LEON HORSTEN

IN DEFENSE OF EPISTEMIC ARITHMETIC*

ABSTRACT. This paper presents a defense of Epistemic Arithmetic as used for a formalization of intuitionistic arithmetic and of certain informal mathematical principles. First, objections by Allen Hazen and Craig Smorynski against Epistemic Arithmetic are discussed and found wanting. Second, positive support is given for the research program by showing that Epistemic Arithmetic can give interesting formulations of Church's Thesis.

1. INTRODUCTION

This paper presents a defense of Epistemic Arithmetic as used for a formalization of intuitionistic arithmetic and of certain informal mathematical principles such as Church's Thesis.

First, I discuss an objection by Craig Smorynski to the effect that Epistemic Arithmetic is unable to capture the full flavor of talk about effective methods, which plays an essential role in the interpretation of intuitionistic mathematics and in the expression of Church's Thesis. This amounts to the objection that the content of Church's Thesis and the meaning of constructivistic arithmetic are *insufficiently analysed* in Epistemic Arithmetic. Second, I discuss an objection by Alan Hazen to the effect that Epistemic Arithmetic does not respect the anti-realist spirit of intuitionism.

The objections by Smorynski and by Hazen are found wanting, but there is a third objection which poses a real difficulty for the project – viz., that the epistemic systems are not given a sufficiently determinate interpretation to enable us to evaluate proposed formalizations, e.g., of Church's Thesis. But I show that some recent developments of Epistemic Arithmetic mitigate this problem to a certain extent. Also, a direction for further progress is proposed in the form of a *theory of presentations* of mathematical objects.

In the final section I turn to a detailed scrutiny of this objection as it bears on Church's Thesis. I find that some formalizations proposed so far are indeed defective, not because they are not sufficiently fine-structured, but because they can *already* be convincingly shown to be incorrect. I refine a version due to Flagg and observe that the question as to its adequacy is not settled by any philosophical or technical results known at

Synthese **116:** 1–25, 1998.

present. I thus conclude generally that the possibility of a successful use of (extensions of) Epistemic Arithmetic to analyse intuitionistic arithmetic and Church's Thesis is still open and promising.

The objections to be discussed turn on general *methodological* questions about the adequacy of proposed formalizations. So I begin with some rather abstract considerations about such matters before turning to the detailed discussion. The criteria for the evaluation of formalizations I urge are necessarily somewhat imprecise, but no more so than analogous things as they are currently understood in application to the evaluation of scientific theories.

Although I have tried to keep this paper self-contained, I found it at times necessary to enter into the intricacies of some of the existing formal systems of Epistemic Arithmetic. For detailed descriptions of these systems, the reader is referred to the original papers in which these systems were proposed.

2. EVALUATING FORMALIZATIONS

2.1. *What is a Formalization?*

A naïve way of thinking about formalizations is the following. Suppose one wants to formalize an informal notion and principles concerning it. A formal language is constructed. This formal language may contain special primitive symbols, as in the case when one wants to formalize classical arithmetic (one will choose, e.g., a name *0* for the number 0, and a function constant s for the successor function). Subsequently one tries to represent the concepts and principles to be formalized in the formal language. Sometimes the representation of the concepts involved will simply involve the addition of new primitive predicates or constants, as when one formalizes the notion of truth by adding a new primitive predicate (T, say) to a formal language. But sometimes this is not so straightforward, for instance when it was attempted to represent the notion of a real number in the language of set theory.

George Boolos has, in a somewhat specific context, emphasized that representation in a formal language in itself is not enough to even raise the question of adequacy of formalizations. As long as the formal language has not been given a fairly definite interpretation, there is no formalization (Boolos 1975, 519–20). This does not mean that there has to be given a formal semantics of the language of the formalization in order to speak of a formalization: the formalization of set theory in the beginning of the century by Zermelo and others was already then a good formalization, long before the Tarskian notion of model. It only means that the formal language has to be an *interpreted* language, i.e., that a fairly definite meaning (in the intuitive sense of the word) has to be assigned to all symbols of the language. Sometimes a formal semantics can actually help, for instance when there is doubt whether the notions and principles to be formalized are coherent, or when a formal semantics can somehow clarify the intended interpretations of the formal language (the inductive technique for building fixed point models for languages containing a truth predicate is an ingenious way of trying to do this). Often an interesting class of models can be isolated by writing down axioms concerning the 'new' primitive symbols that the formal language contains, axioms for which it can be argued that they are valid on the intended interpretations (even if the class of intended interpretations is not yet very clearly circumscribed).

What then is a formalization? A formalization of a concept is a representation of that concept and of statements concerning that concept in an interpreted formal language, considered in the context of a theory that is formulated in that formal language. In the case of epistemic theories of arithmetic, which we will consider in Sections 2 and 3, the interpreted languages will be ordinary formal languages of arithmetic to which an absolute provability operator is added. And the background theories will be familiar axiomatic formalizations of arithmetic in which the absolute provability operator is governed by S4-like principles. The informal concepts to be formalized will be 'constructive provability' and 'effective computability'.[1]

Formalizations serve *epistemological* goals. They yield *theories* of the informal notions in question. These theories should add to the existing knowledge of the informal notions in question. In this respect, a formalization of a notion may be said to be a special kind of analysis of that notion.[2]

Formalizations need not be meaning-preserving. The formalization of the notion of a function as a set of ordered pairs, for instance, is in all likelihood not completely faithful to the original meaning of the concept. Yet it is a very good formalization. Formalizations also need not be unique.[3] The formalization of the function-concept in the untyped λ-calculus is also a good formalization of the concept of a function, even though it is very different from the ordered pairs-formalization. Another illustration is given by modal logic. The S4-formalization, interpreted in a Kripkean possible worlds-semantics, is a good formalization of the notion of possibility. But so is the S5-formalization, interpreted in a Kripkean possible worlds-semantics. Formalizations may to some extent be said to "create" concepts, in the sense of making them more definite, and this can often be done in different ways. This is not to say that any way of doing so is as good as any other. There are *positive criteria* that a formalization has to satisfy in order to be a good formalization. To these criteria we now turn.

2.2. *Proving Truths and Refuting Falsehoods*

One requirement that a formalization has to satisfy in order to be a good formalization is that the formal theory has to prove (representations of) truths about the informal notion and refute (representations of) falsehoods about the informal notion. To the extent that the formal theory succeeds in doing this, the claim that the formalization is a good one receives confirmation. It would be a difficult task to formulate a quantitative theory about the degree of confirmation (or disconfirmation) that a formalization receives in this way. But this is a problem that is not particular to the measurement of goodness of formalizations, it is a problem of confirmation theory in general.

In order to find out whether the theory, formulated in the formal language, proves only truths of the concepts and principles to be formalized, the 'received wisdom' concerning the informal concepts or principles must be consulted. If there is a concensus about the truth (falsehood) of certain properties of the concept or principle to be formalized, then an adequate formalization ought to respect this – or there have to be very good reasons to believe that the literature on the subject is simply mistaken. It may also happen that the formalization claims certain properties or principles about the concepts over which there is disagreement in the literature. That can be reason for concern. For if the intended semantics of the formal system does not make it clear *why* these principles have to be taken to be true, then this is reason to suspect that the formalization does more than describe the "core meaning" of the concepts involved. On the other hand, if the formalization does explain why these controversial principles have to be taken to be true, more power to the formalization! And there is always the further possibility, alluded to earlier, that the process of formalization reveals that where there seemed to be only one concept there really were two, so that what appeared to be controversy over a property of one concept is dissolved by pointing out that the concept is ambiguous: on one definite way of reading the ambiguous notion it indeed has the property in question, whereas on the other way of reading it it hasn't.

