
Can Church's Thesis be Viewed as a Carnapian Explication?

Paula Quinon

Abstract Turing and Church formulated two different formal accounts of computability that turned out to be extensionally equivalent. Since the accounts refer to different properties they cannot both be adequate *conceptual analyses* of the concept of computability. This insight has led to a discussion concerning which account is adequate. Some authors have suggested that this philosophical debate – which shows few signs of converging on one view – can be circumvented by regarding Church's and Turing's theses as *explications*. This move opens up the possibility that both accounts could be adequate, albeit in their own different ways. In this paper, I focus on the question of whether Church's thesis can be seen as an explication in the precise Carnapian sense. Most importantly, I address an additional constraint that Carnap puts on the explicative power of axiomatic systems - an axiomatisation explicates when it is clear which mathematical entities form the theory's intended model - and that implicitly applies to axiomatisations of recursion theory used in Church's account of computability. To overcome this difficulty, I propose two possible clarifications of the pre-systematic concept of "computability" that can both be captured in recursion theory, and I show how both clarifications avoid an objection arising from Carnap's constraint.

Keywords explications · Rudolf Carnap · the Church–Turing thesis · Church's thesis · computability · axiomatic systems · structuralism

Introduction

Mathematicians' project of formalizing the concept of effective computability in the 1930s had various motivations. Turing wanted to solve the *Entscheidungsproblem* – the decision problem regarding provability of first-order sen-

Paula Quinon
International Center for Formal Ontology
Warsaw University of Technology
E-mail: p.quinon@ans.pw.edu.pl

tences – formulated by Hilbert in 1928 [25]. Gödel and Church were interested in specifying the concept of formal system and therefore needed a sharp concept of effective method to account for finite reasoning in such systems. In particular, Church and his group searched for effective methods of defining functions on natural numbers, and thereby, a way of singling out the class of functions that can be effectively computed.¹

Various models of computation were formulated in response to these objectives. Church’s thesis, formulated in 1936, identifies the pre-systematic concept of “effectively computable” or “calculable” with the property of “being generally recursive” defined for functions on natural numbers.² Turing’s thesis, formulated in the same year, translates this pre-systematic concept into “being computable by a Turing machine”. The two definitions were soon shown to be extensionally equivalent. Hence, the “Church-Turing thesis”.³

However, the fact that general recursiveness and Turing computability are extensionally equivalent does not mean that they capture the same properties. This raises the question of which of the two accounts, Church’s or Turing’s, if any, provides an adequate conceptual analysis of the concept of effective computability, where by “conceptual analysis” I mean an attempt to clarify a given concept by identifying its conceptual parts. On this understanding the two theses differ significantly in many aspects. For instance, Church’s thesis states that effective computability can be analyzed in terms of properties of functions defined on natural numbers understood as abstract objects. Turing’s thesis, by contrast, expresses that effective computability can be analyzed in terms of properties of functions defined on strings of symbols. Thus, the two theses provide very different analyses of the concept in question. If one assumes, as is often tacitly done, that only one analysis of a given concept can be correct, once the latter has been properly disambiguated, then Church’s analysis and Turing’s analysis cannot both be adequate.

Gödel, for whom the problem of defining computability was “an excellent example [...] of a concept which did not appear sharp to us but has become so as a result of a careful reflection” (reported by Wang in 1974 [50, page 84]), thought that it was Turing’s analysis that captured the concept of computability in the most adequate way. He claimed that it is “absolutely impossible that anybody who understands the question and knows Turing’s definition should

¹ Turing provides a useful clarification of the concept of effective computability he aimed at explaining in (1936). For an overview of the conceptual motivations of Church, Kleene and Rosser see [44]. Gödel’s idea of how to formalize the concept of computability can be found in [40]. Tennant (2015 [47, page 144]) suggests a “practical” reason why in the early 1930s, the intuitive, pre-formal notion of a computable function was “crying out for explication”, because it was thought to be important in urgent military applications of high-speed cryptography methods.

² Sieg in [44] claims, contra Davis [14], that Church himself never wanted to formulate the thesis for λ -calculus.

³ It is common in general discussions to refer to both theses under the heading of “Church’s thesis”. Church’s thesis then covers other models of computation too, *e.g.*, Church’s lambda calculus (1936) or Post’s combinatory processes (1936).

decide for a different concept" [50, page 84]. Similarly, in (193? [23, page 168])⁴ Gödel wrote that "the correct definition of mechanical computability was established beyond any doubts by Turing". Moreover, Gödel also held that it is exactly the adequacy of Turing's thesis as a conceptual analysis that establishes the correctness of all the other equivalent mathematical definitions of computability [23, page 168]. Even Church himself acknowledged the conceptual advantages of Turing's definition, writing in 1937 that when it comes to conceptual adequacy of his clarification of the concept of computability:

computability by a Turing machine [...] has the advantage of making the identification with effectiveness in the ordinary (not explicitly defined) sense evident immediately. [11, page 43].

Despite Gödel's view on the special status of Turing's thesis as a conceptual analysis, the analysis itself did not attract much attention at the time of its formulation. As Shagrir puts it:

Turing's argument has been fully appreciated only recently. Turing's contemporaries mention the analysis, along other familiar arguments, *e.g.*, the confluence of different notions, quasi-empirical evidence, and Church's step-by-step argument but do not ascribe it any special merit. Logic and computer science textbooks from the decades following the pioneering work the 1930s ignore it altogether. [40, pages 9–10].

The interest in the conceptual issues related to computability was renewed at the end of the last century. The most recent discussions about the primacy of one account of computability over the other can be found in Soare (1996 [46]), Shapiro (1982 [41]), Gandy (1980 [20], 1988 [21]), Rescorla (2007 [37], 2012 [38]), Copeland & Proudfoot (2010 [12]), Quinon (2014 [36]), and also, the most recently, in the volume edited for Turing's 100th anniversary by Juliet Floyd and Alisa Bokulich (2017 [19])⁵. Various arguments are put forward by both camps. Some of the authors formulate arbitrary affirmations, as Gandy (1988 [21]) who writes that Turing's work in providing the definitive meaning of "computable function" is "a paradigm of philosophical analysis: it shows that what appears to be a vague intuitive notion has in fact a unique meaning which can be stated with complete precision" [21, pages 84–86]. Other authors advance sophisticated arguments, such as Rescorla (2007 [37], 2012 [38]), who advocates for the supremacy of Church's thesis by arguing that the problem of defining semantics for Turing computability leads to either extensional inadequacy or circularity. In consequence, claims Rescorla, Turing's thesis fails to be an analysis of computability at all. Rescorla refers to the well known problem, sometimes called "the semantical Halting problem", according to which if there were no "external" constraints on the choice of such a semantics – for instance, constraints on possible denotation functions that can be used for

⁴ This paper, published in the third volume of Gödel's *Collected Works*, is taken from *undated* handwritten notes in English, most probably for a lecture, found in a spiral notebook in the *Nachlass*.

⁵ I am grateful to an anonymous reviewer for reminding me about this publication.

mapping from numerals to numbers – a Turing machine would compute – under some “deviant semantics” – non-computable functions on natural numbers (Shapiro 1982 [41], Copeland & Proudfoot 2010 [12], van Heuveln 2000 [26]).

Disagreements about the proper analysis of important concepts tend to linger on indefinitely, as witnessed by many similar debates in the history of philosophy. A way out of a possible philosophical standoff would be to view Turing’s and Church’s theses not as competing conceptual analyses, but rather as explications – *i.e.*, rational reconstructions or specifications of an intuitive concept – in the sense of Carnap (1950 [7]).⁶ While there can be only one adequate conceptual analysis of a given concept, there may be several equally fruitful explications of the same pre-systematic concept. In consequence, one can defend both attempts – Church’s and Turing’s – as good explications of effective computability, and there is no need to argue for one to the disadvantage of the other.

The idea that the Church-Turing thesis is an explication has already been put forward by several authors. Some do not use the term “explication”, but speak of “rational reconstruction” instead (Mendelson 1990 [32, page 229], Schulz 1997 [39, page 182]); others classify the thesis as an explication, but only in an informal sense (Machover 1996 [31, pages 259-260]).

Tennant [47, chapter 10] presents the Church-Turing thesis as a paradigmatic example of a successful explication in both an informal and a strictly Carnapian sense. The reason that Tennant brings to support his hypothesis is rather traditional, as he says that the explicative power of the Church-Turing Thesis is confirmed by the fact that many formal characterizations of the intuitively computable functions coincide in extension. What is interesting in Tennant’s approach is that the CTT is a “founding instance” of the method, which means for Tennant that “all subsequent explications of other concepts may be measured against it, when assessing their degree of success” [47, page 139] and that philosophers should “emulate this achievement in yet other conceptual domains” [47, page 149]. He says even that Carnap’s “monumental work on the foundations of the theory of probability (Carnap 1950), in which he introduced, for the first time, the very idea of explication” [47, page 150] should be seen as a follow up to the endeavour of the logicians and foundation-alists of the 1930s. From Tennant’s presentation one could get the impression that Carnap may have been inspired by Church and Turing, but Carnap – in fact – never mentions computability in the context of explication⁷. I will have reason to return to this surprising fact and its probable cause below.

