STEWART SHAPIRO

Understanding Church's Thesis, again

The paper tries to show that Church's thesis can be proved mathematically (though neither by a formal nor by a set-theoretic proof) if only we give up the temptation to force absolute certainty and absolute precision onto non-mathematical reality and realize that no given mathematical notion (let alone the notion of computability) enjoys absolute precision. Church/Turing formulations would support rather negative than positive claims about computability: if one shows that a given function f is not recursive, then Church's thesis will provide us with conclusive reason for concluding that f is not computable.

A number of recent articles challenge traditional views concerning Church's thesis (CT). Mendelson [1990] and Gandy [1988] claim that CT is susceptible of rigorous, mathematical proof and Gandy, at least, argues that CT has actually been proved. Turing's [1936] study of a human following an algorithm is cited as the germ of the proof. In fact, Gandy refers to (a version of) CT as "Turing's theorem". Sieg [1992] is a bit more guarded, but the conclusion is similar. "Turing's theorem" is the proposition that if *f* is a number-theoretic function that can be computed by a being satisfying certain determinacy and finiteness conditions, then *f* can be computed by a Turing machine[1].

The Mendelson/Gandy/Sieg arguments are, I believe, substantially correct, and they raise interesting and important questions concerning the nature of computability and, more generally, the relationship between mathematics and non-mathematical reality, as well as questions concerning the nature of proof, the centerpiece of mathematical epistemology. The purpose of this article is to use CT as a case study in order to pursue those questions.

---

[1] Turing argues that humans satisfy some of these conditions, but apparently Sieg considers this text to be less than proof (of which more later).

Acta Analytica 11, 1993; S. 59-77

The idea that there is a blurry boundary between mathematics and empirical science is, of course, widely held on the contemporary philosophical scene. However, the view is typically supported by global considerations, with metaphors like "the web of belief". It is rare for the holism to be illustrated and tested through the detail of specific examples. Moreover, the more traditional view that mathematics and science are somehow different in kind, in both subject matter and methodology, is also widely held (e.g., Field [1980], [1984]).

I take the liberty of using an old article of mine (Shapiro [1981]) as an exemplar of the erstwhile standard view concerning CT[2]:

> Computability is a property related to either human abilities or mechanical devices, both of which are at least prima facie non-mathematical. It is therefore widely agreed that the question of Church's thesis is not a mathematical question, such as the Goldbach conjecture ... That is to say, mathematicians do not seek to show either that CT follows from accepted laws of number theory or that it contradicts such laws. Nevertheless, both mathematicians and philosophers have offered various non-mathematical arguments either for or against the thesis. Goldbach's conjecture can be settled, if at all, only by mathematical argument, but CT can be settled, if at all, only by arguments that are, at least in part, philosophical.
> The general attitude is reflected in a mathematics book about knots:
>
> Mathematics never proves anything about anything except mathematics, and a piece of rope is a physical object and not a mathematical one. So before worrying about proofs, we must have a mathematical definition of what a knot is ... This problem ... arises whenever one applies mathematics to a physical situation. The definition should define mathematical objects that approximate physical objects as closely as possible ... There is no way to prove ... that the mathematical definitions describe the physical situation exactly. (Crowell and Fox [1963,3])

There are at least two groups of questions here. One concerns the metaphysical or semantic status of CT. Does it even have a (bivalent, non-trivial) truth value? The other concerns the epistemic status of CT. If it does have a truth value, is CT the kind of thing that can be proved or refuted mathematically? The questions are, of course, closely related, but I will focus on each one separately. The next section concerns metaphysics/semantics and the following concerns epistemology and proof.

[2] In the same article, I note a minority opinion (once expressed by Harvey Friedman during a panel discussion) that CT can be given a mathematical proof by constructing a set of axioms for computability and showing that these axioms are satisfied by all and only recursive functions. I return to this below.

1. Modality, metaphysics, and truth.

It is widely held that recursiveness, Turing computability, etc. are about as precise as notions can get. They are rigorously defined properties of functions of natural numbers. Is the extension of the informal property of computability just as precise? If not, then, depending on how vague properties are to be handled, either Church's thesis does not have a truth value, it has a non-standard truth value, or it is false. There is no question of a precise property exactly coinciding in extension with a vague one, let alone a question of how one can *prove* such things. Church [1936] himself wrote:

> This definition is thought to be justified by the considerations which follow, so far as positive justification can ever be obtained for the selection of a formal definition to correspond to an intuitive notion.

and the classic Rogers [1967, 20]:

> Church's thesis may be regarded as a proposal ... that we agree heretofore to supply certain previously intuitive terms (e.g., "function computable by algorithm") with certain precise meanings.

Call this the issue of *precision*.

