Bernays Project: Text No. 20

On judging the situation in proof theoretical research (with discussion) (1954)

Paul Bernays

(Zur Beurteilung der Situation in der beweistheoretischen Forschung (mit Diskussion), 1954.)

Translation by: *Dirk Schlimm. Revised by Bill Tait (3/11/03), and Awodey, Schlimm, and Sieg (5/15/03).*

Comments:

The French parts (two paragraphs) of the discussion still need to be translated.

When I speak here in brief about the situation in proof theoretic research, it appears appropriate to remind ourselves of what is characteristic of this research: it is the systematic investigation of the kinds of applications and the consequences of logical reasoning in the mathematical disciplines, in which the concept formations and the assumptions are fixed in such a way that a strict formalization of the proofs is possible with the help of the means of expression of symbolic logic.

As you know, Hilbert stimulated this kind of investigation mainly with regard of the questions of consistency. But he also has envisaged from the beginning the treatment of questions regarding the completeness and decidability in the framework of these investigations, for example already in the lecture *Axiomatic thinking* (1917). He formulated in more detail questions regarding completeness in the lecture *Problems of the founding of mathematics* in Bologna (1928).

To be sure, Hilbert imagined many things regarding both the results to be obtained and the method to be simpler than they eventually turned out to be. The realization of these major difficulties excited the idea in many that proof theoretic research has led to a definitive failure. But a glance at the actual state of affairs shows that this is out of the question: the methods of proof theoretic considerations find themselves in a rich state of development and considerable results have been obtained in various directions. Let me list some noteworthy successes regarding the problems Hilbert formulated:

1. Gödel's Completeness Theorem (proof of the completeness of the first order predicate calculus) together with its related extensions.

2. One succeeded in making the concept of decidability precise in such a way that systematic results could be obtained on the basis of this definition, in particular the proof of the unsolvability of the decision problem for predicate calculus by Church and, in a second way, by Turing.

3. While the methods aforementioned led only to conclusions concerning undecidability, Tarski succeeded, on the other hand, to specify decision procedures for certain mathematically non-trivial domains. In connection with these results as well as through results supplementing Gödel's completeness theorem, there have been applications in mathematics which are of interest not only to mathematicians concerned with foundations.

4. Regarding the questions of consistency, a consistency proof for full analysis has not been achieved from the finite standpoint, but one has been obtained for restricted analysis (for example in Weyl's sense or in the sense of ramified type theory) from a constructive standpoint. Gentzen first supplied such a proof for the number theoretic formalism; but Gentzen already had in mind the extension of his method to ramified analysis. This has then carried through by Lorenzen, Schütte, and Ackermann, whereby the method of proof also became more transparent. Also to be mentioned is a new transparent consistency proof for number theory by Stenius. Furthermore, it is remarkable that the extension of the finite standpoint to the constructive standpoint in a freer sense makes it possible to consider proofs that do not have to be formalized in the full sense, but can contain parts in which metamathematical derivations can be specified which depend each on a syntactical numerical parameter. In this way, one transcends the domain of those systems to which Gödel's incompleteness theorem applies.

By the way, this important theorem is by no means to be judged only as a negative result; rather it plays a role for proof theory similar to that of the discovery of the irrational numbers for arithmetic.

5. Finally, efforts have been made to supplement the statement of consistency with a more general form of question: what can be extracted from the formal provability of a theorem from the constructive standpoint? Kreisel's investigations move in this direction.

After all this it would obviously be totally inadequate to speak of a general fiasco of proof theory. On the other hand it must be acknowledged that not only the most essential work in this domain still has to be done, but also that, regarding the methodology, there is no clear resolution and no unanimity. I would like to raise a few points in this connection.

One speaks today a bit condescendingly about the "naive set theory." But we must, however, remind ourselves that it is, in any case, naive to think that, by a retreat to the axiomatic standpoint, without any contentual approach supporting it, we have at our disposal anything like what we started with. The retreat to the axiomatic in the case of non-Euclidean geometry is less problematic, because we there take as as basis arithmetic and set theory as given knowledge. The discussions about possible geometries, in particular the model theoretic considerations, take place within the framework of arithmetic (analysis). By challenging this framework and assigning to set theory itself the role of an axiomatic theory, it becomes necessary to determine a different underlying framework which has to act as the arithmetic proper. Different views are possible with regard to the choice of this methodological framework.