⁶ My current purpose does require me to account for all the details distinguishing various understandings of what “analysis” and what “explication” mean. For instance, Floyd (2012 [18]) accounts for differences between Turing’s understanding of his own work – that she calls “Turing’s ideal of explication and analysis” – and the way Carnap proposed to introduce new concepts to the scientific language. Similarly, Carnap’s ideal of explication differs from similar notions found in Quine, Strawson, Turing or Wittgenstein. I will not get into details of these discussions as arguments that I present apply already at a higher level of generality [49].

⁷ The only place where Carnap refers to Church’s and Turing work on computability is the note on page 79 of *Meaning and Necessity* (1947 [6]). Carnap highlights equivalence

Floyd (2012 [18]) concentrates on Turing's thesis – as clearly distinguished from Church's thesis – and argues that it satisfies Carnap's criteria for a good explication. She writes that “Turing's explication of this class of functions is distinctively vivid and simple: it offers similarity between explicandum and explicans, precision of formulation within systematic science, fruitfulness in the formulation of general laws, and remarkable simplicity” [18, pages 34].

Murawski & Woleński (2006 [33]) refer explicitly to Carnap's method of explication. They discuss the philosophical status of Church's version of the thesis. They claim that the most adequate way of thinking of the theoretical status of the thesis is in terms of its being a Carnapian explication. They argue that this is more satisfactory than treating it as an empirical hypothesis, or an axiom, or a theorem, or a definition.⁸ Murawski & Woleński contend that, as is required from a Carnapian explication, the thesis consists in replacing “an intuitive and not quite exact concept by a fully legitimate mathematical category” [33, page 322]. In support, they claim that “Church examined the standard (stock) use of the term ‘computable function’ in the language of informal mathematics” [33, page 324]. They state [33, page 311] that “[a] useful notion in providing intuitions concerning effectiveness is that of an algorithm”, so – as I would say following Carnap's methodology – the concept of an algorithm can be seen as a clarification of the pre-scientific concept. They have no doubts that exactness and fruitfulness are satisfied⁹, and state that similarity and simplicity¹⁰ might be more problematic, but it “does not need to worry users of (CT)” [33, page 322].

In this paper, I provide a detailed and careful study of Church's thesis – as distinguished from Turing's thesis and other models of computability – from the point of view of Carnap's methodology. The argument that the axioms of recursion theory explicate the concept of computability is more demanding than the one necessary in the case of Turing's version. The complicating factor is that Carnap committed himself to the view that an axiomatic theory might not be suitable for providing an explication of a concept. This follows from his view that semiformal, semi-interpreted axiomatic systems¹¹ – where non-logical constants are not properly clarified and interpreted – are unsuitable forms for explicating pre-systematic concepts. Since recursion theory is standardly developed in the form of a semiformal, semi-interpreted system, Carnap's view precludes interpreting Church's thesis as an explication at all,

between various formal models of computation, but does not say anything about Church's thesis. I am grateful to an anonymous reviewer for drawing my attention to the latter.

⁸ Murawski & Woleński (2006 [33]) suggest various possible theoretical statuses for the Church-Turing thesis in the context of analyticity/syntheticity and a priority/a posteriority.

⁹ One might object that the authors do not illustrate their investigations by examples.

¹⁰ In section 1 of this paper we explicate in detail all Carnapian terminology. Here a reader less accustomed with the theory of explications can benefit from the presentation while relying on her intuitive understanding of its meanings.

¹¹ Again, full explications of differences between various axiomatic systems are explained later, in section 3.

let alone a good one. This difficulty, which I will refer to as *Carnap's implicit objection*¹², seems to have escaped other authors.

My focus will be on the question of whether Carnap's implicit objection can be overcome in the case of Church's thesis. More precisely, I study Carnap's objection to semiformal, semi-interpreted axiomatic systems and I show how the objection can be overcome to enable recursion theory to provide a Carnapian explication of computability.

In **Section 1** ("Carnap's method of explication") I analyse details of the process of explication (*i.e.*, clarification stage, context choice, specification stage) and the requirements that enable assessment of the adequacy of an explication. In **section 2** ("Church's thesis as a Carnapian explication") I verify that Church's thesis follows a Carnapian process of explication and I show that it satisfies Carnap's requirements of adequacy. In **section 3** ("Axiomatic systems as explications") I search for those features of axiomatic systems that provides them with an explicative power. **Section 4** ("Carnap's implicit objection and arithmetic") is devoted to Carnap's examples of an explicating axiomatic system – Frege's arithmetic – and a non-explicating one – Peano's arithmetic. Using conclusions from the previous section, I argue that Peano's Arithmetic also explicates the concept of natural number. Finally, in **section 5** ("Multiplicity of clarification of computability"), I indicate two possible ways to justify the claim that Church's thesis is an explication.

1 Carnap's method of explication

While going through the state of art in the introductory section, I referred to the Carnapian method of explication in an informal way. This section is devoted to introducing this method in a systematic way.

The method of explication was proposed by Rudolf Carnap, most extensively in his *Logical Foundations of Probability* (1950 [7]), as a procedure for introducing new concepts to scientific or philosophical language. By the method of explication, Carnap explains,

we mean the transformation of an inexact, prescientific concept, the *explicandum*, into a new exact concept, the *explicatum*. [7, page 3].

It has been formulated in the context of what is known as "the paradox of analysis"¹³. Carnap's method of explication resolves the paradox as it consists

¹² I am indebted to Erik J. Olsson for suggesting this term.

¹³ According to Lavers (2013 [30]) the term "the paradox of analysis" comes from Langford (1942/1968 [29]) and was formulated as a commentary on G.E. Moore's philosophy .

[t]he paradox of analysis can be explained as follows. Suppose one gives an analysis of some concept A by saying that to be A is to be B. If A and B have the same meaning, then the analysis is uninformative. If A and B have different meanings, then the analysis is incorrect. While many might see the paradox of analysis as something of a curiosity, Carnap took it to rule out a certain view on analysis. [30, page 226].

in replacing an intuitive notion by a scientific one without requirement to preserve the exact extensional and intensional adequacy. Thus, an explication can be seen as a mapping from an informal (or prescientific, or pre-systematic) domain – where terms usually do not have any unique, precise and specialized meaning – to a formal (or scientific) domain – where terms are carefully defined in the context of other terms that reached the precise meaning earlier in the process of explication, or have been accepted as primitive. To illustrate this process Carnap proposes thinking about a pocketknife:

A natural language is like a crude, primitive pocketknife, very useful for a hundred different purposes. But for certain specific purposes, special tools are more efficient, e.g., chisels, cutting machines, and finally the microtome. If we find that the pocket knife is too crude for a given purpose and creates defective products, we shall try to discover the cause of the failure, and then either use the knife more skilfully, or replace it for this special purpose by a more suitable tool, or even invent a new one. [Strawson's] thesis is like saying that by using a special tool we evade the problem of the correct use of the cruder tool. But would anyone criticize the bacteriologist for using a microtome, and assert that he is evading the problem of correctly using the pocketknife? [9, pages 937-938].

In many ordinary life situations a pocketknife is an extremely useful tool, however it is useless in more specific situations that necessitate a higher degree of precision. When one wants to cut out a sophisticated wooden form, one will rather use, for example, a stainless steel laser woodcutter with a micro fence system. The way in which one should develop the language of sciences, according to Carnap, is analogous to how the pocketknife is, in some contexts, replaced by a more specialized tool. One should adjust the tool and instead of an imprecise explicandum use a precise explicatum well suited for a specific context of use; the explicandum, however, should remain in use outside the targeted context of application. Similarly, one will keep her pocketknife in a pocket and use it daily, but one will use a highly precise woodcutter to make a decoration for one's bedroom.

The method of explication, as Carnap sees it, consists of two steps:

1. The *clarification* of the explicandum.
2. The *specification* of the explicatum.

The rationale for clarification is that a given term may have many different meanings in ordinary language. Unless one of these meanings is clearly picked out from the start and the context of its use is clearly indicated, it is unlikely that the method of explication will yield a useful result. Clarification serves this purpose. As Carnap explains,

[a]lthough the explicandum cannot be given in exact terms, it should be made as clear as possible by informal explanations and examples. [7, page 3].

In this paper, I argue that it is exactly the stage of clarification, which allows us to overcome Carnap's worries relating semiformal, semi-interpreted axiomatic systems and allows us to formulate an explicating implicit definition. I highlight this point, because – despite its huge importance – the clarification stage easily gets underestimated. Already Carnap noted that his commentators show a tendency to think that in the nature of the explicandum there is an inherent inexactitude, that an explicandum can never be given in exact terms, and that – therefore – it does not matter in which words and to what extent it is clarified. According to Carnap such a view is “quite wrong”:

On the contrary, since even in the best case we cannot reach full exactness, we must, in order to prevent the discussion of the problem from becoming entirely futile, do all we can to make at least practically clear what is meant as the explicandum. What X means by a certain term in contexts of a certain kind is at least practically clear to Y if Y is able to predict correctly X's interpretation for most of the simple, ordinary cases of the use of the term in those contexts. It seems to me that, in raising problems of analysis or explication, philosophers very frequently violate this requirement. They ask questions like: ‘What is causality?’, ‘What is life?’, ‘What is mind?’, ‘What is justice?’, etc. Then they often immediately start to look for an answer without first examining the tacit assumption that the terms of the question are at least practically clear enough to serve as a basis for an investigation, for an analysis or explication. [7, page 4].

The fact that the terms in question are “unsystematic and inexact” does not mean that there is no way of approaching intersubjective agreement regarding their “intended” meanings. The procedure suggested by Carnap is to indicate the meaning by some well-chosen examples, perhaps supplemented by, to use his words, “an informal explanation in general terms” [7, page 4].

