THE CONSISTENCY OF ARITHMETIC

Robert K. Meyer

Australian National University

Abstract

This paper offers an elementary proof that formal arithmetic is consistent. The system that will be proved consistent is a first-order theory $R^\sharp$, based as usual on the Peano postulates and the recursion equations for $+$ and $\times$. However, the reasoning will apply to any axiomatizable extension of $R^\sharp$ got by adding classical arithmetical truths. Moreover, it will continue to apply through a large range of variation of the underlying logic of $R^\sharp$, while on a simple and straightforward translation, the classical first-order theory $P^\sharp$ of Peano arithmetic turns out to be an exact subsystem of $R^\sharp$. Since the reasoning is elementary, it is formalizable within $R^\sharp$ itself; i.e., we can actually demonstrate within $R^\sharp$ (or within $P^\sharp$, if we care) a statement that, in a natural fashion, asserts the consistency of $R^\sharp$ itself.

The reader is unlikely to have missed the significance of the remarks just made. In plain English, this paper *repeals* Gödel's famous second theorem. (That's the one that asserts that sufficiently strong systems are inadequate to demonstrate their own consistency.) That theorem (or at least the significance usually claimed for it) was a *mistake*—a subtle and understandable mistake, perhaps, but a mistake nonetheless. Accordingly, this paper *reinstates* the formal program which is often taken to have been blasted away by Gödel's theorems—namely, the Hilbert program of demonstrating, by methods that everybody can recognize as effective and finitary, that intuitive mathematics is reliable. Indeed, the present consistency proof for arithmetic will be recognized as correct by anyone who can count to 3. (So much, indeed, for the claim that the reliability of arithmetic rests on transfinite induction up to ε_0, and for the incredible mythology that underlies it.)

I

Repealing Gödel's theorem is no small matter. That theorem, after all, has often been taken to be the central fact of modern mathematical logic. Recursion theory is its child; much of model theory is at least its step-child. Moreover, if there is an error in Gödel's arguments, or in the significance ordinarily attributed to them, the mistake itself must be of a subtle—in fact, almost of an excusable—sort. For no arguments in the subject have been examined more carefully, or worked out (in a number of alternative forms) more rigorously. I.e., a mistake, if there is one, would have to be of the kind that we are almost programmed to overlook, or even to condone.

Moreover, if we claim to have found a mistake, we must take pains that, in the process of correcting it, we do not ourselves stray into areas that Gödel has correctly asserted to be off-limits. Thus, for example, the present paper offers no correction of Gödel's famous first theorem; to the contrary, we shall demonstrate, as expected, that $R^\sharp$ and all of its consistent axiomatizable extensions are *incomplete.* Nor do we offer a *technical* correction of the second theorem; that there is a formula *CONSIS* of $P^\sharp$, which may be taken to express the consistency of $P^\sharp$ but which is unprovable in $P^\sharp$, is not here in question.

What is offered, rather, is a *philosophical* correction. First, we correct the impression that $P^\sharp$, or anything like it, is the system whose consistency needs to be proved. Second, we establish the consistency of the system $R^\sharp$ (and of related systems) that may be more or less adequate to the needs of mathematics.

Let us hurry on, now, to the crux of the matter. That crux lies in the little sign $\sim$. $\sim$, of course, is supposed to mean 'not.' And the real import of the Gödel arguments, which are so often tied to such delicate matters as numeralwise expressibility, systematic unprovability, perverse self-reference, and the like, may in fact be summed up quite succinctly. In any serious attempt to formalize mathematics, or even arithmetic, effectively, $\sim$ *never* means what it is supposed to mean, *within* a particular system. To be sure, we can claim to give a semantic interpretation of an effectively presented arithmetic, and tell the world that, on this interpretation, $\sim$ means 'not.' The world will then ask how it comes about that, on some occasions on which A is false, we cannot prove $\sim A$ within the system. And we shall perhaps reply that, relative to the interpretation, the system is incomplete. But how much more accurate it would be to reply that, because of certain

formal anomalies that arise in the technical engineering, we just cannot so fix things up that $\sim$ works within the system the way that 'not' is to be taken as working in English. (In view of the Liar Paradox and other natural anomalies, however, we might more candidly admit that we don't know how 'not' really works in English, either.)

We make this point clear, on Gödelian grounds, with respect to the standard first-order Peano arithmetic $\mathsf{P}^\sharp$. We suppose some standard coding (Gödel numbering) that assigns to each formula A a unique natural number I. (In future, we suppose that this coding is an effective bijection from the natural numbers onto formulas and we shall use A^I for the formula with Gödel number I.) We let $\mathbf{N}$ be the set of all natural numbers, and, taking our formal system abstractly, we *identify* natural numbers with the corresponding numerals of the system. Then, as is well-known, there is an open formula Px, with sole free variable x, which serves as a *provability predicate* for $\mathsf{P}^\sharp$ in the following sense:

(1) For all natural numbers I, PI is a theorem of $\mathsf{P}^\sharp$ iff A^I is a theorem of $\mathsf{P}^\sharp$.

Moreover, taking truth with respect to the standard model $\mathbb{N} = \langle \mathbf{N}, \,', +, \times \rangle$ in the usual Tarskian sense, we have also

(2) For all natural numbers I, PI is true iff A^I is a theorem of $\mathsf{P}^\sharp$.

Accordingly, on our semantic understanding of $\sim$ we have immediately from (2),

(3) For all natural numbers I, $\sim PI$ is true iff A^I is *not* a theorem of $\mathsf{P}^\sharp$.

But we do *not* have, on straightforward Gödelian grounds,

(4) For all natural numbers I, $\sim PI$ is a theorem of $\mathsf{P}^\sharp$ iff A^I is *not* a theorem of $\mathsf{P}^\sharp$.

I.e., viewed extrinsically, as in (3), we may perhaps view $\sim$ to be a formal counterpart of 'not,' for the context under consideration. Viewed systemati-

cally, and given (1), what clearer demonstration could we ask than (4) of the proposition that $\sim$ just doesn't work formally in $\mathsf{P}^\sharp$ the way that 'not' works intuitively, given that (4) must fail (on pain of total bankruptcy otherwise of the standard mythology)?

We fix the considerations and notation of the previous paragraph for the rest of the paper (including the standard Gödel numbering, which there is no need to specify further). And we next note that these considerations are central to an examination of the character and import of Gödel's second theorem. It is of little interest that we cannot *prove* the consistency of $\mathsf{P}^\sharp$ within itself, unless $\mathsf{P}^\sharp$ has the vocabulary to say that it is consistent (and, moreover, that what it says in this vocabulary is in fact unprovable). And now we note the following.

In the first place, nobody ever expected $\mathsf{P}^\sharp$ to be muttering introspectively about itself at all. $\mathsf{P}^\sharp$ was constructed to say that $5 + 3 = 8$, that every number greater than 1 has a prime divisor, and the like. It was not constructed to say that nobody ever loved it before Hilbert came along, that it often wishes it were complete, or, for present purposes, that it is consistent. Of course, after Gödel, we are all now prepared to believe that $\mathsf{P}^\sharp$ does introspect, in code. And this has made its psychoanalysis (or whatever the equivalent process is when applied to formal systems) a regular element in the training of mathematical logicians. And we should note that, as in all psychoanalysis, there is a certain indistinctness in the method. When, e.g., $\mathsf{P}^\sharp$ seems to say "Every natural number is the sum of four squares," we may suppose that it is bragging "Show me a sentential tautology that I can't prove" or perhaps complaining "I can only demonstrate Fermat's last theorem for regular primes." And this should lead to a little humility on our part; stripped of the code, $\mathsf{P}^\sharp$ is still *saying*, "Every natural number is the sum of four primes." The rest we read into what it says, and, since $\mathsf{P}^\sharp$ is not itself a complete system, we err even in simple arithmetic if we try to interpret what it says categorically.

Second, we must exercise unusual care, as Feferman pointed out in [1] and as Gödel himself reflected in a recent footnote to [2], in finding a technical form for Gödel's second theorem. We must formalize the statement, uttered by $\mathsf{P}^\sharp$ about itself, "I am consistent." Here, our previous worries about negation are going to come to the fore, in a couple of ways. First, any statement that may be reasonably taken in $\mathsf{P}^\sharp$ to express its own consistency is presumptively going to involve negation in an essential way. There are various intuitively equivalent ways of asserting the consistency of $\mathsf{P}^\sharp$, in English, but

all of them in the end are assertions like, "This is not provable" or "Nothing of this form is provable" or "If this is provable, then that is not provable." Evidently, if we are worried about our capacity to *express* 'not' in $\mathsf{P}^\sharp$, and to express 'not provable' in particular, we are going to have to worry about whether we can express in $\mathsf{P}^\sharp$ a statement that says that $\mathsf{P}^\sharp$ is consistent, even in code.

Third, accordingly, one begins to wonder what Gödel's second theorem adds to his first theorem. No one expects us to be able to prove what we cannot say. And it then seems otiose to claim that any effective, finitary proof of the consistency of formal arithmetic would yield a proof in $\mathsf{P}^\sharp$ of a formula which, so far as $\mathsf{P}^\sharp$ is concerned, is only a dubious candidate for the role of being *the statement* in the vocabulary of $\mathsf{P}^\sharp$ which expresses the consistency of $\mathsf{P}^\sharp$.

We shall not linger over these issues. Indeed, we shall even take, below, a rather orthodox stand with respect to them. But we do note that they are exacerbated when we ask, "What particular form, even in English, should the statement that $\mathsf{P}^\sharp$ (or any formal arithmetic) take?" We devote the next section to some of these problems, again singling out the crucial role of negation.

II

First, let us ask ourselves, "Why should we (or Hilbert) have *cared* whether arithmetic is consistent?" And let us now answer the question. Mainly, we don't care. Put optimistically, we are so strongly convinced that arithmetic is consistent (or at least that any part of it of which serious use is going to be made is consistent), that demonstrating its consistency is just a game—the game of seeing how little, or how much, is required for a formal consistency proof. After all, it was the reliability of mathematical *analysis* that truly worried Hilbert and others. And, since the ultimate effect of the great 19th and early 20th century programs was to substitute insecurity in reasonings about infinitesimals and series, the neck-wringing administered to these programs in 1931 have registered no gain on that main point. Indeed, they have led, since 1931, to a certain abandonment of the main point, all hands being needed (as Reid puts it aptly in [3]) to defend the homeland of arithmetic.

While we are playing the game, however, what we are presumably concerned to show is that our intuitive arithmetic is *reliable*, by establishing

that our carefully chosen formal counterparts of that arithmetic are reliable. So long as we accept, as we do for present purposes, Gödel's first theorem, part of that task remains beyond us. For, according to that theorem, no formal system (in the ordinary first-order vocabulary) that can be effectively presented for the arithmetic of the natural numbers, or for any theory containing that arithmetic, will serve as a fully acceptable formal counterpart of the intuitive arithmetic.

So any formal arithmetic is, intuitively speaking, at best a partial arithmetic. Being partial, however, is not the same thing as being unreliable. And this is the point at which the usual appeals to Gödel's second theorem suggest that the entire situation is absolutely perverse. For it is claimed that no bag of mathematical tricks can be demonstrated to be reliable, except on appeal to some trick that isn't in the bag. I.e., we are confronted with a picture on which the reliability of any mathematical system (save such as are inadequate for whole number arithmetic) can only be demonstrated in some system less reliable, *prima facie*, than the system from which we began.

This picture, if accurate, severs mathematical logic from its chief foundational purpose: Namely, making possible a rigorous reconstruction of intuitive mathematics. Instead, one gets the impression that even the reconstruction of simple arithmetic is dubious enough, and that every step on from there becomes even more dubious. And it is accordingly no wonder that mathematical logicians, by and large, have gone off to live in a world of their own—a world, frankly, that has little relevance to mathematics, even less to the philosophy of mathematics, and almost none to general philosophy. For the depressing picture is that *more* than intuitive mathematics must be assumed in order to reconstruct intuitive mathematics. Chauvinist mathematicians (e.g., Poincaré), who always bridled at the suggestion that their discipline was just pure logic, may find cause to rejoice in this picture. But logicians must weep, for it denigrates exact thought for the sake of the old mumbo-jumbo.

We began to talk about consistency, but we have slipped in this section to talk about reliability. Consistency is a formal property of formal systems (though, depending on the author, it may be any one of several properties, not necessarily related). Reliability is an intuitive property, measuring a formal system against the purposes for which it was designed. Let it be clear, then, from the outset that it is reliability that is most desired; we wish, after all, our formal systems to be adequate to their purposes. And, whatever formal property we decide to identify with the honorific 'consistent,' it is of interest

to us only insofar as possession of this property by a formal system is a guide to, and hopefully a guarantee of, the reliability of the system in question.

With that point in mind, let us think of several things that we might wish to mean by consistent (with, to begin with, immediate application to classical first-order Peano arithmetic $\mathsf{P}^\sharp$, but *mutatis mutandis* to other systems).

(1) $\mathsf{P}^\sharp$ is consistent iff there is a formula A unprovable in $\mathsf{P}^\sharp$.

(2) $\mathsf{P}^\sharp$ is consistent iff $0 = 1$ is unprovable in $\mathsf{P}^\sharp$.

(3) $\mathsf{P}^\sharp$ is consistent iff, for each pair I, J of distinct natural numbers, $I = J$ is unprovable in $\mathsf{P}^\sharp$.

(4) $\mathsf{P}^\sharp$ is consistent iff, for each pair T, U of distinct polynomials, $T = U$ is unprovable in $\mathsf{P}^\sharp$.

(5) $\mathsf{P}^\sharp$ is consistent iff $0 \neq 0$ is unprovable in $\mathsf{P}^\sharp$.

(6) $\mathsf{P}^\sharp$ is consistent iff, for each formula A, at least one of $A, \sim A$ is unprovable in $\mathsf{P}^\sharp$.

(7) $\mathsf{P}^\sharp$ is consistent iff, for each formula A, if A is provable in $\mathsf{P}^\sharp$ then A is arithmetically true.

These suggested characterizations of consistency are arranged in what, in intuitive terms, might be taken as an ever more severe order. For, evidently, and without regard for any deduction-theoretic structure whatsoever, if everything is provable in $\mathsf{P}^\sharp$, so is $0 = 1$. If $0 = 1$ is provable, so is $I = J$ for some distinct numbers I, J. $I = J$, for distinct numbers I, J, is a degenerate case of a bad polynomial. This takes us, in increasing intuitive severity, from (1) through (4).

(5), to some degree, is a horse of a different color, though we show below that, on minimal deduction-theoretic assumptions, any system consistent in sense (5) will be consistent in sense (4) and hence in the lesser senses. Since we presume that $0 = 0$ is provable, together with all good polynomial identities, any system consistent in sense (6) will be consistent in senses (5) and (4), and so on down.

(7), to be sure, is a semantic kind of consistency, whereas (1) through (6) were deduction-theoretic. Moreover, it is conceptually non-constructive, since arithmetic truth (i.e., truth in the standard model $\mathbb{N}$) cannot be effectively characterized. But, we suppose, to the extent that one understands (7), one understands that, for every formula A, not both A and $\sim A$ can be true together, whence (7) implies (6) and so on down.

Possession of any of these properties is a greater index of the reliability of $P^\sharp$ than is possession of the property before it. Accordingly, we should like $P^\sharp$ to have them all, and to have (7) in particular. From another viewpoint, however, (7) is the least hopeful, since, apart from arithmetic faith, we don't even know what it means. And, on the other hand, if we have the faith then it is pretty trivial that (7) holds.

Our concern in this paper, however, is with reason, not faith. Accordingly, we shall scrap (7), save as it plays a simplifying and heuristic role in carrying out certain arguments (which we shall not trouble ourselves to carry out rigorously). But, for $P^\sharp$ in particular, a strange thing happens with (1) through (6). For, though we graded them intuitively on an increasing order of severity, the deduction-theoretic character of $P^\sharp$ is such that, with respect to that system, they are all equivalent. For (1) outright implies (6), in the sense that, if any contradiction is provable in $P^\sharp$, then (by the principle of *ex falso quodlibet*) absolutely everything is provable in $P^\sharp$. So, without any semantic appeals of the sort involved in (7), if $P^\sharp$ is consistent in sense (1) it is already consistent in sense (6), and hence in every intermediate sense.

It is sometimes urged that this equating of various senses of consistency is quite reasonable, on the grounds that a system in which, for any A whatsoever, both of A, $\sim A$ were provable would already be unacceptable. That's as may be; I have disagreed on this point elsewhere (e.g., in [4]), but I shall not press the grounds of that disagreement here. Rather, what I wish to note here is that what is at issue presently is not the *fact* of the consistency of $P^\sharp$ (or of another formal arithmetic) but the *proof* of consistency: i.e., formal demonstration of the reliability of some of our intuitive reasonings.

At this point, the *ex falso quodlibet* of $P^\sharp$ (and, equally, of Heyting's formulation $J^\sharp$ of first-order intuitionist arithmetic) imposes some strain upon this project. For it is a question of *which reasonings* we wish to demonstrate to be reliable. And here we must not confuse with arithmetic itself what belongs more properly to the logical superstructure that we *impose upon* arithmetic, for the purpose of carrying out our arithmetical reasonings.

Again, we revert to the role of negation in formal arithmetic. Classically, we impose upon negation various formal properties—e.g., that the laws of excluded middle, non-contradiction, double negation, etc., hold for it. By no stretch of the imagination can these be considered *arithmetical properties*; indeed, what is valid in the Brouwerian intuitionist critique of classical mathematics can perhaps be seen as a denial of the claim that logical truths like excluded middle shall count *ipso facto* as mathematical truths.

However, the intuitionists quickly err, from the common perspective, in going on to deny that logical truths shall count as *truths*; and, for that matter, to deny specifically that the law of excluded middle, say, shall be available for our arithmetic *reasoning*. To this it may be objected, on good grounds, that excluded middle is built into our very understanding of negation, and that a purported negation (like Heyting's) which lacks this property is not negation at all but only an inferior simulation thereof.

In short, it makes sense to separate what is logical from what is arithmetical. And negation, in particular, must count as logical, not as arithmetical. Arithmetic, at its most elementary level, informs us that $5+3 = 8$. No logical confusions should leave this proposition in any doubt whatsoever, however much our heads begin to swim when we come to think about Gödel formulas. And it is just a defect of the received logic that, if anything is in doubt, $5 + 3 = 8$ is also in doubt (in the sense, e.g., that a demonstration of the Gödel formula leads, in a few more steps, say, to a demonstration of $5+3 = 9$, $5 + 3 = 10$, and so forth).

What is wrong with $\mathsf{P}^\sharp$, in short, and what has led to inflated claims for and interpretations of Gödel's second theorem, is that it fails to account for relevance as a constituent in valid arithmetical argument. Accordingly, $\mathsf{P}^\sharp$ is unable to discriminate between logical anomaly and total arithmetical breakdown. And this is a function, in part, of the way in which $\mathsf{P}^\sharp$ is customarily formulated. A more satisfactory formulation of formal arithmetic, while not offering total safeguards against anomaly, effectively guards us against breakdown.

Specifically, what I shall show here is that a relevantly formulated arithmetic $\mathsf{R}^\sharp$ is trivially consistent in any of the sense (1) through (4) above. Accordingly, with respect to what most reasonably counts as elementary arithmetic—not any formal system, but the simple manipulations that we all learned in school—$\mathsf{R}^\sharp$ is absolutely reliable. Accordingly, Gödelian puzzles about the reliability of arithmetic are relegated to their rightful role—namely, they are puzzles about the logical superstructure of arithmetic, which do not infect the properly arithmetical, and which accordingly do not impose upon the properly arithmetical the degree of doubt (e.g., that $0 = 1$ is unprovable) that they are customarily held to impose.

In lifting a Gödelian restriction, I looked forward above to a certain reassertion of Hilbert's strongly finitistic viewpoint. I can now be more precise about the lines of this reassertion (developing views already set out, in part, in [6]). Hilbert wished finitistic assurance of the reliability of non-constructive

reasoning; in such reasoning, the properties of classical negation play an important role. In so doing, however, he overestimated the importance of consistency in sense (6) above (which, in his framework, he was unable to distinguish from consistency even in the most minimal sense (1)).

For let us not be misled about the importance for our general reasonings of consistency in sense (6). It may be that, as Tarski and others suggest, English and natural languages in general are inconsistent in sense (6), inasmuch as the semantical paradoxes (e.g., the Liar) can be developed therein). In pursuing our natural reasonings, as I learned from Belnap, this fact is, even if true, only a flea bite in the pursuit of knowledge; we simply isolate the putative contradiction and ignore it. Only the *ex falso quodlibet* transforms it into a whale bite, wrecking the pursuit of knowledge.

How much more, then, should we fail to be upset, in our formal reasonings, about the absence of an effective *guarantee* of consistency in sense (6). When we *impose* logic on our arithmetical intuitions, what we really want is a guarantee that this imposition does not cause those intuitions to break down. And note that it is the arithmetical intuitions of which we seek a safeguard— not the logical ones. Elementary arithmetic is safe; we knew that before we began to apply logic to it. What we wish to be sure of, as Hilbert might have put it, is that *excursions through* general logical laws (e.g., excluded middle), either by way of *developing* a simple basis for arithmetic that builds in logic, or of *extending* our arithmetical knowledge, do not render *dubious* what we rightly regard as *indubitable.*

One aspect, accordingly, of a familiar myth must go; namely, that Gödel has consigned us to a winding staircase of increasing epistemological uncertainty about mathematics, each level of which, though it may appear sound from above, seems shaky when one is on it; and where one may always dread that, climbing as we do through the dark, the next level will lead directly into the void. For where Hilbert went wrong was in his assumption that there was a *single* test, which might be expressed *equivalently* in any of the forms (1) through (6), of the reliability of various counterparts of intuitive mathematics. No longer, I hope, will the demonstrable unprovability of $0 = 1$ in a given formal arithmetic be held to be beyond the scope of the methods formalizable in that arithmetic; for this, indeed, is methodological doubt with a vengeance, leaving one no alternative between blind faith in the adequacy of one's formal methods (given that proofs of their adequacy are open to increasing methodological doubt), or hopeless scepticism regarding the adequacy of any formal methods (save for very weak systems).

In place of this aspect of the usual myth, we have instead a task: Namely, to furnish new formal criteria to go with our intuitive fashioning of system to purpose, and to demonstrate effectively that our systems satisfy these criteria. Here we must become more clear about what belongs to the mathematical intuitions themselves and what belongs to the logical superstructure; while I have singled out $\sim$, in particular, as being associated with the superstructure, the degree to which the association can be made precise is of course dependent on definite formal results and particular formal points of view; evidently mathematicians, no less than any of the rest of us, are entitled to say 'not' on occasion when expressing underlying intuitions, and it is a task in itself to attempt the separation of the logical and the mathematical aspects of their so saying.

Beyond that, we must also move on beyond arithmetic; *those* intuitions, after all, are *not* the ones that we are most concerned to safeguard, from an intuitive point of view. From the present viewpoint, the task of furnishing a non-mythological and demonstrably secure reconstruction of all mathematics was interrupted over trivia, and it is time that these trivia were placed once more in proper perspective. Again, I do not propose to *change* the logical superstructure—only to understand it more clearly, by making explicit in a formal way features that have belonged to our intuitive logic all along.

It may be, of course, that there are more fundamental fallacies hidden in the usual logical superstructure, untouched by the present approach. It is still uncomfortable, for example, that another aspect of the Gödelian staircase remains untouched by the present account. For, while we can guarantee *reliability* of formalized mathematics in ways that Gödel is commonly held to have prevented, we remain subject to the Gödelian restraint with respect to the *adequacy* of formal to intuitive mathematics. And this is, in the end, the more important restraint. To be sure, it is up for grabs as to what counts as intuitive mathematics, and what is merely mathematical mythology. In arithmetic, for example, there is certainly a sense in which, if we get the $+$ and $\times$ tables right, we have *done* intuitive mathematics, and further nonsense involving ineffective quantification over the completed infinite domain of natural numbers is to be reserved for tellers of tall tales. But creative mathematics is filled with such tall tales; while it is nice, accordingly, to be rid of the levels of increasing doubt regarding the reliability of the tales that we have managed to spin, there will remain not only graded degrees of reliability for particular stories but also graded degrees of adequacy to the whole story. While it is no small feat, I claim, to have circumvented Gödel

on a point of interpretation of his results that has led to unwarranted pessimism on foundational questions, giving rise to a new foundational program of considerable interest in itself (and which one hopes will serve to check the rampant platonistic fideism of the day), only one of Gödel's theorems can be here repealed. That's a pity, but it will have to do.

III

The reader, I presume, is no longer in any doubt in what sense we are going to prove our preferred formal arithmetic $\mathsf{R}^\sharp$ consistent. To begin with, we shall show that it contains an unprovable formula, making it consistent in sense (1). The unprovable formula itself will be Hilbert's formula $0 = 1$, making $\mathsf{R}^\sharp$ consistent in sense (2). And we shall indicate how, for any distinct I, J under (3), or T, U under (4), we can extend our consistency proof under (2) to refute $I = J$ or $T = U$, offering a certain guarantee of consistency under these points also. There our guarantees stop, for now.

The reader who picked up this paper, and who read its opening sections, in the hope that something truly earth-shaking was about to be demonstrated is to be pardoned, perhaps, a certain let-down at this point. For he is certainly not to be blamed if he took consistency, to begin with, in sense (6). If so, he may be disappointed to discover that not only is effective consistency denied us in this sense in $\mathsf{R}^\sharp$, but that exactly the same barriers exist in $\mathsf{R}^\sharp$ as in $\mathsf{P}^\sharp$ against the proof of such consistency.

Later, I shall outline what is needed to prove consistency in sense (6) for $\mathsf{R}^\sharp$, noting by the way that this will also produce a new proof of consistency for $\mathsf{P}^\sharp$. Meanwhile, however, the reader is urged not to look lightly on the present results. They are exactly what, in his initial ignorance, he may have supposed them to be. For the lasting deficiency of $\mathsf{P}^\sharp$ is, as noted, that it cannot *distinguish* (6)-consistency from any other sort. Indeed, the original form of Gödel's second theorem refers not to (6)-consistency but to (1)-consistency as what cannot be established for $\mathsf{P}^\sharp$ in $\mathsf{P}^\sharp$ itself. And it is this kind of thing that leads to continuing confusion between genuine doubts about the reliability of intuitive arithmetic and mere persisting puzzles.

We now present our first-order Peano arithmetic $\mathsf{R}^\sharp$. We get $\mathsf{R}^\sharp$ by adding to the Peano postulates, together with the usual laws governing $+$, $\times$, $=$, to the first-order logic RQ developed in [7]. RQ is a relevant logic. This makes $\mathsf{R}^\sharp$, I suppose, a relevant arithmetic. But, frankly, I have little time for these

pejorative or honorific adjectives, like 'relevant,' for it is by no means my intention here to formulate a *new kind* of arithmetic, or one which presupposes a *successful formal analysis* of the notion of relevance. To the contrary, while I think that researchers in relevant logics—in particular, Anderson and Belnap—have succeeded in calling attention to features in the logical landscape that had hitherto been overlooked, and which answer in some happy respects to our pre-theoretic intuitions involving relevance, I do not hold with the view (which has apparently come to prevail among the "in" group) that these are "relevance logics," in the sense of providing a "logic of relevance." On that point, I don't particularly think that we need a logic of relevance, and I certainly don't think we have one.

What we do need, and what have been developed, are new ways of looking at logic in general, and of the particular theories that we use logic to build, that eliminate some of the irrelevant anomalies (like the *ex falso quodlibet*) which have hitherto plagued the subject, and which have been allowed to persist, like the office of the Governor-General, in the erroneous belief that they do no harm. (Hmm—Russell should have said that, and it's a pity that he didn't.) Accordingly, I repeat, what is intended here is nothing new, and nothing that belongs to some esoteric branch of philosophical logic with its few dozen initiates. Rather, what is intended is what should have been seen all along; namely, that it is a particular view of logic—indeed, a particular mythology of logic—that has produced the worst of the Gödel anomalies, and that these anomalies cannot survive even rudimentary thought-taking as to what we are really about when we are seeking to formalize arithmetic and other branches of mathematics.

