LOGIC AND LANGUAGE

SYNTHESE LIBRARY

A SERIES OF MONOGRAPHS ON THE

RECENT DEVELOPMENT OF SYMBOLIC LOGIC,

SIGNIFICS, SOCIOLOGY OF LANGUAGE,

SOCIOLOGY OF SCIENCE AND OF KNOWLEDGE,

STATISTICS OF LANGUAGE

AND RELATED FIELDS

LOGIC AND LANGUAGE

STUDIES DEDICATED TO

PROFESSOR RUDOLF CARNAP

ON THE OCCASION OF HIS SEVENTIETH BIRTHDAY

D. REIDEL PUBLISHING COMPANY / DORDRECHT-HOLLAND

Printed in the Netherlands by D. Reidel, Dordrecht

CONTENTS

CONTENTS

YEHOSHUA BAR-HILLEL

A PREREQUISITE FOR RATIONAL PHILOSOPHICAL DISCUSSION

Communication between philosophers has been deteriorating during the last decades. Logical empiricists and British linguistic philosophers have been branding large parts of the output of their speculative colleagues as 'nonsense' and 'literally unintelligible'. Speculative metaphysicians, after having recovered from the first shock, either just disregard these declarations, or else declare, on their part, that the standards of intelligibility employed by the critics are arbitrary.

The breakdown of communication is not always as radical as the situation just described might lead one to believe. An analytic philosopher (let us use this term to cover both logical empiricists and linguistic philosophers) might often try to suggest one or more reinterpretations of his colleague's original formulations, thereby making them intelligible to himself. (This, of course, indicates that he does 'understand' the original formulation; thereby his behavior is not made absurd since this kind of understanding is clearly different from the one he professes not to have.) Unfortunately, this conciliatory action seldom does much good. The reinterpretations have a tendency to become either flagrant truisms or flagrant falsities. Even in case they do neither, more often than not the speculative philosopher will reject them as completely missing the point of his intentions, perhaps adding that this failure could have been foreseen, since there is no reason why his deep insights should be formulable in the analytic philosopher's shallow and arbitrarily-restricted language.

One might have expected that another development would arrest and reverse this trend of deteriorating communication. I refer to the well-known fact that the standards of intelligibility of the late twenties and early thirties – 'whatever can be said at all, can be said in ordinary (thing-, observational) language' – have been undergoing a process of continuous liberalization, the history of which has been told many times. In one of the latest attempts at explicating the empiricist's standard of intelligibility [1]), a discourse is regarded as intelligible, not only if it is formulated wholly in observational language, but also if it is formulated in theoretical or mixed language, if only the theoretical terms occurring in it are connected via theoretical postulates and rules of correspondence with the observational terms. True enough, the degree of intelligibility of such discourse is deemed to be inferior to that of a purely observational one. However, thereby such discourse is not disqualified nor is one entitled to draw the consequence that it is somehow less important or scientific than fully intelligible observational discourse. Clearly, a treatise in theoretical physics, psychology or linguistics is not the worse off because of the fact that its theoretical terms are only partially and indirectly interpretable.

It is nevertheless rather doubtful whether this liberalization will bring about a reunion in philosophy. Though the analytic philosopher may be willing to regard the specifically metaphysical terms used by his colleague as theoretical terms of which he is quite ready not to require more than partial and indirect interpretaticn in observational terms, he will continue to ask his colleague to supply him this interpretation, at least in sufficient outline. But the speculative philosopher will quite often refuse to do this, just as he refused to comply with the earlier demand to supply full and direct interpretation by operational definitions. He might even point out that just as the empiricist will now agree that his earlier demands were unjustified, so he will come to realize in due time that his present demands are still unduly restrictive.

This may indeed turn out to be the case. Most analytic philosophers are today aware that the whole conception of an observational language is rather vague, that the line of demarcation between observational and theoretical terms is blurred, elastic and even to a considerable degree arbitrary, and will therefore be rather careful with their use of the epithets 'meaningless', 'nonsensical', or 'unintelligible'. But they will continue refusing to exert themselves overly in order to supply for their own benefit all those theoretical postulates and rules of correspondence, or at least a sufficiently large outline of these, which they regard as necessary in principle for ensuring a modicum of intelligibility. If an analytic philosopher finds that a certain philosophical text is *underinterpreted*, he may or may not attempt to suggest to the reader of this text

1) R. Carnap, The methodological character of theoretical concepts, The Foundations of Science and the Concepts of Psychology and Psychoanalysis, *Minnesota Studies in the Philosophy of Science*, vol. I, pp. 38–76, University of Minnesota Press, Minneapolis, 1956.

one or more ways of supplementing the missing links, but if the author refuses to follow suit, the analytic philosopher will still know no more rational reaction than to count himself out. The possibility that he himself twenty years hence, or the next generation of analytic philosophers, might conceivably liberalize the standards of intelligibility still further and then rejoin the metaphysical game, will be more or less cheerfully acknowledged but will not influence the present breakdown of communication.

Is this then the end of the conversation? I am not convinced that it must be so. Though I myself, for instance, as an analytic philosopher, see no way of joining many metaphysical discussions because of the hopelessness of remedying their state of underinterpretedness, I am ready to explain why I regard the language of these discussions as underinterpreted, what are my present standards of intelligibility, why I believe that a discussion that does not comply with them holds little or no promise of being fruitful, etc. And I am ready to listen, and listen attentively, if my colleague will challenge my evaluation, criticize my standards of intelligibility and try to persuade me that the language of this or that speculative philosopher, or of all speculative philosophers, is one which it is worthwhile to adopt, at least for certain purposes. *But I am ready to listen and argue with him only if the (meta-) language, in which he explains to me his reasons for challenging my standards, itself complies with these standards.*

This may sound preposterous, but I don't see how it can be helped. My insistence is due to the fact that the situation is objectively asymmetrical. One cannot expect that the analytic philosopher, while endeavoring to persuade by rational means his speculative colleague of the cognitive poverty of his ways of philosophizing, should himself use speculative discourse for this purpose. This type of discourse is unintelligible to him in any capacity, including that of a metadiscourse. This does not mean that he is unable to intelligently manipulate such discourse for other purposes, should the occasion require it. On the other hand, no similar scruples could prevent the speculative philosopher from using scientific (observational plus theoretical) metalanguage to impress his analytical colleague with the importance of using metaphysical object-language.

Here then, it seems, the final *conditio sine qua non* of continued philosophical discussion has been reached. Those *speculative philosophers* who are interested in having analytic philosophers discuss their theses couched in metaphysical language *must use a scientific metalanguage as their* *rational tool of persuasion.* (I shall not discuss here, for obvious reasons, the possible use of extra-rational tools.) The analytic philosopher who is interested in having speculative philosophers stop formulating their theses the way they do, need do nothing beyond stretching his standards of intelligibility to the limit and offering constant reinterpretations in scientific language of the original metaphysical formulations. This includes occasionally reformulating these theses as proposals, comments, exhortations, etc. As I am using the term 'scientific language' now, such languages contain not only declarative sentences but also question sentences, etc.

I am under no illusions as to the effects of my proposal. Even if speculative philosophers should accept it in principle, there is little likelihood that agreement could be reached as to the extent of the *index verborum prohibitorum* for the common philosophical metalanguage. Differences will arise as to which terms of ordinary language are entitled to be considered directly intelligible. As to the theoretical terms, it is only with regard to rigorously-constructed language systems that it is precisely, though not necessarily effectively, determined whether they are or are not empirically significant, relative to the given observational language and after a certain criterion of significance has been adopted. Since nobody, and certainly not myself, seriously requires that a theoretical term of ordinary language should be admitted into philosophical metadiscourse only after this discourse has been completely formalized, it cannot be assumed that agreement would be reached on the admissibility of all candidates, in anticipation of future formalization. It is notorious that no such agreement has been reached as to the status of, say, certain psychoanalytical terms, and this among methodologists who would all of them be regarded as adherents of analytic philosophy. Similar disagreements exist as to the character of such terms in theoretical semantics as 'class', 'proposition' and the like, which some analytic philosophers simply claim not to understand. Add to this that a given term may doubtless be intelligible, perhaps even straight observational, in some of its uses (meanings) but of doubtful status in other uses (meanings), and this even in the discourse of one and the same author – think of what theologians call 'analogical use', for instance –, and the very great difficulties of agreeing upon a common metalanguage in which to discuss the relative merits of the various object-languages used by philosophers should be clear.

Nevertheless, I think that an agreement in principle to use for this purpose a language of about the same structure the analytic philosophers believe that the object-language of science has, should have its beneficial impact. One can only wish that prestige considerations ('having to talk the other fellow's language') will not blind the speculative philosophers in this context. I see no rational reason why they should refuse *a priori* to use a scientific metalanguage in order to justify their conviction that the scientific object-language is not adequate for certain philosophical purposes. Should they contend, however, that for intrinsic reasons such a metalanguage is not up to its purpose, then this would now indeed mean either the end of the conversation, or else the whole issue will just be pushed one step higher the hierarchy of philosophical metalanguages. The simple argument that a scientific metalanguage is unsuitable for the philosopher's use for the *same* reason for which a scientific object-language is unsuitable for this purpose would surely be a rather weak one, as can be shown by innumerable analogies. One can and does show in ordinary non-symbolic (meta-)language that ordinary non-symbolic (object-) language is unsuitable for algebraic purposes. Any contention that a discussion of the relative merits of various proposed notational systems for chemistry should be held in a metalanguage in which these notations are not only mentioned (which they clearly must be) but also used would meet with very strong initial disbelief. Though philosophy is neither chemistry nor algebra, this by itself is not a sufficient reason for rejecting the analogy. My plea is necessarily of a general and vague nature. No recipe follows from what I said. Any single word which the analytical philosopher professes not to understand (sufficiently) in an initial stage of a discussion can be made (sufficiently) intelligible by supplementing a few suitable sentences (though there constantly lurks the danger of *obscurum obscurior* in such situations). However, as to those notorious types of metadiscourse which analytic philosophers tend to regard as definitely underinterpreted *in toto*, it might be simpler to omit them altogether rather than to try giving a sketch of the theory behind them, which would at the best tend to become a very complex and time-consuming affair. But clearly no hard and fast rules of thumb can be expected. Goodwill helps. But unfortunately, though necessary, it is not sufficient.

The Hebrew University of Jerusalem

ABRAHAM A. FRAENKEL

EPISTEMOLOGY AND LOGIC

A. Church in his justified criticism [1]) of A. Koyré's analysis of antinomies without using the logistic method, and A. N. Prior in his rigorous treatment [2]) of the Epimenides paradox, concur in stressing that the point in this paradox is not the supposed emergence of contradictory conclusions but the apparent possibility of *settling an epistemological question on logical grounds.*

The following remarks, which do not pretend to produce an actual novelty, are to point out a somewhat reverse situation: the dependence of solving a logico-mathematical problem on the epistemological attitude in which the problem is viewed. The problem in question is fundamental both in mathematics and logic. Mathematically it is of modern origin, namely: *Cantor's continuum problem* of ascertaining the place of the power of the continuum in the series of alephs, a problem which was named in the first place by Hilbert in his famous lecture on unsolved mathematical problems at the International Congress of Mathematicians in 1900. Its logical aspect mainly corresponds to the problem, discussed from the Eleatic school to our days with increasing profundity but hardly increasing success: of spanning the abyss between discrete sequences of individual members and the amorphous, homogeneous pulp of continuity.

It is well known that, N denoting the set of all positive integers 1, 2, 3, ..., the continuum – for instance the set of all real numbers or the set of all points of a line – can be conceived as the set C the members of which are all subsets of N, i.e. all sets of positive integers. The purpose, then, is to ascertain the power (cardinal number) of the set C. The guess that this power is the second aleph $\aleph_1$ is the *continuum hypothesis.*

Since 1880 mathematicians have been trying in vain to solve the problem; for this purpose we have above all to determine what the subsets of N are

[1]) A. Church, Review of 'Alexandre Koyré, *The liar*'. *The Journal of Symbolic Logic*, vol. 11 (1946), p. 131.

[2]) A. N. Prior, *Epimenides the Cretan. Ibidem*, vol. 23 (1958), pp. 261–266.

and how they can be obtained, whether constructively or not. Disregarding as secondary those subsets, if any, which can be defined only by means of the axiom of choice, the 'normal' type of subsets will depend on a condition which is meaningful for positive integers; any such condition yields the uniquely-defined subset of those integers which satisfy the condition, whereas different conditions – for instance, of equalling 2, and of being an even prime number – may yield the same subset. (The principle of extensionality, whereby sets containing the same members are equal, is presupposed.) For short, the subsets are obtained by *comprehension*. But what are the conditions which furnish the subsets? Skolem's conception of 1923 is now almost generally accepted [1]), meaning the use of a well-formed formula of the first-order functional calculus with a free variable, built up from atomic membership statements. We need not enter into intricacies of this conception to perceive its *impredicative* character. For instance, the set of *all* subsets of N may occur among the constants entering a condition, without a certainty that a predicatively defined condition exists which is equivalent, i.e. which yields the same subset. Of the five attitudes distinguished by Wang [2]), the first four (anthropologism, finitism, intuitionism, predicativism) would refute the possibility of defining a subset in this way. Moreover, the Löwenheim-Skolem theorem states that only denumerably many subsets can be obtained whereas Cantor's theorem maintains that the set of all subsets of any set S has a higher power than S itself; in the present case, that the continuum is not denumerable. (For a profound but non-technical analysis of the Löwenheim-Skolem theorem and its ontological significance, see Myhill [3])).

At least the difficulties of impredicativeness will be 'solved' if we take a platonistic attitude in some sense of a pre-existence of the mathematical concepts, meaning that the mathematician does not *invent* his objects but rather *discover* them. (Cf. Beth [4]), e.g. pp. 407 and 465.) In this light a particular subset of S, determined by means of the totality of all subsets

[1]) Cf. A. A. Fraenkel and Y. Bar-Hillel, *Foundations of set theory* (Amsterdam, 1958), pp. 38–40.

[2]) Hao Wang, Eighty years of foundational studies, *Dialectica*, vol. 12 (1958), pp. 466–497.

[3]) J. R. Myhill, *Symposium: On the ontological significance of the Löwenheim-Skolem theorem. Academic Freedom, Logic and Religion* (Philadelphia, 1953), pp. 57–70.

[4]) E. W. Beth, *The foundations of mathematics*. Amsterdam, 1959.

of S, may be completely *described* though obviously it cannot be *constructed:* it is the relation between the subset and the totality (and possibly other constant objects) which, under suitable conditions, will constitute a unique determination of the subset. Accordingly, the non-predicative character of the definition will, in the eyes of platonists, not prevent the subset in question from being admitted together with subsets constructed in a predicative way.

In point of fact, even this large-minded attitude does not enable us to decide whether Cantor's continuum hypothesis is true or not, and with increasing urgency the conjecture has arisen that the hypothesis might be *independent* of the axioms of classical mathematics; that is to say, that both the continuum hypothesis and its negation were compatible with classical mathematics. The situation would then be somewhat analogous to that of the problem of *parallels* in the beginning of the nineteenth century: after efforts to prove Euclid's axiom of parallels, continued over more than two thousand years, the axiom turned out to be independent, which led to the bifurcation into Euclidean and non-Euclidean (hyperbolic) geometries. Similarly a proof of the independence of the continuum hypothesis should yield two different kinds of mathematics, a Cantorian and a non-Cantorian.

Kurt Gödel [1]) has contributed more than anybody else to a solution of the continuum problem, by showing that Cantor's hypothesis is *compatible* with a certain axiomatic system of classical mathematics. This naturally falls short both of 'solving' the problem and of proving the independence of the hypothesis; the latter would require the additional proof of the compatibility of the negation of the hypothesis.

Yet Gödel, leaning consistently on platonic realism, rather surprisingly denies [2]) the consequences to be drawn from a possible proof of the independence (which proof may be expected just of Gödel). For the concepts and axioms of mathematics, in particular of set theory, says Gödel, 'describe some well-determined reality. In this reality Cantor's conjecture must be either true or false, and its indecidability from the axioms as known today can only mean that these axioms do not contain a

[1]) K. Gödel, *The consistency of the axiom of choice and of the generalized continuum-hypothesis with the axioms of set theory.* Princeton, 1940 (revised ed. 1951).

[2]) K. Gödel, What is Cantor's continuum problem? *Amer. Mathem. Monthly*, vol. 54 (1947), pp. 515–525.

complete description of this reality.' As to this deficiency, Gödel adds that there may exist 'other (hitherto unknown) axioms of set theory which a more profound understanding of the concepts underlying logic and mathematics would enable us to recognize as implied by these concepts.'

Conceived in this light, the inaccessibility of the continuum problem is rather natural; the problem depends on our ability of defining the subsets of a given set, in this case of the set N of all positive integers, and we cannot expect to ascertain the number (power) of a totality of objects (subsets) as long as we do not know what these objects are.

At present we have no idea in what direction the 'missing' axioms required by Gödel might be sought and whether they refer to a mathematical reality yet unknown to us. However, some tentative proposals were made, by a philosophically-minded mathematician [1]) and a mathematically minded philosopher [2]), to take a further and seemingly more radical step towards platonism in order to 'solve' the problem of subsets. When a *finite* set F which contains n members is given then we can enumerate the 2^n subsets of F (including F and the empty set) by a 'combinatorial' procedure, namely by arbitrarily and independently assigning to each member of F one of two values, say 'yes' and 'no', and by relating to any such assignment the subset that contains just those members to which 'yes' has been assigned. The bold idea of transferring this procedure to infinite sets, for instance of obtaining all sets of positive integers by applying the procedure to the set N of all positive integers, was called by Bernays 'quasi-combinatorial' and was imbedded in conceptualism by McNaughton, who proposed hereby to justify the simple theory of types and Zermelo's set theory.

Tempting as such a solution might seem to non-intuitionistic mathematicians, it is thorny in at least three respects. First, there is no indication that this far-reaching description of subsets would at last pave the way for solving the question for which it was designed, namely the continuum problem. Secondly, among the subsets obtained by the quasi-combinatorial procedure, there is no more a discrimination visible between those

[1]) P. Bernays, Sur le platonisme dans les mathématiques. *L'Enseignement Math.*, vol. 34 (1935). pp. 52–69.

[2]) R. McNaughton, Conceptual schemes in set theory. *Philos. Review*, vol. 66 (1957), pp. 66–80.

definable constructively (say, by the principle of comprehension) and those which are only definable by an existential principle, for instance by the axiom of choice. Finally and more generally, assuming that conceptual schemes in set theory do serve logico-mathematical purposes such as the principle of comprehension and the continuum problem: then we have again, only in the reverse direction, a paradoxical situation similar to that pointed out above in the introduction, namely of settling logical questions on epistemological grounds.

The Hebrew University of Jerusalem

PAUL BERNAYS

ZUR ROLLE DER SPRACHE IN ERKENNTNIS-THEORETISCHER HINSICHT

In der Philosophie von Rudolf Carnap nimmt sein Werk 'Logische Syntax der Sprache' eine markante Stellung ein. Die hier entwickelte Konzeption der Wissenschaftslogik als Studium der Wissenschafts-sprache, mit den sich an sie knüpfenden Begriffen, bildet sozusagen den Ausgangsrahmen für Carnaps weitere Untersuchungen. Im Laufe dieser Untersuchungen hat er die Auffassungen, die er in der Logischen Syntax vertritt, erheblich revidiert, und auch jener Rahmen der Betrachtung selbst mit den zugehörigen Begriffsbildungen hat starke Wandlungen erfahren, wozu die Diskussionen mit den Philosophen verwandter Forschungsrichtung Wesentliches beigetragen haben.

Diese Schritte der Revision der Philosophie Carnaps bedeuten eine sukzessive Loslösung von den exklusiven und reduktiven Tendenzen des anfänglichen Programms der Wiener Schule, gegenüber dessen allzu simplifizierenden Thesen ja bereits die Logische Syntax bedeutsame Korrekturen brachte. Hier aber verfocht Carnap ja noch die Ansicht, daß alle Erkenntnistheorie, sofern sie Anspruch auf Wissenschaftlichkeit erheben kann, nichts anderes als Syntax der Wissenschaftssprache, bzw. als solche aufzufassen sei, während er seitdem die Aufgabestellung für die wissenschaftliche Philosophie durch die Hinzunahme der Semantik und der Pragmatik (unter Anknüpfung an C. W. Morris) wesentlich erweiterte und ferner der Unterscheidung des Logischen vom Deskriptiven den anderen Gesichtspunkt der Unterscheidung von theoretischer Sprache und Beobachtungssprache gegenüberstellte. Die Bedeutung, welche die Einführung dieser Erweiterungen des methodischen Rahmens für die Ausgestaltung der Philosophie Carnaps und auch für deren Annäherung an die gewohnteren philosophischen Auffassungen besitzt, soll im folgenden unter einigen Gesichtspunkten beleuchtet werden; zugleich soll dabei auf gewisse sich natürlich anschließende Fragen hingewiesen werden.

1

Die Anlage der Logischen Syntax kann als eine Erweiterung des Ansatzes der Hilbertschen Beweistheorie angesprochen werden. Bei Hilbert erstreckt sich die Methode der Formalisierung nur auf die Mathematik. Allerdings hatte Hilbert in seinem Vortrage 'Axiomatisches Denken' auch gesagt: 'Alles was Gegenstand des wissenschaftlichen Denkens überhaupt sein kann, verfällt, sobald es zur Bildung einer Theorie reif ist, der axiomatischen Methode und damit mittelbar der Mathematik'. Carnap geht in der Logischen Syntax einen Schritt weiter in dieser Richtung, indem er die Wissenschaft im Ganzen als ein axiomatisch-deduktives System betrachtet, welches mittels der Formalisierung zu einem mathematischen Objekt wird: die Syntax der Wissenschaftssprache ist die auf dieses Objekt gerichtete Metamathematik.

Das hierbei benutzte idealisierende Schema der Wissenschaft ist aber gewiß für die Erkenntnistheorie nicht zulänglich. Zunächst einmal stellt es ja nur das fertige Resultat der Wissenschaft dar, nicht den vollen Prozeß des wissenschaftlichen Geschehens. Wohl vermag bei den großen mathematischen Theorien eine axiomatisch-deduktive Präsentation der fertigen Disziplinen das Bedeutsame an ihnen hinlänglich zur Darstellung zu bringen. Doch bereits in der theoretischen Physik ist die Sachlage wesentlich anders, da hier die obersten Grundsätze der Theorie in ihrer mathematisch genauen Fassung für die Forschung meistens das Endergebnis und nicht den Ausgangspunkt bilden.

Außerdem aber ist ja für viele Gebiete der Forschung die Hervorkehrung des Deduktiven gewaltsam. Man verfährt in diesen Gebieten gar nicht deduktiv; vielmehr kommt hier das logische Schließen fast nur für die *heuristischen* Überlegungen zur Anwendung, durch welche die Aufstellung von Hypothesen oder von Tatsachen-Behauptungen motiviert wird. Mit der Hinzunahme der *Pragmatik* kann nun alledem Rechnung getragen werden. In die Pragmatik gehört sicherlich die Erörterung der Entwicklung der Wissenschaften, freilich nicht im Hinblick auf das historisch-Biographische, sondern im Sinne der Herausarbeitung des methodisch Bedeutsamen der Gedankengänge. Hier finden daher die heuristischen Betrachtungen ihre natürliche Einordnung.

Beiläufig sei hier daran erinnert, daß die Heuristik nicht nur in den empirischen Wissenschaften, sondern auch in der rein mathematischen

Forschung eine Rolle spielt, worauf in neuerer Zeit besonders nachdrücklich Georg Pólya hingewiesen hat. Es besteht ja eine methodische Analogie zwischen der mathematischen und naturwissenschaftlichen Forschung in der Hinsicht, daß es auch für die Mathematik eine Art von Empirie und ein Erraten von Gesetzlichkeiten aufgrund einer Reihe von Einzelfeststellungen gibt. Allerdings hat eine solche Art der Aufstellung eines Gesetzes in der Mathematik nur einen provisorischen Charakter, zumal in der Zahlentheorie, wo sich ja der Einzelfall niemals bloß durch unwesentliche Bedingungen (wie solche von Ort und Zeit in der Physik) aussondern läßt, vielmehr jede Zahl ihre besonderen Eigenschaften hat. Daß wir aber selbst in der Zahlentheorie aus unserem Umgehen mit den Zahlen Überzeugungen gewinnen können, zeigt das Beispiel des Satzes von der Eindeutigkeit der Zerlegung der Zahlen in Primfaktoren, den man von dem Zahlenrechnen her (wenn man noch keine zahlentheoretische Beweise kennengelernt hat) geneigt ist, für ganz selbstverständlich zu halten. Erst auf einer fortgeschrittenen Stufe macht man sich das Erfordernis eines Beweises für diesen Satz klar, dem ja dann in befriedigender Weise entsprochen wird.

2

Für die Betrachtung des Verhältnisses von Syntax und Semantik ist es nützlich, wenn wir uns vergegenwärtigen, daß nach der gewohnten Auffassung für eine Sprache als solche wesentlich ist, daß ihre Worte und Sätze eine unmittelbare Sinn-Bezogenheit haben. Wenn wir die Formbildungen einer Sprache losgelöst von der Bedeutung der Ausdrücke zum Gegenstand machen, so ist dieses eine bewußtermaßen vorgenommene, modifizierende Abstraktion.

In Carnaps Logischer Syntax wird die Ausschaltung des Sinnesmäßigen zu einem Teil dadurch ausgeglichen, daß er neben den 'Formbestimmungen' die 'Umformungsbestimmungen' als Regeln der Sprache statuiert. Zu diesen Umformungsbestimmungen für die Sprache einer formalisierten Theorie rechnet er nicht nur solche Regeln, nach denen ein Satz in einen ihm logisch gleichwertigen übergeführt werden kann, sondern allgemeiner alle solchen, nach denen sich logische Abhängigkeiten bestimmen, und ferner auch die Festsetzungen, wonach bestimmte Sätze die Rolle logisch allgemeingültiger Aussagen oder auch *formalisierter Axiome* haben.

Bald hernach hat Carnap, unter dem Einfluß der Untersuchungen von Alfred Tarski und im Zusammenhang mit der Erweiterung seines methodischen Programmes, den Begriff der logischen Folge aus der Syntax in die Semantik verwiesen.

In der Semantik werden den logischen Symbolen mittels der 'rules of truth' ihre Bedeutungen zugeordnet, und an diese Wahrheitsregeln knüpft sich der semantische Folgerungsbegriff. Von da aus kann das formale Deduzieren so eingeführt werden, daß zunächst Folgerungsbeziehungen teils als Sätze, teils als Ableitungsregeln vermerkt werden und dann die Mannigfaltigkeit solcher sich ergebender Sätze und Regeln einer Axiomatisierung unterworfen wird. Auf diese Weise wird der Begriff der Umformungsbestimmungen als ursprünglicher Regeln der Sprache grundsätzlich entbehrlich, während die 'rules of truth' als zur Charakterisierung der Sprache gehörig anzusehen sind.

Die hiermit ermöglichte prägnante Gegenüberstellung des semantischen und des syntaktischen Folgerungsbegriffes hat für die Darstellung der mathematischen Logik – sofern diese nicht von vornherein auf eine konstruktive Methodik ausgerichtet ist – große Vorteile, und besonders Heinrich Scholz hat diesen Gesichtspunkt sehr zur Geltung gebracht.

An der Semantik wird oft als Mangel empfunden, daß sie auf einer nicht konstruktiven Art der Begriffsbildung beruht. Diese Nicht-Konstruktivität ist aber für die Semantik nicht spezifisch. Eine Semantik kann an sich auch in einem elementaren Rahmen der Begriffsbildung betrieben werden. Andrerseits wird sich die Überschreitung der elementaren Begrifflichkeit, ob mit oder ohne Semantik, kaum vermeiden lassen, wenn man, wie Carnap es anstrebt, für die Logik einen solchen Begriff der 'Gültigkeit' festlegen will, aufgrund dessen für jeden rein logischen Satz *A* (d.h. einen Satz ohne außerlogische Bestandteile) nicht nur die Alternative '*A* oder nicht-*A*' logisch gültig ist (im Sinne des Satzes vom ausgeschlossenen Dritten), sondern darüber hinaus entweder die logische Gültigkeit von *A* oder diejenige von nicht-*A* besteht.

Die Semantik wird auch noch in anderer Hinsicht kritisiert, nämlich insofern sie den Bereich der umfangslogischen Betrachtung überschreitet und sich mit den Fragen des Sinnes und insbesondere mit dem Verhältnis des Extensionalen zum Intensionalen befaßt. So macht besonders Willard Quine geltend, daß mit der Einführung der Sinngehalte (Intensionen) von Ausdrücken als Gegenständlichkeiten eine wissenschaftlich unzulässige Hypostasierung vollzogen werde, und daß selbst bei der Reduktion der Fragen des Sinnes auf solche der Sinngleichheit und Sinnverschiedenheit, man sich noch immer in einem Gebiet des schwer Präzisierbaren befinde. Bei dieser Diskussion ist Quine mit Carnap darin einhellig, daß er tendiert, die Sinngleichheit zweier Aussagen als ihre logische Äquivalenz zu erklären und entsprechend auch die Sinngleichheit von Prädikaten und von Kennzeichnungen auf logische Äquivalenzen zurückzuführen. Dadurch tritt der Begriff der Sinngleichheit in enge Beziehung zu dem des Analytischen.

Eine solche Begriffsbestimmung von Sinngleichheit führt aber zu Unzuträglichkeiten, insbesondere sofern man, wie es ja Carnap und viele der heutigen Philosophen tun, die Sachverhalte der reinen Mathematik zu den logischen Gesetzen rechnet. Denn nach dieser Auffassung sind ja je zwei gültige Sätze der reinen Mathematik logisch äquivalent, und es müßten daher, wenn Sinngleichheit dasselbe wäre wie logische Äquivalenz, je zwei zutreffende Sätze der reinen Mathematik, also etwa der Satz, daß es unendlich viele Primzahlen gibt, und der Satz, daß die Zahl π irrational ist, denselben Sinn haben, – oder um ein elementareres Beispiel zu nehmen: der Satz, daß $3 \times 7 = 21$ ist, müßte denselben Sinn haben wie derjenige, daß 43 eine Primzahl ist.

Wir können uns aber für diese Überlegung sogar von der Stellungnahme zu der Frage des rein logischen Charakters der Arithmetik unabhängig machen. Nehmen wir ein Axiomensystem A und zwei ganz verschiedene Lehrsätze S, T, die aus diesem Axiomensystem beweisbar sind. Wir werden dann schwerlich bereit sein zu sagen, die Feststellung 'aus A folgt logisch S' habe denselben Sinn wie die Feststellung 'aus A folgt logisch T', auch wenn diese beiden Aussagen zutreffend, daher auch beide logisch gültig und somit einander logisch äquivalent sind.

Es fällt also keineswegs immer Sinngleichheit von Aussagen mit deren logischer Äquivalenz zusammen. Andrerseits wird man aber doch in vielen Fällen, auch in der Mathematik, eine logische Umformung als nicht Sinn-ändernd betrachten. Zum Beispiel die beiden Aussagen 'wenn a, b, c, n Zahlen der mit 1 beginnenden Zahlenreihe sind und $a^n + b^n = c^n$, so ist $n = 1$ oder $n = 2$' und 'es gibt nicht Zahlen a, b, c, n der mit 1 beginnenden Zahlenreihe, derart daß $n > 2$ und $a^n + b^n = c^n$' wird man als Formulierungen derselben mathematischen Behauptung (des großen Fermatschen Satzes) ansprechen.

An diesen Beispielen tritt uns zunächst die Schwierigkeit der Abgrenzung dessen, was jeweils als sinngleich zu gelten hat, entgegen. Zugleich aber bemerken wir, daß diese Schwierigkeit ihre Ursache in der Unterschiedlichkeit der Abstraktionsweise hat, welche den verschiedenen Untersuchungsgebieten eigentümlich ist. Zwei theoretisch-physikalische Feststellungen, von denen eine aus der anderen durch eine Umrechnung eines in ihr auftretenden mathematischen Ausdruckes hervorgeht, werden wir als sinngleich erklären; wenn es sich aber um mathematische Feststellungen handelt, ist das im allgemeinen nicht statthaft. Von der Formulierung eines mathematischen Satzes werden wir sagen, daß ihr Sinn durch eine elementar-logische Umformung nicht verändert wird; wenn dagegen die elementar-logischen Beziehungen selbst behandelt werden, dann gilt dieses nicht mehr.

Wir haben hier nur die Sinngleichheit von Aussagen betrachtet; jedoch für Prädikate und Kennzeichnungen läßt sich ganz das Entsprechende feststellen. Dabei liefert die Betrachtung mathematischer Kennzeichnungen ein reiches Maß an Beispielen, bei denen die Gegenüberstellung von Extension und Intension ganz im Sinne unserer üblichen wissenschaftlichen Denkweise liegt. Nehmen wir etwa die Darstellung einer positiv-reellen Zahl durch einen Ausdruck der Analysis, z.B. eine unendliche Reihe oder ein bestimmtes Integral. Eine solche Darstellung bildet eine Kennzeichnung der betreffenden reellen Zahl. Die Extension dieser Kennzeichnung ist die reelle Zahl selbst, und die Intension ist eine Regel zur Bestimmung dieser Zahl, d.h. zu ihrer Eingrenzung in beliebig enge Intervalle. Ein und dieselbe positiv-reelle Zahl kann, wie man weiß, durch sehr verschiedenartige solche Regeln bestimmt werden; dann haben wir gleiche Extension bei verschiedener Intension.

Um auch bei Prädikaten ein mathematisches Beispiel verschiedener Extensionen mit gleicher Intension zu erwähnen, so können ja die Primzahlen unter den von 1 verschiedenen Zahlen auf zweierlei Art charakterisiert werden: einerseits als solche, die keinen echten Teiler außer 1 besitzen, andrerseits als solche, die in einem Produkt nur dann aufgehen, wenn sie in mindestens einem Faktor aufgehen. Das ergibt zwei verschiedene Prädikat-Intensionen mit gleicher Extension: die Extension ist die Klasse der Primzahlen, die Intensionen sind die den beiden Charakterisierungen entsprechenden Definitionen des Begriffes 'Primzahl'.

Analoge Beispiele finden sich auch in empirischen Wissenschaften, z.B.

wenn eine Tierart auf verschiedene Weisen charakterisiert werden kann, so daß sich verschiedene Definitionen desselben Artbegriffes ergeben und somit verschiedene Intensionen des Art-Namens bei gleicher Extension.

Unsere Überlegung zeigt uns einerseits, daß es große Klassen von Fällen gibt, in denen der Begriff der Intension eine wissenschaftlich naturgemäße Anwendung hat. Andrerseits sind wir auf die Schwierigkeiten im Begriff der Sinngleichheit aufmerksam geworden, die mit der Verschiedenartigkeit der Einstellung in den verschiedenen Forschungsgebieten zusammenhängen, wobei die Gegenüberstellung bloß des Logischen und des Außerlogischen nicht genügt, um diesen Unterschieden Rechnung zu tragen.

Wir können uns den in dieser Hinsicht vorliegenden Sachverhalt näher bringen, indem wir uns die Art der Abstraktion vergegenwärtigen, auf welche es bei dem Begriff der Intension ankommt. Hier geht man nicht aus von der Absonderung der Sprachausdrücke als Formgebilde von ihrer Ausdrucksfunktion, vielmehr behält man diese geflissentlich bei, und wovon man abstrahiert, das sind nur die für diese Funktion unwesentlichen Besonderheiten der Ausdrucksmittel und die auf ihnen beruhende Vielfältigkeit von Formulierungen, welche für den gleichen Ausdruckszweck verwendbar sind. Diese Mannigfaltigkeit der Möglichkeiten besteht einerseits in konventioneller Hinsicht, durch die Vielheit der Sprachen, andrerseits aber aufgrund begrifflicher und sachlicher Gleichwertigkeiten, wie sie zwischen Bestimmungen, zwischen Eigenschaften und zwischen Beziehungen bestehen können. Eine solche Gleichwertigkeit bewirkt aber nur dann die Vertretbarkeit eines Ausdruckes durch einen anderen, wenn sie im Rahmen der Darlegung oder der Untersuchung, in welcher der Ausdruck verwendet wird, ganz unproblematisch ist, d.h. zu dem Bereiche dessen gehört, worüber man nicht erst diskutiert, sondern was als ausgemacht gilt. In der Tat liegt ja bei unseren Erkenntnisbemühungen, wenigstens im Stadium eines entwickelten Reflektierens, jeweils ein gewisser Vorrat an (teils mehr, teils minder bewußten) Vorstellungen, Ansichten und Überzeugungen zugrunde, an welche wir uns bei unseren Fragen, Überlegungen und Verfahren halten, sei es mit Wissen oder instinktiv. Solche Vorstellungen, Ansichten und Überzeugungen mögen, im Anschluß an Ferdinand Gonseths Begriff *préalable*, als 'vorgängig' bezeichnet werden.

Die Annahme gewisser vorgängiger Vorstellungen und Voraussetzungen für eine jegliche wissenschaftliche Disziplin, und auch für unsere natürliche Einstellung des täglichen Lebens, unterliegt nicht der gleichen Problematik wie die Annahme von Erkenntnissen *a priori*. Es wird nicht behauptet, daß die vorgängigen Voraussetzungen etwas Unumstößliches seien. Eine Wissenschaft, die sich zunächst auf eine Voraussetzung stützt, kann in ihrem weiteren Verlauf uns dazu führen, diese Voraussetzung preiszugeben, wodurch wir eventuell genötigt werden, die Sprache der Wissenschaft zu ändern. Auch bringt es die wissenschaftliche Methodik mit sich, daß wir vorgängige Voraussetzungen uns zum Bewußtsein bringen und sogar zum Gegenstand einer Untersuchung machen, bzw. in das Thema einer Untersuchung einbeziehen können.

Damit verlieren dann diese Voraussetzungen für das betreffende Forschungsgebiet den Charakter der Vorgängigkeit. Im Verlauf der Entwicklung der theoretischen Wissenschaften kommt es so dazu, daß immer mehr von den Voraussetzungen der Thematisierung unterworfen werden, so daß sich der Bereich des Vorgängigen immer mehr verengert.

An die Stelle des früheren, spontanen Vorgängigen treten dann eigens statuierte Ausgangs-Begriffe und Prinzipien.

Im Unterschied zum Begriff des *a priori* ist der des Vorgängigen entweder auf einen Erkenntniszustand oder auf eine Disziplin bezogen; es wird nicht etwas absolut Vorgängiges angenommen.

Wenn man nun diesen Begriff des Vorgängigen akzeptiert, so kann man folgende Definition der Sinngleichheit ansetzen: zwei Aussagen einer Disziplin sind sinngleich, wenn die Äquivalenz zwischen ihnen für die Disziplin vorgängig ist. Entsprechend würde die Sinngleichheit von Prädikaten und die von Kennzeichnungen zu erklären sein. Auch kann in der Definition die Disziplin durch eine Erkenntnislage (Erkenntniszustand) ersetzt werden, in Bezug auf welche man in genügend bestimmter Weise von Vorgängigem sprechen kann.

Es möchte scheinen, daß sich auf diese Art die vermerkten Schwierigkeiten in der Bestimmung von Sinngleichheit beheben lassen. Freilich muß man bei der gegebenen Erklärung der Sinngleichheit in Kauf nehmen, daß die Sinngleichheit von Sätzen von der Disziplin, bzw. der Erkenntnislage abhängt, im Rahmen deren sie betrachtet wird. Das ist aber bei näherem Zusehen nicht so paradox.

3

Wenden wir uns nun zu derjenigen Erweiterung des methodischen Rahmens der Logischen Syntax, welche Carnap durch die Gegenüberstellung von theoretischer Sprache und Beobachtungssprache gewinnt.

Theorie und Experiment einander gegenüberzustellen, ist uns bei der Betrachtung der Methode der Naturwissenschaften geläufig. Doch in der anfänglichen Form des logischen Empirismus kam das Moment des Theoretischen nicht recht zur Geltung; und es haben erst Diskussionen über die anfängliche Auffassung, an denen insbesondere Karl Popper beteiligt war, dazu geführt, daß bei dem revidierten Standpunkt der Logischen Syntax der Ansicht der Vorzug gegeben wurde, wonach die Formulierungen von Naturgesetzen als eigentliche Sätze der Wissenschaftssprache figurieren.

Daß sich hiergegen anfangs eine Resistenz richtete, begreifen wir, wenn wir uns klar machen, daß mit der Anerkennung der Rolle der physikalischen Gesetzesaussagen als eigentlicher Sätze jene Zweiheit von *relations of ideas* und *matters of facts*, wie sie ehemals David Hume als Einteilung aller Gegenstände des Forschens angesetzt hatte, und wie sie die Wiener Schule in etwas präzisierter Form aufrecht zu erhalten bestrebt war, sich als nicht erschöpfend erweist. Die Gesetzesaussagen der Naturwissenschaft sind ja einerseits nicht Aussagen über *relations of ideas*, d.h. nicht Sätze der reinen Logik oder der reinen Mathematik, andrerseits sind sie nicht Feststellungen von Tatsachen, da sie doch die Form allgemeiner hypothetischer Sätze haben.

In der Ausdrucksweise Carnaps besagt diese Konsequenz, daß der Bereich des Deskriptiven (des Außerlogischen) nicht mit dem des Faktischen zusammenfällt, daß vielmehr der Bereich des Faktischen enger ist.

Der gleiche Sachverhalt läßt sich noch von einer anderen Seite her beleuchten. Carnap erklärt in seinem Buche 'Meaning and Necessity' den Begriff des logisch Wahren mit Hilfe von '*state descriptions*'. Dabei knüpft er an den Leibnizschen Gedanken der 'möglichen Welten' an: was notwendig ist, muß in allen möglichen Welten gelten; und die *state descriptions* stellen schematisch die möglichen Weltbeschaffenheiten dar. So definiert nun Carnap: Ein Satz ist logisch wahr, wenn er für jede *state description* zutrifft. Bei dieser Überlegung treten die Begriffe des

Möglichen und des Notwendigen auf. Es ist aber nicht ausgemacht, daß man vom Möglichen und vom Notwendigen nur im logischen Sinne sprechen kann. Carnap selbst erwähnt im Anhang zu seiner 'Introduction to Semantics' (§ 38 d, S. 243) unter den für die Semantik ausstehenden Problemen das Studium solcher nicht-extensionaler Operatoren, welche physikalische oder kausale Modalitäten zum Ausdruck bringen. Physikalische oder kausale Modalitäten betreffen das naturgesetzlich Mögliche und das Naturnotwendige. Wenn nun im Rahmen der Wissenschaftssprache Naturgesetze als gültig statuiert werden, und wenn ferner anerkannt wird, daß die Naturgesetze nicht logisch notwendig sind, so ergibt sich eine Unterscheidung des Notwendigen und des Tatsächlichen, welche von derjenigen zwischen dem Logischen und dem Deskriptiven verschieden ist. Wir können dann *state descriptions* in einem engeren Sinne betrachten, indem wir nur solche zulassen, die den Naturgesetzen gemäß sind, und erhalten damit eine engere Mannigfaltigkeit von möglichen Welten.

Den Feststellungen von Faktischem stehen somit nicht nur diejenigen von logischen Gesetzlichkeiten gegenüber, sondern allgemeiner von Gesetzlichkeiten überhaupt. Diese allgemeinere Entgegensetzung können wir nun mit Hilfe des Begriffs des Theoretischen zum Ausdruck bringen, indem wir den Feststellungen über Tatsächliches die theoretischen Feststellungen gegenüberstellen. Der Bereich des Theoretischen enthält dann als echten Teilbereich den des Logischen.

Das Spezifische des Theoretischen besteht aber gewiß nicht allein in einer Gesamtheit von Aussagen, welche als gültig anerkannt werden, sondern vor allem in einer Begriffswelt, im Rahmen derer die theoretischen Aussagen erfolgen. Innerhalb der Wissenschaftssprache findet die theoretische Begriffsbildung ihren Niederschlag in dem, was Carnap die 'theoretische Sprache' nennt.[1])

Betrachten wir nun des Näheren die Rolle, welche Carnap der theo-

[1]) Wenn hier, in Anlehnung an Carnaps Ausdrucksweise schlechtweg von 'der theoretischen Sprache' die Rede ist, so soll damit nicht auf die Vorstellung von einer Gesamtwissenschaft Bezug genommen sein. Auch in Carnaps eigenen Ausführungen zum Thema der theoretischen Sprache ist das keineswegs der Fall. So spricht er von den 'methodologischen Problemen, die mit dem Aufbau eines theoretischen Systems, etwa eines solchen der theoretischen Physik ... zusammenhängen' ('Beobachtungssprache und theoretische Sprache' Dialectica 47/48, S. 241-242).

retischen Sprache zuweist. Nach seiner Auffassung ist die theoretische Sprache nicht unmittelbar gedeutet, vielmehr erhalten die theoretischen Termini ihre Signifikanz erst in Verbindung mit den 'Korrespondenz-Postulaten', welche zwischen theoretischen Termen und Beobachtungstermen Beziehungen herstellen. Diese Beziehungen sind allerdings nicht als so weitgehend gedacht, daß dadurch alle theoretischen Termini in der Beobachtungssprache definiert würden. Carnap schließt sich vielmehr der Auffassung derer an, welche die Forderung, daß jeder theoretische Terminus sich experimentell definieren lassen müsse und in seiner Verwendung an eine solche Definition gebunden sei, als zu einschränkend für die theoretische Forschung und auch nicht dem tatsächlichen Verfahren der theoretischen Wissenschaft entsprechend erklären, wie es im Kreise des Neupositivismus insbesondere Herbert Feigl und Carl Hempel getan haben.

Hiermit wird ein wesentliches Erfordernis für die Freiheit der theoretischen Gedankenbildung anerkannt. Es bleibt dabei aber doch der Umstand, daß die Theorie nicht als eine Gedankenwelt, sondern bloß sozusagen als eine Sprach-Apparatur angesehen wird. Zu diesem mehr nur technischen Aspekt, den die theoretische Sprache bei Carnap erhält, tritt als ein anderes charakteristisches Moment dasjenige der Reduktion auf das rein Mathematische. Carnap ist bestrebt, nach Möglichkeit die theoretischen Entitäten auf mathematische zu reduzieren. Im Gebiet der Physik zeigt sich diese Möglichkeit in spezieller Art anhand der Vorstellungsweise der Feldtheorie, wonach das physikalische Geschehen in einer Abfolge von Zuständen im Raum-Zeit-Kontinuum besteht. Die Zustandsbestimmung wird durch Skalare, Vektoren und Tensoren gegeben. Z.B. in der reinen Feldtheorie der Gravitation und der Elektrizität erfolgt die Beschreibung des physikalischen Zustandes durch den symmetrischen Tensor des metrischen Feldes, aus dem sich die Längen- und Zeitmessung sowie die Trägheits- und Gravitationskräfte bestimmen, und den antisymmetrischen elektromagnetischen Tensor, der die elektrischen und magnetischen Kräfte bestimmt. Materielle, geladene oder ungeladene Teilchen werden hier als besonders konzentrierte Verteilungen der Feldgrößen in einem räumlich engen Weltgebiet aufgefaßt. Die Komponenten der Tensoren sind Funktionen der Raum-Zeit-Stellen, und bei Einführung eines Koordinatensystems und Wahl von Einheiten werden die Maßzahlen der Komponenten mathematische Funktionen der Raum-Zeit-Koordinaten [1]); nennen wir sie 'Feldfunktionen'. Die physikalische Feldgesetzlichkeit wird durch Differentialgleichungen für diese mathematische Funktionen (in einer gegenüber dem Koordinatensystem invarianten Weise) formuliert, und die Feldfunktionen, welche den Ablauf der Zustände des Systems darstellen, bilden eine Lösung dieses Systems von Differentialgleichungen.

Die Anknüpfung der Theorie an die Erfahrungswirklichkeit wird durch Beziehungen von mehrerlei Art gegeben:

1. solche, auf denen die Einführung von Raum-Zeit-Koordinatensystemen beruht, sowie die Möglichkeiten der Bestimmung von Werten der Feldfunktionen,
2. solche, welche die Auswirkungen von Zuständen des Systems teils auf unsere direkten Wahrnehmungen, teils auf unsere experimentellen Beobachtungen betreffen,
3. solche, die jeweils die Anweisung liefern für die theoretische Übersetzung eines beobachtungsmäßig (sei es nur schematisch oder aber in genauerer experimenteller Bestimmtheit) gegebenen Falles, der mittels der Theorie untersucht werden soll.

Alle diese Beziehungen denkt sich Carnap axiomatisierbar durch Korrespondenz-Postulate in denen Verknüpfungen zwischen den Feldfunktionen und unseren Beobachtungen ausgesagt werden. Ein solches System von Korrespondenz-Postulaten läßt sich jedenfalls dann aufstellen, wenn überhaupt die Mannigfaltigkeit der möglichen Anwendungen der Feldtheorie (der Differentialgleichungen des Feldes) auf Beobachtungen axiomatisierbar ist.

Unter diesem Vorbehalt besteht somit die Möglichkeit, die theoretische Sprache der Physik gänzlich auf mathematische Begriffe zu beschränken und alles spezifisch Physikalische teils in die Beobachtungssprache, teils in die Korrespondenz-Postulate zu verlegen. Die physikalische Theorie sagt dann nichts mehr aus über etwas, das in der physikalischen Natur vorhanden ist, ja sie sagt für sich allein überhaupt nichts aus, sondern liefert nur eine mathematische Handhabe für die Vorausbestimmung von Beobachtungen aufgrund von gegebenen Beobachtungen. Man kann hier streng genommen daher gar nicht von einer theoretischen Sprache reden.

[1]) Die Komponenten des metrischen Feldes sind ja sogar von vornherein unbenannte Zahlen.

Allerdings läßt sich dabei doch eine Art theoretischer Sprache wiedergewinnen, indem man geeignete physikalische Benennungen einführt für gewisse häufig wiederkehrende mathematische Beziehungen und Ausdrücke, in Entsprechung zu den Bedeutungen, welche diese in der inhaltlich aufgefaßten Theorie haben; das Verfahren ist dann analog, wie wenn in einer rein arithmetisch konstituierten (analytischen) Geometrie doch die arithmetischen Beziehungen und Gegenständlichkeiten geometrisch interpretiert und benannt werden.

Was aber an der beschriebenen Methode der Elimination theoretischer Entitäten stutzig macht, ist der Umstand, daß sie ja gleichermaßen auf jedwede Art des Ansetzens von Naturgegenständen anwendbar ist: Wenn in den geläufigen Fällen des täglichen Lebens die Annahme von Naturgegenständlichkeiten sachgemäß ist, und wenn wir ferner unsere geläufigen Methoden der Orientierung über Ort und Zeit zu der Vorstellung der vier-dimensionalen Raum-Zeit-Mannigfaltigkeit extrapolieren, so scheint es nicht angängig zu sein, daß wir in dem Ansetzen von Naturgegenständlichkeit an gewissen Stellen sozusagen abbrechen und hier die Gegenstände durch ihre mathematischen Beschreibungen ersetzen. Dieser Erwägung gegenüber kann jedoch Carnap geltend machen, daß die Unterschiedlichkeit der methodischen Behandlung nicht die Verschiedenheit von Stellen in der Raum-Zeit-Mannigfaltigkeit betrifft, sondern sich auf die Verschiedenheit in der Stufe des Theoretischen bezieht. Was mit einer solchen Verschiedenheit der Stufe gemeint sein soll, läßt sich insbesondere an dem Unterschied der Makro- und der Mikrophysik exemplifizieren. Allgemein liegt eine weitergehende Stufe des Theoretischen bei der Behandlung eines Wissensgebietes da vor, wo die Begriffsbildung zu einer stärkeren Überschreitung des anschaulich Vertrauten nötigt. Ein solcher Schritt der verstärkten Theoretisierung kann erfolgreich sein und sich als sachgemäß erweisen, und es kann sich auch im Umgehen mit den zuerst ungewohnten Begriffen nach reichlicherem Gebrauch eine praktische Sicherheit einstellen. Dabei bleibt aber doch der Unterschied bestehen zwischen dem methodisch mehr und dem weniger Elementaren, d.h. zwischen dem, was dem Konkreten und der Beobachtung näher und dem, was ihnen ferner steht.

Daß die Quantenphysik gegenüber der vorherigen 'klassischen' Physik eine verstärkte Theoretisierung in dem genannten Sinn bedeutet, ist ersichtlich. Auf die Quantenphysik kann allerdings das vorher beschriebene Verfahren der Eliminierung von Entitäten nicht direkt angewendet werden, da bei dieser ja die Vorstellung von der eindeutigen, unabhängig vom Experimentieren objektiv bestimmten Abfolge der Zustände in der Raum-Zeit-Mannigfaltigkeit verloren geht. In anderer Hinsicht kommt aber die Quantenphysik der Absicht des Eliminierungsverfahrens insofern entgegen, als hier die Vorstellung von der Gegenständlichkeit ohnehin eine abgeschwächte ist und das Mathematische der Begriffsbildungen im Vordergrund steht. Die Quantenphysik zeigt uns auch, auf welche Art sich die unterschiedliche methodische Behandlung verschiedener theoretischer Entitäten ohne eine anstößige Zäsur durchführen läßt, indem hier die theoretische Sprache der vorherigen Physik sozusagen die Rolle der Beobachtungssprache erhält.

Hierdurch wird zugleich der Gedanke nahegelegt, daß es wohl angemessen ist, die Unterscheidung zwischen Beobachtungssprache und theoretischer Sprache, anstatt sie absolut zu fassen, auf ein Niveau der Begriffsbildung zu beziehen. Darin werden wir bestärkt, wenn wir uns überlegen, was es in der Wissenschaftspraxis mit der Beobachtungssprache für eine Bewandtnis hat. Wenn sich die Physiker von ihren Experimenten erzählen, so sprechen sie gewiß nicht nur von Objekten der unmittelbaren Wahrnehmung. Man spricht etwa von einem Stück Holz, von einer Eisenstange, einem Gummiring oder einer Quecksilbersäule. In den Bedeutungen solcher Worte sind ja bereits beträchtliche theoretische Momente enthalten. Die Experimentalsprache der Physiker geht aber doch in dieser Hinsicht noch viel weiter.[1]) Bemerkenswert ist auch, daß die Namen der physikalischen Begriffe für einen großen Teil (etwa 'Luftdruck', 'elektrischer Strom') in die gewöhnliche Umgangssprache eingegangen sind.

Im Ganzen läßt sich der Sachverhalt dermaßen charakterisieren, daß

[1]) Allerdings ist ja die These aufgestellt worden, daß alles physikalische Experimentieren auf Feststellungen über Koinzidenzen hinauskomme. Diese Behauptung ist aber gewiß nur *cum grano salis* zu verstehen: Die Feststellung über Koinzidenz (oder Nicht-Koinzidenz) ist jeweils der letzte entscheidende Schritt in dem Gesamtprozeß eines Experimentes, welches überdies erfordert, daß der Experimentator seinen Apparat als solchen erkennt und in der richtigen Weise mit ihm umgeht, ferner daß dieser Apparat sachgemäß angefertigt worden ist, weiter daß der Experimentierende sich eine hinlängliche Überzeugung davon verschafft, daß keine störenden Umstände vorliegen, u.s.w. Daß alles das, was hierzu aufgefaßt werden und eingeübt sein muß, sich auf bloße Feststellungen über Koinzidenzen zurückführen läßt, dürfte wohl schwerlich zutreffen. Aber das ist ja auch wohl mit jener These nicht gemeint.

die Beobachtungssprache einer auf einem bestimmten Niveau befindlichen Wissenschaft Bezug nimmt auf eine vorgängige Vorstellungs- und Begriffswelt, – 'vorgängig' gemäß der in unserm Abschnitt 2 eingeführten Ausdrucksweise. Die vorgängigen theoretischen Begriffe erhalten auf diesem Niveau auch ihre Benennungen in der Beobachtungssprache. Wir brauchen wohl die Beobachtungssprache überhaupt nicht von der Umgangssprache zu trennen. Vielmehr kann vermutlich die Beobachtungssprache als eine durch Hinzufügung einer größeren Reihe von Termini bereicherte Umgangssprache aufgefaßt werden.

Die Relativierung der Beobachtungssprache auf ein begriffliches Niveau wird auch jener Art der Gegenüberstellung des Empirischen und des Theoretischen gerecht, wie sie in Ferdinand Gonseths Prinzip der Dualität intendiert ist. Gemeint ist hier, daß das Empirische und das Theoretische nicht getrennte Bereiche sind, sondern daß in jedem Gebiete und jedem Stadium des Erkennens die beiden Momente zusammenspielen.

Die verschiedenen Gesichtspunkte der im Vorangehenden angestellten Überlegungen: die Eliminierung abstrakter Entitäten, die Unterscheidung von Stufen des Theoretischen und die Relativierung der Beobachtungssprache auf ein begriffliches Niveau haben ihre Anwendung im besonderen für die mathematische Beweistheorie. Diese geht ja aus von der Unterscheidung zwischen der 'klassischen' Methode der Mathematik, wie sie in der Analysis und der Mengenlehre sowie in den neueren abstrakten Disziplinen der Mathematik angewendet wird, und den elementareren Methoden, welche je nach der Art der Abgrenzung als 'finite', 'konstruktive', oder 'prädikative' zu charakterisieren sind. In der beweistheoretischen Untersuchung der klassischen Mathematik wird durch die Methode der Formalisierung der Aussagen und Beweise, wie sie mittels der logischen Symbolik erfolgt, eine Eliminierung abstrakter Entitäten ermöglicht. Diese Eliminierung will man insbesondere dazu verwerten, um die formale Widerspruchsfreiheit klassischer Theorien von einem der elementareren methodischen Standpunkte nachzuweisen. Bisher sind Nachweise der formalen Widerspruchsfreiheit mittels konstruktiver Methoden nur für solche formalen Systeme erbracht worden, die wenigstens einer prädikativen Deutung fähig sind. Neuerdings scheint durch ein Verfahren von Clifford Spector mit einer weiten Fassung des konstruktiven Standpunktes ein Nachweis der Widerspruchsfreiheit für die formalisierte imprädikative Analysis zu gelingen.

Die elementarere 'Metasprache', in der ein solcher Nachweis der Widerspruchsfreiheit geführt wird, hat, wie von Carnap vermerkt worden ist, die Rolle einer Beobachtungssprache. Es war ursprünglich die Idee Hilberts, daß diese Sprache sich ganz im Rahmen der Betrachtung des Konkreten halten, also eine Beobachtungssprache im absoluten Sinne sein sollte. Schrittweise wurde man aber genötigt, mehr und mehr Theoretisches in sie aufzunehmen. Bereits der 'finite Standpunkt' verwendet grundsätzlich mehr, als Hilbert ursprünglich zulassen wollte; doch auch dieser methodische Standpunkt hat sich, aufgrund der Resultate von Kurt Gödel, für den gesetzten Zweck als nicht zulänglich erwiesen. Das Ergebnis dieser Feststellung erscheint als nicht so fatal für die Beweistheorie, wie es anfangs angesehen wurde, wenn man den Gedanken der Bezogenheit der Beobachtungssprache auf ein Begriffsniveau akzeptiert. Die Anerkennung der methodischen Bedeutsamkeit der beweistheoretischen Untersuchungen, und insbesondere derjenigen über formale Widerspruchsfreiheit, ist nicht daran gebunden, daß man die übliche klassische Mathematik für dubios erachtet oder daß man jenen Standpunkt des 'Formalismus' einnimmt, wonach die klassische Mathematik nur als reine Formeltechnik ihre Berechtigung hat. So dachte auch Hilbert im Grunde nie, trotz mancher in solche Richtung weisender Äußerungen von ihm. – Die Aufgabestellung der konstruktiven Nachweise von Widerspruchsfreiheit ist durch die hohe Stufe des Theoretischen motiviert, wie sie in der klassischen Mathematik vorliegt.

Jedenfalls kann ein Angehöriger der konstruktiv-beweistheoretischen Forschungsrichtung sehr wohl die Ansicht vertreten, wie sie auch von Carnap befürwortet wird, daß die Begriffsbildungen der klassischen Mathematik auch als inhaltlich gedeutete ihre berechtigte Anwendung haben. Ob es aber angemessen ist, die sämtlichen von der Mengenlehre eingeführten Entitäten als eigentliche zu akzeptieren, steht auch von diesem Standpunkt zur Diskussion. Auch wird man nicht geneigt sein, die positive Stellung zu den theoretischen Begriffen gerade bloß der mathematischen Begriffsbildung als Privileg zuzuerkennen: Was den mathematischen Klassen und Funktionen recht ist, ist den Entitäten der Naturwissenschaften billig, soweit diese in Verständnis-erzeugender Weise angesetzt sind.

Zürich

JØRGEN JØRGENSEN

SOME REMARKS CONCERNING LANGUAGES, CALCULUSES, AND LOGIC

I am writing this paper lying with a coronal thrombosis in a hospital bed. As I have been told not to make any unnecessary movements Carnap will, I hope, draw the conclusion, that I have found it necessary to try to give a contribution to his 70th birthday as an expression of my gratitude for his stimulating works and for the friendliness he has shown to me during the time, of peace and war, we have known each other.

My present condition gives me plenty of time to think about philosophical problems but, unfortunately, very small possibilities for controlling my thoughts by means of printed matter and oral discussion. It is therefore not without some misgivings that I send this small article to 'Synthese', whose editors have kindly asked me to contribute to this special issue, – and my misgivings are so much heavier as I do not feel quite sure whether the following remarks are in contradiction with Carnap's views or whether they are merely supplementary to it. Anyway, I wish him many happy returns. –

Few subjects have, I think, interested Carnap more than language and the relation between language and logic. And his conception of language (in general) seems to have changed very little since he in his 'Logical Syntax of Language' (London, 1937) defined: 'By a language we mean here in general any sort of calculus, that is to say, a system of formation and transformation rules concerning what are called *expressions*, i.e. finite, ordered series of elements of any kind, namely, what are called *symbols*' (pp. 167–168) – right up to his 'Introduction to Symbolic Logic and its Applications' (New York ,1958) where he writes: '. . . a *language* (i.e. a system of signs and of rules for their use)' (p. 1). In both cases he seems to consider languages as calculuses, and calculuses as systems of something ('symbols' or 'signs') of which one need merely know the formal rules of formation and transformation in order to construct the languages in question.

Now, to my mind, this seems to be a very peculiar way to use the word 'language'. Even if this usage has become habitual to some mathematicians and logicians it is, I am sure, very different from the sense in which the word 'language' is used in everyday usage as well as in the writings of linguists. And as this difference in the use of the word-image 'language' has, in my opinion, caused serious misconceptions as to the relations between languages, calculuses, and logic, I would make a few remarks that may, I hope, contribute to clear up the matter a little – even if they probably leave more questions open than they answer. In passing, I must state that these troubles cannot be eliminated by the distinction between the syntactical, the semasiological, the psychological, and the sociological aspects of language that is mentioned in 'Logical Syntax of Language', p. 5.

First and foremost I would stress that a calculus is *not* a language nor a language a calculus in any of the current senses of these words, – one essential difference between a calculus and a language being that a calculus need not have a meaning, but may be a mere game with shapes or sounds according to rules that are conventional in principle, while it is *essential* to a language to have meanings, to refer to something different from the expressions of the language concerned, or to express some content of consciousness, – to which differences between calculuses and languages may be added that the syntactical rules of languages are always empirical regularities (with many or few exceptions) and not invariable conventional stipulations as are the rules of formation and transformation of any calculus. And, considered from the other side, it is essential to a calculus that its elements and expressions may always be identified as belonging to the calculus concerned by means of its rules of formation or transformation, while such rules (if they exist at all for languages that are not merely interpretated calculuses) are never sufficient to identify the elements or expressions of a language because it is always necessary to take the meaning in consideration.

Before trying to elucidate this a little further I would add the following remark: to my mind it is of fundamental significance to note that what can be sensed or observed (in the usual intersubjective manner) of a language, whether purely logical or not, is merely one side or aspect of it, viz. the expressive aspect, whereas another, viz. the 'content' or 'meaning', can never be observed or discovered by observation of the 'expressions'. Indeed, the 'expressions' are not expressions of anything, if they have not got a content or meaning. And if they do not express anything they have nothing to do with language (or languages) but are mere sensorial non-linguistic phenomena (shapes or sounds).

In the following I will try to say a little more about the content or meaning that is expressed by means of the linguistic expressions. At the present moment I will, however, restrict myself to the suggestion that what is expressed in a linguistic expression is always some content of the consciousness of the speaker (or writer) as e.g. a feeling, a need, a perceptual content, a thought, an attitude, etc. Such contents of consciousness are what the speaker 'means' by his expressions – whether they are understood by others or not, and whether they are intersubjective or not. As a matter of fact, however, many systems of mutual understandable expressions of intersubjective perceptual contents and of common ideas and feelings have developed, and such systems are what we usually call 'natural languages' by means of which the so-called artificial or constructed languages are usually explained and learned.