At this stage, this might be done in relatively crude terms by indicating what is to be included in the scope of reference or extension of the term, and what is intended to be excluded from it. To illustrate this point, Carnap invites the reader to consider several examples. For instance, he says, take the term “salt”. Suppose one wishes to explicate this concept. Then one might say, for example, “I mean by the explicandum ‘salt’, not its wide sense which it has in chemistry but its narrow sense in which it is used in the household”. A possible explicatum for this term might be “sodium chloride”, or NaCl, in the language of chemistry. Or, to take another of Carnap's examples, suppose one wishes to explicate the term “true”. A clarification of the explicandum may involve stating that one does not target the meaning that “true” has in phrases like “a true democracy” or “a true friend”, but rather the meaning it has in phrases like “this sentence is true”, “what he just said is true”, and so on. Then one would specify a formal theory of truth – as Carnap himself suggests – the one formulated by Tarski. Again, in making these choices one has not yet explicated the term “true”; one has only clarified its intended meaning.

All explanations of this kind serve only to make clear what is meant as explicandum; they do not yet supply an explication, say a definition of the explicatum; they belong still to the formulation of the problem, not yet to the construction of an answer. [7, page 4].

A clarification of the explicandum enables the next step of the explication process, a specification of the explicatum and formulation of the exact concept in the targeted context. This exact concept shall in some way “correspond” to the intuitive, everyday concept. But what does this relation of correspondence entail? It is obvious that one cannot hope for complete similarity in meaning. After all, the whole point of explication is, in a sense, to diverge from the meaning of the intuitive concept by introducing a more exact correlate that is hopefully more useful in a given scientific or formal context. This is confirmed by actual scientific practice, for example, analytic geometry reveals little similarity to intuitive geometrical concepts like “line”, “distance”, “angle”, and so on, but it still achieves an adequate explication of those concepts. This is because, it introduces them in the precise framework of arithmetic using a coordinate system, which, in consequence, opens up the possibility of applying algebraic and analytic methods to investigate geometrical concepts and their relations in a spectacularly fruitful manner.

Since the method of explication requires neither extensional nor intensional adequacy between explicandum and explicatum, instead of asking whether the solution is right or wrong one should ask whether it is satisfactory for the purposes for which it is to be used. In this way, generalizing from prototypical examples of scientific definitions, Carnap arrives at the following four general requirements on an explicatum. Those requirements enable an assessment of the adequacy of the explication and also enable a comparison of various explications of the same clarification of the concept.

1. **Similarity.** “The explicatum [the thing that explicates] is to be *similar to the explicandum* [the thing that is explicated] in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used; however, close similarity is not required, and considerable differences are permitted.”
2. **Exactness.** “The characterization of the explicatum, that is, the rules of its use (for instance, in the form of a definition), is to be given in an *exact* form, so as to introduce the explicatum into a well-connected system of scientific concepts.”
3. **Fruitfulness.** “The explicatum is to be a *fruitful* concept, that is, useful for the formulation of many universal statements (empirical laws in the case of a nonlogical concept, logical theorems in the case of a logical concept).”
4. **Simplicity.** “The explicatum should be as *simple* as possible; this means as simple as the more important requirements 1), 2) and 3) permit.” [7, page 7].¹⁴

¹⁴ Carnap adds, referring to the fourth requirement, that “[t]he simplicity of a concept may be measured, in the first place, by the simplicity of the form of its definition and, second,

In this paper, I am concerned with the question whether Church's thesis – as distinguished from Turing's thesis and other models of computability – can be viewed as an explication in Carnap's sense. I will start by checking the structural and substantial adequacy of this thesis against Carnap's four requirements.

2 Church's thesis as a Carnapian explication

The usual formulation of Church's thesis goes as follows:

[Church's thesis] A number-theoretic function is computable if and only if it is general recursive.

Undoubtedly, Church's thesis satisfies the condition of transforming a pre-systematic concept (explicandum) into an exact and scientific concept (explicatum). Church himself was very clear on this point:

The purpose of the present paper is to propose a definition of effective calculability which is thought to correspond satisfactorily to the somewhat vague intuitive notion in terms of which problems of this class are often stated, and to show, by means of an example, that not every problem of this class is solvable. [10, page 346].

Or:

We now define the notion [...] of an *effectively calculable* function of positive integers by identifying it with the notion of recursive of positive integers [...]. This definition is thought to be justified by the considerations which follows, so far as positive justification can ever be obtained for the selection of a formal definition to correspond to an intuitive one. [10, page 356].

Kleene (1952 [28]) wrote, in this connection:

Since original notion of effective calculability of a function (or of effective decidability of a predicate) is a somewhat vague intuitive one, [Church's thesis] cannot be proved [...]

While we cannot prove Church's thesis, since its role is to delimit precisely an hitherto vague conceived totality, we require evidence that it cannot conflict with the intuitive notion which it is supposed to complete; i.e., we require evidence that every particular function which our

by the simplicity of the forms of the laws connecting it with other concepts." [7, page 7]. He also emphasizes that simplicity is of secondary importance, since many scientific theories involve complex concepts that turn out to be highly useful. In general, simplicity comes into consideration only in cases where there is a question of choice among several concepts which achieve about the same similarity and exactness, and seem to be equally fruitful: if these concepts show a marked difference in their degree of simplicity, the scientist will, as a rule, prefer the simplest of them [7, page 7].

intuitive notion would authenticate as effectively calculable is [...] recursive. The thesis may be considered a hypothesis about the intuitive notion of effective calculability, or a mathematical definition of effective calculability; in the latter case, the evidence is required to give the theory based on the definition the intended significance. [28, 317–319].

When it comes to clarification of the intuitive concept of computability that Church was aiming at, the best account can be found in Sieg (1997)¹⁵. Sieg reconstructs the intellectual framework in which Church was aiming to formalise the concept of effective calculability and he thereby reconstructs Church's contribution to the analysis of the informal notion of effective computability. Sieg's reconstruction is the best we have to understand Church's legacy, as Church's own writings – unlike Turing's – lack transparency regarding his theoretical motivations, *i.e.*, lack clarification of the concept of computability he worked with or on [44, page 154].

It is implied by Sieg that Church had two aims in searching for an explication of the intuitive concept of computability. Firstly, Church aimed to find a way to determine what “constructive definability” is in order to formulate what is called today Kleene's representability thesis: “a formula can be found to represent any particular constructively defined function of positive integers whatever”. Church and his group believed that finding such a set of formulas would amount to constructing a logical system that would not be subject of Gödel's Incompleteness Theorems. This led Church to another aspect or clarification of the intuitive concept of computability. It gradually became crucial for him to capture the concept of the minimal step in reasoning.

When it comes to the domain for which an explication is to be formulated, the concept of recursivity is formulated for a precise domain: the structure of natural numbers (understood as abstract mathematical entities). Numerous passages in Church's writing clearly confirm the choice of this domain. For instance, in the passage quoted just at the beginning of this section we read:

We now define the notion [...] of an *effectively calculable* function of positive integers by identifying it with the notion of recursive function of positive integers. [10, page 356, my underline].

Kleene is equally explicit on this point:

We entertain various proposition about natural numbers [...] This heuristic fact [that all recognized effective functions turned out to be general recursive], as well as certain reflections on the nature of symbolic algorithmic processes, led Church to state the [...] thesis. [27, page 60, my underline].

The concept of computability is hence explained as referring to a particular set of functions on natural numbers and its formalization is executed within a well-known mathematical context, namely within axiomatic number theory,

¹⁵ See also Sieg's other papers for additional analyses, *e.g.*, [43] and [45]; and also Martin Davis' [14].

initiated by Peano and Dedekind, and proof theory, as studied by Hilbert. At this time axiomatic number theory still was interpreted materially, *i.e.* axioms were formulated in such a way that their intended interpretation – the domain of natural numbers – was obvious. Hilbertian formalization was a very new way of describing mathematical reasoning and mathematical structures. Church’s understanding of intuitive computability followed - as Sieg suggests - a similar development. Church’s aim was at first to express something specific about natural numbers, and it gradually became more abstractly oriented towards an analysis of minimal steps between stages.

The explicatum in Church’s thesis – expressed with logico-mathematical concepts *i.e.*, recursive functions as given by a semiformal, semi-interpreted system of axioms with recursive definitions – fulfils the Carnapian requirements:

The explicatum is given by explicit rules for its use, for example, by a definition which incorporates it into a well-constructed system of scientific either logicomathematical or empirical concepts. [7, page 3].

It seems clear that the structure of Church’s thesis follows that of an explication and also that it is formulated for a clarified concept and with the intention of inserting the formal definition into a specific scientific framework. I am now in a position to determine the extent to which Church’s thesis satisfies the requirements that Carnap puts on an explicatum.

Concerning the *similarity* of the explicandum to the explicatum, Carnap requires that the explicatum be such that it could be used “in most cases in which the explicandum has so far been used”. When one thinks about the development of the language of science and conceptual clarification, it means that if at an earlier stage the expression “ x ” was used to refer to the concept X in the context Y (“ x ”, X , Y), then after the conceptual development took place and the concept X was succeeded by the concept X^* , the expression “ x ” refers to X^* (“ x ”, X^* , Y). Sometimes “ x ”, instead of shifting meaning, is replaced by “ x^* ” (like in the case of “fish” and “piscis”) (“ x^* ”, X^* , Y). As Carnap puts it:

The former concept has been succeeded by the latter in this sense: the former is no longer necessary in scientific talk; most of what previously was said with the former can now be said with the help of the latter (through often in a different form, not by simple replacement). [7, page 6].

