F(m)$), Con (S_3) (S_3 is consistent: $\forall n - \mathrm{pr}(n, \mathfrak{m}_0)$, $F(\mathfrak{m}_0)$ being the sentence $\forall x_2 \exists x_1(x_1 \in x_2)$ of S_3), $\lambda(n)$ (the number of lines in the nth proof), $\rho(m)$ (the number of general quantifiers standing at the beginning of $F(m)$ and having ranges all extended to the end of $F(m)$).

We note incidentally that we can fix on an arbitrary enumeration of all the proofs of S_3. We assume merely that the enumeration is made in such a way that when the mth proof contains the nth proof as a proper part, then $m > n$ in the enumeration.

From 2.41–2.42 and Df7, we can prove immediately that the following are theorems of S_1.

4.1. $\supset(m, n, k) \supset (gSm \equiv (gSn \supset gSk))$.

4.2. $\equiv(m, n, k) \supset (gSm \equiv (gSn \equiv gSk))$; $\vee(m, n, k) \supset (gSm \equiv (gSn \vee gSk))$.

4.3. $\exists(m, n, k) \supset (gSm \equiv \exists x(t(g, n, x)Sk))$.

4.4. $(\mathrm{inf}\ (m, n, k)\ \&\ m\eta Tr\ \&\ n\eta Tr) \supset k\eta Tr$.

By the definitions (2.18 and 2.19) of g_j and $t(g, n, x)$, we can prove in S_1:

4.5. $(t(g, n, x))_n = x$, $(t(t(g, n, x), m, y))_n = x$, $(t(t(g, n, x), m, y))_m = y$, etc.

By 2.44, we have:

4.6. $-\mathfrak{m}\eta Tr \equiv \mathrm{neg}\ (\mathfrak{m})\eta Tr$.

The first crucial theorem we want to prove in S_1 is:

4.7. $\mathrm{ax}fr(m) \supset m\eta Tr$.

Proof. The axioms of S_3 are of nine kinds falling under two groups: (1) the axioms of quantification theory given by Q1–Q6; (2) the three definite axioms Ax1–Ax3 of set theory. If $\mathrm{ax}fr(m)$, then a closure of $F(m)$ is an axiom of S_3. We prove 4.7 by making induction on $\rho(m)$. First we assume that we have defined nine arithmetic predicates $\mathrm{ax}_1, \cdots, \mathrm{ax}_9$ corresponding to the nine kinds of axioms with their initial quantifiers omitted. Thus, for instance, $\mathrm{ax}_1(m)$ if and only if $F(m)$ is of the form $(p \supset (q \supset p))$.

Case 1. $\rho(m) = 0$. We have then in S_1: $\mathrm{ax}fr(m) \supset (\mathrm{ax}_1(m) \vee \cdots \vee \mathrm{ax}_9(m))$. Since the proofs for all cases in each of the two groups are similar, we prove only one case from each group for illustration.

Case 1a. $F(m)$ is of one of the forms given in Q1–Q6. Then 4.7 can be proved by 2.41–2.42 (together with their consequences such as 4.1) and the number theory of S_1. For example, suppose $F(m)$ is of the form $(\forall x(\phi x \supset \psi x) \supset (\forall x \phi x \supset \forall x \psi x))$. In other words, ax_4 (m). We can prove in S_1: $\exists m_1 \cdots \exists m_7 \exists n(\supset(m, m_1, m_2)\ \&\ Q(m_1, n, m_3)\ \&\ \supset(m_2, m_4, m_5)\ \&\ Q(m_4, n, m_6)\ \&\ Q(m_5, n, m_7)\ \&\ \supset(m_3, m_6, m_7))$. Hence, by repeated applications of 2.42 and 4.1, we have: $gSm \equiv (\forall y(t(g, n, y)Sm_6 \supset t(g, n, y)Sm_7) \supset (\forall y(t(g, n, y)Sm_6) \supset \forall y(t(g, n, y)Sm_7)))$. Therefore, by the quantification theory in S_1, gSm. Hence, by Df7, $m\eta Tr$.

Case 1b. $F(m)$ is of one of the forms Ax1–Ax3. Then 4.7 can be proved

by appealing to 2.40–2.42 (together with their consequences such as 4.1–4.4) and the corresponding axioms Ax1–Ax3 in S_1. For example, suppose $F(m)$ is: $\exists x_2 \forall x_1 (x_1 \in x_2 \equiv (x_1 \in x_3 \vee \forall x_5 (x_5 \in x_1 \equiv x_5 \in x_4)))$. In other words, $\mathrm{ax}_9(m)$. We can prove in S_1: $\exists m_1 \cdots \exists m_8 (\exists(m, 2, m_1) \,\&\, Q(m_1, 1, m_2) \,\&\, \equiv(m_2, m_3, m_4) \,\&\, M(m_3, 1, 2) \,\&\, \vee(m_4, m_5, m_6) \,\&\, M(m_4, 1, 3) \,\&\, Q(m_5, 5, m_6) \,\&\, \equiv(m_6, m_7, m_8) \,\&\, M(m_7, 5, 1) \,\&\, M(m_8, 5, 4))$. Hence, by repeated applications of 2.40, 2.42, 4.2, and 4.3, we have: $gSm \equiv \exists x \forall y ((t(t(g, 2, x), 1, y))_1 \in (t(t(g, 2, x), 1, y))_2 \equiv ((t(t(g, 2, x), 1, y))_1 \in (t(t(g, 2, x), 1, y))_3 \vee \forall z ((t(t(t(g, 2, x), 1, y), 5, z))_5 \in (t(t(t(g, 2, x), 1, y), 5, z))_1 \equiv (t(t(t(g, 2, x), 1, y), 5, z))_5 \in (t(t(t(g, 2, x), 1, y), 5. z))_4)))$. Therefore, by 4.5, writing a in place of $(t(t(g, 2, x), 1, y))$, we have: $gSm \equiv \exists x \forall y (y \in x \equiv (y \in a_3 \vee \forall z (z \in y \equiv z \in (t(a, 5, z))_4)))$. Since a and $(t(a, 5, z))_4$ are by definition again terms of S_3 (taking sets as values), (a closure of) the right-hand side of the equivalence is provable in S_1 by Ax3. Hence, we can prove in S_1: gSm. By Df7, $m \eta Tr$.

Case 2. Assume that 4.7 is true for all m such that $F(m)$ satisfies the condition $\rho(m) = n$. We want to prove the cases of 4.7 where $\rho(m) = n+1$. Suppose given: $\mathrm{ax} fr(m)$ & $\rho(m) = n+1$. By induction hypothesis, we can then prove: $\exists j \exists k (Q(m, j, k) \,\&\, k \eta Tr)$. Therefore, we have by Df7 and 2.19: $\exists j \exists k (Q(m, j, k) \,\&\, \forall x (t(g, j, x) Sk))$. Hence, by 2.42 and Df7, $m \eta Tr$.