The syntactical rules of (natural) languages are neither conventional stipulations nor invariable norms but empirically observable regularities (with a greater or smaller number of exceptions) concerning the formation of the expressions of the various languages. To be true, *if* these regularities are being fixed voluntarily they change into conventional norms that may not be changed except by voluntary decision. But at the same time the respective 'languages' change into calculuses that are not able to 'live' and develop as natural languages do. The syntactical or grammatical rules of a (natural) language are in fact, I think, merely more or less fossilized habits of expressing certain 'meanings' (e.g. cases of nouns, moods of verbs, modalities of sentences, etc.). But they change, although slowly, during the development of the languages, the creating factors presumably being original speakers or writers finding new and apt expressions of old or new thoughts or impressions, or children not yet having learnt to speak 'correctly', or creative minds in the various fields of mental life, while the teaching of the present stage of language (including scientific language), journalism, everyday talk, slogans, etc. are stereotyping and fossilizing factors. So two important tendencies are mutually frustrating and working against each other: the one being the striving for originality, new thoughts and points of view, new discoveries, new impressions, and corresponding new expressions of all these experiences, – the other being the striving for easy mutual understanding and communication, for standardized or unified thinking and feeling, for generalizations as conditions of inferences, and moreover for all that tends towards calculatory thinking. –

If now a language is given, i.e. if a system of expressions of contents of consciousness is given, then it is, of course, possible to abstract from the content and concentrate solely on the formal or expressive aspect, – and this is apparently what Carnap generally does. But this approach seems to be fatal in the respect that it leads to overlooking the fact that the formal laws that may be found ruling the constructions of *expressions* are conditioned by the relations existing between the elements of *content*, and can therefore neither be found by investigation of the expressions alone (one and the same content being expressed in different ways in the different languages) nor be altered conventionally according to some 'Principle of Tolerance'. Such alterations will merely change the linguistic expressions into meaningless shapes or sounds. The whole system of such shapes or sounds, the construction of which is regulated by conventionally fixed rules, is not a *language* at all, but at best a *game* or a *calculus* with uninterpreted (and, so far, meaningless) shapes or sounds, misleadingly called 'symbols' or 'signs'. Not even so-called 'implicit definitions' can change such meaningless shapes into meaningful expressions. What is needed for obtaining an interpretation of such a formal system is *ostensive* definitions of its elements. (Even the truth-tables in the sentential calculus cannot give the so-called 'logical constants' (or 'sentential connectives') any meaning, because the truth-tables themselves have no meaning, if they are not conceived as expressions of mental facts as e.g. rejection, incompatibility, doubt, assertion, etc.).

Before trying to present a short survey of my view of the relations between languages, calculuses, and logic, I would insert yet a preliminary remark that concerns the problem of intercommunication or mutual understanding of a language. As suggested above, any linguistic expression must be an expression of some content of the consciousness of the person (or animal) using the language. But how can a listener know the meaning that the speaker intends to express? In my opinion it would be one of the most difficult tasks philosophers or psychologists have ever undertaken to give a full and satisfactory answer to this question. But I suppose that a first rough approximation to an answer can be suggested in the following way:

Even if we cannot creep into other peoples bodies and observe by their senses or think with their brains, we nevertheless actually use various criteria of what is common to the contents of consciousness of different people (including ourselves), and by these criteria we become convinced from our early childhood that we are living in a common world and that human beings to a rather large extent are feeling and thinking similarly when they are in similar situations. Primarily members of a certain language-community learn to use similar sounds as intersubjective signs for things or phenomena in their external phenomenal world, as e.g. 'mama', 'daddy', 'food', etc. Such sounds function very well not merely as expressions of our needs and wishes but also as expressions of contents of perception of the persons and things named by those sounds. And little by little by means of mutual imitation and adaptation of such usage the words (i.e. the sounds *with* their meanings) become more and more differentiated and precise. During this process it never occurs to us that we should be able to find the common meaning of the sounds either by observation of the sounds themselves or by observation of our language-partners alone. The partners simply show us, by pointing with their fingers, direction of their glances, or by other gestures, that the sound they produce and which we experience as an element in *our* perceptual field is meant to signify some object in *our* perceptual content (to which the partner also belongs). In this way the sound produced by the speaker acquires a meaning for the listener too, and as long as the process of communication goes smoothly no suspicion of a possible difference between the speaker's and the listener's meaning connected with a given sound arises. Indeed, the smoothness of the process of communication and collaboration between speaker and listener seems to be our main, perhaps sole, criterion of the identity of the meanings attached to the sounds used. A suspicion of not-identity of meanings of similar sounds seems first to arise when the listener's reactions on the speaker's utterances differ from the speaker's intentions, and the difficult problems that have resulted from such suspicion are, I think, mainly due to philosophical or epistemological reflexions that may very well depend on untenable presuppositions, as e.g. that my phenomenal external world exists in my brain. This last supposition may not even be true concerning my cenesthetic or other corporal sensations, e.g. my hunger or my tooth-ache that certainly exists in my inside, resp. in my tooth, when they are observed by *my* appropriate interoceptors, even if they cannot be observed at all by other receptors (e.g. eyes or ears), whether my own or others. The difficulty of the problems turning up here are surely not to be underrated, but a preconceived attitude towards them ought not, on the other hand, frustrate the way to a possible fruitful view of the relations between languages, calculuses, and logic. –

After these preliminary remarks I will now try to sketch the main lines of such a view that to my mind seems to be promising, well knowing that I must here restrict myself to a rather concentrated account, leaving to the readers to draw the conclusions of the sketch.

As already remarked, it seems to me evident that a language (natural or artificial) is a means of communication by which a certain person (the speaker) communicates a certain content of his consciousness to another person (the listener). The linguistic means of communication are word-images or statement-images (sentences), primarily sounds or complexes of sounds, secondarily shapes or complexes of shapes. To *use* a language means to use word- or statement-images as means to communicate a content of consciousness, e.g. a content of perception, an idea, a thought, a feeling, an emotion, an intention, a wish, a claim, a decision, etc. Usage is a special form of behaviour in which the speaker produces sounds or shapes as expressions of his content of consciousness in order to communicate (part of) it to the listener, and the listener interprets or understands the expressions of the speaker as expressions of the speakers content of consciousness.

The means of expression are not merely word- or statement-images, but also tone of voice, accentuation, gesture, look, attitude, etc. The content of consciousness that is expressed may be labelled the *speaker's 'meaning'* in the widest sense, and the listener's interpretation of the speaker's expression may be labelled the *listener's 'meaning'* in the widest sense. These two kinds of 'meaning' need not, however, be identical, and perhaps they rarely are so. The listener's interpretation is 'correct' merely when he adequately experiences the means of expression *either* as an expression of the speaker's actual content of consciousness (as it really is) *or* as an expression of the content of consciousness which the speaker (really) intends to express. In the first-mentioned case one may say that the listener has found a *psychologically* correct interpretation, in the last-mentioned case that he has found a *linguistically* correct interpretation.

To find such correct interpretations of the expressions of a speaker is, however, a very difficult matter, and many misunderstandings are possible as well as correct understandings which the speaker had not intended by his expressions.

Mutual understanding thus seems to be a very complicated process which I cannot go deeper into here. What I would stress in this connection is merely the great importance of distinguishing between *word- and statement-images* on the one hand and *'meanings' or concepts* expressed by means of these images on the other. Indeed, in my opinion it is just the connection between the images mentioned and their 'meanings' that make the most fundamental difference between a language and a calculus, it being essential to a language to have a meaning and to a calculus not to need one. While an element of a calculus need not be capable of being interpreted but may function as a mere sound or shape, a word ceases to be a word and degenerates into a mere sound or shape, if it loses its 'meaning' or ceases to be part of an expression of a 'meaning'. When linguists formulate grammatical and syntactical rules for conjugations or for combinations of word-images they always presuppose that these images are images of *words*, i.e. that they have a 'meaning'. But when constructors of calculuses formulate rules for combinations of the symbols of the calculuses they abstract in principle from any possible interpretation of same. The linguistic 'commutation test' is, therefore, so important in linguistics in order to identify or to distinguish between *words* (see e.g. L. Hjelmslev: 'Prolegomena to a Theory of Language', pp. 29–48, especially pp. 46–47), while such a test is quite irrelevant in a calculus. While a calculus remains a calculus whether an interpretation of it exists or not, a language ceases to be a language, if it loses its 'meaning'. And while Carnap's 'Principle of Tolerance' (see 'Logical Syntax of Language', pp. 51–52) may be valid for any calculus, it has no validity at all for languages. The syntactical rules of a calculus, as of any other game, are conventional, whereas the syntactical rules of a language are empirical truths found by observation of the actual linguistic usage. So much concerning the differences between calculuses and languages. And now some remarks concerning the peculiarity of logic.

In my opinion logic has to do neither with relations between words or sentences, nor between calculatory shapes or sounds, but with relations between 'meanings' or concepts. The fundamental *logical* relations are exclusion (either exhaustive, i.e. contradiction, or non-exhaustive, i.e. contrariety) and entailment and its converse. These relations do not exist between word- or statement-images nor between the symbols of a calculus, but solely between 'meanings' or concepts and, secondarily, between statements. The images and 'symbols' are mere sense data, and no sense datum is standing in any *logical* relation to any other sense datum. To be true, some sense data are incompatible with some other sense data while they are compatible with some different ones, and in some cases one can make inferences from some sense data to some others while in other cases no such inferences are possible. But such incompatibilities or compatibilities or inferential connections are *empirical* facts. They are not *logical* relations between sense data, although they may be foundations of logical relations between the *concepts* to the formation of which they may give occasion. A sense datum as such neither contradicts nor entails any other sense datum. In this sense it seems quite right to speak of a 'logical atomism' of sense data. Experience, however, shows that some sense data have always appeared together or that the one has always succeeded the other (or, at least, that such relations between sense data have merely been altered by alteration of the conditions of the sense organs, e.g. by closing of the eyes, or by withdrawing of the fingers, etc.), and these observed constant connections between the sense data seem to be the bases of our concepts of things and of laws, i.e. concepts of constant connections of sense data. If such connections remain constant while the complex that I name 'I' or 'myself' varies, then the connections are independent of the I (the observer) and may be considered *objective*: the independence of the variations of the 'subject' being my criterion of objectivity, – even if the sense data in some sense are 'my' sense data, i.e. are experienced by me. The fact *that* I experience something is, of course, dependent on me, but *what* I experience may be independent of me.

While no *logical* relations exist between my *sense data* such relations do exist between my *concepts* – whatever the nature of concepts may be (they are, in my opinion, a kind of mental attitudes). Some concepts are in fact incompatible while other concepts are compatible, and among some of the compatible concepts there may exist entailment-relations while other compatible concepts are logically independent of each other. Like any other content of consciousness, any concept may be expressed by some sound or shape, and the system of such meaningful expressions are, as already remarked, the (descriptive) languages. Most of those linguistic expressions are not expressions of *logically* related concepts, and so far the language has nothing to do with logic. The *linguistic* syntax of everyday languages is of such a character that it allows the construction of contradictions as well as consistences, of inconsequences as well as logically correct inferences. Therefore, if one formalizes a text of an everyday language, the result will most often be a quite illogical collection of symbols for words, whose 'meanings' are not logically related to each other. It is, however, possible to express logically related concepts (proofs, argumentations, inferences) in everyday language, and where this is done everyday language will contain small selections of linguistic expressions that *are* expressions of *logically* related concepts. By formalization of such selections of everyday language it is possible to reach what may be called a 'logical language' with a 'logical syntax'. But this is quite another thing than an everyday language with a linguistic syntax. Indeed, the logical character of such 'logical language' depends solely on the 'meaning' of the expressions, i.e. on the logical relations between the concepts expressed in the language. As one and the same 'meaning' may be expressed in many ways (in different languages or even within a single language) there are many different linguistic syntaxes, whereas there seems to be merely one logic common to all intelligent people.

Logic is thus, in my opinion, best defined as the study of the logical relations between our *concepts* (and statements). These relations are not in any way conventional, but have a character of evidence and necessity that probably is due to the fact that the formation of our concepts are subjected to a kind of laws of nature, which we must follow willy-nilly. These laws are, evidently, not laws of association of ideas, but they may, perhaps, be analogous to Wertheimer's 'Laws of Organization in Perceptual Forms' (see 'A Source Book of Gestalt Psychology', prepared by Willis D. Ellis, p. 71 f.). This, however, being as it may, I would merely stress that logic is not a conventional calculus but really a science of 'the laws of thought' as Boole said. Two incompatible concepts cannot be made compatible by any definition or other convention, and neither can a statement by convention be made to follow from another from which it does not in fact follow. This being so, I also would like to emphasize once more that it seems impossible in principle to define entailment by means of truth tables, and this time for another reason than the one mentioned before. Indeed, a statement may follow from another statement quite independently of the truth or falsity of the statements. The entailment relation between statements seems to depend on the relation between the *predicates* of the statements and thus on their *intension*, it being quite immaterial whether these concepts have any extension at all. Or, otherwise expressed, a statement may be *necessary* although not being *true*, if by 'truth' we mean correspondence with something different from the statement. Thus 'logical sentences' are expressions of complexes of concepts between which there exists some logical relation which relation it is, as a fact, impossible to deny when it is clearly experienced. In this sense logical connections are empirical, although not sensorial, facts, even if sensorial experiences may form the original basis from which the concepts have developed.

When concepts have arisen they may function as a kind of norms or models for our classification and ordering of perceptions, i.e. we classify and order our perceptions in such a way that they conform to the conceptual schemes available at the given moment. To the extent these schemes function as norms, they necessitate a definite conception of the perceptions. If we have formed a definite concept of a triangle, or a horse, or an atom, these concepts determine whether a given perceived object is or is not what we mean by the words 'triangle', or 'horse', or 'atom'. And we know that objects conforming to these concepts *must* have the qualities that are entailed by (or follow from) these concepts. If they do not have these qualities, the objects would not be triangles, or horses, or atoms. But this necessity concerns merely our present *concepts*, and it is experience only that can assure us of the (perceptual) existence of objects corresponding to the concepts. Thus, the consistency of our concepts can never guarantee that corresponding objects exist. And if we wish that our concepts should correspond to our perceptions, we must form and alter our concepts in such a way that the desired correspondence is obtained. Our conceptual schemes are always hypothetical as to the objects of our perceptions, and the whole development of our conceptual knowledge (our concepts and theories) consists in a running alteration of our hypothetical concepts and theories according to the principle: respect the facts.

To treat in detail the way in which this task is performed (by means of abstractions, idealizations, hypothetical generalizations, etc.) would take us too far, and I must therefore conclude this sketchy article by giving the following summary:

A language is *not* a collection of mere sounds or shapes, or 'strings of marks', according to rules or not, but such a system *connected with* 'meanings', i.e. conceived as a collection of expressions of contents of consciousness. As such contents are continously developing as new external and internal experiences crop up, it is impossible to prescribe eternal or necessary rules for the usage of linguistic symbols (expressions). The regularities actually found by analysis of living (or dead) linguistic systems are relative to a given stage in this development, even if some of those regularities are changing more slowly than others. To define a certain language by means of (in terms of) the regularities existing at a given moment is merely to fix our *concept* of the language concerned but will not stop the development of the language itself, which development will sooner or later make our (synchronic or structural) definition obsolete.

A calculus, on the other hand, is a conventional construction of meaningless shapes or marks that must be mutually combined according to conventional rules. If these rules are arbitrarily changed, the calculus changes into another calculus. Thus it is quite certain that the rules of a given calculus will remain constant or invariable forever. The development of a calculus can, consequently, merely consist in the addition or making of new combinations in conformity with the rules once fixed, but never in an alteration of these rules.

If some interpretation of a calculus is found (or presupposed by the foundation of the calculus), then the calculus becomes a (meaningful) language. But such a language is always of a restricted character and application, its meaningfulness never stretching further than the interpretation found.

Finally, logic is the study of the so-called logical relations between the parts (or elements) of the content of consciousness which are called 'concepts' and 'statements'. Such relations can never be altered conventionally and therefore have a character of necessity. If they are 'formalized', a 'logical language' results, and if the regularities of this language are fixed as *norms* for the construction of the expressions of this language, it may be called a 'logical calculus'. As regards possibilities of application, the status of such a 'logical calculus' is quite the same as the status of any other calculus, i.e. it is meaningful merely to the extent to which an interpretation can be found by internal or external experience. Without such interpretation the logical calculus, as all other calculuses, is merely a conventional game that may be entertaining, and even fascinating, but not useful in any other way – until an interpretation is may be found.

University of Copenhagen (København)

W. V. QUINE

CARNAP AND LOGICAL TRUTH[1)]

I

Kant's question 'How are synthetic judgments *a priori* possible?' precipitated the *Critique of Pure Reason*. Question and answer notwithstanding, Mill and others persisted in doubting that such judgments were possible at all. At length some of Kant's own clearest purported instances, drawn from arithmetic, were sweepingly disqualified (or so it seemed; but see § II) by Frege's reduction of arithmetic to logic. Attention was thus forced upon the less tendentious and indeed logically prior question, 'How is logical certainty possible?' It was largely this latter question that precipitated the form of empiricism which we associate with between-war Vienna – a movement which began with Wittgenstein's *Tractatus* and reached its maturity in the work of Carnap.

Mill's position on the second question had been that logic and mathematics were based on empirical generalizations, despite their superficial appearance to the contrary. This doctrine may well have been felt to do less than justice to the palpable surface differences between the deductive sciences of logic and mathematics, on the one hand, and the empirical sciences ordinarily so-called on the other. Worse, the doctrine derogated from the certainty of logic and mathematics; but Mill may not have been one to be excessively disturbed by such a consequence. Perhaps classical mathematics did lie closer to experience then than now; at any rate the infinitistic reaches of set theory, which are so fraught with speculation and so remote from any possible experience, were unexplored in his day. And it is against just these latter-day mathematical extravagances that

1) This paper was written early in 1954 at the request of Professor Schilpp, for inclusion in a volume on Carnap which he had been planning. The paper has since appeared in Italian translation as 'Carnap e la verità logica', Rivista di Filosofia, vol. 48 (1957), pp. 3–29. Selected portions, running to somewhat less than half, have appeared also in *American Philosophers at Work* (Sidney Hook, ed.), Criterion Books, New York, 1956.

empiricists outside the Vienna Circle have since been known to inveigh,[1] in much the spirit in which the empiricists of Vienna and elsewhere have inveighed against metaphysics.

What now of the empiricist who would grant certainty to logic, and to the whole of mathematics, and yet would make a clean sweep of other non-empirical theories under the name of metaphysics? The Viennese solution of this nice problem was predicated on language. Metaphysics was meaningless through misuse of language; logic was certain through tautologous use of language.

As an answer to the question 'How is logical certainty possible?' this linguistic doctrine of logical truth has its attractions. For there can be no doubt that sheer verbal usage is in general a major determinant of truth. Even so factual a sentence as 'Brutus killed Caesar' owes its truth not only to the killing but equally to our using the component words as we do. Why then should a logically true sentence on the same topic, e.g. 'Brutus killed Caesar or did not kill Caesar', not be said to owe its truth *purely* to the fact that we use our words (in this case 'or' and 'not') as we do? – for it depends not at all for its truth upon the killing.

The suggestion is not, of course, that the logically true sentence is a contingent truth about verbal usage; but rather that it is a sentence which, given the language, automatically becomes true, whereas 'Brutus killed Caeser', given the language, becomes true only contingently on the alleged killing.

Further plausibility accrues to the linguistic doctrine of logical truth when we reflect on the question of alternative logics. Suppose someone puts forward and uses a consistent logic, the principles of which are contrary to our own. We are then clearly free to say that he is merely using the familiar particles 'and', 'all', or whatever, in other than the familiar senses, and hence that no real contrariety is present after all. There may of course still be an important failure of intertranslatability, in that the behavior of certain of our logical particles is incapable of being duplicated by paraphrases in his system or vice versa. If the translation in this sense is possible, from his system into ours, then we are pretty sure to protest that he was wantonly using the familiar particles 'and' and

[1]) An example is P. W. Bridgman, 'A physicist's second reaction to Mengenlehre,' *Scripta Mathematica*, vol. 2, 1933–4, pp. 101–117, 224–234.

'all' (say) where he might unmisleadingly have used such and such other familiar phrasing. This reflection goes to support the view that the truths of logic have no content over and above the meanings they confer on the logical vocabulary.

Much the same point can be brought out by a caricature of a doctrine of Levy-Bruhl, according to which there are pre-logical peoples who accept certain simple self-contradictions as true. Over-simplifying, no doubt, let us suppose it claimed that these natives accept as true a certain sentence of the form 'p and not p'. Or – not to over-simplify too much – that they accept as true a certain heathen sentence of the form 'q ka bu q' the English translation of which has the form 'p and not p'. But now just how good a translation is this, and what may the lexicographer's method have been? If any evidence can count against a lexicographer's adoption of 'and' and 'not' as translations of 'ka' and 'bu', certainly the natives' acceptance of 'q ka bu q' as true counts overwhelmingly. We are left with the meaninglessness of the doctrine of there being pre-logical peoples; prelogicality is a trait injected by bad translators. This is one more illustration of the inseparability of the truths of logic from the meanings of the logical vocabulary.

We thus see that there is something to be said for the naturalness of the linguistic doctrine of logical truth. But before we can get much further we shall have to become more explicit concerning our subject matter.

II

Without thought of any epistemological doctrine, either the linguistic doctrine or another, we may mark out the intended scope of the term 'logical truth', within that of the broader term 'truth', in the following way. First we suppose indicated, by enumeration if not otherwise, what words are to be called logical words; typical ones are 'or', 'not', 'if', 'then', 'and', 'all', 'every', 'only', 'some'. The logical truths, then, are those true sentences which involve only logical words *essentially*. What this means is that any other words, though they may also occur in a logical truth (as witness 'Brutus', 'kill', and 'Caesar' in 'Brutus killed or did not kill Caesar'), can be varied at will without engendering falsity.[1]

[1] Substantially this formulation is traced back a century and a quarter by Yehoshua Bar-Hillel, 'Bolzano's definition of analytic propositions,' *Methodos*, vol. 2, 1950,

Though formulated with reference to language, the above clarification does not of itself hint that logical truths owe their truth to language. What we have thus far is only a delimitation of the class, *per accidens* if you please. Afterwards the linguistic doctrine of logical truth, which is an epistemological doctrine, goes on to say that logical truths are true by virtue purely of the intended meanings, or intended usage, of the logical words. Obviously if logical truths *are* true by virtue purely of language, the logical words are the only part of the language that can be concerned in the matter; for these are the only ones that occur essentially.

Elementary logic, as commonly systematized nowadays, comprises truth-function theory, quantification theory, and identity theory. The logical vocabulary for this part, as commonly rendered for technical purposes, consists of truth-function signs (corresponding to 'or', 'and', 'not', etc.), quantifiers and their variables, and '='.

The further part of logic is set theory, which requires there to be classes among the values of its variables of quantification. The one sign needed in set theory, beyond those appropriate to elementary logic, is the connective 'ε' of membership. Additional signs, though commonly used for convenience, can be eliminated in well-known ways.

In this dichotomy I leave metatheory, or logical syntax, out of account. For, either it treats of special objects of an extralogical kind, viz. notational expressions, or else, if these are made to give way to numbers by arithmetization, it is reducible via number theory to set theory.

I will not here review the important contrasts between elementary logic and set theory, except for the following one. Every truth of elementary logic is obvious (whatever this really means), or can be made so by some series of individually obvious steps. Set theory, in its present state anyway, is otherwise. I am not alluding here to Gödel's incompleteness principle, but to something right on the surface. Set theory was straining

pp. 32–55 (= *Theoria*, vol. 16, 1950, pp. 91–117). But note that the formulation fails of its purpose unless the phrase 'can be varied at will,' above, is understood to provide for varying the words not only singly but also two or more at a time. E.g., the sentence 'If some men are angels some animals are angels' can be turned into a falsehood by simultaneous substitution for 'men' and 'angels', but not by any substitution for 'angels' alone, nor for 'men', nor for 'animals' (granted the non-existence of angels). For this observation and illustration I am indebted to John R. Myhill, who expresses some indebtedness in turn to Benson Mates. – I added most of this footnote in May, 1955; thus one year after the rest of the essay left my hands.

at the leash of intuition ever since Cantor discovered the higher infinites; and with the added impetus of the paradoxes of set theory the leash was snapped. Comparative set theory has now long been the trend; for, so far as is known, no consistent set theory is both adequate to the purposes envisaged for set theory and capable of substantiation by steps of obvious reasoning from obviously true principles. What we do is to develop one or another set theory by obvious reasoning, or elementary logic, from unobvious first principles which are set down, whether for good or for the time being, by something very like convention.

Altogether, the contrasts between elementary logic and set theory are so fundamental that one might well limit the word 'logic' to the former (though I shall not), and speak of set theory as mathematics in a sense exclusive of logic. To adopt this course is merely to deprive 'ε' of the status of a logical word. Frege's derivation of arithmetic would then cease to count as a derivation from logic; for he used set theory. At any rate we should be prepared to find that the linguistic doctrine of logical truth holds for elementary logic and fails for set theory, or vice versa. Kant's readiness to see logic as analytic and arithmetic as synthetic, in particular, is not superseded by Frege's work (as Frege supposed [1])) if 'logic' be taken as elementary logic. And for Kant logic certainly did not include set theory.

III

Where someone disagrees with us as to the truth of a sentence, it often happens that we can convince him by getting the sentence from other sentences, which he does accept, by a series of steps each of which he accepts. Disagreement which cannot be thus resolved I shall call *deductively irresoluble*. Now if we try to warp the linguistic doctrine of logical truth around into something like an experimental thesis, perhaps a first approximation will run thus: *Deductively irresoluble disagreement as to a logical truth is evidence of deviation in usage (or meanings) of words.* This is not yet experimentally phrased, since one term of the affirmed relationship, viz. 'usage' (or 'meanings'), is in dire need of an independent

1) See §§ 87f., 109 of Gottlob Frege, *Foundations of Arithmetic* (New York: Philosophical Library, and Oxford: Blackwell, 1950), a reprint of *Grundlagen der Arithmetik* (Breslau, 1884) with translation by J. L. Austin.

criterion. However, the formulation would seem to be fair enough within its limits; so let us go ahead with it, not seeking more subtlety until need arises.

Already the obviousness or potential obviousness of elementary logic can be seen to present an insuperable obstacle to our assigning any experimental meaning to the linguistic doctrine of elementary logical truth. Deductively irresoluble dissent from an elementary logical truth *would* count as evidence of deviation over meanings if anything can, but simply because dissent from a logical truism is as extreme as dissent can get.

The philosopher, like the beginner in algebra, works in danger of finding that his solution-in-progress reduces to '$0 = 0$'. Such is the threat to the linguistic theory of elementary logical truth. For that theory now seems to imply nothing that is not already implied by the fact that elementary logic is obvious or can be resolved into obvious steps.

The considerations which were adduced in § I, to show the naturalness of the linguistic doctrine, are likewise seen to be empty when scrutinized in the present spirit. One was the circumstance that alternative logics are inseparable practically from mere change in usage of logical words. Another was that illogical cultures are indistinguishable from ill-translated ones. But both of these circumstances are adequately accounted for by mere obviousness of logical principles, without help of a linguistic doctrine of logical truth. For there can be no stronger evidence of a change in usage than the repudiation of what had been obvious, and no stronger evidence of bad translation than that it translates earnest affirmations into obvious falsehoods.

Another point in § I was that true sentences generally depend for their truth on the traits of their language in addition to the traits of their subject matter; and that logical truths then fit neatly in as the limiting case where the dependence on traits of the subject matter is nil. Consider, however, the logical truth 'Everything is self-identical', or '$(x)(x = x)$'. We *can* say that it depends for its truth on traits of the language (specifically on the usage of '$=$'), and not on traits of its subject matter; but we can also say, alternatively, that it depends on an obvious trait, viz. self-identity, of its subject matter, viz. everything. The tendency of our present reflections is that there is no difference.

I have been using the vaguely psychological word 'obvious' non-technically, assigning it no explanatory value. My suggestion is merely that the linguistic doctrine of elementary logical truth likewise leaves explanation unbegun. I do not suggest that the linguistic doctrine is false and some doctrine of ultimate and inexplicable insight into the obvious traits of reality is true, but only that there is no real difference between these two pseudo-doctrines.

Turning away now from elementary logic, let us see how the linguistic doctrine of logical truth fares in application to set theory. As noted in § II, we may think of 'ε' as the one sign for set theory in addition to those of elementary logic. Accordingly the version of the linguistic doctrine which was italicized at the beginning of the present section becomes, in application to set theory, this: among persons already in agreement on elementary logic, deductively irresoluble disagreement as to a truth of set theory is evidence of deviation in usage (or meaning) of 'ε'.

This thesis is not trivial in quite the way in which the parallel thesis for elementary logic was seen to be. It is not indeed experimentally significant as it stands, simply because of the lack, noted earlier, of a separate criterion for usage or meaning. But it does seem reasonable, by the following reasoning.

Any acceptable evidence of usage or meaning of words must reside surely either in the observable circumstances under which the words are uttered (in the case of concrete terms referring to observable individuals) or in the affirmation and denial of sentences in which the words occur. Only the second alternative is relevant to 'ε'. Therefore any evidence of deviation in usage or meaning of 'ε' must reside in disagreement on sentences containing 'ε'. This is not, of course, to say of *every* sentence containing 'ε' that disagreement over it establishes deviation in usage or meaning of 'ε'. We have to assume in the first place that the speaker under investigation agrees with us on the meanings of words other than 'ε' in the sentences in question. And it might well be that, even from among the sentences containing only 'ε' and words on whose meanings he agrees with us, there is only a select species S which is so fundamental that he cannot dissent from them without betraying deviation in his usage or meaning of 'ε'. But S may be expected surely to include some (if not all) of the sentences which contain *nothing* but 'ε' and the elementary logical particles; for it is these sentences, insofar as true, that constitute (pure, or unapplied) set theory. But it is difficult to conceive of how to be other than democratic toward the truths of set theory. In exposition we may select some of these truths as so-called postulates and deduce others from them, but this is subjective discrimination, variable at will, expository and not set-theoretic. We do not change our meaning of 'ε' between the page where we show that one particular truth is deducible by elementary logic from another and the page where we show the converse. Given this democratic outlook, finally, the law of sufficient reason leads us to look upon S as including *all* the sentences which contain only 'ε' and the elementary logical particles. It then follows that anyone in agreement on elementary logic and in irresoluble disagreement on set theory is in deviation with respect to the usage or meaning of 'ε'; and this was the thesis.

The effect of our effort to inject content into the linguistic doctrine of logical truth has been, up to now, to suggest that the doctrine says nothing worth saying about elementary logical truth, but that when applied to set-theoretic truth it makes for a reasonable partial condensation of the otherwise vaporous notion of meaning as applied to 'ε'.

IV

The linguistic doctrine of logical truth is sometimes expressed by saying that such truths are true by linguistic convention. Now if this be so, certainly the conventions are not in general explicit. Relatively few persons, before the time of Carnap, had ever seen any convention that engendered truths of elementary logic. Nor can this circumstance be ascribed merely to the slipshod ways of our predecessors. For it is impossible in principle, even in an ideal state, to get even the most elementary part of logic exclusively by the explicit application of conventions stated in advance. The difficulty is the vicious regress, familiar from Lewis Carroll,[1] which I have elaborated elsewhere.[2] Briefly the point is that the logical truths, being infinite in number, must be given by general conventions rather than singly; and logic is needed then to begin with, in the metatheory, in order to apply the general conventions to individual cases.

[1]) What the tortoise said to Achilles,' *Mind*, vol. 4, 1895, pp. 278ff.
[2]) 'Truth by convention,' in O. H. Lee (ed.), *Philosophical Essays for A. N. Whitehead* (New York, 1936), pp. 90–124. Reprinted in H. Feigl and W. Sellars (eds.), *Readings in Philosophical Analysis* (New York: Appleton, 1949).

'In dropping the attributes of deliberateness and explicitness from the notion of linguistic convention,' I went on to complain in the aforementioned paper, 'we risk depriving the latter of any explanatory force and reducing it to an idle label.' It would seem that to call elementary logic true by convention is to add nothing but a metaphor to the linguistic doctrine of logical truth which, as applied to elementary logic, has itself come to seem rather an empty figure (cf. § III).

The case of set theory, however, is different on both counts. For set theory the linguistic doctrine has seemed less empty (cf. § III); in set theory, moreover, convention in quite the ordinary sense seems to be pretty much what goes on (cf. § II). Conventionalism has a serious claim to attention in the philosophy of mathematics, if only because of set theory. Historically, though, conventionalism was encouraged in the philosophy of mathematics rather by the non-Euclidean geometries and abstract algebras, with little good reason. We can contribute to subsequent purposes by surveying this situation. Further talk of set theory is deferred to § V.

In the beginning there was Euclidean geometry, a compendium of truths about form and void; and its truths were not based on convention (except as a conventionalist might, begging the present question, apply this tag to everything mathematical). Its truths were in practice presented by deduction from so-called postulates (including axioms: I shall not distinguish); and the selection of truths for this role of postulate, out of the totality of truths of Euclidean geometry, was indeed a matter of convention. But this is not *truth* by convention. The truths were there, and what was conventional was merely the separation of them into those to be taken as starting point (for purposes of the exposition at hand) and those to be deduced from them.

The non-Euclidean geometries came of artificial deviations from Euclid's postulates, without thought (to begin with) of true interpretation. These departures were doubly conventional; for Euclid's postulates were a conventional selection from among the truths of geometry, and then the departures were arbitrarily or conventionally devised in turn. But still there was no truth by convention, because there was no truth.

Playing within a non-Euclidean geometry, one might conveniently make believe that his theorems were interpreted and true; but even such conventional make-believe is not truth by convention. For it is not really truth at all; and what is conventionally pretended is that the theorems are true by non-convention.

Non-Euclidean geometries have, in the fullness of time, received serious interpretations. This means that ways have been found of so construing the hitherto unconstrued terms as to identify the at first conventionally chosen set of non-sentences with some genuine truths, and truths presumably not by convention. The status of an interpreted non-Euclidean geometry differs in no basic way from the original status of Euclidean geometry, noted above.

Uninterpreted systems became quite the fashion after the advent of non-Euclidean geometries. This fashion helped to cause, and was in turn encouraged by, an increasingly formal approach to mathematics. Methods had to become more formal to make up for the unavailability, in uninterpreted systems, of intuition. Conversely, disinterpretation served as a crude but useful device (until Frege's syntactical approach came to be appreciated) for achieving formal rigor uncorrupted by intuition.

The tendency to look upon non-Euclidean geometries as true by convention applied to uninterpreted systems generally, and then carried over from these to mathematical systems generally. A tendency indeed developed to look upon all mathematical systems as, *qua* mathematical, uninterpreted. This tendency can be accounted for by the increase of formality, together with the use of disinterpretation as a heuristic aid to formalization. Finally, in an effort to make some sense of mathematics thus drained of all interpretation, recourse was had to the shocking quibble of identifying mathematics merely with the elementary logic which leads from uninterpreted postulates to uninterpreted theorems.[1] What is shocking about this is that it puts arithmetic *qua* interpreted theory of number, and analysis *qua* interpreted theory of functions, and geometry *qua* interpreted theory of space, outside mathematics altogether.

The substantive reduction of mathematics to logic by Frege, Whitehead, and Russell is of course quite another thing. It is a reduction not to elementary logic but to set theory; and it is a reduction of genuine interpreted mathematics, from arithmetic onward.

[1]) Bertrand Russell, *Principles of Mathematics* (Cambridge, 1903), pp. 429f.; Heinrich Behmann, 'Sind die mathematischen Urteile analytisch oder synthetisch?' *Erkenntnis*, vol. 4, 1934, pp. 8ff.; and others.

V

Let us then put aside these confusions and get back to set theory. Set theory is pursued as interpreted mathematics, like arithmetic and analysis; indeed, it is to set theory that those further branches are reducible. In set theory we discourse about certain immaterial entities, real or erroneously alleged, viz. sets, or classes. And it is in the effort to make up our minds about genuine truth and falsity of sentences about these objects that we find ourselves engaged in something very like convention in an ordinary non-metaphorical sense of the word. We find ourselves making deliberate choices and setting them forth unaccompanied by any attempt at justification other than in terms of elegance and convenience. These adoptions, called postulates, and their logical consequences (via elementary logic), are true until further notice.

So here is a case where postulation can plausibly be looked on as constituting truth by convention. But in § IV we have seen how the philosophy of mathematics can be corrupted by supposing that postulates always play that role. Insofar as we would epistemologize and not just mathematize, we might divide postulation as follows. Uninterpreted postulates may be put aside, as no longer concerning us; and on the interpreted side we may distinguish between *legislative* and *discursive* postulation. Legislative postulation institutes truth by convention, and seems plausibly illustrated in contemporary set theory. On the other hand discursive postulation is mere selection, from a preëxisting body of truths, of certain ones for use as a basis from which to derive others, initially known or unknown. What discursive postulation fixes is not truth, but only some particular ordering of the truths, for purposes perhaps of pedagogy or perhaps of inquiry into logical relationships ('logical' in the sense of elementary logic). All postulation is of course conventional, but only legislative postulation properly hints of *truth* by convention.

It is well to recognize, if only for its distinctness, yet a further way in which convention can enter; viz. in the adoption of new notations for old ones, without, as one tends to say, change of theory. Truths containing the new notation are conventional transcriptions of sentences true apart from the convention in question. They depend for their truth partly on language, but then so did 'Brutus killed Caesar' (cf. § I). They come into being through a conventional adoption of a new sign, and they become true through conventional definition of that sign *together with* whatever made the corresponding sentences in the old notation true.

Definition, in a properly narrow sense of the word, is convention in a properly narrow sense of the word. But the phrase 'true by definition' must be taken cautiously; in its strictest usage it refers to a transcription, by the definition, of a truth of elementary logic. Whether such a sentence is true by convention depends on whether the logical truths themselves be reckoned as true by convention. Even an outright equation or biconditional connection the definiens and the definiendum is a definitional transcription of a prior logical truth of the form '$x = x$' or '$p \equiv p$'.

Definition commonly so-called is not thus narrowly conceived, and must for present purposes be divided, as postulation was divided, into legislative and discursive. Legislative definition introduces a notation hitherto unused, or used only at variance with the practice proposed, or used also at variance, so that a convention is wanted to settle the ambiguity. Discursive definition, on the other hand, sets forth a preëxisting relation of interchangeability or coextensiveness between notations in already familiar usage. A frequent purpose of this activity is to show how some chosen part of language can be made to serve the purposes of a wider part. Another frequent purpose is language instruction.

It is only legislative definition, and not discursive definition nor discursive postulation, that makes a conventional contribution to the truth of sentences. Legislative postulation, finally, affords truth by convention unalloyed.

Increasingly the word 'definition' connotes the formulas of definition which appear in connection with formal systems, signalled by some extra-systematic sign such as '$=_{df}$'. Such definitions are best looked upon as correlating two systems, two notations, one of which is prized for its economical lexicon and the other for its brevity or familiarity of expression.[1]) Definitions so used can be either legislative or discursive in their inception. But this distinction is in practice left unindicated, and wisely; for it is a distinction only between particular acts of definition, and not germane to the definition as an enduring channel of intertranslation.

The distinction between the legislative and the discursive refers thus to

[1]) See my *From a Logical Point of View* (Cambridge, Mass.: Harvard, 1953), pp. 26f.

the act, and not to its enduring consequence, in the case of postulation as in the case of definition. This is because we are taking the notion of truth by convention fairly literally and simple-mindedly, for lack of an intelligible alternative. So conceived, conventionality is a passing trait, significant at the moving front of science but useless in classifying the sentences behind the lines. It is a trait of events and not of sentences.

Might we not still project a derivative trait upon the sentences themselves, thus speaking of a sentence as forever true by convention if its first adoption as true was a convention? No; this, if done seriously, involves us in the most unrewarding historical conjecture. Legislative postulation contributes truths which become integral to the corpus of truths; the artificiality of their origin does not linger as a localized quality, but suffuses the corpus. If a subsequent expositor singles out those once legislatively postulated truths again as postulates, this signifies nothing; he is engaged only in discursive postulation. He could as well choose his postulates from elsewhere in the corpus, and will if he thinks this serves his expository ends.

VI

Set theory, currently so caught up in legislative postulation, may some day gain a norm – even a strain of obviousness, perhaps – and lose all trace of the conventions in its history. A day could likewise have been when our elementary logic was itself instituted as a deliberately conventional deviation from something earlier, instead of evolving, as it did, mainly by unplanned shifts of form and emphasis coupled with casual novelties of notation.

Today indeed there are dissident logicians even at the elementary level, propounding deviations from the law of the excluded middle. These deviations, insofar as meant for serious use and not just as uninterpreted systems, are as clear cases of legislative postulation as the ones in set theory. For here we have again, quite as in set theory, the propounding of a deliberate choice unaccompanied (conceivably) by any attempt at justification other than in terms of convenience.

This example from elementary logic controverts no conclusion we have reached. According to §§ II and III, the departure from the law of the excluded middle would count as evidence of revised usage of 'or' and 'not'. (This judgment was upheld in § III, though disqualified as evidence for the linguistic doctrine of logical truth.) For the deviating logician the words 'or' and 'not' are unfamiliar, or defamiliarized; and his decisions regarding truth values for their proposed contexts can then be just as genuinely a matter of deliberate convention as the decisions of the creative set theorist regarding contexts of 'ε'.

The two cases are indeed much alike. Not only is departure from the classical logic of 'or' and 'not' evidence of revised usage of 'or' and 'not'; likewise, as argued at length in § III, divergences between set-theorists may reasonably be reckoned to revised usage of 'ε'. Any such revised usage is conspicuously a matter of convention, and can be declared by legislative postulation.

We have been at a loss to give substance to the linguistic doctrine, particularly of elementary logical truth, or to the doctrine that the familiar truths of logic are true by convention. We have found some sense in the notion of truth by convention, but only as attaching to a process of adoption, viz. legislative postulation, and not as a significant lingering trait of the legislatively postulated sentence. Surveying current events, we note legislative postulation in set theory and, at a more elementary level, in connection with the law of the excluded middle.

And do we not find the same continually in the theoretical hypotheses of natural science itself? What seemed to smack of convention in set theory (§ V), at any rate, was 'deliberate choice, set forth unaccompanied by any attempt at justification other than in terms of elegance and convenience'; and to what theoretical hypothesis of natural science might not this same character be attributed? For surely the justification of any theoretical hypothesis can, at the time of hypothesis, consist in no more than the elegance or convenience which the hypothesis brings to the containing body of laws and data. How then are we to delimit the category of legislative postulation, short of including under it every new act of scientific hypothesis?

The situation may seem to be saved, for ordinary hypotheses in natural science, by there being some indirect but eventual confrontation with empirical data. However, this confrontation can be remote; and, conversely, some such remote confrontation with experience may be claimed even for pure mathematics and elementary logic. The semblance of a difference in this respect is largely due to over-emphasis of departmental boundaries. For a self-contained theory which we can check with experience includes, in point of fact, not only its various theoretical hypotheses of so-called natural science but also such portions of logic and mathematics as it makes use of. Hence I do not see how a line is to be drawn between hypotheses which confer truth by convention and hypotheses which do not, short of reckoning *all* hypotheses to the former category save perhaps those actually derivable or refutable by elementary logic from what Carnap used to call protocol sentences. But this version, besides depending to an unwelcome degree on the debatable notion of protocol sentences, is far too inclusive to suit anyone.

Evidently our troubles are waxing. We had been trying to make sense of the role of convention in *a priori* knowledge. Now the very distinction between *a priori* and empirical begins to waver and dissolve, at least as a distinction between sentences. (It could of course still hold as a distinction between factors in one's adoption of a sentence, but both factors might be operative everywhere.)

VII

Whatever our difficulties over the relevant distinctions, it must be conceded that logic and mathematics do seem qualitatively different from the rest of science. Logic and mathematics hold conspicuously aloof from any express appeal, certainly, to observation and experiment. Having thus nothing external to look to, logicians and mathematicians look closely to notation and explicit notational operations: to expressions, terms, substitution, transposition, cancellation, clearing of fractions, and the like. This concern of logicians and mathematicians with syntax (as Carnap calls it) is perennial, but in modern times it has become increasingly searching and explicit, and has even prompted, as we see, a linguistic philosophy of logical and mathematical truth.

On the other hand, an effect of these same formal developments in modern logic, curiously, has been to show how to divorce mathematics (other than elementary logic) from any peculiarly notational considerations not equally relevant to natural science. By this I mean that mathematics can be handled (insofar as it can be handled at all) by axiomatization, outwardly quite like any system of hypotheses elsewhere in science; and elementary logic can then be left to extract the theorems.

The consequent affinity between mathematics and systematized natural science was recognized by Carnap when he propounded his P-rules alongside his L-rules or meaning postulates. Yet he did not look upon the P-rules as engendering analytic sentences, sentences true purely by language. How to sustain this distinction has been very much our problem in these pages, and one on which we have found little encouragement.

Carnap appreciated this problem, in *Logical Syntax*, as a problem of finding a difference in kind between the P-rules (or the truths thereby specified) and the L-rules (or the L-truths, analytic sentences, thereby specified). Moreover he proposed an ingenious solution.[1] In effect he characterized the logical (including mathematical) vocabulary as the largest vocabulary such that (1) there are sentences which contain only that vocabulary and (2) all such sentences are determinable as true or false by a purely syntactical condition – i.e., by a condition which speaks only of concatenation of marks. Then he limited the L-truths in effect to those involving just the logical vocabulary essentially.[2]

Truths given by P-rules were supposedly excluded from the category of logical truth under this criterion, because, though the rules specifying them are formally stated, the vocabulary involved can also be recombined to give sentences whose truth values are not determinate under any set of rules formally formulable in advance.

At this point one can object (pending a further expedient of Carnap's, which I shall next explain) that the criterion based on (1) and (2) fails of its purpose. For consider, to begin with the totality of those sentences which are expressed purely within what Carnap (or anyone) would want to count as logical (and mathematical) vocabulary. Suppose, in conformity with (2), that the division of this totality into the true and the false is reproducible in purely syntactical terms. Now surely the adding of one general term of an extra-logical kind, say 'heavier than', is not going to alter the situation. The truths which are expressible in terms of just 'heavier than', together with the logical vocabulary, will be truths of only the most general kind, such as '$(\exists x)(\exists y)(x$ is heavier than $y)$', '$(x) \sim (x$ is heavier than $x)$', and '$(x)(y)(z)(x$ is heavier than $y \cdot y$ is heavier than $z \cdot \supset . x$ is heavier than $z)$'. The division of the truths from the falsehoods

[1] Carnap, *Logical Syntax of Language*, § 50.
[2] Cf. § I above. Also, for certain reservations conveniently postponed at the moment, see § IX on 'essential predication.'

in this supplementary domain can probably be reproduced in syntactical terms if the division of the original totality could. But then, under the criterion based on (1) and (2), 'heavier than' qualifies for the logical vocabulary. And it is hard to see what whole collection of general terms of natural science might not qualify likewise.

The further expedient, by which Carnap met this difficulty, was his use of Cartesian coördinates.[1] Under this procedure, each spatio-temporal particular c becomes associated with a class K of quadruples of real numbers, viz., the class of those quadruples which are the coördinates of component point-events of c. Further let us write $K[t]$ for the class of triples which with t appended belong to K; thus $K[t]$ is that class of triples of real numbers which is associated with the momentary state of object c at time t. Then, in order to say e.g. that c_1 is heavier than c_2 at time t, we say '$H(K_1[t], K_2[t])$', which might be translated as 'The momentary object associated with $K_1[t]$ is heavier than that associated with $K_2[t]$.' Now $K_1[t]$ and $K_2[t]$ are, in every particular case, purely mathematical objects; viz., classes of triples of real numbers. So let us consider all the true and false sentences of the form '$H(K_1[t], K_2[t])$' where, in place of '$K_1[t]$' and '$K_2[t]$', we have purely logico-mathematical designations of particular classes of triples of real numbers. There is no reason to suppose that all the truths of *this* domain can be exactly segregated in purely syntactical terms. Thus inclusion of 'H' does violate (2), and therefore 'H' fails to qualify as logical vocabulary. By adhering to the method of coördinates and thus reconstruing all predicates of natural science in the manner here illustrated by 'H', Carnap overcomes the objection noted in the preceding paragraph.

To sum up very roughly, this theory characterizes logic (and mathematics) as the largest part of science within which the true-false dichotomy *can* be reproduced in syntactical terms. This version may seem rather thinner than the claim that logic and mathematics are somehow true by linguistic convention, but at any rate it is more intelligible, and, if true, perhaps interesting and important. To become sure of its truth, interest, and importance, however, we must look more closely at this term 'syntax'.

As used in the passage: 'The terms 'sentence' and 'direct consequence' are the two primitive terms of logical syntax,'[2]) the term 'syntax' is of

[1]) *Logical Syntax of Language*, §§ 3, 15.

[2]) Carnap, *Philosophy and Logical Syntax*, p. 47.

course irrelevant to a thesis. The relevant sense is that rather in which it connotes discourse about marks and their succession. But here still we must distinguish degrees of inclusiveness; two different degrees are exemplified in *Logical Syntax*, according as the object language is Carnap's highly restricted Language I or his more powerful Language II. For the former, Carnap's formulation of logical truth is narrowly syntactical in the manner of familiar formalizations of logical systems by axioms and rules of inference. But Gödel's proof of the incompletability of elementary number theory shows that no such approach can be adequate to mathematics in general, nor in particular to set theory, nor to Language II. For Language II, in consequence, Carnap's formulation of logical truth proceeded along the lines rather of Tarski's technique of truth-definition.[1] The result was still a purely syntactical specification of the logical truths, but only in this more liberal sense of 'syntactical': it was couched in a vocabulary consisting (in effect) of (a) names of signs, (b) an operator expressing concatenation of expressions, and (c), by way of auxiliary machinery, the whole logical (and mathematical) vocabulary itself.

So construed, however, the thesis that logico-mathematical truth is syntactically specifiable becomes uninteresting. For what it says is that logico-mathematical truth is specifiable in a notation consisting solely of (a), (b), *and* the whole logico-mathematical vocabulary itself. But *this* thesis would hold equally if 'logico-mathematical' were broadened (at *both* places in the thesis) to include physics, economics, and anything else under the sun; Tarski's routine of truth-definition would still carry through just as well. No special trait of logic and mathematics has been singled out after all.

Strictly speaking, the position is weaker still. The mathematics appealed to in (c) must, as Tarski shows, be a yet more inclusive mathematical theory in certain respects than that for which truth is being defined. It was largely because of his increasing concern over this self-stultifying situation that Carnap relaxed his stress on syntax, in the years following *Logical Syntax*, in favor of semantics.

[1] *Logical Syntax*, especially §§ 34a-i, 60a-d. 71a-d. These sections had been omitted from the German edition, but only for lack of space; cf. p. xi of the English edition. Meanwhile they had appeared as articles: 'Die Antinomien . . .' and 'Ein Gültigkeits-kriterium . . .' At that time Carnap had had only partial access to Tarski's ideas

VIII

Even if logical truth were specifiable in syntactical terms, this would not show that it was grounded in language. Any *finite* class of truths (to take an extreme example) is clearly reproducible by a membership condition couched in as narrowly syntactical terms as you please; yet we certainly cannot say of every finite class of truths that its members are true purely by language. Thus the ill-starred doctrine of syntactical specifiability of logical truth was always something other than the linguistic doctrine of logical truth, if this be conceived as the doctrine that logical truth is grounded in language. In any event the doctrine of syntactical specifiability, which we found pleasure in being able to make comparatively clear sense of, has unhappily had to go by the board. The linguistic doctrine of logical truth, on the other hand, goes sturdily on.

The notion of logical truth is now counted by Carnap as semantical. This of course does not of itself mean that logical truth is grounded in language; for note that the general notion of truth is also semantical, though truth in general is not grounded purely in language. But the semantical attribute of logical truth, in particular, *is* one which, according to Carnap, is grounded in language: in convention, fiat, meaning. Such support as he hints for this doctrine, aside from ground covered in §§ I-VI, seems to depend on an analogy with what goes on in the propounding of artificial languages; and I shall now try to show why I think the analogy mistaken.

I may best schematize the point by considering a case, not directly concerned with logical truth, where one might typically produce an artificial language as a step in an argument. This is the imaginary case of a logical positivist, say Ixmann, who is out to defend scientists against the demands of a metaphysician. The metaphysician argues that science presupposes metaphysical principles, or raises metaphysical problems, and that the scientists should therefore show due concern. Ixmann's answer consists in showing in detail how people (on Mars, say) might speak a language quite adequate to all of our science but, unlike our language, incapable of expressing the alleged metaphysical issues. (I applaud

(cf. 'Gültigkeitskriterium,' footnote 3), the full details of which reached the non-Slavic world in 1936: Alfred Tarski, 'Der Wahrheitsbegriff in den formalisierten Sprachen,' *Studia Philosophica*, vol. 1, pp. 261–405.

this answer, and think it embodies the most telling component of Carnap's own anti-metaphysical representations; but here I digress.) Now how does our hypothetical Ixmann specify that doubly hypothetical language? By telling us, at least to the extent needed for his argument, what these Martians are to be imagined as uttering and what they are thereby to be understood to mean. Here is Carnap's familiar duality of formation rules and transformation rules (or meaning postulates), as rules of language. But these rules are part only of Ixmann's narrative machinery, not part of what he is portraying. He is not representing his hypothetical Martians themselves as somehow explicit on formation and transformation rules. Nor is he representing there to be any intrinsic difference between those truths which happen to be disclosed to us by his partial specifications (his transformation rules) and those further truths, hypothetically likewise known to the Martians of his parable, which he did not trouble to sketch in.

The threat of fallacy lurks in the fact that Ixmann's rules are indeed arbitrary fiats, as is his whole Martian parable. The fallacy consists in confusing levels, projecting the conventional character of the rules into the story, and so misconstruing Ixmann's parable as attributing truth-legislation to his hypothetical Martians.

The case of a non-hypothetical artificial language is in principle the same. Being a new invention, the language has to be explained; and the explanation will proceed by what may certainly be called formation and transformation rules. These rules will hold by arbitrary fiat, the artifex being boss. But all we can reasonably ask of these rules is that they enable us to find corresponding to each of his sentences a sentence of like truth value in familiar ordinary language. There is no (to me) intelligible additional decree that we can demand of him as to the boundary between analytic and synthetic, logic and fact, among his truths. We may well decide to extend our word 'analytic' or 'logically true' to sentences of his language which he in his explanations has paired off fairly directly with English sentences so classified by us; but this is our decree, regarding our word 'analytic' or 'logically true'.

IX

We had in § II to form some rough idea of what logical truth was supposed to take in, before we could get on with the linguistic doctrine of logical truth. This we did, with help of the general notion of truth [1]) together with a partial enumeration of the logical vocabulary of a particular language. In § VII we found hope of a less provincial and accidental characterization of logical vocabulary; but it failed. Still, the position is not intolerable. We well know from modern logic how to devise *a* technical notation which is admirably suited to the business of 'or', 'not', 'and', 'all,' 'only', and such other particles as we would care to count as logical; and to enumerate the signs and constructions of that technical notation, or a theoretically adequate subset of them, is the work of a moment (cf. § II). Insofar as we are content to think of all science as fitted within that stereotyped logical framework – and there is no hardship in so doing – our notion of logical vocabulary is precise. And so, derivatively, is our notion of logical truth. But only in point of extent. There is no epistemological corollary as to the *ground* of logical truth (cf. § II).

Even this half-way tolerable situation obtains only for logical truth in a relatively narrow sense, omitting truths by 'essential predication' (in Mill's phrase) such as 'No bachelor is married'. [2]) I tend to reserve the term 'logically true' for the narrower domain, and to use the term 'analytic' for the more inclusive domain which includes truths by essential predication. Carnap on the contrary has used both terms in the broader sense. But the problems of the two subdivisions of the analytic class differ in such a way that it has been convenient up to now in this essay to treat mainly of logical truth in the narrower sense.

The truths by essential predication are sentences which can be turned into logical truths by supplanting certain simple predicates (e.g. 'bachelor') by complex synonyms (e.g. 'man not married'). This formulation is not inadequate to such further examples as 'If A is part of B and B is part of C then A is part of C'; this case can be managed by using for 'is part of' the synonym 'overlaps nothing save what overlaps'.[3]) The relevant notion

[1]) In defense of this general notion, in invidious contrast to that of analyticity, see my *From a Logical Point of View*, pp. 137f.

[2]) Cf. M. White, 'The analytic and the synthetic: an untenable dualism,' in Sidney Hook (ed.), *John Dewey: Philosopher of Science and Freedom* (New York: Dial, 1950), pp. 316–330. Reprinted in Leonard Linsky (ed.), *Semantics and the Philosophy of Language* (Urbana: University of Illinois Press, 1952).

[3]) Cf. Nelson Goodman, *The Structure of Appearance* (Cambridge, Mass.: Harvard, 1951).

of synonymy is simply *analytic* coextensiveness (however circular this might be as a definition).

To count analyticity a genus of logical truth is to grant, it may seem, the linguistic doctrine of logical truth; for the term 'analytic' directly suggests truth by language. But this suggestion can be adjusted, in parallel to what was said of 'true by definition' in § V. 'Analytic' means true by synonymy and logic, hence no doubt true by language and logic, and simply true by language *if* the linguistic doctrine of logical truth is right. Logic itself, throughout these remarks, may be taken as including or excluding set theory (and hence mathematics), depending on further details of one's position.

What has made it so difficult for us to make satisfactory sense of the linguistic doctrine is the obscurity of 'true by language'. Now 'synonymous' lies within that same central obscurity; for, about the best we can say of synonymous predicates is that they are somehow 'coextensive by language'. The obscurity extends, of course, to 'analytic'.

One quickly identifies certain seemingly transparent cases of synonymy, such as 'bachelor' and 'man not married', and senses the triviality of associated sentences such as 'No bachelor is married'. Conceivably the mechanism of such recognition, when better understood, might be made the basis of a definition of synonymy and analyticity in terms of linguistic behavior. On the other hand such an approach might make sense only of something like degrees of synonymy and analyticity. I see no reason to expect that the full-width analyticity which Carnap and others make such heavy demands upon can be fitted to such a foundation in even an approximate way. In any event, we at present lack any tenable general suggestion, either rough and practical or remotely theoretical, as to what it is to be an analytic sentence. All we have are purported illustrations, and claims that the truths of elementary logic, with or without the rest of mathematics, should be counted in. Wherever there has been a semblance of a general criterion, to my knowledge there has been either some drastic failure such as tended to admit all or no sentences as analytic, or there has been a circularity of the kind noted three paragraphs back, or there has been a dependence on terms like 'meaning', 'possible', 'conceivable', and the like, which are at least as mysterious (and in the same way) as what we want to define. I have expatiated on these troubles elsewhere, as has White.[1])

[1]) Quine, *From a Logical Point of View*, Essay II; White, *op. cit.*

Logical truth (in my sense, excluding the additional category of essential predication) is, we saw, well enough definable (relatively to a fixed logical notation). *Elementary* logical truth can even be given a narrowly syntactical formulation, such as Carnap once envisaged for logic and mathematics as a whole (cf. § VII); for the deductive system of elementary logic is known to be complete. But when we would supplement the logical truths by the rest of the so-called analytic truths, true by essential predication, then we are no longer able even to say what we are talking about. The distinction itself, and not merely an epistemological question concerning it, is what is then in question.

What of settling the limits of the broad class of analytic truths by fixing on a standard language as we did for logical truth? No, the matter is very different. Once given the logical vocabulary, we have a means of clearly marking off the species logical truth within the genus truth. But the intermediate genus analyticity is not parallel, for it does not consist of the truths which contain just a certain vocabulary essentially (in the sense of § II). To segregate analyticity we should need rather some sort of accounting of synonymies throughout a universal language. No regimented universal language is at hand, however, for adoption or consideration; what Carnap has propounded in this direction have of course been only illustrative samples, fragmentary in scope. And even if there were one, it is not clear by what standards we would care to settle questions of synonymy and analyticity within it.

X

Carnap's present position [1]) is that one has specified a language quite rigorously only when he has fixed, by dint of so-called meaning postulates, what sentences are to count as analytic. The proponent is supposed to distinguish between those of his declarations which count as meaning postulates, and thus engender analyticity, and those which do not. This he does, presumably, by attaching the label 'meaning postulate'.

But the sense of this label is far less clear to me than four causes of its seeming to be clear. Which of these causes has worked on Carnap, if any, I cannot say; but I have no doubt that all four have worked on his readers. One of these causes is misevaluation of the role of convention in con-

[1]) See particularly '*Meaning postulates*.'

nection with artificial language; thus note the unattributed fallacy described in § VIII. Another is misevaluation of the conventionality of postulates: failure to appreciate that postulates, though they are postulates always by fiat, are not *therefore* true by fiat; cf. §§ IV-V. A third is over-estimation of the distinctive nature of postulates, and of definitions, because of conspicuous and peculiar roles which postulates and definitions have played in situations not really relevant to present concerns: postulates in uninterpreted systems (cf. § IV), and definitions in double systems of notation (cf. § V). A fourth is misevaluation of legislative postulation and legislative definition themselves, in two respects: failure to appreciate that this legislative trait is a trait of scientific hypotheses very generally (cf. § VI), and failure to appreciate that it is a trait of the passing event rather than of the truth which is thereby instituted (cf. end of § V).

Suppose a scientist introduces a new term for a certain substance or force. He introduces it by an act either of legislative definition or of legislative postulation. Progressing, he evolves hypotheses regarding further traits of the named substance or force. Suppose now that some such eventual hypothesis, well attested, identifies this substance or force with one named by a complex term built up of other portions of his scientific vocabulary. We all know that this new identity will figure in the ensuing developments quite on a par with the identity which first came of the act of legislative definition, if any, or on a par with the law which first came of the act of legislative postulation. Revisions, in the course of further progress, can touch any of these affirmations equally. Now I urge that scientists, proceeding thus, are not thereby slurring over any meaningful distinction. Legislative acts occur again and again; on the other hand a dichotomy of the resulting truths themselves into analytic and synthetic, truths by meaning postulate and truths by force of nature, has been given no tolerably clear meaning even as a methodological ideal. One conspicuous consequence of Carnap's belief in this dichotomy may be seen in his attitude toward philosophical issues, e.g. as to what there is. It is only by assuming the cleavage between analytic and synthetic truths that he is able e.g. to declare the problem of universals to be a matter not of theory but of linguistic decision.[1]) Now I am as impressed as anyone

[1]) See Carnap, 'Empiricism, semantics, and ontology,' *Revue internationale de Philosophie*, 1950, especially § 3, longest footnote.

with the vastness of what language contributes to science and to one's whole view of the world; and in particular I grant that one's hypothesis as to what there is, e.g. as to there being universals, is at bottom just as arbitrary or pragmatic a matter as one's adoption of a new brand of set theory or even a new system of bookkeeping. Carnap in turn recognizes that such decisions, however conventional, 'will nevertheless usually be influenced by theoretical knowledge.'[1]) But what impresses me more than it does Carnap is how well this whole attitude is suited also to the theoretical hypotheses of natural science itself, and how little basis there is for a distinction.

The lore of our fathers is a fabric of sentences. In our hands it develops and changes, through more or less arbitrary and deliberate revisions and additions of our own, more or less directly occasioned by the continuing stimulation of our sense organs. It is a pale grey lore, black with fact and white with convention. But I have found no substantial reasons for concluding that there are any quite black threads in it, or any white ones.

Harvard University

[1]) *Op cit.*, § 2, fifth paragraph.

EVERT W. BETH

EXTENSION AND INTENSION

I. INTRODUCTION

One of the most important operations in the logic of science as conceived of by R. Carnap is the *explication* of a familiar but vague concept. It consists in replacing that concept by a new and exact concept; the earlier concept is called the *explicandum*, the new concept by which it is replaced is called the *explicatum*.

An operation of this kind can be carried out in two different ways. In the first place, we may start from a critical examination of such definitions or other elucidations of the explicandum as we happen to come across and try to convert these descriptions into a precise definition of a new concept which is then accepted as an explicatum. Secondly we may define the explicatum in a straightforward manner, independently of any traditional elucidations of the explicandum, and then try to show that the new concept thus obtained is indeed a suitable substitute for the earlier concept.

The purpose of the present essay is to apply the second procedure in an attempt to provide explications for the concepts of *extension* and of *intension*. These concepts are currently used in an epistemological context and with respect to terms which occur in certain domains of knowledge. Such domains may consist, for instance, of all cosmological insights of educated adult persons, of all astronomical insights of professional astronomers, or of all chemical insights of a professional chemist.

We shall assume that the insights belonging to each domain can be expressed by means of true sentences in a suitable terminology and that for each domain these sentences form a deductive system in accordance with elementary predicate logic. Of course, these assumptions are highly debatable ones; they are nevertheless justified by the fact that in similar epistemological discussions they are currently made, that they simplify the following observations, and that our main conclusions do not essentially depend on them.