So, near enough, it is just the old arithmetic which is being formalized here. Or, frankly, which we are attempting to formalize here, since it is always possible that the present formalization of $\mathsf{R}^\sharp$ itself runs into anomalies of a sort not here foreseen. Nor need we be wedded to $\mathsf{R}^\sharp$, or to its underlying logic RQ, in particular. There are both logical and arithmetical options, closed in one way here, that others might wish to close differently, with quite similar results. While, both philosophically and technically, I think that there are significant advantages in working with $\mathsf{R}^\sharp$ as formulated here—for one thing, RQ is understood rather more thoroughly than other logics in its neighbourhood, while it incorporates certain classical principles of strength that I think are right in conception and convenient for formulating concrete mathematical theories—I do not wish to commend the familiar mistake of looking for more precision in a uniquely preferred formalization than our

indistinct and confused underlying intuitions will warrant.

The system $\mathsf{R}^\sharp$ is formulated straightforwardly. It is based on a first-order language $\mathscr{L}^\sharp$ built up as usual from the countably infinite sequence of variables x, y, z, x_0, x_1, etc., the individual constant 0, the unary operation $'$, the binary operations $+, \times$, the binary predicate $=$, the unary connective $\sim$, the binary connectives $\wedge, \rightarrow$, and the universal quantifiers $\forall x, \forall y$, etc.

I prefer to view $\mathsf{R}^\sharp$, and other formal systems, as *abstract* in something approaching the sense of Curry's [8]. This means that the constituents of $\mathscr{L}^\sharp$ are not to be taken as pieces of language in the ordinary sense (whatever that mixed-up ordinary sense may be), but as abstract objects and functions. This frees the language we are *using* (Curry's U-language) for its ordinary referential sense, to which the conventions of standard (mathematical) English will apply without further comment. But the reader who is stuck in a fallacious nominalism may easily translate everything into his own tongue, as he is hereby encouraged to do.

We next build up the usual grammatical categories in $\mathscr{L}^\sharp$, as inductive classes. From the individual constant 0 and the successor operation $'$ we build up the *natural numbers* on the following expected definitions.

- $1 =_{\mathrm{DF}} 0'$
- $2 =_{\mathrm{DF}} 1'$, etc.

We shall let $\mathbf{N}$ be the class of all natural numbers, and we shall let 'I,' 'J,' 'i,' 'j,' etc., serve (usually) as intuitive variables ranging over $\mathbf{N}$. (As noted, since our system is abstract, we do not distinguish *numbers* from *numerals*; put otherwise, each numeral *is* the number which, on other accounts, it would be held to *designate*.)

We next build up the class $\mathbf{NT}$ of *number terms* by closing $\mathbf{N}$ under $+$, $\times$, $'$. Examples of number terms are $2 + 2$, $3 \times (4 + 6')'$, and so forth. And we build up the class of $\mathbf{T}$ of *terms* by closing the union of $\mathbf{NT}$ and the set of formal variables under $+$, $\times$, $'$. Thus, e.g., $x + 3$, $y \times (4 + x_7)$ are terms but not number terms.

Let T, U be terms. Then, and only then, $T = U$ is a member of the class $\mathbf{EQ}$ of *equations*. If, in particular, T, U are numbers, $T = U$ is a member of the class $\mathbf{NEQ}$ of *number equations*; if T, U are number terms, $T = U$ is a member of the class $\mathbf{NEQQ}$ of *numerical equations*. We may also refer to equations as *atomic equations* or as *polynomial equations* or as *atomic formulas*.

Equations are member of the class **FORM** of *formulas*; moreover, **FORM** is closed under the primitive connectives $\rightarrow$, $\wedge$, $\sim$ and the quantifiers $\forall x$, etc. (Set-theoretically, $\sim$ and the quantifiers are of course exactly 1-place operations on **FORM**, while $\rightarrow$, $\wedge$ are 2-place operations.) A formula in which no variables occur will be called a *numerical formula*. A formula in which no variables occur *free* will be called a *sentence*, and we let **SENT** be the class of all sentences.

In considering $\mathscr{L}^{\sharp}$, we shall be principally concerned with its set **FORM** of formulas, and we shall often write that $A \in \mathscr{L}^{\sharp}$, or the like, instead of that $A \in$ **FORM**. And we shall also be using 'A,' 'B,' etc., as intuitive variables ranging over **FORM**, 'X,' 'Y,' etc., as intuitive variables ranging over the class **VAR** of formal variables, and 'T,' 'U,' etc., as intuitive variables ranging over the class **T** of terms.

In characterizing $\mathscr{L}^{\sharp}$, we saved for the last some nasty matters that we hitherto trusted the reader to take for granted: e.g., that the various grammatical categories of $\mathscr{L}^{\sharp}$ are built up recursively, without having made explicit what, in our abstract perspective, that might involve. E.g., by definition $0''$ *is* the number 2. But is $1 + 1$ the number 2? Scott may be taken to have suggested, in [8], that it might not hurt too much if it were. Here, however, we follow tradition to the extent of distinguishing, as formal objects, $1 + 1$ from 2. And what this means, essentially, is that each of our primitive operations on terms, like $+$, shall have an *atom property*, which insures that each of our formal terms shall be built up from 0 and the formal variables, and a *generalized ordered pair property*, which insures, say, that if $T + U = V + W$, then $T = V$ and $U = W$; it is unnecessary, surely, to be more explicit here, except to note that similar remarks apply to our primitive predicate $=$ and our primitive connectives and quantifiers. (In short, a formal language, as here conceived, is something like a universally free algebra, or a generalized arithmetic in Kleene's sense, save as syntactical restrictions leave our primitive operations undefined for certain arguments; e.g., $(0 = 0) = 0$, which would count elsewhere as ill-formed, fails here to be a formal object because its first argument is inadmissible as an argument for $=$, conceived as a function from pairs of terms to formulas.)

Having made those points, we can now make some trivial ones. First, $\mathscr{L}^{\sharp}$ does not admit anything like an ill-formed formula, nor has it any use for parentheses. (Why should it?) We, however, will need parentheses, and other grouping conventions, as part of our informal English notation for discussing $\mathscr{L}^{\sharp}$. In particular, we adopt the following conventions. First, unary

operations and connectives shall have minimal scope (consistent with the formal grammar). Second, we shall sometimes drop $\times$ for simple juxtaposition on terms; at any rate, $\times$ will bind more strongly than $+$; both conventions are arithmetically ordinary. Third, binary connectives (including those to be defined) shall be ranked $\wedge$, $\circ$, $\vee$, $\supset$, $\supset$, $\rightarrow$, $\equiv$, $\equiv$, $\leftrightarrow$ in order of *increasing* scope; i.e., $\wedge$ binds more tightly than $\vee$, etc. On what is a minor added convention, we shall, however, use simple juxtaposition between formulas sometimes in place of $\circ$, in which case an invisible $\circ$, unlike a visible, shall bind more tightly than $\wedge$ and, *a fortiori*, more tightly than other connectives in the above list. Fourth, equal operations and connectives shall be associated to the left. Fifth, we use dots (sparingly) as parentheses in accordance with the conventions of [8], which are also (essentially) the familiar conventions of [9] and [7]. If all that fails, or just to be kind to the reader, we use parentheses; in fact, I trust that common sense alone, without reference to the above explicit conventions, will already suffice to resolve almost all possible ambiguities. And common sense also will be relied upon to dissolve conventions too tedious, or too dull, to make explicit; e.g., I trust the reader to resolve $X = Y + Z$ in the direction of sense rather than of nonsense. And, especially, I trust the reader to resolve ambiguities that might otherwise turn on confusion of bound variables; e.g., when $T \leq U$ is schematically defined as $\exists X(T + X = U)$, it is too tiresome to mention that X should be the first variable which is free in neither T nor U, and similarly in other cases.

One sore point of possible ambiguity does require explicit mention, however. Normally, I shall use AX for a formula A in which X may occur free, and AT, AY, etc. for the result of proper substitution of a term T, variable Y, etc., for all free occurrences of free X in A. The persisting convention, which essentially follows Kleene's practice in [5], is that AX shall be chosen first; i.e., the convention is tied visually, in general, to the particular sign 'X.' (It need not be tied to the particular sign 'A,' while exceptions, if clear in context, may be made to the general convention.) So far as proper substitution is concerned, a substitution is proper, following the usage of [10], if it does not involve confusion of bound variables. And, in order that AT may be always defined when AX is given, we follow [8] in first rewriting the bound variables of AX on a definite plan (e.g., the plan of [8]) and then substituting T for free X when necessary to avoid confusion. E.g., let AX be $\forall y(x + y = z)$, and let X be x, Y be y, Z be z. Then AZ is, straightforwardly, $\forall y(z + y = z)$. But AY, to avoid confusion of bound variables, must be something like $\forall x_0(y + x_0 = z)$, and similarly in other cases. (It is

always burdensome to decide which conventions to employ in these matters; the perspicuous ones must be carefully spelled out, lest inaccuracy result, while the accurate ones induce notationitis. I have opted here for perspicuity, supplying as much accuracy as is necessary; but the careful reader must stick to the general convention.)

Next we enter some (schemes of) definition, for the most part familiar.

$D\vee$. $A \vee B =_{\text{DF}} \sim(\sim A \wedge \sim B)$

$D\exists$. $\exists X A =_{\text{DF}} \sim\forall X \sim A$

$D\supset$. $A \supset B =_{\text{DF}} \sim A \vee B$

$D\equiv$. $A \equiv B =_{\text{DF}} (A \supset B) \wedge (B \supset A)$

$D\leftrightarrow$. $A \leftrightarrow B =_{\text{DF}} (A \rightarrow B) \wedge (B \rightarrow A)$

$D\circ$. $A \circ B =_{\text{DF}} \sim(A \rightarrow \sim B)$

$D\top$. $\top =_{\text{DF}} 0 = 0$

$D\perp$. $\perp =_{\text{DF}} \sim\top \ (=_{\text{DF}} 0 \neq 0)$

$D\leq$. $T \leq U =_{\text{DF}} \exists X(T + X = U)$

$D<$. $T < U =_{\text{DF}} \exists X(T + X' = U)$

$D\geq$. $T \geq U =_{\text{DF}} U \leq T$

$D>$. $T > U =_{\text{DF}} U < T$

$D\neq$. $T \neq U =_{\text{DF}} \sim(T = U)$ ($D \nless, D \nleq, D \ngtr, D \ngeqslant$ are similar)

$D==$. $T = U = V =_{\text{DF}} T = U \wedge U = V$ ($D <<, D \leq\leq$, etc., are similar)

$D\leq=$. $T \leq U = V =_{\text{DF}} T \leq U \wedge U = V$ ($D <=, D <\leq, D \leq=$, etc., are similar)

$D\text{F}$. $\text{F} =_{\text{DF}} 0 = 1$

$D\text{T}$. $\text{T} =_{\text{DF}} \sim\text{F} \ (=_{\text{DF}} 0 \neq 1)$

We shall add more definitions later. Note, thus far, that $\rightarrow$, $\circ$, $\leftrightarrow$ are respectively *relevant* implication, fusion, and equivalence, whereas other connectives are classical. $\circ$, in particular, is quite an interesting connective, whose central role for relevant logics and their algebras has been developed in quite a number of papers; e.g., in [11]. Otherwise, our definitions are the standard ones, up to the exercise of certain options (e.g., in defining $\leq$, given $<$). Some ($D\neq$, $D==$, etc.) are purely for convenience, which we may add to without being so explicit about it. Otherwise, the main point of interest

is that we have isolated 2 distinct constant truths, $\top$ and T, and 2 paired constant falsehoods, $\bot$ and F. Systematically, the pair $\top, \bot$ will prove more interesting for our analysis of $\mathsf{R}^\sharp$; $\top$ in particular will play the role of the constant truth **t** which has been of interest through the whole development of relevant logics (cf. [7]). But F, though it is something of an absurd curiosity, has often been taken to express *the proposition* whose underivability is sought to guarantee the reliability of formal arithmetic (cf. (2)-consistency above). Accordingly, we pick F in particular as the number equation that we shall show underivable in $\mathsf{R}^\sharp$.

IV

Our language $\mathscr{L}^\sharp$ having been set up, we now characterize our first-order Peano arithmetic $\mathsf{R}^\sharp$ in familiar axiomatic fashion. Following [5], I divide the deductive principles into three groups, depending on whether they are sentential (R-level), quantificational (proper RQ-level), or number-theoretic (proper $\mathsf{R}^\sharp$-level).

As sentential axiom schemes, the following suffice for R.

$A1.\quad A \to B\ \to.\ B \to C\ \to.\ A \to C$

$A2.\quad A \to.\ A \to B\ \to B$

$A3.\quad A \wedge B \to A$

$A4.\quad A \wedge B \to B$

$A5.\quad (A \to B) \wedge (A \to C)\ \to.\ A \to\ B \wedge C$

$A6.\quad A \wedge (B \vee C) \to A \wedge B\ \vee\ A \wedge C$

$A7.\quad {\sim}{\sim}A \to A$

$A8.\quad A \to {\sim}B\ \to.\ B \to {\sim}A$

$A9.\quad A \to {\sim}A\ \to\ {\sim}A$

I formulate rules, as before, by using $\&, \Rightarrow$ in a self-explanatory metalinguistic way. Rules at the sentential level are

$\to E.\quad A \to B\ \&\ A \Rightarrow B$

$\wedge I.\quad\ A\ \&\ B \Rightarrow A \wedge B$

The quantificational, RQ-level of $\mathsf{R}^\sharp$ is attained economically by adding the following additional axioms scheme and rule.

A10. $\forall X AX \to AT$

C∀I. $A \to . B \vee C \Rightarrow A \to . \forall X B \vee C$

where it is required, for $C\forall I$, that X occur free in neither A nor C.

     As specific number-theoretic axioms, we take

A11. $x = y \to x' = y'$

A12. $x = y \to . x = z \to y = z$

A13. $x' = y' \to x = y$

A14. $x' \neq 0$

A15. $x + 0 = x$

A16. $x + y' = (x + y)'$

A17. $x \times 0 = 0$

A18. $x \times y' = x \times y + x$

To these we add a *rule* of mathematical induction, namely

RMI. $A0 \;\&\; AX \to AX' \Rightarrow AX$

Note that we get an equivalent formulation of $\mathsf{R}^\sharp$ if we drop RMI and replace it with the specific axiom scheme

AMI. $A0 \wedge AX \to AX' \;\to\; AX$

Normally, we shall be indifferent between these two formulations. But, where specificity is desired, let $\mathsf{R}^\sharp$ be the original system with RMI, and $\mathsf{R}^\sharp_A$ be the alternate system with AMI.

     To compare and contrast $\mathsf{R}^\sharp$ with alternative first-order Peano arithmetics, we consider the following further principles.

P0. $\bot \to A$

J0. $A \to . B \to A$

$J\forall I.\quad A \to B \Rightarrow A \to \forall X B$

$RM0.\quad A \to. A \to A$

$RM3.\quad A \vee. A \to B$

We get classical first order Peano arithmetic $\mathsf{P}^\sharp$ (e.g., as developed in [5]) by adding $P0$, or alternatively $J0$, as an additional axiom scheme to the axioms and rules for $\mathsf{R}^\sharp$ above. (Thus the classical system is a straightforward extension of $\mathsf{R}^\sharp$.) We get intuitionist arithmetic $\mathsf{J}^\sharp$ in a more complicated way. From the given basis for $\mathsf{R}^\sharp$, we first subtract $A7$ (double negation) and replace $C\forall I$ with the intuitionistically acceptable rule $J\forall I$. Then we add $P0$ (*and* $J0$) as before; moreover, since $D\vee$, $D\exists$ do not apply intuitionistically, we must take $\vee$ and the $\exists X$ as additional primitives, adopting axioms dual to $A3$, $A4$, $A5$, $A10$ and a rule dual to $J\forall I$ to govern them. (We also need to assure the $\mathsf{J}^\sharp$ provability of $A \to. B \to C :\to. A \to B \to. A \to C$, proved in $\mathsf{R}^\sharp$ using $A7$, but I shall let the reader worry about that.)

More interesting than covering such old ground is the question of formulating first-order arithmetics that are based on other relevant logics and semi-relevant logics. In particular, I propose a system $\mathsf{RM}^\sharp$, got by adding the characteristic sentential axiom $RM0$ above of the Dunn-McCall system RM to the above basis for $\mathsf{R}^\sharp$; and a further system $\mathsf{RM3}^\sharp$, got by adding $RM3$ also to $\mathsf{RM}^\sharp$. $\mathsf{RM}^\sharp$, in particular, may prove to be an interesting formal arithmetic, for reasons related to those on which RM itself is found interesting in [7]. $\mathsf{RM3}^\sharp$, on the other hand, must be viewed as a curiosity; its interest lies in the fact that, sententially, it is a 3-valued logic, and is accordingly easy to deal with (despite the deficiencies pointed out in this and other finitizations of RM, at the sentential level, by Dunn in [14]). Moreover, the basic consistency argument here will apply to $\mathsf{RM3}^\sharp$ also, despite the considerable strengthening that this involves of $\mathsf{R}^\sharp$.

It would also be interesting—though I just haven't thought about it, much—to weaken $\mathsf{R}^\sharp$. One line of weakening is quite obvious. Instead of grafting the arithmetical axioms onto RQ, what about grafting them onto the system EQ of [7], or even onto a system based on the minimal relevant logic B of [15]? On past experience, any interesting results that one gets using the strong and natural logic R hold, *mutatis mutandis*, for its weaker cousins; it would be quite surprising if a system $\mathsf{E}^\sharp$, for example, turned out significantly weaker, in its arithmetical part, than $\mathsf{R}^\sharp$. But I simply leave such questions open, for now, noting that considerable formal drudgery is

involved in making contributions to them.

There are, however, other lines on which $\mathsf{R}^\sharp$ can be weakened, along with other formal arithmetics. Of particular interest are the axioms $A13$, $A14$, and the rule RMI, each of which corresponds rather directly to one of the Peano postulates, and each of which raises its own special problems. Let us look at these problems (or some of them, anyway) in turn.

$A13$ "says" that no 2 numbers have the same successor. More accurately, it says that the successor function, on natural numbers, is 1–1. The trouble is that this is not the sort of thing that we should *have to say*; if we think of the natural numbers as built up from 0 by applying the successor operation, it would seem that we could take for granted that each result of applying the operation produces a new unique number, unless a *specific assumption* is made to the contrary. And this feeling that there is something out of order about $A13$ is confirmed by its specific form; it seems a little silly that we should demonstrate $2 + 1 = 3$ by first demonstrating that $2 + 1 + 1 = 4$ and then subtracting. To be sure, we should certainly expect $A13$ to hold as a *material* implication—as a *fact*, so to speak, about natural numbers. But if we think of occurrences of $\rightarrow$ in *arithmetical postulates* as reflecting, at least in part, an order of *generation* of natural numbers in the inferences that, strictly speaking, we can in the first instance *make about* natural numbers, then $A13$ deserves further thought. In some form, we shall not wish to do without it. But we note that, in the present context, $A13$ has a certain collapsing effect on the structure of arithmetic propositions, as viewed from an $\mathsf{R}^\sharp$ perspective. This collapsing effect can be defended, on the grounds that, on structural considerations, arithmetic propositions are tightly wound together anyway. And, technically, it is to be welcomed, since it simplifies a number of formal arguments. But alternative forms of $A13$, or perhaps a perspective on which it is no longer primitive, should also be explored.

$A14$ "says" that 0 is not the successor of any number. Given that we can put $A13$ in a positive form, it is our only *negative* axiom. As such, given the remarks about $\sim$ above, the reader will readily grasp that I already consider $A14$ suspicious. This is not to say that I consider rejecting it; indeed, if one thinks of negative axioms as embodying *counteraxioms*, in the sense of [8], then, given $A13$, the *counteraxiom scheme* $T' = 0$ is not unreasonably the only such scheme that arithmetic has need of.

However, it requires very little thought to see that the $\mathsf{P}^\sharp$ or $\mathsf{J}^\sharp$ form of $A14$ endows our suggested counteraxioms with surprising properties. For any theorem $\sim A$ of these systems can equivalently be expressed as $A \rightarrow \perp$.

Given the transitivity of provable implication and the *ex falso quodlibet* in the form $P0$, this means that $\sim A$ is a theorem iff, for every formula B, we have $A \to B$ as a theorem. In short, $A14$, as a classical or intuitionist axiom, does not simply deny, say, that 4 is 0; we would all join, after all, in this simple denial. Rather, $A14$ affirms that, if 4 is 0, or if X' is equated with 0 for some variable X, then all hell breaks loose, when $A14$ is taken as a $\mathsf{P}^\sharp$ or $\mathsf{J}^\sharp$ axiom.

In short, what an axiom asserts cannot be separated from the systematic context in which the asserting is done. And note that $\mathsf{R}^\sharp$, while it buys the Peano postulate, does not buy the *logical context* conferred upon that postulate by $\mathsf{P}^\sharp$. I.e., $\mathsf{R}^\sharp$ agrees that it is *false* that 0 is the successor of some number, but not that it is *absurd*. And only, in fact, that kind of theory of negation which holds that *all falsehoods* are absurd could provide *general grounds* for interpreting the Peano postulate $A14$ in the $\mathsf{P}^\sharp$ way.

The last point is not to be pushed too far. For we might have *specific grounds* (e.g., cancellation laws) for preferring the stronger $\mathsf{J}^\sharp$ and $\mathsf{P}^\sharp$ interpretation of $A14$. Here, these specific grounds are not accepted. We accept, indeed $X' = 0 \supset 1 = 0$ as an *arithmetical fact*, signalled by the material $\supset$. But we do not accept it as an *arithmetical law*, that might be signalled by an $\to$.

Let us get clear on this point. I am not concerned with the *practical* inferences which a number theorist will permit himself. So far as practical inferences are concerned, he will no doubt assume that he is working with facts in a numerical universe of discourse, and reason accordingly. Rather, we are concerned here with the *logical framework* from which we build up the natural numbers. (It is not ruled out that there are alternative frameworks, good for different purposes.) We are told that there are certain principles—enunciated by Dedekind and made famous by Peano—on which the natural numbers turn out to be, intuitively and near enough, what we first held them to be. Now certainly this is not the case on first principles alone; unless there are deductive connections between certain number-theoretic statements and others, we shall never get beyond first principles. But we cannot take these deductive connections themselves as part of the discipline that we are formalizing, lest the entire enterprise of providing that discipline with a foundation become viciously circular. (More accurately, we can only take them so in part, trying to find a sense, say, in which RMI becomes properly arithmetical, while $\to E$, perhaps, remains firmly logical.)

We take the study of deductive connections, by and large, to belong to

logic. But a funny thing happens when we try to find hard and fast grounds
to prefer some purported general deductive connections to others. For these
grounds tend to refer back to the theories themselves that we use our logic to
build (some of them reified, in traditional Leibnizian mythology, as "possible
worlds"). Now it stands to reason that all the theories constructed on some
particular logical framework will respect that logic (the "soundness" half of
traditional completeness proofs), and even perhaps that the sum of inferential
principles that are respected by all theories (or perhaps all theories that are
in some further way preferred) constructed out of a given logical framework
are in fact correct inferential principles according to that given logic (the
"completeness" half of completeness proofs). But this just isn't a hard and
fast justification of the logic itself; far less is it a justification for our choice
of such-and-such a logic for building theories. Rather, this kind of (familiar)
justification becomes outright circular; if we impose a certain logic upon
arithmetic, we cannot justify this imposition on the ground that arithmetic
(among other things) respects this logic.

That our choice among alternative proferred formal logics is severely in-
determinate is a proposition that I no longer have any serious grounds to
doubt. All rational ways out of the indeterminacy are viciously circular,
while it is Quinaciously silly to substitute sociological grounds (e.g., certain
systems are "entrenched") for rational ones. But we can still find evidence
for and against particular formal points of view; e.g., the kind of mystical
insecurity enshrined in customary interpretation of Gödel's second theorem
must certainly count as evidence against the thesis that the kinds of deduc-
tive connections on which it depends constitute a rational approach to the
formalization of arithmetic.

We return to the point at issue. We have detected an illicit appeal to logic
in justifying the $\mathsf{P}^\sharp$ theoremhood of, say, $4 = 0 \rightarrow 1 = 0$, as a consequence of
$A14$. However, we have noted the objection that this is not a logical but an
arithmetical inference. After all, assuming that $4 \times 1 = 4 = 0 = 4 \times 0$, we
can just cancel to get $1 = 0$. For it is surely the case, for positive I, that if
$I \times J = I \times K$, then $J = K$.

To this one must respond, more forcefully that in the discussion of $A13$,
that laws and facts must be kept straight. In the natural order of things, it
makes sense to hold that $I' \times J = I' \times K$ *because* $J = K$, not vice versa. I.e.,
in the generation of the natural numbers, we expect to get to $J = K$ *before*
we get to $I' \times J = I' \times K$. If the road to B leads through A, we expect of
someone who shows up in B that he may have been at A; it is unnecessary

to walk backwards from B to A to prove the point.

So, I think, $T' = 0 \to 1 = 0$ is not even arithmetically justified, as a law. And it is quite different from the indubitable arithmetical fact noted above, that $T' = 0$ *materially* implies $1 = 0$, which is an immediate and trivial consequence of $A14$. On the sum total of arithmetical facts, it would be surprising if $R^\sharp$ did not agree with $P^\sharp$; the question, rather, is one of providing an acceptable and reliable logical framework for such facts.

Intuitive arithmetic, at least at its elementary levels, we took to be pre-logically given. When, however, we come to reconstruct that arithmetic from first principles, and to formalize it, we need to attend to the way in which our arithmetic is given, after analysis. For the latter purpose, we assume constructively generated natural numbers and certain fundamental laws concerning them. As we spin out the deductive consequences of these laws, we find ourselves with an ever widening realm of arithmetical facts. We create a mythology concerning these facts—indeed, as facts, they are only mythological facts—and we allow ourselves more and more to reason and to talk about numbers as we do about chairs and tables.

There is always danger in this procedure, quite apart from formal consistency proofs and the like. The danger is almost a theological one; namely, that we shall take too seriously what we have ourselves fashioned, and begin to ascribe to the creatures of our own imagination an independence and a reality which is not to be conferred upon them. Here, paradoxical results are positively salutary; like the strange things that go on in our dreams, they help us to preserve the distinction between what we fancy and what there is. So, though the law-fact distinction on which we are depending is at best heuristic anyway, we must beware of thinking of our generating laws as merely *describing* arithmetical facts; as the story goes, they must be taken moreover to *constitute* the facts. Sometimes, to be sure, this kind of constructive attitude leads to proposals to diminish the class of mathematical facts, eschewing those that cannot be "constructively" proved. However, such proposals would be out of place here; classical mathematics remains, so far as we are concerned, the only game in town, while assuring its reliability by cutting down on its supply of facts seems to me no more hopeful than seeking to assure the reliability of our dreams by dropping certain characters from them. Our dream is the usual dream; if we dream it a little differently, so be it.

A little more light, incidentally, is now thrown on how we intend to interpret the relevant implication $\to$ of $\mathsf{R}^\sharp$. We intend to interpret it, not unnat-

urally, as 'implies.' We do not intend to interpret it as 'relevantly implies,' which would be redundant. We do not intend to interpret it as 'logically entails'; for, having added proper arithmetical axioms, we have theorems $A \rightarrow B$ which belong not to logic but to formal arithmetic proper.

