## Chapter I

## FORMAL SYSTEMS AND FORMAL REASONING

In this chapter the notion of formal system, defined tentatively in the introduction, will be discussed more in detail. The definition in the opening paragraph is illustrated by an example in the second. The later paragraphs contain discussion of various matters connected with the definition, including certain ways of modifying it, and also its relation to language. For this purpose it is necessary to introduce and define a number of notions related to symbolism as such. This terminology throws some light on formal systems and their nature (§7). The chapter closes with a statement of the assumptions and notation on which the later chapters are based.

1. **Definition of a Formal System.**[1] A formal system is defined by a set of conventions called its primitive frame, specifying the following:

I - A set of objects with which the system has to deal. These objects will be called its terms.

II - A set of propositions, called elementary propositions, concerning these terms.

III - Which of the elementary propositions are true - i.e., are theorems.

The specifications in each of these three divisions has a recursive character. More in detail, we have specifications as follows:

I. TERMS
- A. Primitive Terms - a finite or infinite list.
- B. Rules of Term Formation - i.e., rules for forming further terms by means of specified (primitive) operatives.[2]

II. ELEMENTARY PROPOSITIONS. Rules for forming these from primitive terms by means of specified (primitive) predicates.

III. THEOREMS
- A. Axioms - That is a set of elementary propositions stated to be true outright.

1. The definition given here is that of [16] with some minor changes.

2. The word "operative" is used rather than "operation" so as to be consistent with §5.

B. Rules of Derivation - rules specifying how the elementary theorems are to be derived from the axioms.

It is then understood that every term is a primitive term or is obtained from primitive terms by a finite number of operations according to the rules I B; every elementary proposition is formed by applying a predicate to a sequence of the proper number (and perhaps kind) of terms; every elementary theorem is either an axiom or is derived from the axioms by use of a finite number of the rules of derivation.

It is required further that the rules I, II, and III A be definite, in the sense that, in any given case, whether something is a term, elementary proposition, or axiom, can be decided by a finite process; further, in the first two categories a unique method of construction is determined for each case. On the other hand it need not be a definite question whether an elementary proposition is true; but it must be definite whether a proposed derivation of an elementary theorem is or is not correct. (Systems in which this requirement of definiteness is relaxed somewhat have been considered, and may be admitted as indefinite formal systems, but these will not be considered here, except incidentally.) A system in which the concept of truth for the elementary propositions is definite will be called a decidable system.

It is convenient to call the considerations based on I and II alone the morphology of the system, while those based on III are called the theory proper. Thus the morphology of a formal system consists of the rules for determining the elementary propositions. This morphology may be quite complicated; it may be necessary to distinguish several - even infinitely many - categories of terms, relations between terms, and what not. It may also be necessary to introduce auxiliary concepts of a similar nature in stating the theoretical rules; in such cases these auxiliary notions will also be considered as morphological. All this is quite legitimate provided the requirements for definiteness are fulfilled.

The constituents of the primitive frame may also be classified as follows:

1. Primitive Ideas
   a. Primitive terms
   b. Primitive operatives
   c. Primitive predicates
   d. Auxiliary morphological notions
2. Postulates
   a. Morphological rules
   b. Axioms } Theoretical rules.
   c. Rules of derivation

Note that nothing excludes the possibility that there may be infinitely many items under any or all of these heads, provided the requirements for definiteness are satisfied. Infinitely many axioms are frequently expressed by axiom schemes[3] as in the example below.

In the following we shall be concerned with systems which have the following special character. Let $P_1, P_2, \ldots, P_m$, Q be structural characterizations of elementary propositions formed from unspecified terms $X_1, X_2, \ldots, X_n$. Each rule then says: if $X_1, \ldots, X_n$ are determined to be terms such that $P_1, P_2, \ldots, P_m$ are true, then Q is true for these same $X_1, \ldots, X_n$. If all rules have this character, then an elementary proposition B is true if and only if there exists a finite sequence of elementary propositions $A_1, A_2, \ldots, A_n$, such that B is $A_n$, and every $A_k$ is either an axiom or is a consequence of a certain subset of its predecessors by one of the rules.

2. **An Example from Group Theory.** We shall take a formalization of Dickson's postulates for groups.[4] The primitive frame of this system is as follows:

I. TERMS
   A. Primitive Terms.
      An infinite sequence $a_0, a_1, a_2, \ldots$
   B. Rules of Term Formation. There are two primitive operatives:
      1. A binary operative forming from any two terms A and B a term $A \circ B$.
      2. A unary operative forming from a term A a new term $A'$.