The minimal requirement for a sharpened axiomatization is that the objects not be taken from a domain that is regarded as being antecedent, but that they be constituted by generating processes. But the meaning of this could be that these generating processes determine the extension of the objects; this point of view motivates the *tertium non datur*. In fact, the openness of a domain can be understood in two senses: on the one hand, that the processes of construction lead beyond every single element, and on the other hand, that the resulting domain does not represent a mathematically determined manifold at all. Depending on whether the number sequence is understood in the first sense or in the second, one obtains the acknowledgment of the *tertium non datur* with respect to the numbers or the intuitionistic standpoint. For the finite standpoint the requirement is added that the considerations have to be made by means of investigating finite configurations, thus in particular assumptions in the form of general statements are excluded.

Even the maximal requirement for the methodical framework goes beyond that of the finite standpoint. This in fact contains existence assumptions, required for the possibility of systematic considerations, which are not self-evident from the standpoint of the properly concrete. For example, the application of such existence assumptions is necessary, if we want to show the eliminability of complete induction in the sense of Lorenzen. Originally, Hilbert also wanted to adopt the narrower standpoint which does not presuppose the intuitive general concept of numeral. This can be seen from his lecture in Heidelberg (1904) among others. It was already a kind of compromise that he decided in favor of adopting the finite standpoint in his publications. If we make ourselves clear on this, then the need for transition from the finite standpoint to an extended constructive standpoint does not appear to be so catastrophic.

To be sure, this requires a philosophical adjustment. Many think that one either has to accept only absolute evidence, or that evidence has to be generally abandoned as a feature of the sciences. Instead of this "all or nothing" attitude, it appears to be more appropriate to understand evidence as something that is acquired. The human being obtains evidences in the way he learns to walk, or as the birds learn to fly. Hereby one comes to the Socratic acknowledgment of our basic inability to know in advance. We only can try out points of view and standpoints in the theoretical realm and possibly have intellectual success with them.

It is not the opinion that with these points of view the problem of the foundations is already solved in principle. But at least such modesty allows that we not be completely disconcerted whenever new antinomies are discovered. Such antinomies then appear rather to be instructive clues for the right choice of our approaches and methods.

The problematic in the foundational research that still has not yet been overcome consists of different aspects: on the one hand, in respect to the choice of the methodical standpoint in the foundational research, as well as the choice of the deductive framework, on the other hand, in respect to the understanding of mathematics. With regard to this second point a decision is maybe not to be expected by means of the foundational research, but in respect to the first questions it is not too immodest to hope that the comparison of the results of the different directions of research will yield a clear advantage to one of the ways of proceeding in the foreseeable future.

Technische Hochschule, Zürich

Résumé of the address of Alfred TARSKI

Mr. Tarski gives a detailed account of the decision problem in logic and mathematics: clarification of the problem, precise formulation in terms of general recursive sets and a discussion of the adequacy of this formulation from the intuitive point of view, methods applied in studying the problem, main results so far obtained, important unanswered questions, some philosophical aspects of the problem itself and of the known results. The discussion is largely based on the monograph of Tarski, *A Decision Method for Elementary Algebra and*

Geometry (Berkeley, 1951), and the joint monograph of Tarski, Mostowski, and Robinson, *Undecidable Theories* (Amsterdam, 1953).

University of California, Berkeley (Cal.).

DISCUSSION

M. Arnold Schmidt. — My introduction of degrees of consistency, which has been mentioned by Mr. Bernays, was merely meant to emphasize the problem of the role that consistency may play epistemologically. In this connection I may mention a phenomenon that is related to the degrees of consistency: in the newer proof theoretic investigations of Schütte and Lorenzen on mathematical codifications that include the numbers of the second number class the consistency proof succeeds only up to a particular segment of that number class when the metamathematics is precisely restricted.

With regard to the extensions that the finite standpoint has experienced in the course of its development I'd like to remark that the tertium non datur remains excluded at all stages of this development.

With respect to the problem of the evidence one is allowed to say in a certain analogy to the interpretation of the Kantian a priori that the individual can obtain evidence through reflection, but that the criteria for the evidence must be independent of such experience in order to rule out deceptive evidence which can arise by habituation. As much as I acknowledge that the matters of fact which are not evident at first sight can become evident by a thorough clarification, I want to emphasize, on the other hand, that in my opinion there can be *only one* kind of evidence, thus no relative or staged evidence. From this point of view the task of the proof consists in reducing something that is not evident to something that is evident.

M. P. Bernays. — There is no disagreement with regard to the first point. With respect to the second remark I'd like to call attention to the fact that I did not intend to write history. Had this been the case, I would have distinguished five stages of metamathematics: 1. the finite standpoint, 2. the definite standpoint (1 with existence assumptions), 3. Intuitionism, 4. tertium non datur, 5. impredicative concept formation. This ordering gives more and more freedom. While it was possible to point out intimate agreements between Intuitionism (3) and the classical standpoint (4), this has not succeeded for (4) and (5) although Gentzen had struggled with it. Thus the decisive point lies beyond the introduction of the tertium non datur. Finally I'd like to say that one must not just construct the evidence objectively and forget about the subjective determinations.

