# International Library of Psychology Philosophy and Scientific Method

GENERAL EDITOR: C. K. OGDEN, M.A. (*Magdalene College, Cambridge*)

| | |
|---|---|
| PHILOSOPHICAL STUDIES | *by* G. E. MOORE, Litt.D. |
| THE MISUSE OF MIND | *by* KARIN STEPHEN |
| CONFLICT AND DREAM* | *by* W. H. R. RIVERS, F.R.S. |
| TRACTATUS LOGICO-PHILOSOPHICUS | *by* L. WITTGENSTEIN |
| PSYCHOLOGICAL TYPES* | *by* C. G. JUNG, M.D. |
| SCIENTIFIC THOUGHT* | *by* C. D. BROAD, Litt.D. |
| THE MEANING OF MEANING | *by* C. K. OGDEN and I. A. RICHARDS |
| INDIVIDUAL PSYCHOLOGY | *by* ALFRED ADLER |
| SPECULATIONS (*Preface by Jacob Epstein*) | *by* T. E. HULME |
| THE PSYCHOLOGY OF REASONING* | *by* EUGENIO RIGNANO |
| THE PHILOSOPHY OF "AS IF" | *by* H. VAIHINGER |
| THE NATURE OF INTELLIGENCE | *by* L. L. THURSTONE |
| TELEPATHY AND CLAIRVOYANCE | *by* R. TISCHNER |
| THE GROWTH OF THE MIND | *by* K. KOFFKA |
| THE MENTALITY OF APES | *by* W. KÖHLER |
| PSYCHOLOGY OF RELIGIOUS MYSTICISM | *by* J. H. LEUBA |
| THE PHILOSOPHY OF MUSIC | *by* W. POLE, F.R.S. |
| THE PSYCHOLOGY OF A MUSICAL PRODIGY | *by* G. REVESZ |
| PRINCIPLES OF LITERARY CRITICISM | *by* I. A. RICHARDS |
| METAPHYSICAL FOUNDATION OF SCIENCES | *by* E. A. BURTT, Ph.D. |
| THOUGHT AND THE BRAIN* | *by* H. PIÉRON |
| PHYSIQUE AND CHARACTER* | *by* ERNST KRETSCHMER |
| PSYCHOLOGY OF EMOTION | *by* J. T. MACCURDY, M.D. |
| PROBLEMS OF PERSONALITY | *in honour of* MORTON PRINCE |
| THE HISTORY OF MATERIALISM | *by* F. A. LANGE |
| PERSONALITY* | *by* R. G. GORDON, M.D. |
| EDUCATIONAL PSYCHOLOGY | *by* CHARLES FOX |
| LANGUAGE AND THOUGHT OF THE CHILD | *by* J. PIAGET |
| SEX AND REPRESSION IN SAVAGE SOCIETY* | *by* B. MALINOWSKI, D.Sc. |
| COMPARATIVE PHILOSOPHY | *by* P. MASSON-OURSEL |
| SOCIAL LIFE IN THE ANIMAL WORLD | *by* F. ALVERDES |
| HOW ANIMALS FIND THEIR WAY ABOUT | *by* E. RABAUD |
| THE SOCIAL INSECTS | *by* W. MORTON WHEELER |
| THEORETICAL BIOLOGY | *by* J. VON UEXKÜLL |
| POSSIBILITY | *by* SCOTT BUCHANAN |
| THE TECHNIQUE OF CONTROVERSY | *by* B. B. BOGOSLOVSKY |
| THE SYMBOLIC PROCESS | *by* J. F. MARKEY |
| POLITICAL PLURALISM | *by* K. C. HSIAO |
| HISTORY OF CHINESE POLITICAL THOUGHT | *by* LIANG CHI-CHAO |
| INTEGRATIVE PSYCHOLOGY* | *by* W. M. MARSTON |
| THE ANALYSIS OF MATTER | *by* BERTRAND RUSSELL, F.R.S. |
| PLATO'S THEORY OF ETHICS | *by* R. C. LODGE |
| HISTORICAL INTRODUCTION TO MODERN PSYCHOLOGY | *by* G. MURPHY |
| CREATIVE IMAGINATION | *by* JUNE E. DOWNEY |
| COLOUR AND COLOUR THEORIES | *by* CHRISTINE LADD-FRANKLIN |
| BIOLOGICAL PRINCIPLES | *by* J. H. WOODGER |
| THE TRAUMA OF BIRTH | *by* OTTO RANK |
| THE STATISTICAL METHOD IN ECONOMICS | *by* P. S. FLORENCE |
| THE ART OF INTERROGATION | *by* E. R. HAMILTON |
| THE GROWTH OF REASON | *by* FRANK LORIMER |
| HUMAN SPEECH | *by* SIR RICHARD PAGET |
| FOUNDATIONS OF GEOMETRY AND INDUCTION | *by* JEAN NICOD |
| THE LAWS OF FEELING | *by* F. PAULHAN |
| THE MENTAL DEVELOPMENT OF THE CHILD | *by* K. BÜHLER |
| EIDETIC IMAGERY | *by* E. R. JAENSCH |
| THE CONCENTRIC METHOD | *by* M. LAIGNEL-LAVASTINE |
| THE FOUNDATIONS OF MATHEMATICS | *by* F. P. RAMSEY |
| THE PHILOSOPHY OF THE UNCONSCIOUS | *by* E. VON HARTMANN |
| OUTLINES OF GREEK PHILOSOPHY | *by* E. ZELLER |
| THE PSYCHOLOGY OF CHILDREN'S DRAWINGS | *by* HELGA ENG |
| INVENTION AND THE UNCONSCIOUS | *by* J. M. MONTMASSON |
| THE THEORY OF LEGISLATION | *by* JEREMY BENTHAM |
| THE SOCIAL LIFE OF MONKEYS | *by* S. ZUCKERMAN |
| THE DEVELOPMENT OF THE SEXUAL IMPULSES | *by* R. E. MONEY-KYRLE |
| CONSTITUTION TYPES IN DELINQUENCY | *by* W. A. WILLEMSE |
| THE SCIENCES OF MAN IN THE MAKING | *by* E. A. KIRKPATRICK |
| ETHICAL RELATIVITY | *by* E. A. WESTERMARCK |

* *Asterisks denote that other books by the same author are included in the Series.*

*A complete list will be found at the end of the volume*

# THE LOGICAL SYNTAX OF LANGUAGE

By

RUDOLF CARNAP

PROFESSOR OF PHILOSOPHY IN THE UNIVERSITY OF CHICAGO

LONDON

ROUTLEDGE & KEGAN PAUL LTD

BROADWAY HOUSE: 68–74 CARTER LANE, E.C.4

for a sentence of the form $(\mathfrak{v}_1)(\mathfrak{v}_2)\ldots(\mathfrak{v}_n)(\mathfrak{pr}_1(\mathfrak{v}_1,\ldots\mathfrak{v}_n) \equiv \mathfrak{pr}_2(\mathfrak{v}_1,\ldots\mathfrak{v}_n))$ we can always write $\mathfrak{pr}_1 = \mathfrak{pr}_2$; and for a sentence of the form $(\mathfrak{v}_1)\ldots(\mathfrak{v}_n)(\mathfrak{pr}_1(\mathfrak{v}_1,\ldots\mathfrak{v}_n) \supset \mathfrak{pr}_2(\mathfrak{v}_1,\ldots\mathfrak{v}_n))$ we can always write $\mathfrak{pr}_1 \subset \mathfrak{pr}_2$. For this mode of symbolization without arguments, two different translations into word-language are possible. For instance, let 'P' and 'Q' be $\mathfrak{pr}^1$; then we can translate '$P \subset Q$' as: "The property P implies the property Q", or, if we wish, as: "the class P is a sub-class of the class Q"; correspondingly "sub-relation", when it is a question of many-termed $\mathfrak{pr}$. Further, we can interpret the $\mathfrak{Pr}$ '$P \vee Q$' when it is used without arguments as the "*sum* of the classes P and Q", and '$P \cdot Q$' as the "*product* of the classes P and Q"; analogously also the "sum" and "product of relations" in the case of many-termed $\mathfrak{pr}$. '$\Lambda$' and '$V$' used without arguments can be interpreted as "*null class*" and "*universal class*" (or as "null relation" and "universal relation", respectively). As an example of an application of the class symbolism, the Axiom of Selection PSII 21 may be used (the $\mathfrak{p}$ which occur are to be taken from suitable types of at least the second order):

$$[(M \subset \sim \text{Leer}) \cdot (F)(G)([M(F) \cdot M(G) \cdot \sim \text{Leer}(F \cdot G)] \supset (F = G))] \supset (\exists H)(F)[M(F) \supset \text{A1}(F \cdot H)]$$

Hereby 'A1' ("cardinal number 1") is to be defined as follows (compare § 38 *b*):

$$\text{A1}(F) \equiv (\exists x)(y)(F(y) \equiv (y = x))$$

The mode of symbolization whose introduction is indicated in the foregoing is completely analogous to Russell's symbolism of classes; the whole theory of classes and relations of the [*Princ. Math.*] can easily be put into this simplified form. But we shall not go into this here, as it raises no further fundamental problems.

## § 38. The Elimination of Classes

The historical development of the use of class symbols in modern logic contains several noteworthy phases, the examination of which is fruitful for the study even of present-day problems. We select for our consideration the two most important steps in this development, which are due to Frege and Russell. Frege [*Grundgesetze*] was the first to give an exact form to the traditional differentiation between the *content* and the *extent of a concept*. According to his view, the content of a concept is represented by the sentential function (that is to say, by an open sentence in which the



free variables serve to express indeterminateness and not universality). The extent (for instance, in the case of a property concept, i.e. of a one-termed sentential function, the corresponding class) is represented either by a special expression containing the sentential function, or else by a new symbol which is introduced as an abbreviation for this expression. An identity-sentence with class expressions here means the coextensiveness of the corresponding properties (if, for instance, '$k_1$' and '$k_2$' are the class symbols belonging to the $\mathfrak{pr}$ '$P_1$' and '$P_2$', then '$k_1 = k_2$' is equivalent in meaning to '$(x)[P_1(x) \equiv P_2(x)]$'). Later on, Russell proceeded in the same manner. Following the traditional modes of thought, however, Frege made a mistake at a certain point; and this mistake was discovered by Russell and subsequently corrected.

It was a decisive moment in the history of logic when, in the year 1902, a letter from Russell drew Frege's attention to the fact that there was a contradiction in his system. After years of laborious effort, Frege had established the sciences of logic and arithmetic on an entirely new basis. But he remained unknown and unacknowledged. The leading mathematicians of his time, whose mathematical foundations he attacked with unsparing criticism, ignored him. His books were not even reviewed. Only by means of the greatest personal sacrifices did he manage to get the first volume of his chief work [*Grundgesetze*] published, in the year 1893. The second volume followed after a long interval in 1903. At last there came an echo—not from the German mathematicians, much less the German philosophers, but from abroad: Russell in England attributed the greatest importance to Frege's work. In the case of certain problems Russell himself, many years after Frege, but still in ignorance of him, had hit upon the same or like solutions; in the case of some others, he was able to use Frege's results in his own system. But now, when the second volume of his work was almost printed, Frege learned from Russell's letter that his concept of class led to a contradiction. Behind the dry statement of this fact which Frege gives in the Appendix to his second volume, one senses a deep emotion. But, at all events, he could comfort himself with the thought that the error which had been brought to light was not a peculiarity of his system; he only shared the fate of all who had hitherto occupied themselves with the problems of the extension of concepts, of classes, and of aggregates—amongst them both Dedekind and Cantor.

The contradiction which was discovered by Russell is the antinomy which has since become famous, namely that of the class of those classes which are not members of themselves. In his Appendix, Frege examined various possibilities for a way out of the difficulty, but without discovering a suitable one. Then Russell, in an Appendix to his work [*Principles*] which appeared in the same year (1903), suggested a solution in the form of the *theory of types*, according to which only an individual can be an element of a class of the first level, and only a class of the *n*th level can be an element of a class of the $n+1$th level. According to this theory, a sentence of the form '$k \epsilon k$' or '$\sim(k \epsilon k)$' is neither true nor false; it is merely meaningless. Later on Russell showed that this antinomy can also be so formulated as to apply not only to classes but to properties as well (the antinomy of 'impredicable', see § 60*a*). Here, also, the contradiction is eliminated by means of the rule of types; applied to $\mathfrak{pr}^1$ (as symbols for properties) it runs thus: the argument of a ${}^1\mathfrak{pr}$ can only be an individual symbol, and the argument of an ${}^{n+1}\mathfrak{pr}$ can only be an ${}^n\mathfrak{pr}$.