II. ELEMENTARY PROPOSITIONS. There is one binary predicate (equality) forming from two terms A and B an elementary proposition A = B.

III. THEORETICAL RULES
   A. Axiom Schemes. For any terms A,B,C, we have:
      1. $A \circ (B \circ C) = (A \circ B) \circ C$.
      2. $A \circ a_0 = A$.
      3. $A \circ A' = a_0$.
   B. Rules of derivation. For any terms A,B,C we have:
      1. If A = B, then B = A.
      2. If A = B and B = C, then A = C.
      3. If A = B, then $A \circ C = B \circ C$.
      4. If A = B, then $C \circ A = C \circ B$.

This system has infinitely many primitive terms; two primitive operatives, one binary and one unary; one primitive predicate, a binary one. It has a very rudimentary morphology, since there is only one category of terms and no auxiliary

3. A term due to von Neumann [66]; cf. Post [69]§10, Lemma 2.
4. See Dickson's original paper [28].

morphological notions.[5] There are three axiom schemes, each containing infinitely many axioms. The rules are the usual ones for equality. One of the primitive terms, $a_0$, appears explicitly in the axiom schemes 2 and 3, hence it plays the role of a constant; while nothing can be said in an elementary theorem about one of the other terms that cannot be said about any term whatever. It is worth noting that although the above system is abstracted from Dickson's postulate set, none of the four original postulates is an elementary proposition. Three of the postulates turn up as axiom schemes with infinitely many axioms, while the fourth postulate is our first rule of term formation.[6]

The elementary theorems of the system are the equations which hold in any group identically in $a_1, a_2, \ldots$. Examples are:

$$a_1 = a_1$$

$$a_1' o a_1 = a_0$$

$$a_0' = a_0$$

$$(a_1 o a_2)' = a_2' o a_1' \quad \text{etc.}$$

It should be noted that the proof in full of even the second and third of these is quite tedious.

The notion of epitheorem was defined in the introduction. Examples of this notion are the following: If A is any term, then

$$A = A,$$

$$A' o A = a_0 .$$

Further, if A and B are any terms such that A = B then A' = B'.[7] The statement made above to the effect that $a_1, a_2, a_3, \ldots$ function as variables - i.e., that any elementary theorem is valid if arbitrary terms A and B are substituted for $a_1$ and $a_2$ - is also an epitheorem; although its statement and proof requires a recursive definition of substitution and an argument by induction.[8]

3. **Representation and Interpretation.** The primitive frame says nothing about what the terms are. It postulates, so to speak, a class of terms, but leaves them completely undetermined

---

5. In fact it is completely formalized in the sense explained below.

6. Of course the existential aspect of some of Dickson's postulates is not exactly reflected here; it is taken care of by our postulation of $a_0$ as the primitive term and the inverse of a term as a primitive operative, together with the second rule of term formation.

7. Note that these can be expressed as compound propositions by the universal quantifier, e.g.,

$$(x) \cdot x = x$$

$$(x)\ (y) \cdot x = y \cdot \supset \cdot x' = y'$$

8. Cf. [16]; also Chapter III below.

A formal system so considered will be called abstract. If, on the other hand, we specify that the terms are such and such objects, and if the specification is such that there is a distinct object for each primitive term and for each process of construction out of the primitive terms,[9] then we shall say that we have a represented formal system, or a representation of the formal system in the objects concerned.

Given an abstract formal system, one can find a representation for it in infinitely many different ways. One can even find a representation in the "expressions" of any "language" containing at least two distinct symbols. Thus consider the two Greek letters lambda and mu. Let the expressions be the linear rows formed from them, such as the following four examples:

$$\lambda\mu$$
$$\lambda\mu\mu$$
$$\lambda\lambda\mu\mu\lambda\mu\lambda\lambda\mu\lambda\mu\mu\mu$$
$$\lambda\lambda\lambda\mu\lambda\mu\lambda\lambda\lambda\mu$$

Then we can represent our system of group theory in these expressions as follows: let $a_n$ be the expression formed by writing a lambda followed by $n+1$ mu's; given terms A and B, let AoB be the expression formed by writing two lambdas, then two mu's, then A, then B; while A' is formed by writing two lambdas, then one mu, then A. Thus in the above array the first three expressions are $a_0, a_1, a_0 \circ a_2'$, respectively, while the fourth expression is not a term. Then it may be shown that we have a representation in the above sense.