But, in so interpreting $\rightarrow$, I shall not supply any "formal semantics" for this evidently non-truth-functional connective. For that matter, no formal semantics is supplied, save in our passing references to the standard model, even for presumptively truth-functional connectives and quantifiers, like $\wedge$, $\sim$, $\supset$. To be sure, a formal semantics, of a familiar extensional sort, can be supplied for relevant theories. But supplying it, at this stage, would be to miss the point; we are doing honest arithmetic now, which has suffered enough from extensional reduction, and the only semantics intended is the intuitive one; not only shall $\rightarrow$ be taken as 'implies,' but $\sim$ shall be taken as 'not,' $\forall$ shall be taken as 'for all (natural numbers),' and so forth. Since even for $\mathsf{P}^{\sharp}$ this intuitive semantics is at best rough and ready (we already saw in I that $\sim$ means 'not' only by mythological courtesy, which $\mathsf{P}^{\sharp}$ cannot itself systematically support), the *best* that we can do for $\mathsf{R}^{\sharp}$ (or any other serious formal system) is to relate it loosely to intuitive sense-making; for it is intuitive sense-making, after all, that we are eventually about, with or without the useful but presently overrated mediation of "formal semantics."

However, relating $\rightarrow$ in particular to intuitive sense-making requires some further comments; in English, 'implies' is certainly less regimented than 'and,' and perhaps 'not,' whence all the fuss about finding a formal counterpart for it. In fact, among presently overrated pursuits, the utility of natural usage as a guide to logical analysis must also be accorded high place. It is a guide, for sure, to which we should look for clues as to how to construct our formal systems; but formal systems, in general, serve as crude and inadequate models of natural language, while natural language only vaguely and partially reveals the connections of ideas which it is the business of formal systems to lay bare.

This, to be sure, is just logical common sense; in fact, one of the difficulties involved in *rejecting* the reading 'implies,' 'if-then,' or the like for $\supset$ (without some adverbial qualification like 'materially implies') is that one has to explain patiently that one understands all the common sense points that are urged *in defence* of this odd reading of $\supset$, without understanding any better why people are so attached to it. For English usage enters into the rejection only heuristically; 'materially implies,' I suppose, is now part of English, and one understands it fairly well; and one becomes accustomed,

though he may wince a bit, to speakers in whose peculiar idiolect 'materially' gets left out.

The central point, rather, is not the loss of an adverb but the loss of a notion: namely, that some connection of ideas is required for authentic inference. Relative to this point, incidentally, far too much stress has been placed on authentic *logical* inference, as though the only inferential connections that we care about belong to deductive logic. Many of them do, of course, and we sought above to give them their due. But we were also concerned above with authentic *arithmetical* inference, signalled by our proper arithmetical axioms (especially those of the form $A \rightarrow B$). (Such matters became especially confused, in years gone by, by the lumping together of logical and mathematical truths as jointly analytic, suggesting that mathematical inference (in a properly mathematical form) is just a cousin-under-the-skin of logical inference anyway. The result, practically, was that mathematical implication simply got left out of *everybody's* formal considerations (save those of the intuitionists, who had, however, other axes to grind). For some people fixed up their vocabulary so that they asserted that A implies that B when they took $A \supset B$ to be *true*; others, when they took $A \supset B$ to be *logically valid*. But the former is *too weak* a relation to stand in for mathematical implication, in any generally usable sense; and the latter, whatever its other defects, is in certain respects *too strong*; we do not look to *logic alone* to separate the authentic mathematical inferences from the unauthentic ones.)

There is not much that I have to say about authentic mathematical inference. But, when we are building up mathematics from first principles, it seemed to me above that there is a *direction* of inference which is quite independent of truth-values, but which corresponds rather to an *intuitive order* in the way that theorems are built up from axioms. This is not to deny that we may allow ourselves *practical inferences* that do not respect this order; moreover, since there is an *interplay* between guiding intuition and formal derivation—the latter directs the former at least as often as it is directed thereby—formal treatments may differ in seeking such a direction. Nor have I sought it to the exclusion of other considerations here, as the discussion of A13, A14 makes clear.

This point can perhaps be made clearer with reference to Gentzen-style consecution calculi, and particularly with respect to their operational rules. As these rules work, they follow the order of generation of formal objects by building up theorems containing more complex constituents from theorems whose are simpler. And it is a certain measure of a system's normality that

it may be Gentzenized in this fashion, without operational rules (e.g., cut) which violate this principle that the more simple shall precede the more complex.

The example cannot be pushed too far, since an order of generation of formal objects is not necessarily reflected in a natural fashion in an order of generation of theorems into which these objects enter as constituents. And we have, of course, further problems when we turn to subjects like analysis, where the formal objects in which we are primarily interested (i.e., the real numbers) cannot be constructively generated at all. But the example at least encourages us to think, while we are formulating our formal systems, about the relation between *how we formulate them* and Gödel's second theorem. And when we congeal at least some of our rules into the indicative, as implications $A \rightarrow B$, rather than as mere statements of purported fact $A \supset B$, it is likewise presumptively not indifferent which of these implications we take as first principles for a discipline in question. Nor is an implication $A \rightarrow B$ to be lightly reversed to $B \rightarrow A$, even when the presumptive truth-conditions are "right."

Besides $A13$, $A14$, the other primitive number-theoretic principle of $\mathsf{R}^\sharp$ reflecting a Peano postulate is of course the mathematical induction rule RMI. The idea, of course, is that we wish to assure that every property A of natural numbers which holds of 0 and which is passed on from every number to its successor holds of all natural numbers; because of the countable vocabulary of $\mathscr{L}^\sharp$, and the uncountable number of number-theoretic properties, we cannot *say* this in $\mathscr{L}^\sharp$ (as the standard story goes); but we can say it of every property A which can be expressed in the vocabulary of $\mathscr{L}^\sharp$ by some open formula AX with sole free variable X. Hence we approximate, at least in $\mathsf{P}^\sharp$, Peano's inductive characterization of the class of natural numbers, while the deficiency in vocabulary causes us to fall short of that full characterization and hence to get an incomplete system, instead of the full system of arithmetic truth for which the genuine Peano postulates are categorical.

I am not sure whether I believe the standard story or not; until a better one comes along, I suppose that, mythologically at least, I do. But I am at least uneasy about the direction in which it is taken for granted; namely, that while there may be arithmetic properties expressed by no formula of $\mathscr{L}^\sharp$, at least every formula AX of $\mathscr{L}^\sharp$ does express some property P. The effect of this assumption is to build a naive comprehension axiom into $\mathsf{P}^\sharp$, carried above into $\mathsf{R}^\sharp$: namely, that every open formula determines some property.

This suggests that, in second-order relevant arithmetic, we may adopt

without restriction the comprehension scheme

(1) $\exists F \forall X (FX \leftrightarrow A)$

whenever F is not free in A. But there are two reasons to think (1) dubious. The first is just general, and probably unwarranted, suspicion. Unrestricted comprehension axioms cause trouble. To this it may be retorted, naturally, that there are quite elementary proofs of the consistency of second order $\mathsf{P}^\sharp$, which naturally will make second order $\mathsf{R}^\sharp$ also consistent. The reply, citing Gödel on our side, is that these proofs cannot be as elementary as one might wish. Indeed, if the fighting chance that $\mathsf{P}^\sharp$ itself is inconsistent turns out correct, one may take it as certain that (1), as it is reflected in RMI, will be the cause. While it would be pretty sneaky if this happened, and the cause of great consternation throughout mathematical logic, one way that any dream can end is with a rude awakening.

Second, there may be other reasons to impose restrictions on (1) in the particular framework of $\mathsf{R}^\sharp$. For, just as one expects some connection between A and B if $A \to B$ is to be justified as a theorem, so one might expect some connection between each individual ι and the proposition that ι has property P in order to count P as a genuine property of individuals. (I.e., considered as a reified function from individuals to propositions, P *depends upon its argument*, to borrow a phrase from Belnap.)

So put, the point has a metaphysical flavor with which some readers (and I, for that matter) will feel uncomfortable. But it can be given direct syntactical sense. In relevant second-order *logic*, formulated with an unrestricted comprehension scheme (1), some apparent fallacies of relevance become provable. We shall find that $\mathsf{R}^\sharp$ itself admits some first-order versions of these apparent fallacies—e.g., $X = Y \to Z = Z$—though it is an open question whether they shall be counted as fallacious in this context, or whether we should instead appeal to strong relations of mutual relevance between arithmetic propositions. One feels, frankly, a strong pull in each direction; one takes it that the natural numbers are tightly organized, structurally, permitting say, the following argument from $3 = 5$ to $7 = 7$. Assume $3 = 5$. By symmetry of identity, $5 = 3$. By transitivity, $5 = 5$. Adding 2 to each side $7 = 7$. End of valid argument. However, we were also suspicious above of seeking specific arithmetic grounds for principles, like the *ex falso quodlibet*, that are not supported here on general logical grounds.

In the case of RMI, as in the $A13$ case above, we have resolved open ques-

tions in the direction of strength and tradition. Partly, this is just a matter of formal efficiency; it is easiest to re-think formal proofs (which is a tedious enough job anyway) when one changes only what must be changed to get on with the task in an altered perspective. And it was, of course, enticing merely to change the *sentential base* of formal arithmetic to an R-valid form, leaving the properly quantificational and number-theoretic postulates essentially as one found them (e.g., in [5]). As Routley warns, this may be the wrong course; perhaps further insights that would result from a more thoroughgoing application of other investigations into relevant logics are being sacrificed. (If so, the reader at least has substantial hints about how to find some of them, and he is welcome to mention my name if he does so.) Indeed, perhaps, as Routley moreover suggests, the entire project, as undertaken here, is too conservative; perhaps classical logic undermines classical mathematics more thoroughly than I have undertaken here to believe except in noting open possibilities, and that the aim of producing an essentially classical mathematics from the standpoint of relevant logics is itself misguided. Only time, and further research, can resolve such issues; given that alternative philosophical views may underlie different decisions on the questions raised above, it is to be doubted that such research itself will resolve them fully. Meanwhile, let us continue with our "minimum mutilation" of the orthodox standpoint, endeavoring to show that it involves no real mutilation at all.

V

Now let us do some arithmetic. We shall busy ourselves, mainly, with noting that R$^\sharp$ has, roughly, the same substantive theorems as anybody's arithmetic. In doing so, we have [5] particularly in mind, and we note that Kleene's proofs of elementary theorems of arithmetic in P$^\sharp$ and J$^\sharp$ can be straightforwardly adapted to R$^\sharp$ in many cases. (Full proofs can be found in [18], of which this paper is essentially a summary. Also developed in [18] are new natural deduction techniques for relevant theories, which are rather helpful in the actual construction of derivations from axioms and rules above.)

Two cautions, however, must be added to the reader who, with [5] in hand, wishes to try out some of the formal theorems below as exercises. First, despite the claim that nobody really reasons from contradictions to arbitrary formulas, [5] in fact does so quite often, within derivations. Such fallacious arguments cannot be incorporated here. Second, a reasonable amount of at-

tention must be paid to the exact form both of the formulas to be proved and to the deductive machinery that is available to prove them. In applying RMI, for example, one must prove a premiss of the form $AX \to AX'$. It does *not* suffice to prove $AX \supset AX'$ in its place. (Thus, given the intentions behind the relevant $\to$, the conclusion of the induction step of an inductive argument must *depend upon* the hypothesis of the induction. *Prima facie*, this would seem to impose a nasty restriction on the applicability of mathematical induction. Practically, it does no such thing, in demonstrating elementary theorems of the sort listed below. For, as we shall see in part below, many arguments that, strictly speaking, are relevantly invalid can be replaced with valid arguments that yield the same conclusion. The underlying thought, in the RMI case, is that the property A which we are going to demonstrate for all natural numbers (in the sense that every number has A) must be passed on *according to law* for the inductive stage of the argument to go through. If it is not passed on in this law-respecting sense, we are back in mythology, not demonstration. Now it well may be that we cannot show this, in the sense that $AX \to AX'$ is a theorem of $\mathsf{R}^\sharp$, for a formula AX that we take to express A; i.e., A itself may not, in a direct sense, be passed on according to law. However, it often happens that there is another property B expressed by a formula BX, such that having B is a sufficient condition for having A. E.g., $AX \wedge CX$, for any formula CX in which at most X occurs free, would do for such a BX. In this case, we simply carry out the inductive argument on BX, by proving $B0$ and $BX \to BX'$, after which we may infer AX trivially from the conclusion of the induction. And this is, in fact, a very frequent case; actual inductive arguments are often *enthymematic*, in the sense of [19], which just means that some (in general, true) premisses of the inductive stage of the argument have been suppressed. By supplying the premisses explicitly (or, as is more convenient, finding some formal stand-in for the conjunction of all demonstrable truths), we may restore the explicit, and explicitly valid, form of such enthymematic arguments.)

Similar remarks apply to the temptation to use a rule $\supset E$ (like $\to E$ except that its major premiss is $A \supset B$) in carrying out arguments by *modus ponens*. (This rule $\supset E$, otherwise known as γ, is perhaps the singly most discussed topic, and the most thoroughly studied technically, in all of relevant metatheory. We shall run into it again here, too.) As a matter of practical inference, succumbing to this temptation may not be too bad. But it must be avoided in strictly correct formal derivations in $\mathsf{R}^\sharp$. Again, $\supset E$ is a *factual* inference scheme, not a law-based one. While useful, perhaps, in the mythological and

dream-theoretic aspects of our formal arithmetic, it does not belong in building up effective foundations for the subject. (We might have felt explicitly and sufficiently warned on this subject from opening considerations, since the definition of $\supset$ itself involves $\sim$ in an essential way; while, of course, one can avoid this formally by taking $\supset$ as primitive and equipping it with suitable axioms, there is no way to avoid it conceptually. And here we must make our formal philosophy regarding negation quite clear; our $\sim$ is classical (more accurately, it is as classical as it can be, after Gödel, for $\mathsf{R}^\sharp$ and related systems); and it is, here, to be taken as classically as possible (and not, e.g., in the permuted classical sense of [16]). This means, as above, that we endorse in its entirety the *formal* classical story *about* negation, as it is reflected, e.g., in double negation, excluded middle, and sentential and predicate tautologies in general, in the classical vocabulary. But it is only when we suspend disbelief to operate *within* the classical mythology that we endorse the *informal* classical story. But using $\supset E$ as an *inference rule*—either primitive or derivable on strictly finitary grounds—builds our suspended disbelief into the very structure of our formal system itself. Especially when it is taken, as it often is, as the principal or even the sole primitive inference rule, the effect of $\supset E$ is to assure that *derivations in general* are dependent on the reliability of our intuitions about the erratic connective $\sim$. And this renders impossible one very important goal of Hilbert-style enterprises: namely, to show that basic intuitions are not upset by proof-theoretical detours through areas governed by less secure intuitions. It is notable, accordingly, that some of the most important results concerning these enterprises—e.g., Gentzen's consistency proof for arithmetic—are achieved exactly by limiting detours through $\supset E$. But, unlike Gentzen, who hints that his elimination of $\supset E$ is something of a *tour de force*, we view the rule (save, again, as a practical inference scheme) as positively unwanted, and as involving a fact-law confusion. For, in familiar presentations of classical mathematical theories, $\supset$ is saddled with too many responsibilities; it has a primary responsibility to express certain putative mathematical facts, but a secondary responsibility to congeal into the indicative inference schemes on which these facts are spun out of first principles. That's one too many responsibilities, and, by reserving $\to$ for the second, and letting it carry the deductive responsibilities via $\to E$, we can begin to structure arithmetic propositions in a way that reflects our proof-theoretic grounds for holding them. Since, moreover, we can locate the place of particular propositions in the structure—e.g., that $0 = 1$—we can dispose simply and effectively of the worst offenders against arithmetic

sense, without placing the entire enterprise in jeopardy because parts of the entire structure are not effectively vouchsafed. Meanwhile, our powers of classical expression are not in the least reduced; nor need it be the case that we must do without classical theorems, classically expressed; for $\supset$ remains, freed along with other classical particles for their full and ordinary duties.)

In stating theorems of $\mathsf{R}^\sharp$, we note that it is trivially indifferent whether we state them with particular variables x, y, z; as schemes for schematic variables X, Y, Z; or as schemes for schematic terms T, U, V. Ordinarily we choose the second. And, to begin with, identity has its fundamental properties.

B1. $X = X$
B2. $X = Y \wedge Y = Z \to X = Z$
B3. $X = Y \to Y = X$

So do $+$ and $\times$, 0 and 1.

B4. $0 + X = X$
B5. $X + 1 = X'$; $X + 2 = X''$; etc.
B6. $(X + Y) + Z = X + (Y + Z)$
B7. $X + Y = Y + X$
B8. $X = Y \to X + Z = Y + Z$
B9. $X = Y \to Z + X = Z + Y$
B10. $X + Z = Y + Z \to X = Y$
B11. $0 \times X = 0$
B12. $X \times 1 = X$; $X \times 2 = X + X$; etc.
B13. $(X \times Y) \times Z = X \times (Y \times Z)$
B14. $X \times Y = Y \times X$
B15. $X = Y \to X \times Z = Y \times Z$
B16. $X = Y \to Z \times X = Z \times Y$
B17. $X \times (Y + Z) = X \times Y + X \times Z$

So far as logical properties of $\mathsf{R}^\sharp$ are concerned, they include, of course, all those inherited sententially and quantificationally from the underlying sys-

tems R and RQ, as established in [7] and elsewhere; moreover, many properties not directly inherited are easily demonstrated for $\mathsf{R}^\sharp$ by *mutatis mutandis* arguments. Thus, for example, as admissible rules of $\mathsf{R}^\sharp$ we have ordinary principles like universal generalization, replacement in theorems of provable $\mathsf{R}^\sharp$ equivalents, rewriting bound variables, many classical confinement laws, and the like. Truth-functional connectives and quantifiers have their ordinary elementary properties, in the sense (e.g.) that double negation, De-Morgan laws, associativity, commutativity, and idempotence of $\wedge$ and $\vee$, the usual distributive laws for $\wedge$ and $\vee$, and the usual quantifier analogues (e.g., quantifier interchange) of these principles hold in the form of provable equivalences. A particularly useful metatheorem, incidentally, is the following familiar strengthened form of the replacement theorem: Let $A \to B$ and C be theorems of $\mathsf{R}^\sharp$, and let $C^\star$ result from C by replacing positive occurrences of A with B or negative occurrences of B with A. Then $C^\star$ is a theorem of $\mathsf{R}^\sharp$.

More interesting is the fact that identity in $\mathsf{R}^\sharp$ fully classical, in ways not necessarily to be expected from general relevant principles. The key to this classical character of $\mathsf{R}^\sharp$ identity lies in the following four theorems.

$B18.$ $X = Y \to Z = Z$
$B19.$ $X = Y \to \top$
$B20.$ $\top \to . A \to A$
$B21.$ $A \leftrightarrow \top \to A$

We have already noted that $B18$ is slightly surprising, so some indication of its proof is in order. By symmetry and transitivity of identity, $X = Y \to X = X$. $X = X \to 0 + X = 0 + X$, by $B9$. $0 + X = 0 + X \to 0 = 0$, by the subtraction principle $B10$. By transitivity of provable implication and definition of $\top$, we already have $B19$; i.e., $X = Y \to 0 = 0$. But $0 = 0 \to Z + 0 = Z + 0$ by $B8$, whence $\top \to Z = Z$ applying $A15$ and elementary properties of identity. Putting the last together with $B19$, we get $B18$. Crucial to the demonstration, incidentally, is $A13$, which enters via the inductive proof of key principle $B10$.

$B20$, on the other hand, does not depend on $A13$; it is proved by induction on the complexity of A, the key atomic case being covered by $\top \to Z = Z$ and application of the identity axiom $A12$. Given $B20$, we get $B21$ from left to right by permuting antecedents; from right to left, by detaching $\top$ from

$A2$.

$B21$, incidentally, is the thesis needed to show that $\top$ will have for $\mathsf{R}^\sharp$ the properties of a *least theorem* **t**, mentioned above in motivating $D\top$. In fact, a number of alternative definitions might have been considered; $\forall x(x = x)$, for one. But $0 = 0$ has, perhaps, a peculiar arithmetic appropriateness to serve as a representative theorem of $\mathsf{R}^\sharp$ which implies all the other theorems; it is, we might hold, the fundamental arithmetic truth. Moreover, it will continue to so serve on significant weakening of the present system.

We return now to the substitution properties of identity. We let TX, AX be an arbitrary term and an arbitrary formula of $\mathscr{L}^\sharp$ in which X may occur free, and we extend above conventions by letting TY, TZ as well as ZY, AZ stand temporarily for the indicated proper substitution of Y,Z respectively for X. Then we have the following theorem schemes for $\mathsf{R}^\sharp$.

$B22$. $Y = Z \to TY = TZ$

$B23$. $Y = Z \to . AY \leftrightarrow AZ$

Proof of these theorems is by induction on the complexity of TX, AX respectively. Of particular interest in both arguments is the sub-case of the base case in which no replacements are made (e.g., when AX is $0 = 4$ in $B23$). In this sub-case under $B22$, apply $B18$; under $B23$, apply $B19$, $B20$. For the inductive argument under $B22$, apply $A11$, $B8$, $B9$, $B15$, $B16$. All else is straightforward.

$B22$, $B23$, of course, license the corresponding replacement *rules* in $\mathsf{R}^\sharp$; i.e., we can freely replace provably equal terms in theorems. This means, in particular, that we can revert to customary arithmetic looseness in our notation for terms and formulas, applying commutative, associative, distributive laws, etc., tacitly as required. However, although we have the corresponding admissible *rules* also for provable equivalents in $\mathsf{R}^\sharp$, we do *not* have them in strong forms like $B23$, as *theorems*. For example, although we can prove, as an instance of $B23$, the scheme

$B24$. $Y = Z \to . A \leftrightarrow A$

there are, in $\mathsf{R}^\sharp$, *unprovable* instances of the analogous scheme

(1) $B \leftrightarrow C . \to . A \leftrightarrow A$

On the not unreasonable assumption that equals are equals, and should behave analogously, in any grammatical category, and that moreover $\leftrightarrow$ serves as formula equality in the same sense that $=$ serves as term equality, proposals have been made to formulate a first-order predicate logic $\mathsf{RQ}{=}$ without $B24$, and hence without $B23$ in the unrestricted form above. The most promising of these proposals, in my view, comes from Urquhart, and it consists in taking the AX that enters into $B23$ as an *atomic* formula in which X actually occurs free; when $B23$ is taken as an axiom only for these cases, the ordinary deductive rules produce what seem to be the appropriate instances of $B23$ as theorems. (This question, as noted earlier, is not unrelated to the question of formulating acceptable comprehension axioms for second-order relevant theories. It is likely that there will be considerable leeway in answering such questions, even after the present somewhat unsatisfactory semantical state of relevant quantification theories has been further clarified. No harm is done, it should be noted, to underlying sentential insights even if $B22$, $B23$ are taken generally at full strength; conversely, it has been proposed that $B23$ be weakened even further than by Urquhart.)

Let these matters be as they may, however; $B22$, $B23$, $B24$ are *structurally grounded* in $\mathsf{R}^\sharp$, as (1) is not. And they are grounded on what, for the moment, seem reasonable forms of the Peano postulates, despite our flirtation above with various alternative possibilities. (One further point, however, is in our form $A12$ of the transitivity axiom, which, on general principles, is strictly stronger than the conjunction of $B2$ and $B3$. But $A12$ satisfies the minimal Urquhart criterion, and its modification is an unnecessary last resort.)

Enough has been said, now, to form the notion of a *correct equation* and to show that $\mathsf{R}^\sharp$ is adequate to this notion in one direction. Specifically, a *correct number equation* is one of the form $I = I$, where $I \in \mathbf{N}$. All other number equations $I = J$ are *incorrect*. Evidently, among number equations, an equation is correct iff it is intuitively *true*; incorrect, iff it is intuitively *false*.

It is similarly clear which are to be taken as the *correct numerical equations*. For, by the rules of computation which we learned in third grade, each number term K (e.g., $(2 \times 2)' + 3$) may be *reduced* to a number I; we take this reduction process as familiar to all, and we shall write $r(K)$ for the unique number I that results from its application to the number term K. Then, where K, L are number terms, $K = L$ is a *correct numerical equation* iff $r(K) = r(L)$; otherwise, $K = L$ is *incorrect*. Note that 'correct'

and 'intuitively true' still correspond, as do 'incorrect' and 'intuitively false.' Trivially,

Theorem 1. *Let $K = L$ be a numerical equation. Then if $K = L$ is correct, it is a theorem of* $\mathsf{R}^\sharp$

Proof. By intuitive induction on $r(K)$, using $B1$, $B22$, and the recursion axioms $A15$–$A18$ on $+, \times$, as usual. $\qquad\qquad\qquad\qquad\qquad\square$

Theorem 1 is trivial and effective; indeed, its proof requires no appeal to the *proper* Peano postulates incorporated in $\mathsf{R}^\sharp$, namely $A13$, $A14$, RMI. Rather it is its *converse* that causes all the trouble; while it is nice to know that we can prove $2 + 2 = 4$, it isn't much help if we can also prove $2 + 2 = 5$. I.e., we wish to know not only that true numerical equations are provable, but that false ones are unprovable. To do so, we must assure ourselves that there is no (possibly intricate and involved) proof of, e.g., $2 + 2 = 5$. And this raises exactly the reliability questions examined at the outset; arithmetic faith, which takes it as given that $\mathsf{P}^\sharp$ (and hence $\mathsf{R}^\sharp$) is already consistent in the semantic sense (7) of II, disposes immediately of the converse of Theorem 1. But arithmetic reason, we are told, is stymied; for this would involve consistency at least in sense (3) above; and, in view of Gödel's second theorem, of such consistency we putatively lack any immediately effective guarantee, at least for $\mathsf{P}^\sharp$.

The immediate converse of Theorem 1 is as just stated; false equations are unprovable. But there is a sort of converse that is equally trivial, and which may also be proved here as usual. Where $T = U$ is an equation, let $T \neq U$ be the corresponding *unequation*. *Mutatis mutandis*, we speak of *number unequations, numerical unequations*, as above. For the latter, correctness can be defined in the same third grade manner, still coinciding with intuitive truths. So,

Theorem 2. *Let $K \neq L$ be a numerical unequation. Then if it is correct (i.e., if $K = L$ is incorrect), then $K \neq L$ is a theorem of* $\mathsf{R}^\sharp$.

Proof. By Theorem 1, there are distinct numbers (distinct because $K = L$ is incorrect) I, J such that we can prove $K = I$ and $L = J$. One of I, J, is the greater, and we may suppose it is J. So, for some natural number M, we can prove $I + M' = J$, in $\mathsf{R}^\sharp$. Putting in equals for equals, we prove $K = L \to I = I + M'$. By $B10$, $K = L \to M' = 0$, whence by $A14$ and *modus tollens* (for $\to$), $K \neq L$ is a theorem of $\mathsf{R}^\sharp$, which was to be proved. $\qquad\square$

Theorem 2, unlike Theorem 1, depends on some proper Peano postulates, namely $A13$, $A14$. (As proved, it depends on RMI, but this is inessential.) It can be generalized further, along with Theorem 1, by introducing the notion of a *correct numerical formula*, in the obvious fashion. (I.e., starting with the correct and incorrect numerical equation, we lay it down that $A \wedge B$ is correct iff A is correct and B is correct, while $\sim A$ is correct iff A is incorrect. Incorrectness is defined dually, which takes care of all truth-functional compounding of numerical equations. We shall not worry here about extending these notions immediately to numerical formulas $A \to B$, while quantified formulas are excluded by definition of numerical formula.) Correctness still goes with truth, and incorrectness with falsehood. Then

Theorem 3. *Let A be a numerical formula in which $\to$ does not occur. Then if A is correct, A is a theorem of $\mathsf{R}^\sharp$; while if A is incorrect, $\sim A$ is a theorem of $\mathsf{R}^\sharp$. Moreover, $\mathsf{R}^\sharp$ is* complete *on the set of these numerical formulas, in the sense that at least one of A, $\sim A$ is a theorem of $\mathsf{R}^\sharp$ for each A satisfying the first condition of the theorem.*

Proof. By induction on the complexity of A. Theorems 1 and 2 provide the base cases, while the inductive cases are immediate on elementary properties of the truth-functions. $\qquad\square$

Theorems 2 and 3 generalize Theorem 1 in one direction. To generalize in another direction, we admit free variables into the class of equations that we are considering. And we promote the reader from third grade to eighth grade, by letting him mingle a little elementary algebra in with his arithmetic. To begin with, we introduce standard exponential notation (for *number* exponents) via the definitions

$DEX.$ $T^1 =_{\mathrm{DF}} T$; $T^{I'} =_{\mathrm{DF}} T^I \times T$

for all terms T and positive numbers I. (DEX, of course, does not introduce an exponential function into the system $\mathsf{R}^\sharp$ itself; rather it simply provides an alternate notation for given formal objects, into whose construction an explicit exponentiation operator does not enter.)