The deductive systems corresponding to the above-mentioned domains of knowledge will be, respectively, referred to as E, as A and as C.

II. DEDUCTIVE SYSTEMS AND THEIR MODELS

It is essential for our discussion that the systems E, A, and C, as well as all similar ones, are *incomplete*.

An incomplete system S is a deductive system for which we can find two sentences U and $\overline{U}$, formulated in the terminology of S, such that neither U nor $\overline{U}$ is contained in S. Accordingly, the system S will have two kinds of models. For models of both kinds all sentences contained in S are true; but for models of the first kind U is true and $\overline{U}$ is false, and for models of the second kind U is false and $\overline{U}$ is true.

Now let M be any model of S. Those sentences which are formulated in the terminology of S and which are true with respect to M will form a certain deductive system $S°$. We observe that

(i) $S°$ is complete as, for any choice of a sentence U, either U or $\overline{U}$ is true with respect to M and thus must be contained in $S°$;

(ij) $S°$ includes S, as any sentence U in S is true with respect to M and thus must be contained in $S°$.

We can state this observation more briefly by saying that *each model M of S determines a complete extension $S°$ of S.*

It is a well-known fact that *different* models M and M' may determine *the same* complete extension $S°$. In that case (which for our present discussion is not particularly important) the terminology of S clearly does not enable us to describe the difference between M and M'.

The choice of a definite model M of S determines a unique *denotatum* $M(t)$ for each *constant* or *closed term t* in the terminology of S. In particular, as we have seen, it determines a unique *truth value* $M(U)$ for each sentence U.

III. EXTENSION AND INTENSION

Let us now consider a certain specific model $M°$ of S and a certain set $\{M\}$ of models of S; we assume that $M°$ belongs to $\{M\}$. Then we can define, *with respect to $M°$ and $\{M\}$*, the *extension* and *intension* of a constant or closed term t in the terminology of S, as follows.

(i) The *extension* of t is the denotatum $M°(t)$ of t;

(ij) The *intension* of t is the denotatum $M(t)$ of t as a function of M for M in $\{M\}$.

Hence t and t' have *the same extension*, if $M^\circ(t) = M^\circ(t')$, and *the same intension*, if, for every M in $\{M\}$, $M(t) = M(t')$. As M° is in $\{M\}$, sameness of intension implies sameness of extension; the converse implication clearly does not hold true.

In the following discussion, we shall assume that the set $\{M\}$ is the set of all models M of the deductive system S under discussion.

Now let U and V be sentences in the terminology of S. Then U and V will have *the same extension*, if U and V are both *true* or both *false* with respect to M°. Accordingly, both U and V or both $\overline{U}$ and $\overline{V}$ are in the complete extension S° determined by M°; this will be the case, if and only if both $U \rightarrow V$ and $V \rightarrow U$ are contained in S°.

U and V will have *the same intension* if, with respect to any model M of S, U and V have the same truth value, that is, if both $U \rightarrow V$ and $V \rightarrow U$ are true with respect to any model M of S. By the completeness theorem for elementary logic this will be the case, if and only if both $U \rightarrow V$ and $V \rightarrow U$ are contained in S.

IV. APPLICATIONS

Let us apply the above notions to the deductive system E characterized above. M° will be the *real world*, whereas $\{M\}$ will contain both the real world and all *possible worlds*. These 'metaphysical' terms will cause no trouble because of the assumptions which have been stated. Moreover, as E is incomplete, certain differences between possible worlds M and M' can be described by means of the terminology of the system E. In particular, this terminology makes allowance for the description of certain differences between the real world M° and a possible world M.

In asserting a sentence U in the terminology of the system E, we may refer either to all models in the set $\{M\}$ or only to the model M°. In the first case, the assertion will be justified if and only if the sentence U is in E; in the second case, it will be justified if and only if U is contained in the complete extension E° determined by M°.

(i) Let us suppose that the terminology of the system E contains the terms 'Morning Star' and 'Evening Star'. Presumably, the sentence:

$$\textit{Morning Star} = \textit{Evening Star}$$

will be in E° but not in E. Hence, there will be possible worlds in which the Morning Star is different from the Evening Star. Thus, with respect to M° and $\{M\}$, the terms 'Morning Star' and 'Evening Star' will have the same extension, but different intensions.

The situation will change if the system E is replaced by the deductive system A. By our suppositions, M° will still be a model of A, but the set $\{M'\}$ of all possible worlds which are models of A will be a proper part of the above set $\{M\}$. Accordingly, we may assume that the sentence:

Morning Star = Evening Star

is in A. Therefore, with respect to M° and $\{M'\}$, the terms 'Morning Star' and 'Evening Star' will have both the same extension and the same intension.

(ij) As a second example, let us consider a counterfactual conditional, such as:

If the moon were a green cheese, then $2 + 2 = 5$.

If referred to the real world M°, this sentence is true; hence it is contained in the system E° and so its assertion would seem to be justified. However, the subjunctive mood in the sentence makes it clear that it should be referred, not only to the real world M°, but rather to the set $\{M\}$ of all possible worlds. Presumably, the deductive system E contains the sentence:

$$2 + 2 \neq 5,$$

but *not* the sentence:

The moon is not a green cheese.

Hence the above counterfactual conditional cannot be contained in E and so its assertion is not justified.

Again, the situation changes completely if the system E is replaced by the deductive system A.

(iij) It is well known that definitions of the so-called *disposition terms* involve subjunctive clauses of the above kind. More often than not the opinion is defended that such clauses cannot be understood as material implications.

However, the above discussion suggests that subjunctive clauses can be understood as material implications, provided they be referred, not only to the real world M°, but rather to the set $\{M\}$ of all possible worlds. It would carry us too far to consider this point in more detail.

V. CONCLUSION

It has been shown that, for the concepts of *extension* and *intension* with regard to (closed) terms occurring in various domains of knowledge, an acceptable explication can be given by reference to the denotata of these terms with respect to the real world and to certain possible worlds.

In the context of this discussion the introduction of these apparently metaphysical entities will not cause any trouble, as they are nothing else but the *models* of certain deductive systems corresponding to each of the relevant domains of knowledge. It is of essential importance for our discussion that the deductive systems involved are incomplete.

It seems to follow from our discussion that an explication of the above concepts, and of a few related ones, does not require the introduction of a modal or otherwise intensional metalanguage. The explication can be stated in the semantic metalanguage which serves as a means of expression for the theory of models.

The University of Amsterdam

HENRYK MEHLBERG

THE PRESENT SITUATION IN THE PHILOSOPHY OF MATHEMATICS

I. THE PHILOSOPHY OF MATHEMATICS BETWEEN 1930 AND 1960

Thirty years have elapsed since an important international convention took place in Königsberg with a view to surveying the main schools of thought in the philosophy of mathematics. Although this convention was held in the city where Immanuel Kant was born, taught, wrote, and died, his philosophical views of mathematical knowledge, which had enjoyed a virtual monopoly in the nineteenth century, played a negligible role in the 1930 debates. As a matter of fact, these debates were almost entirely dominated by three rival schools of thought usually referred to as mathematical *logicism*, *intuitionism*, and *formalism*. None of these trends was related to Kant's philosophy of mathematics.[1]

The logicist group was represented in Königsberg by R. Carnap [2], the spokesman of intuitionism was A. Heyting,[3] and J. von Neumann [4] stated the case for formalism. The investigators who have been mainly responsible for these three trends took no direct part in the convention. Thus, modern logicism originated with B. Russell [5] and A. N. Whitehead,[5] whose work was decisively influenced by their older contemporary, G. Frege, and could be ultimately traced to Leibniz's philosophy of mathematics. Similarly, both the founder of the formalist school, D. Hilbert,[6] and the originator of modern mathematical intuitionism, L. E. J. Brouwer,[7] chose to be represented in Königsberg by the

[1] On closer analysis, this holds also of mathematical intuitionism in spite of the superficial similarity between the role of our intuition of time in mathematical knowledge as stressed by both Kant and Brouwer. Cf. I. Kant's *Prolegomena zu einer jeden künftigen Metaphysik*, sec. 10 (1783), and L. E. J. Brouwer, *Thesis* (1907).

[2] R. Carnap, 'Die logizistische Grundlegung der Mathematik,' *Erkenntnis*, *II* (1931), pp. 91–105.

[3] A. Heyting, 'Die intuitionistische Grundlegung der Mathematik,' *ibid.*, pp. 106–15.

[4] Johann von Neumann, 'Die formalistische Grundlegung der Mathematik,' *ibid.*, pp. 116–21.

[5] A. N. Whitehead and B. Russell, *Principia Mathematica*, Vol. I-III (1910–1913).

[6] D. Hilbert, '*Gesammelte Abhandlungen*,' III (1935), pp. 146ff.

[7] *Loc. cit.*

aforementioned members of their respective groups. The proceedings of the Königsberg convention made apparent several advantages and drawbacks of each of the three philosophies of mathematics without revealing a decisive superiority of any one of them.

Another international convention largely concerned with the philosophy of mathematics was recently held at Stanford, California.[1]) Ostensibly the three schools of thought labelled logicism, intuitionism, and formalism have still dominated the 1960 debates although not all of these schools were represented by the same investigators as in Königsberg. Actually, only A. Heyting was again the spokesman of the intuitionist movement, testifying thereby to his admirable Dutch perseverance and consistency of endeavor. Naturally, Professor Heyting became more conciliatory and appreciative of co-operation rather than competition among the three schools of mathematical philosophy during the three decades which elapsed between the 1930 and 1960 conventions. Naturally also, his emphasis on the co-operation of the three trends of mathematical philosophy has contributed to making more appealing and effective his 1960 presentation of mathematical intuitionism (as distinct both from logicism and formalism and from more recent 'constructivist' attempts [2]) at laying down a more satisfactory foundation for mathematics by resorting to the theory of general recursive functions developed in the last two decades. This shift in emphasis, however, did not affect the basic tenets of intuitionism and its undeniable ability to persist in its vicissitudes without changing intrinsically to any appreciable extent.

As for logicism and formalism, their 1930 protagonists have been replaced by other spokesmen in 1960, for obvious reasons. *J. von Neumann* passed away a few years ago and the formalist position was presented in a letter addressed to the 1960 convention by P. Bernays, Hilbert's closest associate and co-author of the formalist *magnum opus* on the foundations of mathematics.[3]) This letter made it apparent that Hilbert's initial outlook,

[1]) The 1960 International Congress for Logic, Methodology, and Philosophy of Science, August 24–September 2, 1960.

[2]) Cf. *Constructivity in Mathematics*, Proceedings of the Colloquium held in Amsterdam in 1957 (1959).

[3]) D. Hilbert and P. Bernays, *Die Grundlagen der Mathematik*, Vol. I (1934); Vol. II (1939).

both anti-logicist and anti-intuitionist,[1]) has been considerably 'liberalized' in the subsequent evolution of formalism.

The spokesman of logicism at Stanford was A. Church, one of the world's authorities on mathematical logic and foundational research. R. Carnap, who had represented the logicist philosophy in Königsberg, made an essential contribution to the debate on foundations of probability which took place in another section of the Stanford convention. He refrained, however, from commenting on the differences between various schools of mathematical philosophy, possibly because in the interval between the 1930 and 1960 conventions (actually as early as 1934) he developed a more comprehensive theory of mathematical knowledge, radically different from the logicist outlook. Carnap's new position, condensed in his 'Principle of Tolerance' and expounded in detail in his 1934 work on the philosophy of language,[2]) was subsequently modified to some extent by his adoption, after the 1936 International Philosophical Congress in Paris, of A. Tarski's [3]) 'semantic' approach to the logical analysis of language. The new mathematical philosophy of the leader of logical positivism turned out to be more 'liberal' than logicism in some respects and more restrictive in others.

The restrictive implications of Carnap's new outlook on mathematics followed from the fact that, between 1934 and 1936, he identified the philosophy of mathematics with the study of the *morphological* (or 'syntactical') aspects of the language used by mathematicians in the formulation of mathematical theories, on the understanding that the part played either by the human users of mathematical language, or by the nature of the entities which the language enables its users to talk about or refer to, is irrelevant to the philosophical problems of mathematics. This exclusive emphasis on a linguistic approach to problems in the philosophy of mathematics had not appreciably shifted when Carnap came to realize in 1936 that the philosophy of mathematics includes not only the morphology of mathematical languages but also their semantics as conceived, explored, and placed on a firm foundation by A. Tarski. Tarski's semantic investigations, which centered about the

[1]) D. Hilbert, *loc. cit.*

[2]) R. Carnap, *Die logische Syntax der Sprache* (1934).

[3]) A. Tarski, 'Grundlegung der wissenschaftlichen Semantik.' *Actes du Congrès de philosophie scientifique*, fasc. VII (1936).

extra-linguistic reference of the language under consideration, instead of exploring the syntactical or morphological forms and patterns of the relevant linguistic expressions, have shed new and revelatory light on several problems of great philosophical importance, e.g., the problem concerning the meaning of truth.[1])

After having adopted the 'semantical' viewpoint Carnap made important contributions to the study of many philosophically significant, semantic features of any language including languages used by mathematicians.[2]) He discovered useful, alternative procedures equivalent to Tarski's pioneering semantic method. However, the fact that, from then on, Carnap considered both the semantic and the morphological aspects of mathematical languages as being within the scope of mathematical philosophy did not affect the restrictive nature of his new outlook on mathematics as compared with logicism since he maintained the emphasis on a linguistic approach to the philosophy of mathematics after having extended the subject matter of the latter in the semantic direction.

On the other hand, it is important to realize that the more comprehensive and liberal nature of Carnap's new philosophy of mathematics as compared with logicism was also due to his linguistic approach. The liberalizing tendency of the linguistic approach can be explained as follows: since in Carnap's view the solution to problems in mathematical philosophy is determined by the language the mathematician uses in formulating his theories, any differences of opinion concerning the philosophical aspects of mathematics can be accounted for by assuming that those who hold conflicting views of mathematical knowledge have actually chosen different languages for the formulation of mathematical theories. This explanation holds, in particular, according to Carnap, for the apparent gap between intuitionism and logicism. In his 'Logical Syntax of Language' Carnap has already tried both to show in detail how the mathematical language used by intuitionists [3]) differs from the language preferred by logicists [4]) and to prove that the disagreement between these two philosophical views of mathematics is completely reducible to the linguistic preferences which are implicit in the two

[1]) A. Tarski, *The Concept of Truth in Formalized Languages* (1930, Polish).
[2]) R. Carnap, *Introduction to Semantics* (1946).
[3]) *Ibid.*, Chapter I-II.
[4]) *Ibid.*, Chapter III.

philosophies. Since the choice of a particular linguistic medium with a view to formulating a particular body of knowledge is merely a matter of convenience and manageability, no genuine cognitive question answerable by a true assertion or a true negation is involved. It seems obvious, therefore, that, according to Carnap's new position, the conflict between logicists and intuitionists (or, for that matter, between any rival schools of mathematical philosophy) would not involve any genuine philosophical problem as to who is right or who is wrong.

In a later exposition of his views [1]) Carnap stresses the dichotomy of two types of questions concerned with the language of mathematics or of any other science, viz., questions he calls 'external' and 'internal' with regard to a given language. Thus, in so far as mathematical language is concerned, the 'internal' questions refer to purely mathematical issues answerable within mathematics proper by well-known deductive procedures, e.g., the question of whether there is a finite or infinite number of prime numbers. On the other hand, according to Carnap, 'external' questions, e.g., those about the nature of mathematical existence or the subject matter of mathematical knowledge, are actually issues concerning the advantages or disadvantages of using a particular mathematical language as compared with other languages which the mathematician might prefer for his professional activities. These 'external' questions obviously belong to the philosophy of mathematics rather than to mathematics proper and are decidable in terms of convenience, conciseness, or manageability, without requiring direct evidence in support of a particular answer. In other words, since philosophical questions are of the 'external' variety, they actually concern how to choose the most convenient mathematical language from an infinite variety of possible languages. No such choice can be meaningfully said to be true or false, and this holds also of the mathematical philosophy underlying a linguistic choice of this kind.

In particular, the question of whether intuitionists, logicists, or formalists are right in their respective philosophical views of mathematics should be classified as 'external' and as admitting neither a true nor a false answer. Philosophically speaking, intuitionism, logicism, and formalism are all on a par. This is how Carnap's conciliatory attitude towards all

[1]) R. Carnap, 'Empiricism, semantics, and ontology,' *Revue Internationale de Philosophie*, Vol. IV (1950).

rival schools of mathematical philosophy turned out to be a direct consequence of his linguistic approach to this branch of the philosophy of science. No wonder that he was reluctant to comment, at the 1960 Congress, on the philosophical status of these three philosophies of mathematics.

II. LOGICISM IN 1960

In his address to the 1960 Congress at Stanford A. Church presented an outline of mathematical philosophy which may be termed '*moderate logicism.*' He pointed out that, in the original formulation of the logicist position, due mainly to B. Russell, the relationship between mathematics and logic has been characterized by the following tenets: (1) all mathematical concepts can be defined in terms of purely logical concepts, or, as Professor Church put it, 'the mathematical vocabulary is part of the logical vocabulary'; (2) all mathematical assumptions (axioms, postulates) can be derived from purely logical laws by applying well-established and purely logical modes of reasoning.

If these two initial, logicist claims could be substantiated, then mathematics would form a part of logic, according to the view defended by B. Russell in the first stage of the evolution of logicism. A. Church acknowledged, however, that the second logicist claim had proved to be untenable. He did not elaborate on this point but we may presume that he had in mind mathematical assumptions such as Zermelo's 'axiom of choice' or the 'axiom of infinity' used by Russell and Whitehead in their derivation of Peano's postulates from logical premises. There is hardly any reason either for considering these two axioms as intrinsically logical (or derivable from logical premises) or for questioning the need for these and for similar postulates in the foundations of mathematics. It seems, therefore, that Professor Church's suggestion that the second logicist claim (and, consequently, the claim that mathematics is literally part of logic) be dropped is fully supported by the present situation in the foundations of mathematics.

What about the first logicist claim that the mathematical vocabulary is part of the logical vocabulary? A. Church expressed the opinion that this claim, which has been made ever since the initial phase of the logicist philosophy of mathematics, was corroborated by subsequent developments in foundational research. It is obvious that although this claim cannot establish the view that mathematics is literally a part of logic, it could guarantee that, in a precise sense, logic is prior to mathematics. For suppose that some mathematical terms are clearly part of the logical vocabulary and that all the remaining mathematical terms are definable by means of, and hence replaceable by, an appropriate combination of logical terms (which is just another way of putting the first claim). Then, no mathematical term, proposition, or proof could be meaningful unless the corresponding logical terms were meaningful. Since the converse is not the case, one can see in what precise sense logic would be prior to mathematics provided that the validity of the first claim be assumed or established; the meaningfulness of mathematics would depend upon the meaningfulness of logic, but not vice versa.

To evaluate the present status of *moderate logicism*, i.e., the validity of the first logicist claim on presently available evidence, it is useful to recall briefly how this claim came to be supported by the impressive overall achievement of the 'century of mathematical criticism.' To put it in a nutshell: the successive generalizations of the concept of natural number (integers, rational, irrational, real, complex, and hyper-complex numbers) were first shown to be definable in terms of natural numbers. Similarly, the successive generalizations of the notion of number-theoretical function have proved definable in terms of this notion, parallel to the generalizations of those numbers over which the domain and the range of the generalized functions, respectively, vary. Later on, the decisive role played by the set-theoretical, conceptual framework in establishing the number-theoretical definability of the successively generalized concepts has been noticed. The next step was reached when the role of set-theoretical concepts in the number-theoretical definability of more complex mathematical entities induced several investigators (mainly Frege and Russell) to make a successful attempt at defining the number-theoretical concepts themselves in set-theoretical terms.[1] This was achieved by defining a cardinal number as the set of all sets isomorphic with a single set and a natural number as any cardinal number which is never assignable both to a particular set and to some proper sub-set of this set. The relation of isomorphism of sets, made use of in the aforementioned

[1]) B. Russell, *Introduction to Mathematical Philosophy* (1919).

definitions, has also proved to be definable in terms of sets, once every relation was shown by N. Wiener [1]) to be so definable.

The final problem which then arose was concerned with the definability of set-theoretical concepts in purely logical terms. Accordingly, the outcome of the entire foundational trend just outlined would come to the single question of whether or not the concept of set, and the group of other related set-theoretical notions, are definable in purely logical terms on the assumption that the position of moderate logicism is justifiable.

Some authors feel that the concept of set is intrinsically logical and that, consequently, the over-all result of the 'century of mathematical criticism' comes to the very claim of moderate logicism. As for myself, I find extremely illuminating and impressive the spectacular array of successful attempts at analyzing the conceptual apparatus of mathematics which have been made in the era of 'mathematical criticism' and have culminated in establishing the definability of mathematical concepts in terms of sets. I see little reason, however, for classifying the concept of set under the heading of logical concepts. There is no doubt, either, that many a respectable logician's name could be quoted in support of a non-logical status of the concept of 'set'. However, problems of this kind cannot be solved by counting noses. The important point is that, should the issue between logicism and intuitionism (or, for that matter, between Leibniz's and Kant's philosophies of mathematics) be made conditional only on whether or not the term 'set' qualifies as logical, then hardly any justice would be done to the Leibniz-Kant controversy over the nature of mathematical knowledge. The latter would then be interpreted as pertaining to lexicography and depending upon a somewhat different propensity for the lexicographic classification of terms like 'set'. There seems to be much more involved in the conflict of these two views of mathematical knowledge.

Another, much stronger argument against identifying the logicist philosophy of mathematics with attaching the label 'logical' to the term 'set' can be derived from a closer examination of the very structure of logical systems. I have in mind the undoubtedly logical theories concerning the propositional and the predicate-calculi both of which, including their undefined concepts, have to be utilized in conjunction with set-theoretical

[1]) N. Wiener, 'A simplification of the theory of relations,' *Proceedings of Cambridge Philosophical Society*, Vol. 17 (1912–14), pp. 387–390.

concepts in order to ensure the number-theoretical definability of all less elementary mathematical concepts. The point is that any satisfactory axiomatic foundation for the propositional and predicate-calculi consists of a list or description of all the undefined symbols of the logical calculus under consideration, a list of rules defining the class of 'well-formed' sequences of symbols (i.e., sequences which represent propositions within the calculus), a list of the unproved propositions or axioms of the calculus, and a list of rules of inference which are associated with the calculus and which enable the logician to derive his theorems from his axioms. Sometimes other lists are added to the four lists I have just mentioned, e.g., a list of rules specifying those sequences of symbols which are *terms* (not *propositions*) or a list which determines how a well-formed formula can be *refuted* (rather than demonstrated) on the basis of the listed axioms. The important feature of all such lists, no matter whether they are concerned with the supply of undefined symbols, or with criteria for well-formedness, or with rules of inference, is the circumstance that these lists [1] are not expressible in the purely logical language in which the axioms and theorems of the logical system under consideration are couched but require instead resort to ordinary language and to quite a few unmistakably mathematical concepts, including concepts of set-theory. For example, in most satisfactory systems of the propositional calculus, as exemplified in the system developed by Professor Church,[2] the list of undefined symbols is required to include an *infinite* (strictly speaking, an enumerably infinite) set of propositional variables. The concept of an enumerably infinite set is set-theoretical rather than logical since 'sets' could not possibly be defined at the start of logic and are nevertheless indispensable for the construction and the understanding of the logical calculus of propositions. It must be granted that the deductive system of the classical propositional calculus is susceptible to several distinct axiomatizations which may differ from each other, as well as from the axiomatization constructed by Professor Church, in regard to all the relevant lists. There is no shortage of such equivalent, alternative axiomatizations of the classical propositional

[1] With only the exception of list #3 in regard to those special axiomatic systems which do not involve 'axiom-schemata.' Cf. S. C. Kleene, *Introduction to Metamathematics*, (1952), p. 81ff., concerning the distinction between an 'axiom' and an 'axiom-schema.'
[2] A. Church, *Introduction to Mathematical Logic*, Vol. I, (1955) p. 69ff.

calculus. Yet no matter how much they differ from one another in regard to their respective lists of undefined symbols, of well-formed sequences of symbols, of axiomatic propositions, and of rules of inference, the fact is that the definitions of these four groups of rules, essential for the logical system under consideration, are always couched in ordinary language and involve arithmetical (i.e., in the last analysis, set-theoretical) concepts. Granted, logical systems can be satisfactorily axiomatized so as not to involve any axioms at all; the rules of inference can do any job of which the axioms are capable. But this circumstance only emphasizes the indispensable part played by mathematical, and hence by set-theoretical, concepts in any intelligible axiomatic system of logic. Never mind that, for good reasons, the rules of formation (i.e., the list of symbols and the definitions of well-formed sequences of symbols) and the rules of transformation (i.e., the list of axioms and of rules of inference) are often said to be formulated in the *meta-language* rather than in the *object-language* of the system. This does not prevent these rules from being integral parts of the system and from involving set-theoretical concepts.

The conclusion following from this second argument is clear: it may be the case that all mathematical terms, including the term 'set', are either part of an admittedly logical vocabulary or that they can be defined by appropriate combinations of items contained in this vocabulary. Such a fact, even if granted, would not establish the priority of logic to mathematics in the aforementioned, precise sense, because set-theoretical concepts are definitely needed both for setting up the basic logical theories of the propositional and predicate calculus and for making these two calculi intelligible and meaningful. All one can grant is the symmetrical relation of mutual inter-connection between mathematical and logical concepts. This symmetrical relation could not possibly establish the asymmetrical priority of logic to mathematics in the aforementioned sense.

We may mention that attempts have been made at formalizing the meta-language required to express the aforementioned four groups of rules inherent in logical systems.[1]) It is obvious that, in these special cases, the meta-language is distinct from ordinary language. However, an examination of such meta-languages shows readily that their vocabulary is not part of the logical vocabulary.

[1]) S. Leśniewski, 'Grundzüge eines neuen Systems der Grundlagen der Mathematik,' *Fundamenta Mathematicae*, Vol. XIV (1929).

Finally, let us mention a third argument against granting a logical status to the concept of set. The point is, simply, that abstract set-theory has been axiomatized in so many ways which are obviously incompatible with each other [1]) that the inclusion of the notion of set within logic would make logic a highly speculative and controversial discipline, hardly suitable for laying down firm 'foundations' for mathematics, including the theory of natural numbers: who could possibly doubt that the 'foundations' of statements like '$2 + 3 = 5$' or '$1 = 1$' cannot be supplied by a logic which includes or entails the axiom of infinity; or the claim that the set of all sets is its own element; or, alternatively, the claim that the assertion to the effect that the set of all sets is its own element does not make sense?

Shall we infer from the aforementioned three arguments that, since on presently available evidence the first logicist claim seems to be as untenable as the second claim, the position of logicism, however moderate, must be definitively abandoned? In the sequel, I shall argue that this is only apparently the case. In particular, we shall see that *pluralistic logicism* is capable of providing a satisfactory adjustment to the present situation in foundational research because such a version of logicism supplies a common basis for the main trends in contemporary philosophy of mathematics.

III. FORMALISM IN 1960

1. *The Implications of the Unrealizability of Formalist Objectives.*

The formalist philosophy of mathematics originated with D. Hilbert and consisted initially of a vast research program. The program included the following major objectives: (1) to prove the consistency of classical mathematics by using special 'finitary' methods (i.e., procedures which meet somewhat more stringent conditions than those recommended by intuitionists) within a new discipline called 'metamathematics' or 'proof-theory'; (2) to solve the decision-problem of classical mathematics, i.e.,

[1]) Thus, in some axiomatizations of set-theory, the concept of a self-containing set is meaningless whereas, in other axiomatizations, it has a meaning and can be shown to apply to some sets. Similarly, the set of all sets is meaningless in some axiomatic systems of set-theory although its existence is provable in other set-theoretical systems. Furthermore, in some systems of set-theory, a sharp line is drawn between sets and classes while other systems do without this distinction. Cf. A. Fraenkel and J. Bar-Hillel, *Foundations of Set-Theory* (1958).

to discover effective methods capable either of solving any pre-assigned mathematical problem in a finite number of predetermined steps or of proving that the problem has no solution. This second objective was often formulated somewhat inaccurately as tantamount to the principle that 'every mathematical problem is solvable.'

By 1937 both formidable objectives of the formalist program were admitted everywhere (including Göttingen) to be unrealizable in principle, in view of logical discoveries made by K. Gödel [1]) and A. Church [2]) in 1931 and 1936 respectively. Among Gödel's several 1931 results, one was concerned with the unexpected fact that any non-trivial mathematical theory (I mean, any theory capable of providing proofs for all those properties of natural numbers which are derivable from Peano's classical set of five number-theoretical postulates) cannot be proved to be (simply) consistent without circularity since its consistency-proof would have to make use of these very properties of natural numbers and of additional, independent assumptions transcending number-theory. This shows that Hilbert's first objective was unattainable in principle.

The unsolvability of the problem which underlies Hilbert's second objective is a direct consequence of a result obtained in 1936 by A. Church. He showed then the non-existence of a decision-procedure for the lower predicate-calculus, i.e., the impossibility of a uniform procedure for solving all the problems of this calculus in a finite number of predetermined steps. Church's discovery came as a surprise because of the comparative simplicity of this calculus and in view of the fact that the calculus was known to be complete, in the sense that all its valid formulae (i.e., all formulae expressible in the calculus and true in every non-empty universe of discourse) are provable within the calculus. The impact of the result, however, was due to another circumstance, viz., to the fact that the non-existence of a decision-procedure for a particular axiomatic system was thereby established and, consequently, the unsolvability of the general decision problem shown.

Church's theorem has then been successively extended to several logical

[1]) K. Gödel, 'Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme,' *Monatshefte für Mathematik und Physik*, Vol. XXXVIII, (1931) pp. 173–198.

[2]) A. Church, 'A Note on the Entscheidungsproblem,' *The Journal of Symbolic Logic*, Vol. I (1936), and his abstract in the *Bulletin of the American Mathematical Society*, Vol. XLI, (1935) p. 333.

and mathematical theories such as number-theory, elementary group-theory (i.e., group-theory to the extent to which it is expressible in the lower functional calculus), the theory of topological spaces, the simplified theory of logical types, etc. Moreover, A. Tarski [1]) and his associates have devised general methods for determining whether or not any pre-assigned theory admits of a decision-procedure. The number of theories known to be devoid of a decision-procedure was thereby considerably increased. At this juncture it will suffice to notice that all these results show unmistakably that Hilbert's second objective is unrealizable in principle.

In view of the aforementioned two negative results it may seem that the entire formalist program, which has been initially made up of the problems concerned with the decision-procedure and the consistency of classical mathematics, should be abandoned since both problems have turned out to be unsolvable in principle. It is all too human that, at the beginning, Gödel's 1931 discoveries, which also underlie Church's 1936 finding, have been looked upon suspiciously by the Hilbert group in Göttingen. Professor Bernays told me that, during a brief initial phase, Göttingen presumed that some decisive and elusive mistake was implicit in Gödel's proofs. However, it did not take long for Göttingen to realize that this conjecture was unwarranted. One cannot but admire the way Professor Bernays then faced the situation created by Gödel's discoveries and responded to their challenge. His memorable reaction can be summarized as follows: Since the decision problem is now known to be unsolvable in most cases, we shall have to determine the mathematical theories which *do* admit of a decision-procedure. Moreover, since the problem of proving the consistency of classical mathematics within a finitary metamathematics turned out to be unsolvable, we shall from now on drop the requirement of finitary admissibility. Instead, we shall try to discover whatever is discoverable and important about classical mathematics by using metamathematical methods regardless of whether the latter may be finitary or must be non-finitary. There is no need to elaborate on how fruitful and momentous this courageous response to Gödel's and Church's discoveries has been.

The questions I should like to raise in connection with the 1960 status

[1]) A. Tarski, R. M. Robinson, A. Mostowski, *Undecidable Theories* (1953).

of formalism are concerned with the following points: (a) the implications of the failure of Hilbert's initial program, and (b) the extent to which this program was actually affected by the aforementioned discoveries. Since the unsolvability of the general decision-problem, as discovered by Church in a limited domain and then successively extended by A. Tarski and his associates, seems definitive, I shall confine my comments on this point to a single remark: 'So much the better.' For the effect of this negative result is simply the fact that the creative mathematician and logician could not possibly be replaced by any computer, however ingeniously built, and that, socially speaking, scientific progress will keep these creative thinkers from unemployment regardless of any conceivable advance in automation.

However, the first negative result has none of these advantages. On the contrary, if it were the case that we shall never be able to prove the consistency of the basic mathematical theories then, by the same token, we would be incapable of *knowing* that these theories are *consistent*, let alone that they are *true*. Our ignorance of the consistency and of the truth of these theories would, in turn, entail the impossibility of knowing whether any other theory which presupposes the basic mathematical theories is true. Practically speaking, any and every scientific theory would be involved. Gödel's result would imply that virtually no scientific theory could be classified as knowledge. The only alternative we would be left with in regard to science would be *belief*: since *knowing* the truth of these theories would transcend the scope of man's potentialities, he would have to *believe* blindly in the validity of his general outlook, which is presently based on scientific information; he would also have to believe blindly in the dependability of the individual and social actions which he is presently compelled to take on the basis of scientific information. Thus, *belief*, rather than *knowledge*, would have to be admitted as the basis of our theoretical outlook and practical activity in view of the pervasive part which science has come to play in man's life.

I do not claim that such a conclusion would amount to a catastrophe. But, to begin with (1), I do wonder whether the impossibility of attaining Hilbert's first objective, i.e., of constructing a finitary, metamathematical consistency-proof for classical mathematics, would actually justify the over-all replacement of scientific knowledge with belief. And I also wonder (2) to what extent Hilbert's initial program has actually been shown unrealizable by Gödel's 1931 results. Let us first evaluate briefly (1), i.e., the impact upon the sum-total of scientific knowledge of the alleged impossibility of proving the consistency of basic mathematical theories. The view that the whole of scientific knowledge would be affected by the impossibility of proving the consistency of mathematical (and logical) theories presupposes, obviously, that knowledge based on proof, on the one hand, and belief, on the other, are the only two alternatives available to man. But this is hardly the case. I am not referring to empirical knowledge of facts perceived with or without instruments, or remembered, since in these cases the dichotomy is clearly wrong. However, in the case of logical and mathematical theories, the most typical cases of knowledge are certainly those involving a deductive proof; hence, it seems natural that, in the logico-mathematical field, all knowledge-situations are often considered as involving proof. This, however, is certainly not the case. For the most advanced form of organized knowledge in the fields of mathematics and logic is certainly represented by a successful application of the axiomatic method. And the knowledge of most informational items which belong in a given axiomatic system is certainly derived from a deductive proof for this item of information. But it is obvious that the truth of the axioms of the axiomatic system under consideration could not possibly depend on the availability in the system of a valid deductive proof for them. Hence, unless a derivative system is envisaged, the axioms of which are provable in some other, ultimately non-derivative system, we have to acknowledge that if we know at all that the axioms of the non-derivative system are true, such knowledge is not the result of a deductive proof.

It may be granted that, in quite a few axiomatic systems, the relevant axioms may be claimed to be 'self-evident' and 'self-evidence' may be considered as a second source of reliable information, or knowledge, in the logico-mathematical field, apart from deductive proof. This view has been considerably weakened during the evolution of the axiomatic method in the last century but it is far from being definitively refuted.[1]) However, even if granted, the cognitive significance of self-evidence would be of no avail in regard to the consistency-problem, since the consistency of basic mathematical theories is far from self-evident. The central importance

[1]) H. Mehlberg, *The Reach of Science* (1958), p. 225.

acquired by set-theory in mathematical science illustrates this point sufficiently: nobody would consider as self-evident any one of the major available axiomatizations of set-theory. If set-theory, in turn, is used as a basis for number-theory, or algebra, or analysis, or any other basic mathematical theory, then the self-evidence of the latter would not be compatible with the fact that set-theory is admittedly devoid of self-evidence.

The only sensible way out of this predicament created by Gödel in 1931 which seems open at present is the idea that we should take the difference between mathematics and metamathematics seriously. This is, at bottom, Hilbert's basic attitude. However, this great mind interpreted the distinction between these two inter-related and distinct fields of science as implying the following difference between the kinds of knowledge which are available or desirable in the two fields, respectively: in classical mathematics, knowledge is usually based on a deductive proof from axioms except for the axioms themselves which we may assume as self-evident or intuitive, or as to be chosen so as to insure their intuitiveness. On the other hand, knowledge of how problems in metamathematics are solved should, according to Hilbert, be always of a 'finitary' kind, i.e., of a kind satisfying a more stringent requirement of intuitiveness than the requirement adopted in mathematics proper. This view of Hilbert's is historically intelligible since he was faced with intuitionist competitors. But we may doubt whether the historical explanation of Hilbert's position provides the epistemological justification of this position. The point is that, if metamathematical knowledge is required to satisfy too rigid conditions, we shall find out that there is no metamathematical knowledge. In particular, the metamathematical problem of the consistency of mathematics may prove to transcend the potentialities of human knowledge if the knowledge of a system's consistency were expected to meet the unrealistic conditions which were inherent in the initial phase of the formalist program. Yet the case of empirical sciences shows clearly that knowledge may be *dependable* to a high degree without being deductively established and, therefore, virtually infallible. And the requirement of dependable rather than of deductively established or virtually infallible knowledge can point to a way out of the predicament created by the demonstrable impossibility of a non-circular consistency-proof for basic mathematical theories.

What I am driving at is simply the idea that a mathematical theory can be dependably known to be consistent if a reasonable number of appropriately variegated attempts at deriving a contradictory consequence from this theory have been made over an adequate time by competent investigators and if all these attempts have proved unsuccessful. Granted, their failure does not theoretically preclude the possibility that, in a more or less remote future, a lucky investigator may succeed in deriving such consequences from this theory. This theoretical possibility points simply to the fact that our knowledge of the consistency of the theory is not infallible although it is dependable to a very high degree, But, short of infallibility, an adequate degree of dependability is perfectly sufficient to warrant our classifying the consistency of this theory under the heading of knowledge rather than belief. The situation is exactly the same as in the empirical sciences, where no infallible knowledge is available and where dependable knowledge is carefully distinguished from more belief.

I would like to adduce one more argument in support of the view of metamathematical knowledge just outlined. It may well happen that, with a view to unifying mathematical science, several theories may be axiomatized within a single comprehensive system in spite of the fact that the axioms of this comprehensive system are less dependably known to be true than some theorems of the unified theories and, consequently, less dependable than some theorems of the comprehensive system, too. Thus, elementary arithmetic can be (and sometimes was) axiomatized on a set-theoretical basis. Yet nobody would claim that we know the truth of set-theoretical axioms in any major available axiomatization of the theory of sets more dependably than we do know the truth of grade-school arithmetic. Hence, the axiomatic unification of mathematical knowledge (including the reduction of mathematics to logic, or to set theory, or to both) provides a foundation for mathematics only in those cases where the axioms of the unifying system are more dependable than the axioms and theorems of the set of mathematical theories which are unified or reduced to logic within the single, comprehensive system. Consequently, the failure of Hilbert's grandiose program of unification can hardly affect the dependability of the unified mathematical theories.

The limited effect of the failure of Hilbert's program upon the dependability of the impressive cluster of mathematical theories which he tried to place on a common 'foundation' can be clarified by reference to certain relevant views of Gödel's which he informally conveyed to me, some years ago, during a discussion we had at Princeton, N. J. According to Gödel, an axiomatization of classical mathematics on a logical basis or in terms of set-theory is not literally a foundation of the relevant mathematics, i.e., a procedure aiming at establishing the truth of the relevant mathematical statements and at clarifying the meaning of the mathematical concepts involved in these theories. In Gödel's view, the role of these alleged 'foundations' is rather comparable to the function discharged, in physical theory, by explanatory hypotheses. Thus, in the physical theory of electro-magnetic phenomena, we can explain why the sky looks blue to us under normal circumstances, and we are even able to produce the same phenomenon in the laboratory. Both the explanation of the physical phenomenon under consideration and its production under laboratory conditions are due to the logical fact that the statements describing the blue of the sky or that of an artificially produced area in the laboratory are theorems provable within an axiomatic system the postulates of which are concerned with hypothetical laws governing electro magnetic phenomena, the composition of the atmosphere, etc. It would not occur to a physicist that these electro-magnetic assumptions which enjoy the role of postulates in an axiomatized, or axiomatizable physical theory, are more dependably known to be true than the prescientific phenomena (like the blue of the sky) which are being explained by being shown to be provable theorems in the aforementioned physical theory. Thus, the actual function of postulates or axioms occurring in a physical theory is to *explain* the phenomena described by the theorems of this system rather than to provide a genuine 'foundation' for such theorems. Professor Gödel suggests that so-called logical or set-theoretical 'foundations' for number-theory, or any other well established mathematical theory, is explanatory, rather than really foundational, exactly as in physics.

2. *The Limitations of Formalism due to the Discoveries of K. Gödel and A. Church*

To sum up: the impossibility of constructing a deductive proof for the consistency of classical mathematics by finitary, metamathematical procedures would not entail a wholesale obliteration of the boundary line which separates scientific knowledge from mere belief. We shall now discuss briefly the limitations of Hilbert's initial program which Gödel's and Church's discoveries have made apparent. The extent of these limitations can be determined by analyzing the role played by non-finitary assumptions in the available consistency-proofs for classical mathematics. We shall confine ourselves to consistency-proofs for classical number-theory and shall select (as representative examples) the *absolute* consistency-proof for number-theory discovered by Gentzen by resorting to a non-intuitionist extension of Hilbert's finitary basis, and to the relative, intuitionistic consistency-proof for classical number-theory derivable from several related results of Gödel, Kleene, Mostowski, and other investigators. Let us start with the relative, intuitionist consistency-proof.

The gist of the proof may be put as follows: the fact that any proof valid in Holland (i.e., intuitionistically) is also valid outside of Holland (classically) whereas the converse is not the case was apparent and generally known ever since Professor A. Heyting's formalization of intuitionist logic (1930). However, as a result of an impressive array of results due to Kolmogorov,[1]) Gödel,[2]) Kleene,[3]) Menger,[4]) Kreisel,[5]) and other investigators, it has become equally clear that, whenever a theorem T is classically provable, it suffices to replace T with a slightly modified formulation T' of T in order to ensure the intuitionistic provability of T'. Since T and T' are classically equivalent while T and non-T can be shown to be equivalently transformable into T' and non-T', respectively, it is obvious that the classical provability of T and of non-T would entail the intuitionistic provability of T' and of non-T'. In other words, if classical mathematics were inconsistent then so would be intuitionistic mathemat-

[1]) A. Kolmogorov, 'О принципе tertium non datur', *Recueil Mathématique de la Société Mathématique de Moscou*, Vol. XXXII (1924-5), pp. 646-667. I owe the historically important reference to Kolmogorov's Russian paper to Prof. A. Church.

[2]) K. Gödel, 'Zum intuitionistischen Aussagekalkül,' *Ergebnisse eines mathematischen Kolloquiums*, (1933), p. 40.

[3]) S. C. Kleene, 'On the intuitionistic logic', *Proceedings of the Tenth International Congress of Philosophy*, (1949), pp. 741–743.

[4]) K. Menger, 'Bemerkungen zu Grundlagen,' *D.M.V.*, Vol. XXXLII, (1928), pp. 213–226, 298–323. In contrast to inter-translatability relations referred to in the footnotes ≠2)–3) and concerned with classical and intuitionist logic and number-theory, Menger's result establishes a parallelism of intuitionist set-theory and the classical theory of 'analytic' (or Suslin) sets.

[5]) G. Kreisel, Interpretation of Analysis by means of constructive Functionals of finite Types, cf. Ref. ≠2 on p. 101 of Section I, pp. 101-128.

ics. This may be considered as a relative consistency-proof for classical mathematics and hence as an approximate and partial solution of Hilbert's first foundational problem.

Needless to say, I do not question the cogency of Gödel's 1931 result that classical mathematics cannot be shown to be consistent without circularity. Nor do I doubt that, on its literal interpretation, Hilbert's problem of a finitary metamathematical consistency-proof has been shown by Gödel to be unsolvable. It seems to me, however, that on its literal interpretation Hilbert's first foundational problem refers to an absolute and finitary, metamathematical consistency-proof. What I did stress as a consequence of the inter-translatability relations between classical and intuitionist logic and number-theory as established by Gödel, Kleene, and other investigators, is an intuitionist, relative consistency-proof for classical logic and number-theory. It is hardly questionable, however, that such a relative intuitionist consistency-proof comes so close to a partial solution of Hilbert's first foundational problem as to warrant the claim that, owing to the aforementioned inter-translatability relations, a partial solution to this formalist problem has actually been obtained.

The qualifications of this solution, from a formalist viewpoint, are obvious: the most serious gap is undoubtedly due to the circumstance that only some evidence in support of number-theoretical consistency has been produced. We are, accordingly, left in the dark as to the possibility of an intuitionist and relative consistency-proof, for other basic mathematical theories like analysis,[1]) theory of function-spaces, variational calculus, etc. On the other hand, we are entitled to disregard the distinction between the intuitionist and the finitary viewpoint as described, e.g., by Bernays. Since finitary formalism makes fewer assumptions than intuitionism, a relative intuitionist consistency-proof guarantees *a fortiori*, formalist consistency – on the assumption that intuitionist mathematics is consistent.

To determine the limitations discovered in the finitary positions by Gödel and Church more accurately, we shall now discuss the non-finitary assumptions involved in another consistency-proof for number-

[1]) However, Kreisel succeeded in extending relative intuitionist consistency-proofs to virtually the whole of classical analysis. Says he: "If intuitionist analysis is a-consistent then classical analysis is consistent." (loc. cit., p. 108).

theory, *viz.*, in contrast to our previous example, a case of an absolute consistency-proof. G. Gentzen [1]) succeeded in constructing such an absolute consistency-proof for number-theory. In order to avoid a clash with Gödel's denial of the possibility of such an absolute consistency-proof within the framework of *Principia Mathematica*, Gentzen has introduced an additional, powerful rule of inference which Gödel did not take into consideration, viz., the rule of transfinite induction. To obtain his consistency-proof, he was not compelled to apply the rule of transfinite induction to all transfinite, ordinal numbers. It turned out that it would suffice to stipulate that this rule applies to a segment of Cantor's sequence of transfinite ordinal numbers, *viz.*, to those numbers which are below ε_0. The question which therefore arises in connection with Gentzen's proof, and with similar attempts made by other investigators, is concerned with the extent to which recourse to transfinite induction (limited in the aforementioned way) exceeds finitary requirements. Unfortunately, the requirement of 'finitary' admissibility of any metamathematical proof was never made clear, although it has characterized the formalist philosophy from its very beginning. It is even more regrettable that this undefined requirement became so misleading as to induce several outstanding investigators to put several mutually incompatible interpretations [2]) on the requirement.

Let us start, however, with an official explanation of the meaning of 'finitary admissibility' by a most authoritative exponent. In his paper, 'Sur les questions méthodologiques de la théorie hilbertienne de la démonstration'[3]), Professor Bernays compares the methodological assump-

[1]) G. Gentzen, 'Die Widerspruchsfreiheit der reinen Zahlenlehre,' *Mathematische Annalen*, (1936), Vol. CXII.

[2]) Thus, P. Bernays emphasizes the difference between the finitary and the intuitionistic (or constructivistic) viewpoints, in contrast to J. Herbrand, who simply identifies these viewpoints. Cf. the lucid presentation of the former's position in his contribution to: F. Gonseth, *Les entretiens de Zurich sur les fondements de la méthode des sciences mathématiques*, (1941), pp. 144–152 and Herbrand's explicit identification of the two relevant positions '...an intuitionistic (i.e., a finitary) argument': J. Herbrand, 'Sur la non-contradiction de l'arithmétique.' *Journal für reine und angewandte Mathematik*, (1932), Vol. CLXVI.

Moreover, K. Menger has indicated several significantly distinct degrees of constructivity and has shown the arbitrariness of any choice among those various concepts of constructivity. Cf. K. Menger, 'Bemerkungen zu Grundlagenfragen,' *Jahresbericht D.M.V.*, (1929), Vol. XXXVII.

[3]) F. Gonseth, *op. cit.*

tions inherent in the 'finitary attitude' with the intuitionist assumptions, on the one hand, and those made by Gentzen in his consistency-proof, on the other. Since Bernays admits that the inadequacy of finitary assumptions for a consistency-proof of number-theory was established by Gödel, he is prepared to go beyond Hilbert's initial finitary requirements. He stresses, however, that an extension of Hilbert's assumption which would suffice to secure a consistency-proof for number-theory, as exemplified by Gentzen, need not go as far as intuitionists do.

We thus obtain a sequence of three increasingly liberal requirements, *viz.*, the finitary, the Gentzen, and the intuitionist, requirement. The finitary requirement is more stringent than the intuitionist since the former bans from mathematics the negation of any proposition which arises by universal quantification over an infinite domain. The finitary requirement also prohibits the use of any proposition which is universally quantified over an infinite domain, in the antecedent of a conditional. On the other hand, the intuitionist ban applies only to the application of the Law of Excluded Middle to infinite sets and affects, therefore, only some of the statements banned by the finitary requirement. Finally, Gentzen's additional, non-finitary assumption extends the validity of inductive reasoning into the transfinite domain, in contrast to B. Russell's suggestion of defining natural numbers as verifying the inductive reasoning.[1]) As a matter of fact, to obtain his consistency-proof, Gentzen need not stipulate the validity of mathematical induction for admittedly non-constructible, transfinite, ordinal numbers. It suffices to stipulate transfinite induction for a segment of Cantor's transfinite ordinal numbers of the second class.

We shall not discuss the implications of the aforementioned relationship between finitary and intuitionist requirements; since the former has been officially abandoned, this relationship is of no more than historical interest. It is important, however, to ascertain (1) whether Gentzen's requirements are more liberal than those recommended by intuitionists, and, more importantly, (2) whether his consistency-proof for number-theory takes care of the predicament created by Gödel's 1931 result.

In connection with question (1), we have to observe that if the partial validity of transfinite induction assumed by Gentzen were implied by intuitionist mathematics where the rule of transfinite induction is not being made use of explicitly, then this partial transfinite rule would have

[1]) B. Russell, *Introduction to Mathematical Philosophy*, (1919).

to be demonstrably a derivative inference-rule in the intuitionist calculus. This, however, is not the case since, in view of Gödel's result, the transfinite induction rule cannot be provable in the system of *Principia Mathematica* and is, therefore, unprovable in intuitionist mathematics, *a fortiori*.

As for question (2), it would seem to me that even if Gödel's ban on consistency-proofs for number-theory is overcome by postulating the additional validity of transfinite induction, the circularity of such a consistency-proof remains obvious. For Gentzen has to use Peano's axioms supplemented by his transfinite rule in order to prove the consistency of Peano's axioms. This is hardly illuminating, even if it is not at variance with Gödel's result.

Moreover, the consistency of number-theory, in spite of its intrinsic interest, is a far cry from a consistency-proof for other theories of classical mathematics like analysis or set-theory. As long as no consistency-proof for the core of classical mathematics will be made available, partial consistency-proofs will be of limited interest only.

Another essential qualification follows from another 1931 Gödel result concerning the essential incompletability of any axiomatization of number-theory. This basic result shows that, even if the class of provable theorems in an axiomatic system for number-theory is shown to be consistent, no guarantee of consistency is provided for the non-axiomatizable class of all true number-theoretical statements nor even for the class of those number-theoretical statements which are consequences (in the model-theoretical sense) [1]) of Peano's axioms. Gödel's example of an undecidable number-theoretical proposition which can be shown to be true by metamathematical methods is a case in point; his proposition is a model-theoretical consequence of Peano's axioms and yet undecidable by means of these axioms. The consistency of the class of model-theoretical consequences of number-theoretical axioms is not included in Gentzen's result.

Finally, a difficulty which we shall discuss in some detail in connection with intuitionism should be pointed out in regard to Gentzen's result. It is obvious (and explicitly admitted by Professor Bernays) that both finite and transfinite induction are acceptable from the viewpoint of an

[1]) A. Tarski, 'Über den Begriff der logischen Folgerung,' *Actes du Congrès International de Philosophie Scientifique*, (1936), Vol. VII.

extended finitary attitude only on the assumption that the hereditary property whose universal applicability has to be established must be finitary in some sense. But, as we shall see, attempts at circumscribing in a precise and intelligible way those properties which are eligible for an application of the principle of induction have failed to clarify this concept, so far. Hence, even the claim that the finitary attitude involves weaker assumptions than the intuitionist cannot be tested because the concepts of 'finitary' property, of 'definite' property (Carnap),[1]) of recursive property, etc., have all proved inadequate for clarifying the scope of the inductive principle.[2])

IV. INTUITIONISM IN 1960

I have already referred to the surprising ability of mathematical intuitionism to stand up under a barrage of more or less justifiable criticisms and to endure without undergoing any appreciable intrinsic change. I have also indicated that this intuitionist perseverance is perfectly compatible with a growing readiness to co-operate with other trends in the philosophy of mathematics without making any essential sacrifices in point of doctrine. The question I should like to discuss in this section is whether it is presently possible to define a foundational basis both common to the main trends in the philosophy of mathematics and free from the serious difficulties which the developments of the last three decades have proved to be inherent in the intuitionist position. It seems to me that the answer is a definite 'yes' and that *pluralistic logicism* can be shown to provide such a basis.

Let us notice, to begin with, that the inter-translatability relations between intuitionist mathematics and logic on the one hand and classical mathematics and logic on the other reduce the gap between intuitionism and logicism to a considerable extent. For suppose that all classically provable theorems can be bi-uniquely mapped on intuitionistically provable theorems on the understanding that the mapping is brought

[1]) R. Carnap, *Die logische Syntax der Sprache*, (1934).

[2]) The evaluation of the present status of formalism which I have proposed in the third section of this paper seems to apply without any major change to the independent presentation of the formalist viewpoint by Professor Curry, whose position must not be confused with the Göttingen position. Cf. H. B. Curry, *An Outline of the Formalist Philosophy of Mathematics*, (1951).

about by replacing a classically provable theorem T by another theorem T' where T' is both classically equivalent to T and provable within the intuitionist framework. Suppose also that T' can be obtained from T in a finite number of predetermined steps by applying a set of inter-translatability rules which are fixed in advance for the two associated (classical and intuitionist) theories. This being the case, it seems impossible to escape the conclusion that these two associated theories consist basically of the same results since the only difference between the two theories would then reside in the circumstance that the small number of logically constant symbols (propositional connectives and quantifiers) are somewhat differently interpreted by intuitionists and classicists respectively.

It is obvious that the intuitionists might argue (as Professor Heyting actually did during a discussion we had in Chicago in 1957) that this difference of interpretation is all that matters. It is equally obvious that intuitionists are entitled to determine their evaluation of scientific results in any way they choose, and that their evaluation deserves to be fully appreciated. Let us notice, however, that if the difference separating classicists from intuitionists would affect only their respective interpretations of logical symbols, then intuitionism would become a version of logicism, no matter how much the two interpretations of logical terms differ from each other. More importantly, it would seem that, from a practical point of view, a working mathematician could hardly be expected to appreciate the separate and individual nature of intuitionist mathematics under the circumstances just mentioned. For the mathematician is essentially an *architect of proofs*, no matter whether he lives inside or outside Holland. Suppose he is told that, given a classical mathematical theory, he can quietly proceed with the derivation of new theorems and the solution of new problems by applying the rules of classical logic to this theory. Suppose also that he will be assured of obtaining the intuitionist counterparts of his classical proofs and solutions by using intertranslatability relations which are both fixed in advance for the two associated (classical and intuitionist) theories under consideration and applicable in a finite number of predetermined steps. This being so, we can hardly expect the mathematician who is working on some non-trivial problems to worry over the mechanizable transitions between intuitionist and classical theories.

To support this argument in favor of a narrowing gap between classical and intuitionist philosophies of mathematics which results from the inter-translatability relations, let us point out that such relations are by no means confined to the field of syntax or logical morphology. It has been shown by Prof. Kleene [1]) that Tarski's 'semantic' conception of mathematical truth (and, by the same token, the over-all semantic approach to mathematical philosophy as distinct from the syntactical viewpoint) can be transferred to intuitionist mathematics without any serious difficulties. This result provides another contribution to the *rapprochement* between logicism and intuitionism which has been effected by the findings concerning the inter-translatability relations.

A more powerful argument in support of a kind of logicism which I have already referred to as 'pluralistic' and which seems to reduce to zero the gap between intuitionism and logicism, can be derived from the specific nature of what a mathematician does in his professional capacity, i.e., from the fact of his being primarily an architect of proofs. Once we realize the part played by the construction of proofs in the mathematician's work, we can hardly reject the position of pluralistic logicism provided the latter be construed as asserting the following: 'Any proof which is valid in some particular mathematical theory has a valid *replica* which clearly belongs to some pure logic and can therefore be established by a logician within his own field.' What I am driving at is simply the so-called 'deduction-theorem' of mathematical logic which has been independently discovered by Tarski [2]) and Herbrand [3]) for conventional systems of logic and has subsequently been extended to other logical systems. In substance, this result states that whenever a mathematical theorem T can be proven on the basis of a finite set of axioms A_i(i = 1, 2, ... n) then the conditional $\prod_{i}^{1,n} A_i \supset T$ (where $\prod_{i}^{1,n} A_i$ signifies the logical conjunction of all the axioms A_i, i = 1, 2, ... n) is a special case (a 'substitution-instance') of a conditional 'if C then D' which is both expressible and provable within pure logic. If the set of axioms A_i of the mathematical theory under consideration is infinite (either general

[1]) *Op. cit.*, p. 499ff.

[2]) A. Tarski, 'Über einige fundamentale Begriffe der Metamathematik,' *Comptes Rendus des séances de la Société des Sciences et des Lettres de Varsovie*, Vol. XXIII (1930).

[3]) J. Herbrand, 'Recherches sur la théorie de la démonstration,' *Travaux de la Société des Sciences et des Lettres de Varsovie*, (1928).

recursive or recursively enumerable) then the proof of any particular theorem T will always involve only a finite sub-set of the class of axioms under consideration and the Herbrand-Tarski Deduction Theorem will still hold. Thus, no matter what the cardinality of the set of axioms of a given theory happens to be, he who would know all the laws of logic, the form of which is conditional, would by the same token be in a position to prove any mathematical theorem which is derivable from a specified set of axioms. To effect the transition from logic to mathematics, the logician would merely have to show that the mathematical theorem T is a substitution-instance of a conditional expressible and derivable within logic proper. In other words, the only operation required to provide a transition between the class of logically provable laws and the class of theorems provable in any preassigned mathematical theory consists in an application of the '*dictum de omni et nullo*.'

To realize the relevance of the Herbrand-Tarski result to the intrinsic affinity of intuitionism and logicism, we have to take into account that the Deduction Theorem shows that logic *suffices* to prove whatever the mathematician has proved or could possibly prove, within any axiomatic or axiomatizable mathematical theory. This means that no recourse to an extra-logical 'intuition' is necessary at any stage of a mathematical proof. The adequacy of pure logic for establishing the relevant mathematical results is therefore guaranteed by the Deduction Theorem whenever this theorem applies. This theorem holds, however, not only in classical mathematics associated with classical logic as developed by Russell and Whitehead in their *Principia Mathematica* and subsequently improved upon and simplified considerably by several investigators like Gödel, Tarski, Quine and Bernays, but the same result applies also to intuitionist mathematics associated with a system of intuitionist logic as developed, e.g., by A. Heyting. Thus, if a mathematical theorem is intuitionistically derivable from a set of intuitionistically acceptable postulates by applying the axioms, definitions, and rules of inference of intuitionistic logic, then the entailment of this theorem by the conjunction of these postulates is a substitution-instance of a law of intuitionist logic. Hence, both classical *and* intuitionist mathematics have their exact *replicas* in their respective logical systems. The only difference between these two kinds of mathematics would seem to consist in the circumstance that two different logics are used in these kinds of mathematics.

Finally, a fourth argument, pointing to an intrinsic difficulty facing the intuitionist philosophy of mathematics to the extent to which the latter disassociates itself from logicism, may be indicated. The point is that the fundamental intuitionist concept, viz. that of constructibility, which the leading intuitionists consider as central in their position and even as sufficient to characterize this position completely proves to be both vague and elusive when submitted to a closer analysis. Professor A. Heyting states explicitly that 'the exigency that only constructible objects can be mentioned suffices to explain the particularities of intuitionistic mathematics.'[1]) On the other hand, he refuses to propose an explicit definition of constructibility and even asserts that it is impossible to define this intuitionistically crucial concept. In support of the view condemning any attempt to define the concept of constructibility, he invokes an argument developed by Miss R. Péter in the same volume.[2]) It seems therefore significant that Miss Péter derives, from the same considerations, a conclusion diametrically opposed to Heyting's conclusion. Says she: 'Es hat den Anschein, daß sich der Konstruktivitätsbegriff überhaupt nicht zirkelfrei erfassen läßt.'[3])

As for myself, I cannot help feeling that the conceptual difficulties inherent in the notion of constructibility are essentially similar to those characteristic of the concept of 'empirical verifiability'; the latter has been and is still being considered central in the philosophy of empirical sciences by virtually every major school of philosophy of science like pragmatism, logical positivism, operationism, etc. This may be the reason why Henri Poincaré has sometimes characterized his mathematical intuitionism as 'pragmatism.' The similarity between mathematical constructibility and empirical verifiability seems very significant and I will deal with it in a forthcoming paper of mine. The only point I should like to make at this juncture in connection with the aforementioned similarity is that the ambiguity, vagueness, and elusiveness of the concept of empirical verifiability as used in the philosophy of empirical sciences is notorious in spite of sustained attempts by many outstanding investigators to

[1]) A. Heyting, 'Some Remarks on Intuitionism,' *Constructivity in Mathematics*, (1959), pp. 69–71.

[2]) R. Péter, 'Rekursivität und Konstruktivität, *Constructivity in Mathematics*,' (1959), pp. 226–233.

[3]) R. Péter, *ibid.*, p. 233.

clarify this concept. These repeated failures have even induced some authors to condemn any new attempt as hopeless and to consider the borderline between empirically verifiable and unverifiable assumptions as incapable of being drawn in any significant and non-arbitrary way.

Such a pessimistic conclusion seems to me unwarranted [1]) although I do appreciate the fact that the conclusion may be helpful in pointing out the difficulties involved in clarifying the meaning of 'empirical verifiability.' The necessity and difficulty of a similar clarification in the case of mathematical constructibility can hardly be doubted by anybody who is familiar with the relevant literature. Thus, in discussions concerning empirical verifiability, the distinction between actual possibility of verification and theoretical possibility of verification (actual verifiability vs. verifiability in principle) [2]) continues to play an essential role and nobody doubts the importance of, and the need for, making such a distinction. However, 'mathematical' (or intuitionist) constructibility obviously involves either the actual or the theoretical *possibility* of constructing a mathematical entity rather than the *fact* that such a construction has been effected. Yet as late as 1959 Prof. Heyting wrote: 'By a constructive theory I mean a theory in which an object is only considered as existing *after* it has been constructed. In other words, in a constructive theory there can be no mentioning of other than *constructible* objects.' [3]) The two words which I have italicized show clearly that even the distinction between factually constructed mathematical objects and objects whose construction is merely possible is disregarded by the most lucid and effective exponent of intuitionism; the subtler and more important distinction between objects whose construction is actually possible and the more comprehensive class of mathematical objects whose construction is at least possible in principle is apparently not even considered.

No wonder that the vagueness and ambiguity of the term 'constructibility' has induced several investigators to explain this concept in terms of less objectionable concepts. In the main, these attempts center around the concept of general recursiveness as a satisfactory substitute for 'constructibility.' This substitution would have the advantage of being

[1]) H. Mehlberg, *The Reach of Science*, (1958), p. 221ff.
[2]) H. Mehlberg, *Ibid.*, p. 316ff.
[3]) A. Heyting, *ibid.*

supported by the volume of evidence known to be favorable to 'Church's thesis'[1]) and of replacing an undefined and obscure term by another term which is susceptible to several rigorous and demonstrably equivalent definitions. Unfortunately all such attempts at formulating intuitionism in terms of recursiveness instead of in terms of constructibility have been far from successful. Prof. Heyting, who speaks about intuitionism with an authority both unquestionable and second to none, has serious doubts as to the adequacy of approaching the problem of constructibility in terms of general recursiveness. His presentation of the intuitionist position at the Stanford Congress seems to support his doubts very strongly. Several difficulties which a definition of constructibility in terms of general recursiveness would be confronted with if constructibility were interpreted as effective calculability (which is often the case) were pointed out by Professor L. Kalmař.[2]) Finally, the most telling objection to Heyting's aforementioned identification of intuitionism with the constructibility of every entity in the mathematical universe of discourse was formulated by Professor A. Grzegorczyk. The following passage quoted from his 1957 essay will give an idea of what his objection comes to: [3]) 'The constructive analysis... is not part of the intuitionistic mathematics. All methods of proof are allowed. The constructive tendency consists only in the narrowing of the field of mathematical entities considered in classical analysis. In the constructive calculus, we consider only those objects for which we have constructive methods of approximation, this means methods given by means of constructive functions of integers. The most constructive are the general recursive ones.'