Second, computability is, at least prima facie, a modal notion, as indicated by the suffix, "ability". We say that a function f is computable if one *can* compute all of its values or if it is *possible* for a human or machine to compute f (ignoring finite limitations on memory and lifetime). Recursiveness and Turing computability, on the other hand, at least appear to be non-modal. A Turing machine is a set of ordered quadruples with a certain structure, and a function f is Turing computable if *there is* a Turing machine with a given relation to f. The relation is defined in terms of sequences of configurations, which are ordered n-tuples with a certain structure. No modality here. Similarly, a function f is recursive if *there is* a finite sequence of functions whose last member is f and such that each member of the sequence is either one of a certain class of initial functions or bears a certain relation to earlier functions in the sequence. Again, no modality[3].

Thus, prima facie, CT asserts that a modal notion coincides in extension with a non-modal one. In contemporary philosophy, there are two traditions that would demur at this point. One of them, traced to Quine, is skeptical of modal notions altogether, suggesting that they are too vague or indeterminate for respectable scientific (or quasi-scientific) use:

> We should be within our rights in holding that no formulation of any part of science is definitive so long as it remains couched in idioms of ... modality. (Quine [1986, 33-34])

On such a view, the modal nature of computability underscores the problem of precision, and undermines the assertion that CT has a "definitive" truth value.

The other tradition, while not skeptical of modality as such, doubts that there can be any useful reduction of a modal notion to a non-modal one. Authors of this persuasion claim that modal operators are "primitive" or else they invoke some sort of possible worlds to explicate modality (see Lycan and Shapiro [1986]). Now, CT proposes a reduction, of sorts, of the modal computability to the non-modal recursiveness, and a proof of CT would establish the reduction. Call this the issue of *modality*[4].

[3] For details, see any textbook on computability. The classic Rogers [1967, p. 14] shows how "Turing machine" can be defined in traditional arithmetic terms (although this definition is not used):

> Let $T = \{0,1\}$ and $S = \{0,1,2,3\}$. Then a Turing machine can be defined as a mapping from a finite subset of $N \times T$ into $S \times N$. Here T represents the conditions of a tape cell, S represents operations to be performed, and N gives possible labels for internal states.

One can attempt to define the pre-formal computability in non-modal terms: A function is computable if there is an algorithm that computes it. Notice, however, that one cannot think of an "algorithm" here in terms of actual, concrete tokens. There aren't enough of them. If "algorithm" is short for "possible algorithm", then, of course, computability is still modal. A non-modal alternative is to take algorithms to be abstract objects.

[4] Those, like David Lewis (e.g., [1986]), who believe in the existence of possible worlds, can claim a "reduction" (of sorts) of modal notions to extensional notions. Necessity just *is* truth in all possible worlds, etc. However, these modal realists typically do not give a non-modal analysis of "possible world". In effect, that notion is primitive. Lewis claims that the straightforward alternative to modal realism is to invoke a "primitive modality". The situation with CT, and the other "theses" discussed below, is not like this. Ultimately, the "reducing theory" is just arithmetic and set theory.

Contrary to both of these traditions, the modal nature of computability does not automatically disqualify CT from having a non-trivial, determinate truth value, nor a rigorous proof. In fact, both traditions should be rejected. There are other cases, which are similar to Church's thesis, but are not at all problematic. Define a *chess-game-model* to be a sequence of strings (in standard chess notation) which denotes a series of legal moves of a complete game. One very simple chess-game-model is the string "P-K4, P-Q3; resigns". The notion of a chess-game-model is defined entirely in string-theory. It is not modal. Let the *chess thesis* be the statement that every possible game of chess is represented by a chess-game-model and every chess-game-model represents a possible game of chess[5]. Here we have a proposed reduction of a modal notion (possible game of chess) to a non-modal one (chess-game-model).

Notice that an advocate of the chess-game-thesis need not claim that the notion of "chess game model" somehow captures the meaning of "possible game of chess", or that it is an "analysis" of chess (whatever that would be). It is a claim of extensional equivalence, in the form "every A corresponds to a unique B, and vice versa". This is all the "reducing" that we want or need.

If understood this way, the chess thesis can hardly be denied. To be a little more careful, the chess thesis would only be denied by a nominalist, who holds that either there are no strings at all (since strings are abstract) or there aren't enough string tokens to represent every possible chess game[6].

The chess thesis is an exchange of modality for ontology. Instead of speaking of what is possible, we speak of what occurs in an abstract, but actual

[5] I do not wish to raise issues concerning whether someone who resigns before his second move has really played a game of *chess*. If the phrase "possible game of chess" is bothersome, then one can substitute something like "possible play according to the current rules of chess". The latter is a modal notion, and that is all that matters here.