Now it is a very remarkable fact that Frege himself had already made a similar classification of all sentential functions into levels and kinds which also were arranged according to the kinds of their arguments ([*Grundgesetze*] Vol. I, pp. 37 ff.). In this he had done important preliminary work for Russell's classification of types. But on two points—like traditional logic and Cantor's Theory of Aggregates—he made errors, which were corrected by means of Russell's rule of types. It is because of these errors that, in spite of the perfectly correct classification of functions, the antinomies arise. Frege's first error consisted in the fact that in his system all expressions (or more exactly, all expressions which begin with the assertion symbol) are either true or false. He was thus obliged to count as false, expressions in which an unsuitable argument was attributed to some predicate. It was Russell who first introduced the triple classification into true, false, and meaningless expressions —a classification which was to prove so important for the further development of logic and its application to empirical science and philosophy. According to Russell, those expressions which have unsuitable arguments are neither true nor false; they are meaningless (in our terminology: they are not sentences at all). When this first error of Frege is corrected, then the antinomy of the term 'impredicable' can no longer be set up in his system—for the definition would have to contain the contra-syntactical expression '$F(F)$'. The antinomy which relates to classes, however, can still be constructed in his system. For Frege made a second mistake in not applying the type-classification of the predicates (sentential functions), which he had constructed with such insight and clarity, to the classes corresponding to the predicates; instead of that, he counted the classes—and similarly the many-termed extensions—simply as individuals (objects) quite independently of the level and kind of the sentential function which defined the class in question. And even after the discovery of the contradiction, he still thought that he need not alter his procedure (Vol. II, pp. 254 f.), because he believed the names of objects and the names of functions to be differentiated by the fact that the former have a meaning of their own while the latter remain incomplete symbols which only become significant after being completed by means of other symbols. Now, since Frege held the numerals '0', '1', '2', etc., to be significant in themselves, and since, on the other hand, he defined these symbols as class symbols of the second level, he was compelled to regard class symbols, as opposed to predicates, as individual names. Today we have the tendency to regard all the partial expressions of a sentence which are not sentences in their turn as dependent; and to attribute independent meaning at most to sentences.



In order to define a cardinal number in Frege's sense without making use of classes, we have only to replace Frege's class of properties by a property of properties (designated by a ${}^2\mathfrak{pr}$). It is remarkable that Frege at an earlier stage expressed this view himself ([*Grundlagen*] 1884, p. 80, Note): "I think that [in the definition of 'cardinal number'], instead of 'extent of the concept', we might say simply 'concept'. But then two kinds of objections would be raised:.... I am of the opinion that both these objections could be removed; but that might lead too far at this stage." Later he apparently abandoned this view altogether. Then again—as it appears when one looks back—Russell seemed to be very close to the decisive point of abandoning classes altogether. While for Frege it was important to introduce the class symbols as well as the predicates—since in his system they obey different rules—the whole question had a different aspect for Russell. In order to avoid Frege's error, Russell did not adopt the class symbols as individual symbols but instead he divided them into types which correspond exactly to the types of the predicates. But by this means a quite unnecessary duplication was introduced. Russell himself recognized that it was of no importance for logic whether "classes" —that is to say, anything which is designated by the class symbols —"really exist" or not ("no-class theory"). The further development proceeded ever more definitely in the direction of the standpoint that class symbols are superfluous. In connection with Wittgenstein's statements, Russell himself later discussed the view that classes and properties are the same, but he did not as yet acknowledge it (1925: [*Princ. Math.*], 2nd edition of Vol. I). The whole question is connected with the problem of the Thesis of Extensionality (see § 67). Behmann [*Logik*] introduces the class symbolism merely as an abbreviated method of writing, in which the predicates are given without arguments; he insists, however, on differentiating between extensional and intensional sentences, holding that this method of writing is only admissible for the former. Von Neumann [*Beweistheorie*] and Gödel [*Unentscheidbare*] do not even symbolically make any difference between predicates and the corresponding class symbols; in the place of the latter, they simply use the former. The critique of Kaufmann ([*Unendliche*], [*Bemerkungen*]) concerning Russell's concept of class is also worthy of note. But this criticism is really directed less against the Russellian system itself than against the philosophical discussions by Russell and others of the concept of class, which do not properly belong to the system.

We will summarize briefly the development which we have just been considering. Frege introduced the class expressions in order to have, besides the predicates, something which could be treated like an object-name. Russell recognized the inadmissibility of such a treatment, but, nevertheless, retained the class expressions. The former reason for their introduction having been removed, however, they are now superfluous and therefore have been finally discarded.

## § 38*a*. On Existence Assumptions in Logic

If logic is to be independent of empirical knowledge, then it must assume nothing concerning the *existence of objects*. For this reason Wittgenstein rejected the Axiom of Infinity, which asserts the existence of an infinite number of objects. And, for kindred reasons, Russell himself did not include this axiom amongst the primitive sentences of his logic. But in Russell's system [*Princ. Math.*] as well as in that of Hilbert [*Logik*], sentences such as '$(\exists x)(F(x) \vee \sim F(x))$' and '$(\exists x)(x = x)$', and others like them, in which the existence of at least one object is stated, are (logically) demonstrable. Later on, Russell himself criticized this point ([*Math. Phil.*], Chap. xviii, Footnote). In the above-mentioned systems, not only the sentences which are true in every domain, independently of the number of objects in that domain, but also sentences (for example, the one just given) which are true, not in every domain, but in every *non-empty* domain, are demonstrable. In practice, this distinction is immaterial, since we are usually concerned with non-empty domains. But if, in order to separate logic as sharply as possible from empirical science, we intend to exclude from the logical system any assumptions concerning the existence of objects, we must make certain alterations in the forms of language used by Russell and Hilbert.

We may proceed somewhat as follows: No free variables are admitted in sentences and therefore universality can only be expressed by means of universal operators. The schemata of primitive sentences PSII 18 and 19 are retained (see § 30); PSII 16 and 17 are replaced by rules of substitution: $(\mathfrak{v}_1)(\mathfrak{S}_1)$ can be transformed into $\mathfrak{S}_1\binom{\mathfrak{v}_1}{\mathfrak{A}}$, and $(\mathfrak{p}_1)(\mathfrak{S}_1)$ into $\mathfrak{S}_1\binom{\mathfrak{p}_1(\mathfrak{Arg}_1)}{\mathfrak{S}_2}$. RII 2 disappears; but certain other rules must be laid down instead. In the language thus altered, when an object-name such as 'a' is given, 'P(a)' can be derived from '$(x)(\mathrm{P}(x))$'; and again, '$(\exists x)(\mathrm{P}(x))$' from 'P(a)'.



The important point is that the existential sentence can only be derived from the universal one when a proper name is available; that is to say, only when the domain is really non-empty. In the altered language, as opposed to the languages of Russell and Hilbert, the sentence '$(x)(\mathrm{P}(x)) \supset (\exists x)(\mathrm{P}(x))$' is not demonstrable without the use of a proper name.

In our object-languages I and II, the matter is quite different owing to the fact that they are not *name-languages* but *coordinate-languages*. The expressions of the type 0 here designate not objects but positions. The Axiom of Infinity (see § 33, 5*a*) and sentences like '$(\exists x)(x = x)$' are demonstrable in Language II, as are similar sentences in Language I. But the doubts previously mentioned are not relevant here. For here, those sentences only mean, respectively, that for every position there is an immediately succeeding one, and that at least one position exists. But whether or not there are objects to be found at these positions is not stated. That such is or is not the case is expressed in a co-ordinate language, on the one hand, by the fact that the $\mathfrak{fu}_b$ at the positions concerned have a value which appertains to the normal domain, or, on the other, by the fact that they have merely a trivially degenerate value. But this is stated not by analytic but by synthetic sentences.

*Example*. In the system of the *physical language*, the sentence which states that quadruples of real numbers (as quadruples of co-ordinates) exist is analytic. In its material interpretation it means that spatio-temporal positions exist. Whether something (matter or an electro-magnetic field) is to be found at a particular position is expressed by the fact that at the position in question the value of the density—or of the field-vector, respectively—is not zero. But whether anything at all exists—that is to say, whether there is such a non-trivially occupied position—can only be expressed by means of a synthetic sentence.

If it is a question not of the existence of objects but of the *existence of properties or classes* (expressed by means of predicates), then it is quite another matter. Sentences like '$(\exists F)(F = F)$' ("There exists a property (or class)") and '$(\exists F)(\mathrm{Leer}(F))$' ("There exists a null property (or class)") are true in every possible domain, including the null-domain; they are also analytic and logically demonstrable in the aforesaid system without existence assumptions.

incide. (*b*) The terms 'incomplete language', 'L-incomplete language', 'indeterminate language', 'descriptive language' coincide.

**Theorem 59.13.** If S is d-complete, then it is resoluble; and conversely. By Theorem 48.5.

For the d-terms, no valid theorems analogous to Theorems 11 and 12 exist.

**Theorem 59.14.** (*a*) If S is contradictory, then S is both d-complete and complete. (*b*) If S is inconsistent, then S is complete. By Theorem 1.

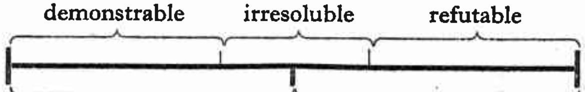
How the properties of languages here defined are transferred from one language to another can be seen from the table on p. 225 (B). The relation of the terms to one another is indicated by the arrows in the table below (as on p. 183).

*Properties of languages*

| L-d-terms: | | d-terms: | | c-terms: | | L-c-terms: |
|---|---|---|---|---|---|---|
| L-contradictory | → | contradictory | → | inconsistent | ← | L-inconsistent |
| L-non-contradictory | ← | non-contradictory | ← | consistent | → | L-consistent |
| L-d-complete<br>L-resoluble | → | d-complete<br>resoluble | → | complete<br>determinate<br>logical | ↔ | L-complete<br>L-determinate |
| L-d-incomplete<br>L-irresoluble | ← | d-incomplete<br>irresoluble | ← | incomplete<br>indeterminate<br>descriptive | ↔ | L-incomplete<br>(L-indeterminate)<br>synthetic |

We shall see that every consistent language which contains a general arithmetic is irresoluble. Only poorer languages are resoluble, for example, the sentential calculus. A richer language, though not resoluble, can yet be determinate and complete, provided that sufficient indefinite rules of transformation are laid down. This is the case, for instance, with the logical sub-languages of I and II. For such an *irresoluble but complete* language, the following classification of sentences holds; it is at the same time the classification of the logical sentences of any irresoluble language whatsoever (for the classification of the descriptive sentences, see p. 185):

(d-terms:) demonstrable irresoluble refutable



(c- and L-terms:) valid *analytic* contravalid *contradictory*



## § 60*a*. The Antinomies

In investigating the non-contradictoriness of a language, the first thing to be asked is whether the familiar so-called antinomies or paradoxes which appeared in earlier systems of logic and of the Theory of Aggregates have definitely been eliminated. This point is an especially critical one when we are concerned with a language which is rich enough to formulate, to any extent, its own syntax, whether in an arithmetized form or with the help of special syntactical designations. The syntactical sentences may sometimes speak about themselves, and the question arises whether this reflexiveness may not possibly lead to contradictions. This question is significant because it is not concerned with calculi of a specially constructed kind but with all systems whatsoever which contain arithmetic. We shall now investigate this question and in doing so we shall avail ourselves of the results obtained by Gödel.

We shall follow Ramsey's example in dividing the antinomies into two kinds, and we shall see that those of the second kind are the ones which come into consideration for our inquiry. These will therefore be examined more closely. In the examples we propose to use partly the word-language and partly a symbolism similar to that which was used in Language II; for the syntactical designations we shall employ in some cases Gothic symbols, and in others inclusion in inverted commas. Let us consider, to begin with, the following two antinomies.

1. **Russell's antinomy** [*Princ. Math.* I]; [*Math. Phil.*]. We define as follows: a property is called *impredicable* when it does not apply to itself. Expressed in symbols: "Impr$(F) \equiv \sim F(F)$". If in this case we substitute 'Impr' itself for '$F$', we get the contradictory sentence: "Impr(Impr) $\equiv \sim$ Impr(Impr)".