Both the analysis of difficulties inherent in the concept of constructibility, which is admittedly crucial in the intuitionist philosophy of mathematics, and the survey of unsuccessful attempts at overcoming these difficulties support the conclusion that the position of an immutable and inflexible intuitionism seems to have become virtually untenable in the face of developments which have taken place in the last three decades. It is worth

[1]) Cf. S. C. Kleene, *Introduction to Metamathematics*, (1952), p. 300ff.

[2]) L. Kalmàr, 'An Argument against the Plausibility of Church's Thesis,' *Constructivity in Mathematics*, (1959), pp. 72–80. However, serious doubts concerning Kalmàr's position have been expressed by R. M. Robinson. Cf. *Journal of Symbolic Logic*, Vol. XXIII pp. 362–363.

[3]) A. Grzegorczyk, 'Some Approaches to Constructive Analysis,' *Constructivity in Mathematics*, (1959), p. 43.

pointing out that intuitionism has nevertheless maintained its initial stimulating power and that the serious objections raised by the intuitionist requirement of constructibility would be overcome automatically if intuitionism were construed as a version of pluralistic logicism.[1]

V. CONCLUSION: PLURALISTIC LOGICISM AND CONTEMPORARY PHILOSOPHY OF MATHEMATICS

There is no need to elaborate on the position outlined in section III that the common basis provided by the Deduction Theorem for logicism, intuitionism, and many other trends in contemporary philosophy of mathematics points towards a pluralistic *logicism* rather than towards *intuitionism*. To bring out the survival of logicist tendencies in such a common basis (and the possibility of thereby freeing intuitionism from the drawbacks which have recently become apparent) I would like to conclude the discussion by making the following points:

1) The pluralist logicism I am proposing resembles, in some respects, R. Carnap's 'Principle of Tolerance.' However, the former differs essentially from the latter for the following reasons: Carnap's Principle states, in substance, that any language involving among its syntactical features (or, according to a later version, among its syntactical *and* semantical features) any logic whatsoever may be chosen for the formulation of any relevant group of mathematical theories, provided that this language contains the facilities required by such a formulation. In contrast to this identification of the logic associated with a given language L with a group of syntactical features of L, (or of syntactico-semantical features of L) pluralistic logicism does not involve any commitment to such a linguistic view of the logic associated with a given language. Nor does pluralistic logicism imply the possibility of identifying a logical system associated with a language used by the members of a given society with certain intrinsic features of this language on the understanding that the way this society makes use of this language can be disregarded completely in so far as the logic associated with the language is concerned. This intralinguistic approach upon which pluralistic logicism confers no monopoly seems to be implicit in the Principle of Tolerance. Needless to say, while pluralistic logicism is not committed to the aforementioned

[1]) Cf. section V of this paper.

nominalist [1]) approach to logic, it does not imply a Platonistic or, at least, a realist attitude towards logical ontology either, although it is compatible with such an attitude. I, for one, would certainly favor a realist view of the ontological subject matter of logic since, to my mind, the difficulties it raises are substantially easier to overcome than are the difficulties inherent in a nominalist view of logical subject matter.

2) One of the main difficulties confronting the philosophy of mathematics which centers around the claim that the mathematical vocabulary is part of the logical vocabulary and abandons Russell's second claim that mathematics is literally part of logic [2]) has been seen to consist in the fact that the issue between logicism and intuitionism would then hinge on the single lexicographic question as to whether or not the concept of set belongs in logic proper. From the viewpoint of pluralist logicism it hardly matters whether the term 'set' is classified under the heading 'logical.' For even if the mathematical character of this term were granted and both claims of Russell's original, radical logicism were dropped, it would still remain a fact that the knowledge of all logical laws would automatically yield the proofs of all valid theorems in all axiomatized or axiomatizable mathematical theories. Consequently, the main objection raised by 'moderate logicism' does not affect pluralist logicism.

3) The feature of mathematical knowledge which pluralistic logicism emphasizes is the fact that for every mathematical theory there exists *some* logic capable of providing the necessary tools for the derivation of all the relevant theorems of this theory without any recourse to extra-logical 'intuition.' The circumstance that for every mathematical theory some logic can supply all the tools required by this theory is our main reason for regarding this conception of mathematical knowledge as a version of logicism. We have qualified this logicism as 'pluralistic' because

1) A nominalist view of logical subject matter is construed here in a 'radical' sense and denies any extra-linguistic status to those entities with which logic is concerned. The more specific question of whether the extra-linguistic subject matter of logic – if granted – consists of individuals only, or includes abstract entities also, is not involved in the thesis of pluralistic logicism. Obviously, only the first member of this alternative to radical nominalism is usually referred to as 'nominalism', and is contrasted with Platonism. This conventional, moderate version of nominalism, compatible with pluralistic logicism, was skilfully developed and defended by Professor L. Henkin at the Stanford meeting.

2) In other words, the philosophy of 'moderate logicism,' according to the terminology proposed in section II.

it does not confer a monopoly upon, or discriminate against, any particular logic. Thus, from the viewpoint of pluralistic logicism the system of 'classical', two-valued and extensional logic developed by Russell and Whitehead in their monumental work is no more legitimate than the non-classical logics which Von Neumann and Von Weiszäcker have respectively constructed in order to adjust classical logic to the quantum-theoretical situation. A comparable legitimacy must be granted, from the viewpoint of pluralistic logicism, to the bunch of many-valued logics discovered by J. Lukasiewicz and E. Post (independently from one another), to the families of modal logics, to C. I. Lewis' 'strict' logics, to H. B. Curry's variable-free (combinatory) logics, to A. Tarski's logics involving infinitely long propositions and/or infinitely long proofs (on the understanding that the respective length of admissible propositions and/or proofs is specifiable by a class of transfinite, ordinal numbers), to logics which possess finite or transfinite hierarchies of logical types, of either the 'ramified' or the 'simplified' or the 'cumulative' variety – in brief to the snowballing assembly of various logical systems which are now being manufactured on the assembly line. The only condition to be satisfied by anyone of these various logical systems is adequacy to the relevant cognitive objective. No 'ideal' status is claimed for any of these logical systems or for the language with which a particular logical system is associated. It goes without saying that this egalitarian rule applies also to the logic and the language of *Principia Mathematica*.

What about the relative roles of logicism, intuitionism, and formalism, in the present situation of the philosophy of mathematics? It is obvious that from the viewpoint of pluralistic logicism, mathematical intuitionism is still to be considered as an expression of the most stimulating and stable attitude in foundational research. It is inevitable, however, from this viewpoint, that the intuitionist position be readjusted so as to overcome the grave difficulties it has been confronted with between 1930 and 1960. In my view, this readjustment cannot fail to bridge the gap between the classical formulations of intuitionism and logicism since both would then be classifiable under the heading of pluralistic logicism.

As for logicism, we have seen that in the present situation, 'radical' logicism, as represented by B. Russell, rests on two clearly untenable claims. On the other hand, the 'moderate' logicism suggested by A. Church does not seem to be moderate enough in order to take care of the difficulties inherent in the 'radical' variety. However, one more step towards liberalizing the logicist position would reach the viewpoint of pluralistic logicism, maintain the common core of the aforementioned varieties of logicism and help overcome the objections to which these varieties are open.

The case of formalism is more difficult to evaluate at present. It is undeniable that both fundamental formalist objectives (viz. the discovery of decision-procedures for all the theories of classical mathematics and the construction of a finitary, metamathematical consistency-proof for these theories) have neither been attained nor could possibly have been attained, in view of Gödel's and Church's results. The failure to find the aforementioned decision-procedures seems to be definitive, and on the whole, beneficial. However, during the last decades, several consistency-proofs for crucial classical theories have been constructed and the Gödel predicament has been taken care of by sacrificing part of the finitary requirements. The most important consistency-proofs are exemplified by the relative, intuitionist consistency-proof (which provides a partial and approximate solution to Hilbert's first foundational problem) and by Gentzen's absolute consistency-proof for number-theory. On closer analysis, the latter seems to be unsatisfactory.

Yet neither the partial successes nor the partial failures of Hilbert's original program determine the actual role of non-finitary metamathematics in the present situation. The essential point is that, by now, non-finitary metamathematics is both scientifically and philosophically one of the most significant aspects of contemporary investigations into the nature of mathematics. Granted, the specifically formalist features of metamathematics have virtually disappeared in the process and there is little to prevent metamathematical research carried out by investigators with a formalist background from being identified with metamathematical investigations creditable to logicists, intuitionists, or constructivists. It seems very doubtful to me whether the common denominator of all these avenues of approach to metamathematics contains anything beyond the basic thesis of pluralistic logicism. In any event, this co-operative feature of contemporary philosophy of mathematics and foundational research is perhaps its most remarkable feature.

Needless to say, no complete picture of the present situation in mathematical philosophy was the object of this study. We have selected three particular features of this situation, (logicism, intuitionism, formalism) as the sole object of our examination, for reasons which should have become apparent in the course of the discussion. But new, promising approaches to mathematical philosophy are noticeable in recent literature: there is no doubt that a more comprehensive presentation of this particularly intriguing branch of the philosophy of science would have to take into account these recent approaches. They include for instance, the philosophical views of mathematics expounded in Wittgenstein's posthumous work,[1] the mathematical philosophy underlying the powerful French Bourbaki group,[2] the numerous versions of constructivism (inside and outside Poland), Goodstein's 'constructive formalism',[3] etc. Professor Saunder MacLane's attempt so to extend the import of various applications of mathematics as to make these applications of mathematics an essential factor in understanding its philosophy, deserves a special interest since a problem is often more likely to be solved if its range has been adequately widened.[4]

A glimpse of these novel attempts suggests that pluralistic logicism can provide a common basis for all of them. But any judgement concerning the adequacy of pluralistic logicism to this extended scope of mathematical philosophy would be premature at present. I have no doubt that no matter which judgment will be supported by coming developments, all these recent avenues of approach to mathematical philosophy are most welcome, in view of the baffling complexity of the relevant problems. The three decades which have elapsed between 1930 and 1960 have taught us at least one lesson: There are more mathematical puzzles between heaven and earth than logicists, intuitionists, and formalists have ever dreamt of.

The University of Chicago

[1] L. Wittgenstein, *Remarks on the Foundations of Mathematics*, (1956).

[2] Bourbaki, 'Foundations of mathematics for the working mathematician,' *The Journal of Symbolic Logic*, (1949), Vol. XIV.

[3] R. L. Goodstein, *Constructive Formalism*, (1951).

[4] S. MacLane, 'A Different Approach to the Foundations of Mathematics,' address delivered to the departmental philosophy-seminar of The University of Chicago.

KARL MENGER

A COUNTERPART OF OCCAM'S RAZOR IN PURE AND APPLIED MATHEMATICS; ONTOLOGICAL USES

INTRODUCTION

One of the most efficient methodological tools, Occam's celebrated razor, is the maxim that *it is vain to do with more what can be done with fewer*. This principle is primarily ontological: *Entities must not be multiplied beyond necessity*. But the razor may also be construed as a semantic maxim opposing the use of synonyms. Occam's principle is often called the *Law of Parsimony*.

A thorough examination of science and mathematics reveals that they include only few superfluous entities but that, on the contrary, many ideas have merged and lost their identity in misconceptions. Perhaps the most important example are variables – a term applied to totally discrepant ideas [1]) to which mathematicians and scientists, overparsimoniously, have denied individual names, thereby making variables in certain ways unmanageable. Thus what is needed is a counterpart to the Law of Parsimony – so to speak, a *Law against Miserliness* – stipulating that *entities must not be reduced to the point of inadequacy* and, more generally, that *it is vain to try to do with fewer what requires more*.

This law, too, has what might be loosely called ontological as well as semantic applications. It condemns gaps in ontology just as Occam's law repudiates redundancies; and it may be construed as a maxim denouncing equivocations just as Occam's law opposes synonyms. The present paper is devoted to the former uses.[2]

1) Cf. the author's papers 'The idea of variable and function,' *Proc. Natl. Acad. Sci. U.S.A.*, 39 (1953), pp. 956–961; 'On variables in mathematics and in natural science' *Br. Journ. Phil. Sci.*, 5 (1954), pp. 134–142; 'Mensuration and other mathematical connections of observable material,' in *Measurement: Definitions and Theories*, ed. Churchman and Ratoosh, John Wiley, New York, 1959, pp. 97–128; 'An axiomatic Theory of functions and fluents,' in *The Axiomatic Method*, ed. Henkin, Suppes, and Tarski. North-Holland Publ. Co., Amsterdam, 1959, pp. 454–473, and the author's book *Calculus. A modern Approach*, Ginn, Boston 1955; mimeo-editions Chicago 1952 and 1953.

2) A forthcoming paper is devoted to semantic uses of the principle.

Since current pure and applied mathematics include little that might be shaved off, but abound in medleys of diverse constituents that pass for homogeneous, what contemporary mathematics calls for is not so much a razor as a device that analyzes mixtures and isolates their various components. Mathematico-scientific methodology is in need – in fact, in urgent need – of a *separator* or a *prism* resolving conceptual mixtures into the spectra of their meanings.

The resolution of equivocal uses of terms into various nuances of meaning is, of course, no novelty. That method has led to the clarification and the improvement of technical expositions in many a branch of knowledge. But the material that will be resolved by the methodological prism on the following pages is absolutely fundamental and not only of technical nature; and it includes conceptual combinations of such incongruity that the analysis will resolve more than mere nuances. The separator will sort altogether diverse *categories*. Some of these categories have been completely ignored in the traditional literature, whence the aforementioned gaps. Therefore, the analysis of the mathematico-scientific material suggests the need for certain extensions of the classical conceptual frame. Such extensions will be outlined in a spirit of parsimony but without miserliness.

1. THE NOTION OF VARIABLES

Logicians mean by a variable a symbol that stands for any element of a certain class (which I will call the *scope* of the variable) or a symbol that may be replaced with the designation of any element of that scope. The first articulate definition of certain variables in this sense (viz., of *number variables*, i.e., variables whose scopes are classes of numbers) seems to have been given by Weierstrass [1]).

Variables in the sense of the logicians are hardly ever used outside of pure and applied mathematics, while they are ubiquitous in those fields. This state of affairs has a sound reason. The reason is that most nonmathematical assertions can be well understood without variables, whereas complicated mathematical assertions in purely verbal form are almost incomprehensible. This reason is, however, purely *practical*. The *basic* situation is as follows:

[1]) Cf. Pincherle, *Giornale di mat.*, 18 (1880), p. 243.

(a) Variables may be used in nonmathematical (as well as mathematical) assertions. Consider, for example, the proposition:

The (paternal) grandfather of the father of any person is the father of the grandfather of that person.

This assertion could be expressed in terms of the abbreviations g and f, and by the use of a variable x – one might say, by the use of a *person variable* – as follows:

$$g(f(x)) = f(g(x)) \text{ for any person } x.$$

(b) Variables may be dispensed with in mathematical (as well as in nonmathematical) assertions. Not a single variable occurs in Alkhowarizmi's *Al Jebr*, the famous book whose title is at the root of the word algebra. Nor are purely verbal assertions by any means inferior to formulations in terms of variables. Compare, e.g., the assertion:

The cube of the square of any number equals the square of the cube of that number with the formulations

$$cu(sq(x)) = sq(cu(x)) \text{ or } (x^2)^3 = (x^3)^2 \text{ for any number } x.$$

The latter are shorter but not otherwise superior.

But if there is only a superficial difference, with regard to variables in the sense of logic, between mathematical and nonmathematical assertions, why has Russell [1]) called variables 'perhaps the most distinctly mathematical of all notions'? What probably prompted Russell's dictum is the fact that pure and applied mathematics do not confine themselves to the use of variables in the sense of logic but abound in totally discrepant uses of the term variable. In the traditional literature, the term is used as a reference to x, y, s, and t in each of the following examples – six variations of the idea of something being the square of something else:

(1) *Comprehensive assertions about numbers*, such as $x^2/x = x$ for any number $x \neq 0$; $(x^2)^2 = x^4$ for any number x; or, for any two numbers x and y, if $y = x^2$, then $y^2 = x^4$; (*number* here and in the sequel means *real number*)

and *descriptions of classes of pairs of numbers*; e.g., the class of all pairs (x, x^2) for any number x; or of all pairs (x, y) such that $y = x^2$.

(2) *Assertions about functions*; e.g., that the function x^2 (or, as some mathematicians say, the function $y = x^2$) is nonconstant.

[1]) *Principles of Mathematics*, Cambridge, 1903, p. 5 and p. 89.

(3) *Physical laws*, such as Galileo's Law $s = \frac{1}{2} gt^2$ or, in proper units, $s = t^2$.
(4) *Discussions in the analytic geometry of curves* such as the parabola $y = x^2$.
(5) *Laws of analysis;* e.g., if $y = x^2$, then $dy/dx = 2x$.
(6) *Algebraic formulae about polynomials* such as $x^2 = x \cdot x$ or $x^2/x = x$.
These examples of which only (1) will appear to include number variables in the sense of Weierstrass, also explain why Russell [1]) has called variable 'certainly one of the most difficult [notions] to understand.' A full understanding presupposes that the mélange illustrated in Examples (1) to (6) be resolved in the prism.

2. THE RESOLUTION OF VARIABLES IN THE PRISM

The prism separates each two of the six listed types of 'variables.'
I. The Separation of (2) from (1). Each of the three assertions in (1) synthesizes countless specific assertions such as $3^2/3 = 3$, $e^2/e = e$ (where e designates the base of the natural logarithms); $(0^2)^2 = 0^4, (3^2)^2 = 3^4$, $(e^2)^2 = e^4$; and if $0 = 0^2$, then $0^2 = 0^4$; if $9 = 3^2$, then $9^2 = 4^4$. Similarly, the definitions in (1) combine countless specific pairs of numbers such as $(0,0^2)$, $(3,3^2)$, (e,e^2). The specific assertions or pairs of numbers are obtained by replacing the variables with designations of elements of their scopes. On the other hand, the statement (2) that the function x^2 is nonconstant is not the synthesis of specific assertions such as
the function e^2 is nonconstant or the function $9 = 3^2$ is nonconstant.
These assertions are not even true. What (2) does assert is the property of one specific entity, called 'the function x^2' and obtained by pairing the number x^2 to any number x, for example, 3^2 to 3. The most concise definition of that entity is as the *result* of the said pairing; in other words, as the class of all ordered pairs (x, x^2) for any number x. (This, incidentally, is the example given in (1) of a definition of a class in terms of a number variable.)
II. The Separation of (3) and (4) from (1). In contrast to the number variables in (1), Galileo's s and t and Descartes' coordinates designate specific entities. One symptom of this profound contrast is the fact that,

[1]) *Principles of Mathematics*, Cambridge, 1903, p. 5 and p. 89.

without changing the meaning of (1), one may replace the number variables with other letters as in

$$(c^2)^2 = c^4 \text{ for any number } c,$$

or even interchange them as in
(1′) the class of all pairs (y, x) such that $x = y^2$.
In Galileo's law, on the other hand, s and t must not be interchanged. It is not true that $t = s^2$. Nor is the parabola $x = y^2$ identical with the parabola $y = x^2$.
But, one might ask, does not Galileo's law imply that $t = s^2$ *provided one lets* 's' *denote the time, and* 't' *the distance?* Indeed it does. But, similarly, $9 = 3^2$ implies that $3 = 9^2$ *provided one lets* '3' *denote the number of the Muses and* '9' *the number of the Graces.* Yet no one will, on this ground, call 3 and 9 number variables. The point is that the definitions (1) and (1′) are tantamount *without anything being renamed,* since, in contrast to the letters in Galileo's law (3), the number variables in (1) do not name or designate anything specific, to begin with.
That the letters in (3) cannot be interpreted as number variables (whose scopes consist of observed distance and time values) appears also from an attempt to use them as such, as in
(3′) $s = t^2$ for any two values s and t of distance and time.
Obviously (3′) is false [1] except in insignificant cases of miniature scopes, e.g., if the scope of t consists only of two numbers, 3 and -3, and that of s includes only 9.
III. The Separation of (5) from (1). The prism separates the concepts in (5) from the number variables in (1). Indeed, if the letters in (5) are replaced with '3' and '9' ,the result is the sentence

$$\text{if } 9 = 3^2, \text{ then } d9/d3 = 2 \cdot 3.$$

[1] Ever since the 18th century, mathematicians have tried to reconcile the interpretation of s and t as variables with the fact that, for all significant scopes of the variables, (3') is false. Their attempts have centered on the idea of so-called *dependent variables.* In (3'), only t is regarded as replaceable with any value of the time (and called an *independent* variable) while s is said to be *dependent* since, after a value of t has been chosen, the value of s is determined. But a variable (i.e., a symbol that may be replaced with the designation of any element of its scope) which is dependent (i.e., cannot be replaced with the designation of any element of its scope) is a plain contradiction in terms.

But this implication is false since the antecedent is true, whereas the consequent, being nonsensical, is not.

IV. The Separation of (5) from (2), (3), and (4). While the letters in assertion (5) are not number variables, they are variables of some kind. The law (5) synthesizes specific assertions that can be obtained by re-. placing the letters, e.g., with designations of Galileo's s and t or Descartes' x and y, the result being

$$\text{if } s = t^2, \text{ then } \mathrm{d}s/\mathrm{d}t = 2t \text{ and if } y = x^2, \text{ then } \mathrm{d}y/\mathrm{d}x = 2x.$$

Thus the concepts in (5) have been separated from Galileo's and Descartes. The latter are, so to speak, special cases of the concepts (5). Being variables of some kind, the concepts (5) are also separated from the specific function in (2).

V. The Separation of (6) from (1) and (5). In contrast to the variables in (1) and (5), the polynomials x and x^2 in the algebraic assertions (6) are not meant to be replaced with designations of either numbers or entities connected in Galileo's laws.

VI. The Separation of (2) from (6). The function in (2) is separated from the polynomials in (6), since the latter are not classes of pairs of numbers. Dividing the polynomial x^2 by the polynomial x yields the polynomial x, just as $3^2/3 = 3$ and $e^2/e = e$. The function x^2, on the other hand, divided by the function x yields only a restriction of the function x, namely, the class of all pairs (x, x) for any number $x \neq 0$, and not the full function x, since $x^2/x = x$ only for $x \neq 0$.

VII. The Separation of (3) from (2) and (6). Galileo's s and t refer to the physical universe, whereas the assertions (1), (2), and (6) belong to pure mathematics.

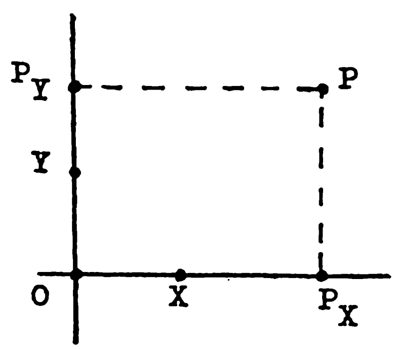
Only (4) has not yet been separated from all the other examples. This separation, too, will be achieved, but only after the socalled variables in (4) themselves have been treated with the prism and resolved in two discrepant types. The separation of all the other types of socalled variables has been completed.

In the medley of entities indiscriminately called variables, the prism thus separates mathematical and nonmathematical entities; it separates classes and elements; it separates designations of entities and symbols not designating anything specific; and it separates particular and general ideas. It would be difficult to find among objects of intellectual studies another group that is both as prominent and as ill-assorted as the lot of so-called variables.

3. FILLING TWO GAPS IN THE TRADITIONAL, NEO-PYTHAGOREAN ONTOLOGY

The prism has revealed that the idea of variables in the sense of something that is common to the examples (1) to (6) has been reduced to the point of inadequacy. What this means is much more, and much more serious, than the mere terminological shortcoming of equivocations – of references to unlike categories by the same term. It appears that several categories that are absolutely basic for mathematical science are not being articulately defined and studied at all – a defect that in the Introduction has been somewhat loosely described as gaps in the traditional ontology. These gaps must be filled by definitions of those suppressed categories rather than covered up by equivocal uses of the term variable. The first step in this direction is the ascertainment of the exact nature of Descartes' coordinates x and y and of Galileo's s and t.

In 1637, Descartes introduced coordinates in a physical plane, such as a sheet of paper or a blackboard, where points are pencil marks or chalk dots. Choose a frame of reference consisting of three noncollinear points, say, O, X, and Y. For any point P in the plane (see the figure), let P_X



denote the projection of P on the X-axis (i.e., the line joining O and X). Measure the distance from O to P_X using the segment from O to X as unit; and let OP_X be the number thus obtained. Define P_Y and the number OP_Y analogously. Then, in a crucial passage of *La Geometrie*, Descartes says verbatim [1]: *pourceque* OP_X *et* OP_Y *sont deux quantités indeterminées & inconnuës, ie les nomme l'vne x et l'autre y.* Descartes here takes it for granted that special letters are reserved for special purposes – an idea that goes back to the great French algebraist Vieta [2], who used capitalized vowels as unknowns. In algebra, Descartes replaced Vieta's vowels by the last letters of the lower case alphabet and introduced, accordingly, also in geometry 'x' and 'y' for his so-called *quantités indeterminées et inconnuës.*

What with Descartes' epochal contributions to mathematics and what with his celebrated *clarté,* few tasks are as difficult as propounding that one basic aspect of mathematics and science has been obfuscated by his legacy. But it actually is the quoted passage in *La Geometrie* to which some defects of the traditional mathematico-scientific ontology can be traced.

To begin with, Descartes' x and y are not really 'indeterminate and unknown quantities.' As actual examples of unknowns, consider the letters in the equations

$$\text{find all numbers } x \text{ such that } 1 - x^2 = 0;$$

$$\text{find all numbers } x \text{ and } y \text{ such that } x + y = 10 \text{ and } x - y = 2.$$

These letters are tentative designations of numbers satisfying the conditions expressed in the formulae, and the word *find* stipulates that the reader produce numerals designating those numbers without reference to those conditions.

The situation in geometry is basically different. A physical plane with a frame of reference is given. Before a point P has been chosen in that plane, there are no unknown numbers x and y; and after P has been chosen, Descartes' procedure determines two numbers, called the abscissa of P and the ordinate of P, which I will denote by the symbols $x(\mathrm{P})$ and $y(\mathrm{P})$, respectively. For example, the numbers associated with the afore-mentioned points O and X are:

$$x(\mathrm{O}) = 0, y(\mathrm{O}) = 0 \text{ and } x(\mathrm{X}) = 1, y(\mathrm{X}) = 0.$$

[1] On p. 321 of the 1637 edition. I am retaining Descartes' spelling, but will use different letters for the points.

[2] *In artem analyticam isagoge*, Paris, 1591.

Even someone who knows Descartes' procedure of determining the numbers x(P) and y(P) for any point P still may ask 'What is x?' or 'What is the abscissa?' His situation is comparable to that of someone who knows the procedure of determining the number $\log x$ for any positive number x but who still may ask 'What is log?' or 'What is the logarithmic function?' He may be told that the logarithmic function is the result of pairing, to each number x, the said number $\log x$; in other words, that it is the class of all pairs $(x, \log x)$ for any number $x > 0$. Similarly, the question concerning the abscissa may be answered by saying that it is the result of pairing, à la Descartes, the number x(P) to any point P; in other words, it is the class of all pairs (P, x(P)) for any point P of the plane. An analogous definition is possible for the ordinate y.

Galileo's t likewise is the result of pairing a number t(T) with each act of reading a timer that was started when the object under observation was dropped; and s is the class of all pairs (S, s(S)) for any act S of reading the mark opposite to the falling object on a vertical rod [1]).

As a collective reference to Galileo's t and s, Descartes' x and y, and the like, I have revived a term coined by Newton and used by him in reference to the said objects studied in physics and geometry, namely, the term *fluents*. Newton, the supreme virtuoso in operating with fluents, sketched his general idea of their nature in the sentence: *fluentes vocabo quantitates has, quas considero tamquam gradatim et indefinité crescentes* [2]) but he did not define specific fluents such as s or t at all. Two centuries later, the term fluent was all but forgotten, even in the British literature. Of

[1]) I have stressed the difference between these concepts and number variables in the Preface to the 1952 edition of my book *Calculus. A Modern Approach* and have elaborated the theory in Chapter VII of the 1953 and 1955 editions as well as in the papers quoted in [1]) on p. 80. Carnap defined what he called *physical quantities* as associations of numbers with quadruples of numbers (thus as 4-place functions) in *The logical syntax of language*, London and New York 1937, pp. 149 sq. and as number variables in *Foundations of Logic and Mathematics*, Encycl. of Unified Science, vol. I, Chicago, 1939. I thus am glad to see that he expresses views more similar to those stressed loc. cit. [1]) on p. 80 in his *Einführung in die symbolische Logik*, Wien, 1954 and *Introduction to symbolic Logic*, New York 1958, especially p. 168 sq.

[2]) *Methodus fluxionum et serierum infinitarum*, 1737. Apparently influenced by Descartes, Newton continues: *hasque representabo per ultimas alphabeti litteras u, x, y, et z ut discerni possint ab aliis quantitatibus quae in equationibus considerantur tamquam cognitae et determinatae.*

course, abandoning Newton's term and replacing it with another word would have been, in itself, harmless. But 'fluent' was almost universally replaced with 'variable'[1]), and this particular terminological change created the danger of confusion. Such confusion has actually developed. Fluents such as time and gas pressure have been explicitly identified with *number variables* by some mathematicians [2]), and with *classes of numbers* by others [3]). These are procedures of the kind that has been denounced in the Introduction. They are attempts to do with fewer what requires more. Accordingly, those identifications leave a serious gap in the mathematico-scientific ontology.

Those attempts seem to indicate an implicit overestimation of the scientific role of numbers by themselves – something of a neo-Pythagorean philosophy. Mathematicians and scientists often take numbers for quantities. But a quantity is not a mere number. A *quantity* is a number paired to what I will call the *object* of the quantity – something extra-mathematical such as a physical point, a falling object, an observer's act of reading a clock, or even a class of observables [4]). The number paired to the object is called the *value* of the quantity.

For two quantities to be *equal* it is necessary (and of course sufficient) that, besides equal values, they have the same object. For instance, if Peter is the oldest man in his village and is 172 cm tall, then

(Peter, his height in cm) and (the oldest man in the village, 172)

1) One may compare Newton's quoted description of fluents with de la Vallée Poussin's description of a variable as une *quantité qui passe par une infinité des valeurs, distinctes ou non* in his *Cours d'Analyse* vol. 1. It should be observed that those descriptions were not elaborated either by rigorous explicit definitions or by postulates defining the concept implicitly.

2) Cf., e.g., Courant, *Differential and Integral Calculus*, vol. 1, p. 16.

3) Cf. Mc Kinsey, Sugar, and Suppes, 'Axiomatic foundations of classical particle mechanics', *Journ. Rat. Mech. Anal.*, 2. 1953, pp. 253–272, and Artin, *Calculus and Analytic Geometry*, 1957, p. 70.

4) What laws, if any, govern the association of the value of a quantity with its object; whether the association must be relative to some frame of reference; and if so, what changes of the value correspond to changes of the frame – these questions will be discussed elsewhere. But whatever the answers to these questions may be, strong emphasis must be laid on the (usually neglected) *objects* of quantities, which remain permanent in the said changes. Weyl's elaborate definition of a quantity in *The Classical Groups*, Princeton 1939, pp. 16 sq. does not seem to include a clear reference to an object.

are equal quantities, whereas

(a certain tree, its height in cm),

even if that height is 172, is a different quantity. If two quantities are either equal or have different objects, then I will call them *consistent*. Thus two quantities are *inconsistent* if they have the same object and unequal values.

In this sense, each of Galileo's and Descartes' fluents t, s, x, and y is *a class of consistent quantities*. The class of all extramathematical objects of the quantities belonging to a fluent will be called the *domain* of the fluent. The class of all values of those quantities is referred to as the *range* of the fluent. While the range, being a class of numbers, is a concept of pure mathematics, the domain is not.

After the fluents t and s, connected with a moving particle, have been defined separately, they may be interrelated. In the case of a falling particle, they are connected by the law $s = t^2$. In discovering this law Galileo did not primarily pair the ranges of s and t or, more precisely, values of s to values of t. Pairing numbers to numbers is the activity of a mathematician defining a function and not that of a physicist discovering a law of nature. What Galileo actually paired were the domains of s and t or, more precisely, acts of mark reading to acts of clock reading. Specifically, he paired *simultaneous* acts. And he discovered that the corresponding values of the fluents satisfied the condition:

$$s(\mathrm{S}) = (t(\mathrm{T}))^2 \text{ for any two simultaneous acts S and T.}$$

This is what the law $s = t^2$ expresses.

The traditional, neo-Pythagorean ontology ignores objects of quantities, domains of fluents, and those pairings of the domains which are the prerequisite for the connection of fluents by functions.

The theory here outlined fills still another gap in the traditional ontology. It clarifies the role played in (5) by the letters x and y, which definitely are not number variables but appeared to be variables of some kind (see p. 85). Since they may be replaced with designations of Galileo's and Descartes' s, t, x, and y, each such replacement resulting in an assertion about specific fluents, the letters in (5) are *fluent variables*. The traditional theory fails to distinguish this important category from either number variables or specific fluents by indiscriminately applying the term variable to all three of them.

4. THE PRISM AS A SOLUTION OF PARADOXES IN ANALYTIC GEOMETRY

As far as analytic geometry is concerned, the prism not only analyzes ideas. It separates two altogether unlike theories: the analytic treatment of *physical planes*, where points are ink marks or chalk dots, and the analytic geometry of the *arithmetical plane*, whose points are ordered pairs of numbers.

There is no doubt as to which set-up Descartes had in mind by dint of his frequent references to compasses and linkages as well as pictorial illustrations in *La Geometrie* of those tools, which produce rows of ink marks or chalk dots but certainly not classes of pairs of numbers.

What Descartes had in mind when speaking of the parabola $y = x^2$ clearly was a class of points in a physical plane in which a frame of reference had been chosen, namely, the class of all points P such that $y(\mathrm{P}) = (x(\mathrm{P}))^2$. It was Russell in the course of arithmetizing geometry and Study [1]) in clarifying the foundations of nonpostulational geometry who developed the analytic geometry of the arithmetical plane as an algebraic theory of numerically defined objects.

There is – at least on the macroscopic level – a correspondence between the latter objects and the observable points in a physical plane. This correspondence is of course not coincidental but purposely achieved by formulating arithmetical definitions that reflect physical facts. The analogue of the points O, X, and Y constituting Descartes' frame of reference are the points (0, 0), (1, 0), and (0, 1) of the arithmetical plane, respectively (even though the latter points cannot be said to constitute a frame of reference of the arithmetical plane). The analogue of the parabola consisting of all points P such that $y(\mathrm{P}) = (x(\mathrm{P}))^2$ is the class of all arithmetical points (x, y) such that $y = x^2$. This class, therefore, is called a parabola in the arithmetical plane. In fact, the much-discussed *isomorphism* between the physical model and the arithmetical construct is so perfect that there seems to be hardly any point in separating the two.

[1]) Study, *Die realistische Weltansicht und die Lehre vom Raum*, Vieweg, Braunschweig, 1914.

Actually, however, dangers loom below the surface. The prism separates the formula $y = x^2$ in Descartes' description of a physical parabola from the apparently identical formula $y = x^2$ in the definition of the arithmetical parabola. The latter x and y are number variables, whereas Descartes' x and y are fluents. Not only have the two pairs of symbols altogether unlike meanings but they follow incompatible rules. For instance, the variables x and y may be interchanged; the fluents x and y must not be interchanged. A mere juxtaposition of the two results leads to what may be called a

PARADOX OF ANALYTIC GEOMETRY. *The two parabolae consisting of all points (x, y) such that $y = x^2$ and of all points (y, x) such that $x = y^2$ are identical. Yet the two parabolae $y = x^2$ and $x = y^2$ are different.*

This paradox cannot be resolved by pointing out that what actually differs from the class of all points (x, y) such that $y = x^2$ is the class of all points (x, y) – and not the class of all (y, x) – such that $x = y^2$; and by stipulating that, accordingly, the number variable x should be reserved for the first members, and y for the second members, of ordered pairs. For this stipulation would make it impossible to express important geometrical facts – to mention only the simplest example, that, for $x \neq y$ the points (x, y) and (y, x) are distinct.

The safest escape from the paradoxes of analytic geometry is by keeping apart what the prism has separated: number variables in the theory of the arithmetical plane, and fluents in Descartes' physical plane – concepts that are *toto caelo* different.

5. OTHER USES OF THE PRISM

The use of the methodological prism is by no means confined to the analysis of the notion of variables. By using this tool one can, for example, separate the idea of the derivative of a function from the idea of the rate of change of one fluent with regard to another fluent [1]) as well as the integral of a function from the cumulation of one fluent with regard to another fluent.

The prism, moreover, analyzes the mixture of ideas traditionally called

[1]) Menger, 'Rates of change and derivatives,' *Fundam. Mathem.*, 46, 1958, pp. 89–102.

random variables. It separates Kolmogorov's purely mathematical concept of a measurable function in whose *domain* a probability measure is given – a concept of basic importance for the theory of probability – from the scientific concept of a random fluent with an extramathematical domain and a frequency distribution or probability measure in its *range* – a concept that seems to be the main object of statistical studies[1]).

Illinois Institute of Technology, Chicago

[1]) Menger, 'Random variables and the general theory of variables,' *Proc. 3rd Berkeley Symposium Math. Stat. Probability*, vol. 2, Berkeley, 1956, pp. 215–229.

RICHARD MONTAGUE

TOWARDS A GENERAL THEORY OF COMPUTABILITY[1]

1. INTRODUCTION

There are nowadays two kinds of computers: *digital computers* and *analogue computers*. The principal distinction is this. The questions and answers of a digital computer are presented in symbolic form, perhaps printed on a tape or on cards. An analogue computer, on the other hand, relies on the dependence of one physical magnitude on another. For example, the question (or the input) may be an electrical voltage, and the answer (or output) another voltage related to the first in a manner determined by the structure of the machine.

There are two corresponding notions of computability. A function of k arguments on and to natural numbers may be called *digitally computable* if, loosely speaking, it is possible to construct a digital computer which, when given any k-tuple of natural numbers as arguments, will compute the value of the function for those arguments. *Analogue computability* may be characterized analogously; here it is natural to deal with functions on and to real numbers. It would be incorrect to suppose that the notions of digital and analogue computability are thus unambiguously reduced to the notion of a digital or analogue computer. Other terms occurring in the characterization above must also be analyzed, notably 'given' and 'will compute'.

A theory of digital computability was first constructed in Turing's paper 'On computable numbers, with an application to the Entscheidungsproblem,' *Proceedings of the London Mathematical Society*, ser. 2, vol. 42, (1936–7), pp. 230–265. Three observations may be made in connection with Turing's development.

[1] Some of the ideas of this paper were obtained in the summer of 1958 while the author was attending a conference supported by the Air Force Office of Scientific Research, under contract AF 49(638)-33. The preparation of the paper was supported in part by NSF grant G-13226, and in part by NSF grants G-6693 and G-14006.
I am grateful also to my student, Mr. David Kaplan, for valuable suggestions.

In the first place, Turing's enterprise may be regarded in part as an explication, in the sense of Carnap, of the term 'digitally computable function on and to natural numbers'.[1]) But this was an explication for which a definite criterion of adequacy did not exist. Thus, despite strong convictions, the question whether Turing's definition is an adequate explication of digital computability has received neither a decisive answer nor a definite sense.

In the second place, Turing made no attempt to analyze the general notion of a digital computer, but confined his attention to computers of a simple and uniform mode of operation, which can plausibly be supposed to suffice for the computation of all functions that are at all digitally computable. No broader class of computers was required; for Turing was concerned not with the notion of a computer but with that of computability.

Finally, let me indicate why Turing's analysis has practical importance. Some functions of natural numbers, for instance addition, are clearly, on intuitive grounds, computable. Other, more involved functions can be shown to be computable by the construction or programming of a suitable machine. Why, then, do we need an exact definition of computability? Would not a rough, intuitive notion suffice in those cases which are of practical interest? It is often important to know that a given function is *not* computable. Without an exact definition of computability, we can compile evidence to this effect, for instance, the failure of attempts to construct an appropriate machine; but such evidence can never be conclusive. Thus Turing's proof of the non-computability of certain functions (as well as of the existence of a universal digital computer) would have been impossible without a precise analysis of digital computability.

Now Turing's notion of computability applies directly only to functions on and to the set of natural numbers. Even its extension to functions defined on (and with values in) another denumerable set S cannot be accomplished in a completely unobjectionable way. One would be inclined to choose a one-to-one correspondence between S and the set of natural numbers, and to call a function f on S computable if the

[1]) I am guilty here of an historical inaccuracy. Turing was actually interested in the notion of a real number whose decimal expansion was effectively computable. But his work can be, and usually has been, reconstrued in the sense indicated in the text.

function of natural numbers induced by f under this correspondence is computable in Turing's sense. But the notion so obtained depends on what correspondence between S and the set of natural numbers is chosen; the sets of computable functions on S correlated with two such correspondences will in general differ. The natural procedure is to restrict consideration to those correspondences which are in some sense 'effective', and hence to characterize a computable function on S as a function f such that, for some effective correspondence between S and the set of natural numbers, the function induced by f under this correspondence is computable in Turing's sense. But the notion of effectiveness remains to be analyzed, and would indeed seem to coincide with computability.[1])

When we consider functions defined on all real numbers (or all real numbers within some interval), it is even clearer that Turing's notion of computability is not directly applicable. Yet there is in this domain an intuitive notion of computability; it is just such functions with which analogue computers are best equipped to deal. Turing's notion can be transferred to functions of real numbers by considering computations of approximations; in this way the notion of a *recursive real function* has recently evolved. We have no guarantee, however, that such functions coincide with the functions that, intuitively speaking, are analogue computable.

It remains then to develop a theory of computability (or effectiveness) which will apply in a natural way to functions of any sort. Such a theory should have as special cases both Turing's theory and a theory of analogue computability. A tentative basis for a theory of this kind is given in the following sections. Only explications are attempted here; no theorems of any interest are either proved or stated.

It is to be expected that an analysis of general computability will resemble Turing's development with respect to the three features noted above.

Besides the kinds of application to which Turing's theory has lent itself,

[1]) The notion of a computable function on expressions *can* be reduced to that of a computable function on natural numbers without the arbitrariness inherent in the usual procedure, which involves the choice of a particular Gödel-numbering. One follows the indications in the text, understanding by an *effective* correspondence one under which the operation of concatenating two expressions induces a computable function of natural numbers. The details of this suggestion, which are quite simple, may be found in Chapter IV of my doctoral dissertation, *Contributions to the axiomatic of set theory* (Berkeley, 1957, unpublished).

one might expect from general computability a new philosophical application of some importance. The hypothesis that Turing's explication is adequate, in other words, that the intuitively computable functions of natural numbers coincide with the Turing computable functions, is known as *Church's thesis*. Discussion of Church's thesis has suffered for lack of a precise general framework within which it could be conducted. The proposed general theory of computability may provide such a framework and lead to an exact evaluation of Church's thesis, as well as the related hypothesis that the intuitively computable functions of real numbers coincide with the recursive real functions, of which mention has earlier been made. It is now, however, too early for this hope to have been realized.

2. GENERAL COMPUTABILITY

By a *state-variable* we understand a function whose domain is the set of real numbers (regarded here as instants of time); a state-variable is understood as giving a complete history of one aspect of the world. The values of state-variables, in the cases most usual in physics, will be real numbers – for example, the temperature of a certain particle or position, or a component of the velocity of a certain object. This is not, however, essential. We permit state-variables to assume any values whatever; thus a state-variable could give the color of a certain object, or the object occupying a certain position.

A *logically possible world history* is an indexed system of state-variables, that is, a function D whose domain is some set U and whose values are state-variables. (The members of U are identified with the aspects of the world, and for each i in U, D_i is regarded as the state-variable giving the history of the aspect i.) We shall generally not be interested in all logically possible world histories; in the light of scientific knowledge, we may exclude some histories as physically impossible. Thus we distinguish a class K of *physically possible histories*, upon which, for the time being, we impose only the following conditions.

(1) K is a class of logically possible world histories.

(2) All histories in K have a common domain, which we may call the *index-set* of K, or I_K. (Thus the aspects of the world are regarded as fixed; they may not vary from one physically possible history to another, though their behavior may.)

(3) K is *closed under temporal translation*, that is, whenever a history D is in K and $\triangle$ is a real number, the history E will also be in K, where

$$E_i(t) = D_i(t + \triangle),$$

for each i in I_K and each real number t.

The definition of computability will be relativized to a class K satisfying these three requirements. We shall consider later how to specify K.

Our computers will consist of four aspects of the world – two inputs which are regarded as controlled by the operator, and two outputs, whose behavior is governed by the machine. (In a physical example, the four aspects might correspond to four wires carrying electric currents to or from the machine.) One of the inputs registers the argument for which a function value is to be computed, and one of the outputs registers this value; these two parts of a computer may be called its argument input and value output respectively. The remaining input, when brought to a designated value, acts as a signal that a computation is to be commenced, and the remaining output, upon attaining a designated value, indicates the completion of the computation; these two parts may be called the signal input and the signal output respectively. Thus a computer may be identified with a sextuple (i, j, k, l, a, b), where i, j, k, l are aspects of the world (that is, elementsof I_K); here i and k are regarded as the signal input and signal output respectively, j and l as the argument input and value output, and a and b as the designated values of the two signal elements.

Such a computer is operated as follows. Let x be the argument for which a function value is to be obtained. The operator brings the argument input j to the value x, and the signal input i to the designated value a. Then he waits. The computer may or may not be able to compute a function value corresponding to x. If it is, then after a certain time it will flash the signal output (that is, the aspect k will assume its special value b); the simultaneous value of the output l will be taken as the function value.

In characterizing a computer, it is natural to impose a condition implying the uniqueness of the computed value. Such a condition, which has as well some other consequences, is included in Definition 2 below.

Definition 1. Let A be an arbitrary set of real numbers. Then t is said to be a *minimal element* of A if t is a member of A and no member of A is less than t.

Clearly, any set of real numbers has at most one minimal element. A set may have no minimal element, even under the assumption that it has a greatest lower bound; an example is the set of positive real numbers.

Definition 2. A *computer* (relative to K) is an ordered sextuple (i, j, k, l, a, b) such that i, j, k, l are in I_K, a, b are any objects, and, for all D, E, t_0, t_1, if the following conditions hold:

(1) D, E are in K (that is, are physically possible histories),

(2) t_0 is a real number,

(3) $D_i(t_0) = E_i(t_0) = a$,

(4) $D_j(t_0) = E_j(t_0)$, and

(5) t_1 is a minimal element of the set of real numbers $t \geqslant t_0$ for which $D_k(t) = b$,

then there exists a real number t_2 such that

(6) t_2 is a minimal element of the set of real numbers $t \geqslant t_0$ for which $E_k(t) = b$, and

(7) $E_l(t_2) = D_l(t_1)$

Definition 3. Let x, y be any objects, let C be a computer (relative to K), and let C = (i, j, k, l, a, b). Then we say that y is a *value computed for the argument* x by the computer C (relative to K), or, in symbols, Val_K (y, x, C), if there are real numbers t_0 and t_1, together with an object D such that

(1) D is in K,

(2) $D_i(t_0) = a$,

(3) $D_j(t_0) = x$,

(4) t_1 is a minimal element of the set of real numbers $t \geqslant t_0$ for which $D_k(t) = b$, and

(5) $D_l(t_1) = y$.

In view of the stipulation that K be closed under temporal translation, we see that, for any computer C (relative to K), there is at most one object y such that $Val_K(y, x, C)$.

With each computer C we may associate the *function computed by* C (relative to the class K of physically possible histories); we shall call this function $f_{C,K}$.

Definition 4. If C is a computer, then $f_{C, K}$ is the set of ordered pairs (x, y) such that $Val_K(y, x, C)$.
Thus $f_{C, K}$ is a function whose domain consists of those objects for which C can compute a value, and for each such object x, $f_{C, K}(x)$ is the unique object y such that $Val_K(y, x, C)$. An object x will fail to be in the domain of $f_{C, K}$ if either (1) there is no time at which it is possible to *ask* C *the question concerning* x (that is, if C = (i, j, k, l, a, b), to bring the aspect i to the value a, and, simultaneously, the aspect j to the value x) or (2) one can ask C the question concerning x, but C is incapable of finding an answer. By the fact that the class of physically possible histories is closed under temporal translation, it is clear that if at some time it is possible to ask C the question concerning x, then this is also possible at any other time.

Definition 5. A function f is *computable* (relative to the class K of physically possible histories) if and only if there is a computer C (relative to K) such that $f = f_{C, K}$.
Notice that we have not restricted our attention to those computers whose behavior, upon the asking of a question, could properly be called deterministic. Although there is only one possible value y which a given computer C can obtain for an element x of the domain of $f_{C, K}$, we have not ruled out the possibility that there may be many different ways in which C can arrive at y. (In other words, there may be many different values of t_1 and D which, for given values of t_0, i, j, k, l, a, b, x, y, satisfy conditions (1) – (5) of Definition 3.) The computers which behave deterministically may be singled out as follows.

Definition 6.[1]) A *deterministic computer* (relative to K) is an ordered sextuple (i, j, k, l, a, b) which is a computer (relative to K) and such that, for all D, E, t_0, t_1, if the following conditions hold:
(1) D, E are in K,
(2) t_0, t_1 are real numbers such that $t_0 \leqslant t_1$,
(3) $D_i(t_0) = E_i(t_0)$, $D_j(t_0) = E_j(t_0)$, $D_k(t_0) = E_k(t_0)$, $D_l(t_0) = E_l(t_0)$,
(4) $D_i(t_0) = a$, and

1) A closely related notion of a *partially deterministic theory* was introduced in my paper, 'Deterministic theories', which is to appear in a collection published by the United States' Air Force Office of Scientific Research, and in *Philosophy of Science*.

(5) for all real numbers t such that $t_0 \leqslant t < t_1$, $D_k(t) \neq b$,
then the following equalities also hold:

$D_i(t_1) = E_i(t_1)$
$D_j(t_1) = E_j(t_1)$,
$D_k(t_1) = E_k(t_1)$,
$D_l(t_1) = E_l(t_1)$.

3. PHYSICALLY POSSIBLE HISTORIES

In specifying the class K of physically possible histories, it is natural to refer to the models of some scientific theory. For the construction of theories, we suppose the following *symbols* to be available:

(1) the *logical constants* of the first-order predicate calculus with identity and definite descriptions, that is, left and right parentheses together with the symbols v, . , ⊃ , ≡ , ~ , ∀ , ∃ , = , ι (respectively read 'or', 'and', 'if... then', 'if and only if', 'not', 'for all', 'for some', 'is identical with', 'the unique object... such that');
(2) a denumerable infinity of (individual) *variables*;
(3) for each non-negative integer n, an infinity of n-place *operation symbols*;
(4) for each positive integer n, an infinity of n-place *predicates*.

In particular, we stipulate that R and I shall be one-place predicates, and + and · two-place operation symbols. (These four symbols are respectively regarded as denoting the class of real numbers, the subclass of integers, and the operations of addition and multiplication.)

An *expression* is a concatenation of finitely many symbols. A *quasi-term* is a variable, or else an expression beginning with ι or an operation symbol. The class of *meaningful expressions* is the smallest class U such that (1) all variables are in U, (2) for any n-place predicate or operation symbol δ, and any quasi-terms $\zeta_1, \ldots, \zeta_n$ in U, the expression $\delta\, \zeta_1 \ldots \zeta_n$ is in U, (3) for any quasi-terms ζ, η in U, the expression $\zeta = \eta$ is in U, (4) for any variable α and any members φ, ψ of U which are not quasi-terms, the expressions $(\varphi \vee \psi)$, $(\varphi \cdot \psi)$, $(\varphi \supset \psi)$, $(\varphi \equiv \psi)$, $\sim \varphi$, $\forall\, \alpha\varphi$, $\exists\, \alpha\varphi$, $\iota\, \alpha\varphi$ are in U.

A *term* is a meaningful expression which is a quasi-term, a *formula* a meaningful expression which is not a quasi-term.

A *model* is an ordered triple (A, E, x) where A is a set containing x as a member and E is a function assigning to each predicate and operation symbol an extension of the proper sort; that is, if δ is an n-place predicate, then $E(\delta)$ is a set of ordered n-tuple of elements of A, and if δ is an n-place operation symbol, then $E(\delta)$ is a function whose domain is the set of ordered n-tuples of elements of A and whose values lie always in A. An *assignment* within a model (A, E, x) is a function whose domain is the set of variables and whose values lie always in A. If f is an assignment within some model, α a variable, and a any object, then f^{α}_{a} is that function g whose domain is the set of variables and such that $g(\alpha) = a$, and $g(\beta) = f(\beta)$ for every variable β other than α.

If φ is a meaningful expression, M = (A, E, x), M is a model, and f is an assignment within M, then the notion of the *value* of φ in M for f, or $\text{Val}(\varphi, M, f)$, is introduced by the following recursive definition. (1) If α is a variable, then $\text{Val}(\alpha, M, f) = f(\alpha)$. (2) Let $\zeta_1, \ldots \zeta_n$ be terms. If δ is an n-place predicate, then $\text{Val}(\delta\zeta_1 \ldots \zeta_n, M, f)$ is 1 or 0 according as the n-tuple $(\text{Val}(\zeta_1, M, f), \ldots, \text{Val}(\zeta_n, M, f))$ is or is not in $E(\delta)$. If, on the other hand, δ is an n-place operation symbol, then $\text{Val}(\delta\zeta_1 \ldots \zeta_n, M, f)$ is $[E(\delta)]\,(\text{Val}(\zeta_1, M, f), \ldots, \text{Val}(\zeta_n, M, f))$. (3) If ζ, η are terms, then $\text{Val}(\zeta = \eta, M, f)$ is 1 or 0 according as $\text{Val}(\zeta, M, f)$ is or is not identical with $\text{Val}(\eta, M, f)$. (4) Let α be a variable and φ, ψ formulas. Then $\text{Val}((\varphi \vee \psi), M, f)$ is 1 if either $\text{Val}(\varphi, M, f)$ is 1 or $\text{Val}(\psi, M, f)$ is 1, and is otherwise 0; the values of $(\varphi \cdot \psi)$, $(\varphi \supset \psi)$, $(\varphi \equiv \psi)$, $\sim \varphi$ are determined similarly. $\text{Val}(\forall\, \alpha\varphi, M, f)$ is 1 if $\text{Val}(\varphi, M, f^{\alpha}_{a})$ is 1 for all a in A, and is otherwise 0; similarly for $\exists\, \alpha\varphi$. If there is a unique object a in A for which $\text{Val}(\varphi, M, f^{\alpha}_{a})$ is 1, then $\text{Val}(\iota\alpha\varphi, M, f)$ is that object a; otherwise $\text{Val}(\iota\alpha\varphi, M, f)$ is x.

Let φ be a meaningful expression whose free variables are $\alpha_1, \ldots, \alpha_n$, M be a model, and M = (A, E, x). Then the *function defined by* φ *in* M, or $F_{\varphi, M}$, is that function whose domain is the set of ordered n-tuples of elements of A and such that, for any $a_1, \ldots, a_n$ in A, $F_{\varphi, M}(a_1, \ldots, a_n)$ is the unique object b such that, for some assignment f within M, $f(\alpha_i) = a_i$ for $i = 1, \ldots, n$, and $\text{Val}(\varphi, M, f) = b$. If φ is a formula, φ is *true in* M just in case $F_{\varphi, M}(a_1, \ldots, a_n) = 1$ for all elements $a_1, \ldots, a_n$ of A.

A *standard model* is one that assigns to mathematical symbols their expected meaning, that is, a model (A, E, x) such that E(R) and E(I) are respectively the set of real numbers (or one-tuples of real numbers) and the set of integers (or one-tuples of integers), and E(+) and E(·) are functions which, when restricted to the set of real numbers, coincide with addition and multiplication respectively. (This notion of a standard model is appropriate when we consider the computability only of functions of real numbers. In other cases, more stringent requirements would be imposed.)

By a *scientific theory* we may understand a pair (T, U), where T is a set of formulas and U a set of terms; for simplicity, we make a rather restrictive assumption: each term in U is to contain exactly one free variable. The members of T are regarded as the axioms of the theory, and the members of U as the observationally meaningful terms of the theory; the free variable of such a term is regarded as referring to time. A *model of* a theory (T, U) is a model in which all members of T are true.

Given a scientific theory (T, U), we may identify the aspects of the world with the members of U. Thus a *physically possible history* relative to (T, U) is a function D with domain U for which there exists a standard model M of (T, U) such that, for all ζ in U, D_ζ is the restriction of the function $F_{\zeta, M}$ to the set of real numbers.

The notion of computability, at least for functions of real numbers, can thus, it seems, be relativized to a scientific theory rather than a class of possible histories.[1] Such a theory could consist of a description of the behavior of a certain type of machine. It is, for example, easy to construct a theory describing the behavior of Turing machines; we can then prove that a function is computable relative to this theory if and only if it is a partially computable function of natural numbers (in the standard sense). It would be more interesting, however, to consider a theory which formalizes physics or some branch of it, and to ask what functions then turn out to be computable.

University of California, Los Angeles and Berkeley

[1] Thus computability relative to (T, U) is defined as computability relative to K, where K is the set of physically possible histories relative to (T, U). The requirement that K be closed under temporal translation corresponds now to a requirement on (T, U).

CARL G. HEMPEL

INDUCTIVE INCONSISTENCIES *)

1. INTRODUCTION

In the philosophical study of induction, no task is of greater importance than that of giving a clear characterization of inductive procedures: only when this has been done can the problem of justification significantly be raised. If induction is conceived as a peculiar type of inferential reasoning, its precise characterization will naturally call for the formulation of distinctive rules of inductive inference. A variety of such rules have indeed been set forth in the philosophical literature. But certain quite familiar types of such rules, though widely countenanced even in recent writings on the subject, can be shown to lead into logical inconsistencies. This is the more serious because the defective rules include some which have been held to represent the most basic types of sound inductive reasoning.

In this article, I propose to exhibit this defect in two familiar types of induction rules and to examine the sources of the 'inductive inconsistencies' they generate. The ideas here set forth are based to a large extent on the work of others, and especially on Carnap's conception of inductive logic and its applications.

2. INCONSISTENCIES GENERATED BY STATISTICAL SYLLOGISMS

One type of inductive inference that leads into inconsistencies is represented by the so-called statistical syllogism and its variants.

A statistical syllogism [1]) is an argument of the form

a is F
(2.1) The proportion of F's that are G is q
Hence, with probability q, a is G

*) This article was written during my tenure, on a United States Government Fellowship, as a Fulbright Research Fellow at the University of Oxford, 1959–60.

[1]) See, for example, Williams, D. C. *The Ground of Induction* (Harvard University Press, 1947); and the discussion of the idea in ch. IV of Barker, S. *Induction and Hypothesis* (Cornell University Press, 1957).

In some variants of this mode of reasoning, the conclusion or also the second premiss is expressed in non-numerical terms. Thus, e.g. Toulmin [1]) puts forward as valid certain types of argument which he calls quasi-syllogisms, and which take forms such as the following:

a is F
(2.2) The proportion of F's that are G is less than 2 per cent
So, almost certainly (or probably) a is not G.

a is F
(2.3) The proportion of F's that are G is minute
So, almost certainly (or probably) a is not G.

The inference patterns here listed are applicable only when the 'reference class' F is finite; for only then has the phrase 'the proportion of F's that are G' a clear meaning. Analogous types of argument which are not subject to this restriction are suggested, however, by the frequency interpretation of statistical probability. In current mathematical theory, statistical probabilities are construed as set-measures governed by certain axioms; and a formula of the form '$p(G, F) = r$', which specifies the statistical probability of set G with respect to set F, asserts, roughly, that the measure of the intersection of G and F, divided by the measure of F, equals r. The application of the mathematical theory to empirical subject matter is effected by the frequency interpretation of statistical probability, which construes '$p(G, F) = r$' as stating the long-run relative frequency, r, with which a 'random experiment' of some specified kind F – performed by man or by nature – tends to yield an outcome of kind G. For the case where r is close to 1, this frequency interpretation is usually expressed in the following form: If $p(G, F)$ is very close to 1, then if an experiment of kind F is performed just once, it is practically certain that a result of kind G will occur.[2]) This principle might be thought to authorize the following inference schema, in which the second premiss no longer requires the reference class F to be finite:

[1]) Toulmin, S., *The Uses of Argument* (Cambridge University Press, 1958), pp. 109ff. (For the conclusion-form 'almost certainly, or probably, a is not G', see p. 139).
[2]) This formulation follows closely those given in Cramér, H. *Mathematical Methods of Statistics* (Princeton University Press, 1946), p. 150, and in Wald, A. *On the Principles of Statistical Inference* (University of Notre Dame, Indiana, 1942), p. 2.

a is *F*
(2.4) The statistical probability for an *F* to be a *G* is nearly 1
So, it is almost certain that *a* is *G*

For convenience, I shall henceforth refer to all the different types of inference just listed, and to certain analogous ones, as *broadly statistical syllogisms*, or briefly as statistical syllogisms. Now it is readily seen that all broadly statistical syllogisms lead into inconsistencies because the individual case *a* which the conclusion assigns to the class *G* (or to which the conclusion attributes the characteristic, or property, *G*) will in fact belong to different reference classes, F_1, F_2, . . . whose members exhibit *G* with different relative frequencies or statistical probabilities. For arguments of form (2.1), an example given by Barker [1]) illustrates this neatly: Suppose that Jones is a Texan, and that 99 per cent of Texans are millionaires; but that Jones is also a philosopher, and that only 1 per cent of these are millionaires. Then rule (2.1) permits the construction of two statistical syllogisms, both with true premisses, which yield the incompatible conclusions that, with probability .99, Jones is a millionaire, and that with probability, .01, Jones is a millionaire. Consider next Toulmin's example of a quasi-syllogism of form (2.2):[2]

Petersen is a Swede
(2.5) The proportion of Roman Catholic Swedes is less than 2 per cent
So, almost certainly, Petersen is not a Roman Catholic.

Suppose that the premisses of this argument are true. Then, as Cooley [3]) has pointed out, the premisses of the following quasi-syllogism may well be equally true:

Petersen made a pilgrimage to Lourdes
(2.6) Less than 2 per cent of those making a pilgrimage to Lourdes are not Roman Catholics
So, almost certainly, Petersen is a Roman Catholic.

Thus, the quasi-syllogistic inference schema can lead from true premisses to incompatible conclusions.[4]

1) Barker, *loc. cit.*, p. 76.
2) Toulmin, *loc. cit.*, p. 109.
3) Cooley, J. 'On Mr. Toulmin's Revolution in Logic', *The Journal of Philosophy* 56: 297–319 (1959), p. 305. The phrasing of Cooley's example has been slightly modified to make it fit the pattern (2.2) more closely.
4) While Toulmin repeatedly emphasizes that quasi-syllogisms are *valid*, he later adds the remark: 'It must of course be conceded that quasi-syllogisms can properly be

To construct an analogous example for the schema (2.4), consider a set of 10,000 balls of which 9,000 are made of glass and are white, while the remaining 1,000 are made of ivory, one of them being white, the other 999, black. Let D be a certain procedure of selecting one of the 10,000 balls. Let us assume that this is a random procedure, so that the statistical probability of obtaining a white ball as a result of D will be $p(W, D) = .9001$. Let the event b be one particular performance of the experiment D. Then (2.4) yields the following argument with true premisses:

b is D

(2.7) $p(W, D) = .9001$

Hence, it is almost certain that b is W (i.e. that b yields a white ball).

Suppose now that b happens to yield an ivory ball. Then b may also be regarded as an instance of another experiment, D^*, which consists in selecting at random one of the ivory balls in the given set. But for this experiment, the probability of selecting a non-white ball is $p(-W, D^*) = .999$; and schema (2.4) now authorizes the argument:

b is D^*

(2.8) $p(-W, D^*) = .999$

Hence, it is almost certain that b is $-W$ (i.e. that b does not yield a white ball).

Again, we have a pair of rival arguments conforming to the same rule and starting with true premisses, and yet leading to incompatible conclusions. Despite its apparent plausibility, then, the construal of certain types of statistical arguments as having the form of broadly statistical syllogisms is untenable; for those syllogisms generate inductive inconsistencies [1] in the following sense: For an argument with true premisses that has the form of a statistical syllogism, there exists, in general, a rival argument of the same form, again with true premisses, whose conclusion is logically incompatible with that of the first argument.

advanced only if the initial data from which we argue state all that we know of relevance to the question at issue' (*loc. cit.*, p. 140). This remark, which implies that the argument (2.5) 'can be properly advanced' only if the premisses of Cooley's quasi-syllogism are not known to be true, will be considered in section 4 below.