[6] Unlike the situation with CT, we don't have problems with "the infinite" here, since there are only a finite number of possible chess games (according to rules for determining draws). Of course, there are still a lot of possible games, more than the number of string tokens in the solar system. Presumably, a nominalist could hold that for every possible game of chess there is a *possible* string that represents it, but this is not a reduction.

mathematical structure, a set of strings on a finite alphabet in this case[7]. As Putnam [1975] notes in a different context,

> [m]athematics has … got rid of *possibility* by simply assuming that, up to isomorphism anyway, all possibilities are simultaneously *actual* — actual, that is, in the universe of "sets".

In the case of Church's thesis, the problem of modality is resolved in similar fashion. The claim behind CT is that Turing machines somehow represent possible algorithms or possible machine programs, and sequences of Turing machine configurations represent possible computations. From this perspective, CT would hold only if, for every possible algorithm, there is a Turing machine that represents an algorithm that computes the same function. The thesis, then, is that the possibilities of computation are reflected accurately in a certain arithmetic, or set-theoretic structure.

One could, I suppose, make a similar claim about the algorithms behind λ-terms and recursive definitions, but if one thinks of CT along present lines, it is more natural to focus on Turing machines (as we just did). Those are supposed to be models of actual computing devices. Moreover, Turing's own arguments — the centerpiece of the purported "proof" of Church's thesis — constitute a study of a human computer following an algorithm, noting what sorts of moves are allowed, what abilities are presupposed, etc. Turing argues that anything such a person does can be simulated on a Turing machine.

Two historical asides: When Church first posed CT, he argued for it by noting that every computable function examined to date had been shown to be recursive, that the class of recursive functions is closed under certain operations, and that a number of different characterizations are extensionally equivalent. This "evidence" might be labeled "quasi-empirical". In a letter to Church around that time, Gödel regarded the proposal of CT as "thoroughly unsatisfactory". Church replied that "if [Gödel] would propose any definition of effective calculability, [Church] would undertake to prove that it was included in lambda-

[7] Incidentally, there is a trend in contemporary philosophy of mathematics that attempts to reverse this "exchange". The plan is to reduce ontology by invoking (primitive) modality. Instead of asserting that *there exists* a number with a given property, one speaks of what is possible, or of what one can construct. See, for example, Hellman [1989], Chihara [1990], and, for a response, Shapiro [1993]. Putnam [1967], [1975] suggests that there is no real difference between certain modal and certain ontological assertions. This is congenial with the present orientation toward CT, and the resolution of the problem of modality.

definability". But this would be more of the same kind of evidence already available, and there was plenty of that. Gödel seemed to prefer the rigor of conceptual analysis to the quasi-empirical methodology. This analysis was provided by Turing's work, which did convince Gödel (see Kleene [1981], [1987], Davis [1982], and my review, Shapiro [1990]).

Second, the foregoing interpretation, with reference to "analysis", is reminiscent of what has been called "Church's superthesis", an assertion that for any algorithm whatsoever, there is a Turing machine that computes the same function *the same way*:

> … the evidence for Church's thesis, which refers to *results*, to functions computed, actually establishes more, a kind of *superthesis*: to each … algorithm … is assigned a … [Turing machine] programme, modulo trivial conversions, which can be seen to define the same computation process as the [algorithm]. (Kreisel [1969, 177])

This is getting close to asking for the *meaning* of computability, demanding some sort of intensional equivalence between computability and Turing computability. For this line to be developed fully, we would need a better articulation of how to identify "computation processes" or, in other words, of what it is for two algorithms to compute the same function "the same way". But perhaps this need not stand in the way of the extensional thesis. Turing and Church do not state the superthesis, and perhaps the argument for CT can proceed without *completing* the "conceptual analysis".

To return to our theme, even if the modal nature of computability does not prevent CT from having a non-vacuous truth value, it still requires further considerations to establish that it does, and that this value is True. It is not self-evident that CT is exactly analogous to, say, the chess thesis. In particular, we need to be shown that the specific modal notions involved in computability are in fact sufficiently determinate to support a clear notion of computability. That's what a proof of CT would have to establish.

The problem of *other idealizations* is an extension of both the problem of precision and the problem of modality. The so-called "trivial" side of CT is that every Turing machine represents a possible algorithm. One could complain about feasibility. Some Turing machines do not represent possible algorithms because no human or machine could compute even one instance before the sun goes cold. The standard response is that we are to ignore (accidental) finite bounds on human lifetimes, the amount of paper or magnetic media in the universe, etc.

This is a standard idealization in mathematics, not unlike what is done in virtually any area of applied mathematics, geometry, etc. However, one can still wonder whether there is a sufficiently determinate property of computability that exactly corresponds to recursiveness. Is the idealization univocal and determinate?