2. **Grelling's antinomy**. Definition: in a language which contains its own syntax, a syntactical predicate (for example, an adjective) is called *heterological* if the sentence which ascribes the property expressed by the predicate to the predicate itself is false. If, for instance, 'Q' is a syntactical predicate, then "Het('Q') $\equiv$ $\sim$Q('Q')" is true. (The fundamental difference between this antinomy and the foregoing, which is disregarded in many presentations, is to be noted, namely, that here the property Q is attributed, not to the property Q but to the predicate, i.e. the

symbol ‘Q’.) *Example:* the adjective ‘monosyllabic’ is heterological because ‘monosyllabic’ is not monosyllabic but penta-syllabic. Now, if instead of the predicate ‘Q’, we take the predicate ‘Het’ itself which has just been defined, we get, from the definition as stated, the contradictory sentence “Het(‘Het’) ≡ ∼Het(‘Het’)”.

In order to avoid antinomies in his language, Russell set up a complicated rule of types, which, particularly in the theory of real numbers, gave rise to certain difficulties, to overcome which he found it necessary to state a special axiom, the so-called Axiom of Reducibility. Ramsey ([*Foundations*] Treatise 1, 1925) has shown that the same object may be attained by a far simpler method. He discovered, namely, that it is possible to differentiate between two kinds of antinomies which may be designated as logical (in the narrower sense) and syntactical (the latter are also called linguistic, epistemological, or semantic). Example (1) belongs to the first category and (2) to the second. Following Peano, Ramsey pointed out that the antinomies of the second kind do not appear directly in the symbolic system of logic, but only in the accompanying text; for they are concerned with the expressions. From this fact he drew the practical conclusion that in the construction of a symbolic system it is not necessary to take note of these syntactical antinomies. Now since the antinomies of the first kind are already eliminated by the so-called simple rule of types, this is sufficient; the branched rule of types and the axiom of reducibility which it necessitates are superfluous.

On the basis of the *simple rule of types* (as in II for instance) the type of a predicate is determined by the type of the appertaining arguments alone. On the basis of Russell's branched rule of types, the form of the chain of definitions of a predicate is also a factor in determining its type (for instance, whether it is definite or not). But the simple rule of types is sufficient to determine that a predicate always belongs to a type other than that of the appertaining arguments (namely, that it always belongs to a type of a higher level). Thus, here, a sentence cannot have the form ‘*F*(*F*)’. And hence a definition of the form given for ‘impredicable’ is obviously impossible. In the same way, the other well-known antinomies of the first kind are obviated by means of the simple rule of types.

The problem of the *syntactical antinomies*, however, obviously reappears when it is a question of a language S in which the syntax of S itself can be formulated, and therefore in the case of every



language which contains arithmetic. There is a prevalent fear that with a syntax of this sort, which refers to itself, either contradictions similar to the syntactical antinomies will be unavoidable, or in order to avoid them, special restrictions, something like the “branched” rule of types, will be necessary. A closer investigation will show, however, that this fear is not justified.

The above-mentioned view is held, for instance, by Chwistek. He had already, before Ramsey, had the idea of stating only the simple rule of types, and thus rendering the axiom of reducibility unnecessary. Later, however, he came to the conclusion that with the rejection of the branched rule of types the syntactical antinomies—that of Richard, for example—would appear (see Chwistek [*Nom. Grundl.*]). In my opinion, however, the indispensability of the branched rule of types in Chwistek's system is due only to the fact that he uses the autonymous mode of speech for his syntax (the so-called Semantics) (see § 68).

Apart from Grelling's, the most important example of a syntactical antinomy is the one which was already famous in antiquity, the *antinomy of the liar* (for the history of this see Rüstow). Someone says: “I am lying”, or more exactly: “I am lying in this sentence”, in other words: “This sentence is false.” If the sentence is true, then it is false; and if it is false, then it is true.

Another antinomy which belongs to the category of the syntactical antinomies is Richard's (see [*Princ. Math.*] I, 61, and Fraenkel [*Mengenlehre*] p. 214 ff.). In its original version it is concerned with the decimals definable in a particular word-language. It can be easily transferred to $\mathfrak{Zpr}^1$ in the following manner. Let S be a language whose syntax is formulated in S. In S there are at most a denumerable number of $\mathfrak{Zpr}$ which are definable. Therefore we can correlate univocally a natural number with every such $\mathfrak{Zpr}^1$ (for instance, by a lexicographical arrangement of the definition-sentences or, in an arithmetized syntax, simply by the term-number of the $\mathfrak{Zpr}^1$). Let ‘c’ be a numerical expression; we will call the number c a Richardian number if c is the number of a $\mathfrak{Zpr}^1$, say ‘P’, which does not appertain to the number c, so that ‘P(c)’ is false (contradictory). Accordingly, the adjective ‘Richardian’ is a defined $\mathfrak{Zpr}^1$, and thus has correlated with it a certain number, say b. Now b must be either Richardian or not. If b is Richardian, then, according to the definition, the property having the number b does not appertain to b; therefore, in this case, in contradiction to our assumption, b is not Richardian. Hence b must be non-Richardian. b must leave the definition of 'Richardian' unfulfilled, and therefore must possess the property having the number b; that is to say, b must be Richardian. This is a contradiction.

It is characteristic of the syntactical antinomies mentioned that they operate with the concepts 'true' and 'false'. For this reason we will examine these concepts more closely before considering the syntactical antinomies any further.

## § 60*b*. The Concepts 'True' and 'False'

The concepts 'true' and 'false' are usually regarded as the principal concepts of logic. In the ordinary word-languages, they are used in such a way that the sentences '$\mathfrak{S}_1$ is true' and '$\mathfrak{S}_1$ is false' belong to the same language as $\mathfrak{S}_1$. *This customary usage of the terms 'true' and 'false' leads, however, to a contradiction.* This will be shown in connection with the antinomy of the liar. In order to guard ourselves against false inferences, we will proceed in a strictly formal manner. Let the syntax of S formulated in S contain three syntactical adjectives, '$\mathfrak{N}$', '$\mathfrak{W}$', '$\mathfrak{F}$', concerning which we will make only the following assumptions (V 1–3). In these, we shall write the sentence: "$\mathfrak{A}_1$ has the property $\mathfrak{N}$" in an abbreviated form, thus: '$\mathfrak{N}(\mathfrak{A}_1)$'. If '$\mathfrak{N}(\mathfrak{A}_1)$' is interpreted as "$\mathfrak{A}_1$ is a non-sentence", '$\mathfrak{W}(\mathfrak{A}_1)$' as: "The expression $\mathfrak{A}_1$ is a sentence, and, specifically, a true sentence", and '$\mathfrak{F}(\mathfrak{A}_1)$' as: "$\mathfrak{A}_1$ is a sentence, and, specifically, a false sentence", then our assumptions V 1–3 are in agreement with the ordinary use of language.

V 1. Every expression of S has exactly one of the three properties $\mathfrak{N}$, $\mathfrak{W}$, $\mathfrak{F}$.

V 2*a*. Let 'A' be any expression whatsoever of S (not: "designation of an expression"); if $\mathfrak{W}$('A'), then A. [For instance: if "this tree is high" is true, then this tree is high.]

V 2*b*. If A, then $\mathfrak{W}$('A').

V 3. For any $\mathfrak{A}_1$, the expressions '$\mathfrak{N}(\mathfrak{A}_1)$', '$\mathfrak{W}(\mathfrak{A}_1)$', '$\mathfrak{F}(\mathfrak{A}_1)$' do not possess the property $\mathfrak{N}$ (hence, they do possess either $\mathfrak{W}$ or $\mathfrak{F}$, according to V 1).

From V 1 and 2*b* it follows that:

If $\mathfrak{F}$('A'), then not $\mathfrak{W}$('A'), and therefore not A. (4)



From V 1 and 2*a* it follows that:

If not A, then not $\mathfrak{W}$('A'), and therefore $\mathfrak{F}$('A'), or $\mathfrak{N}$('A'). (5)

Now in analogy with the assertion of the liar, it is easy to show that the investigation of an expression $\mathfrak{A}_2$ with the text '$\mathfrak{F}(\mathfrak{A}_2)$' leads to a contradiction. The fact that an expression is here designated by a symbol (namely: '$\mathfrak{A}_2$'), which itself occurs in itself, easily has a confusing effect. But we can also establish the contradiction without this direct reflexive relation; it is not, as is so often believed, the reflexiveness which constitutes the error upon which the contradiction depends; the error lies rather in the unrestricted use of the terms 'true' and 'false'. Let us examine the two expressions '$\mathfrak{F}(\mathfrak{A}_1)$' and '$\mathfrak{W}(\mathfrak{A}_2)$'. Obviously these are expressions, at worst non-sentences. We are entirely at liberty as to which expressions we choose to designate by '$\mathfrak{A}_1$' and '$\mathfrak{A}_2$'; let us agree that:

(*a*) $\mathfrak{A}_1$ shall be the expression '$\mathfrak{W}(\mathfrak{A}_2)$'; (*b*) $\mathfrak{A}_2$ shall be the expression

'$\mathfrak{F}(\mathfrak{A}_1)$'. (6)

(Here, as can be seen, no designation of an expression occurs in the expression itself.)

According to V 3:

Either $\mathfrak{W}$('$\mathfrak{F}(\mathfrak{A}_1)$') or $\mathfrak{F}$('$\mathfrak{F}(\mathfrak{A}_1)$'). (7)

We first make the assumption: $\mathfrak{W}$('$\mathfrak{F}(\mathfrak{A}_1)$'). From this, in accordance with V 2*a*, it would follow that: $\mathfrak{F}(\mathfrak{A}_1)$. This, according to (6*a*) is $\mathfrak{F}$('$\mathfrak{W}(\mathfrak{A}_2)$'); from which, according to (4), would follow: not $\mathfrak{W}(\mathfrak{A}_2)$. This is, by (6*b*): not $\mathfrak{W}$('$\mathfrak{F}(\mathfrak{A}_1)$'). Our assumption leads to its own opposite and is therefore refuted.

Thence, according to (7), it is true that:

$\mathfrak{F}$('$\mathfrak{F}(\mathfrak{A}_1)$'). (8)

From this, by (4), follows:

not $\mathfrak{F}(\mathfrak{A}_1)$. (9)

This, according to (6*a*) is:

not $\mathfrak{F}$('$\mathfrak{W}(\mathfrak{A}_2)$'). (10)

By V 3:

$\mathfrak{W}$('$\mathfrak{W}(\mathfrak{A}_2)$') or $\mathfrak{F}$('$\mathfrak{W}(\mathfrak{A}_2)$'). (11)

From (10) and (11):

$\mathfrak{W}$('$\mathfrak{W}(\mathfrak{A}_2)$'). (12)

Thence, in accordance with V 2*a*:

$$\mathfrak{W}(\mathfrak{A}_2). \quad (13)$$

From (8) and (6*b*):

$$\mathfrak{F}(\mathfrak{A}_2). \quad (14)$$

Therefore, in accordance with V 1:

$$\text{not } \mathfrak{W}(\mathfrak{A}_2). \quad (15)$$

(13) and (15) constitute a contradiction.

This contradiction only arises when the predicates 'true' and 'false' referring to sentences in a language S are used in S itself. On the other hand, it is possible to proceed without incurring any contradiction by employing the predicates 'true (in $S_1$)' and 'false (in $S_1$)' in a syntax of $S_1$ which is not formulated in $S_1$ itself but in another language $S_2$. $S_2$ can, for instance, be obtained from $S_1$ by the addition of those two predicates as new primitive symbols and the erection of suitable primitive sentences relating to them, in the following way: 1. Every sentence of $S_1$ is either true or false. 2. No sentence of $S_1$ is at the same time both true and false. 3. If, in $S_1$, $\mathfrak{S}_2$ is a consequence of $\mathfrak{K}_1$, and if all sentences of $\mathfrak{K}_1$ are true, then $\mathfrak{S}_2$ is likewise true. A theory of this kind formulated in the manner of a syntax would nevertheless not be a genuine syntax. *For truth and falsehood are not proper syntactical properties;* whether a sentence is true or false cannot generally be seen by its design, that is to say, by the kinds and serial order of its symbols. [This fact has usually been overlooked by logicians, because, for the most part, they have been dealing not with descriptive but only with logical languages, and in relation to these, certainly, 'true' and 'false' coincide with 'analytic' and 'contradictory', respectively, and are thus syntactical terms.]