[1]) In an essay dealing with the explanatory and predictive use of statistical probability statements, I have referred to this peculiarity as the *ambiguity* of statistical explanation and prediction; cf. 'Deductive-Nomological *vs.* Statistical Explanation', forthcoming in Feigl, H. (ed.) *Minnesota Studies in the Philosophy of Science*, vol. III. (Minneapolis: University of Minnesota Press.)

This is true also of an inductive rule of a slightly different kind, which is among those listed by Black in essays dealing with the justifiability of induction. Black formulates it as follows:

R: To argue from *Most instances of A's examined in a wide variety of conditions have been B* to (probably) *The next A to be encountered will be B.*[1])

Black adds that inductive arguments governed by R vary in 'strength' according to the number and variety of the favorable instances reported in the premiss; so that 'although R permits us to *assert a certain conclusion categorically*, ... the strength of the assertion fluctuates with the character of the evidence'.[2]) In contrast to broadly statistical syllogisms, then, rule R leads to a conclusion which does not contain a modal qualifier like 'probably' or 'certainly'; yet, the conclusion is supposed to be asserted with more or less 'strength'. Our earlier illustrations show readily that an argument which, in accordance with R, leads from true premisses to a very strong assertion of a given conclusion can generally be matched by a rival one, governed by the same rule, which from equally true premisses leads to the strong assertion of the contradictory of that conclusion. In this sense, rule R generates inconsistencies.

Deductive forms of inference never generate inconsistencies, of course. In particular, for an argument of the syllogistic form

a is F
(2.9) All F are G
a is G

whose premisses are true, there exists no rival argument of the same form whose premisses are true as well, and whose conclusion is logically incompatible with that of the given argument: incompatible conclusions can be deduced only from incompatible premiss-sets, and sets of true premisses are not incompatible.

[1]) Black, M. 'Self-Supporting Inductive Arguments', *The Journal of Philosophy* 55: 718–725 (1958), p. 720 (italics in the original); see also the same author's 'The Inductive Support of Inductive Rules', in Black, M. *Problems of Analysis* (Cornell University Press, 1954), p. 196.

[2]) Black, 'Self-Supporting Inductive Arguments', p. 720 (Italics supplied.) Black notes that the rule 'as it stands' is not 'a wholly acceptable rule for inductive inference' (*ibid.*); but he holds that the rule R can be used in a legitimate inductive argument supporting R itself, and it seems fair, therefore, to assume that the faults he finds with this rule do not include so decisive a defect as that of generating inconsistencies.

3. PROBABILITY: MODAL QUALIFIER OR RELATION?

The inconsistencies just noted do not show, of course, that all non-deductive arguments based on statistical information are unsound, but only that the construal of such arguments as quasi-syllogistic is untenable. That construal seems to aim at too close a formal assimilation of non-deductive statistical arguments to deductive inference. Thus, e.g., given that the premisses of the deductive syllogism (2.9) are true, the conclusion '*a* is *G*' will 'necessarily' – i.e., as a logical consequence – be true as well and can therefore be categorically asserted. In the corresponding statistical arguments, however, the truth of the premisses does not thus guarantee the truth of '*a* is *G*'; and if, in analogy to the deductive case, one insists on formulating a sentence which the truth of the premisses would entitle us to assert, it may seem tempting to do so by prefixing to '*a* is *G*' a qualifying phrase such as 'it is practically certain that', 'very probably', or 'with probability *r*'. And this is precisely what is done when statistical arguments are construed as quasi-syllogistic.

That this is a misconstrual becomes clear when we reflect that by the same token we should be able to schematize the deductive syllogism (2.9) in the form

a is *F*
(3.1) All *F* are *G*
Hence, certainly (or, necessarily) *a* is *G*.

In fact, Toulmin does just this when he puts the syllogistic counterpart of one of his quasi-syllogistic arguments into the form

Petersen is a Swede
(3.2) No Swedes are Roman Catholics
So, certainly, Petersen is not a Roman Catholic.[1]

But the certainty here in question is clearly a logical relation between the premisses and the conclusion of a deductive argument: the statement '*a* is *G*' is *certain*, or *necessary*, *relative to the given premisses*, i.e., it is logically implied by them. To treat the term 'certainly 'in the manner of (3.1) and (3.2), as a qualifier applicable to a single statement is incorrect: If the logical force of the argument (2.9) is to be expressed with the help of the term 'certain' or its cognates, then it has to be done in an explicitly relativized form, such as this:

[1] Toulmin, *loc. cit.*, p. 131.

(3.3) '*a* is *G*' is certain relative to (i.e., is logically implied by) '*a* is *F*' and 'All *F* are *G*'.

To say this is not to deny that the word 'certain' and its cognates can also be used as qualifiers of single statements, in contexts of the form 'it is certain that *p*', 'certainly *p*', etc. Let me distinguish three major purposes for which phrases of this kind are used: (i) to claim that the particular statement standing at the place of '*p*', or briefly the *p*-statement, is a logico-mathematical truth or perhaps a nomological one (i.e., a consequence of certain laws of nature), so that we are entitled to assert it categorically and without qualifications; (ii) to claim that the *p*-statement is categorically and unqualifiedly assertable in some more inclusive, and more elusive, sense which is conceived as being governed by objective standards (some would make this claim, for example, for a class of presumptive *a priori* truths thought to include the truths of logic and of mathematics as a proper subclass); (iii) to show – rather than to state – that the utterer of the phrase means to assert the *p*-sentence without qualification, and perhaps with special emphasis. But if 'certainly' is understood in the first of these senses, then arguments such as (3.1) and (3.2) are simply fallacious. The same holds true for the second sense of 'certainly'. If, for example, that qualifier is taken to apply to all and only those sentences which are *a priori* truths in some specified sense, then it may well happen that in an argument of the form (3.1) which has true premisses, the conclusion, though true, is not an *a priori* truth: hence, in arguments of the form (3.1) the premisses then by no means warrant the conclusion. The schema (3.1) could be turned into a sound form of argument by adding the prefix 'certainly' to both of the premisses: But the resulting schema would no longer represent the syllogistic argument whose logical structure (3.1) was intended to exhibit. Finally, if the word 'certainly' is taken in the third sense, then its presence is as irrelevant to the logic of the argument as would be the occurrence of such words as 'emphatically', 'fortunately', or 'unexpectedly' in its place.

In sum, then, it is simply incorrect to represent the logical force of a syllogistic argument in the manner of (3.1) or (3.2), where the word 'certainly' plays the rôle of a modal qualifier of the conclusion: certainty must be construed here as a logical relation, in the manner of (3.3). The fact that the phrasing 'certain relative to . . .', which is used in (3.3), does not occur in ordinary English is not, of course, a flaw of the proposed construal: in fact, it is precisely a too close adherence to phrasings used in everyday discourse which has obscured the logic of the inferences here under consideration.

Analogous remarks apply to statistical arguments of the kind which the notions of statistical syllogism and of quasi-syllogism are intended to illuminate. In the context of such arguments, phrases such as 'it is practically certain that', 'probably', etc., as well as Black's expression 'strength of assertion', must be construed, not as qualifying the conclusion, but as representing a logical relationship between the premisses and the conclusion: they indicate the extent to which the premisses support or confirm the (unqualified) conclusion. Thus, e.g., the arguments whose structure the schema (2.2) was meant to exhibit are not to the effect that from the given premisses we may validly infer 'Almost certainly, a is not G', but rather to the effect that those premisses lend very strong support to the statement 'a is not G', or that the premisses confer upon this statement a very high probability.

Thus, in analogy to (3.3), the arguments which (2.2) was meant to represent might be schematized as follows:

(3.4) 'a is not G' is almost certain (or is highly probable) relative to the two statements 'a is F' and 'Less than 2 per cent of F's are G'.

The concept of probability here invoked is not, of course, the statistical one, which, as we noted, represents a quantitative relation between two kinds or classes of events, F and G; rather, it is what Carnap has called logical or inductive probability, or degree of confirmation – a concept representing a logical relation between statements. This inductive probability is the central concept of the theories of probability developed by Keynes, Mazurkiewicz, Jeffreys, von Wright, and other writers. It is still a controversial question to what extent the inductive support conferred by an evidence statement e upon a hypothesis h can be represented by a precise quantitative concept $c(h, e)$ with the formal characteristics of a probability. At any rate, Carnap has developed a rigorous general method of defining such a concept which is applicable to formalized languages having the structure of a first-order functional calculus.[1])

[1]) See especially his *Logical Foundations of Probability* (The University of Chicago Press, 1950), sec. 110; and the generalization in *The Continuum of Inductive Methods* (The University of Chicago Press, 1952).

But the main point here at issue is independent of the prospects for the development of a precise quantitative theory of inductive logic: If terms such as 'almost certainly', 'probably', and 'with probability r' are to express the force of the inductive statistical arguments we have been considering then they must be understood, not as qualifiers of single statements, but as representing relations between statements. These relations might be expressed in the manner of (3.4); or, in the framework of a quantitative inductive logic such as Carnap's, in formulas of the form

(3.5) $c(h, e_1 e_2 \ldots e_n) = r$

which indicate that the statements (inductive 'premisses') $e_1, e_2, \ldots, e_n$ jointly confer the logical probability r upon the statement (inductive 'conclusion') h.

In conclusion of this brief comparison of deductive and inductive inference, one further point should be noted: The schematizations (3.3), (3.4), (3.5) are concerned only with the logical connections between the premisses and the conclusion and not at all with their truth or falsehood. But since in a deductive argument the conclusion cannot fail to be true if the premisses are true, deductive inference rules can be used to effect a transition from given statements which are known or considered to be true to another statement which has the same status; thus, as Carnap puts it,[1] deductive inference rules permit, as it were, the acquisition of new statements on the basis of statements already possessed. In an inductive inference, on the other hand, the 'premisses' lend only partial support to the 'conclusion', and truth is not, therefore, automatically transferred from the former to the latter. Hence even if the premisses all belong to the class of statements previously accepted or possessed, the conclusion cannot be added to that class; it can only be qualified by a number representing its probability relative to the premisses. In reference to inductive 'inferences' or 'arguments', therefore, one can speak of a 'conclusion' only *cum grano salis:* the conclusion cannot be detached from the premisses and asserted on its own when the premisses are true. The question whether the detachability of the conclusion of a deductive inference with true premisses has at least some weaker analogue in the case of inductive inference will be considered in the final section of this essay.

[1]) Carnap, *Logical Foundations of Probability*, p. 206.

The idea that in the context of inductive arguments probability has to be construed as a relation has recently been criticized at length by Toulmin, who especially takes Carnap and Kneale to task for holding this view, and who insists, on the contrary, 'that 'probably' and its cognates are, characteristically, modal qualifiers of our assertions'; more specifically: 'To say 'Probably *p*' is to assert guardedly, and/or with reservations, that *p*; it is *not* to assert that you are tentatively prepared to assert that *p*.'[1]) Now surely, in ordinary discourse, the word 'probably' and its cognates are often used in this way. We may distinguish here, more precisely, between two purposes which the qualifier in phrases such as 'probably *p*' may serve: (i) It may show – rather than state – to what extent the speaker is willing to commit himself to *p*. (If the qualifier has the form 'With probability *r*', then the quotient $r/(1 - r)$ may indicate the odds at which – for whatever reasons – the speaker is prepared to bet on *p*); or else (ii) the qualifier may indicate the extent to which it is rationally assertable or credible that *p*, where rational assertability or credibility is thought of as governed by objective standards. Toulmin does not seem to opt quite unequivocally for one of these two meanings in which 'probably' and its cognates may be used. The following statement of his, for example, suggests the first meaning: 'When I say '*S* is probably *P*', I commit myself guardedly, tentatively or with reservations to the view that *S* is *P* and (likewise guardedly) lend my authority to that view.'[2]) However, the second meaning appears to be closer to what Toulmin has in mind; as is suggested, for example by his remark: 'Actually, statements about the probability of *p* are concerned, in practice, with the extent to which we are *entitled* to bank on, take it that, subscribe to, put our weight and our shirts on *p* ...'[3])

But when used in the first sense, qualifiers such as 'probably' clearly cannot serve to exhibit the logic of a statistical argument; and if they are understood in the second sense, then they have to be construed as relative to given grounds. For the credibility of an empirical assertion – in sharp contrast to its truth or falsity – depends on the available evidence; the phlogiston theory of combustion, for example, was much more highly credible on the evidence available before Lavoisier's researches than

[1]) Toulmin, *loc. cit.*, pp. 84 and 85. (Italics in the original.)

[2]) Toulmin, *loc. cit.*, p. 53.

[3]) Toulmin, *loc. cit.*, p. 83 (Italics supplied).

afterwards. Hence, a phrase of the form 'It is highly credible that *p*' (or 'probably *p*', in the sense here under discussion) is not a self-contained statement any more than a phrase of the form 'x is a larger number'. Frequently, expressions of the form 'almost certainly, *p*', 'probably *p*', etc., as used in ordinary discourse, can be regarded as elliptical statements referring to the total evidence available at the time of utterance or at some other time suggested by the context. When we say, for example, 'Probably, there is no life on the moon', the tacit reference is presumably to the evidence available at present. But if the qualifier in the conclusion of a statistical syllogism is thus understood as relative to the total evidence available at the time when the syllogism is presented, the argument is of course invalid: The premisses of (2.2), for example, do not warrant the conclusion that on the *total* evidence available, it is very probable that *a* is not *G*. The only reasonable construal left is that suggested in (3.4).

Toulmin emphatically rejects this relativization of probability and insists that reference to the total evidence is required only for estimating or measuring the probability of a hypothesis, which itself is a non-relational characteristic, just as reference to evidence is required to estimate the truth value of a hypothesis.[1]) But this analogy is misleading. The truth values, truth and falsity, are non-relational characteristics of hypotheses; i.e., a phrase of the form 'hypothesis *h* is true' is a self-contained statement which need not be supplemented by specifying some body of evidence. To estimate whether a given hypothesis *h* is true or false, we have to refer to the available evidence, say *e*, which will confer on *h* a more or less high confirmation, $c(h, e)$: the latter represents the probability of *h* – or, what comes to the same, the probability that *h* is true – on the evidence *e*. This probability will normally change with the evidence, whereas the truth value of *h* is completely independent of it. Thus, as we noted before, phrases of the form '*h* is probable', or '*h* has the probability *r*', are not self-contained statements at all, and it makes no sense therefore to speak of measuring or estimating the probability of *h*, any more than it makes sense to speak of estimating whether the number 7 is larger. And though Toulmin has interesting things to say about the ordinary use of words like 'probably', his remarks give no clear meaning at all to the

[1]) Toulmin, *loc. cit.*, pp. 80–81.

notion of probability as a non-relational concept.[1]) In this case, ordinary usage has surely proved to be an unreliable guide.[2])

As soon as, in the schematization of statistical inferences, the concepts of probability, near-certainty, etc., are recognized as relational and the various types of broadly statistical syllogism are accordingly replaced by schemata of the kind suggested in (3.4) and (3.5), one perplexing aspect of statistical arguments vanishes; namely, the impression that statistical arguments make it possible to establish, on the basis of true premisses, pairs of incompatible conclusions of such forms as 'Almost certainly (very probably) *a* is *G*' and 'Almost certainly (very probably) *a* is not *G*'. For example, the two apparently conflicting arguments (2.5) and (2.6) do not, as their quasi-syllogistic construal incorrectly suggests, establish the conclusions that Petersen almost certainly is a Roman Catholic, and that he almost certainly is not: rather, the arguments show that relative to *one* set of premisses, the statement 'Petersen is a Roman Catholic' is highly probable, whereas its contradictory is highly probable relative to *another* set of premisses: and this does not involve a logical inconsistency any more than does the observation that certain sets of premisses deductively imply the statement 'Petersen is a Roman Catholic', whereas other sets deductively imply its contradictory.

4. THE REQUIREMENT OF TOTAL EVIDENCE

But while construal in the manner of (3.4) thus removes one puzzling aspect of statistical arguments, it does not fully dispose of the problem

[1]) He does say (*loc. cit.*, p. 55): 'surely, if I say 'It is probably raining' and it turns out not to be, then ... I was mistaken;' and later he again qualifies as 'paradoxical and inconsistent with our common ways of thinking' the idea that 'if I say, 'it is probably raining', the discovery that no rain was falling would not refute my statement'. (*loc. cit.*, p. 84). These remarks suggest strongly that in Toulmin's view the statement 'it is not raining' implies 'it is not the case that it is probably raining'. But then, by contraposition 'it is probably raining' would imply 'it is raining'. And while this construal would give a strong empirical content to sentences of the form 'probably, *p*', it is of course quite unacceptable; and it also conflicts with Toulmin's general observation that 'one cannot specify any happening which would conclusively verify or falsify a prediction held out as having only a certain probability" (*loc. cit.*, p. 82): thus, his views on the content and on the refutability of non-relativized probability statements remain unclear.

[2]) For further discussion of the relations between truth, probability, and verification, see Carnap, R. 'Truth and Confirmation', and 'The Two Concepts of Probability', sec. VI; both in Feigl, H. and Sellars, W. (eds.) *Readings in Philosophical Analysis.* (New York Appleton-Century-Crofts, 1949.)

raised by the inconsistencies encountered in section 2. The unresolved residual problem is this: If two sets of statements deductively imply contradictory consequences then the statements in the two sets cannot all be true: hence at least one of the arguments is based on some false premisses. But, as we noted, if two sets of statements confer very high probabilities upon contradictory conclusions, the statements in the two sets may all be true. Thus, we face the question: Given two valid inductive arguments whose premisses have been tested and accepted as presumably true, but whose conclusions – pertaining perhaps to some future event – are logically incompatible: on which, if any, of them are we to base our expectations and decisions? Or, more generally: On the basis of different sets of statements that we consider as true, a given hypothesis *h* – e.g., a prediction – can be assigned quite different probabilities: which of these, if any, is to count as a guide in forming our beliefs concerning the truth of *h* and in making decisions whose outcomes depend on whether *h* is true?

An answer is suggested by a principle to which we have alluded before, and which has in fact been tacitly or explicitly accepted by many writers on inductive reasoning. Carnap calls it *the requirement of total evidence* and formulates it as follows: 'In the application of inductive logic to a given knowledge situation, the total evidence available must be taken as basis for determining the degree of confirmation.' [1]) Broadly speaking, we might say that according to this requirement, the credence which it is rational to give to a statement at a given time must be determined by the degree of confirmation, or the logical probability, which the statement possesses on the total evidence available at the time. Alternatively, that credence may be determined by reference to any part of the total evidence which gives to the statement the same support or probability as the total evidence: in this case, the omitted portion of the total evidence is said to be *inductively irrelevant* to the statement, relative to the evidence actually used.

For our residual problem, this principle implies the maxim that the support which the premisses of a statistical argument confer upon its conclusion can serve to determine the credence rationally to be given to

[1]) Carnap, *Logical Foundations of Probability*, p. 211; cf. also Carnap, R. 'On the Application of Inductive Logic', *Philosophy and Phenomenological Research* 8: 133–48 (1947–48), pp. 138–39.

that conclusion or the decisions rationally to be based on it only if the premisses constitute either the total evidence *e* available at the time or else a part of *e* which supports the conclusion to the same extent as does *e*.

Compliance with the requirement of total evidence disposes of our residual problem. For suppose we are confronted with two statistical arguments of which one attributes near-certainty to '*a* is *G*', the other to '*a* is not *G*'. Then these arguments cannot both meet the requirement of total evidence. For if they did, the probabilities which their premisses confer upon '*a* is *G*' and '*a* is not *G*', respectively, would equal the probabilities which the total evidence confers upon those statements: but one and the same body of evidence, e.g. the total evidence, – provided only that it is logically consistent – cannot confer high probabilities on each of two contradictory statements; for the two probabilities add up to 1.

Let us note in passing that the requirement of total evidence is trivially satisfied by any *deductive* argument whose premisses are part of the total evidence. For here the premisses confer certainty, and thus the logical probability 1, upon the conclusion; but so does the total evidence available since, by hypothesis, it includes the premisses of the given argument.[1]

At this point, let us consider briefly a criticism which Ayer [2] has levelled against the principle of total evidence and indeed against the conception of inductive probability as a logical relation between statements. Ayer notes that according to Keynes, Kneale, Carnap, and certain other authors, probability statements are non-empirical: if they are true, they are necessarily true; if false, necessarily false. This feature is especially clear in Carnap's theory of inductive probability, according to which any statement of the form '$c(h, e) = r$' – which is the basic form of an inductive probability statement – is either analytic or self-contradictory. Now, if for *h* we choose some fixed hypothesis, such as that our favorite horse will win tomorrow's race, then, Ayer points out, we can assign to it many different probabilities simply by taking into account more and more of the relevant evidence. But since each of these probability

[1]) On this point, see also Carnap, *Logical Foundations of Probability*, p. 211.

[2]) Ayer, A. J. 'The conception of probability as a logical relation'. In Körner, S. (ed.) *Observation and Interpretation*. Proceedings of the Ninth Symposium of the Colston Research Society. (New York and London, 1957), pp. 12–17.

statements would constitute a necessary truth, none of them can be regarded as superior to the others. 'The addition of more evidence may, indeed, yield a higher or lower probability for the statement in which we are interested. But ... this probability cannot be said to be more, or less, correct than the one which was yielded by the evidence with which we started.[1]) The difficulty here adumbrated is closely related to the residual problem mentioned at the beginning of this section; and Ayer notes Carnap's proposal to meet it by means of the principle of total evidence. But while granting that this principle seems to accord, to some extent, with common sense, Ayer questions the possibility of justifying it 'on Carnap's principles', precisely because a true probability statement concerning *h* which is not based on the total evidence is no less analytic than is one that does meet the requirement of total evidence.

But this demand for a justification of the total-evidence requirement in terms of the principles of inductive logic is beside the point; for, as Carnap notes, the principle of total evidence 'is not a rule of inductive logic, but of the methodology of induction'.[2]) More explicitly, we might say that the principle specifies a necessary, though not sufficient, condition for the rationality of inductive beliefs and decisions. Certain conditions of rationality can be formulated also for the application of deductive reasoning (though, as we noted, the requirement of total evidence is trivially satisfied in this case); for example, rationality of belief requires that if a set of statements is accepted as presumably true, or as expressing presumably true beliefs, then any logical consequence of that set must be accepted as well. This is not a principle of formal logic, of course: formal logic will tell us that if a given set of statements is true then such and such other statements are true as well: but it does not tell us what statements to believe or to act on; indeed, the notion of accepting certain statements, like that of total evidence, is pragmatic in character and cannot be defined in terms of the concepts of formal (deductive or inductive) logical theory.

But if the requirement of total evidence cannot be justified by the principles of formal inductive logic, on what grounds can it be advocated? One might well say that the requirement is simply a partial explication

[1]) Ayer, *loc. cit.*, p. 14.

[2]) Carnap, *Logical Foundations of Probability*, p. 211. On Carnap's conception of the methodology of induction, see also pp. 202–205 of the same work.

of the conditions governing rational belief and rational choice. Thus, Carnap constructs an example in which the requirement is violated and rightly points out that everybody would regard this violation as a serious mistake in inductive reasoning.[1]) It might be added, in the same vein, that if we allowed ourselves to depart from this requirement, we would sometimes be led to give high credence to statements which the available evidence told us were false. For example, we might give high credence to the generalization 'Any egg that hatches yields a chicken' as a result of limiting our evidence to that subset of our total evidence *e* pertaining to hens' eggs only, and thus disregarding further information, also included in *e*, about birds hatched from other kinds of eggs, which would show our generalization to be false. And while, of course, it is to be expected that inductive arguments from available evidence will sometimes lead us to give high credence to statements which, unbeknownst to us, are in fact false, rationality surely demands that high credibility must not be assigned to a statement that is known to be false, or, more precisely, to a statement that is logically incompatible with accepted evidence statements.

The practical application of the requirement of total evidence faces considerable difficulties; for our total information is always so comprehensive and complex that it cannot be expressed in two statements having the simple form of the premisses in schemata such as (3.4), indeed, it is vastly more complex than the kind of evidence contemplated in any of the theorems of inductive logic that are now available. But, as Carnap notes, a theorem of inductive logic – and any such theorem provides a schema for valid inductive arguments – 'can nevertheless be applied indirectly, provided the additional knowledge is, at least approximately, irrelevant for the hypothesis in question'.[2]) I have tried to show in another article [3]) that empirical science does indeed present us with various explanatory and predictive arguments of a fairly simple statistical character which meet the requirement of total evidence at least in an intuitively clear sense.

As was noted earlier, Toulmin, too, invokes a principle of total evidence: While he insists that all quasi-syllogisms in his sense are valid – their

[1]) Carnap, *On the Application of Inductive Logic*, p. 139.
[2]) Carnap, *loc. cit.*, p. 494.
[3]) Hempel, 'Deductive-Nomological *vs.* Statistical Explanation', sec. 11.

validity, like that of deductive syllogisms is said to be 'manifest' and 'surely not open to doubt'[1] – he later remarks that 'quasi-syllogisms can properly be advanced only if the initial data from which we argue state all that we know of relevance to the question at issue: if they represent no more than a part of our relevant knowledge, we shall be required to argue not categorically but bypothetically – 'Given only the information that Petersen is a Swede, we might conclude that the chances of his being a Roman Catholic were slight ...''[2]) It is not made very clear what is meant by validity here nor in what sense and for what reasons a quasi-syllogism, though valid, 'can properly be advanced' only if it meets the requirement of total evidence. The latter part of the passage just quoted seems to suggest that in Toulmin's opinion the conclusion of a quasi-syllogism (including its qualifier 'probably', 'almost certainly', etc.) can be unconditionally asserted if that requirement is met. But then his allegedly non-relative probability statements would seem to amount to elliptically stated relative probability statements referring to the total evidence available: and on this construal, his quasi-syllogisms would normally be invalid, as was shown in section 3. But, as we have noted, Toulmin rejects the interpretation of his probability statements as elliptic and holds instead that the support which the total evidence gives to a hypothesis provides the best *estimate* of *the* probability of the hypothesis. But this leaves us with the question what it is that is supposedly estimated in this manner; and, as was mentioned earlier, it remains obscure precisely what meanings Toulmin attributes to such locutions as '*h* is almost certain', '*h* is probable', 'the probability of *h*', and 'the client's 'real' chance of living to eighty'.[3])

5. INCONSISTENCIES GENERATED BY ELEMENTARY INDUCTION RULES

Let us now turn to another class of presumptive induction rules that generate inconsistencies. These rules are of special interest because they are widely thought to represent the most elementary and fundamental modes of inductive reasoning: we will therefore refer to them as 'elementary induction rules'.

[1]) Toulmin, *loc. cit.*, pp. 131, 132.
[2]) Toulmin, *loc. cit.*, p. 140.
[3]) Toulmin, *loc. cit.*, p. 71.

Here are two examples, the first of which expresses the presumptive form of inductive reasoning by simple enumeration:

(5.1) To argue from *All examined instances of A's have been B* to *All A's are B.*[1])

(5.2) If among the n observed instances of A's, m have been found to be instances of B, expect that m/n A's are B. Meanwhile, however, continue to search for further instances of A and constantly modify the estimated ratio (m/n) as new data accumulate.[2])

Suppose now that in order to ascertain how a certain physical magnitude y (e.g., the length of a metal bar) varies with another physical magnitude x (e.g., the temperature of the bar), the associated values of x and y have been measured in n cases: let $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ be the pairs of associated values thus established. Then the n points whose Cartesian co-ordinates are given by these number pairs can be connected by infinitely many different curves $C_1, C_2, \ldots$, each of which represents the values of y as a certain function of the values of x; let us say, $y = F_1(x)$; $y = F_2(x)$; ... Now let A be the class of all the pairs of physically associated values of the magnitudes x and y; then, on our assumptions, it is true to say:

(5.3) All of the n examined instances of A's satisfy the formula '$y = F_1(x)$'.

Hence, rule (5.1) directs us to infer the general law

(5.3a) All A's satisfy the formula '$y = F_1(x)$'.

But on our assumptions, it is equally true to say:

(5.4) All of the n examined instances of A's satisfy the formula '$y = F_2(x)$'

which, by (5.1), yields the conclusion

(5.4a) All A's satisfy the formula '$y = F_2(x)$';

and so forth.

Thus, on the basis of the same empirical data, namely, the n measurements of physically associated values of x and y, the rule (5.1) yields infinitely many different presumptive laws, each representing y as a certain mathematical function of x. Furthermore, since no two of the considered functions are identical, there are certain values of x to which F_1 and F_2,

[1]) Black, *The Inductive Support of Inductive Rules*, p. 196.

[2]) Black, M. ''Pragmatic' Justifications of Induction', in Black, *Problems of Analysis*, p. 164.

for example, assign different values of y; hence, the generalizations (5.3a) and (5.4a) are logically incompatible with each other; and so are any other two of the generalizations obtainable by means of (5.1).

The rule (5.2) yields inductive consistencies in the same way. To see this, it suffices to note that (5.2) yields (5.1) for the case where $m = n$; but inconsistencies can also be shown to arise when m is less than n.

Essentially the same argument applies to Reichenbach's basic rule of induction:

(5.5) If an initial section of n elements of a sequence x_1 is given, resulting in the frequency f^n, and if, furthermore, nothing is known about the probability of the second level for the occurrence of a certain limit p, we posit that the frequency $f^i (i > n)$ will approach a limit p within $f^n \pm \delta$ when the sequence is continued.[1])

Indeed, let the initial segment consist of our pairs (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) obtained by measurement. Among them, the relative frequency of those exhibiting the functional relationship F_1 is 1; but so is the relative frequency of those pairs exhibiting the functional relationships F_2, F_3, and so forth. Hence, assuming that nothing is known as yet about what Reichenbach calls second-level probabilities, the rule directs us to posit that if the measurement of physically associated values of x and y is continued beyond the initial n cases, the proportion of pairs conforming to F_1 will approach a limit which falls within $1 - \delta$; and that the same is true of the proportion of pairs conforming to F_2, F_3, and so forth. And though it is not the case that each of these limit statements is logically incompatible with each of the others, it can readily be seen that there still are infinitely many pairs of logically incompatible statements among the posits thus obtained. Thus, rule (5.5), too, leads from true premisses to a logically inconsistent set of conclusions.

The inconsistencies here noted are of significance also for the idea that all inductive reasoning presupposes a principle of the uniformity of nature which, when used as a supreme major premiss, can turn inductive arguments into deductive or 'quasi-deductive' ones.[2]) It is well-known

[1]) Reichenbach, H. *The Theory of Probability* (University of California Press, 1944), p. 446.

[2]) The idea, which is familiar from Mill's work, has recently been advocated, for example, by H. G. Alexander in his contribution to the symposium 'Convention, Falsification and Induction' in *The Aristotelian Society*, Supplementary Volume 34 (London, 1960). Alexander stresses, however, that several such presuppositions are

that attempts to give a suitable formulation of the principle in question encounter serious difficulties. The statement, for example, that what has happened in the past will, under the same circumstances, happen again in the future, is clearly inadequate: If this principle is understood to require full identity of all attending circumstances, then the same circumstances simply do not recur, and the rule is inapplicable; if sameness of only the 'relevant' circumstances is required, the principle is a truism, for any apparent departure from it can then be attributed to a difference in some relevant factor not recognized as such. A formulation which avoids these shortcomings and which also would seem to express much more precisely the intent of the uniformity principle is this:

(5.6) A generalization which has been borne out in all instances so far examined will be borne out also in all further instances.

But this principle is self-contradictory. For when applied to our example, it implies that all the pairs of physically associated values of x and y satisfy the formula '$y = F_1(x)$', but also the formula '$y = F_2(x)$', and so forth, since the n pairs so far measured satisfy all of those formulas.

The method we used to generate inconsistencies by means of elementary induction rules is akin to that employed by Goodman in posing his 'new riddle of induction'.[1] One of the examples characteristic of his approach is this: Suppose that according to our total evidence at a certain time t, all emeralds that have been examined so far (i.e., before t) are green. Then, according to standard conceptions of confirmation, the total evidence supports the generalization h_1: 'All emeralds are green'. Now let 'grue' be a predicate that applies to objects examined before t just in case they are green and to other objects just in case they are blue. Then, according to the total evidence at t, all emeralds observed so far are grue; hence, the total evidence also supports the generalization h_2: 'All emeralds are grue'. But when applied to emeralds examined after t, the two hypotheses thus supported yield the conflicting predictions that all those emeralds will be green, and that they will all be grue and hence blue. Goodman remarks: 'Thus although we are well aware which

involved in inductive reasoning, and he suggests that if these are taken into account, inductive reasoning in science would take a 'quasi-deductive form: 'Quasi-deductive' because it is impossible to state these presuppositions in a completely precise form.' (*loc. cit.*, p. 140).

[1]) See Goodman, N. *Fact, Fiction, and Forecast* (Harvard University Press, 1955) pp. 73ff.

of the two incompatible predictions is genuinely confirmed, they are equally well confirmed according to our present definition.'[1]) He suggests that the total evidence genuinely confirms h_1 rather than h_2 because the former is a 'lawlike statement' (i.e., has the characteristics of a law except for possibly being false), whereas the latter is not; and because only a statement that is lawlike is capable of receiving confirmation from its established instances. Thus, there arises the new riddle of induction, namely the problem of stating clearly 'what distinguishes lawlike or confirmable hypotheses from accidental or non-confirmable ones'.[2]) Goodman notes that only to the extent that this problem is solved can we make a distinction between valid and invalid inductive inferences; and he then outlines his 'theory of projection' which distinguishes between confirmable and non-confirmable hypotheses in terms of the 'entrenchment' of the predicates used in their formulation.[3])

While Goodman couches his discussion in terms of confirmation rules, it is readily seen that his hypothesis-pairs can also be used to show that the elementary induction rules mentioned above can lead from a consistent body of total evidence to an inconsistent set of conclusions: herein lies the affinity between Goodman's argument and the one we used at the beginning of the present section. The latter, however, seems to add a new facet to the important problem raised by Goodman. For one may well be inclined to agree that a generalization such as 'All emeralds are grue' is not lawlike, and that its applicability to as yet unexamined cases is not attested to by its previously established instances; but among the conflicting generalizations obtainable in the manner of our earlier example, there are many which would seem to be equally lawlike, and thus equally capable of confirmation by their instances; and if this is so, then none of these incompatible generalizations would be ruled out by restricting permissible inductive conclusions to lawlike statements. (And indeed, Goodman's rules are not intended to arbitrate between well-confirmed but incompatible lawlike hypotheses.)

Suppose, for example, that the pairs of associated values of x and y measured so far are: $(0, -1)$; $(1, 0)$; $(2, 1)$. These satisfy the following generalizations, among others:

[1]) Goodman, *loc. cit.*, p. 75.
[2]) Goodman, *loc. cit.*, p. 80.
[3]) Goodman, *loc. cit.*, ch. IV.

(5.7) $y = (x - 1); y = (x - 1)^3; y = (x - 1)^5; \ldots$

$$y = \cos \pi \left(1 - \frac{x}{2}\right); \; y = (x - 1)^2 \cos \pi \left(1 - \frac{x}{2}\right);$$

$$y = (x - 1)^4 \cos \pi \left(1 - \frac{x}{2}\right); \; \ldots$$

Each of these pairwise incompatible generalizations represents, I think, a perfectly good lawlike statement, capable of confirmation by established instances. Hence by restricting the use of our elementary induction rules to the cases where the conclusion is a lawlike statement, we may well eliminate inductive inconsistencies of the kind constructed by Goodman, but we will still be left with inconsistent sets of hypotheses of the kind illustrated by (5.7).

In philosophical discussions of the justifiability of inductive procedures, rules of the kind considered in this section are often treated as essentially adequate, if perhaps somewhat oversimplified, formulations of norms of inductive reasoning: [1]) we now see that, whatever the merits of the problem of justification may be, that problem does not even arise for those elementary induction rules; for they lead into logical inconsistencies and thus violate what surely is the very minimum requirement that any proposed rule of scientific procedure must meet before the question of its justification can be raised.

Are the inconsistencies here encountered attributable again to a violation of the requirement of total evidence? At first glance, this seems implausible; for the rules (5.1) and (5.2), as well as the principle (5.6), include what appears to be a simple version of that requirement, namely, the proviso that the given information must cover all the instances so far examined; and rule (5.5) may be understood as presupposing that condition as well. And in the examples just considered of contradictions generated by elementary induction rules, the proviso in question was always assumed to be satisfied.

However, as is illustrated by the paradoxes of confirmation [2]), the concept

[1]) Black, for example, formulates his various rules of induction in order to provide a clear characterization of the principles or policies whose justifiability is in question; and Reichenbach's ingenious argument aimed at a justification of induction deals specifically with his rule considered above, which is held to represent the fundamental principle of inductive procedure.

[2]) Cf. Hempel, C. G. 'Studies in the Logic of Confirmation', *Mind* 54: 1–26 and 97–121 (1945); especially, pp. 9–14.

of the 'instances' of a non-singular hypothesis is by no means as clear as it may seem, and there are good reasons to think, therefore, that the requirement of total evidence cannot be adequately expressed by the condition that the evidence must include all the instances so far observed. And indeed, in reference to two other examples constructed by Goodman, Carnap has argued that they do involve a violation of the requirement of total evidence.[1]) For the example mentioned above, Carnap's objection would take this form: In the case of the prediction that the next emerald will be grue, *more* is known than that the emeralds so far observed were all grue, i.e., that they were either examined before t and were green or were not examined before t and were blue: it is known that they were all examined before t. And failure to include this information in the evidence violates the requirement of total evidence.

But an inductive logic constructed in accordance with Carnap's conception would avoid our inconsistencies for yet another reason: according to that conception, as was pointed out in section 3 above, an inductive argument must be construed as showing that the information given in the evidence, which forms the 'premisses', lends more or less strong inductive support to the 'conclusion'; and thus construed, inductive inference does not lend itself to the categorical establishment of the conclusion even if the premisses are known to be, or are accepted as, true statements. Hence, the possibility of positing or accepting incompatible statements as the result of inductive inferences does not arise.

But perhaps, in an inductive logic thus conceived, the difficulty posed by the inconsistencies would simply appear in a different form? For example, if the information on the many emeralds observed so far shows them all to have been both green and grue, does it not stand to reason that this information should confer a high probability on each of the two incompatible predictions 'the first emerald examined after t will be green' and 'the first emerald examined after t will be grue'? Again, the answer is in the negative. For as a consequence of the basic postulates for inductive probability, the sum of the probabilities which a logically consistent set of statements – e.g., the total evidence at t – confers upon

[1]) Cf. Goodman, N. 'A Query on Confirmation', *The Journal of Philosophy* 43: 383–385 (1946); Carnap, 'On the Application of Inductive Logic', sec. 3; and Goodman's reply, 'On Infirmities of Confirmation Theory', *Philosophy and Phenomenological Research* 8: 149–151 (1947).

two logically incompatible hypotheses is at most 1; hence, if one of the probabilities is close to 1, the other must be close to 0.

6. ON RULES OF RATIONAL DECISION AND BELIEF

The elementary induction rules considered in the previous section construe inductive reasoning as leading to the acquisition of new statements on the basis of given ones. In this respect, they accord well with the familiar conception that inductive procedures, at the common sense and at the scientific levels, lead to the *acceptance* of certain empirical hypotheses on the basis of evidence that gives them more or less strong, but not, as a rule, logically conclusive, support. The body of scientific knowledge at a given time would then be represented by the set of all statements accepted by science at that time. Membership in this set would be granted to a hypothesis, however well confirmed, only until further notice, i.e., with the understanding that the privilege may be withdrawn if evidence unfavorable to the hypothesis should appear in the future.

The rejection of our elementary induction rules thus naturally suggests the question whether there is not some consistent alternative way of construing this conception of scientific knowledge and, more specifically, the notion of rules authorizing the addition of sufficiently supported 'new' empirical hypotheses to the set of previously accepted ones. This question clearly belongs to what Carnap calls the methodology of induction: it concerns the application of inductive logic to the formation of rational beliefs. It seems of interest, therefore, to inquire whether the question might not be treated as a special case of another, very general, problem of application which has received a great deal of attention in recent years, namely, the problem of formulating rules for rational choice or decision in the face of serveral alternatives: the acceptance of a hypothesis might then be construable as a case of theoretical choice between alternative hypotheses.

The problem of rational decision rules has recently been dealt with in the statistical theory of decision functions and the theory of games, which do not make use of the concept of inductive probability, and it has also been investigated from the point of view of inductive logic. Here, I will limit myself to a brief consideration of Carnap's approach to the question. On the assumption that a system of inductive logic in Carnap's sense is available, the problem of rational choice can be posed in the following schematic form: An agent X has to choose one out of n courses of action, $A_1, A_2, \ldots, A_n$, which, on his total evidence e, logically exclude each other and jointly exhaust all the possibilities open to him. The agent contemplates a set $O_1, O_2, \ldots, O_m$ of different possible 'outcomes' which, on e, are mutually exclusive and jointly exhaustive (i.e., e logically implies that exactly one of these outcomes will come about). Then, for any one of those actions, say A_j, and any one of those outcomes, say O_k, the given system of inductive logic determines a probability for the hypothesis that, given e, A_j will lead to the outcome O_k. Indeed, if a_j and o_k are statements describing A_j and O_k respectively, that probability is given by $c\,(o_k, e \cdot a_j)$.

What course of action it is rational for X to choose in the given circumstances will depend, of course, on what his objectives are; or, putting it more broadly, what value or disvalue he attaches to the various outcomes that might occur as a result of his action. In many theoretical studies of rational decision-making and in particular in Carnap's treatment of the problem, it is assumed that the values and disvalues in question can be represented by a quantitative concept of utility, i.e., a function u assigning to each possible outcome O_k a real number $u(O_k)$, or briefly u_k, which indicates the utility of outcome O_k for X. The task of specifying operational criteria for this concept of utility – i.e., in effect, of specifying methods of measuring the utilities of possible 'outcomes' (which may be very complex) for a given person – raises difficult problems, which have been the object of much theoretical and experimental work in recent years;[1] in the present context however, we need not enter into these issues.

The problem to be solved now calls for the formulation of a general decision rule such that, given any e and any set of A_j and O_k which meet the conditions mentioned above, and given also the utilities attached to the O_k, the rule will determine which of the available courses of action it is rational to adopt in the given circumstances. Carnap has proposed a rule which directs the agent to choose an action which offers him the highest expectation of utility. The expectation value, or the probability-

[1] For details and further bibliographic references see, for example, Carnap, *Logical Foundations of Probability*, sec. 51; von Neumann, J. and Morgenstern, O., *Theory of Games and Economic Behavior* (Princeton University Press, 2nd ed., 1947); Savage, L. J. *The Foundations of Statistics* (New York: Wiley, 1954), ch. 5; Luce, R. D. and Raiffa, H. *Games and Decisions* (New York: Wiley, 1957), ch. 2; Braithwaite, R. B. *Scientific Explanation* (Cambridge University Press, 1953), ch. VII.

estimate, of the utility associated with action A_j is given by the formula

(6.1) $u'(A_j, e) = c(o_1, e \cdot a_j) \cdot u_1 + \ldots + c(o_m, e \cdot a_j) \cdot u_m,$

and Carnap's rule may be stated as follows:

(6.2) *Rule of maximizing the estimated utility:* In the specified circumstances, choose a course of action for which the estimate of the resulting utility is a maximum, i.e., is not exceeded by the utility estimates associated with any of the alternative courses of action.[1])

In an attempt to apply this maxim to the problem of acceptance rules for scientific hypotheses, let us suppose now that a scientist has at his disposal the set of all statements accepted by science at the time, which we may assume to be expressed in the form of one complicated sentence e; that he has invented, or has been presented with, a set of n hypotheses, $h_1, h_2, \ldots, h_n$, which, on e, are pairwise incompatible while jointly exhausting all possibilities (i.e., e logically implies the negation of the conjunction of any two of the hypotheses, as well as the disjunction of all of them); and that he has to choose one from among the following $n + 1$ courses of action: To accept h_1 and add it to e; ...; to accept h_n and add it to e; to accept none of the n hypotheses and thus to leave e unchanged. The problem is to construct a rule that will determine which choice it is rational to make. Clearly, this approach to the problem of rules for rational inductive acceptance does not involve the kind of narrowly inductivist conception of scientific research which, though hardly espoused nowadays, has been made a flogging horse by some writers on scientific procedure; more specifically, we are not envisaging a rule which, given some empirical evidence, will make it possible inductively to infer 'the', or even a, hypothesis or theory that will account for, or explain, the given evidence. Rather, it is assumed here that a set of rival hypotheses have been proposed: the construction of such hypotheses requires, in general, scientific inventiveness and, in important cases, great genius; it cannot be achieved by the mechanical use of induction rules. The inductive problem here considered is rather that of deciding, on the available evidence – which may include the results of extensive tests –, which, if any, of the proposed hypotheses is to be 'accepted' and thus to be added to the *corpus* of scientific knowledge.

Now, Carnap's decision principle (and analogously also such policies

[1]) Cf. Carnap, *Logical Foundations of Probability*, p. 269.

as the minimax principle developed in the theory of games and statistical decisions [1])) requires, as a basis for a rational decision, a specification both of the total evidence and of the utilities attached to the various possible outcomes of the contemplated actions. In our case, the possible outcomes may be described as: enlarging e by h_1, where h_1 is true; enlarging e by h_1 where h_1 is false; . . . ; enlarging e by h_n where h_n is true; enlarging e by h_n where h_n is false; leaving e unchanged. What utilities are we to assign to these outcomes? This much is clear: the utilities should reflect the value or disvalue which the different outcomes have from the point of view of pure scientific research rather than the practical advantages or disadvantages that might result from the application of an accepted hypothesis, according as the latter is true or false. Let me refer to the kind of utilities thus vaguely characterized as *purely scientific*, or *epistemic, utilities*.

Construing the proverbial 'pursuit of truth' in science as aimed at the establishment of a maximal system of true statements, we might try as a first step to measure the utility of adding a hypotheses h to e in terms of the strength of that part of the information contained in h which is not contained in e, and which thus goes beyond what has been previously established. This 'new' information contained in h is expressed by the sentence $h \vee -e$. (For h is equivalent to $(h \vee e) \cdot (h \vee -e)$; the first of the two conjoined sentences follows from h as well as from e and thus represents information given by h as well as by e; the second of the conjoined sentences follows from h and thus expresses part of the information given by h, but it has no content in common with e since its disjunction with e is a logical truth.)

To represent the amount, or the strength, of the information given by a sentence, we use the concept of a *content measure* for the sentences of a (suitably formalized) language L. By such a content measure, we understand any function m which assings to every sentence s of L a number $m(s)$ in such a way that (*i*) $m(s)$ is a number in the interval from 0 to 1, inclusive of the endpoints; (ii) $m(s) = 0$ just in case s is a logical truth; (iii) if s_1 and s_2 have no common content – i.e., if the sentence $s_1 \vee s_2$, which expresses their common content, is a logical truth – then $m(s_1 \cdot s_2)$

[1]) Cf. Carnap's remarks in sec. 98 of *Logical Foundations of Probability*, and the literature listed in note 1, p. 37.

$= m(s_1) + m(s_2)$. Content measures in this sense can readily be constructed for certain kinds of formalized languages.[1])

Suppose now that m is a content measure for a formalized language suitable for the purposes of empirical science. Then we might tentatively set the utility of adding h to e equal to $m(h \vee - e)$ is h is true, and equal to $- m(h \vee - e)$ if h is false. More generally, taking account of the principle of diminishing marginal utility, we might set the utility of adding h to e directly proportional to the amount of new information provided by h, or to the negative value of that amount, according as h is true or false; and inversely proportional to the amount of information already contained in e. This would yield the following definition:

(6.3) *Relative content measure of purely scientific utility:* The purely scientific utility of adding h to e is $k \cdot m(h \vee - e)/m(e)$ when h is true, and the negative of this value when h is false; k being some positive constant.

Now it can be shown [2]) that if this utility measure is adopted – no matter which of the many possible measure functions m might be – then Carnap's principle of maximizing the estimated utility yields the following decision rule for the case, characterized above, of a choice between the $n + 1$ alternatives of accepting $h_1, \ldots$, accepting h_n, and accepting none of the alternative hypotheses:

(6.4) *Acceptance rule based on relative content measure of utility.* Of the n hypotheses, at most one can have a probability on e which exceeds $\frac{1}{2}$; if there is one, accept it. Otherwise, there may be at most two hypotheses with a probability of $\frac{1}{2}$; in this case, accept one of these, or, alternatively, accept none of the n hypotheses. Finally, if each of the n hypotheses has a probability of less than $\frac{1}{2}$ on e, accept none of them. (In the first case, the estimated utility will be positive, in all other cases, zero.)

[1]) For specific examples see Hempel, C. G. and Oppenheim, P. 'Studies in the Logic of Explanation', *Philosophy of Science* 15: 135–175 (1948), especially secs. 8 and 9; and Carnap, R. and Bar-Hillel, Y. 'An Outline of a Theory of Semantic Information'. Massachusetts Institute of Technology, Research Laboratory of Electronics. Technical Report No. 247 (1952). As background, see also Carnap, *Logical Foundations of Probability*, sec. 73.

[2]) The proof, which will be omitted here, is a generalization of the argument used to establish a more limited result in sec. 12 of my essay 'Deductive-Nomological *vs.* Statistical Explanation'.

Thus, if epistemic utility is construed in the manner of (6.3), then Carnap's general principle of maximizing the estimated utility yields a rule which makes the acceptance of one or none of the *n* rival hypotheses depend solely on the probabilities which these hypotheses possess on the total evidence *e*. This rule cannot lead into inductive inconsistencies since the accepted hypothesis must have a probability of at least $\frac{1}{2}$ on the total evidence and thus cannot be incompatible with the latter: and the total evidence, it will be recalled, represents in our case the set of *all* statements accepted in science at the time. Nevertheless, rule (6.4) is unsatisfactory; in particular, it is much too lenient to be suitable as a general rule of scientific procedure. This must not be taken to prove, however, that Carnap's rule for rational choice simply cannot yield a reasonable acceptance rule for scientific hypotheses: quite likely, our crude definition of epistemic utility is at fault.

And indeed, apart from providing true or false new information, the addition of a hypothesis *h* to *e* has other aspects which are of importance to pure science, and which have to be taken into account in an attempt to define a concept of purely scientific utility. For example, if *h* has the character of a general law or of a theoretical principle, its explanatory power with respect to relevant data included in *e* will strongly influence the potential utility of accepting *h*. A closely related factor would no doubt be the gain in logical simplicity which would accrue to the total system of accepted statements as a result of incorporating *h* into it. If factors such as these are to be taken into account they will have to be given clear and precise definitions. Some initial steps towards this end have been taken in recent years,[1]) but a great deal of further work is needed if a reasonably adequate general concept of epistemic utility is to be attained.

The approach just outlined to the problem of inductive acceptance rules construes the formation of rational empirical belief and the establishment of scientific knowledge as involving the use of certain inductive principles which, under specified conditions, authorize the (provisional) acceptance

[1]) For a definition of the explanatory power of hypotheses expressible in certain simple kinds of formalized languages, see Hempel and Oppenheim, *loc. cit.*, secs. 8 and 9. On the subject of simplicity in the sense here referred to, see Popper, K. *The Logic of Scientific Discovery* (London: Hutchinson, 1959), ch. VII and *passim;* and cf. also the lucid discussion and tentative explication in Barker, S. *Induction and Hypothesis*, where further bibliographic references, especially to the work of Kemeny, will be found.

of a hypothesis on a given body of total evidence, rather than simply determine its degree of confirmation. As an alternative, it would be interesting to investigate possible ways of construing the logic of rational belief and of scientific knowledge without assuming acceptance rules: the only inductive principles invoked in such a construal would then be, broadly speaking, probabilistic in character; for example, they might be inductive rules of the kind envisaged by Carnap and might take such forms as (3.4) and 3.5);[1]) or they might be of some different character, perhaps in accordance with statistical decision theories.

In fact, it has recently been argued, especially by writers on statistical decision procedures and on the theory of games, that it makes no clear sense to speak of the acceptance of a scientific hypothesis *per se*, without specification of a course of action to be based on it; and that, in particular, what in decision theory is sometimes briefly referred to as the acceptance of a given hypothesis always amounts to the adoption of a certain course of action. On this view, one would have to construe the notion of scientific knowledge without using the idea of acceptance at all; or, at best, one would have to construe acceptance as a pragmatic concept that has no counterpart in the logic of science. However, as I have tried to show elsewhere,[2]) this view, though supported by some very plausible arguments, faces difficulties of its own.[3])

At present, it seems to me an open question whether the idea of inductive acceptance of a hypothesis in pure science can be given a clear and methodologically illuminating construal, and correlatively, whether there are any good reasons for preserving the familiar notion of scientific induction rules that authorize the acceptance of a hypothesis on the basis of suitable evidence. For the further clarification of these issues, it will be necessary to elaborate more fully and precisely the alternative

[1]) In this connection, cf. Carnap's remarks in *Logical Foundations of Probability*, p. 206.

[2]) Cf. Hempel, 'Deductive-Nomological *vs.* Statistical Explanation', sec. 12.

[3]) The considerations here outlined seem to me to cast doubt upon the view that the question "whether to *accept* a certain hypothesis, whether to *believe* it – is... easier to answer than the question of whether to *act upon* it". This view is set forth by R. Chisholm in his book *Perceiving: A Philosophical Study* (Cornell University Press, 1957), pp. 10-11 (italics the author's). Part I of this book, entitled "The Ethics of Belief", contains many illuminating observations on issues discussed in the present essay.

conceptions of scientific knowledge briefly considered in this section: and this calls for additional philosophical analysis in Carnap's sense, aimed at an exact logical explication [1]) of the concepts central to the problem.

Princeton University

[1]) See Carnap, *Logical Foundations of Probability*, ch. I.

W. STEGMÜLLER

EINIGE BEITRÄGE ZUM PROBLEM DER TELEOLOGIE UND DER ANALYSE VON SYSTEMEN MIT ZIELGERICHTETER ORGANISATION

1. KAUSALITÄT UND TELEOLOGIE

Nach herkömmlicher Anschauung ist die teleologische oder finale Betrachtungsweise von der kausalen abzugrenzen. Während im letzteren Falle eine zweckfreie Wirklichkeitserkenntnis vorliegt, wird im ersten Fall die Welt oder ein bestimmter Ausschnitt aus ihr unter Zweckgesichtspunkten beschrieben und erklärt. Da der Begriff der Teleologie somit als eine Art von Gegenbegriff zu dem der Kausalität konstruiert werden soll, empfiehlt es sich zwecks Präzisierung des Verhältnisses von kausaler und teleologischer Betrachtungsweise, zunächst die Zusammenhänge aufzusuchen, in denen von Kausalität die Rede ist; denn die verschiedenen mit der Kausalität zusammenhängenden Begriffe sind in den bisherigen Untersuchungen schärfer bestimmt worden als die analogen Begriffe, die der teleologischen Auffassungsweise entsprechen.

Von Kausalität ist hauptsächlich in drei Kontexten die Rede: Bestimmte *Gesetze* werden als kausale ausgezeichnet, ferner werden kausale *Erklärungen* von nichtkausalen unterschieden und schließlich wird bisweilen in naturphilosophischen Betrachtungen das *allgemeine Kausalprinzip* zur Diskussion gestellt.[1])

Ein scheinbar vierter Typus von Kontexten: die alltäglichen *singulären Kausalurteile* ('das Ereignis B ist durch die Umstände A verursacht worden'), wird am besten so interpretiert, daß es sich dabei um eine primitive Vorstufe von kausalen Erklärungen handelt. Der vorwissenschaftliche Charakter dieser Erklärungen äußert sich erstens darin, daß die beteiligten Gesetze in der Regel nicht nur ungenau, sondern überhaupt nicht angegeben werden, sowie zweitens darin, daß die für das zu erklärende Ereignis relevanten Bedingungen nicht näher analysiert, sondern unter dem unklaren Begriff der Ursache summarisch zusammengefaßt werden. Dieser vierte Typus kann daher außer Betracht bleiben. Ebenso

[1]) Für nähere Details vgl. [11] S. 173ff.

dürfen wir im gegenwärtigen Zusammenhang das allgemeine Kausalprinzip unberücksichtigt lassen. Der diesem Prinzip zugrunde liegende Gedanke kann approximativ so formuliert werden, daß grundsätzlich alle Ereignisse der Welt in zutreffender Weise kausal erklärt werden können. Zum Unterschied davon würde ein allgemeines Teleologieprinzip den Gedanken zum Ausdruck bringen, daß sich für alle Vorkommnisse in der Welt eine adäquate teleologische Erklärung finden lasse. In beiden Fällen ist also der fragliche Begriff auf den des betreffenden Erklärungstypus zurückführbar. Und da in dieser Abhandlung nicht das Problem der *Gültigkeit* eines allgemeinen Teleologieprinzips erörtert werden soll, sondern die Frage der *Bedeutung* teleologischer Aussagen, genügt die Beschränkung auf die Gesetzes- und Erklärungstypen.

Was den Gesetzesbegriff betrifft, so stößt man hier gleich zu Beginn auf eine Reihe von Schwierigkeiten. Zunächst ist festzustellen, daß es anscheinend bis heute noch nicht gelungen ist, ein präzises Kriterium für den Unterschied von gesetzesartigen und nichtgesetzesartigen Aussagen anzugeben.[1]) Auf diese Frage soll hier nicht eingegangen werden; denn es handelt sich dabei um kein für den Begriff der Teleologie spezifisches Problem. Es ist vielmehr von mindestens derselben Relevanz für Fragestellungen von ganz anderer Art: z.B. für das Problem der Induktion, das Problem der Explikation des Begriffs der wissenschaftlichen Erklärung oder das Problem der irrealen Konditionalsätze. Es wäre jedenfalls nicht sinnvoll, Untersuchungen über das Problem der Teleologie mit einer bisher ungelösten Frage zu belasten, die von einer viel höheren Allgemeinheitsstufe ist als das vorliegende Problem und die in ganz anderen Zusammenhängen in derselben Gestalt wiederkehrt.

Eine zweite Schwierigkeit bildet die Abgrenzung kausaler Gesetze von nichtkausalen. Beim Versuch einer solchen Abgrenzung treten zahlreiche neue Begriffe wie 'Determinismus', 'Mikrogesetze', 'quantitative Begriffe', 'Nahwirkungsprinzip' usw. auf, die ihrerseits einer Explikation bedürftig sind. Darauf kann an dieser Stelle nicht eingegangen werden.[2]) Angenommen, die Explikation all dieser Begriffe sei in befriedigender Weise erfolgt. Dann ergibt sich noch immer keine Möglichkeit einer einwandfreien Abgrenzung kausaler Gesetze von nichtkausalen Gesetzen; denn der überlieferte Sprachgebrauch ist an dieser Stelle nicht eindeutig.

[1]) Vgl. [4]; siehe auch [12].
[2]) Vgl. dazu [3] und [11], S. 179ff.

Die Abgrenzung wird daher nur durch eine Festsetzung erfolgen können. Es ist nicht erforderlich, die verschiedenen möglichen Abgrenzungen zu diskutieren. Denn für unser Problem gelangen wir sofort zu einer negativen Feststellung: Wie immer auch der Begriff des Kausalgesetzes präzisiert werden mag, die nichtkausalen Gesetze können auf keinen Fall als teleologische Gesetze bezeichnet werden, sondern sind auf eine Weise zu charakterisieren, die zu den teleologischen Begriffen in keiner Beziehung steht. Falls z.B. als auszeichnendes Merkmal von Kausalgesetzen das des Determinismus genommen wird, so sind die nichtkausalen Gesetze nichtdeterministische Gesetzmäßigkeiten, d.h. statistische Gesetze. Wenn man dagegen jene Gesetze kausal nennt, die den Charakter von Nahwirkungsgesetzen haben, so würde es sich bei nichtkausalen Gesetzen um Fernwirkungsgesetze handeln (wie z.B. die Gravitation in der von Newton konzipierten Mechanik). Bei anderen Festsetzungen würde man auf den Unterschied zwischen quantitativen und qualitativen, Mikro- und Makrogesetzen, Sukzessions- und Koexistenzgesetzen stoßen usw.

Wenn es also möglich sein sollte, den Begriff der Teleologie als eine Art von 'Gegenbegriff' zu dem der Kausalität zu bilden, so kann die Abgrenzung jedenfalls nicht auf dem Wege über den Gesetzestypus erfolgen. Die nichtkausalen Gesetzmäßigkeiten sind niemals teleologische, sondern statistische (im Gegensatz zu deterministischen) oder Fernwirkungsgesetze (im Gegensatz zu Nahwirkungsgesetzen) oder qualitative (im Gegensatz zu solchen, die in quantitativer Sprache formuliert sind) oder Koexistenz- bzw. Strukturgesetze (im Gegensatz zu Sukzessionsgesetzen) usw. Um die folgenden Betrachtungen zu vereinfachen, soll angenommen werden, daß als Unterscheidungsmerkmal für die Kausalgesetzlichkeit der Begriff des Determinismus verwendet wird. Da wir nicht teleologische Gesetze von kausalen zu unterscheiden brauchen, können die Naturgesetze im allgemeinen auch Kausalgesetze i.w.S. genannt werden, die dann in die beiden Klassen der Kausalgesetze i.e.S. oder deterministischen Gesetze und der nichtkausalen oder statistischen Gesetze zerfallen.

Für den Zweck einer präzisen Abgrenzung des Begriffs der Teleologie von dem der Kausalität verbleibt als letzte Möglichkeit daher nur der Begriff der Erklärung. Es käme also darauf an, *kausale Erklärungen* von *teleologischen Erklärungen* zu unterscheiden. Bezüglich des allgemeinen

Begriffs der Erklärung kann die Analyse von Hempel-Oppenheim [1]) zugrunde gelegt werden, wonach jede Erklärung eines Phänomens in der Angabe von geeigneten Antecedensbedingungen sowie von relevanten Gesetzen besteht, aus denen zusammen das Phänomen abgeleitet werden kann. Wieder können im gegenwärtigen Zusammenhang alle jene begrifflichen Schwierigkeiten außer Betracht bleiben, die im Rahmen einer präzisen Explikation des Begriffs der Erklärung zu bewältigen sind, da auch diese Schwierigkeiten für das Problem der Teleologie nicht spezifisch sind, sondern in derselben Weise bei der Explikation anderer Erklärungstypen auftreten. Wesentlich ist dagegen der Umstand, daß jede Erklärung eine Erklärung auf Grund von bestimmten *Gesetzen* ist. Kausale Erklärungen i.e.S. sind dann jene, bei denen die verwendeten Gesetze Kausalgesetze i.e.S. darstellen.[2]) Auf Grund der obigen Feststellung ergibt sich daraus aber unmittelbar, daß auch der Erklärungsbegriff keine Basis dafür liefert, um die Kausalität von der Teleologie abzugrenzen; denn teleologische Erklärungen müßten danach im Gegensatz zu den kausalen Erklärungen jene sein, bei denen die verwendeten Gesetze keine kausalen, sondern teleologische Gesetze sind. Da die nichtkausalen Gesetze aber zu einem der eben angeführten Typen gehören – je nach Abgrenzung des Begriffs des Kausalgesetzes zu einem anderen Typ, für unsere Wahl also zum Typ der statistischen Gesetze –, die mit Teleologie nichts zu tun haben, so wird auch dieser letzte Abgrenzungsversuch hinfällig.

Der Gedanke, Kausalität und Teleologie als einander logisch ausschließende und erschöpfende Unterfälle allgemeiner Prinzipien, Gesetze oder Erklärungen zu explizieren und in diesem Sinn die Teleologie als Gegenbegriff zu dem der Kausalität zu konstruieren, muß daher vollkommen preisgegeben werden. Insbesondere können teleologische Erklärungen, was immer man im einzelnen darunter verstehen mag, nicht im Prinzip von kausalen Erklärungen verschieden sein. Sie müssen vielmehr auf jeden Fall zugleich kausale Erklärungen i.w.S. (d.h. Er-

[1]) Vgl. [5].

[2]) Da in der Regel für eine adäquate Erklärung von Phänomenen mehrere Gesetzmäßigkeiten verwendet werden, könnte man Mischformen unterscheiden, je nachdem, ob alle verwendeten Gesetze Kausalgesetze sind ('rein kausale Erklärung') oder nur einige davon ('partiell kausale Erklärung'). Da der Gesetzestypus für das Problem der Teleologie jedoch ohne Relevanz ist, können wir von dieser feineren Unterscheidungsmöglichkeit absehen.

klärungen auf Grund von Bedingungen und Gesetzen) darstellen und u.U. sogar kausale Erklärungen i.e.S.