In a represented formal system the criterion of truth for the elementary propositions is the same as in the corresponding abstract one. Thus, in the above representation of group theory, the relation of equality may be thought of as a class of pairs of expressions which is defined completely and solely by the system itself.

In contrast with representation stands the notion of interpretation. We say we have an interpretation when we put not only the terms but the predicates into correspondence with an external reality. Thus if we say that $a_0$ is the identity transformation while $a_n$ is the transformation $x' = x+n$, and say that AoB and A' are the product of the transformations A and B and the inverse of A respectively, then this is not a representation because distinct processes of construction from the primitive terms correspond to the same transformation. But in this case it is true that whenever A = B, the terms A and B correspond to the same transformation. Thus if we interpret equality as identity of the corresponding transformations, we do have an interpretation and, as we say, a valid one. An interpretation is valid if the

9. In the terminology of § 5 below, if there is a distinct term for each distinct term noun.

intuitive predicate holds whenever the formal one does.

**4. Some Semiotical Concepts. The U-language.** In recent years there has arisen a school of thinkers who maintain that many problems can be solved, or at least illuminated, by examining critically the language in which they are expressed. This has led to a whole science of symbolism which Morris calls semiotics.[10] It is profitable to look at a formal system from that point of view. But in order to do so it is necessary to explain some semiotical concepts first, particularly since writers on this subject are not completely consistent in terminology. This digression will occupy the next two sections;[11] we shall return to the discussion of formal systems in §6.

By a language is meant any system of objects, called symbols, which can be produced in unlimited quantity, like the letters of ordinary print or the phonemes of speech, and combined with one another to form combinations called expressions.[12] Thus symbols and expressions are undefined, but they are defined for any given language by the conventions of that language. Generally the expressions are arbitrary linear series of symbols in which repetitions are allowed; in that case the language will be called a linear language. It is irrelevant whether or not the language is used for communication purposes. If it is so used it will be called a communicative language. A language in the sense of linguistics will be called a natural language.

Whenever we talk about a language it is said that we must do so in a second language. In that case the first language is called the object language, the second the metalanguage. If then one talks about the metalanguage, one has to do so in a third, meta-metalanguage and so on. But this way of speaking ignores one particular fact: viz., that any investigation, whether semiotical or not, has to be conducted, not in an arbitrary metalanguage, but in the communicative language which is mutually understood by speaker and hearer. I shall call this language the U-language, i.e., the language being used.

This U-language has the following characteristics: 1) It is specific. It is not enough to say that the U-language for these lectures is English; for they would be quite unintelligible to many who know English well, and there are doubtless lectures going on in English at this moment which none of us here can make

10. See [64] and [65]. The word is used in similar, but often somewhat different senses, by various other writers.

11. The subject matter of these sections was also dealt with in my paper [20]. I have taken some statements verbatim from that source, but I have made both additions and omissions due to the difference in emphasis.

12. It is not necessary here to go into the distinction between expression - designs and expression - events. For this see Carnap [6] §3. I shall use "symbol" and "expression" in the design sense.

head or tail of. Moreover, there are no if's, and's or but's about it; whether it does or does not contain, for instance, means for expressing modalities, is a question of fact. 2) It is a growing thing. As we proceed we shall modify, add to, and refine it. These changes, however, are gradual, so that the language preserves its identity, in a certain sense, throughout.[13] 3) There is always vagueness inherent in it; but we can, by skillful use, obtain any desired degree of precision by a process of successive approximation. 4) We can never transcend it - whatever we study we study by means of it. 5) It cannot be exhaustively described, and it can lead to contradiction if carelessly used.

This notion of U-language corresponds to what we mean when we say "ordinary language." The point is that everything we do depends on the U-language. Formal systems, semiotical systems, all systems have to be explained in the U-language and studied in the U-language. Moreover, we can significantly make statements about the U-language within the U-language.

It follows from the foregoing that there is no such thing as a meta-U-language. But we can take some segment, say L, of the U-language, isolate it, and introduce into the U-language technical terminology for referring to L as object-language. This technical terminology, together with such symbols from U as we need to make the statements we want, form a segment M of U which we can isolate again, and so on. Such an M is what I shall call a metalanguage over L. The process of setting up metalanguages I shall call metasemiosis; it is a powerful aid to clarification of U. But the language used is the so modified U-language, the new terminology in M being defined partly by reference to L, partly recursively.