Even though 'true' and 'false' do not in general occur in a proper syntax (that is to say, in a syntax which is limited to the design-properties of sentences), yet the majority of ordinary sentences which make use of these words can be translated either into the object-language or into the syntax-language. If $\mathfrak{S}_1$ is 'A', then '$\mathfrak{S}_1$ is true' can, for example, be translated by 'A'. In logical investigation, 'true' (and 'false') appears in two different modes of use. If the truth of the sentence in question follows from the rules of transformation of the language in question, then 'true' can be translated by 'valid' (or, more specifically, by 'analytic', 'demonstrable') and, correspondingly, 'false' by 'contravalid' (or 'contradictory', 'refutable'). 'True' may also refer to indeterminate sentences, but in logical investigations this only happens in the conditional form, as, for example: 'If $\mathfrak{S}_1$ is true, then $\mathfrak{S}_2$ is true (or false, respectively).' A sentence of this kind can be translated into the syntactical sentence: '$\mathfrak{S}_2$ is a consequence of $\mathfrak{S}_1$ (or is incompatible with $\mathfrak{S}_1$, respectively).'



## § 60*c*. The Syntactical Antinomies

We will now return to the question whether, in the formulation of the syntax of S in S, contradictions of the kind known as *syntactical antinomies* may not arise if, in the ordinary phrasing of these antinomies, 'true' and 'false' are replaced by syntactical terms in the manner indicated above.

Let S be a non-contradictory language (and, further, a consistent one), which contains arithmetic, and hence an arithmetized syntax of S itself also. Then a certain method exists whereby it is possible to construct, for any and every syntactical property formulable in S, a sentence of S, $\mathfrak{S}_1$, such that $\mathfrak{S}_1$ attributes this property—whether rightly or wrongly—to itself. This has already been shown in the case of Language II (see § 35). Now, by means of a construction of this kind, we will try to restate the antinomy of the liar. It consists of a sentence which asserts its own falsehood.

First, let us replace 'false' in this antinomy by '*non-demonstrable*'. If we construct a sentence of S, $\mathfrak{S}_1$, which asserts of itself that it is non-demonstrable in S, then we have in $\mathfrak{S}_1$ an analogue to the sentence $\mathfrak{G}$ of Language II which has already been discussed (and to the sentence $\mathfrak{G}_I$ of Language I). Here no contradiction arises. If $\mathfrak{S}_1$ is true (analytic), then $\mathfrak{S}_1$ is not false (contradictory), but is only non-demonstrable in S. This is actually the case (see Theorem 36.2). The properties 'analytic' and 'non-demonstrable' are not incompatible.

Now let us replace 'false' by 'refutable' in the sentence of the liar. Assume that a sentence, $\mathfrak{S}_2$, is constructed in S which asserts that $\mathfrak{S}_2$ is itself refutable (in S). $\mathfrak{S}_2$ is then an analogue to the assertion of the liar. We will now observe whether the contradiction arises in the ordinary way. First let us assume that $\mathfrak{S}_2$ is actually refutable. Then $\mathfrak{S}_2$ will be true, and therefore analytic. On the other hand, however, every refutable sentence is contradictory, and hence not analytic. Therefore the assumption is a false one and $\mathfrak{S}_2$ is non-refutable. From this no contradiction follows. $\mathfrak{S}_2$ is actually non-refutable; since $\mathfrak{S}_2$ means the opposite of this, $\mathfrak{S}_2$ is false, and is therefore contradictory. But the properties 'non-refutable' and 'contradictory' are quite consistent with one another (see the diagram on p. 210); for instance '$\sim ()(\mathfrak{G})$' possesses both.

The impossibility of reconstructing the antinomy of the liar with the help of the terms 'non-demonstrable' or 'refutable' is due to the fact that not all analytic sentences are also demonstrable, and similarly not all contradictory sentences are also refutable. But what would happen if we were to use in place of 'true' and 'false' the syntactical terms 'analytic' and 'contradictory'? Like 'true' and 'false', these two terms constitute a complete classification of the logical sentences. It is easy to show that we can construct contradictions if we assume that 'analytic (in S)' and 'contradictory (in S)' are defined in a syntax which is itself formulated in S. We could then, of course, construct a logical sentence $\mathfrak{S}_3$ which, in material interpretation, would mean that $\mathfrak{S}_3$ was contradictory. $\mathfrak{S}_3$ would correspond exactly to the assertion of the liar. Since it would be a logical sentence, $\mathfrak{S}_3$ would be either analytic or contradictory. Now, if $\mathfrak{S}_3$ were contradictory, $\mathfrak{S}_3$ would be true, therefore analytic, therefore not contradictory. Hence, $\mathfrak{S}_3$ would have to be non-contradictory. But then $\mathfrak{S}_3$ would be false, and therefore contradictory—which would be a contradiction.

On the same assumptions it would be possible also to construct Grelling's antinomy. Let us state the procedure for Language II. Assuming that a predicate 'An' is definable in II in such a way that 'An$(x)$' means: "The $^{\text{SN}}$sentence $x$ is analytic (in II)." 'Heterological' could then be defined as follows: 'Het$(x) \equiv \sim$An(subst $[x, 3, \text{str}(x)]$)'. Let 'Het$(x)$' have the series-number b. Then it is easy to show that, for the sentence 'Het(b)', either assumption—that it is analytical or that it is contradictory—leads to a contradiction.

We have seen that if 'analytic in S' is definable in S, then S contains a contradiction; therefore we arrive at the following result:



**Theorem 60c.1.** If S is consistent, or, at least, non-contradictory, then '*analytic* (*in* S)' *is indefinable in* S. The same thing holds for the remaining c-terms which were defined earlier (in so far as they do not coincide with d-terms), for instance, 'valid', 'consequence', 'equipollent', etc. But it is not true for every c-term which does not coincide with a d-term.

If a syntax of a language $S_1$ is to contain the term 'analytic (in $S_1$)' then it must, consequently, be formulated in a language $S_2$ which is richer in modes of expression than $S_1$. On the other hand, the d-term 'demonstrable (in $S_1$)' can, under certain circumstances, be defined in $S_1$; whether that is possible or not depends upon the wealth of modes of expression which is available in $S_1$. With Languages I and II the situation on this point is as follows: 'analytic in I' is not definable in I, but it is definable in II; 'analytic in II' is not definable in II, but is only definable in a still richer language. 'Demonstrable in I', because it is indefinite, is not definable in I; but 'demonstrable in II' can be defined in II, namely, by means of '$(\exists r)$ [BewSatzII$(r, x)$]'.

The foregoing reflections follow the general lines of Gödel's treatise. They show also why it is impossible to prove the non-contradictoriness of S in S. Closely related to Theorem 1 is the following theorem (a generalization of Theorem 36.7; see Gödel [*Unentscheidbare*], p. 196; Gödel intends to give a proof of this generalized theorem in a continuation of that treatise).

**Theorem 60c.2.** If S is consistent, or at least non-contradictory, then *no proof of the non-contradictoriness* or consistency *of* S *can be formulated in a syntax which uses only the means of expression which are available in* S.

The investigation of Richard's antinomy (p. 213) leads to a similar conclusion. Assume that in S there is an $\mathfrak{Ag}$ by means of which a univocal enumeration of all the $\mathfrak{z}\text{pr}^1$ which are definable in S might be constructed. This could be effected, for example, by means of an $\mathfrak{fu}_1$ such that every full expression $\mathfrak{fu}_1(\mathfrak{z}\text{pr}^1)$ was a $\mathfrak{Z}$. We will use the symbolism of II and write $\mathfrak{fu}_1$ 'num'.

The univocality of the numbering is assumed:

$$(\text{num}(F) = \text{num}(G)) \supset (x)\,(F(x) \equiv G(x)). \qquad (1)$$

With the help of 'num', 'Ri' ("Richardian") could now be defined:

$$\text{Ri}(x) \equiv (F)\,[(\text{num}(F) = x) \supset \sim F(x)]. \qquad (2)$$

Since 'Ri' is a $\mathfrak{zpr}^1$, it has a certain particular number designated by 'num(Ri)'. We assume first that the number of 'Ri' is itself Richardian: 'Ri[num(Ri)]'. Then if we substitute in (2) 'num(Ri)' for '$x$', and 'Ri' for '$F$', '$\sim$Ri[num(Ri)]' easily follows. Since our assumption leads to its opposite, it follows that it is refuted; and therefore it is proved that

$$\sim \mathrm{Ri}[\mathrm{num}(\mathrm{Ri})]. \qquad (3)$$

From (1):

$$(\mathrm{num}(F) = \mathrm{num}(\mathrm{Ri})) \supset (\sim F[\mathrm{num}(\mathrm{Ri})] \equiv \sim \mathrm{Ri}[\mathrm{num}(\mathrm{Ri})]). \qquad (4)$$

From (3), (4):

$$(\mathrm{num}(F) = \mathrm{num}(\mathrm{Ri})) \supset \sim F[\mathrm{num}(\mathrm{Ri})]. \qquad (5)$$

From (2):

$$(F)\,[(\mathrm{num}(F) = \mathrm{num}(\mathrm{Ri})) \supset \sim F[\mathrm{num}(\mathrm{Ri})]] \supset \mathrm{Ri}[\mathrm{num}(\mathrm{Ri})]. \qquad (6)$$

From (5), (6):

$$\mathrm{Ri}[\mathrm{num}(\mathrm{Ri})]. \qquad (7)$$

The proved sentences (3) and (7) contradict one another; S is therefore contradictory. Thence follows:

**Theorem 60c.3.** If S is consistent, or at least non-contradictory, then it is not possible to construct in S either an $\mathfrak{Ag}$ or an $\mathfrak{Fu}$ by means of which a univocal enumeration of the $\mathfrak{zpr}^1$ of S could be constructed.—Although the aggregate of the $\mathfrak{zpr}^1$ which are definable in S is a denumerable aggregate, in accordance with this Theorem an enumeration of them cannot be effected with the means available in S itself. [The condition in this Theorem is only added for the purpose of facilitating understanding; if S is inconsistent, then in S no univocal enumeration of a number of objects is possible at all, since no (non-synonymous) $\mathfrak{Z}$ are available.]

## § 60*d*. Every Arithmetic is Defective

Let $S_1$ contain an arithmetic (in relation to a certain $\mathfrak{Z}$-series), and let the real numbers be represented in $S_1$ by $\mathfrak{zfu}^1$. Let $S_1$ be a conservative sub-language of $S_2$, and let the arithmetized syntax of $S_1$ be formulated in $S_2$. We will show that with the help of the arithmetico-syntactical terms of $S_2$, as referred to $S_1$, a $\mathfrak{zfu}^1$ can be defined in $S_2$ for which there is no $\mathfrak{zfu}^1$ in $S_1$ having the same course of values; this is true for every language $S_1$, however rich it may be, if we take a sufficiently rich language as $S_2$. We define the $\mathfrak{zfu}^1$ 'k' in $S_2$ in the following way: 1. If $x$ is not a term-number of a $\mathfrak{zfu}^1$ of $S_1$, then $\mathrm{k}(x) = 0$; 2. If $x$ is a term-number of a $\mathfrak{zfu}^1$ of $S_1$, let us say 'h', then $\mathrm{k}(x) = \mathrm{h}(x) + 1$. Then every $\mathfrak{zfu}^1$ of $S_1$ deviates from 'k' for a certain argument (namely, for its own term-number); and therefore in $S_1$ there is no $\mathfrak{zfu}^1$ having the same course of values as 'k'. In other words: a real number can be given which is not equal to any real number definable in $S_1$ (see p. 206).

**Theorem 60d.1.** For every language S *a real number which cannot be defined in S can be given.*

The above definition of 'k' corresponds to the so-called *diagonal method* of the Theory of Aggregates. Theorem 1 corresponds to the well-known theorem of the Theory of Aggregates which states that the aggregate of the real numbers is a non-denumerable aggregate. (On the concept of the non-denumerable aggregates see, however, § 71 *d*.) On the other hand, the above line of thought also corresponds to Richard's antinomy.

We will now summarize briefly the results of this investigation of the syntactical antinomies. Let the syntax of a language S be formulated in S. The reconstruction of the syntactical antinomies by means of terms which are defined in S (for instance, in Language II, 'non-demonstrable in II' or 'refutable in II') does not lead to contradictions; but it opens the way to the proof that certain sentences are non-demonstrable or irresoluble in S. With the help of other terms (for instance, 'analytic', 'contradictory', 'consequence', 'correlated number', 'term-number') the reconstruction of the syntactical antinomies is possible. This leads to the proof that these terms (of which the definitions have up to now only been formulated in words and not within a formalized system) cannot be defined in S, if S is consistent, or at least non-contradictory. Since terms and sentences of pure syntax are nothing other than syntactically interpreted terms and sentences of arithmetic, the investigation of the syntactical antinomies leads to the conclusion that *every arithmetic* which is to any extent formulated in any language *is necessarily defective in two respects.*

**Theorem 60d.2.** *For every arithmetical system* it is possible to state: (*a*) *indefinable arithmetical terms* and (*b*) *irresoluble arith-*

*metical sentences* (Gödel [*Unentscheidbare*]). In connection with (*a*) see Theorems 60*c*.1, 3, 60*d*.1. In connection with (*b*) see Theorem 60*c*.2; further irresoluble sentences analogous to $\mathfrak{G}$ in II and $\mathfrak{G}_I$ in I (see § 36) can be constructed.