2.3. *Extending Expressive Power*

Suppose that the formalization only proves recognized truths, and only refutes recognized falsehoods. Then this is by no means sufficient to say that the formalization is a good one. It may just be due to a lack of expressive power that the formalization does not prove certain falsehoods: evidence may be 'hidden' because of lack of expressive power.[4] So, ideally, one ought to attempt to analyze all concepts as far as possible, write down all axioms about them which can be seen to be valid, relate (in the axioms) the concepts to as many other concepts as possible. And in the highly expressive formal language which one thus obtains, one should see if the formal theory does not prove falsehoods or dubious statements.

This process of extending the expressive power of a formal system is potentially endless, and there is a legitimate question how far it should be taken. The answer to this question depends on the situation. First of all, there is Kreisel's law of diminishing returns (see Kreisel 1987). At some point in this process, the formal theories will become so complex that they are difficult to handle, without giving interesting new results in return. In such situations it is pointless to go further in the direction of increasing expressive power. Second, sometimes there are good reasons *not* to take a certain step which would increase expressive power, namely when one does not have a good idea how the resulting formal theory and its semantics should look. An illustration of this is when one resists treating necessity as a predicate because it would bring in the paradoxes and we do not have a reasonably satisfactory theory of them.[5]

2.4. *Theoretical Virtues*

The foregoing conditions are still not sufficient to speak of a good formalization. For consider the following situation. The aim is to formalize a certain concept, and principles involving that concept. So a new symbol is chosen, and it is specified that it ought to be interpreted as the concept in question. All established truths about the concept are translated, using this new symbol, in first-order logic (or second-order logic, or whatever), in the way that it is taught in elementary logic courses. And if there is reason to believe that the concept consists of component concepts, then one introduces symbols to represent them, stipulates that they have to stand for the component concepts, and so on.

Such a formalization would be considered uninteresting. The reason is that the formalization does not tell us anything that we did not already know before: it does not meet the epistemological requirement that was formulated in Section 2.1. A formalization can only satisfy this requirement if it relates the concept in question to other concepts, and to background theories concerning those other concepts. As a theory concerning the informal concept to be formalized, it has to satisfy the same sorts of normative conditions (which in the philosophy of science are called *theoretical virtues*) which empirical theories have to satisfy:[6] it has to clarify conceptual relations that were not explicitly known before the formalization took place; it has to disambiguate confusions in the literature, it has to unify theories which were previously thought to be unrelated, it has to make distinctions that had not been made before, it has to explain certain facts that were in need of explanation ....

This explains why it is often more fruitful to try to approximate a concept indirectly, using concepts that we have a better grip on, than to capture it explicitly. Take for instance the formalization of the theory of effectively computable functions on the natural numbers in the theory of recursive functions. Another possibility would have been to introduce a new symbol for the notion of 'algorithm', and then to formulate, using this new symbol, all the properties of algorithms one can think of. The road that was actually followed turned out to be more fruitful. The theory of algorithmically computable functions that was actually constructed in the 30's relates the notion of algorithm to other notions (such as the concept of definability), explicates the boundaries of the class of algorithmically computable functions, distinguishes important subclasses of the class of algorithmically computable functions (e.g., the subclass of primitive recursive functions), and so on.

These remarks (especially those of Sections 2.3 and 2.4) show that the goodness of formalizations is a matter of degree: there are bad and excellent formalizations, and there is usually a whole spectrum in between. Evaluating a formalization is in most cases a question of weighing positive and negative aspects. Let us now apply these observations to the case of epistemic formalizations of mathematical practice.

3. FIRST APPLICATION: FORMALIZING CONSTRUCTIVISTIC MATHEMATICS

Stewart Shapiro has proposed a formal theory in which both constructive and nonconstructive aspects of arithmetic can be expressed. The main purpose of his theory is to "integrate" classical and intuitionistic arithmetic.

He attempts to accomplish this by adding an epistemic operator (K) to the formal language of first-order arithmetic. S4 deduction principles (without the Barcan formula) are formulated for K:

(1) $K(A) \rightarrow A$

(2) $K(A) \rightarrow K(K(A))$

(3) $K(A) \rightarrow (K(A \rightarrow B) \rightarrow K(B))$

(4) From A, infer to $K(A)$.

The theory that results from adding these to the axioms of first-order Peano arithmetic is called "Epistemic Arithmetic" (EA). The epistemic operator of EA is to be interpreted as: *it is ideally, or potentially knowable that* (Shapiro 1985b, 25). Principles (1)–(4) are easily seen to be intuitively valid for this interpretation of K as 'absolute provability'.

Shapiro then constructs a translation V from formulas of the language of Heyting arithmetic (HA) to the language of EA (Shapiro 1985b, 25). If we agree to indicate by means of a subscript i that a formula belongs to the language of HA, then we can describe this translation V as follows:

(1) for atomic formulas:

$$V(A_i) = KA_i$$

(2) for complex formulas:

$$V(A \wedge B)_i = K(V(A_i)) \wedge K(V(B_i))$$

$$V(A \vee B)_i = K(V(A_i)) \vee K(V(B_i))$$

$$V(A \rightarrow B)_i = K(K(V(A_i)) \rightarrow K(V(B_i))$$

$$V(A \leftrightarrow B)_i = K(K(V(A_i)) \leftrightarrow K(V(B_i)))$$

$$V(\neg A)_i = K(\neg K(V(A_i)))$$

$$V(\forall x A(x))_i = K(\forall x V(A(x)_i))$$

$$V(\exists x A(x))_i = \exists x K(V(A(x)_i))$$

This translation is intended to bring about the integration of classical and intuitionistic arithmetic. In fact, Shapiro argues, EA is able to express forms of *partial constructivity* (of which examples can be found in mathematical practice) that can be expressed neither in HA nor in classical Peano arithmetic (PA). For instance, take a statement which says that if an x with property A can be effectively found, then there must be a y (which perhaps cannot be effectively found) which has property B. This statement cannot be formalized PA, since its antecedent contains a constructive existential quantifier. And it cannot be formalized in HA, because its consequent contains a classical existential quantifier. Yet in EA it can be expressed as: $\exists x K(A(x)) \rightarrow \exists y B(y)$.[7]

Nicolas Goodman has shown that the translation V is *faithful* (Goodman 1984), in the sense that:

For every formula A_i of the language of HA:
$\vdash_{\mathrm{HA}} A_i \Leftrightarrow \vdash_{\mathrm{EA}} V(A_i)$.

This theorem says that an intuitionistic arithmetical sentence is provable in HA if and only if its translation is provable in EA.[8] Goodman's proof of the faithfulness of V was later substantially improved upon by Robert Flagg and Harvey Friedman, who have developed a very elegant and flexible method for obtaining faithfulness theorems for EA and related systems (Flagg and Friedman 1986b).