Next, we set aside certain terms as *product terms* and as *standard polynomial terms*. Roughly, a *product term* (in variables $X_1, ..., X_N$) shall be anything of the form $IX_1^{J_1}...X_N^{J_N}$, where $I, J_1, ..., J_N$ are positive numbers and $X_1, ..., X_N$ are variables (in increasing order in the original enumeration

of variables). Every natural number shall also count as a product term (in no variables). Every product term shall be a *standard polynomial term*; otherwise, standard polynomial terms shall be of the form $T_0 + ... + T_N$, where each of $T_0, ..., T_N$ is a product term distinct from the number 0. Moreover, in order to insure non-redundancy, no two distinct terms in such a sum shall both be natural numbers, nor shall they differ only by a number coefficient (in the sense, e.g., that $2x^3y^5$ and $7x^3y^5$ differ only by a coefficient). Finally, we shall not admit distinct polynomial terms as both standard if they differ only by commuting summands; e.g., in a standard enumeration of terms, only the first among otherwise standard polynomials that so differ shall in fact count as standard.

By hook and crook, we have fixed things so that each ordinary polynomial, in non-negative integral coefficients, may be identified with exactly one standard polynomial term. Simply by applying commutative, associative, and distributive laws, together with the ordinary laws of arithmetic, we may *reduce*, by eighth grade methods, every term T to its associated standard polynomial term $r(T)$. Where T, U are standard polynomial terms, $T = U$ is a *standard polynomial equation*. For each standard polynomial term T, $T = T$ is *correct*. Otherwise, for standard polynomial terms T, U, we have $T = U$ as *incorrect*. And for an arbitrary equation $T = U$, we say that $T = U$ is *correct* just in case the associated standard polynomial terms $r(T)$, $r(U)$ are identically the same. The wanted generalization of Theorem 1, in the present direction, is now clear.

Theorem 4. *Let $T = U$ be an equation. Then if $T = U$ is correct, then $T = U$ is a theorem of $\mathsf{R}^\sharp$.*

Proof. It suffices to note that the transformations which take us into standard polynomial form are justified by theorems of $\mathsf{R}^\sharp$ already proved, whence $B1$ does the rest. $\square$

Since non-numerical equations are by definition open formulas, correctness for equations in general has ceased to coincide with intuitive truth, save in a secondary sense; i.e., an equation is correct iff it is intuitively satisfied by all numerical values of its free variables. (While this is trivial from left to right, it requires a modicum of proof from right to left, which I leave it to the reader to supply [or look up].) This means that, although we do not wish to extend Theorem 2 by proving $T \neq U$ whenever $T = U$ is incorrect, we do wish the following generalization.

Theorem 5. *Let $T = U$ be an equation, with free variables $X_1, ..., X_N$. Then if $T = U$ is incorrect, $\exists X_1, ..., \exists X_N(T \neq U)$ is a theorem of $\mathsf{R}^\sharp$. Moreover, there are natural numbers $I_1, ..., I_N$, whose respective substitution for $X_1, ..., X_N$ in T, U produces number terms $T^\star, U^\star$, such that $T^\star \neq U^\star$ is a theorem of $\mathsf{R}^\sharp$.*

Proof. Suppose $T = U$ is incorrect. By what was observed in the last paragraph, there are numbers $I_1, ..., I_N$ which, if substituted for $X_1, ..., X_N$ respectively in T, U, produce $T^\star, U^\star$ such that $T^\star = U^\star$ is incorrect. By Theorem 2, $T^\star \neq U^\star$ is a theorem of $\mathsf{R}^\sharp$, proving the final statement of the theorem. The preceding statement then follows by formal existential generalization, ending the proof of Theorem 5. □

We left out the hard part of the proof of Theorem 5 (which is buried in the task parenthetically left to the reader above), but enough has now been observed to show that $\mathsf{R}^\sharp$ is adequate, in at least a rudimentary way, to the demands of elementary number theory. Indeed, the main business of this section has been to note that the transition from $\mathsf{P}^\sharp$ to $\mathsf{R}^\sharp$ has produced no weakening of the most basic and essential insights; otherwise, all is thoroughly familiar, and would no doubt be accepted by the reader without even the small amount of proof that has been here provided, as being demonstrable in a totally effective way.

While we have been showing that $\mathsf{R}^\sharp$ is as adequate as $\mathsf{P}^\sharp$ in essential respects, it is fit to note that it shares some of its inadequacies also. For example, we cannot extend our effective characterization of correctness even to unequations in general, far less to formulas compounded truth-functionally from arbitrary equations. Let us appeal, instead, to the standard mythology, calling an unequation $T \neq U$ *valid* just in case it is true (or correct) for all numerical values of its free variables. Then

Theorem 6. *Assume the standard mythology. Then there is a valid unequation $T \neq U$ which is unprovable in $\mathsf{R}^\sharp$. Moreover, there is no recursively axiomatizable extension $\mathsf{S}^\sharp$ of $\mathsf{R}^\sharp$, with the same rules as $\mathsf{R}^\sharp$, in which all valid unequations are provable, and which is consistent in sense* (6).

Proof. Matiásévič has proved in [21] that Hilbert's 10th problem is recursively unsolvable. We may take that problem, here, in the form: Find a decision procedure for *truth* (in the standard model $\mathbb{N}$) for *sentences* of $\mathscr{L}^\sharp$ of the form $\exists X_1, ..., \exists X_N(T = U)$. Suppose that such a sentence is true; then

it is certainly provable in $\mathsf{R}^\sharp$, by Theorem 1 and existential generalization. Let $\mathbf{M}$ be the set of true sentences of the given form, and let $\mathbf{V}$ be the set of all sentences of this form. Then $\mathbf{M}$ is certainly recursively enumerable. If $\mathbf{V} \smallsetminus \mathbf{M}$ is also recursively enumerable, then $\mathbf{M}$ is recursive, contradicting [21].

Suppose now that every arithmetically valid unequation $T \neq U$ is a theorem of $\mathsf{R}^\sharp$, for *reductio*. Then every universal closure of a valid unequation is a theorem of $\mathsf{R}^\sharp$; i.e., every *true sentence* of the form $\forall X_1, ..., \forall X_N (T \neq U)$ is a theorem of $\mathsf{R}^\sharp$. So, by quantifier interchange, for all $A \in \mathbf{V} \smallsetminus \mathbf{M}$, $\sim A$ is a theorem of $\mathsf{R}^\sharp$. Suppose also that whenever $\sim A$ is a theorem of $\mathsf{R}^\sharp$ we have $A \in \mathbf{V} \smallsetminus \mathbf{M}$. Then, since again we can recursively enumerate theorems of $\mathsf{R}^\sharp$, we can enumerate $\mathbf{V} \smallsetminus \mathbf{M}$, which has already been observed to be impossible. So, for some theorem $\sim A$ of $\mathsf{R}^\sharp$, $A \in \mathbf{M}$. Whence, by the last paragraph, A is also a theorem of $\mathsf{R}^\sharp$. I.e., $\mathsf{R}^\sharp$ is inconsistent with respect to negation. But, on the standard mythology being assumed, all theorems of $\mathsf{P}^\sharp$, and *a fortiori* of $\mathsf{R}^\sharp$, are true in the standard model $\mathbb{N}$, whence $\mathsf{R}^\sharp$ cannot be inconsistent. This refutes our initial reductio assumption in the present paragraph, and so establishes the claim of the theorem that some valid unequation $T \neq U$ is unprovable in $\mathsf{R}^\sharp$.

This proves the second statement of the theorem. For the final statement, let $\mathsf{S}^\sharp$ be some recursively axiomatizable extension of $\mathsf{R}^\sharp$ in which all valid unequations are provable. (Note that it is not required that all statements of $\mathsf{S}^\sharp$ be arithmetically true.) Reason as in the last paragraph to show $\mathsf{S}^\sharp$ inconsistent in sense (6). This suffices to complete the proof of Theorem 6. □

So $\mathsf{R}^\sharp$, like $\mathsf{P}^\sharp$, begins to become inadequate where negation is concerned, even on very simple formulas. Moreover, there is no reliable way to make it adequate, even though the class of recursively axiomatizable extensions of $\mathsf{R}^\sharp$ is much richer than the corresponding class for $\mathsf{P}^\sharp$. (Exercise: In what sense?) On previous remarks, this is hardly surprising; indeed, the argument just gone through might be repeated word for word to get the corresponding (well-known) result for $\mathsf{P}^\sharp$. And the upshot is that we seem to be succeeding in one of our principal purposes; we aimed for a classical theory of negation, and, so far as we can tell, we got one, warts and all. What needs now to be shown is that $\mathsf{R}^\sharp$ is not *at the mercy* of its theory of negation; and accordingly that, in general, formal arithmetic shall no longer be held to ransom over every jot and tilde of its fashionable formulation.

In particular, having set out certain adequacy theorems above, we shall wish to turn below to corresponding reliability theorems. For each of our successive characterizations of correctness raises the corresponding problem of avoiding incorrectness. E.g., just as we wished to round off Theorem 1 with an equally effective guarantee that *only* correct numerical theorems shall be provable in $\mathsf{R}^\sharp$, so we wish to round off its generalization to Theorem 4 with an effective guarantee that no incorrect equations at all will be theorems of $\mathsf{R}^\sharp$. This would, of course, be an effective guarantee that no inconsistency can arise in $\mathsf{R}^\sharp$ in sense (4) of II, and hence in senses (1), (2), and (3) as well.

It would be nice, also, to round off Theorems 2, 3, and 5 with corresponding reliability results. But this is less hopeful, in the present perspective, save by relaxing our criteria of reliability. For it is characteristic even of Theorem 2, and hence of its generalizations to Theorems 3 and 5, that it involves formal negation in an essential way. But this, as it is seen here, is exactly where the most striking and fundamental trouble lies. I.e., as it is seen here, there is nothing wrong with seeking strong guarantees, even within formalized mathematics itself, of the reliability of mathematical intuitions. There is something wrong with seeking *mathematical* guarantees of the *logical* apparatus that we bring to bear on these intuitions. For who would expect, by taking sums, to demonstrate thereby that he avoids the Liar Paradox? 'Not,' it is generally conceded, is a logical particle; it ought also to be conceded that it is, by and large, a logical problem, and that it is not part of the enterprise to provide ironclad guarantees against logic running amuck. And so we shall show, in fact, that none but the usual non-constructive guarantees are available in defence of reliability criteria to match the adequacy criteria of Theorems 2, 3, and 5, save in special cases. And this is reasonable; if negation is, in Hilbertian terms, an ideal element, which we *reason through* and do not *reason to*, the guarantee that is wanted is not that we shall get all the *negative facts* straight (even on the somewhat generous philosophical assumption that there are any negative facts at all). It is rather that our reasonings through putative negative facts shall not impede our pursuit of, nor our confidence in, fundamental positive facts. None of this impedes us from appealing to, and buying as much of, the standard mythology as we want to; in our ordinary and everyday tone of voice, we shall probably buy it all. But let us not suppose, either that arithmetic totters if it totters.

VI

In this section, we develop some further properties of relevant arithmetic, not all of them necessarily expected. To begin with, we define

$D{\supset}$. $A \supset B =_{\mathrm{DF}} A \wedge \top \to B$

$D{\equiv}$. $A \equiv B =_{\mathrm{DF}} (A \supset B) \wedge (B \supset A)$

$D{\neg}$. $\neg A =_{\mathrm{DF}} A \supset \mathsf{F}$

$\supset$, $\equiv$, $\neg$ have many of the properties of an intuitionist implication, equivalence, and negation respectively; in particular, they are to be sharply distinguished from their truth-functional counterparts $\supset$, $\equiv$, $\sim$ introduced above. (For comparison in the latter case, note the following theorem scheme of $\mathsf{R}^\sharp$.)

$QD{\sim}$. $\sim A \leftrightarrow A \to \bot$

$QD{\sim}$ serves as a *quasi-definition*, in the sense of [8], of $\sim$ in $\mathsf{R}^\sharp$, and it illustrates nicely the distinction between an inferential analysis of intuitionist and ordinary negation. For the former is characterized by a loose (enthymematic) implication of an absurd proposition, namely $0 = 1$; the latter by a tight (relevant) implication of a standard (but not necessarily absurd) falsehood; while we have chosen $0 \neq 0$ for that purpose here, we might, of course, have introduced $\bot$ as a direct primitive, using $QD{\sim}$ then as a definition.

Negation has been, in one way or another, a constant concern in this paper, and so it is important to set out with utmost clarity our formal attitude toward it. This attitude is guided by two purposes. The first is that our underlying negation should be capable of an inferential reduction, with reference to our underlying implication; $QD{\sim}$ accomplishes this purpose. The second is that the properties of this negation should be, so far as possible, compatible with the standard mythology. One or the other of these purposes is often preferred, and it is variously argued that relevant logics should prefer one of them—e.g., that they should favor a classical semantical negation in the sense of [22], or else that they should drop some classical principles (especially double negation or excluded middle). (Urquhart can be quoted in defence of both preferences; his intentions are classical semantical in [23], and constructive inferential in [24]; and the second approach has been pursued in depth by Pottinger; e.g., in his contribution to [24].)

However, interesting as these alternate approaches are in themselves (the former in particular having been pursued at length in [22] and related papers), they do not seem adequate to a philosophically interesting approach to the role and utility of negation in classical mathematics. Some reconciliation between them is needed (of the sort that accompanies the traditional treatment of $\sim$ in relevant logics, and which is being pursued here). To assume the standard mythology outright is to build one's mathematical house upon sand; in fact, it is to suggest that that mythology is less secure and reliable than in fact it is, just because the mythology itself gets deeply tangled up in the reliability proofs. But not to give the standard mythology its inferential due, in the way that we actually set out formal systems, is to deprive mathematics of some of its central inferential principles. In short, I opt here neither for a classical semantical nor for a constructive inferential approach to formal negation, but rather for what I have called elsewhere the *classical inferential* approach to formal negation—classical because its semantic *intentions* are classical but inferential because I hold that the formal *treatment* of negation belongs not to mythology but to logic—i.e., to the theory of *inference*, and not to semantical fairy tales as to how that theory is grounded.

To return to our "intuitionist" connectives $\supset$ and $\neg$, we can spell out rather precisely the sense in which they are intuitionist in $\mathsf{R}^\sharp$. First, let us characterize an *intuitionist sublanguage* $\mathscr{L}^\circ$ of $\mathscr{L}^\sharp$ in the obvious fashion; the set of formulas of $\mathscr{L}^\circ$ shall be got by closing the set of atomic formulas of $\mathscr{L}^\sharp$ under the (primitive and $\mathsf{R}^\sharp$-defined) connectives $\supset, \wedge, \vee$, and quantifiers $\forall X$, $\exists X$. Note that $\mathscr{L}^\circ$ is automatically closed, by definitions, under $\neg, \equiv$ also. To save further proliferation of notation, we may also sometimes consider $\mathscr{L}^\circ$ as a language in its own right, with the same atomic formulas, but with $\supset, \wedge, \vee, \forall X, \exists X$ primitive and $D\neg$ and $D\equiv$ in force. When $\mathsf{J}^\sharp$ is formulated in this language (as suggested above and made explicit in [5]), we shall call it J°. Similarly, we shall let R° be the system based on $\mathscr{L}^\circ$ whose set of theorems consists exactly (on the obvious transformation) of those theorems of $\mathsf{R}^\sharp$ which belong to the intuitionist sublanguage $\mathscr{L}^\circ$ of $\mathscr{L}^\sharp$.

There are now two interesting questions about the relationship between $\mathsf{R}^\sharp$ and $\mathsf{J}^\sharp$. First, there is a question of the relationship between R° and J°; i.e., among those formulas of $\mathscr{L}^\sharp$ which can be *expressed* in what (from the viewpoint of $\mathsf{R}^\sharp$) is the intuitionist sublanguage $\mathscr{L}^\circ$, which are the relevantly valid and which are the intuitionistically valid? Second, there is the question of the relationship between $\mathsf{R}^\sharp$ as a whole and that part of it, namely R°, that can be directly related to intuitionist insight; e.g., what is the relationship

in $\mathsf{R}^\sharp$ between its own implication $\rightarrow$ and the defined "intuitionist" $\supset$? Or between $\sim$ and $\neg$? Moreover, to what extent can intuitionistically valid lines of reasoning be applied in $\mathsf{R}^\sharp$, and how can they be applied?

Formulating general answers to these questions is not easy, despite the simple and evident relations that obtain between $\mathsf{J}^\sharp$ and $\mathsf{R}^\sharp$ at the sentential level, and which may be gleaned from [7] and [13]. For, first, these relations are not quite so evident at the proper quantificational level (though, as [7] notes, they continue to be of interest). More important, at the proper number-theoretic level, we run into problems like those surveyed earlier with regard to $A14$. As we are setting down the axioms for $\mathsf{R}^\sharp$, we could always include $\neg(x' = 0)$, or something that relevantly implies it, among them; there are, as we have seen, some plausible reasons to do so. The problem is that, relevantly, this can no longer be considered an honest formal counterpart of a Peano postulate; either, then, it is another necessary postulate which Peano overlooked, or else it is just false. And this illustrates very neatly the *logical relativity* of formalized mathematics, and the folly of supposing that entrenched custom will offer us *preferred logics*. For even the form of our formal mathematical *axioms* may be affected by our presuppositions regarding the logic that we are using. In this situation, it does little good to assert that we are employing the same inferential principles that all right-thinking people have always used (or would have used, if only they had been clear enough). For the inferential principles that we are using themselves contribute to the *sense* that is to be made of the very assertions to which they are to be applied in the drawing of inferences. E.g., from an $\mathsf{R}^\sharp$ viewpoint the inferential behavior of $\neg$ *disqualifies it* as a formal counterpart of 'not'; i.e., statements of the form $\neg A$ do not *mean* that it is not the case that A, quite apart from the question whether they are true or false and even from our general worries about formal negation.

These questions have been thought about too simplistically, and too little. For having (or lacking) a principle is too easily confused with having (or lacking) such-and-such a formal theorem. To end the confusion, one must also know that the formal theorem (or non-theorem) expresses the principle in question. And this is exactly the point which is most insecure. The sense in which $\mathsf{R}^\sharp$ contains (or fails to contain) intuitionist arithmetics must be related to the sense that $\mathsf{R}^\sharp$ makes of these arithmetics, and not necessarily to the sense that it makes of the intuitive principles that are supposed to underlie them. And what we find, in fact, are certain convergences, and certain divergences as well, in underlying insight. Oversimplified, insights converge

in that $\mathsf{R}^\sharp$ takes intuitionist deduction as a special case in the general theory of relevant deduction; thus, as $D\supset$ suggests, intuitionist implication may be viewed as a special kind of relevant implication. But insights diverge on negation and the quantifiers; from the $\mathsf{R}^\sharp$ viewpoint, $\mathsf{J}^\sharp$ accepts enough of the standard mythology on such matters that $\mathsf{J}^\sharp$ is no more credible than $\mathsf{P}^\sharp$; while the points on which $\mathsf{J}^\sharp$ rejects the standard mythology render it less useful and interesting than $\mathsf{P}^\sharp$. For $\mathsf{R}^\sharp$, on the other hand, the standard mythology is there to be used *mythologically*; thus, on non-mythological assertions, $\mathsf{R}^\sharp$ is credible in ways that neither $\mathsf{P}^\sharp$ nor $\mathsf{J}^\sharp$ is credible; while $\mathsf{R}^\sharp$ still retains (we hope) the full mythological interest and utility of $\mathsf{P}^\sharp$.

We now provide partial answers to the questions raised above about the relationship of $\mathsf{J}^\sharp$ and $\mathsf{R}^\sharp$; to provide these answers, we consider $\mathsf{J}^\sharp$ in its completely equivalent form J°, based on the intuitionist sublanguage $\mathscr{L}^\circ$. Note, to begin with, that the following rule of *modus ponens* for $\supset$ (unlike the corresponding rule for material $\supset$) is verified without difficulty in $\mathsf{R}^\sharp$, for all formulas A, B of $\mathscr{L}^\sharp$.

$\supset E.\quad A \supset B\ \&\ A\ \Rightarrow\ B$

And we note that $\supset$ can be related in $\mathsf{R}^\sharp$ to standard notions of deducibility. Specifically, where $\mathbf{S}$ is a set of formulas and A is a formula, there is a standard deduction of A from $\mathbf{S}$ (symbolically $\mathbf{S} \vdash A$) in $\mathsf{R}^\sharp$ just in case there is a sequence of formulas $A_1, ..., A_N$ of $\mathscr{L}^\sharp$ such that (i) A_N is A, and (ii) each of the A_I is either a member of $\mathbf{S}$, a theorem of $\mathsf{R}^\sharp$, or the consequence of the predecessors by one of $\to E$, $\wedge I$. (It is convenient, but not necessary, to restrict the rules that may be employed in standard deduction to the primitive sentential rules of $\mathsf{R}^\sharp$. For we can throw in use of the quantificational rule $C\forall I$, or the induction rule RMI, simply by keeping track in familiar ways (e.g., as in [5]) of the variables concerning which $\mathbf{S}$ makes no special assumptions, and hence which are fit to be used as variables of generalization under $C\forall I$ or as induction variables under RMI; alternatively, and without loss of generality, we can limit $\mathbf{S}$ to sentences, again allowing ourselves free use of generalization and induction rules. Finally, we might trade in $\to E$, $\wedge I$ for the single rule $\supset E$ under (ii) of our characterization above. But these suggestions, except possibly the last, are to be taken as unnecessary complications, slightly affecting the form, but not the content, of the deduction theorem to be noted.)

We write, in Gentzenish fashion, $\mathbf{S} \cup \{A\}$ as $\mathbf{S}, A$ in context.

The Standard Deduction Theorem for $\mathsf{R}^\sharp$. For all subsets $\mathbf{S}$ of $\mathscr{L}^\sharp$, and formulas A, B of $\mathscr{L}^\sharp$, we have

(1) $\mathbf{S}, A \vdash B$ iff $\mathbf{S} \vdash A \supset B$

Proof. The methods of [13] suffice. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

To this principle we may add a number of others, which suffice to guarantee standard deductive behavior for $\wedge$, $\vee$, $\forall X$, $\exists X$.

(2) $\mathbf{S} \vdash A \wedge B$ iff $\mathbf{S} \vdash A$ and $\mathbf{S} \vdash B$

(3) $\mathbf{S}, A \vee B \vdash C$ iff $\mathbf{S}, A \vdash C$ and $\mathbf{S}, B \vdash C$

(4) $\mathbf{S} \vdash A$ iff $\mathbf{S} \vdash \forall X A$, provided that X does not occur free in any member of $\mathbf{S}$

(5) $\mathbf{S}, A \vdash B$ iff $\mathbf{S}, \exists X A \vdash B$, provided that X does not occur free in any member of $\mathbf{S}$ or in B

(6) $\mathbf{S} \vdash A$ and $\mathbf{S}, A \vdash B$ yield $\mathbf{S} \vdash B$

(7) If $\mathbf{S} \subseteq \mathbf{T}$ and $\mathbf{S} \vdash A$ then $\mathbf{T} \vdash A$

(8) If $A \in \mathbf{S}$ then $\mathbf{S} \vdash A$

(9) If $A \in \mathsf{R}^\sharp$ then $\varnothing \vdash A$

Again, verifications are straightforward. And the import of (1)–(9) lies in a partial answer to our second question above; namely, we can, at least where intuitionist negation is not involved, apply intuitionistically valid principles of reasoning quite generally in $\mathsf{R}^\sharp$, and not simply to that part R° of $\mathsf{R}^\sharp$ which is, so to speak, in the intuitionist vocabulary $\mathscr{L}^\circ$. Reasoning via mathematical induction counts too.

(10) If $\mathbf{S} \vdash A0$ and $\mathbf{S}, AX \vdash AX'$, then $\mathbf{S} \vdash AX$, provided that X does not occur free in $\mathbf{S}$.

Verification of (10) rests on the following admissible rule and theorem of $\mathsf{R}^\sharp$.

$RMJ.\quad A0 \,\&\, AX \supset AX' \;\Rightarrow\; AX$

$AMJ.\quad A0 \wedge \forall X(AX \supset AX') \supset AX$

RMJ may be got easily by taking $\top \wedge AX$ as the induction formula in an application of the primitive rule RMI, whence the scheme AMJ (corresponding to AMI) follows quickly as a theorem scheme of $\mathsf{R}^\sharp$. (This point is looked into in some detail in [18]; suffice it to note here that it is often convenient in carrying out arguments by mathematical induction formally in $\mathsf{R}^\sharp$ to appeal to RMJ, AMJ rather than RMI directly.)

At this point, the reader must be cautioned against taking *too seriously* the standard deduction theorem for $\mathsf{R}^\sharp$ and its various concomitants. For the standard sense of deduction is not the preferred sense of deduction in $\mathsf{R}^\sharp$, for reasons set out in [7] and elsewhere. However, $\mathsf{R}^\sharp$ *accommodates* standard deduction in the ways set out above, which is germane to the $\mathsf{J}^\sharp$ connection presently being investigated. (This kind of accommodation holds quite generally for systems based on RQ, being specialized here to $\mathsf{R}^\sharp$ simply because that is the system presently being investigated.)

The first systematic conclusion to be drawn for $\mathsf{R}^\sharp$ from (1)–(10) is the following.

Theorem 7. *Let A be any negation-free theorem of first-order intuitionist quantification theory* JQ. *Then every substitution instance of A in the language $\mathscr{L}^\sharp$ is a theorem of $\mathsf{R}^\sharp$.*

Proof. The negation-free theorems of JQ may be axiomatized using the negation-free axioms and rules of [5]. Use (1)–(9) to show that all substitution instances of these axioms hold in $\mathsf{R}^\sharp$, and that $\mathsf{R}^\sharp$ theoremhood is preserved on application of JQ rules, ending the proof of the theorem. (Examples of verification:

(i) Of axiom $A \supset. B \supset A$. By (8), $A, B \vdash A$ in $\mathsf{R}^\sharp$ for arbitrary $A, B \in \mathscr{L}^\sharp$. By (1), $A \vdash B \supset A$. By (1) again, $\varnothing \vdash A \supset. B \supset A$. By the converse of (9), $A \supset. B \supset A$ is accordingly a (schematic) theorem of $\mathsf{R}^\sharp$.

(ii) Of rule $A \supset B \;\Rightarrow\; \exists X A \supset B$, where X is not free in B. Suppose, for arbitrary formulas A, B of $\mathscr{L}^\sharp$, that $A \supset B$ is a theorem of $\mathsf{R}^\sharp$ and that X is not free in B. By (9), (1), $A \vdash B$ in $\mathsf{R}^\sharp$. By (5), $\exists X A \vdash B$ in $\mathsf{R}^\sharp$. By (1), $\varnothing \vdash \exists X A \supset B$ in $\mathsf{R}^\sharp$, completing the verification.) $\qquad\square$

Theorem 7 does not involve negation, but it holds in the strong schematic sense explained in the previous footnote. By sticking to the intuitionist sublanguage $\mathscr{L}^\circ$ of $\mathscr{L}^\sharp$, we get a stronger result.