This defectiveness is not to be understood as if there were, for instance, arithmetical terms which could not be formally (i.e. in a calculus) defined at all, or arithmetical sentences which could not be resolved at all. For every term which is stated in any unambiguous way in a word-language, there exists a formal definition in an appropriate language. Every arithmetical sentence $\mathfrak{S}_1$ which is, for instance, irresoluble in the language $S_1$ is yet determinate in $S_1$; in the first place there exists a richer syntax-language $S_2$, within which the proof either that $\mathfrak{S}_1$ is analytic or that $\mathfrak{S}_1$ is contradictory can be stated; and secondly, there exists an object-language $S_3$ of which $S_1$ is a proper sub-language, such that $\mathfrak{S}_1$ is resoluble in $S_3$. But there exists neither a language in which all arithmetical terms can be defined nor one in which all arithmetical sentences are resoluble. [This is the kernel of truth in the assertion made by Brouwer [*Sprache*], and, following him, by Heyting [*Logik*], p. 3, that mathematics cannot be completely formalized.] In other words, *everything mathematical can be formalized, but mathematics cannot be exhausted by one system*; it requires an infinite series of ever richer languages.

## (*d*) TRANSLATION AND INTERPRETATION

### § 61. TRANSLATION FROM ONE LANGUAGE INTO ANOTHER

We call $\mathfrak{Q}_1$ a *syntactical correlation* between the syntactical objects ($\mathfrak{A}$ or $\mathfrak{K}$) of one kind and those of another when $\mathfrak{Q}_1$ is a many-one relation by means of which exactly one object of the second kind is correlated to every object of the first, and every object of the second kind to at least one of the first. The $\mathfrak{A}$ (or $\mathfrak{K}$) which is correlated to $\mathfrak{A}_1$ (or $\mathfrak{K}_1$, respectively) by means of $\mathfrak{Q}_1$ is called the $\mathfrak{Q}_1$-*correlate* of $\mathfrak{A}_1$ (or of $\mathfrak{K}_1$), and is designated by '$\mathfrak{Q}_1[\mathfrak{A}_1]$' (or '$\mathfrak{Q}_1[\mathfrak{K}_1]$'). Herein the following condition is assumed: if $\mathfrak{A}_n$ has no direct $\mathfrak{Q}_1$-correlate but can be subdivided into the expressions $\mathfrak{A}_1, \mathfrak{A}_2, \ldots \mathfrak{A}_m$, which have such correlates, then $\mathfrak{Q}_1[\mathfrak{A}_n]$ is equal



to the expression composed of $\mathfrak{Q}_1[\mathfrak{A}_1], \mathfrak{Q}_1[\mathfrak{A}_2], \ldots \mathfrak{Q}_1[\mathfrak{A}_m]$. The class which contains all and only the $\mathfrak{Q}_1$-correlates of the sentences of $\mathfrak{K}_1$ is designated by '$\mathfrak{Q}_1[\mathfrak{K}_1]$'. According to this, the correlates of sentences are also determined by means of a correlation between expressions, and the correlates of sentential classes by means of a correlation between sentences. [In a formalized syntax, $\mathfrak{Q}_1$ can, for instance, be either an $\mathfrak{Sg}^2$, a $\mathfrak{Pr}^2$, an $\mathfrak{Ag}^1$, or an $\mathfrak{Fu}^1$.] We say that a certain syntactical relation is transformed into a certain other one by means of $\mathfrak{Q}_1$ if, when the first relation subsists between any two objects, the second subsists between the $\mathfrak{Q}_1$-correlates of these objects.

A syntactical correlation, $\mathfrak{Q}_1$, between all sentential classes (or all sentences, or the expressions of an expressional class $\mathfrak{K}_1$, or all symbols) of $S_1$ and those of $S_2$, is called a **transformance** of $S_1$ into $S_2$ in respect of classes (or of sentences, or expressions, or symbols, respectively) provided that, by means of $\mathfrak{Q}_1$, the consequence relation in $S_1$ is transformed into the consequence relation in $S_2$. For $\mathfrak{K}_1$ it is assumed that no expression of $\mathfrak{K}_1$, but every sentence of $S_1$ which does not belong to $\mathfrak{K}_1$, is univocally analyzable into several expressions of $\mathfrak{K}_1$. $\mathfrak{Q}_1$ is called a *transformance* of $S_1$ into $S_2$ if $\mathfrak{Q}_1$ is a transformance of $S_1$ into $S_2$ of one of the kinds mentioned. '*L-transformance* in respect of classes (sentences, and so on)' is analogously defined, the requirement in this case being the maintenance of the relation 'L-consequence'.

**Theorem 61.1.** If $\mathfrak{Q}_1$ is a transformance of $S_1$ into $S_2$, then $\mathfrak{Q}_1$ is also an L-transformance of $S_1$ into $S_2$.

**Theorem 61.2.** If $\mathfrak{Q}_1$ is a transformance of $S_1$ into $S_2$ in respect of sentences, then by $\mathfrak{Q}_1$ the consequence relation between sentences in $S_1$ is transformed into the consequence relation between sentences in $S_2$. The converse is not universally true.

A transformance of $S_1$ into $S_2$ is called **reversible** when its converse (that is, the relation subsisting in the reverse direction) is a transformance of $S_2$ into $S_1$; otherwise *irreversible*.

**Theorem 61.3.** Let $\mathfrak{Q}_1$ be a transformance of $S_1$ into $S_2$; if $\mathfrak{Q}_1$ is reversible, then $\mathfrak{Q}_1$ is a one-one relation. The converse is not universally true.

*Example* of an *irreversible* transformance in respect of sentences: the transformance given by Lewis [*Logic*], p. 178, of his system of strict implication (without the existential postulate) into the ordinary is true. We can, as previously, alter the formulation of the condition thus: for any '$P_1$' and '$P_2$', '$M(P_1) \equiv M(P_2)$' must always be a consequence of '$(x)\,(P_1(x) \equiv P_2(x))$'. With this as a basis, we now give the following definitions.

*Extensionality in relation to partial expressions.* Let $\mathfrak{Pr}_1$ occur in $\mathfrak{S}_1$; $\mathfrak{S}_1$ is called extensional in relation to $\mathfrak{Pr}_1$ if for any closed $\mathfrak{Pr}_2$, and any $\mathfrak{K}_1$ such that $\mathfrak{Pr}_1$ and $\mathfrak{Pr}_2$ are coextensive in relation to $\mathfrak{K}_1$, $\mathfrak{S}_1$ and $\mathfrak{S}_1\begin{bmatrix}\mathfrak{Pr}_1\\ \mathfrak{Pr}_2\end{bmatrix}$ are always equipollent in relation to $\mathfrak{K}_1$. Let $\mathfrak{Fu}_1$ occur in $\mathfrak{S}_1$; $\mathfrak{S}_1$ is called extensional in relation to $\mathfrak{Fu}_1$ if, for any closed $\mathfrak{Fu}_2$ and any $\mathfrak{K}_1$ such that $\mathfrak{Fu}_1$ and $\mathfrak{Fu}_2$ have the same course of values in relation to $\mathfrak{K}_1$, $\mathfrak{S}_1$ and $\mathfrak{S}_1\begin{bmatrix}\mathfrak{Fu}_1\\ \mathfrak{Fu}_2\end{bmatrix}$ are equipollent in relation to $\mathfrak{K}_1$. If $\mathfrak{S}_1$ is extensional in relation to all the closed $\mathfrak{S}$, $\mathfrak{Pr}$, and $\mathfrak{Fu}$ which occur in $\mathfrak{S}_1$, $\mathfrak{S}_1$ is called *extensional.* An $\mathfrak{Sg}_1$, to which $\mathfrak{Pr}$, $\mathfrak{Fu}$, or $\mathfrak{S}$ are suitable as arguments, is called extensional if every full sentence of $\mathfrak{Sg}_1$ with closed arguments is extensional in relation to every argument. Correspondingly for every $\mathfrak{Sfu}_1$ or $\mathfrak{Pr}_1$ to which $\mathfrak{Pr}$, $\mathfrak{Fu}$, or $\mathfrak{S}$ are suitable as arguments.

If every sentence of S is extensional in relation to every closed partial expression $\mathfrak{Pr}$ (or $\mathfrak{Fu}$) then S is called *extensional in relation to* $\mathfrak{Pr}$ (or $\mathfrak{Fu}$, respectively). If S is extensional in relation to partial sentences, to $\mathfrak{Pr}$, and to $\mathfrak{Fu}$, then S is called **extensional**.

**Theorem 66.1.** (*a*) If S is extensional in relation to $\mathfrak{Pr}$, then two closed $\mathfrak{Pr}$ which are coextensive (absolutely or in relation to $\mathfrak{K}_1$) are always (absolutely or in relation to $\mathfrak{K}_1$, respectively) synonymous. (*b*) If S is extensional in relation to $\mathfrak{Fu}$, then two closed $\mathfrak{Fu}$ which have the same course of values (absolutely or in relation to $\mathfrak{K}_1$) are always (absolutely or in relation to $\mathfrak{K}_1$, respectively) synonymous.

*Examples:* The languages of Russell and of Hilbert and our own Languages I and II are *extensional in relation to partial sentences.* That is shown, for instance, by the criterion of Theorem 65.7*c* (cf. Hilbert [*Logik*], p. 61). The symbols of equivalence in these languages are symbols of proper equivalence and hence, according to Theorem 65.10*b*, they are also symbols of proper identity for $\mathfrak{S}$. The form of the language will be simpler if only one symbol of identity is used (as in I and II, and in contrast with Russell and Hilbert), the same for $\mathfrak{S}$ as for $\mathfrak{Z}$, ${}^0\mathfrak{A}$ and so on. If from Russell's language R we construct a new language R′, by extending the rules of formation to admit of undefined $\mathfrak{pr}_b$ with $\mathfrak{S}$ as arguments, then



R′ is no longer necessarily extensional in relation to partial sentences; in order to guarantee extensionality here also, we can proceed, for example, by admitting $\mathfrak{S} = \mathfrak{S}$ as a sentence, and (in analogy with PSII 22, see below) stating a new primitive sentence as follows: '$(p \equiv q) \supset (p = q)$'. If the extended language II′ is constructed from II in the same way, then it is extensional in relation to partial sentences. Here no new primitive sentence is necessary, since we use the symbol of identity as symbol of equivalence, so that the above sentence of implication is demonstrable.

Languages I and II are also *extensional* in general. In II the extensionality in relation to $\mathfrak{Pr}$ and $\mathfrak{Fu}$ is guaranteed by PSII 22 and 23 (see p. 92). In the case of the other languages, the question of extensionality in relation to $\mathfrak{Pr}$ and $\mathfrak{Fu}$ can only be decided after further stipulations have been made, especially regarding what undefined ${}^n\mathfrak{pr}_b$ (for $n > 1$) are to be admitted.

The languages of Lewis, Becker, Chwistek, and Heyting are *intensional*, for partial sentences as well as for the rest (see § 67).

## § 67. The Thesis of Extensionality

Wittgenstein ([*Tractatus*], pp. 102, 142, 152) put forward the thesis that every sentence is "a truth-function of the elementary sentences" and therefore (in our terminology) extensional in relation to partial sentences. Following Wittgenstein, Russell ([*Introd. Wittg.*], pp. 13 ff.; [*Princ. Math.*] Vol. 1, 2nd edition, pp. xiv and 659 ff.) adopted the same view with regard to partial sentences and predicates; as I also did, but from rather a different standpoint ([*Aufbau*], pp. 59 ff.). In so doing, however, we all overlooked the fact that there is a multiplicity of possible languages. Wittgenstein, especially, speaks continually of "the" language. From the point of view of general syntax, it is evident that the thesis is incomplete, and must be completed by stating the languages to which it relates. In any case it does not hold for all languages, as the well-known examples of intensional languages show. The reasons given by Wittgenstein, Russell, and myself, in the passages cited, argue not for the necessity but merely for the possibility of an extensional language. For this reason we will now formulate the *thesis of extensionality* in a way which is at the same time more complete and less ambitious, namely: *a universal language of science may be extensional*; or, more exactly: for every given intensional language $S_1$, an extensional language $S_2$ may be constructed such that $S_1$ can be translated into $S_2$. In what follows we shall discuss the most important examples of intensional sentences and demonstrate the possibility of their translation into sentences of an extensional language.