Combining Case 1 and Case 2, we prove 4.7 by applying the induction principle 2.5 of S_1.

It should be noted that here again we are applying Ax5 of S_1 in its full generality on account of the indispensable bound class variable involved in the definition of Tr. In other words, the above case of induction cannot be proved in S_2 where only inductions with regard to sentences involving no bound class variables are allowed.

Two immediate corollaries of 4.7 are[22]:

4.8. $\vdash_{S_1} \mathrm{ax}\ (m) \supset m \eta Tr$.

4.9. If a constant $\mathfrak{m}$ is given, then $\vdash_{S_1} \mathrm{ax}\ (\mathfrak{m}) \supset \mathfrak{m} \eta Tr$.

Then we can prove that every theorem of S_3 is true.

4.10. $\vdash_{S_1} (\mathrm{pr}(n, m) \,\&\, \lambda(n) \leqq 1) \supset m \eta Tr$.

Proof. If $(\mathrm{pr}(n, m) \,\&\, \lambda(n) \leqq 1)$, then $(\mathrm{ax}\ (n) \,\&\, n = m)$ and therefore, by 4.8, $m \eta Tr$.

4.11. If $\forall n \forall m ((\mathrm{pr}(n, m) \,\&\, \lambda(n) \leqq k) \supset m \eta Tr)$, then $\forall n \forall m ((\mathrm{pr}(n, m) \,\&\, \lambda(n) \leqq k+1) \supset m \eta Tr)$.

Proof. If $(\mathrm{pr}(n, m) \,\&\, \lambda(n) \leqq k+1)$ then, according to the assumption about the enumeration of proofs, $\mathrm{ax}\ (m) \vee \exists i \exists j\ (\mathrm{inf}\ (i, j, m) \,\&\, \exists n_1 \exists n_2 (\mathrm{pr}(n_1, i) \,\&\, \mathrm{pr}(n_2, j) \,\&\, \lambda(n_1) \leqq k \,\&\, \lambda(n_2) \leqq k))$. Therefore, if $\forall n \forall m ((\mathrm{pr}(n, m) \,\&\, \lambda(n) \leqq k) \supset m \eta Tr)$, then $\mathrm{ax}\ (m) \vee \exists i \exists j\ (\mathrm{inf}\ (i, j, m) \,\&\, i \eta Tr \,\&\, j \eta Tr)$. Hence, by 4.8 and 4.4, $m \eta Tr$.

4.12. $\vdash_{S_1} \mathrm{pr}(n, m) \supset m \eta Tr$. Or $\vdash_{S_1} \mathrm{thm}\ (m) \supset m \eta Tr$.

---

[22] We use the sign $\vdash$ as in Quine [1]. When necessary, we specify the system concerned by $\vdash_{S_1}$, etc.; for instance, $\vdash_S \phi$ if and only if the closure of ϕ is a theorem of S.

Proof. If $\mathrm{pr}(n, m)$, then $\exists k(\lambda(n) \leqq k)$. Therefore, by 4.10, 4.11, and the induction principle 2.5 of S_1, 4.12 is proved.

The consistency of S_3 can now be proved.

THEOREM IV. $\vdash_{S_1} -\mathrm{pr}(n, \mathfrak{m}_0)$. *Or* $\vdash_{S_1} \mathrm{Con}\ (S_3)$.

Proof. By definition, $F(\mathfrak{m}_0)$ is the sentence $\forall x_2 \exists x_1(x_1 \in x_2)$, the denial of which is equivalent to Ax 2 of S_3. Therefore, thm (neg $(\mathfrak{m}_0)$) and, by 4.12, neg $(\mathfrak{m}_0)\eta Tr$. Hence, by 4.6, $-\mathfrak{m}_0\eta Tr$ and, by 4.12, $-\mathrm{pr}(n, \mathfrak{m}_0)$. Therefore, we have also: Con (S_3).

This completes the formalization within S_1 of a consistency proof for S_3. We now make a few remarks on the relation between analogously related systems.

If S is any set theory which has the same notation as S_3 but contains in addition to all the axioms of S_3 also a finite number of other set-theoretical axioms, and S' is related to S as S_1 is to S_3; then we can formalize similarly in S' a consistency proof for S.

However, if S should include an infinite number of set-theoretical axioms (i.e., include axiom schemata beyond quantification theory), then the situation would be different. Thus, let S_4 be the ordinary Zermelo set theory which has the same notation as S_3 but contains beyond Ax1–Ax3, the axioms of infinity, sum set, power set, and the following axiom schema (the Aussonderungsaxiom)[23]:

Ax8. If p is any sentence of S_4 in which x_2 is not free, then $\exists x_2 \forall x_1(x_1 \in x_2 \equiv (x_1 \in x_3\ \&\ p))$.

Let S_5 be the system obtained from S_1 by adding the extra axioms of S_4. Can we prove Con (S_4) (as a sentence of S_5) to be a theorem of S_5?

Professor J. Barkley Rosser has observed in correspondence that we cannot do this with the above method of proving IV. More specifically, the proof for the analogue of Case 1b of 4.7 would break down, the crux being that we would need something like an infinite alternation of sentences. As a result, we can prove only the analogue of 4.9 (but not that of 4.8), and therefore we cannot prove an analogue of 4.10.

Since S_5 contains a truth definition of S_4 as well as all the theorems of S_4, we would expect Con (S_4) to be provable in S_5 by some alternative method. However, our attempts for obtaining such a proof have not been successful. We even suspect that there might be a way of demonstrating the unprovability of Con (S_4) in S_5. In any case, so far as we know, the provability or unprovability of Con (S_4) in S_5 remains an open question.

We have assumed the same axiom Ax8 in both S_4 and S_5, allowing no class variables to occur in the sentence p. If we extend the system S_5 and replace Ax8 by a similar but stronger Ax8′ in which p may be any sentence of S_5, then a proof of Con (S_4) can indeed be obtained in the resulting system

[23] Compare for a description of these other axioms Wang [2].

S_5'. Thus, we can proceed exactly as in the proof of Theorem IV except that in proving an analogue of Case 1b of 4.7, we employ special treatments for the alternative when $F(m)$ is of the form $\exists x_2 \forall x_1 (x_1 \in x_2 \equiv (x_1 \in x_3 \ \& \ p))$ as given in Ax8. There are of course infinitely many sentences of S_4 which are of this form, but for each $F(m)$ of such a form, there exists a natural number j such that $F(m)$ is $\exists x_2 \forall x_1 (x_1 \in x_2 \equiv (x_1 \in x_3 \ \& \ F(j)))$. Hence, by arguments similar to those used in the proof of 4.7, we can prove in S_5' something like: $gSm \equiv \exists x \forall y (y \in x \equiv (y \in a \ \& \ gSj))$, a being a term of S_4 which takes sets as values. But the right-hand side of the equivalence is a case of Ax8′ (although not a case of Ax8 on account of the class variable occurring in gSj through an analogue of Df6). Therefore, we have in S_5': gSm. In this way we can prove:

4.13. $\vdash_{S_5'} \text{Con } (S_4)$.