Ein weiterer naheliegender Gedanke, das Schema der wissenschaftlichen Erklärung der Unterscheidung zwischen Kausalität und Teleologie zugrunde zu legen, muß noch kurz erwähnt werden. In bildhaften Darstellungen des Unterschiedes zwischen kausalen Prozessen und solchen, die nur teleologisch erklärbar sein sollen, wird der kausale Vorgang häufig dadurch charakterisiert, daß die einzelnen Zustände des zugrundegelegten Systems die Folgezustände determinieren, während im Fall der Teleologie das gegenwärtige Geschehen durch das künftige Geschehen bestimmt wird. Wenn man diesen Gedanken durch Anwendung auf das Schema der Erklärung zu präzisieren versucht, so ergäbe dies folgendes: Die außer den Gesetzen in der Erklärung verwendeten Antecedensbedingungen müssen dem zu erklärenden Phänomen zeitlich folgen. Der Ausdruck 'Antecedensbedingungen' ist dann natürlich nicht mehr angemessen. Sprechen wir statt dessen von Realisationsbedingungen. Wie müßte die Anwendung eines solchen teleologischen Erklärungsschemas aussehen, etwa um ein gegenwärtiges Phänomen zu erklären? Wegen des Umstandes, daß das Explanandum den Realisationsbedingungen zeitlich vorangeht, würde die Erklärung darin bestehen, daß das zu erklärende Phänomen mit Hilfe relevanter Gesetze aus geeigneten Realisationsbedingungen, *die durch Betätigung von hellseherischen Fähigkeiten erschlossen werden müßten*, abgeleitet wird. Diese Nötigung zur Abschweifung ins Gebiet der Parapsychologie ist ein drastischer Hinweis darauf, daß auch der faszinierende Gedanke, eine Rollenvertauschung von Vergangenheit und Zukunft für die Explikation des Begriffs der Teleologie nutzbar zu machen, in Wahrheit eine Sackgasse darstellt. Das Problem, ob sich dieser Gedanke eines rein teleologischen Erklärungsschemas überhaupt in logisch konsistenter Weise durchführen läßt, braucht daher nicht weiter verfolgt zu werden.

Es verbleibt noch eine Möglichkeit, die zugleich den einzig gangbaren Weg darstellt. Darin klärt sich auch sofort das Mißverständnis auf, welches dem Bild von der 'Determination der Gegenwart durch die Zukunft' zugrunde liegt. Wenn der Unterschied zwischen teleologischen und kausalen Erklärungen weder im Gesetzestypus noch im zeitlichen Verhältnis von Explanandum und Antecedensbedingungen liegt, so kann er nur *auf der Art der Antecedensbedingungen* beruhen. Da teleologische

Erklärungen sich von den nichtteleologischen dadurch unterscheiden sollen, daß in ihnen der Zweckgesichtspunkt eine entscheidende Rolle spielt, so muß dieser Begriff des Zweckes oder des Zieles daher in bezug auf die Antecedensbedingungen zur Anwendung gelangen. Dieser Fall ist tatsächlich dann gegeben, wenn sich die beabsichtigten Erklärungen auf Vorgänge im menschlichen Bereich beziehen, soweit an diesen Vorgängen handelnde Personen beteiligt sind. Denn dann gehören zu den Antecedensbedingungen des Explanandums die Motive jener handelnden Personen und die darin enthaltenen Pläne zur künftigen Verwirklichung bestimmter Zielsetzungen. Wie von verschiedenen Autoren hervorgehoben worden ist, wird dadurch selbstverständlich nicht zeitlich Früheres durch Späteres determiniert; denn nicht die zeitlich auf das Explanandum folgenden Ereignisse, deren Verwirklichung von den Handelnden bezweckt wird (die zukünftigen Ziele) gehören zu den Antecedensbedingungen, sondern die Ziel*setzungen*, welche dem Handlungsbeginn vorangehen oder zumindest mit ihm gleichzeitig sind. Zielsetzungen brauchen ja überhaupt nicht verwirklicht zu werden – sie können durch den Gang der Ereignisse vereitelt werden –; jenes niemals stattfindende Künftige kann daher auch keine Realisationsbedingung für das Explanandum bilden. Das Vorliegen bewußter Zielsetzungen handelnder Wesen ist auch der einzige Fall, in dem von teleologischen Erklärungen gesprochen werden kann. Hinter dem Begriff einer *objektiven Zweckmäßigkeit*, welche nicht das Resultat bewußten Zweckhandelns ist – z.B. im Bereich der organischen Natur –, steckt nichts. Bereits von früheren Autoren – z.B. von Franz Brentano – ist zutreffend festgestellt worden, daß alle Aussagen über das Bestehen einer objektiven 'Zweckordnung' in der Welt mit solchen Aussagen gleichwertig seien, wonach diese Ordnung auf einen zwecksetzenden Willen zurückgeführt werden müsse.[1]) Wir stehen daher angesichts teleologischer Aussagen von der Gestalt 'der Gegenstand (Vorgang) *A* dient dem Zweck *B*' vor der folgenden Alternative: (a) Entweder diese Aussage läßt sich auf der Grundlage einer präzisen Explikation der darin enthaltenen Begriffe als intensionsgleich mit einer solchen Aussage erweisen, die eine partielle Erklärung für A enthält, zu deren

[1]) Daher bestand z.B. für Brentano die Aufgabe des sog. teleologischen Gottesbeweises nicht darin, aus der Zweckmäßigkeit der Welt auf die Existenz eines göttlichen Urhebers zu schließen, sondern in dem Nachweis, daß die *scheinbare* Zweckmäßigkeit der Welt auf einer *tatsächlichen* beruht.

Antecedensbedingungen die Motive (Ziel- oder Zwecksetzungen) handelnder Wesen gehören; m.a.W.: die vorliegende Aussage läßt sich in eine solche übersetzen, in der, statt einfach vom Zweck von *A* zu reden, explizit *Motive* oder Zweck*setzungen* angegeben werden, die zur Erklärung von *A* oder bestimmter Beschaffenheiten von *A* heranzuziehen sind. (b) Oder aber diese Aussage muß in eine nichtteleologische Aussage übersetzbar sein, in der weder über Zwecke noch über Motive handelnder Wesen gesprochen wird, so daß sich die Rede von den Zwecken des A in der ursprünglichen Aussage als eine bloße 'façon de parler' erweist.
Dieser Alternative könnte man nur dann entgehen, wenn es sich als möglich erweisen sollte, den Begriff des Zweckes ohne Bezugnahme auf ein zwecksetzendes Bewußtsein zu präzisieren, ihn also etwa nach Analogie zu den in der modernen Physik vorkommenden Grundbegriffen als einen theoretischen Begriff im Sinne Carnaps [1]) zu konstruieren. Die Verwendung des Ausdrucks 'Zweck' für eine derartige Konstruktion würde jedoch keineswegs eine Explikation des Begriffs der Teleologie liefern, sondern nichts weiter darstellen als einen Sprachmißbrauch.
Die zuletzt angestellten Erwägungen haben offenbar bedeutsame Konsequenzen für die Frage, inwieweit zur Erklärung von Vorgängen in der organischen Natur auf teleologische Betrachtungsweisen zurückgegriffen werden darf.
Zusammenfassend ergibt sich somit folgendes:
(1) Der Unterschied zwischen Kausalität und Teleologie liegt *weder* im Gesetzestypus *noch* in der Art der Anwendung des Erklärungsschemas.
(2) Teleologische Erklärungen sind dadurch ausgezeichnet, daß die Antecedensbedingungen *Motive* (Zielsetzungen) handelnder Wesen einschließen.
(3) Die Unterscheidung zwischen *objektiver Zweckbetrachtung* (z.B. in der Natur) ohne Rückgang auf Zwecksetzungen und Betrachtung unter dem Gesichtspunkt des *Zweckhandelns* ist undurchführbar, da sich teleologische Erklärungen von kausalen *nur* im Sinne von (2) unterscheiden.
(4) Kausalität und Teleologie bilden keine Gegensätze; vielmehr sind alle teleologischen Erklärungen zugleich kausale Erklärungen i.w.S.: *Teleologische Erklärungen sind kausale Erklärungen aus Motiven.*
(5) Falls die Unterscheidung von Gesetzen in deterministische und

[1]) Vgl. [2].

statistische zugrunde gelegt wird und Erklärungen mittels deterministischer Gesetze als kausale Erklärungen bezeichnet werden, so ergibt sich, daß teleologische Erklärungen sogar kausale Erklärungen i.e.S. sein können. Teleologische Erklärungen gliedern sich dann in die beiden Untergruppen: (a) *kausalteleologische Erklärungen* und (b) *statistisch-teleologische* Erklärungen. Die Zugehörigkeit zu dem einen oder anderen Typus hängt davon ab, ob sich die Motive sowie das Verhalten der in der Erklärung angeführten handelnden Wesen unter deterministische Gesetze subsumieren lassen oder ob sie nur in der Gestalt statistisch beschreibbarer Regelmäßigkeiten erfaßbar sind.

2. DIE DREI SCHICHTEN SCHEINBARER TELEOLOGIE

Auf Grund der bisherigen Betrachtungen scheint sich das ganze Problem der Teleologie in nichts aufzulösen. Denn in der Tatsache, daß es kausale oder statistische Erklärungen von Vorgängen gibt, in denen auch die Motive von handelnden Personen eine Rolle spielen, liegt weder etwas Merkwürdiges noch etwas, das jemals bezweifelt wurde.

Schwierigkeiten ergeben sich erst auf Grund der oben formulierten Alternative. Denn diese zwingt uns, für all jene Aussagen, in denen teleologische Ausdrücke und Wendungen wie 'Zweck', 'Ziel', '...dient dazu, um...' usw. vorkommen, ohne daß handelnde Wesen vorausgesetzt werden können, Übersetzungen in nichtteleologische Formulierungen anzugeben. Das wissenschaftstheoretische Problem der Teleologie kann daher mit der Aufgabe identifiziert werden, die *Regeln für diese Übersetzungen* aufzusuchen und zu formulieren.

Zu beachten ist dabei, daß dieses wissenschaftstheoretische Problem in einem gewissen Sinn an den Stand der einzelwissenschaftlichen Forschung gebunden bleibt; denn über das Zutreffen der in der obigen Wendung 'ohne daß handelnde Wesen vorausgesetzt werden dürfen' enthaltenen Voraussetzung kann nur der Fachwissenschaftler und nicht der Logiker entscheiden. Der Logiker kann allerdings durch vorbereitende Analysen diese Entscheidung erleichtern. Dies sei kurz am Beispiel des neovitalistischen Standpunktes gezeigt; denn vor allem in bezug auf die biologischen Phänomene bekommen all diese Dinge Relevanz.

Verschiedene Autoren, darunter auch Hempel-Oppenheim, haben mit Recht darauf hingewiesen, daß alle Versuche, biologische Phänomene

mit Hilfe des teleologischen Begriffs der Entelechie zu erklären, solange undiskutabel bleiben, als nicht geeignete Gesetze formuliert worden sind; denn nicht mit Hilfe von Begriffen, sondern nur mit Hilfe von Gesetzen lassen sich Phänomene und Vorgänge erklären. Gesetzt jedoch den Fall, daß es gelungen sei, solche Gesetze anzugeben und eine empirisch fundierte Theorie *T* zu entwickeln, welche auf diesen Gesetzen beruht und welche überdies den Begriff der Entelechie als zentralen Grundbegriff enthält. Dann bestehen nur zwei Möglichkeiten: *Entweder* der Begriff der Entelechie steht in bezug auf seine logische Struktur in vollkommener Analogie zu Begriffen der theoretischen Physik wie z.B. 'Gravitation', 'Neutrino' u.dgl. Da nach der früheren Feststellung kein Gegensatz zwischen teleologischen und nichtteleologischen Gesetzen konstruierbar ist, wäre dann die Theorie *T* ihrer Struktur nach von physikalisch-chemischen Theorien ununterscheidbar und die Charakterisierung des Begriffs der Entelechie oder der Theorie *T* durch das Prädikat 'teleologisch' sowie die Bezeichnung der mittels *T* vollzogenen Erklärungen als teleologischer Erklärungen hätten jeglichen Sinn verloren. Die Tatsache, daß *T* eine 'nichtmechanistische' Theorie darstellt, da ihre Grundbegriffe und Gesetze nicht auf die der Mechanik zurückgeführt werden können, würde genauso wenig eine Berechtigung dafür ergeben, *T* eine teleologische Theorie zu nennen, wie die Tatsache, daß die Grundbegriffe und Gesetze der Elektrizitätslehre nicht auf die der Mechanik zurückführbar sind, dazu führen kann, im Falle der Elektrizitätslehre von einer teleologischen Theorie zu sprechen.[1])
Oder die Theorie wird in einer solchen Weise aufgebaut, daß die mit ihr vollzogenen Erklärungen den Charakter teleologischer Erklärungen im früheren Sinne erhalten. Dann müssen offenbar die Entelechien 'bewußtseinsbegabte' Wesen sein, deren Motive (Zielsetzungen) das organische Geschehen lenken. Der Neovitalismus wäre also, *selbst unter der Voraussetzung, daß er geeignete Gesetze zu formulieren und damit eine empirisch überprüfbare Theorie aufzubauen vermöchte*, zu der Annahme gezwungen, daß das Naturgeschehen durch Geister oder Dämonen gelenkt wird, es sei denn, daß diese neovitalistische Richtung den Gedanken einer teleologischen Betrachtungsweise des organischen Naturgeschehens vollkommen preisgibt.

[1]) Diese Analogie zeigt zugleich, daß es gänzlich verfehlt ist, das vorliegende Problem durch die Alternative 'Mechanismus oder Teleologie?' zu formulieren.

Es ist kaum anzunehmen, daß ein moderner Naturforscher die letztere Möglichkeit ernsthaft ins Auge faßt. Obwohl die Entscheidung trotzdem natürlich eine Angelegenheit des biologischen Fachmannes bleibt, wollen wir daher für das Folgende voraussetzen, daß in bezug auf biologische Phänomene die teleologische Betrachtungsweise preiszugeben ist. Dann muß in allen Fällen von Beschreibungen, Analysen und Erklärungen, in denen teleologische Ausdrücke und Wendungen vorkommen, diese Teleologie als eine bloß scheinbare aufgefaßt werden. Damit wird die zu Beginn dieses Abschnittes erwähnte Aufgabe aktuell, die Regeln für die Übersetzung solcher teleologischer Aussagen in nichtteleologische aufzufinden.

Den bisher entscheidendsten Schritt in dieser Richtung dürfte E. Nagel getan haben [1]), der den Begriff der Systeme mit zielgerichteter Organisation in den Mittelpunkt der Untersuchung über teleologische Aussagen rückte und die Struktur solcher Systeme in Umrissen beschrieben hat. Im folgenden soll versucht werden, die Analyse teleologischer Aussagen weiterzuführen und eine detailliertere Charakterisierung von Systemen mit zielgerichteter Organisation zu geben.

Es dürfte sich als zweckmäßig erweisen, die Aussagen über scheinbare Teleologie in *drei Schichten* zu gliedern, die sich durch zunehmende Komplexität voneinander unterscheiden. Die Schichten stehen nicht beziehungslos nebeneinander. Vielmehr sind die Aussagen der niedrigeren Schicht in bestimmter Weise von denen der höheren Schicht abhängig. Darin liegt auch die eigentliche Schwierigkeit für die Lösung des wissenschaftstheoretischen Problems der Teleologie. Denn die Übersetzung der scheinbar teleologischen Aussagen einer niedrigeren Schicht in nichtteleologische kann jeweils nur unter einer bestimmten Voraussetzung (einer Existenzvoraussetzung oder einer genetischen Hypothese) erfolgen, deren Begründung ihrerseits eine teleologische Erklärung von höherer Ordnung zu implizieren scheint. Die erforderliche Analyse kann daher nur dadurch zum Abschluß gebracht werden, daß die Aussagen aller drei Schichten untersucht werden und ihre Übersetzbarkeit in nichtteleologische Aussagen aufgezeigt wird. Das Unbefriedigende an vielen Diskussionen des Problems der Teleologie liegt in der Tatsache, daß darin nur teleologische Aussagen von einer ganz bestimmten Schicht in Be-

[1]) Vgl. [8] und [9], insbes. [9], S. 546–548.

tracht gezogen werden: Selbst wenn es dabei gelungen sein sollte, die erforderliche Übersetzung in nichtteleologische Aussagen vorzunehmen, so bleibt doch der Eindruck bestehen, als sei das Problem der Teleologie selbst noch ganz ungelöst, da teleologische Aussagen von ganz anderem Typus nicht berücksichtigt worden waren.

Zu den scheinbar teleologischen Aussagen der untersten Schicht sollen die *elementar-teleologischen Aussagen* gerechnet werden. Dies sind Sätze, in denen über Zwecke, Funktionen oder Aufgaben von Objekten (z.B. Organen wie Herz oder anderen Bestandteilen eines Organismus wie Blut oder Chlorophyll) oder von Eigenschaften von Objekten (z.B. der Farbe eines Organismus) gesprochen wird, ohne daß dabei gleichzeitig die Existenz eines mit Bewußtsein versehenen zwecksetzenden Wesens angenommen würde. Zwei Beispiele für diesen Typus wären die Aussagen, daß die Funktion der Leukozyten darin besteht, den Organismus gegen eindringende fremde Mikroorganismen zu schützen, oder daß ein Farbmuster auf den Flügeln einer bestimmten Schmetterlingsart (welches etwa raubtieraugenähnlich sein möge) dazu diene, die Art gegen feindliche Vögel zu schützen. In Aussagen von der ersten Art wird die Erhaltung der individuellen Organismen als Zweck angegeben, in Aussagen von der zweiten Art die Erhaltung einer Species.

Diese elementar-teleologischen Aussagen können sofort in nichtteleologische Aussagen übersetzt werden, in denen nur mehr davon die Rede ist, daß etwas eine *notwendige Bedingung* für etwas anderes darstellt. Sprachlich wäre dies häufig durch irreale Konditionalsätze von der Gestalt 'wenn nicht..., so...' wiederzugeben, also z.B.: 'wenn das menschliche Blut nicht eine hinreichende Anzahl von Leukozyten enthielte, so würde der menschliche Organismus durch eindringende Mikroorganismen geschädigt oder sogar vernichtet werden', 'wenn die Einzelexemplare der Schmetterlingsspecies jenes abschreckende Farbmuster auf den Flügeln nicht enthielten, so würde die Species durch Feinde ausgerottet werden'. Es ist allerdings nicht ganz korrekt, wenn Nagel Umformulierungen von dieser Art als Übersetzungen von teleologischen in nichtteleologische *Erklärungen* bezeichnet.[1]) Selbst ungenaue Erklärungen, in denen die beteiligten Gesetze überhaupt nicht und die relevanten Bedingungen nur unvollständig angegeben werden, sind stets

[1]) [9] S. 541.

dadurch charakterisiert, daß gewisse Bedingungen als *hinreichend* für das zu Erklärende ausgezeichnet werden, ohne jedoch als *notwendig* dafür bezeichnet zu werden. Hierin liegt eine gewisse Schwierigkeit, die zugleich dazu führen kann, die so verblüffend einfache Übersetzungsmöglichkeit von elementar-teleologischen Aussagen in nichtteleologische als inadäquat zu betrachten: Die ursprünglichen teleologischen Aussagen waren nämlich als (allerdings unvollständige) *Erklärungen* intendiert. Wenn daher die Übersetzung keine Umformung von teleologischen in nichtteleologische Erklärungen liefert, sondern eine Übersetzung von teleologischen Erklärungen in Aussagen, die nicht in das Erklärungsschema hineinpassen, so könnte dies von Vertretern des Teleologiestandpunktes leicht so interpretiert werden, daß darin die Unmöglichkeit zum Ausdruck komme, teleologische Erklärungen durch kausale zu ersetzen.

Tatsächlich enthalten die elementar-teleologischen Aussagen gegenüber kausalen Erklärungen einen Bedeutungsüberschuß. Dieses Mehr an Bedeutung beruht aber nicht darauf, daß diese teleologischen Aussagen prinzipiell unübersetzbar sind in nichtteleologische, sondern haben ihren Grund darin, daß in den elementar-teleologischen Aussagen eine 'selbstverständliche' Existenzvoraussetzung steckt, die bei Vornahme der Übersetzung explizit gemacht werden muß; denn die üblichen naturwissenschaftlichen Aussagen über Kausalzusammenhänge beinhalten keine derartigen Voraussetzungen. Diese Existenzvoraussetzung besteht darin, daß die betrachteten Objekte und Eigenschaften, über deren Zwecke in elementar-teleologischen Aussagen berichtet wird, als Bestandteile von Dingen oder Systemen von Dingen mit bestimmter Organisation betrachtet werden, *zu deren Erhaltung sie beitragen*. Physiker hingegen haben keinen Anlaß, bei der Untersuchung von physikalischen und chemischen Vorgängen ein bestimmtes System zu bevorzugen und diese Prozesse auf das System zu beziehen. Daher besteht für sie auch kein Grund, diesen physikalisch-chemischen Prozessen eine 'Funktion' oder 'Aufgabe' zuzuschreiben.

Wenn es sich dabei nur um einen Unterschied des Aspektes handelte, unter dem physische Prozesse oder Eigenschaften von Dingen betrachtet werden, so wäre dies für das Teleologieproblem kaum von Relevanz. Entscheidend ist jedoch der Umstand, daß jene Systeme, deren Existenz in elementar-teleologischen Aussagen vorausgesetzt wird, ihrer Struktur und Tätigkeitsweise nach eine '*zielgerichtete Organisation*' haben. Den

Prototyp von solchen Systemen bilden die Organismen mit ihrer Fähigkeit zur Selbsterhaltung, Selbstregulation, Anpassung und Selbstreproduktion. Mit der Einführung solcher Systeme verschiebt sich das ganze Problem auf eine höhere Ebene. Denn der Vertreter des Teleologiestandpunktes wird sagen: 'Daß elementar-teleologische Aussagen sich auf die skizzierte Weise in solche übersetzen lassen, in denen keine teleologischen Ausdrücke vorkommen, kann zugegeben werden. Für das eigentliche Problem der Teleologie ist dies aber ohne jegliche Bedeutung. Denn sowohl für die ursprünglichen wie für die als Übersetzungen vorgeschlagenen Aussagen muß die Voraussetzung gemacht werden, daß Systeme mit zielgerichteter Organisation wie z.B. Organismen zugrunde liegen. Das Funktionieren dieser Systeme kann aber durch nicht-teleologische Ausdrücke nicht adäquat beschrieben und erklärt werden'. Alle jene Aussagen, welche die Funktionsweise von derartigen Systemen, die wir auch weiterhin als 'Systeme mit zielgerichteter Organisation' oder 'zielgerichtete Systeme mit Selbstregulation' (kurz: '*ZO-Systeme*') bezeichnen werden, beschreiben und erklären, rechnen wir zu den *teleologischen Aussagen der zweiten Schicht*. Eine Lösung des wissenschaftstheoretischen Problems der Teleologie setzt voraus, daß auch für diese Aussagen der Nachweis ihrer prinzipiellen Übersetzbarkeit in nicht-teleologische Aussagen erbracht wird. Dazu muß zunächst vor allem der Begriff des Systems mit zielgerichteter Organisation präzisiert werden. Der Behandlung dieses komplizierten Problems soll der folgende Abschnitt gewidmet werden. Zuvor sei aber noch eine kurze Bemerkung darüber gemacht, inwiefern teleologische Aussagen von einer noch höheren Schicht berücksichtigt werden müssen.

Angenommen, es sei gelungen, die Struktur von ZO-Systemen ohne Verwendung teleologischer Begriffe adäquat zu charakterisieren [1]). Dann

[1]) Nur die allgemeine Charakterisierung solcher Systeme kann als wissenschaftstheoretische Aufgabe betrachtet werden. Sie muß allerdings so weit führen, daß daraus die prinzipielle Vermeidbarkeit teleologischer Begriffsbildungen ersichtlich wird. Jede konkrete Analyse empirisch bekannter Spezialfälle von ZO-Systemen muß Aufgabe eines Spezialfachmannes bleiben. Es sei hier vorausgesetzt, daß sowohl die wissenschaftstheoretische Aufgabe gegenüber den Aussagen der zweiten Schicht bewältigt worden ist als auch die Spezialforschung ergeben hat, daß alle Probleme hinsichtlich empirisch bekannter Fälle von ZO-Systemen ohne einen teleologischen Begriffsapparat gelöst werden können.

kann gefragt werden, wie die *Entstehung* solcher Systeme zu erklären sei. Hier muß eine Unterscheidung gemacht werden. Sofern es sich um von Menschen erzeugte ZO-Systeme handelt – denn z.B. auch Gebilde der Technik können die Struktur von Selbstregulatoren haben –, muß die Entstehungsfrage durch eine teleologische Erklärung im früheren Sinn beantwortet werden; d.h. für die Entstehung solcher Gebilde muß eine historisch-kausale Erklärung gegeben werden, zu deren Antecedensbedingungen die Zielsetzungen und Wünsche menschlicher Personen gehören. Was die 'naturgewachsenen' ZO-Systeme betrifft, so ist diesen gegenüber die Frage nicht unmittelbar entscheidbar. Sie muß zunächst dahingehend präzisiert werden, daß es sich um das Problem der erstmaligen Entstehung solcher Systeme überhaupt handelt. Denn das Problem der Entstehung spezieller Fälle naturgewachsener ZO-Systeme durch Reproduktion seitens anderer solcher Systeme gehört noch zur zweiten Schicht: Die Reproduktion bildet ja selbst einen der Vorgänge, die an ZO-Systemen beobachtet werden können; ihre Beschreibung und Erklärung in einer nichtteleologischen Sprache ist daher ein Teil der Aufgabe, teleologische Aussagen der zweiten Schicht in nichtteleologische zu übersetzen.

Dagegen stellt die Frage, wie es überhaupt erstmals zur Entstehung von ZO-Systemen in der Welt gekommen sei, ein Problem sui generis dar. Es ist der Standpunkt denkbar, daß sowohl gegenüber den teleologischen Aussagen der ersten sowie denen der zweiten Schicht die Übersetzbarkeit in eine nichtteleologische Sprache zugegeben, gleichzeitig aber behauptet wird, daß die Entstehung von ZO-Systemen nur mittels teleologischer Begriffe (z.B. durch Heranziehung eines göttlichen Schöpfungsaktes) erklärt werden könne: 'Das Funktionieren von Organismen ist naturwissenschaftlich-kausal erklärbar, die Entstehung von Organismen jedoch nicht.' Alle Aussagen, welche die erstmalige Entstehung von ZO-Systemen beschreiben und erklären, rechnen wir zur *dritten Schicht*. Das Problem der Teleologie kann erst dann als vollständig gelöst betrachtet werden, wenn auch bezüglich dieser zur dritten Schicht gehörenden Aussagen eindeutig entschieden ist, ob sie in einer nichtteleologischen Sprache formuliert werden können. Theologische Betrachtungen, insbesondere die Diskussionen zum teleologischen Gottesbeweis, bewegen sich gewöhnlich auf dieser Ebene der Aussagen der dritten Schicht. Da sich demgegenüber die wissenschaftstheoretischen und meist auch die naturphilosophischen Erörterungen des Teleologieproblems auf die Aussagen der ersten und zweiten Schicht beziehen, mochte der Anschein entstehen, als bestehe zwischen diesen beiden Arten von Betrachtungen kein Zusammenhang.

3. ANALYSE VON ZO-SYSTEMEN

Es kommt darauf an, einen möglichst umfassenden Begriff von ZO-Systemen zu bilden, der auf alle jene Systeme anwendbar ist, die eine Art von 'Selbstregulation' besitzen. Da dieser Begriff auf Dinge anwendbar sein soll, die sich in anderen Hinsichten stark voneinander unterscheiden (naturgegebene Objekte wie Organismen, menschliche Schöpfungen wie die freie Verkehrswirtschaft, aber auch technische Gebilde wie z.B. Temperaturregulatoren), kann die Analyse nur die allgemeinsten Charakteristika solcher Systeme hervorheben. Die spezielle Art der Strukturen und Gesetzmäßigkeiten, die am Prozeß beteiligt sind, muß offen bleiben. Die Analyse muß jedoch so weit vorangetrieben werden, daß daraus ersichtlich wird: In allen Fällen, in denen zur Erklärung eines Vorganges an einem solchen System ein Erklärungsschema zur Anwendung gelangt, ist dies ein Schema vom Typus der kausalen Erklärung i.w.S., zu dessen Antecedensbedingungen keine Zwecke gehören.

Der zugrunde liegende Gegenstandsbereich soll in zwei Teilklassen S und U zerfallen. Die Elemente von U bilden die Umgebung des ZO-Systems, während S das fragliche System selbst darstellt. Das für unsere Betrachtungen wesentliche Merkmal von S soll darin bestehen, daß es eine bestimmte Eigenschaft G gegenüber Störungen, die von U herrühren und den inneren Zustand von S modifizieren, 'zu erhalten trachtet' bzw. auf die Verwirklichung von G 'hintendiert', sofern die von U herrührenden Störungen zu einem unmittelbar vorangehenden Zeitpunkt einen Zustand von S, der die Eigenschaft G besaß, in einen solchen überführte, dem diese Eigenschaft nicht mehr zukommt.

S soll aus n Teilen $S_1, \ldots, S_n$, den S-Teilen, bestehen, wobei n eine natürliche Zahl $\geqslant 1$ sei. Jeder dieser Teile sei verschiedener Zustände fähig. Wir nennen sie die *Teilzustände* von S. Die Anzahl der möglichen Teilzustände lassen wir offen. Sie braucht für die einzelnen S-Teile nicht dieselbe zu sein; es wird auch nicht vorausgesetzt, daß diese Anzahl endlich ist, sie kann vielmehr gegebenenfalls abzählbar unendlich sein und es kann sogar der Fall kontinuierlich unendlich vieler Teilzustände zugelassen werden. Die Teilzustände sollen für die Eigenschaft G von kausaler Relevanz sein. Es wird sich jedoch als nicht erforderlich erweisen, diesen Begriff der kausalen Relevanz zu definieren.

Den n S-Teilen werden die folgenden *Teilzustandsvariablen* zugeordnet: $X_{1_t}, X_{2_t}, \ldots, X_{n_t}$. Die Werte dieser Variablen bestehen aus den Teilzuständen. Daher können die Wertbereiche der Variablen endlich oder unendlich sein. Wir bezeichnen sie mit $I_1, \ldots I_n$. Wir unterscheiden also formal zwischen dem Teil S_j und dem j-ten Wertbereich I_j, obwohl die Elemente von I_j genau die Teilzustände des S-Teiles S_j sind.

Neben den Teilzustandsvariablen verwenden wir auch *Teilzustandskonstanten* von der Gestalt 'C_{i_k}'. Dabei läuft der Index i von 1 bis n und k über die zulässigen Werte der Variablen X_{i_j}, d.h. also über die Teilzustände des i-ten S-Teiles. In 'C_{i_k}' bedeuten i und k stets feste Zahlen. Eine solche Teilzustandskonstante kann aufgefaßt werden als Abkürzung für eine mehr oder weniger komplexe Aussage, die den k-ten Zustand des i-ten S-Teiles so weit vollständig beschreibt, als dies für die zu betrachtenden kausalen Erklärungen erforderlich ist. Die Teilzustände von S werden nämlich als Antecedensbedingungen von kausalen Erklärungen fungieren. Teilzustandskonstante können für Teilzustandsvariable *substituiert* werden. Dabei ist nur die Substitution der 'C_{j_k}' für die Variablen 'X_{j_i}' zulässig.

Unter A_{i_k} verstehen wir den k-ten Teilzustand des i-ten S-Teiles. Auch hier sind i und k stets feste Zahlen. Die A_{j_k} sind also die Elemente von I_j. Da unsere Kenntnis von Teilzuständen nur so weit zu reichen braucht, als sie durch Teilzustandskonstanten darstellbar sind, sagen wir auch: 'C_{i_k}' *bezeichnet* (den Teilzustand) A_{i_k} abgekürzt: *Bez* (C_{i_k}, A_{i_k}), wobei wir die durch 'C_{i_k}' nicht zum Ausdruck gebrachten Merkmale von A_{i_k} vernachlässigen. Es wird vorausgesetzt, daß für jedes A_{i_k}, das in der S-Analyse vorkommt, ein bezeichnendes 'C_{i_k}' zur Verfügung steht.

Zusätzlich zu den angeführten Variablen und Konstanten verwenden wir die Zeitvariable 't' und Zeitkonstante 'c_t', welche bestimmte Zeitpunkte T bezeichnen (*Bez* (c_t, T)). Ebenso wie in dieser letzten Aussage werden wir uns auch im folgenden auf bestimmte Zeitpunkte durch das Symbol 'T' beziehen. 'T_1 ist früher als T_2' kürzen wir ab durch: '$T_1 < T_2$'.

Den Ausdruck '$(X_{1_t}, X_{2_t}, \ldots, X_{n_t}, t)$' nennen wir die *Zustandsmatrix* von *S*. Die einzelnen Zustände des Systems *S*, die für die folgenden Betrachtungen von Relevanz sind, werden wir aus den zulässigen Spezialisierungen dieser Matrix erhalten (d.h. genauer: die zulässigen Spezialisierungen der Matrix bezeichnen jene Zustände). Durch das Prädikat '*G*' bezeichnen wir eine Eigenschaft – in der Regel wird dies eine ziemlich komplizierte Eigenschaft von höherer Ordnung sein –, die das System unter geeigneten Umständen besitzen kann. *G* soll nur von einem *S*-Zustand sinnvoll ausgesagt werden können.

Im Gegensatz zum System *S* ist es nicht erforderlich, die Umgebung *U* in bezug auf ihre Struktur zu analysieren; denn diese Umgebung von *S* ist für uns nur so weit von Belang, als sie einen Störungsfaktor für die Zustände von *S* darstellt, d.h. nur so weit, als sie kausale Einwirkungen auf *S* ausübt, welche *S*-Zustände mit der Eigenschaft *G* in *G*-fremde *S*-Zustände verwandelt. Trotzdem kann es sich natürlich in den konkreten Anwendungen als möglich erweisen, für *U* und seine in bezug auf *S* kausal relevanten Zustände in zwangloser Weise eine analoge Analyse durchzuführen wie für *S*. Die einzelnen *U*-Zustände wären dann als Spezialisierungen einer *U*-Zustandsmatrix zu konstruieren. In diesem Fall müßte in einer Aussage über eine auf *S* einwirkende Störung stets auf *U* in seiner Gesamtheit Bezug genommen werden. Es könnte sich jedoch auch als zweckmäßiger erweisen, in den speziellen Aussagen über Störungen von *S* nicht immer das ganze *U* zum Gegenstand zu nehmen, sondern jeweils nur ganz bestimmte 'Ausschnitte' daraus, die sich gerade für *S* als kausal relevant erweisen. Formal würde dieser Unterschied dadurch zum Ausdruck kommen, daß im ersten Fall für alle Aussagen über *U* ('*U*-Zustandsbeschreibungen') derselbe Gegenstandsbereich zugrundegelegt würde, während in der zweiten Darstellungsweise die Gegenstandsbereiche je nach Situation variieren könnten. Es soll hier keine Entscheidung für die eine oder andere Methode getroffen werden. Daher werden die beiden Begriffe des *U-Zustandes* und der *U-Zustandsbeschreibung* auch nicht durch Definitionen eingeführt, sondern als undefinierte Grundbegriffe verwendet. Erst in der Anwendung auf die Beschreibung konkreter Systeme *S* wären auch diese Begriffe scharf zu definieren. Analog wie die Zustandsbeschreibungen von *S* einen *S*-Zustand, bestehend aus *n* Teilzuständen, in dem Genauigkeitsgrad charakterisieren, der für die kausale Analyse des Funktionierens von *S* erforderlich ist, sollen die Zustandsbeschreibungen von *U* einzelne *U*-Zustände soweit kennzeichnen, daß diese Beschreibungen als Antecedensbedingungen von adäquaten kausalen Erklärungen der Störungen von *S*-Zuständen fungieren können.

Für die folgenden Formulierungen werden einige übliche Symbole der Logik verwendet. Sie dienen lediglich als Abkürzungen umgangssprachlicher Wendungen. Daher braucht keine Entscheidung darüber zu erfolgen, welches Logik-System dabei verwendet wird. Es ist jedoch klar, daß im Falle einer vollständigen Formalisierung die Ausdrucksmittel der einfachen Prädikatenlogik nicht genügen würden, sondern eine höhere Logik benötigt wird. Außer logischen Symbolen verwenden wir noch das Zeichen '→'. Es soll nur dort benützt werden, wo das Erklärungsschema auf einen konkreten Fall angewendet werd. Vor dem '→' stehen die Antecedensbedingungen und relevanten Gesetze als Explanans, hinter dem '→' steht das Explanandum. Sowohl Explanans wie Explanandum können entweder semantisch, d.h. als Sätze, oder in absoluter (sprachunabhängiger) Weise wiedergegeben werden. Die sprachunabhängige Verwendung soll durch Definitionen eingeführt werden. Im Falle einer kausalen Erklärung i.e.S., in der nur deterministische Gesetze beteiligt sind, soll das Symbol '→' die logische Ableitungsbeziehung ausdrücken; im Falle von Erklärungen mit Hilfe statistischer Gesetze kann es auch die Wahrscheinlichkeitsimplikation bedeuten. Der Unterschied zwischen diesen beiden Fällen ist für das Folgende ohne Belang; daher darf dasselbe Symbol zur Deckung beider Fälle verwendet werden.

Zur Wiedergabe sprachlicher Ausdrücke verwenden wir deutsche Großbuchstaben. Zustandsbeschreibungen von *S* bezeichnen wir mit '$\mathfrak{Z}$' (evtl. mit unteren Indizes), dagegen *S*-Zustände mit '*Z*'. Umgebungszustandsbeschreibungen werden mit '$\mathfrak{U}$', Umgebungszustände mit '*U*' bezeichnet. Ein analoger Unterschied wird für Gesetze gemacht; als Symbole verwenden wir hier '$\mathfrak{L}$' und '*L*'. Um auszudrücken, daß eine Umgebungszustandsbeschreibung $\mathfrak{U}$ den Zustand *U* und eine Gesetzesaussage $\mathfrak{L}$ das Gesetz *L* zum Inhalt hat, schreiben wir auch: '*Bez* ($\mathfrak{U}$, *U*)' und '*Bez* ($\mathfrak{L}$, *L*)'. Zur Verdeutlichung werden gelegentlich in Klammern Erläuterungen beigefügt; um eindeutig klarzustellen, daß es sich dabei um keine Bestandteile der Definitionen selbst handelt, beginnen diese Erläuterungen stets mit 'd.h.'.

D1. Z ist ein *S-Zustand* (genauer: ein *T_l-S-Zustand*) $SZu(Z)$ bzw. $SZu(Z, T_l) =_{Df}$ Z ist ein geordnetes $(n+1)$-tupel von der Gestalt $A_{1_{k_1}}; A_{2_{k_2}}; \ldots; A_{n_{k_n}}; T_l$, wobei alle k_i feste Zahlen sind, jedes $A_{i_{k_i}}$ Element von I_i ist und T_l eine Zeitkonstante darstellt.

D2. $\mathfrak{Z}$ ist eine zulässige Spezialisierung der Zustandsmatrix '$(X_{1_i}, X_{2_i}, \ldots, X_{n_i}, t)$' von $S =_{Df}$ $\mathfrak{Z}$ ist ein Ausdruck von der Gestalt '$(C_{1_{j_1}}, C_{2_{j_2}}; \ldots, C_{n_{j_n}}, c_l)$', wobei die j_i feste Zahlen sind, jedes C_{i_j} eine Teilzustandskonstante darstellt und c_l eine Zeitkonstante.

Zulässige Spezialisierungen der Zustandsmatrix von S nennen wir auch *S-Zustandsbeschreibungen.*

D3. *Bez* $(\mathfrak{Z}, Z)$ ('der Ausdruck $\mathfrak{Z}$ bezeichnet den S-Zustand Z') $=_{Df}$ $\mathfrak{Z}$ ist eine zulässige Spezialisierung der Zustandsmatrix von der Gestalt '$(C_{1_{j_1}}, \ldots, C_{n_{j_n}}, c_l)$', Z hat die Gestalt: $A_{1_{k_1}}; \ldots; A_{n_{k_n}}; T_m$, für jedes $C_{i_{j_i}}$ gilt: *Bez* $(C_{i_{j_i}}, A_{i_{k_i}})$ und *Bez* (c_l, T_m) (d.h. die Teilzustandskonstanten '$C_{i_{j_i}}$' bezeichnen die Teilzustände $A_{i_{k_i}}$ und die Zeitkonstante 'c_l' bezeichnet den Zeitpunkt T_m).

Auf Grund der früheren Festsetzung über die verfügbaren '$C_{i_{j_i}}$' gilt der Satz, daß Z genau dann ein S-Zustand ist, wenn es eine zulässige Spezialisierung $\mathfrak{Z}$ der Zustandsmatrix von S gibt, so daß *Bez* $(\mathfrak{Z}, Z)$.

D4. $G(Z; T_l)$ ('Z ist ein T_l-G-Zustand') $=_{Df}$ Z ist ein S-Zustand von der Gestalt $A_{i_{k_1}}; \ldots; A_{n_{k_n}}; T_l$, so daß gilt:
$G(A_{1_{k_1}}; \ldots; A_{n_{k_n}}; T_l)$ (d.h. der durch die $A_{i_{k_i}}$ und T_l festgelegte Zustand Z besitzt die Eigenschaft G).

D5. $Z_1, L \rightarrow Z_2 =_{Df} (E\mathfrak{A}_1)(E\mathfrak{A}_2)(E\mathfrak{L})(SZu(Z_1) \wedge SZu(Z_2) \wedge Bez(\mathfrak{A}_1, Z_1) \wedge Bez(\mathfrak{A}_1, Z_2) \wedge Bez(\mathfrak{L}, L) \wedge \mathfrak{A}_1, \mathfrak{L} \rightarrow \mathfrak{A}_2)$.

Wir treffen ferner die folgende Festsetzung: Wenn Z ein S-Zustand von der Gestalt $A_{1_{l_1}}; \ldots; A_{n_{l_n}}; T_m$ ist, so soll durch '$T'Z$' der Zeitpunkt T_m bezeichnet werden.

D6. L ist ein *S-Gesetz* $SG(L) =_{Df}$ es gibt ein Z_1 und ein Z_2 mit $T'Z_1 < T'Z_2$ so daß $Z_1, L \rightarrow Z_2$.

D7. (der S-Zustand) Z_1 *determiniert* (den S-Zustand) Z_2 $Det(Z_1, Z_2) =_{Df}$ es gibt ein S-Gesetz L, so daß $T'Z_1 < T'Z_2$ und $Z_1, L \rightarrow Z_2$.

D8. S ist ein *abgeschlossenes System* $=_{Df}$ für jeden S-Zustand Z gilt: entweder es gibt kein Z_1, so daß $T'Z_1 < T'Z$, oder es gibt ein Z_2, so daß $T'Z_2 < T'Z$ und Z_2 determiniert Z.

Der in **D**7 und **D**8 verwendete Ausdruck 'determiniert' ist nach den früheren Vorbemerkungen nicht so zu verstehen, als dürften nur kausale

Gesetze im spezielleren Sinne der deterministischen Gesetze zur Anwendung gelangen.

Wie bereits bemerkt, sollen Umgebungszustände einfach mit 'U', Umgebungszustandsbeschreibungen mit '$\mathfrak{U}$' bezeichnet werden. Diese Darstellungsweise ist dadurch gerechtfertigt, daß die Umgebung als Ganze in den formalen Definitionen keine Rolle spielt.

D$_9$. L ist ein *U-Gesetz* bezüglich S $UG(L) =_{Df} (\mathrm{E}\mathfrak{U})(\mathrm{E}U)(\mathrm{E}\mathfrak{L})(\mathrm{E}L)(\mathrm{E}\mathfrak{L}_1)(\mathrm{E}L_1)(\mathrm{E}\mathfrak{Z})(\mathrm{E}Z_1)(\mathrm{E}\mathfrak{Z}_2)(\mathrm{E}Z_2)(Bez\ (\mathfrak{U}, U) \wedge Bez\ (\mathfrak{L}, L) \wedge Bez\ (\mathfrak{L}_1, L_1) \wedge Bez\ (\mathfrak{Z}_1, Z_1) \wedge Bez\ (\mathfrak{Z}_2, Z_2) \wedge SG(L) \wedge T'Z_1 < T'Z_2 \wedge \mathfrak{U}, \mathfrak{Z}_1, \mathfrak{L}, \mathfrak{L}_1 \rightarrow \mathfrak{Z}_2 \wedge \neg (\mathrm{E}L')(\mathrm{E}Z)(SG(L') \wedge T'Z < T'Z_2 \wedge Z, L' \rightarrow Z_2))$.

Die Motivierung für diese Definition ist die folgende: Wenn die Umgebung an der Determination eines S-Zustandes beteiligt ist, so wird sie in der Regel nicht *allein* daran beteiligt sein, sondern dieser S-Zustand wird außerdem mitbestimmt sein durch einen früheren Zustand von S. Damit die Determination durch die Umgebung aber keine 'leere' ist, muß die Möglichkeit ausgeschlossen werden, daß der fragliche S-Zustand bereits durch einen früheren vollständig determiniert ist. Analog zum früheren Fall (vgl. **D**$_5$) soll auch dann, wenn U-Gesetze und U-Zustände beteiligt sind, die nichtsemantische Schreibweise benützt werden. Dann läßt sich z.B. die in **D**$_9$ enthaltene Ableitungsbeziehung auch so wiedergeben: $U, Z_1, L, L_1 \rightarrow Z_2$.

D$_{10}$. (Der S-Zustand) Z ist *U-determiniert* $Udet(Z) =_{Df}$ es gibt ein U_1 und ein Z_1 mit $T'Z_1 < T'Z$, ferner ein U-Gesetz L_1 sowie ein S-Gesetz L_2, so daß $U_1, Z_1, L_1, L_2 \rightarrow Z$, während es für kein Z' mit $T'Z' < T'Z$ ein S-Gesetz L' gibt, so daß $Z', L' \rightarrow Z$.

Die Motivierung für diese Definition ist eine analoge wie im vorigen Fall.

D$_{11}$. $\bar{G}(Z; T_i) =_{Df}$ Z ist ein S-Zustand von der Gestalt $A_{1k_1}; \ldots; A_{nk_n}; T_i$ und es gilt: $\neg G(Z; T_i)$ (d.h. dieser Zustand besitzt nicht die Eigenschaft G).

D$_{12}$. Z wird (bezüglich G) *U-gestört* zu Z_1 $Ugest(Z, Z_1) =_{Df} T'Z < T'Z_1 \wedge G(Z; T'Z) \wedge \bar{G}(Z_1; T'Z_1) \wedge Udet(Z_1) \wedge (Z')(T'Z < T'Z' \wedge T'Z' < T'Z_1 \cdot \supset G(Z'; T'Z'))$ (d.h. durch eine aus der Umgebung stammende kausale Einwirkung wird der G-Zustand Z in den zeitlich darauffolgenden Zustand Z_1 verwandelt, dem die Eigenschaft G nicht mehr zukommt).

Es sind natürlich vielerlei U-Störungen der Zustände des Systems S

denkbar. Doch da das System S nur unter dem Gesichtspunkt analysiert wird, daß es eine bestimmten seiner Zustände zukommende Eigenschaft 'zu erhalten tendiert', können wir uns auf jene U-Störungen beschränken, durch die Zustände mit der Eigenschaft G in G-fremde Zustände, also ohne diese Eigenschaft G, transformiert werden.

Der in der folgenden Definition verwendete und für sich verständliche Begriff des aus bestimmten A_i, B_j und T bestehenden S-Zustandes könnte ohne Mühe durch eine geeignete Definition explizit eingeführt werden.

D13. Das k-tupel von Teilzuständen $A_1, \ldots, A_k$ $(1 \leqslant k \leqslant n)$ mit $A_1 \in I_{s_1}, \ldots, A_k \in I_{s_k}$ ist ein Element der *G-k-Ausschlußklasse* von $S =_{Df}$ es gibt kein $(n-k+1)$-tupel $B_1, \ldots, B_{n-k}, T$ mit $B_1 \in I_{l_1}, \ldots, B_{n-k} \in I_{l_{n-k}}$ (mit $I_{s_j} \neq I_{l_i}$ für alle $1 \leqslant j \leqslant k$ und $1 \leqslant i \leqslant n-k$), so daß für einen nur aus $A_1, \ldots, A_k, B_1, \ldots, B_{n-k}, T$ bestehenden S-Zustand Z gilt: $G(Z; T)$.

Unter der *G-Ausschlußklasse* verstehen wir die Vereinigungsmenge aller G-k-Ausschlußklassen für $k = 1, \ldots, n$. Es möge beachtet werden, daß z.B. das Paar A_{2_i}, A_{5_j} der G-2-Ausschlußklasse (und damit der G- Ausschlußklasse) angehören kann, ohne daß A_{2_i} oder A_{5_j} der G-1-Ausschlußklasse anzugehören brauchen.

Wenn die in **D**13 angeführten Zusatzbedingungen erfüllt sind, so ist Z eine eindeutige Funktion der A_i, B_j und T, so daß wir in einem solchen Fall abkürzend schreiben können: $Z = \varphi(A_1, \ldots, A_k, B_1, \ldots, B_{n-k}, T)$.

D14. X ist ein Glied von Z $Glied(Z, X) =_{Df}$ Z hat die Gestalt: $A_{1_{k_1}}; \ldots; A_{n_{k_n}}; T$ und $(X = A_{1_{k_1}}$ oder $\ldots$ oder $X = A_{n_{k_n}})$.

Man beachte, daß T nicht zu den Gliedern von Z gerechnet wird; dieser Begriff wird also auf die Teilzustände von Z beschränkt.

D15. Das k-tupel von Teilzuständen $A_1, \ldots, A_k$ $(1 \leqslant k \leqslant n)$ mit $A_i \in I_{s_i}$ bildet ein Element der *G-k-Vernichtungsklasse* von $S =_{Df} A_1, \ldots,$ A_k ist ein Element der G-k-Ausschlußklasse und für jedes Z, welches die Bedingung erfüllt: Glied$(Z, A_1), \ldots,$ Glied(Z, A_k), gilt: es gibt kein Z', so daß Det(Z, Z') und $G(Z'; T'Z')$.

Unter der G-Vernichtungsklasse soll die Vereinigungsmenge aller G-k-Vernichtungsklassen verstanden werden.

Die inhaltliche Bedeutung der in **D**13 und **D**15 eingeführten Begriffe ist die folgende: Wenn k Teilzustände zur G-Ausschlußklasse gehören, so ist nicht nur jener S-Zustand Z, dessen Glieder sie bilden, kein G-Zustand, sondern diese Eigenschaft G fehlt auch allen jenen Zuständen, die aus Z dadurch hervorgehen, daß die übrigen Teilzustände durch beliebige andere Teilzustände der betreffenden Teilzustandsbereiche ersetzt werden. Trotzdem könnte in einem solchen Falle die G-Eigenschaft für das System S in der Weise 'wiederhergestellt' werden, daß auf Grund eines S-Gesetzes der Zustand Z einen anderen S-Zustand Z' mit der Eigenschaft G determiniert. Ist auch diese Möglichkeit ausgeschlossen, so gehört der fragliche Zustand bzw. das fragliche k-tupel zur G-Vernichtungsklasse.

D_{16}. Das Glied X von Z *entspricht* dem Glied Y von Z_1 $Ent(Z, X; Z_1, Y)$ $=_{Df}$ Z hat die Gestalt: $A_{1_{i_1}}; \ldots; A_{n_{i_n}}; T$, Z_1 hat die Gestalt: $A_{1_{j_1}}; \ldots; A_{n_{j_n}}; T_1$ und es gilt: $(X = A_{1_{i_1}} \wedge Y = A_{1_{j_1}}) \vee \ldots \vee (X = A_{n_{i_n}} \wedge Y = A_{n_{j_n}})$.

D_{17}. X ist ein *Glied der Abweichung* von Z_2 bezüglich Z_1 $X \in Abw(Z_2, Z_1) =_{Df}$ X ist ein Glied von Z_2 und wenn X dem Glied Y von Z_1 entspricht, so gilt: $X \neq Y$.

D_{18}. Y ist ein Element des *Abweichungsbereiches* zwischen Z_2 und Z_1 $Y \in AB(Z_2, Z_1) =_{Df} (EX)(X \in Abw(Z_2, Z_1) \wedge (s)((1 \leqslant s \leqslant n) \supset (X \in I_s \equiv Y \in I_s)))$.

Wenn X ein Teilzustand von Z_2 ist, der von dem entsprechenden Teilzustand von Z_1 abweicht, so soll also der gesamte Wertbereich, zu dem X gehört, eine Teilklasse des Abweichungsbereiches zwischen Z_2 und Z_1 bilden. Insbesondere ist das im Definiens erwähnte X selbst ein Element dieses Abweichungsbereiches.
Das Gegenstück zum Begriff des Abweichungsbereiches ist der Begriff des Entsprechungsbereiches.

D_{19}. Y ist ein Element des *Entsprechungsbereiches* zwischen Z_2 und Z_1 $Y \in EB(Z_2, Z_1) =_{Df}$ (EX)(Glied(Z_1, X) $\wedge$ Glied(Z_2, X) $\wedge$ $(s)((1 \leqslant s \leqslant n) \supset (X \in I_s \equiv Y \in I_s)))$.

Die für ZO-Systeme charakteristische Zurückführung U-gestörter Zustände in G-Zustände soll jetzt durch drei Begriffe gekennzeichnet werden.

D_{20}. Der Zustand Z_1 von S wird *G-eigenkompensiert* zu $Z_2 =_{Df}$ (EZ)(Ugest(Z, Z_1) $\wedge$ Det(Z_1, Z_2) $\wedge$ $G(Z_2)$ $\wedge$ $(X)(X \in$ Abw.$(Z_2, Z) \supset X \in$ AB$(Z_1, Z)))$.

Daß der S-Zustand Z_2 die Eigenschaft G besitzt, wurde hier abkürzend durch '$G(Z_2)$' wiedergegeben. Inhaltlich besagt das Definiens folgendes: Es gibt einen dem S-Zustand Z_1 zeitlich vorangehenden Zustand Z, der die Eigenschaft G besitzt. Dieser Zustand wurde durch eine U-Störung in den $\bar{G}$-Zustand Z_1 umgewandelt (all dies ist in 'Ugest(Z, Z_1)' ausgedrückt). Auf Grund eines S-Gesetzes wird der $\bar{G}$-Zustand Z_1 wieder in einen G-Zustand Z_2 zurückgeführt. Im letzten Glied des Definiens wird gefordert, daß sich Z_2 und Z höchstens um solche Teilzustände voneinander unterscheiden, die denselben Wertbereichen angehören wie die Glieder der Abweichung zwischen Z und Z_1. Es werden also bei der 'Rückführung' in einen neuen G-Zustand nur jene Teilzustände von Z_1 modifiziert, die durch die vorangehende U-Störung aus andersartigen Teilzuständen von Z hervorgegangen sind. Die durch die U-Störung nicht betroffenen Teilzustände bleiben hingegen während des ganzen Prozesses, d.h. sowohl beim Übergang von Z in Z_1 wie bei dem von Z_1 in Z_2, konstant. Es sei noch ausdrücklich darauf aufmerksam gemacht, daß im letzten Formelglied 'AB' vorkommt und nicht 'Abw'. Es wird keineswegs verlangt, daß der ursprüngliche G-Zustand wiederhergestellt wird; vielmehr werden auch alle jene Fälle zugelassen, in denen einige oder alle Glieder der Abweichung von Z bezüglich Z_1 zugleich Glieder der Abweichung von Z bezüglich Z_2 sind. Die Wiederherstellung des ursprünglichen Zustandes ist dagegen nur ein sehr spezieller Fall der Eigenkompensation. Von E. Nagel ist die Möglichkeit der Eigenkompensation nicht in Erwägung gezogen worden, anscheinend ohne zwingenden Grund (vgl. [9] S. 548).

D21. Der Zustand Z_1 von S wird *G-fremdkompensiert* zu $Z_2 =_{Df}$ $(EZ)(\text{Ugest}(Z, Z_1) \wedge \text{Det}(Z_1, Z_2) \wedge G(Z_2) \wedge (X)(X \in \text{Abw}(Z_1, Z) \supset \text{Glied}(Z_2, X)))$.

Bis auf den letzten Bestandteil ist das Definiens von **D**21 mit dem von **D**20 identisch. Dieser letzte Bestandteil besagt, daß jene Teilzustände von Z_1, durch die sich Z_1 von Z unterscheidet, unverändert in den G-Zustand Z_2 übernommen werden. Die Kompensation stützt sich also ausschließlich auf eine Variation der übrigen Teilzustände von Z_1. Daß eine Variation in bezug auf mindestens einen Teilzustand erfolgt ist, wird durch $G(Z_2)$ und $\bar{G}(Z_1)$ gewährleistet.

D22. Der Zustand Z_1 von S wird *gemischt G-kompensiert* zu $Z_2 =_{Df}$ $(EZ)(\text{Ugest}(Z, Z_1) \wedge \text{Det}(Z_1, Z_2) \wedge G(Z_2) \wedge (EX)(EY)(X \in \text{Abw}(Z_1, Z) \wedge \text{Ent}(Z_2, Y; Z_1, X) \wedge X \neq Y) \wedge (EX)(EY)(X \in \text{EB}(Z_1, Z) \wedge \text{Ent}(Z_2, Y; Z_1, X) \wedge X \neq Y))$.

Durch das Definiens wird diesmal gewährleistet, daß sowohl eine partielle Eigen- wie eine partielle Fremdkompensation vorliegt, d.h. daß Z_2 sich in bezug auf mindestens ein Glied der Abweichung von Z_1 bezüglich Z von Z_1 unterscheidet und von Z in bezug auf mindestens eines der übrigen Glieder. Offenbar hätte im letzten Bestandteil des Definiens statt '$X \in \text{EB}(Z_1, Z)$' auch 'Glied(Z, X)' gewählt werden können.

D23. Der Zustand Z_1 von S wird *G-kompensiert* zu $Z_2 =_{Df}$ Z_1 wird G-eigenkompensiert oder G-fremdkompensiert oder gemischt G-kompensiert zu Z_2.

Es gelten die folgenden Sätze:

(1) Ein S-Zustand Z_1, der zu einem früheren Zustand Z in der Relation steht: Ugest(Z, Z_1), wird dann und nur dann zu einem S-Zustand Z_2 G-kompensiert, wenn kein k-tupel von Gliedern von Z_1 Element der G-Vernichtungsklasse ist.

(2) Wenn Ugest(Z, Z_1) und das k-tupel, bestehend aus den Gliedern der Abweichung von Z_1 bezüglich Z, ein Element der G-Ausschlußklasse bildet, so gibt es keinen S-Zustand Z_2, so daß Z_1 G-fremdkompensiert wird zu Z_2.

(3) Wenn Ugest(Z, Z_1) und das k-tupel, bestehend aus den Gliedern der Abweichung von Z_1 bezüglich Z, ein Element der G-Ausschlußklasse bildet, so wird Z_1 dann und nur dann zu einem S-Zustand Z_2 G-kompensiert, wenn Z_1 zu Z_2 G-eigenkompensiert oder gemischt G-kompensiert wird.

(4) Es seien die folgenden Bedingungen erfüllt:

(*a*) Ugest(Z, Z_1);

(*b*) das k-tupel (von Gliedern von Z_1), bestehend aus den Gliedern der Abweichung von Z_1 bezüglich Z, ist kein Element der G-Ausschlußklasse;

(*c*) alle übrigen k-tupel von Teilzuständen, die den Abweichungsbereichen zwischen Z und Z_1 angehören, sind Elemente der G-k-Ausschlußklasse, ausgenommen jenes k-tupel von Teilzuständen, das ganz in Z liegt, d.h. dessen sämtliche Glieder zugleich Glieder von Z sind;

(*d*) $\neg$ Det(Z_1, Z);

dann gilt: Z_1 wird genau dann zu einem S-Zustand Z_2 G-kompensiert, wenn Z_1 zu Z_2 G-fremdkompensiert wird.

Die Ausnahmebestimmung in (*c*) muß hinzugenommen werden, weil der ursprüngliche *S*-Zustand *Z* nach Voraussetzung ein *G*-Zustand ist und daher kein *k*-tupel von Teilzuständen von *Z* ein Element der *G-k*-Ausschlußklasse bilden kann. Die Voraussetzung (*d*) ist nicht vollkommen korrekt formuliert, weil darin die Zeitbestimmung außer acht gelassen wurde. Genauer müßte es heißen: Durch Z_1 wird kein (zeitlich späterer) *S*-Zustand Z' determiniert, der in bezug auf sämtliche Glieder mit den entsprechenden Gliedern von *Z* übereinstimmt (wobei wieder zu bedenken ist, daß das *T* von *Z* nicht zu den Gliedern von *Z* gehört).

(5) Es seien die folgenden Bedingungen erfüllt:
(*a*) und (*b*) analog Satz **(4)**;
(*c*) zur *G*-Ausschlußklasse gehören alle jene von *Z* verschiedenen *S*-Zustände Z_i, für die gilt: Z_i stimmt in bezug auf die Glieder aus dem Entsprechungsbereich zwischen *Z* und Z_1 mit *Z* überein;
(*d*) $\neg$ Det(Z_1, Z);
dann gilt: Z_1 wird genau dann zu einem *S*-Zustand Z_2 *G*-kompensiert, wenn Z_1 zu Z_2 *G*-fremdkompensiert oder gemischt *G*-kompensiert wird.

Bezüglich (*d*) gilt die analoge Zusatzbemerkung wie für Satz **(4)**.

(6) Es seien die folgenden Bedingungen erfüllt:
(*a*) Ugest(Z, Z_1);
(*b*) alle *k*-tupel von Teilzuständen, deren Glieder den Abweichungsbereichen zwischen *Z* und Z_1 angehören, wobei *k* die Anzahl der Abweichungsbereiche sein möge, sind Elemente der *G-k*-Ausschlußklasse, vorausgesetzt, daß diese *k*-tupel nicht ganz in *Z* liegen;
(*c*) Z_1 wird nicht zu einem Zustand Z' gemischt *G*-kompensiert, der in sämtlichen Gliedern der Abweichung von *Z* bezüglich Z_1 mit *Z* übereinstimmt;
dann gilt: Z_1 wird genau dann zu einem *S*-Zustand Z_2 *G*-kompensiert, wenn Z_1 eigenkompensiert wird zu Z_2 und überdies alle Glieder von Z_2 mit den entsprechenden Gliedern von *Z* identisch sind.

Bezüglich der Wendung 'ganz in *Z* liegen' vgl. **(4)** (*c*). In der Bedingung (*c*) werden diesmal nur jene sehr speziellen Fälle der gemischten Kompensation ausgeschlossen, welche die gestörten Teilzustände in die ursprünglichen zurückführen, darüber hinaus aber, sozusagen 'überflüssig', andere Teilzustände variieren.

(7) Es seien die folgenden Bedingungen erfüllt:
(*a*) Ugest(Z, Z_1);
(*b*) das k-tupel, bestehend aus den Gliedern der Abweichung von Z_1 bezüglich Z, ist Element der G-k-Ausschlußklasse;
(*c*) alle von Z verschiedenen S-Zustände, welche mit Z_1 jene Glieder gemeinsam haben, die nicht den Abweichungsbereichen zwischen Z_1 und Z angehören, sind $\bar{G}$-Zustände;
(*d*) Z_1 wird nicht eigenkompensiert zu einem Zustand Z', der in bezug auf alle Glieder mit den entsprechenden Gliedern von Z übereinstimmt (d.h. Z_1 wird nicht in den ursprünglichen Zustand zurückgeführt);
dann gilt: Z_1 wird genau dann zu einem S-Zustand Z_2 G-kompensiert, wenn Z_1 gemischt G-kompensiert wird zu Z_2.

(8) Es seien die folgenden Bedingungen erfüllt:
(*a*) Ugest(Z, Z_1);
(*b*) das k-tupel, bestehend aus den Gliedern der Abweichung von Z_1 bezüglich Z, ist Element der G-k-Ausschlußklasse;
(*c*) alle jene S-Zustände sind $\bar{G}$-Zustände, die sich von Z_1 erstens in bezug auf mindestens ein Glied unterscheiden, das dem Abweichungsbereich zwischen Z und Z_1 angehört, und zweitens in bezug auf mindestens ein Glied, das dem Entsprechungsbereich zwischen Z und Z_1 angehört;
dann gilt: Z_1 wird genau dann zu einem S-Zustand Z_2 G-kompensiert, wenn Z_1 G-eigenkompensiert wird zu Z_2.

(9) Es seien die folgenden Bedingungen erfüllt:
(*a*) Ugest(Z, Z_1);
(*b*) analog (*c*) von **(8)**;
dann gilt: Z_1 wird genau dann zu einem S-Zustand Z_2 G-kompensiert, wenn Z_1 G-eigenkompensiert oder G-fremdkompensiert wird zu Z_2.

(10) Der Fall der G-Kompensation kann in S dann und nur dann eintreten, wenn nicht jeder $\bar{G}$-Zustand ein k-tupel von Teilzuständen enthält, das ein Element der G-k-Vernichtungsklasse ist.