5. **An Attempt at Mathematical Grammatics.**[14] The semiotical study of formal systems is greatly facilitated if we have suitable terminology for dealing with symbols and combinations of symbols, and the various grammatical distinctions between them. An attempt at devising a suitable terminology will be made here.[15]

So far as symbols and expressions are concerned it is now standard practice to use an expression in quotation marks as a name for that expression.[16] Thus "λ" is Greek lower case lambda,

---

13. One can object that the notion of U-language is still somewhat vague. But one can introduce further refinements as necessary. Thus it will be specified below that the phrase "U-language" shall mean the U-language before the A-language is adjoined to it. If further precision is necessary we can make additional stipulations. The situation is, in a way, analogous to that in physics. The physicist does not know his experimental constants absolutely; he knows them to such accuracy as serves the purpose.

14. The word "grammatics" was introduced in this sense in [21].

15. Cf. Ajdukiewicz [1].

16. This usage is ascribed to Frege. However this usage occurs, in combination with italics, in [85], and doubtless even earlier.

a "hat" is a word of three letters, but a hat is a covering for the head, etc. This is a real help in avoiding the confusions of the "autonymous mode of speech"[17]- i.e., the usage where a specimen of an expression is used as a name for that expression without typographical indication - when the expression has a meaning of its own in the U-language. But if one were mentioning some Russian words, written in the Cyrillic alphabet, in an English context, the autonymous mode of speech would be quite harmless. I shall use quotation marks when I think they add to clearness, but not where other standard conventions of the U-language are sufficient.

In a communicative language, however, the forming of symbols into expressions is not of much significance. The significant combinations are another class of combinations which I shall call phrases. Thus in the sentence

"I see both red and blue dahlias"

the following are samples of expressions:

"see both red"
"and blue"
"th re"

while the following are phrases:

"I see"
"both red and blue"
"both...and"

where in the last one the dots indicate a blank to be filled. This last example shows that phrases are not a subclass of expressions (if the language is regarded as linear).[18]

We shall need a classification of phrases according to their grammatical nature. The primary classification is into nouns, sentences, and functors.[19] Nouns and sentences together may be called closed phrases in contrast to functors. A noun names something,[20] a sentence makes a statement. As to functors, these are phrases which combine other phrases. A functor combines one or more phrases, called its arguments, to form a phrase called its value. To get complete generality we must admit functors whose arguments are other functors. As for the value

---

17. This term was introduced by Carnap [8], which should be consulted in connection with the possibilities of confusion. The English edition is [7]. For changes in Carnap's position see the Appendix to [6].

18. Such phrases consisting of detached parts are particularly common in German. Cf. Ajdukiewicz [1].

19. The name "functor" is said to be due to Kotarbinski [55] but I have never seen his work. See Tarski [82] footnote 7, p. 274. Carnap [6] uses "functor" in the sense of what is below called "nominal functor."

20. Thus "noun" has a different meaning from what it does in linguistics. Perhaps "name" would be better, but that has difficulties also.

we have two choices: either we can require all functors to have one argument and allow the value to be a functor,[21] or we can admit functors of several arguments and (without loss of generality) require the value to be closed. We shall do the latter and we shall also use the word "closure" for the value. Then functors may be classed as nominal or sentential according to the nature of the closure. This classification can be extended to phrases, a nominal phrase being either a noun or a nominal functor. As to the arguments we can distinguish functors as junctors if all arguments are nominal, and as nectors if at least one is sentential; in the latter case it is a pure nector if all arguments are sentential, otherwise mixed.[22] Various other classes of functors will be defined according to the scheme in Table 2. Names in parentheses are not used here and are to be regarded as tentative. In any of these classes we distinguish primary and secondary functors; a primary functor is one all of whose arguments are closed, a secondary functor is one for which at least one argument is a functor.

For the more important types of phrases a convenient terminology is exhibited in Table 3. Here the name for the phrase itself is given in Column 1. In Column 2 there is a name for the significance[23] of the phrase. In Column 3 there is listed, for functors only, a name for the significance of the combination of the phrase and its argument. This last I have found convenient in discussing structure. Thus a predication is a proposition, but not every proposition is a predication - it may be a connection. Where two or more names appear in Columns 1-3 they are alternatives. In the fourth column there are some examples from ordinary mathematics; here numbers are taken as terms, so that an ordinary functional sign is an adjunctor. These examples are given under the narrowest category which applies.