Recall, for example, that there are more "realistic" mathematical definitions of possible computing devices. The crucial difference between a Turing machine and a finite-state machine is that a Turing machine has an unlimited "tape" which can be used for storage, scratch work, etc., while a finite state machine does not. There is no bound in advance on how much "work space" a finite state machine can have, but for each such machine, there is a limit to its space. By contrast, a Turing machine cannot run out of work space. Clearly, real computers, such as PC's or mainframes, are more like finite state machines than Turing machines. Real computers are being made with larger and larger storage capacities, both in terms of working memory and disk space, but for each such machine or each such network, there is a fixed limit to its work space — even if this limit is measured in gigabytes[8].

The point is that recursiveness or Turing computability represent more idealized models of computability than what may be called finite-state computability. And there are other models, such as push-down automata computability, and Turing computability in deterministic polynomial time (or space, or both). If we get even more "realistic" and put a fixed bound, in advance, on the amount of material available for *any* computation, we would get something like "computable by a finite state machine with at most N states" or "computable by a Turing machine with less than N squares of tape", for some fixed N. At least mathematically, these do not appear to be very interesting notions. For the most part, only finite, partial functions would be "computable". But, of course, only

[8] Turing's original [1936] analysis did not focus on mechanical computing devices, but rather on a human following an algorithm. In these cases, the idealization of Turing machines is that for any such person (or any such possible person), there is no limit on the amount of paper and pencils she can use in the course of the calculation and no bound on her lifetime and attention span. If we were to build in the assumption that, no matter what, there will always be some limit on the available materials (even if no specific limit is set for all cases), we would be closer to finite state computations.

finite, partial functions *are* computable, in the literal sense of "capable of computation". Again, the modality is non-trivial[9].

The problem of other idealizations will be addressed after we deal briefly with the epistemic side of our question.

2. What do we prove and what does a proof show?

In contemporary mathematical logic, there are several models of mathematical deduction, or proof. The most common construes a deduction to be a sequence of well-formed-formulas in a formal language, constructed according to certain rules. Such a deduction is a proof if the deductive system is sound and if the premises are interpreted as statements previously known to be true. Call such a sequence a "formal proof". Sometimes a proof is taken to be a derivation in Zermelo-Fraenkel set theory (ZF), or a sequence of statements that can be "translated" into a derivation in ZF. Call this a "ZF-proof".

Clearly, in claiming that CT is capable of proof, Mendelson and Gandy are not asserting the existence of a formal proof or a ZF-proof:

> The fact that it is not a proof in ZF or some other axiomatic system is no drawback; it shows that there is more to mathematics than appears in ZF. (Mendelson [1990, 233])

This is an insightful consequence of the Mendelson/Gandy/Sieg position. No doubt, the study of computability in Turing [1936], or anywhere else for that matter, could be cast in a formal deductive system or in ZF. But that would not be the end of the matter. The issue would then be to determine that the resulting derivation is a good "translation" of the informal arguments. Can *that* be established with a formal proof or a ZF-proof? The problem of evaluating the adequacy of formal translations of informal arguments is a wide and deep one, not to be solved here, but some remarks on the present case are in order.

[9] Recall that, even in ordinary cases, Wittgenstein, at least, was skeptical of there being any fact of the matter concerning which function a given person or mechanical device is computing (see, for example, Wittgenstein [1958], [1978] and Kripke [1982]). Functions are infinite, and no human or machine will ever compute more than a finite number of values. There is no fact of the matter concerning which features of the organism or device are relevant to the execution of the abstract algorithm, which features are incidental, and which features interfere with the execution (such as the decay of the parts).

The only non-logical term in the language of ZF is "$\in$", the sign for membership. To echo Crowell and Fox [1963], before worrying about whether CT can be proved in ZF, we would need a formulation of "computability" and either "recursiveness", "Turing computability", "λ-definability", etc. in the language of ZF. There are, or could be, good formulations of the latter notions in ZF, by following a number-theoretic formulation of Turing computability (see note 3) with a translation of number theory into ZF. There would be little room for doubt here, I presume (at this point in history). Formulating "computability" in the language of ZF is another story. How could we be sure that the proposed set-theoretic predicate really is an accurate formulation of the intuitive, pre-theoretic computability? Would we prove that? In effect, a statement that the proposed predicate does in fact coincide with computability would be the same sort of thing as CT, in that it would propose that a precise (now set-theoretic) property is equivalent to an intuitive one. We would be back where we started, philosophically[10].

A formal proof of CT would consist of a direct formulation of an argument in a formal deductive system. Presumably, one would begin with a formalization of number-theory, and add a predicate for computability, together with some axioms for it (see note 2). Then one would show in the deductive system that this predicate holds of all and only recursive functions. Here the "translation problem" would focus on the axioms for computability. There would be no *formal* guarantee that the axioms are both necessary and sufficient for computability. This question wouldn't be settled by a derivation in ZF or a formal deductive system, not without begging another question.