Similarly, if a system S' is related to a system S as a predicate calculus of the $(n+1)$th order founded on natural numbers is to one of the nth order, we can prove Con (S) in S'; but if we weaken S' by stipulating that no variables of the highest type are allowed in defining sets of lower types, then the question whether Con (S) is provable in the resulting system seems to remain open.

Another point worthy of some consideration is the question of proving Con (S_3) in S_2. In the proof of Con (S_3) (as a sentence of S_1) in S_1, the induction principle 2.5 is applied at four places (viz. in the proofs of 2.33, 2.37, 4.7, and 4.12) in such a manner that analogous arguments do not seem formalizable in S_2. Hence, a similar proof of Con (S_3) cannot be carried out in S_2, although, as we mentioned before, we have not been able to obtain a proof for the assertion that Con (S_3) as an arithmetic sentence in S_2 is not provable in S_2.

If we want to prove Con (S_3) (with the arithmetic notions involved as defined in S_2) in S_2, we must try to avoid the applications of induction to sentences containing bound class variables, such as in the proofs of the theorems (answering to) 2.33, 2.37, 4.7, and 4.12. But it does not seem possible to avoid all these applications.

One crucial point seems to be the presence of a rule of inference (the rule Q7 of modus ponens) in S_3 which permits us to infer a shorter sentence from longer ones. If we could so reformulate S_3 that all its rules of inference are such that we can only infer a longer sentence from some definite finite number of shorter ones, then, no matter whether the number of axioms be finite or infinite, we would be able to prove Con (S_3) for the reformulated S_3 without applying those inductions. Indeed, if that were possible, we would have a decision procedure for S_3, and the proof of Con (S_3), as can be expected, would be quite simple.

It may be of interest to note that we can prove in place of Con (S_3) the following metatheorem about S_2 which tells us that no given definite proof

of S_3 can be a proof of the sentence $\forall x_2 \exists x_1(x_1 \in x_2)$ of S_3:

4.14. If $\mathfrak{n}$ is a constant, then $\vdash_{S_2} -\mathrm{pr}(\mathfrak{n}, \mathfrak{m}_0)$.

This depends on the fact that if we consider only the individual numbers one by one, we need not make the inductions. To see this, we merely observe that 2.45–2.47 (instead of 2.40–2.42) are sufficient for proving in place of two theorems answering to 4.10 and 4.11 two metatheorems:

4.15. For given $\mathfrak{m}$ and $\mathfrak{n}$, $\vdash_{S_2}(\mathrm{pr}(\mathfrak{n}, \mathfrak{m}) \,\&\, \lambda(\mathfrak{n}) \leqq 1) \supset \mathfrak{m}\eta Tr$.

4.16. For each given $\mathfrak{k}$ and each given $\mathfrak{m}$, if for every given $\mathfrak{n}$, $(\mathrm{pr}(\mathfrak{n}, \mathfrak{m}) \,\&\, \lambda(\mathfrak{n}) (\leqq \mathfrak{k}) \supset \mathfrak{m}\eta Tr$, then for every given $\mathfrak{n}$, $(\mathrm{pr}(\mathfrak{n}, \mathfrak{m}) \,\&\, \lambda(\mathfrak{n}) \leqq \mathfrak{k}+1) \supset \mathfrak{m}\eta Tr$.

Moreover, if S_6 is the system obtained from S_3 by adding Ax8 as a new axiom and $\mathrm{pr}_6(n, m)$ (an arithmetic sentence of S_2) represents that the nth proof of S_6 is a proof of $F(m)$, we can also prove for S_2 a metatheorem analogous to 4.14:

4.17. If $\mathfrak{n}$ is a constant, then $\vdash_{S_2} -\mathrm{pr}_6(\mathfrak{n}, \mathfrak{m}_0)$.

The connections between S_2 and S_6 are of special importance because they are similarly related as the von Neumann-Bernays set theory and the Zermelo-Fraenkel. It is known[24] that given any two systems S'' and S related to each other as S_2 is to S_6, the relative consistency of S'' to S can be proved. Furthermore, the proof can be formalized in each ordinary system which is roughly as strong as S_1 or a second order predicate calculus founded on natural numbers. In other words[25],

4.18. If S'' is related to S as S_2 is to S_6, and S''' is an ordinary system roughly no weaker than S_1, then $\vdash_{S'''} \mathrm{Con}\,(S) \supset \mathrm{Con}\,(S'')$, Con (S) and Con (S'') being arithmetic sentences of S''' representing respectively the consistency of S and S''.

Therefore, if we choose S and S'' in such a way[26] that arguments which can be carried out in S_1 can also be carried out in S'', then $\mathrm{Con}\,(S) \supset \mathrm{Con}\,(S'')$ as a sentence of S'' becomes a theorem of S'' and therefore, by Gödel's second theorem, Con (S) cannot be provable in S'' unless S'' is inconsistent[27].

4.19. If S and S'' are related as before and S'' contains S_1, then Con (S) is not a theorem of S'' unless S'' is inconsistent.

However, if the arithmetic sentence $Pc(n)$ of S'' represents that the nth proof of S is the proof of a definite refutable sentence of S (i.e., a sentence such as $0=1$ whose denial is known to be provable in S), then $\forall n(-Pc(n))$ may be taken as Con (S) and we can prove:

4.20. If S and S'' are as in 4.19, then $\vdash_{S''} -Pc(\mathfrak{n})$, $\mathfrak{n}$ being any given

[24] This was first proved by Dr. Novak (see Novak [1]). Later on a different proof was presented in Rosser-Wang [1]. Compare also Wang [2].

[25] Such a result is implicit in Novak [1]. Recently, Dr. Robert F. McNaughton gave a careful proof of 4.18 in his doctoral thesis entitled *On establishing consistency of systems* (Harvard Library, April 1951).

[26] For example, S'' would be like that if we choose as S the system S_4 or the system N in §2 of Wang [4].

[27] See footnote 31 below.

number.

In other words, in such a case although Con (S) is demonstrably unprovable in S'', we can prove by elementary considerations about S'' that no proof of S can be given which is a proof of a previously fixed refutable sentence of S. Indeed, such considerations can be formalized in number theory and we can prove in number theory an arithmetic sentence Con* representing that for every number n, if it is for a certain numeral $\mathfrak{m}$ of S'' the number of a sentence $-Pc(\mathfrak{m})$ of S'', then the nth sentence of S'' is a theorem of S''. Since S'' is assumed to be at least as strong as S_1 which contains number theory, Con* as a sentence of S'' can also be proved in S'' although Con (S) cannot.

4.21. If S and S'' are as in 4.19, then $\vdash_{S''}$ Con*.

If we assume that all theorems of S'' are *true* in some definite sense, Con* may also be taken as expressing indirectly the consistency of S[28]. But, it has to be admitted, this is not very clear.