In regard to these tables the following remarks are relevant. First, in many languages there are phrases which belong to more than one category. Thus quotation marks, in the above usage, are a functor which converts an expression of any kind into a noun; when applied to a sentence they are a subnector, when applied to a noun, an adjunctor. Second, no attempt has been made to take account of the difference between variables and constants. That is another cross-classification of phrases. From the present viewpoint "x" as used in elementary algebra is as much a noun as "0", and "x = y" as much of a sentence as "1 = 2". It is not denied that there is a distinction between phrases containing variables and those which do not. Third, the categories given

21. This is essentially done in grammar; it is also the idea back of the Schönfinkel application operator (cf. [74], [17]).

22. The words "junctor" and "nector" are suggested by the use of "junction" and "nexus" in Jespersen [50]. However, the usage of the terms is quite different.

23. In the sense of Morris's "significatum." It is unnecessary to go into a precise analysis of this concept here.

Table 2

CLASSIFICATION OF FUNCTORS

| Arguments \ Value | Phrase | Noun | Sentence |
|---|---|---|---|
| Phrase | Functor | Nominal Functor | Sentential Functor |
| All Nominal | Junctor | Adjunctor or Operator | Predicator |
| Some Sentential | Nector | Subnector | (Pronector) |
| Mixed | Mixed Nector | (Mixed Subnector) | (Adnector) |
| All Sentential | (Pure Nector) | (Pure Subnector) | Connector |

here may require subdivision in other respects. Fourth, auxiliary symbols, such as parentheses, punctuation marks, and some particles have not been explicitly provided for. In general these serve to show how symbols are to be combined, and so are parts of functors, but this may not account for all uses. Fifth, it is admitted that there are great difficulties in applying these notions to the natural languages, and the matter is worthy of further study. Sixth, it is not possible to be completely consistent with ordinary terminology, because that terminology is not always consistent with itself.[24]

There is a need for a notation for referring to phrases which is as efficient as quotation marks for referring to expressions. This is particularly so in the case of functors, which in mathematics frequently consist of detached parts. A functor is naturally regarded as comprising all the auxiliary

24. The relations between the various categories, which are quite complex, can be formalized by the theory of combinatory functionality. See [17]§6 and [15]§2. Cf. also Ajdukiewicz [1].

Table 3

| 1<br>Name of Phrase | 2<br>Significatum | Closure | Examples |
|---|---|---|---|
| Phrase | Concept | | |
| Noun | Term,Entity,Object | | $0, \pi, x$ |
| Sentence | Proposition,Statement | | $x = y$ |
| Functor | Functive | Function | |
| Junctor | Junctive | Junction | |
| Nector | Nective | Nexus | |
| Adjunctor,Operator | Adjunctive,Operative | Adjunction,Operation | d/dx |
| do.,primary | do.,primary | do.,primary | $\vert --- \vert$, $(---_1)+(---_2)$ |
| Predicator | Predicate | Predication | ---is continuous,<br>$---_1$ converges to $---_2$ |
| -do.,primary | -do.,primary | -do.,primary | = , < , is a prime, $\vdash$ |
| Subnector | Subnective | Subnexus | the x such that |
| Adnector | Adnective | Adnexus | uniformly |
| Connector | Connective | Connection | and,or,(x). |

symbols - commas, parentheses, etc., necessary to symbolize the closure except for the insertion of the arguments. This means there must be blanks to show where the arguments are inserted. Modifying the usage of Quine[25] I shall indicate blanks in functors thus: $---_1$, $---_2$, $---_3$, etc., the subscript to be attached when there is more than one argument being used to show where the different arguments go. Thus addition, absolute value, equality, and (one notation for) the ordered triple would be: "$(---_1)+(---_2)$", "$|---_1|$", "$(---_1)=(---_2)$", and "$[---_1, ---_2, ---_3]$" respectively. These are to be understood as referring to the functors, not the expressions containing dots and subscripts.[26] Sometimes the blanks may be omitted, but in such cases they are understood. I shall use this notation without the quotation marks when referring to the functives.

When these functors are in the U-language, or some segment of it, there are standard conventions in regard to parentheses which apply. This takes care of primary functors in ordinary mathematics. That is sufficient for our present purposes.

6. **Semiotical Aspect of a Formal System. The A-Language.** Let us return to the notion of a formal system and look at it frankly from a semiotical point of view.