An *intended* interpretation of a formal system is an interpretation that respects *all* the restrictions on the meaning of the formal system, even those which are not expressed in the language of the formal system. Many people believe that PA, e.g., has exactly one intended interpretation: the natural number structure. The intended interpretation of EA is like the intended interpretation of PA, except that there now also is the operator K, which is to be interpreted as absolute provability. It is a controversial matter whether there is exactly one intended interpretation of HA: many constructivistic logicians feel that there is a rich variety of interpretations of intuitionistic logic and arithmetic, none of which can claim superiority over all other interpretations.[9] In any case, even if there is a unique intended interpretation of HA, Goodman's faithfulness theorem does not guarantee that the intended interpretation of an intuitionistic sentence A coincides with the intended interpretation of its epistemic translation $V(A)$.[10] Consider for instance the so-called *negative translations* from classical to intuitionistic languages.[11] Even though there are faithfulness theorems for these translations, I am aware of no logician who claims that they are *meaning-preserving*.

There are two objections that can be made to the claim that V is meaning-preserving, to which we will now turn. The first of these is not persuasive. But the second one is, and this is explicitly acknowledged by Shapiro. So Shapiro does not claim that V is meaning-preserving.

First, if we look at the clauses of Heyting's proof interpretation, we see that they do not contain an overt modal component: instead of "it is provable in principle that", these clauses contain occurrences of "I have a proof that".[12] But here it seems that to the extent that they do not contain a modal component, Heyting's proof conditions fail to be faithful to the meanings of the intuitionistic connectives, at least as they are understood nowadays, rather than there being a semantic deficiency in Shapiro's translation functions. Dummett for instance argues that the meanings of the intuitionistic connectives implicitly contain a modal component (see Dummett 1982, 118–19).

A more telling objection to the thesis that V is meaning-preserving arises when we look at Heyting's clause for the intuitionistic implication.[13] This clause says (roughly): "I have a proof of $A \to B$ iff I have a *method* which, given a proof of A, produces a proof of B". Shapiro admits that the notion of "transformations of proofs" is not captured by V, and cannot be captured in the language of EA (Shapiro 1985b, 25). This observation also seems to be behind Smorynski's objection against V to the effect that it "does not capture the full flavor of talk about methods" (Smorynski 1991, 1497). On the strength of this observation, he then passes a harsh judgement on epistemic mathematics:

> The justification of the study of "epistemic mathematics" as a joint codification of classical and intuitionistic mathematics is really nothing more than a pretense, a lip-service rather than a genuine – or even plausible – explanation. What one has emerges as an exercise in formalism – a nice system in which two initially incompatible formal systems can be both formally (and faithfully) interpreted. Just as institutions take on a lift of their own, so too do formal systems in logic. One forgets that formal systems are simply codifications of some portion of mathematical practice and begins to identify them with that practice. (Smorynski 1991, 1497)[14]

But this does not follow from the fact (which I think should simply be conceded) that V is not completely meaning-preserving. Granting that V does not succeed in exactly expressing the meanings of the intuitionistic connectives and statements, one could argue that this translation comes more or less close to doing this. And even if it does not even come close it is possible that the proposed formalization of intuitionistic arithmetic fulfills the requirements for being a good formalization that were laid down in Section 2. We will now investigate to what extent that is so.

It should be emphasized that the question of the existence of *faithfulness theorems* is relevant to the evaluation of Shapiro's formalization proposal. HA is a representation of the 'received wisdom' concerning the informal concept of constructive arithmetical provability. It is the common core of all (or at least most) constructivistic schools. Goodman's faithfulness theorem guarantees that EA confirms all (translations of) statements which are recognized to be true in HA, refutes all (translations of) statements which are recognized to be false by HA, and remains neutral on the (translations of) statements on which HA remains neutral.

But, as noted in Section 2.3, this is not nearly enough. If the epistemic approach is to give us a decent theory of the constructivistic enterprise, then we need faithfulness theorems also for stronger systems and more expressive languages. Some work has been done in this direction. Goodman's theorem has been extended to higher-order arithmetic (Flagg 1986a) and to set theory. There are a few faithfulness theorems for translations to epistemic systems of constructivistic systems containing principles about the so-called 'lawlike universe', such as intuitionistic versions of Church's thesis or Markov's principle (Horsten 1997). But many questions remain open. For instance, there exists at present no formalization in epistemic arithmetic of the intuitionistic theory of lawless sequences. In short, it would have to be established that all (or at least most) constructivistic theories can be faithfully translated into the epistemic framework.[15] Otherwise there remains a suspicion that the faithfulness theorems are due only to an artificial restriction of the expressive power of the languages in question. In sum, so far the faithfulness theorem has proved to be reasonably stable under extension of expressive power and under strengthenings of the systems involved.[16] But it remains to be seen whether this continues to be so.

But we want more than just faithfulness theorems. Aside from the truths that are recognized by formal systems of constructivistic arithmetic, there are truths about constructive provability that are not recognized by these systems. Ideally, we want the epistemic formalizations to respect *all* informal established truths about constructivistic truth. With respect to this requirement, Shapiro notes (Shapiro 1985b, 23–5) that the clauses of the definition of V closely resemble Heyting's explication of the meaning of the intuitionistic connectives in terms of proof conditions, which was constructed by him *in tempore non suspectu*, i.e., before variants of the translation V were first constructed by Gödel.[17] And here the situation is very different from the situation concerning the negative translations from classical to constructivistic systems. Reading the clauses of such translations as "giving the content of the classical connectives" (or coming close to doing that) amounts (very roughly) to the doctrine that what it means for a classical sentence to be true is that it is provably irrefutable (or, a little more precisely, that any proof of its falsehood can be transformed into a proof of an absurdity). But that explication of classical truth is considered by most classical logicians and mathematicians to be highly controversial at best.

In Section 2 it was claimed that formalizations are theories. Interesting formalizations have theoretical virtues, much like interesting empirical theories do. To conclude this section, let us then look at the theoretical virtues that epistemic formalizations of arithmetic have or lack.

First of all, it seems fair to say that the epistemic framework has *unifying power*. Epistemic arithmetic "integrates" classical and intuitionistic arithmetic (Flagg 1986a).

Secondly, the epistemic approach allows us to *explicate distinctions* which could not be formally expressed before. We have mentioned that in epistemic languages it is possible to express forms of partial constructivity, which cannot be expressed in the language of Peano arithmetic, nor in the language of Heyting arithmetic (Shapiro 1985a, 2; Shapiro 1985b, 40).

Thirdly, it seems that epistemic formalizations can play a role in the unraveling of certain *confusions in the literature*. The following is an example of this. Consider the following argument by Alan Hazen against the meaning-preservingness of V:

> Shapiro and Goodman are interested solely in recapturing within a classical context the distinctions intuitionistic mathematics makes; they are not sympathetic to the critical and anti-realist side of intuitionism as a philosophy of mathematics. It would be nice, however, if a technical explanation could give at least some suggestion of this side of things, if only to help understand why intuitionistic *mathematics* (including logic) should be thought of as congenial to intuitionistic philosophy. (Hazen 1990, 186)

More specifically, what he takes to be the trouble with Shapiro's epistemic system of arithmetic is this:

> Shapiro and Goodman have, between them, shown that a sentence of first-order arithmetic is a theorem of Heyting Arithmetic if and only if its translation is a theorem of the modal system [i.e., EA]. A non-realist interpretation of the modal system is not possible, however, since the (non-modal) sentences saying, e.g., that the natural numbers are strictly linearly ordered by *greater than* and that for every natural number there is a greater, are provable: any model for the system must have an infinite domain ... (Hazen 1990, 188)

But it is not clear what is objectionable about this. Since Shapiro and Goodman give a translation of the language of HA *in a classical system of arithmetic*,[18] the theory in which HA is interpreted will assert the existence of an infinite structure. The intuitionist will believe only the provable sentences of EA *which are translations of sentences of HA*. And from the translations of the axioms of HA, using only translations of intuitionistic rules of inference, no actual infinity will be provable to exist. And in this sense, the translation V *does* appreciate the anti-realist side of intuitionism.