5. **Relativity of number theory and in particular of induction.** In each of many different forms of set theory, we say that we can develop the ordinary number theory. Sometimes within a same set theory we can also develop number theory in more than one way. Naturally the number theory which we obtain is in each case relative to the axioms of the set theory as well as to the definitions we adopt for the arithmetic notions. If we consider each set theory as a theory for the set concept, then the number theory and the arithmetic concepts we obtain in each case are also relative to the underlying set concept. In particular, in each system of set theory which contains number theory, the principle of induction becomes a set-theoretical principle derivable from the axioms of the system; and whether induction is applicable to a certain sentence of the system depends both on the strength of the axioms of the system and the definitions for the arithmetic notions such as those for the number zero, for the successor function, and for the predicate of being a natural number (or for the class of all natural numbers). In this section we shall illustrate the connection between number theories and their underlying set theories with some rather striking examples.

Let us consider first a system R which has the same notations as S_3 (or as Zermelo set theory) and contains Q1–Q7, plus certain stronger proper axioms in place of Ax1–Ax3. These axioms are, roughly speaking, such as to guarantee the development of ordinary number theory, the existence of infinite sets of natural numbers and predicative classes of such sets. In strength R amounts to a system related to a second-order predicate calculus founded on natural numbers as the von Neumann-Bernays set theory is to the Zermelo-Fraenkel. However, to facilitate considerations about the system, we are presenting it in a rather unnatural form. Thus the general variables $x, y, z, \cdots$ of R are understood roughly as ranging over natural numbers,

(28) Compare in this connection footnote 35 below.

infinite sets of them, as well as classes of such sets. From these we select by contextual definitions the sets $a, b, c, \cdots$ which are capable of being elements of classes:

5.1. $\forall a \phi a$ for $\forall x(\exists y(x \in y) \supset \phi x)$.

5.2. $\exists a$ for $-\forall a -$.

And then variables $r, s, t, \cdots$ of the lowest type are introduced by further restricting the domain:

5.3. $\forall t \phi t$ for $\forall a(\exists b(a \in b) \supset \phi a)$.

5.4. $\exists t$ for $-\forall t -$.

Identity is defined as in S_3:

5.5. $x = y$ for $\forall z(z \in x \equiv z \in y)$.

The proper axioms of R can now be stated.

R1. Axiom of extensionality. $x = y \supset (x \in z \supset y \in z)$.

R2. Existence of denumerably many entities of the lowest type.

$$\exists t \forall s(s \in t \equiv (s = c \vee s = b)).$$

R3. Existence of infinite sets of them. $\exists a \forall t(t \in a \equiv t \in x)$.

R4. Existence of predicative classes of such sets. If y is not free in ϕ and all bound variables in ϕ are element variables (a, b, etc.), then $\exists y \forall a(a \in y \equiv \phi)$.

It should be emphasized that R4 can actually be replaced by a small finite number (7 being sufficient) of axioms so that the number of the proper axioms of R become finite. We shall assume such a reduction has actually been made so that R contains only a finite number of axioms (besides the quantification axioms Q1–Q6). Consequently, the method of consistency proof for S_3 is applicable to R.

We know that number theory can be developed in R in different ways. To be explicit, we assume that number theory is developed in R in the following manner([29]).

By R3 and R4, $\exists a \forall t(t \in a \equiv -t \in t)$. Therefore, $\exists a \forall t(a \neq t)$. Hence, by R2, if we substitute a for both c and b, $\exists t \forall s(s \in t \equiv s \neq s)$. Take this set t as the number zero.

By R2, for every t, there exists an s (the unit set of t) which contains t as the only member. For every natural number t, let the unit set of t be its successor $t+1$.

Let the set of natural numbers t be the intersection of all sets which contain zero and the successor of each of its members.

5.6. $\mathrm{Nn}(t)$ for $\forall x((0 \in x \ \& \ \forall s(s \in x \supset s+1 \in x)) \supset t \in x)$, or $t \in \mathrm{Nn}$ or $\mathrm{Nn}(t)$ for $\forall a((0 \in a \ \& \ \forall s(s \in a \supset s+1 \in a)) \supset t \in a)$.

We note that on account of R3, the two alternative ways of defining Nn are really the same.

According to 5.6 and R4, if $\forall m \phi m$ stands for $\forall t(\mathrm{Nn}(t) \supset \phi t)$, then we have

([29]) Compare Quine [1] and Wang [1].

immediately the induction principle:

5.7. If ϕ is as in R4, then $(\phi 0 \,\&\, \forall n(\phi n \supset \phi(n+1))) \supset \forall m \phi m$.

Let R' be a system with the same notation as S_1 and related to R as S_1 is to S_3. In other words, R' is exactly like S_1 except for containing R1–R4 in place of Ax1–Ax3; or what is the same, R' contains Ax4–Ax5 in addition to the axioms of R. Number theory can be developed in R' in exactly the same manner as in S_1, for instance, by defining zero and successor as in R but defining $\mathrm{Nn}(t)$ as in 2.3 (or, what is the same, by replacing x by X in 5.6). Using such a definition for Nn, we can prove in R' a principle of induction applicable to all sentences of R' (just like 2.5).

Let us assume that the syntax of R and that of R' have both been arithmetized in the usual manner, yielding two arithmetic statements Con (R) and Con (R') which express respectively the consistency of R and R'. If the arithmetization is carried out in the framework of a set theory, then the exact expansion of the arithmetic statement Con (R) or Con (R') depends on the definitions we adopt for the arithmetic notions. Thus, since in R and R' we just assumed different definitions for Nn, the arithmetic statements expressing the consistency of R and R' are different in the two systems. Let $\mathrm{Con}_1(R)$ and $\mathrm{Con}_1(R')$ be the arithmetic statements of R which express respectively the consistency of R and R', and $\mathrm{Con}_2(R)$ and $\mathrm{Con}_2(R')$ be those of R'. Let further $R\#$ be the system obtained from R by adding $\mathrm{Con}_1(R)$ a a new axiom, then there are also arithmetic statements $\mathrm{Con}_1(R\#)$ and $\mathrm{Con}_2(R\#)$ respectively of R and R' which express the consistency of $R\#$.

Using the method of proving Con (S_3) in S_1, we can prove:

5.8. $\vdash_{R'} \mathrm{Con}_2(R)$.

Moreover, applying the method of a previous paper([30]), we can also prove:

5.9. $\vdash_{R'} \mathrm{Con}_2(R\#) \supset \mathrm{Con}_2(R')$.

As $R\#$ is a natural extension of R with an additional axiom whose truth is guaranteed by the consistency of R, we would expect the provability of the relative consistency of $R\#$ to R as expressed by: (1) $\mathrm{Con}_2(R) \supset \mathrm{Con}_2(R\#)$. However, if that were provable in R', we would obtain from 5.8 and 5.9: $\vdash_{R'} \mathrm{Con}_2(R')$. Then R' would be inconsistent([31]). Therefore (1) cannot be provable in R'. But why?