Die angeführten Sätze enthalten u.a. notwendige und hinreichende Bedingungen für das Eintreten reiner und gemischter Fälle der drei Kompensationsarten. Die Beweise folgen ohne Schwierigkeiten aus den vorangehenden Definitionen. Zur Illustration werde **(7)** bewiesen: Aus **(7)** (*a*) folgt, daß Z ein G-Zustand und Z_1 ein $\bar{G}$-Zustand ist. Wenn Z_1

gemischt G-kompensiert wird zu Z_2, so wird Z_1 nach **D**23 G-kompensiert zu Z_2. Es möge daher umgekehrt Z_1 zu Z_2 G-kompensiert werden. Wegen **D**23 muß dazu gezeigt werden, daß es sich dabei weder um eine Fremd- noch um eine Eigenkompensation handeln kann. Falls Z_1 zu einem Z_2 fremdkompensiert würde, müßte Z_2 hinsichtlich der Glieder der Abweichung von Z_1 bezüglich Z mit Z_1 übereinstimmen. Nach Voraussetzung **(7)** (*b*) ist aber das k-tupel, bestehend aus diesen Gliedern, ein Element der G-k-Ausschlußklasse. Nach **D**13 wäre dann Z_2 ein $\bar{G}$-Zustand, im Widerspruch zu der Annahme, daß Z_1 zu Z_2 fremdkompensiert worden ist; denn nach dieser Annahme müßte Z_2 auf Grund von **D**21 ein G-Zustand sein. Es bleibt noch der Fall der Eigenkompensation zu behandeln. Wegen **(7)** (*d*) ist die Wiederherstellung des ursprünglichen G-Zustandes ausgeschlossen. Die Eigenkompensation könnte daher nur zu einem Zustand führen, dessen Glieder von den entsprechenden Gliedern von Z (ganz oder teilweise) verschieden sind. Dabei würden wegen **D**20 die Glieder, in bezug auf welche Z_1 mit Z übereinstimmt, unverändert bleiben. Alle auf diese Weise herbeigeführten S-Zustände aber wären nach Voraussetzung **(7)** (*c*) $\bar{G}$-Zustände. Also kann der Fall der Eigenkompensation überhaupt nicht eintreten.

Für gewisse Zusammenhänge ist es ratsam, die Wendung zu gebrauchen 'Z_2 ist der Kompensationszustand von Z_1', wenn gilt: $\bar{G}(Z_1)$, $G(Z_2)$ und Z_1 wird G-kompensiert zu Z_2.

D24. S ist ein *partielles ZO-System* $=_{Df}$ es gibt mindestens ein Z_1 und ein Z_2, so daß gilt: wenn Ugest(Z_1, Z_2), so gibt es ein Z_3, zu dem Z_2 G-kompensiert wird (d.h. mindestens ein G-Zustand muß, wenn er durch eine Umgebungsstörung in einen $\bar{G}$-Zustand verwandelt wurde, durch ein S-Gesetz wieder in einen G-Zustand transformiert werden).

Um den leeren Fall auszuschließen, könnte hier gefordert werden, daß eine derartige Störung mit nachfolgender Kompensation tatsächlich stattgefunden hat. Da die Eigenschaft, ein partielles ZO-System zu sein, eine Dispositionseigenschaft ist, treten bezüglich des 'wenn . . ., dann ——' im Definiens die bekannten Schwierigkeiten auf. Die Diskussion der irrealen Konditionalsätze ist daher an dieser Stelle von Relevanz.

(11) S ist genau dann ein partielles ZO-System, wenn nicht jeder $\bar{G}$-Zustand von S ein k-tupel von Teilzuständen enthält, das Element der G-Vernichtungsklasse ist.

D_{25a}. S ist ein *partielles F-ZO-System* (ein partielles ZO-System mit ausschließlicher Fremdkompensation) $=_{Df}$ S ist ein partielles ZO-System und ein $\bar{G}$-Zustand wird dann und nur dann zu einem G-Zustand Z_2 kompensiert, wenn Z_1 G-fremdkompensiert wird zu Z_2.

Analog zu **D_{25a}** lauten **D_{25b}** bis **D_{25g}** für die Begriffe: *E-ZO-System* (ausschließlich Eigenkompensation), *Ge-ZO-System* (ausschließlich gemischte Kompensation), *FE-ZO-System* (ausschließlich Fremd- oder Eigenkompensation), *FGe-ZO-System* (ausschließlich gemischte Kompensation oder Fremdkompensation), *EGe-ZO-System* (ausschließlich gemischte Kompensation oder Eigenkompensation), *EFGe-ZO-System* (jede Art von Kompensationsmöglichkeit).

(12) S ist genau dann ein partielles ZO-System, wenn S ein partielles EFGe-ZO-System ist.

D_{26}. S ist ein *vollständiges ZO-System* $=_{Df}$ für jedes Z_1 von S, zu dem es ein Z gibt, so daß Ugest(Z, Z_1), existiert ein Kompensationszustand Z_2.

Auch hier könnte man die Analoga zu den in **D_{25b}** bis **D_{25g}** eingeführten Begriffen bilden.

(13) Wenn jeder $\bar{G}$-Zustand von S durch eine U-Störung hervorgerufen werden kann, so ist S genau dann ein vollständiges ZO-System, wenn die G-Vernichtungsklasse leer ist.

D_{27}. S ist ein *vollständig – totales ZO-System* $=_{Df}$ S ist ein vollständiges ZO-System und kein $\bar{G}$-Zustand von S enthält ein k-tupel von Teilzuständen, welches ein Element der G-k-Ausschlußklasse ist.

(14) Wenn jeder $\bar{G}$-Zustand von S durch eine U-Störung hervorgerufen werden kann und wenn außerdem für ein Z_1 mit Ugest(Z, Z_1) stets gilt, daß Z_1 G-fremdkompensiert wird zu einem S-Zustand Z_2, so ist S ein vollständig-totales ZO-System.

Wir brechen an dieser Stelle die Analyse von ZO-Systemen ab. Alle Typen von solchen Systemen sind dadurch charakterisiert, daß sie eine Eigenschaft G 'zu erhalten trachten'. Die Beschreibung dieser Tendenz erfordert keinen Rückgang auf teleologische Prinzipien. Es genügt, daß das fragliche System eine kausal beschreibbare Kompensationsvorrichtung besitzt. Diese Vorrichtung braucht selbstverständlich kein 'Mechanismus' in dem Sinne zu sein, daß ihre Tätigkeit allein mit Hilfe von Gesetzen der Mechanik erklärt werden kann.

Die Analyse könnte in verschiedener Weise fortgesetzt werden. So z.B. könnte man im Falle, daß Ugest(Z, Z_1) und Z_2 Kompensationszustand von Z_1 ist, den Begriff der *Kompensationsbereiche* zwischen Z_2 und Z_1 einführen und darunter die Wertbereiche verstehen, zu denen die Glieder der Abweichung von Z_2 bezüglich Z_1 gehören. Eine Anwendung dieses Begriffs wäre die folgende: Wenn Ugest(Z, Z_1), so wird Z_1 nur dann zu einem Zustand Z_2 gemischt G-kompensiert, wenn die Klasse der Kompensationsbereiche zwischen Z_2 und Z_1 weder identisch ist mit der Klasse der Abweichungsbereiche zwischen Z und Z_1 noch mit der Klasse der Entsprechungsbereiche zwischen Z und Z_1. Man könnte ferner *Teilsysteme von ZO-Systemen* betrachten, die selbst den Charakter von F-, E-, Ge-ZO-Systemen usw. haben und unter gewissen Bedingungen das gesamte System ausmachen. Schließlich könnte auch die *Zeitanalyse* berücksichtigt werden. Es wären zu unterscheiden: der Störungszeitraum (d.h. die Zeitspanne zwischen $T'Z_1$ und $T'Z$ bei Ugest(Z, Z_1)), der Kompensationszeitraum (d.h. die Zeitspanne zwischen dem T von Z_1 und dem des Kompensationszustandes Z_2) und der Operationszeitraum (die Summe von Störungs- und Kompensationszeitraum). Unter der *Konsistenz* eines Systems, insbesondere eines ZO-Systems, kann die Tatsache verstanden werden, daß jeder S-Zustand höchstens einen anderen S-Zustand determiniert.

4. TELEOLOGISCHE AUSSAGEN DER DRITTEN SCHICHT

Die vorangehenden Analysen haben gezeigt: Soweit es sich um das Funktionieren von Systemen mit zielgerichteter Organisation handelt, scheint kein Anlaß zu bestehen, seine Zuflucht zu zwecksetzenden Prinzipien zu nehmen; die Struktur von ZO-Systemen ist eine rein kausale, nichtteleologische Struktur. Diese Struktur – insbesondere die jeweilige Regulationsvorrichtung – im Detail zu beschreiben, ist nicht mehr Aufgabe philosophischer Analyse, sondern empirischer Spezialuntersuchungen. Wir setzen für das Folgende voraus, durch solche Untersuchungen sei gezeigt worden, daß auch Systeme mit Selbstregulation, mit Selbstreproduktion ('Fortpflanzung') usw. unter den Begriff des kausal beschreibbaren ZO-Systems subsumiert werden können, sowie daß auch dort, wo im Verhältnis zwischen Individuum und Umwelt eine höhere Zweckmäßigkeit zu walten scheint, nur ein kausal analysierbares

ZO-System von höherer Ordnung vorliegt. Dann ist auch für den Bereich der Organismen die Teleologie als eine scheinbare nachgewiesen, allerdings nur, *soweit das bloße Funktionieren der Organismen zur Diskussion steht.*

Die Situation ändert sich von neuem, wenn die *Entstehungsfrage* aufgeworfen wird. Teleologische Aussagen, die diese genetische Frage zu beantworten suchen, rechneten wir früher zur dritten Schicht. Hier von einer ganz neuen Schicht von teleologischen Aussagen zu sprechen, ist dadurch gerechtfertigt, daß die eben skizzierte 'Kausaltheorie' des Funktionierens von ZO-Systemen keineswegs eine Kausaltheorie der Entstehung von ZO-Systemen zur Folge hat.

Es liegt zunächst sogar nahe, auf Grund von Analogiebetrachtungen die Notwendigkeit teleologischer Erklärungen der Entstehung von ZO-Systemen anzunehmen. So ist nicht nur der menschliche Organismus in bezug auf sein Blut ein Temperaturregulator, der diese Temperatur äußeren Einflüssen gegenüber konstant zu halten sucht, sondern z.B. auch ein Kühlschrank. Gebilde der Technik können ebenso ZO-Systeme sein, wie naturgewachsene organische Gebilde. Und gerade weil bei allen von Menschen geschaffenen Maschinen die Entstehungsfrage durch eine echte teleologische Erklärung beantwortet werden muß – natürlich im früher geschilderten Sinn einer kausalen Erklärung i.w.S., unter deren Antecedensbedingungen Motive zielstrebiger Wesen vorkommen –, ist die Verführung außerordentlich groß, im Falle biologischer Gebilde oder anderer naturgeschaffener Selbstregulatoren (zu deren primitiven Formen man vielleicht bereits Atome oder Planetensysteme rechnen könnte) die Entstehungsfrage durch eine Erklärung aus Motiven zu beantworten. Diese Analogiebetrachtung wäre eine ähnliche wie jene, auf Grund deren bereits Aristoteles lange vor dem Ausbau der exakten Naturwissenschaften sein teleologisches Weltbild entwickelt hat. Nur daß diese Analogiebetrachtung jetzt auf eine höhere Ebene gehoben worden wäre; denn während Aristoteles bereits Analogien zwischen den Produkten einfacher Handwerkertätigkeit und den Erzeugnissen der Natur zugrunde legte, würde nunmehr erst die Beantwortung der Frage nach der Entstehung von ZO-Systemen derartige Analogien rechtfertigen.

Zwei Faktoren scheinen es vor allem zu sein, die an dieser Stelle eine teleologische Betrachtungsweise aufzwingen: erstens die Tatsache, daß eine 'zufällige' erstmalige Entstehung eines Systems mit zielgerichteter Organisation ungeheuer unwahrscheinlich wäre, und zweitens das Phänomen der Höherentwicklung der Organismen. Vor allem die durch zahllose empirische Daten gestützte und praktisch nicht zu bezweifelnde Hypothese, daß die höheren, d.h. differenzierteren Formen von Organismen aus niedrigeren Formen entstanden sind, das Leben sich also im Sinne zunehmender Differenzierung weiterentwickelt, ist oft geradezu als ein Beweis für das Walten eines lenkenden Geistes angesehen worden. Ein charakteristisches Beispiel aus der Literatur sei hier angeführt. In dem Werk von A. Kastil über die Philosophie von Franz Brentano heißt es anläßlich der Erörterung des teleologischen Gottesbeweises: 'Ebensowenig läßt sich die Fortentwicklung [1]) über eine bereits erreichte Stufe von hoher Vollkommenheit hinaus ohne Zuhilfenahme neuer und scheinbar höchst teleologischer Annahmen verstehen. ...Der Gedanke sei durch einen Vergleich illustriert. Es gelänge, eine Maschine zu konstruieren, die sich nicht bloß als Individuum ständig erneuerte, sondern auch aus der umgebenden Natur Stoffe aufnehmend, eine kleine Maschine aus sich heraus entwickelte, die dann zur Größe der ersten heranwüchse, also eine sich fortpflanzende Maschine. Wird sie sich immer weiter fortentwickeln? Und werden die abstammenden Maschinen noch vollkommener sein? Jeder gesunde Sinn wird das Gegenteil erwarten. Blinde Abweichungen werden wohl vorkommen, aber in verschlechterter Auflage. Immerhin könnte auch eine Verbesserung darunter sein. Jedenfalls aber weit mehr Verschlechterungen, und darum wird die Maschine, wenn sie überhaupt variieren kann, degenerieren, bis sie schließlich gar nicht mehr gehen wird. Ganz analog liegt der Fall bei den höheren Organismen, ja bei jedem Organismus...' [2]) Wenn die kybernetische Forschung zu der vorliegenden Frage überhaupt etwas beitragen kann, so sicherlich dies, daß äußerste Vorsicht am Platze ist hinsichtlich solcher Äußerungen, wie daß 'jeder gesunde Sinn' eine selbständige Weiterentwicklung von Maschinen ablehnen müsse. Der Umstand, daß es sich als möglich erwiesen hat, Maschinen zu erbauen, die sich in bezug auf bestimmte Fähigkeiten weiterentwickeln können und darin schließlich sogar ihren Erbauer zu übertreffen vermögen, bildet vielmehr einen deutlichen Hinweis darauf, daß auch der Vorgang der Selbstverbesserung und

[1]) Gemeint ist die Fortentwicklung der Organe von Lebewesen.
[2]) [6], S. 280.

Selbstdifferenzierung etwas darstellt, zu dessen Erklärung kein weiteres Rüstzeug erforderlich ist als physikalisch-chemische Gesetzmäßigkeiten. Vorläufig liegt hier allerdings erst eine noch zu bewältigende Aufgabe vor. Sie wird am besten wieder in eine philosophische und eine empirische zerlegt: So wie auf der zweiten Schicht der Begriff des ZO-Systems im Mittelpunkt stand und die philosophische Analyse diesen Begriff zu präzisieren suchte und die prinzipielle Struktur von ZO-Systemen zu beschreiben hatte, während die Analyse konkreter Spezialtypen solcher Systeme eine einzelwissenschaftliche Aufgabe blieb, so würde auf der dritten Schicht der Begriff des '*Selbstverbesserers*' oder '*Selbstdifferenzierers*' im Mittelpunkt stehen. Auch hier müßte sich die philosophische Analyse darauf beschränken, diesen Begriff zu präzisieren und zu zeigen, daß die in einem solchen Objekt stattfindenden Prozesse prinzipiell eine nichtteleologische kausale Erklärung gestatten, während die Beschäftigung mit den bekannten Fällen von Selbstdifferenzierern der empirischen Analyse überlassen bleiben müßte.

Angenommen, dieses Projekt sei in allen erforderlichen Einzelheiten realisiert worden. Dann ergibt sich auch eine neue Ausgangsbasis zur Beurteilung der anderen noch offenstehenden Frage, wie sich die erstmalige, ohne Beteiligung einer zwecksetzenden Instanz scheinbar ungeheuer unwahrscheinliche Entstehung eines Systems mit zielgerichteter Organisation erklären lasse. Denn *diese Frage ist jetzt zurückgeschoben worden auf die der Entstehung jener primitiven Formen solcher Systeme, bei denen der ganze Entwicklungsprozeß einsetzt.* Nach Voraussetzung sollen ja dieser Prozeß selbst und das Funktionieren der in seinem Verlauf entstehenden 'höheren' Formen von Organismen in nichtteleologischer Weise erklärbar sein. Obwohl die Frage solcher Entstehung eine rein empirische Angelegenheit bleibt, ebenso wie die andere Frage, ob wirklich an einer Stelle in abrupter Weise eine ganz neue Form von Systemen eintritt oder ob die Entwicklung von Organismen nur als die Fortsetzung eines weiter zurückreichenden kontinuierlichen Prozesses anzusehen ist, mögen doch zwei Dinge erwähnt werden, die hier von Relevanz sein dürften: Erstens ist die '*Zufallsentstehung*' jener Primitivformen von Organismen bei weitem nicht mehr mit jenem Maß an Unwahrscheinlichkeit behaftet wie es bei der zufälligen Entstehung von höheren Formen der Fall wäre. (Wenn hier von zufälliger Entstehung gesprochen wird, so ist damit natürlich dies gemeint, daß derartige Systeme aus Konstellationen der anorganischen Natur hervorgehen, wobei nur physikalisch-chemische Gesetzmäßigkeiten, jedoch kein zwecksetzendes Bewußtsein beteiligt sind). Zweitens lehrt die Erfahrung, daß Leben offenbar nur an einem sehr kleinen Teil des materieerfüllten Raumes zur Entstehung und Entwicklung gelangt und daß auch dort – wie etwa auf unserem Planeten – die Entstehung (jedoch nicht unbedingt die Weiterentwicklung) auf eine sehr kurze Zeitepoche beschränkt bleibt. Während ein häufigeres Auftreten jener höheren Systemformen im Universum tatsächlich Anlaß zu Verwunderung geben müßte, wenn man auf echte teleologische Erklärung verzichten wollte, dürfte dagegen dieser große kosmische Seltenheitswert des Lebens mit der statistischen Unwahrscheinlichkeit seines zufälligen Entstehens in Einklang zu bringen sein.

Hinsichtlich der teleologischen Aussagen der ersten und zweiten Schicht kann bereits heute die prinzipielle Übersetzbarkeit in nichtteleologische Aussagen behauptet werden. Dagegen bestehen für einen analogen Nachweis bezüglich der teleologischen Aussagen der dritten Schicht noch erhebliche Lücken. Bevor man sich anschicken kann, diese Lücken zu schließen, muß man sich Klarheit darüber zu verschaffen suchen, worin sie bestehen. Die vorangehenden Bemerkungen sollten als Hinweise in dieser Richtung aufgefaßt werden. Da es hier nicht darauf ankam, eine Theorie zur Lösung der dabei auftretenden Schwierigkeiten zu entwerfen, konnte auf eine Präzisierung verschiedener relevanter Begriffe ('Selbstverbesserer', 'Höherentwicklung', 'zufällige Entstehung') verzichtet werden.

Um teleologische Erklärungen von Vorgängen, an denen keine bewußten menschlichen Zielsetzungen beteiligt sind, generell als eine nachweislich bloße 'façon de parler' bezeichnen zu dürfen, muß somit gezeigt worden sein, daß die folgenden fünf Bedingungen erfüllt sind:

(1) Organismen und sonstige Naturgebilde mit zielgerichteter Organisation sind unter einen präzisierbaren Begriff des ZO-Systems subsumierbar;

(2) Die Struktur von ZO-Systemen und die in ihnen ablaufenden Prozesse können in einer nichtteleologischen Sprache vollständig und adäquat beschrieben und erklärt werden; insbesondere genügen für die Erklärungen (und eventuellen Prognosen) kausale Betrachtungsweisen, in denen Kausalgesetze i.w.S. zur Anwendung gelangen, jedoch niemals auf Motive zwecksetzender Wesen Bezug genommen wird;

(3) Alle natürlichen Prozesse der 'Höherentwicklung' können unter einen präzisierbaren Begriff des Selbstverbesserers oder Selbstdifferenzierers subsumiert werden;
(4) Die Struktur von und die Prozesse in Selbstdifferenzierern sind in derselben Weise einer rein kausalen Analyse zugänglich wie Struktur und Vorgänge in ZO-Systemen;
(5) Die Entstehung der einfachsten Fälle von ZO-Systemen, bei denen der zu höheren Formen führende Prozeß der Selbstdifferenzierung einsetzt, ist kausal erklärbar. Die für die Entstehung dieser Gebilde erforderliche Ausgangskonstellation besitzt eine hinreichende statistische Wahrscheinlichkeit, um an bestimmten Raum-Zeit-Gebieten des Universums zur Verwirklichung zu führen.
(5) bleibt eine rein empirische Angelegenheit, die in den Punkten (1) bis (4) angeführten Aufgaben dagegen zerfallen in wissenschaftstheoretische und erfahrungswissenschaftliche Teile.
Wenn im Vorangehenden immer wieder von kausalen Erklärungen i.w.S. bzw. von Kausalgesetzen i.w.S. die Rede war, so darf dies natürlich nicht so verstanden werden, als werde damit verlangt, daß die physikalischen Gesetze *in ihrer gegenwärtigen Fassung* ausreichen müßten, um alle scheinbar teleologischen Vorgänge zu erklären. Wie allgemein anerkannt wird, sind diese Gesetze keine verifizierten Tatsachen, sondern stellen Hypothesen dar, die durch das vorliegende Erfahrungsmaterial mehr oder minder gut bestätigt werden. Auch eine sehr gut bestätigte Hypothese kann sich aber auf Grund neuer Erfahrungsdaten als unhaltbar erweisen. Die Frage, ob eine 'kausale Theorie des Lebens' auf der Basis der heute akzeptierten Naturgesetze möglich sei, ist daher für das Problem der Teleologie genau so wenig von Relevanz, wie für die prinzipielle Voraussagbarkeit astronomischer Vorgänge die Frage von entscheidender Bedeutung ist, ob die bis heute bekannten Gesetze für alle diese Voraussagen ausreichen. Theoretisch ist sogar die Situation denkbar, daß zwei Klassen von physikalischen Theorien T_1 und T_2 für das anorganische Naturgeschehen dasselbe leisten, T_2 jedoch darüber hinaus die Erklärung bestimmter biologischer Phänomene ermöglicht, während T_1 keine derartigen Erklärungen liefert. Hier wäre der Theorie T_2 der Vorzug zu geben. Es würde sich bei einer derartigen Situation nur um einen speziellen Fall eines allgemeinen wissenschaftstheoretischen Sachverhaltes handeln: Wenn von zwei Theorien die erste die Erklärung

und Prognose bestimmter Arten von Vorgängen ermöglicht, die zweite jedoch nicht, und im übrigen beide Theorien dieselbe Leistungsfähigkeit besitzen, so ist die erste Theorie mit dem größeren Voraussagegehalt vorzuziehen.

Da im Zusammenhang mit den teleologischen Aussagen der dritten Schicht auch der 'teleologische Gottesbeweis' erwähnt worden ist, so möge abschließend zu diesem Punkt eine Bemerkung angefügt werden. Von empiristischer Seite würde dieses Argument vermutlich aus demselben Grunde zurückgewiesen werden wie alle anderen theologischen Argumente, nämlich weil darin mindestens ein Ausdruck verwendet wird, der gegen das empiristische Sinnkriterium verstößt. Angesichts der Problematik dieses Kriteriums würde ein solcher Einwand nur eine sehr geringe Durchschlagskraft besitzen. Es ist auch gar nicht erforderlich, sich bei der Kritik dieses Argumentes auf die empiristische Grundthese zu berufen; denn der eigentliche Mangel des Beweises liegt an anderer Stelle: Zunächst kann überhaupt nicht von einer bewiesenen Behauptung gesprochen werden, sondern nur von einer Hypothese. Diese soll allerdings dadurch zu einem beweisbaren Satz verschärft werden, daß man alle übrigen Hypothesen logisch ausschließt. Es bedarf keiner komplizierten wissenschaftstheoretischen Analysen, um einzusehen, daß ein derartiges Verfahren undurchführbar ist. Logisch ausgeschlossen darf nur dasjenige werden, was kontradiktorisch ist; die Ausschaltung der unübersehbaren Zahl logisch möglicher anderer Hypothesen und der Übergang zu der Behauptung, daß nur eine einzige dieser Hypothesen endgültig gesichert sei, stellt daher auf alle Fälle eine Voreiligkeit dar. Im vorliegenden Falle kann genau angegeben werden, worin diese Voreiligkeit besteht: Sie liegt im Übergang vom '*bisher* nicht auf nichtteleologische Weise kausal erklärbar' zum '*prinzipiell* nicht auf nichtteleologische Weise kausal erklärbar'. Der Fehler dieses Überganges besteht unabhängig davon, wie es mit dem Problem der Übersetzbarkeit teleologischer Aussagen aller Schichten in nichtteleologische Sätze steht. Sollte jedoch der Nachweis dieser prinzipiellen Übersetzbarkeit aller teleologischen Aussagen in nichtteleologische erbracht worden sein, so wäre damit zugleich gezeigt worden, daß es sich bei dem fraglichen Argument nicht nur um einen logisch nicht zu rechtfertigenden Schluß aus zutreffenden Voraussetzungen handelt, sondern um einen Fehlschluß aus unrichtigen Prämissen.

BIBLIOGRAPHIE

[1] Braithwaite, R. B., 'Teleological Explanation', *Arist. Proc.*, 47, 1947.

[2] Carnap, R., 'The Methodological Character of Theoretical Concepts', *Minnesota Studies in the Philosophy of Science*, 1956. Deutsche Übersetzung von A. Scheibal in: *Ztschr. f. Phil. Forschg.*, XIV/2, XIV/4, 1960.

[3] Feigl, H., 'Notes on Causality', *Readings in the Philos. of Sc.*, 1953.

[4] Goodman, N., '*Fact, Fiction & Forecast*', Cambridge, Mass. 1955.

[5] Hempel, C. G., und Oppenheim, P., 'The Logic of Explanation', *Philos. of Sc.*, 15, 1948. Teil I–III abgedruckt in: *Readings in the Philos. of Sc.*, 1953.

[6] Kastil, A., '*Die Philosophie Franz Brentanos*', München 1951.

[7] Macovschi, E., 'Wie die ersten Proteine auf der Erde entstanden sind', *Medizin. Monatsspiegel*, VII/I, 1958.

[8] Nagel, E., 'Mechanistic Explanation and Organismic Biology', *Philos. and Phen. Res.*, 9, 1951.

[9] Nagel, E., 'Teleological Explanation and Teleological Systems', *Readings in the Philos. of Sc.*, 1953.

[10] Rosenblueth, A., Wiener, N., und Bigelow, J., 'Behaviour, Purpose and Teleology', *Philos. of Sc.*, 10, 1943.

[11] Stegmüller, W., 'Das Problem der Kausalität', *Probleme der Wissenschaftstheorie* (Festschrift für V. Kraft), Wien 1960.

[12] Stegmüller, W., 'Conditio Irrealis, Dispositionen, Naturgesetze und Induktion', *Kant-Studien*, 50/3, 1959.

KARL DÜRR

BELEUCHTUNG VON ANWENDUNGEN DER LOGISTIK IN WERKEN VON RUDOLF CARNAP

Von Anwendungen der Logistik sprechen wir dann und nur dann, wenn entweder einzelne Sätze oder Systeme von Sätzen gebildet werden, in denen Zeichen einer logistischen Sprache in Verbindung mit nichtlogischen Konstanten auftreten.

In den Werken von R. Carnap finden sich zahlreiche Anwendungen der Logistik. Es seien hier drei seiner Werke genannt, in denen in systematischer Weise Anwendungen der Logistik dargestellt sind, nämlich:

1) Abriß der Logistik 1929
2) Einführung in die symbolische Logik 1954
3) Einführung in die symbolische Logik
   Zweite umgearbeitete und erweiterte Auflage 1960

Jedes der drei Werke enthält einen zweiten Teil, der als 'angewandte Logistik' oder als 'Anwendungen der symbolischen Logik' bezeichnet ist. Diese Werke werden uns hier als Quellen dienen. In Hinweisen auf diese Werke benutzen wir als Abkürzung für den Titel 1): 'Abriß', für 2): 'Einführung 1' für 3): 'Einführung 2'.

Es sei hier betont, daß die Möglichkeit von Anwendungen der Logistik in außerlogischen Gebieten bedeutsam ist für die Stellung, welche dieser Disziplin in dem gesamten Bereich der Wissenschaften zukommt; die Realisierungen solcher Möglichkeiten, die in den genannten Werken zu finden sind, sind deshalb für die Stellung der Logistik im Bereich der Wissenschaften wesentlich.

Die Übersetzung einzelner Sätze einer Wortsprache, etwa der deutschen Sprache, in die logistische Sprache kann als eine Anwendung der Logistik betrachtet werden; und man findet im zweiten Teil der genannten Werke eine Fülle von Fällen, in denen dieses Verfahren zur Anwendung kommt.

In den Übersetzungen gegebener Sätze, die im Abriß dargestellt sind, treten lediglich logistische Zeichen und nicht-logische Konstanten auf, und es sind die nicht-logischen Konstanten in eigentümlicher Weise gebildet. Sofern eine nicht-logische Konstante, die in dem betreffenden Satz der Wortsprache auftritt, ein Zeichen für ein Individuum ist, wird dieses Zeichen entweder bei der Übersetzung beibehalten, aber als eine undeklinierbare Partikel behandelt, oder es wird ein kleiner lateinischer Buchstabe, etwa 'a' oder 'b', in der logistischen Sprache als eine Konstante, die ein Individuum bezeichnet, verwendet: dagei bleibt es zweifelhaft, ob man sich zu denken hat, daß der betreffende Buchstabe schon in dem Satz der Wortsprache, der zu übersetzen war, auftrat, oder ob der gegebene Satz, der übersetzt werden soll, selbst schon as als ein Satz zu betrachten ist, der eigentlich nicht mehr der Wortsprache angehört, sondern nur den Übergang von einem Satz der Wortsprache zu einem Satz der logistischen Sprache erleichtern soll. Nicht-logische Konstanten des gegebenen Satzes, die entweder Klassenausdrücke oder Ausdrücke zwei- oder mehrstelliger Relationen sind, werden bei der Übersetzung ersetzt durch Ausdrücke, die nach einer bestimmten Regel, die im Abriß (S. 92) dargestellt ist, zu bilden sind. Dieser Regel zufolge bleibt der betreffende Ausdruck der Wortsprache bei der Übersetzung in die logistische Sprache in gewissem Sinn erhalten, wird dabei aber doch in solcher Weise transformiert, daß durch gewisse beigefügte Schriftzeichen kenntlich gemacht wird, ob der betreffende Ausdruck ein einstelliges oder ein zweistelliges oder ein mehrstelliges Prädikat ist. In dieser Regel verkörpert sich die Tendenz, Sätze in der logistischen Sprache so darzustellen, daß ihre logische Struktur erkennbar wird.

Als Beispiel eines Satzes, der im Abriß dargestellt und in der angedeuteten Weise behandelt wird, sei hier folgender Satz genannt: 'Ich habe denselben Lehrer wie du'. Es wird zunächst darauf hingewiesen, daß dieser Satz der Wortsprache zwei verschiedene Bedeutungen haben kann, und es werden dann zwei Sätze der logistischen Sprache dargestellt, die eindeutig sind und von denen ein jeder mit einer und nur einer der beiden Bedeutungen verknüpft ist. In jedem der beiden Sätze, die als Übersetzungen des Satzes der Wortsprache in die logistische Sprache gelten können, treten lediglich logistische Zeichen und eine nicht-logische Konstante auf; die nicht-logische Konstante wird gebildet, indem ein Wort der deutschen Sprache, nämlich das Wort 'Lehrer' in solcher Weise dargestellt wird, daß sich erkennen läßt, daß dieser Ausdruck in der logistischen Sprache als Bezeichnung einer zweistelligen Relation aufzufassen ist. (vgl. Abriß S. 93).

In Einführung 1 und Einführung 2 wird die Frage, wie Sätze der Wortsprache in Sätze der logistischen Sprache zu übersetzen sind, in streng systematischer Weise behandelt; dieser systematischen Behandlungsweise liegt das Prinzip zugrunde, daß in diesem Zusammenhang die Verschiedenheit möglicher Sprachformen zu berücksichtigen ist. Es empfiehlt sich, hier von der Unterscheidung von Dingsprachen und Koordinatensprachen auszugehen; doch ist zu sagen, daß innerhalb jeder dieser Sprachformen feinere Unterteilungen möglich sind. (vgl. Einführung 1, S. 135 und S. 138 und Einleitung 2, S. 159 und S. 163) Im folgenden beschränken wir uns auf die Betrachtung von Sätzen, die den Dingsprachen angehören. Unter diesen Sätzen finden sich auch solche, in denen quantitative Begriffe auftreten, doch werden wir hier nicht auf die diesen Sätzen eigentümliche Beschaffenheit eingehen.

Wir betrachten die verschiedenen Arten von Dingsprachen als Arten der logistischen Sprachen und stellen die logistischen Sprachen den Wortsprachen gegenüber. Den Sätzen der Wortsprachen entsprechen Sätze der logistischen Sprachen, die als ihre Übersetzungen zu bezeichnen sind. Aber die Relation, die durch das Wort 'Übersetzung' ausgedrückt wird, ist nicht eine eineindeutige, sondern eine mehreindeutige, d. h. es können mehrere Sätze der logistischen Sprache Übersetzungen desselben Satzes der Wortsprache sein. Man kann dies auch so ausdrücken, daß man sagt, daß ein Satz der Wortsprache mehrere Bedeutungen haben kann, und daß eine Übersetzung dieses Satzes in die logistische Sprache immer nur eine dieser Bedeutungen hat.

Die Unterscheidung der Hauptarten von Dingsprachen und der Unterarten dieser Hauptarten, die in den genannten Werken durchgeführt wird, stützt sich auf zwei Gesichtspunkte. Der eine dieser Gesichtspunkte betrifft die Festsetzungen, durch die bestimmt wird, was im Rahmen der Sprachformen, die man im Auge hat, als Individuum zu betrachten ist. Dieser Gesichtspunkt ist der fundamentale, auf den insbesondere die Unterscheidung der drei Hauptarten von Dingsprachen begründet ist. Der zweite Gesichtspunkt betrifft die Deutung des Begriffes 'Ding'. Auf diesen Gesichtspunkt gründet sich die Unterscheidung gewisser Unterarten von Dingsprachen.

Es sei hier zunächst hervorgehoben, daß sich unter den Unterarten der Dingsprachen eine Sprachform findet, der in gewisser Hinsicht eine Sonderstellung einzuräumen ist. Es ist dieser Sprachfoım anderen Sprachformen gegenüber eigentümlich, daß sie nur dann als Ausdrucksmittel für die Beschreibung von Tatsachen dienlich ist, wenn der Wandelbarkeit der Dinge nicht Rechnung zu tragen ist. Es sei uns erlaubt, diese Sprachform hier kurz die primitive Dingsprache zu nennen; wir können dann die übrigen Sprachformen der Dingsprachen als die nicht-primitiven kennzeichnen.

In Einführung 1 und Einführung 2 werden zwei Sätze dargestellt, die als Beispiele dienen können, nämlich als Beispiele von Sätzen der deutschen Wortsprache, denen in jeder Form der Dingsprachen, die zu den nicht-primitiven zu zählen ist, genau ein Satz entspricht, der als eine Übersetzung des betreffenden Satzes der deutschen Wortsprache in die logistische Sprache gelten kann. (vgl. Einführung 1, S. 137 und Einführung 2, S. 161) Die beiden Sätze der deutschen Wortsprache, die als Beispiele dienen, nennen wir kurz die Paradigmen.

Zwei Sätze der logistischen Sprache, die Übersetzungen desselben Paradigmas sind, können verschieden sein, sind es aber nicht in allen Fällen. Von Interesse für uns sind hier insbesondere die Fälle, in denen die beiden Übersetzungen nicht zu identifizieren sind.

In diesen Fällen rührt die Verschiedenheit der Übersetzungen daher, daß demselben Wort des Paradigmas in den beiden Übersetzungen verschiedene Zeichen entsprechen, nämlich Zeichen, von denen sich etwa sagen ließe, daß sie zwar derselben Wortfamilie angehören, aber nicht Ausdrücke desselben Typus sind. Dieser Fall liegt beispielsweise dann vor, wenn das eine der beiden logistischen Zeichen, die hier zu vergleichen sind, eine Konstante ist, die ein Individuum bezeichnet und das andere Zeichen ein Prädikat ist; doch auch dann, wenn beide Zeichen Prädikate sind, können sie Ausdrücke sein, die nicht denselben Typus haben. Sind etwa die beiden Prädikate, die zu vergleichen sind, Prädikate, die nicht dieselbe Stellenzahl haben, oder sind es Prädikate verschiedener Stufe, so werden wir auch sagen, daß sie Ausdrücke sind, die nicht denselben Typus haben. In allen diesen Fällen können die beiden Sätze, die Übersetzungen desselben Paradigmas sind, aber verschiedenen Formen der Dingsprachen angehören, nicht identifiziert werden. Diese Feststellung ist deshalb bedeutsam, weil sich daraus folgern läßt, daß in manchen, wenn auch nicht notwendigerweise in allen Fällen ein Satz der deutschen Wortsprache sich in mehr als nur einer Weise deuten läßt.

Es braucht hier kaum betont zu werden, daß, was von der deutschen Wortsprache gilt, auch von anderen Wortsprachen, beispielsweise von der englischen und der französischen, gelten wird.

Ein sehr bedeutsames und instruktives Beispiel der Anwendung der Logistik auf ein außerlogisches Gebiet ist ein System, das als ein Axiomensystem der Raum-Zeit-Topologie, genauer als das 'K-Z-System' bezeichnet wird. Dieses System ist in den drei Werken, die unsere Quellen sind, dargestellt. In jeder dieser drei Darstellungen wird verwiesen auf eine frühere Abhandlung von R. Carnap, die betitelt ist 'Über die Abhängigkeit der Eigenschaften des Raumes von denen der Zeit'. (Kantstudien, XXX, 331–345, 1925) Wir nennen im folgenden diese Schrift kurz 'die Abhandlung' und zitieren sie unter dem Titelwort 'Abhandlung'. Es ist offensichtlich, daß ein enger thematischer Zusammenhang besteht zwischen dieser Abhandlung und den drei späteren Darstellungen. Zum Beweise dafür sei darauf hingewiesen, daß in der Abhandlung in gleicher Weise wie in den späteren Darstellungen die beiden Symbole 'K' und 'Z' als Zeichen der beiden außerlogischen Grundrelationen auftreten. (a. a. O., S. 336) Wie schon angedeutet wurde, wird das System, das hier zu betrachten ist, als das K-Z-System bezeichnet; daran erkennt man, daß die Relationen K und Z zur Charakteristik des Systemes geeignet sind.

Unter dem Titel 'Literaturhinweis' bringt die Abhandlung folgende Ankündigung: 'Die Durchführung des Systems der Raum-Zeit-Topologie auf Grund der Relationen K und Z soll in einer späteren Abhandlung gegeben werden unter dem Titel 'Topologie der Raum-Zeit-Welt, axiomatisch dargestellt mit den Mitteln der symbolischen Relationstheorie.' (a. a. O., S. 345) Die mir zugänglichen Bibliographien der Schriften R. Carnaps bringen keinen Hinweis darauf, daß eine solche Abhandlung je erschienen sei; doch läßt sich sagen, daß das Thema der geplanten Abhandlung in je einem Abschnitt der drei Werke, auf die wir verwiesen haben, zur Behandlung gekommen ist.

Die Darstellungen, die das K-Z-System in den drei Werken gefunden hat, sind einander nicht in jeder Hinsicht gleich. Hier ist insbesondere auf die Tatsache hinzuweisen, daß in den beiden neueren Darstellungen die einzelnen Sätze vielfach doppelt dargestellt werden, nämlich zuerst in der Sprache A, dann in der Sprache C. Die Sprache C kann der Sprache A gegenüber als eine erweiterte Sprache bezeichnet werden, da sie über mehr Ausdrucksmittel verfügt als die Sprache A. Die Unterscheidung zweier Sprachen, in denen die einzelnen Sätze dargestellt werden können, findet sich noch nicht im Abriß.

Die Sprachen A und C sind symbolische Sprachen. Jedes elementare Zeichen dieser Sprachen ist entweder ein logisches Zeichen oder eine nicht-logische Konstante; zu den nicht-logischen Zeichen dieser Sprachen gehören insbesondere die beiden Zeichen der Grundrelationen, d. h. die Zeichen 'K' und 'Z'. Diesen Sprachen entspricht im Abriß eine einzige Sprache, die auch eine symbolische Sprache ist. Es empfiehlt sich kürzehalber zur Bezeichnung dieser Sprache ein Zeichen einzuführen; wir nennen sie die die Sprache 'D'. Es ist nun aber an dieser Stelle noch auf eine vierte Sprache zu verweisen, die man im Gegensatz zu den Sprachen A, C und D, welche die Objektsprachen sind, die Metasprache nennen kann. Diese Metasprache ist die deutsche Sprache, die aber hier als eine Sprache zu betrachten ist, deren Zeichenbestand nicht abgeschlossen ist, sondern erweitert werden kann, wenn Anlaß zur Erweiterung gegeben ist. Erweitert wird der Wortschatz der deutschen Sprache im vorliegenden Fall dadurch, daß gewisse Ausdrücke eingeführt werden, die eigentlich Übersetzungen bestimmter Ausdrücke der Objektsprachen sind. Um ein Beispiel dafür zu geben, sei auf folgenden Umstand hingewiesen. Der Ausdruck 'trans' ist ein Ausdruck der Sprache D und 'Trans' ein Ausdruck der Sprache C; ihnen entspricht in der erweiterten Metasprache als ihre Übersetzung der Ausdruck 'transitiv', der zwar in grammatikalischer Bedeutung auch schon zuvor der deutschen Sprache angehörte, aber in der Metasprache in neuer Bedeutung verwendet wird. Es läßt sich nun sagen, daß sowohl im Abriß als auch in Einführung 1 und Einführung 2 Sätze oder Formeln der Objektsprachen vielfach noch in der erweiterten Metasprache dargestellt sind. Dadurch, daß in manchen Fällen den Formeln der symbolischen Sprache eine Übersetzung in die deutsche Metasprache beigegeben ist, wird das Verständnis der Formeln erleichtert; es muß aber doch gesagt werden, daß streng genommen die Formel und Ihre Übersetzung in die deutsche Metasprache nicht in jeder Hinsicht zu identifizieren sind.

Die Sätze des K-Z-Systems lassen sich in drei Klassen aufteilen, nämlich in die Klasse der Axiome, die Klasse der Theoreme und die Klasse der Definitionen. Die Klasse der Axiome bildet das Axiomensystem im engeren Sinn. Wir beschränken uns hier auf die Betrachtung der Definitionen. Die Definitionen bilden für sich wieder ein System, das nach einer strengen Methode aufgebaut ist. Das ist die Methode, die hier als eine Anwendung der Logistik zu beleuchten ist.

Es sei uns erlaubt, an dieser Stelle eine kurze historische Digression einzuschalten. Heinrich Scholz hat in einem Artikel, der betitelt ist 'Pascals Forderungen an die mathematische Methode' (Festschrift zum 60. Geburtstag von Prof. Dr. Andreas Speiser 1945) hingewiesen auf Pascals Methodenlehre, die in dem Werk 'De l'esprit géométrique' dargestellt ist. Für uns sind insbesondere zwei Regeln bedeutsam, die in dem zweiten der beiden Fragmente, die in dem genannten Werk vereinigt sind, formuliert werden. Wir werden diese Regeln in der Übersetzung von Wolfgang Struve wiedergeben; doch werden wir dabei auf die Wiedergabe derjenigen Teile des Textes, die für uns nicht wesentlich sind, verzichten. Der Text erhält dann folgenden Wortlaut:
Notwendige Regeln für die Definitionen.
Bei den Definitionen nur vollkommen bekannte oder schon erklärte Begriffe verwenden.
Notwendige Regeln für die Beweise.
Alle Behauptungen beweisen, indem man zu ihrem Beweis nur Axiome oder schon bewiesene Behauptungen verwendet. (Blaise Pascal, 'Vom Geist der Geometrie', Übersetzt und eingeleitet von Wolfgang Struve, 1948, S. 73; nicht mehr im Buchhandel erhältlich)
Unter vollkommen bekannten Begriffen sind hier diejenigen Begriffe zu verstehen, für die Pascal als terminus technicus den Ausdruck 'Mots primitifs' eingeführt hat. (ebd. S. 22)
H. Scholz hat in dem genannten Artikel hervorgehoben, daß Pascal der erste gewesen ist, der der Axiomatik im engeren Sinn eine Theorie der Definitionen zur Seite gestellt hat im vollen Bewußtsein um den Parallelismus, der das Verhältnis der Theorie der Definitionen zur Axiomatik beherrschen muß. (a. a. O., S. 26).
Es sei hier bemerkt, daß die Analogie, die H. Scholz im Auge hat, in schönster Weise in den Ausführungen über die axiomatische Methode im Abriß hervortritt. (a. a. O., S. 70)
Wesentlich ist für uns die in Pascals Regel der Definitionen liegende Vorschrift, in Definitionen, d. h. im Definiens von Definitionen nur Begriffe zu verwenden, die entweder Grundbegriffe sind oder zuvor schon definiert worden sind. Indem wir den modernen Ausdruck 'Kettendefinition', der im Abriß gebildet wird, verwenden, erklären wir, daß Pascals Vorschrift dazu führt, Kettendefinitionen aufzustellen. (vgl. Abriß, S. 70) Es liegt in Pascals Regel der Definitionen ein Element, das

auch in der modernen Axiomatik und insbesondere im K-Z-System, mit dem wir uns hier beschäftigen, zur Geltung kommt.
Wir führen an dieser Stelle die beiden Ausdrücke 'Koinzidenz' und 'Eigenzeit' ein und setzen im Anschluß an die Terminologie R. Carnaps fest, daß 'K' als das Zeichen der Koinzidenz und 'Z' als das Zeichen der Eigenzeit zu betrachten sei. (vgl. Abhandlung, S. 336 und S. 338)
Da K und Z als Grundrelationen gelten, müssen streng genommen die Zeichen 'K' und 'Z' als undefiniert, d. h. als Grundbegriffe betrachtet werden. Es schiene nun aber doch unbefriedigend, wenn nicht vor der Entwicklung des begrifflichen Systems eine mögliche Deutung der Grundrelationen gegeben würde. Eine solche mögliche Deutung der Grundrelationen findet man in der Tat durchweg in den Schriften, die hier heranzuziehen sind, d. h. im Abriß, in Einführung 1, in Einführung 2 und auch schon in der Abhandlung. Um deutlich hervortreten zu lassen, daß diese Erklärungen der Grundrelationen nicht zu den Definitionen des begrifflichen Systems gehören, wollen wir sie nicht Definitionen, sondern Erläuterungen nennen. In diesen Erläuterungen werden nun die Ausdrücke 'Weltpunkt', 'Weltlinie' und 'genidentisch' zur Erklärung der beiden Grundrelationen verwendet, während im definitorischen System die Begriffe 'genidentisch' und 'Weltlinie' mit Hilfe des Zeichens einer der beiden Grundrelationen bestimmt werden. Der Begriff 'Weltpunkt' wird im Abriß auch mit Hilfe des Zeichens einer der Grundrelationen definiert; in Einleitung 1 und in Einleitung 2 wird er nicht definiert und lediglich in der Metasprache verwendet.
Das eigentümliche Verfahren, das hier bei der Einführung der Zeichen der beiden Grundrelationen angewandt wird, erfährt seine Rechtfertigung dadurch, daß auf diesem Weg ein Zugang zu einer möglichen Deutung und damit zum Verständnis des begrifflichen Systems eröffnet wird.
Wir fassen nun die zum K-Z-System gehörigen Ketten von Definitionen ins Auge.
Die Unterscheidung logischer und nicht-logischer Zeichen wird in diesem Zusammenhang wesentlich. Von dieser Unterscheidung haben wir schon vielfach Gebrauch gemacht; wir heben an dieser Stelle hervor, daß wir uns im Gebrauch dieser Ausdrücke an die Terminologie, die in den genannten Werken R. Carnaps eingeführt und befolgt wird, anschließen.
Es sei hier noch bemerkt, daß wir im folgenden das System der Definitionen

des K-Z-Systems, das in Einführung 2 dargestellt ist, vor Augen haben. In dem vorliegenden Abschnitt sind acht Definitionen dargestellt und es sind Anweisungen gegeben, welche die Bildung weiterer sieben Definitionen ermöglichen. Es ist nun aber festzustellen, daß als Glieder von Ketten von Definitionen nur die acht Definitionen, die dargestellt sind, in Betracht kommen. Die Unterscheidung, die hier zu machen ist, gründet sich auf die Unterscheidung logischer und nicht-logischer Symbole. Denn bei der Anordnung von Definitionen in Ketten zählen hier nur die nicht-logischen Konstanten. Die sieben Definitionen, die nicht dargestellt sind, sich aber nach den Anweisungen, die gegeben werden, bilden lassen, enthalten nun keine nicht-logischen Konstanten und können darum hier nicht zu Gliedern von Ketten von Definitionen gemacht werden.

Die acht Definitionen, die dargestellt und numeriert sind, sind nicht Glieder ein und derselben Kette von Definitionen, sondern es bilden die erste, zweite und dritte Definition, durch welche die drei Begriffe 'genidentisch', 'Weltlinie' und 'Weltlinienreihe' bestimmt werden, die erste Kette von Definitionen; und es bilden die vierte bis achte Definition, durch welche die fünf Begriffe 'Wirkungsrelation', 'gleichzeitig', 'Raum' 'Wirkungsgebiet' und 'Umgebung' bestimmt werden, eine zweite Kette von Definitionen.

Diese Feststellung läßt sich in folgender Weise begründen.

Die erste Definition bestimmt den Begriff 'genidentisch'; im Definiens dieser Definition erscheint außer logischen Zeichen nur das Zeichen der Eigenzeit; da dieses Zeichen eine nicht-logische Konstante ist, ist 'genidentisch' ein nicht-logischer, d. h. deskriptiver Begriff.

Die zweite Definition bestimmt den Begriff 'Weltlinie'; im Definiens dieser Definition erscheint außer logischen Zeichen nur der Begriff 'genidentisch'; da dieser Begriff deskriptiv ist, ist auch der Begriff 'Weltlinie' deskriptiv.

Die dritte Definition bestimmt den Begriff 'Weltlinienreihe'; im Definiens dieser Definition erscheint außer logischen Zeichen nur der Begriff 'Weltlinie'; da dieser Begriff deskriptiv ist, ist auch der Begriff 'Weltlinienreihe' deskriptiv.

Auf Grund dieser Feststellungen ergibt sich, daß die erste, zweite und dritte Definition eine Kette von Definitionen bilden, deren erstes Glied die erste, deren zweites Glied die zweite und deren drittes Glied die dritte Definition bildet.

Die vierte Definition bestimmt den Begriff der Wirkungsrelation. Im Definiens dieser Definition erscheinen außer logischen Zeichen die Zeichen der Koinzidenz und der Eigenzeit. Da diese beiden Zeichen deskriptiv sind, ist auch der Begriff der Wirkungsrelation deskriptiv.

Die Beschreibung der vier noch folgenden Definitionen darf jetzt wohl kürzer gefaßt werden als die Beschreibungen im Vorangehenden gefaßt worden sind. Wir heben darum lediglich hervor, daß in allen Fällen im Definiens einer Definition das Definiendum der ihr unmittelbar vorangehenden Definition erscheint und daß darum, weil das Definiens der vierten Definition ein deskriptiver Begriff ist, alle Begriffe, die durch die vierte bis achte Definition bestimmt werden, deskriptiv sind. Es ergibt sich nun, daß die fünf letzten Definitionen Glieder einer Kette von Definitionen sind. Es ist auch offensichtlich, daß sich nicht eine neue Kette von Definitionen in der Weise bilden läßt, daß man auf das letzte Glied der ersten Kette das erste Glied der zweiten Kette unmittelbar folgen läßt; denn das Definiendum des letzten Gliedes der ersten Kette erscheint nicht im Definiens des ersten Gliedes der zweiten Kette.

Es ist nun aber im Hinblick auf die acht Definitionen, die wir hier im Auge haben, noch auf ein Element der Kunst des Definierens aufmerksam zu machen; das ist die Art der Verwendung der logischen Zeichen beim Aufbau des Definiens der Definitionen. Daß im Definiens von Definitionen, deren Definiendum ein deskriptiver Begriff ist, außer den deskriptiven Begriffen auch logische Zeichen vorkommen, ist ein wesentlicher Umstand, und in diesem Punkt geht nun die moderne Kunst des Definierens, die sich auf die Logistik stützt, weit über das hinaus, was die alte Kunst des Definierens, wie sie etwa Pascal vorschweben mochte, zu leisten vermochte.

Es ist hier auf den Umstand hinzuweisen, daß sich die Klasse der logischen Zeichen ebensowohl wie die Klasse der deskriptiven Zeichen in die Klasse der undefinierten und die Klasse der definierten Zeichen aufteilen läßt. Es ist offensichtlich, daß sich im Gebiet der logischen Zeichen ebenso wie im Gebiet der deskriptiven Zeichen Ketten von Definitionen bilden lassen; nur werden in diesem Fall im Definiens und darum auch im Definiendum keine deskriptiven Zeichen mehr auftreten. Aber bei der Deutung des Definiens einer Definition, die deskriptive Zeichen enthält, wird man vielfach gezwungen sein, auf die Definitionen der logischen Zeichen, die im Definiens auftreten, einzugehen und eventuell wird man

sogar Ketten von Definitionen verfolgen müssen, um definierte logische Zeichen, die im Definiens auftreten, sukzessive zu eliminieren.
Ein Beispiel einer Definition, deren Definiens in der angedeuteten Weise zu behandeln ist, ist die Definition des Begriffes 'Wirkungsrelation', welche das erste Glied der zweiten Kette von Definitionen, die wir beschrieben haben, bildet.
Die betreffende Definition ist dargestellt in Einführung 2 (S. 203) und dort mit dem Zeichen 'D 4' bezeichnet. Wir erwähnen, daß im Definiens dieser Definition ein zusammengesetzter Ausdruck auftritt, der sich etwa als eine unbestimmte positive Potenz eines Ausdruckes, der selbst als eine Verkettung der Zeichen der beiden Grundrelationen ist, kennzeichnen läßt. Die Deutung dieses Ausdruckes wird möglich mit Hilfe einer Kette von Definitionen, die drei Glieder hat. Das Definiendum des ersten Gliedes der Kette ist der Begriff 'Erblichkeit einer Eigenschaft in bezug auf die Relation R', wobei das Zeichen der Relation R mit der gegebenen Verkettung der beiden Grundrelationen zu identifizieren ist. Das Definiendum des dritten Gliedes der Kette ist der zusammengesetzte Ausdruck, dessen Bedeutung durch die dreigliederige Kette von Definitionen zu bestimmen ist.
Wenn man sich nun das System der beiden ersten Ketten von Definitionen und den Aufbau des Definiens der einzelnen Glieder dieser Ketten vergegenwärtigt, so wird man zu der Überzeugung gelangen, daß das axiomatische System der Definitionen des K-Z-System als ein Werk hoher logischer Präzision gelten darf.

Universität Zürich

ARNE NAESS

TYPOLOGY OF QUESTIONNAIRES ADOPTED TO THE STUDY OF EXPRESSIONS WITH CLOSELY RELATED MEANINGS

If carried out with an eager and open mind, painstaking empirical research leads us into vast uncharted regions of facts and relations. The more one penetrates into the thickness of such regions, the more one is fascinated. One is – often against one's will – drawn further and further into the study of details and intricate structures revealed by the data found or collected. Bystanders are often astonished at this: What has gradually broadened into a whole world is, seen from outside, only a secondary and special field or at least a field of no importance for any great problems. And the outsider is right: it is only very rarely that a piece of empirical research obtains a great weight in solving or clarifying central problems.

Much that is being done, and especially in the methodologically less advanced departments of science, is not very interesting or attractive to more than a few people working in exactly the same region. This holds good of some studies carried out by Socratic interviews or by (standardized) questionnaires intended to reveal more and more about the relations between words of closely related (cognitive) meaning. It is my contention, however, that in spite of the in many ways appalling crudity of the questionnaire techniques and in spite of the manifest inability of many subjects to enter into difficult linguistic or other fields, the data gathered are often apt to reveal or suggest as much to the researcher as do penetrating meditations or introspections based on data found in one's own head or gathered in an informal way. But first of all it is the very richness of the materials, the surprising differences among 'ordinary people' (non-philosophers and non-linguists) in how they conceive terms, phrases, texts and functions of language – together with uniformities at places one could not have foretold – which spurs one to renewed efforts.

In this article I shall try to give a condensed survey of most kinds of questionnaires used in my studies of synonymity and of some kinds used by other authors (Revie, L. Lövestad, H. Tennessen). The list forming the main part of this article will, I hope, facilitate the work of other researchers who wish to contribute to empirical semantics. There is much confusion about how questionnaires can be used and an underestimation of the variety of approaches or techniques.

In particular, there are three old misconceptions to be fought: (1) If one wishes to know how a person uses a term, the questionnaire method consists in asking the person directly how he uses it. (2) If the formulations of the questionnaire are misinterpreted, there is nothing one can do with it in order to find out how and why. (3) When a researcher asks a subject a set of questions, this set is identical with the questions the researcher asks himself in his research program.

'Questionnaire' is in this paper taken in a wide sense: a set of formulations of questions in a definite order which are intended to be placed before an identified group of people under conditions – verbal and non-verbal – which are to some extent known or tested or standardized. It is a research instrument which is adapted to the need of asking more than one person the same questions one or more times under various conditions.

If controversial questions are asked, this does not imply that the researcher believes he can get the correct answer by using questionnaires containing just those questions. Very often the researcher is interested primarily in questions which are not put into the questionnaire. Thus, only a minority of our questionnaires pose direct questions to the subjects such as 'Does T and U express the same meaning to you in this context?', and if direct questions are asked, the answers are not taken at face value. The material obtained may furnish evidence in an indirect way, it may strengthen or weaken an argument which is relevant (but scarcely decisive) for the questionnaire. Last, but not least, there is nothing in the technique of questionnaires which precludes that competent people are asked. Professionals and experts do, however, tend to refuse answering isolated questions and insist upon reformulating, modifying and expanding the formulations of the questionnaire in a way that make *comparison* with other respondents very difficult. This is not any obstacle of major importance, but it certainly makes it unwise to use professionals as subjects except when valuable for the specific aim which has lead the researcher to frame the questionnaire.

The following classificatory system has suggested itself as the number of synonymity questionnaires which I developed in the course of research, increased and it began to be difficult to survey their similarities and differences. It is reasonable to expect that other classifications will be more convenient or theoretically important at later stages of research.

In studying closeness in meaning by means of questionnaires it has been one of our aims to develop at least one kind of questionnaire corresponding to each kind of well-known philological and philosophical criterion or definition of (cognitive) synonymity. One might have expected that this aim would result in questionnaires with 'crucial' questions reminding of the formulations used in those criteria or definitions. The indirect character of most (good) questionnaires and the non-operational character of the criteria or definitions make the relation between professional definitions rather intricate, however.

An example will illustrate this. Carnap introduces a concept *intentional isomorphism* which may be said to be a kind of very strong relation of sameness of meaning. Why is there no questionnaire directly corresponding to it, that is, a standardized set of questions including, and only including, what must be answered in order to identify a case of logical isomorphism? Firstly, the concept of intentional isomorphism is constructed in relation to sets of semantical *rules*, not *use*. Just as the concepts of L-implication and L-equivalence, they have no meaning except for formal languages. Now, questionnaires are instruments for finding out how people *use* certain expression, not to ascertain which are the formation- and transformation rules of calculi (which possibly never have been used by anybody). Secondly, if the concept of logical isomorphism is somewhat changed so as to be relevant for investigations of use, there will not be one characteristic questionnaire corresponding to it, but a characteristic *set* of ordinary questionnaires: if two complex expressions $x_1y_1z_1$ and $x_2y_2z_2$ can be analyzed each into three components, x_1, y_1, z_1 and x_2, y_2, z_2, the concept of logical isomorphy can be said to require not only synonymity of the pairs $x_1y_1z_1$ and $x_2y_2z_2$ but also of the three pairs of minimum expressions x_1x_2, y_1y_2 and z_1z_2.[1]) This means that we may use questionnaires of any of the families surveyed in the following, but in characteristic quadruplets.

Terminology

Individual copies of questionnaires have the same status semantically as

[1]) Cp. A. Naess, 'Logical Equivalence, Intentional Isomorphism and Synonymity as studied by Questionnaires', *Synthese*, vol. Xa, p. 476.

individual occurrences of sentences. There may be a vast number of copies of them. We shall think of questionnaires as written. It is of importance, however, to remember that when used orally, they give rise to acoustic processes, each one with its peculiarities.

Each questionnaire (Q) has a name; those used in studying synonymity (s), in a vague and broad sense of closeness in meaning, have the name 'Qs' and two numbers, as in 'Qs 4 No. 3', 'Qs 4' is to be considered as an abbreviation for 'Qss4'', the second 's' for 'statement'. Questionnaires of the class Qss study likeness of meaning between statements. Closely corresponding questionnaires concerning synonymity of designation (terms) get the name 'Qsd 4' etc. Similarly, 'i' in 'Qsi4' and 'q' in 'Qsq4' stand for 'imperative sentence' and 'question sentence' respectively.

In order to avoid unnecessary complications, questionnaires identical with Qs 4 but for very minute or peripherical modifications of wording, have been called 'Qs 4-questionnaires'.

Each synonymity questionnaire concerns a pair of sentences or designations, the so-called 'crucial expressions', the synonymity of which is to be investigated, for instance 'true' and 'perfectly certain'. The second number in the questionnaire name changes with changing expressions to be investigated or with changing order of expressions, Qsdl No. 11 investigates the use of 'true' – 'perfectly certain'. Qsd1 No. 12 investigates 'perfectly certain' – 'true', Qsd1 No. 13 'the case' – 'perfectly certain', Qsd1 No. 14 'the case' – 'true', etc.

Just as in the case of sentences, their meaning and the attitude towards them change with the context, therefore the order in which the questions are posed may affect the answer.

The name 'Qsd 1' is (1) a class-designation denoting all questionnaires which are generated by inserting definite designations into two open places in a sentence sequence. The sentence sequence is the same for all the particular questionnaires, the open places are those of the crucial expressions. 'Qsd 1' is also (2) a name for the skeletal sentence systems (sentence schemes) we get by taking away the crucial expressions from the particular questionnaires Qsd1 No. 1, Qsd1 No. 2, etc. If nothing else is explicitly said, 'Qsd 2' is intended used as a class-designation. Broader classes are designated by such symbols as Qsxy – the class of all kinds of synonymity questionnaires, including those concerning question sentences, etc.

Next in importance to the classification of synonymity questionnaires into such as are concerned with declarative sentences, imperative sentences, question sentences, and designations, we have found the classification according to the terms used in, and the meaning intended by, the questions of the questionnaire. This gives the *fundamentum divisionis* of the following classification. Five families of questionnaires are distinguished.

The first family includes all the questionnaires by which the subject is asked, roughly speaking, whether two given expressions mean the same or express the same. The *crucial questions* include what we call *synonymity expressions*: 'express the same assertion', 'mean the same', etc. Nearly all questionnaires of this family have asked whether a pair of expressions mean the same *for the respondent* as they occur *in a definite context* given in the questionnaire. More general questions are, of course, possible within the range of variation of the questionnaire-family.
Using the terms of the questionnaires one may say, roughly speaking, that it is required of the answers in order to be taken as confirmations of synonymity in some sense, that they *state* that certain statements express the same or mean the same. In the classification, this requirement is named a *synonymity requirement*.
Characteristic of the first family is the close relation between crucial question and synonymity requirement. A subject is *asked* e.g. whether two sentences T and U express the same, and the answer yes is the only *requirement* for classing T and U as synonymous (i.e. Qs-family-I-synonymous) for the subject relative to the contexts.[1]) The 'level of directness' of the questionnaires is maximal. They are used primarily in order to see how answers undergo variation with variations of subjects and – often more important – with change of contexts.

The second family and the rest of the questionnaire families avoid the terms 'mean the same' etc., and are with few exceptions more complicated and judged more difficult by the respondents. Roughly speaking, a first

[1]) The objection that Qs-family-I-synonymity is not real synonymity is of the same nature as the objection that intelligence as measured by this or that test is not real intelligence. (Cp. *Author*, 'Synonymity as revealed by intuition', *Philosophical Review*, vol. 56, 1957).

group of questionnaires of family II invite the test-person to try to conceive variations in conditions affecting the truth value of sentences. The concepts of synonymity intended to be studied by these questionnaires, are related to the truth-value concepts proposed by Carnap and others. The synonymity requirement is inconceivability of a difference in acceptance. A second group avoids the appeal to the imagination and thus *decrease the task-character of the questionnaires.* Roughly speaking, various conditions are indicated in the questionnaire, and the test-person is invited to answer whether under those conditions the crucial sentences would be true. A third group is a plain true-false questionnaire, by which the test-person is invited to judge each sentence of a series containing the crucial expressions, as to its truth-value. Synonymity-requirement: each pair of sentences, the members of which differ only in crucial expressions, have the same truth-value.