The conclusion, so far, is that if there is to be a mathematical proof of CT, the proof cannot be fully captured with a formal proof or a ZF-proof. If one *identifies* mathematical proof with formal proof or ZF-proof[11], then one can invoke *modus tollens* and accept the conclusion that CT is not a mathematical question. There is an essential "quasi-empirical" or "philosophical" side to it.

[10] If there were a theorem in ZF equating recursiveness with a set-theoretic formulation of computability, we would have more evidence for CT, or else evidence that the set-theoretic formulation is correct (or both). The indicated theorem would be the same sort of thing as the equivalence of recursiveness and Turing computability.

[11] The proposed identification is a lot like CT, equating a pre-formal notion (proof) with a precise, mathematical one (formal proof or ZF-proof). Moreover, the pre-formal notion of "proof" is at least prima facie modal. So, to be consistent, a holder of the received view should also hold that *this* identification is quasi-empirical or philosophical.

Against the received views, and with Mendelson, Gandy, and Sieg (and, in the general matter, Putnam [1975]), I submit that this is a false dilemma. CT is, in part, a quasi-empirical or philosophical question but that does not prevent it from being a mathematical question as well, capable of demonstration (or refutation) with whatever standards of rigor are operative in live mathematics. The proper conclusion of the foregoing considerations is that CT is not a *formal* (or ZF) matter, but, with Mendelson, there is more to mathematics, and to mathematical proof, than is dreamt of in ZF and in other formal deductive systems.

So what is "proof"? This is, of course, a deep problem in the philosophy of mathematics, and I have no new (positive) insights to convey here. For present purposes, let us define a "proof" to be a rationally compelling argument, one that a mathematician (*qua* mathematician) should find thoroughly convincing. It is not a question, of course, of what a given mathematician does find convincing. Errors, gaps, and fallacies abound, and people are often not convinced when they should be. Like any other epistemic notion, "proof" is inherently normative. There is no consensus on the nature of normativity, and (again) I have no plans on changing that situation. Normativity remains one of the more sticky problems on the agenda of contemporary philosophy. However, the notion of mathematical proof is one where some progress has been made. The present purpose is to lend some perspective to that progress.

Notice that, as presently construed, "proof" depends on context. What someone should find convincing depends on her training, on what is already known, etc. Moreover, a proof is not something that is immune from all conceivable skeptical challenges. Even if there is some notion of absolute rigor, independent of social context, mathematics as practiced does not need to adhere to such a standard. Moreover, the notion of "proof" is not necessarily precise.

In a collection of notes entitled "What does a mathematical proof prove?" (published in his [1978]), Imre Lakatos makes a distinction between the pre-formal development, the formal development, and the post-formal development of a branch of mathematics (see also my [1989]). Lakatos observes that even after a branch of mathematics has been successfully formalized, there are residual questions concerning the relationship between the formal deductive system (or the definitions in ZF) and the original, pre-formal mathematical ideas. How can we be sure that the formal system accurately reflects the original mathematical structures? These questions cannot be settled with a derivation in a further formal deductive system, not without begging the question or starting a regress — there would be new, but similar, questions concerning the new deductive system.

The crucial point here is that it does not follow that the residual questions are (merely) philosophical or quasi-empirical, or that the questions are in any way non-mathematical. In some cases one can and should regard the questions as settled. Moreover, this is the normal situation in mathematics. There is nothing unusual about CT.

To bolster this claim, Mendelson mentions other situations in mathematics which are like CT in the relevant respects, but which are not subject to the same doubts, prima facie (see also Shapiro [1981]). Here, I will only discuss a few cases.

Consider, first, the chess thesis, the assertion that for every possible chess game, there is an appropriate sequence of strings that represents it. It seems as clear as anything that this thesis is true, by whatever standards of rigor are prevalent in modern mathematics. Consider, for example, a "theorem" about possible chess games based on something like the chess thesis: It is not possible to force a checkmate with two nights and a king against a lone king. I submit that this is as certain as anything in mathematics, despite the modality. After grasping such a proof, it would be irrational to doubt this claim about possible chess games, just as irrational as doubting things correctly proved in informal number theory.

Near the top of this paper, I mentioned a mathematics book about knots (Crowell and Fox [1963]). The authors prove that a "figure-eight knot" cannot be transformed into an "overhand knot" without "tying" or "untying". All of the quoted expressions are given careful, topological definitions. The issue concerns the relationship between these definitions and pieces of rope. The authors may be right that there "is no way to prove … that the mathematical definitions describe the physical situation exactly", but it is quite clear that *something* about real knots has been proved beyond rational doubt. One can see that the mathematical "models" do in fact correspond to the structure of real (physical) knots — enough so that it would be irrational for someone to ignore the result and keep on trying to transform the one knot into the other. Furthermore, suppose that someone did claim that after ten hours of hard, concentrated work, he did in fact transform a figure-eight knot into an overhand knot. The rational conclusion (for us) to draw would be that he had (perhaps unknowingly) untied

and retied the rope. Why? Because we have *proved* that the proposed task is impossible[12].