Although we do not say what the terms are, yet in setting up the primitive frame we do introduce a notation for them and for certain notions connected with them. In the system of § 2 this new notation consists of the following:

nouns: "$a_0$", "$a_1$", "$a_2$", ...
binary adjunctor: "$(---_1) \circ (---_2)$"
unary adjunctor: "$(---)'$"
binary predicator: "$(---_1) = (---_2)$";

All other phrases used in the primitive frame are in the U-language and have significance there. This applies to the letters "A", "B", "C", etc., which are used as variables in the U-language and will therefore be called U-variables; also the words like "proposition," "operative" (="adjunctive") etc.

Now the terminology so introduced is intended to be adjoined to the U-language and used therein. Thus when the frame says that $a_0, a_1, \ldots$ are terms, that is tantamount to saying that "$a_0$", "$a_1$", ... are to be members of a new category of nouns – call them term-nouns – in the U-language. The rest of the frame says in effect that the adjunctors form term-nouns from term-nouns, and the predicators sentences from the term-nouns. This introduction of symbols into the U-language, without saying what

25. See [71]. The usage already occurred in [72].

26. The formalism of Church [13], or some other notation involving apparent variables, can also be used, but these notations are thought of as ways of representing functives rather than functors.

they signify, is the characteristic semiotical process in mathematics.[27]

The terminology so adjoined to U forms a language I shall call the A-language. This A-language is sufficient to express the elementary propositions of the system; but it cannot express any epitheorems, or even its own rules. These require that the A-language be embedded in the rest of U. In the following the phrase "U-language" shall mean the U-language before introduction of the A-language.

A formal system frequently originates in the following way. We start with an unformalized theory expressed in a language L, which is a segment of the U-language. By making assumptions explicit, and similar analysis, we transform the theory so that L, or some modification of it, becomes the A-language of a formal system. In that case we say we have formalized the theory.

This sense of formalization is to be contrasted with another which has been emphasized by Carnap In this case we form a metalanguage M over L. This M is so formed that the criterion of truth in M is based on recursive definitions and reference to the L-expressions regarded as physical shapes (or structures) only. In that case I shall say M is syntactical[28] with respect to L. Such an M-language is not quite formal in our sense, because it may be necessary to inspect the L-expressions. However, it has been shown by Hermes and others[29] that a formalization (in our sense) can be accomplished. When it is obtained we end up with a situation where M, rather than L, is the A-language of a formal system. If this process is applied to a case where L is already formal (in our sense), then it will turn out, since L-truth is dependent only on recursive definitions, that it is definable in M. Thus the two formalizations will be equivalent to one another.

This discussion shows that the two concepts of formalization, - the abstract way, by means of a formal system, and the syntactical

---

27. In ordinary mathematics we frequently do not distinguish sharply between symbols used and symbols mentioned. Thus we say that x occurs in such and such an expression. But this is only an appearance; it can be avoided by the methods of Chapter III below. Such usage occurs because certain elements in the structure of our mathematical expressions reflect a structure in the concepts symbolized. We never say, for instance, that "x" is composed of two crossed lines. In mathematics we do not talk about our symbols, we use them.

28. This is Carnap's "formal"([6],p. 10). However, if M mentions the significata of certain L-expressions, but only in a manner which is analogous to the representation, as opposed to the interpretation, of a formal system, then I would still consider M as syntactical in the sense of the text; whereas Carnap apparently would not call it formal. Thus practically all of Carnap's semantics, which he contrasts with syntactics as being radically different, is syntactical in the present sense.

29. See Hermes [43]. Cf. also Tarski [77], Schröter [75].

way - are essentially equivalent. The main difference between them is that in one we end up with a formal system in which a form of the original L is the A-language, in another a totally different M is the A-language. The former method requires only one set of symbols, and these are more or less familiar; so that the symbolism is more perspicuous and closer to the interpretation in reality which was the reason for making the study in the first place. The second method requires a dual symbolism; but since the language used is a new one, pernicious linguistic habits are more effectively broken up; further it is concrete and agrees with certain philosophical doctrines. It is natural that one of these methods should appeal more strongly to mathematicians, the other to philosophers. But work is being done from both viewpoints, and the reader should understand both, and their relations to one another.