The similarity requirement means that the transformation of X to X^* is such that extension of X^* significantly overlaps with the extension of X . For instance, when the explicatum is the concept of truth (X), clarified as a property of sentences (Y), Tarski’s theory of truth (X^*) regulates the use of the term “truth” (“ x ”), sometimes replaced by “ T -truth” (“ x^* ”). In most cases, a sentence that is informally true will also be true in Tarski’s sense, and the other way round. Similarly, the functions that are effectively computable

in an intuitive sense will correspond closely to the functions singled out by Church's explication of that same concept.

The second requirement that Carnap puts on the explicatum is that "the rules of its use are to be given in an *exact* form". This is the consequence of Carnap's demand that the explicatum be embedded in an already existing theory and in an accepted manner. In this way, the rules of its use can be specified in an exact way by means of precise definitions. Similarly, the axiomatic theory of recursion, together with associated rules for defining new functions, provides very specific information about how to use its terms and clearly determines which number-theoretical functions the formal predicate "being a recursive function" applies to.

When it comes to *fruitfulness*, there can be no doubt that the formalization of the concept of an effectively computable function has been tremendously valuable, giving rise to the whole new field of study in which many substantial results could be conclusively proved. Due to the axiomatization of the theory of recursive functions, it is possible not only to specify which functions are recursive, but also to study them from a meta-level, for example within a theory of computability, formal arithmetic, or complexity theory.

As Tennant (2015 [47]) explains this situation:

One of the benefits of (the presumed truth of) the Turing-Church Thesis is that we can re-visit the diagonal argument [...]. We can now *reprove* the claim that not all computable functions are total, *no longer making any use of [intuitive principles]*. This is because, courtesy of the Turing-Church Thesis, we can effectively enumerate the computable functions by effectively enumerating, 'instead', the *Turing-Machine computable* ones. The latter task can be accomplished with absolute precision, with no hostages to fortune in judging, intuitively, whether a given finite set of English instructions really does guarantee that the function in question is genuinely computable. This is because Turing machines themselves can be effectively coded *as numbers*, and, given any such number, one can effectively determinate *whether* it encodes a Turing machine, and, if so, exactly, *which* Turing machine it encodes. This removes all doubt from the claim, in the proof above, that *one can effectively enumerate all and only the (Turing-machine)-computable functions*. The diagonal argument in question now goes through as before. Not all computable functions are total. [47, page 149].

Another significant benefit of the Turing-Church Thesis is that it affords an exceptionally simple proof of Gödel's famous result that first-order arithmetical truth is not axiomatizable. [...]. The argument to be presented there will involve yet another application of the diagonal method. [47, page 149].

The functioning of the fruitfulness is depicted in Figure 1. It illustrates how we can gain new knowledge about computable function via the formalization provided by Church.

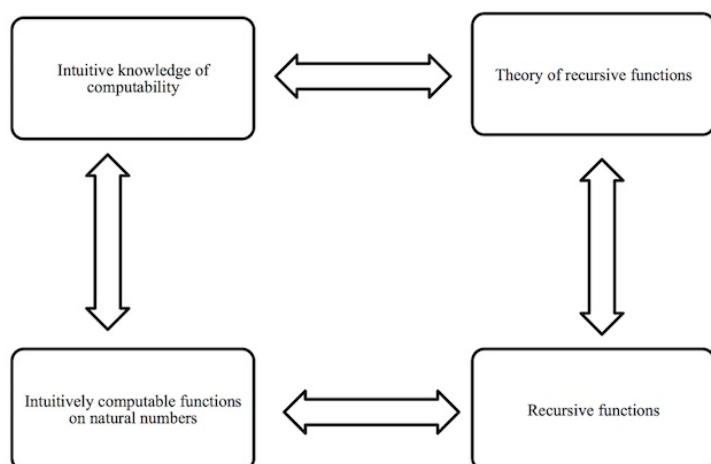


Fig. 1 Fruitfulness. There is an intuitive concept of what can be computed with the use of natural numbers and also an informal theory associated with this concept. Once the concept is formalized, it starts functioning in a formal theory that, ideally, should enable generalization of many intuitive laws. This schema comes from discussions with Erik J. Olsson.

Finally, *simplicity* is verified by (i) observing the striking simplicity of the definition of the class of functions that recursive functions are composed from, and also (ii) observing the simplicity of rules that are used to formulate new functions, and hence to prove new general laws (that opens up a possibility of connecting the precise concept of recursive function with other important concepts). Church, like Turing, instead of trying to capture all possible processes, proposes a small number of constraints from which all complex computations can be obtained. General Recursive Functions are most frequently defined as consisting of zero, successor, projection functions together with operations of composition, primitive recursion and search function (μ -operator), see Ender-ton (1972 [17, 18–20]).

3 Axiomatic systems as explications

In the face of the above analysis, it is difficult not to agree with Tennant and those others who claim that Church’s thesis can be seen as a Carnapian explication, indeed, a paradigmatic example of one. Surprisingly, however, there are elements of Carnap’s view that run counter to this – otherwise very natural – conclusion. Carnap was reluctant to accept without reservation what he refers to as “semiformal, semi-interpreted axiomatic systems” as a properly scientific form of explication. The theory of recursive functions, which is used in Church’s explication, belongs to this class. Accordingly, Carnap never mentions the Church-Turing thesis as an example of explication. Taking that into

account, I need to explicitly address his objection regarding the explicative power of semiformal, semi-interpreted axiomatic systems.

Most generally, Carnap explains his interest in the axiomatic method by saying that the axiomatic method can be used to introduce new concepts into the language of science, which is exactly the objective of the method of explication [7, page 15]. This section is devoted to reconstructing Carnap's analysis of the adequately explicating axiomatic systems and those, which – despite playing other important epistemic roles – do not explicate.

Carnap starts his discussion by highlighting that the axiomatic method consists of two steps, formalization and interpretation:

1. *Formalization* consists in the formulation of formal axioms indicating relational properties of the theory under study, as Carnap puts it:

The formalization (or axiomatization) of a theory or of the concepts of a theory is here understood in the sense of the construction of a formal system, an *axiom system* (or postulate system) for that theory. [7, page 15].

2. *Interpretation* consists of indicating which entities the theory refers to:

The *interpretation* of an axiom system consists in the interpretation of its primitive axiomatic terms. This interpretation is given by rules specifying the meaning, which we intend to give to these terms; hence the rules are of a semantical nature. [7, page 16].

Then Carnap distinguishes three types of axiomatic systems, which are – or have been – used in mathematical practice [7, page 15–16]. As I see it following Carnap's presentation, these types of axiomatisation can be differentiated by the way in which the two steps are implemented.

The first type is a *strictly formal axiomatic system*, by which he means:

a formal system in the strict sense, sometimes called a calculus (in the strict sense) or a syntactical system; in a system of this kind all rules are purely syntactical and all signs occurring are left entirely uninterpreted [...]. [7, page 15].

In *Logical Foundations of Probability* (1950 [7]) strictly formal systems of axioms are simply listed as a possible type of axiomatisation without further details, but an extensive discussion of their properties can be found in *Introduction to Semantics* (1942 [4]) and also in *Logical Syntax of Language* (1937 [2]). In [4] Carnap introduces the idea that the whole science of language (called *semiotic*) divides into three fields:

If in an investigation explicit reference is made to the speaker, or, to put it in more general terms, to the user of a language, then we assign it to the field of *pragmatics*. (Whether in this case reference to designata is made or not makes no difference for this classification.) If we abstract from the user of the language and analyze only the expressions and their designate, we are in the field of *semantics*. And if, finally, we abstract from the designate also and analyze only the relations between the expressions, we are in (logical) *syntax*. [4, page 9].

Furthermore, both syntax and semantics can be studied from two perspectives [4, page 12]:

- a *descriptive* approach, strongly entwined with pragmatics, consisting of empirical investigations of the semantical or syntactical features of historically given languages, or
- a *pure* approach, fully independent from all pragmatic considerations, consisting of laying down “definitions for certain concepts, usually in form of rules”, and studying “the analytic consequences of these definitions”. These rules can be formulated with intention of providing a model of existing “pragmatical facts”, but they can also be chosen in an arbitrary manner as a consistent set of sentences. [4, page 13, compare also page 155].

The *pure* approach to syntax opens up a possibility to abstract not only from the users of the language, but also from meaning, and in consequence study syntax without any appeal to semantics. The construction of a pure syntactical system of rules (called sometimes by Carnap a “calculus K”) consists of well-defined and strictly formal steps, including:

a classification of signs, the *rules of formation* (defining “sentence in K”), and the *rules of deduction*. The rules of deduction usually consist of primitive sentences and rules of inference (defining “directly derivable in K”). Sometimes, K contains also rules of refutation (defining “directly refutable in K”). If K contains definitions they may be regarded as additional rules of deduction. [4, page 155].

It should be now clear that in Carnap’s view a purely syntactical system cannot play the role of an explication as “for a genuine explication [...] an interpretation is essential” [7, page 16], whereas “in a system of this kind all rules are purely syntactical and all signs occurring are left entirely uninterpreted” [7, page 15]. The only function of a purely syntactical system is to “determine the procedure of formal deduction, *i.e.*, of the construction of proofs and derivations” [4, page 155].