M. W. Quine. — Tarski expressed some doubts whether every general recursive function is mechanically computable in an intuitive sense of those words. A function F is general recursive if and only if the following conditions hold: there is a finite set S_0. of elementary equations such that: a) for every argument n (given by numeral), there is a sequence S of equations, so that the last equation in S gives the value of $F(n)$ and b) each equation in S either belongs to S_0 or is obtained from earlier equations in S by substitution for variables or by substitutivity of identity.

Then, for any recursive function F there is the following method for computing $F(n)$ for any n. Let S be the class of all finite sequences fulfilling (b) above, omitting purely alphabetical variations. The sequences in S, can be enumerated, by consideration of alphabetical order and increasing length. Then examine each sequence in S, in order of enumeration, until one is encountered which ends by evaluating $F(n)$. This process is mechanical, and it is bound to terminate if F is general recursive.

M. A. Tarski. — I suggest the following clarification: For any given formalized system, in which, technically speaking, the set of provable sentences is recursively enumerable, we can distinguish between two kinds of machines: a proof machine and a decision machine. The proof machine will produce all the provable sentences—in an amount of time which, for each sentence, will be finite but which in general cannot be evaluated in advance. The decision machine will permit us to determine whether any given sentence is provable or not, and will do this in amount of time which, for each sentence, can be evaluated in advance. Imagine now the following situation: For some formalized system we have a proof machine but no decision machine. On the other hand, we have proved by non-constructive methods that this system is complete and consistent, and that consequently the set of provable sentences is general recursive. (Interesting examples of such non-constructive proofs were pointed out recently by R. L. Vaught in the *Bulletin of the American Mathematical Society*, vol. 69, 1953, pp. 396-397.) Then, for any given sentence, we are sure that the proof machine will produce in a finite amount of time either this sentence or its negation, and hence will inform us whether or not the sentence is provable. However, the amount of time required cannot be evaluated in advance. Thus we are confronted with a case in which it is really not clear whether we have a positive solution to the decision problem. It seems that the situation can be summarized as follows: To obtain a negative solution to the decision problem it suffices to show that the set of provable sentences is not general recursive. To obtain what would intuitively be a satisfactory positive solution it is necessary not only to show that the set of provable sentences is general recursive, but to show this in a constructive way. Of course, a precise discussion of the matter would require a clarification of the notion of a constructive proof.

M. H. Behmann. — I'd like to come back to the question regarding the evidence of geometry. As Helmholz argues plausibly in his popular scientific talks, the different "evidence" of the various geometries — in particular of the elliptic (Riemann), the parabolic (Euclid), and the hyperbolic (Bolyai-Lobatschewsky) — is by no means a quality that belongs to the respective geometry "by itself," but something that is peculiar to them in different degrees merely "for us" — depending on the degree of how familiar and how accustomed we are in fact —, just as it is not a property of the language whether it is intelligible for somebody and unintelligible of somebody else. To prove this practically, I have occasionally reproduced in a drawing the optical impression of a simple object, namely a street with three houses on each side, in a hyperbolical world. (The houses are erected on a horizontal plane with congruent equally angled quadrangles as ground plan alongside two parallel curves with equal distances to a straight line; the construction is based on the spherical or the projective model with the center of the sphere of reference as the point of view.) At first one has the impression of a distorted perspective — compared to the one we are used from the objectiveness of our world —, but soon hereafter it is possible to envision to some extent how one would go about to familiarize oneself with such a world, if one were continously presented by suitable means (trick films) with the optical experiences which correspond to a hyperbolical world, in order to "represent" already in advance possible experiences in it, as, e.g., the sequence of the optical impressions when driving trough a street. The character of the evidence of the geometries is thus largely subjective, moreover it varies individually and it can be influenced actively.

M. P. Bernays. — Although differently constituted beings could have a different evidence, it is our concern to determine what evidence is for us.

M. H. Dingler. — Evidence is always the renunciation of actual foundation. Since "something is learned" represents a theory, an insecure sentence that has to be verified, it is impossible for this sentence to lie at the basis of our knowledge, as it is the case of the evidence. Thus learned evidence does not occur. It is not possible to justify something by having learned it, because it is possible to learn everything.