Since the A-language of a formal system is being used, and since the terms are undetermined, it is natural to use the terminology of the second column of Table 3 in referring to the significata. But this usage does not commit us to an idealistic philosophy. Thus, if one understands a proposition as what a sentence expresses, one is at liberty to say that a sentence expresses only itself. Thus a position of nominalism is quite compatible with this terminology. That an idealistic position is also compatible is quite clear. This flexibility in the "absolute" terminology is a good reason for using it.

Because a formal system is not a linguistic system, it is not necessary to be absolutely precise in distinguishing between use and mention. In this respect we are on a par with ordinary mathematics.

7. **Modifications of a Formal System.** The system of § 2 had the peculiarity that there was only one kind of term, that there were no auxiliary notions necessary for the rules, and that the morphology consisted solely of the grammatical character of the primitive A-phrases with respect to the categories term-noun and sentence. Such a system may be called <u>completely formalized</u>. It is plausible that any formal system can be reduced to that form by introducing additional primitive ideas so that the morphological and auxiliary notions can be expressed in the A-language.

Again a formal system can always be reduced to one in which there is only one primitive predicate and that a unary one. To do this, introduce for each primitive predicate a corresponding new primitive adjunctive and let there be a new primitive predicator "$\vdash$---" (as is customary). Then if A is an elementary proposition, let a be the corresponding term obtained by using the adjunction in place of the predication to which it corresponds. Then the proposition

$$\vdash a$$

can replace A.

If this reduction is carried out, then the terms arising from the original elementary propositions will constitute a new class of terms. Let us call these formulas. Generally it will be necessary to revise the morphology so as to treat these as a special category of terms. Thus if we apply this reduction to the example of § 2 with "$(---_1) \square (---_2)$" as the adjunctor replacing "$(---_1) = (---_2)$", then we have to have two categories of terms, basic terms with the same rules as the original terms, and formulas with the rule:

If A and B are basic terms, $A \square B$ is a formula.

Thus the original system is equivalent to a system with a single unary predicate and an additional category of terms. Further, various notions originally formulated in terms of propositions can be reformulated in terms of formulas. Thus we can think of axioms as a class of formulas, of rules of the kind mentioned at the end of §1 as relations of derivability between formulas, and so on. Indeed we can revise the whole conception of a formal system in this way. The conception is essentially that of Hilbert and his followers. It is particularly natural in the case of a syntactical theory, where the formulas are the sentences of the object-language; but it is somewhat artificial in connection with systems like that of § 2.[30]

All this suggests a reinterpretation of the word "proposition." That is, we could retain the name "proposition" instead of using "formula," adopting "elementary statement" for a concept

$$\vdash A$$

where A is a proposition. This is indicated in bringing the propositional algebra under the definition of §1, where it is desirable to use "proposition" for the terms of the system on account of the intended interpretation.

On the other hand one might go in exactly the opposite direction. Why not introduce primitive propositions[31] and primitive connectives, and rules of formation from elementary propositions out of simpler ones? If we do this for the propositional algebra we are led to a formulation in which there are no terms, which is akin to what Quine prefers. However, such a system can be brought under the concept of §1 by changing its ontology but not its mathematical content.

There are, thus, many different ways of conceiving a formal system, and a choice between them is somewhat arbitrary.

---

30. This way of looking at a formal system was mentioned by Kleene in his review [54]. It was also contained in [23]. In principle it is certainly much older.

31. By this I mean primitive ideas which are propositions, not postulates as in the Principia Mathematica [86].

8. **The System 𝔖.** In the sequel we shall suppose that we have to do with a formal system 𝔖 concerning which we shall make the following ASSUMPTIONS:

A1. 𝔖 is a formal system, according to the definition of §1, which is completely formalized. We shall however admit the possibility of primitive propositions in 𝔖 (Cf. close of §7.)

A2. The theoretical rules in 𝔖 are of the form stated at the end of §1; i.e., every application of a rule allows us to infer an elementary theorem from a finite number of elementary premises.

We shall also use the following conventions regarding NOTATION (a more detailed statement is made in III §2):

| Letters | Meaning |
|---|---|
| Latin l.c. | Terms. |
| Latin Caps. | Propositions, elementary or compound. |
| German l.c. | Classes or sequences of terms. |
| German Caps. | Classes or sequences of propositions. |
| Greek Caps. | Statements and classes of statements as defined later. |
| Roman Numerals. | Refer henceforth to the chapters in these lectures. |

Most of the symbols in this list will be used as U-variables. But the following will have constant significance.

| | |
|---|---|
| 𝔄 | The axioms of 𝔖. |
| 𝔈 | The elementary propositions. |
| 𝔖 | The system; also its elementary theorems. |
| 𝔓 | The propositions (elementary and compound) of 𝔖. |
| ℜ | The rules of 𝔖. |
| 𝔗 | Introduced as functor in II §8. |

Other constants will be introduced in Chapter III (see §2).