The second type of axiomatic system that Carnap discusses in his (1950, [7]) is called a *semiformal, semi-interpreted axiomatic system*. Axiomatic systems of this type, can – if adequately developed – serve to introduce new terms to the language in such a way that these terms explicate.

The introduction of new concepts into the language of science - whether as explicata for prescientific concepts or independently - is sometimes done in two separate steps, formalization and interpretation. [...] The two steps are the two phases of what is known as the axiomatic (or postulational) method in its modern form. [...] We are not speaking here of a formal system in the strict sense, sometimes called a calculus (in the strict sense) or a syntactical system; in a system of this kind all rules are purely syntactical and all signs occurring are left entirely uninterpreted [...]. On the other hand, we are not speaking of axiom systems of the traditional kind, which are entirely interpreted.

In the discussions of this book we are rather thinking of those semiformal, semi-interpreted systems which are constructed by contemporary authors, especially mathematicians, under the title of axiom systems (or postulate systems). In a system of this kind the axiomatic terms (for instance, in Hilbert's axiom system of geometry the terms "point", "line", "incidence", "between", and others) remain uninterpreted, while for all or some of the logical terms occurring (e.g., "not", "or", "every") and sometimes for certain arithmetical terms (e.g., "one", "two") their customary interpretations is – in most cases tacitly – presupposed. [7, page 15–16].

However, not all semiformal and semi-interpreted axiomatic systems explicate, but only those – as we can read from Carnap's analysis – which formalization is formulated in such a way that it directs us smoothly towards an interpretation. Carnap's idea reminds us here of the Fregean constraint on formulation of first principles: any successful foundations of a mathematical theory must explicitly account, even at the most fundamental (for example, axiomatic) level, for applications of the entities forming the intended model of this theory.¹⁶ In more Carnapian terms we can say that the clarification that one targets while formulating axioms is such that it leaves no doubts which elements from the domain of the targeted scientific context are the objects that shall fall under the term being explicated. If the formalization does not fulfil this requirement, a semiformal, semi-interpreted axiomatic system does not explicate, but rather plays – in exactly the same way as purely syntactic systems – a deductive role and provides support in inferential contexts.

Carnap presents his (quite standard) view on systems of semiformal, semi-interpreted axioms in [3, SS16]. It consists of two parts: a basic *logical* calculus and a *non-logical* part.

- The basic logical calculus:

could be approximately the same for all those calculi; it could consist of the sentential calculus and a smaller or greater part of the functional calculus as previously outlined. [3, page 37].

And it is:

[...] essentially the same for all the different specific calculi, it is customary not to mention it at all but to describe only the specific part of the calculus. [3, page 38].

- The non-logical part contains terms specific to the subject matter of the theory, as Carnap explains:

The specific partial calculus does not usually contain additional rules of inference but only additional primitive sentences, called *axioms*. [...] What usually is called an *axiom system* is thus the second part of a calculus whose character as a part is usually not noticed. [3, page 37–38].

Later in the same section he also writes:

¹⁶ For a discussion of Frege's constraint see, for instance, [51].

An axiom system contains, besides the logical constants, other constants which we may call its specific or axiomatic constants. [3, page 38].

A system of semiformal, semi-interpreted axioms plays an explicative role if there is a straightforward manner of interpreting its logical and non-logical terms. Logical terms are most frequently interpreted in – as Carnap calls it – a “normal” way.

Not only is a basic logical calculus tacitly presupposed in the customary formulation of an axiom system but so also is a special interpretation of the logical calculus, namely, that which we called the normal interpretation. [3, page 38].

Non-logical terms in axiomatic systems are interpreted explicitly through semantical rules and explicit definitions using primitive signs and logical signs.

Some of [the specific or axiomatic constants] are taken as primitive; others may be defined. The definitions lead back to the primitive specific signs and logical signs. An interpretation of an axiom system is given by semantical rules for some of the specific signs, since for the logical signs the normal interpretation is presupposed. If semantical rules for the primitive specific signs are given, the interpretation of the defined specific signs is indirectly determined by these rules together with definitions. But it is also possible — and sometimes convenient, as we shall see — to give the interpretation by laying down semantical rules for another suitable selection of specific signs, not including the primitive signs. If all specific signs are interpreted as logical signs, the interpretation is a logical and L-determinate¹⁷ one¹⁸; otherwise it is a descriptive one. (Every logical interpretation is L-determinate; the converse does not always hold.) [3, page 38].

As I said earlier, there are several ways in which a semiformal, semi-interpreted axiomatic system can be formalized. In order to exhibit an explicative power, an axiomatic system needs to be formalized in such a way that there is an interpretation of its non-logical terms that points out the objects targeted at the stage of clarification.

Carnap (1950 [7]) illustrates the difference between an explicating semiformal, semi-interpreted axiomatic system and a non-explicating one with the

¹⁷ “We call a sentence of a semantical system S (logically true or) *L-true* if it is true in such a way that the semantical rules of S suffice for establishing its truth. We call a sentence (logically false) or *L-false* if it is false in such a way that the semantical rules suffice for finding that it is false. The two terms just defined and all other terms defined on their basis we call *L-semantical terms*. If a sentence is either L-true or L-false, it is called *L-determinate*, otherwise (*L-indeterminate* or) *factual*.” (Carnap 1939 [3, page 13]).

¹⁸ The class of *descriptive* signs are “those which designate things or properties of things”; the class of *logical* signs “serve chiefly for connecting descriptive signs in the construction of sentences but do not themselves designate things, properties of things, etc. Logical signs are, e.g., those corresponding to English words like “is”, “are”, “not”, “and”, “or”, “if”, “any”, “some”, “every”, “all”.” (Carnap 1939 [3, page 7]).

example of arithmetic. He considers – respectively – Peano Arithmetic (PA) and Frege Arithmetic (FA). Peano's axioms account for the behaviour of basic arithmetical functions (*i.e.*, successor, addition, multiplication) and it is only of further concern, on which mathematical entities these functions are defined. Clarification in the case of PA aims at explaining behaviour and mutual relations between elements of the natural number progression and not individual properties of these elements. Since various progressions of different mathematical entities satisfy these axioms, PA does not – according to Carnap – enable us to single out any specific interpretation of arithmetical terms and, in consequence, it does not provide an explication of the concept of natural number. It rather belongs – says Carnap – to the logical or proof-theoretic part of the theory. By contrast, Frege's axioms are formulated in such a way that individual mathematical entities are targeted already at the stage of clarification. The formalization of Frege's axioms opens up a natural way of identifying arithmetical terms with cardinalities of finite sets and, in consequence, enables an explication of the concept of natural number.

The third type of axiomatic system that Carnap indicates is a *fully interpreted axiom system of the traditional kind*, called also by many authors: *material axioms*. In such a system the stage of interpretation precedes the stage of formalization. All the symbols – logical or not – are assumed to be explicated or have definite meanings in advance. This axiomatisation does not play a role of an explication, with any explicative role played by explicit definitions used to define the non-logical terms beforehand. For instance, in the case of Euclidean geometry – the paradigmatic example of a material axiomatisation – primitive axiomatic terms “point”, “line” and “plane” are explicitly defined in the first place, and only then axioms regulating their behaviour are introduced. This is most probably the reason why Carnap does not pursue any discussion of the explicative power of material axiomatisations in his 1950.

According to Carnap, only semiformal, semi-interpreted axiomatic systems can – if carefully formalised – play a role of an explication. The main point of what I earlier called “Carnap's implicit objection” refers precisely to the necessity of carefulness in formalization. In order to determine which general features an axiomatic system should possess in order to have explicative power, it is crucial to understand the difference between explicating and non-explicating systems. In the next section, I will follow Carnap's example and go through the details of how PA and FA differ in this respect.

4 Carnap's implicit objection and arithmetic

Carnap ascribes explicative power only to those semiformal, semi-interpreted axiomatic systems that formalise a mathematical concept already explicitly targeted already at the stage of *clarification*. Only then – he claims – can the non-logical symbols be *interpreted* in such a way that the axiomatic system indeed *explicates* this concept. Carnap illustrates the difference between an explicating and a non-explicating semiformal, semi-interpreted axiomatic system

by discussing axiomatisations of the concept of natural number from the end of the XIX century: Second-Order Peano Arithmetic¹⁹ and Frege Arithmetic. He claims that the first of them does not explicate the concept of natural number, whereas the second perfectly fulfils all the criteria required from a successful explication of this concept. In this section, I analyse Carnap's example to understand which features of a semiformal, semi-interpreted axiomatic systems confer explicative power. On this basis I will assess the explicative power of the axiomatic system of recursion theory.

The idea that Frege's account of natural numbers is an exemplary explication appears in various places where Carnap presents his method. Lavers (2013 [30]) suggests that Carnap strongly believed that Frege's project did not only achieve a successful explication of the concept of natural number, but that the whole of Frege's endeavor can be reconstructed as an exemplary explication process. Indeed, Carnap shows great enthusiasm for Frege's project, for instance when he says in *Logical Foundations of Probability* (1950 [7]):

The first exact explications for the ordinary arithmetical terms have been given by G. Frege and later in a similar way by Bertrand Russell. Both Frege and Russell give explicata for the arithmetical concepts by explicit definitions on the basis of a purely logical system whose primitive terms are presupposed as interpreted. On the basis of this interpretation of the arithmetical terms, Peano's axioms become provable theorems in logic. It is a historically and psychologically surprising fact that this explication was such a difficult task and was achieved so late, although the explicanda, the elementary concepts of arithmetic, are understood and correctly applied in every child and have been successfully applied and to some extent also systematized for thousands of years. [7, page 17].