Theorem 8. *Let A be any theorem of intuitionist quantification theory* JQ, *in the language $\mathscr{L}^\circ$. Then A is a theorem of* R$^\circ$, *and hence of* R$^\sharp$.

Before proving Theorem 8, we note that R$^\sharp$ admits the following form of the *ex falso quodlibet*. First, let the *positive sublanguage* of $\mathscr{L}^\sharp$, which we call $\mathscr{L}^+$, contain all atomic formulas of $\mathscr{L}^\sharp$ and be closed under $\wedge$, $\vee$, $\rightarrow$, $\exists X$, $\forall X$; i.e., $\mathscr{L}^+$ contains negation $\sim$ only insofar as it may be eliminated by one of the definitions $D\vee$, $D\exists$. Note also that, by attending to the appropriate definitions, $\mathscr{L}^+$ also contains $\top$, F, $\supset$, $\neg$, etc., and so it has the entire intuitionist sublanguage $\mathscr{L}^\circ$ as its proper part. We shall use R$^{\sharp+}$ for the set of theorems of R$^\sharp$ in the positive sublanguage $\mathscr{L}^+$. Now for the EFQ.

EFQ Lemma for R$^{\sharp+}$. For all formulas A of $\mathscr{L}^+$, we have F $\rightarrow A$ as a theorem of R$^{\sharp+}$.

Proof. By induction on the length of A. In the atomic case, it suffices to show $0 = 1 \rightarrow X = Y$, for distinct variables X, Y, using RMI. Where A is compound, it is of one of the forms $B \wedge C$, $B \vee C$, $B \rightarrow C$, $\forall XB$, $\exists XB$, where we may assume on inductive hypothesis that $0 = 1 \rightarrow B$, $0 = 1 \rightarrow C$ are theorems of R$^{\sharp+}$. Every case except that in which A is of the form $B \rightarrow C$ may be settled on elementary properties of classical logic particles. We borrow from below the result that, since B is a formula of $\mathscr{L}^+$, $B \rightarrow \top$ is a theorem of R$^\sharp$. Then on inductive hypothesis and $B21$, we have F $\rightarrow$. $\top \rightarrow C$ as a theorem of R$^\sharp$; whence by the borrowed result and the strengthened replacement theorem for R$^\sharp$, we have F $\rightarrow$. $B \rightarrow C$ as a theorem of R$^\sharp$, completing the inductive argument and establishing the *ex falso quodlibet* for R$^{\sharp+}$ in the form just set out. $\qquad\square$

So we may add to our list of interesting R$^\sharp$ theorems

$B25$. F $\rightarrow A$, where $A \in \mathscr{L}^+$
$B26$. F $\supset A$, where $A \in \mathscr{L}^+$

For that matter, we might as well record, as an R$^\sharp$ theorem scheme

$B27.\quad B \to C \to.\ B \supset C$

whence the passage from $B25$ to $B26$ is clear. And we may now add to our principles of standard deduction that are $\mathsf{R}^\sharp$ valid

(11) If $\mathbf{S} \vdash \mathsf{F}$ then $\mathbf{S} \vdash A$, for each set $\mathbf{S}$ of formulas of $\mathscr{L}^\sharp$ and each formula A in $\mathscr{L}^+$.

Then, resuming our interrupted business,

Proof of Theorem 8. Suppose that B is a theorem of JQ, in the language $\mathscr{L}^\circ$. We may suppose that JQ is formulated in F-formulation, as in [8], with $\mathsf{F} \supset A$ as the sole special axiom scheme governing F and with negation introduced by a quasi-definition $\neg A \equiv A \supset \mathsf{F}$, taken also as an axiom scheme. The negation-free axioms of JQ may be taken as before, and the rules as well. Then, since our language is $\mathscr{L}^\circ$ to begin with, we simply show that every step in a JQ proof of B, in that language, is a theorem of $\mathsf{R}^{\sharp+}$. All steps except those that come by axioms $\mathsf{F} \supset A$ may be handled as before (taking care of the quasi-definition by our outright definition $D\neg$), whereas $\mathsf{F} \supset A$ holds directly by $B26$ for all formulas A in $\mathscr{L}^\circ$, thus completing the proof of Theorem 8. □

The reader may well have wondered why we bothered to state Theorem 7 as a separate result, inasmuch as Theorem 8 seems more satisfying; he may also wonder why we are so stuck on JQ theorems, since except in the passing observation (10) we have not yet attended to such theorems of $\mathsf{J}^\sharp$ as involve the proper number-theoretic axioms. Again, our concern for negation—both the intuitionistic negation $\neg$ of J° and the ordinary negation $\sim$ of $\mathsf{R}^\sharp$—is the cause. For, while we have *defined* $\neg$ on all formulas of $\mathscr{L}^\sharp$, it only has the intuitionist EFQ property on *positive* formulas. So the distinction between Theorem 7 and Theorem 8 is important; $A \supset.\ B \supset A$, and every other negation-free theorem scheme of JQ, *remains* a theorem scheme in $\mathsf{R}^\sharp$; that is the burden of Theorem 7. But $\mathsf{F} \supset A$, and like schemes, is only a *partial* theorem scheme of $\mathsf{R}^\sharp$; e.g., it holds for all formulas A of $\mathscr{L}^\sharp$ which *translate directly* corresponding formulas of $\mathscr{L}^\circ$, but it does not necessarily hold on formulas A that are not admitted into $\mathscr{L}^\circ$ at all; such formulas, needless to say, involve the negation $\sim$ of $\mathsf{R}^\sharp$ in an essential way.

In fact, by altering our definition of F in $\mathscr{L}^\sharp$, we can get a translation of J° into $\mathsf{R}^\sharp$ to which Theorem 7 can be extended directly, without the cautions attached to Theorem 8. For let us define

$D\mathtt{F}_*$. $\mathtt{F}_* =_{\mathrm{DF}} 0 \neq 1 \to 0 = 1$ (i.e., $\mathtt{T} \to \mathtt{F}$)

$D\neg_*$. $\neg_* A =_{\mathrm{DF}} A \supset \mathtt{F}_*$

We now have, in $\mathtt{F}_*$, a formula that relevantly implies all formulas of $\mathscr{L}^\sharp$ in the system $\mathsf{R}^\sharp$.

EFQ Lemma for $\mathtt{F}_*$. For all formulas A of $\mathscr{L}^\sharp$, we have $\mathtt{F}_* \to A$ as a theorem of $\mathsf{R}^\sharp$.

Proof. We show, by induction on the length of A, that both $\mathtt{F}_* \to A$ and $A \to \sim\mathtt{F}_*$ are theorems of $\mathsf{R}^\sharp$. It is clear from the reductio axiom $A9$ and double negation that $\mathtt{F}_* \to \mathtt{F}$ is a theorem of $\mathsf{R}^\sharp$, whence by the previous EFQ Lemma we have $\mathtt{F}_* \to A$ for all $A \in \mathscr{L}^+$, including in particular atomic A. Of course we have $\sim\mathtt{F}_*$, and so $\top \to \mathtt{F}_*$ by $B21$; whence, for atomic A, we have $A \to \sim\mathtt{F}_*$ applying $B19$. This establishes our induction proposition for the case where A is atomic; we must consider in addition the case where A is of one of the forms $B \wedge C$, $\sim B$, $\forall X B$, $B \to C$, where on inductive hypothesis we have $\mathtt{F}_* \to B$, $\mathtt{F}_* \to C$, $B \to \sim\mathtt{F}_*$, $C \to \sim\mathtt{F}_*$ as theorems of $\mathsf{R}^\sharp$, and where we must show $\mathtt{F}_* \to A$, $A \to \sim\mathtt{F}_*$. Using contraposition, the case where A is $\sim B$ is already built into the inductive hypothesis (appealing also to double negation), while the $B \wedge C$, $\forall X B$ cases remain easy to verify; similarly, using the fact that $B \to C \to \sim(B \wedge \sim C)$ is relevantly valid, the inductive hypothesis quickly delivers $B \to C \to \sim\mathtt{F}_*$ as a theorem. So the interesting case is still $\mathtt{F}_* \to. B \to C$; this cannot be handled as before, since we are no longer assured that $B \in \mathscr{L}^+$. However, it will suffice for the latter (by the strengthened replacement theorem for $\mathsf{R}^\sharp$ and inductive hypothesis) that we can show $\mathtt{F}_* \to. \sim\mathtt{F}_* \to \mathtt{F}_*$ to be a theorem of $\mathsf{R}^\sharp$. For this we establish theorems of $\mathsf{R}^\sharp$ as follows:

- $(\mathtt{T} \to \mathtt{F}) \circ \mathtt{T} \to \mathtt{F}$
- $(\mathtt{T} \to \mathtt{F}) \circ \mathtt{T} \to \top$
- $(\mathtt{T} \to \mathtt{F}) \circ \mathtt{T} \to. \mathtt{T} \to \mathtt{F}$
- $\mathtt{T} \to \mathtt{F} \to. \mathtt{T} \to. \mathtt{T} \to \mathtt{F}$
- $\mathtt{T} \to \mathtt{F} \to. \mathtt{T} \to. \mathtt{T} \to. \mathtt{T} \to \mathtt{F}$
- $\mathtt{T} \to \mathtt{F} \to. \mathtt{T} \to. \sim(\mathtt{T} \to \mathtt{F}) \to \mathtt{F}$
- $\mathtt{T} \to \mathtt{F} \to. \sim(\mathtt{T} \to \mathtt{F}) \to. \mathtt{T} \to \mathtt{F},$

leaving the reader to fill in details. The final formula is what is to be shown, and it completes the EFQ Lemma for F_*. $\qquad\square$

The above lemma suffices for the extension of Theorem 7 to all of JQ on the suggested translation.

Theorem 7$_*$. Let A be any theorem of first-order intuitionist quantification theory JQ. Then every substitution instance of A in the language $\mathscr{L}^\sharp$ is a theorem of $\mathsf{R}^\sharp$, on straightforward translation into $\mathscr{L}^\sharp$ of intuitionist $\wedge$, $\vee$, $\forall X$, $\exists X$ and using $D{\supset}$, $D\neg_*$, $D\mathsf{F}_*$ for intuitionist implication, negation, and an intuitionistically absurd proposition, respectively.

Proof. As of Theorem 7, using the EFQ Lemma for F_* to complete the proof on axioms $\mathsf{F} \supset A$ of JQ. $\qquad\square$

In our particular $\mathsf{J}^\sharp$ context, however, it may be doubted that Theorem 7$_*$, despite its increased generality, really improves Theorem 8. For F was defined so as to correspond naturally to an arithmetically absurd proposition, namely $0 = 1$. As such, F is firmly within an intuitionistically acceptable vocabulary to begin with. On the other hand, our F_* is defined using the intuitionistically unacceptable connective $\sim$. It adds nothing to the soundness of the translation of the intuitionistically acceptable sentences of $\mathscr{L}^\circ$ themselves. And it subtracts from the adequacy of the translation by forcing upon us, at the *logical level*, some notoriously objectionable theses (to intuitionists). E.g., as in [13], we have $\neg_* A \vee \neg_* \neg_* A$ as a theorem scheme of $\mathsf{R}^\sharp$, as it certainly is not in JQ (or even in $\mathsf{J}^\sharp$, if $\mathsf{J}^\sharp$ is consistent).

We have, of course, no particular brief here either for intuitionist affirmations or for intuitionist denials, in themselves. $\mathsf{R}^\sharp$ already diverges from $\mathsf{J}^\sharp$ on some points; while one can decrease, or perhaps even eliminate, key divergences by fiddling with one system or the other, there is no particular reason to do so here. But the *delicacy* of these matters of translation is again apparent; for intuitionist and relevant insight more nearly coincide when we translate the intuitionist absurdity F by a formula which, relevantly speaking, is not absurd. And this is reasonable; from the $\mathsf{R}^\sharp$ viewpoint, the intuitionist theory of implication is generally acceptable, when understood correctly. But the intuitionist theory of negation is not generally acceptable; as a theory about absurd propositions, it is acceptable, more or less, but only subject to the constraint that such propositions, like $0 = 1$ in the arithmetic case, be absurd (i.e., imply everything) not in the whole relevant vocabulary, but only in that part of it which directly reflects intuitionist concerns.

So, in thinking generally about the relation between J and R, here is my position. At the sentential level, J-insights are exactly contained in R-insights, in the way spelled out formally in [13]. And this is not merely, I think, a matter of owning an exact translation from J to R, at this level, but also a matter of coincidence of insight reflected by the translation; the intuitionist treatment of ¬, while it is relevantly strange and must be approached with the caution just noted, does not upset this coincidence. At the quantificational level, there appears, on the other hand, to be an important divergence between JQ and RQ over the status of the confinement law that allows distribution of $\forall X$ over $\vee$. I doubt that this divergence can be reconciled without serious motivational impact on each system; the less serious impact, I think, comes in adding it to JQ; under certain circumstances, an intuitionist can perhaps accept this confinement law, though he must be presumed to reject it in general. On the other hand, it is hard to see how a relevant viewpoint can do without this confinement law, while preserving its classical inferential theory of negation. (Without that, it would be easy, though the whole character of the project changes thereupon.)

Finally, whether we have brought them quantificationally into coincidence or not, let us think about the relation between J and R as a vehicle for the development of particular theories. What is to be expected here, I think, is a core of coincidence, with increasing divergence on the periphery. The coincidental core again reflects points of agreement in the underlying theory of deduction; but divergence is to be expected, in the $A14$ case (for example), in how non-logical assumptions are spelled out. A nice result would lie in the isolation of sub-theories with respect to which J-formulations and R-formulations agree exactly, in the sense of having exactly the same theorems in these sub-theories. Such results might yield information from the R-viewpoint (e.g., that a constructive proof of $\exists X A X$ is available by a demonstration of AT for some term T, for certain choices of $\exists X A X$), while conferring relevant reliability on the J-viewpoint. (For a large class of theories, such isolation is trivial by the methods of [6], but it will not prove so in general. As an example here of an isolated sub-theory on which $J^{\sharp}$ and $R^{\sharp}$ agree, we can take the class of numerical equations, presumably, a result perhaps to be improved by isolating a somewhat larger and less trivial theory.)

Let us now take up this last point with particular reference to the relations between $J^{\sharp}$ and $R^{\sharp}$. First, it requires nothing more than $B27$ to verify all the number-theoretic axioms of J° *except* $A14$ on direct $R^{\sharp}$ translation. (This includes mathematical induction, given AMJ above.) So we may improve

Theorem 8 as follows.

Theorem 9. *Let A be any theorem of $\mathsf{J}°$ provable (e.g., in the system of [5]) without using the intuitionist form of $A14$. Then A is a theorem of $\mathsf{R}°$, and so of $\mathsf{R}^\sharp$.*

Proof. If A is a theorem of $\mathsf{J}°$, then there is a proof of it in $\mathsf{J}°$. On the proof of Theorem 8 and the observation just above, if $A14$ is not used in this proof then all axioms used in the proof are theorems of $\mathsf{R}^\sharp$, while the rules used in the proof are admissible in $\mathsf{R}^\sharp$ translation, thus completing the proof of Theorem 9. □

Theorem 9 is a handy vehicle for verifying particular number-theoretic principles in $\mathsf{R}^\sharp$, including many of those set out above. For example, general algebraic laws, of the sort reflected in Theorem 4, do not sensibly depend on the fact that the natural numbers start with 0, but that in counting them we never return to 0. Indeed, if we did return to 0, as in the natural numbers modulo some natural number, these algebraic laws would still hold. So it is quite unnecessary to prove such laws retail—e.g., that addition is associative, and the like—since we get them wholesale from Theorem 9. Of course, in particular cases we must back up our formal intuitions by actual $\mathsf{J}°$ proofs, before applying Theorem 9; usually, however, the intuitions will do, while the proofs may be taken from [5] and elsewhere.

The intuitionist form of $A14$ is unprovable in $\mathsf{R}^\sharp$, nor is the slack necessarily taken up by its relevantly valid counterpart. So there are some theorems of $\mathsf{J}°$ that are not theorems of $\mathsf{R}°$. However, as a simple corollary of Theorem 9, we have

Theorem 10. *Suppose that A is a theorem of $\mathsf{J}°$. Then $\forall x \neg x' = 0 \supset A$ is a theorem of $\mathsf{R}°$, and hence of $\mathsf{R}^\sharp$.*

Proof. Evidently A is a theorem of $\mathsf{J}°$ iff $\forall x \neg x' = 0 \supset A$ is a theorem of $\mathsf{J}°$, by the standard deduction theorem for $\mathsf{J}°$, and the $\mathsf{J}°$ form of $A14$. Moreover, $\forall x \neg x' = 0 \supset A$, since it has $A14$ in its antecedent, intuitionistically speaking, never requires $A14$ for its proof, whence the theorem follows immediately from Theorem 9, ending its proof. □

Theorems 9 and 10 show that $\mathsf{J}°$ is contained, in a well-defined way, in $\mathsf{R}°$. The containment, though not completely straightforward, would become so if we were to add (as we have contemplated but rejected) the intuitionist form of $A14$ as a new $\mathsf{R}^\sharp$ axiom. For reasons set out above, this is worth thinking

about, as a cancellation principle if not as a negation principle. And it may be that, for reasons of stability, we shall have to have this principle, in which case we shall (grudgingly) add it; but our story is that we do not need it, despite appearances to the contrary, for we have its effect in any case, in the sense already explained above.

We may now ask the converse question. To what extent is R° contained in J°? Had we chosen the alternative translation into $\mathsf{R}^\sharp$, based on $D\mathsf{F}_*$, $D\neg_*$, we could have immediately exhibited instances $\neg_* A \vee \neg_* \neg_* A$ of excluded middle that are intuitionistically unprovable but theorems of $\mathsf{R}^\sharp$. (E.g., choose a Gödel formula of the form $\neg A$ for J°. By the intuitionist disjunction property, if we could prove $\neg A \vee \neg\neg A$ in J° then we can prove one of $\neg A$, $\neg\neg A$, rendering J° inconsistent by Gödel's theorem. But $\neg_* A \vee \neg_* \neg_* A$ is a theorem scheme of $\mathsf{R}^\sharp$, whence, if J° is consistent, then some non-theorems of J° are taken into theorems of $\mathsf{R}^\sharp$ on the translation based on $\neg_*$.)

However, we have rejected the translation based on $\neg_*$ and have preferred the one based on $D\neg$. And at this point the relation between R° and J° becomes interesting. Forgetting about our reflections on $A14$, to what extent do the strong classical negation and quantificational principles of $\mathsf{R}^\sharp$ render provable theorems of first-order Peano arithmetic to which an intuitionist would object? Again, we are not concerned with theorems containing the classical negation $\sim$ to which an intuitionist might object; for he does not even recognize that connective. Rather, we are concerned with the impact of classical deductive methods on the theorems in what an intuitionist might concede to be, more or less, his vocabulary. While we are not particularly concerned with observing intuitionist strictures, it is not unlikely that, to some extent, we may observe them anyway, insofar as there are alternative constructive proofs of theorems more easily obtained using the classical principles recognized by $\mathsf{R}^\sharp$. It would be surprising if this always happened. But it must happen sometimes, and, having in $\mathsf{R}^\sharp$ a vocabulary which can be taken to comprehend both the intuitionist and the classical vocabulary, we can compare and contrast the proof procedures which $\mathsf{R}^\sharp$ makes available in each vocabulary.

VII

In the last section, we investigated some connections between intuitionist and relevant arithmetic. In this section, we do the same for connections between

classical and relevant arithmetic. We again begin with some convenient definitions.

$D\nabla$. $\quad \nabla A =_{\mathrm{DF}} A \wedge \top$

$D\triangle$. $\quad \triangle A =_{\mathrm{DF}} A \vee \bot$

These definitions are trivial, but useful. (Mnemonic: ∇ points down, like $\top$.) One thing that makes them useful is that ∇ has some of the properties associated with a modal necessity operator, while $\triangle$ is similarly linked to modal possibility. Let us also, while we are at it, add an intensional disjunction which will be dual to $\circ$ as $\vee$ is dual to $\wedge$.

$D+$. $\quad A + B =_{\mathrm{DF}} \sim A \to B$

$+$ does not enter often, while context serves to distinguish it from our arithmetic operation $+$. (But, for the record, note that our F_* above was just $\mathsf{F} + \mathsf{F}$.) Now let us note some rather trivial theorem schemes and admissible rules of $\mathsf{R}^\sharp$ for our newly defined operations ∇, $\triangle$.

$B28$. $\quad \nabla A \to A$

$B29$. $\quad A \;\Rightarrow\; \nabla A$

$B30$. $\quad A \to \triangle A$

$B31$. $\quad \nabla A \supset B \leftrightarrow \nabla A \to B$

$B32$. $\quad \nabla A \circ \nabla B \leftrightarrow \nabla A \wedge \nabla B$

$B33$. $\quad \triangle A + \triangle B \leftrightarrow \triangle A \vee \triangle B$

$B34$. $\quad \triangle(A \wedge B) \leftrightarrow \triangle A \wedge \triangle B$

$B35$. $\quad \nabla(A \vee B) \leftrightarrow \nabla A \vee \nabla B$

$B36$. $\quad \nabla(A \wedge B) \leftrightarrow \nabla A \wedge B$, etc.

$B37$. $\quad \triangle(A \vee B) \leftrightarrow \triangle A \vee B$, etc.

$B38$. $\quad \nabla\nabla A \leftrightarrow \nabla A$

$B39$. $\quad \triangle\triangle A \leftrightarrow \triangle A$

$B40$. $\quad \sim\nabla A \leftrightarrow \triangle\sim A$, etc.

$B41$. $\quad \nabla\forall X A \leftrightarrow \forall X \nabla A; \; \nabla\exists X A \leftrightarrow \exists X \nabla A$

$B42$. $\quad \triangle\forall X A \leftrightarrow \forall X \triangle A; \; \triangle\exists X A \leftrightarrow \exists X \triangle A$

$B43$. $\quad \nabla\triangle A \to \triangle\nabla A$

$B44.$ $\triangle\triangledown\triangle A \leftrightarrow \triangle\triangledown A;\ \triangledown\triangle\triangledown A \leftrightarrow \triangledown\triangle A$

$B45.$ $A \to B\ \Rightarrow\ \triangledown A \to \triangledown B;\ A \to B\ \Rightarrow\ \triangle A \to \triangle B$

$B46.$ $\triangle A \wedge \triangle(A \supset B) \to \triangle B$

$B47.$ $\triangledown A \to \triangle B \leftrightarrow \triangledown A \supset \triangle B$

$B48.$ $A \supset \triangle B \leftrightarrow A \supset \triangle B$

$B49.$ $\triangledown B \to \top$

$B50.$ $\bot \to \triangle B$

As a modal structure, it must be confessed that our $\triangledown$, $\triangle$ do not supply anything terribly interesting; for example, the non-equivalent proper modalities are just $\triangledown$, $\triangle$, $\triangledown\triangle$, $\triangle\triangledown$. Moreover, together with reasonable modal properties like $B28$, $B29$, $B30$, $B34$, $B35$, $B38$, $B39$, $B40$, we have strange ones like $B36$, $B37$. Indeed, we might note, for example, that all of $\triangle(A \supset B)$, $A \supset \triangle B$, $\triangledown A \supset B$, $\triangledown A \supset \triangle B$, $A \supset \triangle B$, and $\triangledown A \to \triangle B$ are provable $\mathsf{R}^\sharp$ equivalents, using $B37$, $B47$, $B48$ and trivial properties of the truth-functional connectives. Since our modalities arise in a truth-functional way this is not surprising; while that the structure is at all interesting rests on the relevant denial of the classical equivalences

$P1.$ $A \leftrightarrow \triangledown A$

$P2.$ $A \leftrightarrow \triangle A$

In fact, $P2$ is just the *ex falso quodlibet* in another form, while $P1$ is its dual.

     Although $P1$ and $P2$ do not *hold* in $\mathsf{R}^\sharp$, they offer an interesting way of *classifying* formulas. For we note, in the first place, that we do have in $\mathsf{R}^\sharp$ the theorem scheme

$B51.$ $T = U \leftrightarrow \triangledown(T = U)$

in virtue of $B19$ of Section V, $D\triangledown$, and elementary properties of conjunction. I.e., $P1$ holds for all *atomic formulas* of $\mathsf{R}^\sharp$. Immediately, by $B40$, $P2$ holds for all *negations* of atomic formulas; i.e.,

$B52.$ $T \neq U \leftrightarrow \triangle(T \neq U)$

is another theorem scheme of $\mathsf{R}^\sharp$.

The atomic formulas of $\mathsf{R}^\sharp$ are just the *equations*. $P1$, of course, is a theorem of $\mathsf{R}^\sharp$ for some formulas A which are not equations. Accordingly, let us call any formula A for which $P1$ is a theorem of $\mathsf{R}^\sharp$ a *secondary equation*. Similarly, inasmuch as $P2$ holds for all unequations, let us call any A for which it holds in $\mathsf{R}^\sharp$ a *secondary unequation*. It is evident and trivial that A is a secondary equation iff $A \to \top$ is a theorem of $\mathsf{R}^\sharp$; similarly, A is a secondary unequation iff $\bot \to A$ is a theorem of $\mathsf{R}^\sharp$.

Can any formula be *both* a secondary equation and a secondary unequation?

Observation. Some formula A is both a secondary equation and a secondary unequation iff $\bot \to \top$ is a theorem of $\mathsf{R}^\sharp$.

Verification. If $\bot \to \top$ is a theorem of $\mathsf{R}^\sharp$, $\top$ is both an equation and a secondary unequation. Conversely, if both $\bot \to A$, $A \to \top$ are theorems of $\mathsf{R}^\sharp$, so by transitivity of provable implication is $\bot \to \top$, completing the verification. □

We shall observe below that $\bot \to \top$ is not a theorem of $\mathsf{R}^\sharp$, whence the classes of secondary equations and secondary unequations are disjoint. Let us call the union of these classes the class of *secondary formulas*. Further question: is every formula a secondary formula?

Second Observation. Suppose that A, B are secondary equations, and that C is any formula. Then each of $A \wedge C$, $A \vee B$, $A \circ B$, $A + B$, $\forall X A$, $\exists X A$ is a secondary equation; moreover, $\sim A$, $A \supset C$ are secondary unequations.

Verification. By definition, the secondary equations A, B are equivalent respectively to $\triangledown A$, $\triangledown B$. Use $B36$, $B35$, $B32$, $B41$, $B40$ in the verification. For the $+$ case, we have $A \to \top$, $B \to \top$. $A + B \to \top + \top$ follows; moreover, by the reductio axiom and $D+$, $\top + \top \to \top$ is a theorem of $\mathsf{R}^\sharp$. Other details are left to the reader. □

Third Observation. Suppose that A, B are secondary unequations, and that C is any formula. Then each of $A \wedge B$, $A \vee C$, $A \circ B$, $A + B$, $\forall X A$, $\exists X A$, $C \supset A$ is a secondary unequation; so is $C \supset A$, while $\sim A$ is a secondary equation.

Verification. By definition, $A \leftrightarrow \triangle A$, $B \leftrightarrow \triangle B$ are theorems of $\mathsf{R}^\sharp$, whence particular verifications are dual to those in the last observation. □

Fourth Observation. Suppose that A, B are secondary formulas. Then so are $\sim A$, $A \wedge B$, $A \vee B$, $A \supset B$, $A \equiv B$, $\forall X A$, $\exists X A$.

Verification. Straightforward from previous observations. $\square$

Our observations, it will be noted, do not serve to answer the question that we asked earlier. But they answer it in part. Since the atomic formulas are secondary formulas, and since by our final observation the class of secondary formulas is closed under the truth-functions and quantifiers, it is clear that every formula which, from the viewpoint of $\mathsf{R}^\sharp$, is a *classical formula* is also a secondary formula.