Let us enumerate some of the most important *examples of intensional sentences*. 'A' and 'B' are abbreviations (not designations) for sentences, e.g. "It is raining now in Paris", etc. 1. Russell ([*Princ. Math.*], Vol. 1, p. 73 and [*Math. Phil.*], pp. 187 ff., and similarly Behmann [*Logik*], p. 29) gives examples of approximately the following kind: "Charles says A", "Charles believes A", "it is strange that A", "A is concerned with Paris". Incidentally Russell himself later, influenced by Wittgenstein's opinions, rejected these examples, and asserted that their intensionality was only apparent ([*Princ. Math.*], Vol. 1, 2nd edition, Appendix C). We prefer to say instead that these sentences are genuinely intensional but are translatable into extensional ones. 2. Intensional sentences concerning being-contained-in and substitution in relation to expressions: "(The expression) Prim (3) contains (the expression) 3"; "Prim (3) results from Prim ($x$) by substituting 3 for $x$". Sentences of this kind (but written in symbols) occur in the languages of Chwistek and Heyting. 3. Intensional sentences of the logic of modalities: "A is possible"; "A is impossible"; "A is necessary"; "B is a consequence of A"; "A and B are incompatible". Sentences of this kind (in symbols) occur in the systems of the logic of modalities constructed by Lewis, Becker, and others. 4. The following intensional sentences are akin to those of the logic of modalities: "Because A, therefore B"; "Although A, nevertheless B"; and the like. That any sentence $\mathfrak{S}_1$ of the examples given is intensional in relation to 'A' and 'B' follows easily from the criterion of Theorem 65.8*a*. If, for instance, 'A' is analytic and 'C' is synthetic, then '$A \equiv C$' is a consequence of 'C'; but the false sentence "A is necessary $\equiv$ C is necessary" is not a consequence of 'C'. These examples will be discussed in greater detail in what follows.

The above examples appear at first glance to be very different in kind. But, as a closer examination will show, they agree with one another in one particular feature, and this feature is the *reason for their intensionality: all these sentences are quasi-syntactical sentences* and, in particular, they are quasi-syntactical with respect to those expressions in relation to which they are intensional. With the establishment of this characteristic, *the possibility of their translation into an extensional language* is at once given, inasmuch, namely, as *every quasi-syntactical sentence is translatable into a correlative syntactical sentence*. That the syntax of any language (even an intensional one) can be formulated in an extensional language is easy to see. For arithmetic can be formulated to any desired extent in an extensional language, and hence an arithmetized syntax also. Incidentally this is equally true of a syntax in axiomatic form.



What we have said holds for all examples of intensional sentences so far known. Since we are ignorant of whether there exist intensional sentences of quite another kind than those known, we are also ignorant of whether the methods described, or others, are applicable to the translation of all possible intensional sentences. For this reason the *thesis of extensionality* (although it seems to me to be a fairly plausible one) is presented here *only as a supposition*.

## § 68. Intensional Sentences of the Autonymous Mode of Speech

Some of the known examples of intensional sentences belong to the autonymous mode of speech. When translated into an extensional language, they are transformed into the correlated syntactical sentences. We will first of all examine the converse process, namely, the construction from an extensional syntactical sentence of an intensional sentence with an autonymous expression. By this means the nature of these intensional sentences will become clear.

Let $S_1$ and $S_2$ be extensional languages; and let $S_2$ contain $S_1$ as a sub-language and the syntax of $S_1$ by virtue of $\mathfrak{Q}_1$. Let $\mathfrak{A}_1$ be an $\mathfrak{S}$, $\mathfrak{Pr}$, or $\mathfrak{Fu}$ of $S_1$, and $\mathfrak{S}_2$ (in $S_2$) have the form $\mathfrak{Pr}_2(\mathfrak{Q}_1[\mathfrak{A}_1])$. In material interpretation: $\mathfrak{Q}_1[\mathfrak{A}_1]$ is a syntactical designation of $\mathfrak{A}_1$; $\mathfrak{S}_2$ ascribes to $\mathfrak{A}_1$ a certain syntactical property expressed by $\mathfrak{Pr}_2$. $\mathfrak{Pr}_2(\mathfrak{A}_1)$ is in general not a sentence of $S_2$. Now, out of $S_2$, we construct an extended language $S_3$ (that is to say, $S_2$ is a proper sub-language of $S_3$). The rules of formation are extended as follows: in $S_3$, for every $\mathfrak{A}_3$ which is isogenous with $\mathfrak{A}_1$ in $S_1$, $\mathfrak{Pr}_2(\mathfrak{A}_3)$ is a sentence, and hence $\mathfrak{Pr}_2(\mathfrak{A}_1)$ also (let this be $\mathfrak{S}_1$); further, the rules of transformation are extended as follows: in $S_3$, for every $\mathfrak{A}_3$ which is isogenous with $\mathfrak{A}_1$ in $S_1$, $\mathfrak{Pr}_2(\mathfrak{A}_3)$ is equipollent to $\mathfrak{Pr}_2(\mathfrak{Q}_1[\mathfrak{A}_3])$, and therefore $\mathfrak{S}_1$ is also equipollent to $\mathfrak{Pr}_2(\mathfrak{Q}_1[\mathfrak{A}_1])$ (this is $\mathfrak{S}_2$). Then, according to the criterion given earlier (p. 238), $\mathfrak{A}_1$ is *autonymous* in $\mathfrak{S}_1$. A sentence which is formulated like $\mathfrak{S}_1$ is in general intensional in respect of $\mathfrak{A}_1$.

*Example:* Let $S_1$ be I. As syntax-language in $S_2$ we will take the word-language. Let the $\mathfrak{Q}_1$-correlates (the syntactical designations) be formed with inverted commas. Let $\mathfrak{A}_1$ be '$0^{||}=2$', and accordingly $\mathfrak{A}_2$, "'$0^{||}=2$'". Let $\mathfrak{S}_2$ be "'$0^{||}=2$' is an equation'. Then $\mathfrak{S}_1$ is '$0^{||}=2$ is an equation'. For $S_3$ we stipulate that $\mathfrak{S}_1$ and $\mathfrak{S}_2$ be mutual consequences of one another; and likewise, corresponding other sentences with the same $\mathfrak{Pr}$. Then '$0^{||}=2$' is autonymous in $\mathfrak{S}_1$, and, according to Theorem 65.8*b*, $\mathfrak{S}_1$ is intensional in relation to '$0^{||}=2$'. For let $\mathfrak{A}_3$ be 'Prim(3)'; then $\mathfrak{A}_2 \equiv \mathfrak{A}_3$ is analytic but 'Prim(3) is an equation' ($\mathfrak{S}_4$), because it is equipollent to "'Prim(3)' is an equation', is contradictory; hence, since $\mathfrak{S}_1$ is analytic, $\mathfrak{S}_1 \equiv \mathfrak{S}_4$ is contradictory.

Now some of the examples of intensional sentences previously mentioned have the same character as the intensional sentences constructed in the way here described: their intensionality is due to the occurrence of an autonymous expression. We will cite some examples of this, at the same time giving the correlated syntactical sentences. The latter may belong to an extensional language. [Sentences 1*b* and 2*b* belong to descriptive syntax, 3*b*, 4*b*, and 5*b* to pure syntax. The preceding investigations and definitions have all been given in relation to pure syntax only; they may, however, be correspondingly extended to apply to descriptive syntax.] To interpret these sentences as belonging to the autonymous mode of speech seems to me to be the natural thing, especially in the case of 4*a* and 5*a*. However, if anyone prefers not to ascribe one of them (say 2*a* or 3*a*) to the autonymous mode of speech, he is at liberty to do so; the sentence in question will then belong to the material mode of speech. The only essential points are: (1) these intensional sentences are quasi-syntactical; and (2) they can (together with all other sentences of the same language) be translated into extensional sentences, namely, into the correlated syntactical sentences.

| *Intensional sentences* of the autonymous mode of speech | *Extensional sentences* of syntax |
|---|---|
| Let 'A' be an abbreviation (not a designation) of some sentence. | |
| 1*a*. Charles says (writes, reads) A. | 1*b*. Charles says 'A'. |
| 2*a*. Charles thinks (asserts, believes, wonders about) A. | 2*b*. Charles thinks 'A'. |
| 3*a*. A has to do with Paris. | 3*b*. 'Paris' occurs in a sentence which results from 'A' by the elimination of defined symbols. |
| 4*a*. Prim(3) contains 3. | 4*b*. '3' occurs in 'Prim(3)'. |
| 5*a*. Prim(3) results from Prim(*x*) by the substitution of 3 for *x*. | 5*b*. 'Prim(3)' results from 'Prim(*x*)' by the substitution of '3' for '*x*'. |

[Of the same kind is the following: "it is astounding that...", that is to say: "many wonder about the fact that...".]



We have here interpreted the previously mentioned (p. 246) examples of intensional sentences put forward by Russell, Chwistek, and Heyting, as sentences of the *autonymous mode of speech*. This interpretation is suggested by the relevant indications given by the authors themselves. Russell's sentences are already presented in the word-language; and for the sentences of Chwistek and Heyting, which are formulated in symbols, the authors themselves give paraphrases in the word-language corresponding to 4*a* and 5*a*.

Chwistek's system of so-called *semantics* is, on the whole, dedicated to the same task as our syntax. But Chwistek throughout employs the autonymous mode of speech (apparently without being aware of it himself). He uses as the designation of an expression with which a sentence of semantics is concerned either this expression itself or, alternatively, a symbol which is synonymous with it (and is thus, originally, not a designation but an abbreviation for it). As a result of the employment of the autonymous mode of speech, many sentences of Chwistek's semantics are intensional. Because of this, he has come to the conclusion that every formal (Chwistek says "nominalistic") theory of linguistic expressions must make use of intensional sentences. This view is refuted by the counter-example of our syntax, which, although strictly formal, is consistently extensional (this is most clearly seen in the formalized syntax of I in I, in Part II). The fact that Chwistek believed himself forced to abandon the simple rule of types for his semantics and to return to the branched rule (see § 60*a*), was also, in my opinion, only a consequence of his use of the autonymous mode of speech.

Heyting gives as the word-translation of certain symbolic expressions of his language: "the expression which results from *a* when the variable *x* is replaced wherever it appears by the combination of symbols *p*" ([*Math.* 1], p. 4) and: "*g* does not contain *x*" ([*Math.* 1], p. 7). Such formulations, like our examples 4*a* and 5*a*, belong, without any doubt, to the autonymous mode of speech. But even the sentential calculus of Heyting's system [*Logik*] contains intensional sentences; sentential junctions which can be shown to possess no characteristic are used (see p. 203). These circumstances make it natural to suppose not only that the whole system can be translated by us into a system of syntactical sentences, but also that this was in a certain sense the author's intention. "In a certain sense" only, because the distinction between the object- and the syntax-languages is nowhere explicitly made; so that it is not even clear which language it is whose syntax is supposed to be represented in the system. According to [*Grundlegung*], p. 113, the assertion of a sentence (which is formulated symbolically by placing the symbol of assertion in front of the sentence) is "the establishment of an empirical fact, namely the fulfilment of the intention expressed by the sentence" or of the expectation of a possible experience. Such an assertion may mean, for example, the historical circumstance that I have a proof of the proposition in question lying in front of me. According to this, the assertions in Heyting's system should be interpreted as sentences of descriptive syntax. On the other hand, Gödel [*Kolloquium* 4], p. 39, gives an interpretation of Heyting's system in which the sentences of the system would be purely syntactical sentences about demonstrability; ''A' is demonstrable' is formulated by means of 'BA', and consequently in the autonymous mode of speech.