Or, when he states in his *Introduction to Symbolic Logic and Its Applications* (1958 [8]):

the concept of the inductive cardinal numbers [...] is an explicatum for the concept of finite number that has been widely used in mathematic, logic and philosophy, but never exactly defined prior to Frege. [8, page 2].

Similarly, the same idea guides his reply to Strawson on linguistic naturalism (1963 [9]):

With respect to the numerical words 'one', 'two', *etc.* [f]or thousands of years many people used these words adequately for all practical purposes, and for several centuries the mathematicians have had a systematically constructed theory involving these words. But even in this case, complete clarity was lacking. Before Frege, nobody was able to give an

¹⁹ The fact that Carnap means rather PA2 and not PA1 is confirmed by what Carnap writes while explaining what he means by Peano's axioms in his other writings, *eg.*, 1939 [3, SS17].

exact account of the meanings of these words in non-arithmetical terms. [...] Therefore we have to say that in spite of practical skill in usage, people in general, and even mathematicians before Frege, were not completely clear about the meaning of numerical words. [9, page 935].

This is not the place to assess Frege's intentions; Lavers [30] subjects Frege's intentions to close scrutiny²⁰, what I am interested in is Carnap's reconstruction of Frege's account. In what follows, I will indicate those passages from Carnap where he refers to Frege's account in terms corresponding to different parts of the explication process, which I identified in section 1.

The first step of this process is the *clarification* of the explicandum. Carnap sees Frege's declared objective in explicating the following prescientific concept:

the ordinary meaning of the word 'three' as it appears in every-day life and in science (*Two Concepts of Probability*, 1945 [5, page 513])

or

the concept of finite number that has been widely used in mathematics, logic and philosophy (*Introduction to Symbolic Logic and Its Applications*, 1958 [8, page 2]).

An explication always has a specific structure, consisting of "the transformation of an inexact, prescientific concept, the *explicandum*, into a new exact concept, the *explicatum*" [7, page 3]. It also involves indicating a domain within which the *explicatum* will be formulated. The concept of natural number shall, according to Carnap, to be explicated – as Frege wanted – in the context of set theory, where natural numbers are identified with finite cardinals. To explicate "three" Carnap recalls:

the definition of the cardinal number three [formulated] by Frege and Russell as the class of all triples. [5, page 513].

The second step of the explication process is the *specification* of the *explicatum*. Specification aims, most importantly, at a specific scientific context and is to be expressed using the part of the scientific language belonging to this context. Carnap thinks that in the context of set theory the *explicata* of individual numerical concepts are achieved by appealing to logical language. For instance, the concept of "three" is achieved by:

the concept of the class of all triples [which is] defined not by means of the word 'triple' but with the help of existential quantifiers and the sign of identity [...]. [5, page 513].

He also says:

²⁰ It is surprising that Lavers (2013 [30]), in his paper devoted to Carnap's justification of Frege's definition of natural number as explication, never refers to [7] where Frege's definition and Peano's axioms are contrasted, and never mentions Carnap's explicit rejection of Peano's axioms as explications.

By Frege’s explication of the numerical words, which I regard as one of the greatest philosophical achievements of the last century, the logical connection between these words and logical particles like ‘there is’, ‘not’, ‘or’, and ‘the same as’ became completely clear for the first time. [9, page 935].

This is for Carnap “the first exact explication for the ordinary arithmetical terms” and “[i]t is a historically and psychologically surprising fact that this explication was such a difficult task and was achieved so late” [7, page 17].

Carnap explicitly rejects Peano Arithmetic as explicating the concept of natural number. The way in which Peano’s axioms are formalized does not correspond to the clarification stage where natural numbers were targeted. Therefore, there exists no single intended interpretation that could be prioritized at the interpretation stage.

Peano’s axiom system, by furnishing the customary formulas of arithmetic, achieves in this field all that is to be required from the point of view of formal mathematics. However, it does not yet achieve an explication of the arithmetical terms ‘one’, ‘two’, ‘plus’, etc. In order to do this, an interpretation must be given for the semiformal axiom system. There is an infinite number of true interpretations for this system, that is, of sets of entities fulfilling the axioms, or, as one usually says, of models for the system. One of them is the set of natural numbers as we use them in everyday life. But it can be shown that all sets of any entities exhibiting the same structure as the set of natural numbers in their order of magnitude – in Russell’s terminology, all progressions – are likewise models of Peano’s system. From the point of view of the formal system, no distinction is made between these infinitely many models. However, in order to state the one interpretation we are aiming at, we have to give an explication for the terms ‘one’, ‘two’, etc., as they are meant when we apply them in everyday life. [7, page 17].

In this passage, Carnap claims that there is no distinguished interpretation of non-logical symbols of Peano Arithmetic that would single out a *unique* model satisfying Peano’s axioms. On the contrary, there always exist many possible interpretations, due to the fact that Peano’s axioms explicate the behaviour of individual mathematical objects named by arithmetical terms, but do not explicate terms themselves. In different words, Peano Arithmetic explicates relations between natural numbers, but it neither explicates what individual natural numbers are, nor explicates characteristics of the objects forming the class of natural numbers.

Carnap’s understanding of the concept of natural number is no doubts convergent with the Fregean one. However, assuming the state of knowledge today, this is not the only possible understanding, and there is a way to argue that Second Order Peano Arithmetic is actually one of the possible alternative explications. It seems to me – and I agree here with Lavers [30] – that Carnap, unlike Frege, would be open to accept other explications of the concept of natural number formulated in different scientific contexts.

Consider the claim that numbers are objects. Carnap would see introducing the numbers as objects as one possible way to give a systematic account of arithmetic. However, Carnap might also see introducing an axiom of infinity and then associating numbers with second order properties as another way to accomplish the same goal. The decision between these options Carnap might treat as arbitrary. But once a decision is taken here this does not make the propositions of the chosen system subjective, or dependent on us, or merely linguistic. It is no the objectivity of arithmetic that separated Frege and Carnap. What separated them was Frege's belief that anyone who attempted the same goal – providing a systematic account of arithmetic – would have to make essentially the same pragmatic choices as he did. Carnap, however, would understand Frege as holding this view because of the lack, at the time, of any alternative equally systematic treatment of number. Carnap would, that is, see Frege's fault as being too far ahead for his time, and not as being metaphysically misguided. [30, page 240].

Indeed, Carnap himself recognized that Peano's and Frege's arithmetics have different expressive powers. Peano's axioms enable the expression of information about structural properties of the natural number progression, whereas Frege's axioms allow us to speak of individual numbers. As Carnap puts it:

It is important to see clearly the difference between Peano's and Frege's systems of arithmetic. Peano's system, as mentioned, does not go beyond the boundaries of formal mathematics. Only Frege's system enables us to apply the arithmetical concepts in the description of facts; it enables us to transform a sentence like 'the number of fingers on my right hand is 5' into a form which does not contain any arithmetical terms. Peano's system contains likewise the term '5', but only as an uninterpreted symbol. It enables us to derive formulas like ' $3 + 2 = 5$ ', but it does not tell us how to understand the term '5' when it occurs in a factual sentence like that about the fingers. Only Frege's system enables us to understand sentences of this kind, that is to say, to know what we have to do in order to find out whether the sentence is true or not. [7, 17–18].

The idea that axiomatic systems understood as implicit definitions can be used as explications of mathematical concepts dates to Gergonne's paper from (1818 [22]), however it took till 1965 and the paper by Benacerraf "What Natural Numbers Could Not Be" [1], where the author argues against Frege-Russell view of numbers as objects, and the subsequent "structuralist turn", that it made its way to the mainstream. According to mathematical structuralism Peano's axioms explicate when the concept of natural number is targeted in clarification and it is assumed that what clarifies this concept are structural properties between natural numbers. Since mathematics is – according to mathematical structuralism – a science of structures, and every progression is an equally good representation of natural numbers and natu-

ral numbers are adequately defined by PA2.²¹ Unlike in Carnap's time, when many remained sceptical of the expressiveness of implicit definitions, today structuralism is one of the most studied positions in philosophy of mathematics. In consequence, PA2 is considered to be as successful explication of the concept of natural number as the one proposed by Frege.

The endeavour of explaining what is crucial for constructing an explicating semiformal, semi-interpreted axiomatic system guided me back towards the question of the explicative power of Church's thesis. In section 2, I showed that it satisfies all the requirements on a Carnapian explication. What remains to be done in section 5, is to verify that there is a way to overcome Carnap's implicit objection, which means, a way to find a clarification that motivates the formalization.

5 Multiplicity of clarifications of computability

The possibility of accepting a multiplicity of clarifications for one concept closely corresponds to Carnap's ideal of scrutinizing the language of science relative to various scientific frameworks.²² Accordingly, the same intuitive or presystematic concept can be clarified, and then explicated, in different manners depending on the context. Earlier in this paper, I indicated two possible clarifications of the concept "natural number": the first, within set theory, as cardinalities of collections, further formalised with use of purely logical tools²³; the second, from a structuralist perspective, as a sequence of elements that can be identified by their relational dependencies, again, formalised with purely logical tools²⁴ as an implicit definition. The first clarification was strongly supported by Carnap, whereas he did not consider the second at all. Moreover, Carnap explicitly opposed the idea that "natural numbers" can be explicated with an implicit definition that does not distinguish between various progressions of discrete elements. However, as I suggested in the previous section, Carnap operated before mathematical structuralism became mainstream in philosophical thinking, and there is no reason to think that he would refrain from accepting the second clarification today.