Let us, to spell this out, characterize now the *classical sublanguage* $\mathscr{L}^*$ of $\mathscr{L}^\sharp$ as we previously characterized the intuitionist sublanguage $\mathscr{L}^\circ$. Specifically, $\mathscr{L}^*$ is got from $\mathscr{L}^\sharp$ by dropping $\rightarrow$ and all its works; i.e., $\mathscr{L}^*$ is the result of closing the set of atomic formulas under the connectives $\sim$, $\wedge$, and the quantifiers $\forall X$ and with such definitions as are applicable from these primitives remaining in force. P^* will be like $\mathsf{P}^\sharp$ in this vocabulary, and may be formulated by changing $\rightarrow$ to $\supset$ everywhere among the axioms and rules set out for $\mathsf{P}^\sharp$ above; in particular, on this formulation, P^* is equipped with a primitive rule $\supset E$. And R^*, as before, shall be the fragment of $\mathsf{R}^\sharp$ in the classical sublanguage; i.e., insofar as we may identify systems with the sets of their theorems and languages with the sets of their formulas, $\mathsf{R}^* = \mathscr{L}^* \cap \mathsf{R}^\sharp$. (We shall find fault with that identification below, but it is convenient for the present.)

P^*, clearly, is just like $\mathsf{P}^\sharp$, in two respects. First, its set of theorems is exactly the set of arrow-free theorems of $\mathsf{P}^\sharp$. Second, simply by trading in primitive $\rightarrow$ for defined $\supset$, the two systems are definitionally equivalent. On the other hand, R^* is merely the classical fragment of $\mathsf{R}^\sharp$, in what $\mathsf{R}^\sharp$ considers the classical vocabulary. And, as in the last section, we are faced with the question, "How classical is R^*?" And, we might add, "To what extent do classical properties of R^* rub off on the surrounding system $\mathsf{R}^\sharp$?"

In one direction, the rubbing off is obvious.

Theorem 11. *Suppose that A is a theorem of R^*. Then A is a theorem of P^*. I.e., all theorems of $\mathsf{R}^\sharp$ in the classical vocabulary are theorems of $\mathsf{P}^\sharp$.*

Proof. Trivial. We get $\mathsf{P}^\sharp$ from $\mathsf{R}^\sharp$ by adding an axiom. So theorems of $\mathsf{R}^\sharp$ are automatically theorems of $\mathsf{P}^\sharp$; the same holds for fragments R^*, P^*, ending the proof. $\square$

It would be nice to establish the converse of Theorem 11, by showing that all theorems of P^* are already theorems of R^*; i.e., to show that $\mathsf{R}^\sharp$ contains, as an exact subsystem in the classical sublanguage, classical first-order Peano arithmetic. It would be even nicer to find some good reason why this is not the case; e.g., that $0 = 1$ is provable in $\mathsf{P}^\sharp$, though not in $\mathsf{R}^\sharp$. However, we content ourselves here with another sort of exact translation from P^* to R^*, which will serve some though not all of the purposes of a demonstration that R^* and P^* coincide in the set of their theorems.

Theorem 12. *The following conditions are equivalent, for $A \in \mathscr{L}^*$.*

(i) A is a theorem of $\mathsf{P}^\sharp$

(ii) $\triangle A$ is a theorem of $\mathsf{R}^\sharp$

(iii) $\triangle A$ is a theorem of $\mathsf{P}^\sharp$

(iv) $\triangle \triangledown A$ is a theorem of $\mathsf{P}^\sharp$

(v) $\triangle \triangledown A$ is a theorem of $\mathsf{R}^\sharp$

(vi) $\triangledown \triangle A$ is a theorem of $\mathsf{P}^\sharp$

(vii) $\triangledown \triangle A$ is a theorem of $\mathsf{R}^\sharp$

Before proving Theorem 12, let us comment upon it. Its import is that, for every formula A which $\mathsf{R}^\sharp$ recognizes as classical, A is a theorem of classical first-order Peano arithmetic iff $A \vee \bot$ is a theorem of $\mathsf{R}^\sharp$. From the $\mathsf{R}^\sharp$ viewpoint, this involves a hedge; so to speak, whenever $\mathsf{P}^\sharp$ asserts A, $\mathsf{R}^\sharp$ asserts, on the translation set out in (ii), either that A or else $\bot$—i.e., or else $\mathsf{P}^\sharp$ is inconsistent with respect to negation. However, it may be that we can show also that

(viii) A is a theorem of $\mathsf{R}^\sharp$

is also equivalent, for all formulas A built out of the classical connectives and quantifiers, to each of (i)–(vii). If so we shall not, on a segment of the conventional wisdom not under present challenge, be able to demonstrate this equivalence, in general, by strictly finitary means. For, as we shall see, to be able to do so would be to provide a strictly finitary demonstration of the consistency of $\mathsf{P}^\sharp$ itself. (Caution: this last remark may seem incompatible with opening remarks; it is not, however, since Gödel's second theorem is

under attack here not with respect to its technical content for $\mathsf{P}^\sharp$ and related systems, but with respect to its purported philosophical significance. Put otherwise—there are deficiencies in the way that $\mathsf{P}^\sharp$ is formulated and conceived which lead to serious and unnecessary inflation of the philosophical claims that are associated with the Gödel result. The most important of these deficiencies lies in the confusion of a guarantee of negation-consistency with a guarantee of mathematical reliability. Opponents of Hilbert's program argued from the outset that the former guarantee was not *sufficient* for the latter guarantee. What is argued here is that it is not *necessary*; or, to put it otherwise, that there are various sorts of mathematical reliability, and various accompanying guarantees thereof. Gödel's result does rule out one kind of guarantee, for $\mathsf{R}^\sharp$ no less than for $\mathsf{P}^\sharp$, so far as I have been able to think the matter through. But it is sheer confusion to suppose that the guarantee ruled out is, in itself, all that important, or that guarantees of more fundamental sorts of reliability go down with the ship. To be sure, one may have these fundamental guarantees by simply omitting all but the most rudimentary logical machinery, or by translating them into far-fetched Formalese counterparts. (Cf. [1] on the latter point, even for negation-consistency.) The innovation here is the combination of strong, classically oriented logical machinery, straightforward Formalese translations of the fundamental guarantees, and finitary demonstration that these guarantees in fact hold.)

The important equivalence in Theorem 12 is between (i) and (ii); this is what enables us to isolate $\mathsf{P}^\sharp$ as a well-defined subsystem of $\mathsf{R}^\sharp$, and accordingly to apply to $\mathsf{R}^\sharp$ many well-known metamathematical results that have been obtained for $\mathsf{P}^\sharp$. But (iii)–(vii), though rather trivial, should not be overlooked either. For their import is that R^* and P^* do indeed coincide on formulas of the form $\triangle A$, $\triangledown\triangle A$, $\triangle\triangledown A$, or on formulas equivalent in $\mathsf{R}^\sharp$ to a formula of one of these forms. Thus, for example, what we have called secondary unequations, including in particular all formulas of the form $T \neq U$, which are expressible in the classical sublanguage $\mathscr{L}^*$ are theorems of $\mathsf{R}^\sharp$ iff they are theorems of $\mathsf{P}^\sharp$. In fact, we noted above that $\mathsf{R}^\sharp$, like $\mathsf{P}^\sharp$, was deficient with respect to the unequations themselves, in that there are some formulas $T \neq U$ that, on the standard mythology, are arithmetically valid, but which are unprovable in $\mathsf{R}^\sharp$. Given Theorem 12, we see now that $\mathsf{R}^\sharp$ is exactly as satisfactory, and exactly as deficient, in this respect as is $\mathsf{P}^\sharp$; for exactly the same unequations are theorems of both systems. We now turn to the actual proof of Theorem 12, which is by straightforward verification.

Proof of Theorem 12. By $P1$, $P2$, all of (i), (iii), (iv), (vi) are immediate

equivalents. By $B29$, (ii) implies (v); by $B43$, (v) implies (vii). By Theorem 11, (vii) implies (vi). So we may complete the proof of the theorem by showing that (i) implies (ii).

Suppose, then, that A is in the classical vocabulary $\mathscr{L}^*$, and that A is a theorem of $\mathsf{P}^\sharp$. Then A is a theorem of P^*, which we may take as formulated with all classical predicate tautologies in the language $\mathscr{L}^*$, and with the classical counterparts (with $\supset$ for $\rightarrow$) of our number-theoretic axioms $A11$–$A18$. As rules, we may take just $\supset E$, a simple rule $A \Rightarrow \forall X A$ of universal generalization, and RMI in the form $A0 \ \& \ AX \supset AX' \Rightarrow AX$. (This is an alternate formulation of P^*, designed to simplify the proof; its ineffective character is inessential, but, if the reader objects to it, he may easily transform the proof so that it applies to P^* as originally formulated above.)

We now show, by induction on the length of proof of A in P^*, that $\triangle A$ is a theorem of R^*, and so of $\mathsf{R}^\sharp$. Suppose, first, that A is a logical axiom; Anderson and Belnap show in [7] that all classical predicate tautologies are theorems of RQ, so that A is *a fortiori* a theorem of $\mathsf{R}^\sharp$; so then is $\triangle A$, by $B30$. Similarly, since $B \rightarrow C \rightarrow . B \supset C$ is a theorem of $\mathsf{R}^\sharp$, it is immediately verified that the P^* forms of $A11$–$A18$ are theorems of $\mathsf{R}^\sharp$, since they are either the same as, or follow trivially from, their $\mathsf{R}^\sharp$ counterparts.

Suppose, then, that A comes in P^* by predecessors by application of one of the rules $\supset E$, universal generalization, RMI, where we may suppose on inductive hypothesis that, for each premiss B of this application of the appropriate rule, $\triangle B$ is a theorem of $\mathsf{R}^\sharp$. There are three rules, so there are three cases.

Case 1. A comes by $\supset E$. On inductive hypothesis, there is a formula B of $\mathscr{L}^*$ such that $\triangle(B \supset A)$, $\triangle B$ are theorems of $\mathsf{R}^\sharp$. By $\wedge I$, $B46$ in $\mathsf{R}^\sharp$, $\triangle A$ is a theorem of $\mathsf{R}^\sharp$, as desired.

Case 2. A comes by universal generalization. Then A is of the form $\forall X B$, where by inductive hypothesis $\triangle B$ is a theorem of $\mathsf{R}^\sharp$. Universal generalization is admissible for $\mathsf{R}^\sharp$ also, whence $\forall X \triangle B$ is a theorem of $\mathsf{R}^\sharp$. By $B42$, $\triangle \forall X B$, which is $\triangle A$, is a theorem of $\mathsf{R}^\sharp$, as desired.

Case 3. A comes by the P^* form of RMI. Then A is of the form AX, where, on inductive hypothesis, $\triangle A0$ and $\triangle(AX \supset AX')$ are theorems of $\mathsf{R}^\sharp$. In the paragraph following $B28$–$B50$, we noted the equivalence of the latter to $AX \supset \triangle AX'$ in $\mathsf{R}^\sharp$, where a useful intuitionist-style $\supset$ has replaced

material $\supset$. Since $\bot \supset \triangle AX'$ by $B50$, $B27$, we may use Theorem 8 to reason intuitionistically to $AX \vee \bot \supset \triangle AX'$; i.e., $\triangle AX \supset \triangle AX'$. With $\triangle A0$, this sets up our relevantly valid intuitionist form RMJ of the rule of mathematical induction, whence we conclude $\triangle AX$ therefrom—i.e., $\triangle A$, as desired.

This completes the inductive argument, the proof that (i) implies (ii), and hence the proof of Theorem 12.

$\square$

Theorem 12 completes the proof that, in one satisfying sense, $\mathsf{R}^\sharp$ exactly contains the classical first-order Peano arithmetic $\mathsf{P}^\sharp$. But, as we have seen, this sense could be made more satisfying, by adding (viii) above to our collection (i)–(vii) of equivalent statements relating P^* theoremhood and R^* theoremhood. Can we add (viii)? Half the time, at least, to make official what was observed informally above.

Theorem 13. *Suppose that A is a secondary unequation in the language $\mathscr{L}^*$. Then A is a theorem of $\mathsf{P}^\sharp$ iff A is a theorem of $\mathsf{R}^\sharp$.*

Proof. Immediate from the equivalence of (i), (ii) of Theorem 12, and the definition of a secondary unequation. $\square$

What makes Theorem 13 interesting is Observation 4; every formula of $\mathscr{L}^*$ is either a secondary equation or a secondary unequation. Moreover, by the preceding observations, negation takes us from one of these classes to the other, making sense of the claim just preceding the theorem. And it turns out that a rather large number of formulas in which we are interested turn out to be secondary unequations, and so immediately are relevantly provable iff they are classically provable. Some examples are

$B53.$ $X \leq Y \wedge Y \leq X \supset X = Y$

$B54.$ $X \neq 0 \supset. XY = XZ \supset Y = Z$

$B55.$ $X' = 0 \supset 1 = 0$

$B56.$ $X + Y = 0 \supset X = 0 \wedge Y = 0$

$B57.$ $XY = 0 \supset X = 0 \vee Y = 0$

$B58.$ $XY + Z = XU + V \wedge Z < X \wedge V < X \supset Y = U \wedge Z = V$

$B59.$ $\forall Y (\forall Z (Z < Y \supset AZ) \supset AY) \supset AX$

$B60.$ $AX \supset \exists Y (AY \wedge \forall Z (Z < Y \supset \sim AZ))$

(To make it a consequence of Theorem 13, we must of course choose AX in the classical vocabulary in B59, B60.) I leave it to the reader to verify that each of B53–B60 is in fact both a secondary unequation and a theorem of $\mathsf{P}^\sharp$ (for the latter task, he may look it up in [5], as I did), whence $\mathsf{R}^\sharp$ theoremhood is immediate from $\mathsf{P}^\sharp$ theoremhood by Theorem 13.
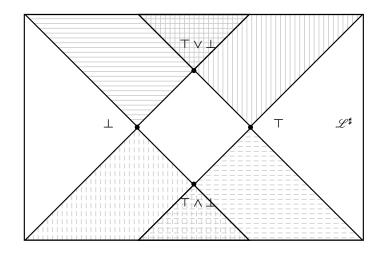
The particular formulas chosen have independent interest. They are paradigm cases for the claim that we can distinguish *materially true* from *lawlike* generalizations in arithmetic. I.e., that, starting from a narrow class of arithmetic laws that reflect rather directly the building up of the natural numbers from 0 under successor, we may establish anyway a wide class of arithmetic facts, which need not themselves be taken as laws (even though we may establish some other laws also). We have already discussed the cancellation principle B54, and in particular its specialization to B55, under this concern. So the failure of cancellation is redeemed after all; while B55 does not hold in $\mathsf{R}^\sharp$ in the form of a relevant, or even of an intuitionist, implication, it is materially true; so is the more general principle B54 to which one might appeal for its justification. Note also that only in the material sense does B55 have anything to do with negation, or with implying an absurdity, or the like.

Similar remarks can be made about the other principles listed. For example, B58 asserts the uniqueness of quotients and remainders, while B60 is the least number principle. It would be disturbing if such arithmetical facts could not be established, as we wend our way from generating considerations deeper and deeper into the standard mythology. But it does not seem to be so disturbing to do without some of the corresponding relevant implications; e.g., while 2 is in fact the least prime, it seems a bit superfluous to insist that one must be able to *infer* this fact, or even that there exists a least prime, from the primeness of 17.

Still, these are delicate questions. We can get a great deal of classical arithmetic relevantly. But Theorem 13 is still half a loaf, and we have only partial results concerning the other half—e.g., under Theorems 1, 3, and 4. We now outline three strategies for improving these results, both by making more efficient use of what we already know and by seeking new results not presently in hand.

First, in an algebraically familiar sense, the class of secondary formulas divides into the *ideal* of secondary equations and the *filter* of secondary unequations; let us call these, respectively, the $\top$-*ideal* and the $\bot$-*filter*. By definition, the $\top$-ideal consists exactly of those formulas of $\mathscr{L}^\sharp$ that provably

imply $\top$ in $\mathsf{R}^\sharp$, while the $\bot$-filter consists exactly of those formulas that $\bot$ implies. (For comparison, note that we are normally more interested in the opposite notions: the $\top$-*filter*, consisting of exactly the things that $\top$ provably implies, namely the *theorems* of $\mathsf{R}^\sharp$; and the $\bot$-*ideal*, consisting of just those formulas that provably imply $\bot$, namely the *formally refutable* formulas.) A snapshot is perhaps in order.
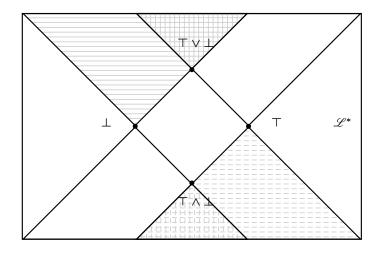

For purposes of the snapshot, let us suppose that provably equivalent formulas have indeed been identified (so that, e.g., the point marked '$\top$' may stand indifferently for all the formulas provably equivalent to $\top$: $0 = 0$, $1 = 1$, and so forth.) Put otherwise, we may suppose that our snapshot is not a snapshot of formulas but of arithmetic propositions, on the identification of formulas that (according to $\mathsf{R}^\sharp$) have the same arithmetic content. Then, in particular, the $\bot$-filter is marked by the solid horizontal lines; these indicate the secondary unequations; and the $\top$-ideal is marked by the dotted horizontal lines, indicating the secondary equations. Similarly, the theorems in the $\top$-filter are indicated by solid vertical lines, while the formally refutable propositions in the $\bot$-ideal are indicated by dotted vertical lines.

Built into our snapshot are 2 assumptions: (1) that $\bot$ does not imply $\top$, which is easy to show; (2) that $\top$ does not imply $\bot$, which embodies the faith that $\mathsf{R}^\sharp$ is negation-consistent; if this faith is wrong, $\bot$ gets identified with $\top \vee \bot$, and $\top$ with $\top \wedge \bot$, but otherwise the snapshot remains as is. (The corresponding $\mathsf{P}^\sharp$ snapshot has $\bot$ at the bottom, $\top$ at the top, and everything in both the $\top$-ideal and the $\bot$-filter, while both the $\top$-filter and

the $\bot$-ideal collapse to the single points $\top$ and $\bot$ respectively.)

Let us now specialize our snapshot to formulas of $\mathscr{L}^*$. These are the classical formulas and, by Observation 4, all of the associated propositions, being expressed by secondary formulas, are in the part of the snapshot marked by solid or dotted horizontal lines. I.e., specialized to classical propositions, the snapshot looks like this.

Note that, although there are no more vertical lines unaccompanied by corresponding horizontal lines, there remain horizontal lines unaccompanied by vertical ones. And this is, exactly, the Gödelian incompleteness of formal arithmetic; i.e., there are formulas (or propositions, if one prefers) that belong neither to the $\top$-filter nor to the $\bot$-ideal. Note moreover that, if a classical formula is a theorem of $\mathsf{R}^\sharp$, then either it is provably equivalent to $\top$, or else it belongs in the $\top \vee \bot$-filter; i.e., it's a theorem of R^* also. And this leaves as an open question whether, and to what extent, theorems of P^* that are in the $\top$-ideal are also theorems of R^*; note that such theorems of P^*, to be theorems of R^*, must be provably equivalent to $\top$ in $\mathsf{R}^\sharp$.

Some special interest attaches to formulas of the form $\triangle\triangledown A$, $\triangledown\triangle A$. Let us call any formula equivalent to one of the former a $\triangle\triangledown$-*formula*; to one of the latter, a $\triangledown\triangle$-*formula*. And, if a formula is either a $\triangle\triangledown$-formula or a $\triangledown\triangle$-formula, it will be a *strong secondary formula*. Note that, on the same considerations (and faith) as before, no formula is *both* a $\triangle\triangledown$-formula and a $\triangledown\triangle$-formula.

Let us think briefly about $\triangledown\triangle$-formulas. Expanding definitions, these are

of the form $\top \wedge (A \vee \bot)$. Evidently, $\top \wedge \bot$ is a least bound for $\triangledown\triangle$-formulas, provably implying them all; moreover, since $\top \wedge \bot$ is equivalent to $\triangledown\triangle\bot$, it is itself a $\triangledown\triangle$-formula. It is also clear that $\top$ is a greatest bound for $\triangledown\triangle$-formulas, being provably implied by them all in view of $B49$. Moreover, $\top$ is also a $\triangledown\triangle$-formula, being provably equivalent to $\triangledown\triangle\top$ by lattice properties. So, on our snapshots above of $\mathscr{L}^{\sharp}$ and $\mathscr{L}^{*}$, all $\triangledown\triangle$-formulas fall in the section of the snapshot bounded below by $\top \wedge \bot$, and bounded above by $\top$, inclusive. (However, it must not be assumed that $\triangledown\triangle$-formulas are *linearly ordered* by the relation of provable implication in $\mathsf{R}^{\sharp}$, as the diagram perhaps misleadingly suggests.)

We note the following additional theorem schemes of $\mathsf{R}^{\sharp}$.

$B61.$ $\triangle\triangledown\top \leftrightarrow \bot \vee \top$

$B62.$ $\triangle\triangledown\bot \leftrightarrow \bot$

$B63.$ $\triangledown\triangle\top \leftrightarrow \top$

$B64.$ $\triangledown\triangle\bot \leftrightarrow \bot \wedge \top$

$B65.$ $\triangle\triangledown A \wedge \triangle\triangledown B \leftrightarrow \triangle\triangledown(A \wedge B); \ \triangle\triangledown A \vee \triangle\triangledown B \leftrightarrow \triangle\triangledown(A \vee B)$

$B66.$ $\triangle\triangledown A \wedge \triangledown\triangle B \leftrightarrow \triangledown\triangle(A \wedge B); \ \triangle\triangledown A \vee \triangledown\triangle B \leftrightarrow \triangle\triangledown(A \vee B)$

$B67.$ $\triangledown\triangle A \wedge \triangledown\triangle B \leftrightarrow \triangledown\triangle(A \wedge B); \ \triangledown\triangle A \vee \triangledown\triangle B \leftrightarrow \triangledown\triangle(A \vee B)$

$B68.$ $\sim\triangle\triangledown A \leftrightarrow \triangledown\triangle\sim A; \ \sim\triangledown\triangle A \leftrightarrow \triangle\triangledown\sim A$

$B69.$ $\forall X \triangle\triangledown A \leftrightarrow \triangle\triangledown\forall X A; \ \exists X \triangle\triangledown A \leftrightarrow \triangle\triangledown\exists X A$

$B70.$ $\forall X \triangledown\triangle A \leftrightarrow \triangledown\triangle\forall X A; \ \exists X \triangledown\triangle A \leftrightarrow \triangledown\triangle\exists X A$

It is clear, then, that the class of strong secondary formulas is closed under the truth-functional connectives and quantifiers. Moreover, given $B61$, $B62$, $B68$, we can make exactly dual remarks about the $\triangle\triangledown$-formulas to those made above about the $\triangledown\triangle$-formulas; i.e., all $\triangle\triangledown$-formulas fall in the sections of our snapshots bounded below by $\bot$ and above by $\bot \vee \top$, inclusive. Note that these observations enforce strict limits upon the propositions expressed by strong secondary formulas that are theorems of $\mathsf{R}^{\sharp}$; for the $\triangledown\triangle$-theorems, being anyway secondary equations, must be provably equivalent to $\top$, while the $\triangle\triangledown$-theorems, being bounded above by $\bot \vee \top$, must in fact be equivalent to $\bot \vee \top$. Dual remarks apply to the formally refutable strong secondary formulas; so if a strong secondary formula is formally decidable in $\mathsf{R}^{\sharp}$, its associated proposition must lie at one of the endpoints of the admissible intervals; i.e., at one of $\bot, \top, \bot \vee \top, \bot \wedge \top$. And the formally undecidable

propositions, which Gödel assures us that we must have, on pain of negation-inconsistency, go between these endpoints when they are secondary formulas, of the strong sort.

Let us return, now, to our questions about the relative strength of R^* and P^*. Our main problem is, "Are there secondary equations that are theorems of P^* but not of R^*?" We now see that we shall answer this question in the negative if we can show that every secondary equation which is a theorem of P^* is also, according to $\mathsf{R}^\sharp$, a strong secondary formula—specifically, a $\triangledown\triangle$-formula. We know that this is the case for all equations, by Theorem 4; but we do not yet know whether it is the case for all secondary equations.

Meanwhile, let us note that our ruminations about strong secondary formulas give us another translation of P^* into R^*, on which P^* is an exact subsystem of $\mathsf{R}^\sharp$. We may define this translation ϕ recursively as follows, on all formulas A of $\mathscr{L}^*$.

(i) If A is an equation $A\phi$ is $\triangle A$

(ii) $(A \wedge B)\phi$ is $A\phi \wedge B\phi$

(iii) $(\sim A)\phi$ is $\sim A\phi$

(iv) $(\forall X A)\phi$ is $\forall X A\phi$

I.e., given a classical formula A, we get $A\phi$ simply by putting in $\triangle(T = U)$ for each atomic component $T = U$ in the formula A.

Theorem 14. *Let the translation ϕ be defined as above on $\mathscr{L}^*$. Then the following conditions are equivalent, for each formula A in $\mathscr{L}^*$.*

(i) *A is a theorem of $\mathsf{P}^\sharp$*

(ii) *$A\phi$ is a theorem of $\mathsf{P}^\sharp$*

(iii) *$A\phi$ is a theorem of $\mathsf{R}^\sharp$*

Proof. In view of $P2$, it is again trivial that (i) and (ii) are equivalent statements, for any A in $\mathscr{L}^*$. We shall finish the proof by establishing the equivalence of (ii) and (iii). First, it is clear that $A\phi$ must be a strong secondary formula. For, if A is atomic, $A\phi$ is of the form $\triangle(T = U)$, which is provably equivalent in $\mathsf{R}^\sharp$ to $\triangle\triangledown(T = U)$ by $B51$. So in this case $A\phi$ is a strong secondary formula. Otherwise, by definition of ϕ, $A\phi$ is of one of the forms $\sim B$, $B \wedge C$, $\forall X B$, where we may assume on a suitable inductive

hypothesis that each of B, C is a strong secondary formula, whence by $B65$–$B70$ we may conclude that $A\phi$ is a strong secondary formula.

Now, however, we are through. For, by clauses (iv), (v), (vi), and (vii) of Theorem 12, strong secondary formulas of $\mathscr{L}^*$ are theorems of $\mathsf{R}^\sharp$ iff they are theorems of $\mathsf{P}^\sharp$. This establishes the equivalence of (ii) and (iii) of the theorem, and ends its proof.

$\square$

Theorem 14 is in one sense an improvement on Theorem 12, in that, unlike the translations directly suggested by that theorem, our translation ϕ is a *homomorphism* from the algebra of formulas of $\mathscr{L}^*$ into itself, preserving both theoremhood and non-theoremhood in P^* on translation into R^*. Accordingly, it is convenient for direct application of some of the Gödel arguments and results to $\mathsf{R}^\sharp$, as we shall see in the next section. And it is particularly useful in that it is an effective containment of P^* in $\mathsf{R}^\sharp$; for even if, as we hope, P^* and R^* coincide in their theorems, we shall give reasons below to suspect that any purported demonstration thereof is likely to contain objectionable fideistic elements. But Theorem 12 remains in other respects our best result, since it localizes all possible differences between P^* and R^* in a mere disjoined $\perp$, otherwise leaving the internal structure of formulas to be compared alone.

Let us return now to the consideration of secondary formulas in general; on analogy with the above, it will be convenient to identify the secondary equations in the $\top$-ideal as ∇-formulas, and the secondary unequations in the $\perp$-filter as $\triangle$-formulas. We now lay down a simple test that will enable us to distinguish ∇-formulas from $\triangle$-formulas in concrete cases.