M. A. Tarski. — I would like to add two historical remarks to the exposition of Mr. Bernays: 1. The first deep metamathematical results (obtained in addition by specifically metamathematical methods) are due to L. Löwenheim (*Mathematische Annalen*, vol. 76, 1915, pp. 447-470); his work does not seem to have been influenced by Hilbert. 2. As an essential contribution of the Polish school to the development of metamathematics one can regard the fact that from the very beginning it admitted into metamathematical research all fruitful methods, whether finitary or not. Restriction to finitary methods seems natural in certain parts of metamathematics, in particular in the discussion of consistency problems, though even here these methods may be inadequate. At present it seems certain, however, that exclusive adherence to these methods would prove a great handicap in the development of metamathematics. Furthermore I should like to remark that there seems to be a tendency among mathematical logicians to overemphasize the importance of consistency problems, and that the philosophical value of the results obtained so far in this direction seems somewhat dubious. Gentzen's proof of the consistency of arithmetic is undoubtedly a very interesting metamathematical result, which may prove very stimulating and fruitful. I cannot say, however, that the consistency of arithmetic is now much more evident to me (at any rate, perhaps, to use the terminology of the differential calculus, more evident than by epsilon) than it was before the proof was given. To clarify a little my reactions: Let G be a formalism just adequate for formalizing Gentzen's proof, and let A be the formalism of arithmetic. It is interesting that the consistency of A can be proved in G; it would perhaps be equally interesting if it should turn out that the consistency of G can be proved in A.

M. P. Bernays. — My thought has not been rightly interpreted. I did not wish to say that Gentzen's proof made arithmetic or truths about arithmetic more evident. But I tried to stress that some mathematical methods allow simultaneously to show deductibility and validity.[1]

[1] Editorial footnote: This paragraph is in English in the original.

M. A. Tarski. — To supplement the concluding remarks of Mr. Bernays' talk, I should like to point out a new direction of metamathematical research—the study of the relations the exposition between models of formal systems and the syntactical properties of these systems (in other words, the semantics of formal systems). The problems studied in this domain are of the following character: Knowing the formal structure of an axiom system, what can we say about the mathematical properties of the models of this system; conversely, given a class of models having certain mathematical properties, what can we say about the formal structure of postulate systems by means of which we can define this class of models? As an example of results so far obtained I may mention a theorem of G. Birkhoff (*Proceedings of the Cambridge Philosophical Society*, vol. 31, 1935, pp. 433-454), in which he gives a full mathematical characterization of those classes of algebras which can be defined by systems of algebraic identities. kn outstanding open problem is that of providing a mathematical characterization of those classes of models which can be defined by means of arbitrary postulate systems formulated within the first-order predicate calculus. As a last remark, I should like to express my full agreement with the opinion of Mr. Bernays that the state of knowledge in the realm of foundations is still *etwas sehr unabgeschlossenes*. I am far from deploring this situation, for I am inclined to believe that a scientific theory is only then *abgeschlossen* when it has entered a period of senility and is in fact close to its death. I do not understand why it is advisable to wait for the death of a theory in order to draw from its results more general, say philosophical, conclusions. Such conclusions may be drawn whenever this activity is justified by the character of the results obtained; of course, these conclusions may be subjected to revision in the light of later developments.

M. Ch. Perelman. — Si, d'après la conception classique, cartésienne et pascalienne, l'évidence nous dispense de la preuve, il est également vrai qu'il faut une forme d'6vidence pour accepter qu'il y a preuve. Pour éviter un cercle, nous sommes bien obligés de distinguer, dans le sens des remarques de M. Bernays, des degrés dans l'évidence ou, du moins, des formes d'évidence.

M. H. Dingler. — It is correct that a proof cannot be effectuated without evidence; but I distinguish between two kinds of evidences: one is justifiable from the knowledge of what has been intended implicitly (knowing evidence), the other does not have such a justification (empty evidence). The latter is useless for science as means of proof.

M. Y. Bar-Hillel. — It seems to me that our discussion has shown signs of a well known confusion between two different, though related, meanings of "proof": the syntactical (or semantical) one, in which a proof is a certain series of sentences, and the pragmatical one, in which a proof is a certain action by a human being, taken in order to convince another such being of something or other. Proof, in its first sense, has no connection whatsoever with such psychological notions as evidence.

I have also the impression as if another confusion has crept into our discussion: that between *a proof with premises* (or a *derivation*, in the terminology of Carnap's Logical Syntax) and a *proof without premises* (a *derivation without premises* or simply *proof*). It should be of importance to keep these two notions apart from now on.

M. Ch. Perelman. — Pour qu'on puisse parler d'une conception syntaxique de la preuve, il faudrait que cette forme de preuve soit convaincante pour quelqu'un; une transformation syntaxique n'est donc une preuve que grâce à ses propriétés pragmatiques.

Le président, le professeur **A. Heyting**, clôture la séance, après avoir constaté que la discussion a montré la présence d'éléments intuitifs dans les mathématiques.