In this section, I investigate the possibility of finding a clarification, or clarification, of the intuitive concept of computability that could justify formalization by using the axioms of recursion theory.

Discussions about what was a correct clarification of the concept of computability were vivid from the very beginning of search for the formalization of

²¹ In general, according to Shapiro's structuralism (1991 [42]), any coherent second-order theory (which is categorical, *i.e.*, all its models are isomorphic) adequately defines a mathematical structure.

²² I will not be exploring this path any further, but it is worth noting that this inclusive attitude is certainly related to Carnap's conventionalist understanding of language. See also, [49].

²³ Frege's understanding of which tools are "logical" is obviously debatable, but not necessary for my argument.

²⁴ Again, I will not enter the discussion.

this concept. Gödel famously criticized Church's original explication of computability based on the λ -calculus and his criticism was mainly directed at the clarification of the explicandum. As reported by Church himself in a letter to Kleene (dated November 29, 1935, but referring presumably to events from 1934):

In regard to Gödel and the notions of recursiveness and effective calculability, the history is the following. In discussion [sic] with him the notion of lambda-definability, it developed that there was no good definition of effective calculability. My proposal that lambda-definability be taken as a definition of it he regarded as thoroughly unsatisfactory. I replied that if he would propose any definition of effective calculability which seemed even partially satisfactory I would undertake to prove that it was included in lambda-definability. His only idea at the time was that it might be possible, in terms of effective calculability as an undefined notion, to state a set of axioms which would embody the generally accepted properties of this notion, and to do something on that basis. Evidently it occurred to him later that Herbrand's definition of recursiveness, which has no regard to effective calculability, could be modified in the direction of effective calculability, and he made this proposal in his lectures. At that time he did specifically raise the question of the connection between recursiveness in this new sense and effective calculability, but said he did not think that the two ideas could be satisfactorily identified "except heuristically." [14, page 9].

As suggests Sieg [44], against Davis [14], Church himself was skeptical regarding explication of effective calculability by λ -definability.

The fact that the thesis was formulated in terms of recursiveness indicates also that λ -definability was at first, even by Church, not viewed as one among equally natural definitions of effective calculability: the notion just did not arise from an analysis of the intuitive understanding of effective calculability. [44, page 157].

It is also a well-known fact that Gödel believed that Turing's approach to computability was much more convincing, because it was aiming at Gödel's notion of *absolute* computability, or – in different words – computability independent of the formal system for which it is defined.

In this paper, I claim that there can exist more than one clarification of the intuitive concept, that leads to its adequate explication. It is also possible that multiple clarifications support use of the same formalization in the explication of the given concept. This is actually the case of the axioms of recursion theory in the context of Church's thesis.

Like any other semiformal, semi-interpreted axiomatic system, the axioms of recursion theory need a strategy to overcome *Carnap's implicit objection*. Overcoming *Carnap's implicit objection* requires pointing out such a clarification and such a context of formalization that provides a justification for use of

an implicit definition. For the case of Church's thesis, I will indicate two possible clarifications of the intuitive or presystematic concept of computability that are successfully captured by the formal concept of recursive function. The first clarification states that intuitive computability corresponds to the class of computable functions on natural numbers, the second states that intuitive computability corresponds to a domain-independent procedural understanding of effective computation. As Sieg (1997 [44]) points out, both were considered by Church. Sieg, while reconstructing Church's conceptual analysis of the concept of effective computability, suggests that Church's approach changed from the original *material* interpretation, towards a "step-by-step" *procedural* interpretation of recursion.

The two clarifications of the concept of computability that enable the axioms of recursion theory to overcome *Carnap's implicit objection* correspond exactly to those two ways in which one can look at the system of recursive functions. Firstly, the axioms of recursion theory can be seen as accounting for certain feature of a certain class of functions defined on *natural numbers*, namely *computability*, or more precisely *constructive definability* that is able "to represent any particular constructively defined function of positive integer whatever". I accounted for this approach while introducing Church's original motivation for formulating his thesis.

Under the first understanding, the theory of recursive functions relies on the concept of natural number adopted as primitive, or adequately defined beforehand. Therefore, the explication targets a specific class of functions defined on a specific domain, and achieves this objective in a successful way. The concept of effectively computable is clarified by the concept of a computable function on natural numbers. In consequence, Carnap's implicit objection is obviously easily avoidable.

The fact that the theory of computation is always domain dependent is often seen as a drawback. This is exactly what Gödel objected to Church's thesis while calling for *absoluteness* of the concept of computation. Today we know that even the model of computation favoured by Gödel, namely Turing's one, is strictly dependent on the domain (entities that can be the subject of computations).²⁵ What makes an important intensional difference between the two models is the starting point of clarifying the intuitive concept of computation. Church, at least initially, was interested in characterising a class of computable functions on natural numbers, whereas Turing was aiming at formalising effective procedures and wanted to explicate computability as minimal processes that can be carried by an idealised human computer.

The procedural approach is most often associated with Turing's account²⁶, however, when Sieg [44] reconstructs Church's analysis of computability, he indicates that even if Church's first motivation was "quasi-empirical" (tar-

²⁵ There exist strong arguments that that, in general, computability and an ω -progression are tightly related. See [24] and [35].

²⁶ It is apparent in both, his work on *Entscheidungsproblem* and his work on the concept of calculation in a logic. I am grateful to an anonymous reviewer for directing my attention towardon the latter.

getting the idea of representability of computable functions in arithmetic), it gradually (under the influence of Gödel and Turing) became more oriented towards capturing what the minimal constructive steps that can be performed on natural number are, and which functions are the constructive ones.

Under this second understanding, the theory of recursive functions belongs to the framework of uninterpreted axiomatic systems where no definition of non-logical symbols is given in explicit manner. In consequences, all interpretations of terms and functions are relative to the model and it opens up a possibility of studying computations as pure procedures in a domain-independent manner. It corresponds to the idea, explored increasingly often, that the axioms of recursion theory can be studied from the contemporary model theoretic perspective as one of many entwined axiomatic systems. This is strictly related to a procedural understanding of effective computation. It would be an overstatement to say that Church himself fully abstracted from natural numbers, but one can risk a thesis that its continuation can be seen in model-theoretical skepticism, that today is a widely studied position in philosophy ([34], for overview see [15]).

Whichever line of defence we chose, recursion theory is saved in its role of explicating the intuitive concept of computability. Moreover, the possibility of providing different clarifications of computability for different scientific contexts is a good thing for those authors, quoted in the introductory part of this paper, who advocate distinguishing Church's thesis from Turing's thesis.

6 Conclusions

In this paper, I examined the idea that Church's thesis is an explication in the Carnapian sense. I claimed that a positive answer to this inquiry necessitates not only a successful verification of Carnapian requirements, but that clarification of computability needs to explicitly aim at the intended interpretation of the non-logical symbols of the axioms of the theory of recursion.

Therefore, my first assignment was in understanding the role of "the clarification of the explicandum" in the case of an axiomatic semiformal, semi-interpreted system. I relied on Carnap's own example, namely axiomatic accounts of natural numbers. According to Carnap, when the concept of natural number is clarified by Frege-Russell understanding of natural numbers, and explicated by a semiformal, semi-interpreted axiomatic system that targets accounting for these (by Frege Arithmetic), this system fulfils the requirements of an explication. By contrast, Peano's semiformal, semi-interpreted axiom system cannot be used as an explication of the concept of natural number, because individual constants and other denoting arithmetical terms will always lack an explicit interpretation in that system.

However, as I also argued, from a structuralist perspective Carnap's judgment on Peano Arithmetic seems unnecessarily restrictive. A structuralist thinks of the natural numbers not as individual mathematical entities, but as positions in a structure. This idea can be combined with the observations

that PA2, while being a semiformal, semi-interpreted system, is categorical: all models that satisfy the axioms share the same standard structure. The structuralist could now argue that while PA2 is not itself an interpreted system, in the sense that the denoting terms are given explicit interpretations, it still singles out a particular structure in this more abstract sense. It is a natural next step to think that the intuitive concept of natural number can be “explicated” in terms of this abstract structure singled out by the axioms of PA2: “one” denotes the first position, “two” the second, and so on. For in this way, we achieve a “transformation of an inexact, prescientific concept . . . into a new exact concept . . .”, which is Carnap’s own high-level description of what explication is all about [7, page 3]. In this sense, a semiformal, semi-interpreted axiom system can – contrary to what Carnap writes – achieve an explication so long as that axiom system is categorical and we adopt a suitable version of structuralism in mathematics.

The same line of reasoning can be applied to recursive functions: in order to defend them as explicating one needs to find a suitable clarification. In this paper I suggested two possible ways of proceeding. Firstly, Church’s thesis, as it is originally stated, is embedded in number theory and hence the model of its interpretation is fixed from the beginning in a way which straightforwardly satisfies Carnap’s criteria for a good explication. Secondly, I observe that thanks to the emergence of mathematical structuralism and of model-theoretic scepticism as full-fledged paradigms for clarification of mathematical concepts, many semiformal, semi-interpreted axiomatic systems, including recursion theory or Peano Arithmetic, can be seen as explicating.