Closer to home, Mendelson observes that there is little doubt that the so-called "trivial half" of CT, all recursive functions are computable, is established:

> The so-called initial functions are clearly … computable; we can describe simple procedures to compute them. Moreover, the operations of substitution and recursion and the least-number operator lead from … computable functions to … computable functions. In each case, we can describe procedures that will compute the new functions.

Mendelson concludes that this "simple argument is as clear a proof as I have seen in mathematics, and it is a proof in spite of the fact that it involves the intuitive notion of … computability". As an aside, notice that one bonus of this proof is that it allows us to see where the idealization from actual human or machine abilities comes in. There is no limit on the sequence of functions used to define a recursive function. We ignore, or reject, the possibility of a sorites situation.

For a final example, consider a small portion of Turing's study of a human following an algorithm. Turing shows that the alphabet involved in executing the algorithm must be finite. First, there is an upper limit to the size of a single symbol--any human (or mechanical device) will have *some* limit on the amount of space it can scan at one time[13]. This strikes me as a good premise, one that is clear and hard to doubt. It follows that "if we were to allow an infinity of symbols, then there would be symbols differing to an arbitrarily small extent". In a footnote, Turing suggests that under reasonable assumptions, there is a natural topology for the space of symbols, under which they form a conditionally compact space. The conclusion that there are only finitely many symbols follows from another premise (not stated) that there is *some* limit to the ability

[12] Suppose that someone transformed the knots in a few seconds, right in front of us. What would we conclude? Probably that it was a sleight of hand — the person cheated. I am not saying, however, that it is impossible for us to change our mind. It is conceivable (barely) that someone could get us to see that we were wrong about the transformation theorem, and thus that we had made a mistake in the topological definitions. It is also (barely) conceivable that someone could show us how to force checkmate with two knights and a king against a lone king. I don't think this undermines the present considerations. Infallibility is not required in mathematics.

[13] This is among the "finiteness conditions" mentioned by Sieg [1992].

to discriminate symbols. Again, the argument seems about as good as anything in mathematics[14].

Gandy [1988, 82] draws a deep conclusion about the general matter at hand:

> Turing's analysis does much more than provide an argument for Church's thesis; *it proves a theorem* ... The proof is quite as rigorous as many accepted mathematical proofs — it is the subject matter, not the process of proof, which is unfamiliar. However — as with published mathematical proofs — there are gaps which need to be filled in.

One need not go that far, at least not yet. The foregoing considerations show that CT is the *kind* of thing that can be proved mathematically. It does not follow that CT has in fact been proved. Given the above rough characterization of "proof" as something like "rationally compelling argument", one should expect there to be borderline cases of proofs. Perhaps the argument for CT is such a borderline.

3. What to make of this.

I conclude with a few remarks on what it is that would be proved if CT were in fact established. This depends on what CT is, and, more generally, on the relation between mathematics and the physical world, not to mention the nature of mathematics itself. We return to the problem of precision and the problem of other idealizations.

It is widely held, often implicitly, that mathematics deals exclusively with absolutely precise notions and concepts. No vagueness abounds here. One way to square this with the provability of CT would be to hold that computability is also a precise property of number-theoretic functions. In Shapiro [1981], I note that this view is congenial with a structuralist account of the application of mathematics. In this case, the idea is that there are definite mathematical structures underlying possible computing machines and/or the human ability to calculate. Rogers [1967, 1] may have had something like this in mind, when he re-

[14] Kreisel [1967] is a detailed, insightful account of mathematical arguments that involve intuitive, pre-theoretic notions. Kreisel calls this "informal rigor". The most well-known instance of this is Kreisel's proof that, for first-order logic, the pre-theoretic notion of validity coincides with its model-theoretic formulation in set theory.

ferred to computability as "the informal *mathematical* notion of function computable by algorithm" (emphasis mine)[15]. In writing about CT, Post [1941] suggested that

> ... for full generality a complete analysis would have to be made of all the possible ways in which the human mind could set up finite processes.
>
> ... we have to do with a certain activity of the human mind as situated in the universe. As activity, this logico-mathematical process has certain temporal properties; as situated in the universe it has certain spatial properties.

On a view like this, a proof of CT would establish that the precise computability is coextensive with recursiveness.

It is, of course, more common, and perhaps more natural, to think of computability as a vague property from ordinary language. This, together with the precision of mathematics, entails the received view that CT is not a mathematical matter.