But perhaps there is a more serious problem for Shapiro's translation V that Hazen's objection points to. Take the epistemic translation of the intuitionistic successor postulate: $K\forall x\exists yK(y = s(x))$. Since the operator K is governed by the S4 rules, it seems that it follows *by classical (epistemic) logic alone* that $\forall x\exists y(y = s(x))$: in particular, we need two instances of the axiom $K(A) \rightarrow A$. Perhaps this should not be so: the intuitionistic successor postulate is usually taken to assert only the *potential* existence of a successor for each number; pure logic alone does not have the power of making mere potentialities actual!

To show that even this objection is not convincing, let us consider the finer analysis of the absolute provability operator that was carried out in

Horsten (1994). In EA the sentential operator K is a primitive or unanalyzed operator. But the notion of provability in principle seems to contain both a modal component (prova*bility*) and an epistemic component (mathematical *proof*). Therefore it seems useful to attempt to construct a formal system in which Shapiro's operator K is regarded as a complex operator $\Diamond P$, where $\Diamond$ is the familiar modal operator, and the operator P should be read as "some (not further specified) mathematician has a proof that ...". Some of the logical properties which Shapiro postulates for the notion of absolute provability (i.e., the axioms of S4) can then be *derived* from more basic logical properties of the notion of possibility and (having a) proof.

This results in a modal-epistemic formalization of arithmetic, called MEA. The language of MEA contains a modal operator ($\Diamond$) and an epistemic operator (P), and the symbols of the first-order language of arithmetic. The system MEA is based on a standard formalization of classical first-order logic. The modal operator of MEA is governed by the S5 axioms, plus the (platonistic) principle:

(M) $\Diamond A \rightarrow A$ for all sentences not containing any occurrences of $\Diamond$ or P.

The intended domain of MEA is the natural number structure. The principle M asserts that all arithmetical truths are necessarily true, and all arithmetical falsehoods are necessarily false.

MEA has the following two epistemic axioms:

(P1) $PA \rightarrow A$

(P2) $PA \rightarrow PPA$,[19]

and the following two modal-epistemic principles:

(ME1) From A, infer $\Diamond PA$

(ME2) $(\Diamond PA \wedge \Diamond P(A \rightarrow B)) \rightarrow \Diamond PB$.

As arithmetical axioms, we have as before the Peano axioms for elementary arithmetic. Again, arguments can be given to the effect that these principles are sound for the intended interpretation of MEA.

Then a translation V^* from the language of HA to the language of MEA is constructed which is just like V, except that all occurrences of K in V-translations of formulas are replaced by $\Diamond P$. It is shown that V^* is also a faithful translation (Horsten 1994, 287). But more important for our present purposes is the following. First, one can prove a lemma which shows that for each formula A in which P only occurs when it is immediately preceded by an occurrence of $\Diamond$ (such formulas are called *$\Diamond P$-formulas*), MEA proves $\Diamond A \rightarrow A$. Subsequently it can be shown that MEA proves $\Diamond PA \rightarrow A$ for all $\Diamond P$-formulas A (Horsten 1994, 287).[20] Since the complex operator $\Diamond P$ can be taken as a gloss of Shapiro's absolute provability operator K, this principle is the counterpart of the principle $KA \rightarrow A$ ("if a sentence is provable, then it is true"), which was used essentially in the supposedly objectionable deduction of $\forall x \exists y(y = s(x))$ from $K\forall x \exists y K(y = s(x))$. The principle M is used *essentially* in the derivation in MEA of $\Diamond PA \rightarrow A$ for $\Diamond P$-formulas A. Although M may very well be a true axiom (if a certain form of platonism concerning mathematical objects is true), *it is not a logical truth*. Therefore truths of the form $\Diamond PA \rightarrow A$ are in general not logical truths. And this implies that the principle $KA \rightarrow A$ is, on closer inspection, not a truth of epistemic *logic*.

So the situation is as follows. On the one hand, it seems unreasonable to impose as a condition on formalizations of intuitionistic arithmetic that they use a theory that postulates no actual infinity as their background theory – just as it would be unreasonable to require that they use in their background theory only intuitionistic logic. On the other hand, we have found that the existence of a platonistic number structure can only be derived from constructivistic assumptions *if certain platonistic principles are used in the process* ("Platonism In, Platonism Out").[21] In sum, then, I am unable to discover any for our purposes significant sense in which epistemic languages fail to respect the anti-realist motivations of intuitionism.

4. SECOND APPLICATION: EXPRESSING CHURCH'S THESIS

Church's Thesis (CT) is often expressed (somewhat loosely) as: "Every computable function is recursive".[22] It is well-known that in Peano arithmetic it is not possible to express CT. The consequent of CT can be expressed, using Kleene's T-predicate and U function symbol.[23] But its antecedent contains the "informal" notion of an *algorithm*, and this cannot be expressed in the language of Peano arithmetic.[24] One could attempt to introduce new primitives in the language to talk directly about algorithms, write down basic axioms concerning them, and attempt to *prove* CT from these axioms (Shapiro 1981, 384).[25] But it appears that this proposal has so far not been worked out.

It seems that epistemic arithmetical languages can do somewhat better. Several principles which seem to somehow approximate CT have been proposed in the literature on epistemic arithmetic. I will now investigate these proposed formalizations, and argue that they are defective. But I am not ready to admit that this is due to an inherent impossibility of EA to give a good formalization of CT, although it does underscore the need for a more explicit semantics of EA (especially of the higher-order version of it) than has been given until now. I will propose a new putative formalization of CT, and argue that it is not vulnerable to the charges that I bring against the other candidates.

4.1. *First-Order Formalizations of Church's Thesis*

Let us start by making a few simplifying assumptions, to be relaxed later on. We consider CT for *total* functions, denoted as CT_T. And to avoid for the time being quantification over functions, we content ourselves with approximating CT_T by means of a first-order epistemic *schema* rather than by means of a second-order epistemic sentence.

Now given that we can express "$\phi(x, y)$ expresses a total recursive function" already in the language of Peano arithmetic in terms of the U function symbol and the T-predicate, John Myhill proposes the following formalization of CT_T (Myhill 1985, 47):

$$\text{(M)} \qquad \forall x \exists y K \phi(x, y) \rightarrow \exists e \forall x \exists y (\phi(x, U(y)) \wedge T(e, x, y))$$

Very roughly, (M) says that if for every x, it is possible to find a y which can be shown to stand in the relation ϕ to x, then ϕ determines a total recursive function.