The notation

$$A_1, \ldots, A_m \vdash B$$

shall denote that B follows from A by a single application of a rule of derivation (for $m > 0$) or that B is an axiom (for $m = 0$).

We shall consider also two sorts of extensions of 𝔖. First we may adjoin additional axioms to 𝔖. If 𝔛 is a class or series of propositions, the system obtained by adjoining all the members of 𝔛 to the axioms of 𝔖 will be denoted by 𝔖(𝔛). Such an extension will be called a propositional extension.

Next we must make a special agreement in regard to the word "variable." The word is used in two senses. On the one hand a symbol of the U-language will be called a _U-variable_ if it is used in the U-language in the same way as variable letters are used in ordinary mathematics. On the other hand there are senses in which a term can be a variable. Let u be a term of $\mathfrak{S}$. If the primitive frame for $\mathfrak{S}$ specifies nothing about u, other than that it be a term, then u is an _indeterminate_ in $\mathfrak{S}$. Now let $\mathfrak{S}^*$ be a formal system which is exactly like $\mathfrak{S}$ except that $\mathfrak{S}^*$ contains infinitely many primitive terms not found in $\mathfrak{S}$. Then each of these additional primitive terms will be called a _term variable_. A term variable is then an indeterminate in $\mathfrak{S}^*$, but the converse does not necessarily hold, since there may be indeterminates in $\mathfrak{S}$. Note that if u is a term variable, the "u" is a U-constant - since it is in the A-language of $\mathfrak{S}^*$.[32]

Then we can define our second type of extension, the _term extension_ thus: Let $\mathfrak{x}$ be a class of term variables, then the $\mathfrak{x}$-extension of $\mathfrak{S}$, viz. $\mathfrak{S}(\mathfrak{x})$, is the system obtained by adjoining all the elements of $\mathfrak{x}$ to the primitive terms in $\mathfrak{S}$.

When we make either of these extensions, the elementary sentences in the A-language of $\mathfrak{S}^*$ express different propositions in the different extensions (since the truth-conditions are different). But in the following we shall speak of propositions so related as being the same proposition. This means that we are interpreting "proposition" as meaning a propositional function depending on the extension.[33]

In the following we shall construct formal systems which formalize the relations among the compound propositions of $\mathfrak{S}$. These systems will be called _episystems_. In the episystems,

32. When we introduce a term variable we are really adjoining a new symbol to the A-language as a primitive term noun. A requirement of our notion of formalization is that such a symbol be used as if it had a fixed (but unspecified) meaning. This hypothetical significatum of the new symbol is what we call the term variable.

Of course this notion is artificial. There is no such idea in ordinary mathematics (although the notion of indeterminate in abstract algebra is close to it). But it is no more artificial than the distortions to which we are forced in formalization by metasemiosis. I regard its artificiality not as evidence for the contention that in mathematics we are talking about symbols, but rather that the notion of variable is a derived notion, and can be eliminated in a really fundamental investigation. This has led to the theory of combinators [74, 17] and to the later, but apparently independent, development of the "algebra of analysis" by Menger [63].

For ordinary purposes I believe the autonymous mode of speech used by mathematicians generally involves no serious confusion. I shall not hesitate to use it in the more explanatory portions below. But in the more formal portions I shall carry through the conventions in the text.

33. Incidentally, this shows that the same sentence may express different propositions according to the context in which it is embedded.

much as in one aspect of ordinary propositional algebra, the propositions of $\mathfrak{S}$ will function as terms. We thus have two levels of formalization; that of $\mathfrak{S}$ itself and that of the episystems. On the $\mathfrak{S}$ level propositions are what they are stated to be; on the episystem level they are terms - perhaps they had better be called formulas - while their place as significata of U-sentences is taken by what we shall call *statements*. Again a functive like $---_1 \dot{\supset} ---_2$ is a connective on the $\mathfrak{S}$ level, but becomes an adjunctive when on the episystem level. To avoid confusion in this situation we shall use the terminology of Table 3 with reference to the $\mathfrak{S}$-level, except that "statement" and certain categories expressly stated will refer to the episystem level.