It should not be necessary to go into detail in characterizing the families III, IV and V. Rough characterization is given in the next pages.

FAMILY I

Use of synonymity expressions such as 'synonym', 'mean the same', 'express the same (proposition, assertion)'.
Synonymity requirement: sameness of meaning.
Associated philosophical concepts: Interchangeability salva significatione, synonymity in intention.
Level of directness: Maximum: if N.N. *says* that T and U means the same for him, T and U means the same for him.

GENUS I

Synonymity expression: 'express the same assertion'
Synonymity requirement: expressiveness of same assertion
Mode of presentation: invitation to imagine interchange of expressions
Prototype: Qs1 (*IP*, p. 360) [1]

[1]) For those questionnaires which have been used in some study, the results of which have been published, a reference will be made to this publication. *IP* is in this case and the following an abbreviation for Arne Naess, *Interpretation and Preciseness,* Skrifter utgitt av Det Norske Videnskaps-Akademi, Hist.–Filos. Klasse 1953, No. 1, Oslo, 1953.

Species 1

Synonymity question related to a definite occurrence in a definite text and to a definite event of interpretation by the test person.
Prototype: Qs1. (*IP*, p. 360)

Species 2

Synonymity question more general than in the case of Qs1. Generalization in respect of occurrences, texts, events of interpretation.
Prototype: Qs2. (*IP*, p. 364)

Species 3

As species 2, but furnished with an introduction requiring a definite criterion of sameness of cognitive meaning to be used by the subject. (The criterion is, roughly, sameness of conceived conditions of truth.)
Prototype: Qs30 (*Emp.*, p. 37).[1])

GENUS II

Synonymity expression: 'mean the same'
Synonymity requirement: sameness of meaning
Mode of presentation: invitation to imagine interchange of expressions
Prototype: Qs3 (*IP*, p. 366)

Species 1

Synonymity question related to a definite occurrence in a definite text and to a definite event of interpretation by the subject.

Species 2

Same synonymity expression as for species 1, but question of a more general nature.
Prototype: Qs3 (*IP*, p. 366).

GENUS III

Synonymity expression: 'synonym'
Synonymity requirement: synonymity

[1]) *Emp.* is an abbreviation for *An empirical Study of the Expressions 'true', 'perfectly certain', and 'extremely probable'*, Skrifter utgitt av Det Norske Videnskaps-Akademi, Hist. Filos. Klasse, 1953, No. 4, Oslo, 1954.

Mode of presentation: subjects asked to list synonyms of a given word
Such a questionnaire is never used by the author, but in a mimeographed doctorial thesis by Revie (Univ. of California). It is not classed as synonymity questionnaire 'proper', because then the phrasing of many theorems on likeness of meaning would be more complicated.

Species 1

No reference to texts or persons, straitforward invitation to list synonyms
Prototype, see dissertation by Revie, Univ. of California

FAMILY II

Synonymity requirement: non-existence or inconceivability of a difference (of some kind) under variation of imagined or stated conditions
Associated philosophical concept: inconceivability of logical inequivalence

GENUS I

Synonymity expression: 'same conditions of acceptance (and rejection) of some kind'
Synonymity requirement: inconceivability of difference in conditions of acceptance
Mode of presentation: invitation to conceive of different conditions of acceptance
Prototype: Qs10

Species 1

Synonymity expression: 'same conditions of acceptance', no specification of kind of acceptance
Synonymity question related to a definite occurrence in a definite text [1])
Prototype: Qs5 (*IP*, p. 368)

Species 2

Same synonymity expression as for species 1

[1]) Cp. species 1 of genus I, family I. – No definite event of interpretation is presumed analyzed. Under strange, imagined conditions, the text may get other meanings, and these may be conceived by the subject. Qs5-synonymity is present if T and U make similar changes in meaning with similar changes of conditions.

Synonymity question of a more general character: texts not given, authors said to be different
Prototype: Qs10

Species 3

Synonymity expression: 'same conditions of acceptance *as true* (vs. rejection as false)'
Synonymity requirement: inconceivability of difference in conditions of acceptance *as true*
Synonymity question related to definite occurrence in a definite text
Prototype: Qs11

GENUS II

Synonymity expression: no synonymity expression
Synonymity requirement: interchangeability *salva veritate*
Mode of presentation: invitation to answer whether certain statements are true or false. The whole questionnaire is formulated in the object-language.
Prototype: Qs29

FAMILY III

Synonymity expressions: no synonymity expression, or '(same) necessary or sufficient conditions', etc.
Synonymity requirement of (two) statements: if the first, then the second, and *vice versa*. Of designations: if subsumability under the first, then under the second, and vice versa.
Associated philosophical concept: logical equivalence

GENUS I

Synonymity expression: '(same) necessary, sufficient (condition) of''
Synonymity requirement: truth (or tenability etc.) of 'a' being necessary and sufficient condition of truth of 'b'
Mode of presentation: Texts with 'a' presented. Subjects invited to answer pairs of questions: 'Do you consider it a necessary (sufficient) criterion of the truth of 'a' that 'b_i' is true?'
Prototype: Qs4 (*IP*, p. 369)

Species 1

Synonymity expression: 'necessary and sufficient criterion (condition) of the truth of'
Synonymity requirement: Truth of 'a' necessary and sufficient criterion of truth of 'b'
Prototype: Qs4 (*IP*, p. 369)

Species 2

Same synonymity expression as for species 1
Synonymity requirement as for species 1 plus 'truth' of 'b' necessary and sufficient condition of truth of 'a'
Prototype: Qs20

Species 3

As species 1, but modified in such a way that answers are *use occurrences* (E.g., 'Yes' as answer to 'Do you accept that if N.N. loves his neighbour as himself, he is a good man?' rather than 'Do you accept the statement 'If N.N. –''.)
Mode of presentation: invitation to answer questions about the relation between the crucial expressions and each member of a list of statements V_1, V_2, ... V_i ... The questions have the forms 'Do you consider it a necessary (or: sufficient) condition of the truth of T (or: U) that V_i is true?'
Prototype: Qs25

GENUS II

Synonymity expression: 'if V_i then (not-) T and if V_i then (not-) U' etc.
Synonymity requirement: For all i and j: if V_i then T and if V_i then U, if V_j then not T and if V_j then not U
Prototype: Qs6 (*IP*, p. 369)

FAMILY IV

Argumentational synonymity, closely related to Family III, but marks the transition from formal logic to logic of argumentation (*pro et contra dicere*)

GENUS I

Synonymity expression: none
Synonymity requirement: 1) sameness of pro-arguments and sameness of contra-arguments, 2) Same strength of pro's and con's
Mode of presentation: invitation to imagine arguments that would count as pro (or con or be neutral) in relation to T (or U) but not in relation to U (for T)
Prototype: Qs8 (*IP*, p. 374)

GENUS II

Synonymity expression: none
Synonymity requirement: same as for Genus I
Mode of presentation: List of statements offered as possible arguments. Subjects invited to answer questions concerning the relation of the statements to T and U

FAMILY V

Synonymity expression: none
Synonymity requirement: sameness of extension, sameness of subsumption, sameness of class membership
Associated philosophical concepts: synonymity concepts for extensional languages

GENUS I

List of expressions offered.
Crucial question: 'is this an example of...?'
Synonymity requirement: sameness of example subsumptions
Prototype: Qs18

GENUS II

No initial list of expressions offered. Subjects invited to give instances. Instances offered by subject A presented to B and those by B to A etc.
Prototype: Qs19

FAMILY M

Metaquestionnaires

The aim of these questionnaires is to test the questionnaires already presented to a subject by studying how the subject interpreted the crucial question (etc.) of the questionnaire. (Metaquestionnaires thus defined need not be synonymity questionnaires).

GENUS I

Mode of presentation: a subject is asked whether he would have given a different answer to a (definite) questionnaire if the crucial question had not been what it was, but one of the questions of a list presented to the subject.

Do these so-called synonymity questionnaires *really* measure or somehow indicate synonymity? The answer must be negative, because there is no such thing as synonymity as such. In philological and philosophical literature the term has been defined in various ways and many (mostly inapplicable) formulae have been put forth as criteria of synonymity. Reformulating the criteria in such ways that they can be applied to concrete instances of use of certain sentences or other units of speech, the answers to the questionnaires are seen to be *relevant* as material. There is on the other side no criterion applicable to concrete cases of verbal communication which deserves the name of *the* criterion of synonymity in use.

In various publications Rudolf Carnap has pointed out that the analysis of synonymity in natural languages by questionnaire procedures and by other techniques outlined in *IP* supports his 'intensionalist thesis.' [1]) This thesis 'says that the assignement of an intension is an empirical hypothesis which, like any other hypothesis in linguistics, can be tested by observations of language behavior. On the other hand, the extensionalist thesis asserts that the assignment of an intension, on the basis of the previously determined extension is not a question of fact but merely a matter of

[1]) See especially note 7, in his 'Meaning and Synonymy in Natural Languages' (*Philosophical Studies*, 1955) reprinted in R. Carnap, *Meaning and Necessity*, 2. ed. Chicago, 1956 and in *American Philosophers at Work*, ed. Sidney Hook, N.Y., 1956.

choice. The thesis holds that the linguist is free to choose any of those properties which fit to the given extension; he may be guided in his choice by a consideration of simplicity, but there is no question of right or wrong.'

Arguing for his intensionalist thesis, Carnap [1]) invites the reader to suppose 'that one linguist, after an investigation of Karl's speaking behavior, writes into his dictionary the following:

(1) *Pferd*, horse,

while another linguist writes:

(2) *Pferd*, horse or unicorn.

Since there are no unicorns, the two intensions ascribed to the word 'Pferd' by the two linguists, although different, have the same extension. If the extensionalist thesis were right, there would be no way for empirically deciding between (1) and (2). Since the extension is the same, no response by Karl, affirmative or negative, with respect to any actual thing can make a difference between (1) and (2). But what else is there to investigate for the linguist beyond Karl's responses concerning the application of the predicate to all the cases that can be found? The answer is, he must take into account not only the actual cases, but also possible cases. The most direct way of doing this would be for the linguist to use, in the German questions directed to Karl, modal expressions corresponding to 'possible case' or the like.'

The texts and reference lists used in many questionnaires have just this function to elicit verbal reactions towards objects or relations which the subjects may not have reacted to before and possibly never will react to except verbally.

The main weakness of the extensionalist thesis as formulated (and rejected) by Carnap seems to me to be an implicit assumption that empirical studies of *extensions* could end in a definite hypothesis H_0 about sameness of extension of two expressions T and U of such a kind that there would be no group of hypothesis H_1, H_2, ... covering equally well the materials of the observational journals. But if hypothetico-deductive methodology is applied to observations of concrete instances of use of T and use of U,

[1]) R. Carnap, op. cit., p. 63 in *American Philosophers at Work.*

there will always be a wide range of choice in formulating hypotheses of a general character.[1])

In order to adapt semantical concepts of extension and intension to pragmatics (in the sense of Carnap and others) detailed analysis of concrete empirical procedures is highly desirable. In terms of operations, concepts of intension are not necessarily more vague or speculative than those of extension. The objection raised against intentional concepts may indeed be effectively met as suggested by Carnap: 'If for a given semantical concept there is already a familiar, though somewhat vague, corresponding pragmatical concept and if we are able to clarify the latter by describing an operational procedure for its application, then this may indeed be a simpler way for refuting the objections and furnish a practical justification at once for both concepts.'[2])

University of Oslo

[1]) The choice of hypotheses leads to the main problems of meaning analysis of occurences of terms in natural languages. (A. Naess, 'Forekomstanalysens grunnproblemer', Nordisk Sommeruniversitet 1953, Copenhagen, 1954).
[2]) R. Carnap, *op. cit.*, p. 60.

H. TENNESSEN

PERMISSIBLE AND IMPERMISSIBLE LOCUTIONS[1])

'Principle of Tolerance' and 'Ordinary Language Philosophy'

In the olden days when foreheads were wrinkled and thoughts profound, the philosophers were rarely in doubt about their proper mission: To synthesize all available – or even all conceivable – significant knowledge into one universal theory or system. The polyhistoric system-builders became, as we know, gradually obsolete and finally extinct when human knowledge increased explosively during the enormous expansion and differentiation of the natural, 'hard' sciences in the seventeenth, eighteenth and nineteenth centuries: They drowned helplessly in a true Amazon-flood of unsurveyable, isolated data. And their philosophical descendants, totally incapable as they were of coping with the new situation, made the fatal choice of seeking for themselves a small, secluded, still part of the stormy river where they, undisturbed by the onrushing sciences could devote themselves to their encapsulated 'philosophical' problem-and-theory-formulations: 'does matter (really) exist?', 'does man (really) have a free will?', 'what is the essence of truth?', 'everything is really mind (or matter)', 'cogito ergo sum', 'esse est percipi', 'die Welt is meine Vorstellung', 'l'existence précède l'essence', 'the nothing is the simple negation of the totality of being'... It was considered 'eine Wende der Philosophie' when one 'discovered' that these alleged problems and theories were, as one claimed: '*nonsensical*,' 'a meaningless play with words', 'a systematic abuse (or misuse) of language'... With this 'revolution in philosophy' philosophers were finally brought out of the intellectual dead-water of traditional, academic philosophy and given a more limited but respectable assignment: logical analysis, linguistic clarification, conceptual elucidation, pursuit of meanings, examination of the ways words are ordinarily used, to chart the actual features of everyday discourse... This avenue of escape from permanent stagnation was in many ways a pleasant one. For once, it permitted philosophers to perform what they were bound to feel as

[1]) This article is a modified version of a paper read before the annual meeting of the American Philosophical Association, Yale University, New Haven, Connecticut, December 29, 1960.

philosophically relevant activity, without demonstrating their scientific illiteracy, and without trespassing on the cultivated fields of well-established scientific disciplines: they were operating in scientific no man's land! Unfortunately or not: this is no longer so certain. After the more recent rise and sudden growth of such 'soft' sciences as psychology (including psycho-linguistics and psycho-semantics), social sciences (with socio-semantics), empirical semantics and linguistics, this wasteland of the nomadic philosophers has gradually come to be inhabited and systematically cultivated by scientifically more reliable homesteaders. In this way contemporary analytically oriented philosophers, with a study of 'natural' (or even: 'ordinary' 'natural') language as a major interest, are facing a predicament not altogether different from the one which their ancestors, the polyhistoric system builders and other academic philosophers, so unsuccessfully tried to evade. By disclaiming connections with relevant empirical sciences, there is a great risk that modern analytical philosophy will fossilize anew into a sterile system of encapsulated problem formulations. Needless to say, a language problem – be it a general problem (of, say, the existence of a 'syntax', a 'grammar' or certain alleged 'structures' or 'patterns' of language), or a problem of the actual use or usage of some particular linguistic locutions – it is in any event as much of an empirical problem as is a non-linguistic empirical problem. And it seems preposterous to try to throw light upon – let alone to solve or 'dissolve' – any empirical problem without plunging into the relevant sciences, here *e.g.* empirio-semantics or psycho-linguistics. – A particularly interesting situation arises, when *p.t.* philosophically interesting problems – linguistic or non-linguistic – have not yet been tackled by the scientists within any ramiculated branch of existing science disciplines. The prim and proper philosopher, then, who insists on an *a priori* attitude, has to choose between keeping his hands clean, at the cost of ignorance on relevant matters, or to engage in empirical research himself. It seems that confronted with this choice-situation, most analytical philosophers, and in particular the so-called 'ordinary language' oriented philosophers, choose ignorance as the lesser of the two evils. – The present paper is partly meant as an attempt to indicate what may be gained for *philosophy*[1])

[1]) Attempts are made in a paper read at the *International Congress for Logic, Methodology, and Philosophy of Science* (Stanford University, Stanford, California, August 27, 1960) to show how *social science methodology* may profit by cooperation with

by choosing the more earthly, *a posteriori*, attitude, employing empirical investigations after the pattern of the social (and other 'soft') sciences, and developing the available methods and techniques to fit within a philosophical frame of reference. However, the main endeavor will be to demonstrate how over-zealous philosophical prudishness and self-complacent 'Besserwissen' prevent philosophers from dealing effectively with language problems and hypotheses.[1])

It is today still one of the more obvious aims of analytic philosophizing to distinguish meaningful or, more generally *permissible* from nonsensical or *impermissible* linguistic locutions in order to be able to weed out any locution, X, which *does not make sense* – in at least one communicational event, C_i [2]) – and *eo ipso* is bound to lead to all sorts of absurdities and perplexities (at least in C_i). However, as we all know: there is, unfortunately, no plain and uniform prescription for how to achieve this ever so commendable objective.[3]) I shall in the following briefly touch upon three vaguely dissimilar principal endeavors in this direction, all of which are taintless products of immaculate, inspirational, language lucubrations, totally uncorrupted by any kind of earthly, empirical, considerations.

1. A popular and patent way to establish whether or not *X makes sense* (in C_i), is to ask oneself: 'does X make sense (in C_i)?' At least two considerations may prevent one from putting too much confidence in a flat answer to such a question, viz. a) the many glaring ambiguities in the key expression, 'make sense', and b) the tremendous *individual and situational variations* in tolerance and sensitivity to language ambiguities, as well as in requirements for meaningfulness.

Recent empirio-semantical enquiries – at the University of California – into the ambiguities of T_0: 'It does not make sense to say X. . . ,' revealed at least nine main directions of (more precise) interpretations of T_0. Furthermore it seemed as if X was only to be disqualified as '*eine sinnlose Aneinanderreihung von Worten*' if T_0 is interpreted in the direction of

empirically oriented, research minded analytical philosophers. *Vide:* 'Empirical Semantics and the Soft Sciences' in: *Proceeding of the International Congress for Logic, Methodology and Philosophy of Science 1960.*

1) *Vide* also: Arne Naess, 'Philosophers and Research in the Soft Sciences', *Actes du XIème Congrès Internationale de philosophie*, pp. 255–259.

2) *I.e.:* for at least one person, P_i in one situation, S_i, within one, 'natural', language L_i.

3) *Vide* the author's: 'On Making Sense' (Abstract) *The Journal of Philosophy*, vol. LVII, No. 24, November 24, 1960, pp. 764 and 765.

T_1 or T_2 below; whereas T_0, for the given variable of X, is only expressing a true statement provided T_0 is interpreted in one of the other, cognitively entirely different main direction of (more precise) interpretations of T_0. T_1 and T_2 may here tentatively be formulated as follows:

T_1 One is expressing a logical oddity, a logical, grammatical or syntactical inconsistency, a negative or positive analytic statement – *i.e.* (self-) contradiction or tautology – or a pleonasm, a redundancy – when uttering X.

T_2 One is uttering a random cluster of words, a haphazard conglomerate of irrelative linguistic expressions or (in principle) unverifiable and unfalsifiable statements, totally devoid of cognitive content... when uttering X.

Yet it seems that many analytic philosophers are freely employing this very 'method' with great confidence and apparent success: one simply 'sees' or 'hears' the sense or the nonsense of X. A native speaker, it has been maintained,[1]) can never (or rarely) go wrong. He *perceives* X's cognitive permissibility (impermissibility, respectively) directly, instantaneously, in a flash of revelation, by some sort of linguistic instinct, logical sense or hermeneutical clairvoyance. It seems, however, and regrettably so, that different, presumably competent seers come out with different, mutually exclusive visions. Thus, as pointed out by Benson Mates,[2]) agreement cannot be reached even within so restricted a sample of seers as the class of Philosophy Professors at the university of Oxford, Oxford, England (*viz.* John Austin and Gilbert Ryle). Furthermore, it can be shown that such visions may vary intra-personally as well. For example: the same students who in a logic seminar would 'see' the exclusive sense (aut-junction) as the only 'logical', 'intelligible', 'meaningful' sense in which the expression '*either/or*' could possibly be used, will find it preposterous if they were interpreted to have employed 'either/or' in this way under any more trivial, every-day circumstances.[3]) Empirio-semantic investigations have indicated a pronounced tendency towards

1) *Vide: e.g.*, Stanley Cavell: 'Must we mean what we say?' *Inquiry*, Vol. I, No. 3, September, 1958.
2) *Vide:* 'On the Verification of Statements about Ordinary Language', *Inquiry*, Vol. I, No. 3, 1958, p. 165.
3) Cf. 'On Worthwhile Hypotheses', *Inquiry*, Vol. II, No. 3, 1959, pp. 189 ff.

logical and linguistic rigidity whenever informants are placed in a classroom frame of reference, or confronted with naive, unrefined questionnaires concerning language usage or the like; whereas in most other, 'ordinary', 'normal' situations the same subjects would display the most admirable, semantic latitudinarianism. – A follow-up study was made as to whether it might be possible to *condition* informants to get groups of both kinds: a) one rigidity-group of logical pedants and b) one flexibility-group of latitudinarians. Of 147 students, who up to then had been relatively uninfected with logic or semantics in any form, 75 (Group I) were exposed to a short lecture phrased with the purpose of eliciting what we have described as a 'logico-maniacal' attitude, and 72 (Group II) to another lecture with what we call a more 'common sensical' bias.[1]) The results of these experiments showed quite convincingly that the first lecture (I) sufficed to awaken the presumably dormant hermeneutical clairvoyance and to turn perfectly normal, mentally healthy human beings into rigid, hairsplitting, pedantic logico-maniacs; whereas the second lecture (II) was apparently enough to induce a more common-sensical, latitudinarian attitude. When confronted with three sets of up to 48 sentences, the *logical pedants* intuited, without difficulties or exceptions, the classification of all the given sentences in either one of three categories of impermissible sentences: 1) (self-) contradictions, 2) tautologies, 3) nonsense sentences, whereas the latitudinarians insisted on interpreting all the same sentences to transmit plausible, reasonable, fairly interesting, worthwhile, more or less tenable hypotheses, or more or less advisable proposals.

This outcome is calculated to discourage the predominant, naive-optimistic belief that it is easy to 'see' or 'hear' whether a given sentence or expression 'makes sense' in a specified (or in any?) communicational event, C_i. It seems safe to assume that: given a linguistic locution, X, which off-hand appears 'to make sense' (to a native speaker) in C_i, one can always imagine a different communicational event, C_j, in which X is found 'meaningless', 'nonsensical' or otherwise impermissible. In other words: Any locution, X, will always admit of an unlimited set of

[1]) For the actual formulations of the two lectures, *vide* the authors: 'Logical Oddities and Locutional Scarcities', *Synthese*, Vol. XI, No. 4, December, 1959, pp. 376–78; and: 'What should we say?', *Inquiry*, Vol. II, No. 4, p. 288, footnote no. 13. Cf. also the forthcoming article in the *Journal of Philosophy:* 'Whereof one has been silent...

plausible interpretations of X which make X a 'meaningful' locution, *and* of an equally unlimited set of plausible interpretations which would make X 'nonsensical', 'meaningless'... – Take again the hackneyed example: 'It is raining outside, and I don't believe it'. Not only is there nothing logically odd or otherwise nonsensical and impermissible about this sentence, commonly interpreted: 'It is raining, how extraordinary!' – but one could easily imagine situations where the actual state of affairs would be most adequately conveyed by: 'It is raining outside and (or: but) I don't believe it.' A few hours ago, shortly after I started writing the present paper, it was, contrary to all reasonable expectations actually raining outside. Absorbed in my work, I did not notice; and had I been asked, 'Do you believe it is raining outside?' I would honestly and sincerely have answered: 'No,' that I did not believe it was raining outside. Thus, if a few hours ago I had said to myself; 'It is raining outside while I don't believe that it is raining outside,' I would have made a true assertion. How I could possibly have gained such an insight into the discrepancy between my beliefs and the actual state of things, is another problem. But one thing I know: Whenever I have any such extraordinary insights and feel the necessity for conveying them to myself or to my fellow beings, I shall also feel free to make use of any locutions – however weird or bizarre they may sound or look to a pedantic logician – if they only provide an adequate, verbal transmission.[1]) By discarding such sentences as nonsensical, logically odd or otherwise suspect or impermissible, we prevent language from fulfilling its major purpose (*i.e.*, to increase inter- and intra-personal communicability) by preventing adequate locutions from being employed when effective communication is entirely dependent upon their availability.

2. Another rather commonly accepted contrivance to secure a swift and expedient revelation of the (im)permissibility of a given locution, X, is the 'method' of asking: '*Can* we (say) X?', 'is it *possible* to (say) X?'. The general assumption seems to be that certain 'things' are 'impossible' and

[1]) Another somewhat more far-fetched type of illustration is furnished for example by a person who has long suffered from a lack of sense of reality; at a time where he is gaining some insights in his own psycho-pathological picture, he might well have experiences which would must aptly be conveyed by such sentences as: 'It is raining outside and (but) I don't believe it'... etc. – 'Today is Thursday and (but) I don't believe it.'

therefore (?) 'unsayable' (and/or 'unsayable' and therefore (?) 'impossible'?). One *cannot*, or *it is impossible* to utter, *e. g.*: 'My necktie has a cause,' 'she unintentionally tied a string across the top of the staircase,' 'he yawned (in-)voluntarily (heartily, disgustingly, sleepily, normally, ordinarily, in a standard, paradigmatic manner... etc.),' 'I've been recognizing you for at least three minutes, and I think I am about half done,'[1] 'I promised my chair to stop smoking'... etc. What most plausibly may be meant by: 'cannot', 'it is impossible' in such and similar contexts is quite problematical if not completely obscure. However, it has been claimed that there may be some thin connections between, this linguistic ineffability approach to the (dis-)solution of presumably philosophical problems, and what may be labeled 'linguistic rigidity'. 'Linguistic rigidity' is used to designate certain attitudes to the use of language alleged to be predominant in so-called 'primitive' people as well as in pre-school children. It is generally characterized by an inability to utter certain words or sentences, 'to play with words,' 'to look on language as a word game,' and 'make words mean what we want them to mean.' Humboldt tells a story about a peasant who, after listening to two students of astronomy talking about the stars, said to them: 'I can see that with the help of instruments men could measure the distance from the earth to the remotest start and find out their position and motion. But I should like to know: how did you find out the *names* of the stars?' He assumed that the names of the stars could be found out only from the stars themselves. Vigotsky [2] claims that similar dispositions may be found in every one of us, if we go back to the pre-school age. My experiments with pre-school children [3] show, however, that neither children nor grown-ups have inhibitions in playing with words and making words mean what we choose them to mean, *e.g.*, in saying about a dog: 'it is a cow,' or about a basketball: 'it is a bicycle,' and so forth, as soon as it is made clear to the informants that by 'can you call (say)?' is meant: 'are you capable of making the following sounds?' (*i.e.*, 'uttering the

[1] *Vide:* Lawrence Resnick: 'Words and Processes', *Analysis*, Oct., 1959, pp. 19–24, and H. Tennessen: 'Dialogue on Duration of Recognition' (forthcoming *ibid.*)
[2] L. S. Vigotsky, 'Thought and Speech,' *Psychiatry*, Vol. II, No. 1, February, 1939, p. 36. The Humboldt anecdote is borrowed from the same article.
[3] *Vide e.g.:* 'What should we say', *Inquiry*, Vol. II, No. 4 and 'Vindication of Humpty Dumpty', *Inquiry*, Vol. III, No. 3.

following sentences.') In the pre-text to the above mentioned experiments this was easily accomplished by the following procedure: When the (pre-school) child entered the room, the investigator said: 'Can you call a basketball 'a bicycle'?' Child *(e.g.):* 'No!' Investigator: 'Watch me: *I can,*' and pointing in the direction of a basketball he uttered: 'Bicycle!' Then, turning toward the child again: 'See how easy it is to call a basketball 'a bicycle'? Now you try!' [1]) Of course, one found absolutely no trace of any so-called linguistic rigidity in the pre-school children or in any of the tested age groups of normal informants. What one found was a rather obvious ambiguity in such formulations as: 'we can (not) say 'call, convey, or utter),' 'it is (im-)possible to say (call, convey, or utter)', etc. Reactions classified by Vigotsky, Piaget, Frazer and others as symptoms of linguistic rigidity seem to be due to nothing more exciting than a tendency – among investigators as well as respondents – to oscillate in their interpretation of the mentioned formulations between the following two directions of interpretation: a) 'it is (im-)permissible, (dis-)advantageous, (in-)advisable, (un-)fitting, (un-)fortunate, (un-)reasonable... to say (call, convey, or utter),' and: b) 'it is (un-)achievable, (un-)attainable, within (resp. beyond) my power and capacities... to say (call, convey, or utter).' The former direction of precization (a) seemed in most cases to be preferred to the latter (b).

Moral: There are *no* limitations upon what one *can* mean by the utterance of a sentence, just as there are *no* limitations upon what a sentence *can* mean.[2])

3. The aforementioned empirio-semantic investigations showed *inter alia* that the sentence, T_0 'Should we (under the described circumstances, ever or ordinarily) say X?', may also quite plausibly be interpreted in the direction of T_1 '*Is it likely* that we (under the described circumstances ever or ordinarily) would utter X?' In other words, *it seems as if the frequency with which X is* estimated to occur in a certain situation, S, is taken to indicate the degree of *permissibility of X in S.* If the estimated frequency of occurrence is zero for X in S, then X is totally impermissible in S.

[1]) For the actual experiments see the preliminary report in: 'What Should We Say?', pp. 268–272.

[2]) *Vide:* Jerry A. Fodor: 'What do you mean?', *The Journal of Philosophy,* Vol. LVII, No. 15: July 21, 1960, pp. 499–507, and my forthcoming criticism, *ibid.:* 'People mean, sentences don't'.

When the late John Austin was in Norway, October, 1959, he is reported to have made the following statements: [1]) 'That there are rules prohibiting something from being said, does not say more than that we never say so and so.' And: 'We say there is a rule against saying X, when X is never said.' Ample material from several empirio-semantic investigations point to some rather obvious fallacies in any attempt to reject a locution (as nonsensical, logically odd or otherwise impermissible) on grounds of its linguistic non-occurrence – *without taking into considerations the most likely reason for its scarcity*. It seems in general plain, of course, that if a locution, X, has never occurred in a situation, S, this may be due to the fact that nobody has so far found it *worthwhile* to utter X in S. True enough: One of the reasons why one has so far never found mentioning X in S worthwhile could be that X has hitherto not conveyed, or contributed to the conveyance of, any intelligible statement when uttered in S.[2]) But this would 'not make sense' in S in one or more of the in S. It is more likely that X would 'not make sense' in S in one or more of the many other senses of 'not making sense'. For instance: The statement conveyed by X – or the sentence in which X occurs – may be an undoubtedly false hypothesis (or inadvisable proposal). And insofar as one feels that such statements should not be set forth at all, one may refrain from making them. The point is, however, that what at a time t may seem to everybody to be a false hypothesis, may well at another time t' appear to be very true or even trivial. Consequently, if there were linguistic rules prohibiting sentences which at t were unanimously interpreted to express untenable hypotheses, the same rules would prohibit effective communication at t' when the world had changed in such a way that certain plain (but interesting) matters of fact would be most adequately described by employing some of the at t interdicted sentences.[3])

What here is said about sentences conveying obviously false hypotheses

[1]) The report is only available in mimeographed form. *Vide* the pamphlet: *John Austin and Arne Naess on Herman Tennessen's experimental warning:* 'What Should We Say?'. Oslo, 1959, Berkeley, 1960.

[2]) *I.e.*, X is nonsensical in the sense: X is '*eine sinnlose Aneinanderreihe von Worten*'.

[3]) Consider *e.g.*, the expression: 'to split an atom,' less than 100 years ago this expression would be 'seen' as nonsensical, as a *contradictio in adjecto, ein hölzernes Eisen*. It would, by the same token, be silly of us to ban forever expressions as: 'a rate of speed exceeding the velocity of light,' and the like. (*Vide: The Scientific American*, August 1960).

(and inadvisable proposals) will, of course, apply as well to such sentences which 'it does not make sense' to utter because they are unanimously interpreted to express platitudes or otherwise idle and uncalled for, obviously, undoubtedly true or advisable propositions. Thus if, in a perfectly normal situation a normal person is yawning a normal yawn, it is clear that we should neither say: 'That person is *not* yawning now,' *nor* should we (ever or ordinarily) say: 'That person is yawning.' If such sentences are uttered (under such circumstances), the audience would wonder: Why on earth does he want to say this? What is he up to? He must have some special purpose. What is it that he wants to prove? In other words, the sentences have ceased to have symbol function. Consequently: 'He is (not) yawning' is never, or not ordinarily, said unless there is something fishy or otherwise remarkable about the yawn or we have other good reasons (or plausible pretexts) for making such a remark.[1])

Common to all the above-mentioned 'methods' for determining the desirability of a locution, X (within C_i), is the underlying assumption that *natural language* is in a way, at least '*in its ordinary employment*', *perfect and unimpeachable*, *consistent and complete in itself – incorrigible*, so to speak. From this follows first and foremost that there is no need for constructing *formal* systems of *artificial* languages. On the contrary 'the constructors of calculi' are not only themselves suffering from 'radical misconceptions', they are, with their ill-placed emphasis, apt to avert others from the *Via Triumphalis* of 'ordinary language philosophy'. Furthermore, it follows that, since the perfect and consistent language *is there* to be discovered by the first (qualified?) native speaker that comes along to catch a sight of it, there is no call for broad, empirical investigations of how the natives use, or believe that they use, their language. The chances are that most natives are wrong anyway. In fact they may be as wrong as Ryle was when he reported that 'beneficial actions cannot be voluntary.'[2])

[1]) To illustrate and support this general rule that occurrence or non-occurrence of a linguistic locution, X (within C_i) may be due to nothing more exciting than differences in the demand for X (in C_i) – 6,270 adult Californians were asked: 'Should we ever say, 'Today I met a man who was... tall''? '...could be given values from 1′2″ – 15′. For part of the results *vide: Synthese*, Vol. XI, No. 4, Dec. 1959, pp. 386–87.

[2]) *Vide:* Stanley Cavell: 'Must we mean what we say?', *Inquiry*, Vol. I, No. 3.

An interesting problem arises in connection with the methods by which one native speaker – say, J. A. – of a natural language, L, is supposed to convince another native speaker, of L, *e.g.* G. R., that G. R.'s is wrong in his 'deviating' linguistic revelations as to whether a certain locution, X, is permissible ('makes sense') or not (in C_i). However, the main crank in this linguistic *perpetuum mobile* seems to be an unshakable faith in (and devotion to) the infallibility of certain directions for use of L, subtly enshrined, one is told, in the 'ordinary' use of L. And as the case may well be with many an unshakable faith, the dough-faced devotion serves as a psychological compensation for the conspicuous lack of empirical indicia: *credo quia absurdum!* The natural *secondary* effects are: intolerance towards heresy, and a condemnatory attitude *vis à vis* deviating language revelations.[1])

Since there is no available evidence to support the notion that 'ordinary' language is really such a marvel and miracle, other empirically oriented analytic philosophers tend towards the meaning that it is a mess, (or at least they assume with Wittgenstein 'the *untidy* character of ordinary language' *[Tractatus]*). From this platform it seems obvious that whenever a philosopher finds it important to propound hypotheses concerning the use or usage of a locution, X, within, say, a language society, L_a, it would be methodologically hazardous to let the tenability of the hypotheses rest solely upon the shaky foundation of one single native speaker's *a priori* intuitions of his own and fellow native's 'ordinary employment' of X (in L_a).

Moreover the empirically oriented language philosopher has no quarrel with the 'calculus constructors'. On the contrary: the more empirical investigations reveal of the 'untidiness' of natural languages, the more apparent becomes the need for more precise and consistent formal systems, which permit one to cope with philosophically relevant problems, undisturbed by deceptive ambiguities and other potential bewilderments and perplexities enshrined in the natural languages, particularly in their ordinary employment.

Last, but not least, the empirical approach to the philosophy of language not only permits, but *presupposes* an extremely tolerant attitude towards

[1]) The necessary consolidation of the sect is sought by means of increased sensitivity to the in-group's revelation expectancies.

any uses (usages) of X which may off-hand sound exotic, preposterous, logically odd, meaningless, absurd, or otherwise impermissible. In fact it seems that Carnap's *Principle of Tolerance* may most aptly be applied here: in *The Logical Syntax of Language*, after having discussed several examples of so-called 'negative requirements' (*e.g.* of Brouwer, Kaufmann, and Wittgenstein) 'by which certain forms of language – methods of expression and of inference – would be excluded', Carnap goes on to say: 'Our attitude to requirements of this kind is given a general formulation in the *Principle of Tolerance. It is not our business to set up prohibitions but to arrive at conventions.*[1])

First, it goes without saying that the Principle of Tolerance will here chiefly or exclusively apply to *the receiver* of a piece of communication. It would be disastrous if the Principle of Tolerance were used as a pretext by a sender for not trying to live up to the necessary level of preciseness, required of him in a given communicational event.

Secondly, 'it is our business to arrive at conventions'. It can hardly be questioned that effective, objective *(sachlich)* verbal communication is dependent upon the use of linguistic locutions which are: a) suitable for some special purposes *i.q.* b) clear (*i.e.*, having a satisfactorily high degree of subsumability), and c) in accordance with some ordinary (*i.e.*, frequently occurring) language usages. *Only insofar as point c is concerned is a study of actual language usage of (indirect) value to philosophers.* And this holds true regardless of whether one's underlying assumption is that ordinary language is perfect, or that ordinary language is a mess. In any case, one needs to know a little about the most ordinary usages to prevent unnecessarily drastic deviations from them. Drastic deviations may mislead the sender, as well as the receiver, create communicational disturbances, misunderstandings, and confusion (*vide* Strawson's use of 'presupposition'). However, considerations of a) suitability for special purposes, and b) clarity (subsumability) will most often, if not always, prevent a communicator from flatly adopting any one of the existing language usages of a given, important, linguistic locution. He would feel the need for: 'explications,' 'rational reconstructions' or conceptual alterations of one kind or another. In fact, there are instances where the

[1]) *The Logical Syntax of Language*, Section 17. *Vide* also: *Introduction to Semantics*, Section 39, and compare sections 12 and 36.

sender finds it most advantageous to disregard completely ordinary language (*vide:* Einstein's use of 'simultaneity', Russell's concept of 'evidence', Strawson's notion of 'presupposition,' etc.). He 'makes words mean what he wants them to mean'. There are innumerable cases in the philosophy of language where we realize that what we off-hand may have interpreted to be a language hypothesis, is more readily understood as a verbal recommendation, as a convention or as any other type of *normative* statement, say, a proposal for how to use a given linguistic expression within a specified or unspecified context, or a general, explicit program for how to deal, systematically, consistently, effectively with certain bothersome *types* of formulations. Most *definoform* sentences may more or less obviously be interpreted in this normative direction. The same goes for *e.g.* many so-called *theories of description;* Russell's, Frege's, Hilbert-Bernays'... I want to make it clear that in all such normative cases empirical investigations of actual language usage are either of secondary import or totally uncalled for.[1]) They may, of course, shed some light on the *practicability* of a language proposal, insofar as they reveal the extent to which a proposed language usage is in accordance with certain existing usages from which one (may or) may not want to deviate too drastically, But under no circumstances will the *advisability* of a language *proposal solely* be determined by the *tenability* of language *hypotheses.* And surely it does not argue for the appropriateness of informations about the so-called 'actual features of ordinary speech' that the linguistic 'data' are collected by means of methods as primitive, naive and optimistic as those employed *e.g.* by F. P. Strawson in his criticism of Russell's suggestion for how to handle sentences like, 'The present king of France is wise.' Strawson's recipe is, as everybody knows, simple. One merely a) considers a given sentence in which there occurs a description for which one knows that no descriptum exists, then b) one supposes that someone actually uttered the sentence, and finally c) one asks oneself whether one should (ordinarily) say, 'That's false (untrue)!' If one wouldn't, one has *ipso facto* evidence that in ordinary speech the statement concerned is neither true nor false.

To sum up, it seems obvious that there should be no rule against saying X only because X sounds nonsensical, logically odd, absurd, preposterous

[1]) *Vide: Inquiry,* Vol. III, No. 3, pp. 185–88.

to our logical sense or linguistic instincts. And this is particularly the case if X sounds nonsensical, etc., merely because X has so far never occurred, never been uttered before within a given language society. By the same token, the mere fact that a certain way of speaking and thinking has 'made sense' for a number of millennia creates but a feeble presumption of present or future desirability. We need a language with which we, if necessary, can describe unforeseen, unimagined, hardly imaginable, maybe even unimaginable, inconceivable phenomena. On the other hand, language must permit one to express the most boring, futile, idle, and therefore 'unsayable' commonplace platitudes whenever that seems to be important... or if one wants, *seems to make sense*,[1]) – in still another sense of 'making sense'. It is obvious that locutions are in general *most correctly* applied in communicational events which are the ordinary, standard, paradigmatic cases of their most frequent application. But, after all, what is correctness? As Friedrich Waismann once said, 'I have always suspected that correctness is the last refuge of those who have nothing to say.' And thus spoke Zarathustra, 'Ich sage euch: man muß noch *Chaos* in sich haben, um einen tanzenden Stern gebären zu können.'

Department of Speech, University of California, Berkeley

[1]) A study of the locution: T_0 *It does not make sense (e.g.) to say X* (in S), permitted the investigators (Goldstine and Tennessen) to distinguish at least nine main directions of interpretation of T_0, two of which, T_1 and T_2, were such that most informants offhand would say, 'If (but *only* if) by T_0 one wants to express anything in the direction of T_1 or T_2, then one should exhort against employing X (in S).' (For all other interpretations of T_0 the advisability of employing X (in S) would depend upon further informations of various kinds). *Vide:* 'Vindication of Humpty Dumpty,' *Inquiry*, Vol. III, No. 3.

DAVID RYNIN

NON-COGNITIVE SYNONYMY AND THE DEFINABILITY OF 'GOOD'

It was argued many years ago by G. E. Moore, in his *Principia Ethica*, that 'good' is (in a sense) indefinable. Some critics of Moore who find his analysis faulty still go along with him at least to the extent of holding, as does Charles Stevenson in his *Ethics and Language* (p. 82), that "Good' is indeed indefinable... if a definition is expected to preserve its customary emotive meaning.' And a leading representative of more recent linguistic philosophy, R. M. Hare (in his *The Language of Morals*, London, 1952) in much the same spirit tells us that 'it is not the case that there is any conjunction C of descriptive characteristics such that to say a man has C entails that he is morally good. For *if this were the case we should be unable to commend any man for having these* characteristics...' (p. 145 op. cit.). 'Value terms have a special function in language, that of commending; and so *they plainly cannot be defined in terms of other words which themselves do not perform this function*, for if this is done, we are deprived of a means of performing the function.' (p. 91; I have italicized the question-begging parts of the above quotations). In this paper I wish to examine some of the ideas and assumptions underlying statements of the kind illustrated above, in order, if possible, to come to some decision regarding their correctness.

I begin with some elementary remarks on definition. At least two important senses are attached to this term: 1) definition by stipulation and 2) definition by empirical synonymy determination. It is, as we all know, sometimes useful to replace a longer by a shorter expression, and this may be done by the speaker stipulating that the shorter is to be used in place of the longer. This stipulation is neither true nor false, but, functionally, an imperative sentence or an expression of a decision, sometimes masquerading as a genuine statement. Clearly no issue of importance arises in relation to a definition by stipulation (assuming it is consistent) except perhaps that of convenience.

Contrasted with such a stipulatory definition is a definition based on an empirical synonymy determination. It will be, unlike the former, a true or false statement, and will tell us that a certain expression has the same meaning as another; it will not express the speaker's decision as to usage, but will presuppose it and then give some genuine information about the identity of meaning of the defined and the defining terms. If knowledge consists in recognizing identity amidst diversity, such a definition will give us knowledge. In short, what may perhaps merit the name 'real definition' will thus report the discovery of a relation of synonymy between two terms. Whether this is all that is necessary for what is vaguely called 'a good definition', or whether some other conditions must first be satisfied, we need not here seek to determine – suffice it to say that it is generally agreed that such a relation of synonymy is a necessary condition for a definition. But this is not very enlightening unless we have given a clear meaning to the term 'synonymous with', and to this I now turn.

I shall assume, for want of space in which to defend the view, that the basic conception of synonymy is that of sentence and in particular statement synonymy, and that synonymy of words and other expressions and non-statement sentences is to be, and can be, explained in terms of this fundamental conception. Roughly speaking, my view is that two statements are to be called 'synonymous' if and only if they have identical truth conditions. The more usual formulation of statement synonymy in terms of L-equivalence (sometimes qualified, as with Carnap's Intensional Isomorphism), in the sense, roughly, that two statements S and S' are L-equivalent if and only if it is impossible for them to have different truth values, will coincide with this notion or not depending on certain decisions. For example, as I use 'truth conditions,' 'A is an equilateral triangle' is not synonymous with 'A is an equiangular triangle.' If it is claimed that these statements are synonymous, then the properties *being L-equivalent* or even *being intensionally isomorphic* and *having identical truth conditions* are not the same. Thus, as I understand it, if two statements are synonymous in the sense of having identical truth-conditions, they are L-equivalent; but whether the converse holds depends on a more careful explication of the notion, in relation to cases like that referred to above.

However, it is not necessary for my purpose in this paper to enter into a discussion of this topic, since it will suffice for my argument to deal only with necessary conditions for synonymy. Thus, two statements in order to be synonymous will have to satisfy the condition of L-equivalence. And two terms T and T' will require for their synonymy that they be substitutable one for the other in statements without altering their truth values.

I call the statement of an equivalence of this sort 'an analysis' when offered as a step in the explication of the meaning of a statement. Following Langford (on the analogy of definition, where we speak of the 'definiendum' and the 'definiens'), I shall speak of the statement to be analyzed as the 'analysandum' and of the analyzing statement as the 'analysans.' We shall then speak of analyzing the meaning of a sentence, but of defining the meaning of words. It is clear that the notion of analysis is basic, for we determine that one term can be a synonym for another, and hence possibly a definition, by establishing the L-equivalence of statements which differ only in the manner described above. The term 'synonym' is ambiguous; we overcome this ambiguity by distinguishing between sentence-synonymy and term-synonymy.

Now this term, 'sentence-synonymy,' would appear to apply directly only to statements; for non-statement sentences have no truth-value and hence cannot be equivalent to each other, in the sense of necessarily having the same truth-value. (I reject as a rather bad joke the view sometimes expressed that sentences having no truth-value have the same truth-value, since having none they do not have different truth-value and if not different then the same.)

What holds for non-statement sentences holds as well for what I call 'complex sentences,' that is, sentences having non-statement (as well as statement) sentences as components. Thus for example, in Stevenson's well-known *Pattern-I* analysis of ethical judgments: 'This is good' $=_{df}$ 'I like this; do so as well!', the right-hand expression, 'I like this; do so as well!', is a complex-sentence, having no truth-value; unlike, say, 'I like this; I hope you do, too,' which would be a compound statement, having the truth-value determined by the usual rules for the truth of conjunctions.

If then we require L-equivalence of the analysandum and the analysans in forming a correct non-stipulatory definition, it will be obvious that no definition of 'good' is possible if the analysans is a complex sentence. We are therefore faced with a decision. We may insist on the requirement of L-equivalence as a condition for any correct analysis and hence definition; but in that event we foreclose any possibility of analyzing the meaning of any non-statement sentences including complex sentences. Or abandoning this strict requirement, we may expand the notion of equivalence to apply in some way to sentences that are not statements. On the first alternative we should have to hold that Stevenson's *Pattern-I* analysis is illegitimate, and that therefore no definition of 'good' is possible which rests on the requirement that 'This is good' be L-equivalent to some complex sentence. Clearly, if we wish to legitimize a definition based on such an analysis we shall have to abandon our synonymy requirement of L-equivalence, or in some way transform the complex sentence offered as the analysans into some kind of compound statement which we may use in determining synonymy as before.

In order to make clear what is involved in such a transformation as has just been mentioned I introduce the notion of 'propositional correlate.' I call a statement 'a propositional correlate' 1) of an imperative sentence if, for instance, it is a statement describing sanctions that follow on compliance (or non-compliance) with the imperative; 2) of an interrogative sentence if it is a statement giving an answer to the interrogative; 3) of an exclamatory sentence if it is a statement describing a state of mind or feeling expressed by the exclamation. (A state of mind or feeling is expressed by a sentence if the utterance of the sentence is an effect of this state of mind or feeling.)

Thus a propositional correlate (a P.C.) of the imperative 'Do so as well!' will be, for example, the statement, 'You will be happy if you do so as well;' a P.C. of the interrogative 'Is he a good man?' will be, say, 'He is a good man;' a P.C. of the exclamation 'How good a man he is!' will be, say, 'I greatly admire his goodness.'

In terms of this proposal, when analyzing Stevenson's *Pattern-I*, we put in place of 'do so as well!' (in the analysis: 'This is good' = 'I like this; do so as well)' the statement 'You will be happy if you do so as well.' By virtue of this substitution the analysans becomes: 'I like this; you will be happy if you do so as well,' a compound statement having a truth value. Hence the revised analysis containing this substitution will be of such a character that the requirement of L-equivalence once more becomes at least possible. We then say that the sentence 'This is good' is L-equivalent to 'I like this; do so as well' (and hence possibly synonymous with it) only if 'This is good' is L-equivalent to 'I like this; you will be happy if you do so as well.'

Let me say at once that I do not think that they are even equivalent, let alone L-equivalent, and therefore I do not offer or accept this as an analysis-basis for a definition of 'good.' Indeed I have earlier indicated that I should not necessarily consider them to be synonymous even if they were L-equivalent. In any case, it will be evident from what has gone before that no analysis of 'this is good' will be possible unless the analysandum has a truth value, and this would appear to presuppose that we have already a sufficiently clear understanding of the meaning of 'This is good' for it to have a truth-value (perhaps given in some prior stipulation or in the extension of 'true-statement').

This, perhaps not very convincing, end to our attempt to preserve the requirement of L-equivalence in the analysis of statements whose analysans contains a non-statement component may persuade us to see what can be done by abandoning the requirement in its strict form.

The obvious alteration is to introduce some kind of requirement analogous to L-equivalence, but not identical with it, which will allow us to say that some complex sentence is L-equivalent to some statement. In such a case 'synonymous with' would undergo a corresponding change in meaning. Such an expanded conception of equivalence applicable also to non-statement sentences might be formulated as follows:

1) For statements, the usual definition of L-equivalence is adopted, i.e., necessary identity of T (truth)-value. For imperatives we substitute for identity of T-value, identity (or likeness) of reaction or stimulus affects of the sentences. Thus if two imperatives have the same (or roughly the same, or possibly approximately the same, tendency to produce) effects in the form of responses to them, then we say they have the same P value, i.e., persuasive value. 3) We say two interrogatives are equivalent, have the same I value, if they elicit (or have a tendency to elicit) the same statements as answers. 4) And for exclamatory sentences, we say they are equivalent, have the same E value, if they express the same feelings or attitudes, or possibly have the same tendency to do so.

In terms of these conceptions we may then hope to determine whether an analysans is equivalent to its analysandum, not merely in the special case when the analysans is itself a statement, but also when the analysans is indubitably a complex sentence.

For example, consider again the sentence, 'This is good', uttered on a particular occasion. Having determined the truth value in the usual way [1]) we proceed to ascertain the response to such a judgment of an act or object. We determine, that is, what the P value and the E value of this sentence are in some particular instance of its occurrence, or the distribution of responses in a particular sequence of such occurrences. Instead, that is, of determining merely the truth-value of the statement, we determine a value compounded of the T value, P value, and E value, say, $T_a \cdot P_b \cdot E_c$. This done we turn to a consideration of some supposed synonym in the form of a complex sentence. Let us assume that it is the analysans in Stevenson's *Pattern-I*, that is, 'I like this; do so as well!' We first determine the T value of the statement component, namely, 'I like this,' and then determine the P value of 'Do so as well!' (In this analysis these are the only components. If in some other proffered analysis there were contained an exclamatory sentence as a component of the analysans, say: 'I like this; do so as well!; hurrah for this!' we should have to determine in addition another value, the E value of this exclamatory sentence.) On the basis of these determinations and in terms of the revised and expanded notion of equivalence we should say that 'I like this; do so as well!' is equivalent to (and hence possibly synonymous with) 'This is good' if and only if 'I like this; do so as well!' has the same T value and P value as does 'This is good' (and similarly in more involved cases such as that suggested above). We should, in short, have a method of determining non-cognitive synonymy of terms and sentences.[2]) If such a method should appear adequate, we can consider the question whether indeed there are any synonyms in this extended sense of 'synonymy,' based, so to speak, on a multidimensional analysis instead of the usual one-dimensional analysis.

We now raise the question whether there is any good reason to agree with those writers who hold, say, that 'good' is indefinable if a definition is expected to preserve its customary emotional meaning, or that to call a man 'kind-hearted and considerate' cannot serve to commend him. (I assume here that our analysis of what we may call 'non-cognitive

[1]) If this usual way is rather unusual, this is because we frequently do not bother to fix the meaning of such sentences with sufficient precision to make possible the ascertainment of their truth value; but this oversight can be overcome.

[2]) We imagine here, of course, that psychology has worked out the techniques and concepts required.

meaning' is an approximation to a plausible interpretation of the notion of emotive or evaluative meaning; we need not in this schematic treatment try to present any very exact equivalent.)

From what has been said above, this view (that 'good' is indefinable, etc.), if true, is, at best, a contingent truth. We have described, to be sure in rough outline only, a method of determining whether such non-cognitive synonymy holds between sentences and hence between terms, and we see thereby that it might hold, even if in fact it should not. Stevenson's own view appears to be based on the opinion that the emotive meaning of terms is a function of their emotional histories, and that since each different term has a different such history, it will have a different emotive meaning. ('Although our language affords many terms that have the same descriptive meaning, it is more economical with its emotive terms. Each term bears the characteristic stamp of its emotional history.' (p. 82, op. cit.)

Now, that some terms having the same descriptive meaning have different emotive-evaluative meaning and hence in our terms different P and E values is beyond dispute; but the indefinability of 'good' cannot be deduced from this fact. The question reduces simply to whether or not there are any terms which in addition to being cognitively (descriptively) synonymous with 'good' also have the same P and E values, that is, have the same (or possibly the same tendency to produce the same) persuasive-emotive effects as does 'good.' I have little doubt that the answer is in the affirmative, but having no statistics at hand (has anyone ever made a study of this matter in empirical terms?) I shall resort to an indirect method of showing this. The method involves two steps: 1) a definition of 'a moral man' and 2) asserting an empirical premise, namely, that there are some moral men. I trust that 2) will be granted; I turn then to the definition.

I suppose it will not be considered far-fetched or seriously off the track if we define 'a moral man' as follows: *X is a moral man if and only if X loves the good.* And I lay it down that X loves the good if and only if X responds to the knowledge or belief that Y is good with feelings of approbation of Y, and acts or tends to act so as to support and commend Y. Whatever a moral person conceives to be good he will approve and strive to support and make prosper. If he be a utilitarian, he will love whatever he supposes makes for the greatest happiness of the greatest number; if he be a member of a sect whose conception of the good is what is commanded by the Lord, he will love and support what is supposed by him to be commanded by or beloved of the Lord, and so on. But this evidently amounts to saying that a moral man is one who responds to the definiens of 'good' (determined by the moral values he happens to hold) in the same manner in which he responds to the definiendum, i.e., to 'good' itself. That is, he will interpret 'good' in such a way that 'A is good' will be L-equivalent to 'A is conducive to the greatest happiness of the greatest number' or 'A is beloved of the Lord' or whatever be the cognitive-descriptive synonym of 'good' on which he bases his definition. And not merely will he deem the definiendum to be cognitively synonymous with its definiens, but he will react to both statements in the manner we have described if they are to be non-cognitively synonymous. Similarly, if he is commending certain persons, it will be indifferent for him and his (properly chosen) auditors whether he commends them by means of calling them 'good' or by means of calling them 'God fearing' or 'lovers of mankind', or whatever be the appropriate term describing the kind of man deemed by him and them to be a morally good man. The effects will be the same.

Now I take it that it is not to flatter mankind unduly to suppose that there are some good men in this sense. All of us in fact from time manage to qualify for this title, and this is, as I have attempted to make clear, equivalent to saying that from time to time we find ourselves responding to our favorite definientia for 'good' with the same non-cognitive, i.e., emotive-evaluative, reactions with which we respond to the term 'good' itself.

If I am not mistaken, I have shown some reasons for rejecting the view that 'good' is indefinable if a definition is expected to preserve its customary emotive meaning. Whether this be a reasonable expectation or not is a question whose answer lies outside the scope of this paper. I happen to think it is not a reasonable expectation – but if anyone is anxious to know why, I must refer him to a chapter in one of my as yet unwritten works.

University of California, Berkeley

CHARLES MORRIS

ON THE HISTORY OF THE INTERNATIONAL ENCYCLOPEDIA OF UNIFIED SCIENCE

Rudolf Carnap accepted a professorship in philosophy at the University of Chicago in 1936. I had first met him during a stay in Prague in 1934. Otto Neurath made a visit to Chicago in October 1936, and the three of us met with Donald P. Bean of the University of Chicago Press to propose the publication of an *International Encyclopedia of Unified Science*. A contract with the Press to this end was signed in February 1937. Until Carnap left the University of Chicago in 1952 we worked closely together on the affairs of the *Encyclopedia*, and until Neurath's death in 1945 we each had a very voluminous correspondence with him, and exchanged copies of our own letters. On my desk as I write this is a pile of letters from those years well over a foot high. I have re-read most of them for this occasion of paying homage to Carnap. Some historical facts which I salvaged – especially concerning the *Encyclopedia* – seem to me of sufficient interest to record briefly.

The original idea of the *Encyclopedia* was Otto Neurath's. In a letter of 1935 he wrote that he was at work on the project at least as early as 1920. He wrote that he first talked it over with Einstein and Hans Hahn, and had early discussions about it with Carnap and Philipp Frank. In the 1930's Neurath was with the Mundaneum Institute of the Hague. He had set up 'The Unity of Science Institute' in 1936 as a department of this Institute, and in 1937 this was renamed 'The International Institute for the Unity of Science', with Neurath, Frank, and Morris forming the executive committee. There was also set up an 'Organization Committee of the International Encyclopedia of Unified Science', composed of Neurath, Carnap, Frank, Jørgen Jørgensen, Morris, and Louis Rougier. (Also formed was an 'Organization Committee of the International Congresses for the Unity of Science', composed of the same persons plus L. Susan Stebbing.) The general project of the *Encyclopedia* was discussed at length at the First International Congress for the Unity of Science, Paris, September 1935, and the Congress voted approval of the project. As the preceding details make clear, the idea of the *Encyclopedia* long preceded this Congress.

My correspondence contains three extensive statements of Neurath's ideas about the *Encyclopedia:* a five page outline in May 1936, a four page statement in February 1937, and another four page discussion in June 1938. There were many letters back and forth concerning these proposals. As will be seen, Neurath's ideas ranged very far.

In addition to the two introductory volumes (each to contain ten monographs) with the section title of *Foundations of the Unity of Science*, Neurath thought of two (and at times of three) other, and larger, sections. The three of us came to no agreement as to the proper title for these sections, but the general plan was clear. Section 2 was to deal with *methodological* problems involved in the special sciences and in the systematization of science, with particular stress to be laid upon the confrontation and discussion of divergent points of view. Section 3 was to concern itself with the *actual state of systematization* within the special sciences and the connections which obtained between them, with the hope that this might help toward further systematization. Neurath in 1938 was thinking of six volumes (60 monographs) for Section 2 and eight volumes (80 monographs) for Section 3. But there was even more in his bag of ideas than this.

Neurath had long planned a comprehensive *Visual Thesaurus* (sometimes he called it the Pictorial and sometimes the Isotype Thesaurus) which would be 'eine Weltübersicht in Bilden'. At times he thought that this might be an adjunct of the *Encyclopedia*, in which case he proposed an additional Section 4 for the *Encyclopedia* which would exemplify and apply the methods and results of the preceding three Sections to such fields as education, engineering, law, and medicine. Neurath proposed ten volumes for this Section, each of which would have a pictorial companion in the *Visual Thesaurus*. So at the most elaborate range of his proposals Neurath was thinking of a 26 volume (260 monograph) *Encyclopedia* supplemented by a ten volume *Visual Thesaurus*.

At one time Neurath had contemplated English, French, and German editions of the work. He conceived of it as genuinely international in scope, with writers from Asiatic contries as well as from the West. It is clear from a number of his letters that he had the great French *Encyclopedia* often in mind, both with respect to the historical importance which he envisaged for the work he planned, and with respect to the difficulties the two enterprises encountered. A letter of November 18, 1944, gives a long nine page discussion of Neurath's post-War plans for the various agencies of the unity of science movement (the *Encylopedia*, the *Journal*, the *Library*, and the Congresses). My last letter from him was dated November 7, 1945. He died in London on December 22, 1945, at the age of 63.

Carnap and I did not correspond at any length with Neurath concerning his proposed fourth section for the *Encyclopedia*, though both of us had doubts about it. Section 3 was discussed only in very general terms. Nor was any agreement reached on the details of Section 2, the methodology section. However, a general statement concerning this Section was agreed upon in 1939; it was to be used to obtain advance subscriptions, though because of the War it was never so used. It may be of historical interest to quote some paragraphs from this statement.

'Volumes III–VIII of the *International Encyclopedia of Unified Science*. The second unit will be composed of 6 volumes of 10 monographs each. This unit will stress the problems and procedures involved in the progressive systematization of science.

Differences of opinion exist among those interested in the analysis and integration of the language of science, and it is part of the purpose of the *Encyclopedia* to present such differences. It becomes necessary at the present stage of development to stress the problems encountered and the techniques relevant to their solution, – to take stock, as it were, of the contemporary situation in the analysis and the unification of scientific knowledge. This is the task of the second unit of the *Encyclopedia*. It thus exhibits in a new perspective the problems and procedures somewhat vaguely indicated in such phrases as 'the logic of science' and 'the methodology of science.' Its six volumes of ten monographs each offer the opportunity for treatment of the subject with a comprehensiveness not found elsewhere.

Many persons will be interested in knowing the various results and opinions which exist in this field. Hence Volumes III–VIII will especially stress the controversial differences in regard to special sciences (physics, psychology, etc.), in regard to the possibilities and limitations of scientific unification, and in regard to the methods involved in scientific progress and systematization ... Representatives of controversial opinions will be given a chance to present their views. A special Editorial Committee will plan each of the volumes.

The content of the separate volumes will be roughly as follows:

Volume III:	the most general problems and procedures which an ever-expanding unification of science encounters (construction and confirmation of theories, induction, historical attempts at scientific integration, probability, etc.).
Volume IV:	the nature of logic and mathematics, and their role and place in the structure of science.
Volume V:	physics.
Volume VI:	biology and psychology.
Volume VII:	the social and humanistic sciences.
Volume VIII:	history of the scientific attitude.'

We had hoped in 1939 to complete Section 2 by 1944. And here it is 1960, and the first Section, *Foundations of the Unity of Science*, is still not quite complete!

All of the authors of volume 1 as originally announced in the first monograph completed their monographs. But the second volume ran into many difficulties, and for one reason or another the studies proposed by Federigo Enriques, Jan Lukasiewicz, Arne Næss, Louis Rougier, and Louis Wirth never appeared. The monographs came in slowly, and at one time the Press even suggested suspension of the series until after the War.

The detailed work on the *Encyclopedia* (after the stage of planning) fell largely on Carnap and me. Carnap gave himself without stint to the reading of early drafts of the monographs and to the making of detailed suggestions for their revision. During a number of my long absences from the University, Carnap handled all relevant matters with the Press. He remained, with Hans Reichenbach, editor of *Erkenntnis* from its beginning in 1930 to its lapse in 1939 due to the War (then under the title of *Journal of Unified Science* which had been adopted in 1939). Volume IX was to have contained the abstracts of the papers presented at the Fifth International Congress for the Unity of Science, held at Harvard University, September 5–10, 1939, but the composition of this volume was destroyed by the Nazi invasion of Holland. Carnap was among those who endeavored to continue the *Journal* in the United States, but failure to find financial support made this continuance impossible. In all these ways Carnap was a faithful collaborator through years of hard but rewarding work.

The *Encyclopedia of Unified Science*, though now only a fragment of what had been planned, has had historical significance. The monographs are still very much alive. The movement of which the *Encyclopedia* was a part continues to develop vigorously in its own way. The Institute for the Unity of Science continues its activity in the United States under the leadership of Philipp Frank. Whether the larger plans for the *Encyclopedia* are ever to be resumed is a problem for another generation. But the years of the Congresses and the *Encyclopedia* were memorable years. It is a privilege to have worked with such men as Otto Neurath, Hans Reichenbach, Philipp Frank, and Rudolf Carnap on our common enterprises.

University of Florida