The first thing to notice is that the number y associated with any given number x ought to be *unique* (an algorithm yields a unique output value for each input value), otherwise (M) asserts something that is not implied in CT_T. But this is easily fixed. One way in which it can be done is the following:

$$\begin{aligned} \text{(M}'\text{)} \qquad & \text{“}\phi \text{ determines a function”} \rightarrow \\ & [\forall x \exists y K \phi(x, y) \rightarrow \exists e \forall x \exists y (\phi(x, U(y)) \wedge T(e, x, y))] \end{aligned}$$

There are strong reasons for doubting that the principle (M′) comes close to capturing the content of CT_T. Take a nonrecursive total function $\psi(x, y)$, the halting function, say. It may well be that the following is the case: for every m, there is an n such that $\psi(m, n)$ can be proved; but this infinite collection of proofs cannot be "compressed" into one single algorithm. If such is the case, then the antecedent of (M′) is true, whereas its consequent is false. For this to be a *proof* that (M′) is false it has to be shown that for every m, there is an n such that $\psi(m, n)$ can be proved. It is needless to say that I have no such proof to offer. But the situation I have sketched just now is not intended to be excluded by CT_T. Myhill himself, in an earlier publication, said as much. If we take $\phi(x, 1)$ to be true if x is the Gödel number of a first-order logical validity, and $\phi(x, 0)$ to be true for all other x, then Church's theorem, combined with the contraposition of (M′) gives us $\exists x \forall y \neg K\phi(x, y)$, i.e., there is a sentence of which it is absolutely undecidable whether it is valid. To this, Myhill replies:

> But there seems to be no reason to suppose that any particular problem in the theory of propositional functions will prove especially refractory, just as from each man's disability to see all women we could not infer that any one woman would be invisible to all men. "There is no technique that will test all (such) arguments" is true, while there is no special reason to suppose that there are (such) arguments that no techniques will test (even if that means anything). (Myhill 1952, 171).

Let us then turn to a variant of a formalization of CT_T that is discussed by Robert Flagg (Flagg 1985, 166):[26]

(F1) "ϕ determines a function" $\rightarrow$
$[K\forall x \exists y K\phi(x, y) \rightarrow \exists e \forall x \exists y(\phi(x, U(y)) \wedge T(e, x, y))]$

This principle differs from (M′) only in that it has $K\forall x \exists y K\phi(x, y)$ where (M′) has $\forall x \exists y K\phi(x, y)$. In other words, the antecedent now says that there is a *single proof* which demonstrates that for each x, there is a y which can be shown to stand in the relation ϕ to x. It may be Flagg's intention to thereby exclude counterexamples of the kind that we have raised in response to Myhill's proposal. But (F1) only succeeds in doing this if the absolute proof witnessing the initial occurrence of the provability operator of $K\forall x \exists y K\phi(x, y)$ somehow *guarantees* the existence of an algorithm computing ϕ. Perhaps the underlying philosophical thesis is that this absolute proof somehow *gives* an algorithm for computing ϕ. In other words:

THESIS 1. The only way in which a statement of the form $\forall x \exists y K\phi(x, y)$ can be proved is by giving an algorithm for computing ϕ.

It is not at all obvious that this thesis is true. One might wonder, e.g., whether there are statements of the form $\forall x \exists y K\phi(x, y)$ for which a proof by reductio ad absurdum can be found, but which do not admit of a constructive proof. Nevertheless, it seems very difficult to envisage a concrete candidate for such a counterexample.

But even if Thesis 1 is true, there is a strong reason for doubting that (F1) succeeds in capturing the content of CT_T. We have noted earlier that sometimes CT_T is taken as having the structure of a *biconditional*. And even when CT_T is formalized as a conditional statement, it is commonly assumed that the converse of this statement is true (Mendelson 1990, 228). But the converse of (F1) seems implausible. Suppose that there is an absolutely undecidable arithmetical sentence φ (i.e., an arithmetical φ such that $\neg K\varphi$ and $\neg K\neg\varphi$). Then define a function θ in the following way:

For every x:

$$\theta(x, 1) \text{ if } \varphi;$$

$$\theta(x, 0) \text{ if } \neg\varphi;$$

$$\neg\theta(x, y) \text{ for all numbers } y \text{ which are not identical to 1 or 0.}$$

θ denotes a recursive function; yet $\forall x \exists y K\phi(x, y)$ is false. Again, for this to be a refutation of the converse of (F1), it has to be *proved* that there are absolutely undecidable arithmetical sentences. And there is no consensus about whether there are absolutely undecidable sentences, or even about whether the notion of absolute provability is sufficiently determinate for this question to have a determinate answer. Nonetheless, in the absence of a convincing argument that there are no such sentences, the converse of (F1) is doubtful. And if the converse of CT_T is "obviously true" (Mendelson 1990, 232), then one should be wary about the claim that (F1) comes close to capturing the content of CT_T.

4.2. *Higher-Order Formalizations of Church's Thesis*

There are better epistemic candidates for a formalization of CT_T. But these are essentially higher-order principles. So we have to look at higher-order epistemic formalizations of epistemic arithmetic. In particular, I will briefly describe Flagg's system of epistemic type theory (see Flagg 1986a), and investigate the possibility of expressing Church's thesis in the (more expressive) language of this system.

Flagg's system is based on a typed language. There are two basic types: N (the type of the natural numbers, which are the ground objects), and Ω (the type of the formulas of the language). Furthermore, there are two operations for forming complex types out of simple ones: a product-forming type operation $(\times)$ and a powerset-forming type operation (P). These operations are used, in the familiar way, to form complex types out of simple ones. Corresponding to this, the language of Flagg's system contains unique names for the numbers (obtained from *0* and the successor symbol s), variables of all types, a pairing symbol, an abstraction operator, an epsilon sign (ϵ), and of course the absolute provability operator K.

Flagg's system then contains the usual rules and axioms of classical higher-order logic[27] (with the unrestricted comprehension axiom), the standard second-order formulation of the axioms of Peano Arithmetic, and the S4 axioms governing the absolute provability operator. In sum, Flagg's system is almost exactly like EA, except that it is a higher-order system.

Flagg is not very explicit about the intended interpretations of his system.[28] In particular, it is not immediately obvious how the quantifiers should be interpreted.[29] Nevertheless, the following is a sketch of a consistent interpretation of Flagg's system. Quantifiers range over mathematical objects, *given under a presentation*. Or, equivalently, one can think of the quantifiers as ranging over *ordered pairs of mathematical objects and presentations*. Shapiro suggests that in general a presentation can be taken to be an interpreted *linguistic* expression (Shapiro 1985b, section 7). Certainly this is the most natural construal but there seems to be no compelling reason for such a restriction. I want to be as liberal as possible about what is allowed to count as a presentation of a mathematical object. Even the graph of a total function, for instance, is allowed to count as a presentation.[30] But even the idealized mathematician is finite. If there are infinite presentations (such as the graph of a total function), the mathematician can only have *finite access* to them: she cannot have an entire graph of a total function before her consciousness at once. There is just one restriction that I do impose on what is allowed to count as a possible presentation:[31] the *first-order* variables range over the numbers, *given in a canonical way* (by a finite number of successor symbols prefixed to *0*, say). So one cannot always substitute coextensive higher-order presentations for each other in epistemic contexts *salva veritate* – for first-order presentations there is no such problem. As an illustration of this, suppose one presents a function as "the function which is the constant *1*-function if Goldbach's conjecture is true, and the constant *0*-function otherwise". If Goldbach's conjecture is true, then this presentation is coextensive with "the constant *1*-function". But if Goldbach's conjecture is in addition absolutely unprovable, then it cannot always be substituted *salva veritate* for "the constant *1*-function" in the context of the absolute provability operator. In sum, in extensional contexts the associated presentations do no real work, but in intensional contexts it can make all the difference in the world how a mathematical object is presented.