In consequence, in this paper I shed an additional light on what Carnap is actually saying in his most important text on explications, about semiformal, semi-interpreted axiomatic systems and the reasons for which they can or cannot play the role of an explication. Under my understanding, Carnap claims that such systems explicate only when *explicandum* is *clarified* in a way that it enables such a formalization that the intended interpretation of the non-logical predicates, placeholders of explicata, is straightforward. I observe that this is very close to Frege’s foundational principle, the – so called – Frege’s constraint according to which any successful foundations of a mathematical theory must explicitly account, even at the most fundamental level (*e.g.*, axiomatic), for applications of the entities forming the intended model of this theory.

Acknowledgements I would like to thank Erik J. Olsson (Philosophy Department, Lund University) for many inspiring conversations about the explicative role of CTT, and for the feedback on earlier versions of this paper. I am indebt to Marianna Antonutti and Benedict Eastaugh (Munich Center for Mathematical Philosophy) for all the comments that greatly improved the clarity of my argument. I am very grateful to Hannes Leidegeb (MCMP) for discussing with me Carnap’s philosophy. Finally, thanks to all the participants of the MCMP seminar, to my colleagues from the Philosophy Department in Lund and to my colleagues from the International Center of Formal Ontology at Warsaw Technical University for feedback and encouragements. Last, but not least, I would like to thank two anonymous referees who made sure that the paper is sound and complete.

References

1. Benacerraf, Paul (1965), What Numbers Could Not Be. *Philosophical Review* 74 (1): 47–73.
2. Carnap, Rudolf (1937), *The Logical Syntax of Language*. Open Court Publishing.
3. Carnap, Rudolf (1939), *Foundations of Logic and Mathematics*. University of Chicago Press.
4. Carnap, Rudolf (1942), *Introduction to Semantics*. Harvard University Press.
5. Carnap, Rudolf (1945), Two Concepts of Probability. *Philosophy and Phenomenological Research* 5 (4): 513–532.
6. Carnap, Rudolf (1947), *Meaning and Necessity*. The University of Chicago Press.
7. Carnap, Rudolf (1950), *Logical Foundations of Probability*. Routledge and Kegan Paul.
8. Carnap, Rudolf (1958), *Introduction to Symbolic Logic and Its Applications*. Dover Publications.
9. Carnap, Rudolf (1963), Reply to P. F. Strawson on Linguistic Naturalism. In: Schilpp, Paul Arthur (ed.), *The Philosophy of Rudolf Carnap*. Open Court Publishing.
10. Church, Alonzo (1936), An unsolvable problem of elementary number theory. *American Journal of Mathematics* 58: 345–363.
11. Church, Alonzo (1937), Review of Turing [1936]. *Journal of Symbolic Logic* 2: 42–43.
12. Copeland, Jack & Proudfoot, Diana (2010), Deviant encodings and Turing's analysis of computability. *Studies in History and Philosophy of Science* 41: 247–252.
13. Davis, Martin (1965), *The Undecidable*. Dover Publications Inc.
14. Davis, Martin (1982), Why Gödel Didn't Have Church's Thesis. *Information and Control* 54: 3–24.
15. Dean, Walter (2014), Models and Computability. *Philosophia Mathematica* 22 (2): 143–166.
16. Dutilh Novaes, Catarina & Reck, Erich (2017), Carnapian explications, formalisms as cognitive tools, and the paradox of adequate formalizations. *Synthese* 194: 195–215.
17. Enderton, Herbert B. (1972), *A Mathematical Introduction to Logic*. Academic Press.
18. Floyd, Juliet (2012), Wittgenstein, Carnap and Turing: Contrasting Notions of Analysis. In (Wagner-Beaney 2012): 34–47.
19. Floyd, Juliet & Bokulich, Alisa (eds.) (2017), *Philosophical Explorations of the Legacy of Alan Turing: Turing 100*. Springer 2017.
20. Gandy, Robin (1980), Church's Thesis and Principles for Mechanisms. In: Barwise, Jon; Keisler H. Jerome & Kunen, Kenneth (eds.) *The Kleene Symposium*. North-Holland Publishing Company: 123–148.
21. Gandy, Robin (1988), The confluence of ideas in 1936. In: Herken Rolf (ed.), *The Universal Turing Machine: A Half-Century Survey*. Oxford University Press: 55–111.
22. Gergonne, Jose Diez (1818), Essai sur la théorie des définitions. *Annales de mathématique pure et appliquée* 9: 1–35.
23. Gödel, Kurt (1937?), Undecidable Diophantine Propositions. In: Gödel, K., *Collected Works, Volume III, Unpublished essays and lectures*, Feferman S., et al. (eds.), Oxford University Press 1995: 164–175.
24. Halbach, Volker & Hørsen, Leon (2005), Computational Structuralism. *Philosophia Mathematica* 13 (2): 174–186.
25. Hilbert, David & Ackermann, Wilhelm (1928), *Grundzüge der theoretischen Logik (Principles of Mathematical Logic)*. Springer-Verlag.
26. van Heuveln, Bram (2000), *Emergence and consciousness: Explorations into the Philosophy of Mind via the Philosophy of Computation*. Ph.D. thesis. State University of New York at Binghamton.
27. Kleene, Stephen Cole (1943), Recursive Predicates and Quantifiers. *Translations of the American Mathematical Society* 53 (1): 41–73.
28. Kleene, Stephen Cole (1952), *Introduction to Metamathematics*. North-Holland.
29. Langford, Cooper Harold (1942/1968), The notion of analysis in Moore's philosophy. In: Schilpp, Paul Arthur (ed.), *The Philosophy of G.E. Moore*. Open Court Publishing: 321–342.
30. Lavers, Gregory (2013), Frege, Carnap, and Explication: "Our Concern Here is to Arrive at a Concept of Number Usable for the Purpose of Science". *History and Philosophy of Logic* 34 (3): 225–241.

31. Machover, Moché (1996), *Set Theory, Logic and their Limitations*. Cambridge University Press.
32. Mendelson, Elliott (1990), Second Thoughts about Church's Thesis and Mathematical Proofs. *The Journal of Philosophy* 87 (5): 225–233.
33. Murawski, Roman & Woleński, Jan (2006), The Status of Church's Thesis. In: Olszewski, Adam; Woleński, Jan, & Janusz, Robert (Eds.), *Church's Thesis after 70 Years*. Ontos Verlag: 310–330.
34. Putnam, Hilary (1980), Models and Reality. *Journal of Symbolic Logic* 45(3): 464–482.
35. Quinon, Paula & Zdanowski, Konrad (2007), Intended Model of Arithmetic. Argument from Tennenbaum's Theorem. In: Cooper, S. Barry; Kent, Thomas F.; Löwe, Benedikt & Sorbi, Andrea (eds.), *Computation and Logic in the Real World*, CiE: 313–317.
36. Quinon, Paula (2014), From Computability over Strings of Characters to Natural Numbers. In: Olszewski, Adam; Brożek, Bartosz & Urbańczyk, Piotr (eds.), *Church's Thesis, Logic, Mind & Nature*. Copernicus Center Press: 310–330.
37. Rescrola, Michael (2007), Church's thesis and the conceptual analysis of computability. *Notre Dame Journal of Formal Logic* 48 (2): 253–280.
38. Rescrola, Michael (2012), Copeland and Proudfoot on computability. *Studies in History and Philosophy of Science Part A* 43 (1): 199–202.
39. Schulz, Klaus-Dieter (1997), *Die These von Church. Zur erkenntnistheoretischen und sprachphilosophischen Bedeutung der Recursionstheorie*. Peter Lang Verlag.
40. Shagrir, Oron (2006), Gödel on Turing on Computability. In: Olszewski, Adam; Woleński, Jan & Janusz, Robert (eds.), *Church's Thesis after 70 years*. Ontos-Verlag: 393–419.
41. Shapiro, Stewart (1982), Acceptable Notation. *Notre Dame Journal of Formal Logic* 23 (1): 14–20.
42. Shapiro, Stewart (1991), *Foundations without Foundationalism. A Case for Second-order Logic*. Oxford University Press.
43. Sieg, Wilfried (1994) Mechanical procedures and mathematical experiences. In: George, Alexander (ed.), *Mathematics and Mind*. Oxford University Press: 71–117.
44. Sieg, Wilfried (1997), Step by Recursive Step: Church's Analysis of Effective Calculability. *The Bulletin of Symbolic Logic* 3 (2): 154–180.
45. Sieg, Wilfried (2002), Calculations by Man & Machine: Conceptual analysis. In: Sieg, Wilfried; Sommer, Richard & Talcott, Carolyn (eds.), *Reflections on the Foundations of Mathematics*: 396–415.
46. Soare, Richard (1996), Computability and Recursion. *Bulletin of Symbolic Logic* 2: 284–321.
47. Tennant, Neil (2015), *Introducing Philosophy. God, Mind, World, and Logic*, Routledge.
48. Turing, Alan (1936), On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* 42: 230–265; correction in (1937) 43: 544–546; reprinted in (Davis 1965: 115–154); page numbers refer to the (1965) edition.
49. Wagner, Pierre & Beaney, Michael (eds.) (2012), *Carnap's Ideal of Explication and Naturalism*. Palgrave Macmillan.
50. Wang, Hao (1974), *From Mathematics to Philosophy*. Routledge Kegan & Paul.
51. Wright, Crispin (2000), Neo-Fregean Foundations for Real Analysis. *Notre Dame Journal of Formal Logic* 41(4): 317–33.