([30]) See the argument for proving the relative consistency of N' to N in §2 of Wang [4].

([31]) As we mentioned in the introduction, a reasoning roughly like this was what motivated the investigations reported in this paper. Our interest in this problem was first caused by Professor Rosser's remark to us that if we could prove the relative consistency of the von Neumann-Bernays set theory to the Zermelo (without the axiom of substitution) with fairly elementary means, the former system would be inconsistent. From the summer of 1949 on, we have tried to combine this with the proof in Wang [4] referred to in the preceding footnote. (For an explanation of the relation between results in this paper and similar conclusions presented by myself and others elsewhere, see last section of the present paper, added after the other sections were completed.)

Indeed, using an argument which Dr. John G. Kemeny told us in conversation, we can prove the following:

5.10. $\vdash_{R'} \mathrm{Con}_1(R) \supset \mathrm{Con}_2(R\#)$.

Thus, since R' contains a truth definition for R, we can derive from $\mathrm{Con}_1(R)$ and the truth schema that the number of $\mathrm{Con}_1(R)$, like the numbers of all the theorems of R, belongs to the truth class Tr. Therefore, we would be able to prove $\mathrm{Con}_2(R\#)$ in the same way as $\mathrm{Con}_2(R)$.

In short, the difficulty lies in the inference from $\mathrm{Con}_2(R)$ to $\mathrm{Con}_1(R)$([32]):

THEOREM V. *If we can derive* $\mathrm{Con}_1(R)$ *from* $\mathrm{Con}_2(R)$ *in* R', *then* R' *is inconsistent.*

The reason why no such derivation is available seems to be the following. Let us call a set or class inductive if it contains 0 and the successor of each of its members. In R and R' the class of natural numbers is the intersection of all the inductive classes of R and R' respectively. Since R' contains more classes and therefore more inductive classes, their intersection is smaller than the corresponding intersection in R. Hence, it is possible that there exists some nonstandard model for R which contains more natural numbers than any model of R'([33]). Accordingly, as $\mathrm{Con}(R)$ amounts to an assertion that no natural number represents a proof of contradiction, it is conceivable that although no natural number of R' does so, some natural numbers of R do. At any rate, there is no obvious reason to think that we can prove in R' such is impossible.

Alternatively, we may want to use in R' the same definitions for the arithmetic notions as in R. Then we have instead of the two arithmetic statements, merely the one statement $\mathrm{Con}_1(R)$ for both systems. But then we can no longer prove $\mathrm{Con}_1(R)$ as we proved $\mathrm{Con}(S_3)$ in S_1, since we have in R', with such definitions, only the induction principle 5.7, while a stronger induction principle is needed for proving analogues of 2.40–2.42. Indeed, if we could prove $\mathrm{Con}_1(R)$ in R', R' would be inconsistent either by Theorem V or by the following 5.11 and 5.12. Thus, using proofs similar to those for 5.9 and 5.10, we can prove:

5.11. $\vdash_{R'} \mathrm{Con}_1(R\#) \supset \mathrm{Con}_1(R')$.

5.12. $\vdash_{R'} \mathrm{Con}_1(R) \supset \mathrm{Con}_1(R\#)$.

Since the only hindrance in the way of proving $\mathrm{Con}_1(R)$ in R' is the two applications of induction on sentences containing large variables (as we have

---

([32]) Theorems V and VI are due essentially to Professor Rosser who, in criticizing our attempts to prove the inconsistency of R' with the methods of the present paper, made the crucial points clear.

([33]) The notion of nonstandard models was first introduced by Henkin (cf. Henkin [1; 2] and Rosser-Wang [1]). Previous works on related problems include Skolem [1; 2] and Malcev [1].

already discussed in the last section), we have the following theorem([34]):

THEOREM VI. *If R' is consistent and we use the same definitions for natural numbers as in R, then the principle of induction in its full generality is independent of the axioms of R' and, in particular, there exists some sentence ϕi of R' containing large variables such that $\phi 0$ and $\forall n(\phi n \supset \phi(n+1))$ are provable in R' but $\forall m \phi m$ is not.*

It follows that if R' is ω-consistent, such a sentence $\forall m \phi m$ must be undecidable in R'.

Another example of undecidable sentences can be obtained if we use considerations similar to those used for 4.20. Let $Pc_1(n)$ be an arithmetic sentence of R and R' which represents that the nth proof of R is the proof of a definite refutable sentence of R. We have:

5.13. If R' [etc. as in Theorem VI], then for each $\mathfrak{n}$, $-Pc_1(\mathfrak{n})$ is provable in R' but $\forall m(-Pc_1(m) \supset -Pc_1(m+1))$ is not.

If the latter were provable, we would have by 5.7 a proof for $\mathrm{Con}_1(R)$ in R'.

Moreover, we know that even when we use in R' the same definition of Nn as in R, the induction principle in its generality can be derived from the following form of reducibility principle:

R5. $\exists y \forall m(m \in y \equiv m \eta X)$, or $\exists a \forall m(m \in a \equiv m \eta X)$.

Therefore, we have:

5.14. The axiom R5 is independent of the axioms of R'.

For similar reason, if $m \geqq n$ stands for $\forall X((n \eta X \ \& \ \forall k(k \eta X \supset (k+1)\eta X)) \supset m \eta X)$, then the following statement is also independent of the axioms of R':

R6. $\forall m(m=0 \vee \exists n(m=n+1)) \ \& \ \forall X \exists m \forall n(n \eta X \supset n \geqq m)$.

Of course in R5 and R6 we are assuming that the variables m, n, etc. are introduced in R' with the same definitions as in R.

With regard to Theorem V, Dr. E. Specker has observed that it illustrates how we can express the consistency of a system in different ways and asks the question whether we might find some arithmetic statement in a given system which both expresses the consistency of the system and yet (in spite of Gödel's theorem) is provable in the system([35]).

We should like to suggest as a possible example the arithmetic statement $\mathrm{Con}_2(L)$ of R' and L, where L has the same notation as R' but contains only the axioms of R plus the quantification theory for the large variables. Al-

([34]) This theorem and a number of other conclusions of the present paper are summarized in Wang [3]. However, in Theorem 9 of Wang [3], which answers to the present theorem, the example in parentheses should be deleted. Moreover, the arguments in the lines fourteen to twenty on p. 451 of Wang [3] are also in error and should be corrected according to the more detailed discussions of the present paper.