Specifically, we define a *canonical assignment* α, which will be a partial function from the set of formulas of $\mathscr{L}^\sharp$ to the set $\{+1, -1\}$, using $\wedge$ for arithmetic minimum, $\vee$ for arithmetic maximum, and $-$ for arithmetic inverse, as follows:

(i) $\alpha(A) = -1$, if A is an atomic formula $T = U$

(ii) $\alpha(A \wedge B) = \alpha(A) \wedge \alpha(B)$

(iii) $\alpha(\sim A) = -\alpha(A)$

(iv) $\alpha(A \vee B) = \alpha(A) \vee \alpha(B)$

(v) $\alpha(A \supset B) = -\alpha(A) \vee \alpha(B)$

(vi) $\alpha(\forall X A) = \alpha(A)$

(vii) $\alpha(\exists X A) = \alpha(A)$

(viii) $\alpha(A \circ B) = \alpha(A)$, if $\alpha(A) = \alpha(B)$

(ix) $\alpha(A + B) = \alpha(A)$, if $\alpha(A) = \alpha(B)$

(x) $\alpha(A \rightarrow B) = \alpha(B)$, if $\alpha(A) \neq \alpha(B)$

Strictly speaking, the reader should regard these specifications as holding only where primitive connectives are involved, but he can quickly check that they work out as indicated on defined connectives also. Where α is not defined by the above specifications, e.g., on $A \rightarrow B$ when $\alpha(A) = \alpha(B)$, it is to be regarded as undefined (subject to some further refinements that we shall note below). Note, however that α is defined on all formulas of $\mathscr{L}^*$, and that it works truth-tabularly on these formulas (on, so to speak, an initial "false" assignment to the atomic formulas). We use, however, '+1' instead of 't,' and '-1' instead of 'f,' since otherwise the fact that $\alpha(\top)$ is -1 and $\alpha(\bot)$ is $+1$ would be confusing. The utility of the canonical assignment α lies in the following theorem.

Theorem 15. *Let A be any formula of $\mathscr{L}^\sharp$, and let α be the canonical assignment. Then*

(i)   *if $\alpha(A) = -1$, then A is a ∇-formula*

(ii)   *if $\alpha(A) = +1$, then A is a $\triangle$-formula*

Suppose moreover that A is a formula of $\mathscr{L}^*$. Then $\alpha(A)$ is defined. Moreover,

(iii) *$\alpha(A) = -1$ iff A is a ∇-formula*

(iv)   *$\alpha(A) = +1$ iff A is a $\triangle$-formula*

(v)   *if $\alpha(A) = +1$, then A is a theorem of $\mathsf{P}^\sharp$ iff A is a theorem of $\mathsf{R}^\sharp$*

Proof. We prove first (i) and (ii), by induction on the complexity of A. Note that α is defined on A only if α is defined on all subformulas of A. If A is atomic, (i) holds trivially by definition of α and $B51$. Suppose now that (i) and (ii) hold for B, C, on inductive hypothesis, and that A is of one of the forms $B \wedge C$, $\sim B$, $\forall X B$, $B \rightarrow C$. Then we may rely on the definition of α and the second and third observations preceding Theorem 11 to establish that (i) and (ii) hold for A also, thus completing the inductive proof of (i) and (ii).

On the further supposition of the theorem, A is a (classical) formula of $\mathscr{L}^*$. By clauses (i), (ii), (iii), (vi) in the definition of α, $\alpha(A)$ is defined. Then (iii), (iv) follow from (i), (ii), and our promise to prove that no $\triangle$-formula is a $\triangledown$-formula, which accompanied the initial observation of this section. And (v) then follows immediately from (iv) and Theorem 13, ending the proof of Theorem 15. $\qquad\square$

So, we have a simple truth-tabular check, on formulas of $\mathscr{L}^*$ in particular, that will enable us to decide immediately, for a given classical formula A, whether or not it is a $\triangle$-formula. This is a great aid in applying Theorem 13. Let us borrow again from [5] to illustrate the method, which the reader may then apply to our previous borrowings. While we are at it, let us introduce the defined relation *divides* by

$D|.\quad T|U =_{\mathrm{DF}} \exists X (TX = U)$

And we may note that $T|U$, like $T \leq U$, $T < U$, etc., is a secondary equation, getting the value -1 on the canonical assignment α. Here are two of Kleene's theorems involving $|$.

$B71.\quad X > 1 \supset \sim(X|Y \wedge X|Y')$
$B72.\quad Y \neq 0 \supset (X|Y \supset 0 < X \leq Y)$

If Kleene can prove them, so can we. E.g.,

$$
\begin{aligned}
\alpha(X > 1 \supset \sim(X|Y \wedge X|Y')) \quad &= \alpha(X > 1) \supset -\alpha(X|Y \wedge X|Y') \\
&= -1 \supset +1 \\
&= +1
\end{aligned}
$$

$$
\begin{aligned}
\alpha(Y \neq 0 \supset (X|Y \supset 0 < X \leq Y)) \quad &= \alpha(Y \neq 0) \supset. \; \alpha(X|Y) \supset \alpha(0 < X \leq Y) \\
&= +1 \supset. \; -1 \supset -1 \\
&= +1 \supset +1 \\
&= +1
\end{aligned}
$$

So, since $B71$, $B72$ are formulas of $\mathscr{L}^*$ and theorems of $\mathsf{P}^\sharp$, they are theorems of $\mathsf{R}^\sharp$ by (v) of Theorem 15.

This concludes discussion of our first strategy for making better use of

the results in Theorems 12 and 13—namely, attending with care to the properties in $\mathsf{R}^\sharp$ of our "modalities" $\triangle$, ∇, $\triangle\nabla$, $\nabla\triangle$. While, pursuing this strategy, we have found an efficient translation of P^* into $\mathsf{R}^\sharp$, and efficient proof procedures in $\mathsf{R}^\sharp$ for many of the theorems of P^*, the strategy must be deepened considerably before we may hope to prove, if we can prove, that the theorems of R^* and P^* coincide. Let us look briefly at two other strategies that may yield these results.

One way to show that R^* and P^* coincide is to find some formulation of P^* all of whose axioms are theorems of $\mathsf{R}^\sharp$ and whose rules are admissible in $\mathsf{R}^\sharp$. Essentially, this is the strategy by which Anderson and Belnap showed, as reported in [7], that the theorems of first-order relevant logics such as RQ, in the truth-functional connectives and quantifiers, are exactly the predicate tautologies of the first-order classical logic PQ. A Gentzen consecution calculus formulation of PQ, for example, works very nicely for this purpose. E.g., consider Gentzen's system LK, as set out in [8]. When interpreted in RQ, the axioms of this system become theorems $A \supset A$ of RQ, while the structural and operational rules of LK do become admissible rules of RQ. Accordingly, since all predicate tautologies are theorems of LK, simply by mimicking their Gentzen-style normal form proofs in RQ, we can prove them in this system as well. (Indeed, such considerations are already part of Ackermann's seminal [25], which may be taken to have created relevant logic.)

The nice thing about choosing a Gentzen formulation of PQ is that it avoids difficulties about cut, since this rule may be proved to be admissible in LK without having been taken as primitive. This is important, because the cut problems show up in RQ as problems about the admissibility of γ—i.e., of *modus ponens* for material implication. In fact, the problem of γ for RQ was not solved until [26], long after the proof strategy just summarized had disposed of the question whether RQ exactly contained PQ.

The alluring possibility is that this kind of normal-form strategy will also work to show that P^* is exactly contained in $\mathsf{R}^\sharp$. However, a formulation of P^* which is cut-free in the necessary sense is hard to come by. For Gentzen-style work on formal arithmetic, in particular, has most often been fashioned to the task of proving that formal arithmetics are contradiction-free, and only in a subsidiary sense to the task of normalizing proof procedures. For the former purpose, a simple strategy (employed, in effect, by Schütte in [27]) is to apply normalized proof procedures to some non-finitary system P^{**} that contains P^*; one then argues, in a partially effective way, that P^{**} is consistent, from which it follows that its subsystem P^* is consistent.

Obviously, however, that sort of result isn't going to help us here. To the extent that we believe any sort of non-finitary argument that establishes that $\mathsf{P}^\sharp$ is free of contradiction, we shall automatically believe that $\mathsf{R}^\sharp$ is free of contradiction, by Theorem 11. But a central reason for our interest in $\mathsf{R}^\sharp$ is that we can establish its reliability in ways that are not only strictly finitary but which are comprehensible to any diligent third-grader. If, accordingly, we are going to attend to non-finitary arguments after all, we shall do so only for the purpose of showing that R^* is *as strong as* P^*. In short, as Dunn observes, we need appeal to classical faith for rather more restricted purposes in the R^* framework than in the P^* framework. $\mathsf{P}^\sharp$, without classical faith, is utterly useless and unreliable. $\mathsf{R}^\sharp$, without classical faith, is ordinary basic arithmetic, with perhaps some logical anomalies that do not infect fundamental arithmetic insights; with classical faith, it lacks, we hope, the anomalies, too; but that must be proved.

The third strategy to be considered is quite simple; we can just prove γ. In fact, we can *express* γ for $\mathsf{R}^\sharp$ in either of the following two equivalent forms:

$\supset E.\quad A \supset B\ \&\ A\ \Rightarrow\ B$

$\triangle E.\quad \triangle A\ \Rightarrow\ A$

The equivalence between $\supset E$ and $\triangle E$, in the sense that one of these rules is admissible in $\mathsf{R}^\sharp$ iff the other one is, is quickly established. In the first place, suppose that $\triangle E$ is admissible, and that $A \supset B$, A are both theorems of $\mathsf{R}^\sharp$. By $B30$, $B46$, so is $\triangle B$. By $\triangle E$, B is then a theorem. So if $\triangle E$ is admissible so is $\supset E$. Conversely, suppose $\supset E$ admissible. Let $\triangle B$ be a theorem of $\mathsf{R}^\sharp$. By definitions, this is almost immediately equivalent to $\top \supset B$. But $\top$ is a theorem, whence B by $\supset E$. This completes the argument that $\supset E$ and $\triangle E$ march together; henceforth, we indifferently refer to either as γ, in conformity to an old Ackermann tradition.

Theorem 16. *Suppose that γ is an admissible rule of $\mathsf{R}^\sharp$. Then*

(i)   *Classical arithmetic P^* is exactly contained in $\mathsf{R}^\sharp$ on the direct transformation.*

(ii)   *For no formula A are both A and $\sim A$ provable in $\mathsf{R}^\sharp$; i.e., $\mathsf{R}^\sharp$ is negation-consistent.*

(iii) *Classical arithmetic P^* is negation-consistent.*

Proof. (i) is immediate from Theorem 12; for if A is a theorem of P^* then $\triangle A$ is a theorem of $\mathsf{R}^\sharp$; given $\triangle E$, then A is a theorem of $\mathsf{R}^\sharp$. The converse being trivial by Theorem 11, if γ holds for $\mathsf{R}^\sharp$ then P^* and R^* have the same theorems.

(ii) requires our proof below that something, say F, is a non-theorem of $\mathsf{R}^\sharp$; that, as we have repeatedly noted, is trivial. Suppose nevertheless that, for some A, each of A, $\sim A$ is a theorem of $\mathsf{R}^\sharp$. But, since $\mathsf{R}^\sharp$ contains classical *sentential* logic, $A \wedge \sim A \supset \mathsf{F}$ is a theorem of $\mathsf{R}^\sharp$, whence so by γ is F, which is impossible. So, given γ, $\mathsf{R}^\sharp$ is negation-consistent.

(iii) is immediate from (i) and (ii), ending the proof of Theorem 16. □

In principle, Theorem 16 has been known ever since [28], at least. Indeed, it has even been known, thanks to Dunn, that if one added a sentential variable p to the vocabulary of $\mathsf{R}^\sharp$, p would remain unprovable in the system thus extended. The innovation here, accordingly, is the discovery in the arithmetic vocabulary itself of a statement, e.g., $0 = 1$, that is surely unprovable, thus permitting proof of (ii) for $\mathsf{R}^\sharp$ itself, together with a concrete formulation and study of $\mathsf{R}^\sharp$ that makes it at least plausible that γ should hold. But it must not be thought that proving γ, and hence finding a new approach to an old classical arithmetic, is the chief aim of this study. While I conjecture that γ holds for $\mathsf{R}^\sharp$, I should not be all that shocked, and might even be pleased, if it fails; all would depend on *how* it fails. For our main thesis, after all, is that classical Peano arithmetic is misleadingly formulated, in a way that gives rise to pseudo-problems, and that the *assumption* of γ, which confuses the myths that we are creating and the laws that we use to create them, is from the outset at the heart of the resulting philosophical mess.

In certain respects, the technical claims that are being made, or conjectured, here are by no means new. Kreisel, for example, notes in [29] that the demonstration of the cut-rule for the kinds of Gentzen-style formulations of type theory introduced by Takeuti means that these systems, whose negation-consistency is elementary, are exempt in one sense from the purported Gödel restrictions against internal proofs of their own consistency. And Kreisel draws also the correct conclusion that what cannot then be elementary for these systems is the proof of the admissibility of the cut-rule itself.

However, there is this difference between the approach being taken here and what one may take to be the standard story about Gentzen-style consecution calculi. As noted, Gentzen-style concerns for Cut show up as relevant

concerns for γ. But, though some of Gentzen's own remarks tend to the conclusion that Cut is *philosophically* superfluous, one does not take this as the common opinion. Indeed, Gentzen introduces Cut in [31] as a primitive rule that is clearly justified on his intended interpretation of his calculi, but which, almost miraculously, one can do without in actually formulating some of these calculi.

On this point, the relevant logics have been distinctive almost from the beginning: ever since, in fact, Anderson and Belnap reformulated Ackermann's Π' by dropping γ to get the system E of entailment, for reasons set out in [7] and [32]. For γ, it must be seen, is positively unwanted as a first principle—not, indeed, on the superficial ground that it leads to paradoxes of implication (in fact, in Ackermann's system, it does *not* lead to such paradoxes), but because it creates a philosophical mishmash, which makes almost inevitable the replacement of reason with mythology in, say, discussion of the foundations of mathematics. And here we must refute also another platitudinous lie; namely, that the paradoxes of implication only infect inconsistent systems, in which we are anyway uninterested. For it is not, so far as we know, *actual inconsistency* that raises Gödelian problems for $\mathsf{P}^\sharp$ and related systems, but the *threat of inconsistency*. On account of γ, there is no way in $\mathsf{P}^\sharp$ to *localize* this threat; accordingly, when formalized in the $\mathsf{P}^\sharp$ way, all of arithmetic stands under the *same* threat. When given a metaphysical cast, the result is either an unwarranted scepticism or an unwarranted fideism; either all of mathematics stands in epistemological peril, or else it is all to be taken on speculative trust, grounded perhaps on insights, intuitions, perceptions, and the like that are not further reducible to rational analysis.

The choice is false, because the dichotomy is false. In mathematics, as elsewhere, some propositions are more certain than others. And it was just wrong to suppose that mathematics was given *analytically*, as a domain of human inquiry somehow exempt from those considerations that govern rational men and women in their other pursuits. Rather, as elsewhere, of the more certain propositions of mathematics we can be more certain, and that certainty is unaffected by the confusions into which we *might* fall as we seek to explore the mathematical Universe ever more widely. We know that intuitively; indeed, we know that we can prove, say, that $2 + 2$ is 4 in systems much weaker than $\mathsf{P}^\sharp$, whose consistency is not in doubt even if $\mathsf{P}^\sharp$ turns out to be inconsistent. But it is just silly that we don't know it formally; to that degree, $\mathsf{P}^\sharp$ is a pretty *bad* formalization of the intuitive situation, for it suggests that we entertain doubts that we do not in fact entertain; and it

makes these suggestions, not ineluctably, but out of defects in its own logical machinery.

We spoke, at the outset, of *repealing* Gödel's theorem. But that last discussion was pretty Gödelian, in that we allowed (as a consequence of Gödel's first theorem) that not all of mathematics was equally certain. In objecting, however, to the customary interpretation of Gödel's second theorem, we are merely drawing the obvious corollary of the first theorem: not all of mathematics is equally dubious. But the exact effect of the customary interpretation—for any particular formal system (e.g., $\mathsf{P}^\sharp$) is in fact to make all of the mathematics formalized therein equally dubious; it is as though, when we begin to do analysis, we can no longer trust arithmetic. And the whole burden of the present argument will be recognized when it is seen that *extension* of previous insight ought no more, systematically, to put previous insights in peril by stuffing the universe with theorems incompatible with those insights than it should put these insights in peril by failing to preserve the old theorems that express them. E.g., everybody recognizes that, when we move up to analysis, we are still going to be able to prove that $2 + 2$ is 4; how silly, though, to suppose that there will be a new threat, which we cannot effectively guard against, that we shall prove also $2 + 2 = 5$. That's analytic number theory, with a vengeance.

So, while many of the claims that have been made here could already have been made, on technical grounds, for known systems, it has not been the custom to make them. For it is rather difficult to suppose that γ is a bad rule when the material $\supset$ is the only stand-in for implication that one's system offers. And so, e.g., cut-free systems that lack γ get viewed, in effect, as aberrations, which escape Gödelian constraint (in one direction) only because they formulate artificially what would be subject to that constraint when formalized naturally. To the contrary, the view here (as in [7]) is that mixing up implication with $\supset$ is itself unnatural, and that it might be expected to have anomalous consequences that are not part of Nature but which result instead from egregious error. It was after all a *thesis*, put forth gingerly by Frege and not part of the mainstream of logical analysis until Whitehead and Russell had championed it in *Principia Mathematica*, that the mathematical sense of 'implies' could be naturally reduced to the fragment of that sense preserved in the material $\supset$. The ink was hardy dry on the pages of *PM* before that view came under philosophical challenge by C.I. Lewis and others. But, curiously, it has been taken for granted by almost everybody (save Brouwer and his school, perhaps, though the Heyting $\supset$, anyway, is

open to the same sorts of complaint that are lodged against ⊃), that these philosophical objections had much to do with formalized mathematics. Indeed, formalized mathematics, on the whole, has proceeded as though mere philosophical objections to reigning dogma could not possibly be relevant to the mathematical enterprises within which that dogma is congealed. Even stranger is the fact that the philosophers who made the objections have also treated actual systems of formalized mathematics as though the fact that, according to these philosophers, these systems are based on a wrong notion of implication made no difference to the mathematical content of the systems, or even to the ways in which this content may be accurately formulated. In short, there are serious allegations of error, extant now (in various forms) for over 60 years, which those who claim to have found the error feel it unnecessary to bring to the attention of those who are making the error, and which those who are purportedly making the errors feel safe in ignoring completely; indeed, the prosecution in this case tends to congratulate itself on making its implication "extensionally intelligible," as though a necessary condition for mounting a *challenge* to the role that the material ⊃ plays in systems like P♯ lay in expressing that challenge in terms of the material ⊃ itself. (The → of R, incidentally, is "extensionally intelligible," as reported in [16], lest suspicious readers suppose that R is indifferent to such sociologically necessary conditions as it is its mission to destroy.)

So the entire matter, to say the least, has been very strange. And Anderson and Belnap were quite right in locating a central issue (though not necessarily *the* central issue) in the problem of γ. For, rather generally, one might as well allow that any connective in the implicative family for which, under any circumstances, one will admit *modus ponens* counts indeed as an implication. On this allowance, there is also considerable practical benefit in taking as a *preferred* implication the *weakest* connective for which *modus ponens* is on general grounds admissible. This favors ⊃; moreover, as a practical matter, it must also be allowed that *modus ponens* for ⊃ accords with our usual intuitions and our common practice. If the resultant paradoxes of material implication do not accord either with our usual intuitions or our common practice, there are grounds for puzzlement, but there are not grounds for cancelling our ordinary understandings.

Rather, however, we must ask ourselves, "Why do we hold our common opinions?" And it is not fair to discount some of our common opinions, as all parties to the dispute are fond of doing, so that the others may be held without restriction. In fact, it is just by discounting our common opinion

that we can localize contradiction that we may read into Gödel's second theorem metaphysical speculation which is no proper part of it. But we must give its due to *modus ponens* for material implication also. That due, as all recognize, is at root truth-functional; we are not prepared for A to be true, and for either A to be false or B to be true, without also taking it for granted that B is true. But the difficulty is, despite p. 1 of the elementary text, that *truth* and *falsity* do not offer a sufficiently intelligible framework upon which to build the theory of inference (but it is, I agree, the best place to start the logic book). Especially in mathematics, where we are creating our truths as we go along, *modus ponens* for material implication exposes us to all the difficulties that are wrapped up in formal negation, and builds in from the outset a significant part of the standard mythology that we are in the process of creating. Small wonder, then, that only mythological proofs are available to guarantee the reliability of systems formulated with this rule primitive, however reliability is characterized.

But there is a middle way between what seem to be clashing intuitions: namely, that γ is after all truth-preserving and hence is to be desired as a primitive principle of inference on the one hand; and that, on the other hand, even the threat of contradiction is enough to void all finitary guarantees of reliability, if γ is accepted as a rule. It is just the solution that one would expect, on the Dunn-style motivating considerations urged above. For, when one thinks about it, it is γ itself which lacks the wanted finitary guarantee; its semantical justification is that it preserves truth, but—in arithmetic, anyway—the notion of truth itself is highly ineffective. About particular truths, on the other hand, we can be particularly certain. So let us not mix up the ineffective question of γ for arithmetic with the class of questions—like those involved in the multiplication table—to which we can give immediately effective answers.

In short, as suggested in [28], let us make the admissibility of γ for a system a *normality condition* on that system, not a condition which is built in with the primitive rules. We should not expect, in general, to settle effectively the question whether or not a system is normal. The semantic intrusions into this question are too deep, and, at best, we shall ordinarily have to settle for one of the watered-down criteria of effective proof to which we are condemned on philosophically correct application of Gödelian constraint. I.e., it is not unreasonable that we shall have to answer ineffective questions ineffectively, and even that we should sometimes have to settle for negative answers to such questions where we had hoped for positive ones. But it is most unreasonable

that we should have to answer effective questions ineffectively. And it is, moreover, intolerable that formalized mathematics itself should have been erected on a foundation of cascading doubt, and that this should have come to be taken as the normal state of the project, without a close look at some of the philosophical presuppositions that have been uncritically absorbed into that project.

In leaving the question of γ open for $\mathsf{R}^\sharp$, let us note that, while an affirmative solution would *suffice* to establish (i)–(iii) of Theorem 16, it is not *necessary* for these results. In fact, if $\mathsf{R}^\sharp$ is negation-*inconsistent*, it is already negation-inconsistent in its classical part R^*. This is a direct consequence of the fact that $A \wedge \sim A \to \bot$ is a theorem of $\mathsf{R}^\sharp$, whence, if $\mathsf{R}^\sharp$ is inconsistent, both $\bot, \top$ will be theorems of its classical part R^*.

Accordingly, to establish (ii) and (iii) of Theorem 16, it suffices to establish (i), which, as suggested earlier, may be an easier task than finding a proof of γ for all of $\mathsf{R}^\sharp$. For that matter, it is conceivable that γ fails for $\mathsf{R}^\sharp$, while holding when restricted to its classical part R^*. (On the other hand, in view of Theorem 12, that γ holds when restricted to R^* is completely equivalent to the assertion that P^* and R^* coincide, while continuing to imply directly the consistency of these systems.)

VIII

In this section, we shall apply the results of the last to do some Gödelian metatheory for $\mathsf{R}^\sharp$. We shall do so in a sketchy and intuitive way, making full use of the standard mythology (including the standard mythology of recursive function theory, including Church's thesis) as is convenient.

In one sense, setting out many of the standard metamathematical results for $\mathsf{R}^\sharp$ is superfluous; e.g., simple incompleteness, recursive undecidability, and so forth follow immediately from the corresponding results for $\mathsf{P}^\sharp$. Nonetheless, since we claim that $\mathsf{R}^\sharp$ dodges the second Gödel theorem in important respects, it is of some interest to attend to the ways in which its machinery reflects the standard mythology on these matters. We shall rely on our $\mathsf{P}^\sharp$ translations in so attending, but we shall take our metamathematical notions to be those appropriate to $\mathsf{R}^\sharp$, introducing corresponding notions for $\mathsf{P}^\sharp$ only for purposes of comparison and contrast.

First, for the purposes of this section, we are going to introduce a multiple-use program for natural numbers. We have already reserved natural numbers

themselves for use, essentially, as *names* of numbers—i.e., as numerals. There is no harm in using them also as *formulas*; since we are relying, as set down early in I, on an effective coding which associates with each natural number I a unique corresponding formula A^I, we may just as well take the formula A^I *to be* the number I. (With inessential differences, this is the plan of [1]; and, for that matter, Gödel's original motivating plan, near enough.) Note that this does not get in the way of the previous identification of numerals and numbers, since numerals and formulas belong to distinct syntactic categories.

We may also associate natural numbers, in an effective many-one way, with the recursively enumerable sets of natural numbers. Our plan will be *system-dependent*, being relativized for present purposes to $\mathsf{R}^\sharp$, but applying *mutatis mutandis* to $\mathsf{P}^\sharp$, P^*, etc. Specifically, each natural number I is, on present conventions, also a formula $A^I x$, where x is as above the first variable of $\mathscr{L}^\sharp$. On our substitution conventions, $A^I J$ is the result of replacing all free occurrences (if any) of x with J in $A^I x$; we shall always specialize those conventions to replacement of x in this section. Accordingly, let us define a function r, on $\mathbf{N}$ with subsets of $\mathbf{N}$ as values, by setting, for each $I \in \mathbf{N}$, $J \in \mathbf{N}$

(i) $J \in r(I)$ iff $\vdash_{\mathsf{R}^\sharp} A^I J$

Let now $\mathbf{S}$ be any set of natural numbers, and that $\mathbf{S} = r(I)$ for some $I \in \mathbf{N}$. Then we say that $\mathbf{S}$ is *weakly expressed by the number I* (indifferently, the formula A^I, the formula $A^I x$), or that I (or A^I, etc.) *weakly expresses* $\mathbf{S}$. A set $\mathbf{S}$ is *weakly expressible* in $\mathsf{R}^\sharp$ just in case there is an $I \in \mathbf{N}$ such that A^I weakly expresses $\mathbf{S}$. *Mutatis mutandis*, we speak of weak expressibility of sets $\mathbf{S} \subseteq \mathbf{N}$ in $\mathsf{P}^\sharp$, etc. The following theorem is evident.

Theorem 17. *Let $\mathbf{S}$ be any set of natural numbers. Then the following conditions are equivalent.*

(a) $\mathbf{S}$ *is recursively enumerable*
(b) $\mathbf{S}$ *is weakly expressible in* $\mathsf{P}^\sharp$
(c) $\mathbf{S}$ *is weakly expressible in* P^*
(d) $\mathbf{S}$ *is weakly expressible in* $\mathsf{R}^\sharp$

Proof. The equivalence of (a) and (b) is well-known (e.g., from [5]). (c) merely restates (b). We show that (d) implies (a), and that (c) implies (d),

to complete the proof.

First, suppose that **S** is weakly expressible in $\mathsf{R}^\sharp$. Then there is some formula $A^I x$ that expresses **S** in $\mathsf{R}^\sharp$; i.e., $\mathbf{S} = r(I)$, and (i) holds, for this choice of I. But the set of theorems of $\mathsf{R}^\sharp$ is clearly recursively enumerable, and one may use this enumeration to construct a recursive enumeration of the theorems of $\mathsf{R}^\sharp$ of the particular form $A^I J$, for the I selected above. But this amounts evidently to a plan on which **S** itself may be recursively enumerated, whence (d) implies (a).