## § 69. Intensional Sentences of the Logic of Modalities

We shall now give some further examples of *intensional* sentences together with their *translation into extensional syntactical sentences*. By means of this translation the *intensional sentences* are shown to be *quasi-syntactical*. Sentences 1*a* to 4*a* contain terms that are usually known as *modalities* ['possible', 'impossible', 'necessary', 'contingent' (in the sense of 'neither necessary nor impossible')]. Sentences 5*a* to 7*a* contain terms that are similar in character to these modalities, and are therefore treated by the newer systems of the logic of modalities (Lewis, Lukasiewicz, Becker, and others) together with them. In these systems, the modal sentences are symbolically formulated in approximately the same way as our examples 1*b* to 7*b*. Examples 8*a* are intensional sentences of the ordinary word-language which we add here because, as the syntactical translation shows, they are akin to the modal sentences. 'A' and 'B' are here sentences—i.e. abbreviations (not designations) of certain sentences (such as synthetic sentences) either of the word-language or of a symbolic language.

| *Intensional* sentences of the *logic of modalities* | | *Extensional* sentences of *syntax* |
|---|---|---|
| 1*a*. A is *possible*. | 1*b*. P(A). | 1*c*. 'A' is not contradictory. |
| 2*a*. A • ∼A is *impossible*. | 2*b*. I(A • ∼A); ∼P(A • ∼A). | 2*c*. 'A • ∼A' is contradictory. |



| | | |
|---|---|---|
| 3*a*. A ∨ ∼A is *necessary*. | 3*b*. N(A ∨ ∼A); ∼P∼(A ∨ ∼A). | 3*c*. 'A ∨ ∼A' is analytic. |
| 4*a*. A is *contingent*. | 4*b*. ∼N(A) • ∼I(A); P(A) • P(∼A). | 4*c*. 'A' is synthetic. ('A' is neither analytic nor contradictory; neither 'A' nor '∼A' is contradictory.) |
| 5*a*. A *strictly implies* B; B is a consequence of A. | 5*b*. A < B. | 5*c*. 'B' is an L-consequence of 'A'. |
| 6*a*. A and B are strictly equivalent. | 6*b*. A = B. | 6*c*. 'A' and 'B' are L-equipollent (i.e. mutual L-consequences). |
| 7*a*. A and B are compatible. | 7*b*. C(A, B); ∼(A < ∼B). | 7*c*. 'A' and 'B' are L-compatible. ('∼B' is not an L-consequence of 'A'.) |
| 8*a*. Because A, therefore B; A, hence B. | | 8*c*. 'A' is analytic, 'B' is an L-consequence of 'A', 'B' is analytic. ('A' is valid, 'B' is a consequence of 'A', 'B' is valid.) |

Since the terms used in the logic of modalities are somewhat vague and ambiguous, it is also possible to choose other syntactical terms for the translations; in 2*c*, for instance, instead of 'contradictory' we may put 'contravalid', 'L-refutable', or 'refutable'. Similarly in the other cases, instead of the L-c-term we can take the general c-term, the L-d-term, or the d-term. With regard to 8*c*, in the majority of cases the general c-term (or the P-term) is perhaps more natural as an interpretation of 8*a* than the L-term. The difference between the so-called *logical* and the so-called *real* modalities can be represented in the translation by the difference between L- and general c-terms (or even P-terms):

| | |
|---|---|
| 9*a*. A is logically impossible. | 9*c*. 'A' is contradictory. |
| 10*a*. A is really impossible. | 10$c_1$. 'A' is contravalid. |
| | 10$c_2$. 'A' is P-contravalid. |

The translation of 10*a* depends upon the meaning of 'really impossible'. If this term is so meant that it is also to be applied to cases of logical impossibility, then the translation 10$c_1$ must be chosen; otherwise 10$c_2$. Analogous translations may be given for the three other modalities—for 'logically (or "really", respectively) possible', 'necessary', and 'contingent'.

That sentences 1*a* to 10*a* and 1*b* to 7*b* are *intensional* is easily seen. [*Example:* Let 'Q' be an undefined $\mathfrak{pr}_b$, and '≡' a symbol of proper equivalence. Let $\mathfrak{S}_1$ be 'Prim(3) ≡ Q(2)'; $\mathfrak{S}_2$ be: 'Prim(3) is necessary'; and $\mathfrak{S}_3$: 'Q(2) is necessary'. Then $\mathfrak{S}_2 \equiv \mathfrak{S}_3$ cannot be a consequence of $\mathfrak{S}_1$ (for $\mathfrak{S}_1$ is synthetic, $\mathfrak{S}_2$ analytic, and $\mathfrak{S}_3$ contradictory, and hence $\mathfrak{S}_2 \equiv \mathfrak{S}_3$ is contradictory). Therefore (by Theorem 65.7*b*) $\mathfrak{S}_2$ is intensional in relation to 'Prim(3)'.]

Since the sentences given here are quasi-syntactical, we can interpret them as sentences either of the autonymous or of the material mode of speech. In the case of the sentences of § 68, the verbal formulations, or the verbal paraphrases given by the authors, suggest interpretation in the autonymous mode of speech. On the other hand, in the case of the symbolic sentences 1 *b* to 7 *b*, it is not clear which of the two interpretations is intended—in spite of the fact that paraphrases (of the same kind as sentences 1 *a* to 7 *a*), and sometimes even detailed material explanations as well, are given by the authors. In relation to a particular example, the decisive question (as formulated in the material mode) is the following: Are 'I (A)' and 'A is impossible' to refer to the sentence 'A', or to that which is designated by 'A'? In the formal mode: Is ''A' is impossible' also to be a sentence? [If so, it must undoubtedly be equipollent to 'A is impossible.'] If the answer is in the affirmative, then 'I (A)' and 'A is impossible' both belong to the autonymous mode of speech; if in the negative, then they belong to the material mode of speech. The authors do, it is true, say that the sentences of modality are concerned with propositions, but this assertion would decide the question only if it were quite clear what was meant by the term 'proposition'. We will discuss the two possibilities separately.

1. Suppose that by the term 'proposition' the authors mean what we mean by 'sentence'. Then the term 'proposition' is a syntactical term, namely, the designation either of certain physical objects in descriptive syntax or of certain expressional designs in pure syntax. Then 'A is impossible' is concerned with the sentence 'A', hence is equipollent to ''A' is impossible', and belongs to the *autonymous mode of speech*. In this case the intensionality of the modal sentences does not depend upon the fact that they speak about expressions (in the examples, about sentences, in other cases, also about predicate-expressions) but upon the fact that they do so according to the autonymous and not according to the syntactical method.

2. Suppose that by a 'proposition' the authors mean not a sentence (in our sense) but that which is designated by a sentence. [For instance, in Lewis's [*Logic*], pp. 472 ff., the distinction between 'proposition' and 'sentence' is possibly to be understood in this way.] We will leave aside the question of what it is that is designated by a sentence (some people say thoughts or the content of thoughts, others, facts or possible facts); it is a question that easily leads to philosophical pseudo-problems. So we shall simply say neutrally "that which is designated by a sentence". In this interpretation, the sentence 'A is impossible' ascribes impossibility not to the sentence 'A' but to the A which is designated by the sentence. Here the impossibility is not a property of sentences. ''A' is impossible' is not a sentence; it is therefore a case not of the autonymous but of the *material* mode of speech. 'A is impossible' ascribes to the A which is designated by the sentence a quasi-syntactical property, instead of to the sentence 'A' the correlated syntactical property (here 'contradictory'). [In this example, the second interpretation is perhaps the more natural. It is the only possible one in the case of the formulation 'the process (or: state of affairs, condition) A is impossible'; see § 79, Examples 33 to 35. On the other hand, we are perhaps more inclined to relate a sentence about the consequence-relation or about derivability to sentences rather than to that which is designated by them, and accordingly to choose the first interpretation.] We shall see later that, in general, the use of the material mode of speech, though it is not inadmissible, brings with it the danger of entanglement in obscurities and pseudo-problems that are avoided by the application of the formal mode. So also here, the systems of the logic of modalities are (on the whole) formally correct. But if they are (in the accompanying text) interpreted in the second way, that is, in the material mode of speech, then pseudo-problems easily arise. This may perhaps explain the strange and, in part, unintelligible questions and considerations which are to be found in some treatises on the logic of modalities.



C. I. Lewis was the first to point out that in Russell's language [*Princ. Math.*] there is no way of expressing the fact that a certain proposition necessarily holds or that a particular proposition is a consequence of another. As against this, Russell can rightly maintain that, in spite of it, his system is adequate for the construction both of logic and of mathematics, that in it necessarily valid sentences can be proved and a sentence which follows from another can be derived from the former.

Although Lewis's contention is correct, it does not exhibit any lacuna *within* Russell's language. The requirement that a language be capable of expressing necessity, possibility, the consequence-relation, etc., is in itself justifiable; it is fulfilled by us for instance in the case of our Languages I and II, not by means of anything supplementary to these languages, but by the formulation of their syntax. On the other hand, both Lewis and Russell—they are agreed on this point—look upon the consequence-relation and implication as terms on the same footing as sentential connections, of which the first is the narrower. For this reason, Lewis found himself obliged to extend Russell's language by introducing, in addition to Russell's symbol of implication '⊃' (so-called material implication; in our terminology: proper implication), a new symbol '<' for what is called *strict implication* (in our terminology: an intensional symbol of improper implication without characteristic). This is intended to express the consequence-relation (or derivability-relation), that is to say, in Lewis's language, 'A < B' is demonstrable if 'B' is a consequence of 'A'. Lewis rightly pointed out that Russell's implication does not correspond to this interpretation, and that, moreover, none of the so-called truth-functions (in our terminology: the extensional sentential junctions) can express the consequence-relation at all. He therefore believed himself compelled to introduce intensional sentential junctions, namely, those of strict implication and of the modality-terms. In this way his system of the logic of modalities arose as an intensional extension of Russell's language. The system is set forth by Lewis in [*Survey*], pp. 291 ff., following MacColl, and later presented in an improved form in [*Logic*], pp. 122 ff., profiting by the researches of Becker and others. To Russell's system are added, as new primitive symbols, symbols for 'possible' and 'strictly equivalent', and with the help of these, 'impossible', 'necessary', 'strict implication', 'compatible', etc., are defined. Similar systems have been constructed by Lewis's pupils—by Parry ([*Koll.*], p. 5), for example, and Nelson ([*Intensional*]). Becker ([*Modalitätslogik*]), starting out from Lewis's [*Survey*], has made some interesting investigations using the same method. Before this Lukasiewicz had already worked out so-called many-valued systems of the sentential calculus (see his [*Aussagenkalkül*]). In [*Mehrwertige*] he interprets the sentences of the three-valued calculus by a translation into the modal sentences; these are, as are Lewis's, formulated in accordance with the quasi-syntactical method.

It is important to note the *fundamentally different nature of implication and the consequence-relation.* Materially expressed: the consequence-relation is a relation between sentences; *implication is not a relation between sentences*. [Whether, for example, Russell's opinion that it is a relation between propositions is erroneous or not, depends upon what is to be understood by a "proposition". If we are going to speak at all of 'that which is designated by a sentence', then implication is a relation between what is so designated; but the consequence-relation is not.] 'A ⊃ B' ($\mathfrak{S}_1$)—as opposed to



the syntactical sentence "'B' is a consequence of 'A'" ($\mathfrak{S}_2$)—means, not something about the sentences 'A' and 'B', but, with the help of these sentences and of the junction-symbol '⊃', something about the objects to which 'A' and 'B' refer. Formally expressed: '⊃' is a symbol of the object-language, and 'consequence' a predicate of the syntax-language. Of course, between the two sentences $\mathfrak{S}_1$ and $\mathfrak{S}_2$ there is an important relation (see Theorem 14.7). $\mathfrak{S}_2$ cannot, however, be inferred from $\mathfrak{S}_1$ but only from the (equally syntactical) sentence '$\mathfrak{S}_1$ is valid (or analytic)'. The majority of the symbolic languages (for example, Russell's [*Princ. Math.*]) are (after a suitable extension of the rules of inference) logical languages, and therefore contain no indeterminate sentences. Hence, in these systems, $\mathfrak{S}_2$ can be inferred from $\mathfrak{S}_1$. This explains why the sentences of implication are in general erroneously interpreted as sentences about consequence-relations. [This is one of the points which shows clearly how unfortunate it is that the indeterminate sentences have, for the most part, been disregarded in logical investigations.] The relation of the *intensional symbols of implication* in the systems of the logic of modalities, for instance that of the symbol of strict implication to '⊃' and to 'consequence', will become clear with the aid of the earlier example on p. 235; this relation corresponds exactly to that subsisting between 'LImp', 'Imp', and 'consequence'. [We can ignore here the differences between the intensional implications in the various systems; they correspond to the different definitions of the syntactical concept of 'consequence'.]