([35]) Such a question was first raised by Henkin (see the last part of Henkin [1]) in connection with pathological models in general.

though it follows from Gödel's theorem that $\mathrm{Con}_1(L)$ is not provable in L unless L is inconsistent, it is not obvious that Gödel's theorem also implies the unprovability of $\mathrm{Con}_2(L)$ in L. Indeed, since $\mathrm{Con}_1(L)$ is related to $\mathrm{Con}_2(L)$ as $\mathrm{Con}_1(R)$ is to $\mathrm{Con}_2(R)$, by Theorem V ($\mathrm{Con}_2(R)$ being easily derivable from $\mathrm{Con}_1(R)$), $\mathrm{Con}_1(L)$ may be said to be stronger than $\mathrm{Con}_2(L)$. On the other hand, viewed from the number theory developed in R', it seems completely justifiable to say that $\mathrm{Con}_2(L)$ as a statement of L also expresses the consistency of L.

Another remark relates to the possibility of proving ω-consistency. By 5.9 and Gödel's theorem, it follows that $\mathrm{Con}_2(R\#)$ is not provable in R'. But we know[36] that if R is ω-consistent, then $R\#$ is consistent. Therefore, either (1) the ω-consistency of R is not provable in R', or (2) we cannot formalize in R' the proof for the assertion that $R\#$ is consistent, if R is ω-consistent. Which of the two alternatives is the case?

The answer seems to depend again on how we define the notion of ω-consistency which is closely related to the natural numbers and pseudo-natural numbers allowable by the axioms and definitions of the system. Let us consider merely the simple case where we assume R contains only the normal natural numbers 0 (the empty set), $0+1$ or 1 (the unit set of 0), $1+1$ or 2 (the unit set of 1), etc. and no more.

5.15. The arithmetic statement $w\ \mathrm{Con}_2(R)$ of R' (R is ω-consistent) expresses that for every ϕ of R if $\phi 0$, $\phi 1$, $\phi 2$, etc. are all provable in R, then $-\forall m \phi m$ is not provable in R.

Using this definition, we can prove with the known arguments[37]:

5.16. $\vdash_{R'} w\ \mathrm{Con}_2(R) \supset \mathrm{Con}_2(R\#)$.

Thus let us take $\forall n - Pc_1(n)$ (see 5.13) as $\mathrm{Con}_1(R)$. If $-\mathrm{Con}_2(R\#)$, then $-\forall n - Pc_1(n)$ would be a theorem of $R\#$ and, as $\forall n - Pc_1(n)$ is the only additional axiom of $R\#$, also a theorem of R. If $-Pc(0)$, $-Pc(1)$, $-Pc(2)$, etc. are all theorems of R, then, by 5.15, $-w\ \mathrm{Con}_2(R)$. If there is a numeral $\mathfrak{n}$ such that $Pc(\mathfrak{n})$ is provable in R, then $-\mathrm{Con}_2(R)$ and therefore, by 5.15, $-w\ \mathrm{Con}_2(R)$. It is not hard to formalize the argument in R'.

Therefore, we can also infer the following conclusion:

THEOREM VII. *Although R' contains a normal truth definition for R and we can prove the consistency of R (viz., $\mathrm{Con}_2(R)$), we cannot prove the ω-consistency of R (viz. $w\mathrm{Con}_2(R)$) in R' unless R' is inconsistent.*

We may take this opportunity to state a few simple observations regarding the connection between truth definitions and consistency proofs. Tarski often stresses the importance of the truth schema. Given two systems S and S', he often asks whether there is a class or predicate Tr of S' such that every statement which falls under the following schema is a theorem of S':

[36] See Wang [4, §2].

[37] Ibid.

(T) p if and only if x belongs to Tr. (Or, alternatively, p if and only if $Tr(x)$).

In this schema the letter p can be replaced by any statement of the system S and the letter x by the metalogical designation (name, Gödel number, etc.) of this statement.

Let us say that S' contains a truth definition for S if and only if we can find a class or predicate Tr in S' and prove all the cases of (T), and that S' contains a normal truth definition for S if S contains both a truth definition for S and derivatively a consistency proof for S (or in other words, roughly speaking, S' contains a truth definition for S according to which all the theorems of S are true). Obviously,

5.17. There exist systems S and S' such that S' contains a truth definition for S but no normal one.

For example, take S to be the full Zermelo set theory with all its axioms, and S' to be the system S_1 (see §2); since the former is easily seen to be "stronger" than the latter, there can be no consistency proof for S in S'.

Moreover, if we call a truth definition for S abnormal if some theorems of S come out false according to the definition, we can also find systems S and S' such that S' contains an abnormal truth definition for S. For example, this would be the case if we take again the full Zermelo theory as S and take S_1 plus a contradictory of the axiom of infinity for the elements of S_1 (values of the small variables) as S'.

There is the question whether S' is stronger than S or whether S' is translatable into S, if S' contains a truth definition for S. To answer these questions, we must of course first make clear what we mean by being stronger than or being translatable into another system. Let us assume the definitions we employed on a previous occasion([38]), which do not appear far removed from our ordinary use of such words as modelling, translation, and strength of systems.

As we have shown there, it then follows from Gödel's theorem that if S' is sufficient for ordinary number theory and contains a *normal* truth definition for S, then S' is not translatable into S and S' has no model in S. If further S' contains S as a part then S' is stronger than S.

However, using the same definitions, it is perfectly possible that S' contains a truth definition for S but is both weaker than and translatable into S. For example, S_1 contains a truth definition for the full Zermelo theory, but it is easily shown that S_1 is translatable into the latter but the latter is not translatable into S_1. Although Tarski has shown([39]) that no system S' with the same notation as a system S can contain a truth definition (normal or not) for S, we cannot infer that S' must be stronger than S if S' contains any truth definition for S at all.

([38]) See Wang [4, §1].

([39]) See Tarski [1].

On the other hand, it is possible that S' and S have the same linguistic forms, and yet S' contains some "transformed" truth definition for S in the sense that there is a correlation of all the sentences of S with some sentences of S of certain special forms and for which latter there is a truth definition in S' (normal or not). For example, this seems to be what is happening when we say that a (transformed) normal truth definition for one system of the Zermelo set theory (for example, the original Zermelo system as refined by Skolem) can be found in another (for example, the Zermelo-Fraenkel), which has the same linguistic forms but contains additional axioms (the axiom of substitution in the case of our example)([40]).

6. **Explanatory remarks**([41]). We tabulate below in one place the characteristics of the principal systems studied above, for reference; and take into account the following two recent publications by Professor Mostowski:

Mostowski$_1$. *Some impredicative definitions in the axiomatic set-theory*, Fund. Math. vol. 37 (1950) pp. 111–124.

Mostowski$_2$. Review of Wang [3], J. Symbolic Logic vol. 16 (1951) pp. 142–143.

These two items, which I had no opportunity of seeing while I was preparing the main parts of this paper, call for explanations of the extent to which Mostowki's independent results in Mostowski$_1$ overlap with those presented in Wang [3] and the present paper.

First we give in summary brief descriptions of the principal systems considered in the preceding sections.