Finally, suppose that **S** is recursively enumerable, and so that **S** is weakly expressed by some formula Ax in P^*. Then, by Theorem 14, $Ax\phi$ will weakly express **S** in $\mathsf{R}^\sharp$. Thus (c) implies (d), ending the proof of Theorem 17. □

Let us adapt [33] to call a formula Ax a *class formula* provided that (1) Ax is in the classical sublanguage $\mathscr{L}^*$, (2) x occurs free in Ax, and (3) no variable distinct from x occurs free in Ax. Then we may add to the equivalent conditions (a)–(d) of the theorem,

(e) **S** is weakly expressible in $\mathsf{R}^\sharp$ by some class formula.

(f) **S** is weakly expressible in P^* by some class formula.

For it is trivial that (c) may be specialized to (f), whence the argument from (c) to (d) may be directly transformed into an argument from (f) to (e). We take these added equivalents, (e) in particular, as part of the content of Theorem 17.

Let us use Theorem 17 to establish the existence of a *Gödel formula* for $\mathsf{R}^\sharp$, which may be taken to assert its own unprovability in $\mathsf{R}^\sharp$. In the first place, we form a *Gödel set* **G** of natural numbers by setting, for each $J \in \mathbf{N}$,

(ii) $J \in \mathbf{G}$ iff $J \in r(J)$

I.e., by (i),

(iii) $J \in \mathbf{G}$ iff $\vdash_{\mathsf{R}^\sharp} A^J J$

On past considerations, **G** is evidently a recursively enumerable set, whence, by our addendum (e) to Theorem 17, there is a class formula Gx which weakly expresses **G** in $\mathsf{R}^\sharp$; i.e., for all $J \in \mathbf{N}$,

(iv) $J \in \mathbf{G}$ iff $\vdash_{\mathsf{R}^\sharp} GJ$

Consider now the formula $\sim Gx$. This has a Gödel number; indeed, on our conventions, it *is* a number, which may replace x in the class formula $\sim Gx$. I.e.,

(v) Let $\sim G\sim Gx$ be the Gödel formula.

Despite appearances, the Gödel formula is notationally well-formed on our conventions, as just observed. Then, by (iii)

(vi) $\sim Gx \in \mathbf{G}$ iff $\vdash_{\mathsf{R}^\sharp} \sim G\sim Gx$

But by (iv),

(vii) $\sim Gx \in \mathbf{G}$ iff $\vdash_{\mathsf{R}^\sharp} G\sim Gx$

Evidently, by (vi) and (vii),

(viii) $\vdash_{\mathsf{R}^\sharp} G\sim Gx$ iff $\vdash_{\mathsf{R}^\sharp} \sim G\sim Gx$

So, clearly, from (viii),

(ix) If $\mathsf{R}^\sharp$ is negation-consistent, neither the Gödel formula nor its negation is a theorem of $\mathsf{R}^\sharp$.

Moreover,

(x) If $\mathsf{R}^\sharp$ is negation-consistent, $\sim G \notin G$,

from (ix) and either of (vi), (vii).

Inasmuch as we are assuming the standard mythology, in this section, we have

(xi) $\mathsf{R}^\sharp$ is negation-consistent

whence

(xii) $\sim Gx \notin \mathbf{G}$ & not $\vdash_{\mathsf{R}^\sharp} \sim G\sim Gx$ & not $\vdash_{\mathsf{R}^\sharp} G\sim Gx$

(xii) establishes, on an absolutely standard argument, that a standard Gödel

formula for $\mathsf{R}^\sharp$ has the expected properties. Let us take a peek behind the mythological facade to see what substantive assumptions we are making. We may isolate the following.

(A) Every recursively enumerable set is weakly expressible in the classical arithmetic P^*

(B) The set of theorems of $\mathsf{R}^\sharp$ of the form AJ, for each fixed formula Ax, is recursively enumerable

(C) $\mathsf{R}^\sharp$ is negation-consistent

To these we might add our underlying assumption of an effective correspondence between natural numbers and formulas, our abstract treatment of formal systems in general, and the naive comprehension axiom for sets of natural numbers that enters into (ii). However, none of these latter assumptions is essential to the treatment, though some vestige of them will presumably remain on any treatment. So far as the explicit assumptions (A)–(C) are concerned, we note that (B) hardly counts as an assumption at all; for there is, in fact, no real appeal to Church's thesis in (B), since, with a little patience, we can actually construct a recursive sequence of theorems (i.e., numbers) AJ, for fixed Ax, which will include all such theorems in $\mathsf{R}^\sharp$. As for (C), it is already implied by (A). For if $\mathsf{R}^\sharp$ is negation-*inconsistent*, so is P^*. But then all formulas of P^* are theorems, whence the only recursively enumerable set weakly expressible in P^* is $\mathbf{N}$ itself. Surely there are others–e.g., the set of odd numbers. So if we believe (A), we already believe (C).

Ought we to believe (A)? Well, maybe not. Even to understand (A), we have to know what a recursively enumerable set of natural numbers is. While we don't need the standard mythology for this purpose—a constructive mythology, for example, will also make sense of this notion—it is unlikely that we can make this notion mythology-free. More to the immediate point, the *truth* of (A), even if it is allowed that it is tolerably clear, requires certain assumptions on P^* that cannot be effectively guaranteed. Gödel's assumption in [33], it will be recalled, was that P^* is ω-consistent.

Let us, however, not nitpick with the standard mythology, or seek out degrees of credibility within it, in this section. For we are merely seeking out here the metamathematical properties of $\mathsf{R}^\sharp$ from a standard point of view; there are, I suspect, deeper anomalies to be uncovered yet in P^* and its kin, including $\mathsf{R}^\sharp$, but our purpose here is to find terrain—for $\mathsf{R}^\sharp$, anyway—into which these anomalies cannot penetrate. Meanwhile, there is no doubt on

the standard mythology that $\mathsf{R}^\sharp$, P^*, etc., are consistent, ω-consistent, and what you will, settling for the moment any difficulties with (A).

In looking at the above argument, it is to be noted that the role of the negation $\sim$ of $\mathsf{R}^\sharp$ is in fact very limited. What has happened simply, in Curryesque terms, is that we have found a way to secure within $\mathsf{R}^\sharp$ the effect of the *fixed point* (or *paradoxical*) combinator of [20]. Any other definable unary connective in $\mathsf{R}^\sharp$ (or, for that matter, definable unary function on natural numbers, since formulas are just natural numbers) may be applied to the same effect. For let σ be any such connective, and consider the class formula $\sigma G x$. The reasoning so far as (viii) proceeds in exactly the same way, yielding the following generalized counterpart of (viii).

(xiii) $\vdash_{\mathsf{R}^\sharp} G\sigma G x$ iff $\vdash_{\mathsf{R}^\sharp} \sigma G\sigma G x$

If the conjunction of the two sides of (xiii) together imply something objectionable in $\mathsf{R}^\sharp$, then it had better be the case that neither of them is a theorem. Let us use, for our generalized negation σ, σ-consistent and σ-complete in the obvious sense. Then, on the obvious generalization of (ix), for arbitrary σ,

(xiv)   If $\mathsf{R}^\sharp$ is σ-consistent, then $\mathsf{R}^\sharp$ is σ-incomplete

I.e., if for no sentence A of $\mathscr{L}^\sharp$ are both A and σA theorems of $\mathsf{R}^\sharp$, then for some sentence B neither B nor σB are theorems of $\mathsf{R}^\sharp$; in fact, we may choose B as $G\sigma G x$.

To apply (xiv), let us set

(a) $\sigma_0 A =_{\mathrm{DF}} \neg A$
(b) $\sigma_1 A =_{\mathrm{DF}} A \to \forall y \exists z (y = 2z)$
(c) $\sigma_2 A =_{\mathrm{DF}} A \to. \perp \to \top$
(d) $\sigma_3 A =_{\mathrm{DF}} A \to \top$

In each of the cases (a)–(c), we have a finitary proof of the σ-consistency of $\mathsf{R}^\sharp$; e.g., since we can refute $\exists z(1 = 2z)$, we may be assured that for no A are both A and $\sigma_1 A$ theorems of $\mathsf{R}^\sharp$, and similarly in cases in (a) and (c). Thus there is a sentence, by (xiv), which is itself unprovable, and which does not provably imply $\forall y \exists z (y = 2z)$. Similarly, $\mathsf{R}^\sharp$ is incomplete with respect to the intuitionist negation $\neg$, without invoking in a formal proof of this fact in $\mathsf{R}^\sharp$

with respect to its own negation $\sim$.

This much, however, we may view with equanimity. For each of (a)–(c) invoke an incompleteness that, since $\rightarrow$ is a *relevant* implication, we might well expect. (Note, incidentally, that, as we should also expect, σ-consistency is not necessary for σ-incompleteness. E.g., $\mathsf{R}^\sharp$ is σ_3-inconsistent, in that both $\top$ and $\sigma_3\top$ are theorems; but it is also σ_3-incomplete, we trust, having neither $\bot$ nor $\sigma_3\bot$ as theorems.) For there are some formulas that fail to imply other formulas, irrespective of the truth or the falsity of the components of such an implication; there is no reason, e.g., why we should expect $\mathsf{R}^\sharp$ to assert at least one of $\bot$, $\bot \rightarrow. \bot \rightarrow \top$. So far as $\mathsf{R}^\sharp$ is concerned, both are false, and hence both *deserve* to be unprovable.

There is a bit more bite in the original Gödel result, which we have not yet come to. For the Gödel formula is distinguished not only by being formally undecidable, but by being *true*. And we may well ask whether our generalized Gödel formulas, constructed using arbitrary σ, likewise give rise to true but unprovable formulas.

By way of comment on this point, there are 2 more segments of the conventional wisdom on Gödelian topics that are worth thinking about. One line of this wisdom, suggested in [33] itself and worked out in considerable detail by Wang, is that it is the *semantic* paradoxes that Gödelian arguments rest upon, and that moreover one may choose his favorite semantic paradox (Gödel liked Richard's, with a dose of The Liar) and develop therefrom Gödelian consequences. It is a little hard to see, however, what is particularly *semantic* about the argument above. In fact, if anybody's paradox seems especially germane to it, it is Russell's; **G**, on our multiple use program for natural numbers, is just the collection of r.e. sets that belong to themselves. We do our best, in constructing our Gödel formula, to say that the complement of this set does not belong to itself; we are saved, we hope, from the Russellian contradiction only because this complement it not itself r.e. In fact, if the argument has a semantic component, it seems to lie in our inability to *say* what we intended to say; as we noted essentially in Section I, $\sim Gx$ does *not* express, even weakly, the complement of **G**. At the very least, the *line* between set-theoretic and semantic paradox, which almost all have respected since Ramsey in [34] gave invalid arguments for drawing it, does not seem so clearly drawn after all, once we have got use to Gödel-style coding, which disrupts the categorical distinctions that semantics intends.

And this brings us back to a more central segment of the conventional wisdom on these matters, which we have alternately asserted and viewed

with misgiving. For what warrant do we have, anyway, to think that the Gödel formula is *true* though unprovable, or that, intensionally, it asserts *its own* unprovability in $\mathsf{R}^\sharp$? Here, we must keep in mind, and keep distinct, 2 levels of interpretation, which are often meshed together. One level of interpretation is merely with respect to the natural numbers, as we pointed out in I. The classical part $\mathscr{L}^*$ of $\mathscr{L}^\sharp$, in which we are doing our metatheory, makes presumptively straightforward assertions about the natural numbers, which are to be taken as either true or false, when sentences of $\mathscr{L}^*$, in the standard model $\mathbb{N}$. From this viewpoint, we have indeed produced a true but unprovable formula; for both of $G{\sim}Gx$ and ${\sim}G{\sim}Gx$ are unprovable, and exactly one of them is true in the standard model. We are not yet, however, in a position to say which one, while we are in absolutely no position to say that either of them has, on these considerations, any metamathematical *meaning* whatsoever.

Metamathematical meaning is provided by our second level of interpretation; namely, we *picked* Gx because it weakly expresses the set $\mathbf{G}$. On this level of interpretation, GJ "says" that the formula A^J does in fact belong to the set $\mathbf{G}$; so, presumably, ${\sim}GJ$ *denies* that A^J belongs to $\mathbf{G}$. Well, ${\sim}Gx$, in particular, does *not* belong to $\mathbf{G}$, as we saw in (xii). So, on the second level of interpretation, ${\sim}G{\sim}Gx$ is true. Moreover, on this level, it is reflexively true; when $\mathbf{G}$ is unpacked, the formula asserts its own unprovability.

There is, however, a *gap* between these two levels of interpretation. In one sense, this gap can easily be filled; in another, it is unfillable. For, thus far, nothing prevents ${\sim}G{\sim}Gx$ from being *false* on our first level, as a sheer statement of arithmetic, while being *true*, so to speak, on our second, since it still correctly asserts its non-theoremhood in $\mathbf{G}$. And, in the classical context, anyway, we can *insist* upon its truth, even if it is arithmetically false, by taking it as an extra axiom; since its negation is not classically provable, by the completeness theorem of Gödel for first-order classical logic, no new inconsistency will result by making this move.

That such a *clash of interpretations* might arise, however, is a consequence of our not having chosen Gx very carefully. We asked only that it weakly express $\mathbf{G}$, and were satisfied with any Gx (the first, say, in an enumeration of class formulas) that had this property, having been assured by Theorem 17 that some Gx weakly expresses $\mathbf{G}$. We can make, however, a tighter choice of Gx by tightening our characterization of expressibility.

Specifically, let $\mathbf{S}$ be a set of natural numbers, and let $A^I x$ be a class formula of $\mathscr{L}^*$. Then we shall say that $A^I x$ *truly expresses* $\mathbf{S}$ provided that

$r(I) = \mathbf{S}$ (i.e., that condition (i) above holds, for all J in $\mathbf{N}$, for $\mathbf{S}$ in place of $r(I)$), and moreover, for all $J \in \mathbf{N}$,

(xv) $\vdash_{\mathsf{R}^\sharp} A^I J$ iff $\vDash_{\mathbb{N}} A^I J$

where, for each formula B in the classical sublanguage $\mathscr{L}^*$, $\vDash_{\mathbb{N}} B$ means that B is *semantically valid* (i.e., *true*, if B has no free variables) in the standard model $\mathbb{N}$ of the natural numbers.

We note now the following:

Theorem 18. *Let $\mathbf{S}$ be any set of natural numbers. Then the following conditions are equivalent.*

(a) $\mathbf{S}$ *is recursively enumerable*

(g) $\mathbf{S}$ *is truly expressible in* P^* *by some class formula* $A^I x$

(h) $\mathbf{S}$ *is truly expressible in* $\mathsf{R}^\sharp$ *by a class formula* $A^I x$ *of* $\mathscr{L}^*$

Proof. Again, the equivalence of (a) and (g) is well-known, given the standard proofs of Theorem 17 and the standard mythology. And (g) implies (h) by hauling out once more the ϕ translation from P^* to $\mathsf{R}^\sharp$, while (h) implies (a) by Theorem 17, ending the proof of Theorem 18. $\qquad\square$

So we may partly fill the gap between our two levels of interpretation, by choosing Gx so that it truly expresses $\mathbf{G}$ in $\mathsf{R}^\sharp$. Then, in particular, $G{\sim}Gx$ is arithmetically true iff it is provable in $\mathsf{R}^\sharp$; as we know, it is unprovable, and hence false. Accordingly, our Gödel sentence ${\sim}G{\sim}Gx$ is, as desired, a standard arithmetical truth which is unprovable in $\mathsf{R}^\sharp$.

What happens on our alternative choices of σ; e.g., σ_1? Even though $\neg Gx$ is outside of the classical sublanguage $\mathscr{L}^*$ (for which truth in $\mathbb{N}$ makes immediate mythological sense), $G\neg Gx$ is, as a statement of the form GI, within that sublanguage. Choosing Gx to truly express $\mathbf{G}$, $G\neg Gx$, since it is unprovable, is likewise false in $\mathbb{N}$, whence ${\sim}G\neg Gx$ is true. Is then, on a suitable extension of our semantic ideas, $\neg G\neg Gx$ likewise true but unprovable? There is no reason to think so; for it is not ruled out that both of $G\neg Gx$ and $\neg G\neg Gx$ are unprovable falsehoods.

This, however, underlines remarks made above about the insufficiency of formal $\sim$ to intuitive 'not' in formalized arithmetic, whatever our sincere intentions. For nothing prevents us from taking $\neg$, rather than $\sim$, as the negation with respect to which we characterize truth in the standard model

$\mathbb{N}$. If so, $\neg G \neg Gx$ gets promoted from unprovable falsehood to unprovable truth. In short, we cannot measure the success of our formalism on sincere intentions alone; for these intentions may *admit* favored interpretation, but they cannot *compel* them. And, in fact, no theorems of $\mathsf{R}^\sharp$ will turn out false if we favor $\neg$ over $\sim$ as our preferred negation; we may wince at some of the *non-theorems* of $\mathsf{R}^\sharp$ on this perverse preference, which neither classicist nor even intuitionist could approve.

Wincing aside, we are stuck with the same problems with our preferred classical $\sim$. In fact, there are not two levels of interpretation induced by thinking about $\mathsf{R}^\sharp$, but three. For true-in-$\mathsf{R}^\sharp$, if it means anything (except by mythological consent), must surely mean provable-in-$\mathsf{R}^\sharp$. On this perspective $\sim G \sim Gx$ is as false-in-$\mathsf{R}^\sharp$ as is $\neg G \neg Gx$. As verification, we may, after all, reverse our previous move by adding $G \sim Gx$ consistently to $\mathsf{R}^\sharp$; this upsets our intention that $\mathsf{R}^\sharp$ should have the standard model $\mathbb{N}$, but, systematically, it does not upset $\mathsf{R}^\sharp$ at all. I.e., even after we have moved up to true-expressibility of the r.e. sets, we are still stuck with the non-categoricity of the first-order Peano postulates.

So our move up to true expressibility is not as successful as we might have hoped. On a sheer extensional level, to be sure, it is quite satisfactory; $\sim G \sim Gx$ is true-in-$\mathbb{N}$ while unprovable-in-$\mathsf{R}^\sharp$, contrary to what was hoped and expected before [33]. But this continues to leave out our intensional level of interpretation, on which $\sim G \sim Gx$ asserts-truly-that-it-is-unprovable-in-$\mathsf{R}^\sharp$. At best, we have so far merely an extensional coincidence. I.e., $\sim G \sim Gx$ is arithmetically-true-in-$\mathbb{N}$, and moreover $\sim G \sim Gx$ is true-in-$\mathbb{N}$ iff $\sim G \sim Gx$ is unprovable in $\mathsf{R}^\sharp$. But, clearly, any unprovable arithmetic truth in $\mathsf{R}^\sharp$ will satisfy these criteria, at least in a material sense.

So let us not be too sure that our move from weak expressibility to true expressibility was an improvement. The former, at any rate, is syntactically clear, and involves less mythology; moreover, it does not require us to look over our shoulder at intended extensions of $\mathsf{R}^\sharp$ (given that $\mathsf{R}^\sharp$ itself, like any formal system, cannot distinguish internally its intended from its unintended extensions). And our real problem, even in an extensional sense, is that our success in expressing $\mathbf{G}$ in $\mathsf{R}^\sharp$ does not imply success in expressing the complement of $\mathbf{G}$; indeed, if $\mathsf{R}^\sharp$ is consistent, it implies failure, not only in $\mathsf{R}^\sharp$ but in all of its negation-consistent and recursively axiomatized extensions. (Note that we do not solve this problem simply by taking $\sim G \sim Gx$ as an extra axiom; for this still leaves infinitely many sentences of the form $\sim GI$ unprovable, though true; i.e., we are in no better shape than before to suppose

that $\sim GI$ *says*, except on extrinsic mythology (that may be controverted by extending $\mathsf{R}^\sharp$ non-standardly), that $I \notin \mathbf{G}$.)

Again, the strong suggestion is that it is not *arithmetic* which is here to blame, but logical particles like $\sim$. Similar problems, lest it seem like negation is getting too much of the heat, are raised by the quantifiers. More interesting here, they are raised also by our relevant implication $\rightarrow$, as we saw above on our σ-generalizations of the basic Gödel-style argument. In the $\rightarrow$ case, reflected in the specialization to the $\neg$ case above, we found a way out, in that, for general σ, nothing prevents A and σA from both being false. What hurts, then, where $\sim$ is essentially involved is our *semantic* faith that both of A, $\sim A$ shall not turn out false together. And what gets mixed up here is that we tend to wish for $\sim$ both a *fact-expressing* and an *inferential* role, as we have seen; once material implication has been dumped, on the other hand, we can be content to make the inferential role primary for $\rightarrow$, allowing simply that some arguments are bad arguments regardless of the truth-values of premiss and conclusion.

With respect to $\sim$, on the other hand, we are stuck. If A is bad, $\sim A$ is good, as a matter of semantic faith, whether we can prove it or not. Note, incidentally, the concentration in the above argument in showing that $G{\sim}Gx$ is bad. And so it is sometimes said that the import of Gödel's incompleteness theorem is that we cannot encompass, in the same system, all intuitively valid proof procedures. But that, I think, is an overstatement, at least as it involves formal negation. As $\sim$ enters into *arguments*, its role remains steadfastly inferential; the semantic faith embodied in our understanding of negation is not part of the argument, save as this faith is codified in particular inferential principles. Indeed, it is to be seriously doubted (even in the first-order case, despite the mythology attached to the classical completeness proof) whether *there exists* a totality of valid arguments. Again, we must not confuse semantic faith with deductive reason. Faith is not a proof procedure, valid or otherwise; rather it embodies our guiding ideals, to which we seek to make rational proof procedures conform.

So, while we may invoke guiding ideals to *supply* new proof procedures—indeed, Gödel supplied one, and it counts these days as a *proof* of $\sim G{\sim}Gx$, for suitably chosen Gx—this is a little different from supposing that such proof procedures already exist, in embryo, in virtue of the guiding ideals themselves. And, again, the confusions engendered by material implication cut hard against this point. For the consistent error of modern mathematical logic has been to assume that true (or valid) material implications already

license some admissible (or universally admissible) proof procedure. This, combined with our semantic faith in the truth or falsity of atomic propositions, and in truth-functional compounding of such propositions, does suggest that all the good arguments already *exist*, and that logic is an exercise in *finding* them (blighted, alas, by the fact that there is no effective way to do so, or, past first-order logic, even to tell what they are). The view, indeed, seems to be that God keeps most of the good arguments for himself, but that, in his infinite kindness, he lets us find a few of them.

God, as I have noted elsewhere, has no need of any arguments, even good ones. So it is about time to abandon theological validity as a criterion of good reasoning. Our semantic faith is ours, and it does not come from above; the proof procedures that we admit are ours also, and it is sufficient for the *invalidity* of an alleged proof procedure that we do not admit it; the angels need not be consulted, nor is it relevant that, in a better world, we would admit it. So, when we turn our semantic faith upon our stock of undecidable propositions, it is not that there has been all along an intuitively valid proof procedure which has been thus far overlooked; it is that we have invoked our guiding ideals to *enlarge* our stock of proof procedures. There is danger in this course; guiding ideals do not come with divine guarantees, either, and the enlargement may lead to something silly. But it is especially silly to build the semantic faith so deeply into the proof procedures that the entire deductive enterprise is in constant danger of going completely wacky. And we must not suppose that our guiding ideals themselves—e.g., the negation-completeness of ideal mathematical theories—are themselves capable of proof. That much we should have learned from Gödel. For what, after all, might count as a proof, save as it rests on more comprehensive guiding ideals.

Let us sum up the situation thus far in a theorem.

Theorem 19. $\mathsf{R}^\sharp$ *is negation-incomplete. In particular, we may choose the Gödel sentence $\sim G \sim Gx$ so that it is arithmetically true, and so that, in an intensional sense, it may be taken to express its own unprovability. Neither $\sim G \sim Gx$ nor $G \sim Gx$, when the Gödel sentence is so chosen, is a theorem of $\mathsf{R}^\sharp$. Similarly, letting σ be any defined 1-place connective on formulas of $\mathsf{R}^\sharp$, if $\mathsf{R}^\sharp$ is σ-consistent then neither $G\sigma Gx$ nor $\sigma G \sigma Gx$ is a theorem of $\mathsf{R}^\sharp$, and $\sim G\sigma Gx$ is arithmetically true.*

Proof. Subject to the mythological assumptions, standard results, and philosophical caveats declared above, this theorem already stands proved. $\qquad\square$

We shall also draw some of the usual corollaries. Henceforth, we use Rx for a class formula that weakly expresses the set of theorems of $\mathsf{R}^\sharp$. (Since this set is r.e., there are many formulas that will do this job; for the moment, any one of them will do.) We use Px to weakly express the set of theorems of $\mathsf{P}^\sharp$, as above. And we suppose that each of Rx, Px (and also Gx, when we are interested in it) is a sentence $A\phi$, where ϕ is as in Theorem 14; evidently we may do this without loss of generality, and we reap the benefit, by Theorem 14 and consideration of ϕ, that these formulas then weakly express the *same* sets in both $\mathsf{R}^\sharp$ and $\mathsf{P}^\sharp$.

We have noted the following two theorems, but they are worth making explicit.

Theorem 20. $\mathsf{R}^\sharp$ *is negation-consistent in sense* (6) *of II iff it is negation-consistent in sense* (5). *I.e.,* $\perp$ *is a theorem of* $\mathsf{R}^\sharp$ *iff, for some formula A, each of A,* $\sim A$ *is a theorem of* $\mathsf{R}^\sharp$. *Moreover, this equivalence is effective, in the sense that proof of a contradiction in* $\mathsf{R}^\sharp$ *may be trivially lengthened to become a proof of* $\perp$, *and vice versa.*

Proof. Trivial, since $A \wedge \sim A \rightarrow \perp$ is a theorem scheme of $\mathsf{R}^\sharp$, while $\sim\perp$ is a theorem. $\qquad\square$

Henceforth, we shall take *negation-consistent* to mean indifferently consistent-in-sense-(5) or consistent-in-sense-(6) of II, either for $\mathsf{R}^\sharp$ or for $\mathsf{P}^\sharp$; we shall prefer the former, as somewhat simpler. We also note [TYPESCRIPT ENDS.]

References

[1] S. Feferman. Arithmetization of metamathematics in a general setting. *Fundamenta Mathematicae*, 49:35–92, 1960.

[2] K. Gödel. On formally undecidable propositions of *Principia Mathematica* and related systems I. In S. Feferman, J.W. Dawson, S.C. Kleene, G.H. Moore, R.M. Solovay, and J. van Heijenoort, editors, *Kurt Gödel: Collected Works*, volume I, pages 145–195. Oxford University Press, Oxford, 1986.

[3] C. Reid. *Hilbert*. Springer, New York, 1970.

[5] S. C. Kleene. *Introduction to Metamathematics*. North-Holland, New York, 1952.

[8] H. B. Curry. *Foundations of Mathematical Logic.* McGraw Hill, New York, 1963.

[19] A. R. Anderson and N. D. Belnap. Enthymemes. *Journal of Philosophy*, 58(23):713–723, 1961.

[23] A. Urquhart. Semantics for relevant logics. *Journal of Symbolic Logic*, 37(1):159–169, 1972.

[24] J. Norman and R. Sylvan. *Directions in Relevant Logic.* Kluwer, Boston, 1989.

[25] W. Ackermann. Bedründung einer strengen Implikation. *Journal of Symbolic Logic*, 21:113–128, 1956.

[26] J. M. Dunn, H. Leblanc, and R. K. Meyer. Completeness of relevant quantification theories. *Notre Dame Journal of Formal Logic*, 15(1):97–121, 1974.

[29] G. Kreisel. A survey of proof theory. *Journal of Symbolic Logic*, 33(3):321–388, 1968.

[31] G. Gentzen. Über die Existenz unabhängiger Axiomensysteme zu unendlichen Satzsystemen. *Mathematische Annalen*, 107:329–350, 1932.

[33] K. Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38:173—198, 1931.

[34] F. P. Ramsey. The foundations of mathematics. *Proceedings of the London Mathematical Society*, s2-25(1):338–384, 1926.