This argument at least suggests that there is a difference in kind between mathematics and the rest of our intellectual activity. As above, this distinction is rejected here. Mathematical language has its roots in ordinary language, and those roots cannot be severed. The present concern is to square this holism with the vagueness of ordinary language. One can, I suppose, argue that ordinary or scientific language is not vague, appearances notwithstanding. A more intriguing option is to reject, or at least temper, the precision of mathematical notions. Mendelson does just that:

> The concepts and assumptions that support the notion of partial-recursive function are, in an essential way, no less vague and imprecise than the notion of effectively computable function; the former are just more familiar and part of a respectable theory with connections to other parts of logic and mathematics.

Mendelson does not develop this idea further, and if the claim is taken literally, one can certainly challenge it. I don't know of a border-line case of, say, a

[15] In Section 1 above, there is a passage from Rogers [1967] indicating that computability is vague. There may be some ambivalence here, or else Rogers rejects the view that mathematics deals exclusively with precise notions.

natural number, or a number-theoretic function[16]. However, the notions involved in recursion theory — or any other branch of mathematics for that matter — have not always been so clear and precise. The central notions of "set", "function", and "infinite" all have long and troubled histories. Even a cursory look at the growth of mathematical ideas reveals a lot of uncertainty, ambivalence, vagueness, and plain unclarity. Nevertheless, it is clear, if anything is, that the current formulations of these notions are correct. If nothing else, current definitions capture an important and non-arbitrary "natural kind" underlying the previous *mathematical* discussions. That is, the original vagueness of the notions did not preclude mathematicians, *qua* mathematicians, from discussing the notions, "defining" them, and proving theorems about them. In large part, that's what mathematics is all about.

Similar remarks apply to the "theses" alluded to above, and in Mendelson [1990] and Shapiro [1981], which are claimed to be analogous to CT. In each case, there are arguments, which are often thoroughly convincing, that a given mathematical notion is a clear "natural kind" underlying a notion or concept from everyday language, from natural science, or from other parts of mathematics (a possible chess game, a knot, a limit, a symbol involved in an algorithm, etc.). The definitions are in no way arbitrary, and one is completely justified in accepting theorems about the "definiens" to represent facts about "definiendum"[17]. We can and do prove theses a lot like CT.

This does not preclude the possibility that the pre-theoretic notions may have other definitions, incompatible with the accepted ones. The alternatives may even be useful for some purposes. The notion of continuity comes to mind, with its separate formulations as uniform continuity and pointwise continuity. In a sense, both are correct. There may also be other formulations of "possible game of chess" that take feasibility into account. For the purpose of devising

[16] George Schumm suggests the following: Consider a room that contains only two (clearly) bald men and one borderline case of a bald man. Let n be the denotation of "the number of bald men in the room". Isn't n a borderline case of a natural number? I would say that although it is indeterminate which number the expression denotes, we do not have a borderline natural number here. Unlike the case with "bald", the vagueness here is not in the property "natural number", but rather in the denotation relation.

[17] To take one more example, Lakatos [1976] is a sketch of the historical development of the notion of "polyhedron". The examples make it plain that the original notion was quite vague — borderline cases abound — and yet the current formulation, in set-theoretic terms, clearly captures an essential property underlying the original mathematical ideas.

and ruling out strategies against human opponents, such a formulation may be more useful. But it would not undermine the chess thesis.

Turning to computability, I would suggest that the Church/Turing (et al.) formulations support *negative* claims about computability. If one shows that a given function f is not recursive, then, with CT, that is conclusive reason to conclude that f is not computable. However, since feasibility is ignored, the current formulation is not as useful for establishing positive claims. If one shows that a given function is recursive, that does not, by itself, give us a good reason to think that the function "can" be computed, in any realistic sense. Other formulations, like finite state computability, or Turing computability in polynomial time, may be better for this. But this observation does not undermine CT.

In short, even if one holds that computability is vague, and notes that there are several incompatible "precisifications", one should not conclude that CT is beyond the purview of mathematics. One can think of CT as an assertion that recursiveness represents *the* clear "natural kind" underlying computability, or else *a* clear natural kind underlying computability. One can, and rationally should, use the notion of recursive function to clear the vagueness from computability, and one can know this with whatever certainty anything in mathematics enjoys. Moreover, one can know that facts about recursiveness represent facts about computability, and vice-versa.

In fact, the identification goes both ways. If one shows that a function f is computable (by giving an algorithm, for example), one can conclude without further ado that f is recursive. Such an inference, sometimes called "argument by Church's thesis", is the contrapositive of the above statement concerning negative results about computability (see Shapiro [1983]). The technique was proposed as early as Post [1944]. Rogers [1967] is built around argument by Church's thesis and, to invoke the theme of this paper, no one doubts that Rogers [1967] is a *mathematics* book.