(1) S_3 is a very weak set theory in which we assume merely the null set and the finite sets constructed out of it; S_3 has the same notations as the ordinary Zermelo set theory (one primitive predicate and one kind of variable only) and contains merely the axiom of extensionality, the axiom of null set, and an axiom saying that by adding a new member to a given set, we have again a set. As we know, S_3 has a simple model in the elementary theory of numbers.

(2) S_1 is a second-order predicate calculus with S_3 as its theory of individuals; S_1 contains both predicative and impredicative classes of sets of S_3. (Compare the systems of Quine [2] and Wang [1].) S_1 is as strong as a second-order predicate calculus with the ordinary theory of numbers as its theory of individuals.

(3) S_6 is obtained from S_3 by adding the Aussonderungsaxiom (a schema) guaranteeing the existence of every subset of a given set. S_6 has the same arithmetic model as S_3.

(4) S_2 is related to S_6 as S_1 is to S_3 except that S_2 contains only predicative classes (and no impredicative ones) of the sets of S_6 and that the Aussonderungsaxiom in S_2 becomes a single axiom (the intersection of a set and a class

([40]) See Tarski [2, p. 110], and Rosser-Wang [1, p. 128].

([41]) This section was added in April, 1952.

is again a set) involving a free class variable; S_2 is the partial system of the Neumann-Bernays system as determined by the axioms of the groups I–III and Va of Bernays [1], and S_6 is related to S_2 as the Zermelo-Fraenkel system is to the Neumann-Bernays.

(5) S_4 is roughly the ordinary Zermelo set theory (including, beyond the axioms of S_6, the axioms of infinity, power set, and sum set), and S_5 is related to S_4 as S_3 is to S_1. We note that in S_5 the Aussonderungsaxiom remains the same as in S_4 and no references to classes are allowed in defining sets (or elements). S_5' is a further extension of S_5 where the restriction is removed and the new Aussonderungsaxiom states (as in S_2) that the intersection of a set and a class is again a set.

(6) R is a system which is formulated in the notation of ordinary Zermelo set theory but is as strong as an extension of S_1 related to S_1 as S_2 is to S_6 or, alternatively, as a third order predicate calculus with only predicative classes on the highest level. An important feature of R is that it contains only a finite number of proper axioms (i.e., axioms beyond quantification theory). R' is an extension of R related to R as S_1 is to S_3; $R\#$ is obtained from R by adding Con (R) as a new axiom.

We note that S_2 is related to S_6 in the same way as the Neumann-Bernays system (cf. Bernays [1], to be referred to as the system NB) is to the Zermelo-Fraenkel (obtained from S_4 by adding the Ersetzungsaxiom and the Fundierungsaxiom, to be referred to as the system ZF). When two systems S'' and S are related in the same way as NB is to ZF, we say that S'' is a predicative extension of S. Thus, Theorem II (in §2) and the results 4.17–4.21 (in §4) are concerned with the relations between systems and their predicative extensions.

S_1 and S_3, S_5 and S_4, R' and R are all related to each other in the same way. When S' and S are thus related, we say that S' is an impredicative extension of S. Thus, S_1 is an impredicative extension of S_3, S_5 is one of S_4, R' is one of R. The main results in this paper (including the Theorems I, IV, V, VI) are all concerned with the relations between systems and their impredicative extensions. The interest and validity of these results depends largely on the relative consistency of a system and its impredicative extension, first established by the present author (see Wang [3] and Wang [4]). It seems proper to say that the relative consistency of a system and its predicative extension is much less surprising than that of a system and its impredicative extension. Indeed, the former seems to have been widely accepted even before rigorous proofs by Dr. Novak and others appeared; and usually we assume that Con (S) cannot be proved in a predicative extension S'' because impredicative classes are needed, even when we still have no exact formalization of the matter. On account of these circumstances, we believe that the main results of the present paper have no analogues in studies where merely relations between a system and its predicative extension are considered.

This leads us immediately to the question as to the relationship between the results in the present paper and those obtained in Mostowski$_1$ which are concerned exactly with such relations. Thus, to make things definite, Mostowski confines his attention to the system ZF and its predicative extension NB, and proposes to prove in Mostowski$_1$ the following three theorems. (MI) This contains two parts: (a) there is a predicate Tr in NB such that if $\mathfrak{m}$ is the Gödel number of an arbitrary statement $F(\mathfrak{m})$ of ZF, we can prove in NB: $Tr(\mathfrak{m}) \equiv F(\mathfrak{m})$; (b) if $F(\mathfrak{m})$ is a theorem of ZF, then $Tr(\mathfrak{m})$ is provable in NB but (if NB is consistent) the following general theorem is not provable in NB: $\forall m((m$ is the Gödel number of a theorem of ZF$) \supset Tr(m))$. (MII) There is a definite sentence $H(x)$ of NB such that (if NB is consistent) we can prove in NB both $H(1)$ and $\forall n(H(n) \supset H(n+1))$, but not $\forall n H(n)$. (MIII) There is a definite expression $H'(x)$ of NB such that (if NB is consistent) we cannot prove in NB: $\exists X \forall x(x \eta X \equiv H'(x))$. Since we know that the relative consistency of NB to ZF can be proved in NB and therefore that the consistency of ZF cannot be proved in NB, the above three theorems are intended to explain more precisely why Con (ZF) is not provable in NB.

(MIa) is similar to our Theorem II (in §2) and calls for a similar proof. (Compare also footnote 5 of Wang [3] where the possibility of a theorem like (MIa) was remarked.) But our theorem is stronger and Mostowski does not seem to realize(42) that sometimes we can find in S' a truth definition for S although S' is weaker than S in the usual sense of "being weaker than." Our discussion in §§2 and 5 stress, among other things, this point.

(MII) is an analogue of our Theorem VI (in §5). (Compare also the first half of Theorem 9 in Wang [3].) But they are concerned with different systems and call for completely different proofs.

These observations should suffice to dispel the doubts which Mostowski has expressed(43) regarding the validity of the results summarized in Wang [3] and to explain the extent to which the results of Mostowski$_1$ are similar to those of Wang [3] and the present paper. With regard to the few analogous results, Mostowski graciously credits(44) priority to me, but I believe that he probably reached his conclusions(45) at nearly the same time as Professor Rosser and myself.

Among the systems we tabulated under (1)–(6), the relation between S_5' and S_4 is again of a different sort. S_5' differs from a predicative extension of S_4 in that it contains in addition also the impredicative classes; while it

(42) See the sentence in lines 10–12 on p. 123 of Mostowski$_1$.

(43) See lines 20–18 from bottom on p. 142 and lines 12–21 from top on p. 143 of Mostowski$_2$.

(44) Ibid., line 24.

(45) Professor Mostowski's explanation, in the middle of Mostowski$_1$, p. 118, of the reason why the consistency of ZF is not provable in NB is in error. Since this mistake in exposition has led to serious misunderstandings, he is planning to publish a note of correction in Fundamenta Mathematicae.

differs from the impredicative extension S_5 of S_4 in that class variables are allowed in defining sets. In other words, it actually contains all the axioms of both extensions of S_4. When S''' and S are related in such a way, we shall say that S''' is an *irreducible* extension of S. Thus S_5' is an irreducible extension of S_4.