These remarks determine an interpretation of the formulas of the language of Flagg's system. As an example, consider the expression $\exists f K \forall x (f(x) = x)$, where f is a function variable. This expression is to be read as: there is a function, given under a certain presentation, such that it is provable of that function, under that presentation, that it is the identity function. This way of reading higher-order formulas of Flagg's system can be straightforwardly extended to arbitrary arity and to arbitrary order. And it can be argued (although I will not do so here) that Flagg's system is sound for this interpretation. I realize that all this is still rather vague. To make it precise, a *theory of presentations of mathematical objects* would have to be constructed.[32] But for present purposes, this will have to suffice by way of description of the suggested interpretation of Flagg's system.

Given that we can express the notion of a total recursive function in the language of epistemic type theory, a variant of Flagg's second-order formulation of CT_T can be formulated as follows:[33]

(F2) $\forall v^{P(N\times N)}$ ("v determines a function" $\rightarrow \{K\forall x^N \exists y^N K(\langle x, y\rangle \in v) \rightarrow$ "v expresses a total recursive function"$\})$

The superscripts in this formula indicate to which type the variables in question belong. Roughly, (F2) can then be read as follows: If you can prove of a function f, given under a certain presentation (Q), that for every x, there is a y such that it can be shown that $Q(x, y)$, then f is a total recursive function.

Now the same objection that was made against (F1) can be made against (F2): not all instances of the converse of (F2) are obviously true. Nevertheless, we will show how we can modify (F2) so that the converses of its instances are all obviously true. Let us first focus on CT for *partial* recursive functions (CT_P). The analogue of (F2) for partial recursive functions is:

(F2P) $\forall v^{P(N\times N)}$ ("v determines a function" $\rightarrow \{K\forall x^N(\exists y^N(\langle x, y\rangle \in v) \rightarrow \exists y^N K(\langle x, y\rangle \in v)) \rightarrow$ "v expresses a partial recursive function"$\})$

Again, not all instances of the converse of (F2P) are obviously true. But this is now easily remedied. Consider the following proposed formalization of CT_P:

(CTP) $\forall v^{P(N\times N)}$ ("v determines a function" $\rightarrow \{\exists w^{P(N\times N)}[\forall z^{N\times N}(z \in v \leftrightarrow z \in w) \wedge K\forall x^N(\exists y^N(\langle x, y\rangle \in w) \rightarrow \exists y^N K(\langle x, y\rangle \in w))] \rightarrow$ "v expresses a partial recursive function"$\})$

If we now replace the main occurrence of '$\rightarrow$' in the consequent of (CTP) by a reversed implication '$\leftarrow$', the result still is clearly true. So we might as well replace this nested occurrence of '$\rightarrow$' by a biconditional. The resulting principle would then say, roughly, that the class of partial recursive functions is identical with the class of partial functions having a relatively accessible presentation, viz., they each have a presentation w such that $K\forall x^N(\exists y^N(\langle x, y\rangle \in w) \rightarrow \exists y^N K(\langle x, y\rangle \in w))$. Obvious candidates for such accessible presentations would be codes of suitable Turing machines.

Along the lines of (CTP) it is also possible to formalize CT_T:

(CTT) $\forall v^{P(N\times N)}$ ("v determines a total function" $\rightarrow$ $\{\exists w^{P(N\times N)}[\forall z^{N\times N}(z \in v \leftrightarrow z \in w) \wedge K\forall x^N(\exists y^N(\langle x, y\rangle \in w) \rightarrow \exists y^N K(\langle x, y\rangle \in w))] \rightarrow$ "v expresses a total recursive function"$\}$)

Here again, if we reverse the nested implication, the result is clearly true.

Note that the following formalization proposal for CT_T, which is closer in spirit to Flagg's (F2), is deficient:[34]

(CTT*) $\forall v^{P(N\times N)}$ ("v determines a function" $\rightarrow$ $\{\exists w^{P(N\times N)}[\forall z^{N\times N}(z \in v \leftrightarrow z \in w) \wedge K\forall x^N \exists y^N K(\langle x, y\rangle \in w)] \rightarrow$ "v expresses a total recursive function"$\}$)

The reason is that the converse of (CTT*) is (again) *not* obviously true. For all we can tell, there may be total recursive functions v of which there exists no sufficiently accessible presentation w such that it is knowable of w that the recursive function it expresses is *total*. This is the reason why the antecedent of (CTT) does not entail that it is knowable that w determines a total function.

Finally, there remains the analogue of the objection that was raised by Smorynski against the epistemic translations of constructivistic implications and quantifiers: "To say that a function is computable is to say that there exists an algorithm that computes it; algorithms are methods or sets of instructions; and epistemic languages 'do not capture the full flavor of talk about methods' ". Again, I think that this simply has to be conceded. Therefore even (CTP) and (CTT) do not capture the exact meaning of CT_P, and CT_T, respectively. But if a formalization need not aim at being meaning-preserving (see Section 2.1), then the fact that (CTP) and (CTT) are not meaning-preserving is not in itself cause for concern.

The philosophical thesis behind the formalizations (CTP) and (CTT) is the following analogue of Thesis 1 of Section 4.1:

THESIS 2. The only way in which a statement of the form $\forall x^N(\exists y^N(\langle x, y\rangle \in w) \rightarrow \exists y^N K(\langle x, y\rangle \in w))$ can be proved is by giving an algorithm for computing w

If Thesis 2 is true, then CT_P is true if and only if (CTP) is true and CT_T is true if and only if (CTT) is true. So whether Thesis 2 is true seems to be an interesting question. And *if* it could be established that CT_P is true if and only if (CTP) is true and CT_T is true if and only if (CTT) is true, then this would be an advance in our overall theory of Church's Thesis.

With respect to the question whether the formalization proposals of CT that we have reviewed satisfy the criteria of Sections 2.2 and 2.3, few solid results are known. Flagg proved that a variant of (F1) is consistent with EA (Flagg, 1985),[35] and he proved a variant of (F2) to be consistent with his epistemic type theory. But one wants to know much more. For instance, are (F1), (F2) *conservative* over intuitionistic arithmetic (intuitionistic type theory) under the translation V, or even over classical arithmetic (classical type theory)? With respect to the requirements of Section 2.4, it seems that epistemic formalizations of CT fare relatively well. They relate notions that were previously not clearly perceived as related (namely the concept of algorithm and the concept of absolute provability). And they are of relevance to the philosophical literature on CT. To conclude this section, I now argue for this latter claim.

4.3. *Shapiro on Church's Thesis*

Let us turn briefly to Shapiro's views on Church's Thesis (see Shapiro 1980; Shapiro 1985b, 41–3). Shapiro seems to believe that CT cannot be directly captured in terms of the absolute provability operator. CT concerns the notion of computability: a function is computable if there exists an algorithm that computes it. So computability is objective, extensional, and does "not involve reference to a knowing subject" (Shapiro 1985b, 41). But closely related to the notion of computability there is a *pragmatic* notion, which Shapiro calls *calculability* (or "effectiveness", in the terminology of Shapiro, 1980). Calculability is a property of presentations of algorithms: a function presentation F is calculable if there is an algorithm P such that it can be established that F represents P (Shapiro 1985b, 43). Shapiro suggests that his theory of epistemic arithmetic can be used to shed light on this latter notion.