In conclusion, the resolution of the issue of Church's thesis is not to put it outside of mathematics, nor to force absolute certainty and absolute precision onto non-mathematical reality. Rather, one realizes that mathematics itself does not enjoy absolute certainty and its notions do not enjoy absolute precision.

Acknowledgements:

I would like to thank Michael Detlefsen, Jill Dieterle, Penelope Maddy, and George Schumm for useful and insightful criticism of an earlier version of this project.

References

Chihara, C [1990], *Constructibility and mathematical existence*, Oxford, Oxford University Press.

Church, A [1936], "An unsolvable problem of elementary number theory", *American Journal of Mathematics 58*, 345-363; also in Davis [1965], 89-107.

Crowell, R, and R Fox [1963], *Introduction of knot theory*, Boston, Ginn and Company.

Davis, M [1965], *The undecidable*, Hewlett, New York, The Raven Press.

Davis, M [1982], "Why Gödel didn't have Church's thesis", *Information and control 54*, 3-24.

Field, H [1980], *Science without numbers*, Princeton, Princeton University Press.

Field, H [1984], "Is mathematical knowledge just logical knowledge?", *The Philosophical Review 93*, 509-552.

Gandy, R [1988], "The confluence of ideas in 1936", in R. Herken ed., *The universal Turing machine*, New York, Oxford University Press, 55-111.

Hellman, G [1989], *Mathematics without numbers*, Oxford, Oxford University Press.

Kleene, S [1981], "Origins of recursive function theory", *Annals of the History of Computing 3* (no. 1), 52-67.

Kleene, S [1987], "Reflections on Church's thesis", *Notre Dame Journal of Formal Logic 28*, 490-498.

Kreisel, G [1967], "Informal rigour and completeness proofs", *Problems in the philosophy of mathematics*, ed. by I. Lakatos, Amsterdam, North Holland, 138-186.

Kreisel, G [1969], "Some reasons for generalizing recursion theory", *Logic colloquium '69*, Amsterdam, North Holland, 138-198.

Kripke, S [1982], *Wittgenstein on rules and private language*, Cambridge, Massachusetts, Harvard University Press.

Lakatos, I [1976], *Proofs and refutations*, ed. by J. Worrall and E. Zahar, Cambridge, Cambridge University Press.

Lakatos, I [1978], *Mathematics, science and epistemology*, ed. by J. Worrall and G. Currie, Cambridge, Cambridge University Press.

Lewis, D [1986], *On the plurality of worlds*, Oxford, Basil Blackwell Ltd.

Lycan, W and S Shapiro [1986], "Actuality and essence", *Midwest Studies in Philosophy 11*, 343-377.

Mendelson, E [1990], "Second thoughts about Church's thesis and mathematical proofs", *The Journal of Philosophy 87*, 225-233.

Post, E [1941], "Absolutely unsolvable problems and relatively undecidable propositions", in Davis [1965], 338-433.

Post, E [1944], "Recursive sets of positive integers and their decision problems", *Bulletin of the American Mathematical Society 50*, 284-316; also in Davis [1965], 305-337.

Putnam, H [1967], "Mathematics without foundation", *Journal of Philosophy 64*, 5-22.

Putnam, H [1975], "What is mathematical truth?" in *Mathematics, matter and method: Philosophical Papers Volume 1*, by Hillary Putnam, Cambridge, Cambridge University Press, 60-78.

Quine, W V O [1986], *Philosophy of logic*, second edition, Englewood Cliffs, New Jersey, Prentice-Hall.

Rogers, H [1967], *Theory of recursive functions and effective computability*, New York, McGraw-Hill.

Shapiro, S [1981], "Understanding Church's thesis", *Journal of Philosophical Logic 10*, 353-365.

Shapiro, S [1983], "Remarks on the development of computability", *History and Philosophy of Logic 4*, 203-220.

Shapiro, S [1989], "Logic, ontology, mathematical practice", *Synthese 79*, 13-50.

Shapiro, S [1990], review of Kleene [1981], Kleene [1987], and Davis [1982], *Journal of Symbolic Logic 55* (1990), 348-350.

Shapiro, S [1993], "Modality and ontology", *Mind*, forthcoming.

Sieg, W [1992], "Mechanisms and search: Aspects of proof theory", Carnegie Mellon Report.

Turing, A [1936], "On computable numbers, with an application to the *Entscheidungsproblem*", *Proceedings of the London Mathematical Society 42*, 230-265; also in Davis [1965], 116-153.

Wittgenstein, L [1958], *Philosophical investigations*, tr. by G. E. M. Anscombe, New York, MacMillan Publishing Company.

Wittgenstein, L [1978], *Remarks on the foundations of mathematics*, tr. by G. E. M. Anscombe, Cambridge, Massachusetts, The MIT Press.