Russell's choice of the designation 'implication' for the sentential junction with the characteristic TFTT has turned out to be a very unfortunate one. The words 'to imply' in the English language mean the same as 'to contain' or 'to involve'. Whether the choice of the name was due to a confusion of implication with the consequence-relation, I do not know; but, in any case, this nomenclature has been the cause of much confusion in the minds of many, and it is even possible that it is to blame for the fact that a number of people, though aware of the difference between implication and the consequence-relation, still think that the symbol of implication ought really to express the consequence-relation, and count it as a failure on the part of this symbol that it does not do so. If we have retained the term 'implication' in our system, it is, of course, in a sense entirely divorced from its original meaning; it serves in the syntax merely as the designation of sentential junctions of a particular kind.

## § 70. The Quasi-Syntactical and the Syntactical Methods in the Logic of Modalities

All the foregoing systems of the logic of modalities (within the province of modern logic, in symbolic language) have, it seems, applied the *quasi-syntactical method.* This is not a matter of conscious choice between syntactical and quasi-syntactical methods; rather the method applied is held to be the natural one. All intensional sentences of the previously existing systems of the logic of modalities are, in any case, quasi-syntactical sentences, independently of which of the two interpretations earlier discussed is intended or (by a suitable incorporation in a more comprehensive language) carried into effect. [Incidentally, it should be noted that for each of the systems one of the two interpretations can be arbitrarily chosen and carried out, provided no attention is paid to the authors' indications regarding interpretation. Accordingly, it is, in particular, possible to interpret every sentence $\mathfrak{S}_1$ of the logic of modalities that is intensional in respect of a partial expression $\mathfrak{A}_1$, in such a way that $\mathfrak{A}_1$ is autonymous in $\mathfrak{S}_1$.] *Every intensional system of the logic of modalities* (and that even when synthetic sentences are admitted as arguments) *can be translated into an extensional syntactical language*, whereby every intensional sentence, since it is quasi-syntactical, is translated into the correlated syntactical sentence. In other words: *syntax already contains the whole of the logic of modalities*, and the construction of a special intensional logic of modalities is not required.

Whether, for the construction of a logic of modalities, the quasi-syntactical or the syntactical method is chosen is solely a question of expedience. We will not here decide the question but will only state the properties of both methods. The use of the quasi-syntactical method leads to intensional sentences, while the syntactical method can also be carried into effect in an extensional language. In a certain sense, the quasi-syntactical method is the simpler; and it may be that it will prove to be the appropriate one for the solution of certain problems. It will only be possible to pronounce judgment on its fruitfulness as a whole when the method is further developed. Hitherto, if I am not mistaken, it has in the main only been applied to the domain of the sentential



calculus which, on account of the resolubility of its sentences, is quite a simple one (see Parry [*Koll.*], pp. 15 f.). It cannot be said that the logic of modalities does not necessitate any syntactical terms and is therefore simpler. For the construction of every calculus, and therefore also of the logic of modalities, a syntax-language is required in which the statement of the rules of inference and of the primitive sentences is formulated (see § 31); it is usual simply to take the word-language for this purpose. Now, as soon as this syntax-language is obtained, everything that it is desired to express by the sentences of modality—and, in general, far more—can be defined and formulated within it. That is the reason why we have here given preference to the syntactical method. It is, however, in any case, a worth-while task to develop the quasi-syntactical method in general, and its use in the logic of modalities in particular, and to investigate its possibilities in comparison with the syntactical method.

Even if in the construction of a logic of modalities we wish to use, not the syntactical but the ordinary method hitherto employed, the realization that this method is a quasi-syntactical one can help us to overcome a number of uncertainties. These, for example, have manifested themselves at various points in the fact that, wishing to start from evident axioms, logicians have found themselves in doubt about the evidence of certain sentences; it has even happened that sentences which had previously been individually regarded as evident have turned out later to be incompatible. As soon, however, as it is seen that the concepts of modality—even when they are formulated quasi-syntactically—are concerned with syntactical properties, their relativity is recognized. They must always be referred to a particular language (which may be other than that in which they are formulated). In this way the problems regarding the evident character of *absolute* relations between the modality-concepts disappear.

## § 71. Is an Intensional Logic Necessary?

Some logicians take the view that the ordinary logic (for instance, that of Russell) is deficient in some respects and must therefore be supplemented by a new logic, which is designated as intensional logic or the logic of meaning (e.g. Lewis, Nelson

[*Intensional*], Weiss, and Jörgensen [*Ziele*], p. 93). Is this requirement justified? A close examination shows that two different questions, which should be treated separately, are here involved.

1. Russell's language is an extensional language. It is required that it be supplemented by an *intensional language* for the purpose of expressing the concepts of modality ('consequence', 'necessary', etc.). We have dealt with this question before, and have seen that the concepts of modality may also be expressed in an extensional language, and that their formulation only led to intensional sentences because the quasi-syntactical method was used. Neither for an object-language concerned with any domain of objects nor for the syntax-language of any object-language is it necessary to go outside the framework of an extensional language.

2. As opposed to the ordinary formal logic, a *logic of content* or a *logic of meaning* is demanded. And, further, it is believed that this second requirement also will be fulfilled by the construction of an intensional logic of modalities; thus it often happens that the designations 'intensional logic' and 'logic of meaning' are used synonymously. It is thought, that is, that the concepts of modality, since they are not dependent merely upon the truth-values of the arguments, are therefore dependent upon the *meaning* of the arguments. This is often especially emphasized in connection with the consequence-relation (e.g. Lewis [*Survey*], p. 328: "Inference depends upon meaning, logical import, intension"). If all that is meant by this is merely that, if the meanings of two sentences are given, the question of whether one is a consequence of the other or not is also determined, I will not dispute it (although I prefer to regard the connection from the opposite direction, namely, the relations of meaning between the sentences are given by means of the rules of consequence; see § 62). But the decisive point is the following: *in order to determine whether or not one sentence is a consequence of another, no reference need be made to the meaning of the sentences. The mere statement of the truth-values is certainly too little; but the statement of the meaning is, on the other hand, too much. It is sufficient that the syntactical design of the sentences be given.* All the efforts of logicians since Aristotle have been directed to the formulation of the rules of inference as *formal rules*, that is to say, as rules which refer only to the form of the sentences (for the development of the formal character of logic, see Scholz [*Geschichte*]). It is theoretically possible to establish the logical relations (consequence-relation, compatibility, etc.) between two sentences written in Chinese without understanding their sense, provided that the syntax of the Chinese language is given. (In practice this is only possible in the case of the simpler artificially constructed languages.) The two requirements (1) and (2), which are usually blended into one, are entirely independent of one another. Whether we wish to speak merely of the forms of the language $S_1$ or of the sense (in some meaning of the word) of the sentences of $S_1$, in either case an intensional language may be used; but we can also use an extensional language for both these purposes. *The difference between the extensionality and intensionality of a language has nothing to do with the difference between the formal and the material treatment.* Now, is it the business of logic to be concerned with the sense of sentences at all (no matter whether they are given in extensional or in intensional languages)? To a certain extent, yes; namely, in so far as the sense and relations of sense permit of being formally represented. Thus, in the syntax, we have represented the formal side of the sense of a sentence by means of the term 'content'; and the formal side of the logical relations between sentences by means of the terms 'consequence', 'compatible', and the like. All the questions which it is desired to treat in the required logic of meaning are nothing more than questions of syntax; in the majority of cases, this is only concealed by the use of the material mode of speech (as is demonstrated by many examples in Part V). Questions about something which is not formally representable, such as the conceptual content of certain sentences, or the perceptual content of certain expressions, do not belong to logic at all, but to psychology. All questions in the field of logic can be formally expressed and are then resolved into syntactical questions. *A special logic of meaning is superfluous*; 'non-formal logic' is a *contradictio in adjecto*. Logic is syntax.



Sometimes the demand for an intensional logic is made in a third connection: it is maintained that hitherto logic has only dealt with the *extension of concepts*, whereas it should also deal with the intension of concepts. But, actually, the newer systems of logic (Frege, as early as 1893, followed by Russell and Hilbert) have got far beyond the stage of development of the mere logic of extension in this sense. Frege himself was the first to define in an exact way the old distinction between the intension and the extension of a concept (namely, by means of his distinction between a sentential function and its course of values). One can rather maintain the reverse, that modern logic, in its latest phase of development, has completely suppressed extension in favour of intension (cf. the elimination of classes, § 38). This misunderstanding has already been cleared up many times (see Russell [*Princ. Math.*], I, p. 72; Carnap [*Aufbau*], p. 58, Scholz [*Geschichte*], p. 63); it is always reappearing, however, amongst philosophers who are not thoroughly acquainted with modern logic (and amongst psychologists, who, in addition, confuse the logical and the psychological content of a concept).

## (*f*) RELATIONAL THEORY AND AXIOMATICS

### § 71*a*. Relational Theory

In the *theory of relations*, the properties of relations are investigated, particularly the structural properties—that is to say those which are retained in isomorphic transformance. A theory of this kind is nothing more than the syntax of many-termed predicates. We have abandoned the usual distinction between the one-termed predicates and the class-symbols appertaining to them, and designate both class and property by $\text{pr}^1$ (see §§ 37, 38). Similarly we no longer differentiate the $n$-termed predicates for $n > 1$ from the relational symbols which have hitherto been correlated with them as symbols of extension. In this section, we shall indicate briefly how the most important terms of the theory of relations may be incorporated in the general syntax of the predicates.

With regard to the terms used in the theory of relations (such as 'symmetrical', 'transitive', 'isomorphic', etc.), it is important to distinguish between their formulation in the object-language and their formulation in the syntax-language. By means of this distinction—the necessity of which is usually disregarded—certain paradoxes in connection with the question of the multiplicity of the transfinite cardinal numbers and the possibility of non-denumerable aggregates are, as we shall see, clarified.

We will call an $n$-termed predicate **homogeneous** when, from a sentence constructed from it and $n$ arguments, another sentence always arises as a result of any permutation of the arguments. The majority of the terms of relational theory refer to homogeneous two-termed predicates.



The relational properties of symmetry, reflexiveness, and so on are expressed, according to the ordinary method introduced by Russell, by means of predicates of the second level (or, in Russell's own symbolism, by class symbols of the second level). We will write the definitions in the following form (employing the symbolism of Language II, but leaving open the question as to whether the expressions of the zero level are numerical expressions or designations of objects):

(Non-emptiness):* $\mathrm{Erf}(F) \equiv (\exists x)(\exists y)(F(x,y))$ (1)

(Emptiness): $\mathrm{Leer}(F) \equiv \sim \mathrm{Erf}(F)$ (2)

(Symmetry):

$$\mathrm{Sym}(F) \equiv [\mathrm{Erf}(F) \bullet (x)(y)(F(x,y) \supset F(y,x))] \quad (3)$$

(Asymmetry): $\mathrm{As}(F) \equiv (x)(y)(F(x,y) \supset \sim F(y,x))$ (4)

(Reflexiveness):

$$\mathrm{Refl}(F) \equiv [\mathrm{Erf}(F) \bullet (x)(y)((F(x,y) \vee F(y,x)) \supset F(x,x))] \quad (5)$$

(Total reflexiveness):

$$\mathrm{Reflex}(F) \equiv [\mathrm{Erf}(F) \bullet (x)(F(x,x))] \quad (6)$$

(Irreflexiveness): $\mathrm{Irr}(F) \equiv (x)(\sim F(x,x))$ (7)

(Transitivity):

$$\mathrm{Trans}(F) \equiv [(\exists x)(\exists y)(\exists z)(F(x,y) \bullet F(y,z)) \bullet (x)(y)(z)((F(x,y) \bullet F(y,z)) \supset F(x,z))] \quad (8)$$

(Intransitivity):

$$\mathrm{Intr}(F) \equiv (x)(y)(z)[(F(x,y) \bullet F(y,z)) \supset \sim F(x,z)] \quad (9)$$

We have altered the usual forms of the definitions (see Russell [*Princ. Math.*]; Carnap [*Logistik*]) by introducing in the definiens of (3), (5), (6), and (8) an existential sentence or '$\mathrm{Erf}(F)$' as a conjunction-term. According to the definitions hitherto given, transitivity and intransitivity do not exclude one another; and similarly, neither do symmetry and asymmetry, reflexiveness and irreflexiveness. If, for instance, a relation has no intermediary term (that is to say, no term which occurs in one pair of the relation as second term, and in another pair as first term) then it is simultaneously both transitive and intransitive (because the implicans in the definiens of (9) is always false); and for the same reason a null relation is at the same time transitive, intransitive, symmetrical, asymmetrical, reflexive and irreflexive. On this account we introduce conditions which require for symmetrical, reflexive, and transi-

* *Erfülltheit.*