It will be clear from the foregoing that whereas I agree with Shapiro that computability is extensional and objective, I do think that the notion of computability involves reference to a knowing subject (since the notion of an algorithm does: an algorithm is a method that can be used *by humans*). And I do not believe it to be a foregone conclusion that CT (as opposed to its pragmatic counterpart) cannot be formalized using the notion of absolute provability.

5. CONCLUDING REMARKS

Our overall judgement about the success of attempts to formalize in Epistemic Arithmetic parts of informal mathematics (in particular constructive provability and algorithmic computability) must be a balanced one.

The weakest aspect of Epistemic Arithmetic as a whole is the lack of an illuminating and precise model-theoretic semantics for epistemic systems, especially for higher-order epistemic theories.[36] In particular, what is needed is an explicit theory of presentations of mathematical objects. With respect to faithfulness to the 'received wisdom', in the case of the formalizations of constructive provability it remains to be shown that constructivistic theories that extend Heyting Arithmetic (be it in the 'lawlike' direction or in the 'lawless' direction) can also be modeled in a natural way in classical Epistemic Arithmetic (to a large extent, the proof of the epistemic program is in the eating). In the case of the formalizations of CT, we have barely scratched the surface with respect to the conditions of Sections 1.2 and 1.3. On the positive side, the epistemic languages have so far withstood challenges to the claim that they verify only (representations of) truths and falsify only (representations of) falsehoods. In the case of the formalization of CT, it took some work to see how this can be done. We had to concede that Myhill and Flagg's epistemic formalizations of CT are defective. But it was also shown that variants of Flagg's proposals can be constructed which are immune against the criticism that was raised against the former formalization proposals.

Moreover, at present there seem to be no convincing *philosophical* arguments for the thesis that the program of trying to find epistemic representations of constructive provability and of Church's Thesis is doomed to failure. Smorynski's argument is unconvincing because it imposes an unreasonable requirement on the formalizations in question. The epistemic program would indeed be hopeless if epistemic formalizations of constructive provability and of algorithmic computability would have to be meaning-preserving. But, granting that the proposed formalizations are not meaning-preserving, it can be maintained that they do give us interesting analyses or theories about the concepts in question. And that is all that we should require from a formalization anyway. Hazen's objection that epistemic arithmetic does not respect the anti-realist motivation of constructivistic logic and mathematics was shown to be unconvincing using a finer analysis of the concept of provability in principle. The formalization proposals of CT that were reviewed in this paper challenge Shapiro's thesis that the notion of computability does not involve reference to a knowing subject.[37] In any case, his thesis does not follow from the objectivity and the extensionality of the notion of computability.

NOTES

* I am indebted to Tony Anderson for suggestions for improvement and for valuable and pleasant discussions on the subject matter of this paper. Comments by Igor Douven, Herman Roelants, John Burgess, Tony Martin and an anonymous referee also lead to significant improvements. The research for this paper was financially supported by the Flemish Fund for Scientific Research, which is gratefully acknowledged.

1 There may be more concepts of which interesting formalizations can be given in epistemic languages, e.g., epistemic notions of randomness.

2 I draw much of my inspiration here from (Anderson 1993), to which the reader is referred for further discussion and argumentation.

3 I am indebted to Igor Douven for this point.

4 As an example, consider the familiar propositional modal logics. Quine has argued that the notion of necessity is philosophically unsound, and that therefore no 'good' formalization of the notion of necessity can exist. However, he believes that the problematic aspects of the notion of necessity only come to the surface when we consider *quantified* modal logic. So in his view, most of the problems related to modal logic remain hidden if we only consider the propositional logic of necessity.

5 This is the justification that was given by Richard Montague for representing necessity by means of a propositional operator, rather than by means of a predicate that takes names of sentences as arguments. For a discussion of these matters, see (Montague 1963).

6 Concerning the difference between analyses (of which formalizations form a particular case) and empirical theories, see (Anderson 1993, Section 5).

7 There is a question whether epistemic arithmetic allows the expression of *all* forms of partial constructivity (see Lifschitz 1985; and Horsten 1993).

8 The only-if-direction of the theorem is straightforward; the if-direction is nontrivial.

9 See for instance (Troelstra and van Dalen 1988b, 839).

10 This was emphasized by Alan Hazen (Hazen 1990, 179).

11 For a discussion of negative translations and accompanying faithfulness theorems, see (Troelstra and van Dalen 1988a).

12 See for example (Heyting 1930).

13 Something analogous can be said for the clause for the intuitionistic universal quantifier.

14 The last sentence of the quote by Smorynski, by the way, is less applicable to Shapiro than to Smorynski, who seems to have missed the philosophical point of EA. Smorynski thinks that Shapiro is arguing for the replacement of classical and intuitionistic languages by epistemic languages, and claims that this will never happen (Smorynski 1991, 1496–1497). But surely Shapiro is not arguing for this (admittedly unreasonable) thesis. The point is simply this. Intuitionistic mathematics is a part of mathematics which is understandable for the classical mathematician. A formalization of mathematics which wants to be encompassing has to be able to account for this fact. The conjecture of Shapiro is that classical mathematicians, when they learn the meanings of the intuitionistic connectives, or do intuitionistic mathematics, *implicitly* use an "absolute" notion of provability, and a translation function which resembles V. Hence he incorporates these in his formalization.

15 Also, one wants more information about the stability of the faithfulness theorems under relatively small modifications of the modal logic on which the epistemic formal systems are based.

16 To some extent negative translations from classical to constructivistic languages also pass this test. Nevertheless, negative translations do not work well when one wants to translate theories about lawless sequences to a classical context.

17 There is a dispute in the literature about whether the logical connectives that occur in Heyting's explication of the meaning of the logical connectives can be taken to be constructive (see Hellman, 1989).

18 One could attempt to explicate the proof-conditions of HA in an *intuitionistic* theory, but this is another matter, and it is not what Hazen has in mind, since his favorite interpretation is an interpretation in a classical theory too.

19 In Horsten (1994) I consider the possibility that the principle P2 should be weakened to $PA \rightarrow \Diamond PPA$.

20 The reason for the restriction to $\Diamond P$-formulas is (roughly) that one needs to exclude the cases where A is of the form PB for some B. The sentence $\Diamond PP(5 + 7 = 12) \rightarrow P(5 + 7 = 12)$, for instance, makes an inadmissible inference from what is proved in some possible world to what is proved in the actual world. Even though $\Diamond PP(5 + 7 = 12)$ is arguably logically valid (and provable in MEA), it is not *logically* valid that *anything* has been proved in the actual world.

21 In Horsten (1994) it is also shown how V can be quite naturally modified into a translation from intuitionistic arithmetic to a formalization of Geoffrey Hellman's modal-structural interpretation of arithmetic, which is also anti-realist in spirit.

22 Sometimes it is also expressed as having the structure of a biconditional: "The class of computable functions coincides with the class of recursive functions".