As we have mentioned above (cf. 4.13 in §4), the consistency of S_4 is provable in S_5'. Moreover, since we can prove in S_5' that there exists an inductive set (cf. remark after Theorem V in §5)([46]), and that the intersection of a class and a set is again a set, the intersection of all inductive classes of S_5' is the same as the intersection of all the inductive sets of S_5'. Therefore, if we define in S_4 the set Nn of all natural numbers as the intersection of all its inductive sets and use the same formal definition in S_5', we can in S_5' still make induction on all classes and therefore all sentences of S_5' (compare the remarks about R5 in §5). Hence, by arguments similar to those for 5.10 (in §5), we can derive Con ($S_4\#$) from Con (S_4) in S_5', where $S_4\#$ is related to S_4 as $R\#$ is to R. Hence, Con ($S_4\#$) is also a theorem of S_5'. But we know (compare the proof of 5.11 in §5) that we can also derive Con (S_5) from Con ($S_4\#$) in S_5. Therefore, Con (S_5) is also provable in S_5' and Con ($S_4\#$) is not provable in S_5. Hence, S_5' is demonstrably stronger (in the sense of Wang [4]) than S_5. (Compare Theorems 6, 11, and 12 of Wang [3].)

Let NQ be the system obtained from NB by adding all the impredicative classes. Then NQ is an irreducible extension of ZF, related to ZF as S_5' is to S_4. By reasoning similar to those in the preceding paragraph, we can prove Con (ZF) and Con (ZF#) in NQ (ZF# is to ZF as $R\#$ is to R). Since we know that Con (ZF) cannot be proved in NB, it also follows that there must be certain impredicative classes which cannot be proved to exist in NB. Indeed, since the relative consistency of NB to ZF can be proved in NQ, Con (NB) can also be proved in NQ([47]). It is not clear whether Con (NB) might also be provable in the impredicative extension of ZF, which is demonstrably weaker than NQ (just as S_5 is weaker than S_5').

The most common examples of irreducible extensions seem to be the cases where S is the ordinary nth order (n being 2, 3, 4, $\cdots$) predicate calculus and S''' is the $(n+1)$th. If we take natural numbers as the individuals (the entities of the first or lowest type) of these systems, then we see that what we have said about S_4, S_5, S_5' all apply (mutatis mutandis) to the systems S, S' (the impredicative extension of S) and S''', respectively.

A similar but slightly different case is the following. Let R^* be the system obtained from R' by adding the new axiom R5 (stating that every class of

([46]) This of course depends on the definitions of zero and the successor function, and the particular form of the axiom of infinity. For instance, if we use the original Zermelo axiom of infinity and define zero and the successor function as in 2.1 and 2.2 (of §2), then the sum set of the postulated infinite set is an inductive set.

([47]) This possibility is asserted in Mostowski[1], last footnote on p. 113, without proof.

natural numbers is equivalent to a set of natural numbers, see §5), then we see that R^* is again related to R and R' as S_5' is related to S_4 and S_5.

BIBLIOGRAPHY

PAUL BERNAYS

1. *A system of axiomatic theory*, J. Symbolic Logic, Part I, vol. 2 (1937) pp. 65–77; Part II, vol. 6 (1941) pp. 1–17.

KURT GÖDEL

1. *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme* I, Monatshefte für Mathematik und Physik vol. 38 (1931) pp. 173–198.

DAVID HILBERT and PAUL BERNAYS

1. *Grundlagen der Mathematik*, vol. 1, 1934, vol. 2, 1939, Berlin, Springer.

LEON HENKIN

1. *The completeness of formal systems*, a Thesis at Princeton University accepted in October, 1947. (Abstract appeared on p. 61 of J. Symbolic Logic vol. 13 (1948)).
2. *Completeness in the theory of types*, J. Symbolic Logic vol. 15 (1950) pp. 81–91.

A. MALCEV

1. *Untersuchungen aus dem Gebiete der mathematischen Logik*, Recueil Mathématique N.S. vol. 1 (1936) pp. 323–336.

JOHN VON NEUMANN

1. *Zur Einführung der transfiniten Zahlen*, Acta litterarum ac scientiarum Regiae Universitatis Hungaricae Francisco-Josephinae, Sectio scientiarum mathematicarum vol. 1 (1923) pp. 199–208.

ILSE L. NOVAK

1. *A construction for models of consistent systems*, Fund. Math. vol. 37 (1950) pp. 87–110.

JOHN G. KEMENY

1. *Type theory vs. set theory*, a Thesis at Princeton University (1949). (Abstract appeared on p. 78 of J. Symbolic Logic vol. 15 (1950)).

W. V. QUINE

1. *Mathematical logic*, New York, 1940; second printing, Cambridge, Mass., 1947.
2. *Element and number*, J. Symbolic Logic vol. 6 (1941) pp. 135–149.

J. BARKLEY ROSSER and HAO WANG

1. *Non-standard models for formal logics*, J. Symbolic Logic vol. 15 (1950) pp. 113–129.

ARNOLD SCHMIDT

1. *Über deduktive Theorien mit mehreren Sorten von Grunddingen*, Math. Ann. vol. 115 (1938) pp. 485–506.

THORALF SKOLEM

1. *Über einige Grundlagenfragen der Mathematik*, Skrifter utgitt av Det. Norske Videnskaps-Akademi, I, no. 4, 1929, 49 pp.
2. *Über die Nicht-charakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen mit ausschliesslich Zahlenvariablen*, Fund. Math. vol. 23 (1934) pp. 150–161.

ALFRED TARSKI

1. *Der Wahrheitsbegriff in den formalisierten Sprachen*, Studia Philosophica vol. 1 (1936) pp. 261–405. (Original in Polish, 1933.)
2. *On undecidable statements in enlarged systems of logic and the concept of truth*, J. Symbolic Logic vol. 4 (1939) pp. 105–112.
3. *The semantic conception of truth and the foundations of semantics*, Readings in Philosophical Analysis, selected and edited by Herbert Feigl and Wilfred Sellars, New York, 1949, pp. 52–84. (Original appeared in Philosophy and Phenomenological Research vol. 4 (1944).)

HAO WANG

1. *A new theory of element and number*, J. Symbolic Logic vol. 13 (1948) pp. 129–137.
2. *On Zermelo's and von Neumann's axioms for set theory*, Proc. Nat. Acad. Sci. U.S.A. vol. 35 (1949) pp. 150–155.
3. *Remarks on the comparison of axiom systems*, ibid. vol. 36 (1950) pp. 448–453.
4. *Arithmetic translations of axiom systems*, Trans. Amer. Math. Soc. vol. 71 (1951) pp. 283–293.

ZÜRICH, SWITZERLAND