23 Kleene's T-predicate is true of a triple $\langle e, x, y\rangle$ if and only if e is the code of a Turing machine which, when started on an input with value x, yields a computation which is coded as the number y. $U(y)$ is the number which is the output of the computation with code y.

24 In the language of constructivistic arithmetic we can come closer to expressing CT (see Troelstra and van Dalen 1988a, chapter 4). But even there one cannot quite express it, because CT has the form of a *classical* implication.

25 This possibility was apparently suggested by Harvey Friedman (Shapiro 1981, 364, footnote 3).

26 Flagg's actual formalization of CT_T is:

(F′) $K\forall x\exists y K\phi(x, y) \rightarrow$ "there Turing machine *of which it can be proved* that it computes ϕ"

I have emphasized earlier the need to restrict the principle to functions. It is puzzling why Flagg takes the consequent of CT_T to be an epistemic sentence. Perhaps it is related to Shapiro's *pragmatic* version of CT_T (cf. Section 4.3 below).

27 The principle of substitution of identicals has to be restricted somewhat, but we can ignore this complication for our present purposes.

28 This is to be deplored. We have emphasized in Section 2 the need for explicitation of the intended models of the epistemic systems that have been proposed in the literature.

29 Shapiro discusses the difficulties involved in higher-order quantification in epistemic contexts (Shapiro 1985b, Section 7). These difficulties are not confined to epistemic mathematics; one runs into the same problems when one sets out to do set-theoretical quantified provability logic (see Boolos, 1993, 226).

[30] This guarantees that every function has a presentation, and that if f is a function variable formulas of the form $\forall f \varphi$ quantify over *all* functions.
[31] And some such restriction has to be made for the interpretation to be *sound* for Flagg's system.
[32] The need for such a theory also becomes pressing when one attempts to construct a theory of highly intensional constructivistic objects such as lawless sequences.
[33] Here I make similar modifications to Flagg's actual proposal as I have made in my statement of Flagg's first-order proposal.
[34] This was pointed out to me by Tony Martin.
[35] This proof was subsequently simplified by Goodman (1986).
[36] The situation here resembles that of modal logic during the 50's. There are all these systems, but in the absence of a clear and unifying semantic framework there is the suspicion (that was voiced, in the case of modal logic, most strongly by Quine) that we really don't know what we are talking about. I am indebted to Tony Anderson for this observation.
[37] But of course, in line with what was said in Section 2.1, I do not want to exclude the possibility that *another* good formalization of CT might be proposed which makes no reference at all to a knowing subject.

REFERENCES

Anderson, C. A.: 1993, 'Analyzing Analysis', *Philosophical Studies* **72**, 199–222.

Boolos, G.: 1975, 'On Second-Order Logic', *Journal of Philosophy* **82**, 509–27.

Boolos, G.: 1993, *The Logic of Provability,* Cambridge University Press, Cambridge.

Dummett, M.: 1982, 'The Philosophical Basis of Intuitionistic Logic', reprinted in P. Benacerraf, and H. Putnam (eds.), *Philosophy of Mathematics. Selected Readings*, 2nd edition, Cambridge University Press, Cambridge, pp. 97–129.

Flagg, R.: 1985, 'Church's Thesis is Consistent with Epistemic Arithmetic', in S. Shapiro (ed.), *Intensional Mathematics*, pp. 121–72.

Flagg, R.: 1986a, 'Integrating Classical and Intuitionistic Type Theory', *Annals of Pure and Applied Logic* **32**, 27–51.

Flagg, R. and H. Friedman: 1986b, 'Epistemic and Intuitionistic Formal Systems', *Annals of Pure and Applied Logic* **32**, 53–60.

Goodman, N.: 1984, 'Epistemic Arithmetic is a Conservative Extension of Intuitionistic Arithmetic', *Journal of Symbolic Logic* **49**, 192–203.

Goodman, N.: 1986, 'Flagg Realizability in Epistemic Arithmetic', *Journal of Symbolic Logic* **51**, 387–92.

Hazen, A.: 1990, 'The Myth of the Intuitionistic "Or" ', in M. Dunn and A. Gupta (eds.), *Truth or Consequences?*, Kluwer, Dordrecht, pp. 177–95.

Hellman, G.: 1989, 'Never Say "Never"! On the Communication Problem between Intuitionism and Classicism', *Philosophical Topics* **17**, 47–67.

Heyting, A.: 1930, 'Die formalen Regeln der intuitionistischen Logik', *Sitzungsberichte der Preussischen Akademie von Wissenschaften. Physikalisch-Mathematische Klasse*, pp. 42–56.

Horsten, L.: 1993, 'Note on an Objection of Lifschitz against Shapiro's "Epistemic Arithmetic" ', in J. Czermak (ed.), *Proceedings of the 15th International Wittgenstein-Symposium. Part I: Philosophy of Mathematics*, Hölder-Pichler-Tempsky, Vienna, pp. 289–96.

Horsten, L.: 1994, 'Modal-Epistemic Variants of Shapiro's System of Epistemic Arithmetic', *Notre Dame Journal of Formal Logic* **35**, 284–91.

Horsten, L.: 1997, 'Provability in Principle and Controversial Constructivistic Principles', *Journal of Philosophical Logic* **26**, 635–60.

Kreisel, G.: 1987, 'Church's Thesis and the Ideal of Informal Rigour', *Notre Dame Journal of Formal Logic* **28**, 499–519.

Lifschitz, V.: 1985, 'Calculable Natural Numbers', in S. Shapiro (ed.), *Intensional Mathematics*, pp. 173–90.

Mendelson, E.: 1990, 'Second Thoughts about Church's Thesis and Mathematical Proofs', *Journal of Philosophy* **87**, 225–33.

Montague, R.: 1963, 'Syntactical Treatments of Modality, with Corollaries on Reflection Principles and Finite Axiomatizability', *Acta Philosophica Fennica* **16**, 153–67.

Myhill, J.: 1952, 'Philosophical Implications of Mathematical Logic', *The Review of Metaphysics* **6**, 165–98.

Myhill, J.: 1985, 'Intensional Set Theory', in S. Shapiro (ed.), *Intensional Mathematics*, pp. 47–62.

Shapiro, S.: 1980, 'On the Notion of Effectiveness', *History and Philosophy of Logic* **1**, 209–30.

Shapiro, S.: 1981, 'Understanding Church's Thesis', *Journal of Philosophical Logic* **10**, 353–65.

Shapiro, S. (ed.): 1985a, *Intensional Mathematics*, Amsterdam, North-Holland.

Shapiro, S.: 1985b, 'Epistemic and Intuitionistic Arithmetic', in S. Shapiro (ed.), *Intensional Mathematics*, pp. 11–46.

Smorynski, C.: 1991, 'Review of S. Shapiro, "Intensional Mathematics"', *Journal of Symbolic Logic* **56**, 1496–9.

Troelstra, A. S. and D. van Dalen: 1988a, *Constructivism in Mathematics. An Introduction*, Vol. I, North-Holland, Amsterdam.

Troelstra, A. S. and D. van Dalen: 1988b, *Constructivism in Mathematics. An Introduction*, Vol. II, North-Holland, Amsterdam.

Center for Logic, Philosophy of Science and Philosophy of Language
Institute of Philosophy
University of Leuven
Kardinaal Mercierplein 2
B-3000 Leuven
Belgium
E-mail: leon.horsten@hiw.kuleuven.ac.be